Investigating cofactor transfer for a B_{12}-dependent enzyme

by

Alexander T. Duong

B.S. Chemistry
University of Florida, 2022

Submitted to the Department of Chemistry
In Partial Fulfilment of the Requirements for the Degree of

Master of Science in Chemistry

at the
Massachusetts Institute of Technology

May 2024

Signature of Author _____

Alexander T. Duong
Department of Chemistry
May 17, 2023

Certified by _____

Catherine L. Drennan
Professor of Biology and Chemistry
Investigator and Professor, Howard Hughes Medical Institute
Thesis Supervisor

Accepted by _____

Adam P. Willard
Professor of Chemistry
Graduate Officer

Investigating cofactor transfer for a B_{12}-dependent enzyme

By

Alexander T. Duong

Submitted to the Department of Chemistry on May 17, 2024
in partial fulfilment of the requirements for the Degree of Master of Science in Chemistry

Abstract

The metallocofactors utilized by enzymes can range in complexity from single metal ions to organometallic cofactors well over 1000 Da. These cofactors enable these metalloenzymes to accomplish a diverse set of unique and challenging chemistry that are critical to core life functions. One of these metallocofactors, adenosylcobalamin (AdoCbl), has only one cognate enzyme in humans: methylmalonyl-CoA mutase (MCM), which is involved in the catabolism of several amino acids, cholesterol, and odd-chain fatty acids. MCM relies on two other proteins, a G-protein metallochaperone called methylmalonic aciduria type A protein (MMAA) and a protein called adenosyltransferase (ATR), to load and off-load cofactor. Mutations or deletions of the gene for MCM, or in any of the genes corresponding to accessory proteins which interfere with cofactor delivery and removal, can lead to a potentially lethal inborn error in metabolism. If the cofactor becomes damaged in the active site of MCM, ATR unloads the cofactor, repairs it, and reloads the regenerated AdoCbl onto the mutase. A molecular understanding of this process has been challenging to obtain due to the difficulty of structurally characterizing a three-protein MCM-MMAA-ATR complex that is transient in nature. An orthologous protein from *C. metallidurans* in which the G-protein metallochaperone is naturally fused to its target mutase isobutyryl-CoA mutase (IcmF) provides an alternative two-protein IcmF-ATR system for structural and biochemical characterization. Recent work has shown that the IcmF system utilizes a mechanism of active site opening similar to non-fused systems like that of humans. However, the mechanisms by which ATR recognizes the presence of damaged cofactor and then removes it remains unclear. In this thesis, we discuss the development of an assay based on UV-Vis spectroscopy to monitor cofactor transfer between IcmF and ATR. We also discuss efforts to substitute histidine residues in IcmF suspected of serving as intermediate binding sites during cofactor transfer, with the goal of using the developed assay as a means of observing potential changes in transfer efficiency by perturbing these histidine residues. This work seeks to improve our understanding of AdoCbl-dependent enzyme maturation, and inform our ability to harness their unique reactivity.

Thesis Supervisor: Catherine L. Drennan

Title: Professor of Biology and Chemistry; Investigator and Professor, Howard Hughes Medical Institute

Table of Contents

Author Contributions

. The crystallization, data collection, processing, and refinement of the apo-*Cm*ATR structure was done by Daphne A. Faber and Francesca A. Vaccaro. The alignments of ATR and the docking experiments, as well as Figures 4, 5, and 6, are from the thesis work of F. A. Vaccaro. Processing and partial refinement of the IcmF with extra-cobalamin dataset, collected in 2014 by David A. Born and Marco Jost, was done by D. A. Faber. The stopped-flow experiments and data analysis were done with the equipment, expertise, and guidance of Maria-Eirini Pandelia at Brandeis University. All other experimental work and written sections were done by Alexander T. Duong. Catherine L. Drennan, F. A. Vaccaro, A. T. Duong, Alison E. P. Biester, and Gisele A. Andree contributed to edits.

I.   *Introduction*

B12-dependent enzymes in humans

Cobalamin (Cbl or Vitamin B_{12}) has only two coenzyme forms, methylcobalamin and adenosylcobalamin (AdoCbl), that are known cofactors in mammalians (**Figure 1A**) [1]. Of the B_{12}-dependent enzymes in mammals, only one uses AdoCbl: methylmalonyl-CoA mutase (MCM), which catalyzes the 1,2-rearrangement of (R)-methylmalonyl-CoA to succinyl-CoA, ultimately affecting the overall catabolism of several amino acids, cholesterol, and odd-chain fatty acids (**Figure 1B**) [2,3]. MCM is composed of two domains: a triose-phosphate isomerase (TIM) barrel domain and a Rossmann domain (**Figure 2A**) [4]. Similar to many other Cbl-dependent enzymes with a Rossmann fold, such as glutamate mutase and methionine synthase, the Rossmann domain in MCM is the Cbl-binding domain [5,6]. The dimethylbenzimidazole (DMB) tail of AdoCbl, which is normally the lower axial ligand, is displaced by a histidine residue when bound to MCM [4]. The TIM barrel is the substrate-binding domain, with the enzyme active site residing at the interface of the two domains. When the two domains are clamped shut, the substrate and cofactor are brought near one another and the reaction can proceed.

The capability for MCM to accomplish this challenging chemistry comes from the highly reactive 5′-deoxyadenosyl primary radical generated following homolytic cleavage of the cobalt-carbon bond of AdoCbl [7,8]. If the radical on the deoxyadenosyl moiety is lost or if the cofactor is poisoned through oxygenation, the reactivity of AdoCbl, and thus the mutase, is lost. In this event, the catalytically dead cofactor must be removed, regenerated, and reinserted into the active site of the mutase. This task is undertaken by an enzyme called adenosyltransferase (ATR), a homotrimeric protein which removes the damaged cofactor, re-adenylates it to AdoCbl using ATP, and reloads the repaired cofactor onto the mutase [9].

To prepare the mutase active site for cofactor loading and off-loading, a P-loop G-protein metallochaperone known as methylmalonic aciduria type A protein (MMAA) in humans is utilized [10,11]. In the human system and others like it, the G-protein metallochaperone functions as a dimer [12,13]. Upon GTP binding, conformational changes in the G-protein chaperone dimer wedge open the mutase, exposing the active site that lies between the Rossmann and TIM barrel domains [14]. The DMB tail of AdoCbl, which is more solvent exposed when bound to ATR, provides a molecular handle to facilitate cofactor delivery onto MCM [15]. After delivery of the AdoCbl by the ATR, GTP hydrolysis facilitates a conformational change in G-protein chaperone [10,11,16,17] that then clamps shut the Rossmann and TIM barrel domains, sealing away the AdoCbl (**Figure 3A**) [14,18]. Mutations that cause loss of activity in MCM, either directly in the mutase itself or by affecting the AdoCbl delivery pathway, cause an inborn error of metabolism and a potentially lethal accumulation of methylmalonic acid in the blood, medically referred to as methylmalonic aciduria [19–21].

Isobutyryl-CoA mutase fused – a model system for studying B12-dependent enzyme maturation

There are several structures of the G-protein chaperone in complex with its target AdoCbl-dependent mutase.[14,18,22,23]. However, there are no structures of a mutase-G-protein chaperone-ATR complex, either during AdoCbl loading or off-loading of damaged cofactor. As a tripartite system with inherently high heterogeneity, the human system is challenging to isolate and characterize using current structural biology techniques. There exists an orthologous, two-part system in the gram-negative bacteria *Cupriavidus metallidurans*, reducing the flexibility of the system for structural characterization. The AdoCbl-dependent mutase, isobutyryl-CoA mutase fused (IcmF) contains a Cbl-binding Rossmann domain and a substrate-binding TIM barrel domain, but additionally contains a P-loop GTPase domain between these domains which has high sequence similarity to MMAA (**Figure 2B**) [12,17,23,24,24,25]. IcmF catalyzes several similar rearrangements of isobutyryl-CoA to *n*-butyryl-CoA, pivalyl-CoA to 2-methylbutyryl-CoA, and

isovaleryl-CoA to valeryl-CoA [24,26,27] (**Figure 1B**). Furthermore, *Cm*ATR is a PduO-type ATR, the same as human ATR [28], simplifying mechanistic comparisons.

Previous work has shown that the maturation mechanism for the fused *C. metallidurans* system only differs by how the G-protein metallochaperone – fused to the mutase – reaches an active state [12–14,18,23]. MMAA - and a homolog in bacterial AdoCbl-dependent mutase systems, MeaB - exist as dimers in solution and open the active site of their cognate mutases by conformational change following GTP binding. However, IcmF requires oligomerization into a supramolecular structure to achieve the "dimerization" of the G-protein domain that can prop open the Rossmann and TIM barrel domains upon GTP binding [18]. Data have been presented to indicate that IcmF oligomerization plays a critical role in cofactor off-loading and repair, as this G-protein domain interface will be necessary to reopen the mutase domains for cofactor removal (**Figure 3**). As the cofactor is moved from IcmF to its partner ATR, cob(II)alamin goes from 5-coordinate base-off/His-on to 4-coordinate base-off, where the axial H39 ligand on IcmF is replaced by the non-ligating F110 on ATR [23,29]. The naturally occurring fusion of the AdoCbl-dependent mutase and its G-protein metallochaperone, in addition to a human-like ATR, makes the two-part system from *C. metallidurans* an attractive model to study the maturation of AdoCbl-dependent enzymes. Insights into the maturation mechanism of the *C. metallidurans* system has shed light on how standalone systems, such as that of humans, function in both wild-type conditions and in disease states associated with methylmalonic aciduria. Elucidating the off-loading and cofactor repair mechanism can similarly further our understanding of the molecular underpinnings of AdoCbl-dependent mutase functionality.

The structure of adenosyltransferase from C. metallidurans

Recent work from previous members of the Drennan lab has yielded the structure of apo-*Cm*ATR [29]. Alignments using ClustalW with other PduO-type ATRs, including from humans (PDB: 6D5X) and *Mycobacterium tuberculosis* (6WGV) show high sequence homology and conserved

7

sequence motifs (**Figure 4**) [15,29–31]. There are structures with various substrates and products bound that help elucidate the various states of ATR. These structures include: *Hs*ATR with ATP and cob(II)alamin or AdoCbl bound; and the structure of *Mtb*ATR with AdoCbl and triphosphate (PPP$_i$) bound [15,31]. The high conformational similarity despite the differences in bound molecule indicate that *Cm*ATR, like other ATRs, is unlikely to undergo significant structural changes upon ATP and cob(II)alamin binding (**Figure 5A**) [29]. Additionally, these comparisons indicate that conserved residues in other ATRs, which form the solvent exposed, hydrophobic pocket necessary to bind an unligated cob(II)alamin are also present in *Cm*ATR. The conserved active site phenylalanine, F110 in *Cm*ATR (**Figure 5C**), is in the appropriate conformation to block solvent from coordinating the cobalt, enabling the unique 4-coordinate base-off state of cob(II)alamin that precedes cobalt reduction to cob(I)alamin and adenylation to form AdoCbl [9]. N-terminal residues known to bind ATP and coordinate Mg^{2+} ions in *Mtb*ATR are present in *Cm*ATR as well, but are disordered in the structure. Resolving these residues would likely require an ATR structure with ATP bound, which would allow for more structural comparisons with other PduO-type ATRs.

Docking experiments with the C. metallidurans system

Although no experimentally derived structures of the mutase-G-protein chaperone-ATR complex exist, there have been attempts to computationally predict the structure. Previous attempts to predict the structure of the ATR + IcmF complex used alphafold2 multimer and RosettaDocking4.0 [29,32,33]. Alphafold2 multimer was unable to predict a structure for the complex but was able to recapitulate the individual structures of the IcmF dimer and ATR trimer. The RosettaDocking4.0 docking experiments used the solved structure of *Cm*ATR and a previously solved structure of the apo-*Cm*IcmF dimer bound to GDP (PDB: 4XC8) [23] to predict by computation a structure for the ATR + IcmF complex. In this IcmF structure, the mutase active site is in an open conformation in one protomer of the dimer, making it competent for cofactor

delivery or removal; the other protomer has a closed active site, representing a state competent for catalysis [23]. In a complex suitable for cofactor delivery and removal, the active sites of IcmF and ATR are expected to be positioned near one another. However, the top docking solution with the maximum amount of contact between IcmF and ATR also positioned the active sites the furthest away from each other, over 50 Å apart (**Figure 6A**). Three other solutions, which together represented the other top solutions produced, place the active sites between 20 – 30 Å apart but have significantly reduced contact between the two proteins (**Figure 6B, C, and D**) [29]. The shortest distance between the active sites in the docking solutions is over 20 Å, likely still too far for a direct handoff of the cofactor to occur. From these predicted binding modes, additional conformational changes of IcmF would be required to reposition the Cbl-binding Rossmann domain closer to ATR. However, the existence of the G-protein domain linking the Rossmann domain to the TIM barrel domain conformationally restricts any potential rearrangements, and previous structural snapshots of the IcmF dimer have shown that Rossmann domain and the G-protein domain form a continuous β-sheet that moves in lockstep as one rigid body [23]. However, if the cofactor transfer is a multistep process and uses intermediary residues to bind the cofactor before delivering it to the active site, this distance becomes more tenable. If such putative intermediary binding sites exist, it is likely that they are located on IcmF, given that it is significantly larger than ATR. Additionally, it is now known that a higher order oligomeric state of IcmF – at minimum, a tetramer – is necessary to create the G-protein domain "dimer" interface required for opening IcmF for cofactor delivery or retrieval, changing the available surfaces for ATR to interact with IcmF [18]. Future computational predictions for an IcmF + ATR complex will need to incorporate this consideration before proceeding.

Residues that could play a role in cob(II)alamin off-loading

Despite the lack of structural information for the mutase-G-protein chaperone-ATR complex, there are possible insights into which residues may be involved in the cofactor transfer

pathway. An X-ray crystallographic dataset collected in 2014 by David Born with Marco Jost on an IcmF crystal that was soaked with hydroxocobalamin provided a starting point for investigating the potential intermediary cofactor binding sites. More recently, the dataset has fully been reprocessed and partially refined by a previous group member, Daphne Faber. In this dataset, there existed a number of non-canonical Cbl-binding histidine residues with additional density near them which could fit a Cbl molecule: H104, 765, 812, 872, 993 and 1032 (**Figure 7**). Given this dataset came from a crystal which was soaked in excess hydroxocobalamin, it is possible these extra Cbl densities are an artifact of sample preparation. Investigating any potential role these histidine residues play in cofactor transfer between IcmF and ATR would require a means of measuring cofactor transfer, then perturbing these histidine residues by substitution to observe any changes to transfer efficiency and completion.

Basis for observing cofactor transfer by UV-Vis spectroscopy

The basis for observing cofactor transfer between an AdoCbl-dependent mutase enzyme and ATR by UV-Vis spectroscopy is the change in coordination number of the cobalt in cobalamin. From IcmF to ATR, cob(II)alamin goes from 5-coordinate base-off/His-on to 4-coordinate base-off, where the axial H39 ligand on IcmF is replaced by the non-ligating F110 on ATR [23,29]. This coordination change is reflected in the UV-Vis spectrum by a λ_{max} shift from 474 nm to 464 nm in non-fused systems [31]. An assay for observing this transfer in the human MCM-ATR system by UV-Vis has been described in the literature [34]. The assay works by incubating the mutase with cob(II)alamin, adding a "repair solution" consisting of ATR, ATP, and GTP to induce cofactor transfer, then monitoring the transfer by UV-Vis spectroscopy. Since the presence of oxygen will generate hydroxocobalamin from cob(II)alamin, the samples must be kept under anaerobic conditions. Once ATR is bound with ATP and cob(II)alamin, reduction of cob(II)alamin would normally occur to generate the cob(I)alamin supernucleophile [9]. Subsequent nucleophilic attack on the 5′-carbon of ATP would then produce PPP$_i$ and regenerate AdoCbl [9]. With no

electron source present, the ATR + cob(II)alamin + ATP complex stalls at the unique 4-coordinate base-off state [35,36]. Additionally, IcmF will bind cob(II)alamin passively from solution, which can be verified by centrifuging an IcmF + cob(II)alamin sample through a low molecular weight cut off (MWCO) spin filter and checking the flowthrough for the presence or absence of unbound cofactor. This check is necessary as the 5-coordinate base-off/His-on cob(II)alamin bound to IcmF is indistinguishable by UV-Vis from the 5-coordinate base-off cob(II)alamin in solution, where water acts as the 5th and axial ligand [37]. Removal of the bound cob(II)alamin from IcmF into solution is unfavorable, even with GTP present. ATR + ATP must be present for the removal of cob(II)alamin to become favorable [26]. Back-transfer of cob(II)alamin from ATR to IcmF is disfavored as well [26]. These factors together suggest that structural changes caused by GTP hydrolysis and the presence ATP-bound ATR gate the ejection of cob(II)alamin from IcmF.

Here, we describe the development of a UV-Vis spectroscopy-based assay for observing the off-loading of cob(II)alamin from IcmF to ATR based on the assay used by Banerjee et al [34]. We also describe the design of IcmF variants in which histidine residues have been substituted and report our progress toward purification of the variants. The results of this work reveal the off-loading of Cbl to be a fast process, occurring <1 s and requiring the use of stopped-flow instrumentation to fully capture the transfer. Difficulties in obtaining well-folded and behaved IcmF variants point toward the protein being sensitive to perturbation, with tendency to aggregate.

II. *Results*

Characterizing the loaded and unloaded states using the fused IcmF system

Off-loading of damaged AdoCbl from the fused IcmF enzyme system has not been extensively studied. Thus, we modified an assay used to study off-loading of damaged AdoCbl from MCM to ATR for repair for use with IcmF [34]. This assay takes advantage of the fact that Cbl bound to a mutase is base-off/His-on whereas Cbl bound to ATR is base-off, and that there are

11

different absorption spectra for base-off/His-on Cbl and base-off Cbl (**Scheme 1**). Several control experiments were done to verify if the UV-Vis spectroscopy assay would work similarly for the *C. metallidurans* IcmF-ATR system, since the assay was originally designed for the human MCM-ATR system [34]. Two control spectra, one of IcmF + cob(II)alamin and one of ATR + cob(II)alamin + ATP, were collected to establish the λ_{max} of the 5-coordinate base-off/His-on of cob(II)alamin bound to IcmF and the 4-coordinate base-off cob(II)alamin bound to ATR. The λ_{max} for IcmF + cob(II)alamin was 476 nm and the λ_{max} for the ATR + cob(II)alamin was 465 nm (**Figure 8**). To show that ATP is necessary for cob(II)alamin binding in *Cm*ATR like other PduO-type ATRs [38,39], two additional spectra were collected: one of ATR + cob(II)alamin and of ATR + cob(II)alamin + PPP$_i$. The λ_{max} for both spectra were 476 nm indicating that cob(II)alamin was not 4-coordinate and thus not bound to ATR but was rather it was in solution and 5-coordinate base-off (**Figure 8 and 9**).

Initial attempts at measuring cofactor-offloading fail to capture the full transfer

After determining the λ_{max} for the various states of cobalamin, UV-Vis spectroscopy was employed to monitor the cofactor off-loading from IcmF. The first two trials of the off-loading assay were done using a buffer containing no glycerol, and only collected data every 15 s over 20 m. Plotting the first time point at t = 0 m and the last time point at t = 20 min demonstrates there is no change in the UV-Vis spectra and a constant λ_{max} of 465 nm (**Figures 10 and 11**). Repeating the assay with the same conditions, but now using 5 % glycerol in the buffer to slow cofactor transfer, a new initial λ_{max} = 467 nm was observed, with an end point at λ_{max} of 465 nm after t = 1 h (**Figure 12**). When the glycerol content was increased to 30% v/v, an initial λ_{max} of 469 nm was observed, with an endpoint λ_{max} of 466 nm after t = 1 h (**Figure 13**). The high glycerol content of the buffer in this trial caused significant loss of protein during spin filtering as well as issues mixing the samples in the cuvette, so in the next assay attempt, the glycerol content was dropped back down to 5% v/v. As an alternative means to slow cofactor transfer, this next assay attempted was

performed at a lower temperature of 4 °C instead of 25 °C. Different ATR concentrations of 10 µM, 8 µM, and 6 µM were also utilized (**Figures 14, 15, and 16**). In these three experiments, the initial λ_{max} was 466 nm and the λ_{max} at t = 2 h was 465 nm. To address issues with mixing the samples at higher glycerol content, the stir bar functionality of the spectrophotometer was used for additional experiments, which included ATR concentrations of 10 µM and 6 µM with 10% v/v glycerol and were run at 4 °C. In the spectra for these experiments, the λ_{max} was 467 nm and the λ_{max} at t = 2 h was 465 nm (**Figures 17 and 18**). In all of these attempts at the assay, we were unable to see the expected initial λ_{max} of 476 nm corresponding to cob(II)alamin bound to IcmF, i.e. the transfer was too fast to capture the starting state. Our attempts to slow down the cofactor transfer by using glycerol and/or lower temperatures and/or low protein concentrations were insufficient.

Anaerobic stopped-flow UV-Vis spectroscopy can be used to measure cofactor off-loading

Since our attempts to slow down the cofactor transfer were insufficient, we sought to collect data at significantly earlier time points and turned to anaerobic stopped-flow UV-Vis spectroscopy. After replicating the IcmF + cob(II)alamin control and observing the λ_{max} of 476 nm (**Figure 19**) consistent with spectrum previously observed (**Figure 8**), the full off-loading assay was performed using a UV-Vis spectrophotometer attached to an anaerobic stopped-flow setup (see Methods for instrumentation details). The UV-Vis data were collected on a logarithmic timescale and at 25 °C, with most collection points at <1 s and continuing to 250 s. In these experiments, the full λ_{max} shift from 476 to 465 nm was observed (**Figures 20 and 21**), indicating that we were able to capture the full cofactor off-loading process starting with cob(II)alamin on IcmF and ending on ATR. Plotting the change in absorbance at 465 nm over time, corresponding to cob(II)alamin bound to ATR, we can see that there is a significant amount of cob(II)alamin being transferred <1 s, far faster than what could be measured in previous attempts at the assay (**Figures 22 and 23**). Using this change in absorbance at 465 nm, along with $\Delta\varepsilon_{465nm}$ = 5,900 M^{-1}

cm^{-1} and a path length of 1 cm, the estimated percent of cob(II)alamin transferred to ATR was 11 and 25% for these assays, for Experiments 10 and 11, respectively [34]. These low values for percent transfer could be an indication that the concentration of GTP used in these experiments was too low. Previous studies have shown a strong dependence of GTP concentration on cob(II)alamin release from IcmF [26].

