Combined Spectroscopic and Computational Studies of Free and Enzyme-Bound B_{12} Species

By

Laura D. Elmendorf

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2023

Date of final oral examination: 15 December 2023

The dissertation is approved by the following members of the Final Oral Committee:

Thomas C. Brunold, Professor, Inorganic Chemistry
Judith N. Burstyn, Professor, Chemical Biology
Daniel C. Fredrickson, Professor, Inorganic Chemistry
Brian G. Fox, Professor, Biochemistry

Combined Spectroscopic and Computational Studies of Free and Enzyme-Bound B_{12} Species

Laura D. Elmendorf

Under the supervision of Professor Thomas C. Brunold at the University of Wisconsin-Madison

Abstract

Vitamin B_{12} and its derivatives, known collectively as cobalamins (Cbls) or the B_{12} family, play an essential role in catalyzing a diverse range of reactions across all areas of life. Understanding the versatile reactivity of these cofactors requires a detailed understanding of their geometric and electronic structures, as well as how these change upon enzyme binding and the addition of substrates. Spectroscopic techniques like electronic absorption, resonance Raman (rRaman), and magnetic circular dichroism spectroscopies are invaluable in this area, particularly when combined with computational techniques such as density functional theory that can bridge the gap between spectral features and structural implications. This dissertation presents a collection of projects that used these methods to delve into the properties and reactivity of free and enzyme-bound B_{12} species, as well as the mononuclear non-heme iron enzyme cysteamine dioxygenase (ADO). This work includes a study of vibronic coupling in vitamin B_{12}, which afforded new insight into the choice of functional for quantum mechanical modeling of B_{12} and enabled the identification of the elusive Co–C stretching mode in this molecule's rRaman spectrum. It also includes an investigation of Co–C bond activation in ethanolamine ammonia-lyase (EAL), an adenosylcobalamin (AdoCbl)-dependent enzyme that catalyzes the conversion of ethanolamine to acetaldehyde and an ammonium ion. As with other AdoCbl-dependent enzymes, EAL's catalytic cycle begins with homolytic cleavage of the cofactor's Co–C bond, and the rate of homolysis is increased by 12 orders of magnitude when the cofactor is enzyme-bound in the presence of

substrate. We identified conformational changes and cofactor-protein electrostatic interactions that occur upon substrate binding as key factors for promoting controlled and dramatically accelerated Co–C bond homolysis. Finally, two computational studies of PduO-type adenosyltransferases (ATRs), one with the native cobalamin substrate and the other with a rhodium-substituted analogue, explored key steps of the enzyme's catalytic cycle.

Acknowledgements

It is difficult to convey in words how thankful I am for the many people who made this dissertation possible, but I am going to do my best here.

To my advisor, Thomas Brunold, I truly cannot thank you enough for your unwavering support over the last five and a half years. I have learned so much from you, and your guidance and expertise helped me grow both as a scientist and a person. I am really going to miss my time in your group! To Judith Burstyn, thank you for the years of excellent advice and for always asking the best questions during group meetings. Thank you also to the other members of my committee, Daniel Fredrickson and Brian Fox, for your time and feedback.

To the rest of the Brunold and Burstyn groups, I am eternally grateful for your friendship. Rebeca, Elizabeth, and Mickie, you are all such an inspiration to me, and you have been fantastic mentors. Joshua, your company was such an important part of my time in graduate school, both in the lab and outside of it, and I look forward to many years of friendship yet to come. Ryan, I am so glad that you joined team B_{12}; I cannot imagine tackling the EAL project without you. Becky, Zach, and Cristina, you made the office a genuinely fun place to be. Madeleine, I have always admired your hard work and your dedication to improving the world around you. Thank you all for creating such a wonderful environment to learn and work in.

To my other friends, I love you all, and I am so thankful for your presence in my life. To Megan and Kyana, thank you for being the world's best roommates! You are like family to me, and I am so excited to cheer you on as you finish your graduate work. To Peyton, Brady, Joe, Cassie, Amber, Danica, Aislynn, Helen, and all of the many people who brightened my days in Madison and elsewhere, I have absolutely loved the time we have spent together. Thank you also to the wonderful people I met at Williams College who helped shape the scientist I am today. To

Patrick Barber, thank you for encouraging my early foray into chemistry research and for instilling in me a lifetime fascination with lanthanides. To Sarah Goh and the Ford Lab Beta Team, the semester I spent working with you was hugely influential in leading me down the path of chemistry research, so thank you for making chemistry so fun!

To my family, thank you so much for your endless love and support. Thomas and Mih-Ho (and yes, I am including you in the family section!), I am so happy to have you in my life. Aunt Heidi, thank you for always encouraging my love of science. To the rest of my extended family, including Clare, Uncle Randy, Aunt Nancy, Scout, and Will, I miss you all and hope to see you more often in the coming years. To my parents and my sister Caroline, your encouragement means the world to me. Thank you for listening to me obsess over chemistry for so many years! Pips, it has been incredibly cool watching everything you have accomplished, and I am very proud of you. I could not have done any of this without you. Love you all!

Table of Contents

Chapter 1

Electronic Structure Studies of Free and Enzyme-Bound B_{12} Species by Magnetic Circular Dichroism and Complementary Spectroscopic Techniques

This chapter has been published under the following: Elmendorf, L. D.; Brunold, T. C. Electronic Structure Studies of Free and Enzyme-Bound B_{12} Species by Magnetic Circular Dichroism and Complementary Spectroscopic Techniques. In *Methods in Enzymology*; Academic Press, **2022**; *669*, 333–365. https://doi.org/10.1016/bs.mie.2022.02.002.

Chapter 1: Electronic Structure Studies of Free and Enzyme-Bound B_{12} Species by Magnetic Circular Dichroism and Complementary Spectroscopic Techniques

1.1. Overview

The rich history of electronic absorption (Abs) and circular dichroism (CD) spectroscopic studies of B_{12} species goes back over half a century, and by the 1960s much of the nomenclature used to describe characteristic B_{12} spectral features had been established.[1,2] Since then, magnetic circular dichroism (MCD) spectroscopy and other complementary techniques have provided an increasingly detailed understanding of the electronic structure of cobalamins. CD spectroscopy measures the difference in the absorption of left- and right-circularly polarized light.[3] MCD spectroscopy adds the application of a magnetic field parallel to the direction of light propagation. MCD spectroscopy is an extremely powerful tool for studying enzymes and cofactors containing transition metal ions; it does not require isotope enrichment, works with paramagnetic and diamagnetic species, probes both the ground and excited states, and, because all electronic transitions in the visible/near-IR region originate from the metalloenzyme active site or metallocofactor, is highly site selective.[3] Additionally, the intensities of the electronic transitions dominating CD and MCD spectra are governed by different selection rules than those dominating Abs spectra, often allowing for electronic transitions that are too weak or too close in energy to be resolved in Abs spectra to be probed using these complementary methods.[3–5]

For an electronic transition from the ground state to an excited state, the Abs intensity is proportional to the square of the transition moment integral:[4]

$$I^{ABS} \propto \langle \Psi_{GS} | \hat{\mu} | \Psi_{ES} \rangle^2$$

Ψ_{GS} and Ψ_{ES} are the ground and excited state wavefunctions, respectively, and $\hat{\mu}$ is the transition moment operator. This can be broken down into two terms of significance, electric dipole intensity (I^{ED}) and magnetic dipole intensity (I^{MD}):[3]

$$I^{ED} \propto \langle \Psi_{GS} | \hat{\mu}_e | \Psi_{ES} \rangle^2$$

$$I^{MD} \propto \langle \Psi_{GS} | \hat{\mu}_m | \Psi_{ES} \rangle^2$$

$\hat{\mu}_e$ and $\hat{\mu}_m$ are the electric dipole and magnetic dipole moment operators, respectively. Electric dipole transitions are usually several orders of magnitude more intense than magnetic dipole transitions, so these dominate Abs spectra. Evaluating the transition moment integral for electric dipole transitions leads to the following selection rules: (i) transitions between states of different spin are forbidden (spin selection rule) and (ii) transitions between wavefunctions of the same parity are forbidden, i.e. only g $\rightarrow$ u and u $\rightarrow$ g transitions are allowed (orbital or Laporte selection rule). Importantly, ligand field (LF or d $\rightarrow$ d) transitions are parity forbidden g $\rightarrow$ g excitations and thus have low Abs intensity.[3]

The CD signal intensity is governed by the rotational strength (R), which is proportional to the imaginary (Im) part of the dot product of the electric and magnetic dipole transition moments:[3]

$$R \propto \mathrm{Im}\{\langle \Psi_{GS} | \hat{\mu}_e | \Psi_{ES} \rangle \cdot \langle \Psi_{GS} | \hat{\mu}_m | \Psi_{ES} \rangle\}$$

Obtaining rotational strength therefore requires transitions that are both electric and magnetic dipole allowed, and the corresponding transition dipole moments must be colinear. These requirements are only fulfilled by certain low-symmetry point groups in which the electric dipole and magnetic dipole operators transform as the same irreducible representation. CD signal therefore requires a lack of symmetry of the molecule investigated.

While the presence of a CD signal requires molecular chirality (more precisely, the lack of an improper rotation axis), an MCD signal is observed for all matter as a consequence of the

Faraday effect.[3] Transitions that are formally forbidden according to the Abs and CD selection rules, such as LF transitions, can gain MCD intensity through spin-orbit coupling. Three different terms contribute to the MCD signal intensity, which is measured as the difference in extinction coefficients for left-circularly polarized (LCP) and right-circularly polarized (RCP) light, $\Delta\varepsilon = \varepsilon_{\mathrm{LCP}} - \varepsilon_{\mathrm{RCP}}$:[3,6]

$$\frac{\Delta\varepsilon}{\mathrm{E}} = \frac{\varepsilon_{\mathrm{LCP}} - \varepsilon_{\mathrm{RCP}}}{\mathrm{E}} \propto \mathrm{B}\left\{-\mathrm{A}_1\left(\frac{\partial f(\mathrm{E})}{\partial \mathrm{E}}\right) + \left(\mathrm{B}_0 + \frac{\mathrm{C}_0}{k\mathrm{T}}\right) f(\mathrm{E})\right\}$$

where E is the energy of the incident radiation, B is the magnetic flux density, f(E) is the absorption bandshape, k is the Boltzmann constant, T is the temperature, and the A_1, B_0, and C_0 terms depend on the transition under study. The magnetic field results in a Zeeman splitting of orbitally degenerate excited states (A_1, or the "A-term"), field-induced mixing of all electronic states (B_0, or "B-term"), and a Zeeman splitting of the ground state of paramagnetic species, causing a population difference among the spin sublevels (C_0, or "C-term"). The temperature-dependent C-term contributions dominate low-temperature MCD spectra of paramagnetic species.

MCD spectroscopy is particularly effective when used in combination with other spectroscopic techniques and computational tools, such as electron paramagnetic resonance (EPR) spectroscopy and density functional theory calculations, both discussed in other chapters. Another technique that nicely complements MCD spectroscopy is resonance Raman (rR) spectroscopy, which involves laser excitation in the visible or near-UV spectral region to enhance vibrational modes along which the chromophore distorts when undergoing an electronic transition in resonance with the laser wavelength. Because the intense bands in the visible region of the Abs spectra of B_{12} species arise from corrin $\pi \rightarrow \pi*$ transitions, rR spectroscopy provides an excellent probe of conformational differences of the corrin ring of cobalamins in different enzyme environments.[7–14] The information it yields on vibronic coupling and the polarization of electronic

transitions has been extremely helpful for assigning electronic transitions associated with key Abs, CD, and MCD spectral features displayed by B_{12} species in all three accessible Co oxidation states.[15–19]

This chapter summarizes representative MCD studies of free and enzyme-bound B_{12} species. Complementary spectroscopic and computational data are also presented and discussed where appropriate.

1.2. Spectroscopic Studies of Free B_{12} Species

1.2.1. Co^{III}orrinoids

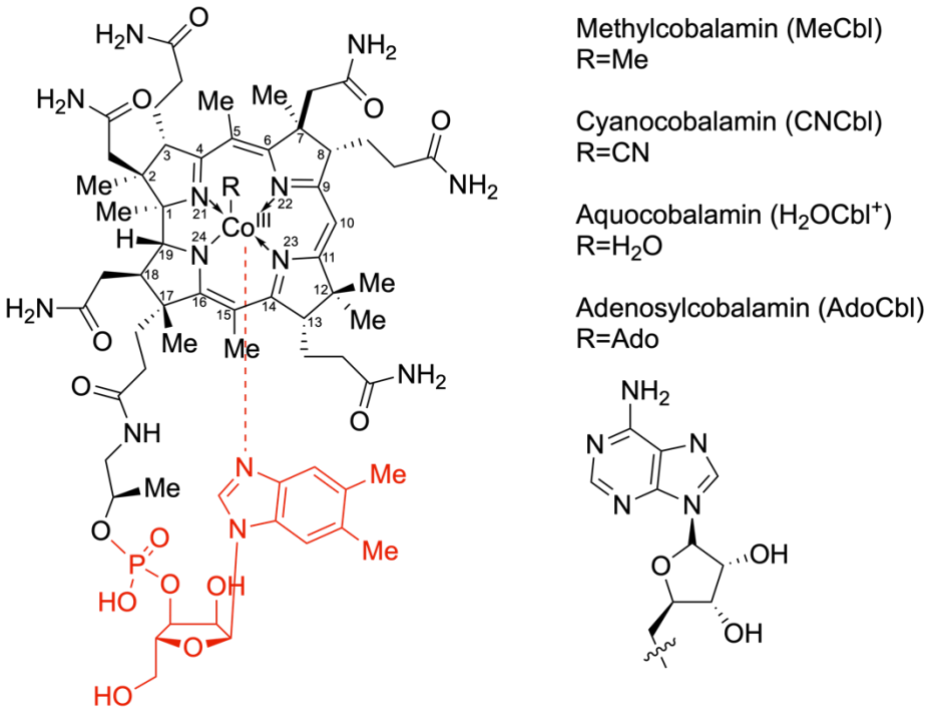


Figure 1.1. Chemical structure of Co^{III}Cbls, where **R** indicates the upper axial ligand. Relevant atom numbers are included. Note that in Co^{III}cobinamides, the nucleotide loop (shown in red) including the coordinating dimethylbenzimidazole (DMB) is absent and a water molecule occupies the lower axial position.

The Abs, CD, and MCD spectra of Co^{III}Cobalamins (Co^{III}Cbls, Figure 1.1), which are six-coordinate, low spin (S=0) $3d^6$ species, are dominated by corrin $\pi \rightarrow \pi^*$ transitions. Because they

are diamagnetic, Co^{III}Cbls lack the intense, temperature-dependent C-term MCD features observed for paramagnetic species. Nonetheless, their MCD spectra are still quite informative. Electronic communication between the axial ligands and corrin macrocycle causes the Abs, CD, and MCD spectra of Co^{III}Cbls to be sensitive to the axial coordination environment.[15] The identity of the upper axial ligand is particularly impactful; early spectroscopic studies led to a distinction between "typical" Abs spectra displayed by the majority of Co^{III}Cbls, such as CNCbl and H_2OCbl^+, and "unique" spectra exhibited primarily by alkylcobalamins like AdoCbl and MeCbl.[1]

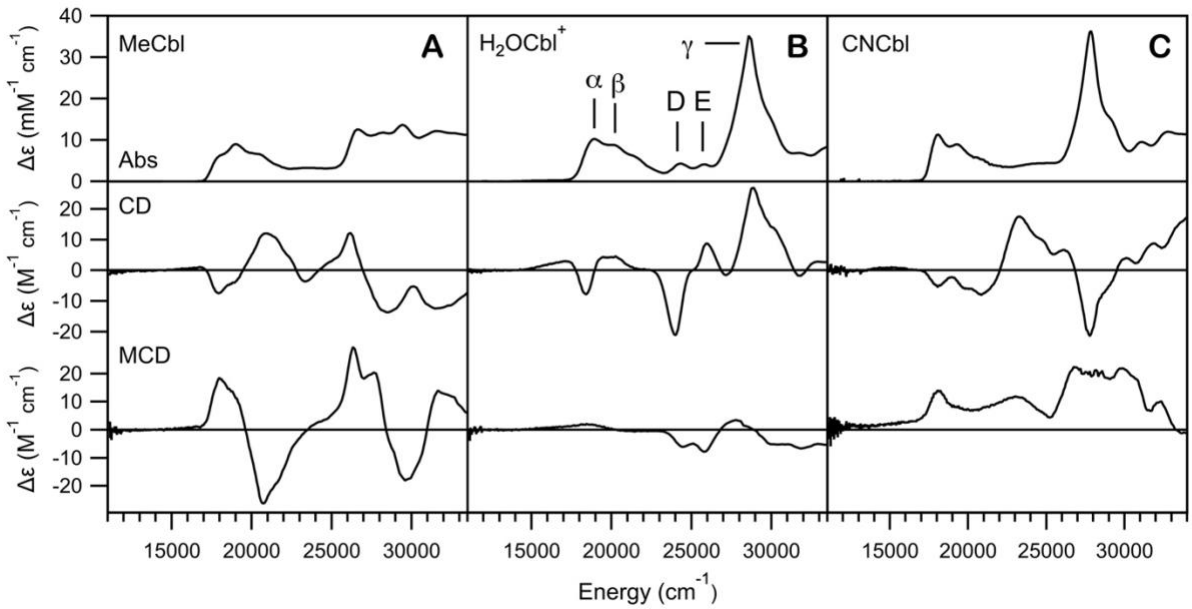


Figure 1.2. Absorption (top), CD (center), and 7 T MCD (bottom) spectra recorded at 4.5 K of (A) MeCbl, (B) H_2OCbl^+, and (C) CNCbl. Band designations are given for H_2OCbl^+. Adapted from ref. 15.

Co^{III}Cbl Abs spectra can be divided into three regions: (i) the α/β bands at lowest energy in the visible region, (ii) the D/E bands toward higher energy, and (iii) the γ band in the UV region.[15,20] The α band was definitively assigned as a corrin $\pi \rightarrow \pi^*$ transition polarized along the long corrin axis (the $C^5 \cdots C^{15}$ vector in Figure 1.1), while the identity of the nearby β band – attributed to a vibronic sideband of the α transition or the origin of a separate electronic transition – has been a matter of debate.[20,21] Time-dependent density functional theory (TD-DFT)

computations failed to definitively settle this question, as the choice of functional greatly impacted the number of electronic transitions predicted in this region.[22–24] However, recent experimental studies using triply resonant sum frequency (TRSF) spectroscopy[25] and ultrafast Abs and fluorescence spectroscopy[26] confirmed that the β feature in "typical" Abs spectra arises from a vibrational progression rather than a distinct electronic transition. This assignment is also supported by the presence of a single weak feature coinciding with the α band in the MCD spectra of CNCbl and H_2OCbl^+.[15] While the α/β peaks in the "unique" Abs spectra observed for MeCbl and AdoCbl were also attributed to the origin and vibrational sideband of a single corrin $\pi \rightarrow \pi^*$ transition,[15] at least one additional electronic transition must be present in close energetic proximity, as the corresponding MCD spectra exhibit two oppositely signed features in this region (Figure 1.2). Notably, "typical" and "unique" Abs spectra differ with respect to the energy of the α/β bands, which red-shift as the σ-donor strength of the upper axial ligand increases from H_2OCbl^+ to CNCbl to MeCbl (Figure 1.2).

The D/E and γ regions of typical Co^{III}Cbl Abs spectra further demonstrate the power of CD and MCD spectroscopy to resolve overlapping Abs features. The two prominent bands in the D/E region of the Abs spectrum of H_2OCbl^+ arise from two separate electronic transitions, as their spacing (~1,500 cm^{-1}) exceeds the energy of the corrin-based vibrational modes coupled to these electronic transitions.[15] The presence of two intense CD features offset from those bands indicates that at least four transitions actually contribute to this region. TD-DFT calculations revealed that the most intense Abs peak in the D/E region can be attributed to a corrin HOMO-1 → LUMO ($\pi \rightarrow \pi^*$) transition.[15] Meanwhile, the features in the γ region are particularly distinctive for different Co^{III}Cbls. In "typical" Abs spectra, the γ region is dominated by a single intense band arising from a corrin $\pi \rightarrow \pi^*$ transition polarized along the short corrin axis (the $Co\cdots C^{10}$ vector in Figure

1.1).[21] In "unique" Abs spectra, the intensity of the γ band is redistributed over several weaker features of similar intensity in the UV spectral region.

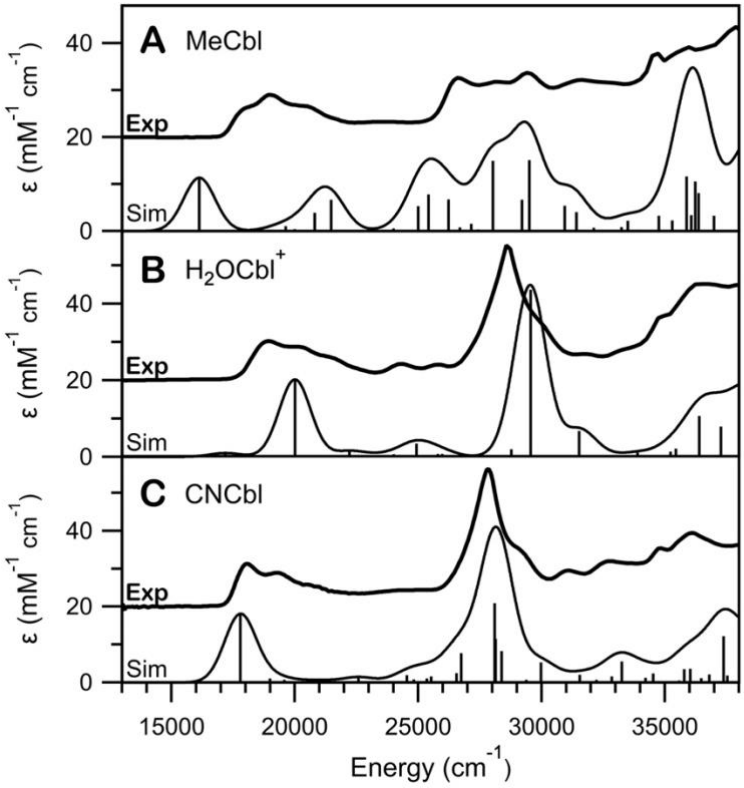


Figure 1.3. Experimental and TD-DFT computed Abs spectra for (A) MeCbl, (B) H_2OCbl^+, and (C) CNCbl. The computed spectra were uniformly red shifted by 5500 cm^{-1} to facilitate a comparison with the experimental data. Adapted from ref. 15.

The remarkable differences observed between the Abs, CD, and MCD spectra of different Co^{III}Cbl species are the result of subtle but important contributions of the axial ligands to the corrin-based frontier orbitals. TD-DFT computations, which accurately reproduced the key experimental features and trends (Figure 1.3), revealed that the distinction between "typical" and "unique" Abs spectra can be understood in terms of differences in σ-donation from the upper axial ligand to the Co^{III} ion.[15] With increasing σ-donor strength of the upper axial ligand, the contribution from the formally unoccupied Co $3d_{z2}$ orbital to the HOMO increases, which induces a stronger σ-antibonding interaction between the Co^{III} ion and lower axial base. This raises the energy of the

HOMO and red-shifts the HOMO → LUMO transition. Increased σ-donation also decreases the effective nuclear charge of the Co^{III} ion, raising the relative energies of the filled Co 3d orbitals and thus leading to more extensive mixing of these orbitals with the corrin π-based frontier orbitals. As a result, the number of filled MOs directly beneath the HOMO increases, causing the appearance of multiple, similarly intense features in the γ region of "unique" Abs spectra.[15]

Substitution of the lower axial DMB ligand also has a significant effect on the Abs, CD, and MCD spectra. This is nicely illustrated by Co^{III}cobinamides ($Co^{III}Cbi^{+}$s), natural precursors of cobalamin species that lack the nucleotide loop including the dimethylbenzimidazole (DMB) moiety and instead bind H_2O as the lower axial ligand in aqueous solution, which makes them ideal models of "base-off" Co^{III}Cbls at neutral pH.[15] In the Abs spectrum of $AdoCbi^{+}$, the α band is blue-shifted by ~3,000 cm^{-1} from its counterpart in the AdoCbl Abs spectrum (Figure 1.4).[27] This shift was attributed to the elimination of the moderate Co–N(DMB) σ* interaction in the HOMO upon replacement of the DMB ligand with a more weakly σ-donating water molecule, which results in a sizeable stabilization of the donor MO for this transition.[15]

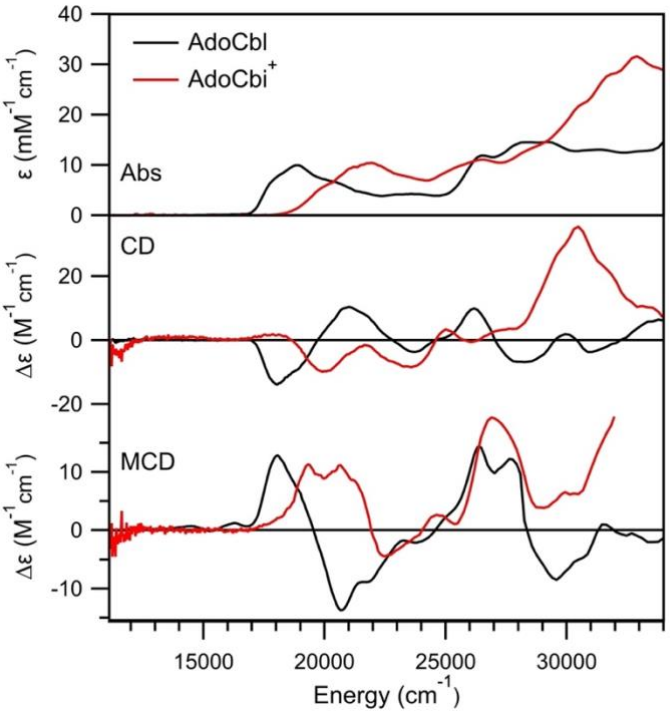


Figure 1.4. Comparison of the 4.5 K Abs (top), CD (center), and 7 T MCD (bottom) spectra of AdoCbl and $AdoCbi^+$. Adapted from ref. 15.

1.2.2. Co^{II}Corrinoids

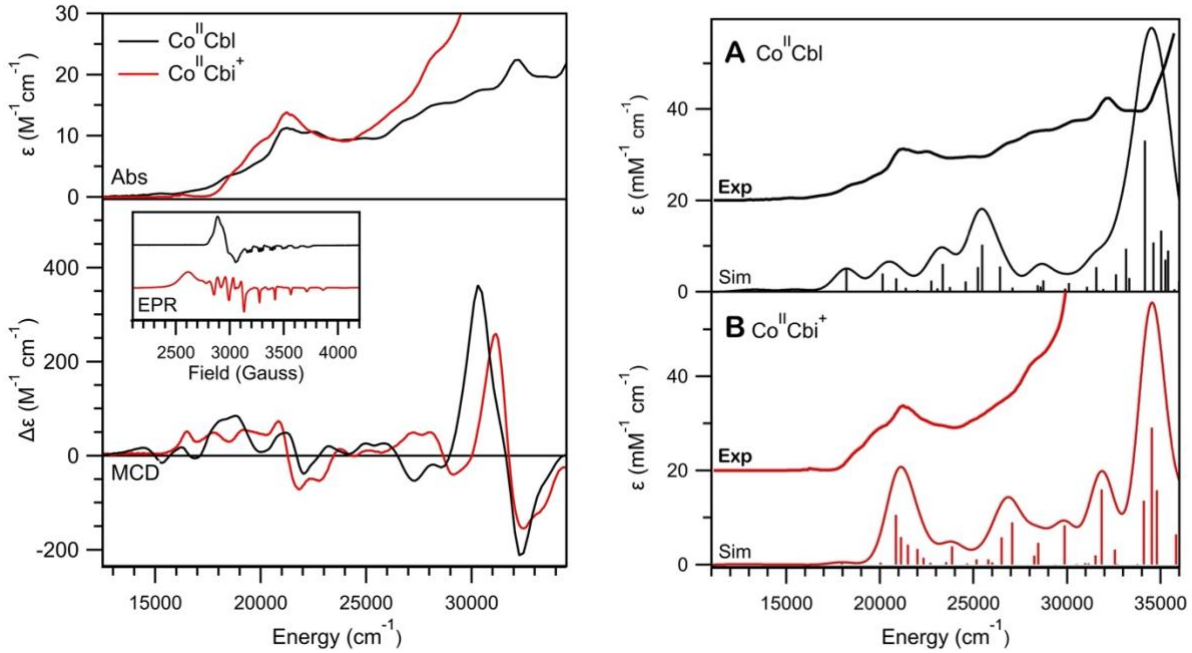


Figure 1.5. Left: Abs and 7 T MCD spectra at 4.5 K of $Co^{II}Cbl$ (black) and $Co^{II}Cbi^+$ (red). The corresponding X-band EPR spectra at 40 K are shown in the inset. Right: Experimental 4.5 K and TD-DFT-computed Abs spectra of (A) $Co^{II}Cbl$ and (B) $Co^{II}Cbi^+$. Adapted from ref. 16.

Co^{II}corrinoids can formed in several different ways, such as via (i) homolytic cleavage of the Co–C bond of alkylcobalamins, as is the case in the catalytic cycles AdoCbl-dependent enzymes, (ii) one-electron reduction of a Co^{III}corrinoid species, and (iii) oxidation of Co^{I}Cbl, chemically or accidentally in the catalytic cycles of MeCbl-dependent enzymes.[16] Co^{II}corrinoids are low-spin (S = 1/2) $3d^7$ species that are usually five-coordinate, lacking an upper axial ligand. Because of their paramagnetic ground states, Co^{III}corrinoids are excellent targets for MCD studies. Distinct transitions dominate the different regions of the Abs, CD, and MCD spectra of Co^{II}corrinoids (Figure 1.5): d $\rightarrow$ d transitions below ~17,000, metal-to-ligand charge transfer (MLCT) transitions from ~17,000 to 20,000 cm^{-1}, and corrin $\pi \rightarrow \pi^*$ transition above ~20,000 cm^{-1}.[28] The Co d $\rightarrow$ d transitions are parity forbidden and therefore contribute minimally to the Abs spectra but gain significant MCD intensity through spin-orbit coupling. The opposite effect is observed for corrin $\pi \rightarrow \pi^*$ transitions; e.g., the transition responsible for the dominant Abs band in the visible region with a maximum at ~21,000 cm^{-1} carries very little intensity in the MCD spectrum, a testament to the different selection rules for these techniques.[14,16,19]

The prominent feature at ~21,000 cm^{-1} in the Co^{II}Cbl Abs spectrum originates from a corrin $\pi \rightarrow \pi^*$ transition involving the same donor and acceptor orbitals as the transition associated with the α/β bands of Co^{III}Cbls. However, because the donor MO for this transition contains less Co 3d character in Co^{II}Cbl than in Co^{III}Cbl, the transition energy is less affected by changes in the lower axial ligand.[16] As a result, the Abs spectra of Co^{II}Cbl and $Co^{II}Cbi^+$ are quite similar. A much greater difference is observed, however, for the LF transitions, only discernable in the MCD spectra. The weaker σ-donation of H_2O compared to the DMB ligand leads to a sizable stabilization of the occupied Co $3d_{z2}$-based MO, causing a significant blue-shift of the LF transitions originating from this orbital. MCD spectroscopy is therefore much better suited than Abs spectroscopy for studying

the axial environment of Co^{II}Cbls. EPR spectroscopy complements MCD spectroscopy particularly well in this context, as it is also sensitive to the relative energies of the Co 3d-based MOs and the composition of the singly occupied, Co $3d_{z2}$-based MO (Figure 1.5, inset).

1.2.3. Co^{I}Corrinoids

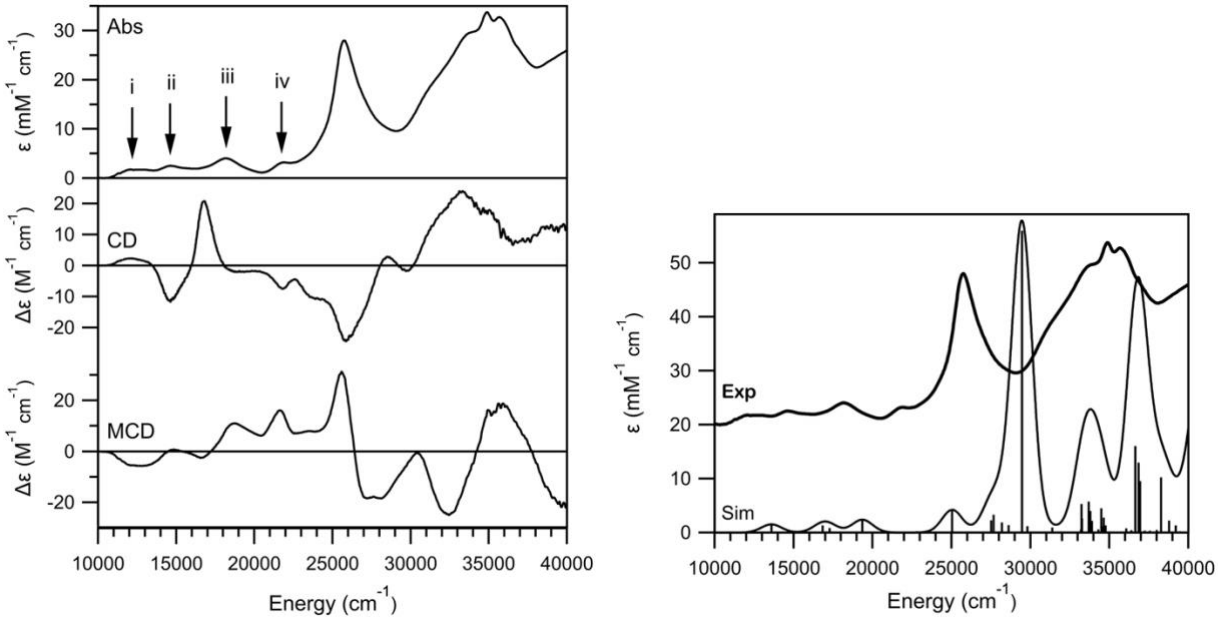


Figure 1.6. Left: Absorption (top), CD (center), and 7 T MCD (bottom) spectra recorded at 4.5 K of Co^{I}Cbl. The four low-energy Abs features in the near-IR/visible region are labelled i – iv. Right: Experimental 4.5 K and TD-DFT-computed Abs spectra of Co^{I}Cbl. Adapted from ref. 17.

Co^{I}Cbl, the "superreduced" form of B_{12}, is formed (i) via reduction of Co^{II}Cbl (e.g. using Zn dust), (ii) as a part of the catalytic cycles of MeCbl-dependent enzymes and reductive dehalogenases, and (iii) during AdoCbl biosynthesis.[17] Co^{I}Cbl is a four-coordinate, low-spin (S=0) $3d^8$ species in which the Co^{I} ion is ligated solely by the corrin macrocycle; consequently, Co^{I}Cbl and Co^{I}cobinamide are indistinguishable by Abs, CD, and MCD spectroscopy. As in the case of Co^{III}Cbls, the diamagnetic ground state of Co^{I}Cbl results in a temperature-independent and relatively weak, but nonetheless quite well-resolved MCD spectrum (Figure 1.6).[17] The Abs spectrum of Co^{I}Cbl contains four moderately intense bands in the near-IR/visible region (at ~12,500, 15,000, 18,000, and 22,000 cm^{-1}) and a much more intense feature in the UV region (at

~26,000 cm^{-1}). TD-DFT computations revealed that the HOMO of Co^{I}Cbl has primarily Co $3d_{z2}$ orbital character, while the three lowest unoccupied MOs are corrin π*-based.[17] As a result, the four low-energy transitions observed in the near-IR/visible regions of the Co^{I}Cbl Abs spectrum (bands i – iv in Figure 1.6) formally arise from MLCT transitions, though they also possess significant $\pi \rightarrow \pi^*$ character due to the mixing between the Co 3d and corrin π/π^* frontier orbitals. Toward higher energy, the dominant Abs feature at ~26,000 cm^{-1} arises from a similar corrin $\pi \rightarrow \pi^*$ transition as the one responsible for the α/β bands in the Co^{III}Cbl Abs spectra. Interestingly, the energy of this transition increases from MeCbl to Co^{II}Cbl and Co^{I}Cbl. Computational studies revealed that the blue-shift of this transition with decreasing formal oxidation state of the Co ion can be attributed to an increase in π-backbonding from the filled Co 3d orbitals to the empty corrin π*-based acceptor MO and the consequent destabilization of this orbital.[17]

The different binding modes, Co oxidation states, and axial ligand environments are key to the distinct reactivities displayed by different B_{12} species, and MCD spectroscopy provides a uniquely sensitive probe of these characteristics. The detailed understanding of the correlation that exists between structural perturbations and spectral changes provided the necessary foundation for carrying out Abs, CD, and MCD spectroscopic studies of corrinoids in various enzyme active sites. The following section outlines examples of how this approach was used to obtain molecular-level insight into the reaction mechanisms of enzymes involved in the biosynthesis of B_{12} and B_{12}-dependent enzymes.

1.3. Application of MCD Spectroscopy to the Study of Enzyme-Bound B_{12} Species

1.3.1. ATP:Co^{I}Corrinoid Adenosyltransferases (ACATs)

Because AdoCbl-dependent enzymes are susceptible to inactivation via oxidation of the Co^{II}Cbl intermediate generated during turnover,[29] all organisms that employ these enzymes in their

metabolism are believed to produce ATP:Co^{I}corrinoid adenosyltransferases (ACATs).[30] ACATs catalyze the replacement of the upper axial ligand of exogenous cobalamins or incomplete corrinoids with the 5'-deoxyadenosyl (Ado) group of ATP.[30,31] Three non-homologous families of ACATs have been identified and classified according to their biological roles in *Salmonella enterica*.[31] The CobA ACAT is involved in the *de novo* synthesis of AdoCbl and in the scavenging of exogenous corrinoids, whereas the PduO and EutT ACATs are required for the growth of *S. enterica* on 1,2-propanediol and ethanolamine (via supply of AdoCbl to diol dehydratase and ethanolamine ammonia-lyase, respectively).[32–35]

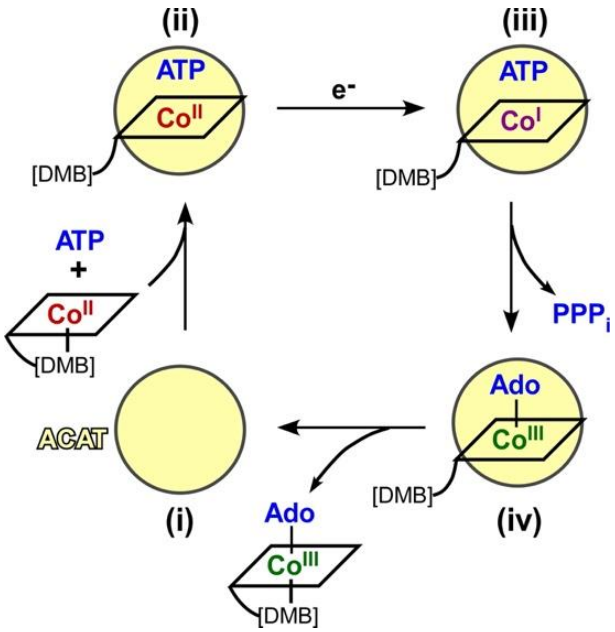


Figure 1.7. Generalized reaction cycle of ACATs. Binding of ATP and Co^{II}Cbl to substrate-free enzyme (i) yields an unusual four-coordinate Co^{II}Cbl species (ii) that is activated for reduction to the Co^{I} state (iii). This strongly nucleophilic Co^{I}Cbl intermediate then attacks the 5'-carbon of ATP to form the Co–C bond of AdoCbl (iv) and triphosphate (PPP$_i$). Release of AdoCbl product or transfer to the respective acceptor enzyme completes the catalytic cycle.

The general mechanism proposed for ACATs involves the formation of a Co^{I}corrinoid intermediate that performs a nucleophilic attack on the 5'-carbon of co-substrate ATP (Figure 1.7).[11,30,31,36,37] A particularly intriguing step in this mechanism is the thermodynamically challenging Co^{II} → Co^{I}corrinoid reduction (step ii → iii) that must precede the adenosylation (step

iii → iv), given that the reduction potentials of free Co^{II}corrinoids (e.g., E^0 vs NHE of –610 mV for Co^{II}Cbl and –490 mV for $Co^{II}Cbi^+$)[38] are well below those of readily available reducing agents in the cell. However, when bound to an ACAT active site in the presence of ATP, Co^{II}corrinoids can be reduced by free dihydroflavins (E^0 = –228 mV for FMN at pH 7.5).[32,36]

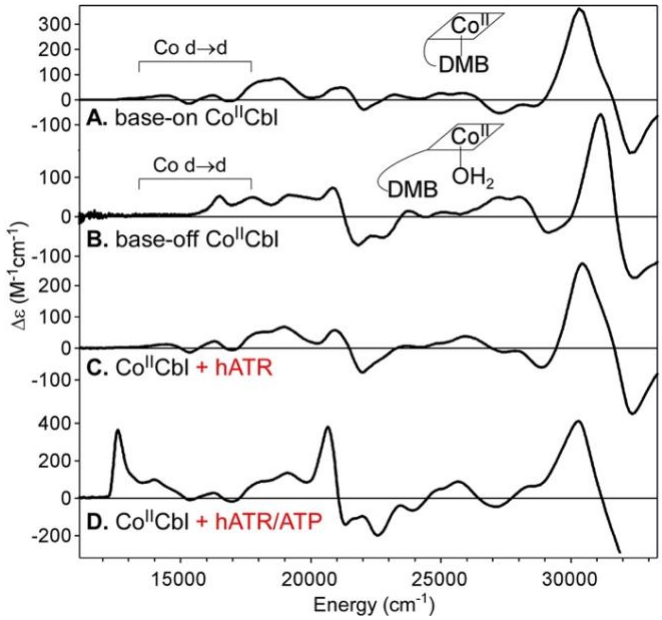


Figure 1.8. MCD spectra at 4.5 K and 7 T of (A) base-on Co^{II}Cbl, (B) base-off Co^{II}Cbl, (C) Co^{II}Cbl in the presence of *h*ATR, and (D) Co^{II}Cbl in the presence of *h*ATR and ATP. Adapted from ref. 39.

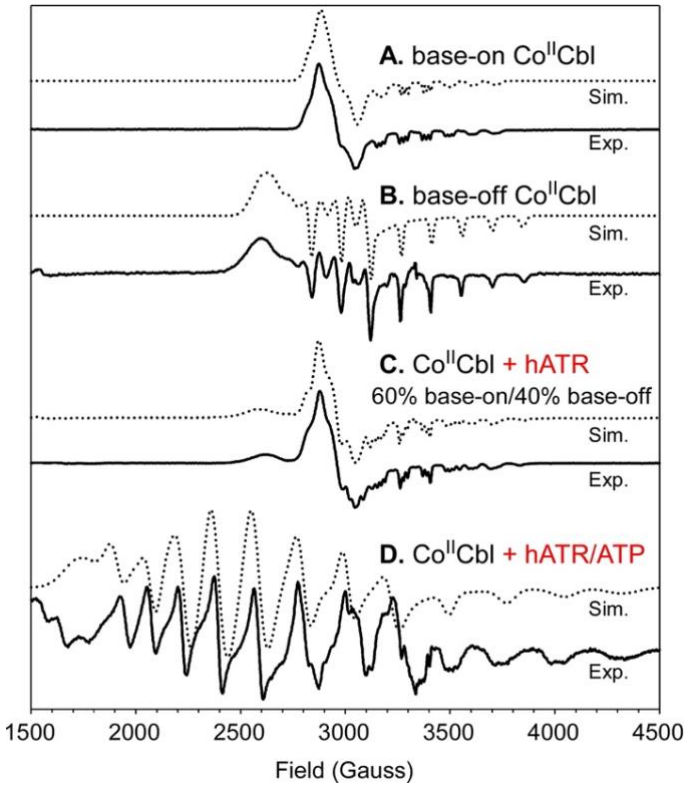


Figure 1.9. X-band EPR spectra at 40 K of (A) base-on Co^{II}Cbl, (B) base-off Co^{II}Cbl, (C) Co^{II}Cbl in the presence of *h*ATR, and (D) Co^{II}Cbl in the presence of *h*ATR and ATP. Simulated spectra (dotted lines) are shown above the experimental spectra (solid lines). Adapted from ref. 39.

Our initial ACAT target was the PduO-type human adenosyltransferase *h*ATR.[39] The MCD spectrum of Co^{II}Cbl bound to *h*ATR in the absence of co-substrate ATP was found to exhibit only small differences from that of free Co^{II}Cbl (cf. Figure 1.8C and 1.8A). Likewise, our X-band EPR spectra of free and *h*ATR-bound Co^{II}Cbl were similar (cf. Figure 1.9C and 1.9A), though a quantitative analysis indicated a sizeable (~40%) contribution from base-off Co^{II}Cbl (Figure 1.9B) to the protein spectrum. Much more dramatic changes were observed upon Co^{II}Cbl binding to *h*ATR complexed with ATP (Figure 1.8D). Most noticeably, the lowest-energy transition exhibited a large red-shift and a substantial increase in intensity, signaling the formation of an unprecedented Co^{II}Cbl species. Similarly, a very unusual EPR spectrum was obtained for Co^{II}Cbl bound to the *h*ATR/ATP complex (Figure 1.9D). A fit of the EPR spectrum associated with the newly formed Co^{II}Cbl species revealed unusually large $g_{\perp}$ and ^{59}Co hyperfine values. On the basis of these

results, we concluded that *h*ATR accomplishes the thermodynamically challenging $Co^{II} \rightarrow Co^{I}$Cbl reduction by generating an effectively 4-coordinate (4c), square-planar Co^{II} species that lacks any significant axial bonding interactions. Formation of this unusual species only in the presence of co-substrate ATP suggested that the enzyme exerts exquisite control over the timing of Co^{II}Cbl reduction, thereby protecting itself from deleterious side-reactions by the transiently formed Co^{I}Cbl "supernucleophile".[40]

Even though no evolutionary relationship exists among the CobA, PduO, and EutT families of ACATs,[31] spectroscopic studies of CobA from *Salmonella enterica* (*Se*CobA),[41] the PduO-type ACAT from *Lactobacillus reuteri* (*Lr*PduO),[11,12,14] and the EutT-type ACAT from *S. enterica* (*Se*EutT)[13,42] provided compelling evidence that all three families use the same basic mechanism for the reduction of their Co^{II}corrinoid substrates, as outlined above for *h*ATR. However, our results also revealed some intriguing mechanistic differences among the ACAT families, especially with regards to the order of substrate binding and the effectiveness of the conversion of Co^{II}Cbl and $Co^{II}Cbi^{+}$ to 4c species. While *Se*CobA and *Lr*PduO achieve higher 5c $\rightarrow$ 4c conversion yields with $Co^{II}Cbi^{+}$,[12,43] which is expected given the fact that the Co–OH_2 bond is considerably weaker than the Co–N(DMB) bond, *Se*EutT fails to generate 4c $Co^{II}Cbi^{+}$ and is thus unable to adenosylate this cobalamin precursor.[13,42]

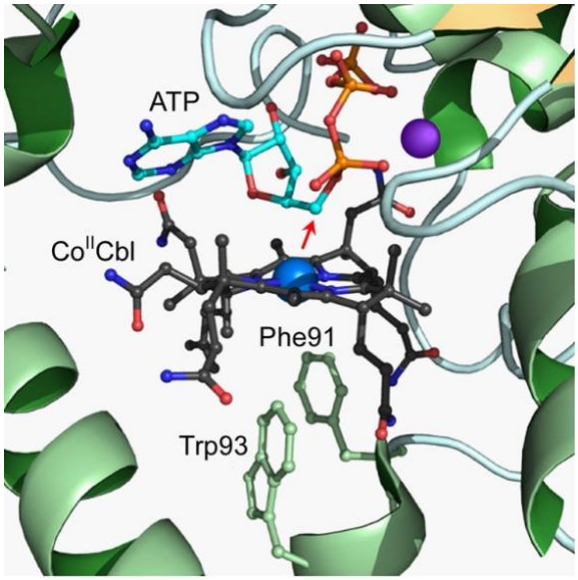


Figure 1.10. X-ray crystal structure of 4c Co^{II}Cbl (black) in the active site of *Se*CobA, with Mg-ATP located above the corrin ring. The Co and Mg atoms are shown as blue and purple spheres, respectively. Key hydrophobic residues that block off the lower axial position are depicted in pale green. The 5'-carbon of ATP that is highlighted by a red arrow.

X-ray crystallographic studies of Co^{II}Cbl bound to *Se*CobA[44] (Figure 1.10) and *Lr*PduO[45] confirmed our prediction that 4c Co^{II}Cbl species are formed in these enzymes when ATP is present. In both structures, co-substrate ATP occupies a position near the vacant "upper" axial coordination site of the Co^{II}Cbl, while enzyme-derived hydrophobic residues block access to the "lower" site to prevent the coordination of a water molecule. Although EutT has thus far eluded characterization by X-ray crystallography, kinetics and spectroscopic studies of the *S. enterica* EutT revealed that this enzyme exhibits an additional metal cofactor requirement not observed for the other ACAT families.[32,42] Intriguingly, though, some EutT enzymes from other organisms, including *Listeria monocytogenes*, lack this divalent metal cofactor.[46]

1.3.2. AdoCbl-Dependent Enzymes: Isomerases and Eliminases

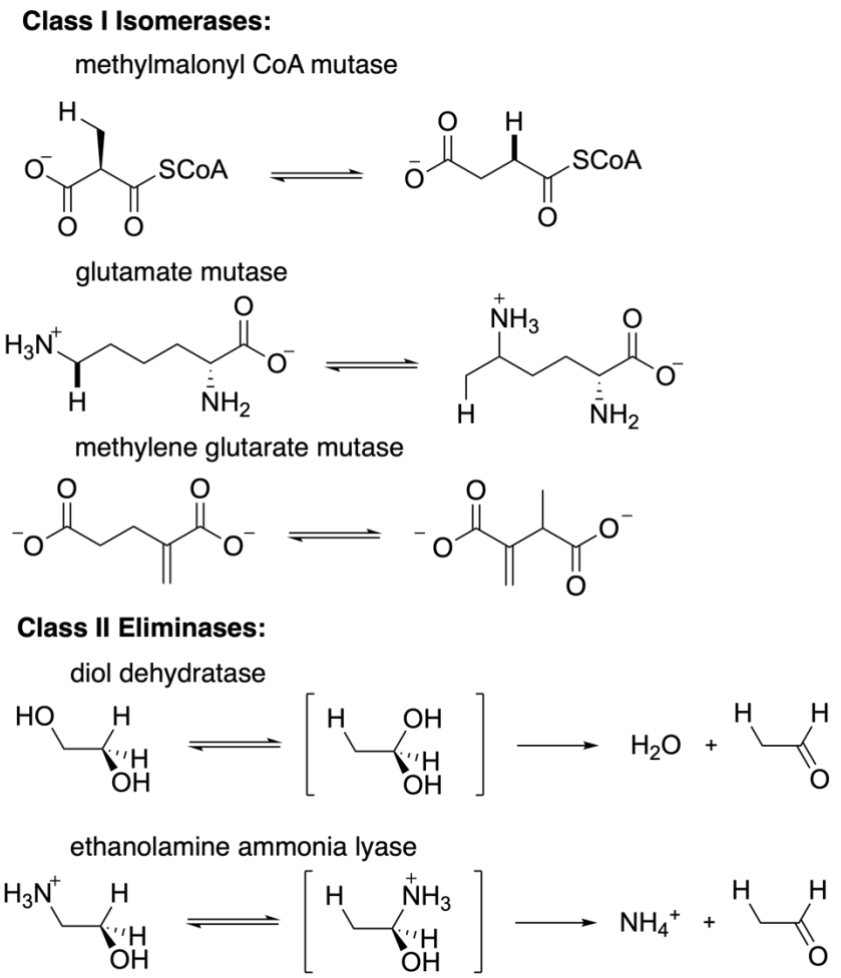


Figure 1.11. Classification of representative AdoCbl-dependent enzymes with their substrates and products.

AdoCbl-dependent enzymes use the unique reactivity of the Co–C(Ado) bond to perform a range of isomerization reactions.[29,47–50] Differences in their catalytic mechanisms allow for the classification of these enzymes into two subclasses: (i) Class I isomerases, which catalyze the 1,2-exchange of a functional group, such as an alcohol, amine, or carbon-containing fragment, and (ii) Class II eliminases, for which substrate rearrangement is followed by the spontaneous elimination of a leaving group, such as water or ammonia (Figure 1.11).[29,47] Interestingly, the two subclasses differ in their modes of AdoCbl binding. Crystal structures and EPR spectroscopic studies revealed that Class I isomerases bind AdoCbl in the base-off/His-on mode, where the DMB ligand is

displaced by a protein-derived histidine.[51–54] In contrast, Class II eliminases bind AdoCbl base-on, with the DMB ligand still bound to the Co^{III} ion.[55–57] Class I isomerases include the only AdoCbl-dependent enzyme found in humans, methylmalonyl coenzyme-A mutase (MMCM), which catalyzes the conversion of methylmalonyl-CoA to succinyl-CoA. The role this reaction plays in the catabolism of branched-chain amino acids, odd-chain fatty acids, and cholesterol makes MMCM vital to human health.[29,58,59]

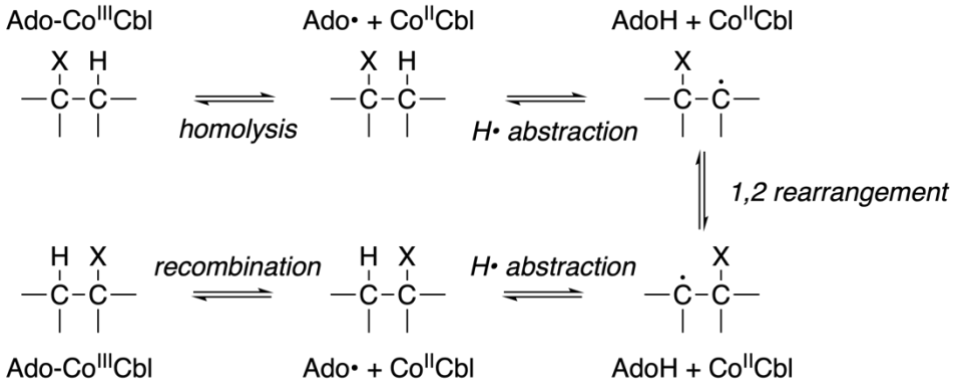


Figure 1.12. Basic mechanism of substrate rearrangement catalyzed by AdoCbl-dependent enzymes.

Following substrate binding, the catalytic cycles of all of AdoCbl-dependent enzymes are initiated through homolytic cleavage of the cofactor's Co–C bond, which generates Co^{II}Cbl and an organic radical centered on the 5'-carbon of the Ado moiety (Ado•) (Figure 1.12). Remarkably, the rate of Co–C bond homolysis is increased by 12 orders of magnitude (corresponding to a ~17 kcal/mol decrease in activation energy) when the cofactor is bound to the enzyme and substrate is added.[60,61] Elucidating the mechanism by which AdoCbl-dependent enzymes accomplish this trillion-fold rate enhancement for Co–C bond homolysis while suppressing potentially harmful side reactions involving the post-homolysis radical species has been an area of research for decades.[47,49,62] In general, this Co–C bond activation could be achieved via destabilization of the AdoCbl ground state, stabilization of the post-homolysis products Ado• and Co^{II}Cbl, or a combination of the two. MCD spectroscopy is well-suited for evaluating these possibilities, as it

permits a detailed assessment of the cofactor/enzyme active site interactions in both the AdoCbl and Co^{II}Cbl states. With our understanding of the origin of the key spectral features of Cbls, changes in the MCD spectra can be correlated with geometric and electronic perturbations of the cofactor by the active site.

This MCD-based approach was originally used for studying the Class I isomerase MMCM. Abs and MCD spectra were collected for MMCM-bound AdoCbl and Co^{II}Cbl both in the absence and presence of substrate (analogue) and compared to the spectra obtained for the free cofactors.[28,63,64] Importantly, only minor spectral changes were observed upon incorporation of AdoCbl into the enzyme active site and the subsequent addition of substrate (Figure 1.13).[63] The energies of all features in the α/β and γ regions remained largely unchanged, and the slight increase in Abs intensity of the γ features could be attributed to a change in the dielectric environment of the cofactor. These data argued against a significant enzyme-induced weakening of the Co–C bond in the AdoCbl ground state. In contrast, the electronic structure of Co^{II}Cbl was found to be notably perturbed upon protein binding and, to an even greater extent, following the addition of substrate. MCD spectra revealed a substantial blue-shift of the Co 3d $\rightarrow$ corrin π^* MLCT transitions in the ~17,000 – 20,000 cm^{-1} region, while the LF and corrin-centered $\pi \rightarrow \pi^*$ transitions were only weakly affected (Figure 1.14). An analysis of these results in the framework of DFT computations led to the proposal that MMCM induces a fairly uniform stabilization of the Co 3d-based MOs of Co^{II}Cbl and, because these orbitals are filled, a stabilization of the enzyme-bound cofactor as a whole.[28,64] Nearly identical results were obtained in a subsequent study of another member of the Class I isomerase family, glutamate mutase (GM), which catalyzes the interconversion of L-glutamate and L-threo-3-methylaspartate.[64]

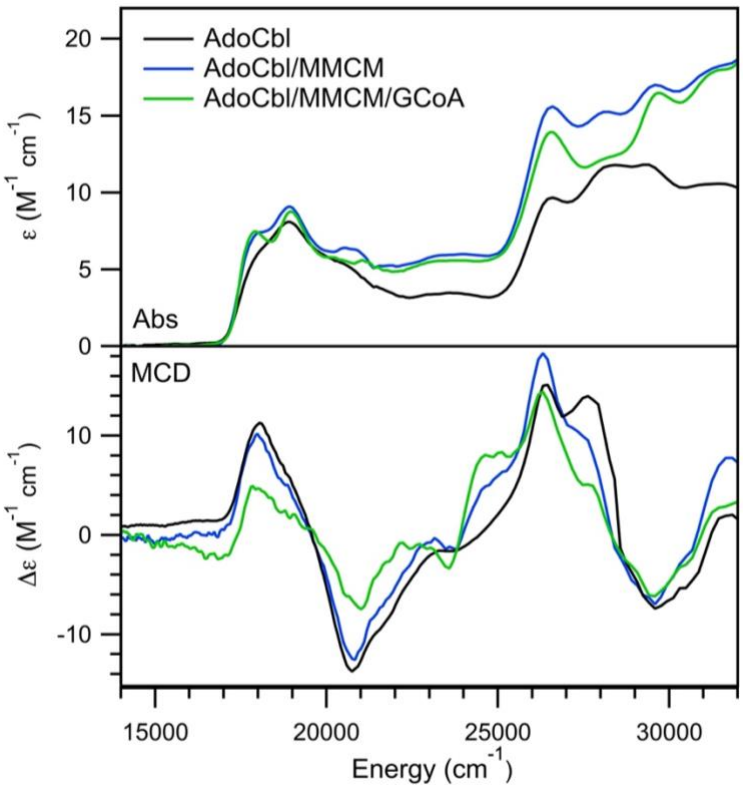


Figure 1.13. Abs spectra at 4.5 K (top) and MCD spectra (bottom) at 25 K and 7 T of free AdoCbl, AdoCbl bound to MMCM (AdoCbl/MMCM), and AdoCbl bound to MMCM complexed with the "slow" substrate glutaryl-CoA (AdoCbl/MMCM/GCoA). Adapted from ref. 63.