Isobutyryl-CoA mutase fused variant design and production fails to produce suitable full-length protein

To begin to probe the potential role of the histidine residues with extra density observed in the hydroxycobalamin-soaked IcmF structure, variants were produced (**Figure 7**). Using standard site-directed mutagenesis techniques, first a construct for an H872A variant was produced. During the protein purification, several irregularities compared to the purification of the wild-type IcmF were noted. When running the anion exchange (AEX) column, there was an A$_{280}$ protein peak at the start with no percent high salt elution buffer, as well as at the conductivity expected from wild-type IcmF (**Figure 24**). Pooling fractions from the peak at the expected conductivity the protein seemed to behave normally for the size exclusion chromatography (SEC) column (**Figure 25**). The yield of 3.72 mg/L was comparable to the ~4 mg/L often obtained with the wild-type protein (**Figure 26**). However, when analyzing the quality of the protein form the purification with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), the overloaded band at ~130 kDa present in the samples from immobilized metal affinity chromatography (IMAC) and AEX, which was assumed to correspond to the 123 kDa IcmF, was absent in the lane for the SEC sample (**Figure 27**).

Two subsequent attempts at purifying more H872A variant protein from the same growth as the initial attempt yielded intact protein, based on the SDS-PAGE gels (**Figures 28 and 29**). However, both purifications at the AEX step showed the same irregularity where the variant

14

protein eluted off significantly earlier at lower conductivities than wild-type IcmF. Additionally, more conservative H872D and H872S variants were produced and an H872C variant, where the cysteine could ligate the cob(II)alamin, was also produced. A plasmid construct for an H993A variant was also produced, but purification has not yet been attempted.

III.   *Discussion*

In this work, we sought to develop an assay capable of measuring the off-loading of damaged Cbl from IcmF, a fused mutase system. The control spectra of IcmF + cob(II)alamin and ATR + cob(II)alamin and ATP established λ_{max} values of 476 and 465 nm for cob(II)alamin bound to IcmF and ATR, respectively. These values are similar to those of 474 and 464 nm recorded for MCM and human ATR, respectively [34]. The control spectra of ATR + cob(II)alamin and ATR + cob(II)alamin + PPP$_i$ show that *Cm*ATR behaves similarly to other PduO-type ATRs in agreement with published data for CmATR [26]. ATRs from humans and *Lactobacillus reuteri* both highly favor ATP to be bound before cob(II)alamin can be bound [38,39]. PPP$_i$ has previously been shown to be insufficient to promote cob(II)alamin binding and generate the unique 4-coordinate ATR + cob(II)alamin state which precedes reduction to cob(I)alamin and nucleophilic attack [39]. Since both samples lacked ATP, *Cm*ATR should be unable to bind cob(II)alamin. From the λ_{max} of 476 nm of both spectra, cob(II)alamin is indeed unbound and 5-coordinate base-off in solution.

The initial attempts at applying the assay described in Mascarenhas *et. al.* 2022 for the *C. metallidurans* system failed to recapitulate the expected λ_{max} shift from 476 to 465 nm [34]. After we noticed a change in initial λ_{max} when glycerol was included in the reaction buffer (**Figure 12**), we considered if the two-part *C. metallidurans* system could be facilitating cofactor off-loading faster than the three-part human system. Further modifications to the assay were tested to investigate potential ways of slowing down the transfer enough to observe the expected initial λ_{max} corresponding to cob(II)alamin bound to IcmF. Although the increase to 30% v/v glycerol showed a promising λ_{max} of 469 nm (**Figure 13**), closer to the expected λ_{max} of 476 nm, the increase in

time between repair solution injection to spectra collection resulting from longer manual mixing times likely resulted in loss of data in the early time points of the reaction. The large loss of protein from spin filtering with a high glycerol buffer was a significant hinderance as well. Returning to 5% v/v glycerol but lowering the reaction temperature to 4 °C did not show any appreciable change in λ_{max}, each only recording a λ_{max} shift of 466 to 465 nm. Additionally, lowering the ATR concentration to see if this slowed the transfer down did not have a noticeable effect (**Figures 14, 15, and 16**). A stir bar was introduced with the hope that, without a need for manual mixing, earlier time points in the transfer could be observed. Introduction of the stir bar appeared to only have the effect of causing fluctuations in the baseline of the UV-Vis spectra (**Figures 17 and 18**). Furthermore, the increase in glycerol content to 10% v/v only brought the initial λ_{max} to 467 nm.

If cofactor off-loading was indeed a fast process, the set-up and procedure used to perform these assays using a standard UV-Vis spectrophotometer was likely too slow to fully capture the transition. This complication led towards moving the assay to stopped-flow instrumentation. A stopped-flow instrument in an anaerobic chamber, with UV-Vis spectrometer capabilities, enabled rapid spectra collection following mixing of IcmF + cob(II)alamin with the repair solution. Compared to manual mixing and spectra collection, this instrumentation enabled time points far earlier than ~5 s to be collected. Indeed, the majority of the observed cofactor transfer occurred in less than 1 s, indicating a necessity for using the stop-flow apparatus (**Figures 22 and 23**). From the stopped flow data, the estimated transfer efficiencies of ~11% and 25% are comparable to the ~20% value recorded in the literature when similar amounts of GTP is used for the assay [26].

When plotting the change in absorbance at 465 nm over time, the reaction initially gives the appearance of a biphasic process composed of an initial fast process <1 s, a slower process, followed by a much slower process >10 s (**Figures 22 and 23**). However, the presence of an intermediate state is not supported by the existence of isosbestic points at 321 and 371 nm. The

16

biphasic nature of the data displayed in **Figures 22 and 23** could be explained by structural heterogeneity within the sample. IcmF requires the formation of supramolecular structures for Cbl loading and unloading [18]. Thus, heterogeneity in IcmF samples is well established [18,23].

With a functioning assay in hand to observe cofactor off-loading activity, we proceeded to investigate the histidine residues identified from the partially refined structure of IcmF with extra Cbl density are involved in cofactor transfer. H872 was chosen as the first target as it was the only residue of interest which appeared to bind cob(II)alamin in a bis-histidine fashion with H39, the canonical binding histidine [23]. However, this variant proved difficult to purify following the protocol for wildtype IcmF. For such a large protein, it seems unlikely that the single mutation of H872A could so drastically change the charge of the protein. Since ion exchange chromatography generally separates species based on charge, it was unexpected that this residue change resulted in such a drastic change in elution behavior on the AEX column. Furthermore, at pH 7.5, histidine is neutrally charged. Therefore, a mutation to alanine should not affect protein surface charge. The mutations to aspartic acid and serine would create a change in charge, while preserving some of capacity for coordination of cofactor. For all the variants purified in this work, a similar issue as to what was observed with H872A at the AEX step was encountered. Since IcmF is known to oligomerize for cofactor loading [18], unloading, this enzyme may be extra sensitive to residue substitutions.

To examine the potential that the histidine residues are a part of the cofactor transfer pathway, the sequences of IcmF from several different organisms were compared (**Figure 30**). In addition to the canonical Cbl-binding histidine in the active site, H39 in *Cm*IcmF, the residues H812, H872, and H993 are also conserved across other organisms. However, the IcmFs from the chosen organisms were already highly similar in sequence, so it is unclear if these histidine residues have mechanistic significance. The sequence for *Cm*IcmF was then aligned with the MCMs of several different organisms (**Figure 31**). The canonical Cbl-binding histidine was

conserved, but the histidine residues identified as potential Cbl interacting residues were not. There are, however, other histidine residues which are conserved. It is possible that, if there are residues that help facilitate cofactor transfer, they may differ in relative location in the MCMs. It is also possible that the structure of the ATR-MCM-G-protein chaperone complex places the ATR and MCM active sites near enough to each other to enable direct cofactor transfer. Further work, including more information regarding the ATR-MCM-G-protein chaperone complex as well as the ongoing mutational work and assay described here, is necessary to determine if the absence of these histidine residues in MCMs indicates an alternative mechanism or not.

What remains is to generate usable IcmF histidine variants with which to perform the assay. There are several other histidine residues which are of potential interest that were not targeted in the mutational work described here: H104, H765, H812, and H1032. More conservative mutations that still preserve some metal ligation properties could be attempted with H993 as well. As this protein is needed for a cofactor-transfer assay, the purity of the sample need not necessarily be high as any contaminating proteins are unlikely to be able to bind Cbl. Taking this viewpoint, the purification could be simplified down to just the first column for the immobilized metal affinity chromatography (IMAC) step. It will also be important to confirm that residues substitutions are not changing the overall protein structure. CD could be used for verification. If His to Ala substitutions are problematic in terms of structural changes, more conservative mutations can be prepared.

Ongoing structural efforts, not described in this work, also provide avenues for investigating the cofactor off-loading mechanism for IcmF. Previous work which captured the supramolecular structure necessary to create the G-protein domain active state "dimer" using cryogenic-electron microscopy (cryo-EM) along with the nonhydrolyzable GTP analog guanosine-5'-[(β,γ)-methyleno]triphosphate (GMPPCP) [18]. We could consider use of GMPPCP to stabilize IcmF in an open state in order to capture a IcmF + cob(II)alamin + ATR off-loading complex. The

same strategy could be applied to capture the IcmF +AdoCbl + ATR loading complex. Additional experiments with a cofactor loading assay, not described in this work, show that AdoCbl loading onto IcmF by ATR proceeds on the order of minutes, rather than seconds. This slower loading may make isolation of a loading complex a more feasible task by Cryo-EM. Higher resolution Cryo-EM structures of the IcmF supramolecular structure obtained using GMPPCP will additionally enable more detailed comparison of the G-protein domain "dimer" interface with that of standalone G-protein chaperones found in humans and other bacterial systems. Additional work to obtain the structure of *Cm*ATR bound to ATP and cob(II)alamin would allow for additional comparisons to other PduO-type ATRs and provide an additional snapshot for the cofactor off-loading mechanism.

The two-protein system from *C. metallidurans* is an attractive model to study the maturation of AdoCbl-dependent enzymes. Understanding how the IcmF-ATR system is able to recognize damaged cofactor and unload it across long distances will illuminate a key mechanism for such a biologically important set of enzymes. Understanding these metalloenzymes and how they function and interact with their cofactors will greatly inform our understanding of how they can accomplish such challenging chemistry.

Materials

E. coli Bl21 T7 Express and *E. coli* DH5α chemically competent cells were purchased from New England Biolabs (NEB). The Q5 Site-Directed Mutagenesis and Monarch plasmid miniprep kits were purchased from NEB. Luria Broth (LB) medium components were purchased from Fisher BioReagents. Ampicillin, kanamycin, and isopropyl-b-D-thiogalactopyranoside (IPTG) were purchased from GoldBio. NaCl was purchased from Fisher Chemical. The EDTA-free protease inhibitor cocktail tablets and GTP were purchased from Roche. The Ni-NTA 1 mL HisTrap HP and 5mL HisTrap FF columns, and Superdex200 16/60 size exclusion column (SEC) were purchased from GE Healthcare. The MonoQ and the Capto HiRes Q anion exchange columns (AEX) were

purchased from Cytiva. The airtight quartz cuvettes were purchased from Starna Cells. The airtight syringes were purchased from Hamilton. The gel filtration standards, polyacrylamide gels, sodium dodecyl sulfate solution and were purchased from BioRad. All other materials were obtained from Sigma-Aldrich.

Methods

Cloning

The pET28a plasmid containing the *Cupriavidus metallidurans* gene *icmF* was used as the template for site-directed mutagenesis (**Table 1**) [23]. Using the primers listed below (**Table 2**), variants of IcmF were generated using the Q5 Site-Directed Mutagenesis kit. *E. coli* DH5α competent cells were transformed by heat shock, and plasmids were extracted from 5 mL cultures using the Monarch plasmid miniprep kit. Whole sequences of the purified plasmids (**Table 1**) were verified by nanopore sequencing performed by Plasmidsaurus.

Isobutyryl-CoA mutase fused: protein expression and purification

The protocols for protein expression and purification for wild-type IcmF and IcmF variants are described below. *E. coli* BL21 T7 Express chemically competent cells were transformed with the appropriate plasmid by heat shock, and 100 mL of LB with 50 µg/mL kanamycin was inoculated from a single colony of transformed cells. The culture was grown overnight at 37 °C in an orbital incubator shaker at 220 RPM. The overnight culture was split between 6 flasks of 1L of LB fortified with 50 µg/mL kanamycin and held in an orbital incubator shaker at 37 °C and 220 RPM. At OD_{600} = 0.5-0.6, the flasks were held at 4 °C until they reached around 16 °C and were then placed in an orbital shaker prechilled to 16 °C and set to 220 rpm. After 2 h, expression was induced with a final IPTG concentration of 0.1 mM, and following induction, the culture was incubated at 16 °C and 220 RPM for 10 to 12 h. Cells were then harvested by centrifugation at 5,000g for 25 min at 4 °C, and flash frozen in liquid nitrogen (LN_2) for long-term storage at -80 °C.

For purification, a cell pellet from 2L of culture was resuspended in 50 mL of lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine-HCl, 1 tablet of EDTA-free protease inhibitor (Roche), and 5 µL benzonase. Cells were lysed by sonification, and then clarified by centrifugation at 28,000g for 30 min at 4 °C. The lysate was passed through a 0.2 µm filter before being loaded onto a 5 mL Ni-NTA HisTrap FF column equilibrated with lysis buffer attached to an FPLC (BioRad NGC System). The column was washed with 10 CVs of buffer (50 mM HEPES pH 7.5 and 500 mM NaCl) containing 20 mM, then 40 mM imidazole, and then eluted into fractions at 210 mM imidazole. Fractions containing protein were pooled, buffer exchanged into 50 mM HEPES pH 7.5, and concentrated using a 50 MWCO centrifugal filter. The concentrated fractions were then filtered using a 0.2 µm filter and loaded onto either a MonoQ or Capto HiRes Q AEX column equilibrated with 50 mM HEPES pH 7.5 and 5 mM NaCl attached to an FPLC. The protein was then eluted using a linear gradient of 10-80% of 50 mM HEPES pH 7.5 and 500 mM NaCl. Elution occurs sharply at a conductivity of around 41.3 cS/cm, corresponding to 250-300 mM NaCl. A second peak immediately followed the first, containing aggregated and truncated protein. Fractions from the initial peak were concentrated with a 50 MWCO centrifugal filter and then passed through a 0.2 µm filter. The concentrated protein was then loaded onto a Superdex S200 16/60 SEC column equilibrated with 20 mM HEPES pH 7.5 and 50 mM NaCl. Eluted fractions containing protein were concentrated with a 50 MWCO centrifugal filter and then flash frozen in LN_2 for long term storage at -80 °C. Purity of the protein was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), using a 4-20% w/v gel. Concentration of the protein was determined by measuring the absorbance at 280 nm using a NanoDrop UV-Vis spectrophotometer (Thermo Scientific) and molecular weight (MW = 123 kDa) and extinction coefficient (ε = 84,600 M^{-1} cm^{-1}) calculated by the ProtParam tool [40].

Adenosyltransferase: protein expression and purification

E. coli BL21 T7 express chemically competent cells were transformed by heat shock with a pMCSG7 plasmid containing the *C. metallidurans* gene *cobO* [41]. A culture of 25 mL of LB supplemented with 100 mg/µL of ampicillin was inoculated with a single colony of transformed cells and grown at 37 °C and 220 RPM overnight. The overnight culture was then split between 6 flasks of 1 L LB media fortified with 100 mg/µL of ampicillin. The flasks were kept at 37 °C and 220 RPM, until an OD_{600} = 0.6 was reached. Flasks were cooled to ~15 °C and then placed in a prechilled incubator shaker at 15 °C and 220 RPM. Expression was induced with a final concentration of 0.1 mM IPTG. Cells were harvested after 18 h by centrifugation at 5,000g for 25 min at 4 °C, then flash frozen with LN_2 for long-term storage at -80 °C.

For purification, a cell pellet from 2 L of culture was resuspended in lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole) supplemented with 1mM PMSF and 1/2 tablet of EDTA-free protease inhibitor. Cells were lysed by sonication, then clarified by centrifugation at 28,000g for 30 min at 4 °C. The lysate was passed through a 0.2 µm filter, then loaded onto a 1 mL Ni-NTA HisTrap HP attached to an FPLC and equilibrated with lysis buffer. Protein was eluted using a gradient of elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole), and fractions from the peaks centered around 60- and 200-mM imidazole were pooled and concentrated with a 30 MWCO centrifugal filter. The concentrated protein was then passed through a 0.2 µm filter and loaded onto a Superdex S200 16/60 SEC column equilibrated with 50 mM HEPES pH 7.5 and 500 mM NaCl. Eluted fractions containing protein were concentrated with a 30 MWCO centrifugal filter and then flash frozen in LN_2 for long term storage at -80 °C. Purity of the protein was assessed by SDS-PAGE using a 4-20% w/v gel. Concentration of the protein was determined by measuring the absorbance at 280 nm using a NanoDrop UV-Vis spectrophotometer and extinction coefficient (ε = 15,595 M^{-1} cm^{-1}) and molecular weight (MW = 19.9 kDa) calculated by the ProtParam tool [40].

UV-Vis off-loading assay

Buffer A (20 mM HEPES pH 7.5, 50 mM NaCl, 2 mM MgCl$_2$, and 2 mM tris(2-carboxyethyl)phosphine (TCEP)) was sparged with argon gas to remove oxygen then transferred to a N$_2$ anaerobic chamber in an airtight vial. For some experiments, Buffer A contained 5-30% v/v glycerol, specific amount is noted by assay below (**Table 3**). Pure cob(II)alamin was prepared by dissolving AdoCbl in Buffer A and then illuminating the resulting solution with a 23 W lamp for 3 hours under an N$_2$ atmosphere. Purity of the generated cob(II)alamin solution was checked by collecting a UV-Vis spectra. AdoCbl in solution will have a broad peak with a λ_{max} of 525 nm whereas cob(II)alamin in solution will have this peak shifted to a λ_{max} of 476 nm [34]. ATR and IcmF samples were filtered through a Costar 0.2 µm filter after thawing on ice. Two solutions were then prepared under an N$_2$ atmosphere using the sparged Buffer A: enzyme solution containing IcmF and cob(II)alamin in Buffer A and the repair solution containing ATR, ATP, and GTP in Buffer A. See table below for solution compositions (**Table 3**). The enzyme solution was allowed to incubate in an airtight quartz cuvette with a rubber septum (Starna Cells) for 15 minutes. The repair solution was then transferred to an airtight syringe (Hamilton), and both solutions were brought out of the anaerobic chamber to a Cary UV-Vis spectrophotometer (Agilent). The repair solution was injected into the cuvette containing the enzyme solution, mixed by inversion (or by stir-bar for Assays 8 + 9), and spectra from 300-800 nm were collected every 15 s over 20 min – 2 h, with specific times noted below (**Figures 10-18**). For the experiments conducted at 4 °C, samples and reagents were kept cold by metal cold beads, and the temperature of the Cary instrument was set to 4 °C. The control off-loading experiment was performed by injecting buffer A instead of the repair solution. A control to check Cbl(II) binding to IcmF was done by incubating Cbl(II) with IcmF for 15 min, then taking the solution and centrifuge filtering with a low MWCO spin filter. The concentrated solution was a deep reddish orange, and the flowthrough was clear, indicating that the majority of Cbl(II) was bound to IcmF. Additional control spectra of ATR + cob(II)alamin + ATP, ATR + cob(II)alamin, and ATR + cob(II)alamin + PPP$_i$ were collected.

Stopped-flow experiments

A SX20 stopped-flow spectrometer (Applied Photophysics) equipped with a photodiode array detector and photomultiplier detector was used for these experiments. Buffer A (20 mM HEPES pH 7.5, 50 mM NaCl, 2 mM MgCl$_2$, 2 mM TCEP, and 5% v/v glycerol) was sparged with argon to remove oxygen then transferred to a N$_2$ anaerobic chamber in an airtight vial. Cob(II)alamin was prepared by dissolving AdoCbl in Buffer A and then illuminating with a 23 W lamp for 3 hours under a N$_2$ atmosphere. Purity of the generated cob(II)alamin solution was checked by collecting a UV-Vis spectra. AdoCbl in solution will have a broad peak with a λ_{max} of 525 nm where cob(II)alamin in solution will have this peak shifted to a λ_{max} of 476 nm [34]. The enzyme solution (IcmF, cob(II)alamin, and Buffer A) and the repair solution (ATR, ATP, GTP, and Buffer A) were prepared under an N$_2$ atmosphere using the sparged Buffer A (**Table 4**). Both solutions were allowed to incubate for 15 min in the injection syringes. Each shot from the syringes was set to 180 µL, for a total reaction volume of 360 µL. For the control, syringe 2 instead contained only Buffer A. Data were collected logarithmically for 250 s, scanning through 0-800 nm, at 25 °C.

Tables

Table 1: Plasmid constructs used in this work.

Name	Notable Features	Source
pET28a_*Cm*IcmF	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	Ref.[23]
pMCSG7_*Cm*ATR	N-terminal 6x His-tag, TEV cleavage site, AMP^R, T7 promoter	Ref.[41]
pET28a_*Cm*IcmF_H872A	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	This study
pET28a_*Cm*IcmF_H872D	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	This study
pET28a_*Cm*IcmF_H872S	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	This study
pET28a_*Cm*IcmF_H872C	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	This study
pET28a_*Cm*IcmF_H993A	N-terminal 6x His-tag, thrombin cleavage site, KAN^R, T7 promoter	This study

Table 2: Primers used for site-directed mutagenesis.

Name	Sequence (5' to 3')
IcmF_H872A_F	CCGGTCACTGGCCGCGCAGGAGA
IcmF_H872A_R	CCCGATGTCTGGATGTGG
IcmF_H872D_F	GTCACTGGATGCGCAGG
IcmF_H872D_R	CGGCCCGATGTCTGGATG
IcmF_H872S_F	GTCACTGAGCGCGCAGG
IcmF_H872S_R	CGGCCCGATGTCTGGATG
IcmF_H872C_F	GTCACTGTGCGCGCAGG
IcmF_H872C_R	CGGCCCGATGTCTGGATG
IcmF_H993A_F	ACAGCTCAAGGCCGACGGCACCT
IcmF_H993A_R	TCGTAGTACAGCGATTCC

Table 3: Reaction conditions for cofactor transfer controls and assays.

Reagent	Control 1	Control 2	Control 3	Control 4	Assay 1	Assay 2	Assay 3	Assay 4	Assay 5	Assay 6	Assay 7	Assay 8	Assay 9
IcmF (µM)	6	0	0	0	6	6	6	6	6	6	6	6	6
Cob(II)-alamin (µM)	10	10	10	10	10	10	10	10	10	10	10	10	10
ATR (µM)	0	10	10	10	10	10	10	10	10	8	6	10	6
GTP (µM)	0	0	0	0	67	67	67	67	67	67	67	67	67
ATP (µM)	0	333	0	0	333	333	333	333	333	333	333	333	333
PPP_i (µM)	0	0	0	333	0	0	0	0	0	0	0	0	0
Glycerol (%v/v)	0	0	0	0	0	0	5	30	5	5	5	10	10
Temperature (°C)	25	25	25	25	25	25	25	25	4	4	4	4	4
Total volume (µL)	300	300	300	300	300	300	400	400	385	400	400	600	600
Figure	**8**	**8**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**

Table 4: Reaction compositions for the off-loading assay using stopped-flow instrumentation.

Reagent	Syringe 1 Conc (µM)	Syringe 2 Conc (µM)	Final Rxn Conc (µM)
Cob(II)alamin	30	0	15
IcmF	18	0	9
ATR	0	60	30
GTP	0	200	100
ATP	0	600	300

Figures and Schemes
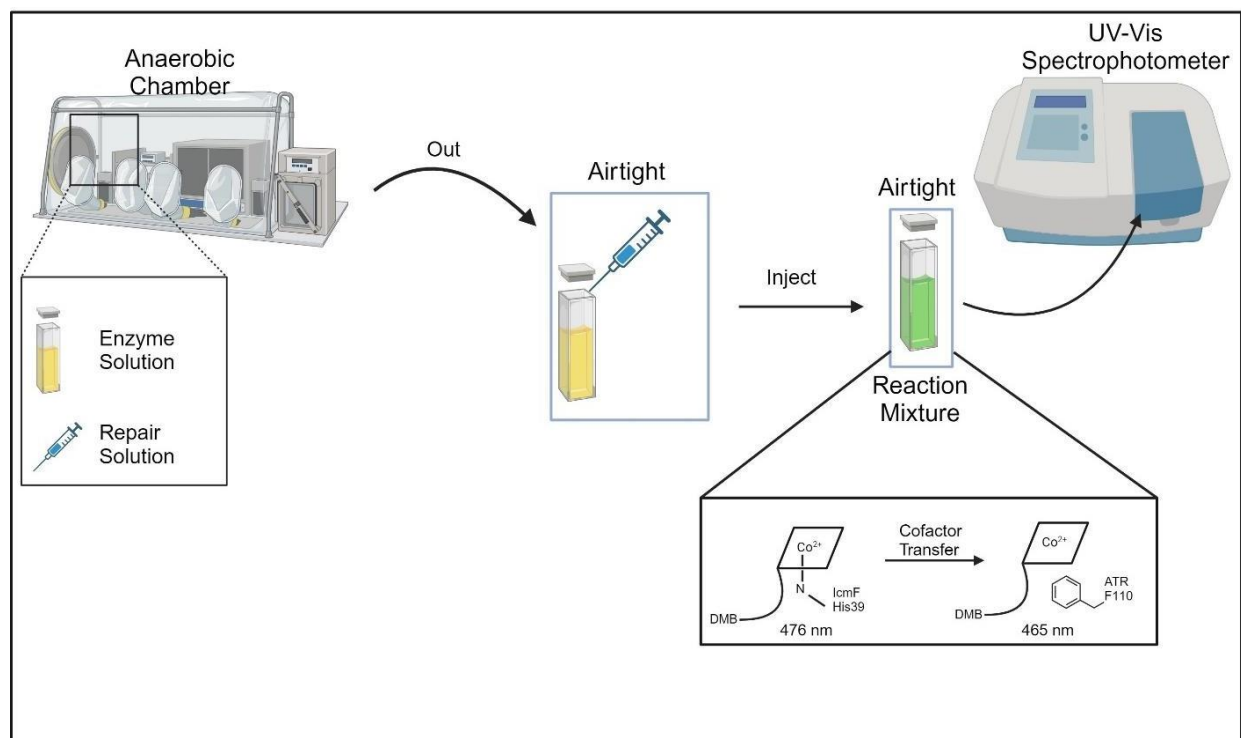
Scheme 1: Overview of the off-loading cofactor transfer assay. As described in the Methods section, the Enzyme solution contains IcmF and cob(II)alamin, and the Repair solution contains ATR, ATP, and GTP. Both are prepared in an anaerobic chamber, with the enzyme solution kept in an airtight cuvette with a rubber septum, and the repair solution kept in an airtight syringe. Both are brought out of the chamber, and the Repair solution is injected into the cuvette through the septum. The reaction mixture is mixed by inversion, and then the cuvette placed in a UV-Vis spectrophotometer for data collection. The cob(II)alamin is 5-coordinate base-off/His-on on IcmF, where H39 is the lower axial ligand. As it is transferred to ATR, F110 blocks ligation by solvent, leaving cob(II)alamin only 4-coordinate base-off. Figure created using BioRender.
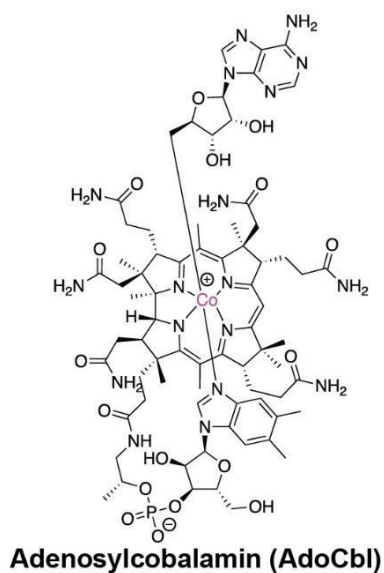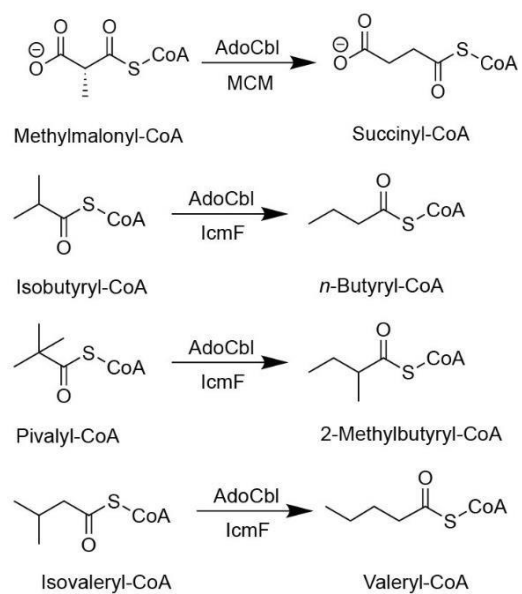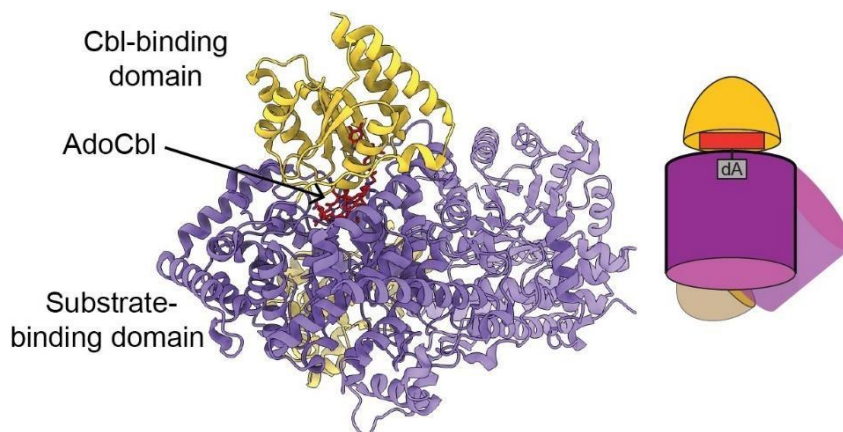
Figure 1: Adenosylcobalamin and catalyzed reactions of AdoCbl-dependent enzymes. A.
The structure of the adenosylcobalamin coenzyme. **B.** The rearrangements catalyzed by MCM
and IcmF.

A.

Methylmalonyl-CoA mutase
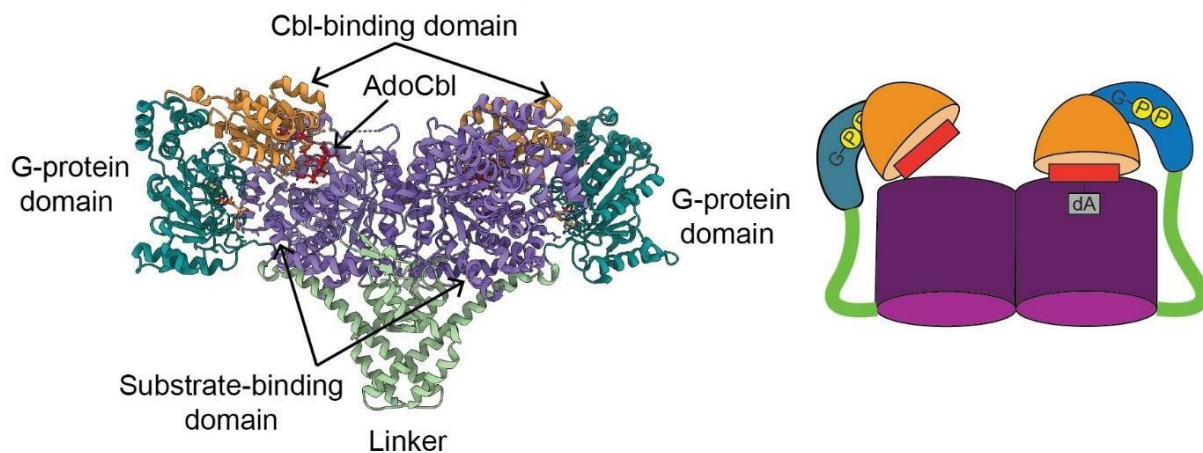## (MCM)

B.

Isobutyryl-CoA Mutase Fused
(IcmF)

Figure 2: Structures and domain architecture of MCM and IcmF. A. The structure and simplified domain architecture of MCM bound to AdoCbl from *P. shermanii* (PDB: 4REQ), consisting of the Cbl-binding Rossmann domain (yellow ribbons), the AdoCbl (red sticks), and the substrate-binding TIM barrel domain (purple ribbons) [42]. The MCM dimer from *P. shermanii*, as with other bacterial MCMs, is composed of a large mutase subunit and a small mutase subunit. The MCM dimer from humans is a homodimer [43]. **B.** The structure and simplified domain architecture of IcmF bound to GDP and AdoCbl from *C. metallidurans* (PDB: 4XC6), consisting of

the Cbl-binding Rossmann domain (yellow ribbons), the AdoCbl (red sticks), the substrate-binding

TIM barrel domain (purple ribbons), the G-protein domain (teal ribbons), the linker (green ribbons),

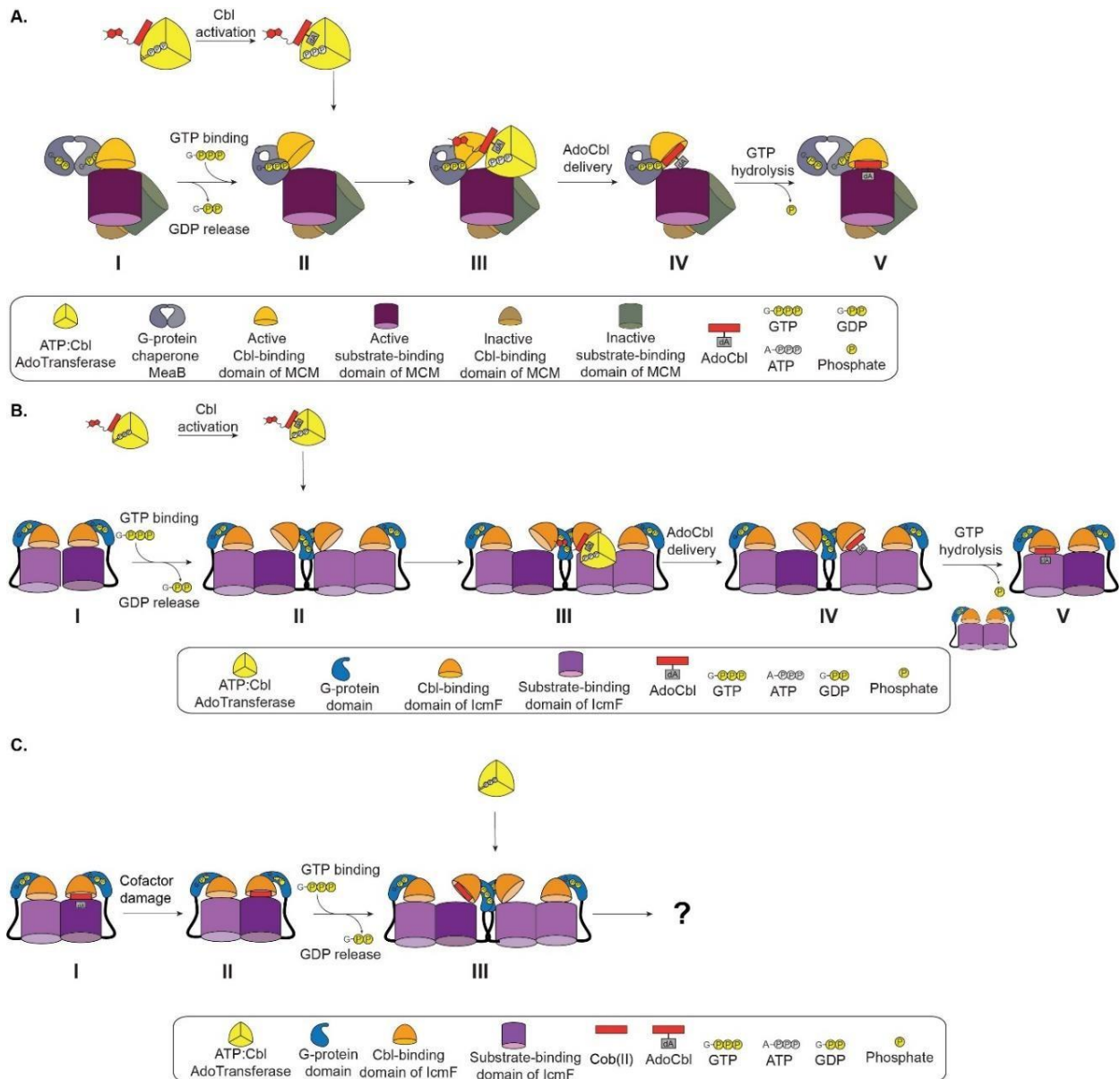and GDP (sticks colored by heteroatom) [23].

Figure 3: Cofactor transfer mechanisms of non-fused and fused AdoCbl-dependent enzyme systems. A. Maturation of the non-fused MCM system from *M. extorquens*. GTP binding to the G-protein chaperone causes conformational changes in the G-protein chaperone dimer, which pry open the mutase domains of the active subunit of the MCM heterodimer. ATR, loaded with AdoCbl, is then able to deliver the cofactor to the open active site of MCM. GTP hydrolysis then facilitates another structural shift which clamps shut the mutase, locking in the cofactor. **B.** Maturation of IcmF. GTP binding leads to the formation of supramolecular structures of IcmF. With

the tetrameric structure, the G-protein domain can form a dimeric interface, similar to that of standalone G-protein chaperones, allowing for the opening of the mutase domains. ATR is then able to deliver AdoCbl to the open active site. GTP hydrolysis causes dissociation of the higher order oligomeric structure, separating the G-protein domains and clamping shut the mutase. **C.** A partial mechanism for cofactor repair in the IcmF system. The G-protein domain binds GTP and dimerizes, forming the supramolecular structure of IcmF competent for cofactor removal. ATR binds ATP, which would allow it to bind the damaged cofactor, cob(II)alamin, and re-adenylate it following reduction to cob(I)alamin. The details of cofactor ejection and handoff to ATR are unknown, as well as what factors allow IcmF and ATR to sense the presence of damaged cofactor.

```
H.sapiens         MAVCGLGSRLGLGSRLGLRGCFGAARLLYPRFQSRGPQGVEDGDRPQPSSKTPRIPKIYT    60
C.metallidurans   ------------------------------------------------MGNRLSKIAT    10
B.thailandensis   ------------------------------------------------MGNRLSKIAT    10
M.extorquens      -------------------------------------------------MVKLNRIYT     9
M.tuberculosis    ------------------------------------------------MAVHLTRIYT    10
P.shermanii       -------------------------------------------------MVNITRVYT     9
                                                                   .: :: *

                  ATP binding
H.sapiens         KTGDKGFSSTFTGERRPKDDQVFEAVGTTDELSSAIGFALELVTEKGHTFAEELQKIQCT   120
C.metallidurans   RTGDAGTTGLGDGSRVGKNSLRIVAIGDVDELNSHIGLLLTEPD-LPEDVRAALLHIQHD    69
B.thailandensis   RTGDDGTTGLGDGSRVRKDDARIAAIGDVDELNSQIGVLLAEP--LPDDVRAALSAIQHD    68
M.extorquens      RTGDQGTTGLANGERRSKADLRVEAYGTVDETNACIGLARLTAE-P--ALDAMLARIQND    66
M.tuberculosis    RTGDDGTTGLSDMSRVAKTDARLVAYADCDEANAAIGAALALGH-PDTQITDVLRQIQND    69
P.shermanii       RTGDAGTTRLSNNEVAPKTDPRVQAYGQVDETNCTIGVALTLDP--SDDMQKVLAIVQNE    67
                  :*** * :        .    * .  * .  ** .. **          .   * :*

                                                            Hydrophobic pocket
H.sapiens         LQDVGSALATPCSSAREAHLKYTTFKAGPILELEQWIDKYTSQLPPLTAFILPSGGKISS   180
C.metallidurans   LFDLGGELSIPGYT---------LLKAPQVAQLDDWLAHYNAALPRLAEFILPGGSRPAA   120
B.thailandensis   LFDLGGELCIPGHA---------AITDAHLARLDGWLAHYNGQLPPLEEFILPGGARGAA   119
M.extorquens      LFDLGADLATPPSDKPLGY-EPLRIVPAQVQRLETEIDALNANIPPLKSFVLPGGSAAAA   125
M.tuberculosis    LFDAGADLSTPIVENP-KH--PPLRIAQSYIDRLEGWCDAYNAGLPALKSFVLPGGSPLSA   127
P.shermanii       LFDVGADLSSPVVSDP-KF--RPVRVDQTSVDRLEKWIDEFGEDLPALRSFILPGGSPLAA   125
                  * * *. *. *           .   :  .*:        :* *  *:**.*.  ::

                      ATP binding                           ATP binding
H.sapiens         ALHFCRAVCRRAERRVVPLVQ-----------MGETDANVAKFLNRLSDYLFTLARYAAM   229
C.metallidurans   QAHICRTVCRRAERALVELGA-----------AEALNEAPRQYLNRLSDLLFVLARVLNR   169
B.thailandensis   LAHVCRTVCRRAERSIVALGA-----------SEPLNAAPRRYVNRLSDLLFVLARVLNR   168
M.extorquens      ALHLARTVCRRAERLVVALSGVE---------SEAISGEALQYLNRLSDFLFVASRAANR   176
M.tuberculosis    LLHVARTVVRRAERSAWAAVDAH--------PEGVSVLPAKYLNRLSDLLFILSRVANP   178
P.shermanii       QLHVARSTCRRAERAAWEAVEAFGGEDGSSEPKGGVSLIAVKYLNRLSDLLFNLSRHANY   185
                  *..*:. *****                .   :::*****  ** :*

H.sapiens         KEGNQEKIYMKNDPSAESEGL   250
C.metallidurans   AGGGSDVLWQRERES------   184
B.thailandensis   AAGGADVLWDRTRAH------   183
M.extorquens      DG-ADDVLWVPGQNR------   190
M.tuberculosis    ---DGDVLWRPGGDRTAS---   193
P.shermanii       EAEHDEVLWVPDGEREV----   202
                     : ::
```