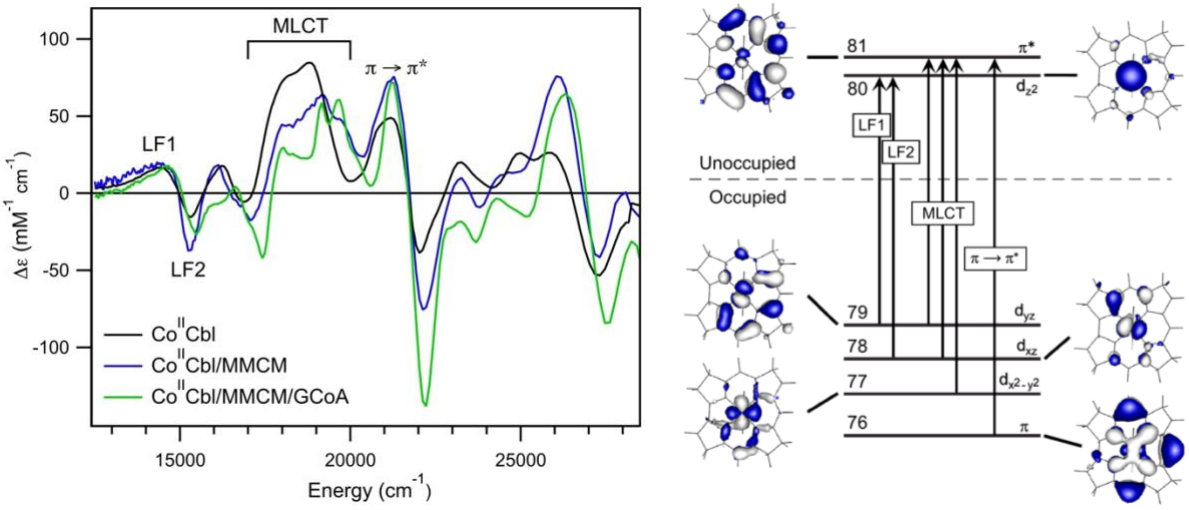


Figure 1.14. Left: MCD spectra at 4.5 K and 7 T of free Co^{II}Cbl, Co^{II}Cbl bound to MMCM (Co^{II}Cbl/MMCM), and Co^{II}Cbl bound to MMCM complexed with the "slow" substrate glutaryl-CoA (Co^{II}Cbl/MMCM/GCoA). Right: Relative energies and isosurface plots of the relevant spin-down MOs of Co^{II}Cbl based on a spin-unrestricted DFT calculation. Electronic transitions that have been correlated to bands in the experimental Co^{II}Cbl MCD spectra on the left are indicated by vertical arrows. Adapted from ref. 28.

X-ray crystal structures reported for MMCM and GM provide clues as to how these enzymes stabilize the Co^{II}Cbl post-homolysis product to promote Co–C bond homolysis. Displacement of the DMB ligand by a protein-derived histidine offers the Class I isomerases a means of influencing the electronic properties of Co^{II}Cbl. Additional tuning of the Co–N(His) bond is facilitated through a conserved DXHXXG motif, which encodes a network of H-bonding residues involving the coordinating His and conserved aspartate and lysine/asparagine residues.[51,65,66] Substituting any of these residues dramatically reduces catalytic activity, emphasizing their importance for enzyme function.[67,68] Coupling proton uptake by the H-bonding triad to Co–C bond homolysis would be expected to reduce charge donation from the His ligand in the Co^{II}Cbl state, causing a stabilization of the Co 3d orbitals and thus the cofactor as a whole, as suggested by our MCD results.[28]

Further evidence that the mechanism of Co–C bond activation employed by Class I isomerases relies primarily on a stabilization of the post-homolysis products, rather than a destabilization of the AdoCbl ground state, was obtained by a number of experimental and computational studies. A rR spectroscopic study of MMCM revealed that vibrational modes involving the Co–C(Ado) bond of AdoCbl were essentially unaffected upon cofactor binding to MMCM, even following the addition of substrate, ruling out any significant Co–C bond lengthening prior to homolysis.[8,10] Similarly, a computational study of MMCM led to the conclusion that AdoCbl binding to the enzyme active site does not induce any significant changes in Co–C bond length or corrin ring conformation.[69] Other computational studies of MMCM and GM showed that electrostatic interactions between Co^{II}Cbl and the ribose oxygens of the Ado moiety contribute to the stabilization of the post-homolysis products, along with favorable interactions between the Ado moiety and nearby protein residues.[70–73] Conformational changes of

the Ado moiety prior to Co–C bond homolysis were noted in one of these studies, but the corrin ring was found to be minimally perturbed.[73]

Interestingly, the Co–C bond activation mechanism proposed for Class I isomerases cannot be the same for Class II eliminases such as ethanolamine ammonia-lyase (EAL), which catalyzes the conversion of ethanolamine to acetaldehyde and ammonia.[47,74–76] Because Class II eliminases bind AdoCbl base-on, they are unable to activate the Co–C bond through preferential stabilization of the Co^{II}Cbl state via H-bonding interactions involving the lower axial ligand or proton uptake. Advanced EPR techniques were used to compare the reaction intermediates and key structural features of the active sites of isomerases and eliminases in some detail, which revealed interesting differences between these two subclasses.[53,54,77–84] For example, in Class I isomerases, the substrate binds ~6 to 7 Å from the Co ion, and the cleaved Ado moiety undergoes a ribose "pseudorotation" to orient itself toward the substrate.[51,85] In Class II eliminases, the substrate binds farther from the Co ion, at a distance of ~10 to 12 Å, and the ribosyl moiety rotates about the glycosidic bond in order to span the greater distance.[77–80,83] Possible protein-cofactor interactions that contribute to this mechanism have been explored for EAL through the use of cofactor analogues with modifications to the upper and lower axial ligands.[86,87] Despite these efforts, the mechanism of Co–C bond activation employed by Class II eliminases remains unclear.

1.4. Reductive Dehalogenases

Organisms that produce reductive dehalogenases (RDases) utilize halogenated aromatic and aliphatic substances as terminal electron acceptors in a process termed organohalide respiration.[88] These organisms can couple the reduction of halogenated substances with the production of ATP. Early EPR studies of as-isolated and reduced RDases revealed the presence of Fe-S clusters and a Co^{II}corrinoid cofactor.[89,90] Indeed, all dehalogenases examined so far contain

two Fe-S clusters, either a pair of [4Fe-4S] clusters or one [4Fe-4S] and one [4Fe-3S] cluster, which are expected to serve as electron-transfer centers.[89,91–94] Moreover, all but one of the characterized RDases contain a corrinoid cofactor, the exception being the *Desulfomonile tiedjei* DCB-1 enzyme, which appears to possess a heme cofactor.[95]

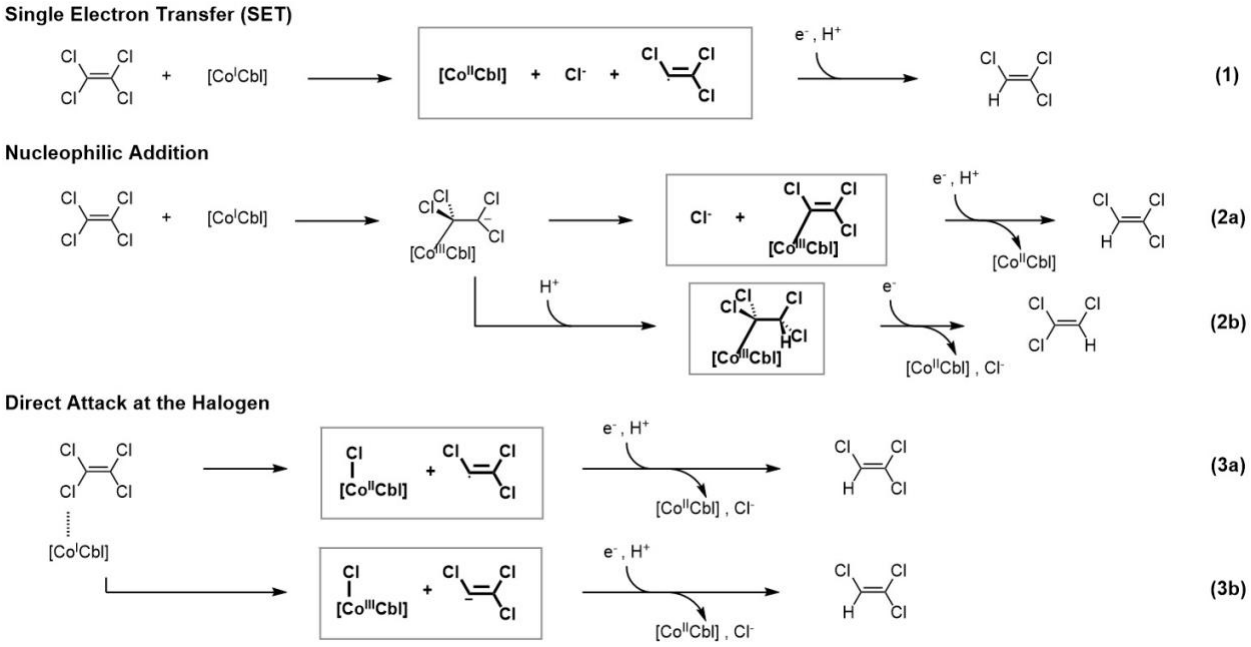


Figure 1.15. Possible mechanisms for the initial step of the conversion of PCE to TCE catalyzed by PceA: (1) single electron transfer; (2a) nucleophilic addition followed by Cl^- elimination; (2b) nucleophilic addition followed by protonation; and direct attack at the halogen with the Cl–C bond cleaved (3a) homolytically or (3b) heterolytically. The intermediates highlighted in bold and shown in a box were evaluated computationally. Adapted from ref. 96.

Tetrachloroethylene reductive dehalogenase (PceA) uses norpseudo-Co^{II}Cbl (Co^{II}Cbl*) – an unusual cobalamin that lacks the methyl group found on the intramolecular loop adjacent to the phosphate moiety and features an adenine base in lieu of the DMB – to catalyze the reductive dehalogenation of per- and trichloroethylenes (PCE and TCE, respectively) to primarily *cis*-dichloroethylene (DCE).[97] The enzymatic conversion of PCE to TCE (and subsequently DCE) was proposed to proceed via a mechanism in which the first step involves one of several options: single electron transfer, nucleophilic addition followed by chloride elimination or protonation, or direct

attack at the halogen (Figure 1.15).[96] Difficulties with producing adequate quantities of PceA greatly hampered direct experimental studies of the reaction mechanism. To overcome these challenges, we collected MCD spectra of PceA in the absence and presence of substrate, generated computational models of these states using the quantum mechanics/molecular mechanics (QM/MM) method, and validated our models on the basis of our MCD data and X-ray crystal structures.[96] Computational models of viable reaction intermediates were then constructed using the same computational methodology.

Consistent with published X-ray crystal structures,[97] the MCD spectrum of as-isolated PceA indicated that $Co^{II}Cbl^{*}$ is five-coordinate, with the intramolecular base replaced by a water molecule in one of the axial positions (Figure 1.16).[96] The lowest energy feature, which arises from a LF transition that is especially sensitive to perturbations of the axial bonding interaction,[16] was found to be red-shifted by ~320 cm^{-1} from its counterpart in the MCD spectrum of $Co^{II}Cbi^{+}$, indicating a weaker Co–O(H_2) bond in the protein-bound $Co^{II}Cbl^{*}$. Addition of TCE to PceA caused only minor changes to the MCD spectrum, ruling out any major conformational changes of the $Co^{II}Cbl^{*}$ species (Figure 1.16). Most importantly, this result provided compelling evidence that the PceA-bound $Co^{II}Cbl^{*}$ species remained five-coordinate even in the presence of substrate. Thus, the mechanism for the thermodynamically challenging $Co^{II} \rightarrow Co^{I}Cbl^{*}$ reduction used by PceA must differ fundamentally from that utilized by the ACATs (discussed above). In contrast to ACATs, PceA contains two low potential [4Fe-4S] clusters that serve as an internal reducing system.[98] While the physiological reductant remains unknown, these $[4Fe\text{-}4S]^{1+}$ clusters are sufficiently potent to reduce PceA bound $Co^{II}Cbl^{*}$.[99]

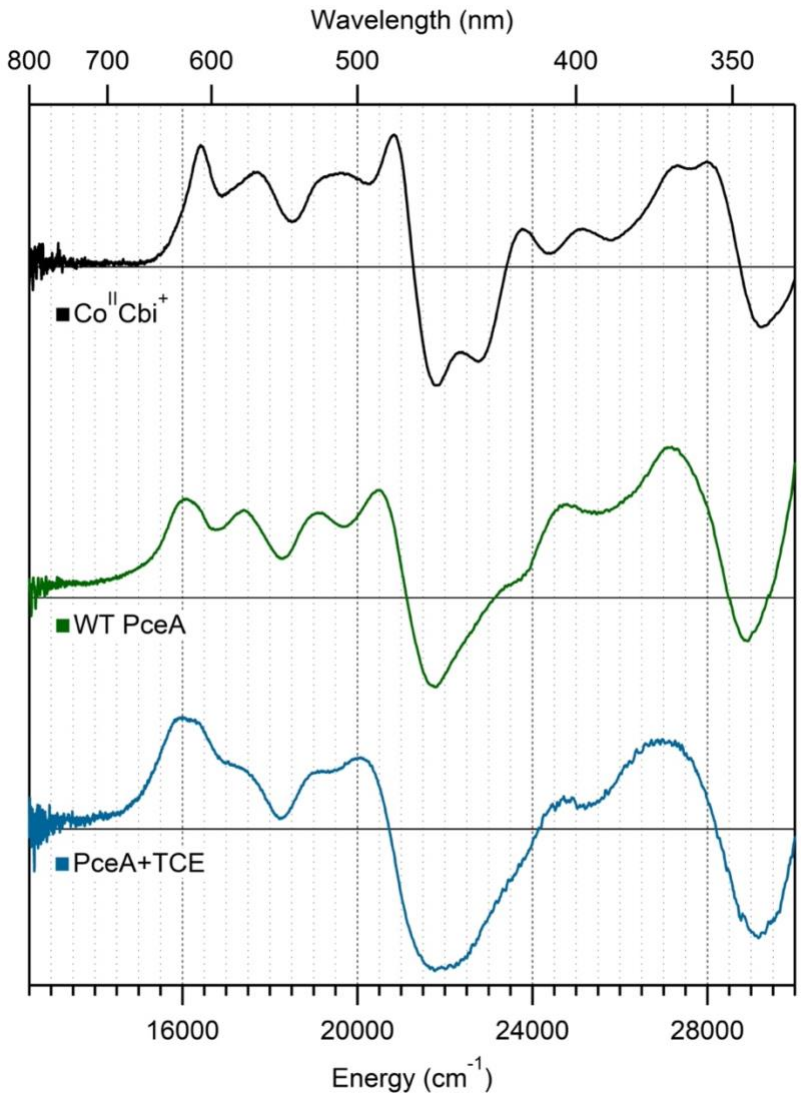


Figure 1.16. MCD spectra at 4.5 K and 7 T of (from top to bottom) $Co^{II}Cbi^{+}$; as-isolated PceA, and PceA incubated with TCE. Adapted from ref. 96.

Importantly, our QM/MM optimized models of substrate-free and TCE-bound PceA both featured five-coordinate $Co^{II}Cbl^{*}$ with a water molecule occupying the lower axial position, as stipulated by the close resemblance of the MCD spectra obtained for these species and free $Co^{II}Cbi^{+}$ (Figure 1.16).[96] The same QM/MM computational methodology was then applied to viable reaction intermediates in the catalytic cycle of PceA (Figure 1.15). The intermediate predicted to possess the lowest energy was the one resulting from single electron transfer (SET) from $Co^{I}Cbl^{*}$ to the substrate to yield $Co^{II}Cbl^{*}$, a chloride ion, and a vinylic radical (Figure 1.15).

While earlier studies led to conflicting conclusions as to the mechanism of enzymatic reductive dehalogenation, we were able to rule out mechanisms invoking a Co–Cl adduct. Instead, our computations strongly favored the SET initiated reaction.

1.5. Research Overview

This dissertation describes a collection of spectroscopic and computational studies of B_{12} cofactors and B_{12}-dependent enzymes, as well as computational work on the non-heme mononuclear iron enzyme cysteamine dioxygenase (ADO).

- Chapter 2 describes a combined spectroscopic and computational investigation of vibronic coupling in CNCbl. In addition to advancing our understanding of how popular functionals B3LYP and BP86 model CNCbl's electronic properties, this work identified the elusive Co–C stretching peak in the cofactor's rRaman spectrum.
- Chapter 3 investigates the mechanism through which the AdoCbl-dependent enzyme ethanolamine ammonia-lyase (EAL) activates the cofactor's Co–C bond. Spectroscopic data revealed that activation is not achieved through stabilization of Co^{II}Cbl, as is the case for Class I AdoCbl-dependent enzymes. Rather, computational models generated using molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) identify enzyme-cofactor-substrate interactions that activate the AdoCbl "ground state."
- Chapter 4 presents a computational study of the catalytic cycle of a PduO-type ATR, with a focus on the thermodynamically challenging Co^{II}Cbl → Co^{I}Cbl reduction step and the formation of the Co–C bond. These calculations predict an increase in Co^{II}/Co^{I}Cbl reduction potential when bound to PduO in the presence of cosubstrate ATP.

- Chapter 5 describes progress made towards the computational modeling of free and enzyme-bound adenosylrhodibalamin (AdoRhbl), a rhodium-substituted B_{12} analogue with enormous potential to advance structural studies of AdoCbl-containing enzymes.
- Chapter 6 presents the first X-ray crystal structure of cysteamine dioxygenase (ADO), a non-heme iron enzyme that catalyzes the conversion of cysteamine to hypotaurine, as well as the oxidation of N-terminal cysteine peptides. MD simulations and QM/MM optimizations were performed with this structure in order to validate conclusions on how ADO's structure contributes to its unique substrate specificity.
- Chapter 7 was written as part of the Communicating PhD Research to the Public project led by the Wisconsin Initiative for Science Literacy (WISL). It discusses my work on vibronic coupling in CNCbl (Chapter 2) for an audience of non-scientists.
- Chapter 8 summarizes the research progress I have achieved and proposes future areas of study that would build on these results.

References

(1) Firth, R. A., Hill, H. A. O., Pratt, J. M., Williams, R. J. P. & Jackson, W. R. The Circular Dichroism and Absorption Spectra of Some Vitamin B_{12} Derivatives. *Biochemistry* **6**, 2178–2189 (1967).

(2) Toohey, J. I. A Vitamin B_{12} Compound Containing No Cobalt. *Proc. Natl. Acad. Sci. USA.* **54**, 934–942 (1965).

(3) Johnson, M. K. CD and MCD Spectroscopy. in *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism* (ed. Que, L. J.) 233–285 (University Science Books, 2000).

(4) Solomon, E. I., Kirk, M. L., Gamelin, D. R. & Pulver, S. Bioinorganic Spectroscopy. *Methods Enzymol.* **246**, 71–110 (1995).

(5) Solomon, E. I., Neidig, M. L. & Schenk, G. Magnetic Circular Dichroism of Paramagnetic Species. in *Comprehensive Coordination Chemistry II* (eds. McCleverty, J. A. & Meyer, T. J.) vol. 2, 339–349 (Elsevier Ltd, 2004).

(6) Neese, F. & Solomon, E. I. MCD C-term Signs, Saturation Behavior, and Determination of Band Polarizations in Randomly Oriented Systems with Spin $S \geq 1/2$. Applications to $S = 1/2$ and $S = 5/2$. *Inorg. Chem.* **38**, 1847–1865 (1999).

(7) Dong, S., Padmakumar, R., Banerjee, R. & Spiro, T. G. Resonance Raman Co−C Stretching Frequencies Reflect Bond Strength Changes in Alkyl Cobalamins, but Are Unaffected by Trans Ligand Substitution. *J. Am. Chem. Soc.* **118**, 9182–9183 (1996).

(8) Dong, S., Padmakumar, R., Maiti, N., Banerjee, R. & Spiro, T. G. Resonance Raman Spectra Show That Coenzyme B_{12} Binding to Methylmalonyl-Coenzyme A Mutase Changes the Corrin Ring Conformation but Leaves the Co–C Bond Essentially Unaffected. *J. Am. Chem. Soc.* **120**, 9947–9948 (1998).

(9) Dong, S., Padmakumar, R., Banerjee, R. & Spiro, T. G. Co–C Force Constants from Resonance Raman Spectra of Alkylcobalamins: Insensitivity to Dimethylbenzylimidazole Coordination. *Inorganica Chim. Acta* **270**, 392–398 (1998).

(10) Dong, S., Padmakumar, R., Banerjee, R. & Spiro, T. G. Co–C Bond Activation in B_{12}-Dependent Enzymes: Cryogenic Resonance Raman Studies of Methylmalonyl-Coenzyme A Mutase. *J. Am. Chem. Soc.* **121**, 7063–7070 (1999).

(11) Park, K., Mera, P. E., Escalante-Semerena, J. C. & Brunold, T. C. Kinetic and Spectroscopic Studies of the ATP:Corrinoid Adenosyltransferase PduO from Lactobacillus reuteri: Substrate Specificity and Insights into the Mechanism of Co(II)corrinoid Reduction. *Biochemistry* **47**, 9007–9015 (2008).

(12) Park, K., Mera, P. E., Escalante-Semerena, J. C. & Brunold, T. C. Spectroscopic Characterization of Active-Site Variants of the PduO-type ATP:Corrinoid Adenosyltransferase from *Lactobacillus Reuteri*: Insights into the Mmechanism of Four-Coordinate Co(II)corrinoid Formation. *Inorg. Chem.* **51**, 4482–4494 (2012).

(13) Park, K., Mera, P. E., Moore, T. C., Escalante-Semerena, J. C. & Brunold, T. C. Unprecedented Mechanism Employed by the Salmonella enterica EutT ATP:Co^{I}rrinoid

Adenosyltransferase Precludes Adenosylation of Incomplete CoIIrrinoids. *Angew. Chem., Int. Ed.* **54**, 7158–7161 (2015).

(14) Park, K., Mera, P. E., Escalante-Semerena, J. C. & Brunold, T. C. Resonance Raman Spectroscopic Study of the Interaction between Co(II)rrinoids and the ATP:corrinoid Adenosyltransferase PduO from *Lactobacillus reuteri*. *J. Biol. Inorg. Chem.* **21**, 669–681 (2016).

(15) Stich, T. A., Brooks, A. J., Buan, N. R. & Brunold, T. C. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **125**, 5897–5914 (2003).

(16) Stich, T. A., Buan, N. R. & Brunold, T. C. Spectroscopic and Computational Studies of Co^{2+}Corrinoids: Spectral and Electronic Properties of the Biologically Relevant Base-on and Base-Off Forms of Co^{2+}Cobalamin. *J. Am. Chem. Soc.* **126**, 9735–9749 (2004).

(17) Liptak, M. D. & Brunold, T. C. Spectroscopic and Computational Studies of Co^{1+}Cobalamin: Spectral and Electronic Properties of the "Superreduced" B_{12} Cofactor. *J. Am. Chem. Soc.* **128**, 9144–9156 (2006).

(18) Brunold, T. C., Conrad, K. S., Liptak, M. D. & Park, K. Spectroscopically Validated Density Functional Theory Studies of the B_{12} Cofactors and their Interactions with Enzyme Active Sites. *Coordination Chemistry Reviews* **253**, 779–794 (2009).

(19) Park, K. & Brunold, T. C. Combined Spectroscopic and Computational Analysis of the Vibrational Properties of Vitamin B12 in its Co^{3+}, Co^{2+}, and Co^{1+} Oxidation States. *J. Phys. Chem. B* **117**, 5397–5410 (2013).

(20) Pratt, J. M. The Roles of Co, Corrin, and Protein. II. Electronic Spectra and Structure of the Corrin Ligand: Molecular Machinery of the Protein. in *Chemistry and Biochemistry of B_{12}* (ed. Banerjee, R.) 113–164 (Wiley, 1999).

(21) Salama, S. & Spiro, T. G. Visible and Near-Ultraviolet Resonance Raman Spectra of Photolabile Vitamin B_{12} Derivatives with a Rapid-Flow Technique. *J. Raman Spectrosc.* **6**, 57–60 (1977).

(22) Andruniów, T., Jaworska, M., Lodowski, P., Zgierski, M. Z., Dreos, R., Randaccio, L. & Kozlowski, P. M. Time-Dependent Density Functional Theory Study of Cobalt Corrinoids: Electronically Excited States of Methylcobalamin. *J. Chem. Phys.* **129**, 085101 (2008).

(23) Kornobis, K., Kumar, N., Wong, B. M., Lodowski, P., Jaworska, M., Andruniów, T., Ruud, K. & Kozlowski, P. M. Electronically Excited States of Vitamin B_{12}: Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated ab Initio Methods. *J. Phys. Chem. A* **115**, 1280–1292 (2011).

(24) Solheim, H., Kornobis, K., Ruud, K. & Kozlowski, P. M. Electronically Excited States of Vitamin B_{12} and Methylcobalamin: Theoretical Analysis of Absorption, CD, and MCD Data. *J. Phys. Chem. B* **115**, 737–748 (2011).

(25) Handali, J. D., Sunden, K. F., Thompson, B. J., Neff-Mallon, N. A., Kaufman, E. M., Brunold, T. C. & Wright, J. C. Three Dimensional Triply Resonant Sum Frequency Spectroscopy Revealing Vibronic Coupling in Cobalamins: Toward a Probe of Reaction Coordinates. *J. Phys. Chem. A* **122**, 9031–9042 (2018).

(26) Salerno, E. V., Miller, N. A., Konar, A., Li, Y., Kieninger, C., Kraütler, B. & Sension, R. J. Ultrafast Excited State Dynamics and Fluorescence from Vitamin B_{12} and Organometallic [Co]-C≡C-R Cobalamins. *J. Phys. Chem. B* **124**, 6651–6656 (2020).

(27) Reig, A. J., Conrad, K. S. & Brunold, T. C. Combined Spectroscopic/Computational Studies of Vitamin B_{12} Precursors: Geometric and Electronic Structures of Cobinamides. *Inorg. Chem.* **51**, 2867–2879 (2012).

(28) Brooks, A. J., Vlasie, M., Banerjee, R. & Brunold, T. C. Co-C Bond Activation in Methylmalonyl-CoA Mutase by Stabilization of the Post-homolysis Product Co^{2+} Cobalamin. *J. Am. Chem. Soc.* **127**, 16522–16528 (2005).

(29) Banerjee, R. & Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **72**, 209–247 (2003).

(30) Warren, M. J., Raux, E., Schubert, H. L. & Escalante-Semerena, J. C. The Biosynthesis of Adenosylcobalamin (Vitamin B_{12}). *Nat. Prod. Rep.* **19**, 390–412 (2002).

(31) Mera, P. E. & Escalante-Semerena, J. C. Multiple Roles of ATP:cob(I)alamin Adenosyltransferases in the Conversion of B_{12} to Coenzyme B_{12}. *Applied Microbiology and Biotechnology* vol. 88 41–48 (2010).

(32) Moore, T. C., Mera, P. E. & Escalante-Semerena, J. C. The EutT Enzyme of Salmonella enterica Is a Unique ATP:Cob(I)alamin Adenosyltransferase Metalloprotein That Requires Ferrous Ions for Maximal Activity. *J. Bacteriol.* **196**, 903–910 (2014).

(33) Johnson, C. L. V, Buszko, M. L. & Bobik, T. A. Purification and Initial Characterization of the *Salmonella enterica* PduO ATP:Cob(I)alamin Adenosyltransferase. *J. Bacteriol.* **186**, 7881–7887 (2004).

(34) Johnson, C. L. V, Pechonick, E., Park, S. D., Havemann, G. D., Leal, N. A. & Bobik, T. A. Functional Genomic, Biochemical, and Genetic Characterization of the Salmonella PduO Gene, an ATP:Cob(I)alamin Adenosyltransferase Gene. *J. Bacteriol.* **183**, 1577–1584 (2001).

(35) Buan, N. R., Suh, S. J. & Escalante-Semerena, J. C. The EutT Gene of *Salmonella Enterica* Encodes an Oxygen-Labile, Metal-Containing ATP:Corrinoid Adenosyltransferase Enzyme (vol 186, pg 5711, 2004). *J. Bacteriol.* **186**, 7826 (2004).

(36) Mera, P. E. & Escalante-Semerena, J. C. Dihydroflavin-Driven Adenosylation of 4-Coordinate Co(II) Corrinoids: Are Cobalamin Reductases Enzymes or Electron Transfer Proteins? *J. Biol. Chem.* **285**, 2911–2917 (2010).

(37) Fonseca, M. V & Escalante-Semerena, J. C. Reduction of Cob(III)alamin to Cob(II)alamin in *Salmonella enterica* Serovar Typhimurium LT2. *J. Bacteriol.* **182**, 4304–4309 (2000).

(38) Faure, D., Lexa, D. & Savéant, J.-M. Electrochemistry of Vitamin B_{12}. *J. Electroanal. Chem. Interfacial Electrochem.* **140**, 297–309 (1982).

(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.* **127**, 7660–7661 (2005).

(40) Schrauzer, G. N. & Deutsch, E. Reactions of Cobalt(I) Supernucleophiles. The Alkylation

of Vitamin B_{12}s, Cobaloximes(I), and Related Compounds. *J. Am. Chem. Soc.* **91**, 3341–3350 (1969).

(41) Stich, T. A., Buan, N. R., Escalante-Semerena, J. C. & Brunold, T. C. Spectroscopic and Computational Studies of the ATP:Corrinoid Adenosyltransferase (CobA) from *Salmonella enterica*: Insights into the Mechanism of Adenosylcobalamin Biosynthesis. *J. Am. Chem. Soc.* **127**, 8710–8719 (2005).

(42) Pallares, I. G., Brunold, T. C., Moore, T. C. & Escalante-Semerena, J. C. Spectroscopic Studies of the EutT Adenosyltransferase from *Salmonella enterica*: Mechanism of Four-Coordinate Co(II)Cbl Formation. *J Am Chem Soc* **138**, 3694–3704 (2016).

(43) Pallares, I. G., Moore, T. C., Escalante-Semerena, J. C. & Brunold, T. C. Spectroscopic Studies of the *Salmonella enterica* Adenosyltransferase Enzyme SeCobA: Molecular-Level Insight into the Mechanism of Substrate Cob(II)alamin Activation. *Biochemistry* **53**, 7969–7982 (2014).

(44) Moore, T. C., Newmister, S. A., Rayment, I. & Escalante-Semerena, J. C. Structural Insights into the Mechanism of Four-Coordinate Cob(II)alamin Formation in the Active Site of the *Salmonella enterica* ATP:Co(I)rrinoid Adenosyltransferase Enzyme: Critical Role of Residues Phe91 and Trp93. *Biochemistry* **51**, 9647–9657 (2012).

(45) 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* **47**, 5755–5766 (2008).

(46) Stracey, N. G., Costa, F. G., Escalante-Semerena, J. C. & Brunold, T. C. Spectroscopic Study of the EutT Adenosyltransferase from *Listeria monocytogenes*: Evidence for the Formation of a Four-Coordinate Cob(II)alamin Intermediate. *Biochemistry* **57**, 5088–5095 (2018).

(47) Toraya, T. Radical Catalysis in Coenzyme B_{12}-Dependent Isomerization (Eliminating) Reactions. *Chem. Rev.* **103**, 2095–2127 (2003).

(48) Toraya, T. Cobalamin-Dependent Dehydratases and a Deaminase: Radical Catalysis and Reactivating Chaperones. *Arch. Biochem. Biophys.* **544**, 40–57 (2014).

(49) Banerjee, R. Radical carbon skeleton rearrangements: Catalysis by Coenzyme B_{12}-Dependent Mutases. *Chem. Rev.* **103**, 2083–2094 (2003).

(50) Greenhalgh, E. D. & Brunold, T. C. Biorelevant Chemistry of Cobalamin. in *Comprehensive Coordination Chemistry III* (eds. Que Jr, L. . & Lu, Y.) 812–824 (Elsevier, 2021).

(51) Gruber, K., Reitzer, R. & Kratky, C. Radical Shuttling in a Protein: Ribose Pseudorotation Controls Alkyl-Radical Transfer in the Coenzyme B_{12} Dependent Enzyme Glutamate Mutase. *Angew. Chemie - Int. Ed.* **40**, 3377–3380 (2001).

(52) Mancia, F., Smith, G. A. & Evans, P. R. Crystal Structure of Substrate Complexes of Methylmalonyl-CoA Mutase. *Biochemistry* **38**, 7999–8005 (1999).

(53) Padmakumar, R., Taoka, S., Padmakumar, R. & Banerjee, R. Coenzyme B_{12} Is Coordinated by Histidine and Not Dimethylbenzimidazole on Methylmalonyl-CoA

Mutase. *J. Am. Chem. Soc.* **117**, 7033–7034 (1995).

(54) Zelder, O., Beatrix, B., Kroll, F. & Buckel, W. Coordination of a Histidine Residue of the Protein-Component S to the Cobalt Atom in Coenzyme B_{12}-Dependent Glutamate Mutase from *Clostridium cochlearium. FEBS Lett.* **369**, 252–254 (1995).

(55) Shibata, N., Tamagaki, H., Hieda, N., Akita, K., Komori, H., Shomura, Y., Terawaki, S.-I., Mori, K., Yasuoka, N., Higuchi, Y. & Toraya, T. Crystal Structures of Ethanolamine Ammonia-Lyase Complexed with Coenzyme B_{12} Analogs and Substrates. *J. Biol. Chem.* **285**, 26484–26493 (2010).

(56) Shibata, N., Sueyoshi, Y., Higuchi, Y. & Toraya, T. Direct Participation of a Peripheral Side Chain of a Corrin Ring in Coenzyme B_{12} Catalysis. *Angew. Chemie Int. Ed.* **57**, 7830–7835 (2018).

(57) Abend, A., Bandarian, V., Nitsche, R., Stupperich, E., Rétey, J. & Reed, G. H. Ethanolamine Ammonia-Lyase has a 'Base-On' Binding Mode for Coenzyme B_{12}. *Arch. Biochem. Biophys.* **370**, 138–141 (1999).

(58) Katz, J. & Chaikoff, I. L. The Metabolism of Propionate by Rat Liver Slices and the Formation of Isosuccinic Acid. *J. Am. Chem. Soc.* **77**, 2659–2660 (1955).

(59) Flavin, M., Ortiz, P. J. & Ochoa, S. Metabolism of Propionic Acid in Animal Tissues. *Nature* **176**, 823–826 (1955).

(60) Chowdhury, S. & Banerjee, R. Thermodynamic and Kinetic Characterization of Co−C Bond Homolysis Catalyzed by Coenzyme B_{12}-Dependent Methylmalonyl-CoA Mutase. *Biochemistry* **39**, 7998–8006 (2000).

(61) Hay, B. P. & Finke, I. R. G. Thermolysis of the Co–C Bond in Adenosylcorrins: 3: Quantification of the Axial Base Effect in Adenosylcobalamin by the Synthesis and Thermolysis of Axial Base-Free Adenosylcobinamide: Insights into the Energetics of Enzyme-Assisted Cobalt–Carbon Bond Homolysis. *J. Am. Chem. Soc.* **109**, 8012–8018 (1987).

(62) Kräutler, B., Keller, W. & Kratky, C. Coenzyme B_{12} Chemistry: The Crystal and Molecular Structure of Cob(II)alamin. *J. Am. Chem. Soc.* **111**, 8936–8938 (1989).

(63) Brooks, A. J., Vlasie, M., Banerjee, R. & Brunold, T. C. Spectroscopic and Computational Studies on the Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase: Evaluation of Enzymatic Contributions to Co–C Bond Activation in the Co^{3+} Ground State. *J. Am. Chem. Soc.* **126**, 8167–8180 (2004).

(64) Brooks, A. J., Fox, C. C., Marsh, E. N. G., Vlasie, M., Banerjee, R. & Brunold, T. C. Electronic Structure Studies of the Adenosylcobalamin Cofactor in Glutamate Mutase. *Biochemistry* **44**, 15167–15181 (2005).

(65) Mancia, F. & Evans, P. R. Conformational Changes on Substrate Binding to Methylmalonyl CoA Mutase and New Insights into the Free Radical Mechanism. *Structure* **6**, 711–720 (1998).

(66) Marsh, E. N. G. & Holloway, D. E. Cloning and Sequencing of Glutamate Mutase Component S from *Clostridium tetanomorphum* Homologies with Other Cobalamin-Dependent Enzymes. *FEBS Lett.* **310**, 167–170 (1992).

(67) Chen, H. P. & Marsh, E. N. G. How Enzymes Control the Reactivity of Adenosylcobalamin: Effect on Coenzyme Binding and Catalysis of Mutations in the Conserved Histidine-Aspartate Pair of Glutamate Mutase. *Biochemistry* **36**, 7884–7889 (1997).

(68) Vlasie, M., Chowdhury, S. & Banerjee, R. Importance of the Histidine Ligand to Coenzyme B_{12} in the Reaction Catalyzed by Methylmalonyl-CoA Mutase. *J. Biol. Chem.* **277**, 18523–18527 (2002).

(69) Freindorf, M. & Kozlowski, P. M. A Combined Density Functional Theory and Molecular Mechanics Study of the Relationship between the Structure of Coenzyme B_{12} and Its Binding to Methylmalonyl-CoA Mutase. *J. Am. Chem. Soc.* **126**, 1928–1929 (2004).

(70) Jensen, K. P. & Ryde, U. How the Co–C Bond is Cleaved in Coenzyme B_{12} Enzymes: A Theoretical Study. *J. Am. Chem. Soc.* **127**, 9117–9128 (2005).

(71) Dölker, N., Maseras, F. & Siegbahn, P. E. M. Stabilization of the Adenosyl Radical in Coenzyme B_{12} - A Theoretical Study. *Chem. Phys. Lett.* **386**, 174–178 (2004).

(72) Sharma, P. K., Chu, Z. T., Olsson, M. H. M. & Warshel, A. A New Paradigm for Electrostatic Catalysis of Radical Reactions in Vitamin B_{12} Enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9661–9666 (2007).

(73) Kwiecien, R. A., Khavrutskii, I. V., Musaev, D. G., Morokuma, K., Banerjee, R. & Paneth, P. Computational Insights into the Mechanism of Radical Generation in B_{12}-Dependent Methylmalonyl-CoA Mutase. *J. Am. Chem. Soc.* **128**, 1287–1292 (2006).

(74) Mori, K., Oiwa, T., Kawaguchi, S., Kondo, K., Takahashi, Y. & Toraya, T. Catalytic Roles of Substrate-Binding Residues in Coenzyme B12-Dependent Ethanolamine Ammonia-Lyase. *Biochemistry* **53**, 2661–2671 (2014).

(75) Bradbeer, C. The clostridial fermentations of choline and ethanolamine I. Preparation and Properties of Cell-Free Extracts. *J. Biol. Chem.* **240**, 4669–4674 (1965).

(76) Chang, G. W. & Chang, J. T. Evidence for the B_{12}-Dependent Enzyme Ethanolamine Deaminase in *Salmonella*. *Nature* **254**, 150–151 (1975).

(77) LoBrutto, R., Bandarian, V., Magnusson, O. T., Chen, X., Schramm, V. L. & Reed, G. H. 5′-Deoxyadenosine Contacts the Substrate Radical Intermediate in the Active Site of Ethanolamine Ammonia-Lyase: ^{2}H and ^{13}C Electron Nuclear Double Resonance Studies. *Biochemistry* **40**, 9–14 (2001).

(78) Bandarian, V. & Reed, G. H. Analysis of the Electron Paramagnetic Resonance Spectrum of a Radical Intermediate in the Coenzyme B_{12}-Dependent Ethanolamine Ammonia-Lyase Catalyzed Reaction of S-2-Aminopropanol. *Biochemistry* **41**, 8580–8588 (2002).

(79) Bender, G., Poyner, R. R. & Reed, G. H. Identification of the Substrate Radical Intermediate Derived from Ethanolamine During Catalysis by Ethanolamine Ammonia-Lyase. *Biochemistry* **47**, 11360–11366 (2008).

(80) Ke, S. C. & Warncke, K. Interactions of Substrate and Product Radicals with Co(II) in Cobalamin and with the Active Site in Ethanolamine Deaminase, Characterized by ESE-EPR and ^{14}N ESEEM Spectroscopies. *J. Am. Chem. Soc.* **121**, 9922–9927 (1999).

(81) Ke, S. C., Torrent, M., Museav, D. G., Morokuma, K. & Warncke, K. Identification of Dimethylbenzimidazole Axial Coordination and Characterization of ^{14}N Superhyperfine and Nuclear Quadrupole Coupling in Cob(II)alamin Bound to Ethanolamine Deaminase in a Catalytically-Engaged Substrate Radical-Cobalt(II) Biradical State. *Biochemistry* **38**, 12681–12689 (1999).

(82) Warncke, K., Schmidt, J. C. & Ke, S. C. Identification of a Rearranged-Substrate, Product Radical Intermediate and the Contribution of a Product Radical Trap in Vitamin B_{12} Coenzyme-Dependent Ethanolamine Deaminase Catalysis. *J. Am. Chem. Soc.* **121**, 10522–10528 (1999).

(83) Warncke, K. & Utada, A. S. Interaction of the Substrate Radical and the 5′-Deoxyadenosine-5′-Methyl Group in Vitamin B_{12} Coenzyme-Dependent Ethanolamine Deaminase. *J. Am. Chem. Soc.* **123**, 8564–8572 (2001).

(84) Robertson, W. D., Wang, M. & Warncke, K. Characterization of Protein Contributions to Cobalt−Carbon Bond Cleavage Catalysis in Adenosylcobalamin-Dependent Ethanolamine Ammonia-Lyase by using Photolysis in the Ternary Complex. *J. Am. Chem. Soc.* **133**, 6968–6977 (2011).

(85) Bothe, H., Darley, D. J., Albracht, S. P. J., Gerfen, G. J., Golding, B. T. & Buckel, W. Identification of the 4-Glutamyl Radical as an Intermediate in the Carbon Skeleton Rearrangement Catalyzed by Coenzyme B_{12} -Dependent Glutamate Mutase from *Clostridium cochlearium*. *Biochemistry* **37**, 4105–4113 (1998).

(86) Chen, Z.-G. G., Zietek, M. A., Russell, H. J., Tait, S., Hay, S., Jones, A. R., Scrutton, N. S., Ziętek, M. A., Russell, H. J., Tait, S., Hay, S., Jones, A. R. & Scrutton, N. S. Dynamic, Electrostatic Model for the Generation and Control of High-Energy Radical Intermediates by a Coenzyme B_{12}-Dependent Enzyme. *ChemBioChem* **14**, 1529–1533 (2013).

(87) Fukuoka, M., Yamada, S., Miyoshi, S., Yamashita, K., Yamanishi, M., Zou, X., Brown, K. L. & Toraya, T. Functions of the D-Ribosyl Moiety and the Lower Axial Ligand of the Nucleotide Loop of Coenzyme B_{12} in Diol Dehydratase and Ethanolamine Ammonia-lyase Reactions. *J. Biochem.* **132**, 935–943 (2002).

(88) John, M., Rubick, R., Schmitz, R. P., Rakoczy, J., Schubert, T. & Diekert, G. Retentive Memory of Bacteria: Long-Term Regulation of Dehalorespiration in *Sulfurospirillum multivorans*. *J Bacteriol* **191**, 1650–1655 (2009).

(89) van de Pas, B. A., Smidt, H., Hagen, W. R., van der Oost, J., Schraa, G., Stams, A. J. & de Vos, W. M. Purification and Molecular Characterization of Ortho-Chlorophenol Reductive Dehalogenase, a Key Enzyme of Halorespiration in *Desulfitobacterium dehalogenans*. *J Biol Chem* **274**, 20287–20292 (1999).

(90) Schumacher, W., Holliger, C., Zehnder, A. J. & Hagen, W. R. Redox Chemistry of Cobalamin and Iron-Sulfur Cofactors in the Tetrachloroethene Reductase of *Dehalobacter restrictus*. *FEBS Lett* **409**, 421–425 (1997).

(91) Maillard, J., Schumacher, W., Vazquez, F., Regeard, C., Hagen, W. R. & Holliger, C. Characterization of the Corrinoid Iron-Sulfur Protein Tetrachloroethene Reductive Dehalogenase of *Dehalobacter restrictus*. *Appl Env. Microbiol* **69**, 4628–4638 (2003).

(92) Wohlfarth, G. & Diekert, G. Anaerobic Dehalogenases. *Curr Opin Biotechnol* **8**, 290–295

(1997).

(93) Krasotkina, J., Walters, T., Maruya, K. A. & Ragsdale, S. W. Characterization of the B_{12}- and Iron-Sulfur-containing Reductive Dehalogenase from *Desulfitobacterium chlororespirans*. *J. Biol. Chem.* **276**, 40991–40997 (2001).

(94) Neumann, A., Siebert, A., Trescher, T., Reinhardt, S., Wohlfarth, G. & Diekert, G. Tetrachloroethene Reductive Dehalogenase of *Dehalospirillum multivorans*: Substrate Specificity of the Native Enzyme and its Corrinoid Cofactor. *Arch Microbiol* **177**, 420–426 (2002).

(95) Ni, S., Fredrickson, J. K. & Xun, L. Purification and Characterization of a Novel 3-Chlorobenzoate-Reductive Dehalogenase from the Cytoplasmic Membrane of *Desulfomonile tiedjei* DCB-1. *J. Bacteriol.* **177**, 5135–5139 (1995).

(96) Greenhalgh, E. D., Kunze, C., Schubert, T., Diekert, G. & Brunold, T. C. A Spectroscopically Validated Computational Investigation of Viable Reaction Intermediates in the Catalytic Cycle of the Reductive Dehalogenase PceA. *Biochemistry* **60**, 2022–2032 (2021).

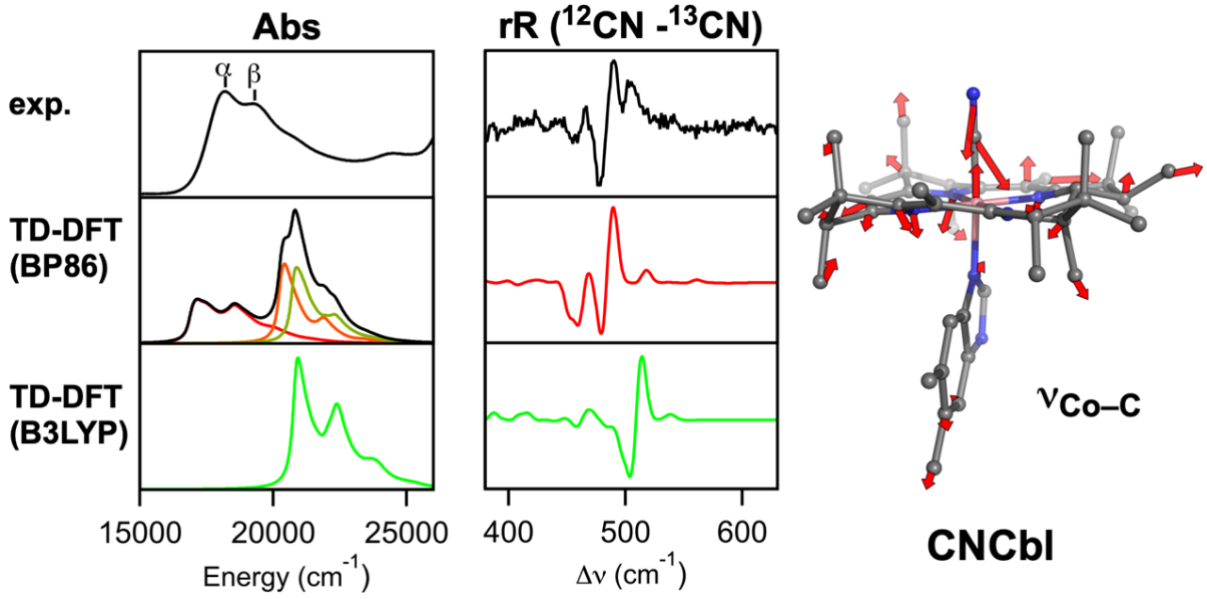
(97) Bommer, M., Kunze, C., Fesseler, J., Schubert, T., Diekert, G. & Dobbek, H. Structural Basis for Organohalide Respiration. *Science (80-.).* **346**, 455–458 (2014).

(98) Schubert, T., Adrian, L., Sawers, R. G. & Diekert, G. Organohalide Respiratory Chains: Composition, Topology and Key Enzymes. *FEMS Microbiol Ecol* **94**, (2018).

(99) Kräutler, B., Fieber, W., Ostermann, S., Fasching, M., Ongania, K.-H. H., Gruber, K., Kratky, C., Mikl, C., Siebert, A. & Diekert, G. The Cofactor of Tetrachloroethene Reductive Dehalogenase of *Dehalospirillum multivorans* is Norpseudo-B_{12}, a New Type of a Natural Corrinoid. *Helv. Chim. Acta* **86**, 3698–3716 (2003).

Chapter 2

Vibronic Coupling in Vitamin B_{12}: A Combined Spectroscopic and Computational Study



This chapter has been published under the following: Elmendorf, L. D.; Brunold, T. C. Vibronic Coupling in Vitamin B_{12}: a Combined Spectroscopic and Computational Study. *Inorganic Chemistry*, **2023**, *62* (32), 12762-12772. https://doi.org/10.1021/acs.inorgchem.3c01305.

Chapter 2: Vibronic Coupling in Vitamin B_{12}: A Combined Spectroscopic and Computational Study

2.1. Introduction

The complexity and versatile reactivity of cobalamins have fascinated scientists for almost a century, since the initial discovery of an "antipernicious anemia factor" by Robscheit-Robbins and Whipple in 1925.[1] That factor, now known as vitamin B_{12} or cyanocobalamin (CNCbl), is a member of the cobalamin family of cofactors, each of which contains a Co ion equatorially ligated by four nitrogen atoms from a tetrapyrrole macrocycle termed the corrin ring. In the Co^{III} state, the "lower" axial position is usually occupied by an intramolecular 5,6-dimethylbenzimidazole (DMB) base that is tethered to the corrin ring by a nucleotide loop, while the "upper" axial position can be occupied by a variety of different ligands, such as cyanide (CN), water (H_2O), methyl (Me), and adenosyl (Ado) moieties. Cobalamins are remarkably adaptable cofactors, catalyzing a diverse range of reactions in both eukaryotes and prokaryotes.[2–5]

Several spectroscopic techniques have proven invaluable for the investigation of how the cobalamin cofactors perform their functions. The most prevalent among them is electronic absorption (Abs) spectroscopy; cobalamins display rich Abs spectra, as expected given their vivid colors, with features arising primarily from corrin-based $\pi \rightarrow \pi^*$ transitions.[6,7] A "typical" Abs spectrum is characterized by two main features: (i) a pair of closely spaced bands in the visible region termed the α/β peaks, which were attributed to one or multiple corrin $\pi \rightarrow \pi^*$ transition(s) polarized along the corrin long axis (the $C^5 \cdots C^{15}$ vector in Figure 2.1) with some metal-to-ligand charge transfer (MLCT) character, and (ii) an intense, relatively sharp feature in the near-UV region called the γ peak, which was assigned to a corrin $\pi \rightarrow \pi^*$ transition polarized along the corrin short axis (the $Co \cdots C^{10}$ vector).[6,8] Alkylcobalamins such as adenosylcobalamin (AdoCbl)

and methylcobalamin (MeCbl) exhibit "unique" Abs spectra in which the β peak is more intense than the α peak and the γ peak is split into several weaker bands. The cobalamin Abs features are sensitive to a multitude of factors, including corrin ring folding, ligand exchange, changes in axial bond lengths, and the oxidation state of the Co ion, making Abs spectroscopy an excellent tool for B_{12} research.

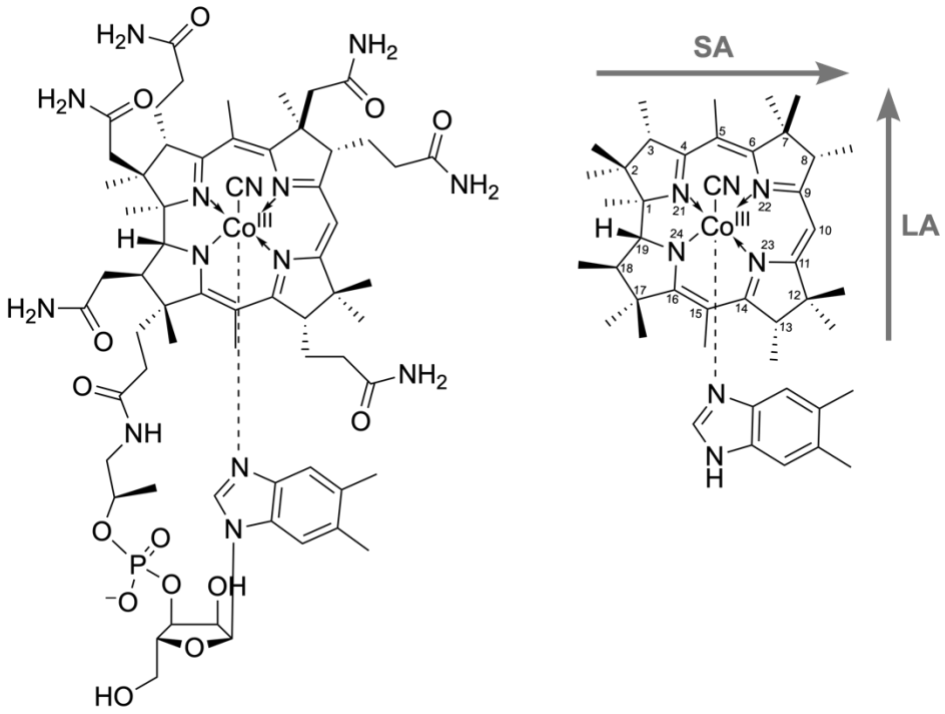


Figure 2.1. Chemical structure of CNCbl (left) and the computational model used in the present study for calculating vibronic coupling (right). The short axis (SA) and long axis (LA) of the corrin ring, along with the numbering scheme used for the relevant atoms, are shown for the computational model.

Another technique that has proven particularly effective in the study of B_{12} cofactors is resonance Raman (rR) spectroscopy,[7–11] wherein laser excitation is used to preferentially enhance vibrational modes along which the chromophore distorts in the resonant electronic excited state. Because cobalamin Abs spectra are dominated by corrin $\pi \rightarrow \pi^*$ transitions, the greatest resonance enhancement occurs for vibrational modes that involve distortions of the conjugated C–C and C–N bonds of the corrin ring.[12] rR spectra of Cbl species are thus dominated by corrin-centered

stretching modes in the high-frequency region (~1500-1625 cm^{-1}). These modes are sensitive to geometric parameters like corrin ring folding, as well as to the Co oxidation state and the identity of axial ligands, all of which impact the extent to which the Co 3d valence orbitals mix with the corrin frontier orbitals.[7,13] rR spectroscopy has been employed, for example, to assess the extent of corrin ring perturbations imposed on Co^{II}cobalamin (Co^{II}Cbl) by the adenosyltransferase PduO.[14] The binding of Co^{II}Cbl to PduO in the presence of co-substrate ATP was found to affect both the frequencies and relative intensities of corrin stretching modes. Through a combination of spectroscopic and computational methods, it was demonstrated that these changes arise from the tight binding imposed by PduO to prevent repositioning of its Co^{II}Cbl substrate.

As a result of electronic coupling between the corrin ring and the axial ligands, laser excitation in resonance with corrin $\pi \rightarrow \pi^*$ transitions also leads to rR enhancement of vibrational modes associated with the axial ligands. Thus, by collecting rR spectra for AdoCbl, MeCbl, and ethylcobalamin (EtCbl) and using isotopic labeling, the modes associated with the Co–C bond stretching and other upper axial ligand vibrations could be identified.[15,16] These data provided invaluable insights into the structural changes accompanying the binding of AdoCbl to B_{12}-dependent enzymes. For instance, rR studies were instrumental in assessing the extent to which the AdoCbl-dependent isomerases methylmalonyl-CoA mutase (MMCM)[13,15–17] and glutamate mutase (GM)[18] activate the Co–C bond in the AdoCbl "ground" state. It was observed that enzyme binding does not significantly affect the frequencies of modes associated with the Ado ligand, indicating that neither MMCM nor GM induces a lengthening of the Co–C bond or otherwise greatly perturbs the cofactor's structure, as would be observed if these enzymes activated the Co–C bond through protein-induced ground state destabilization of the cofactor. Despite these successes, previous rR studies have failed to identify the Co–C stretching mode of CNCbl.

Using spectroscopic data to their full effect requires a theoretical framework within which to interpret spectral signatures. Spectroscopic methods like rR spectroscopy are therefore particularly powerful when combined with computational techniques such as density functional theory (DFT), a quantum mechanical method that has been enormously successful at predicting ground and excited state properties of cobalamins.[19–24] A major drawback of DFT, however, is that different functionals yield different descriptions of the electronic structures of cobalamins. For instance, the popular hybrid functional B3LYP consistently underestimates the Co–C bond dissociation energies (BDEs) of alkylcobalamins,[23,25–27] a useful theoretical benchmark for ground state calculations. This error was suggested to arise from the inclusion of Hartree-Fock exchange and inherent problems with the LYP functional.[27] In support of this proposal, gradient-corrected, pure DFT functionals like BP86 produce BDEs that are in better agreement with experimentally obtained values.[22,27]

While these results appear to suggest that BP86 is an ideal functional for ground state calculations, the appropriate choice for excited state calculations performed using time-dependent DFT (TD-DFT) is less obvious. One notable difference between hybrid and pure generalized gradient approximation (GGA) functionals is that they predict a different number of intense transitions in the α/β region of cobalamin Abs spectra. This is particularly problematic given disputes over the identity of the β band, which has been assigned as either a vibronic sideband of the α transition or as a separate electronic transition. Some recent studies have favored the use of BP86 for excited state calculations for all cobalamin species, based on benchmark studies that compared calculated and experimental Abs, circular dichroism (CD), and magnetic circular dichroism (MCD) spectra.[22,28–32] Other studies have found that B3LYP faithfully reproduces experimental trends like relative band positions and the progressive red-shifting of the α/β peaks

from aquacobalamin (H_2OCbl^+) to CNCbl and MeCbl.[7,33–36] B3LYP also appears to provide a more accurate representation of Abs features of alkylcobalamins like MeCbl and AdoCbl, particularly for transitions with significant $\pi \rightarrow \pi^*$ character.[28,37] Consequently, it has been suggested that although the B3LYP functional is well suited for describing corrin $\pi \rightarrow \pi^*$ transitions, BP86 provides a better description of $d/\pi \rightarrow \pi^*/d$ transitions.[24,28,31,37]

While the combination of spectroscopic and computational studies has afforded a thorough understanding of B_{12} spectral features, lingering uncertainties remain, particularly with regard to vibronic coupling. To address this issue, we employed the recently developed advanced spectral analysis tool implemented in the Orca software package,[38] which can calculate the numerical gradient of excited states at the ground state equilibrium geometry and therefore simulate Abs and fluorescence bandshapes, rR spectra, and rR excitation profiles. Previous computational analyses of the vibrational spectra of cobalamins have either neglected resonance enhancement of vibrational modes[39–42] or considered the relative resonance enhancement of only a select number of corrin stretching modes.[12,14] In this paper, we present a detailed analysis of rR spectra and vibronic coupling in CNCbl as predicted by the B3LYP and BP86 functionals. We used these data to guide further rR spectroscopic studies and identified, for the first time, isotope-sensitive features that can be attributed to the Co–C stretching and Co–C–N bending modes of CNCbl. This finding allowed us to determine the Co–C stretching force constant for CNCbl and compare this value to force constants reported for similar M–C species.

2.2. Materials and Methods

2.2.1. Calculations

Initial coordinates for the complete CNCbl computational model were obtained from a high resolution crystal structure,[43] and the model was solvated in a cubic water box (~35 Å in length)

using the SPC/E water model in GROMACS version 2019.6.[44] A geometry optimization was performed using the quantum mechanics/molecular mechanics (QM/MM) approach as implemented in the ONIOM method of Gaussian 16.[45] The QM region was defined as the Co atom, the corrin ring, the first carbon of each side chain, the DMB ligand, and the CN upper axial ligand. All other atoms, including water molecules, were placed in the MM region. The DFT portion of the calculation was performed using the BP86 functional,[46,47] with the TZVP basis set[48] for Co and all ligating atoms and the def2-SVP basis set[49] for all other QM atoms. The remainder of the cofactor was treated with molecular mechanics, using AMBER-compatible B_{12}-specific parameters taken from the literature.[50]

The coordinates of the QM region of the QM/MM-optimized CNCbl model (including capping hydrogen atoms used to span the QM/MM boundaries) were extracted for TD-DFT calculations with the ORCA 4.2.1 program package.[51] All atoms not explicitly modeled (i.e. atoms placed in the MM region in the geometry optimization step) were included in TD-DFT calculations as point charges. Two series of calculations were performed, one using the BP86 functional[46,47] and the other using the B3LYP functional;[52,53] these calculations were identical in all other respects. The TZVP basis set was employed for Co and its coordinating atoms, while def2-SVP was used for all other atoms. Vertical excitation energies were calculated by the TD-DFT method within the Tamm-Dancoff approximation. For each calculation, the 60 lowest-energy excited states within an energy window of ±4 hartrees with respect to the highest occupied molecular orbital (HOMO)/lowest unoccupied molecular orbital (LUMO) energies were calculated. These data were used to simulate Abs spectra, with each electronic transition assumed to give rise to a Gaussian-shaped band with a full width at half maximum of 1500 cm^{-1}.

Normal mode descriptions and vibrational frequencies were computed by numerical differentiation, with two sets of calculations performed using the same functional and basis set as their respective TD-DFT calculations. For a selection of representative excited states, dimensionless displacements along all normal modes were calculated using the normal mode gradient technique implemented through ORCA's advanced spectral analysis program (ORCA_ASA).[38] These data were used to simulate rR spectra, along with Abs spectra that incorporated vibrational progressions calculated for the excited states of interest. The computed rR frequencies were scaled to align with experimental spectra at the most intense feature at 1501 cm^{-1}, requiring scaling factors of 0.994 and 0.984 for BP86- and B3LYP-simulated spectra, respectively. Vibrational modes were visualized using the PyMOL[54] and Jmol[55] programs. Two simulated isotopic substitutions, H/D exchange at the C10 position and ^{13}C substitution of the Co-bound carbon atom of CN, were performed by manually editing atom masses in the Hessian files and repeating the advanced spectral analysis steps.

2.2.2. Spectroscopy

CNCbl, the chloride salt of aquacobalamin ([H_2OCbl]Cl), and ^{13}C labeled potassium cyanide ($K^{13}CN$) were purchased from Sigma-Aldrich and used as obtained. ^{13}CNCbl was prepared by adding 1 molar equivalent of ^{13}CN to 0.9 molar equivalents of [H_2OCbl]Cl in aqueous solution (**Caution!** KCN is extremely toxic. Great care should be taken to avoid ingestion or skin contact. One should conduct reactions with cyanide in a fume hood and avoid the addition of an acid, as this evolves toxic HCN gas). The formation of ^{13}CNCbl was confirmed using Abs spectroscopy. Frozen rR sample pellets were prepared by injecting drops of 1.0 mM aqueous CNCbl solution into liquid N_2.

Room temperature Abs spectra were collected using a Varian Cary 5e spectrophotometer. rR spectra were obtained upon excitation at 514.5 nm and 363.8 nm with a Coherent I-305 Ar^+ ion laser, with power at the sample kept to ~20 mW to avoid photodecomposition. The ~135° backscattered light was dispersed by an Acton Research triple monochromator (equipped with 1200 and 2400 grooves/mm gratings) and analyzed with a Princeton Instruments Spec X:100 BR deep depletion, back-thinned CCD camera. Spectra were collected at 77 K by placing sample pellets into a quartz finger dewar filled with liquid N_2. The well-resolved shoulder of the ice peak at 314 cm^{-1} was used as an internal standard. The difference spectrum for ^{12}CNCbl and ^{13}CNCbl was obtained after normalizing the intensities of the prominent 634 cm^{-1} peak.

2.2.3. Normal coordinate analysis

A normal coordinate analysis (NCA) was performed on the trinuclear Co–C–N core of CNCbl, using the coordinates obtained from the QM/MM optimization: r(Co–C) = 1.86 Å, r(C–N) = 1.17 Å, and Co–C–N angle = 179.3°. This analysis was based on the Wilson FG matrix method[56] and performed in *MATLAB* (version R2018a; Mathworks, Natick, MA). In addition to the isotope-sensitive rR features at 488 and 503 cm^{-1} identified by us through $^{12}C \rightarrow {}^{13}C$ isotopic substitution, we used the previously reported C–N stretching frequency of 2127 cm^{-1}.[57] See Appendix 2 in the Supporting Information for more details.

2.3. Results

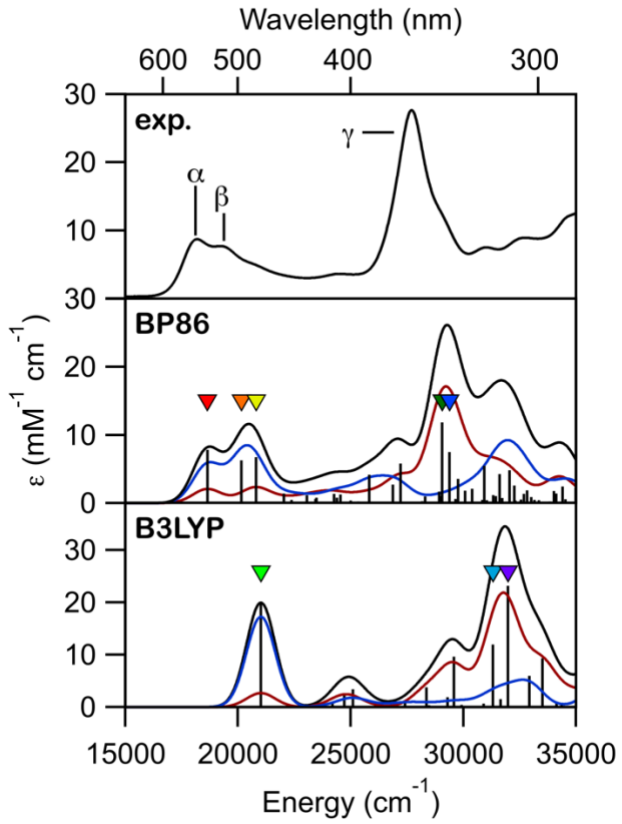


Figure 2.2. Experimental 298 K (top) and TD-DFT computed (center and bottom) Abs spectra of CNCbl (black lines). Vertical sticks indicate individual computed transitions. Separate contributions from LA-polarized and SA-polarized transitions to the computed spectra are shown by blue and red lines, respectively. Colored triangular markers indicate excited states chosen for computing rR spectra. In order of increasing energy, these markers correspond to the following states (i) for BP86, red = α/β(1), orange = α/β(2), yellow = α/β(3), dark green = γ(1), blue = γ(2) and (ii) for B3LYP, green = α/β(1), light blue = γ(1), purple = γ(2).

2.3.1. Electronic Abs Spectroscopy

Figure 2.2 shows the experimental and TD-DFT computed Abs spectra of CNCbl. Both BP86 and B3LYP reproduce key experimental features reasonably well, with one or multiple LA-polarized transition(s) in the α/β region and a single prominent γ feature arising from numerous SA-polarized transitions. The Abs features predicted by B3LYP are uniformly blue-shifted from their experimental and BP86-predicted counterparts, consistent with the known tendency of B3LYP to overestimate transition energies.[7] Notably, BP86 predicts that the α/β region contains

contributions from three similarly intense electronic transitions, all of which are primarily corrin $\pi \rightarrow \pi^*$ in character with a variable degree of $d/\pi \rightarrow \pi^*$ metal-to-ligand charge transfer (MLCT) contributions (Figure 2.3). In contrast, in the B3LYP-computed spectrum, a single feature is present in the α/β region that arises from an almost pure corrin $\pi \rightarrow \pi^*$ excitation, with a much smaller MLCT contribution. Thus, when vibronic coupling is neglected, B3LYP appears to predict an incorrect number of electronic transitions contributing to the α/β region, which was previously noted by Kozlowski and coworkers and used to argue against the use of B3LYP for evaluating excited state properties of Cbls.[24]

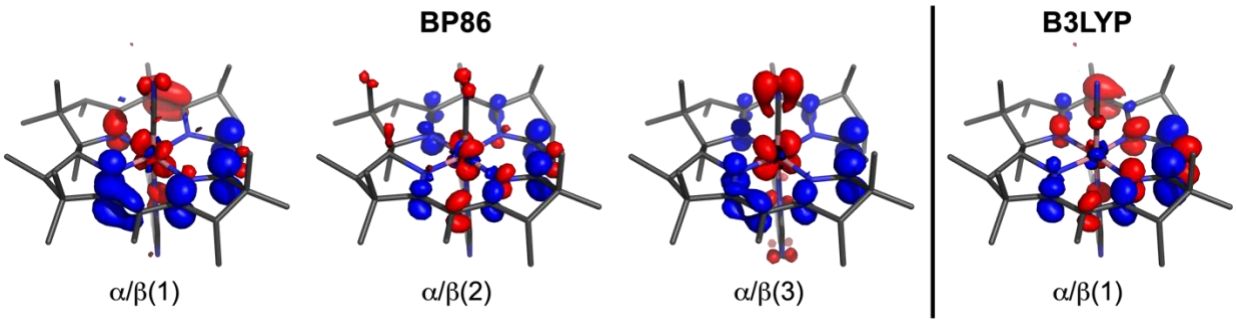


Figure 2.3. Electron density difference maps (EDDMs) for electronic transitions in the α/β region of TD-DFT calculated Abs spectra for CNCbl. Blue and red indicate gain and loss of electron density, respectively.

However, a different picture emerges when vibronic coupling is accounted for through the calculation of excited state distortions along the different normal modes (Figure 2.4). The excited states associated with the three distinct α/β band transitions predicted by BP86 produce a large collection of vibronic features in the computed Abs spectrum, which consequently differs substantially from the experimental spectrum in terms of both the number of peaks and their relative intensities. In contrast, the B3LYP-predicted (with vibronic coupling) spectrum reproduces the overall Abs envelope in the α/β region of the experimental spectrum remarkably well, exhibiting two major peaks and a prominent shoulder separated by ~1500 cm^{-1}.

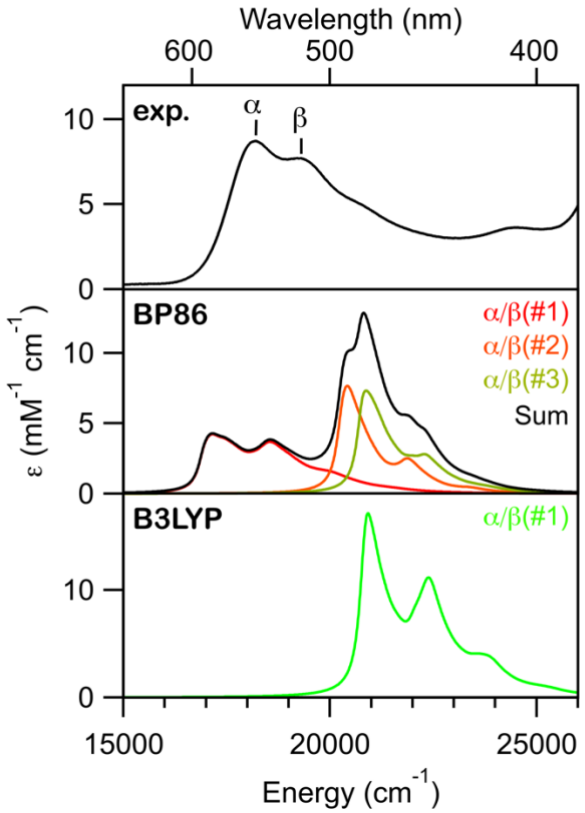


Figure 2.4. α/β region of the experimental (top) and TD-DFT computed with vibronic coupling (center and bottom) Abs spectra of CNCbl. Contributions from individual excited states to the Abs envelopes in the computed spectra are shown in different colors that match the colored markers in Figure 2.2.

2.3.2. rR Spectroscopy

To further assess which functional better accounts for vibronic coupling in CNCbl, rR spectra were computed and compared to our experimental data (Figure 2.5). For excitation in the α/β region, the simulated spectra are in good agreement with the experimental data. Both functionals reproduce the dominant experimental feature at 1501 cm^{-1} and the collection of weaker features observed from 1545 to 1625 cm^{-1}. Of the two functionals, BP86 predicts more significant enhancement of vibrations in the low-frequency region (300-800 cm^{-1}) of the rR spectrum, with numerous prominent low-frequency features that carry as much as ~50% of the intensity of the dominant 1501 cm^{-1} peak. In comparison, the B3LYP-predicted spectrum adheres more closely to the experimental enhancement pattern, with the most intense low-frequency peak reaching only

~20% of the height of the 1501 cm^{-1} peak. This difference in predicted enhancement patterns becomes even more noticeable when comparing the computed rR spectra for excitation in resonance with the two most intense transitions in the γ region, γ(1) and γ(2). In the BP86-predicted spectra, some low-frequency modes are more strongly enhanced than any of the high-frequency modes. In contrast, the high-frequency modes continue to dominate in the B3LYP-predicted spectra, which is in much better agreement with our experimental spectra. BP86 thus appears to perform quite poorly in its description of vibronic coupling for electronic transitions in the γ region of the CNCbl Abs spectrum.

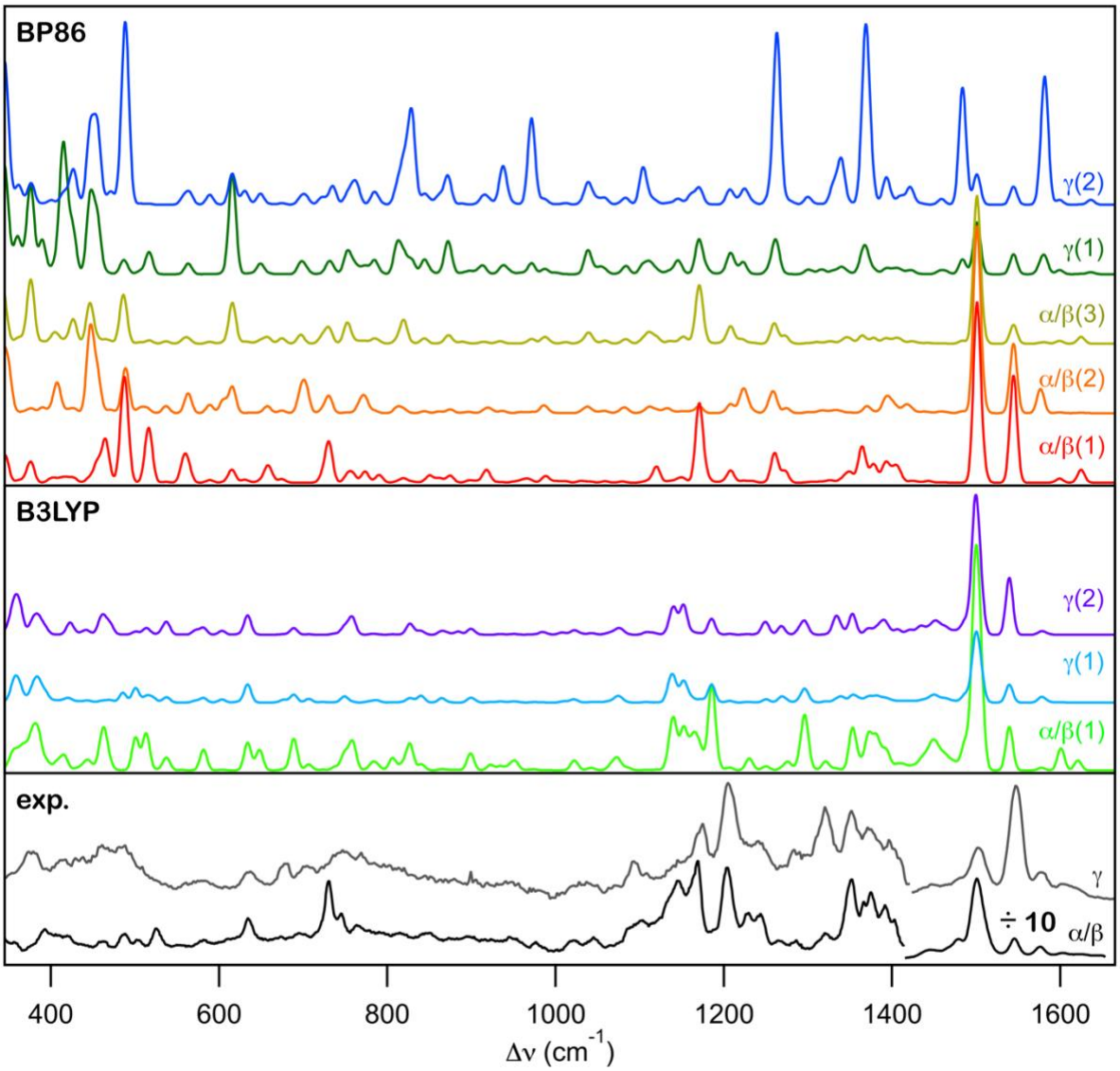


Figure 2.5. BP86- (top) and B3LYP-computed (center) rR spectra of CNCbl, along with experimental spectra (bottom) obtained at 77 K with 514.5 and 363.8 nm excitation (labeled α/β and γ, respectively). Computed frequencies were scaled to align at the intense experimental peak at 1501 cm^{-1} (Δν scaling factors of 0.994 and 0.984 for BP86- and B3LYP-simulated spectra, respectively). Computed spectra are labeled according to the laser excitation energies indicated by the colored markers in Figure 2.2.

Inspection of the compositions of the vibrational modes predicted in the ~1500-1625 cm^{-1} range reveals that they primarily involve stretching motions of the conjugated C–C and C–N bonds of the corrin ring, although slight differences are predicted by the two functionals (see Table 2.1 and Figures 2.6, S4, and S5). Because previous computational studies have provided detailed

assignments of these modes[12,14,40] that are consistent with our results, the present analysis focuses on the predicted rR intensities of the two features at 1501 and ~1545 cm^{-1} that dominate our experimental spectra, designated ν_1 and ν_2, respectively. Experimentally, ν_1 is most significantly enhanced for laser excitation in resonance with the LA-polarized transition(s) in the α/β region (λ_{ex} = 514.5 nm), while ν_2 is preferentially enhanced upon excitation in resonance with the primarily SA-polarized transitions in the γ region (λ_{ex} = 363.8 nm).[8–10] Consistent with these data, both BP86 and B3LYP predict that the relative intensity of ν_1 decreases as the excitation wavelength is moved from the α/β region to the γ region. BP86, however, fails to reproduce the concomitant increase in the intensity of ν_2 that is observed experimentally. In fact, in the BP86-computed rR spectrum for excitation in resonance with the $\gamma(1)$ transition, ν_2 is only weakly enhanced, matched in intensity by bands at 1576 cm^{-1} (ν_3) and 1581 cm^{-1} (ν_{DMB}, which is associated with a mode that mainly involves stretching along the π system of the DMB ligand). For excitation in resonance with the $\gamma(2)$ transition, BP86 predicts the most strongly enhanced feature to be ν_{DMB}. The large enhancement of this mode is unsurprising given that the $\gamma(2)$ transition primarily involves a transfer of electron density from a corrin π-based MO to a DMB π^*-based MO (Figure S2). In comparison, the B3LYP-computed rR spectra for excitation in resonance with the $\gamma(1)$ and $\gamma(2)$ transitions are in much better agreement with the experimental data. Although ν_1 remains the most intense feature, ν_2 becomes the second most intense band overall, with no significant enhancement of any additional modes that lack counterparts in the experimental rR spectrum.

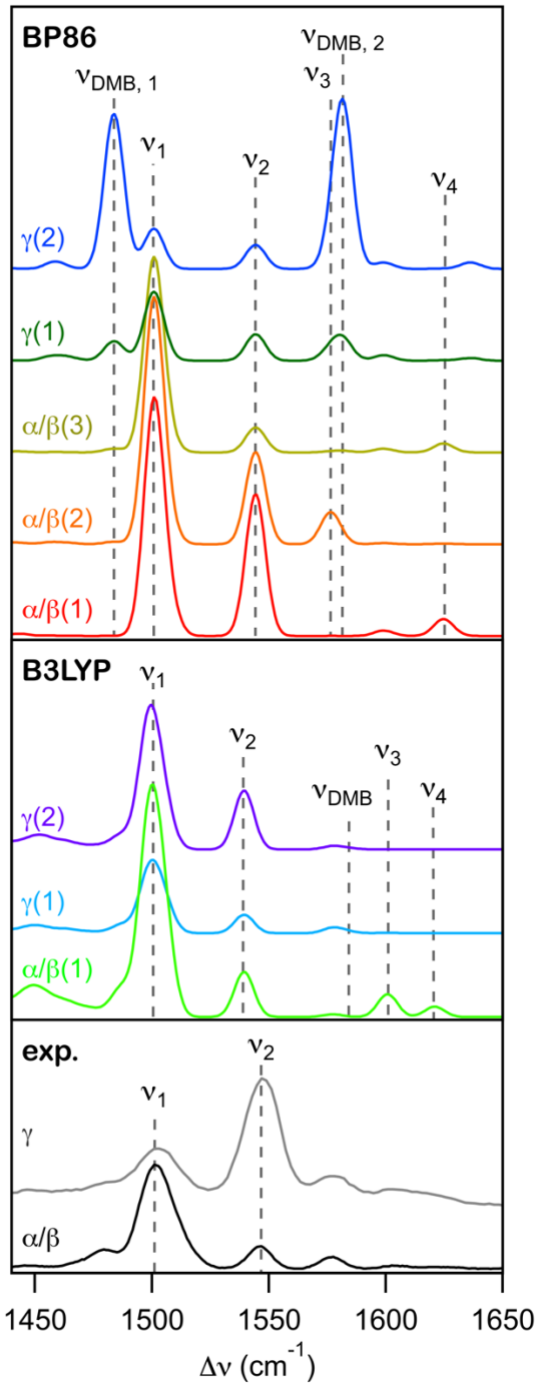


Figure 2.6. High-frequency regions of rR spectra of CNCbl predicted by BP86 (top) and B3LYP (center), as well as experimental spectra (bottom) obtained at 77 K with 514.5 and 363.8 nm excitation (α/β and γ, respectively). Relevant vibrational modes are labeled.

Table 2.1. Computed frequencies and isotope shifts for H/D exchange at C10 of high-frequency vibrational modes with significant rR intensity[a]

BP86		
Mode	Frequency (cm^{-1})[b]	Assignment[c]
$\nu_{DMB,\,1}$	1484 (0)	Stretching mode primarily localized on the DMB ligand
ν_1	1501 (-4)	ν_s with in-phase coupling of SA, LA, and methine stretching
ν_2	1545 (-3)	ν_s with in-phase coupling of LA and methine stretching, out-of-phase with SA stretching
ν_3	1576 (-3)	ν_{as}
$\nu_{DMB,\,2}$	1582 (0)	Stretching mode primarily localized on the DMB ligand
ν_4	1625 (-4)	ν_{as}
B3LYP		
Mode	Frequency (cm^{-1})[b]	Assignment[c]
ν_1	1501 (-3)	Series of connected C–C and C–N stretches along corrin ring, with a minor contribution from DMB stretching
ν_2	1539 (-4)	ν_s with in-phase coupling of LA and methine stretching, out-of-phase with SA stretching
ν_3	1601 (-1)	ν_s with in-phase coupling of LA and SA stretching, out-of-phase with methine stretching
ν_4	1621 (-6)	ν_{as}

[a]See Figures S4 and S5 for DFT-computed eigenvector representations of these modes. [b]Isotope shifts are shown in parentheses. [c]Totally symmetric corrin stretching modes (ν_s) are described as combinations of SA stretching (along the N22–C9 and N23–C11 bonds in Figure 2.1), LA stretching (along the N21–C4 and N24–C16 bonds) and methine stretching (along the C5–C6 and C15–C14 bonds). Antisymmetric corrin stretching modes are denoted ν_{as}.

An interesting pattern is revealed by an *in silico* H/D exchange at the C10 position, which is expected to downshift modes involving SA stretching motions.[12] Experimentally, this substitution impacts only the ν_2 feature at 1545 cm^{-1}, which downshifts by 2 cm^{-1}.[12] In our calculations, however, both BP86 and B3LYP predict a downshift of between 1 and 6 cm^{-1} for all

corrin stretching modes with significant rR intensity (Figure S3 and Table 2.1), possibly due to the truncation of side chains of the corrin ring in our CNCbl computational model.

Turning to the low-frequency region (300-800 cm^{-1}) of the computed rR spectra, of particular interest was the identification of features associated with the Co–C stretch, which to our knowledge have never been observed experimentally for CNCbl.[7,8,11,12] We used *in silico* $^{12}CN \rightarrow {}^{13}CN$ isotopic labeling and simulated rR difference spectra for excitation in resonance with the $\alpha/\beta(1)$ transition in order to identify vibrational modes involving substantial movement of the CN ligand. Rather than a single mode solely associated with the Co–C stretch, both B3LYP and BP86 predict a collection of vibrational modes in the 400-510 cm^{-1} region that involve significant coupling between Co–C stretching (ν_{Co-C}), Co–C–N bending (δ_{Co-C-N}), and corrin ring distortion motions. Some of these modes carry sufficient intensity to cause a visible isotopic shift in the computed difference spectra (Figure 2.7).

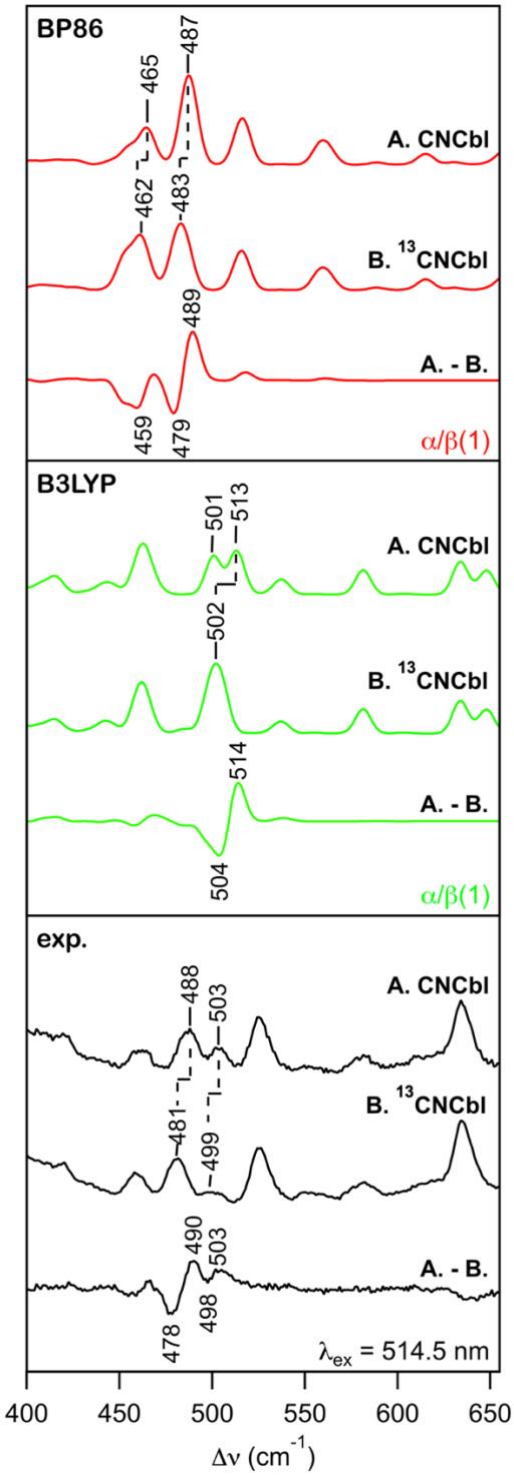


Figure 2.7. Low-frequency regions of computed (top and center) and experimental (bottom) rR spectra of ^{12}CNCbl and ^{13}CNCbl, and their difference spectra. The computed spectra are for excitation in resonance with the α/β(1) transition. The experimental spectra were collected with 514.5 nm excitation and the difference spectrum was generated after normalizing the intensities of the prominent 634 cm^{-1} peak. The energies of rR features that display a visible shift in response to isotopic labeling are indicated.

Animations and vector representations of the computed normal modes allowed for the identification of a common pattern among those modes that display a visible shift in the computed ^{12}CNCbl → ^{13}CNCbl rR difference spectra. The magnitude of the isotopic shift was found to correlate with the amount of Co–C–N bending character, with the largest shifts (5-10 cm^{-1}) involving modes where the Co–C–N angle decreases from 179° to ~140-150° over the course of the vibration. Significant resonance enhancement, meanwhile, was found to require substantial Co–C bond stretching character. For example, B3LYP predicts a relatively pure Co–C–N bending mode (depicted in Figure S7) at 518 cm^{-1} that downshifts by 10 cm^{-1} upon ^{12}CN → ^{13}CN substitution, and although this mode contributes to the corresponding off-resonance Raman spectrum computed using B3LYP (Figure S6), its complete lack of resonance enhancement means that it does not contribute to the computed rR spectra. The visible isotopic shifts of features in the computed rR difference spectra therefore arise from vibrational modes that contain both Co–C bond stretching and Co–C–N bending character (Figure 2.8). This finding emphasizes the importance of accounting for resonance enhancement to facilitate a direct comparison between computed and experimental rR data.

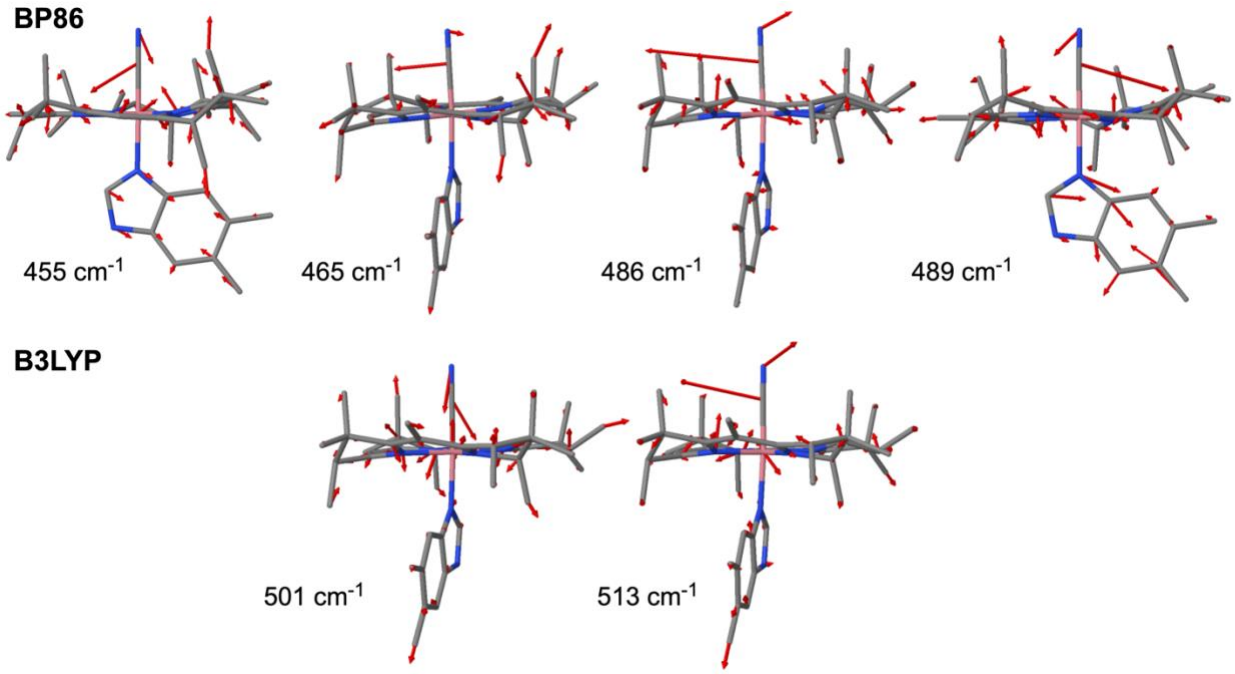


Figure 2.8. DFT-computed eigenvector representations and frequencies of vibrational modes of CNCbl that are predicted by BP86 (top) and B3LYP (bottom) to display visible shifts in ^{12}CN → ^{13}CN rR difference spectra.

The presence of relatively intense features in the 400-510 cm^{-1} region of the computed rR difference spectra predicted by both functionals prompted us to also collect experimental rR spectra for ^{13}CNCbl using 514.5 nm excitation. Two features were found to be sensitive to ^{12}CN→ ^{13}CN isotopic substitution, one at 488 cm^{-1} that downshifts by 7 cm^{-1} and another at 503 cm^{-1} that downshifts by 4 cm^{-1} (Figure 2.7). On the basis of a normal coordinate analysis (see below), these closely spaced features can be assigned as Co–C stretching and Co–C–N bending modes, respectively. The low intensity of these features compared to those associated with the Co–C stretching modes of MeCbl and AdoCbl indicates that the α/β transition(s) in CNCbl are largely localized on the corrin ring, with very little change in electron density along the Co–C bond. Collectively, these data suggest that the experimental identification of the Co–C stretching mode of CNCbl had been hindered by the weak resonance enhancement, the large extent of mixing of

vibrational motions, and the presence of numerous, similarly intense features in the low-frequency region of rR spectra.

2.4. Discussion

Despite the central role that Abs and rR spectroscopies have played in the study of B_{12} and B_{12}-dependent enzymes, the interpretation of spectra obtained with these techniques has remained a matter of debate. In the present study we have used the ORCA_ASA program, which computes excited state gradients and dimensionless displacements along normal modes, to predict Abs bandshapes and rR spectra for CNCbl. These computational data, in conjunction with high quality experimental spectra and $^{12}CN \rightarrow {}^{13}CN$ substitution, have afforded us a unique opportunity to assess whether BP86 or B3LYP, two of the most popular functionals, provides a better description of the excited states of CNCbl.

2.4.1. Impact of vibronic coupling on Abs bandshape

One particularly disputed aspect of the Abs spectra of CNCbl and its derivatives is the origin of the β band, found at ~19,300 cm^{-1} for CNCbl, which has been attributed to either a distinct electronic transition or a vibrational sideband of an electronic transition for which the α band serves as the electronic origin. rR studies of CNCbl revealed that a mode at ~1500 cm^{-1} is strongly coupled to the α-band transition, lending credence to the proposal of a vibrational progression.[12] Other, more advanced spectroscopic techniques have been used to address this question, with mixed conclusions, but also favoring the vibronic progression model. For example, triply resonant sum frequency (TRSF) spectroscopy, a multidimensional spectroscopic technique capable of identifying vibronic coupling, yielded results consistent with a vibrational progression, although the authors were unable to definitively distinguish between the two models.[58] In another recent study, ultrafast transient Abs spectroscopy was used to demonstrate that the bulk of the Abs

intensity in the α/β region is due to a vibrational progression associated with a single electronic transition, with the possibility of a minor contribution from another electronic transition.[59]

Although these experimental data favor the existence of a vibrational progression in the α/β region, previous computational investigations of this region were complicated by the fact that the choice of functional greatly impacts the description of TD-DFT predicted excited states. Consistent with the extensive literature on this subject, the pure GGA functional BP86 predicts three electronic transitions of significant intensity in the α/β region, all of which can be characterized as d/π → π* (MLCT) and π → π* excitations (Figure 2.3). The substantial MLCT character is due in large part to significant (between 5 and 57%) contributions of the Co $3d_{xz}$, $3d_{yz}$, and $3d_{z2}$ orbitals to the three highest occupied molecular orbitals (MOs). B3LYP, by contrast, predicts a single π → π* transition with very little MLCT character. Because several features are present in the α/β region of the experimental Abs spectrum of CNCbl, this apparent discrepancy has been used to argue that BP86 is a better choice than B3LYP for describing the excited states of B_{12} species.

However, traditional TD-DFT computations of Abs spectra consider only the Franck-Condon (i.e., vertical) transitions, ignoring any vibrational progressions. By accounting for vibronic coupling, a different—and more complete—picture of Abs bandshapes is obtained. In the case of CNCbl, the three electronic transitions in the α/β region predicted by BP86 produce a complex pattern of vibrational progressions that is in poor agreement with our experimental Abs data (Figure 2.4). By contrast, the single transition in the α/β region predicted by B3LYP yields an Abs bandshape that is in remarkable agreement with the experimental spectrum. It is apparent that, when vibronic coupling is considered, only B3LYP successfully reproduces the CNCbl Abs

envelope in the α/β region, and thus we conclude that the β band arises from a vibrational progression rather than a separate electronic transition.

2.4.2. Incorporating resonance enhancement into computed Raman spectra

Interpreting the complex vibrational spectra of cobalamins has been greatly aided by the development of new computational techniques over the past few decades. Kozlowski and coworkers developed a DFT-based approach for the calculation of vibrational frequencies and off-resonance Raman intensities using a DFT-scaled quantum mechanical forcefield.[39–42] These studies successfully reproduced experimental spectra and afforded new insight into the origin of some vibrational features, particularly those associated with the upper axial ligand. In the case of AdoCbl, for example, it was shown that a rR peak traditionally attributed to the Co–C stretch actually arises from a Co–C–C bending mode, and that the Co–C stretching motion is strongly coupled with ribose deformation modes.[41] However, the direct comparison of these computational predictions with experimental data was hindered by the fact that resonance enhancement was not accounted for. In an earlier attempt to address this issue, we evaluated excited state distortions through DFT calculations with the PBE functional in order to estimate relative rR intensities for a subset of corrin-based vibrational modes of CNCbl, Co^{II}Cbl, and Co^{I}cobalamin (Co^{I}Cbl).[12,14] While this approach was used successfully to reproduce experimental trends regarding the frequencies and relative rR intensities of these modes, it is not suitable for calculating complete rR spectra of cobalamins.

In the present study, we used the ORCA_ASA tool to compute, for the first time, complete rR spectra for any cobalamin species; i.e., CNCbl. The most conspicuous difference between the BP86- and B3LYP-computed rR spectra is the predicted enhancement of the low-frequency (300-800 cm^{-1}) modes relative to the high-frequency (~1500-1625 cm^{-1}) modes (Figure 2.5). Both

functionals reproduce the experimental observation that excitation in the α/β region most strongly enhances a mode at 1501 cm^{-1}, but the relative intensities of low-frequency features are considerably overestimated in the BP86-computed rR spectra. We attribute this discrepancy to the large amount of MLCT character of the transitions in the α/β region of the BP86-predicted Abs spectrum (Figure 2.3). The tendency of BP86 to overestimate the rR intensities of low-frequency features is even more pronounced for excitation in the γ region. Most notably, in the BP86-computed rR spectrum for excitation in resonance with the γ(2) transition, the dominant feature arises from a DMB-centered vibration, ν_{DMB}, that has no counterpart in the experimental spectrum. By comparison, the B3LYP-computed rR spectra for excitation in the γ region are in much better agreement with the experimental data.

2.4.3. Nature of the Co–C bond

To identify vibrational modes associated with the upper axial CN ligand of CNCbl, we performed *in silico* $^{12}CN \rightarrow {}^{13}CN$ isotopic labeling and evaluated the effect of this substitution on computed rR spectra. Although labeling of the Co-bound C atom was previously used experimentally to identify the Co–C stretching modes of several alkylcobalamins,[15,16] this approach has never been employed successfully in past investigations of CNCbl. It was therefore intriguing to discover that, for excitation in the α/β region, both BP86 and B3LYP predict reasonably intense isotope-sensitive rR features (at 465 and 487 cm^{-1} for BP86 and at 501 and 513 cm^{-1} for B3LYP). These data suggested that it should be possible to identify experimental rR features of CNCbl that are sensitive to $^{12}CN \rightarrow {}^{13}CN$ isotopic labeling. We therefore collected rR spectra of ^{12}CNCbl and ^{13}CNCbl, using sample pellets rather than an NMR tube to reduce quartz background contributions to the low-frequency region. Isotope-sensitive peaks were identified at 488 and 503 cm^{-1}, downshifting by 7 and 4 cm^{-1}, respectively, upon isotopic labeling (Figure 2.7).

Distinguishing between the Co–C stretching and Co–C–N bending modes is no straightforward task. rR studies of cyanide complexes of hemeproteins revealed that the order of the Fe–C stretching and Fe–C–N bending modes (i.e., whether $\nu_{Co–C}$ has a higher or lower frequency than $\delta_{Co–C–N}$) can vary.[60] A normal coordinate analysis (NCA) demonstrated that the order depends on the Fe–C–N bond angle,[61] with the lower frequency mode assuming greater $\nu_{Fe–C}$ stretching character as the Fe–C–N angle decreases. Thus, drawing conclusions about the Co–C bond strength of CNCbl from our rR spectra requires an approach that can account for the possibility of mixing among isotope-sensitive vibrational modes.

To that end, we performed a NCA within the framework of Wilson's F and G matrix formalism.[56] Unlike a basic, isolated harmonic oscillator approximation, this method accounts for mechanical coupling between different motions. For this analysis, we only considered the (nearly) linear Co–C–N unit, which contains three internal coordinates: the Co–C and C–N stretching coordinates and the Co–C–N bending coordinate. The bond lengths and angle required for the G matrix, which encompasses the masses and spatial relationships among atoms, were taken from the QM/MM-optimized structure of CNCbl. The force constants were fitted to three experimental frequencies, namely, the isotope-sensitive modes observed here at 488 and 503 cm^{-1} along with the C–N bond stretching frequency previously identified in the infrared (IR) spectrum of ^{12}CNCbl at 2127 cm^{-1}.[57]

Table 2.2. Experimental and calculated vibrational frequencies, calculated force constants, and potential energy distribution (PED) contributions for normal modes associated with the Co–C–N unit

Mode	Exp. Freq. (cm^{-1})	Calc. Freq. (cm^{-1})	Force constant[a]	PED contributions (%)		
				Co–C	C–N	Co–C–N
$\nu_{Co–C}$	488	488	2.67	75	24	1
$\nu_{C–N}$	2127	2126	16.4	1	99	0
$\delta_{Co–C–N}$	503	502	0.677	15	5	81

[a]Stretching and bending force constants are given in mdyn/Å and mdyn Å/rad^2, respectively.

As shown in Table 2, this NCA reproduces the three relevant experimental frequencies very well. Interestingly, the potential energy distribution (PED) contributions to the different normal modes indicate that while the Co–C–N bending mode, $\delta_{Co–C–N}$, involves substantial Co–C stretching motion, as expected on the basis of our DFT-based normal mode analysis (Figure 2.8), the Co–C stretching mode, $\nu_{Co–C}$, has negligible Co–C–N bending character. This difference between the NCA-predicted and DFT-computed normal mode descriptions is not surprising, considering that the former ignores the complexity and low symmetry of the corrin ring, which creates an asymmetric geometric and electronic environment for the Co–C–N unit.

The fitted force constants obtained from the NCA are $k_{Co–C}$ = 2.67 mdyn/Å, $k_{C–N}$ = 16.4 mdyn/Å, and $k_{Co–C–N}$ = 0.677 mdyn Å/rad^2. To assess the impact of mechanical coupling, the force constant $k_{Co–C}$ was used to predict the Co–C bond stretching frequency within the harmonic oscillator approximation,

$$k = (5.89 \times 10^{-7})\mu\tilde{\nu}^2 \qquad (1)$$

Here, k is the force constant in mdyn/Å, μ is the reduced mass in g/mol, and $\tilde{\nu}$ is the vibrational frequency in cm^{-1}. By treating CN as a single dynamic unit, the effective mass of CN can be approximated as being equal to the total mass of CN, yielding μ = 18.04 g/mol. By using this value

in Eq 1, a value of ν_{Co-C} = 501 cm^{-1} is obtained, which is marginally higher than our experimental value of 488 cm^{-1}. Mixing of the Co–C stretching, C–N stretching, and Co–C–N bending motions thus only slightly depresses the Co–C stretching frequency.

With a force constant of 2.67 mdyn/Å, the Co–C bond is much stronger in CNCbl than in the alkylcobalamins MeCbl, EtCbl, and AdoCbl, for which Co–C force constants of 1.85, 1.77, and 1.50 mdyn/Å, respectively, were determined experimentally.[16] This finding is consistent with the shorter Co–C bond of 1.868 Å in CNCbl[43] compared to 1.979 Å in MeCbl,[43] 2.023 Å in EtCbl,[62] and 2.030 Å in AdoCbl.[63] Although it is tempting to attribute the increased Co–C bond strength in CNCbl to π-backbonding interactions between occupied Co 3d orbitals and empty CN π* orbitals, previous IR and rR studies of other tetrapyrroles challenged this assumption. For example, it was found that the C–N stretching mode has a higher frequency in siroheme-CN adducts than in free CN^-, indicating an increase in C–N bond strength.[64] Because π-donation into the CN π* MOs would weaken the C–N bond, the authors concluded that there is considerably less π-backbonding in the cyanide adducts than in the corresponding carbon monoxide (CO) and nitric oxide (NO) adducts. Similarly, the C–N stretching mode frequency is higher in CNCbl (2127 cm^{-1})[57] than in free CN^- (2079 cm^{-1}),[65] which likewise rules out the possibility of significant π-backbonding. The strengthening of the C–N bond in CNCbl is likely due, at least in part, to σ donation from CN to the Co ion, given that ν_{C-N} also increases from free CN^- to hydrogen cyanide (2093 cm^{-1}).[65]

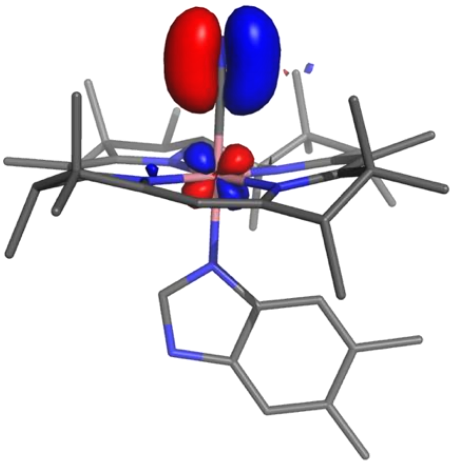


Figure 2.9. Isosurface plot of a representative DFT/B3LYP-computed occupied MO of CNCbl (HOMO-6) that shows a π-antibonding interaction between a Co 3d orbital and a CN π orbital.

Further evidence against π-backbonding in CNCbl is provided by the computed MO compositions for our CNCbl model. DFT computations with neither the BP86 nor the B3LYP functional predict any MOs indicative of π-backbonding, which would be expected to show a π-bonding interaction between a Co 3d orbital and a CN π* orbital. Instead, both predict several occupied MOs that show a sizable π-antibonding interaction between a Co 3d orbital and a CN π orbital (Figure 2.9). This interaction does not ultimately result in a net increase in Co–C bond strength, as both the π-bonding and π-antibonding combinations of the Co 3d and CN π orbitals are occupied, but it sheds light on why π-backbonding is not observed in CNCbl. The strong interaction between Co 3d and CN π orbitals requires that they have similar energies, which implies that the empty CN π* orbitals are too high in energy to mix with the relevant filled Co 3d orbitals. This finding provides an interesting contrast to our previous work on nitrosylcobalamin (NOCbl),[66] for which an occupied Co-based π-backbonding MO was readily identifiable. Because π-backbonding can be ruled out as a significant contributor, we attribute the strong Co–C bond in CNCbl to electrostatic interactions. It was demonstrated through a natural bond order (NBO) analysis that the Co–C bond of CNCbl is more ionic in nature than that of MeCbl,[67] which is

expected to result in an enhanced electrostatic attraction between the cyanide anion and the more positively charged Co ion.

2.5. Summary and Conclusions

The results obtained in the present study indicate that when vibronic coupling is incorporated into TD-DFT computations for CNCbl, B3LYP outperforms BP86 in two important ways: (i) it successfully reproduces the distinctive Abs envelope in the α/β region and (ii) it correctly predicts the relative intensities of low-frequency and high-frequency features in computed rR spectra. These findings have important implications for future computational investigations of cobalamins, particularly because some studies have argued that B3LYP is not an appropriate choice for such calculations. Computing Abs bandshapes in the α/β region also sheds light on the much-disputed origin of the β band, providing substantial computational evidence that this feature is not a distinct electronic transition, but rather the first member of a vibrational progression for which the α band serves as the electronic origin.

Importantly, introducing resonance enhancement into the computation of Raman spectra also indicated that a $^{12}CN \rightarrow {}^{13}CN$ isotopic substitution should result in visible shifts of vibrational modes associated with the axial Co–C bond. This finding guided experimental rR experiments that allowed us to identify features associated with the Co–C stretching (488 cm^{-1}) and Co–C–N bending (503 cm^{-1}) modes in the rR spectrum of CNCbl, key experimental features that have eluded identification until now. From an analysis of these data within the framework of a normal coordinate analysis, we determined that the force constant of the Co–C bond in CNCbl is 2.67 mdyn/Å, and thus considerably larger than the Co–C bond force constant reported for MeCbl (1.85 mdyn/Å).[16]

References

(1) Robscheit-Robbins, F. S.; Whipple, G. H. Blood Regeneration in Severe Anemia: Favorable Influence of Liver, Heart and Skeletal Muscle in Diet. *Am. J. Physiol. Content* **1925**, *72* (3), 408–418. https://doi.org/10.1152/ajplegacy.1925.72.3.408.

(2) Banerjee, R.; Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **2003**, *72*, 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828.

(3) Brown, K. L. Chemistry and Enzymology of Vitamin B_{12}. *Chemical Reviews*. June 2005, pp 2075–2149. https://doi.org/10.1021/cr030720z.

(4) Banerjee, R. Radical Carbon Skeleton Rearrangements: Catalysis by Coenzyme B_{12}-Dependent Mutases. *Chem. Rev.* **2003**, *103* (6), 2083–2094. https://doi.org/10.1021/cr0204395.

(5) Bridwell-Rabb, J.; Drennan, C. L. Vitamin B_{12} in the Spotlight Again. *Curr. Opin. Chem. Biol.* **2017**, *37*, 63–70. https://doi.org/10.1016/J.CBPA.2017.01.013.

(6) Pratt, J. M. The Roles of Co, Corrin, and Protein. II. Electronic Spectra and Structure of the Corrin Ligand: Molecular Machinery of the Protein. In *Chemistry and Biochemistry of B_{12}*; Banerjee, R., Ed.; Wiley: New York, 1999; pp 113–164.

(7) Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914. https://doi.org/10.1021/ja029328d.

(8) Salama, S.; Spiro, T. G. Visible and Near-Ultraviolet Resonance Raman Spectra of Photolabile Vitamin B_{12} Derivatives with a Rapid-flow Technique. *J. Raman Spectrosc.* **1977**, *6* (2), 57–60. https://doi.org/10.1002/jrs.1250060202.

(9) George, W. O.; Mendelsohn, R. Resonance Raman Spectrum of Cyanocobalamin (Vitamin B_{12}). *Appl. Spectrosc.* **1973**, *27* (5), 390–391. https://doi.org/10.1366/000370273774333326.

(10) Mayer, E.; Gardiner, D. J.; Hester, R. E. Resonance Raman Spectra of Vitamin B_{12} and Some Cobalt Corrinoid Derivatives. *J. Chem. Soc. Faraday Trans. 2 Mol. Chem. Phys.* **1973**, *69*, 1350–1358. https://doi.org/10.1039/F29736901350.

(11) Galluzzi, F.; Garozzo, M.; Ricci, F. F. Resonance Raman Scattering and Vibronic Coupling in Aquo- and Cyano-cobalamin. *J. Raman Spectrosc.* **1974**, *2* (4), 351–362. https://doi.org/10.1002/jrs.1250020405.

(12) Park, K.; Brunold, T. C. Combined Spectroscopic and Computational Analysis of the Vibrational Properties of Vitamin B_{12} in Its Co^{3+}, Co^{2+}, and Co^{1+} Oxidation States. *J.*

Phys. Chem. B **2013**, *117* (18), 5397–5410. https://doi.org/10.1021/jp309392u.

(13) Dong, S.; Padmakumar, R.; Maiti, N.; Banerjee, R.; Spiro, T. G. Resonance Raman Spectra Show That Coenzyme B_{12} Binding to Methylmalonyl-Coenzyme A Mutase Changes the Corrin Ring Conformation but Leaves the Co-C Bond Essentially Unaffected. *J. Am. Chem. Soc.* American Chemical Society 1998, pp 9947–9948. https://doi.org/10.1021/ja981584w.

(14) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Resonance Raman Spectroscopic Study of the Interaction between Co(II)rrinoids and the ATP:Corrinoid Adenosyltransferase PduO from Lactobacillus Reuteri. *J. Biol. Inorg. Chem.* **2016**, *21* (5–6), 669–681. https://doi.org/10.1007/s00775-016-1371-x.

(15) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Resonance Raman Co−C Stretching Frequencies Reflect Bond Strength Changes in Alkyl Cobalamins, but Are Unaffected by Trans Ligand Substitution. *J. Am. Chem. Soc.* **1996**, *118* (38), 9182–9183. https://doi.org/10.1021/ja962003a.

(16) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Force Constants from Resonance Raman Spectra of Alkylcobalamins: Insensitivity to Dimethylbenzylimidazole Coordination. *Inorganica Chim. Acta* **1998**, *270* (1), 392–398. https://doi.org/https://doi.org/10.1016/S0020-1693(97)05994-X.

(17) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Bond Activation in B_{12}-Dependent Enzymes: Cryogenic Resonance Raman Studies of Methylmalonyl-Coenzyme A Mutase. *J. Am. Chem. Soc.* **1999**, *121* (30), 7063–7070. https://doi.org/10.1021/ja982753f.

(18) Huhta, M. S.; Chen, H. P.; Hemann, C.; Hille, C. R.; Marsh, E. N. Protein-Coenzyme Interactions in Adenosylcobalamin-Dependent Glutamate Mutase. *Biochem. J.* **2001**, *355* (1), 131–137. https://doi.org/10.1042/0264-6021:3550131.

(19) Kozlowski, P. M. Quantum Chemical Modeling of Co-C Bond Activation in B_{12}-Dependent Enzymes. *Current Opinion in Chemical Biology*. Elsevier Current Trends December 1, 2001, pp 736–743. https://doi.org/10.1016/S1367-5931(01)00273-3.

(20) Mamun, A. Al; Toda, M. J.; Kozlowski, P. M. Can Photolysis of the Co–C Bond in Coenzyme B_{12} -Dependent Enzymes Be Used to Mimic the Native Reaction? *J. Photochem. Photobiol. B Biol.* **2019**, *191*, 175–184. https://doi.org/10.1016/j.jphotobiol.2018.12.018.

(21) Brunold, T. C.; Conrad, K. S.; Liptak, M. D.; Park, K. Spectroscopically Validated Density Functional Theory Studies of the B_{12} Cofactors and Their Interactions with Enzyme Active Sites. *Coordination Chemistry Reviews*. March 2009, pp 779–794. https://doi.org/10.1016/j.ccr.2008.09.013.

(22) Kuta, J.; Patchkovskii, S.; Zgierski, M. Z.; Kozlowski, P. M. Performance of DFT in Modeling Electronic and Structural Properties of Cobalamins. *J. Comput. Chem.* **2006**, *27*

(12), 1429–1437. https://doi.org/10.1002/jcc.20454.

(23) Jensen, K. P.; Ryde, U. Cobalamins Uncovered by Modern Electronic Structure Calculations. *Coordination Chemistry Reviews*. Elsevier March 1, 2009, pp 769–778. https://doi.org/10.1016/j.ccr.2008.04.015.

(24) Toda, M. J.; Kozlowski, P. M.; Andruniów, T. Assessing Electronically Excited States of Cobalamins via Absorption Spectroscopy and Time-Dependent Density Functional Theory. In *Transition Metals in Coordination Environments: Computational Chemistry and Catalysis Viewpoints*; Broclawik, E., Borowski, T., Radoń, M., Eds.; Springer International Publishing, 2019; pp 219–258. https://doi.org/10.1007/978-3-030-11714-6_8.

(25) Andruniow, T.; Zgierski, M. Z.; Kozlowski, P. M. Theoretical Determination of the Co–C Bond Energy Dissociation in Cobalamins. *Journal of the American Chemical Society*. American Chemical Society 2001, pp 2679–2680. https://doi.org/10.1021/ja0041728.

(26) Kozlowski, P. M.; Kumar, M.; Piecuch, P.; Li, W.; Bauman, N. P.; Hansen, J. A.; Lodowski, P.; Jaworska, M. The Cobalt-Methyl Bond Dissociation in Methylcobalamin: New Benchmark Analysis Based on Density Functional Theory and Completely Renormalized Coupled-Cluster Calculations. *J. Chem. Theory Comput.* **2012**, *8* (6), 1870–1894. https://doi.org/10.1021/ct300170y.

(27) Jensen, K. P.; Ryde, U. Theoretical Prediction of the Co-C Bond Strength in Cobalamins. *J. Phys. Chem. A* **2003**, *107* (38), 7539–7545. https://doi.org/10.1021/jp027566p.

(28) Andruniów, T.; Jaworska, M.; Lodowski, P.; Zgierski, M. Z.; Dreos, R.; Randaccio, L.; Kozlowski, P. M. Time-Dependent Density Functional Theory Study of Cobalt Corrinoids: Electronically Excited States of Methylcobalamin. *J. Chem. Phys.* **2008**, *129* (8), 085101. https://doi.org/10.1063/1.2956836.

(29) Kornobis, K.; Kumar, N.; Wong, B. M.; Lodowski, P.; Jaworska, M.; Andruniów, T.; Ruud, K.; Kozlowski, P. M. Electronically Excited States of Vitamin B_{12}: Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated Ab Initio Methods. *J. Phys. Chem. A* **2011**, *115* (7), 1280–1292. https://doi.org/10.1021/jp110914y.

(30) Kornobis, K.; Ruud, K.; Kozlowski, P. M. Cob(I)Alamin: Insight into the Nature of Electronically Excited States Elucidated via Quantum Chemical Computations and Analysis of Absorption, CD and MCD Data. *J. Phys. Chem. A* **2013**, *117* (5), 863–876. https://doi.org/10.1021/jp310446c.

(31) Solheim, H.; Kornobis, K.; Ruud, K.; Kozlowski, P. M. Electronically Excited States of Vitamin B_{12} and Methylcobalamin: Theoretical Analysis of Absorption, CD, and MCD Data. *J. Phys. Chem. B* **2011**, *115* (4), 737–748. https://doi.org/10.1021/jp109793r.

(32) Navizet, I.; Perry, C. B.; Govender, P. P.; Marques, H. M. Cis Influence in Models of Cobalt Corrins by DFT and TD-DFT Studies. *J. Phys. Chem.* **2012**, *116* (30), 8836–8845.

https://doi.org/10.1039/c3dt50336d.

(33) Jaworska, M.; Lodowski, P. Electronic Spectrum of Co-Corrin Calculated with the TDDFT Method. *J. Mol. Struct. THEOCHEM* **2003**, *631* (1–3), 209–223. https://doi.org/10.1016/S0166-1280(03)00249-5.

(34) Conrad, K. S.; Brunold, T. C. Spectroscopic and Computational Studies of Glutathionylcobalamin: Nature of Co-S Bonding and Comparison to Co-C Bonding in Coenzyme B_{12}. *Inorg. Chem.* **2011**, *50* (18), 8755–8766. https://doi.org/10.1021/ic200428r.

(35) Reig, A. J.; Conrad, K. S.; Brunold, T. C. Combined Spectroscopic/Computational Studies of Vitamin B_{12} Precursors: Geometric and Electronic Structures of Cobinamides. *Inorg. Chem.* **2012**, *51* (5), 2867–2879.