Figure 4: Sequence alignments of ATRs from selected organisms. The sequence alignment of ATRs from *H. sapiens, C. metallidurans, B thailandensis, M. extorquens, M. tuberculosis,* and *P. shermanii* indicates several sequences of highly conserved residues. Conserved residues known to participate in binding ATP, PPP$_I$, and/or adenosine are highlighted in orange. The conserved residues known to form a hydrophobic pocket in the active site of ATR are highlighted in red. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions of conservation with strongly similar properties. Periods (.) denote positions of conservation with weakly similar properties. Alignments performed using ClustalW [30]. Figure from the thesis of Francesca Vaccaro [29].
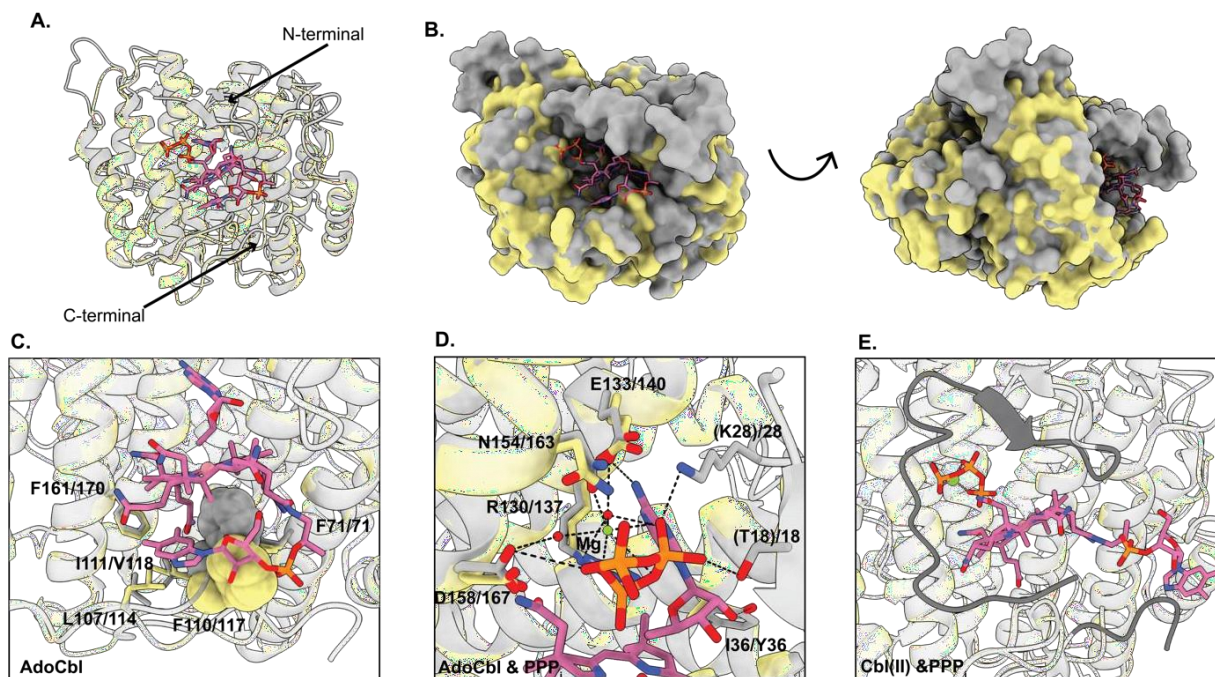
Figure 5: Structural alignments of apo-*Cm*ATR with *Mtb*ATR bound to AdoCbl and PPP$_i$ A.
Alignment of the apo-*Cm*ATR structure (yellow ribbons) with the *Mtb*ATR structure (PDB: 6WGV) (grey ribbons), with AdoCbl and PPP$_i$ bound to *Mtb*ATR (carbons shown in red sticks and sticks colored by heteroatom, respectively) [15,31]. **B.** Alignment of the apo-*Cm*ATR structure (yellow ribbons) with the *Mtb*ATR structure (grey ribbons), with AdoCbl bound to *Mtb*ATR (red sticks). The active site for ATR and binding pocket for AdoCbl is solvent exposed but tucked into a cavity within the protein. **C.** The conserved phenylalanine in the cofactor binding pocket (F110 in *Cm*ATR and F117 in *Mtb*ATR)(PDB: 6WGS) displayed using a space-filling model (*Cm*ATR in yellow ribbons and *Mtb*ATR in grey ribbons) [15]. Residues (*Cm*ATR/*Mtb*ATR) composing the hydrophobic pocket are shown in sticks (*Cm*ATR in yellow and *Mtb*ATR in grey, both in ribbons and surfaces). The space occupied by F110/F117 prevents solvent from coordinating the cobalt in the lower axial position. **D.** AdoCbl (sticks, carbon in pink), PPP$_i$, (sticks colored by heteroatom) and Mg^{2+} (green spheres) in the binding sites of *Mtb*ATR (grey sticks and ribbons) aligned to *Cm*ATR (yellow sticks and ribbons). Residue numbers labeled as *Cm*ATR/*Mtb*ATR. Disordered residues are in parentheses. The similarity between the bound residues from *Mtb*ATR and the residues predicted

to bind from *Cm*ATR indicates the binding sites are pre-formed, and little change occurs upon binding. **E.** The active site of *Mtb*ATR (grey ribbons) with PPP$_i$ (sticks colored by heteroatom) and cob(II)alamin (sticks, carbons in pink) bound (PDB: 6WH5) aligned on *Cm*ATR (yellow ribbons) shows an extended conformation of the DMB tail, which is pushed out by the N-terminal residues and further stabilized by C-terminal residues (dark grey ribbons) [15]. The extended conformation of the DMB tail is believed to provide a molecular handle to facilitate cofactor transfer. Figure is from the thesis of Francesca Vaccaro [29].
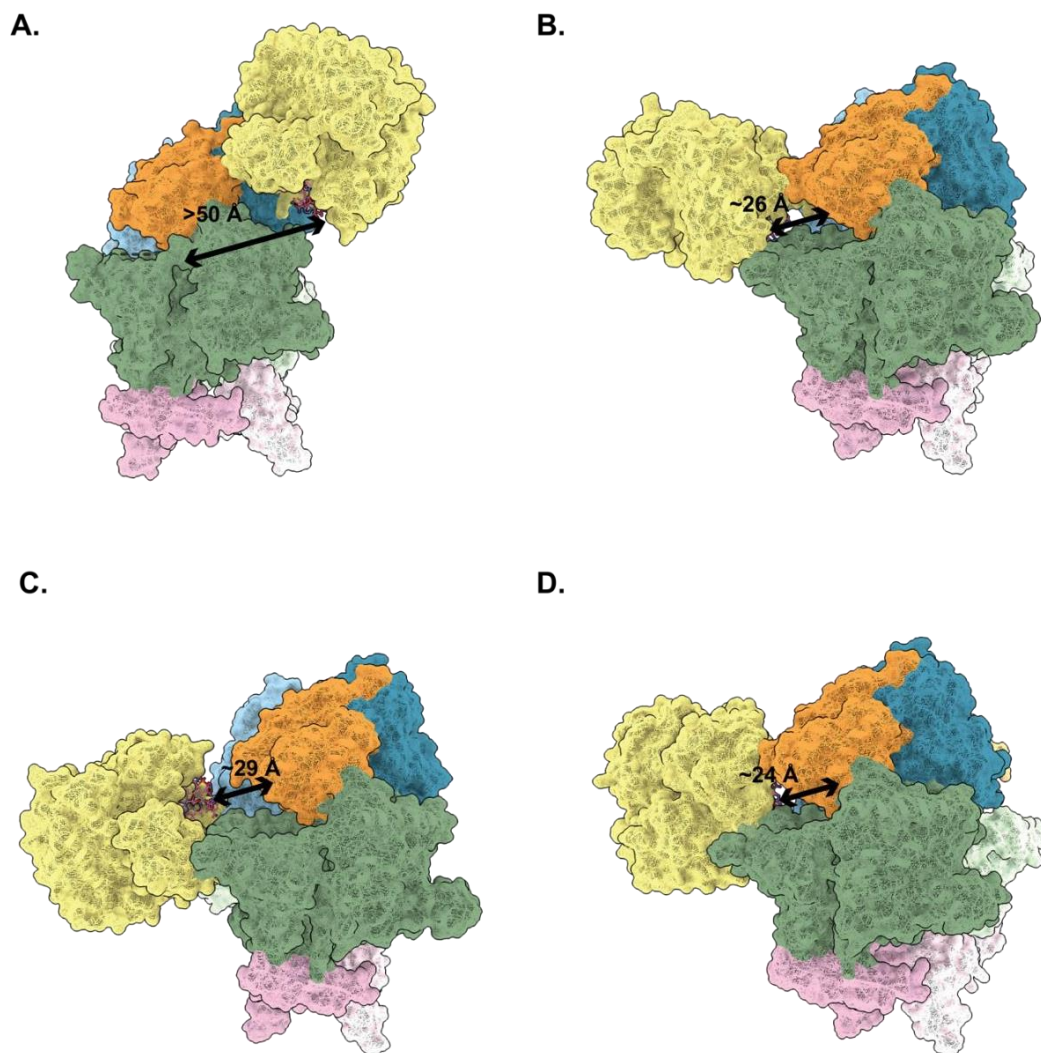
Figure 6: Docking solutions of the IcmF + ATR complex from *C. metallidruans*. A. The docking solution with the maximal amount of contact between IcmF (PDB: 4XC8) and apo-*Cm*ATR also has the furthest distance between the active sites of the two proteins. **B, C, and D.** The docking solutions representing the other top solutions have shorter distances between the active sites, but decreased contact between IcmF and ATR. In all solutions, it is unlikely that a direct hand-off of the cofactor is possible (G-protein domain: blue surface; Cbl-binding domain: orange surface; substrate-binding domain: green surface; linker: pink surface). Figure from the thesis of Francesca Vaccaro [29].
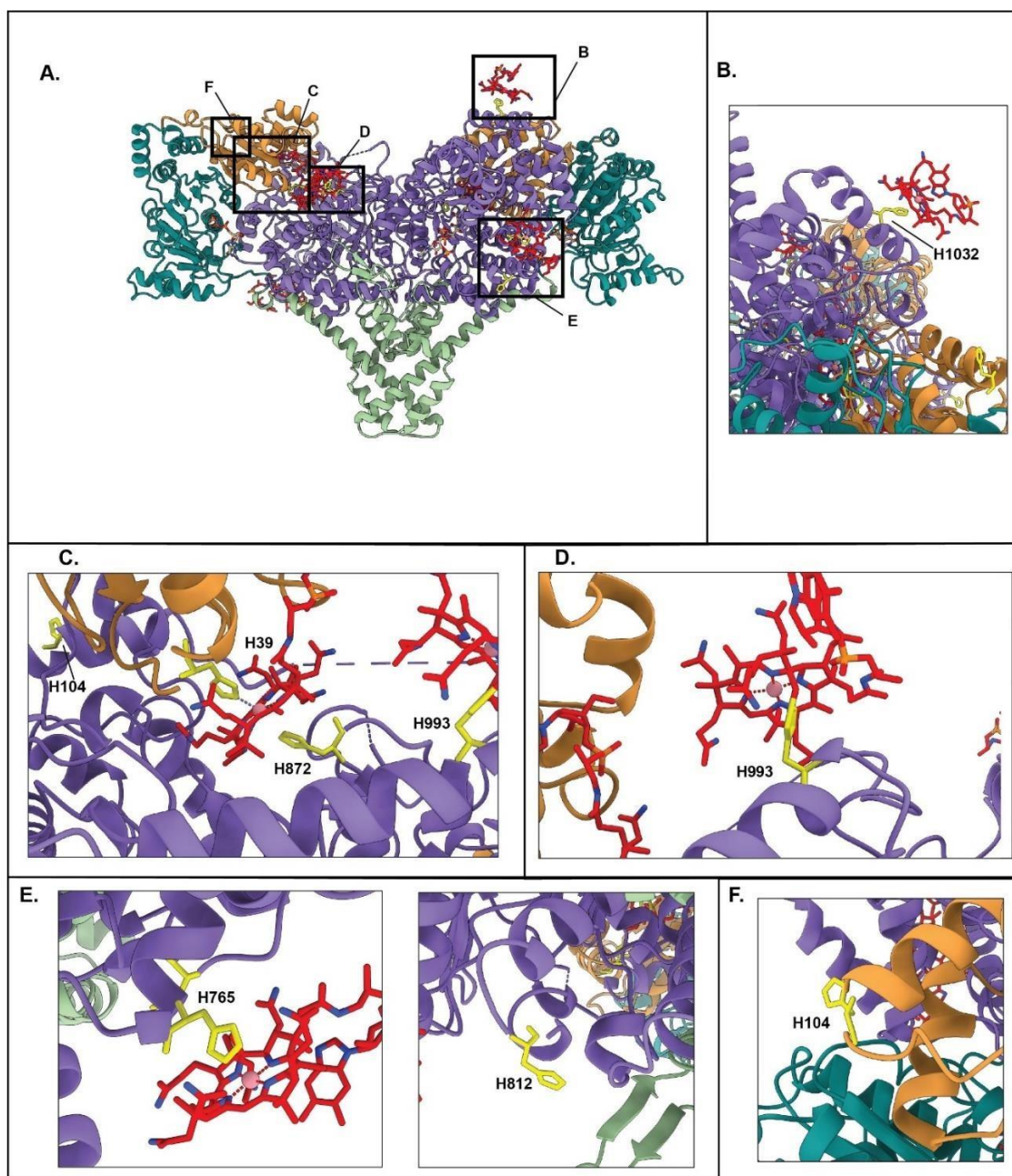
Figure 7: Partially refined structure of the IcmF dimer showing positions that had extra density consistent with Cbl binding. A. Overview of the IcmF dimer with locations for the following insets. **B.** H1032 is near extra density that is consistent with bound Cbl (carbons in red). **C.** H104, H39, H872, and H993 are near Cbl binding site. H39, the canonical Cbl-binding histidine, appears to be binding Cbl in a bis-His manner along with H872. **D.** H993 is also near extra density that is consistent with Cbl. **E.** Left: H765, with Cbl modeled in nearby. Right: H812 (model for Cbl

is not shown). **F.** H104 (model for Cbl is not shown). Color scheme: Histidine residues of interest are shown as yellow sticks; substrate-binding TIM barrel domain is shown as purple ribbons' Rossmann Cbl-binding domain is shown as orange ribbons; G-protein domain is shown as teal ribbons; linker is shown as green ribbons; AdoCbl is shown in sticks (carbons in red, cobalt in pink sphere); GDP is shown sticks colored by heteroatom. Dataset collected by David Born and Marco Jost, fully reprocessed and partially refined by Daphne Faber.
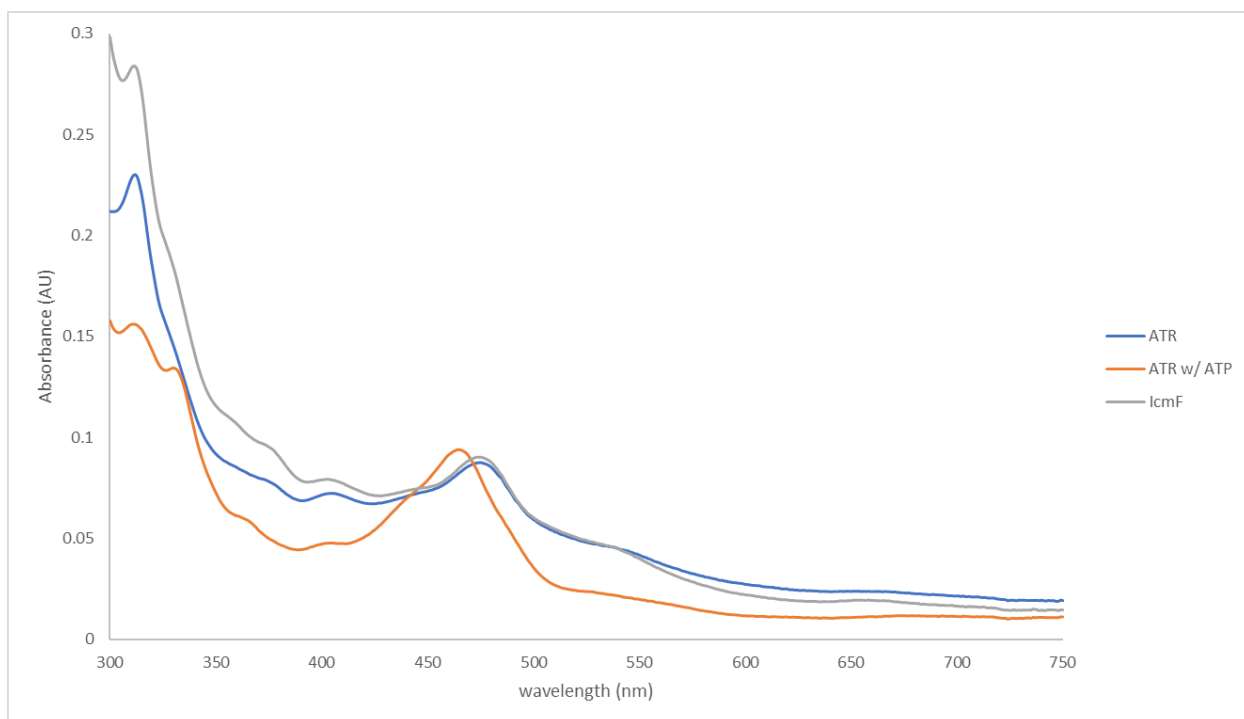
Figure 8: UV-Vis spectra for control experiments of the cofactor transfer assay. The blue spectrum labeled "ATR" is ATR incubated with cob(II)alamin, the orange spectrum labeled "ATR w/ ATP" is ATR incubated with cob(II)alamin + ATP, and the grey spectrum labeled "IcmF" is IcmF incubated with cob(II)alamin; the λ_{max} values are 476 nm, 465 nm, and 476 nm, respectively. The IcmF + cob(II)alamin sample establishes the initial λ_{max} for the transfer. The ATR + cob(II)alamin + ATP sample establishes the endpoint λ_{max} of the transfer. The λ_{max} value of 476 nm indicates cob(II)alamin either in solution and 5-coordinate base-off or bound to IcmF and 5-coordinate base-off/His-on. The λ_{max} of 465 nm indicates cob(II)alamin bound to ATR and 4-coordinate base-off.
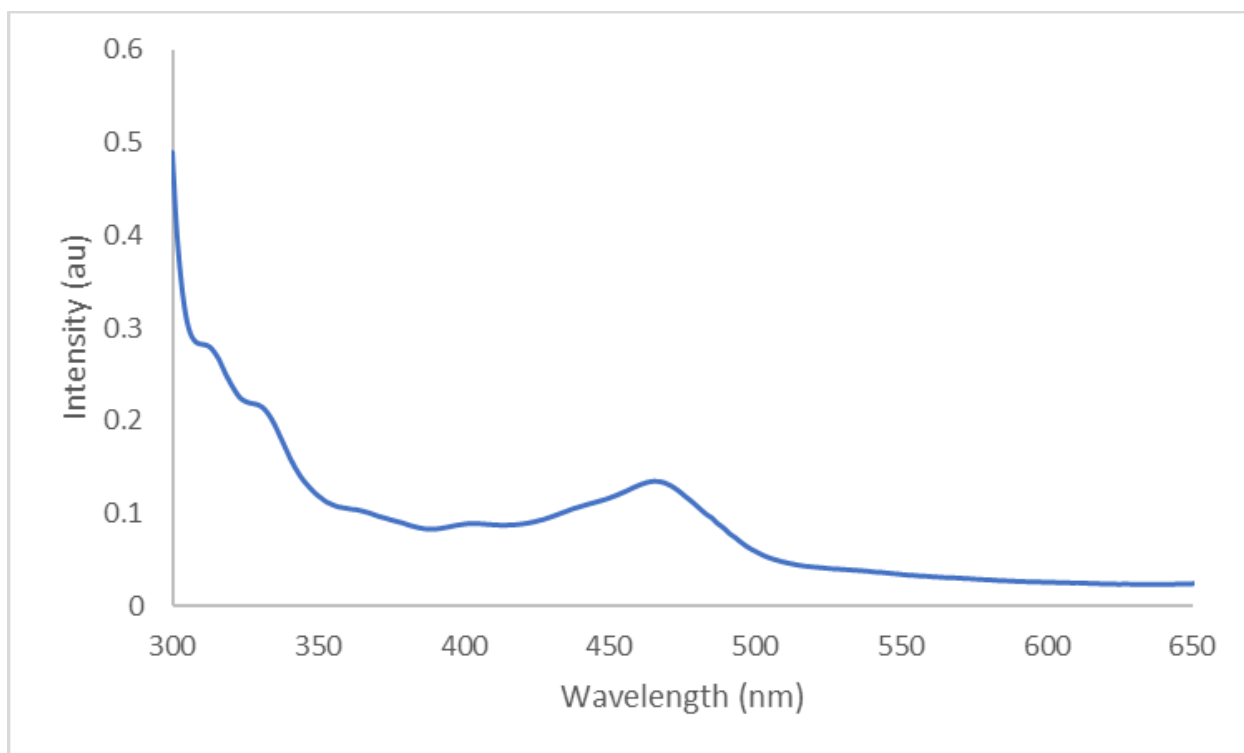
Figure 9: UV-Vis spectrum of ATR with Cob(II)alamin **and PPP$_i$ to verify *CmATR* behavior.** ATR was incubated with cob(II)alamin and PPP$_i$. Other PduO-type ATRs are unable to bind cob(II)alamin without ATP present, and PPP$_i$ is not sufficient to enable cob(II)alamin binding. The λ_{max} is 476 nm, indicating that cob(II)alamin, is unbound and in solution, 5-coordinate base-off.
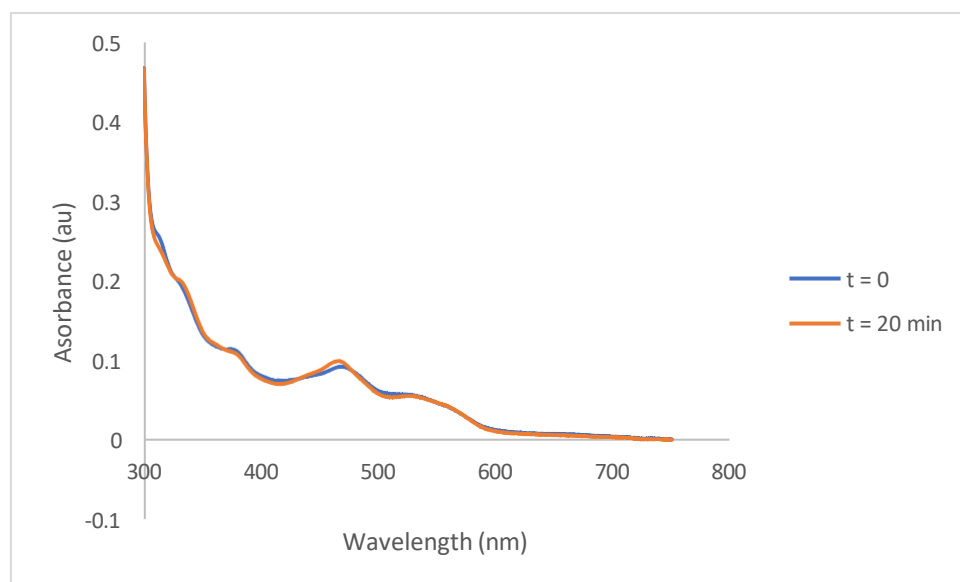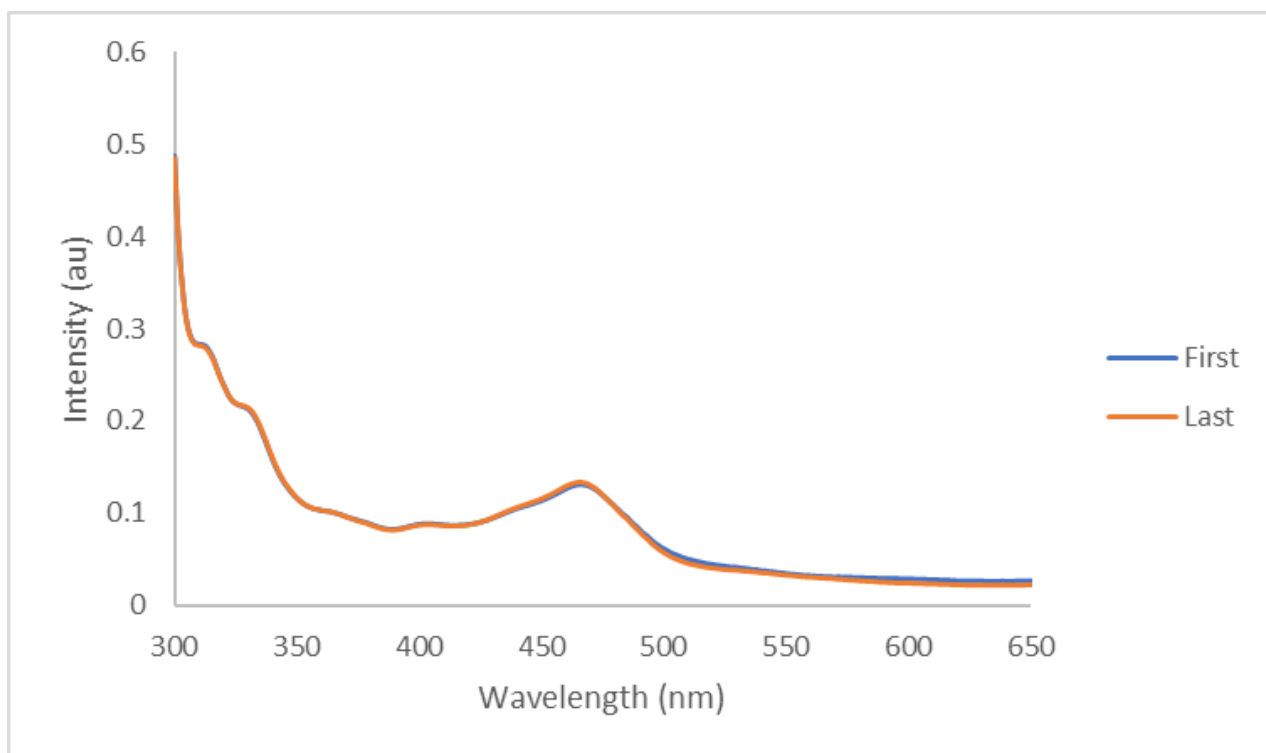
Figure 10: UV-Vis spectra, off-loading assay Experiment 1. The buffer contained no glycerol, conducted at 25 °C, with scans every 15 s for 20 min. The blue spectrum labeled "t=0" corresponds to the first scan taken, and the orange spectrum labeled "t=20 min" corresponds to the last scan taken; the λ_{max} values for both were 465 nm, indicating cob(II)alamin bound to ATR. Scans, not shown on this figure, taken in between the first and last time points do not show any change in λ_{max}.