(36) Conrad, K. S.; Jordan, C. D.; Brown, K. L.; Brunold, T. C. Spectroscopic and Computational Studies of Cobalamin Species with Variable Lower Axial Ligation: Implications for the Mechanism of Co–C Bond Activation by Class I Cobalamin-Dependent Isomerases. *Inorg. Chem.* **2015**, *54* (8), 3736–3747. https://doi.org/10.1021/ic502665x.

(37) Andruniów, T.; Jaworska, M.; Lodowski, P.; Zgierski, M. Z.; Dreos, R.; Randaccio, L.; Kozlowski, P. M. Time-Dependent Density Functional Theory Study of Cobalt Corrinoids: Electronically Excited States of Coenzyme B_{12}. *J. Chem. Phys.* **2009**, *131* (10), 105105. https://doi.org/10.1063/1.3190326.

(38) Petrenko, T.; Neese, F. Analysis and Prediction of Absorption Band Shapes, Fluorescence Band Shapes, Resonance Raman Intensities, and Excitation Profiles Using the Time-Dependent Theory of Electronic Spectroscopy. *J. Chem. Phys.* **2007**, *127* (16), 164319. https://doi.org/10.1063/1.2770706.

(39) Andruniów, T.; Zgierski, M. Z.; Kozlowski, P. M. DFT-SQM Force Field for Cobalt Corrinoids. *Chem. Phys. Lett.* **2000**, *331* (5–6), 502–508. https://doi.org/10.1016/S0009-2614(00)01210-0.

(40) Andruniow, T.; Zgierski, M. Z.; Kozlowski, P. M. Vibrational Analysis of Methylcobalamin. *J. Phys. Chem. A* **2002**, *106* (7), 1365–1373. https://doi.org/10.1021/jp013271k.

(41) Kozlowski, P. M.; Andruniow, T.; Jarzecki, A. A.; Zgierski, M. Z.; Spiro, T. G. DFT Analysis of Co-Alkyl and Co-Adenosyl Vibrational Modes in B_{12}-Cofactors. *Inorg. Chem.* **2006**, *45* (14), 5585–5590. https://doi.org/10.1021/ic052069j.

(42) Mieda-Higa, K.; Mamun, A. Al; Ogura, T.; Kitagawa, T.; Kozlowski, P. M. Resonance Raman Investigation of Dithionite-Reduced Cobalamin. *J. Raman Spectrosc.* **2020**, *51* (8), 1331–1342. https://doi.org/10.1002/jrs.5909.

(43) Randaccio, L.; Furlan, M.; Geremia, S.; Šlouf, M.; Srnova, I.; Toffoli, D. Similarities and

Differences between Cobalamins and Cobaloximes. Accurate Structural Determination of Methylcobalamin and of LiCl- and KCl-Containing Cyanocobalamins by Synchrotron Radiation. *Inorg. Chem.* **2000**, *39* (15), 3403–3413. https://doi.org/10.1021/ic0001199.

(44) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. E.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision C.01. *Gaussian, Inc., Wallingford CT,*. 2016.

(46) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098. https://doi.org/10.1103/PhysRevA.38.3098.

(47) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33* (12), 8822–8824. https://doi.org/10.1103/PhysRevB.33.8822.

(48) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.* **1994**, *100* (8), 5829–5835. https://doi.org/10.1063/1.467146.

(49) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1998**, *97* (4), 2571. https://doi.org/10.1063/1.463096.

(50) Marques, H. M.; Ngoma, B.; Egan, T. J.; Brown, K. L. Parameters for the AMBER Force Field for the Molecular Mechanics Modeling of the Cobalt Corrinoids. *J. Mol. Struct.* **2001**, *561* (1–3), 71–91. https://doi.org/10.1016/S0022-2860(00)00920-0.

(51) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73–78. https://doi.org/10.1002/wcms.81.

(52) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(53) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(54) PyMOL Molecular Graphics System, Ver. 2.5; Schrödinger, LLC.

(55) Jmol: An Open-Source Java Viewer for Chemical Structures in 3D. Http://Www.Jmol.Org/. Accessed 2 Feb 2020.

(56) Wilson, E. B. J.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New

York, 1955.

(57) Brooks, A. J.; Fox, C. C.; Marsh, E. N. G.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Electronic Structure Studies of the Adenosylcobalamin Cofactor in Glutamate Mutase. *Biochemistry* **2005**, *44* (46), 15167–15181. https://doi.org/10.1021/bi051094y.

(58) Handali, J. D.; Sunden, K. F.; Thompson, B. J.; Neff-Mallon, N. A.; Kaufman, E. M.; Brunold, T. C.; Wright, J. C. Three Dimensional Triply Resonant Sum Frequency Spectroscopy Revealing Vibronic Coupling in Cobalamins: Toward a Probe of Reaction Coordinates. *J. Phys. Chem. A* **2018**, *122* (46), 9031–9042. https://doi.org/10.1021/acs.jpca.8b07678.

(59) Salerno, E. V.; Miller, N. A.; Konar, A.; Li, Y.; Kieninger, C.; Kraütler, B.; Sension, R. J. Ultrafast Excited State Dynamics and Fluorescence from Vitamin B_{12} and Organometallic [Co]-C≡C-R Cobalamins. *J. Phys. Chem. B* **2020**, *124* (30), 6651–6656. https://doi.org/10.1021/acs.jpcb.0c04886.

(60) Das, T. K.; Couture, M.; Guertin, M.; Rousseau, D. L. Distal Interactions in the Cyanide Complex of Ferric *Chlamydomonas* Hemoglobin. *J. Phys. Chem. B* **2000**, *104* (46), 10750–10756. https://doi.org/10.1021/jp000452y.

(61) Kim, Y.; Babcock, G. T.; Surerus, K. K.; Fee, J. A.; Dyer, R. B.; Woodruff, W. H.; Oertling, W. A. Cyanide Binding and Active Site Structure in Heme-Copper Oxidases: Normal Coordinate Analysis of Iron-Cyanide Vibrations of A_3^{2+} CN^- Complexes of Cytochromes Ba_3 and Aa_3. *Biospectroscopy* **1998**, *4* (1), 1–15. https://doi.org/10.1002/(SICI)1520-6343(1998)4:1<1::AID-BSPY1>3.0.CO;2-A.

(62) Hannibal, L.; Smith, C. A.; Smith, J. A.; Axhemi, A.; Miller, A.; Wang, S.; Brasch, N. E.; Jacobsen, D. W. High Resolution Crystal Structure of the Methylcobalamin Analogues Ethylcobalamin and Butylcobalamin by X-Ray Synchrotron Diffraction. *Inorg. Chem.* **2009**, *48* (14), 6615–6622. https://doi.org/10.1021/ic900590p.

(63) Ouyang, L.; Rulis, P.; Ching, W. Y.; Nardin, G.; Randaccio, L. Accurate Redetermination of the X-Ray Structure and Electronic Bonding in Adenosylcobalamin. *Inorg. Chem.* **2004**, *43* (4), 1235–1241. https://doi.org/10.1021/ic0348446.

(64) Han, S.; Madden, J. F.; Thompson, R. G.; Strauss, S. H.; Siegel, L. M.; Spiro, T. G. Resonance Raman Studies of *Escherichia Coli* Sulfite Reductase Hemoprotein. 1. Siroheme Vibrational Modes. *Biochemistry* **1989**, *28* (13), 5461–5471. https://doi.org/10.1021/bi00439a022.

(65) Yoshikawa, S.; O'Keeffe, D. H.; Caughey, W. S. Investigations of Cyanide as an Infrared Probe of Hemeprotein Ligand Binding Sites. *J. Biol. Chem.* **1985**, *260* (6), 3518–3528. https://doi.org/10.1016/s0021-9258(19)83653-0.

(66) Pallares, I. G.; Brunold, T. C. Spectral and Electronic Properties of Nitrosylcobalamin. *Inorg. Chem.* **2014**, *53* (14), 7676–7691.

(67) Lodowski, P.; Jaworska, M.; Kornobis, K.; Andruniów, T.; Kozlowski, P. M. Electronic and Structural Properties of Low-Lying Excited States of Vitamin B_{12}. *J. Phys. Chem. B* **2011**, *115* (45), 13304–13319. https://doi.org/10.1021/jp200911y.

Chapter 3

Spectroscopic and Computational Investigation of the AdoCbl-Dependent Enzyme Ethanolamine Ammonia-Lyase: Destabilization of Protein-Bound AdoCbl Contributes to Accelerated Co–C Bond Homolysis

Chapter 3: Spectroscopic and Computational Investigation of the AdoCbl-Dependent Enzyme Ethanolamine Ammonia-Lyase: Destabilization of Protein-Bound AdoCbl Contributes to Accelerated Co–C Bond Homolysis

3.1. Introduction

Adenosylcobalamin (AdoCbl), also called coenzyme B_{12}, is an organometallic cofactor vital to enzymatic isomerization and elimination reactions across all areas of life. AdoCbl is a member of the B_{12} (cobalamin) family, characterized by a cobalt ion ligated equatorially by four nitrogens of a tetrapyrrole macrocycle termed the corrin ring. In the six-coordinate Co^{III} form, a pendant 5,6-dimethylbenzimidazole (DMB) base tethered to the coring ring by an intramolecular loop typically binds in the lower axial position, while the upper axial ligand varies across the B_{12} family. In the biologically active forms, methylcobalamin (MeCbl) and AdoCbl, this position is occupied by a methyl group and an adenosyl moiety, respectively, and these cofactors represent rare biological examples of an organometallic bond.

The Co–C bond lies at the very heart of AdoCbl's function. Homolytic cleavage of this bond generates Co^{II}cobalamin (Co^{II}Cbl) and an organic radical centered on the 5'-carbon of the adenosyl moiety (Ado$^{\bullet}$), the latter of which initiates a radical-based substrate isomerization reaction by abstracting a hydrogen atom from the substrate (see Figure 3.1).[1] In Class I isomerases, this reaction involves the 1,2-exchange of an alcohol, amine, or carbon-based functional group, while in Class II eliminases, rearrangement is followed by spontaneous elimination of a leaving group. Notably, in the presence of substrate, the rate of Co–C bond homolysis of enzyme-bound AdoCbl is increased by a startling 12 orders of magnitude compared to that of the free cofactor, corresponding to a ~17 kcal/mol decrease in activation energy.[2] Exploring the mechanism through

which AdoCbl-dependent enzymes achieve this trillion-fold rate acceleration while avoiding deleterious side reactions by the resulting radical species has long been a thriving area of research.

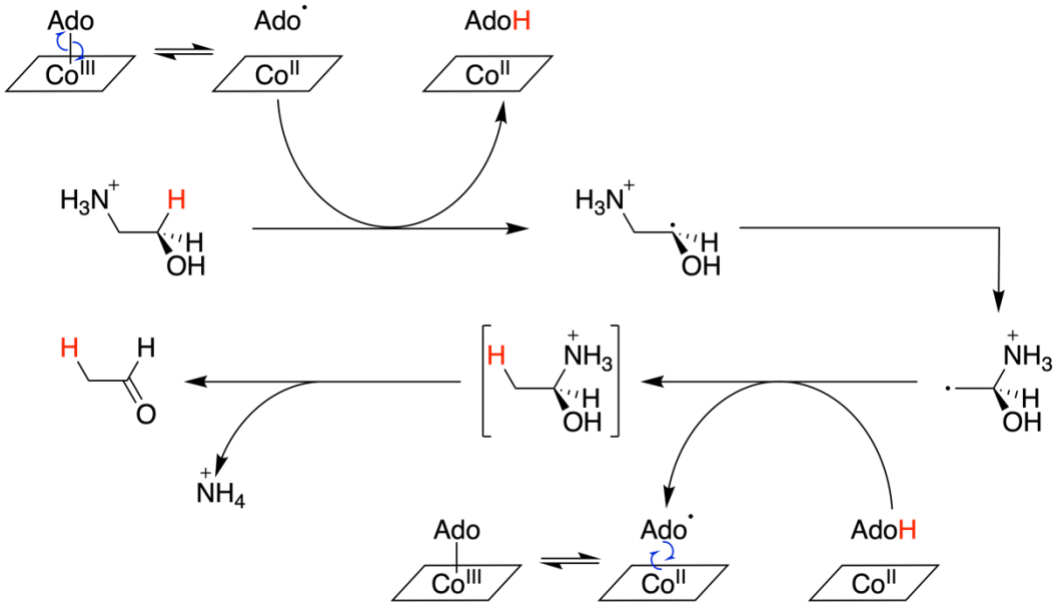


Figure 3.1. Reaction mechanism for EAL-catalyzed conversion of ethanolamine to acetaldehyde and an ammonium ion.

One such AdoCbl-dependent enzyme is ethanolamine ammonia-lyase (EAL), a Class II AdoCbl-dependent eliminase that catalyzes the conversion of ethanolamine to acetaldehyde and an ammonium ion.[3] As a result, EAL enables certain bacteria to use ethanolamine as the sole source of carbon and nitrogen.[4,5] One particularly notable feature of EAL is that its native substrate binds farther away from the cofactor than is observed for Class I AdoCbl-dependent isomerases; electron paramagnetic resonance (EPR) studies indicate a ~9 Å distance between Co^{II} and the substrate-like radical species in EAL,[6] compared to 6-7 Å in Class I isomerases.[7,8] EPR spectroscopy and related methods have afforded an impressively detailed understanding of the adenosyl radical (Ado$^{\bullet}$) trajectory following Co-C bond homolysis, showing that the substrate remains largely stationary relative to the active site while the Ado$^{\bullet}$ moiety swings upwards (by ~6 Å) to bring the radical carbon center in close proximity to the substrate (~2 Å away from the H atom to be abstracted).[9–13] This is in contrast to Class I isomerases like glutamate mutase (GM) and

methylmalonyl CoA mutase (MMCM), in which a change in ribose conformation, sometimes referred to as a "pseudorotation", is enough to move the C5' sufficiently close to the substate.[14–16] Migration of the Ado• radical across a larger distance in Class II eliminases suggests a need for specific enzyme-cofactor interactions that can guide the movement of this moiety.

Another curious facet of EAL is that it binds AdoCbl in the base-on conformation, with the intramolecular DMB remaining bound as the lower axial ligand, as shown through EPR[17–19] and X-ray crystallographic[20,21] studies. In contrast, Class I AdoCbl-dependent isomerases bind the cofactor in the so-called base-off/His-on conformation, whereby the DMB moiety is displaced by a protein-derived histidine (His) that participates in a network of conserved H-bonding residues. This conformation offers Class I isomerases an increased level of control over the cofactor's electronic structure by tuning the donor strength of the axially coordinated His residue through H-bonding. Based on the negligible spectral changes observed upon binding of AdoCbl to Class I isomerases by using electronic absorption (Abs), magnetic circular dichroism (MCD), and resonance Raman (rR) spectroscopies, the ground state of AdoCbl is largely unperturbed upon cofactor binding to the enzyme active site and the addition of substrate for this class of enzymes.[22–25] The electronic structure of Co^{II}Cbl, however, was found to be markedly perturbed in these enzymes. A substantial blue shift of the Co 3d $\rightarrow$ corrin π^* charge transfer transitions observed by MCD spectroscopy was attributed to a uniform stabilization of the filled Co 3d orbitals, and thus an overall stabilization of the enzyme-bound cofactor in the Co^{II} state.[25,26] Coupling of proton uptake by the network of H-bonding residues upon Co–C bond cleavage was proposed to play a key role in this process, given that the resulting decrease in charge donation from His to the nascent Co^{II}Cbl species would lead to the observed stabilization of the Co 3d orbitals. The axial His ligand is therefore crucial to the mechanism of Co–C bond activation employed by Class I enzymes.

In the absence of a His ligand to tune the AdoCbl cofactor's properties, the mechanism by which EAL and other Class II eliminases activate the Co–C bond for homolysis remains unclear. Circular dichroism (CD) spectroscopy revealed that EAL does not undergo any major conformational changes upon binding of AdoCbl and substrate.[27] Similarly, by using time-resolved Abs spectroscopy in combination with pulsed-laser photolysis of the EAL/AdoCbl/substrate ("ternary") complex, no significant perturbation of the AdoCbl ground state or stabilization of the $Ado^{\bullet}/Co^{II}Cbl$ radical pair was observed.[28] These data suggested that more subtle enzyme-cofactor-substrate interactions promote Co–C bond cleavage in Class II eliminases. Identifying those interactions is complicated by the fact that no crystal structures of EAL exist that contain the intact, native cofactor, likely due to the susceptibility of the Co–C bond to cleavage by free electrons generated during X-ray data collection.

Kinetic studies in conjunction with site-directed mutagenesis have shed some light on active site EAL residues and interactions crucial for catalytic activity. Substitutions of an active site glutamate (Gluα287) shown to interact with the substrate and Ado moiety greatly diminished the turnover number and substrate affinities.[29] In another study,[30] a series of EAL variants with a single substitution of an active-site residue believed to hydrogen bond with the substrate (Argα160, Glnα162, Asnα193, Gluα287, or Aspα362) allowed for the identification of the role of each of these five residues. These roles include suppressing undesirable side reactions, stabilizing the transition state, guiding the migrating NH_3 group, and positioning the substrate and cofactor. Modifications to the adenosyl ligand, such as replacement of the ribose group with alkyl chains, greatly reduced or entirely eliminated enzyme activity.[31] Meanwhile, modifications to the lower axial DMB ligand also resulted in reduced activity, but some activity was retained with

analogues that did not significantly reduce the ligand's steric bulk, suggesting that the DMB base may play a role in cofactor positioning.[32]

In the present study, we conducted a spectroscopic and computational investigation of the mechanism by which EAL activates AdoCbl's Co–C bond in the presence of substrate. Abs, MCD, and rRaman spectroscopy were used to investigate changes in the cofactor's geometric and electronic structure upon enzyme binding and the addition of substrate. Models of AdoCbl-bound EAL in the absence and presence of substrate were generated using molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) methods. Experimental validation for these structural models was achieved by simulating their Abs spectra using time-dependent density functional theory (TD-DFT). A relaxed potential energy scan was then used to simulate Co–C bond lengthening and cleavage, highlighting the interactions between cofactor, substrate, and active-site site residues that trigger Co–C bond homolysis and position the Ado moiety for H atom extraction. The scan also allowed for the calculation of Co–C bond dissociation energies (BDEs) in free AdoCbl and enzyme-bound AdoCbl in the absence and presence of substrate, with results that align well with experimental data.

3.2. Methods

3.2.1. Materials

Adenosylcobalamin (AdoCbl), (S)-2-aminopropanol (2AP), and ethanolamine (EA) were purchased from Sigma and used as obtained.

3.2.2. Protein Preparation and Purification

Gene expression and protein purification of *Salmonella enterica* EAL was performed by Dr. Flavia Costa.

3.2.3. Sample preparation

Samples of EAL for spectroscopic studies were prepared in 30 mM phosphate buffer (pH 7.5) with 1 mM tris(2-carboxyethyl)phosphine (TCEP), with the exception of rR samples, which were instead prepared using 50 mM Tris buffer with 1 mM TCEP (pH 7.5). MCD samples also contained 60% (v/v) polyethylene glycol 200 (PEG 200) to ensure glass formation upon freezing. EAL was complexed with AdoCbl in a ~1:0.9 protein:cofactor ratio. Samples containing substrate (analogues) were prepared by mixing the holoenzyme species with an approximately 100-fold molar excess of substrate in buffer.

Samples used to produce Co^{II}Cbl via photolysis were prepared in an oxygen-free glovebox, and any samples containing AdoCbl were handled under red light to prevent premature photolysis of the Co–C bond. For collecting MCD spectra of photolyzed species, samples were injected into MCD sample cells, photolyzed using the 514.5 nm line of an Ar^+ laser at a power of 75 mW for one minute, and then immediately frozen in N_2(l) to prevent oxidation. Frozen rR sample pellets were prepared by injecting drops of sample solution (~170 μM) into N_2(l),

3.2.4. Spectroscopy

Room-temperature Abs spectra of AdoCbl-containing samples were collected on a Cary 4 Bio spectrophotometer.

rR spectra were obtained with excitation at 514.5 nm by a Coherent 1-305 Ar^+ ion laser with ~20 mW of laser power at the sample. The scattered light was collected using a ~135° backscattering arrangement, dispersed by an Acton research triple monochromator (equipped with 1200 and 2400 grooves/mm gratings), and analyzed by a Princeton Instrument Spec X:100 BR deep depletion, back-thinned CCD camera. Spectra were collected at 77 K on pellets placed directly into a quartz finger dewar filled with N_2(l). The well-defined shoulder of the ice peak at 314 cm^{-1} was used as an internal standard to calibrate Raman shifts.

MCD spectra were collected on a Jasco J-715 spectropolarimeter in conjunction with an Oxford Instruments 7T SpectromagPT cryofree magnetocryostat. All MCD spectra reported here were obtained by taking the difference between spectra obtained with the magnetic field parallel and antiparallel to the direction of light propagation to eliminate contributions from the natural CD (and glass strain for frozen samples).

3.2.5. Computational modeling of Co–C bond cleavage

Computational models of AdoCbl, AdoCbl/EAL (holoEAL), and AdoCbl/EAL/EA (the ternary complex) were constructed using published X-ray crystallographic data of *E. coli* EAL with bound AdoCbl (PDB: 5YSN).[21] Molecular dynamics (MD) simulations were performed using the GROMACS 2019.6 software with the AMBER ff14SB force field,[33] supplemented by AMBER-compatible parameters developed specifically for cobalamins.[34] A detailed description of how each structure was prepared for MD simulations is provided in the Supporting Information (Appendix 1). In brief, each model was solvated in a 103.5 nm^3 TIP3P[35] water box and subjected to energy minimization, successive equilibrations under the NVT (isothermal-isochoric) and NPT (isothermal-isobaric) ensembles, and a 5 ns MD simulation at 300 K using the GROMACS 2019.6 software.[36] The GROMACS RMSD-based clustering tool was used to select a single structure from each 5 ns MD run for subsequent geometry optimization.

Geometry optimizations were performed using the quantum mechanics/molecular mechanics (QM/MM) approach as implemented with the ONIOM method in Gaussian 16.[37] The QM region was defined as the corrin ring, the entirety of the side chain that interacts with the Ado ligand, the first carbon of all other side chains, the DMB ligand, and the entire Ado moiety, as well as key active site residues (R160α, Q162α, N193α, S247α, E287α, D362α), relevant water molecules, and the substrate when applicable. These atoms were treated with the B3LYP hybrid

functional[38,39] and Grimme's D3 empirical dispersion corrections,[40] using the TZVP basis set[41] for Co and its six coordinating atoms and the 6-31G* basis set[42] for all other QM atoms. All atoms not in the QM region, including additional water molecules and the remainder of the EAL α/β dimer, were modeled with the AMBER force field[43] as implemented in Gaussian 16 and the AMBER-compatible cobalamin parameters described above.[34] Covalent bonds spanning the QM/MM boundary were capped with H atoms with C–H bond lengths scaled by 0.709. Geometry optimizations of all models were performed spin restricted (S = 0) with QM/MM loose convergence criteria. Subsequently, a relaxed potential energy surface scan was carried out by elongating the Co–C bond in 0.1 Å increments using Gaussian's ModRedundant functionality, with a full QM/MM geometry optimization performed after each 0.1 Å elongation. To allow for the homolytic cleavage of the Co–C bond, these computations were performed for the open-shell singlet state (M_S = 0). Potential energy curves (PECs) associated with the homolytic cleavage of the Co–C bond were constructed by plotting the QM energy for each stationary point during incremental Co–C bond elongation.

3.2.6. Single Point (SP) QM/MM Energy Calculations

A series of SP QM/MM calculations were performed for every stationary point along the three PECs, using the same QM region across all models, to permit a direct comparison of the computed energies. This QM region, hereafter referred to as the "cofactor QM region," was identical to that described above, but excluding any enzyme residues, active-site water molecules, and substrate. For these SP computations, the larger def2-TZVP basis set[44] was used for all QM atoms along with the same functional and dispersion corrections as described above.

3.2.7. TD-DFT Computed Abs Spectra

Abs spectra for the optimized ground state structures of AdoCbl, holoEAL, and the ternary complex were calculated using time-dependent density functional theory (TD-DFT) within the Tamm-Dancoff approximation[45] as implemented in the ORCA 4.2.1 software package.[46] In each case, atoms that had been included in the QM region for the geometry optimization (including capping hydrogens) were used explicitly in the TD-DFT calculation, while atoms from the MM region were included as point charges. Calculations were performed with the B3LYP/G functional, using the def2-TZVP basis set for Co and its six coordinating atoms and the def2-SV(P) basis set for all other atoms. The resolution of identity (RI) approximation for electron correlation[47] was applied using the def2-SVP/C auxiliary basis set.[48] The resulting TD-DFT-calculated transition energies and oscillator strengths were used to simulate Abs spectra by assuming that each transition gives rise to a Gaussian band with full width at half maximum of 1500 cm^{-1}.

3.3. Results and Analysis

3.3.1. Spectroscopic Data

3.3.1.1. Free and EAL-bound AdoCbl

Figure 3.2 shows Abs, CD, and MCD spectra obtained for free AdoCbl, EAL-bound AdoCbl (holoEAL), and EAL-bound AdoCbl in the presence of the slow substrate (S)-2-aminopropanol (2AP). Consistent with previous reports,[27–29,32] the Abs and CD spectra of AdoCbl are insignificantly perturbed by binding of the cofactor to EAL, even upon the addition of substrate. The only change of note is an increase in peak intensity in the γ region of the Abs spectra of the EAL-bound cofactor, as has been observed previously for other AdoCbl-dependent enzymes[24,25] and was attributed to changes in the cofactor's dielectric environment upon enzyme binding. To complement these data, we also collected MCD spectra free AdoCbl, holoEAL, and the AdoCbl/EAL/2AP ternary complex. Because at low temperatures, features arising from these

diamagnetic species were obscured by temperature-dependent features from a small fraction of paramagnetic Co^{II}Cbl (*vide infra*), we only report the 250 K MCD spectra. Consistent with our CD and Abs spectra, the MCD spectra of the holoEAL and the ternary complex only show minor differences from that of AdoCbl. The splitting of the broad, negatively-signed MCD feature at ~20,700 cm^{-1} can be attributed to a slight increase in spectral resolution, suggesting that AdoCbl adopts a more constrained conformation in the EAL active site than in solution. The other MCD features are largely unchanged in energy and intensity.

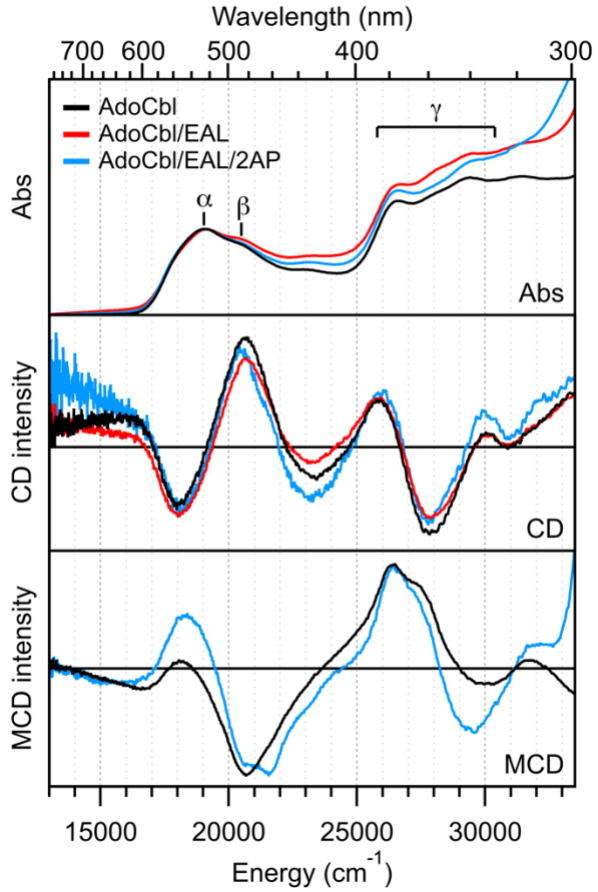


Figure 3.2. Abs (top) and CD (middle) spectra at 298 K and MCD spectra at 250 K and 7 T (bottom) of AdoCbl (black), EAL-bound AdoCbl (red), and EAL-bound AdoCbl in the presence of the slow substrate 2AP (light blue). To facilitate a comparison of the different spectra, intensities were normalized to the most intense transition in each spectrum. Band designations are provided above the Abs spectra.

Differences between free and EAL-bound AdoCbl were investigated further using resonance Raman (rR) spectroscopy (Figure S1). Isotopic labeling and computational methods have previously been used to identify several rR features between 390 and 570 cm^{-1} that arise from heavily mixed stretching, bending, or torsional motions involving the Ado ligand.[22,23,49] Although the resolution of our spectra was insufficient to resolve some of these features, we clearly identified a peak at 428 cm^{-1} that was previously assigned as a Co–C–C bending mode ($\nu_{Co–C}$).[49] Upon AdoCbl-binding to EAL, this feature displays a downshifts of 4 cm^{-1}. Similarly minimal shifts were observed in rR studies of MMCM, which led the authors to conclude that AdoCbl ground state destabilization is not a major contributor to Co–C bond activation.[22,23] Collectively, our Abs, CD, MCD, and rR spectroscopic data indicate that the geometric and electronic structures of AdoCbl are not dramatically perturbed in the EAL active site in the absence or presence of a substrate.

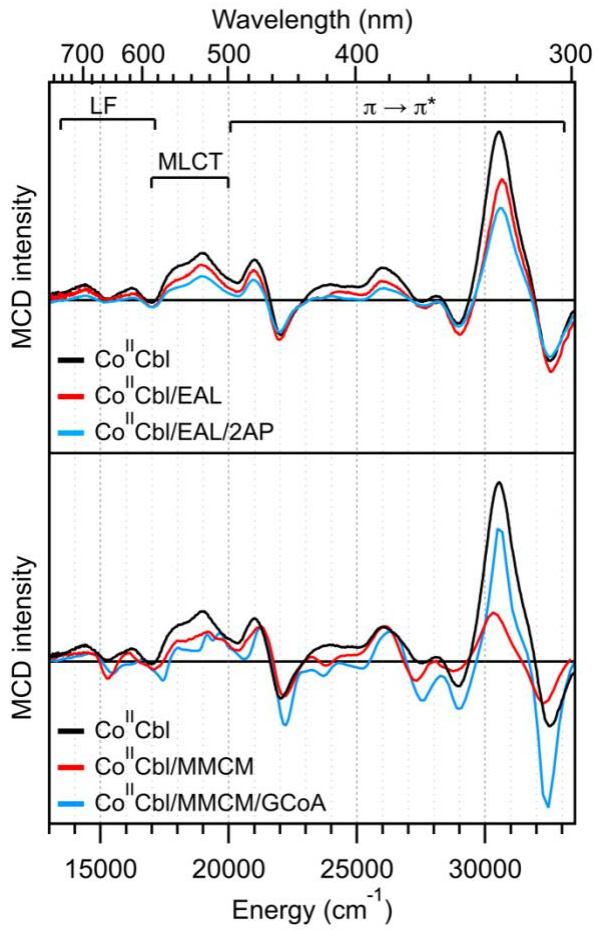


Figure 3.3. Top: MCD spectra at 4.5 K and 7 T of free Co^{II}Cbl (black), EAL-bound Co^{II}Cbl (red), and EAL-bound Co^{II}Cbl in the presence of the slow substrate 2AP (light blue). Bottom: Analogous spectra collected for free Co^{II}Cbl (black), MMCM-bound Co^{II}Cbl (red), and MMCM-bound Co^{II}Cbl in the presence of slow substrate glutaryl-CoA (GCoA) (light blue). Band assignments are provided in the top panel.

3.3.1.2. Co^{II}Cbl Post-homolysis Product

We next examined perturbations induced by EAL on the Co^{II}Cbl post-homolysis product in the absence and presence of substrate. The low-temperature MCD spectra of Co^{II}Cbl, EAL-bound Co^{II}Cbl, and EAL-bound Co^{II}Cbl in the presence of 2AP, all generated through photolysis of the corresponding AdoCbl-containing species, are virtually identical to one another (Figure 3.3) and MCD spectra of Co^{II}Cbl reported previously.[50] This represents a clear departure from the marked differences observed between the analogous MCD spectra of the Class I AdoCbl-dependent isomerases MMCM (reproduced in Figure 3.3 for comparison) and GM. In brief, the

MCD spectrum of Co^{II}Cbl can be divided into three regions that are dominated by contributions from distinct types of transitions: d → d ligand field (LF) transitions below ~17,000 cm^{-1}, metal-to-ligand charge transfer (MCLT) transitions from 17,000 cm^{-1} to 20,000 cm^{-1}, and corrin π → π* transitions above 20,000 cm^{-1}.[50] For MCMM and GM, enzyme binding and the addition of a substrate did not significantly perturb the LF and π → π* regions of Co^{II}Cbl but did induce a uniform blue-shift of features in the MLCT region. When analyzed within the framework of TD-DFT calculations, these shifts were attributed to a sizeable stabilization of the filled Co 3d orbitals by the protein-derived axial histidine ligand in the Co^{II} state due to proton uptake by this His ligand, making that residue an integral component of the mechanism by which MMCM promotes Co–C bond homolysis. The lack of any significant differences among the MCD spectra of free and EAL-bound Co^{II}Cbl in the absence and presence of substrate implies that EAL employs a fundamentally different mechanism of Co–C bond activation, which is not surprising given the DMB-on binding mode of AdoCbl by this enzyme and the fact that the DMB ligand is unable to participate in any H-bonding interactions.

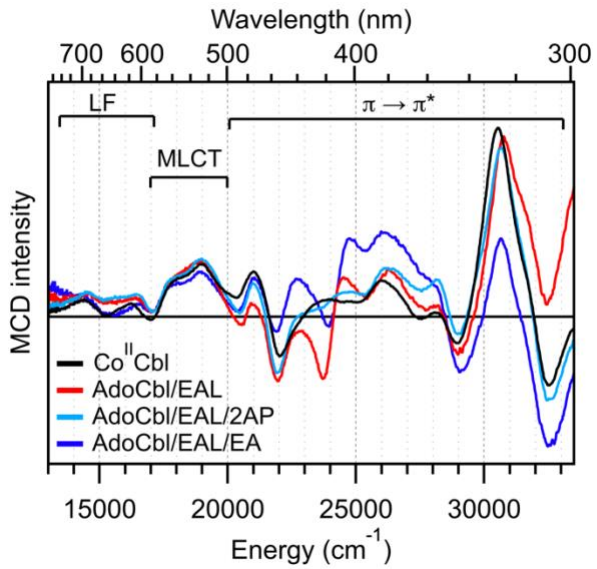


Figure 3.4. MCD spectra at 4.5 K and 7 T of free Co^{II}Cbl (black), EAL-bound AdoCbl (red), EAL-bound AdoCbl in the presence of slow substrate 2AP (light blue), and EAL-bound AdoCbl in the presence of native substrate ethanolamine (dark blue). Band assignments are provided along the top.

Interestingly, MCD spectra of the "intact" AdoCbl/EAL and AdoCbl/EAL/substrate samples used to collect the spectra in Figure 3.2 were found to contain temperature-dependent contributions from a small fraction of Co^{II}Cbl (see Figure 3.4). At 4.5 K, the temperature-dependent C-terms from paramagnetic Co^{II}Cbl are over 10-fold more intense than the temperature-independent B-terms from diamagnetic AdoCbl,[50,51] so even a small fraction of Co^{II}Cbl obscures signals from AdoCbl. We therefore concluded that these "intact" samples contain a mixture of species, predominantly in the AdoCbl state (as demonstrated by the Abs spectra in Figure 3.2) but with a small fraction of Co^{II}Cbl-containing active sites. Although unexpected for the holoenzyme sample, this is unsurprising for the two ternary complexes given that Co^{II}Cbl is known to accumulate during enzyme turnover, with 58% and 90% of active sites containing Co^{II}Cbl during the steady-state phase of EAL's reaction with EA and 2AP, respectively.[52] Importantly, most features are almost identical for the free and EAL-bound Co^{II}Cbl species, precluding significant perturbation of the Co^{II} coordination environment (Figure 3.3).

Interestingly, a new derivative-shaped, temperature-dependent feature is observed at 24,000 cm^{-1} in the MCD spectra of the AdoCbl/EAL and AdoCbl/EAL/EA samples. Because the other Co^{II}Cbl features remain unchanged, this feature most likely arises from a separate paramagnetic species, possibly a tyrosyl radical.[53] Previous EPR studies of EAL incubated with EA revealed the formation of a narrow EPR signal with very low intensity that could not be attributed to the substrate- or product-like radicals.[54,55] Notably, this feature matches the EPR signal displayed by tyrosyl radicals.[56] Therefore, we suggest that under the experimental conditions used in our study, a holoEAL and the native ternary complex contain a small fraction of a tyrosyl radical. In our computational models (detailed more extensively below), Tyrα404 protrudes into the substrate binding pocket, and in the ternary complex, the carbon atom harboring the unpaired electron in in the product-like radical from EAL's catalytic cycle is 3.23 Å from the tyrosine's abstractable hydrogen (see Figure S14). It is possible that in a small fraction of cases, the product-like radical abstracts a hydrogen from Tyr, producing a Tyr radical. Crucially, the slow substrate is less well-suited for this, as its relevant carbon atom is further from Tyrα404 than in the native substrate (3.90 Å), and its additional methyl group prevents it from moving closer. This may explain why a similar species is not observed for AdoCbl/EAL/2AP. In the holoenzyme, meanwhile, Tyrα404 may behave like a radical trap; in the rare event that the Co–C bond cleaves in the absence of substrate and does not undergo fast recombination, Tyrα404 can react with Ado$^{\bullet}$, preventing the dangerously reactive species from diffusing out of the active site. We believe that a Tyr radical is not observed in the photolyzed species because, under these aggressive non-turnover conditions, the Ado radical is immediately quenched, leaving Co^{II}Cbl and AdoH.

Altogether, the lack of substantial perturbations in these MCD spectra suggest that binding to the EAL active site and the addition of substrate to the holoenzyme do not significantly alter the

electronic structure of either the AdoCbl ground state or the Co^{II}Cbl post-homolysis product. These data provide compelling evidence that the Co–C bond activation mechanism employed by EAL is fundamentally different from that used by the Class I isomerases MMCM and GM. We therefore turn to computational techniques that can probe interactions of the cofactor, enzyme, and substrate that cannot be directly explored through Abs and MCD spectroscopy.

3.3.2. Computational modeling of Co–C bond cleavage

To explore EAL-induced AdoCbl structural changes and the role of substrate in triggering Co–C bond homolysis, we generated QM/MM-optimized models of solvated AdoCbl, the AdoCbl/EAL holoenzyme, and the AdoCbl/EAL/EA ternary complex. Building initial structural models of the holoenzyme and ternary complex was complicated by the absence of an X-ray crystal structure of EAL-bound AdoCbl containing an intact (i.e., ~2.0 Å) Co–C bond. Because this study focuses on changes to the AdoCbl geometry and Co–C bond strength imposed by the EAL protein matrix, a thorough computational approach was needed to (i) build the Co–C bond of EAL-bound AdoCbl with minimal user bias and (ii) explore the conformational space of the Ado moiety after Co–C bond formation. Therefore, we employed MD simulations in conjunction with an RMSD-based clustering algorithm to generate structures for subsequent QM/MM geometry optimization. The optimized models were then used for a relaxed potential energy curve scan, in which the Co–C bond was elongated in 0.1 Å increments and a full QM/MM geometry optimization was performed after each step. The resulting QM energies at each stationary point allowed us to compute the Co–C BDE for AdoCbl in each model. A more detailed methodology is provided in the SI (Appendix 1 and Figure S2).

3.3.2.1. QM/MM-optimized models of solvated AdoCbl, holoEAL, and the ternary complex

We first validated our combined MD + QM/MM approach using solvated AdoCbl, in which the AdoCbl cofactor was placed in a large (~100 nm^3) water box. The geometric and electronic structures of solvated AdoCbl have been thoroughly characterized and serve as a benchmark for successful computational modeling.[57,58] After MD simulation and QM/MM geometry optimization, the key metric parameters of the solvated AdoCbl model agree well with those observed in the crystal structure of AdoCbl,[58] with Co–C bond lengths of 2.04 and 2.02 Å, respectively (Table 1). Minor differences, including a decrease in Co–N_{DMB} bond length from 2.24 to 2.14 Å upon optimization, can be attributed, at least in part, to crystal packing forces.[59] To validate an accurate determination of the electronic structure of solvated AdoCbl, TD-DFT was used to compute the corresponding Abs spectrum. This spectrum reproduces the major experimental absorption features reasonably well (Figure 3.6). Furthermore, the computed Co–C homolysis energy for this model (*vide supra*) was 27.8 kcal/mol. Although this value is slightly lower than the experimentally-determined Co–C BDE of ~31 kcal/mol,[60] this is consistent with the known tendency of the B3LYP functional to underestimate the Co–C BDE in alkylcobalamins,[61,62] and so we can conclude that our computational model of AdoCbl successfully reproduces experimental data.

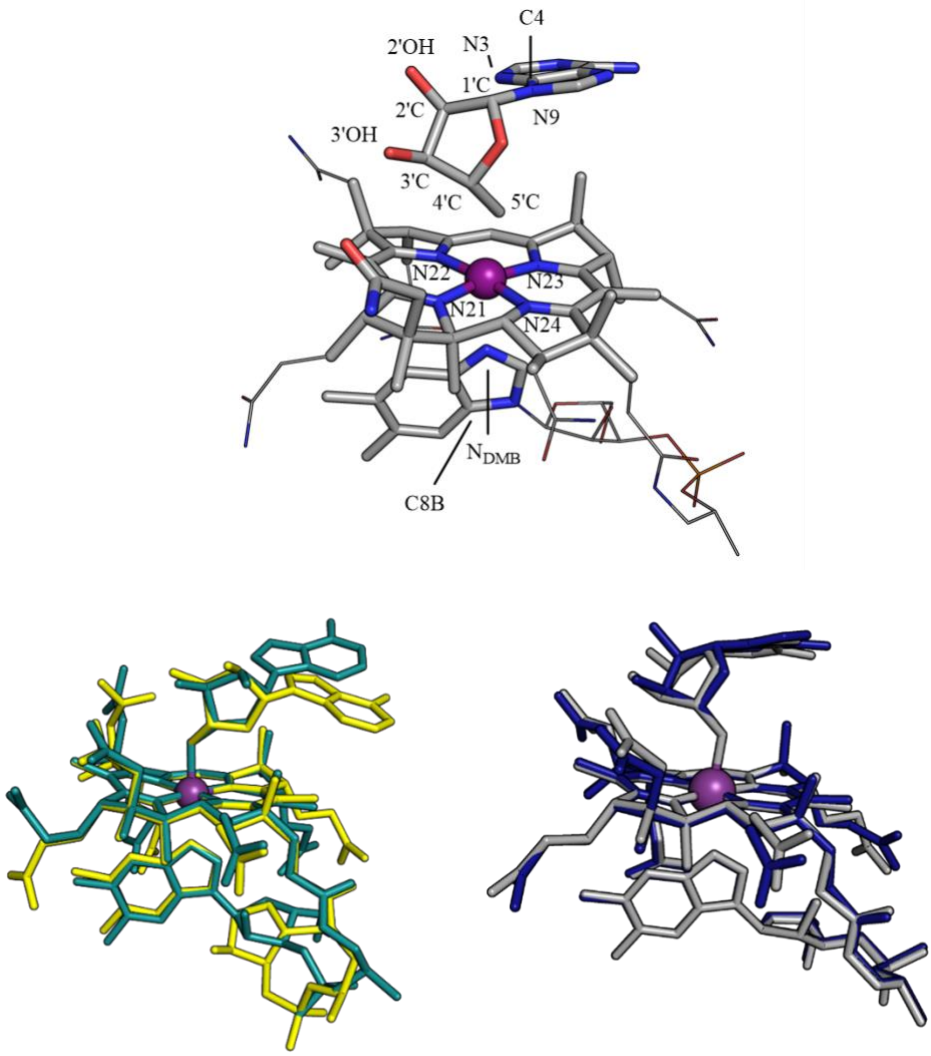


Figure 3.5. Top: AdoCbl numbering scheme with the cofactor QM region and MM region shown as thick and thin sticks, respectively. Bottom left: QM/MM optimized model of solvated AdoCbl (green) and AdoCbl crystal structure (yellow). Bottom right: AdoCbl portion of QM/MM optimized models of AdoCbl/EAL (blue) and AdoCbl/EAL/EA (gray).

Table 3.1. Key metric parameters derived from the crystal structure of AdoCbl, the three QM/MM-optimized AdoCbl-containing models in this work, and the crystal structure of the EAL-bound AdoCbl derivative AdePeCbl

	AdoCbl (X-ray)[a]	**AdoCbl (QM/MM)**	**AdoCbl/EAL (QM/MM)**	**AdoCbl/EAL/EA (QM/MM)**	**AdePeCbl/EAL (X-ray)[b,c]**
Co–C (Å)	2.02	2.04	2.06	2.10	2.00
Co–N_{DMB} (Å)	2.24	2.14	2.50	2.45	2.57
Co-N1 (Å)	1.87	1.89	1.90	1.90	1.82
Co-N2 (Å)	1.92	1.96	1.96	1.97	1.84
Co-N3 (Å)	1.91	1.93	1.93	1.94	1.83
Co-N4 (Å)	1.88	1.90	1.91	1.90	1.83
DMB Tilt[d] (°)	155.2	157.5	157.5	159.1	159.3
Co–5'C–4'C (°)	122.0	125.6	126.3	128.7	N/A
N22–Co–5'C (°)	84.5	83.9	91.3	94.8	90.8
N_{DMB}–Co–5'C (°)	171.7	170.8	163.2	166.3	176.4
C4–N9–1'C–2'C (°)	127.8	149.9	11.9	28.4	N/A

[a.] Values from AdoCbl crystal structure[58] [b.] Average of the two AdePeCbl structures from PDB: 3ABS.[20] [c.] Co-bound methylene unit labelled equivalently to the 5'C atom of the Ado ribose for comparison [d.] DMB Tilt: Co–N_{DMB}–C8B.

Warranted by the close agreement between the optimized and experimental AdoCbl structures, we applied the same MD + QM/MM methodology to the EAL-bound AdoCbl models. The holoEAL model, which consisted of one AdoCbl in an EAL α/β dimer in a ~100 nm^3 water box, was used to investigate changes in the AdoCbl structure imposed by EAL in the absence of substrate. Notably, both Co axial bonds are lengthened in holoEAL compared to solvated AdoCbl, with a slight increase in Co–C bond length from 2.04 Å to 2.06 Å and a more substantial increase

in Co–N_{DMB} bond length from 2.14 Å to 2.50 Å (Table 1). In the absence of an EAL-bound AdoCbl crystal structure with an intact Co–C bond, it is best to compare our model to the crystal structure of EAL containing the cofactor derivative adeninylpentylcobalamin (AdePeCbl) with an intact Co–C bond measuring 2.00 Å (PDB: 3ABS).[20] In this structure, the AdePeCbl Co–N_{DMB} bond length is 2.57 Å and thus even slightly larger than the AdoCbl Co–N_{DMB} bond length in our QM/MM optimized holoEAL model. Thus, both experimental and computational data indicate that a significant lengthening of the Co–N_{DMB} bond occurs in response to AdoCbl binding to EAL. Though the AdePeCbl-bound EAL crystal structure contains an intact Co–C bond, substrate EA is also present in this structure and any conclusions beyond Co–N_{DMB} lengthening are likely not warranted, as active site changes are likely imposed through upper axial interactions with substrate.

In the optimized model of the ternary complex, two active site water molecules that are present in the holoenzyme are displaced by the native substrate EA. The five substrate-binding residues identified in the crystal structure of AdePeCbl complexed with EA (PDB: 3ABS) all maintain H-bonding interactions with the substrate after geometry optimization of the ternary complex model; the RMSD of the heavy atom positions of the substrate and five substrate-binding residues between EAL/AdePeCbl/EA (PDB: 3ABS) and optimized ternary complex is merely 0.55 Å (Figure S9). The optimized AdoCbl structure in the ternary complex shows only minor atomic displacements compared to holoEAL (Figure 3.5). The RMSD of the heavy atom positions of the cofactor QM region between these two EAL-bound AdoCbl models is 0.26 Å and increases to only 0.46 Å when comparing all heavy atoms of AdoCbl. Both optimized EAL-bound AdoCbl models show a displacement of the Co^{III} ion towards the Ado moiety (Figure 3.5) and elongated Co–N_{DMB} and Co–C bonds. Notably, the Co–C bond in the ternary complex is lengthened by an additional

0.04 Å, to 2.10 Å. The computational models, therefore, predict a slight weakening of the Co–C bond upon EA binding to holoEAL.

3.3.2.2. Validation of computational models on the basis of spectroscopic results

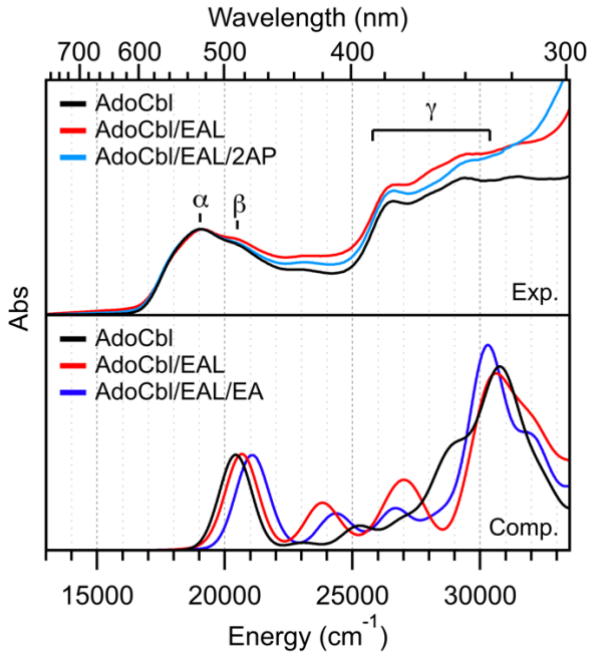


Figure 3.6. Experimental 298 K (top) and TD-DFT-computed (bottom) Abs spectra of AdoCbl (black), holoEAL (red), and the ternary complex (blue).

To experimentally validate our computational models beyond structural comparisons, we used TD-DFT to simulate Abs spectra for the three AdoCbl-containing models. The strengths of the axial bonding interactions in alkylcobalamins are known to affect the position of the low energy, α/β peaks in the Abs spectrum.[57,63] These peaks correspond to the electronic origin and vibrational sideband of a corrin π (HOMO) → π* (LUMO) transition. Because the HOMO also contains small contributions from the Co–C σ and Co–NDMB σ* bonding interactions, the energy of this transition is sensitive to axial ligand perturbations.[57,63] As the Co–C bond lengthens from solvated AdoCbl → holoEAL → ternary complex, the Co–C σ bonding interaction weakens and a red shift of the α/β peaks is expected. Conversely, as the Co–N_{DMB} σ* bond lengthens upon AdoCbl binding to EAL, a blue shift of the α/β peaks is expected. Thus, the simultaneous lengthening of

the Co–C and Co–N_{DMB} bonds should result in minimal shifts of the α/β peaks. Indeed, in the TD-DFT computed Abs spectra for AdoCbl, holoEAL, and the ternary complex, the feature corresponding to the α/β peaks only minimally blue shifts by 235 cm^{-1} and 645 cm^{-1} for holoEAL and the ternary complex, respectively. The absence of more pronounced spectral changes in the computed spectra is consistent with our experimental data, lending further credence to our whole-protein models of holoEAL and the ternary complex.

3.3.2.3. PEC scans of Co–C bond elongation

After validating our QM/MM-optimized structures on the basis of X-ray crystallographic and spectroscopic data, we performed potential energy curve (PEC) scans for Co–C bond lengthening to explore the impact of EAL-binding and the addition of substrate on Co–C bond activation. Figure 3.7A shows the QM energy for each stationary point during the 0.1 Å incremental Co–C bond elongation for all three AdoCbl-containing models investigated. Since the computational models contained a different number of atoms in the QM region, a direct comparison of the absolute QM energies between them is not possible. All QM energies in Figure 3.7A are therefore plotted relative to the lowest QM energy for each model, corresponding to the ground state structure prior to Co–C bond elongation. Importantly, for solvated AdoCbl, the energy increases by 27.8 kcal/mol upon elongation of the Co–C bond to a final distance of 4.24 Å, while for the ternary complex the energy increase is merely 12.7 kcal/mol at a final Co···C distance of 4.30 Å. A decrease in QM energy of 15.1 kcal/mol is therefore computed for a Co–C bond elongation beyond 4 Å when comparing the ternary complex to the solvated cofactor, in excellent agreement with known experimental energy decrease for Co–C bond homolysis in AdoCbl-dependent enzymes. While the early program termination during Co–C bond elongation for holoEAL hinders a direct QM energy comparison for Co–C distances over 3.16 Å, the PEC for

this species is nearly identical to that of the solvated cofactor. Thus, our computational data correctly reproduce the experimental finding that Co–C bond activation by EAL requires the presence of substrate.

To permit a comparison of the absolute QM energies for free and EAL-bound AdoCbl, SP calculations were performed for all steps along each PEC using the same QM region (referred to hereafter as the "cofactor QM region"). Using the lowest QM energy of free AdoCbl as the reference, the cofactor QM energies of the QM/MM-optimized EAL-bound AdoCbl structures are higher (Figure 3.7B), indicating that the EAL protein matrix destabilizes AdoCbl. While the QM energies of both EAL-bound AdoCbl models are higher than that of solvated AdoCbl, the destabilization of AdoCbl in the ternary complex is seemingly less pronounced than in holoEAL (4.5 and 11.7 kcal/mol, respectively). However, inspection of the H-bonding interactions involving the Ado moiety revealed an intramolecular H-bond between the ribose 2'OH hydroxyl group and the adenine N3 atom in the ternary complex (Figure 3.9) that is absent in free AdoCbl and holoEAL. Instead, in holoEAL the same ribose 2'OH group donates an H bond to E287α. While this cofactor-protein H-bonding interaction is accounted for when constructing PECs of the larger QM regions (Figure 3.7A and S16), it does not contribute to the energies of the cofactor QM region (Figure 3.7B). To compensate for the stabilization of AdoCbl in the ternary complex due to the intramolecular H-bond, Figure 3.7B additionally shows the ternary complex PEC offset by +7.2 kcal/mol (which is a reasonable value for this type of H-bond) to align the initial cofactor QM energies for holoEAL and the ternary complex.

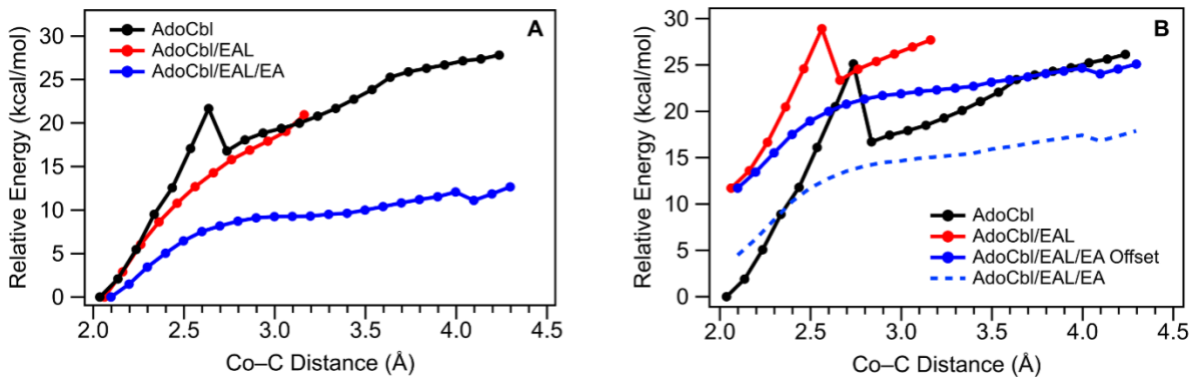


Figure 3.7. Computed PECs for Co–C bond cleavage of free and EAL-bound AdoCbl. Filled circles represent stationary points of QM/MM optimized geometries. Discontinuities are due to a change in spin state from a closed-shell singlet at short Co–C bond distances to an open-shell singlet at larger Co–C bond distances. A: Relative QM energies during Co–C bond elongation obtained using the TZVP/6-31G* split basis set. B: QM energies of the cofactor QM region for all stationary points in Panel A using the def2-TZVP basis set on all atoms.

While our computations predict a significant destabilization of EAL-bound AdoCbl (Figure 3.7B), no significant stabilization of the post-homolysis state is observed, even in the presence of substrate. After adjusting the cofactor QM energy of the ternary complex by +7.2 kcal/mol, there is a mere 1.0 kcal/mol energy difference between the solvated cofactor and ternary complex at a Co···C distance over 4 Å. The increase in energy for free AdoCbl upon elongation of the Co–C bond to a final distance of 4.24 Å is slightly lowered by 1.7 kcal/mol to 26.1 kcal/mol when the larger basis set is used (Figure 3.7), still in good agreement with experimental data. For the ternary complex, the cofactor QM energy increases by 13.4 kcal/mol upon Co–C bond elongation to 4.30 Å; thus, the use of a larger basis set and removal of substrate and active site residues from the QM region have an insignificant effect on the computed PEC (a mere 0.7 kcal/mol difference). The cofactor QM energy difference between the solvated AdoCbl and ternary complex is therefore 12.7 kcal/mol, again in excellent agreement with the experimental data. Increasing the basis set and removing substrate and active site residues from the QM region also has a minor effect on the PEC for holoEAL, with the cofactor QM energies for this species and free AdoCbl showing nearly the same dependence on the Co–C bond distance. Collectively, our results indicate that the ternary

complex has a reduced energy barrier to Co–C bond elongation compared to both holoEAL and solvated AdoCbl regardless of basis set size and the presence or absence of active site residues in the QM region, and that this decreased barrier can be attributed to AdoCbl ground state destabilization.

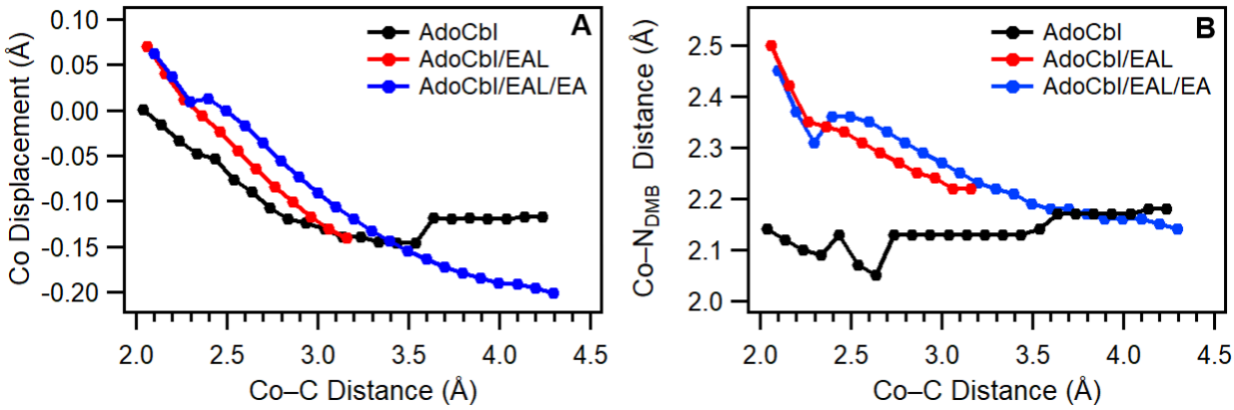


Figure 3.8. A: Co displacement from the corrin macrocycle as a function of Co–C bond distance. Positive values correspond to Co displacements towards the Ado moiety. B: Co–N_{DMB} distance as a function of Co–C bond distance.

To investigate the structural origin of the differences between the PECs in Figure 3.7, we first examined how Co metric parameters change as a function of Co–C bond elongation. The most drastic impact that AdoCbl-binding to EAL has on the Co coordination environment is an elongation of the Co–N_{DMB} bond by over 0.30 Å (Table 1). The geometry optimized models for both holoEAL and the ternary complex show that this lengthening is accomplished by favorable protein-cofactor interactions and a displacement of the Co ion from the plane of the corrin ring towards the Ado moiety (Figure 3.9A). This Co displacement is not observed for free AdoCbl and is therefore a consequence of EAL-induced structural changes. The observed Co displacements of 0.07 and 0.06 Å in holoEAL and the ternary complex, respectively, must be energetically more favorable than an additional elongation of the Co–C bond. These results suggest that there are stabilizing interactions between AdoCbl and the EAL active site that compensate for the energetic

penalty associated with the displacement of the Co ion from the corrin ring by up to 0.07 Å, before additional strain lengthens the Co–C bond from 2.06 Å in holoEAL to 2.10 Å in the ternary complex.

The Co–N_{DMB} lengthening of EAL-bound AdoCbl would only contribute to an increase in the Co–C bond homolysis rate if a less strained Co–N_{DMB} bond is present in the post-homolysis Co^{II}Cbl state. For free AdoCbl, a Co–C bond elongation to 4.24 Å leads to a minor increase in Co–N_{DMB} distance from 2.14 to 2.18 Å and a Co displacement of 0.12 Å towards the DMB moiety. In the ternary complex, the Co–N_{DMB} distance decreases from 2.45 to 2.14 Å when the Co–C bond is lengthened to 4.24 Å, which closely matches the 2.18 Å bond length in free AdoCbl. This change in Co–N_{DMB} distance is accomplished mainly by the displacement of the Co ion toward the DMB moiety, as the net 0.26 Å movement of the Co atom relative to the plane of the corrin ring accounts for most of the total 0.31 Å decrease in Co–N_{DMB} bond distance.

While the Co–N_{DMB} bond elongation of AdoCbl in the ternary complex and formation of a stronger Co–N_{DMB} bond in the post-homolysis Co^{II}Cbl product would be expected to contribute to the overall rate enhancement of Co–C bond homolysis, the same trend in computed Co displacement and Co–N_{DMB} bond length is observed for holoEAL, despite the early program termination during Co–C bond elongation. We therefore conclude that the movement of the Co ion relative to the corrin ring, which allows for the formation of stronger Co–N_{DMB} bond following Co–C bond cleavage, contributes minimally to the lowered barrier to Co–C bond homolysis in the ternary complex. Instead, changes in the interactions between the Ado moiety of AdoCbl and active site residues upon substrate binding must be responsible for Co–C bond activation in EAL.

3.3.2.4. Active site changes after substrate binding

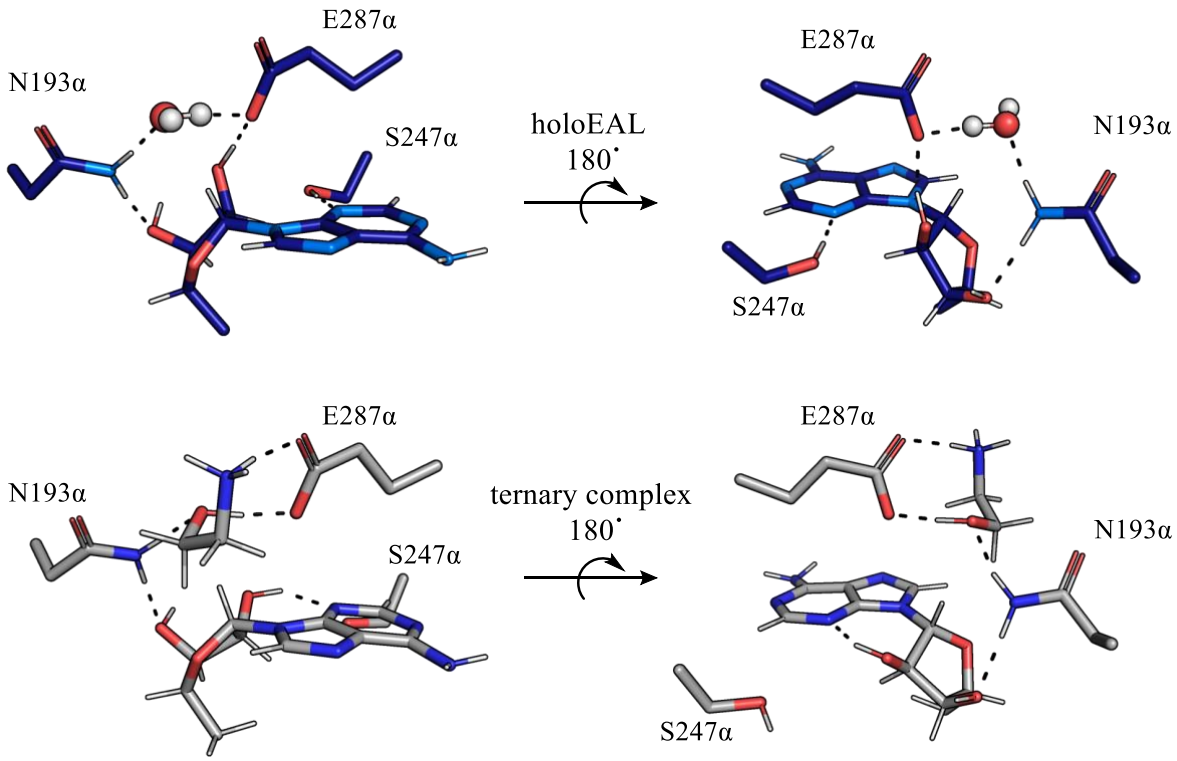


Figure 3.9. Protein–adenosyl and select protein–water/substrate H-bond interactions in QM/MM-optimized AdoCbl/EAL (top) and AdoCbl/EAL/EA (bottom) models, shown from two different angles. Aliphatic hydrogen atoms on side chains are omitted for clarity. Two active site water molecules are displaced upon substrate binding; only one of these is shown for AdoCbl/EAL for clarity.

In the optimized holoEAL model, three H-bonding interactions exist between active site residues and the Ado moiety (Figure 3.9). N193α donates an H-bond to the ribose 3'OH, E287α accepts an H-bond from the ribose 2'OH, and S247α donates an H-bond to the adenine N3. These protein–Ado interactions are maintained throughout the Co–C elongation process. Movement of the Ado• radical toward substrate following Co–C bond homolysis is required, as the substrate radical is positioned ~9 Å from the Co^{II} atom after H atom abstraction by Ado•.[64] Thus, the constrained Ado orientation in holoEAL, enforced through these H-bonding interactions, likely aids in suppressing undesirable Co–C bond homolysis in the absence of substrate. Additionally, they may contribute to the +11.7 kcal/mol ground-state destabilization computed for the cofactor QM region of holoEAL, as discussed below.

In the ternary complex, the number of protein–Ado H-bonding interactions is reduced from three to one. E287α participates in two short H-bonds with EA, measuring 1.7 Å to the amino group and 1.5 Å to the hydroxyl group, consistent with X-ray crystallographic data. The importance of E287α for catalysis has been established with kinetic and Abs data, as substitution of this residue leads to a dramatic decrease in enzymatic efficiency.[65] This residue has been implicated as a dynamic "substrate trigger" for Co–C bond homolysis, with its movement in response to substrate binding and H-bonding to EA promoting Co–C bond cleavage.[66,67] Our computational results support a slightly different mechanism of action for the substrate trigger hypothesis. Upon substrate binding, E287α no longer forms an H-bond to the ribose 2'OH, while S247α now donates an H-bond to the c-acetamide side chain of AdoCbl instead of the adenine N3 atom. As a result, in the ternary complex the 2'OH of ribose forms an intramolecular H-bond to the adenine N3 atom. This interaction allows the Ado• radical to rotate more freely, thus reducing the energetic penalty for positioning this species near substrate. An analogous intramolecular H-bonding interaction has been observed in the crystal structure of the AdoCbl-dependent enzyme glutamate mutase complexed with AdoCbl and substrate, although with a Co–C bond distance of 3.2 Å, this structure may not provide an accurate representation of the key H-bonding interactions that exist in the holoenzyme prior to Co–C bond homolysis.[14]

Torsional strain across the Ado glycosidic linkage has been hypothesized to serve as an important contributor to the destabilization of EAL-bound AdoCbl, as a significant rotation around this bond is required to move the 5'C of Ado• sufficiently close to substrate to allow for H-atom abstraction.[68] Our computational models support this hypothesis, as judged on the differences in the C4–N9–1'C–2'C torsional angle. While for the free AdoCbl model this angle is 149.9°, it is drastically reduced to 11.9° in our holoEAL model, comparable to 37.1° in the AdoCbl/EAL crystal

structure (PDB: 5YSN), and to 28.4° in the model of the ternary complex. In the crystal structure of the ternary complex with a cleaved Co–C bond (PDB: 5YSR), the torsional angle is much larger, at 101.3°, than in our model of the ternary complex.

Since in response to substrate EA binding to holoEAL, significant changes to the H-bonding interactions occur at the ribose 2'OH position, along with an increase in the glycosidic torsional angle and Co–C bond distance, it is possible that some of this torsional strain contributing to the AdoCbl destabilization in the ternary complex is caused by the 2'OH–E287α H-bonding interaction. According to this scenario, the distortions imposed by the ribose 2'OH interactions will no longer contribute to the same ground state destabilization and accelerated Co–C homolysis upon EA binding to the E287αX variants. Indeed, while the E287αX variants investigated previously[30] displayed some trace catalytic activity, in agreement with our results that indicate that ribose conformational strain is not the only factor accelerating Co–C bond homolysis, these variants exhibited a similarly elongated Co–N_{DMB} bond regardless of the ribose distortion.

Though changes occur to the H-bonding network involving S247α and E287α upon EA binding, there is a single H-bonding interaction between the ribose 3'OH and N193α that does not change. Intriguingly, this H-bonding interaction is the only one expected to provide a direct link between substrate and Ado positioning, as N193α serves as an H-bond donor to both the ribose 3'OH and the hydroxyl group of EA. This interaction is preserved during Co–C bond cleavage, allowing EAL to exert a level of control over the Ado• trajectory in the ternary complex that is coupled to EA binding. Support for this hypothesis is provided by a prior experimental study of the N193αD EAL variant, which was shown to exhibit some catalytically activity while undergoing inactivation four times faster than WT EAL.[30]

While the kinetic properties of EAL variants with substitutions at the S247α position have yet to be determined, the role of the analogous residue in the AdoCbl-dependent class II eliminase diol dehydratase (DD), S224α, was previously investigated.[69] The greatly diminished activity of S224αA and S224αN variants led the authors to propose that H bonding between this residue and the substrate is essential for substrate positioning and for protecting highly reactive intermediates from undesirable side reactions. This conclusion is consistent with our findings. In EAL, the H-bond donation from S247α to adenine N3 may serve to contribute to the Ado strain and disfavor formation of the intramolecular H-bond when substrate is absent. Because formation of this intramolecular H-bond appears to contribute to the accelerated Co–C homolysis, S247α likely plays a key role in the Co–C bond activation mechanism employed by EAL.

3.4. Conclusions

This study provides experimental and computational evidence that different mechanisms of Co–C bond activation are employed by the AdoCbl-dependent class I isomerases and class II eliminases. Unlike MMCM and GM, which promote the rate of Co–C bond homolysis through stabilization of the post-homolysis products, EAL displays no appreciable stabilization of Co^{II}Cbl or Ado$^{\bullet}$. Instead, our computed structural models of AdoCbl, holoEAL, and the ternary complex indicate that EAL activates the AdoCbl ground state by elongation of the Co–axial bonds and through Ado distortions. PECs of these structural models indicate a +12.7 kcal/mol ground-state destabilization of AdoCbl when enzyme-bound in the presence of substrate, consistent with published experimental data.

References

(1) Bandarian, V.; Reed, G. H. Isotope Effects in the Transient Phases of the Reaction Catalyzed by Ethanolamine Ammonia-Lyase: Determination of the Number of Exchangeable Hydrogens in the Enzyme-Cofactor Complex. *Biochemistry* **2000**, *39* (39), 12069–12075. https://doi.org/10.1021/bi001014k.

(2) Chowdhury, S.; Banerjee, R. Thermodynamic and Kinetic Characterization of Co−C Bond Homolysis Catalyzed by Coenzyme B_{12}-Dependent Methylmalonyl-CoA Mutase. *Biochemistry* **2000**, *39* (27), 7998–8006. https://doi.org/10.1021/bi992535e.

(3) Bradbeer, C. The Clostridial Fermentations of Choline and Ethanolamine I. Preparation and Properties of Cell-Free Extracts. *J. Biol. Chem.* **1965**, *240* (12), 4669–4674.

(4) Chang, G. W.; Chang, J. T. Evidence for the B_{12}-Dependent Enzyme Ethanolamine Deaminase in *Salmonella*. *Nature* **1975**, *254* (5496), 150–151. https://doi.org/10.1038/254150a0.

(5) Roof, D. M.; Roth, J. R. Ethanolamine Utilization in Salmonella Typhimurium. *J. Bacteriol.* **1988**, *170* (9), 3855–3863. https://doi.org/10.1128/jb.170.9.3855-3863.1988.

(6) Bender, G.; Poyner, R. R.; Reed, G. H.; R. Poyner, R.; H. Reed, G. Identification of the Substrate Radical Intermediate Derived from Ethanolamine during Catalysis by Ethanolamine Ammonia-Lyase. *Biochemistry* **2008**, *47* (43), 11360–11366. https://doi.org/10.1021/bi801316v.

(7) Bothe, H.; Darley, D. J.; Albracht, S. P. J.; Gerfen, G. J.; Golding, B. T.; Buckel, W. Identification of the 4-Glutamyl Radical as an Intermediate in the Carbon Skeleton Rearrangement Catalyzed by Coenzyme B_{12}-Dependent Glutamate Mutase from Clostridium Cochlearium. *Biochemistry* **1998**, *37* (12), 4105–4113. https://doi.org/10.1021/bi971393q.

(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) Ke, S. C.; Warncke, K. Interactions of Substrate and Product Radicals with Co(II) in Cobalamin and with the Active Site in Ethanolamine Deaminase, Characterized by ESE-EPR and ^{14}N ESEEM Spectroscopies. *J. Am. Chem. Soc.* **1999**, *121* (43), 9922–9927. https://doi.org/10.1021/ja990395q.

(10) Warncke, K.; Utada, A. S. Interaction of the Substrate Radical and the 5′-Deoxyadenosine-5′-Methyl Group in Vitamin B_{12} Coenzyme-Dependent Ethanolamine Deaminase. *J. Am. Chem. Soc.* **2001**, *123* (35), 8564–8572. https://doi.org/10.1021/ja003658l.

(11) LoBrutto, R.; Bandarian, V.; Magnusson, O. T.; Chen, X.; Schramm, V. L.; Reed, G. H. 5′-Deoxyadenosine Contacts the Substrate Radical Intermediate in the Active Site of Ethanolamine Ammonia-Lyase: ^{2}H and ^{13}C Electron Nuclear Double Resonance Studies. *Biochemistry* **2001**, *40* (1), 9–14. https://doi.org/10.1021/bi001865s.

(12) Bandarian, V.; Reed, G. H. Analysis of the Electron Paramagnetic Resonance Spectrum of a Radical Intermediate in the Coenzyme B_{12}-Dependent Ethanolamine Ammonia-Lyase Catalyzed Reaction of S-2-Aminopropanol. *Biochemistry* **2002**, *41* (27), 8580–8588. https://doi.org/10.1021/bi0201217.

(13) Canfield, J. M.; Warncke, K. Geometry of Reactant Centers in the Co^{II}-Substrate Radical Pair State of Coenzyme B_{12}-Dependent Ethanolamine Deaminase Determined by Using Orientation-Selection-ESEEM Spectroscopy. *J. Phys. Chem. B* **2002**, *106* (34), 8831–8841. https://doi.org/10.1021/jp0207634.

(14) Gruber, K.; Reitzer, R.; Kratky, C. Radical Shuttling in a Protein: Ribose Pseudorotation Controls Alkyl-Radical Transfer in the Coenzyme B_{12} Dependent Enzyme Glutamate Mutase. *Angew. Chemie - Int. Ed.* **2001**, *40* (18), 3377–3380. https://doi.org/10.1002/1521-3773(20010917)40:18<3377::AID-ANIE3377>3.0.CO;2-8.

(15) Brunk, E.; Kellett, W. F.; Richards, N. G. J.; Rothlisberger, U. A Mechanochemical Switch to Control Radical Intermediates. *Biochemistry* **2014**, *53* (23), 3830–3838. https://doi.org/10.1021/bi500050k.

(16) Kumar, N.; Bucher, D.; Kozlowski, P. M. Mechanistic Implications of Reductive Co–C Bond Cleavage in B_{12}-Dependent Methylmalonyl CoA Mutase. *J. Phys. Chem. B* **2019**, *123* (10), 2210–2216. https://doi.org/10.1021/acs.jpcb.8b10820.

(17) Padmakumar, R.; Banerjee, R. Evidence from Electron Paramagnetic Resonance Spectroscopy of the Participation of Radical Intermediates in the Reaction Catalyzed by Methylmalonyl-Coenzyme A Mutase. *J. Biol. Chem.* **1995**, *270* (16), 9295–9300. https://doi.org/10.1074/jbc.270.16.9295.

(18) Abend, A.; Bandarian, V.; Nitsche, R.; Stupperich, E.; Rétey, J.; Reed, G. H. Ethanolamine Ammonia-Lyase Has a "Base-On" Binding Mode for Coenzyme B_{12}. *Arch. Biochem. Biophys.* **1999**, *370* (1), 138–141. https://doi.org/10.1006/abbi.1999.1382.

(19) Ke, S. C.; Torrent, M.; Museav, D. G.; Morokuma, K.; Warncke, K. Identification of Dimethylbenzimidazole Axial Coordination and Characterization of ^{14}N Superhyperfine and Nuclear Quadrupole Coupling in Cob(II)Alamin Bound to Ethanolamine Deaminase in a Catalytically-Engaged Substrate Radical-Cobalt(II) Biradical State. *Biochemistry* **1999**, *38* (39), 12681–12689. https://doi.org/10.1021/bi983067w.

(20) Shibata, N.; Tamagaki, H.; Hieda, N.; Akita, K.; Komori, H.; Shomura, Y.; Terawaki, S.-I.; Mori, K.; Yasuoka, N.; Higuchi, Y.; et al. Crystal Structures of Ethanolamine Ammonia-Lyase Complexed with Coenzyme B_{12} Analogs and Substrates. *J. Biol. Chem.* **2010**, *285* (34), 26484–26493. https://doi.org/10.1074/jbc.M110.125112.

(21) Shibata, N.; Sueyoshi, Y.; Higuchi, Y.; Toraya, T. Direct Participation of a Peripheral Side Chain of a Corrin Ring in Coenzyme B_{12} Catalysis. *Angew. Chemie Int. Ed.* **2018**, *57* (26), 7830–7835. https://doi.org/10.1002/anie.201803591.

(22) Dong, S.; Padmakumar, R.; Maiti, N.; Banerjee, R.; Spiro, T. G. Resonance Raman Spectra Show That Coenzyme B_{12} Binding to Methylmalonyl-Coenzyme A Mutase Changes the Corrin Ring Conformation but Leaves the Co-C Bond Essentially Unaffected. *J. Am. Chem. Soc.* American Chemical Society 1998, 9947–9948. https://doi.org/10.1021/ja981584w.

(23) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Bond Activation in B_{12}-Dependent Enzymes: Cryogenic Resonance Raman Studies of Methylmalonyl-Coenzyme A Mutase. *J. Am. Chem. Soc.* **1999**, *121* (30), 7063–7070. https://doi.org/10.1021/ja982753f.

(24) Brooks, A. J.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Spectroscopic and Computational Studies on the Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase: Evaluation of Enzymatic Contributions to Co–C Bond Activation in the Co^{3+} Ground State. *J. Am. Chem. Soc.* **2004**, *126* (26), 8167–8180. https://doi.org/10.1021/ja039114b.

(25) Brooks, A. J.; Fox, C. C.; Marsh, E. N. G.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Electronic Structure Studies of the Adenosylcobalamin Cofactor in Glutamate Mutase. *Biochemistry* **2005**, *44* (46), 15167–15181. https://doi.org/10.1021/bi051094y.

(26) Brooks, A. J.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Co-C Bond Activation in Methylmalonyl-CoA Mutase by Stabilization of the Post-Homolysis Product Co^{2+} Cobalamin. *J. Am. Chem. Soc.* **2005**, *127* (47), 16522–16528. https://doi.org/10.1021/ja0503736.

(27) Krouwer, J. S.; Holmquist, B.; Kipnes, R. S.; Babior, B. M. The Mechanism of Action of Ethanolamine Ammonia-Lyase, an Adenosylcobalamin-Dependent Enzyme Evidence That Carbon-Cobalt Bond Cleavage Is Driven in Part by Conformational Alterations of the Corrin Ring. *BBA - Enzymol.* **1980**, *612* (1), 153–159. https://doi.org/10.1016/0005-2744(80)90288-0.

(28) Robertson, W. D.; Wang, M.; Warncke, K. Characterization of Protein Contributions to Cobalt−Carbon Bond Cleavage Catalysis in Adenosylcobalamin-Dependent Ethanolamine Ammonia-Lyase by Using Photolysis in the Ternary Complex. *J. Am. Chem. Soc.* **2011**, *133* (18), 6968–6977. https://doi.org/10.1021/ja107052p.

(29) Chen, Z.-G. G.; Zietek, M. A.; Russell, H. J.; Tait, S.; Hay, S.; Jones, A. R.; Scrutton, N. S.; Ziętek, M. A.; Russell, H. J.; Tait, S.; et al. Dynamic, Electrostatic Model for the Generation and Control of High-Energy Radical Intermediates by a Coenzyme B_{12}-Dependent Enzyme. *ChemBioChem* **2013**, *14* (13), 1529–1533. https://doi.org/10.1002/cbic.201300420.

(30) Mori, K.; Oiwa, T.; Kawaguchi, S.; Kondo, K.; Takahashi, Y.; Toraya, T. Catalytic Roles of Substrate-Binding Residues in Coenzyme B_{12}-Dependent Ethanolamine Ammonia-Lyase. *Biochemistry* **2014**, *53* (16), 2661–2671. https://doi.org/10.1021/bi500223k.

(31) Fukuoka, M.; Nakanishi, Y.; Hannak, R. B.; Kräutler, B.; Toraya, T. Homoadenosylcobalamins as Probes for Exploring the Active Sites of Coenzyme B_{12}-Dependent Diol Dehydratase and Ethanolamine Ammonia-Lyase. *FEBS J.* **2005**, *272* (18), 4787–4796. https://doi.org/10.1111/j.1742-4658.2005.04892.x.

(32) Fukuoka, M.; Yamada, S.; Miyoshi, S.; Yamashita, K.; Yamanishi, M.; Zou, X.; Brown, K. L.; Toraya, T. Functions of the D-Ribosyl Moiety and the Lower Axial Ligand of the Nucleotide Loop of Coenzyme B_{12} in Diol Dehydratase and Ethanolamine Ammonia-Lyase Reactions. *J. Biochem.* **2002**, *132* (6), 935–943.

(33) A. Maier, J.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; E. Hauser, K.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from

Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(34) Marques, H. M.; Ngoma, B.; Egan, T. J.; Brown, K. L. Parameters for the AMBER Force Field for the Molecular Mechanics Modeling of the Cobalt Corrinoids. *J. Mol. Struct.* **2001**, *561* (1–3), 71–91. https://doi.org/10.1016/S0022-2860(00)00920-0.

(35) William L. Jorgensen, Jayaraman Chandrasekhar, J. D. M. Comparison of Simple Potential Functions for Simulating Liquid Water Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(36) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. E.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16, Revision C.01; Gaussian, Inc.: Wallingford CT, 2016.

(38) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377. https://doi.org/10.1063/1.464304.

(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(40) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.

(41) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97* (4), 2571–2577. https://doi.org/10.1063/1.463096.

(42) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A Complete Basis Set Model Chemistry. I. The Total Energies of Closed-Shell Atoms and Hydrides of the First-Row Elements. *J. Chem. Phys.* **1988**, *89* (4), 2193–2218. https://doi.org/10.1063/1.455064.

(43) Bayly, C. I.; Merz, K. M.; Ferguson, D. M.; Cornell, W. D.; Fox, T.; Caldwell, J. W.; Kollman, P. A.; Cieplak, P.; Gould, I. R.; Spellmeyer, D. C. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197. https://doi.org/10.1021/ja00124a002.

(44) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305. https://doi.org/10.1039/b508541a.

(45) Hirata, S.; Head-Gordon, M. Time-Dependent Density Functional Theory within the Tamm-Dancoff Approximation. *Chem. Phys. Lett.* **1999**, *314* (3–4), 291–299. https://doi.org/10.1016/S0009-2614(99)01149-5.

(46) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22). https://doi.org/10.1063/5.0004608.

(47) Weigend, F.; Köhn, A.; Hättig, C. Efficient Use of the Correlation Consistent Basis Sets in Resolution of the Identity MP2 Calculations. *J. Chem. Phys.* **2002**, *116* (8), 3175–3183. https://doi.org/10.1063/1.1445115.

(48) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065. https://doi.org/10.1039/b515623h.

(49) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Resonance Raman Co−C Stretching Frequencies Reflect Bond Strength Changes in Alkyl Cobalamins, but Are Unaffected by Trans Ligand Substitution. *J. Am. Chem. Soc.* **1996**, *118* (38), 9182–9183. https://doi.org/10.1021/ja962003a.

(50) Stich, T. A.; Buan, N. R.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{2+}corrinoids: Spectral and Electronic Properties of the Biologically Relevant Base-on and Base-off Forms of Co^{2+}Cobalamin. *J. Am. Chem. Soc.* **2004**, *126* (31), 9735–9749. https://doi.org/10.1021/ja0481631.

(51) Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C.; A. Stich, T.; J. Brooks, A.; R. Buan, N.; C. Brunold, T. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914. https://doi.org/10.1021/ja029328d.

(52) Hollaway, M. R.; White, H. A.; Joblin, K. N.; Johnson, A. W.; Lappert, M. F.; Wallis, O. C. A Spectrophotometric Rapid Kinetic Study of Reactions Catalysed by Coenzyme-B_{12}-Dependent Ethanolamine Ammonia-Lyase. *Eur. J. Biochem.* **1978**, *82* (1), 143–154. https://doi.org/10.1111/j.1432-1033.1978.tb12005.x.

(53) Mitić, N.; Clay, M. D.; Saleh, L.; Bollinger, J. M.; Solomon, E. I. Spectroscopic and Electronic Structure Studies of Intermediate X in Ribonucleotide Reductase R2 and Two Variants: A Description of the Fe^{IV}-Oxo Bond in the Fe^{III}–O–Fe^{IV} Dimer. *J. Am. Chem. Soc.* **2007**, *129* (29), 9049–9065. https://doi.org/ 10.1021/ja070909i.

(54) Warncke, K.; Schmidt, J. C.; Ke, S. C. Identification of a Rearranged-Substrate, Product Radical Intermediate and the Contribution of a Product Radical Trap in Vitamin B_{12} Coenzyme-Dependent Ethanolamine Deaminase Catalysis. *J. Am. Chem. Soc.* **1999**, *121* (45), 10522–10528. https://doi.org/10.1021/ja984005w.

(55) Sun, L.; Groover, O. A.; Canfield, J. M.; Warncke, K. Critical Role of Arginine 160 of the EutB Protein Subunit for Active Site Structure and Radical Catalysis in Coenzyme B_{12}-Dependent Ethanolamine Ammonia-Lyase. *Biochemistry* **2008**, *47* (20), 5523–5535.

(56) Ivancich, A.; Jouve, H. M.; Sartor, B.; Gaillard, J. EPR Investigation Of-Compound I in Proteus Mirabilis and Bovine Liver Catalases: Formation of Porphyrin and Tyrosyl Radical Intermediates. *Biochemistry* **1997**, *36* (31), 9356–9364. https://doi.org/10.1021/bi970886s.

(57) A. Stich, T.; J. Brooks, A.; R. Buan, N.; C. Brunold, T. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914.

https://doi.org/10.1021/ja029328d.

(58) Mebs, S.; Henn, J.; Dittrich, B.; Paulmann, C.; Luger, P. Electron Densities of Three B_{12} Vitamins. *J. Phys. Chem. A* **2009**, *113* (29), 8366–8378. https://doi.org/10.1021/jp902433x.

(59) Rovira, C.; Kozlowski, P. M. First Principles Study of Coenzyme B_{12}. Crystal Packing Forces Effect on Axial Bond Lengths. *J. Phys. Chem. B* **2007**, *111* (12), 3251–3257. https://doi.org/10.1021/jp0660029.

(60) Finke, B. P. H. and R. G. Thermolysis of the Co–C Bond of Adenosylcobalamin. 2. Products, Kinetics, and Co–C Bond Dissociation Energy in Aqueous Solution. *J. Am. Chem. Soc.* **1986**, *108*, 4820–4829. https://doi.org/10.1021/ja00276a020.

(61) Hirao, H. Which DFT Functional Performs Well in the Calculation of Methylcobalamin? Comparison of the B3LYP and BP86 Functionals and Evaluation of the Impact of Empirical Dispersion Correction. *J. Phys. Chem. A* **2011**, *115* (33), 9308–9313. https://doi.org/10.1021/jp2052807.

(62) Jensen, K. P.; Ryde, U. Theoretical Prediction of the Co–C Bond Strength in Cobalamins. *J. Phys. Chem. A* **2003**, *107* (38), 7539–7545. https://doi.org/10.1021/jp027566p.

(63) J. Brooks, A.; Vlasie, M.; Banerjee, R.; C. Brunold, T. Co−C Bond Activation in Methylmalonyl-CoA Mutase by Stabilization of the Post-Homolysis Product Co^{2+}Cobalamin. *J. Am. Chem. Soc.* **2005**, *127* (47), 16522–16528. https://doi.org/10.1021/ja0503736.

(64) Bender, G.; R. Poyner, R.; H. Reed, G. Identification of the Substrate Radical Intermediate Derived from Ethanolamine during Catalysis by Ethanolamine Ammonia-Lyase. *Biochemistry* **2008**, *47* (43), 11360–11366. https://doi.org/10.1021/bi801316v.

(65) D. Román-Meléndez, G.; von Glehn, P.; N. Harvey, J.; J. Mulholland, A.; Neil G. Marsh, E. Role of Active Site Residues in Promoting Cobalt–Carbon Bond Homolysis in Adenosylcobalamin-Dependent Mutases Revealed through Experiment and Computation. *Biochemistry* **2013**, *53* (1), 169–177. https://doi.org/10.1021/bi4012644.

(66) Chen, Z. G.; Zietek, M. A.; Russell, H. J.; Tait, S.; Hay, S.; Jones, A. R.; Scrutton, N. S. Dynamic, Electrostatic Model for the Generation and Control of High-Energy Radical Intermediates by a Coenzyme B_{12}-Dependent Enzyme. *ChemBioChem* **2013**, *14* (13), 1529–1533. https://doi.org/10.1002/cbic.201300420.

(67) Jones, A. R.; Hardman, S. J. O.; Hay, S.; Scrutton, N. S. Is There a Dynamic Protein Contribution to the Substrate Trigger in Coenzyme B_{12}-Dependent Ethanolamine Ammonia Lyase? *Angew. Chemie - Int. Ed.* **2011**, *50* (46), 10843–10846. https://doi.org/10.1002/anie.201105132.

(68) Shibata, N.; Toraya, T. Structural Basis for the Activation of the Cobalt-Carbon Bond and Control of the Adenosyl Radical in Coenzyme B_{12} Catalysis. *ChemBioChem*. John Wiley & Sons, Ltd March 14, 2023, p e202300021. https://doi.org/10.1002/cbic.202300021.

(69) Ogura, K. I.; Kunita, S. I.; Mori, K.; Tobimatsu, T.; Toraya, T. Roles of Adenine Anchoring and Ion Pairing at the Coenzyme B_{12}-Binding Site in Diol Dehydratase Catalysis. *FEBS J.* **2008**, *275* (24), 6204–6216. https://doi.org/10.1111/j.1742-

4658.2008.06745.x.

Chapter 4

Molecular-Level Insight into the Catalytic Cycle of a PduO-type Adenosyltransferase

Chapter 4: Molecular-Level Insight into the Catalytic Cycle of a PduO-type Adenosyltransferase

4.1. Introduction

Adenosylcobalamin (AdoCbl, Figure 4.1), also called coenzyme B_{12}, serves as an essential cofactor for a diverse range of enzyme-catalyzed reactions across all areas of life, from heteroatom eliminations to carbon skeleton rearrangements to amino group migrations.[1–3] Notably, these reactions include the conversion of methylmalonyl-CoA to succinyl-CoA by the human enzyme methylmalonyl CoA-mutase (MMCM), a crucial step in the catabolism of branched-chain amino acids, cholesterol, and odd-chain fatty acids.[4] AdoCbl is a member of the extended family of cobalamins (Cbls) that feature a cobalt (Co) ion equatorially ligated by four nitrogens from a tetrapyrrole macrocycle termed the corrin ring. In the Co^{III} form, Cbls are primarily six-coordinate (6C), featuring a pendant dimethylbenzimidazole (DMB) base as the lower axial ligand and a variable upper axial ligand that distinguishes the different members of this family. The enzymatically competent forms of B_{12}, AdoCbl and methylcobalamin (MeCbl), provide rare biological examples of an organometallic bond. The Co–C bond allows AdoCbl to act as a radical reservoir; homolytic cleavage of this bond produces Co^{II}cobalamin (Co^{II}Cbl) and a 5'-deoxyadenosyl radical (Ado$^{\bullet}$), the latter of which abstracts a hydrogen atom from the substrate to trigger a radical-based rearrangement reaction that leads to product formation.[5] Alternatively, heterolytic cleavage of the Co–C bond of MeCbl to yield Co^{I}Cbl and a methyl cation allows enzymes to catalyze methyl transfer reactions.

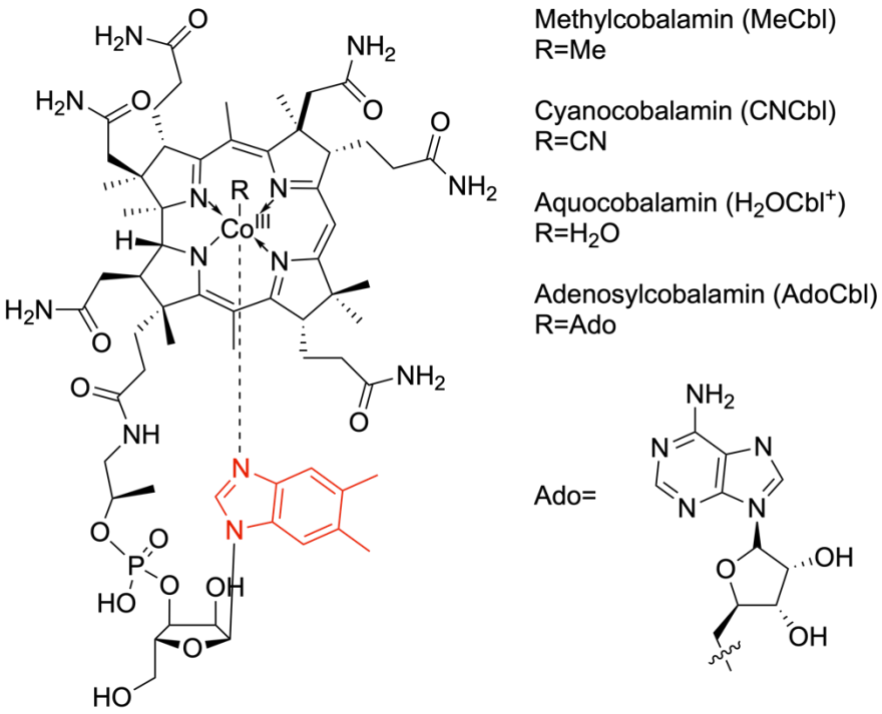


Figure 4.1. Chemical structure of Co^{III}cobalamins (Co^{III}Cbls), where **R** indicates the variable upper axial ligand. The coordinating dimethylbenzimidazole (DMB, shown in red) at the lower face can be displaced upon enzyme binding.

The critical work performed by AdoCbl-dependent enzymes is made possible by a group of enzymes called adenosyltransferases (ATRs) that catalyze the transfer of an adenosyl group to an AdoCbl precursor.[6] There are several contexts in which this adenosylation is required: (i) the *de novo* synthesis of AdoCbl, performed solely by prokaryotes, (ii) the replacement of a different upper axial ligand of exogenous Cbls, as is necessary in all species that lack the enzymatic machinery to synthesize AdoCbl independently, and (iii) the reactivation of inactivated AdoCbl in the infrequent event that the Ado• moiety escapes the active site of an AdoCbl-dependent enzyme during turnover.[7,8] Some ATRs have been shown to act as chaperones as well, delivering the AdoCbl product directly to their target enzymes, rather than relying on diffusion-based trafficking of the valuable and reactive cofactor.[7–10] Disruptions of either the catalytic or the chaperone role of ATRs can have devastating effects. Malfunction of the human adenosyltransferases (hATR), for example, leads to the metabolic disorders methylmalonic aciduria and ketoacidosis.[11,12]

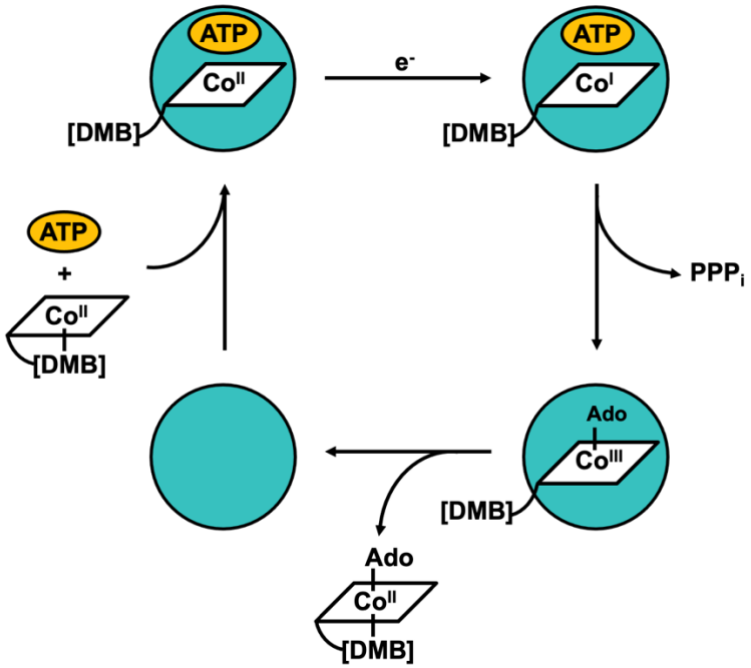


Figure 4.2. Mechanism for the adenosylation of Co^{II}Cbl by ATRs. Co^{II}Cbl is bound in a rare four-coordinate, base-off conformation by the ATR (teal circle), along with cosubstrate ATP (yellow oval), and is subsequently reduced to Co^{I}Cbl. Direct nucleophilic attack from the "supernucleophile" Co^{I}Cbl on ATP yields AdoCbl and tripolyphosphate (PPP$_i$).

Three structurally unrelated and evolutionarily distinct classes of ATRs exist that play slightly different biological roles. CobA, known as a "housekeeping" enzyme in *Salmonella*, has the broadest substrate scope and can adenosylate a wider range of both scavenged and *de novo* synthesized AdoCbl precursors.[13,14] Alternatively, PduO and EutT are central to the metabolism of propanediol and ethanolamine, respectively, in certain bacteria. The PduO-type ATRs also include hATR.[15,16] Despite differences in amino acid sequence, structure, and substrate scope, the three ATR classes share a common catalytic cycle (Figure 4.2). In brief, when cosubstrate ATP is present, ATRs bind Co^{II}Cbl and reduce it to the Co^{I}Cbl, long known to be one of the most reactive nucleophiles in biology.[17] This Co^{I}Cbl species then attacks the 5' carbon of ATP to produce AdoCbl and tripolyphosphate (PPP$_i$). The thermodynamically challenging reduction step of this cycle is particularly intriguing, as the Co^{II}/Co^{I}Cbl reduction potential is -610 mV vs SHE, well outside the range of readily available biological reductants.[18] When ATR-bound in the presence of cosubstrate

ATP, however, Co^{II}Cbl can be reduced by flavodoxin A (E (SHE) = -440 mV),[19] meaning that ATRs must therefore raise the reduction potential of the Co^{II}/Co^{I}Cbl redox couple by several hundred mVs. Indeed, recent work by Gouda et al. revealed that ATR-bound Co^{II}Cbl has a reduction potential of -325 mV.[21]

The mechanism by which ATRs raise the Co^{II}/Co^{I}Cbl reduction potential and lower the activation energy barrier for nucleophilic attack on ATP has been the subject of much research over the past two decades. A key mechanistic component was first uncovered through magnetic circular dichroism (MCD) and electronic paramagnetic resonance (EPR) studies of Co^{II}Cbl bound to hATR and CobA. Novel MCD features and unusually large EPR g shifts and hyperfine coupling constants all pointed towards the formation of an unusual four-coordinate (4C), square-planar Co^{II}Cbl species.[22,23] Because the redox-active molecular orbital (MO) of Co^{II}Cbl is Co $3d_{z2}$-based and thus pointing toward the axial ligand, it was proposed that removal of the DMB ligand allows ATRs to raise the Co^{II}/Co^{I}Cbl reduction potential into the physiologically accessible range. Computational and spectroscopic studies of Co^{II}Cbl confirmed that the Co $3d_{z2}$-based MO is stabilized in the 4C conformation, reducing the thermodynamic barrier for the reduction of the cofactor.[24] Since these early studies, similar 4C Co^{II} species have been documented for other PduO-type as well as two EutT-type ATRs through spectroscopic and, in certain cases, X-ray crystallographic methods.[25–28] The 4C Co^{II}Cbl conformation is imposed, at least in part, by a noncoordinating hydrophobic protein-derived residue that displaces the DMB ligand from the lower axial coordination site and precludes water binding. Substitutions of this residue have been shown to result in drastic decreases in catalytic efficiency because the Co^{II}Cbl is no longer forced to adopt a 4C conformation, precluding the reduction to the Co^{I}Cbl state.[27,29]

This chapter describes a computational investigation of the Co^{II} → Co^{I}Cbl reduction step and subsequent adenosylation of Co^{I}Cbl by PduO. Published crystallographic data were used in conjunction with quantum mechanics/molecular mechanics (QM/MM) computations to generate whole-protein models for several states in the PduO catalytic cycle. Using the optimized structure of PduO/Co^{I}Cbl/ATP as the starting point, a relaxed potential energy scan for Co–C bond formation was performed to assess how the energy and composition of the redox-active Co $3d_{z2}$-based MO change as the Co–C distance is shortened. Single-point calculations demonstrate that in going from free, 5C Co^{II}Cbl to PduO-bound, 4C Co^{II}Cbl, the removal of the sizable Co–N_{DMB} σ* contribution leads to a substantial stabilization of the Co $3d_{z2}$-based redox-active MO. These calculations predict an increase in Co^{II}/Co^{I}Cbl reduction potential by +89 mV from 5C to 4C Co^{II}Cbl, in qualitative agreement with experimental data.[21]

4.2. Materials and Methods

Structural models of free and PduO-bound cobalamins were optimized using the QM/MM approach, as implemented through the ONIOM method in Gaussian 16.[30] For the enzyme-bound models, initial coordinates were obtained from two high-resolution crystal structures of WT PduO from *Lactobacillus reuteri*,[26] one at 1.90 Å resolution containing bound Co^{II}Cbl and ATP (PDB: 3CI1) and the other at 1.11 Å resolution containing AdoCbl and PPP_i (PDB: 3CI3). The former was used to construct models of PduO/Co^{II}Cbl/ATP and PduO/Co^{I}Cbl/ATP, while the latter was used to model PduO/AdoCbl/PPP_i. It should be noted that in both crystal structures, the lack of well-resolved electron density beyond the bridging phosphate of the nucleotide loop indicated that the DMB ligand was disordered in these structures. Consistent with this observation, a third crystal structure obtained in this study of PduO complexed with ATP and Co^{II}cobinamide ($Co^{II}Cbi^{+}$), a truncated B_{12} derivative that lacks the DMB moiety and serves as a model of base-off Co^{II}Cbl,

showed almost identical binding modes of Co^{II}Cbl and $Co^{II}Cbi^{+}$. This suggests that the DMB ligand has no significant impact on the structure of the cofactor or active site. The computational models used in the present study were therefore constructed without the *in silico* addition of DMB. Published X-ray crystallographic data of 5C Co^{II}Cbl were used as the initial coordinates for models of free Co^{II}Cbl.[31] A truncated model of free Co^{I}Cbl was generated starting from coordinates extracted from the 3CI1 crystal structure in which the DMB moiety is not resolved because repeated attempts at geometry optimizations of free Co^{I}Cbl starting from the crystallographic Co^{II}Cbl coordinates invariably failed to converge due to continued movement of the unbound nucleotide loop and DMB moiety.

Protein and water H atoms were added to each structure using the program PDBtoPQR,[32] while cofactor H atoms were added via PyMOL.[33] As PduO is a trimer with three independent active sites, two copies of the cofactor in each model were treated entirely with MM. The QM region of all PduO models included relevant portions of the third active site and was defined as the cobalt atom, corrin ring, first carbon of each side chain, F112, and the adenosyl and phosphate moieties (either as ATP or in the form of the Ado ligand and PPP_i). In free Co^{II}Cbl and Co^{I}Cbl, the QM region included the cobalt atom, corrin ring, first carbon of each side chain, and the DMB moiety. The QM region was treated with density functional theory (DFT) using the B3LYP functional,[34,35] the TZVP basis set[36] for Co and all coordinated atoms, and the 6-31G* basis set[37–46] for all other QM atoms. The MM region was treated using the AMBER force field, along with AMBER-compatible parameters reported in the literature for Cbls[47] and ATP.[48] RESP partial charges for the Cbl portion were also taken from the literature.[49] The QM/MM-optimized model of PduO/Co^{I}Cbl/ATP was then used as the starting structure for a relaxed potential energy scan (PES) of Co–C bond formation by reducing the Co···5’ C distance in 0.05 Å increments using

Gaussian's ModRedundant functionality, which re-optimizes the positions of all other atoms after each step.

The QM/MM geometry-optimized structures obtained for Co^{II}Cbl, Co^{I}Cbl, PduO/Co^{II}Cbl/ATP, PduO/Co^{I}Cbl/ATP, PduO/AdoCbl/PPP_i, and each step of the Co–C bond formation PES were subsequently used in single point (SP) DFT calculations with Orca 4.2.1[50] using the TZVP basis set for all atoms. For open shell (i.e., Co^{II}Cbl-containing) models, unrestricted natural orbitals (UNOs) were calculated in addition to canonical orbitals from SCF calculations in order to identify the redox-active MO. For all enzyme-containing structures, several models of different sizes were tested, containing either the entirety of the QM region or only the atoms from Cbl and residue F112 included in the QM region. Isosurface plots of MOs were generated using an isodensity value of .04 au and visualized in PyMOL.[33]

Using the approach outlined by Noodleman and coworkers,[51] the free energies of these QM/MM-optimized models were used to calculate the difference in Co^{II}/Co^{I}Cbl reduction potential between the free and PduO-bound cofactor. This original method included a proton transfer step, which can be omitted in the present case, leading to the following simplified equation:

$$E^{\circ} = \mathrm{IP}_{(g)} + \Delta\varepsilon_{\mathrm{sol-ET}} - 4.43\ \mathrm{V} \quad (1)$$

where $IP_{(g)}$ is the gas-phase ionization potential for the reduced species, $\Delta\varepsilon_{\mathrm{sol-ET}}$ is the difference in solvation energies of the oxidized and reduced species, and –4.43 V is a correction term for the standard hydrogen electrode. Because of how these single point calculations were performed in Orca, the $IP_{(g)}$ and $\Delta\varepsilon_{\mathrm{sol-ET}}$ terms could not be evaluated separately. However, because the goal of the present study was to determine differences in the Co^{II}/Co^{I}Cbl reduction potential rather than to compute absolute E^0 values, the $\Delta\varepsilon_{\mathrm{sol-ET}}$ term could be ignored, as this term should be similar

for free and PduO-bound Co^{II}Cbl given the solvent-exposed nature of the PduO active site. The approximate redox potentials were therefore computed by simply taking the difference in total energies between the oxidized and reduced systems, with 4.43 V subtracted. In order for these differences in energy to be meaningful, the oxidized and reduced species in each case must contain the same set of atoms. Thus, for the free cofactor, the energy of the reduced species was defined as the sum of the energy of the truncated Co^{I}Cbl model, which did not include the DMB moiety, and the energy of a QM-optimized DMB model. For the PduO-bound cofactor, SP energies for models that included only the Cbl and F112 atoms were used to calculate the reduction potential.

4.3. Results and Discussion

4.3.1. Active site structures of QM/MM-optimized PduO models

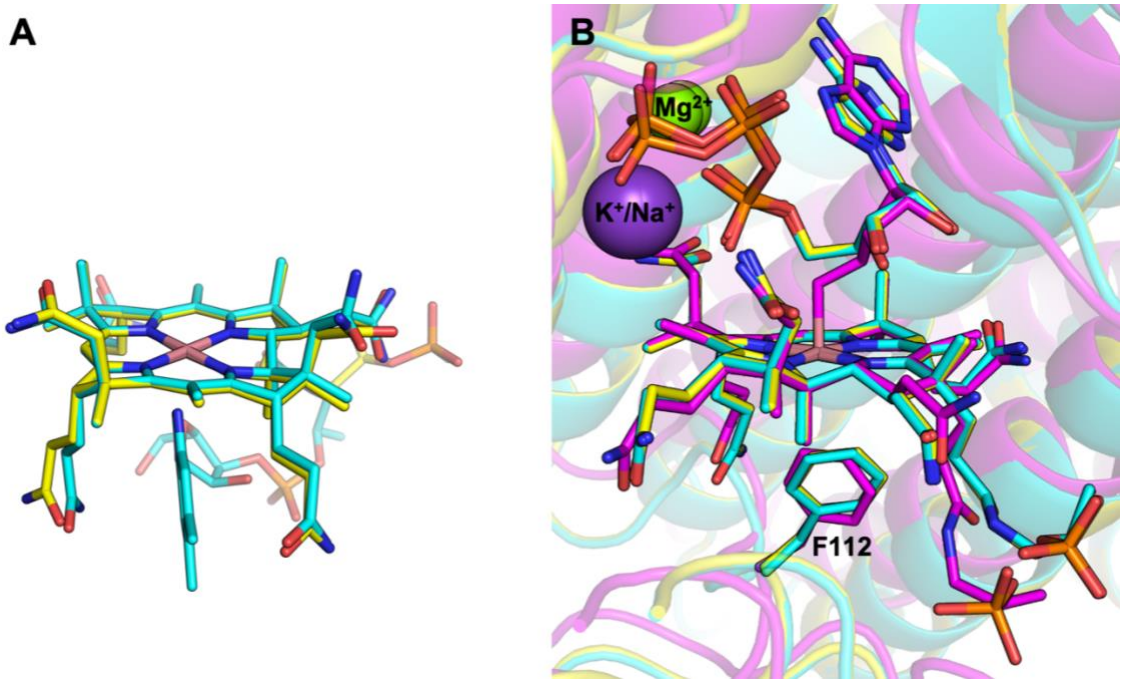


Figure 4.3. A. QM/MM-optimized structures of free Co^{II}Cbl (cyan) and truncated Co^{I}Cbl (yellow). B. QM/MM-optimized structures of PduO/Co^{II}Cbl/ATP (cyan), PduO/Co^{I}Cbl/ATP (yellow), and PduO/AdoCbl/PPP_i (pink).

Table 4.1. Relevant structural parameters derived from the crystal structures and QM/MM-optimized structures of PduO-bound Cbls. Distances are in Å and angles are in degrees

Model	**Corrin fold angle[a]**	**Co···5' C distance**	**5' C···O distance**	**Co···C_{F112} distance[b]**	**RMSD of protein backbone atoms[c]**
3CI1 crystal structure	17.7	2.97	1.46	3.57	-
PduO/Co^{II}Cbl/ATP	14.2	3.02	1.42	3.43	0.711
PduO/Co^{I}Cbl/ATP	13.4	3.09	1.44	3.43	0.708
3CI3 crystal structure	3.75	2.05	2.49	3.59	-
PduO/AdoCbl/PPP_i	13.8	2.05	2.54	3.69	0.793

[a.] The angle between the normal vectors of two fit planes, (C4, C5, C6, C9, C10, N21, N22) and (C10, C11, C14, C15, C16, N23, N24). [b.] The distance between the Co atom and closest C atom of F112. [c.] Determined for each optimized model in comparison with the crystal structure used as starting coordinates.

Complete models of PduO/Co^{II}Cbl/ATP, PduO/Co^{I}Cbl/ATP, PduO/AdoCbl/PPP_i, free Co^{II}Cbl, and free Co^{I}Cbl were constructed through QM/MM optimizations of crystallographic coordinates. For each of the PduO models, the RMSD of the protein backbone atoms is well below the resolution of the starting crystal structure, indicating that the QM/MM-optimized structures are in good agreement with the crystallographic data (see Table 4.1). Inspection of the relevant structural parameters of the optimized PduO models reveals that that the corrin fold angles of the Co^{II}Cbl and Co^{I}Cbl substrated are larger than in free Co^{II}Cbl (by 11.2° and 7.8°, respectively) and in excellent agreement with the X-ray crystal structure of PduO/Co^{II}Cbl/ATP (PDB: 3CI1). This result is also consistent with published resonance Raman (rRaman) data, which indicated that PduO binding perturbs the geometric structure of the corrin ring.[15] It is worth noting that reduction of PduO/Co^{II}Cbl/ATP produces only minor changes in the structural parameters of the cofactor, with a slight decrease in corrin fold angle from 14.2° to 13.4 ° and an increase in Co···5' C distance

from 3.02 Å to 3.09 Å. These minute changes are expected to lead to a small reorganization energy, and thus a high rate, for the Co^{II}Cbl reduction.

The data in Table 4.1 also demonstrate that key characteristics of the active site are maintained, including the positioning of residues that are known to be essential to the enzyme's normal catalytic and chaperone activities. A salt bridge between R128 and D35 (equivalent to R186 and E193 in hATR) is required for normal enzymatic activity, demonstrated through both in vitro studies and the fact that the most common gene mutation associated with the disease methylmalonyl aciduria gives rise to the R186W substitution.[29,52,53] R132 (R190 in hATR), another residue known to be substituted in some cases of methylmalonyl aciduria, interacts directly with the ATP cosubstrate and has been shown to play a role in both the catalytic and chaperone functions of PduO-type ATRs.[10,29,54] The hydrogen-bonding interactions of these active site residues with each other and with the two substrates remain intact in all QM/MM-optimized structures (Figure 4.4)

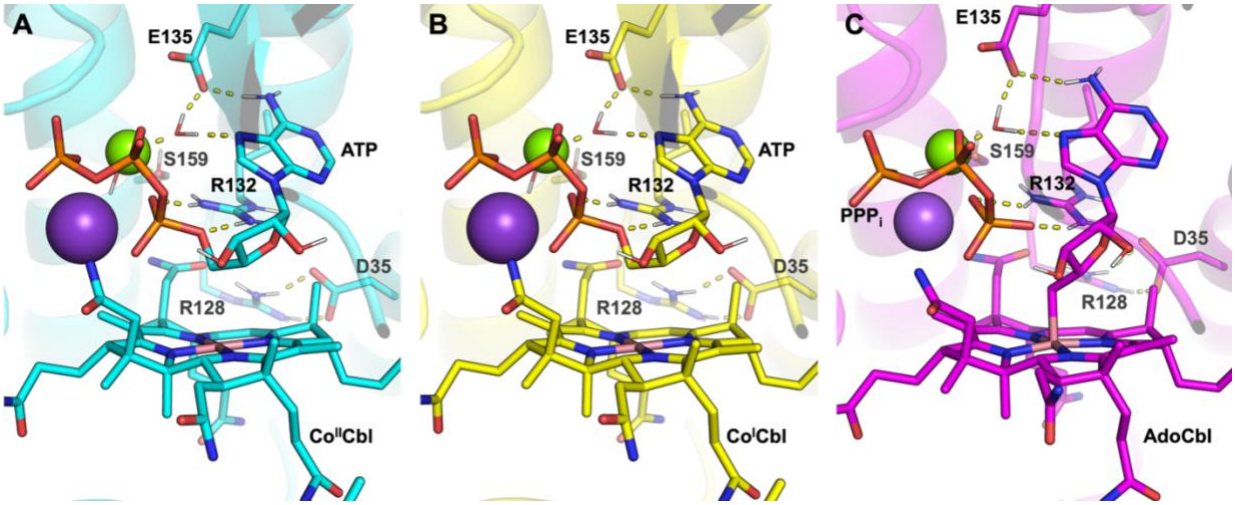


Figure 4.4. Active site region of QM/MM-optimized structures of (A) PduO/Co^{II}Cbl/ATP, (B) PduO/Co^{I}Cbl/ATP, and (C) PduO/ATP/PPP_i with key hydrogen-bonding interactions displayed by dashed lines.

4.3.2. Computed redox potential

The SP energies of the relevant QM/MM-optimized models were used to calculate the change in Co^{II}/Co^{I}Cbl potential upon Co^{II}Cbl binding to PduO. As described above, $E°$ was calculated as the difference in free energy between the oxidized and reduced species minus a correction term of 4.43 V. While the absolute computed values of $E°$ will contain systematic errors because the $\Delta\varepsilon_{sol-ET}$ in Eq 1 was neglected, differences in computed $E°$ values should reproduce experimental trends due to cancelation of this term.[57] Through this approach, the reduction potentials of the free and PduO-bound cofactor were determined to be equal to –8.538 V and –8.449 V respectively. PduO is therefore predicted to raise the Co^{II}/Co^{I}Cbl reduction potential by 0.089 V, which is in good agreement with the experimentally determined value of 0.180 V.[21]

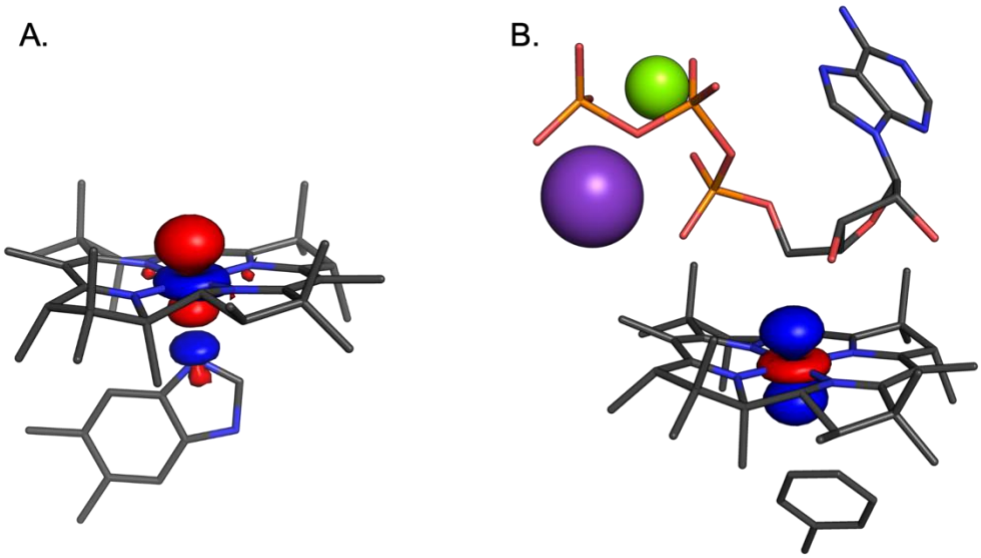


Figure 4.5. Isosurface plots of the Co $3d_{z2}$-based singly occupied unrestricted natural orbitals (UNOs) of (A) free 5C Co^{II}Cbl and (B) PduO/Co^{II}Cbl/ATP.