Figure 11: UV-Vis spectra, off-loading assay Experiment 2. The buffer contained no glycerol, conducted at 25 °C, with scans every 15 s for 20 min. The blue spectrum labeled "First" corresponds to the first scan taken, and the orange spectrum labeled "Last" corresponds to the last scan taken; the λ_{max} values for both were 465 nm, indicating cob(II)alamin bound to ATR. Scans, not shown on this figure, collected in between the first and last time points do not show any change in λ_{max}.
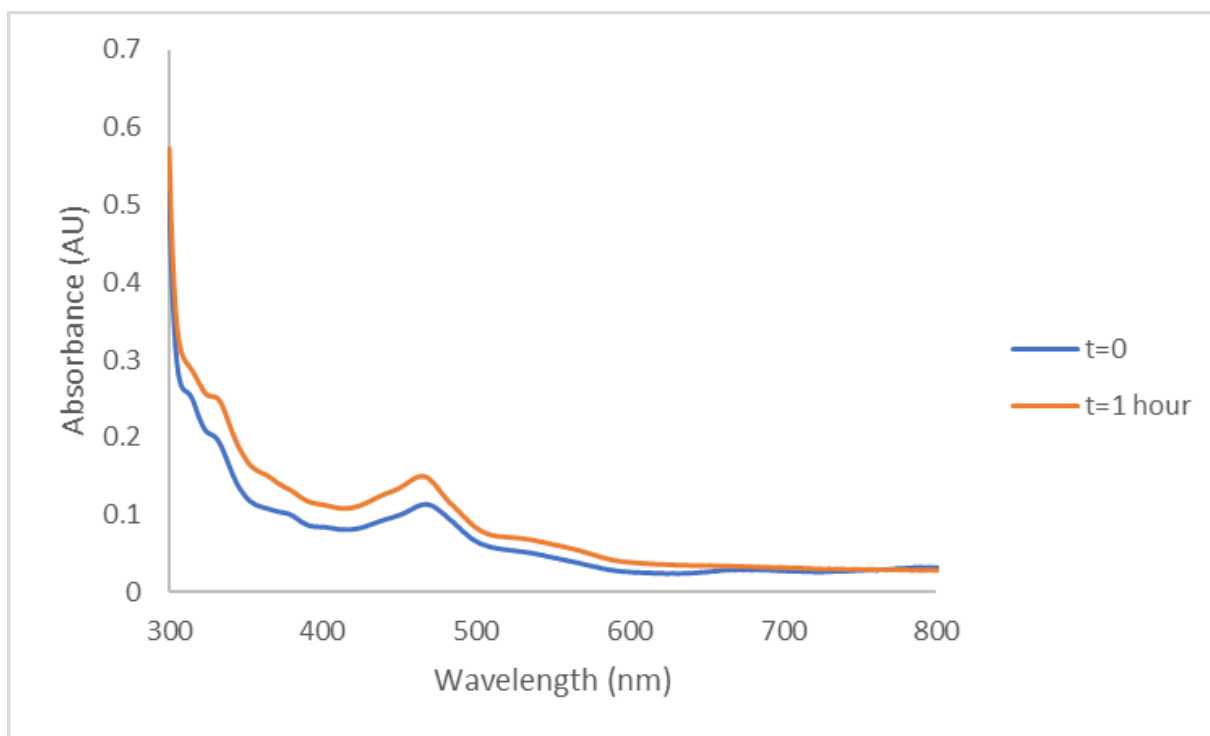
Figure 12: UV-Vis spectra, off-loading assay Experiment 3. The buffer contained 5% v/v glycerol, conducted at 25 °C, with scans every 15 s for 1 h. The blue spectrum labeled "t=0" corresponds to the first scan taken, and the orange spectrum labeled "t=1 hour" corresponds to the last scan taken; the λ_{max} values were 467nm and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 465 nm indicates cob(II)alamin bound to ATR.
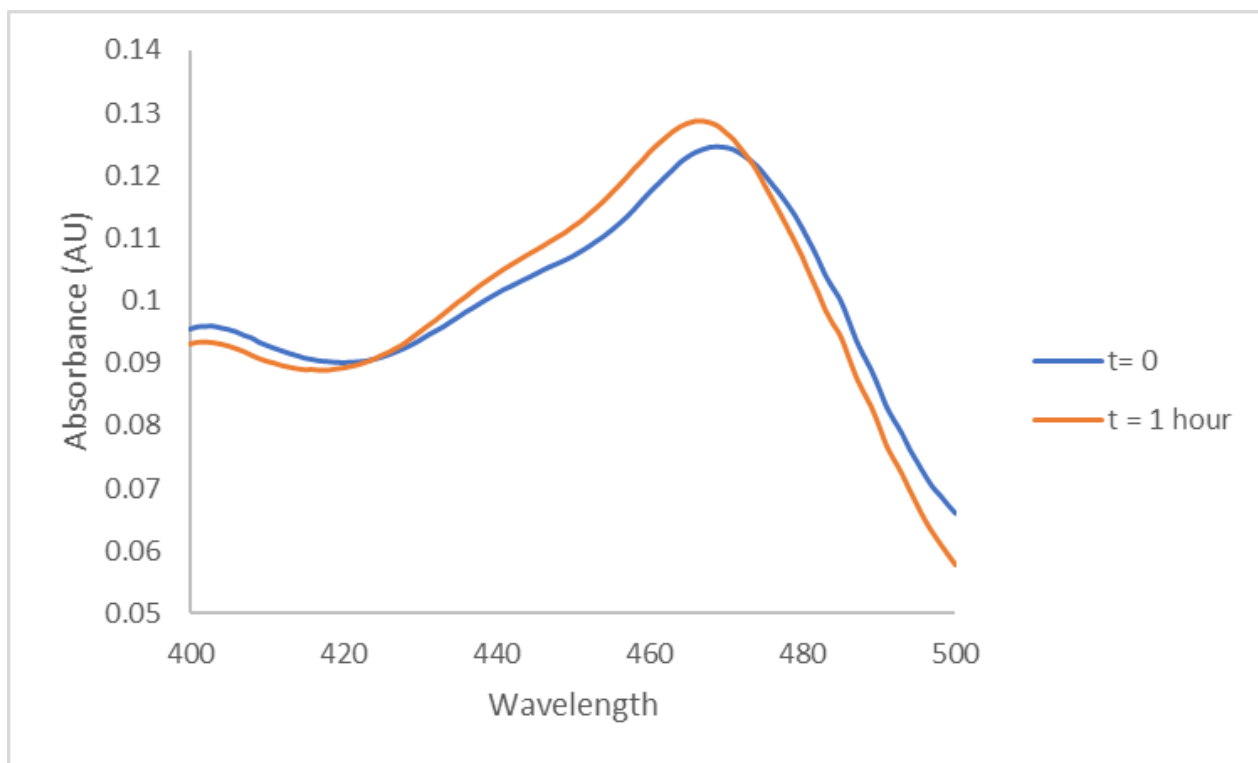
Figure 13: UV-Vis spectra, off-loading assay Experiment 4. The buffer contained 30% v/v glycerol, conducted at 25 °C, with scans every 15 s for 1 h. The blue graph labeled "t= 0" corresponds to the first scan taken, and the orange graph labeled "t = 1 hour" corresponds to the last scan taken; the λ_{max} values were 469 and 466 nm, respectively. The λ_{max} of 469 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 466 nm indicates cob(II)alamin bound to ATR.
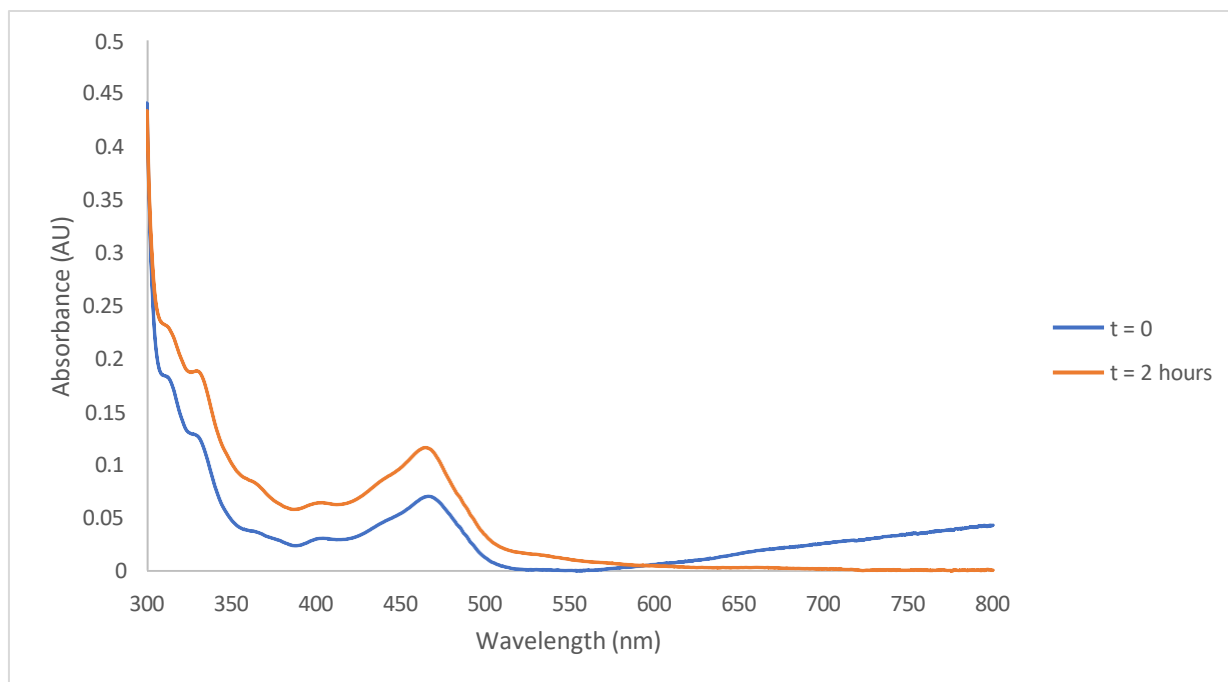
Figure 14: UV-Vis spectra, off-loading assay Experiment 5. The buffer contained 5% v/v glycerol, conducted at 4 °C, with scans every 15 s for 2 h. The final concentration of ATR was 10 µM. The blue spectrum labeled "t = 0" corresponds to the first scan taken, and the orange spectrum labeled "t = 2 hours" corresponds to the last scan taken; the λ_{max} values were 466 nm and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, while 465 nm indicates cob(II)alamin bound to ATR.
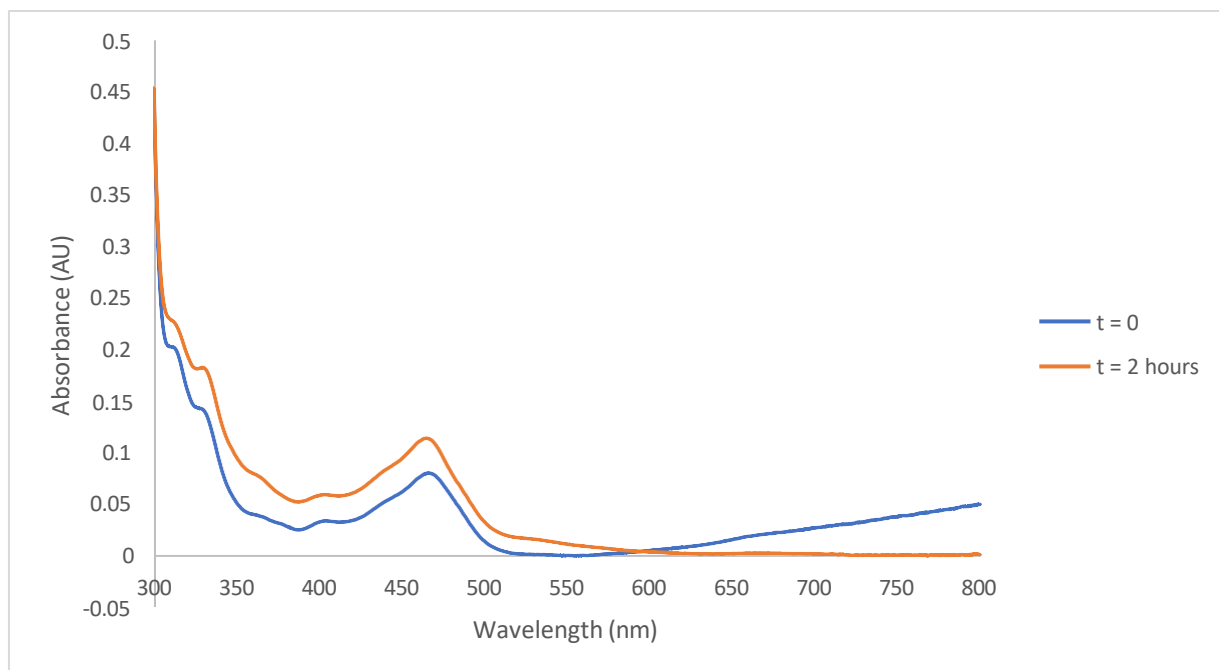
Figure 15: UV-Vis spectra, off-loading assay Experiment 6. The buffer contained 5% v/v glycerol, conducted at 4 °C, with scans every 15 s for 2 h. The final concentration of ATR was 8 µM. The blue spectrum labeled "t = 0" corresponds to the first scan taken, and the orange spectrum labeled "t = 2 hours" corresponds to the last scan taken; the λ_{max} values were 466 and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 465 nm indicates cob(II)alamin bound to ATR.
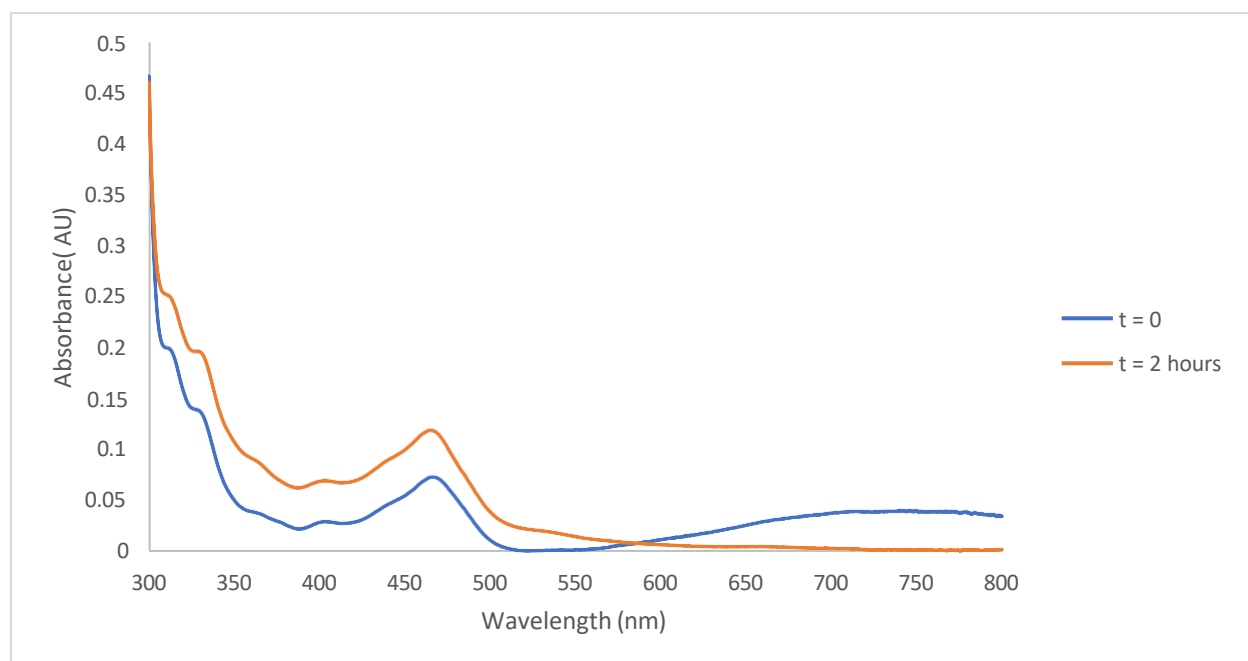
Figure 16: UV-Vis spectra, off-loading assay Experiment 7. The buffer contained 5% v/v glycerol, conducted at 4 °C, with scans every 15 s for 2 h. The final concentration of ATR was 6 µM. The blue spectrum labeled "t = 0" corresponds to the first scan taken, and the orange spectrum labeled "t = 2 hours" corresponds to the last scan taken; the λ_{max} values were 466 nm and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 465 nm indicates cob(II)alamin bound to ATR.

Figure 17: UV-Vis spectra, off-loading assay Experiment 8. The buffer contained 10% v/v glycerol, conducted at 4 °C using stirring, with scans every 15 s for 2 h. The final concentration of ATR is 10 µM. The blue spectrum labeled "[ATR] = 0.01 mM t = 0" corresponds to the first scan taken, and the orange spectrum labeled "[ATR] = 0.01 mM t = 2" hours corresponds to the last scan taken; the λ_{max} values were 467 nm and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 465 nm indicates cob(II)alamin bound to ATR.
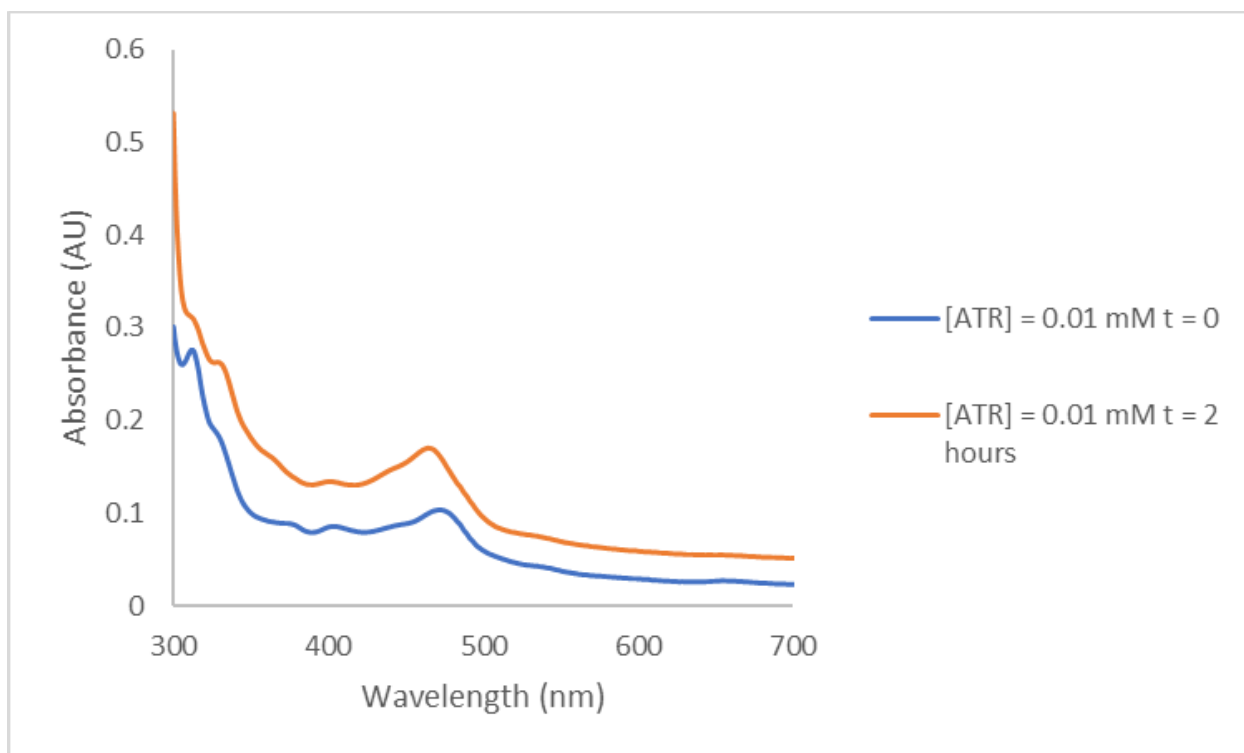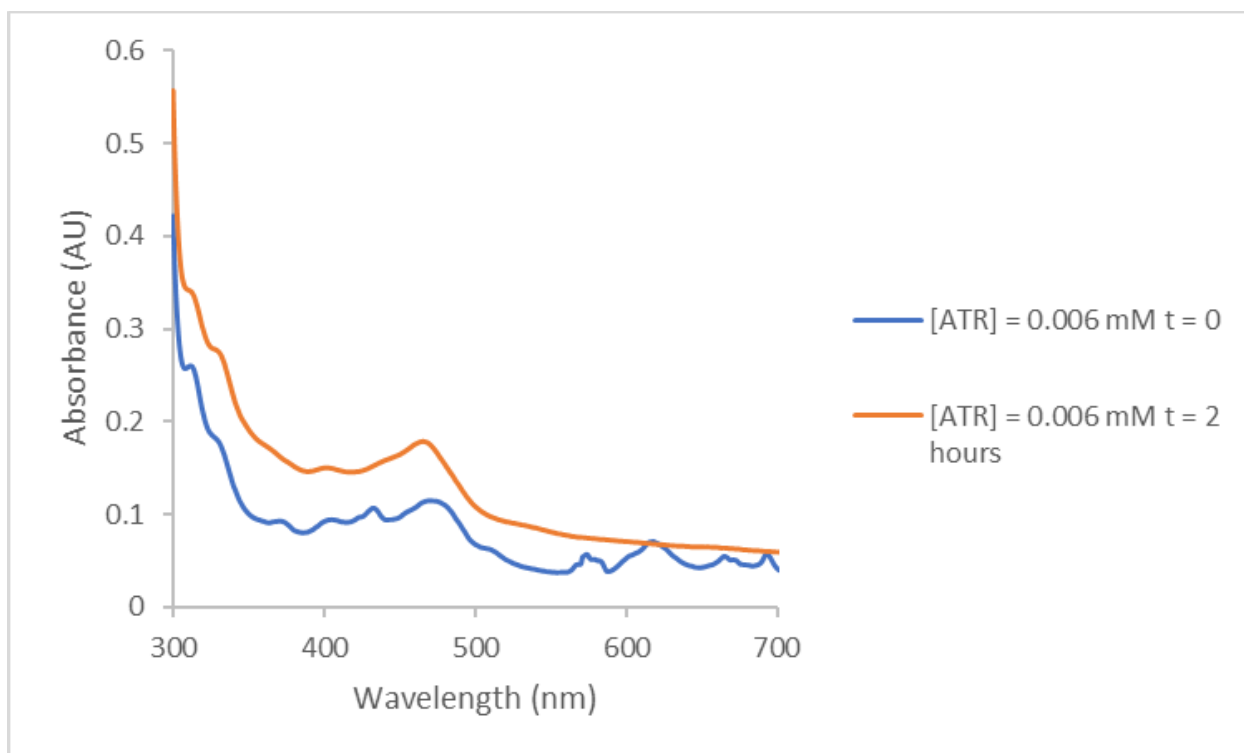
Figure 18: UV-Vis spectra, off-loading assay Experiment 9. The buffer contained 10% v/v glycerol, conducted at 4 °C using stirring, with scans every 15 s for 2 h. The final concentration of ATR is 6 µM. The blue spectrum labeled "[ATR] = 0.006 mM t = 0" corresponds to the first scan taken, and the orange spectrum labeled "[ATR] = 0.006 mM t = 0.2 hours" corresponds to the last scan taken; the λ_{max} values were 467 nm and 465 nm, respectively. The λ_{max} of 467 nm likely represents a mix of cofactor bound to ATR and IcmF, whereas 465 nm indicates cob(II)alamin bound to ATR.