The origin of this difference in reduction potential can be rationalized in terms of the different MO compositions and energies of free Co^{II}Cbl and PduO/Co^{II}Cbl/ATP. Because upon $Co^{II}\rightarrow Co^{I}$Cbl reduction an electron is added to the singly occupied Co $3d_{z2}$-based MO, this orbital is of particular interest. However, a direct analysis of this orbital is complicated by the spin unrestricted formalism that must be applied to these open shell calculations, which produces

slightly different energies and compositions of the spin up and spin down MOs, particularly for those with predominant Co 3d orbital character. To address this issue, the canonical orbitals were transformed into unrestricted natural orbitals (UNOs), which allowed us to readily identify the Co $3d_{z2}$-based orbital as the redox active, singly occupied MO (SOMO), consistent with literature precedent.[22,58] Plots of the SOMO (Figure 4.9) highlight the difference in orbital composition for the free 5C species and the PduO-bound 4C species. With DMB present, the SOMO contains a substantial Co–N_{DMB} σ* interaction, raising its energy to +1.263 eV above the corrin π*-based LUMO, compared to +0.338 eV in the 4C PduO-bound Co^{II}Cbl.

4.3.3. Potential Energy Scan (PES)

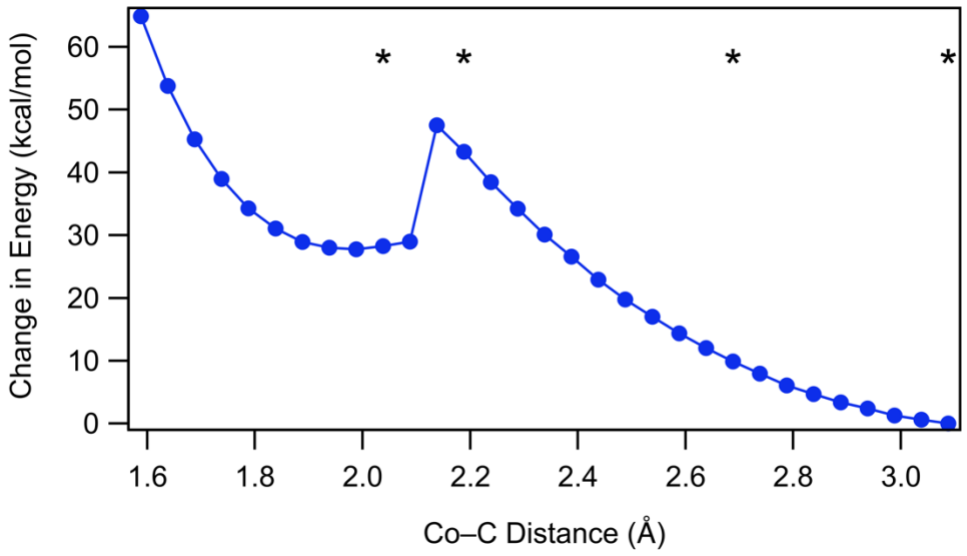


Figure 4.6. Computed PES for Co–C bond formation in PduO. Energies are from DFT SP calculations using all atoms included in the QM region of the whole-protein, relaxed QM/MM potential energy scan. Asterisks indicate structures referenced in Figures 4.7.

In order to explore the mechanistic details of the PduO-catalyzed nucleophilic attack of Co^{I}Cbl on ATP in greater detail, a relaxed potential energy scan (PES) was performed, starting with the optimized structure of PduO/Co^{I}Cbl/PPP_i (with a Co···C distance of 3.09 Å) and decreasing the Co···5' C distance by 0.05 Å per step for 30 steps. The energy of the QM region curves upwards for the first 19 steps, to a Co–C distance of 2.14 Å and an energy 47.5 kcal/mol

higher than the initial structure, before dropping down to a lower potential energy curve. The fact that the computed PES shows a sudden decrease in energy between steps 19 and 20 (reminiscent of a transition between two potential energy surfaces) indicates that the Co–C bond formation is likely a multiconfigurational problem that cannot be accurately modeled using DFT. Nevertheless, this process successfully recreated our 3CI3-based structure of PduO/AdoCbl/PPP_i in the 21st step of the scan, with a Co–C bond length of 2.04 Å and almost identical positioning of PPP_i, key active site residues, and AdoCbl (Figure 4.6). This resemblance leads us to conclude that the calculated relaxed potential energy surface is experimentally meaningful, representing a relevant pathway from the PduO-bound reactants to AdoCbl and PPP_i. While the structure obtained in step 21 was calculated to be 28.3 kcal/mol higher in energy than the starting model of PduO/Co^{I}Cbl/ATP, this energetic penalty may be offset by the formation of favorable interactions in the product state that are not accounted for when only the QM region is considered.

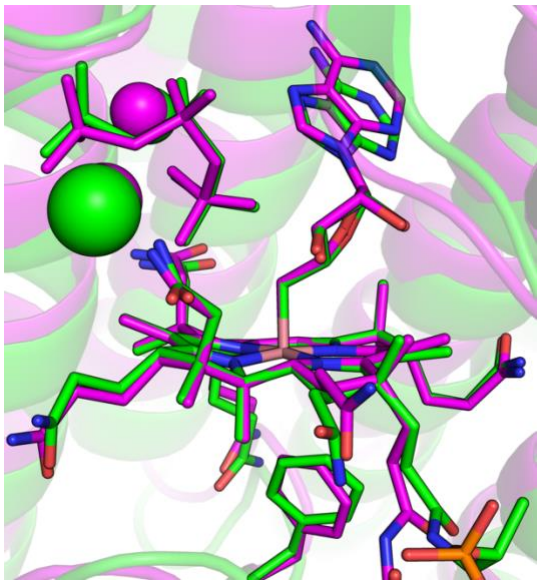


Figure 4.7. Active site regions of QM/MM-optimized PduO/AdoCbl/PPP_i model (pink) and the structure corresponding to step 21 along the Co–C bond formation PES (green).

The computed MOs for free Co^{I}Cbl (Figure 4.7, left) are in excellent agreement with those obtained in a previous DFT/PBE computational study of Co^{I}Cbl.[55] Among the occupied Co 3d-

based frontier orbitals, the Co $3d_{x2-y2}$-derived MO is essentially nonbonding and lowest in energy, at -4.03 eV below the corrin π*-based lowest unoccupied molecular orbital (LUMO). Because the composition of the corrin π*-based LUMO remains essentially unchanged upon cofactor binding to PduO and Co oxidation state changes, MO energies are reported relative to the energy of this orbital for all species investigated. Binding of Co^{I}Cbl to PduO in the presence of ATP causes a destabilization of the Co $3d_{x2-y2}$-derived MO by 1.33 eV but has no significant effect on the relative energies of the other Co 3d-based frontier orbitals. As the Co···5' C distance is decreased along the Co–C bond formation PES, the energies and compositions of most Co 3d-based orbitals remain largely unchanged, with the notable exception of the Co $3d_{x2-y2}$-based MO. For this orbital, the antibonding contribution from the 5' C of ATP increases with decreasing Co···5' C distance, raising the orbital's energy by 0.64 eV from PduO/Co^{I}Cbl/ATP to step 18. Aside from this perturbation, the MOs for all species in Figure 4.7 are characteristic of Co^{I}Cbl, supporting our conclusion that steps 1 through 19 (Figure 4.5) are located on the Co^{I}Cbl PES.

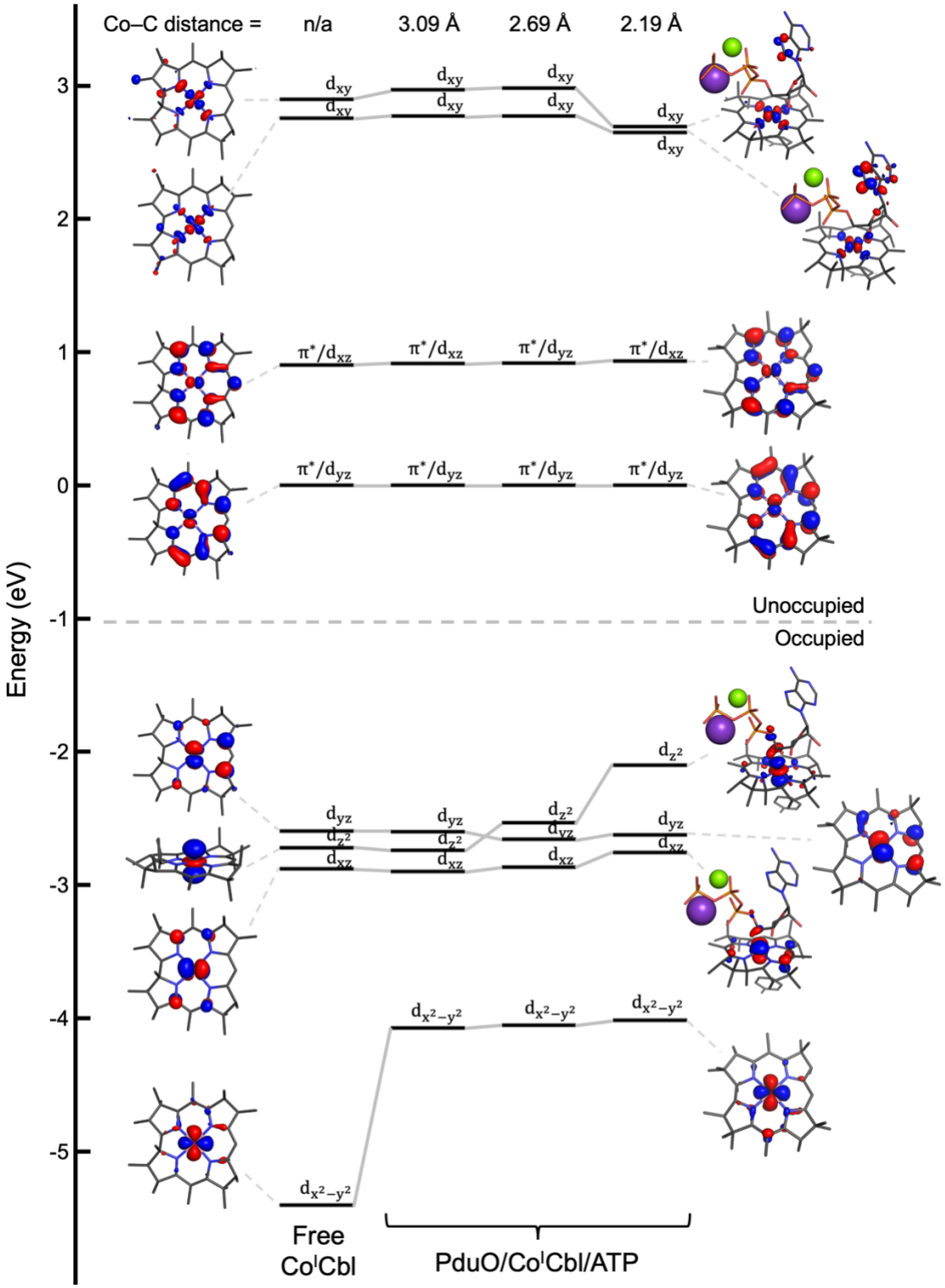
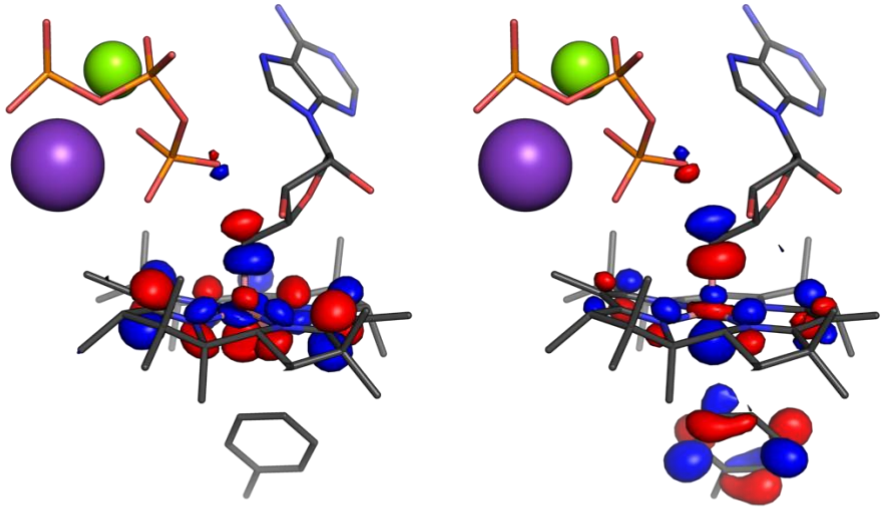


Figure 4.8. Relative energies and isosurface plots of the key MOs of Co^{I}Cbl, PduO/Co^{I}Cbl/ATP, and two structures (steps 8 and 18) along the Co–C bond formation PES. For clarity, ATP, F112, and the two cations are hidden unless the MO contains a sizable contribution from those species. Energies are relative to that of the corresponding lowest energy unoccupied corrin-based MO.

By contrast, the computed energies and compositions of the MOs for the structure corresponding to step 21 of the PES, which represents PduO/AdoCbl/PPP_i, are much more similar to those of AdoCbl (Figure S1). Alkylcobalamins like AdoCbl are known to exhibit a high degree of mixing between the Co orbitals and energetically proximate corrin-based orbitals, resulting in a large number of occupied frontier MOs with significant contributions from the Co 3d orbitals.[56] A similarly high degree of orbital mixing is observed for AdoCbl bound to the PduO/ATP active site, with nine occupied MOs located within 5 eV of the LUMO containing significant Co 3d orbital character. Some of these MOs also contain substantial contributions from the adenine (Ade) and PPP_i moieties. Most notably, two unoccupied MOs are present in this species with predominant Co $3d_{z2}$ orbital character (Figure 4.8), both of which contain a strong Co–C σ* contribution. Collectively, our results indicate that as the Co···5' C distance decreases, the occupied Co $3d_{z2}$-based MO becomes increasingly destabilized. which ultimately results in the formal transfer of two electrons to an unoccupied lower energy MO of the adenosyl cation (corresponding to step 19 → 20 in Figure 4.5). This leads to Co–C bond formation and the formal oxidation of Co^{I} to produce the Co^{III} species AdoCbl.



MO:	LUMO + 2	LUMO + 3
Relative Energy:	0.82 eV	1.21 eV
Co Character:	16.3 %	16.9 %

Figure 4.9. Isosurface plots of the two unoccupied MOs with significant Co $3d_{z2}$ character of the PduO/AdoCbl/PPP_i model corresponding to step 21 along the Co–C bond formation PES. Energies are relative to the lowest energy unoccupied corrin-based MO.

4.4. Summary and Conclusions

QM/MM optimizations of whole-protein models derived from the crystal structure of PduO, bound either with cosubstrates ATP and Co^{II}Cbl or with products AdoCbl and PPP_i,[26] produced useful snapshots of key steps in PduO's catalytic cycle. The electronic structures of these models shed new light on the mechanism through which PduO reduces Co^{II}Cbl to Co^{I}Cbl and activates Co^{I}Cbl for nucleophilic attack on ATP. PduO-bound Co^{II}Cbl contains a significantly stabilized Co $3d_{z2}$-based SOMO, due to the removal of the strong Co–N_{DMB} σ* interaction that is present in the 5C species, thereby facilitating reduction to Co^{I}Cbl. The computed models were used to calculate the difference in Co^{II}/Co^{I}Cbl redox potential in PduO's active site compared to the free cofactor, which was determined to be +0.089 V, in good agreement with published experimental data. Exploration of the Co–C bond formation step through a computed PES reveals that the decrease in Co···5' C distance for nucleophilic attack on ATP leads to a sizable

destabilization of the occupied Co $3d_{z2}$-based MO of Co^{I}Cbl and, eventually, the transfer of the two electrons from this orbital to an unoccupied lower energy MO of the adenosyl cation.

References

(1) Toraya, T. Radical Catalysis of B_{12} Enzymes: Structure, Mechanism, Inactivation, and Reactivation of Diol and Glycerol Dehydratases. *Cell. Mol. Life Sci. C.* **2000**, *57* (1), 106–127.

(2) Banerjee, R.; Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **2003**, *72*, 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828.

(3) Banerjee, R. Radical Carbon Skeleton Rearrangements: Catalysis by Coenzyme B_{12}-Dependent Mutases. *Chem. Rev.* **2003**, *103* (6), 2083–2094. https://doi.org/10.1021/cr0204395.

(4) Banerjee, R.; Chowdhury, S. Methylmalonyl-CoA Mutase. In *Chemistry and Biochemistry of B12*; Banerjee, R., Ed.; Wiley: New York, 1999; pp 707–729.

(5) Marsh, E. N. G.; Patterson, D. P.; Li, L. Adenosyl Radical: Reagent and Catalyst in Enzyme Reactions. *ChemBioChem*. **2010**, *11* (5), 604–621. https://doi.org/10.1002/cbic.200900777.

(6) Mera, P. E.; Escalante-Semerena, J. C. Multiple Roles of ATP:Cob(I)Alamin Adenosyltransferases in the Conversion of B_{12} to Coenzyme B_{12}. *Applied Microbiology and Biotechnology*. **2010**, 88, 41–48. https://doi.org/10.1007/s00253-010-2773-2.

(7) Banerjee, R. B_{12} Trafficking in Mammals: A Case for Coenzyme Escort Service. *ACS Chemical Biology*. **2006**, *1* (3), 149–159. https://doi.org/10.1021/cb6001174.

(8) Banerjee, R.; Gouda, H.; Pillay, S. Redox-Linked Coordination Chemistry Directs Vitamin B_{12} Trafficking. *Acc. Chem. Res.* **2021**, *54* (8), 2003–2013. https://doi.org/10.1021/acs.accounts.1c00083.

(9) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. Adenosyltransferase Tailors and Delivers Coenzyme B_{12}. *Nat. Chem. Biol.* **2008**, *4* (3), 194–196. https://doi.org/10.1038/nchembio.67.

(10) Gouda, H.; Mascarenhas, R.; Pillay, S.; Ruetz, M.; Koutmos, M.; Banerjee, R. Patient Mutations in Human ATP:Cob(I)Alamin Adenosyltransferase Differentially Affect Its Catalytic versus Chaperone Functions. *J. Biol. Chem.* **2021**, *297* (6). https://doi.org/10.1016/j.jbc.2021.101373.

(11) Dobson, C. M.; Wai, T.; Leclerc, D.; Kadir, H.; Narang, M.; Lerner-Ellis, J. P.; Hudson, T. J.; Rosenblatt, D. S.; Gravel, R. A. Identification of the Gene Responsible for the CbIB Complementation Group of Vitamin B_{12}-Dependent Methylmalonic Aciduria. *Hum. Mol. Genet.* **2002**, *11* (26), 3361–3369. https://doi.org/10.1093/hmg/11.26.3361.

(12) Leal, N. A.; Park, S. D.; Kima, P. E.; Bobik, T. A. Identification of the Human and Bovine ATP:Cob(I)Alamin Adenosyltransferase CDNAs Based on Complementation of a Bacterial Mutant. *J. Biol. Chem.* **2003**, *278* (11), 9227–9234. https://doi.org/10.1074/jbc.M212739200.

(13) Escalante-Semerena, J. C.; Suh, S. J.; Roth, J. R. CobA Function Is Required for Both de Novo Cobalamin Biosynthesis and Assimilation of Exogenous Corrinoids in *Salmonella typhimurium*. *J. Bacteriol.* **1990**, *172* (1), 273–280. https://doi.org/10.1128/jb.172.1.273-280.1990.

(14) Moore, T. C.; Newmister, S. A.; Rayment, I.; Escalante-Semerena, J. C. Structural Insights into the Mechanism of Four-Coordinate Cob(II)Alamin Formation in the Active Site of the *Salmonella Enterica* ATP:Co(I)rrinoid Adenosyltransferase Enzyme: Critical Role of Residues Phe91 and Trp93. *Biochemistry* **2012**, *51* (48), 9647–9657. https://doi.org/10.1021/bi301378d.

(15) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Resonance Raman Spectroscopic Study of the Interaction between Co(II)Rrinoids and the ATP:Corrinoid Adenosyltransferase PduO from *Lactobacillus Reuteri*. *J. Biol. Inorg. Chem.* **2016**, *21* (5–6), 669–681. https://doi.org/10.1007/s00775-016-1371-x.

(16) Costa, F. G.; Greenhalgh, E. D.; Brunold, T. C.; Escalante-Semerena, J. C. Mutational and Functional Analyses of Substrate Binding and Catalysis of the *Listeria Monocytogenes* EutT ATP:Co(I)rrinoid Adenosyltransferase. *ACS Appl. Mater. Interfaces* **2020**, *59* (10), 1124–1136. https://doi.org/10.1021/acs.biochem.0c00078.

(17) Corwin, A. H.; Chiwis, A. B.; Poor, R. W.; Whitten, D. G.; Baker, E. W.; Schrauzer, G. N.; Deutsch, E.; Smith, E. L.; Mervyn, L.; Muggleton, P. W.; et al. Reactions of Cobalt(I) Supemucleophiles. The Alkylation of Vitamin B_{12S}, Cobaloximes(I), and Related Compounds. *J. Am. Chem. Soc.* **1959**, *91* (12), 3341–3350.

(18) Lexa, D.; Saveant, J. M. The Electrochemistry of Vitamin B_{12}. *Acc. Chem. Res.* **1983**, *16*, 235–243. https://doi.org/10.1021/ar00091a001.

(19) Hoover, D. M.; Jarrett, J. T.; Sands, R. H.; Dunham, W. R.; Ludwig, M. L.; Matthews, R. G. Interaction of *Escherichia Coli* Cobalamin-Dependent Methionine Synthase and Its Physiological Partner Flavodoxin: Binding of Flavodoxin Leads to Axial Ligand Dissociation from the Cobalamin Cofactor. *Biochemistry* **1997**, *36* (1), 127–138. https://doi.org/10.1021/bi961693s.

(20) Mera, P. E.; Escalante-Semerena, J. C. Dihydroflavin-Driven Adenosylation of 4-Coordinate Co(II) Corrinoids: Are Cobalamin Reductases Enzymes or Electron Transfer Proteins? *J. Biol. Chem.* **2010**, *285*, 2911–2917. https://doi.org/10.1074/jbc.M109.059485.

(21) Gouda, H.; Li, Z.; Ruetz, M.; Banerjee, R. Coordination Chemistry Controls Coenzyme B_{12} Synthesis by Human Adenosine Triphosphate:Cob(I)Alamin Adenosyltransferase. *Inorg. Chem.* **2023**, *62* (32), 12630–12633. https://doi.org/10.1021/acs.inorgchem.3c02163.

(22) Stich, T. A.; Buan, N. R.; Escalante-Semerena, J. C.; Brunold, T. C. Spectroscopic and Computational Studies of the ATP:Corrinoid Adenosyltransferase (CobA) from *Salmonella enterica*: Insights into the Mechanism of Adenosylcobalamin Biosynthesis. *J. Am. Chem. Soc.* **2005**, *127* (24), 8710–8719. https://doi.org/10.1021/ja042142p.

(23) 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.

(24) Liptak, M. D.; Fleischhacker, A. S.; Matthews, R. G.; Telser, J.; Brunold, T. C. Spectroscopic and Computational Characterization of the Base-off Forms of Cob(II)Alamin. *J. Phys. Chem. B* **2009**, *113* (15), 5245–5254. https://doi.org/10.1021/jp810136d.

(25) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Kinetic and

Spectroscopic Studies of the ATP:Corrinoid Adenosyltransferase PduO from *Lactobacillus reuteri*: Substrate Specificity and Insights into the Mechanism of Co(II)Corrinoid Reduction. *Biochemistry* **2008**, *47* (34), 9007–9015. https://doi.org/10.1021/bi800419e.

(26) 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.

(27) Mera, P. E.; Maurice, M. S.; Rayment, I.; Escalante-Semerena, J. C. Residue Phe112 of the Human-Type Corrinoid Adenosyltransferase (PduO) Enzyme of *Lactobacillus reuteri* is Critical to the Formation of the Four-Coordinate Co(II) Corrinoid Substrate and to the Activity of the Enzyme. *Biochemistry* **2009**, *48* (14), 3138–3145. https://doi.org/10.1021/bi9000134.

(28) Park, K.; Mera, P. E.; Moore, T. C.; Escalante-Semerena, J. C.; Brunold, T. C. Unprecedented Mechanism Employed by the *Salmonella enterica* EutT ATP:Co^{I}rrinoid Adenosyltransferase Precludes Adenosylation of Incomplete Co^{II}rrinoids. *Angew. Chem., Int. Ed.* **2015**, *54* (24), 7158–7161. https://doi.org/10.1002/anie.201501930.

(29) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Spectroscopic Characterization of Active-Site Variants of the PduO-Type ATP:Corrinoid Adenosyltransferase from *Lactobacillus reuteri*: Insights into the Mechanism of Four-Coordinate Co(II)Corrinoid Formation. *Inorg. Chem.* **2012**, *51* (8), 4482–4494. https://doi.org/10.1021/ic202096x.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. E.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision C.01. *Gaussian, Inc., Wallingford CT,*. 2016.

(31) Kräutler, B.; Keller, W.; Kratky, C. Coenzyme B_{12} Chemistry: The Crystal and Molecular Structure of Cob(II)Alamin. *J. Am. Chem. Soc.* **1989**, *111* (24), 8936–8938. https://doi.org/10.1021/ja00206a037.

(32) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32*, W665–W667.

(33) PyMOL Molecular Graphics System, Ver. 2.5; Schrödinger, LLC.

(34) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(35) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(36) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.* **1994**, *100* (8), 5829–5835. https://doi.org/10.1063/1.467146.

(37) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54* (2), 720–723. https://doi.org/10.1063/1.1674902.

(38) Hehre, W. J.; Ditchfield, K.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII.

Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261. https://doi.org/10.1063/1.1677527.

(39) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28* (3), 213–222. https://doi.org/10.1007/BF00533485.

(40) Hariharan, P. C.; Pople, J. A. Accuracy of AHn Equilibrium Geometries by Single Determinant Molecular Orbital Theory. *Mol. Phys.* **1974**, *27* (1), 209–214. https://doi.org/10.1080/00268977400100171.

(41) Gordon, M. S. The Isomers of Silacyclopropane. *Chem. Phys. Lett.* **1980**, *76* (1), 163–168. https://doi.org/10.1016/0009-2614(80)80628-2.

(42) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77* (7), 3654–3665. https://doi.org/10.1063/1.444267.

(43) Binning, R. C.; Curtiss, L. A. Compact Contracted Basis Sets for Third-Row Atoms: Ga–Kr. *J. Comput. Chem.* **1990**, *11* (10), 1206–1216. https://doi.org/10.1002/JCC.540111013.

(44) Blaudeau, J. P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. Extension of Gaussian-2 (G2) Theory to Molecules Containing Third-Row Atoms K and Ca. *J. Chem. Phys.* **1997**, *107* (13), 5016–5021. https://doi.org/10.1063/1.474865.

(45) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K through Zn. *J. Chem. Phys.* **1998**, *109* (4), 1223–1229. https://doi.org/10.1063/1.476673.

(46) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* Basis Set for Third-Row Atoms. *J. Comput. Chem.* **2001**, *22* (9), 976–984. https://doi.org/10.1002/jcc.1058.

(47) Marques, H. M.; Ngoma, B.; Egan, T. J.; Brown, K. L. Parameters for the AMBER Force Field for the Molecular Mechanics Modeling of the Cobalt Corrinoids. *J. Mol. Struct.* **2001**, *561* (1–3), 71–91. https://doi.org/10.1016/S0022-2860(00)00920-0.

(48) Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.* **2012**, *8* (11), 4405–4412. https://doi.org/10.1021/ct300613v.

(49) Pavlova, A.; Parks, J. M.; Gumbart, J. C. Development of CHARMM-Compatible Force-Field Parameters for Cobalamin and Related Cofactors from Quantum Mechanical Calculations. *J. Chem. Theory Comput.* **2018**, *14* (2), 784–798. https://doi.org/10.1021/acs.jctc.7b01236.

(50) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73–78. https://doi.org/10.1002/wcms.81.

(51) Han, W. G.; Lovell, T.; Noodleman, L. Coupled Redox Potentials in Manganese and Iron Superoxide Dismutases from Reaction Kinetics and Density Functional/Electrostatics Calculations. *Inorg. Chem.* **2002**, *41* (2), 205–218. https://doi.org/10.1021/ic010355z.

(52) Mera, P. E.; St. Maurice, M.; Rayment, I.; Escalante-Semerena, J. C. Structural and Functional Analyses of the Human-Type Corrinoid Adenosyltransferase (PduO) from

Lactobacillus reuteri. *Biochemistry* **2007**, *46* (48), 13829–13836. https://doi.org/10.1021/bi701622j.

(53) 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. Sacrificial Cobalt–Carbon Bond Homolysis in Coenzyme B_{12} as a Cofactor Conservation Strategy. *J. Am. Chem. Soc.* **2018**, *140* (41), 13205–13208.

(54) St. Maurice, M.; Mera, P. E.; Taranto, M. P.; Sesma, F.; Escalante-Semerena, J. C.; Rayment, I. Structural Characterization of the Active Site of the PduO-Type ATP:Co(I)Rrinoid Adenosyltransferase from *Lactobacillus reuteri*. *J. Biol. Chem.* **2007**, *282* (4), 2596–2605. https://doi.org/10.1074/jbc.M609557200.

(55) Liptak, M. D.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{1+}Cobalamin: Spectral and Electronic Properties of the "Superreduced" B_{12} Cofactor. *J. Am. Chem. Soc.* **2006**, *128* (28), 9144–9156. https://doi.org/10.1021/ja061433q.

(56) Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914. https://doi.org/10.1021/ja029328d.

(57) Grove, L. E.; Xie, J.; Yikilmaz, E.; Miller, A. F.; Brunold, T. C. Spectroscopic and Computational Investigation of Second-Sphere Contributions to Redox Tuning in *Escherichia coli* Iron Superoxide Dismutase. *Inorg. Chem.* **2008**, *47* (10), 3978–3992. https://doi.org/10.1021/ic702412y.

(58) Stich, T. A.; Buan, N. R.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{2+}corrinoids: Spectral and Electronic Properties of the Biologically Relevant Base-on and Base-off Forms of Co^{2+}cobalamin. *J. Am. Chem. Soc.* **2004**, *126* (31), 9735–9749. https://doi.org/10.1021/ja0481631.

Chapter 5

Computational Studies of a PduO-type Adenosyltransferase Containing a Rhodium-Substituted B_{12} Analogue

Chapter 5: Computational Studies of a PduO-type Adenosyltransferase Containing a Rhodium-Substituted B_{12} Analogue

5.1. Introduction

The cobalamin (Cbl or B_{12}) family of cofactors is distinctive for both its structural complexity and its vital role in the catalysis of reactions in all kingdoms of life. The diverse collection of reactions catalyzed by B_{12}-dependent enzymes includes reductive dehalogenation, isomerization, elimination, and methyl-transfer reactions.[1] Cbls consist of a cobalt (Co) ion equatorially chelated by a tetrapyrrole macrocycle termed the corrin ring, which features a pendant nucleotide loop terminating in a dimethylbenzimidazole (DMB) base that can bind to the Co ion in the lower axial position. In its Co^{III} form, B_{12} is typically six-coordinate (6C), with varying axial ligands. The lower position is commonly occupied by the DMB (the "base-on" conformation), a water molecule ("base-off"), or a protein-derived histidine ("base-off/His-on"), while the upper axial position can be occupied by a range of ligands such as methyl (Me) and adenosyl (Ado) groups. The corrin ring and assorted axial ligands impart a unique, versatile reactivity that makes B_{12} an essential part of the biological world.

The precise role of the cobalt ion in this distinctive architecture is not well-explored, as it cannot be removed from B_{12} without complete destruction of the corrin ring.[2] This is a departure from macrocycles like porphyrin ring, another tetrapyrrole, which can be reconstituted with non-native metals with relative ease; the first Cu, Ni, Co, Zn, and metal-free porphyrins were reported almost a century ago,[3] and a great deal of research since then has used metal-substituted heme analogues for catalysis and the study of heme-containing enzymes.[4] The first metal-substituted B_{12} analogue, meanwhile, was not reported until 1974, when Koppenhagen et al. described the syntheses of methylrhodibalamin (MeRhbl) and adenosylrhodibalamin (AdoRhbl).[5] These were

obtained through an arduous process involving the isolation of metal-free hydrogenocobalamin (Hbl) from *Chromatium thermoaceticum* grown in Co-free media, followed by a series of synthetic chemistry steps.

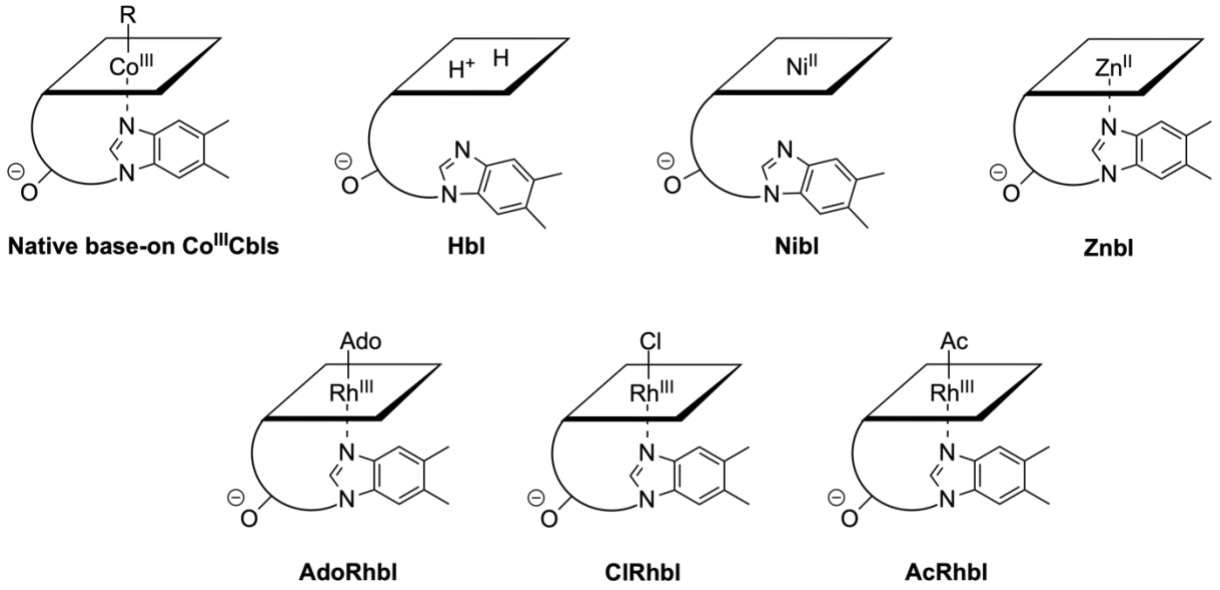


Figure 5.1. Abbreviated structure of native base-on Co^{III}Cbls, metal-free hydrogenobyric acid (Hby), and several metal-substituted Cbl analogues: nibalamin (Nibl), zincobalamin (Znbl), adenosylrhodibalamin (AdoRhbl), chlororhodibalamin (ClRhbl), and acetylrhodibalamin (AcRhbl).

In recent years, much progress has been made towards assembling a wider variety of transition metal B_{12} analogues. A higher yield synthetic pathway to adenosylrhodibalamin was developed, again using a combination of chemical and biological steps, which allowed for a more thorough characterization of AdoRhbl's structural and biochemical properties.[6] Soon after, the development of an elegant, bioengineered synthesis of hydrogenobyric acid (Hby),[2] a Hbl precursor that lacks the nucleotide loop, paved the way for the syntheses of nibalamin (Nibl),[7] zincobalamin (Znbl),[8] chlororhodibalamin (ClRhbl),[9] and acetylrhodibalamin (AcRhbl).[10] ClRhbl in particular shows great potential as a synthetic platform for the preparation of other rhodibalamins, so these may be the first of many new B_{12} derivatives.

The interest in B_{12} derivatives arises, in large part, from their potential use as so-called "antivitamins." The concept of antivitamins relies on the fact that metal-substituted B_{12} molecules and other inert B_{12} analogues can be taken up through normal B_{12} metabolic pathways, which allows them to bind and inhibit B_{12}-dependent enzymes.[11] Antivitamins have applications both as therapeutics, targeting cancers and bacterial infections with an elevated demand for B_{12}, and as research tools that can induce a B_{12} deficiency in otherwise healthy organisms.[12,13] Thus far, antivitamin activity has been demonstrated for several metal-substituted B_{12} analogues. AdoRhbl behaves as an inhibitory analogue of adenosylcobalamin (AdoCbl) in assays with the B_{12}-dependent dependent methionine synthase (MetH) and diol dehydratase (DD).[6] Similarly, Nibl was found to bind to a bacterial adenosyltransferase (ATR) and inhibit its activity.[7] Being an inactive structural mimic of Co^{II}Cbl, Znbl has also been suggested as a viable antivitamin candidate.[8]

In addition to their potential use as antivitamins, metal-substituted B_{12} analogues provide the opportunity to learn more about the structure and function of B_{12}-dependent enzymes by serving as structural mimics of short-live Cbl intermediates. Nibl, for example, is isoelectronic and roughly isostructural to Co^{I}Cbl, the highly reactive "supernucleophile" intermediate that forms during the catalytic cycles of ATRs and methylcobalamin (MeCbl)-dependent enzymes.[14] Similarly, AdoRhbl is isoelectronic to AdoCbl, and its ability to bind to the active site of B_{12}-dependent enzymes without turnover makes it ideal for characterization through X-ray crystallography. This is the purpose that AdoRhbl serves in the work described here, which concerns a crystal structure of an ATR from *Mycobacterium tuberculosis (Mtb)*, the causative pathogen of tuberculosis,[15] that contains AdoRhbl along with cosubstrate ATP.

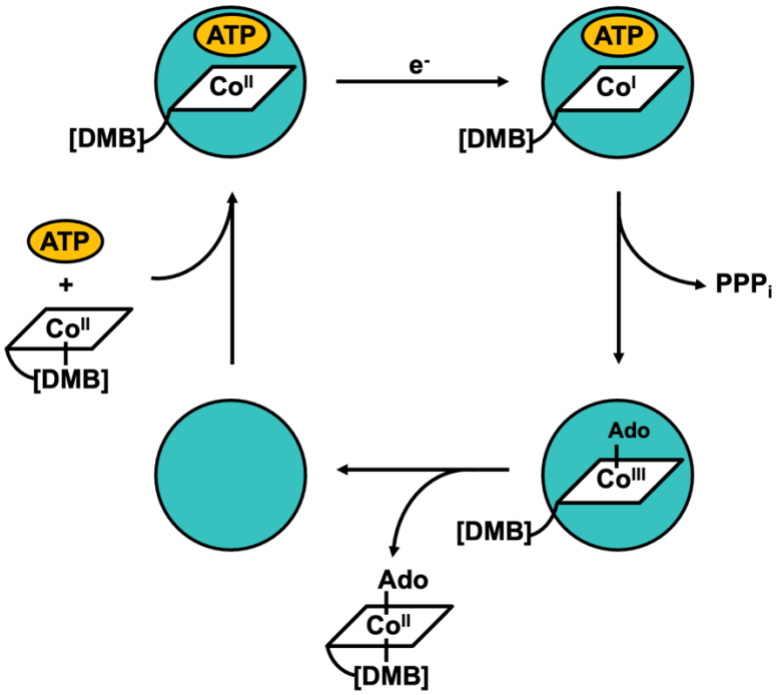


Figure 5.2. Mechanism for the adenosylation of Co^{II}Cbl by ATRs. Co^{II}Cbl is bound in a rare four-coordinate, base-off conformation by the ATR (teal circle), along with cosubstrate ATP (yellow oval).

As discussed in Chapter 4, ATRs are crucial to the activity of B_{12}-dependent enzymes, catalyzing the adenosylation of exogenous cobalamins and incomplete corrinoids.[16–18] Like the human ATR (hATR), *Mtb* ATR is a member of the PduO family of ATRs, and it consists of a homotrimer that binds its Cbl and ATP substrates at subunit interfaces.[15] The general catalytic cycle of ATRs involves (i) binding of Co^{II}Cbl in a four-coordinate (4C) conformation to the ATR in the presence of cosubstrate ATP, (ii) reduction of Co^{II}Cbl to Co^{I}Cbl, (iii) nucleophilic attack of Co^{I}Cbl on ATP to produce AdoCbl and triphosphate (PPP_i), and (iv) release or transfer of the reaction products (Figure 5.2). Banerjee, Kräutler, and coworkers obtained an X-ray crystal structure *Mtb* ATR containing AdoRhbl and PPP_i, essentially representing the ATR's product-bound state.

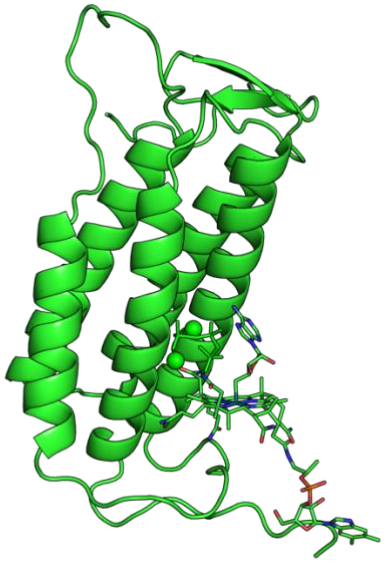


Figure 5.3. X-ray crystal structure of AdoRhbl-bound *Mtb* ATR provided for computational modeling, consisting of one monomer of the homotrimer structure (PDB: 8D32).

In the work described here, computational models of AdoRhbl-bound *Mtb* ATR were generated through quantum mechanics/molecular mechanics (QM/MM) geometry optimizations. Of particular interest was the elongated Rh–C bond observed in the crystal structure, which at 2.23 Å is considerably longer than the 2.07 Å Rh–C bond observed in the previously published crystal structure of free AdoRhbl.[6] Additional QM/MM computations were performed for free AdoRhbl to identify the best functional for QM calculations.

5.2. Materials and Methods

A series of QM geometry optimizations of free AdoRhbl was performed using the atomic coordinates from a 1.85 Å resolution X-ray crystal structure of AdoRhbl as the starting point.[6] Optimizations were performed in Gaussian 16[19] using a series of different functionals; namely, BP86,[20,21] B3LYP,[22,23] and PBE0.[24,25] For each functional, optimizations were performed both in vacuum and with implicit solvation by utilizing the polarizable continuum model (PCM). The LanL2DZ basis set[26–28] was used for the Rh atom, while the 6-31G* basis set[29–38] was employed for all other atoms. Optimized structures of AdoRhbl were visualized in PyMOL[39] and assessed

on the basis of several structural parameters (Figure 5.4). The corrin fold angle, defined as the angle between the planes of best fit for two sets of atoms, (C4, C5, C6, C9, C10, N21, N22) and (C10, C11, C14, C15, C16, N23, N24),[40] was calculated using MATLAB (version R2018a; Mathworks, Natick, MA). In brief, the covariance matrix of each set of atoms with respect to their centroid was calculated, and the cross product of the first two eigenvectors of the matrix was used to calculate the normal vector of that plane. The fold angle was then calculated between the two normal vectors.

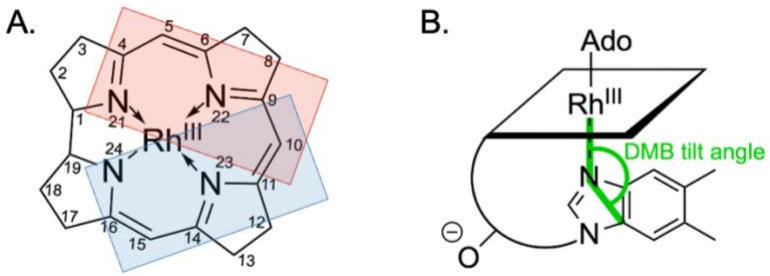


Figure 5.4. Visual representation of two structural parameters used here to assess the QM/MM optimized AdoRhbl structures. (A) The two planes of best fit used to calculate the corrin fold angle. (B) Definition of the DMB tilt angle.

The crystal structure of one monomer of AdoRhbl-containing *Mtb* ATR was provided by Banerjee, Kräutler, and coworkers. The full biological assembly for this structure has since been deposited in the protein data bank (PDB 8D32). This structure was optimized using the quantum mechanics/molecular mechanics (QM/MM) approach as implemented through the ONIOM method in Gaussian. The QM region was defined as the Rh atom, corrin ring, the first carbon of each side chain, the Ado moiety, the polyphosphate (PPP_i), and two Mg^{2+} ions. The results obtained in these initial QM/MM computations were used to guide subsequent calculations using different QM regions and freezing the positions of select atoms. The QM region was treated with the PBE0 functional, using the LanL2DZ basis set for Rh and the 6-31G* basis set for all other QM atoms. The remainder of the system was treated using MM, with the AMBER forcefield[41] and B_{12}-specific

AMBER-compatible parameters.[42] Optimizations were performed both with and without freezing the MM and Mg atoms. PyMOL was used to calculate the root-mean-square deviation (RMSD) of atoms in the enzyme's backbone between the crystallographic starting coordinates and the optimized structures.

5.3. Results and Discussion

The first step of this project was to determine an appropriate DFT functional for the structural modeling of AdoRhbl. Recent DFT studies of Rh-containing complexes suggested B3LYP,[43,44] BP86,[45] and PBE0[46,47] as suitable candidates, so these functionals were used to perform QM optimizations of the structure of free AdoRhbl. In previous work, I have found that performing geometry optimizations of free B_{12} in the absence of explicit solvent molecules or an implicit solvation model can sometimes impede convergence or have a large impact on the resulting structure, favoring intramolecular hydrogen bonds between side chains and the adenosyl moiety. The three functionals were therefore tested both with and without an implicit water solvation model. It should be noted that, in the time since this work was performed, DFT calculations with AcRhbl have been reported that used the dispersion-corrected BP86 and the def2-SVP basis set with an implicit water model.[10]

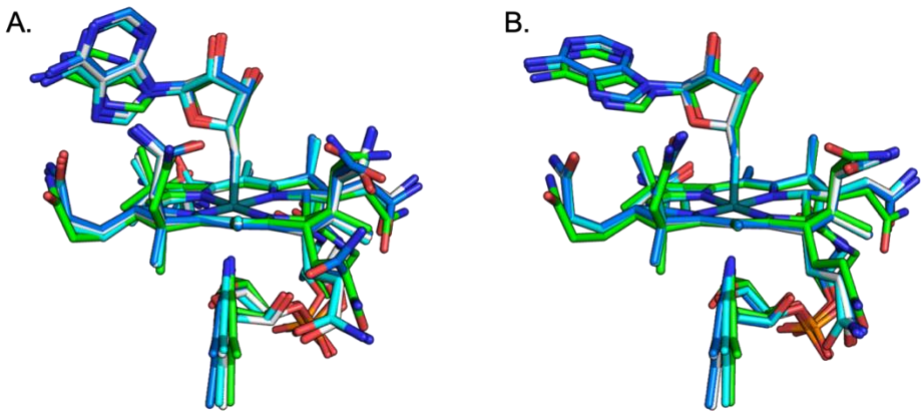


Figure 5.5. Models of AdoRhbl in the crystal structure (green) and QM-optimized models obtained using B3LYP (white), BP86 (blue), and PBE0 (cyan), including (A) optimizations performed in vacuo and (B) optimizations performed with the PCM solvation model using water as the solvent.

Table 5.1. Key AdoRhbl structural parameters derived from the published crystal structure and six QM-optimized computational models. Distances are in Å and angles are in degrees

Structure	**Rh–C**	**Rh–N_{DMB}**	**Rh–C–C**	**DMB tilt**	**Corrin fold angle**	**Helicity[a]**
Crystal structure	2.07	2.27	117.7	167.8	5.97	-3.1
B3LYP	2.09	2.41	120.6	167.8	9.70	-5.6
B3LYP + PCM	2.09	2.46	120.5	167.8	8.77	-5.2
BP86	2.10	2.36	120.5	168.3	8.83	-5.8
BP86 + PCM	2.10	2.40	120.5	167.2	9.78	-5.4
PBE0	2.06	2.34	120.0	167.3	15.0	-5.5
PBE0 + PCM	2.06	2.36	120.0	166.2	8.75	-5.1

a. Defined as the N21–N22–N23–N24 dihedral angle.

All QM-optimized structures display a very slight elongation of the Rh–N_{DMB} bond and an increase in the corrin fold angle compared to the crystal structure of AdoRhbl, changes that can likely be attributed to crystal packing forces.[48] In the absence of the implicit solvation model, all the three functionals predict an altered conformation of the adenosyl moiety in which the amine group of the adenine moiety forms a hydrogen bond with one of corrin side chains. However, the absence or presence of the solvation model has no effect on the Rh–C bond length, which is the primary focus of the present study. We chose to move forward with the PBE0 functional because it most closely reproduces the length of this bond observed in the crystal structure of AdoRhbl.

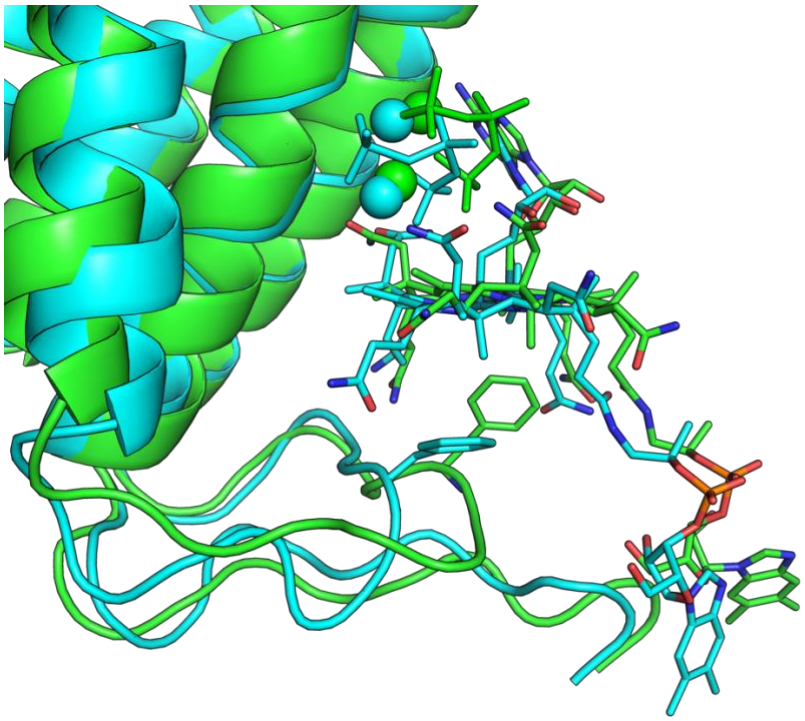


Figure 5.6. Comparison of the AdoRhbl-bound *Mtb* ATR crystal structure (green) and QM/MM-optimized model (cyan).

Initial QM/MM computations revealed that a traditional geometry optimization of a single AdoRhbl-bound ATR monomer would not effectively maintain the structure of the active site (Figure 5.6). The RMSD of atoms in the protein backbone between the optimized structure and starting coordinates was found to be 3.06 Å, well above the resolution of the crystal structure. Even more worrying, during the QM/MM optimization, residue F117 underwent a dramatic shift away from its original position at the lower face of the corrin ring. This residue plays a well-established role in enabling the 4C Cbl conformation by preventing water from binding in the lower axial position, which is crucial to ATR's catalytic cycle by raising the Co^{II}/Co^{I}Cbl reduction into the physiologically accessible range.[15] These drastic changes are perhaps unsurprising when considering that the ATP and Cbl substrates are bound at the interface of two monomers. Previously published X-ray crystal structures of *Mtb* ATR show that the Lys28, Tyr36, and Gly19 from another subunit interact with the adenine ring to position it correctly.[15] Without these and other residues

from the second subunit to clamp ATP and the corrin ring in place, a great deal of rearrangement is allowed.

In order to address this issue and explore the feasibility of an elongated Rh–C bond, two additional optimizations were performed with sections of the AdoRhbl-bound *Mtb* ATR model kept frozen. In the first, the enzyme and crystallographic water molecules were frozen, while AdoRhbl, PPP_i, and the Mg^{2+} ions were permitted to move freely. However, optimization of this model led to a substantial movement of all unconstrained atoms, failing to even maintain the close interactions with residues D167 and N163 that appear to hold the Mg^{2+} ions in place in the crystal structure (Figure 5.7).

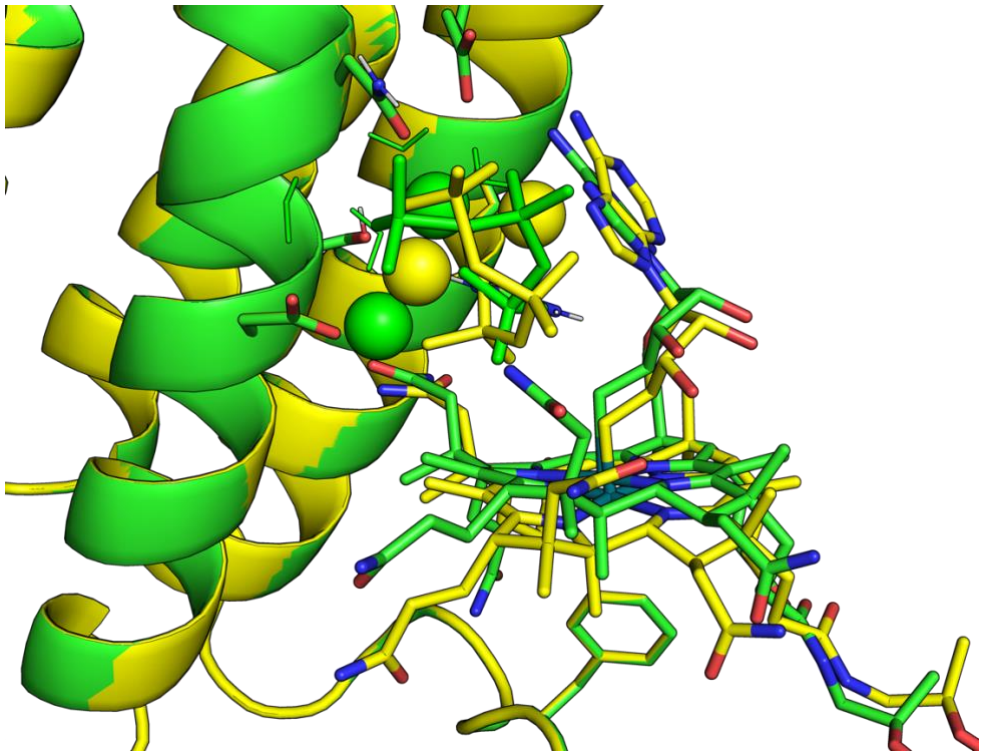


Figure 5.7. Comparison of the AdoRhbl-bound *Mtb* ATR crystal structure (green) and QM/MM-optimized model in which enzyme atoms and crystallographic water molecules were kept frozen (yellow).

In the final optimization, the Mg^{2+} ions were frozen along with the enzyme atoms and crystallographic water molecules. This required moving the Mg^{2+} ions to the MM region, because

the manner in which QM/MM calculations are implemented in Gaussian precludes freezing of only one or two atoms in the QM region. The resulting optimization did not run smoothly, requiring almost 800 steps for all convergence criteria to be met (Figure 5.8). Examining the convergence parameters revealed unusually high internal forces for an extended number of steps, which at ~375 steps led to fluctuations in maximum displacements. Visualizing the structure throughout and at the end of the optimization revealed cleavage of both the Rh–C bond and a P–O bond in the polyphosphate group.

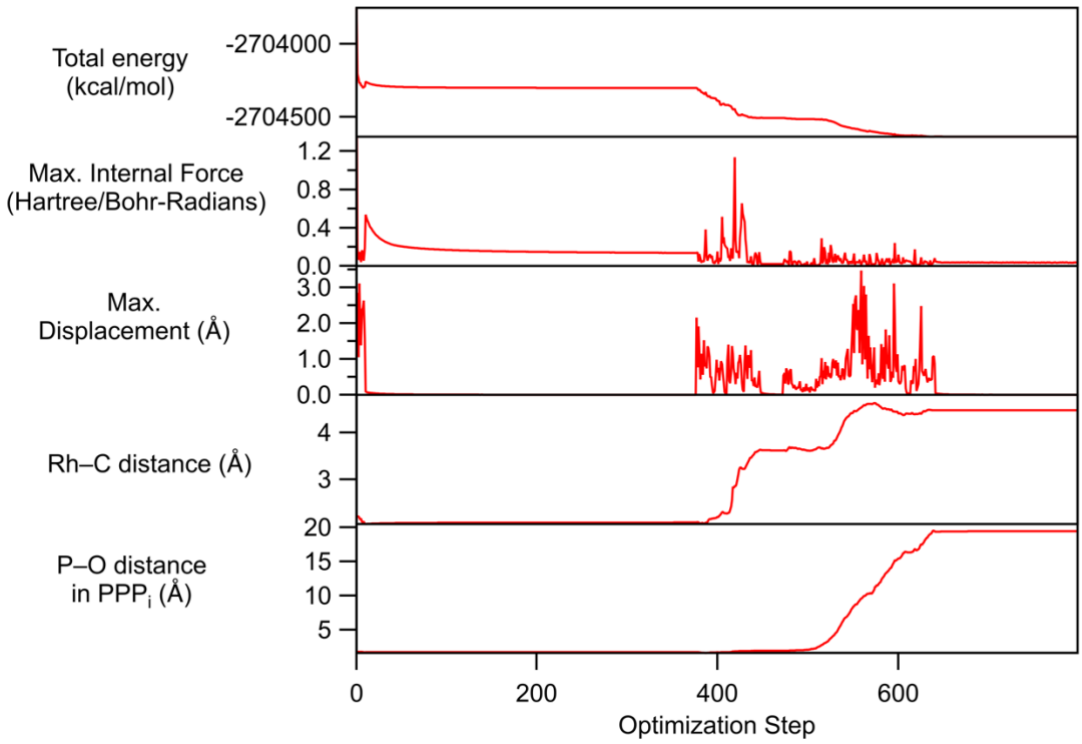


Figure 5.8. Convergence and structural parameters from the optimization of AdoRhbl-bound *Mtb* ATR that was performed with all enzyme atoms, water molecules, and Mg^{2+} ions kept frozen.

The unexpected cleavage of a strong P–O bond suggests that the optimized model does not represent a biologically relevant state. However, it is interesting nonetheless that the Rh–C bond cleaved entirely, forming a bond instead between that carbon and the nearby triphosphate oxygen, rather than maintain the elongated Rh–C bond. Furthermore, examining the earlier steps prior to Rh–C bond cleavage shows that the Rh–C bond shortens to ~2.06 Å within the first ten steps of

the optimization and maintains a length of 2.06 to 2.08 Å for several hundred additional steps. An example of one such intermediate step is shown along with the final optimized structure in Figure 5.9.