Figure 19: Stopped-flow UV-Vis spectra, control experiment. Buffer contained 5% v/v glycerol, with data collected on a logarithmic timescale for 250 s, conducted at 25 °C. Sample only contained IcmF incubated with cob(II)alamin. The λ_{max} of 476 nm indicates that cob(II)alamin is bound to IcmF and is 5-coordinate base-off/His-on. Between the first and last time points, there is no significant change in the spectra, which is expected in the absence of the repair solution.
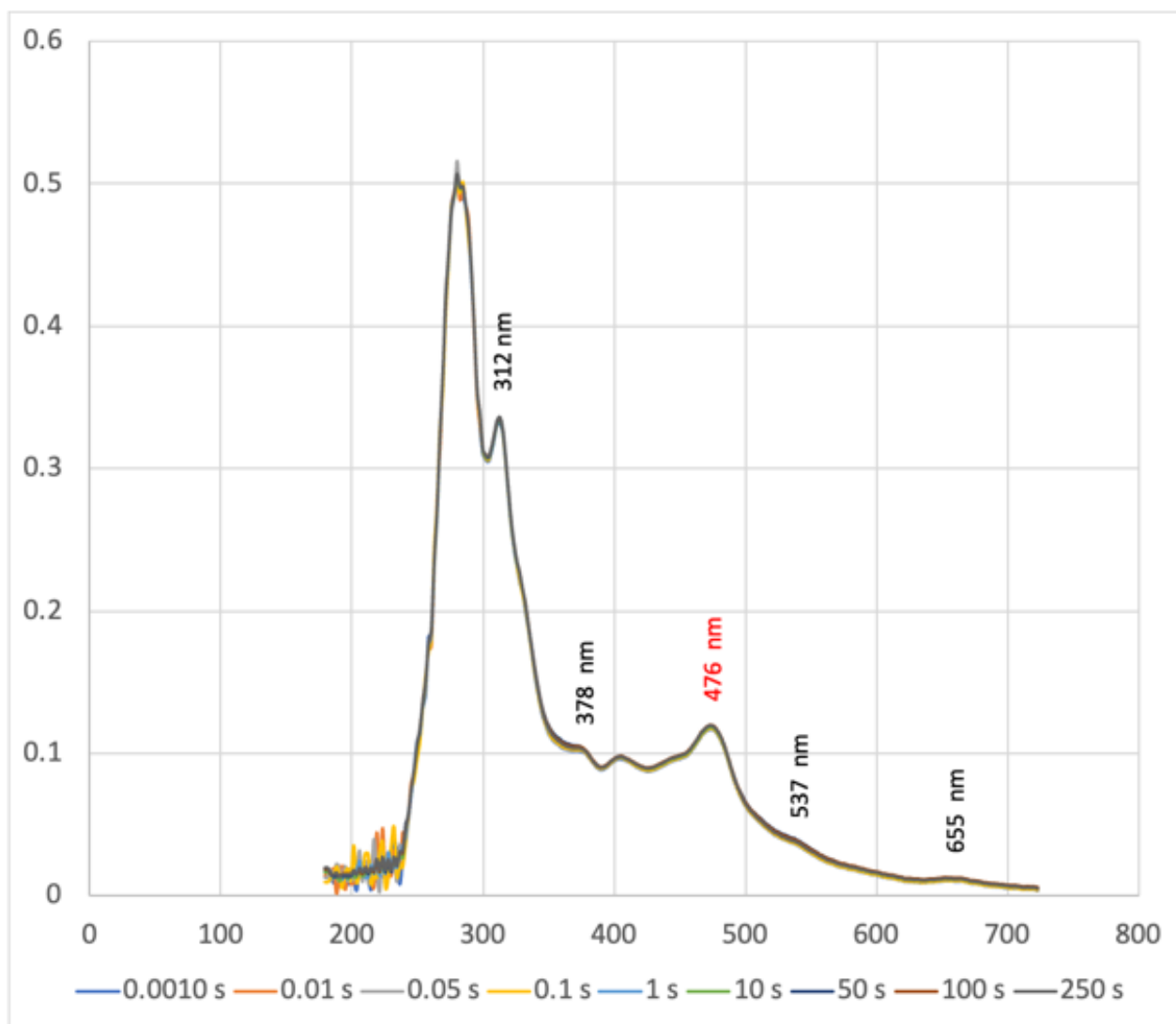
Figure 20: Stopped-flow UV-Vis spectra, off-loading assay Experiment 10. Buffer contained 5% v/v glycerol, with data collected on a logarithmic timescale for 250 s, conducted at 25 °C. The λ_{max} shifts from 476 to 465 nm. The λ_{max} of 476 nm indicates that cob(II)alamin bound to IcmF and is 5-coordinate base-off/His-on, and the λ_{max} of 465 nm indicates that cob(II)alamin bound to ATR and is 4-coordinate base-off.
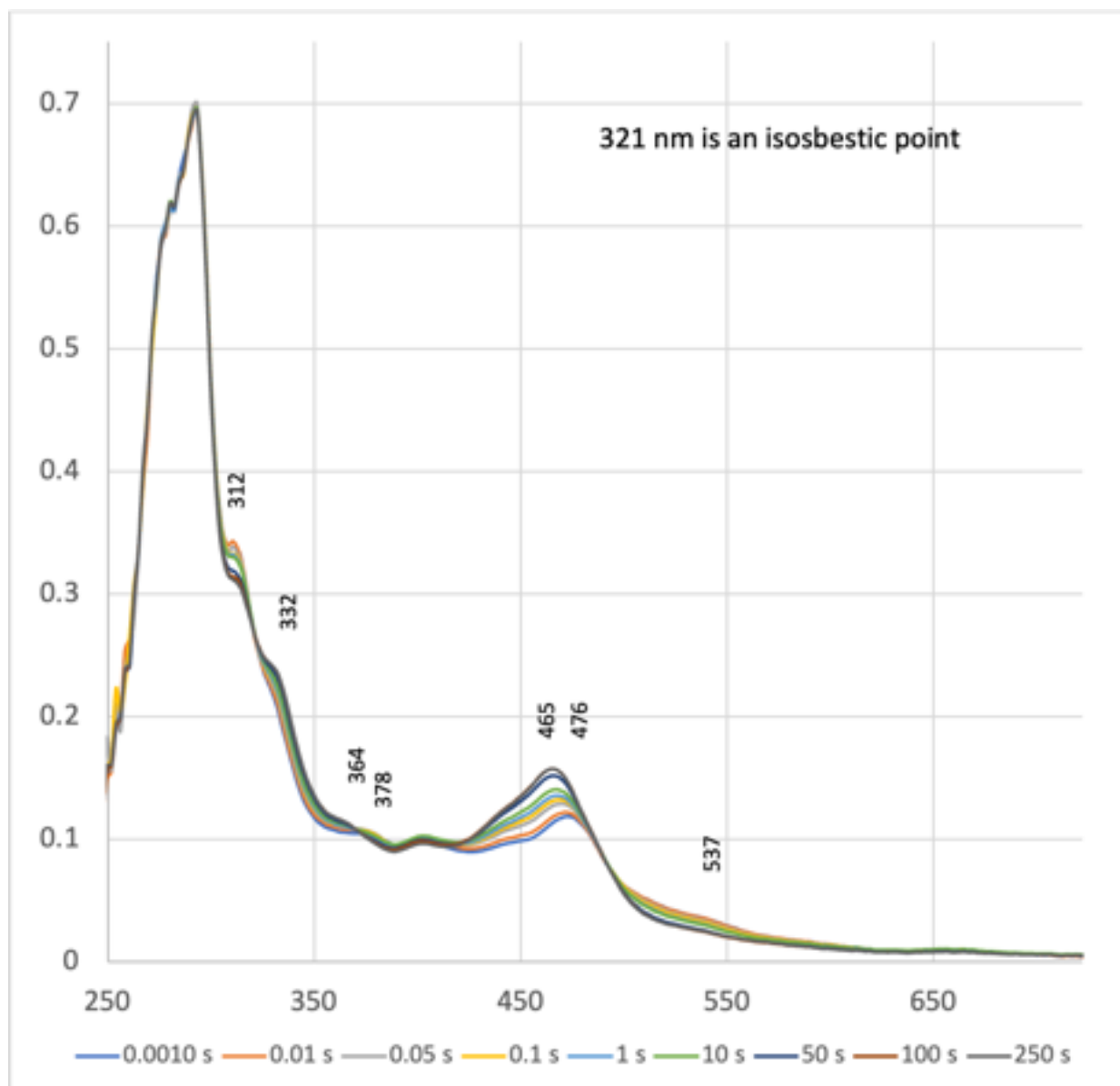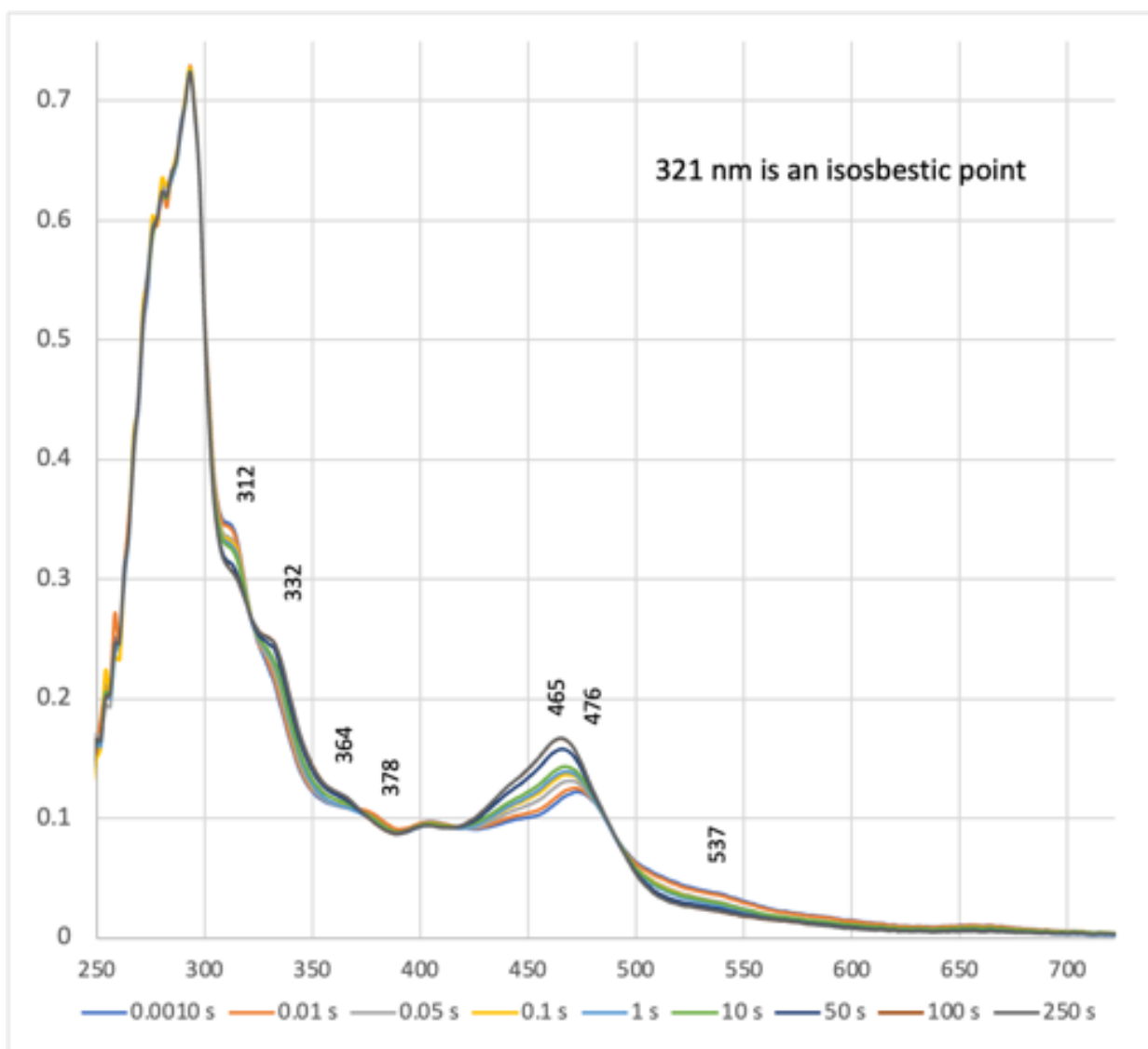
Figure 21: Stopped-flow UV-Vis spectra, off-loading assay Experiment 11. Buffer contained 5% v/v glycerol, with data collected on a logarithmic timescale for 250 s, conducted at 25 °C. The λ_{max} shifts from 476 to 465 nm. The λ_{max} of 476 nm indicates that cob(II)alamin bound to IcmF and is 5-coordinate base-off/His-on, and the λ_{max} of 465 nm indicates that cob(II)alamin bound to ATR and is 4-coordinate base-off.
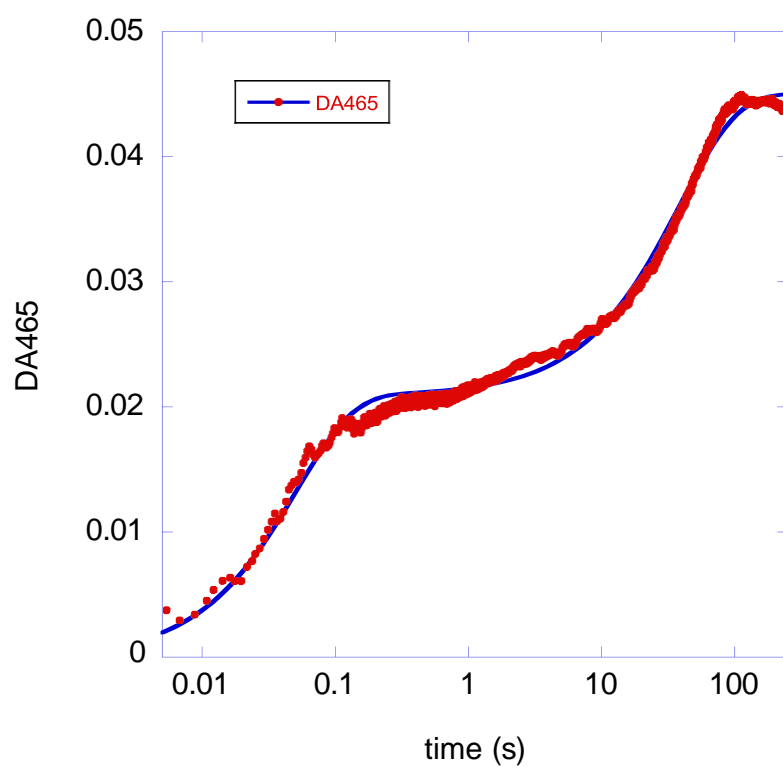
Figure 22: ΔA$_{465nm}$ for stopped flow assay Experiment 10. The red dots indicate the datapoints and the blue line is the fitted curve for the data.
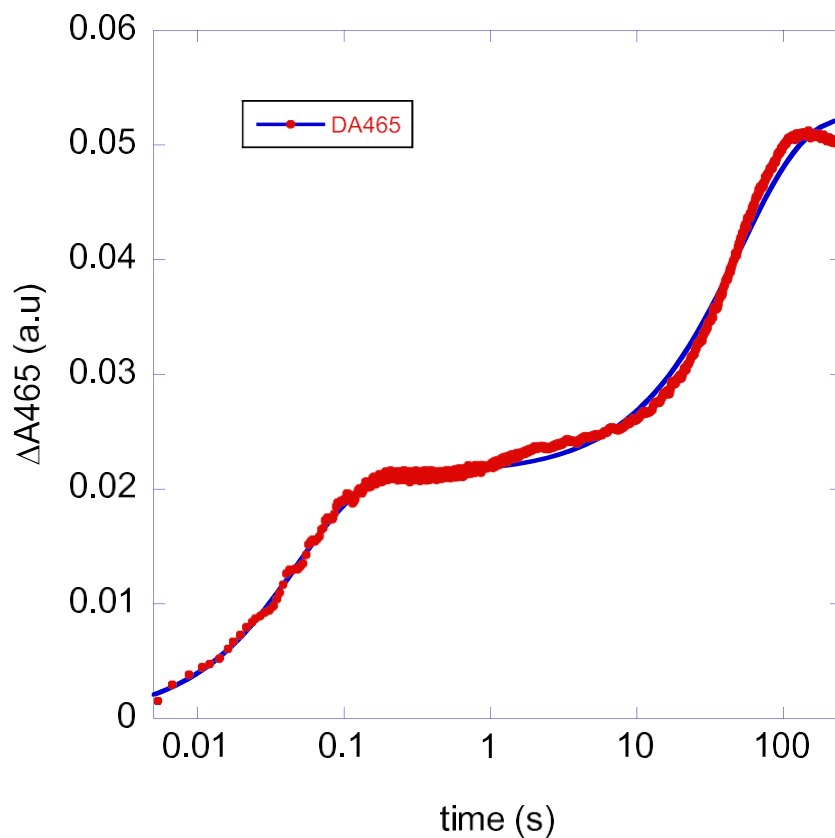
Figure 23: ΔA$_{465nm}$ for stopped flow assay Experiment 11. The red dots indicate datapoints and the blue line is the fitted curve for the data.
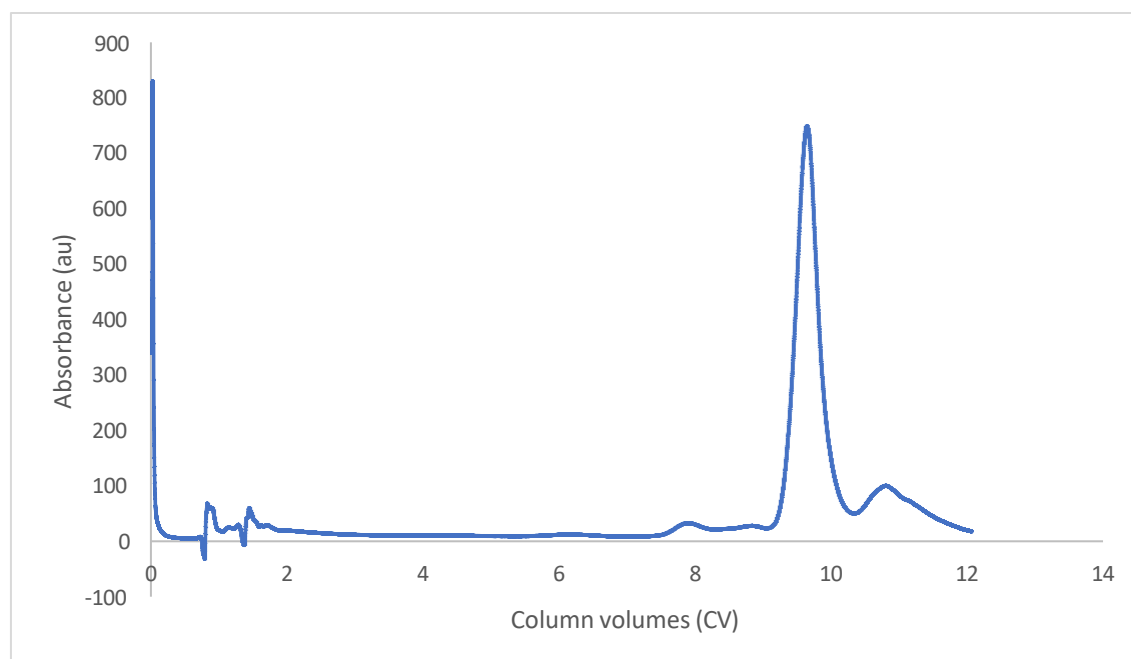
Figure 24: FPLC chromatogram trace, H872A purification attempt 1, from AEX step. The AEX step was performed using a MonoQ column, and the absorbance at 280 nm is being monitored.
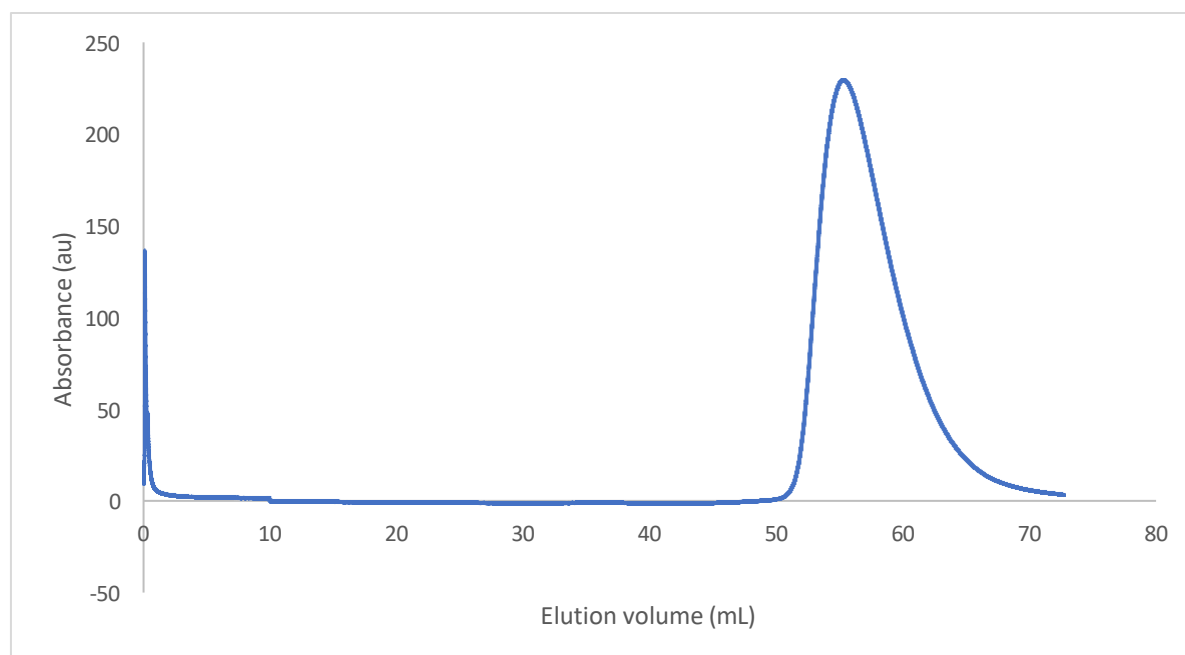
Figure 25: FPLC chromatograph trace, H872A purification attempt 1, from SEC step. SEC was done using a Superdex200 16/60 SEC dolumn, and the absorbance monitored at 280 nm.
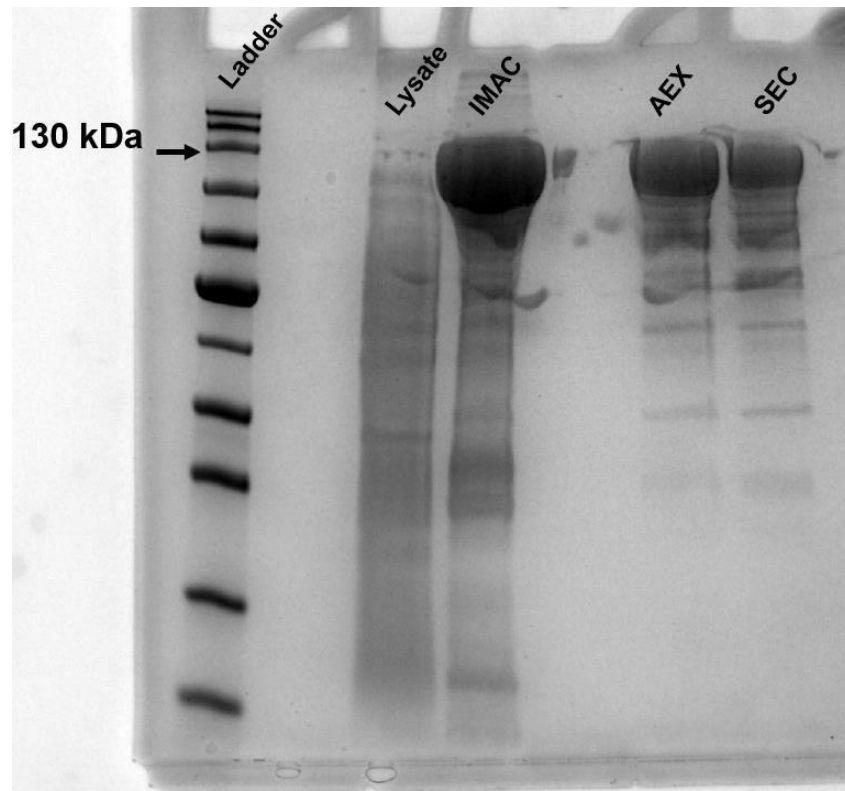
Figure 26: SDS-PAGE gel, wild-type IcmF. Lysate is from the soluble fraction of the cell pellet following sonification and centrifugation.
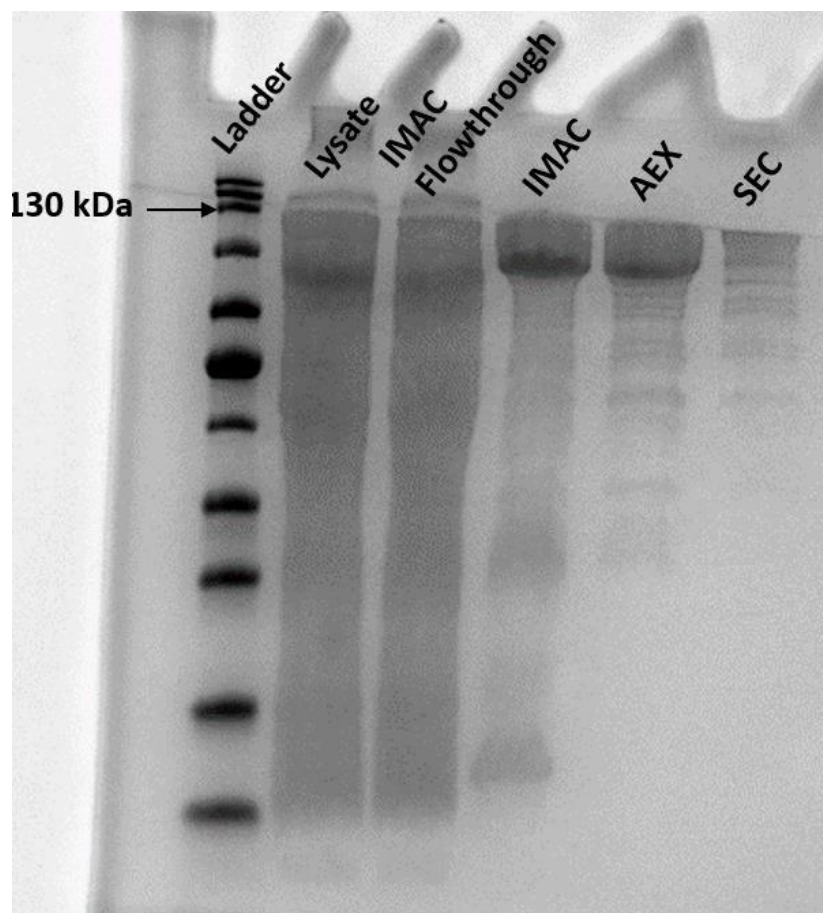
Figure 27: SDS-PAGE gel, H872A purification attempt 1. Lysate is from the soluble fraction of the cell pellet following sonification and centrifugation. IMAC flowthrough is from buffer washing the column with a low imidazole buffer following equilibration of the filtered lysate onto the column.
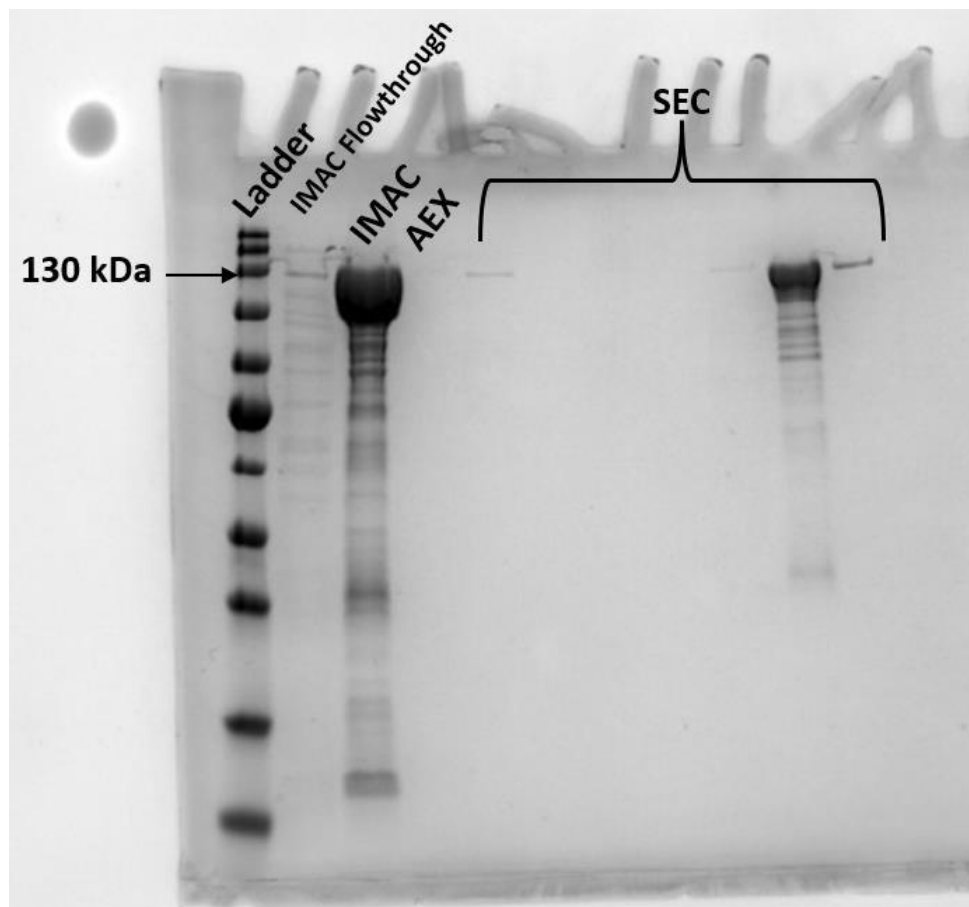
Figure 28: SDS-PAGE gel, H872A purification attempt 2. IMAC flowthrough is from buffer washing the column with a low imidazole buffer following equilibration of the filtered lysate onto the column. It is concerning that there was significantly less protein visible in this lane, as there should be large amounts of protein in this sample which do not bind or weakly bind the HisTrap column. It is also puzzling that there is no AEX protein peak, but we again see protein in SEC fractions.

Figure 29: SDS-PAGE gel, H872A purification attempt 3. Lysate is from the soluble fraction of the cell pellet following sonification and centrifugation. IMAC flowthrough is from buffer washing the column with a low imidazole buffer following equilibration of the filtered lysate onto the column. Overloading of the lanes makes evaluation of the bands inconclusive due to smearing.
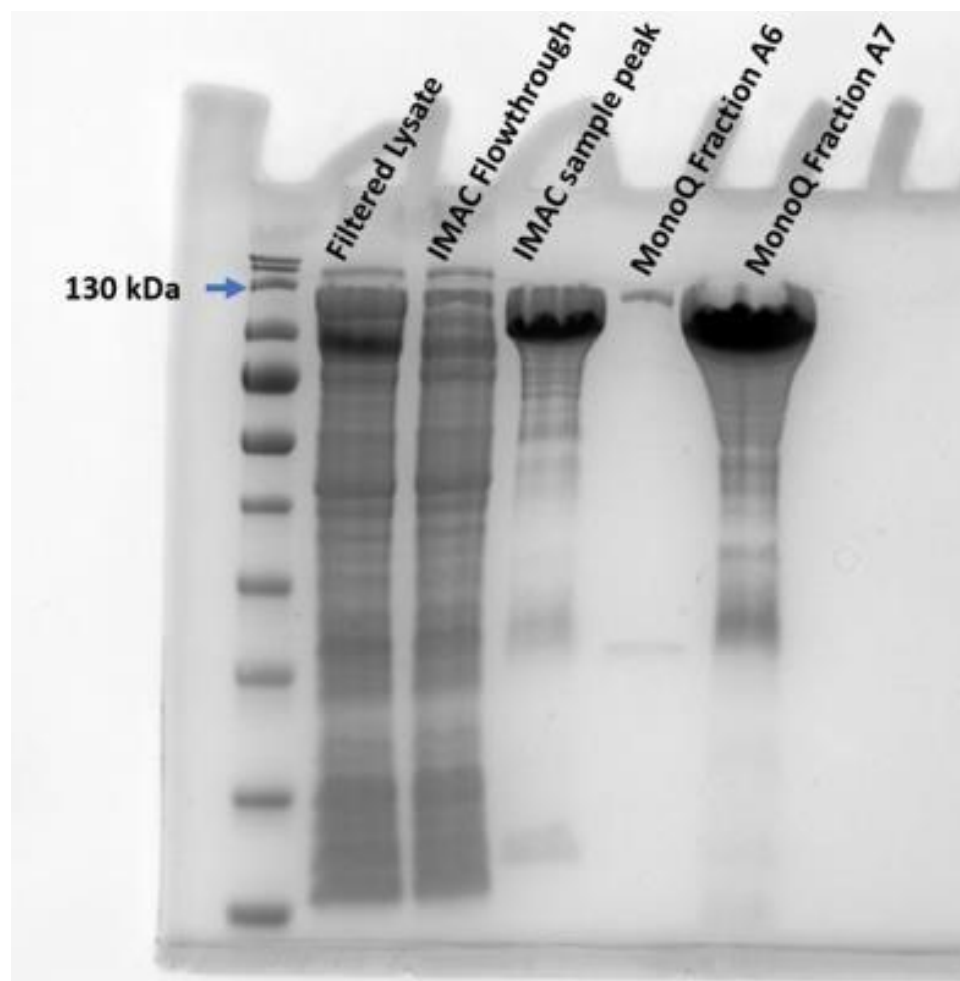
```
                                                           Canonical Cbl-binding histidine
G.kaustophlilus   ----------------MAHIYRPKHHVRFVTASSLFDGHDASINIMRRILQASGAEVIHL      44
N.farcinica       --------------MADSTLHQPAYPVRFVTSAALFDGHDAAINIMRRILQSQGAEVIHL      46
C.metallidurans   MTDLSDVSRTAAAKPPAVPGRGPANKVRFVTAASLFDGHDASINIMRRILQSQGCEVIHL      60
P.xenovorans      MTDLSTPQRAGSHKL------PAGRRLRFVTAAALFDGHDASINIMRRILQASGVEVIHL      54
                                  :****:::*******:*********:.* *****
```

....

```
                                                             CmIcmF H812
G.kaustophlilus   RNYYSVSISGYHIAEAGANPITQLAFTLANGFTYVEYYLSRGMHIDDFAPNLSFFFSNGL      819
N.farcinica       RNFYSVSISGYHIAEAGANPISQLAFTLANGFTYVEAYLARGMHIDDFAPNLSFFFSNGM      807
C.metallidurans   RNFYSVSISGYHIAEAGANPISQLAFTLANGFTYVEAYLARGMHIDDFAPNLSFFFSNGM      828
P.xenovorans      RNFYSVSISGYHIAEAGANPISQLAYTLANGFTYVEAYLARGMSIDDFAPNLSFFFSNGM     1007
                  **:*******************:***.********* **:*** ************** :

                                                             CmIcmF H872
G.kaustophlilus   DPEYSVIGRVARRIWAIVMREKYGANERSQKLKYHIQTSGRSLHAQEIDFNDIRTTLQAL      879
N.farcinica       DPEYSVIGRVARRIWAIALRDKYGAAERSQKLKYHVQTSGRSLHAQEMNFNDIRTTLQAL      867
C.metallidurans   DPEYSVLGRVARRIWAVTMRDKYGANDRSQKLKYHIQTSGRSLHAQEIDFNDIRTTLQAL      888
P.xenovorans      DPEYTVLGRVARRIWAVAMRERYGANERSQKLKYHVQTSGRSLHAQEIDFNDIRTTLQAL     1067
                  ****:*:***********::.*.:*** :********:********** :;**********

G.kaustophlilus   LAIYDNCNSLHTNAYDEAITTPTEESVRRAMAIQLIITKEFGLTKNENPLQGSFIIEELT      939
N.farcinica       IAIYDNCNSLHTNAYDEAVTTPTEDSVRRALAIQLIINREWGLAMNENPLQGSFIIDELT      927
C.metallidurans   IAIYDNCNSLHTNAYDEAITTPTAESVRRALAIQLIINREWGVAKCENPNQGSFLIEELT      948
P.xenovorans      IAIYDNCNSLHTNAFDEAITTPTEESVRRAVAIQLIINREWGLAKNQNPNQGSFVIEELT     1127
                  :*************:***:**** :*****:******.:*:*::   :** ****:*:***

                                                             CmIcmF H993
G.kaustophlilus   DLVEEAVLQEFERLNDRGGVLGAMEMQYQRGKIQDESLYYETKKHTGELPIIGVNTFLNP      999
N.farcinica       DLAEEAVLTEFERISERGGVLGAMETGYQRGKIQDESMLYEHRKHDGSLPIIGVNTFRNP      987
C.metallidurans   DLVEEAVLQEFERIAERGGVLGAMETGYQRGKIQEESLYYEQLKHDGTLPIIGVNTFRNP     1008
P.xenovorans      DLVEEAVLAEFDRLTERGGVLGAMETGYQRGRIQDESMLYEHRKHDGSYPIVGVNTFLSA     1187
                  **.***** **:*: :********* ****:**:**:**: **  ** * **:*****  .
```

....

Figure 30: Sequence alignment of *Cm*IcmF with IcmFs from other organisms. Alignments of IcmFs from *G. kaustophilus, N. farcinica, C. metallidurans, and P. xenovorans.* Sequence positions of interest are highlighted in yellow boxes. There is high sequence similarity between the different IcmFs. Several histidine residues that are of potential interest are conserved. Dotted lines indicate where portions of the alignment were removed for brevity. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions of conservation with strongly similar properties. Periods (.) denote positions of conservation with weakly similar properties. Alignments performed using Clustal Omega [44].

```
                                                        CmIcmF H765
C.metallidurans  VLQNVRGTVQADILKEDQGQNTCIFSTEFSLKVMGDIQEYFVHHQVRNFYSVSISGYHIA   205
H.sapiens        -KEKLTGTIQNDILKEFMVRNTYIFPPEPSMKIIADIFEYTAK-HMPKFNSISISGYHMQ   267
P.shermanii      -PEQLAGTIQNDILKEFMVRNTYIYPPQPSMRIISEIFAYTSA-NMPKWNSISISGYHMQ   246
M.extorquens     -PEKLAGTIQNDILKEFMVRNTYIYPPKGSMRIISDIFGYTSK-NMPKFNSISISGYHMQ   237
M.tuberculosis   -PEQLAGTIQNDILKEFMVRNTYIYPPKPSMRIISDIFAYTSA-KMPKFNSISISGYHIQ   262
                  :::  **:* *****    :** *:   : *::::.:*   *     ::  :: *:******:

                                                   CmIcmF H812
C.metallidurans  EAGANPISQLAFTLANGFTYVEAYLARGMHIDDFAPNLSFFFSNGMDPEYSVL-GRVARR   264
H.sapiens        EAGADAILELAYTLADGLEYSRTGLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRR   327
P.shermanii      EAGATADIEMAYTLADGVDYIRAGESVGLNVDQFAPRLSFFWGIGMNFFMEVAKLRAARM   306
M.extorquens     EAGATQDLELAYTLADGVEYIKAGLAAGLTIDQFAPRLSFFWAIGMNFFMEIAKMRAARL   297
M.tuberculosis   EAGATADLELAYTLADGVDYIRAGLNAGLDIDSFAPRLSFFWGIGMNFFMEVAKLRAGRL   322
                 ****     ::*:***:*. * .:     *: :*.***.****:. **:    .:   *..*

                                                  CmIcmF H872
C.metallidurans  IWAVTMRDKYG-ANDRSQKLKYHIQTSGRSLHAQEIDFNDIRTTLQALIAIYDNCNSLHT   323
H.sapiens        LWAHLIEKMFQPKNSKSLLLRAHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHT   387
P.shermanii      LWAKLVH-QFGPKNPKSMSLRTHSQTSGWSLTAQDVYNNVVRTCIEAMAATQGHTQSLHT   365
M.extorquens     IWAKLVK-EFEPKSDKSLPLRTHSQTSGWSLTAQDVFNNVTRTCIEAMAATQGGTQSLHT   356
M.tuberculosis   LWSELVA-QFAPKSAKSLSLRTHSQTSGWSLTAQDVFNNVARTCIEAMAATQGHTQSLHT   381
                 :*:   :    :   . :*  *:  * **** **  *:   *   ** ::*: *  .  :****

........................................................................................

                                                   CmIcmF H993
C.metallidurans  RIAERGGVLGAMETGYQRGKIQEESLYYEQLKHDGTLPIIGVNTFRNPNGDPTPQTLELA   443
H.sapiens        EIEEMGGMAKAVAEGIPKLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEV-LAID   506
P.shermanii      EVEKVGGMAKAIEKGIPKMRIEEAAARTQARIDSGRQPLIGVNKYRLEHEPPLDV-LKVD   484
M.extorquens     EVEALGGMAKAIEAGIPKLRIEEAAARAQARIDSGRQTIVGINKYKPDDEMKIDL-LRVD   475
M.tuberculosis   EVAEHGGMAQAISDGIPKLRIEEAAARTQARIDSGQQPVVGVNKYQVPEDHEIEV-LKVE   500
                 .: *   **: *: *  :  : :*:* :     :  ..*   ::*:*.::  .      *  :

                 CmIcmF H1032
C.metallidurans  RSSEDEKQSQLHRLTEFH-GAHQADAEAMLARLRQAVID---------NRNVFAVLMDAV   493
H.sapiens        NTS--VRNRQIEKLKKIKSSRDQALAERCLAALTECAAS-------GDGNILALAVDAS   556
P.shermanii      NST--VLAEQKAKLVKLRAERDPEKVKAALDKITWAAGNP---DDKDPDRNLLKLCIDAG   539
M.extorquens     NAD--VRAKQIDKLKRLRAERNQADVDAALAALTKAADGE---------GNLLELAVNAA   524
M.tuberculosis   NSR--VRAEQLAKLQRLRAGRDEPAVRAALAELTRAAAEQGRAGADGLGNNLLALAIDAA   558
                 .:.      *   :* .::  .    .   * :  ..       *:: : ::*

........................................................................................

                 Canonical Cbl-binding histidine
C.metallidurans  RFVTAASLFDGHDASINIMRRILQSQGCEVIHLGHNRSVQEVVTAALQEDVQGIAISSYQ   603
H.sapiens        RLLVAKMGQDGHDRGAKVIATGFADLGFDVDIGPLFQTPREVAQQAVDADVHAVGISTLA   675
P.shermanii      RILLAKMGQDGHDRGQKVIATAYADLGFDVDVGPLFQTPEETARQAVEADVHVVGVSSLA   658
M.extorquens     RILVAKMGQDGHDRGQKVIASAFADLGFDVDIGPLFATPDEAARQAVENDVHIVGVSSLA   644
M.tuberculosis   RILIAKMGQDGHDRGQKVIATAFADIGFDVDVGSLFSTPEEVARQAADNDVHIVGVSSLA   677
                 *:: *     ****  . :::     . * :*      :   *.. *  : **: :.:*:
```

Figure 31: Sequence alignment of IcmF and selected MCMs. The alignment for IcmF and MCM sequences from *H. sapiens, P. shermanii, M. extorquens,* and *M. tuberculosis*. Sequence positions of interest are highlighted in yellow boxes. The canonical Cbl-binding histidine in the active site, H39 in IcmF, is well conserved. Histidine residues suspected of potentially binding Cbl in a transient manner for cofactor off-loading do not appear to be conserved among MCMs. To align the conserved active site histidine properly, sequences annotated for the G-protein domain

and linker were removed. Residues were ordered 578-1093, 1-169. Dotted lines indicate where portions of the alignment were removed for brevity. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions of conservation with strongly similar properties. Periods (.) denote positions of conservation with weakly similar properties. Alignments performed using Clustal Omega [44].

References

(1) Watkins, D.; Rosenblatt, D. S. Inherited Disorders of Folate and Cobalamin Transport and Metabolism. In *The Online Metabolic and Molecular Bases of Inherited Disease*; Valle, D. L., Antonarakis, S., Ballabio, A., Beaudet, A. L., Mitchell, G. A., Eds.; McGraw-Hill Education: New York, NY, 2019.

(2) Banerjee, R.; Vlasie, M. Controlling the Reactivity of Radical Intermediates by Coenzyme B12-Dependent Methylmalonyl-CoA Mutase. *Biochem. Soc. Trans.* **2002**, *30* (4), 621–624. https://doi.org/10.1042/bst0300621.

(3) Kräutler, B. Vitamin B12: Chemistry and Biochemistry. *Biochem. Soc. Trans.* **2005**, *33* (4), 806–810. https://doi.org/10.1042/BST0330806.

(4) Mancia, F.; Keep, N. H.; Nakagawa, A.; Leadlay, P. F.; McSweeney, S.; Rasmussen, B.; Secke, P. B.; Diat, O.; Evans, P. R. How Coenzyme B12 Radicals Are Generated: The Crystal Structure of Methylmalonyl-Coenzyme A Mutase at 2 å Resolution. *Structure* **1996**, *4* (3), 339–350. https://doi.org/10.1016/S0969-2126(96)00037-8.

(5) Tollinger, M.; Konrat, R.; Hilbert, B. H.; Marsh, E. N. G.; Kräutler, B. How a Protein Prepares for B12 Binding: Structure and Dynamics of the B12-Binding Subunit of Glutamate Mutase from Clostridium Tetanomorphum. *Structure* **1998**, *6* (8), 1021–1033. https://doi.org/10.1016/S0969-2126(98)00103-8.

(6) Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Ludwig, M. L. How a Protein Binds B$_{12}$: A 3.0 Å X-Ray Structure of B$_{12}$-Binding Domains of Methionine Synthase. *Science* **1994**, *266* (5191), 1669–1674. https://doi.org/10.1126/science.7992050.

(7) Marsh, E. N. G.; Meléndez, G. D. R. Adenosylcobalamin Enzymes: Theory and Experiment Begin to Converge. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2012**, *1824* (11), 1154–1164. https://doi.org/10.1016/j.bbapap.2012.03.012.

(8) Mansoorabadi, S. O.; Padmakumar, R.; Fazliddinova, N.; Vlasie, M.; Banerjee, R.; Reed, G. H. Characterization of a Succinyl-CoA Radical−Cob(II)Alamin Spin Triplet Intermediate in the Reaction Catalyzed by Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase. *Biochemistry* **2005**, *44* (9), 3153–3158. https://doi.org/10.1021/bi0482102.

(9) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. Adenosyltransferase Tailors and Delivers Coenzyme B12. *Nat. Chem. Biol.* **2008**, *4* (3), 194–196. https://doi.org/10.1038/nchembio.67.

(10) Takahashi-Íñiguez, T.; García-Arellano, H.; Trujillo-Roldán, M. A.; Flores, M. E. Protection and Reactivation of Human Methylmalonyl-CoA Mutase by MMAA Protein. *Biochem. Biophys. Res. Commun.* **2011**, *404* (1), 443–447. https://doi.org/10.1016/j.bbrc.2010.11.141.

(11) Takahashi-Iñiguez, T.; González-Noriega, A.; Michalak, C.; Flores, M. E. Human MMAA Induces the Release of Inactive Cofactor and Restores Methylmalonyl-CoA Mutase Activity through Their Complex Formation. *Biochimie* **2017**, *142*, 191–196. https://doi.org/10.1016/j.biochi.2017.09.012.

(12) Froese, D. S.; Kochan, G.; Muniz, J. R. C.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the Human GTPase

MMAA and Vitamin B12-Dependent Methylmalonyl-CoA Mutase and Insight into Their Complex Formation. *J. Biol. Chem.* **2010**, *285* (49), 38204–38213. https://doi.org/10.1074/jbc.M110.177717.

(13)   Hubbard, P. A.; Padovani, D.; Labunska, T.; Mahlstedt, S. A.; Banerjee, R.; Drennan, C. L. Crystal Structure and Mutagenesis of the Metallochaperone MeaB. *J. Biol. Chem.* **2007**, *282* (43), 31308–31316. https://doi.org/10.1074/jbc.M704850200.

(14)   Vaccaro, F. A.; Born, D. A.; Drennan, C. L. Structure of Metallochaperone in Complex with the Cobalamin-Binding Domain of Its Target Mutase Provides Insight into Cofactor Delivery. *Proc. Natl. Acad. Sci.* **2023**, *120* (8), e2214085120. https://doi.org/10.1073/pnas.2214085120.

(15)   Mascarenhas, R.; Ruetz, M.; McDevitt, L.; Koutmos, M.; Banerjee, R. Mobile Loop Dynamics in Adenosyltransferase Control Binding and Reactivity of Coenzyme B_{12}. *Proc. Natl. Acad. Sci.* **2020**, *117* (48), 30412–30422. https://doi.org/10.1073/pnas.2007332117.

(16)   Padovani, D.; Banerjee, R. A G-Protein Editor Gates Coenzyme B_{12} Loading and Is Corrupted in Methylmalonic Aciduria. *Proc. Natl. Acad. Sci.* **2009**, *106* (51), 21567–21572. https://doi.org/10.1073/pnas.0908106106.

(17)   Padovani, D.; Banerjee, R. Assembly and Protection of the Radical Enzyme, Methylmalonyl-CoA Mutase, by Its Chaperone. *Biochemistry* **2006**, *45* (30), 9300–9306. https://doi.org/10.1021/bi0604532.

(18)   Vaccaro, F. A.; Faber, D. A.; Andree, G. A.; Born, D. A.; Kang, G.; Fonseca, D. R.; Jost, M.; Drennan, C. L. Structural Insight into G-Protein Chaperone-Mediated Maturation of a Bacterial Adenosylcobalamin-Dependent Mutase. *J. Biol. Chem.* **2023**, *299* (9), 105109. https://doi.org/10.1016/j.jbc.2023.105109.

(19)   Dobson, C. M. Identification of the Gene Responsible for the cblB Complementation Group of Vitamin B12-Dependent Methylmalonic Aciduria. *Hum. Mol. Genet.* **2002**, *11* (26), 3361–3369. https://doi.org/10.1093/hmg/11.26.3361.

(20)   Lerner-Ellis, J. P.; Dobson, C. M.; Wai, T.; Watkins, D.; Tirone, J. C.; Leclerc, D.; Doré, C.; Lepage, P.; Gravel, R. A.; Rosenblatt, D. S. Mutations in the MMAA Gene in Patients with the cblA Disorder of Vitamin B12 Metabolism. *Hum. Mutat.* **2004**, *24* (6), 509–516. https://doi.org/10.1002/humu.20104.

(21)   Fuchshuber, A.; Mucha, B.; Baumgartner, E. R.; Vollmer, M.; Hildebrandt, F. Mut0 Methylmalonic Acidemia: Eleven Novel Mutations of the Methylmalonyl CoA Mutase Including a Deletion-Insertion Mutation. *Hum. Mutat.* **2000**, *16* (2), 179–179. https://doi.org/10.1002/1098-1004(200008)16:2<179::AID-HUMU17>3.0.CO;2-R.

(22)   Mascarenhas, R.; Ruetz, M.; Gouda, H.; Heitman, N.; Yaw, M.; Banerjee, R. Architecture of the Human G-Protein-Methylmalonyl-CoA Mutase Nanoassembly for B12 Delivery and Repair. *Nat. Commun.* **2023**, *14* (1), 4332. https://doi.org/10.1038/s41467-023-40077-4.

(23)   Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a Radical B_{12} Enzyme with Its G-Protein Chaperone. *Proc. Natl. Acad. Sci.* **2015**, *112* (8), 2419–2424. https://doi.org/10.1073/pnas.1419582112.

(24)   Cracan, V.; Padovani, D.; Banerjee, R. IcmF Is a Fusion between the Radical B12 Enzyme Isobutyryl-CoA Mutase and Its G-Protein Chaperone. *J. Biol. Chem.* **2010**, *285* (1), 655–666. https://doi.org/10.1074/jbc.M109.062182.

(25)   Padovani, D.; Labunska, T.; Banerjee, R. Energetics of Interaction between the G-Protein Chaperone, MeaB, and B12-Dependent Methylmalonyl-CoA Mutase. *J. Biol. Chem.* **2006**, *281* (26), 17838–17844. https://doi.org/10.1074/jbc.M600047200.

(26)   Li, Z.; Kitanishi, K.; Twahir, U. T.; Cracan, V.; Chapman, D.; Warncke, K.; Banerjee, R. Cofactor Editing by the G-Protein Metallochaperone Domain Regulates the Radical B12 Enzyme IcmF. *J. Biol. Chem.* **2017**, *292* (10), 3977–3987. https://doi.org/10.1074/jbc.M117.775957.

(27)   Cracan, V.; Banerjee, R. Novel Coenzyme B12-Dependent Interconversion of Isovaleryl-CoA and Pivalyl-CoA. *J. Biol. Chem.* **2012**, *287* (6), 3723–3732. https://doi.org/10.1074/jbc.M111.320051.

(28)   Padovani, D.; Banerjee, R. A Rotary Mechanism for Coenzyme B$_{12}$ Synthesis by Adenosyltransferase. *Biochemistry* **2009**, *48* (23), 5350–5357. https://doi.org/10.1021/bi900454s.

(29)   Vaccaro, Francesca A. Structural Investigations of Adenosylcobalamin-Dependent Enzyme Maturation, PhD Thesis Massachusetts Institute of Technology, 2023.

(30)   Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7* (1), 539. https://doi.org/10.1038/msb.2011.75.

(31)   Campanello, G. C.; Ruetz, M.; Dodge, G. J.; Gouda, H.; Gupta, A.; Twahir, U. T.; Killian, M. M.; Watkins, D.; Rosenblatt, D. S.; Brunold, T. C.; Warncke, K.; Smith, J. L.; Banerjee, R. Sacrificial Cobalt–Carbon Bond Homolysis in Coenzyme B$_{12}$ as a Cofactor Conservation Strategy. *J. Am. Chem. Soc.* **2018**, *140* (41), 13205–13208. https://doi.org/10.1021/jacs.8b08659.

(32)   Lyskov, S.; Chou, F.-C.; Conchúir, S. Ó.; Der, B. S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B. D.; Renfrew, P. D.; Sripakdeevong, P.; Borgo, B.; Havranek, J. J.; Kuhlman, B.; Kortemme, T.; Bonneau, R.; Gray, J. J.; Das, R. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS ONE* **2013**, *8* (5), e63906. https://doi.org/10.1371/journal.pone.0063906.

(33)   Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; Ronneberger, O.; Bodenstein, S.; Zielinski, M.; Bridgland, A.; Potapenko, A.; Cowie, A.; Tunyasuvunakool, K.; Jain, R.; Clancy, E.; Kohli, P.; Jumper, J.; Hassabis, D. Protein Complex Prediction with AlphaFold-Multimer. October 4, 2021. https://doi.org/10.1101/2021.10.04.463034.

(34)   Mascarenhas, R.; Gouda, H.; Ruetz, M.; Banerjee, R. Human B12-Dependent Enzymes: Methionine Synthase and Methylmalonyl-CoA Mutase. In *Methods in Enzymology*; Elsevier, 2022; Vol. 668, pp 309–326. https://doi.org/10.1016/bs.mie.2021.12.012.

(35)   Shevell, M. I.; Rosenblatt, D. S. The Neurology of Cobalamin. *Can. J. Neurol. Sci. J. Can. Sci. Neurol.* **1992**, *19* (4), 472–486. https://doi.org/10.1017/S0317167100041676.

(36)   Hsu, J. M.; Kawin, B.; Minor, P.; Mitchell, J. A. Vitamin B12 Concentrations in Human Tissues. *Nature* **1966**, *210* (5042), 1264–1265. https://doi.org/10.1038/2101264a0.

(37)   Van Doorslaer, S.; Jeschke, G.; Epel, B.; Goldfarb, D.; Eichel, R.-A.; Kräutler, B.; Schweiger, A. Axial Solvent Coordination in "Base-Off" Cob(II)Alamin and Related Co(II)-Corrinates Revealed by 2D-EPR. *J. Am. Chem. Soc.* **2003**, *125* (19), 5915–5927. https://doi.org/10.1021/ja021218j.

(38)   St. Maurice, M.; Mera, P.; Park, K.; Brunold, T. C.; Escalante-Semerena, J. C.; Rayment, I. Structural Characterization of a Human-Type Corrinoid Adenosyltransferase Confirms That Coenzyme B $_{12}$ Is Synthesized through a Four-Coordinate Intermediate. *Biochemistry* **2008**, *47* (21), 5755–5766. https://doi.org/10.1021/bi800132d.

(39)   Stich, T. A.; Yamanishi, M.; Banerjee, R.; Brunold, T. C. Spectroscopic Evidence for the Formation of a Four-Coordinate Co $^{2+}$ Cobalamin Species upon Binding to the Human ATP:Cobalamin Adenosyltransferase. *J. Am. Chem. Soc.* **2005**, *127* (21), 7660–7661. https://doi.org/10.1021/ja050546r.

(40)   *2-D Proteome Analysis Protocols*; Link, A. J., Ed.; Methods in molecular biology; Humana Press: Totowa, NJ, 1999.

(41)   Craean, V. F. Structure, Function and Metabolic Roles of IcmF-a Fusion Between the Radical B12 Enzyme and Its G-Protein Chaperone, PhD Thesis, University of Michigan, 2012.

(42)   Mancia, F.; Evans, P. R. Conformational Changes on Substrate Binding to Methylmalonyl CoA Mutase and New Insights into the Free Radical Mechanism.

(43)   Fenton, W. A.; Hack, A. M.; Willard, H. F.; Gertler, A.; Rosenberg, L. E. Purification and Properties of Methylmalonyl Coenzyme A Mutase from Human Liver. *Arch. Biochem. Biophys.* **1982**, *214* (2), 815–823. https://doi.org/10.1016/0003-9861(82)90088-1.

(44)   Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50* (W1), W276–W279. https://doi.org/10.1093/nar/gkac240.