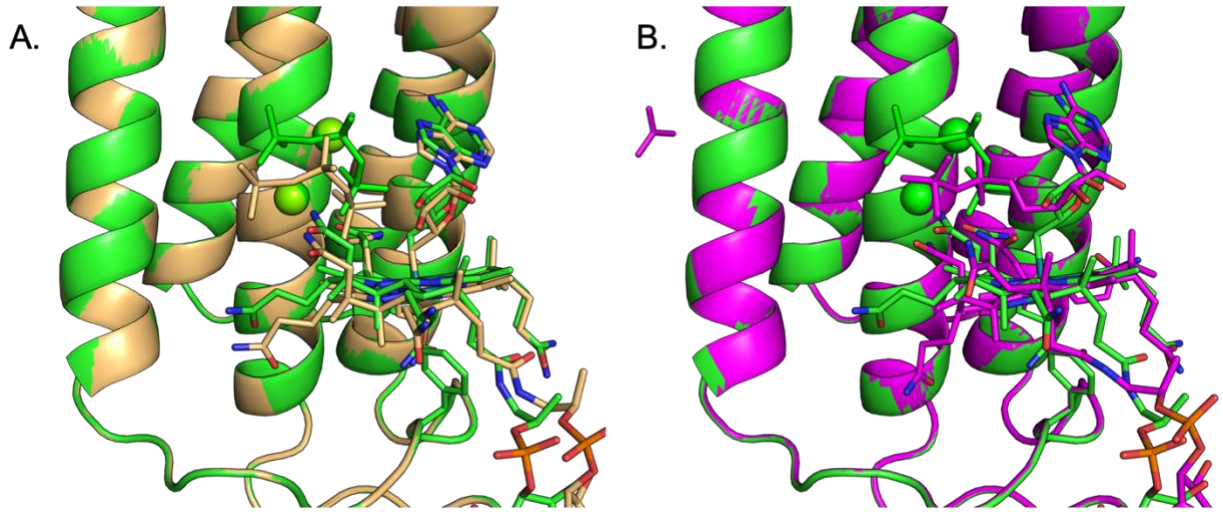


Figure 5.9. Comparison of the AdoRhbl-bound *Mtb* ATR crystal structure (green) with (A) step 120 in the QM/MM-optimization (beige) and (B) the final optimized model (pink). During the QM/MM optimization, all enzyme atoms, water molecules, and Mg^{2+} ions were kept frozen.

Key structural parameters for all three QM/MM-optimized models are compared to the crystal structure in Table 5.2. Most notably, none of the computational models retained the lengthened Rh–C bond, with optimizations producing either a significantly shorter bond or complete Rh–C bond cleavage. The differences in structure depending on whether the Mg^{2+}ions were frozen throughout the optimization emphasize the role that these ions play in positioning the PPP_i group.

Table 5.2. Key AdoRhbl structural parameters in the crystal structure of *Mtb* ATR and the three QM/MM-optimized computational models. Distances are in Å and angles are in degrees.

Structure	**Rh–C**	**Rh–C–C**	**Corrin fold angle**	**Helicity[a]**
Crystal structure	2.23	114.9	8.02	0.7
Optimized with no frozen atoms	2.06	119.0	8.32	-4.2
Optimized with frozen enzyme and water molecules	2.06	115.7	9.88	-2.1
Optimized with frozen enzyme, water molecules, and Mg^{2+}ions	4.47	78.7	6.59	-4.7

a. Defined as the N21–N22–N23–N24 dihedral angle.

5.4. Conclusion and Future Directions

Ultimately, none of these optimizations confirm that a drastically lengthened M–C bond is an accurate representation of an intermediate in the *Mtb* ATR catalytic cycle. It is possible that the 2.23 Å Rh–C bond length is an artifact of the X-ray diffraction data collection or the effect of crystal packing forces. However, the deposition of the full structure of AdoRhbl-bound *Mtb* ATR provides a valuable opportunity for future computational work in this area. A natural next step would be to perform QM/MM optimizations of the entire trimer, with parts of one active site assigned to the QM region and two other active sites placed entirely in the MM region. Additionally, electronic absorption spectroscopy could be used to probe if an elongated Rh–C bond in AdoRhbl-bound *Mtb* ATR also exists in solution, as in this case the dominant absorption features in the visible region should be significantly blue-shifted based on previous studies of Co^{III}Cbl species.

References

(1) Banerjee, R.; Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **2003**, *72*, 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828.

(2) Kieninger, C.; Deery, E.; Lawrence, A. D.; Podewitz, M.; Wurst, K.; Nemoto-Smith, E.; Widner, F. J.; Baker, J. A.; Jockusch, S.; Kreutz, C. R.; et al. The Hydrogenobyric Acid Structure Reveals the Corrin Ligand as an Entatic State Module Empowering B_{12} Cofactors for Catalysis. *Angew. Chemie - Int. Ed.* **2019**, *58* (31), 10756–10760. https://doi.org/10.1002/anie.201904713.

(3) Hill, R.; Holden, H. F. The Preparation and Some Properties of the Globin of Oxyhaemoglobin. *Biochem. J.* **1926**, *20* (6), 1326–1339. https://doi.org/10.1042/bj0201326.

(4) Scholler, D. M.; Wang, M. Y. R.; Hoffman, B. M. Metal-Substituted Hemoglobin and Other Hemoproteins. *Methods Enzymol.* **1978**, *52* (C), 487–493. https://doi.org/10.1016/S0076-6879(78)52053-3.

(5) Koppenhagen, V. B.; Elsenhans, B.; Wagner, F.; Pfiffner, J. J. Methylrhodibalamin and 5' Deoxyadenosylrhodibalamin, the Rhodium Analogues of Methylcobalamin and Cobalamin Coenzyme. *J. Biol. Chem.* **1974**, *249* (20), 6532–6540. https://doi.org/10.1016/s0021-9258(19)42189-3.

(6) Widner, F. J.; Lawrence, A. D.; Deery, E.; Heldt, D.; Frank, S.; Gruber, K.; Wurst, K.; Warren, M. J.; Kräutler, B. Total Synthesis, Structure, and Biological Activity of Adenosylrhodibalamin, the Non-Natural Rhodium Homologue of Coenzyme B_{12}. *Angew. Chemie - Int. Ed.* **2016**, *55* (37), 11281–11286. https://doi.org/10.1002/anie.201603738.

(7) Kieninger, C.; Wurst, K.; Podewitz, M.; Stanley, M.; Deery, E.; Lawrence, A. D.; Liedl, K. R.; Warren, M. J.; Kräutler, B. Replacement of the Cobalt Center of Vitamin B_{12} by Nickel: Nibalamin and Nibyric Acid Prepared from Metal-Free B_{12} Ligands Hydrogenobalamin and Hydrogenobyric Acid. *Angew. Chemie* **2020**, *132* (45), 20304–20311. https://doi.org/10.1002/ange.202008407.

(8) Kieninger, C.; Baker, J. A.; Podewitz, M.; Wurst, K.; Jockusch, S.; Lawrence, A. D.; Deery, E.; Gruber, K.; Liedl, K. R.; Warren, M. J.; et al. Zinc Substitution of Cobalt in Vitamin B_{12}: Zincobyric Acid and Zincobalamin as Luminescent Structural B_{12}-Mimics. *Angew. Chemie - Int. Ed.* **2019**, *58* (41), 14568–14572. https://doi.org/10.1002/anie.201908428.

(9) Widner, F. J.; Kieninger, C.; Wurst, K.; Deery, E.; Lawrence, A. D.; Warren, M. J.; Kräutler, B. Synthesis, Spectral Characterization and Crystal Structure of Chlororhodibalamin: A Synthesis Platform for Rhodium Analogues of Vitamin B_{12} and for Rh-Based Antivitamins B_{12}. *Synth.* **2021**, *53* (2), 332–337. https://doi.org/10.1055/s-0040-1707288.

(10) Wiedemair, M.; Kieninger, C.; Wurst, K.; Podewitz, M.; Deery, E.; Paxhia, M. D.; Warren, M. J.; Kräutler, B. Solution, Crystal and in Silico Structures of the Organometallic Vitamin B_{12}-Derivative Acetylcobalamin and of Its Novel Rhodium-Analogue Acetylrhodibalamin. *Helv. Chim. Acta* **2023**, *106* (2). https://doi.org/10.1002/hlca.202200158.

(11) Ruetz, M.; Koutmos, M.; Kräutler, B. Antivitamins B_{12}: Synthesis and Application as Inhibitory Ligand of the B_{12}-Tailoring Enzyme CblC. In *Methods in Enzymology*; Academic Press, 2022; Vol. 668, pp 157–178. https://doi.org/10.1016/bs.mie.2021.12.016.

(12) Zelder, F. Recent Trends in the Development of Vitamin B_{12} Derivatives for Medicinal Applications. *Chem. Commun.* **2015**, *51* (74), 14004–14017. https://doi.org/10.1039/c5cc04843e.

(13) Kräutler, B. Antivitamins B_{12}—Some Inaugural Milestones. *Chem. - A Eur. J.* **2020**, *26* (67), 15438–15445. https://doi.org/10.1002/chem.202003788.

(14) Liptak, M. D.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{1+}Cobalamin: Spectral and Electronic Properties of the "Superreduced" B_{12} Cofactor. *J. Am. Chem. Soc.* **2006**, *128* (28), 9144–9156. https://doi.org/10.1021/ja061433q.

(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. U. S. A.* **2020**, *117* (48), 30412–30422. https://doi.org/10.1073/pnas.2007332117.

(16) Mera, P. E. & Escalante-Semerena, J. C. Multiple Roles of ATP:cob(I)alamin Adenosyltransferases in the Conversion of B_{12} to Coenzyme B_{12}. *Applied Microbiology and Biotechnology* vol. 88 41–48 (2010). https://doi.org/10.1007/s00253-010-2773-2.

(17) Banerjee, R. B_{12} Trafficking in Mammals: A Case for Coenzyme Escort Service. *ACS Chemical Biology*. **2006**, *1* (3), 149–159. https://doi.org/10.1021/cb6001174.

(18) Banerjee, R.; Gouda, H.; Pillay, S. Redox-Linked Coordination Chemistry Directs Vitamin B_{12} Trafficking. *Acc. Chem. Res.* **2021**, *54* (8), 2003–2013. https://doi.org/10.1021/acs.accounts.1c00083.

(19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. E.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision C.01. *Gaussian, Inc., Wallingford CT*. 2016.

(20) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098. https://doi.org/10.1103/PhysRevA.38.3098.

(21) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33* (12), 8822–8824. https://doi.org/10.1103/PhysRevB.33.8822.

(22) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(23) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(24) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105* (22), 9982–9985. https://doi.org/10.1063/1.472933.

(25) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.

(26) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82* (1), 299–

310. https://doi.org/10.1063/1.448975.

(27) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82* (1), 270–283. https://doi.org/10.1063/1.448799.

(28) Wadt, W. R.; Hay, P. J. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985**, *82* (1), 284–298. https://doi.org/10.1063/1.448800.

(29) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54* (2), 720–723. https://doi.org/10.1063/1.1674902.

(30) Hehre, W. J.; Ditchfield, K.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261. https://doi.org/10.1063/1.1677527.

(31) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28* (3), 213–222. https://doi.org/10.1007/BF00533485.

(32) Hariharan, P. C.; Pople, J. A. Accuracy of AHn Equilibrium Geometries by Single Determinant Molecular Orbital Theory. *Mol. Phys.* **1974**, *27* (1), 209–214. https://doi.org/10.1080/00268977400100171.

(33) Gordon, M. S. The Isomers of Silacyclopropane. *Chem. Phys. Lett.* **1980**, *76* (1), 163–168. https://doi.org/10.1016/0009-2614(80)80628-2.

(34) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77* (7), 3654–3665. https://doi.org/10.1063/1.444267.

(35) Binning, R. C.; Curtiss, L. A. Compact Contracted Basis Sets for Third-row Atoms: Ga–Kr. *J. Comput. Chem.* **1990**, *11* (10), 1206–1216. https://doi.org/10.1002/jcc.540111013.

(36) Blaudeau, J. P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. Extension of Gaussian-2 (G2) Theory to Molecules Containing Third-Row Atoms K and Ca. *J. Chem. Phys.* **1997**, *107* (13), 5016–5021. https://doi.org/10.1063/1.474865.

(37) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K through Zn. *J. Chem. Phys.* **1998**, *109* (4), 1223–1229. https://doi.org/10.1063/1.476673.

(38) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* Basis Set for Third-Row Atoms. *J. Comput. Chem.* **2001**, *22* (9), 976–984. https://doi.org/10.1002/jcc.1058.

(39) PyMOL Molecular Graphics System, Ver. 2.5; Schrödinger, LLC.

(40) Pavlova, A.; Parks, J. M.; Gumbart, J. C. Development of CHARMM-Compatible Force-Field Parameters for Cobalamin and Related Cofactors from Quantum Mechanical Calculations. *J. Chem. Theory Comput.* **2018**, *14* (2), 784–798. https://doi.org/10.1021/acs.jctc.7b01236.

(41) Bayly, C. I.; Merz, K. M.; Ferguson, D. M.; Cornell, W. D.; Fox, T.; Caldwell, J. W.;

Kollman, P. A.; Cieplak, P.; Gould, I. R.; Spellmeyer, D. C. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197. https://doi.org/10.1021/ja00124a002.

(42) Marques, H. M.; Ngoma, B.; Egan, T. J.; Brown, K. L. Parameters for the AMBER Force Field for the Molecular Mechanics Modeling of the Cobalt Corrinoids. *J. Mol. Struct.* **2001**, *561* (1–3), 71–91. https://doi.org/10.1016/S0022-2860(00)00920-0.

(43) Adhikari, S.; Sutradhar, D.; Shepherd, S. L.; Phillips, R. M.; Chandra, A. K.; Rao, K. M. Synthesis, Structural, DFT Calculations and Biological Studies of Rhodium and Iridium Complexes Containing Azine Schiff-Base Ligands. *Polyhedron* **2016**, *117*, 404–414. https://doi.org/10.1016/j.poly.2016.06.001.

(44) Dinda, S.; Patra, S. C.; Ganguly, S. Rhodium(III) Complex with Pyrene-Pyridyl-Hydrazone: Synthesis, Structure, Ligand Redox, Spectral Characterization and DFT Calculation. *J. Chem. Sci.* **2019**, *131* (3), 1–9. https://doi.org/10.1007/s12039-019-1598-5.

(45) Haneder, S.; Da Como, E.; Feldmann, J.; Lupton, J. M.; Lennartz, C.; Erk, P.; Fuchs, E.; Molt, O.; Münster, I.; Schildknecht, C.; et al. Controlling the Radiative Rate of Deep-Blue Electrophosphorescent Organometallic Complexes by Singlet-Triplet Gap Engineering. *Adv. Mater.* **2008**, *20* (17), 3325–3330. https://doi.org/10.1002/adma.200800630.

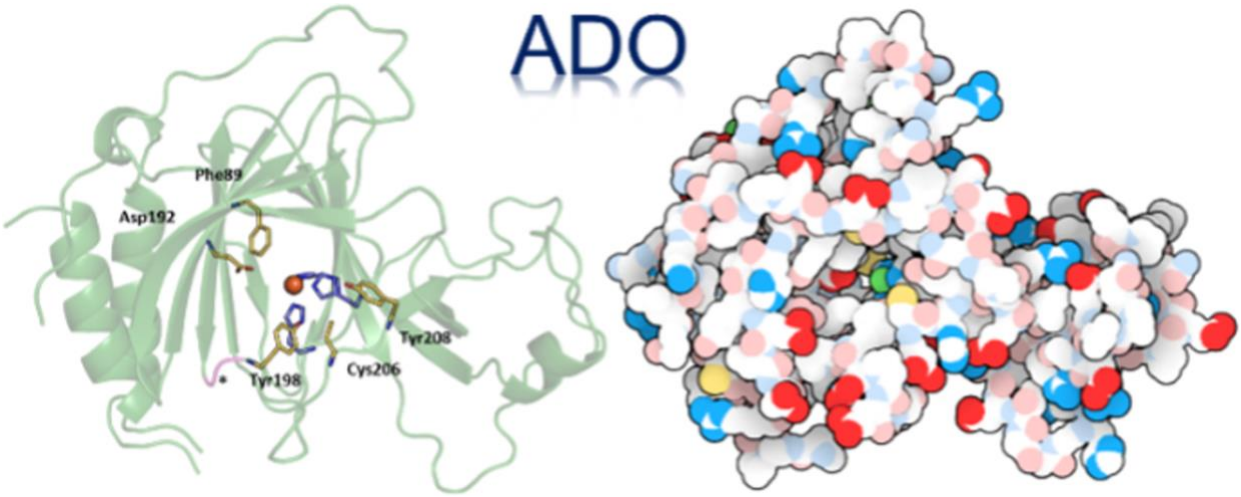
(46) Brahim, H.; Haddad, B.; Brahim, S.; Guendouzi, A. DFT/TDDFT Computational Study of the Structural, Electronic and Optical Properties of Rhodium (III) and Iridium (III) Complexes Based on Tris-Picolinate Bidentate Ligands. *J. Mol. Model.* **2017**, *23* (12), 1–11. https://doi.org/10.1007/s00894-017-3517-3.

(47) Moussa, J.; Chamoreau, L. M.; Degli Esposti, A.; Gullo, M. P.; Barbieri, A.; Amouri, H. Tuning Excited States of Bipyridyl Platinum(II) Chromophores with π-Bonded Catecholate Organometallic Ligands: Synthesis, Structures, TD-DFT Calculations, and Photophysical Properties. *Inorg. Chem.* **2014**, *53* (13), 6624–6633. https://doi.org/10.1021/ic500232w.

(48) Rovira, C.; Kozlowski, P. M. First Principles Study of Coenzyme B_{12}. Crystal Packing Forces Effect on Axial Bond Lengths. *J. Phys. Chem. B* **2007**, *111* (12), 3251–3257. https://doi.org/10.1021/jp0660029.

Chapter 6

The Crystal Structure of Cysteamine Dioxygenase Reveals the Origin of the Large Substrate Scope of This Vital Mammalian Enzyme



This chapter has been published under the following: Fernandez, R. L.; Elmendorf, L. D.; Smith, R. W.; Bingman, C. A.; Fox, B. G.; Brunold, T. C. The Crystal Structure of Cysteamine Dioxygenase Reveals the Origin of the Large Substrate Scope of This Vital Mammalian Enzyme. *Biochemistry* **2021**, *60* (48), 3728–3737. https://doi.org/10.1021/acs.biochem.1c00463.

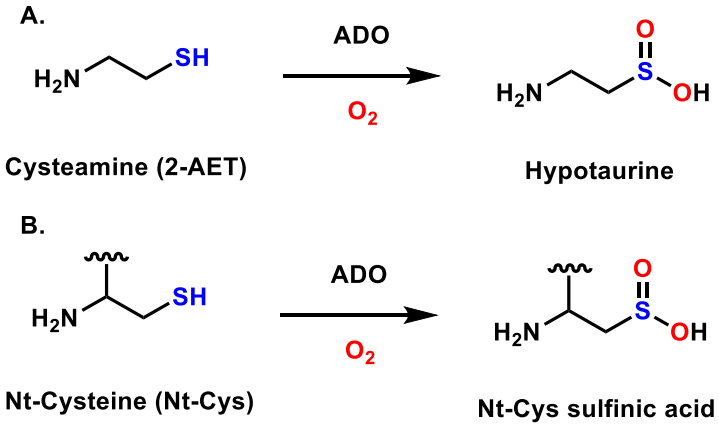
Chapter 6: The Crystal Structure of Cysteamine Dioxygenase Reveals the Origin of the Large Substrate Scope of This Vital Mammalian Enzyme

6.1. Introduction

The metabolism of thiol-containing compounds is vital to mammalian homeostasis. Regulation of hypotaurine, taurine, and cysteine levels preserves cardiac and vascular functions and protects neural cells from excitotoxicity. Additionally, elevated cysteine levels have been associated with Alzheimer's and Parkinson's diseases.[1] To regulate intracellular levels of these thiolate-containing metabolites, mammals use two distinct thiol dioxygenases (TDOs); i.e., cysteine dioxygenase (CDO) and cysteamine dioxygenase (ADO).[2,3] The definitively established function of CDO is to catalyze the oxidation of cysteine (Cys) to cysteine sulfinic acid (CSA) via the incorporation of both oxygen atoms from molecular oxygen.[4] CSA can then be catabolized into pyruvate, sulfate, and hypotaurine.[5] Under normal conditions, rats or mice convert about 70-90% of CSA to hypotaurine and taurine.[6] Hypotaurine has been identified as a contributor to the growth and progression of aggressive high-grade gliomas in the brain.[7] Additionally, taurine has been found to serve as an osmoregulator in the heart,[8] though this compound also plays a role in several essential body functions, such as regulating calcium levels, creating bile cells, balancing electrolytes in the body, and supporting the development of the nervous system.[9]

ADO was once believed to exclusively convert cysteamine (2-aminoethanethiol, 2-AET) to hypotaurine (**Scheme 6.1A**). However, Ratcliffe et al. proposed that the native function of ADO may instead be to catalyze the O_2-dependent conversion of peptides featuring an amino-terminal cysteine (Nt-Cys, **Scheme 6.1B**) to their CSA derivatives.[10] They found that mammalian ADO could replace one of the five plant thiol dioxygenases in *Arabidopsis thaliana (At)*, termed plant cysteine oxidase 4 (PCO4), by oxidizing the plant Nt-Cys peptides both *in vitro* and *in vivo*.

Oxidation of the Nt-Cys peptide substrates by PCO promotes arginylation and degradation in the N-degron pathway,[11] while under anoxic conditions, the peptide persists and enhances Gα-catalyzed GTP hydrolysis, leading to an attenuation of G-protein coupled signaling. Newly identified ADO substrates, regulators of G protein signaling RGS4 and RGS5, are part of the mammalian N-degron pathway and serve as negative regulators of cardiovascular function controlled by G-protein signaling.[12,13]



Scheme 6.1. Reactions catalyzed by ADO.

Five classes of TDOs have been identified to date with differing native functions: namely, CDO, ADO, PCO, 3-mercaptopropionate dioxygenase (MDO), and mercaptosuccinate dioxygenase (MSDO), with CDO and ADO being the only mammalian TDOs. All TDOs belong to the cupin superfamily, which is typified by a common architecture.[14] While cupin proteins tend to have low overall sequence identity, all feature a β-barrel fold and two conserved sequence motifs: $G(X)_5HXH(X)_6G$ and $G(X)_5PXG(X)_2H(X)_3N$. In addition to its classification as a cupin protein, ADO belongs to the PFam family PF07847 (PCO_ADO), which is distinct from the PF05995 (CDO_I) family that includes CDO and MDO. As is typical for members of the cupin superfamily, *Mus musculus* ADO (*Mm*ADO) shares little overall sequence identity with either *Rattus norvegicus* CDO (~14%, *Rn*CDO) or PCO4 (~21%). Nevertheless, all TDOs that have

previously been characterized by X-ray crystallography (CDO, MDO, and PCO) feature the same non-heme iron coordination environment, consisting of a relatively rare 3-histidine (3-His) facial triad. The impact on enzyme activity of removal of one of the coordinating His via site-directed mutagenesis and spectroscopic studies provided compelling evidence that ADO also contains the 3-His binding motif but differs with respect to secondary sphere residues that determine its substrate specificity.[3,15]

An alignment of the amino acid sequences of *Mm*ADO and *Rn*CDO highlights crucial differences in the active site pocket.[3] Perhaps most noticeable, the Cys93 involved in forming the Cys-Tyr cross-link of mammalian CDOs is not conserved in ADO. In *Rn*CDO, this unusual thioether linkage increases activity by properly positioning the Cys and O_2 substrates while also suppressing the coordination of a water molecule that competes with O_2 binding.[16,17] A Cys-Tyr cross-link motif was identified in human ADO via genetic incorporation of an unnatural amino acid, 3,5-difluoro-tyrosine, in conjunction with mass spectrometry and NMR experiments.[18] While the cross-link in *Rn*CDO forms between Cys93 and Tyr157, in human ADO the residues involved, Cys206 and Tyr208 (*Mm*ADO numbering), are separated by only one amino acid. Based on sequence alignment, the ADO cross-link motif is plausible in PCOs; however, neither X-ray crystallography nor tandem MS/MS analyses provided evidence for the formation of a Cys-Tyr thioether bond in PCO4.[19] Consequently, the extent of cross-link formation in *Mm*ADO and its physiological role remain uncertain.

The lack of an X-ray crystal structure of mammalian ADO and the low overall sequence identity between ADO and other TDOs have rendered comparisons with other TDOs uncertain. While the presence of a 3-His facial triad coordinating the Fe cofactor was confirmed on the basis of spectroscopic and site-directed mutagenesis experiments, little has previously been established

about ADO beyond its first coordination sphere. Thus, a structural characterization of ADO was pivotal to identifying key residues involved in substrate binding and other salient features of the protein. Here, we describe the 1.9 Å X-ray crystal structure of *Mm*ADO. As proposed on the basis of previous studies, ADO features a 3-His facial triad that coordinates the iron center. Both the X-ray crystal structure of ADO and a sequence alignment of various TDOs reveal that the structure of ADO is more similar to that of PCO than CDO. As reported for PCO,[19] cross-link formation between Cys206 and Tyr208 was not observed in the structure of *Mm*ADO. Interestingly, a pair of Cys residues is identified that could form a disulfide bond, thus blocking a secondary access tunnel and controlling catalytic activity via regulating delivery of the co-substrate O_2. Finally, molecular dynamics in combination with quantum mechanics/molecular mechanics calculations, validated on the basis of the X-ray crystal structure of resting ADO, were used to generate models of 2-AET and Nt-Cys bound ADO.

6.2. Materials and Methods

6.2.1. Preparation of Recombinant *Mm*ADO

ADO was expressed as previously described[15] with minor additions to the purification protocol. Briefly, *Escherichia coli* Rosetta 2(DE3) cells were transformed with a pET SUMO expression vector containing a codon-optimized *Mus musculus* ADO gene and protein expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.2 mM along with ferrous ammonium sulfate to a final concentration of 100 μM. Cells were harvested 4 hrs post-induction and flash frozen.

Soluble protein was purified via immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) as previously described with additional steps.[15] All buffers included 5 mM tris(2-carboxyethyl)phosphine (TCEP) as a reducing agent and were degassed and

bubbled with argon gas to minimize oxygen presence in buffer. Post elution from the SEC column, the ADO/SUMO fusion protein was dialyzed via centrifugation into 200 mM Tris-HCl, 150 mM NaCl, 5 mM TCEP, 5 mM imidazole, pH 8.0 IMAC buffer. Thin-layer chromatography was used to ascertain that the purified ADO protein was able to catalyze the conversion of cysteamine to hypotaurine (work is underway to develop a quantitative activity assay). The protein was then incubated overnight at room temperature with SUMO protease (ThermoFisher). ADO was separated from the His_6SUMO tag via reverse IMAC. Protein aliquots to be used for crystallization were flash frozen under a stream of argon gas and stored at -80 °C.

6.2.2. Crystallization and structure determination

Purified *Mm*ADO was screened for crystallization response using commercial screens, MRC SD2 microplates, and a Mosquito crystallization robot. Initial diffraction studies indicated almost no crystalline diffraction, even though the crystals seemed well-formed by light microscopy. Progressive exclusion of oxygen during protein purification and crystallization environment yielded progressively better diffraction. The exceptional crystal yielding the refined diffraction data was grown by hanging drop vapor diffusion at 293 K in a Coy anaerobic chamber. The reservoir solution was allowed to degas overnight in the anaerobic chamber, and consisted of 15% PEG 3350, 300 mM $MgCl_2$, and 0.1 M Tris-HCl pH 8.0. The plate was withdrawn from the chamber, and the crystal was quickly cryoprotected against reservoir solution supplemented with 20% ethylene glycol and cooled by direct immersion in liquid nitrogen.

Diffraction data were collected at sectors 21 and 23 of the Advanced Photon Source, Argonne National Laboratory. The refinement data set was collected on an Eiger 9M detector at Life Sciences Collaborative Access Team beamline 21ID-D. A full 360 degree sweep of data was collected at 1.127 Å, 155 mm sample to detector distance, 0.2 degrees/frame, and 0.04 s exposure

time. Data were reduced using XDS[20] and autoPROC.[21] Crystallographic structure solution and refinement were conducted within the Phenix suite of programs.[22] The structure was solved by molecular replacement as implemented in Phaser,[23] using an edited homology model from SWISS-MODEL[24] derived from a PDB:6SBP template, covering residues 10-121 and 153-210 of *Mm*ADO. Molecular replacement initially produced the location and orientation of three copies of *Mm*ADO. The location and orientation of the fourth and final copy was uncovered by molecular replacement using a partially refined model based on the first three positions discovered. The structure was iteratively rebuilt in Coot[25] and refined using Phenix.refine.[26] Figures for all protein structures were created using PyMOL.[27]

6.2.3. Molecular Dynamics (MD) Simulations

MD simulations were performed using the GROMACS (versions 5.1.4 and 2019.6) software package[28] with the AMBER ff19SB force field.[29] Initial coordinates were taken from chain D of the crystal structure reported here, keeping all crystallographic waters within 6 Å of that chain. We chose chain D for our computational studies because this is the most complete chain in the structure. Very similar results were obtained using the second most complete chain in the structure, chain B. Protonation was performed using phenix.reduce and manually adjusted, as necessary. A missing segment spanning Thr26 to Glu29 was constructed manually in PyMOL. MD parameters were generated in AmberTools,[30] using the metal center parameter builder (MCPB.py) modelling tool for Fe(II) and its ligands (His100, His102, His179, Fe-bound waters, and later the Fe-bound substrates).[31] The system was solvated with the SPC/E water model,[32] and its total charge was neutralized by the addition of Na^+ ions. Then, an energy minimization step was performed, followed by two equilibration steps: a 100 ps run under the NVT (isothermal-isochoric) ensemble and a 100 ps run under the NPT (isothermal-isobaric) ensemble. For each

model, structures from the MD trajectory were clustered based on the root-mean-square deviation (RMSD) of the protein backbone. The average structures of the five most populated clusters were chosen for QM/MM optimization.

Modifications to this workflow were required for some models. In the case of the disulfide model, a bond between Cys120 and Cys169 was added to the crystal structure *in silico*. This system required an additional 10 ns added to its MD simulation, as the RMSD of the protein backbone was found to still be rising substantially at the end of the first 10 ns. For the substrate-bound models, starting coordinates were taken from the QM/MM-optimized ADO model, with the substrate placed into the active site using the docking software AutoDock Vina.[33]

6.2.4. Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations

Geometry optimizations were performed using the quantum mechanics/molecular mechanics (QM/MM) approach as implemented in the ONIOM method of Gaussian 16.[34] The QM region was defined as the Fe ion, the side chains of its three histidine ligands (His100, His102, His179), Fe-bound water molecules, and 2-AET or the terminal cysteine of the peptide substrate if applicable. The density functional theory (DFT) portion of the calculations was performed with the unrestricted Becke, 3-parameter, Lee-Yang-Parr (UB3LYP) functional,[35,36] using the triple-ζ valence plus polarization (TZVP) basis set[37] for Fe and its coordinating atoms, and the 6-31G basis set for all other atoms.[38] For all atoms outside of the QM region, the MM calculation was performed with the AMBER force field.[39]

6.3. Results

6.3.1. Crystallization and structure determination

The protein was produced as an ADO/SUMO fusion and crystallized under anaerobic conditions upon removal of the His_6SUMO solubility tag. The structure of ADO was solved to a

nominal 1.89 Å resolution (crystallographic statistics, Supporting Information Table S1). The crystallographic asymmetric unit comprises four independent copies of the protein. In all four independent monomers, a short loop of 4-8 residues between Phe21 and Pro31 was too disordered to model. In two subunits, a second 4-6 residue loop between Pro217 and Ala224 was also disordered. Results from a SuperPose analysis for the four different ADO monomers in the unit cell are presented in Table S2, and an overlay of the four independent monomers is shown in the Supporting Information Fig. S1. Coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with PDB ID code 7LVZ.

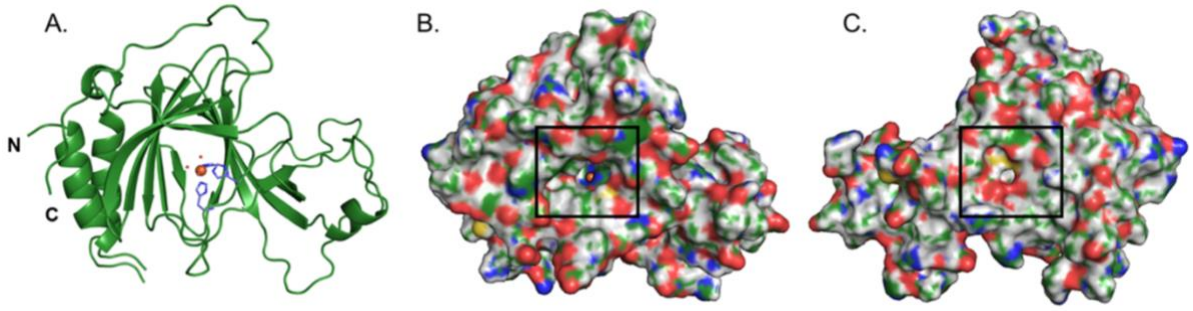


Figure 6.1. *(A)* Cartoon representation of the overall structure of ADO. The active site consists of an Fe ion (orange sphere) that is coordinated by a 3-His triad (purple) and 3 water molecules (red spheres). *(B)* Surface representation of the protein with a black box highlighting the peptide substrate access tunnel to the active site. The RBG coloring is as follows: Fe, orange; carbon, green; hydrogen white, oxygen, red; nitrogen, blue; and sulfur, yellow. *(C)* Same as in B but rotated 180° to reveal the entrance to the putative co-substrate (O_2) tunnel, which is highlighted by a black box.

6.3.2. Protein architecture

The ADO protein adopts a β-barrel structure typical of the cupin superfamily. The 3-His facial triad coordinating the catalytically relevant Fe ion lies at the center of the protein flanked by β-sheets (**Figure 6.1A**). Interestingly, in all four independent copies of the protein in the unit cell, the His130 side chain of a different ADO monomer is present in the active site (Supporting Information Fig. S2). In two cases, His130 is supplied by another protein within the asymmetric unit. In the other two cases, His130 is supplied by a symmetry related protein. In PCO,[19] a His

residue from the N-terminal His_6 tag also protrudes into the active site. However, in contrast to what is observed for PCO, the His residue in the ADO structure does not bind directly to the Fe cofactor. Additionally, in ADO there are no obvious interactions between the His side chain and secondary sphere residues.

The surface representation of ADO highlights a prominent tunnel into the active site through which substrate is likely to bind (**Figure 6.1B**). Although such a tunnel also exists in CDO, the amino acids that define the CDO substrate tunnel are not conserved in ADO, and the large cavity in ADO is in a structurally distinct location as compared to CDO. Interestingly, the surface representation of the "backside" of ADO features a secondary tunnel from the protein surface to the active site (**Figure 6.1C**). It is tempting to speculate that once cysteamine or a bulky Nt-Cys substrate is bound, this much smaller tunnel could provide access for co-substrate O_2 to the Fe cofactor. The entrance to this tunnel is guarded by two Cys residues pointed directly at each other with an S···S distance of ~4 Å (further discussed in *Distinguishing ADO motifs*).

6.3.3. Active site

The Fe cofactor resides in an octahedral coordination environment composed of His100, His102, and His179 (**Figure 6.2**). Because the protein used to obtain this structure was produced and purified semi-anaerobically and crystallized under anaerobic conditions in the presence of 2-AET, the Fe center is almost certainly in the 2+ (i.e., ferrous) oxidation state. In three out of the four crystallographically independent copies in the unit cell, three well-resolved water molecules occupy the remaining coordination sites. In the fourth copy, only two coordinating water molecules are clearly resolved while additional electron density is present in two different regions, 2.18 and 3.65 Å from the Fe center. Although this electron density may seem consistent with a diatomic (e.g., O_2-derived; apparent O–O distance of 1.89 Å) ligand, it is presumably associated with a third

water molecule that is disordered over two positions given that ADO was crystallized under anaerobic conditions. Notably, in all four ADO monomers in the unit cell, one of the coordinated water molecules is within hydrogen-bonding distance of the hydroxyl group of Tyr198. A similar hydrogen-bonding interaction between the equivalent Tyr182 residue and a water ligand exists in PCO.[19]

The active site is surrounded by a hydrophobic pocket with few residues, primarily Leu and Ile, closer than 5 Å to the Fe center. Asp192 of *Mm*ADO corresponds to the PCO4 residue Asp176, which has been shown to be critical for PCO function. Specifically, in the as-isolated Asp176Asn PCO4 variant, only 10% of the active sites contained Fe and the enzymatic activity was decreased ~10-fold even after iron supplementation.[19] The next set of residues near the ADO active site are Phe89, Cys206, and Tyr208, all located ~7 Å from the cofactor. Although these residues are not close to the Fe cofactor, their locations bordering the substrate access tunnel suggest that they could aid in positioning substrate.

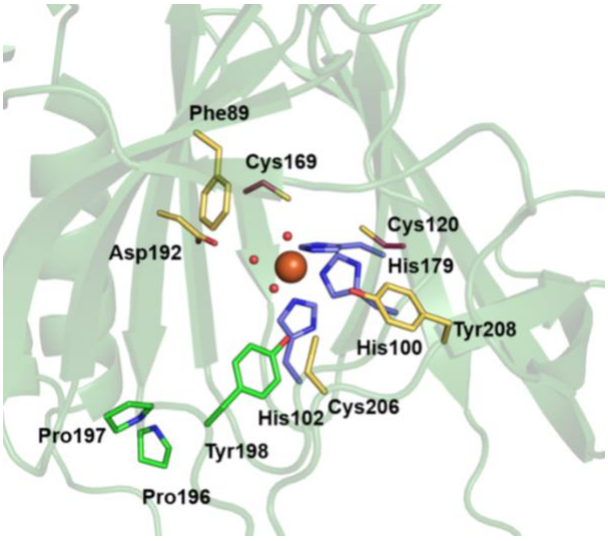


Figure 6.2. Active site region of *Mm*ADO viewed through the substrate tunnel. The Fe ion (orange sphere) is bound by His100, His102, His179 (purple), and three water molecules (red spheres). Other amino acids in the active site are highlighted in yellow (Phe89, Asp192, Cys206, and Tyr208) or green (Pro196, Pro197, and Tyr198). A stereo view of the active site is shown in the Supporting Information Fig. S3.

In human ADO, Cys220 and Tyr222 (corresponding to Cys206 and Tyr208 of *Mm*ADO) have been shown to be capable of forming a thioether cross-link.[18] In the absence of a crystal structure, these two residues were presumed to be located near the Fe cofactor, as cross-link formation required the presence of 2-AET and O_2. In the *Mm*ADO structure, there is no electron density supporting the existence of a cross-link between Cys206 and Tyr208. The lack of a cross-link in this structure is perhaps not surprising, because the protein used to obtain this structure was produced and purified semi-anaerobically and crystallized under anaerobic conditions in the presence of 2-AET (see Methods section). However, the fraction of cross-linked ADO purified aerobically appears to be very small based on published high-resolution mass spectra.[18] Moreover, in the case of PCO4, no cross-link was observed in the X-ray crystal structure or via incubation of protein with substrate and tandem MS/MS analysis.[19] Thus, the physiological relevance of the Cys-Tyr cross-link in ADO remains in question.

6.3.4. Distinguishing ADO motifs

X-ray crystallographic studies revealed that the first coordination spheres of TDOs are identical, with the PCO, CDO, MDO, and, as shown here, ADO active sites all featuring an Fe center coordinated by a 3-His triad. It is the secondary sphere of the active site that boasts unique residues and sequence motifs to tailor the specific reactivity of each enzyme (e.g., the Ser153-His155-Tyr157 catalytic triad in CDO and Gln62 in MDO). The distinguishing MDO and CDO secondary sphere features are neither conserved in the ADO amino acid sequence nor replicated in the X-ray crystal structure. Contrastingly, the ADO and PCO4 active site architectures are remarkably similar, perhaps indicative of closely analogous, or even identical, functions.[19] An alignment of *Mm*ADO, *At*PCO4, *Pa*MDO, *Mm*CDO, and *Bs*CDO was constructed (**Figure 6.3**) to

guide the identification of key residues and their location in the ADO crystal structure (**Figure 6.2**).

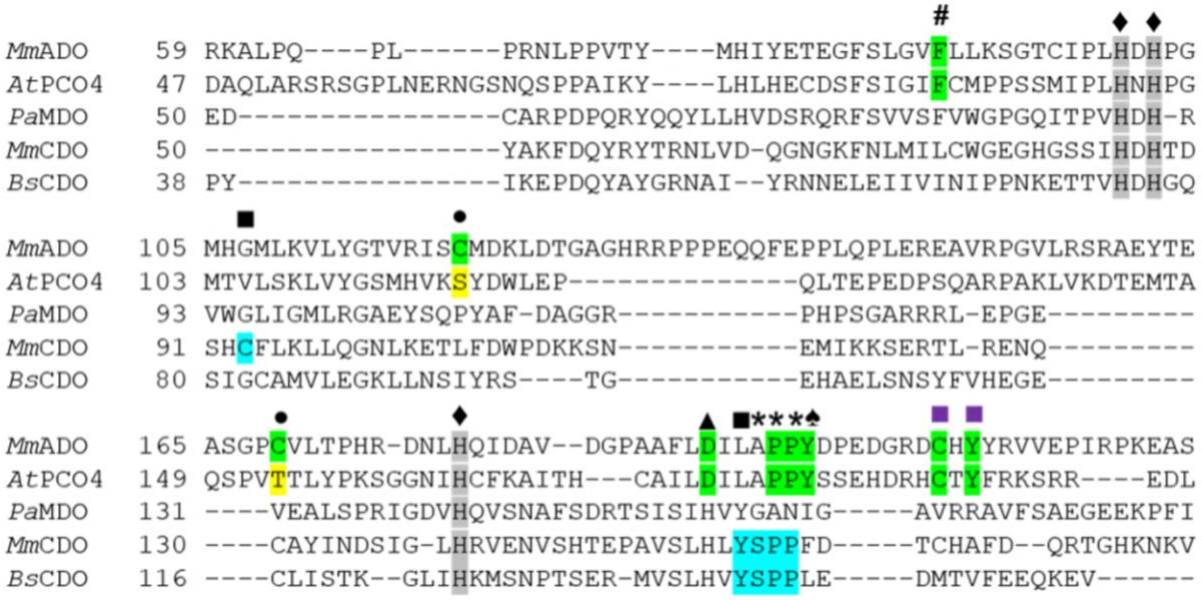


Figure 6.3. Sequence alignment showing conserved residues within TDOs. The sequences of the PCO4 from *Arabidopsis thaliana* (*At*PCO4), the MDO from *Pseudomonas aeruginosa* (*Pa*MDO), and the CDOs from *Mus musculus* (*Mm*CDO) and *Bacillus subtilis* (*Bs*CDO) are compared to the *Mus musculus* ADO (*Mm*ADO) sequence. Important ADO and PCO residues are highlighted in green, for CDO in blue, and key differences in yellow. All proteins possess a 3-His metal-binding motif marked by a ♦ and highlighted in gray. Additional motifs discussed in the text are denoted as follows: # Phe89 (*Mm*ADO numbering), ■ Cys-Tyr cross-link in *Mm*CDO, ■ putative ADO and PCO cross-link motifs, • Cys120 and Cys169 in *Mm*ADO are replaced by Ser118 and Thr153 in PCO4, ▲ Asp192, ***** denotes a *cis*-peptide bond, and ♠ Tyr198 in *Mm*ADO and Tyr182 in *At*PCO4 adjacent to the *cis*-peptide bond.

TDOs are relatively small proteins with sequences containing 200 to 300 amino acids. *Mm*ADO, composed of 256 amino acids, boasts a remarkable 33 Pro residues, compared to 23 and 7 found in *At*PCO4 and *Rn*CDO, respectively. Interestingly, the ADO structure shows a *cis*-peptide bond between Pro196 and Pro197. Jameson et al. established that the presence or absence of a *cis*-peptide bond between Ser158 and Pro159 results in a different positioning of the key Tyr157 residue (*Rn*CDO numbering) involved in the thioether cross-link of CDO.[40,41] Although in CDO the Tyr157 residue precedes this *cis*-peptide motif, in ADO a Tyr residue (Tyr198) is present directly after Pro196 and Pro197; thus, the orientation of Tyr198 is also controlled by a *cis*-peptide

motif. Intriguingly, Tyr198 and Tyr208 reside equidistant and in a similar position from Cys206 (**Figure 6.2**, yellow). This Tyr198-Cys206-Tyr208 sandwich borders the substrate tunnel near the protein-solvent barrier. An analogous loop is conserved in PCO4, where it was postulated to play a role in peptide substrate recognition and binding.[19] On the basis of these observations, it is tempting to speculate that the Tyr198 residue in ADO plays a larger role in substrate binding and enzyme function than does Tyr208.

The backside tunnel to the active site (**Figure 6.1C**) is lined by two surface-exposed Cys residues, Cys120 and Cys169 (**Figure 6.2,** yellow). While not bonded in this crystal structure, the residues are 4.8 Å apart and gate a potential secondary access tunnel to the active site. We hypothesize that under proper oxidizing conditions, these residues could form a disulfide bond so as to control O_2 access to the active site, and thus enzyme activity. Computational studies aimed at assessing the feasibility of this hypothesis, as well as the generation of faithful models of substrate-bound ADO are presented next.

6.3.5. Computational analysis of resting and substrate-bound ADO

Repeated attempts were made to crystallize ADO with substrate 2-AET bound. As these attempts were unsuccessful, computational methods were used to generate whole-protein models of ADO complexed with 2-AET and the CKGL tetramer, representative of RGS5. Initially, a computational model of ADO in the absence of substrate was constructed starting from the crystal structure reported in this study to assess the feasibility of the computational approach chosen. The conformational landscape was sampled by first performing an unrestrained molecular dynamics (MD) simulation in GROMACS,[28] after which a clustering algorithm was used to select five representative ADO structures that encompassed the majority of conformations accessed during the simulation. The geometries of these structures were optimized using a quantum

mechanics/molecular mechanics (QM/MM) approach. The optimized model with the lowest energy was used for further analysis and as the foundation for subsequent calculations of substrate-bound ADO. Importantly, the QM/MM optimized lowest-energy ADO model shows minor deviations from the crystal structure, with the root mean square deviation (RMSD) of the backbone N and C atomic positions being 1.29 Å (for 614 atoms, using the default outlier cut-off value in PyMOL of 2 Å) even though a single ADO monomer was used in our calculations (Supporting Information Fig. S4). Additionally, both the prominent substrate channel and smaller secondary channel are preserved in the computational model, as are the bond lengths and angles of the active site.

In the QM/MM optimized ADO model, the Cys120 and Cys169 residues remain in close proximity, with an S···S distance of 3.5 Å (compared to ~4.8 Å in the crystal structure), hinting at the possibility of disulfide bond formation under oxic conditions. To further evaluate this possibility, a model of ADO containing a disulfide bond between these two residues was generated by altering the crystal structure *in silico* to install an S–S bond and using the computational workflow described above. Interestingly, the optimized ADO models with and without Cys120···Cys169 disulfide display only minor structural differences, both overall and in the active site region (Supporting Information Figs. S5-8). However, a significant narrowing of the backside substrate channel occurs in response to disulfide bond formation, which lends credence to the possibility of a gating mechanism that modulates O_2 access based on cellular oxidative conditions.

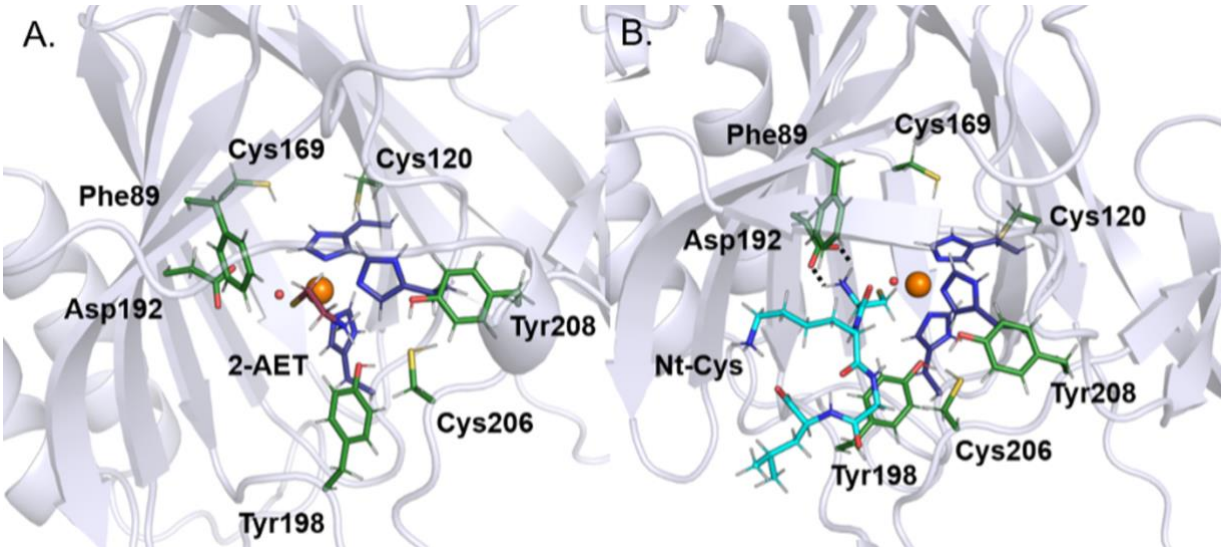


Figure 6.4. Active site regions of the QM/MM-optimized structures of ADO complexed with *(A)* 2-AET (burgundy sticks) and *(B)* the CKGL peptide (cyan sticks), representative of RGS5. The hydrogen-bonding interaction between the amine group of the peptide and Asp192 is indicated by broken lines.

Initial coordinates for substrate-bound ADO models were generated using the docking software AutoDock Vina[42] to place substrate into the active site of the QM/MM-optimized ADO model lacking the Cys120···Cys169 disulfide bond. Both 2-AET and the CKGL peptide were bound to Fe in a monodentate fashion with thiol-only coordination, as stipulated by previous spectroscopic studies of substrate-bound ADO.[15,43] Despite monodentate substrate binding, two water ligands were removed from the Fe center, one trans to His179 to allow for sulfur coordination and one trans to His100 to generate an open coordination site that would be required for O_2 binding and enzymatic turnover. These substrate-bound models were then subjected to MD simulations and QM/MM optimizations, as described above for resting ADO. In the case of 2-AET-bound ADO, the large size of the substrate cavity allowed for substantial conformational changes of this small substrate during the MD simulation, with the amine group pointing toward Asp192, Cys206, or nearby water molecules at different time points. In the lowest-energy QM/MM optimized model (**Figure 6.4A**), the amine group of 2-AET is positioned roughly equidistant from Tyr198 and Cys206. Notably, upon 2-AET binding to ADO, the putative secondary substrate

access channel becomes more defined, now extending from the surface to the active site so that it connects to the larger substrate channel.

Visualization of the MD trajectory for the ADO model complexed with the relevant portion of the RGS5 peptide (CKGL) makes it apparent that important differences exist in how 2-AET and peptide substrates interact with the secondary sphere of ADO. Compared to 2-AET, the peptide interacts with a much larger number of residues lining the active site cavity. After an initial equilibration period during the MD simulation, the orientation of the terminal cysteine remained largely unchanged, with the amine group always pointing toward Asp192. In the lowest energy QM/MM optimized structure (**Figure 6.4B**), the distance between these groups is 1.98 Å, suggesting that Asp192 engages in a hydrogen bonding interaction to properly situate the peptide substrate for catalysis. Another notable residue in the model of peptide-bound ADO is Tyr198, the residue situated next to the Pro196 and Pro197 *cis*-peptide motif. This residue likely participates in stabilizing interactions with substrate, as it is positioned closer than 4 Å from several side chains of the peptide. Finally, while the small backside channel lined by Cys120 and Cys169 appears more restricted than in the 2-AET-bound model, with a small occlusion to the protein's surface, it does not disappear entirely.

6.4. Discussion

Oxidation of 2-AET and Nt-Cys peptides are important biological transformations; yet, little is known about the enzyme, ADO, that catalyzes these reactions. The geometric and electronic structures of substrate-bound ADO have been spectroscopically investigated and both 2-AET and Nt-Cys have been found to coordinate to the Fe via the terminal thiolate moiety.[15,43] While kinetic studies demonstrated that ADO is capable of turning over both 2-AET and Nt-Cys peptides,[3,10] the substrate scope and specificity remain to be established. This work confirms that

ADO adopts both the canonical cupin architecture and 3-His facial triad typical of TDOs. However, unique structural motifs and secondary sphere elements are observed in the X-ray crystal structure of ADO that are not present in CDO or MDO. In addition, this structure provides the necessary foundation for future investigations of the mechanisms of dioxygen activation and peptide substrate oxidation employed by this vital mammalian TDO.

As ADO could not be crystallized with substrate bound, MD and QM/MM computations were employed to investigate the interactions between secondary sphere residues and Fe-bound substrate. Upon binding of CKGL (representative of the RGS5 peptide) to Fe(II)ADO, the amine group of the peptide adopts a position close enough to hydrogen bond with Asp192. Interestingly, the CKGL peptide is positioned closer to Tyr198 than Tyr208, the latter of which has been proposed to form a cross-link with Cys206. The small size of 2-AET allows for greater conformational flexibility within the active site, such that the amine does not interact with the same residues as the representative RGS5 peptide model. In particular, Tyr208 is too far away from 2-AET to affect the Fe–S bonding interaction in 2-AET-bound Fe(III)ADO. In support of this computational prediction, electron paramagnetic resonance spectra obtained for the cyanide/2-AET adducts of WT Fe(III)ADO and its Tyr208Phe variant are superimposable.[44]

Computational modeling of peptide-bound ADO shows that substrate RGS5 would likely fill the larger active site tunnel and force co-substrate O_2 delivery through a separate channel. To explore this possibility, MOLE2.5 was used to identify potential tunnels, cavities, and pores.[45–47] Upon lowering the tunnel radius from the default value of 1.25 to 0.49 Å (and using an internal threshold of 0.9 Å and a cutoff ratio of 0.5 while keeping all other parameters at their default values), a tunnel was identified that starts at residues Cys120 and Cys169 and ends at the Fe atom (Supporting Information Fig. S9). Thus, our MOLE2.5 analysis corroborates the proposal that this

tunnel could selectively allow O_2 to access the substrate-bound active site. Furthermore, QM/MM optimized models of ADO with Cys120 and Cys169 in their dithiol and disulfide forms display minor global structural differences, implying that these residues could form a disulfide bond under oxidizing conditions. As such, Cys120 and Cys169 could serve as a redox sensor to regulate O_2 access to the active site in response to changes in the intracellular redox potential. As ADO is capable of turning over a variety of Cys-containing substrates (though free L-cysteine inhibits ADO catalyzed turnover with cysteamine), it lacks the high substrate specificity characteristic of other TDOs, like CDO. The oxidation of Nt-Cys by ADO marks RGS5 for arginylation by ATE1, a poorly understood post-translational modification.[48] RGS5 is up-regulated in ADO and ATE1 deficient cells, establishing that ADO acts upstream of ATE1.[10] The presence of easily oxidizable residues such as Cys120/Cys169 and the use of O_2 as a co-substrate could connect the redox potential of the cell to enzyme activity, coupling arginylation to oxidative signaling and precluding ADO from depleting cells of Nt-Cys substrates, as well as Cys and 2-AET, in the presence of high O_2 levels.

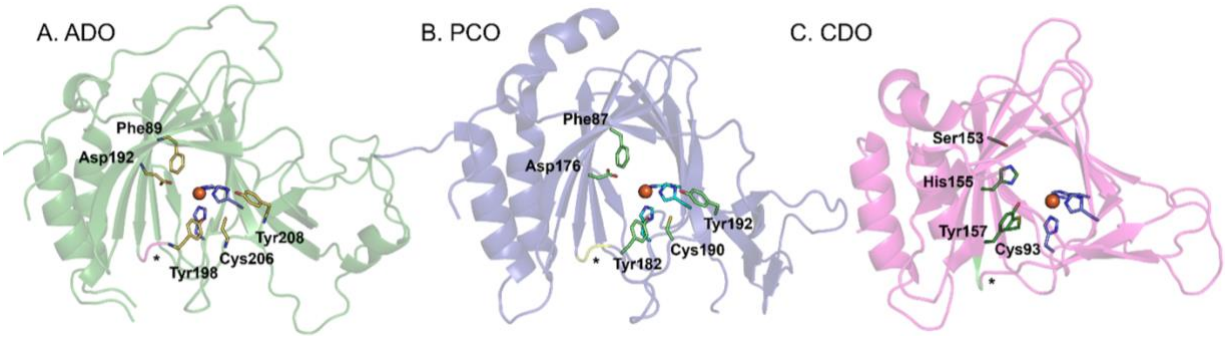


Figure 6.5. Comparison of the protein folds and active site regions of *(A) Mm*ADO (PDB: 7LVZ), *(B) At*PCO4 (PDB:6S7E), and *(C) Rn*CDO (PDB: 4JTO). The unusual *cis*-peptide bond that is preserved in all three enzymes is denoted by an asterisk (*).

6.4.1. Comparison to other thiol dioxygenases

A sequence alignment of relevant TDOs (**Figure 6.3**) demonstrates that ADO is most closely related to *At*PCO4 (21% sequence identity). Comparison of the *Mm*ADO and the *At*PCO4

X-ray crystal structures demonstrates that these enzymes share similar secondary structures (the RMSD of the backbone atomic positions between chain D of *Mm*ADO and *At*PCO4 is 0.90 Å for 472 atoms, using the default outlier cut-off value in PyMOL of 2 Å), as well as identical active site residues (**Figure 6.5A and B**). Intriguingly, ADO has been shown to complement the function of *At*PCO4 both in vivo and in vitro, although sequence similarity analyses have established that a greater than 70% sequence identity is typically needed to extrapolate conservation of function.[49,50]

Closer inspection of the *Mm*ADO and *At*PCO4 structures discloses distinct substrate and co-substrate access tunnels. While ADO contains one larger and one smaller access tunnel, PCO4 features a single large tunnel that extends from one side of the protein all the way to the other side and past the active site. Additionally, residues Cys120 and Cys169, which line the smaller access channel in ADO, are replaced by Ser118 and Thr153 in PCO4. Thus, the coupling of enzymatic activity to the intracellular redox potential that we propose to be important for the mammalian enzyme ADO is unlikely to occur in the plant enzyme PCO4. Considering that intracellular O_2 levels in mammalian cells are highly variable among different tissues, an extra layer of regulation of enzyme activity may be necessary.

The ADO metallocofactor resides in a hydrophobic pocket with one polar residue, Asp192, positioned within hydrogen bonding distance of an Fe-bound water (**Figure 6.5A**). To assess the functional role of the corresponding residue in PCO4, Asp176 (**Figure 6.5B**), the D176N variant was produced and kinetically characterized. The variant had 10% Fe incorporation and showed minimal activity even upon the addition of exogenous iron in the activity assay.[19] CDO also features a polar residue in the active site, Arg60 (**Figure 6.5C**), which has been shown to promote substrate binding via the formation of a salt bridge to the Cys carboxylate tail.[51] It is likely that the

negatively-charged Asp192 in ADO engages in a stabilizing interaction with a positively-charged amine group of Nt-Cys substrate.

Eukaryotic CDO features a cross-link between residues Cys93 and Tyr157 (**Figure 6.5C**). The formation of this cross-link enhances enzymatic activity by repositioning Cys93 and Tyr157 so as to allow for more favorable interactions with the Cys and O_2 substrates, while also preserving an open coordination site for O_2 binding. Although a thioether cross-link involving Cys206 and Tyr208 (*Mm*ADO numbering) has been identified in human ADO by Liu and coworkers,[18] in the crystal structure of *Mm*ADO these two residues reside ~7 Å from the Fe cofactor and no thioether bond is observed (as expected given the largely anaerobic conditions under which the enzyme was produced, purified, and crystallized). While the large separation of Cys206 and Tyr208 from the active site suggests that cross-link formation plays a minor structural or functional role in ADO, small differences between the mouse and human ADO sequences (85% sequence identity) could result in more pronounced differences between these two enzymes than one would expect. Intriguingly, as highlighted in **Figure 6.5**, the unusual *cis*-peptide bond that aids in the proper positioning of Tyr157 in CDO (Ser158 and Pro159) is also present in ADO (Pro196 and Pro197). In the QM/MM optimized structure of CKGL-bound ADO, Tyr198 is positioned close to the peptide substrate. The absence of a Cys206···Tyr208 cross-link in the crystal structure of *Mm*ADO, along with the lack of a direct interaction between Tyr208 and Nt-Cys or 2-AET in the computational models of substrate-bound ADO indicate that Tyr198 may be functionally more relevant than Tyr208.

In conclusion, the crystallographic and computational data of ADO presented here have revealed residues that could influence small molecule and peptide positioning within the active site. Although a thioether cross-link is key to increased turnover in mammalian CDO, a

comparable contribution remains elusive in the *Mm*ADO as the Cys206 and Tyr208 residues implicated in this cross-link reside ~7 Å from the Fe cofactor. Of note, computational modeling of Nt-Cys binding to the ADO indicates that Tyr198 may play a key role in positioning a peptide substrate in the active site. Importantly, a backside tunnel is identified that could serve to deliver O_2 to the peptide-bound ADO active site. This tunnel is lined by Cys120 and Cys169, with a distance and orientation compatible with redox-dependent control of O_2 access through disulfide bond formation. The unique combination of separate peptide substrate and O_2 binding channels, along with the potential for redox-sensitive control of O_2 access to the Fe- and 3-His active site represent a new model for understanding the function of mammalian thiol dioxygenases.

References

(1) Heafield, M. T., Fearn, S., Steventon, G. B., Waring, R. H., Williams, A. C., and Sturman, S. G. (1990) Plasma cysteine and sulphate levels in patients with motor neurone, Parkinson's and Alzheimer's disease. *Neurosci. Lett. 110*, 216–220.

(2) Simmons, C. R., Hirschberger, L. L., Machi, M. S., and Stipanuk, M. H. (2006) Expression, purification, and kinetic characterization of recombinant rat cysteine dioxygenase, a non-heme metalloenzyme necessary for regulation of cellular cysteine levels. *Protein Expr. Purif. 47*, 74–81.

(3) Dominy, J. E., Simmons, C. R., Hirschberger, L. L., Hwang, J., Coloso, R. M., and Stipanuk, M. H. (2007) Discovery and characterization of a second mammalian thiol dioxygenase, cysteamine dioxygenase. *J. Biol. Chem. 282*, 25189–25198.

(4) Stipanuk, M. H., Dominy, J. E., Ueki, I., and Hirschberger, L. L. (2008) Measurement of cysteine dioxygenase activity and protein abundance. *Curr. Protoc. Toxicol.*

(5) Stipanuk, M. H., Dominy, J. E., Lee, J.-I., and Coloso, R. M. (2006) Mammalian cysteine metabolism: new insights into regulation of cysteine metabolism. *J. Nutr. 136*, 1652S-1659S.

(6) Stipanuk, M. H. (1986) Metabolism of sulfur-containing amino acids. *Annu. Rev. Nutr. 6*, 179–209.

(7) Shen, D., Tian, L., Yang, F., Li, J., Li, X., Yao, Y., Lam, E. W. F., Gao, P., Jin, B., and Wang, R. (2021) ADO/hypotaurine: a novel metabolic pathway contributing to glioblastoma development. *Cell Death Discov. 7*.

(8) Schaffer, S. W., Ju Jong, C., KC, R., and Azuma, J. (2010) Physiological roles of taurine in heart and muscle. *J. Biomed. Sci. 17*, 101–163.

(9) Ripps, H., and Shen, W. (2012) Review: Taurine: A "very essential" amino acid. *Mol. Vis. 18*, 2673–2686.

(10) Masson, N., Keeley, T. P., Giuntoli, B., White, M. D., Lavilla Puerta, M., Perata, P., Hopkinson, R. J., Flashman, E., Licausi, F., and Ratcliffe, P. J. (2019) Conserved N-terminal cysteine dioxygenases transduce responses to hypoxia in animals and plants. *Science (80-.). 365*, 65–69.

(11) White, M. D., Klecker, M., Hopkinson, R. J., Weits, D. A., Mueller, C., Naumann, C., O'Neill, R., Wickens, J., Yang, J., Brooks-Bartlett, J. C., Garman, E. F., Grossmann, T. N., Dissmeyer, N., and Flashman, E. (2017) Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat. Commun. 8*.

(12) Lee, M. J., Tasaki, T., Moroi, K., An, J. Y., Kimura, S., Davydov, I. V., and Kwon, Y. T. (2005) RGS4 and RGS5 are in vivo of the N-end rule pathway. *Proc. Natl. Acad. Sci. U. S. A. 102*, 15030–15035.

(13) Tamirisa, P., Blumer, K. J., and Muslin, A. J. (1999) RGS4 inhibits G-protein signaling in cardiomyocytes. *Circulation 99*, 441–447.

(14) Stipanuk, M. H., Simmons, C. R., Karplus, P. A., and Dominy, J. E. (2011) Thiol dioxygenases: Unique families of cupin proteins. *Amino Acids 41*, 91–102.

(15) Fernandez, R. L., Dillon, S. L., Stipanuk, M. H., Fox, B. G., and Brunold, T. C. (2020)

Spectroscopic Investigation of Cysteamine Dioxygenase. *Biochemistry 59*, 2450–2458.

(16) Driggers, C. M., Kean, K. M., Hirschberger, L. L., Cooley, R. B., Stipanuk, M. H., and Karplus, P. A. (2016) Structure-Based Insights into the Role of the Cys–Tyr Crosslink and Inhibitor Recognition by Mammalian Cysteine Dioxygenase. *J. Mol. Biol. 428*, 3999–4012.

(17) Davies, C. G., Fellner, M., Tchesnokov, E. P., Wilbanks, S. M., and Jameson, G. N. L. (2014) The Cys-Tyr cross-link of cysteine dioxygenase changes the optimal ph of the reaction without a structural change. *Biochemistry 53*, 7961–7968.

(18) Wang, Y., Griffith, W. P., Li, J., Koto, T., Wherritt, D. J., Fritz, E., and Liu, A. (2018) Cofactor Biogenesis in Cysteamine Dioxygenase: C−F Bond Cleavage with Genetically Incorporated Unnatural Tyrosine. *Angew. Chemie Int. Ed. 57*, 8149–8153.

(19) White, M. D., Carbonare, L. D., Puerta, M. L., Iacopino, S., Edwards, M., Dunne, K., Pires, E., Levy, C., McDonough, M. A., Licausi, F., and Flashman, E. (2020) Structures of Arabidopsis thaliana oxygen-sensing plant cysteine oxidases 4 and 5 enable targeted manipulation of their activity. *Proc. Natl. Acad. Sci. U. S. A. 117*, 23140–23147.

(20) Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr. D66*, 125–132.

(21) Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr. 67*, 293–302.

(22) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr. 66*, 213–221.

(23) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.

(24) Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res. 46*, W296–W303.

(25) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr. 66*, 486–501.

(26) Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr. 68*, 352–367.

(27) The PyMOL Molecular Graphics System, Version 2.0 Schrodinger, LLC.

(28) Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindah, E. (2015) Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1–2*, 19–25.

(29) Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Migues, A. N., Bickel,

J., Wang, Y., Pincay, J., Wu, Q., and Simmerling, C. (2020) Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput. 16*, 528–552.

(30) D.A. Case, Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., T.E. Cheatham, I., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Jhala, K. K.-, Kovalenko, A., Krasny, R., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Man, V., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., Onufriev, A., Pan, F., Pantano, S., Qi, R., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Skrynnikov, N., Smith, J., Swails, J., Walker, R. C., Wang, J., Wilson, L., Wolf, R. M., Wu, X., York, D. M., and Kollman, P. A. (2020) Amber 2020. University of California, San Francisco.

(31) Li, P., and Merz, K. M. (2016) MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model. 56*, 599–604.

(32) Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem. 91*, 6269–6271.

(33) Trott, O., and Olson, A. J. (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31*, NA-NA.

(34) Frisch, M. J., Trucks, G. W., Schlegel, H. E., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomberts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., J., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., and Fox, J. D. (2016) Gaussian 16, Revision C.01. *Gaussian, Inc., Wallingford CT,*.

(35) Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys. 98*, 5648.

(36) Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B 37*, 785–789.

(37) Schäfer, A., Huber, C., and Ahlrichs, R. (1994) Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys. 100*, 5829–5835.

(38) Hehre, W. J., Ditchfield, K., and Pople, J. A. (1972) Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys. 56*, 2257–2261.

(39) Bayly, C. I., Merz, K. M., Ferguson, D. M., Cornell, W. D., Fox, T., Caldwell, J. W., Kollman, P. A., Cieplak, P., Gould, I. R., and Spellmeyer, D. C. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc. 117*,

5179–5197.

(40) Tchesnokov, E. P., Fellner, M., Siakkou, E., Kleffmann, T., Martin, L. W., Aloi, S., Lamont, I. L., Wilbanks, S. M., and Jameson, G. N. L. (2015) The cysteine dioxygenase homologue from Pseudomonas aeruginosa is a 3-mercaptopropionate dioxygenase. *J. Biol. Chem. 290*, 24424–24437.

(41) Aloi, S., Davies, C. G., Karplus, P. A., Wilbanks, S. M., and Jameson, G. N. L. (2019) Substrate Specificity in Thiol Dioxygenases. *Biochemistry* 2398–2407.

(42) Trott, O., and Olson, A. J. (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31*, NA-NA.

(43) Wang, Y., Davis, I., Chan, Y., Naik, S. G., Griffith, W. P., and Liu, A. (2020) Characterization of the nonheme iron center of cysteamine dioxygenase and its interaction with substrates. *J. Biol. Chem. 295*, 11789–11802.

(44) Fernandez, R. L., Juntunen, N. D., Fox, B. G., and Brunold, T. C. Spectroscopic Investigation of Iron (III) Cysteamine Dioxygenase in the Presence of Substrate (Analogues): Implications for the Nature of Substrate-Bound Reaction Intermediates. *J. Biol. Inorg. Chem.* **2021,** *in press***.**

(45) Sehnal, D., Vařeková, R. S., Berka, K., Pravda, L., Navrátilová, V., Banáš, P., Ionescu, C. M., Otyepka, M., and Koča, J. (2013) MOLE 2.0: Advanced approach for analysis of biomacromolecular channels. *J. Cheminform. 5*, 1–13.

(46) Jones, J. C., Banerjee, R., Shi, K., Aihara, H., and Lipscomb, J. D. (2020) Structural Studies of the Methylosinus trichosporium OB3b Soluble Methane Monooxygenase Hydroxylase and Regulatory Component Complex Reveal a Transient Substrate Tunnel. *Biochemistry 59*, 2946–2961.

(47) Banerjee, R., and Lipscomb, J. D. (2021) Small-Molecule Tunnels in Metalloenzymes Viewed as Extensions of the Active Site. *Acc. Chem. Res. 54*, 2185–2195.

(48) Van, V., and Smith, A. T. (2020) ATE1-Mediated Post-Translational Arginylation Is an Essential Regulator of Eukaryotic Cellular Homeostasis. *ACS Chem. Biol.*

(49) Tian, W., and Skolnick, J. (2003) How well is enzyme function conserved as a function of pairwise sequence identity? *J. Mol. Biol. 333*, 863–882.

(50) Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta - Proteins Proteomics 1854*, 1019–1037.

(51) Dominy, J. E., Hwang, J., Guo, S., Hirschberger, L. L., Zhang, S., and Stipanuk, M. H. (2008) Synthesis of amino acid cofactor in cysteine dioxygenase is regulated by substrate and represents a novel post-translational regulation of activity. *J. Biol. Chem. 283*, 12188–12201.

Chapter 7

Using Light to Study Vitamin B_{12}

This chapter was written in fulfillment of the requirements for Wisconsin Initiative for Science Literacy Graduate Thesis Award for Communicating Research to the Public.

Chapter 7: Using Light to Study Vitamin B_{12}

7.1. Preface

Working as a teaching assistant (TA) for introductory chemistry courses has really shaped how I think about communicating my research. TAs have a lot of information they need to convey and a limited time to do so. Each course faces certain questions: What are the most important topics to cover in a semester? How can we best prepare students both for their careers and for their lives more generally? Over the past few years, I've found that there are two overall messages that I want to convey. First, chemistry is important. You've probably heard this before, if you've taken a high school chemistry class, but it's worth emphasizing. I've taught courses that covered subjects like climate change, air quality, nuclear energy, ethanol tax credits, and the chemical elements found in smartphones, to name a few. Chemistry is an excellent tool for understanding how our society functions, and learning the basics can go a long way to demystifying topics relevant to our daily lives.

The second message is that chemistry can be fun! My favorite part of teaching chemistry labs is watching students figure out what parts of chemistry they enjoy. Sometimes it's the satisfied feeling of solving a tricky problem, and other times it's a more specific project or technique that we cover. In chemistry, it can be helpful to have a knack for memorization and complicated math, but it's equally valid if you get into the field because you like vivid colors or dry ice or lighting things on fire. I was in the latter camp, and I ended up in a great research group for that. I get to work, for example, with lasers and powerful magnets, and those things never stopped feeling cool—although my childhood self would probably picture something quite different if she heard that I now work with force fields (a computational modeling term!). I hope that this chapter conveys some of the joy I felt doing this research.

I owe many people thanks for the existence of this chapter. First, I want to thank my Aunt Heidi, who helped me discover both the wonders of science and the importance of teaching it well. When I wrote this chapter, I thought a lot about a book she gave me, a blend of history and organic chemistry that delves into some of the molecules that shaped history, from spices and vitamins to antimalarial drugs.[1] It was the first chemistry book I really loved, and it's a great example of how learning a bit of chemistry can teach you a lot about the world. A big thank you also to my parents and sister, who have been a (mostly) willing audience to my chemistry ramblings these last few years. I hope this chapter clarifies some of those ramblings! Finally, I owe an enormous thanks to the Wisconsin Initiative for Science Literacy (WISL) at UW-Madison, and particularly to Professor Bassam Shakhashiri, Elizabeth Reynolds, and Cayce Osborne for their help and guidance as I worked on this chapter. I'm very grateful for the opportunity to learn how to communicate my science more effectively.

7.2. Vitamin B_{12}

It might surprise you to know how many different metals can be found in the human body. Some are required in abundance, like the calcium in your bones or the sodium that keeps your muscles and nerves working. Others are required in smaller amounts, such as copper, nickel, iron, and cobalt. Long before early humans started using metals for tools and weapons and jewelry, nature put those metals to work inside living things.

Many of the chemical reactions required to keep your body functioning are facilitated by molecules called enzymes, biological catalysts that speed up chemical reactions without being consumed themselves. Enzymes are built from mass-produced molecular building blocks that are strung together and folded into three-dimensional shapes. The atoms involved are primarily non-metal elements that are found in relative abundance in both the human body and the world around

us: carbon, hydrogen, nitrogen, oxygen, and sulfur. However, those elements are not always sufficient to accomplish particularly difficult or unusual reactions, and that is where metals come in. Metalloenzymes incorporate metal atoms into the mix, using the unusual chemical properties of metals to perform different types of reactions. A classic example of a metalloenzyme-catalyzed reaction is nitrogen fixation, through which bacteria convert inert atmospheric nitrogen (N_2) into ammonia (NH_3), a form of nitrogen that other living species can use to build more complicated molecules.[2]

Some metalloenzymes integrate only a single metal atom into their structures, but in other cases, the metal atom is packaged into a larger chemical structure, such as the cobalt-containing molecule known as vitamin B_{12}. Vitamin B_{12} has the distinction of being both immensely complicated and vitally important to human health. The term "pernicious anemia" was coined in the mid-19th century to describe a disease characterized by fatigue, weakness, and a slow decline in general health, usually culminating in serious digestive and neurological symptoms.[3] At that time, the disease was invariably a death sentence. In 1925, Frieda Robscheit-Robbins and George Whipple reported the successful treatment of anemic dogs with cooked liver,[4] kicking off a decades-long quest to isolate the factor producing these effects. His work on dietary treatments of anemia later earned Whipple, along with George Minot and William Murphy, the 1934 Nobel Prize in Physiology and Medicine. We now understand pernicious anemia to arise from problems absorbing or transporting vitamin B_{12}, often as a result of autoimmune inflammation of the stomach.[5]

Whipple, Minot, and Murphy were the first prominent B_{12} scientists to win the Nobel prize, but not the last; Dorothy Hodgkin headed the 1956 discovery of the molecule's chemical structure,[6] using a technique for which she later received a Nobel Prize in Chemistry. This work provided the

first blueprint to this essential vitamin, which turned out to be, as the paper described it, “very beautifully composed.” For one thing, vitamin B_{12} is quite large for a vitamin! Figure 7.1 compares the complexity of its structure to the relative simplicity of some of the other vitamins required in your diet. At the molecule’s center, a cobalt (Co) atom is enclosed within a mostly flat ring of carbon-carbon and carbon-nitrogen bonds. Chains of atoms strung together trail off the sides of the ring, and one of those side chains loops around the bottom of the ring to bind again to the cobalt atom. Binding to the top of the cobalt atom is a group denoted “R”, which simply means that there are different possible structures in that position. While *vitamin* B_{12} refers to one specific structure (specifically, when R is C≡N, a carbon atom triple-bonded to a nitrogen atom), B_{12} more generally refers to the larger family of molecules with different groups in the R position. Notably, for several of the R groups, the molecule contains a bond between cobalt and a carbon atom, a very rare example of a metal-carbon bond in nature!

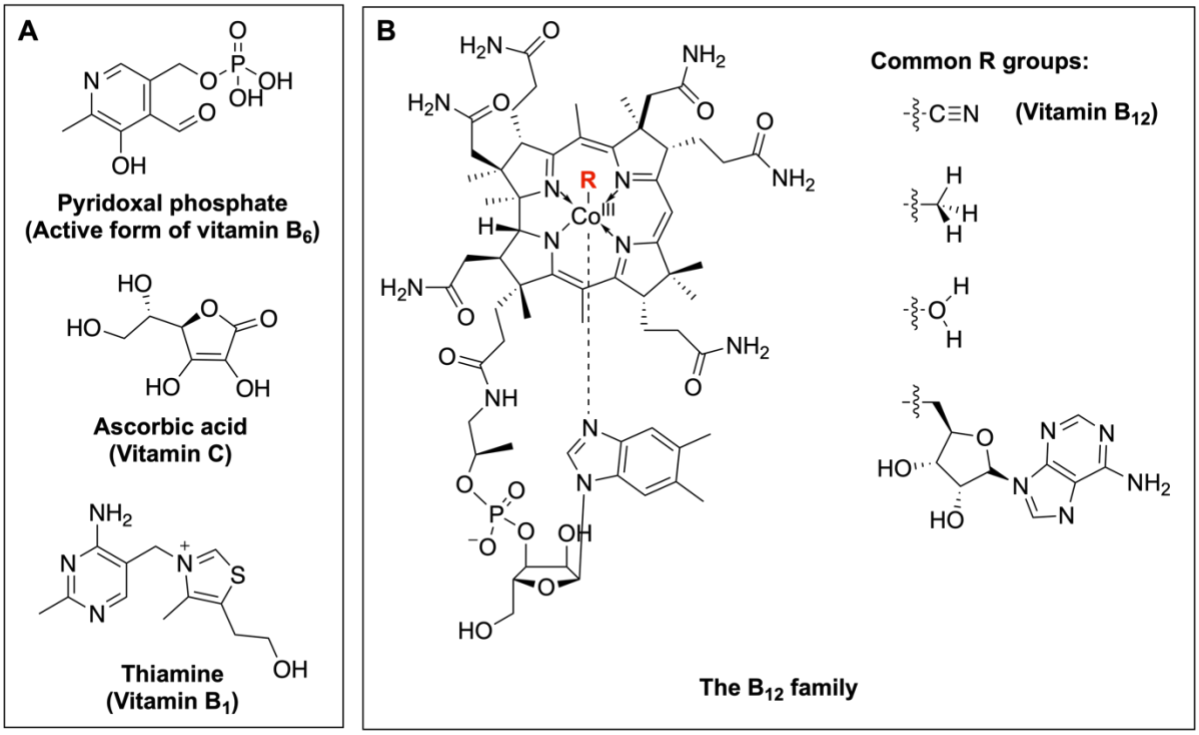


Figure 7.1. The chemical structures of (A) several common vitamins and (B) the B_{12} family. In drawings like these, each line represents a chemical bond, and each vertex or end of a line represents a carbon atom. Hydrogen atoms attached to carbon atoms are omitted. "R" (labeled in red) is used to denote the part of the B_{12} molecule that can have several different structures.

This variability in structure contributes to the B_{12} family's versatility. A true jack of all trades, B_{12} helps enzymes perform many essential chemical reactions for species across all areas of life.[7] Its work can be thought of like this: if a living organism was a factory, then its enzymes would be factory workers and B_{12} would be a tool that they use, useful for a diverse assortment of jobs in all sorts of different factories. Two enzymes in the human body use B_{12} for their work, and impairment of their function has disastrous health effects, including genetic disorders like methylmalonyl aciduria and homocystinuria.[8,9] However, the overall role of the B_{12} family in biology is much broader. Even now, almost a century after Robscheit-Robbins and Whipple published their seminal work, scientists are discovering new facets of the role that vitamin B_{12} holds in the natural world. Recent work, for example, uncovered that a member of the B_{12} family serves as a light sensor in certain bacteria.[10] Chemists are particularly interested in the fact that B_{12}

can accomplish such a wide array of fast, controlled, and specialized chemical reactions, which is the sort of skillset that chemists would love to recreate without nature's help.

Much of my research has focused on enzymes that either build B_{12} or use B_{12} to perform chemical reactions. The project described here, however, focuses on vitamin B_{12} itself, and on expanding the tools (both lab-based and computer-based) that scientists can use to understand how it works. This started as a pandemic project, some computational work I could perform at my sparsely decorated basement desk during the early COVID lockdowns, and I had no idea that it would become such a fascinating and rewarding journey. I hope that this chapter can show you a glimpse of vitamin B_{12} as I see it, as one of nature's most enigmatic and beautiful inventions.

7.3. The building blocks of the universe: atoms and atomic structure

To understand why molecules work the way they do, it is useful to talk briefly about what they are made out of. To that end, there are a handful of chemistry principles worth mentioning here:

1. The most basic, fundamental unit of matter in chemistry is the atom. All matter is built out of atoms.
2. Within an atom, there are three "subatomic" particles: protons, neutrons, and electrons. Protons and neutrons are packed into the tiny, dense core of the atom, called the nucleus, while the electrons travel around them. Although electrons weigh only a minuscule fraction of the mass of the nucleus, the electron cloud takes up most of the space.
3. The specific location of an electron within the electron cloud is fundamentally impossible to know, but probable locations are defined by specific three-dimensional shapes called orbitals. We therefore describe electrons as being "stored" in orbitals, which have specific shapes associated with them.

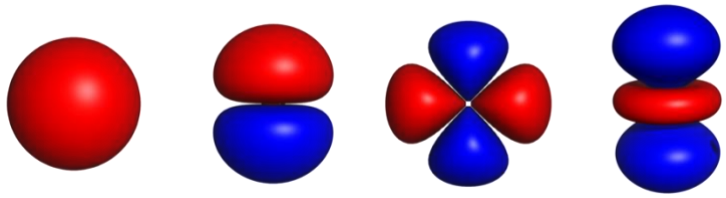


Figure 7.2. Several common types of atomic orbitals.

4. Chemical bonds between atoms are built by the sharing of electrons. When a bond forms, *atomic orbitals* from different atoms combine to form *molecular orbitals*. This means that the size and shape of atomic orbitals define how an atom reacts with other atoms.

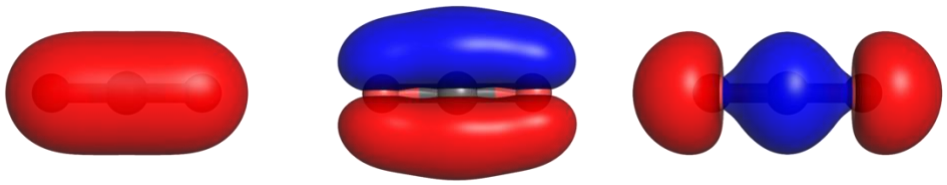


Figure 7.3. Three molecular orbitals of the molecule carbon dioxide (CO_2). The structure of CO_2 is linear, with central carbon double-bonded to two oxygens (O=C=O).

5. Bonds are flexible, albeit to varying degrees. A useful metaphor is to think of them as coiled springs that can be distorted with sufficient force.
6. Molecules have "ground" states that describe the most stable way that atoms and electrons can be positioned, but adding energy can briefly push the system into an "excited" state. This includes electronic excited states (wherein an electron is bumped from a stable molecular orbital into a less stable one) and vibrational excited states (wherein distances and angles between atoms are distorted).

7.4. Spectroscopy: using light as a tool

The arrangement of atoms and electrons defines the way a molecule behaves, but atomic structure is difficult to study directly. I rely on a tool called spectroscopy, which refers to techniques that measure how materials interact with or emit light.

Understanding the full potential of this idea requires expanding the definition of light beyond what is visible to the human eye. First, consider white light, which is a combination of all colors of visible light. This is most readily apparent in phenomena like rainbows and the refraction of light through glass prisms, in which the different speeds at which colors travel through water droplets or glass causes the light to split into its full spectrum of colors. The colors of the rainbow differ in energy, with violet being highest in energy and red being lowest in energy. In this context, energy is not the same thing as intensity. One red laser pointer is lower in *intensity* than one hundred red laser pointers, but an equally intense blue laser pointer is greater in *energy* than the red one.

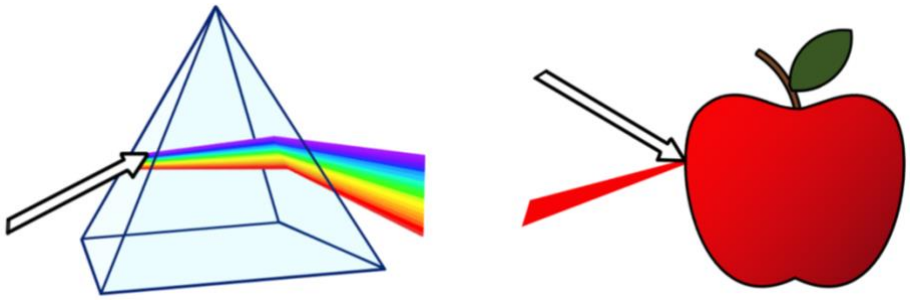


Figure 7.4. Left: White light refracting through a glass prism to produce a rainbow. Right: White light shining on a red apple.

This brings us to the next important fact about light, which is that visible light is one part of a larger whole. Much like blue light is a portion of the visible light spectrum, visible light is a portion of the electromagnetic spectrum. This spectrum covers an enormous range of energies, with innocuous radio waves at one end and Hulk-inducing gamma rays at the other. Different regions of the electromagnetic spectrum interact with matter differently, and these interactions can tell us a lot about different aspects of a molecule's structure or properties. The two forms of spectroscopy used in this project deal with three regions in particular: ultraviolet (UV), visible light, and infrared (IR).

7.5. Experimental and computational investigation of vitamin B_{12}

Ultraviolet-visible (UV-Vis) spectroscopy measures how much UV and visible light is absorbed by a sample, and specifically, how much is absorbed for different energies of light. The human eye actually makes a decent attempt at this form of spectroscopy, registering which visible colors are reflected by an object and which colors are not. When white light hits a red apple, for example, all colors are absorbed except for red, and the reflected red light gives the apple its color (Figure 7.4). UV-Vis spectroscopy uses a similar idea, but with much greater precision and for a large range of energies, expanding into the ultraviolet region.

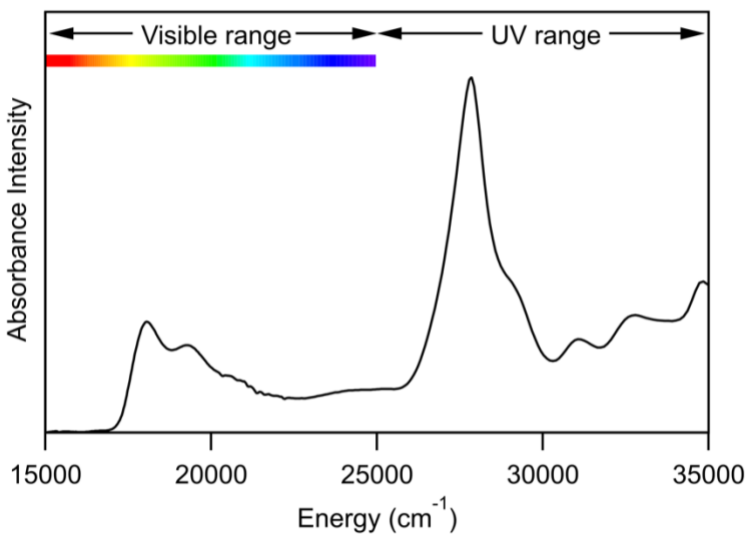


Figure 7.5. Left: UV-Vis spectrum of vitamin B_{12}. Right: Vitamin B_{12} dissolved in water.

Why are only some colors absorbed, and why do we care? Molecules cannot absorb just any amount of energy, only specific amounts that match the energy gaps between the ground state and an excited state. UV and visible light are in the correct range to cause electronic excitations, wherein an electron is briefly bumped from a stable molecular orbital into a less stable one. Measuring the energies of the light that is absorbed provides useful information about the molecule's orbitals, which in turn conveys information about the molecule's bonds and therefore its properties. Consider, for example, the UV-Vis spectrum of vitamin B_{12} (Figure 7.5). Like a red apple, vitamin B_{12} absorbs red light the least, causing it to appear red in color. Decades of study

have demonstrated that the peaks at the low-energy end of the spectrum are very sensitive to changes in the cobalt atom's environment.[11] We can tell, for example, if an enzyme binds B_{12} in a way that changes the length of B_{12}'s Co–C bond or displaces the side chain that binds at the lower face of the ring. This is valuable information to know about an enzyme! To return to the body-as-a-factory metaphor, watching how somebody holds a tool is a great first step to understanding how to use it yourself.

The other forms of spectroscopy used for this project are in another region of the electromagnetic spectrum: the infrared (IR) range. IR radiation is lower in energy than visible light, not enough energy to cause electronic excitations, but enough to cause more vigorous molecular vibrations. In this context, vibrations refer to certain combinations of movements that a molecule can undergo, such as stretching and bending. Once again, the light absorbed needs to have the correct energy to match a specific excitation (in this case, a vibrational motion). Figure 7.6 shows computer-simulated spectra for two forms of vibrational spectroscopy, IR and Raman spectroscopies, for the molecule CO_2. Both types of spectroscopy involve the infrared region, but their differing experimental setups cause them to measure different types of vibrational motion. The common convention for IR spectra is to use transmittance of light on the y axis rather than absorbance, which essentially means that the peaks are upside down.

Each peak in Figure 7.6 is labeled with the vibrational motion associated with it. The peak at 1399 cm^{-1} in the Raman spectrum, for example, corresponds to a symmetric stretching motion of the CO_2 molecule, wherein the C=O bonds lengthen and shorten in unison. The location of that peak (i.e., the energy required to cause a symmetric stretching motion) is dependent on physical attributes like atom mass and bond strength. If you think of a bond as a coiled spring connecting two atoms, bond strength is like the stiffness of the spring; stronger bonds require more force to

distort, while weaker bonds can stretch and break more easily. Vibrational spectroscopies like IR and Raman provide a very direct way to determine the strength of a chemical bond—but only if you know which peaks on the spectrum correspond to vibrational motions involving that bond.

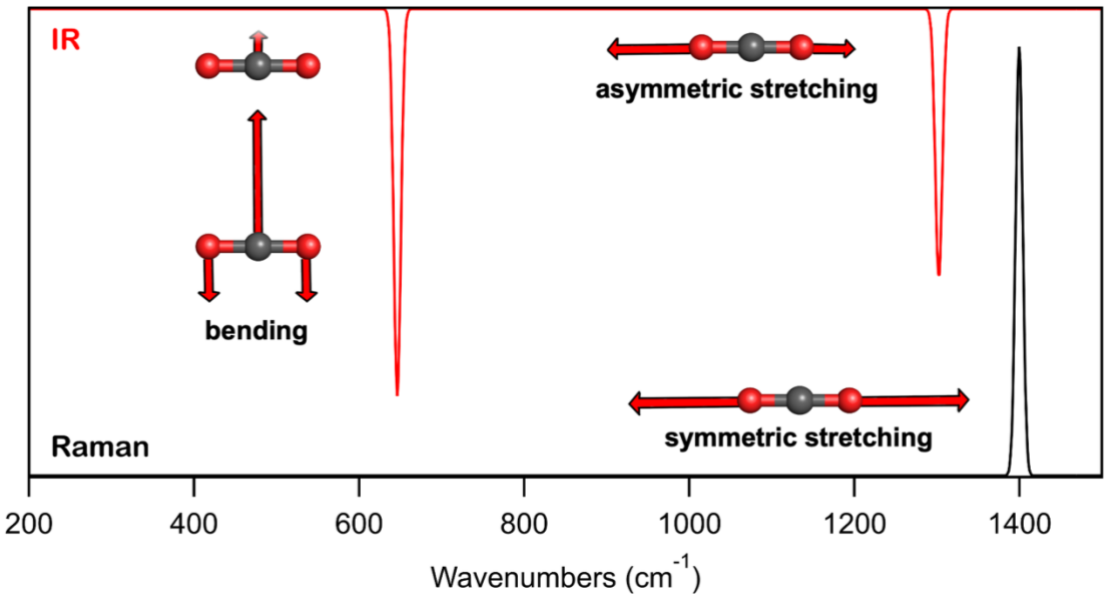


Figure 7.6. Calculated IR (red) and Raman (black) spectra of CO_2, with labeled vector representations of the vibrational mode(s) associated with each feature. CO_2's ability to absorb IR radiation is what makes it such a potent greenhouse gas.

Rather than using traditional Raman spectroscopy, my research uses something called resonance Raman (rRaman) spectroscopy, in which a laser is used to irradiate the sample with a particular color of visible or UV light while the Raman spectrum is being collected. The laser increases the signal for any peaks that involve vibrations of whichever part of the sample is absorbing the light of the laser. In a B_{12}-containing sample, the brightly colored B_{12} molecule absorbs light much more strongly than its surroundings, whether those surroundings are an enzyme that it is bound to or simply the water in which it is dissolved. This is especially useful for experiments in which B_{12} is bound to an enzyme, as it guarantees that the features of the rRaman spectrum arise from B_{12} itself and not from the thousands of other atoms undergoing their own vibrations.

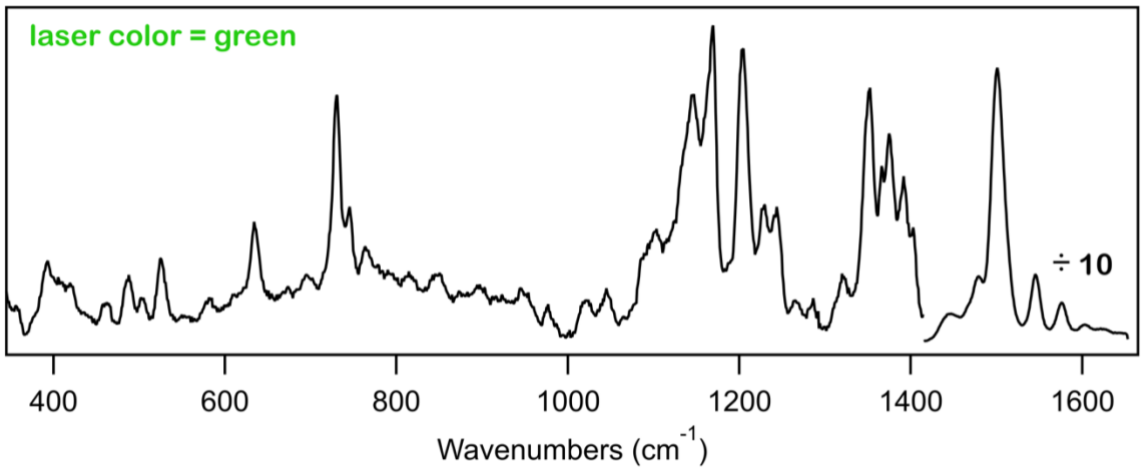


Figure 7.7. rRaman spectrum of vitamin B_{12} obtained using a green laser (energy ~19,400 cm^{-1}). Features above 1400 cm^{-1} are scaled down by a factor of 10 for ease of visualization.

The rRaman spectrum of vitamin B_{12} is incredibly cool to look at, but equally daunting to interpret. Vibrational modes are easier to categorize in a molecule like CO_2, as there are only so many ways a linear, three-atom molecule can vibrate (four ways, to be exact). The number of vibrational modes increases with increasing size, however, and vitamin B_{12} has 181 atoms. Some peaks will overlap, some will be too weak to notice, and all of this with a technique where even humidity can mess with your measurements. How do we figure out what rRaman feature corresponds to a particular vibrational motion?

One way is to swap out one of the atoms of interest for an atom of the same element but with a different mass (called an isotope), which causes rRaman peaks associated with that atom to shift to the left on the spectrum (lower in energy). Subtracting the new spectrum from the original makes it easy to identify any changes that have occurred. Others have used this technique successfully with several vitamin B_{12} derivatives in order to identify Co–C stretching peaks, which are of great interest given that the strength of the Co–C bond is directly relevant to function of those molecules.[12,13] However, until my work, nobody had ever found a feature like that in the rRaman spectrum of vitamin B_{12}.

Another useful tool for interpreting spectra is computer modeling. The equations that describe the shape of an orbital are called “wavefunctions,” due to the (probably unintuitive) fact that electrons exhibit wave-like behavior. Equations describing the properties of a molecule are functions that depend on the wavefunctions of its atoms and are therefore functionals (i.e., functions of other functions). These functionals are the basis of a computational technique that is used to predict the properties and reactivity of chemicals. The main challenge is that the math gets complicated *really* quickly. There is no single correct way to approach the trade-off between accuracy and computational cost. As a result, there are hundreds of functionals that have been developed for use in this type of calculation, each with a slightly different approach to the math. Choosing the appropriate functional for a given molecule requires testing and a comparison to experimental data.

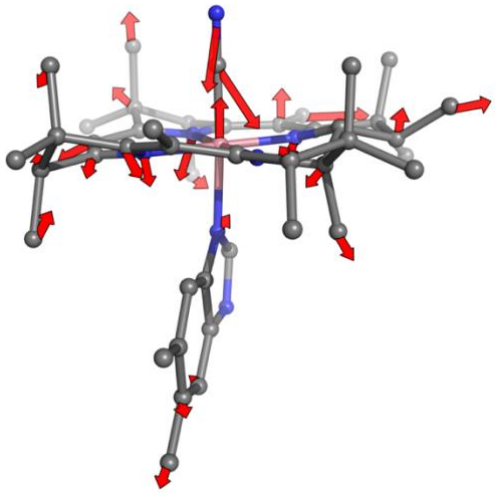


Figure 7.8. Vector representation of a computed rRaman vibration that involves stretching of the Co–C bond and causes a noticeable isotopic shift in the computed rRaman spectrum.

For this project, I evaluated the performance of two popular functionals at simulating UV-Vis and rRaman spectra for B_{12}-containing models. Although the proper choice of functional had been explored extensively,[14–16] earlier analyses were limited by the fact that traditional techniques were able to simulate electronic transitions and vibrational transitions separately, but not together.

Modeling aspects like the shape of UV-Vis peaks or the effect of laser enhancement on Raman spectra require that the calculation consider both types of excitations simultaneously. I was able to achieve this using a recently developed software tool,[17] in conjunction with the two different functionals. The differences in their performance will not be discussed here, but there was something equally valuable that I found about what the two functionals had in common: both predicted that I should be able to identify a Co–C stretching peak in the rRaman spectra of vitamin B_{12} that shifts upon isotopic substitution!

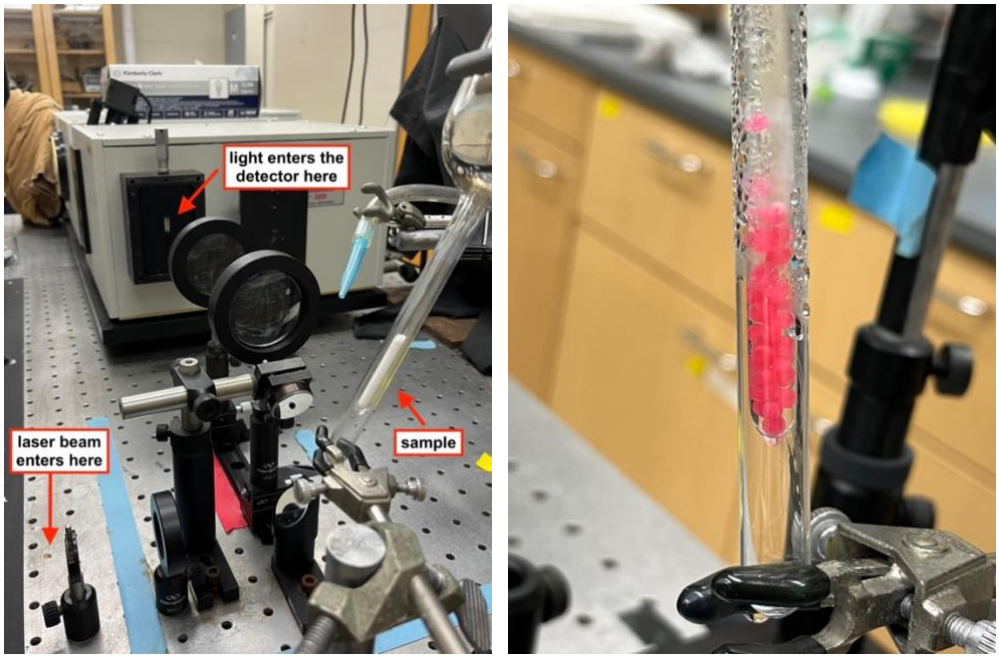


Figure 7.9. Left: experimental rRaman set-up. Right: Vitamin B_{12} pellets in the rRaman sample holder.

So, I gave it a try. Collecting high quality rRaman data is both a science and an art, as it involves very careful adjustments of the experimental set-up. A laser at one end of a table generates a laser beam, which crisscrosses the table through a sequence of mirrors and irises to ensure that a single powerful beam hits the sample that rests in a quartz sample holder. The light scattered from the sample then passes through two lenses to reach the detector. I prepared rRaman samples in the form of frozen sample pellets (Figure 7.9), using a syringe to inject the sample one drop at a time

into liquid nitrogen, which has a temperature of -320 °F. If you notice a resemblance between this sample and a particular frozen treat, you're not wrong—this is similar to how Dippin' Dots are made! Making samples in the form of pellets means that they can be poured directly into the sample holder, rather than needing to freeze the sample in a glass tube. Glass warps the baseline of rRaman spectra, making it difficult to distinguish signal from noise, so removing the glass significantly improves spectra.

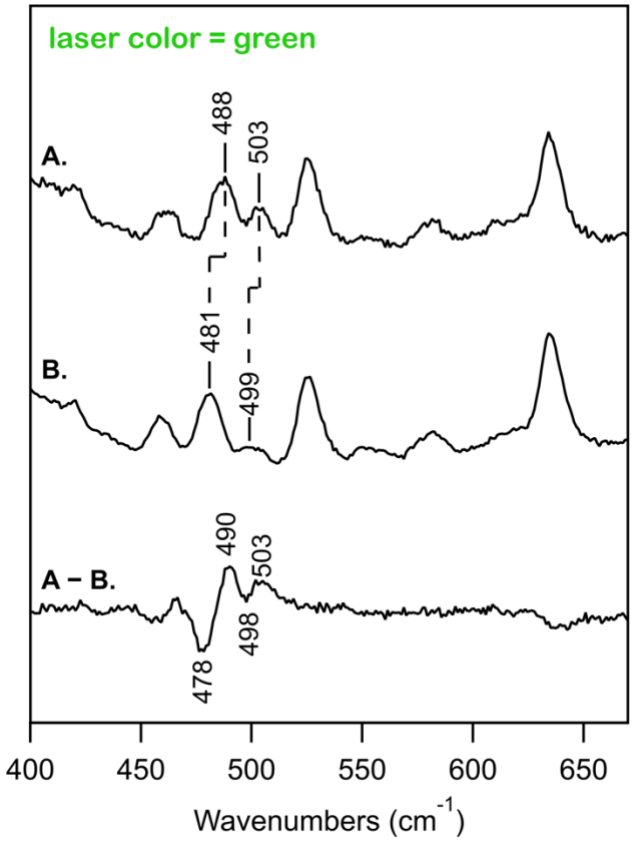


Figure 7.10. rRaman spectra of vitamin B_{12} using a green laser, where (A) corresponds to normal vitamin B_{12}, (B) corresponds to a form of vitamin B_{12} containing a heavier isotope of carbon bound to Co, and (C) is the difference between the two.

I used this approach to collect rRaman spectra for both regular vitamin B_{12} and vitamin B_{12} with a heavier isotope of carbon bound to the Co ion, and I was excited to find that there was a noticeable downshift of two rRaman peaks at 488 and 503 cm^{-1}. Finally, the elusive Co–C bond stretching mode of vitamin B_{12} could be identified! Using computational modeling and some math,

I found that the peak at 488 cm^{-1} corresponds to a Co–C bond stretch and the peak at 503 cm^{-1} corresponds to a Co–C≡N bend (involving the nitrogen atom in vitamin B_{12}'s "R" group). This allowed me to calculate the strength of the Co–C bond, which turns out to be substantially stronger than the Co–C bond in other members of the B_{12} family.

7.6. Conclusions

The results obtained here represent a valuable step forward in our understanding of how to model and interpret spectra of vitamin B_{12}. For one thing, the new computational tool that I tested turned out to provide accurate and valuable predictions of experimental spectra, suggesting that it can be applied to future B_{12} studies. Being able to simulate the combination of electronic and vibrational excitations will facilitate more accurate interpretations of UV-Vis and rRaman spectra of B_{12}-containing systems. I was also able to identify the Co–C stretching peak in vitamin B_{12}'s rRaman spectrum, a piece that has been missing from the puzzle for a long time, which enabled me to calculate the strength of that bond. I feel very lucky to have uncovered such a fundamental aspect of vitamin B_{12}'s nature, and I'm proud to contribute to the foundation upon which future B_{12} researchers can build their work.

References

(1) Le Couteur, P.; Burreson, J. *Napoleon's Buttons: How 17 Molecules Changed History*; Penguin Publishing Group: New York, 2003.

(2) Hoffman, B. M.; Lukoyanov, D.; Yang, Z. Y.; Dean, D. R.; Seefeldt, L. C. Mechanism of Nitrogen Fixation by Nitrogenase: The next Stage. *Chemical Reviews*. American Chemical Society April 23, 2014, pp 4041–4062. https://doi.org/10.1021/cr400641x.

(3) Graner, J. L. Addison, Pernicious Anemia and Adrenal Insufficiency. *Can. Med. Assoc. J.* **1985**, *133* (9), 855-857+880.

(4) Robscheit-Robbins, F. S.; Whipple, G. H. Blood Regeneration in Severe Anemia: Favorable Influence of Liver, Heart and Skeletal Muscle in Diet. *Am. J. Physiol. Content* **1925**, *72* (3), 408–418. https://doi.org/10.1152/ajplegacy.1925.72.3.408.

(5) Toh, B. H. Autoimmune Gastritis and Pernicious Anemia. In *The Autoimmune Diseases*; Academic Press, 2019; pp 833–847. https://doi.org/10.1016/B978-0-12-812102-3.00044-0.

(6) Hodgkin, D. C.; Kamper, J.; MacKay, M.; Pickworth, J.; Trueblood, K. N.; White, J. G. Structure of Vitamin B_{12}. *Nature* **1956**, *178* (4524), 64–66. https://doi.org/10.1038/178064a0.

(7) Banerjee, R.; Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **2003**, *72*, 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828.

(8) Watkins, D.; Rosenblatt, D. S. Inborn Errors of Cobalamin Absorption and Metabolism. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **2011**, *157* (1), 33–44. https://doi.org/10.1002/ajmg.c.30288.

(9) Gherasim, C.; Lofgren, M.; Banerjee, R. Navigating the B_{12} Road: Assimilation, Delivery, and Disorders of Cobalamin. *Journal of Biological Chemistry*. May 10, 2013, pp 13186–13193. https://doi.org/10.1074/jbc.R113.458810.

(10) Bridwell-Rabb, J.; Drennan, C. L. Vitamin B_{12} in the Spotlight Again. *Curr. Opin. Chem. Biol.* **2017**, *37*, 63–70. https://doi.org/10.1016/j.cbpa.2017.01.013.

(11) Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914. https://doi.org/10.1021/ja029328d.

(12) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Force Constants from Resonance Raman Spectra of Alkylcobalamins: Insensitivity to Dimethylbenzylimidazole Coordination. *Inorganica Chim. Acta* **1998**, *270* (1), 392–398. https://doi.org/https://doi.org/10.1016/S0020-1693(97)05994-X.

(13) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co–C Bond Activation in B_{12}-Dependent Enzymes: Cryogenic Resonance Raman Studies of Methylmalonyl-Coenzyme A Mutase. *J. Am. Chem. Soc.* **1999**, *121* (30), 7063–7070. https://doi.org/10.1021/ja982753f.

(14) Kuta, J.; Patchkovskii, S.; Zgierski, M. Z.; Kozlowski, P. M. Performance of DFT in Modeling Electronic and Structural Properties of Cobalamins. *J. Comput. Chem.* **2006**, *27* (12), 1429–1437. https://doi.org/10.1002/jcc.20454.

(15) Jensen, K. P.; Ryde, U. Cobalamins Uncovered by Modern Electronic Structure Calculations. *Coordination Chemistry Reviews*. Elsevier March 1, 2009, pp 769–778. https://doi.org/10.1016/j.ccr.2008.04.015.

(16) Brunold, T. C.; Conrad, K. S.; Liptak, M. D.; Park, K. Spectroscopically Validated Density Functional Theory Studies of the B_{12} Cofactors and Their Interactions with Enzyme Active Sites. *Coordination Chemistry Reviews*. March 2009, pp 779–794. https://doi.org/10.1016/j.ccr.2008.09.013.

(17) Petrenko, T.; Neese, F. Analysis and Prediction of Absorption Band Shapes, Fluorescence Band Shapes, Resonance Raman Intensities, and Excitation Profiles Using the Time-Dependent Theory of Electronic Spectroscopy. *J. Chem. Phys.* **2007**, *127* (16), 164319. https://doi.org/10.1063/1.2770706.

Chapter 8

Conclusions and Outlook

Chapter 8: Conclusions and Outlook

8.1. Overview of Research Progress

The work outlined in Chapter 2 explored the electronic and vibrational properties of CNCbl, combining electronic absorption (Abs) and resonance Raman (rR) spectroscopies with a recently developed computational tool capable of incorporating vibronic coupling into DFT-simulated spectra. This research addressed several ongoing conversations in the field of B_{12} research. The appropriate choice of functional for DFT studies of B_{12} excited states, for example, has been a matter of debate, with a particular focus on the relative performance of the hybrid functional B3LYP and the pure generalized gradient approximation (GGA) functional BP86.[1–4] There have also been lingering uncertainties over the assignment of CNCbl's spectral features, including the characteristic β band in its Abs spectrum. Incorporating vibronic coupling into computed spectra shed new light on these issues, as it enabled the simulation of Abs bandshapes and rR enhancement. When compared to experimental Abs and rR data, B3LYP consistently outperformed BP86. Only B3LYP reproduced the characteristic shape of the α/β region, with its single electronic transition and vibrational progression. Furthermore, B3LYP more accurately predicted rR spectra, with BP86 greatly overestimating the resonance enhancement of peaks in the low-frequency region. In addition to supporting the use of B3LYP for excited state calculations of CNCbl—a valuable insight for future theoretical studiees—these data were instrumental in guiding further experimental work, ultimately leading us to identify a peak in CNCbl's rR spectrum that arises from the elusive Co–C bond stretching mode.

The research described in Chapter 3 investigated the mechanism of Co–C bond activation in ethanolamine ammonia-lyase (EAL), an AdoCbl-dependent enzyme that catalyzes the conversion of ethanolamine to acetaldehyde and an ammonium ion. This study complemented

earlier work performed in this group on the AdoCbl-dependent enzymes glutamate mutase (GM) and methylmalonyl CoA mutase (MMCM). A key distinction between these enzymes is that MMCM and GM are Class I isomerases, which bind AdoCbl in a base-off/histidine (His)-on conformation whereby the DMB ligand is displaced by a protein-derived histidine, while EAL is a Class II isomerase, binding AdoCbl in a base-on conformation.[5] Our previous work determined that the His ligand is an essential component of the Co–C bond activation mechanism employed by MMCM and GM, as it grants these enzymes a measure of control over the cofactor's electronic structure.[6–8] Using magnetic circular dichroism (MCD) spectroscopy, these studies revealed that cofactor binding to MMCM and GM causes a uniform stabilization of the Co 3d orbitals of Co^{II}Cbl, thus promoting homolysis through stabilization of the post-homolysis Co^{II}Cbl product. However, as EAL and other Class I eliminases bind the cofactor in a base-on conformation, they must necessarily employ a different mechanism of Co–C bond activation. In this work, it was discovered that, in the presence of substrate, EAL distorts the adenosyl (Ado) moiety and Co axial bonds, thus activating the Co–C bond through destabilization of the AdoCbl "ground state".

The insights gained in our computational investigation of EAL demonstrate the utility of molecular dynamics (MD) calculations in the study of metalloenzymes. For this project, we incorporated MD calculations into our computational workflow to aid in the *in silico* reconstruction of the Co–C bond, which is cleaved in the X-ray crystal structure of EAL that we used to extract initial coordinates for our computational studies.[9] Early work on this project demonstrated that building this bond through less rigorous means (e.g., manually shifting the Ado moiety in PyMOL) introduced significant user bias, whereas MD simulations allowed us to explore the full conformational space of the Ado moiety. This approach may be similarly useful in computational studies of other AdoCbl-dependent enzymes, as the cleaved Co–C bond in the

crystal structure of EAL is a common occurrence; the sensitivity of the cofactor's Co–C bond to visible light and X-rays means that that there are few crystal structures of AdoCbl-dependent enzymes that contain an intact, native cofactor.[10] This method also proved useful in the study of cysteamine dioxygenase (ADO), where it was used to assess whether substrate binding or the formation of a critical disulfide bond produces large conformational changes of the enzyme (Chapter 6).

In Chapters 4 and 5, we investigated the catalytic cycle of PduO-type adenosyltransferases (ATRs), which catalyze the adenosylation of an AdoCbl precursor. Previous work from this group explored the mechanism by which PduO accomplishes the thermodynamically challenging Co^{II}Cbl → Co^{I}Cbl reduction, demonstrating that the formation of an unusual four coordinate (4C) ATR-bound Co^{II}Cbl species stabilizes the redox-active Co $3d_{z2}$ orbital so as to raise the reduction potential into the physiologically accessible range.[11–16] In the present work, structural models of various steps in PduO's catalytic cycle were constructed through quantum mechanics/molecular mechanics (QM/MM) optimizations starting from published crystallographic data. These models predict that binding to PduO in the presence of cosubstrate ATP increases the Co^{II}/Co^{I}Cbl redox potential by +89 mV, in reasonable agreement with the recently published experimental value of 285 mV.[17] A relaxed potential energy scan provided detailed insight into the changes in electronic structure of the substrates that occur during the Co–C bond formation step.

These projects used a wide range of spectroscopic and computational tools to further our understanding of how the B_{12} family performs its essential biological roles, providing an excellent foundation for future experimental and theoretic work. The following sections outline potential next steps that would build on this research.

8.2. Proposed theoretical investigations of alkylcobalamins

Orca's advanced spectral analysis tool (ORCA_ASA)[18] yielded enormous insight into the electronic and vibrational properties of CNCbl, including a significant advance in our understanding of how the functionals BP86 and B3LYP describe the molecule's electronic structure. Future work could extend these findings by applying the ORCA_ASA tool to the study of alkylcobalamins like methylcobalamin (MeCbl) and AdoCbl. There are two good reasons to pursue this area of research, one related to the choice of functional for TD-DFT calculations and the other related to the assignment of vibrational modes in the rR spectrum of AdoCbl.

Regarding this first point, there are key differences in electronic structure between CNCbl and its biologically active derivatives, AdoCbl and MeCbl. Chapter 1 discusses how this impacts their respective Abs spectra: CNCbl produces a "typical" Cbl spectrum, while AdoCbl and MeCbl produce "unique" Abs spectra that display a change in relative intensity of the α and β features, as well as a redistribution of the intense γ band across several weaker features. This difference is attributed to the greater σ-donor strength of the alkyl ligands, which decreases the effective nuclear charge of the Co^{III} ion and thus raises the relative energies of the filled Co 3d orbitals, producing much more extensive mixing of these orbitals with the corrin π-based orbitals.[19] This fundamental difference in the amount of mixing of Co and corrin orbitals has implications for the choice of functional for TD-DFT calculations, as evidence shows that BP86 overestimates the covalency of metal-ligand bonds.[20] Relatedly, it has been proposed that BP86 provides a better description of d/π → π*/d transitions and B3LYP provides a better description of corrin π→ π* transitions.[3,4,21,22] It may therefore be enlightening to explore how B3LYP and BP86 perform at predicting vibronic coupling in an alkylcobalamin.

The second reason to continue this work is that it would help to definitively assign the isotope-sensitive peaks in AdoCbl's rR spectrum. Spiro and coworkers collected rR spectra of

AdoCbl and performed $^{12}C \rightarrow {}^{13}C$ substitution of the Co-bound carbon and (separately) deuteration of the hydrogens at that carbon atom to reveal a collection of isotope-sensitive modes between 360 and 570 cm^{-1} that were assigned as various stretching, bending, and torsional modes of the adenosyl moiety.[23–26] Crucially, several rR peaks shift upon AdoCbl binding to methylmalonyl CoA mutase (MMCM), including in the absence and presence of substrate and product analogues.[26] These features could therefore serve as sensitive probes of active site/AdoCbl interactions, provided their specific assignments were known. A later study used DFT/B3LYP-computed Raman spectra to establish a new set of assignments, including in-phase and out-of-phase combinations of the Co–C stretching and Co–C–C bending motions.[27] However, the work described in Chapter 2 of this thesis demonstrates that a direct comparison between experimental and computational Raman data is only warranted if the simulated spectra incorporate the effects of resonance enhancement, particularly when the vibrational motion of interest is distributed across several normal modes. Simulating rR spectra of AdoCbl within the advanced spectral analysis framework may allow for a more conclusive assignment of the experimental features of interest.

8.3. Proposed spectroscopic and computational studies of metal-substituted cobalamins

Chapter 5 of this dissertation details computational work performed with a crystal structure of *Mtb* ATR containing adenosylrhodibalamin (AdoRhbl), a structural mimic of AdoCbl. Ultimately, these calculations could not validate the catalytic relevance of the greatly elongated Rh–C bond (2.23 Å), perhaps because these calculations include only one monomer of the trimeric ATR structure. A very natural next step for this project would be to perform QM/MM calculations with the full biological assembly, which has now been deposited in the protein data bank (PDB: 8D32). More broadly, this project highlights the incredible potential of metal-substituted B_{12} analogues in the study of B_{12}-dependent enzymes. In fact, Chapter 3 provides an excellent

demonstration of the complications that arise in crystal structures containing native AdoCbl, which rarely preserve an intact Co–C bond. Recreating the Co–C bond *in silico* required MD calculations to explore the full conformational space of the adenosyl moiety. AdoRhbl provides an opportunity for sidestepping this complication, as it retains the crucial structural characteristics of the native cofactor while remaining intact during X-ray diffraction data collection. It could therefore be quite useful to use AdoRhbl in the study of EAL or other AdoCbl-dependent enzymes, such as the vital human enzyme methylmalonyl CoA mutase (MMCM). Furthermore, making full use of AdoRhbl as in investigational tool would benefit from a more in-depth spectroscopic characterization, including the collection of Abs and rR spectra.

This work also hints at potential studies with other metal-substituted analogues. The nickel-substituted analogue nibalamin (Nibl) is isostructural and isoelectronic to the Co^{I}Cbl, a critical intermediate in the catalytic cycle of all ATRs and MeCbl-dependent methyltransferases. Nibalamin was found to be a very effective inhibitor of a bacterial ATR, binding to its active site even in the presence of excess Co^{I}Cbl.[28] Spectroscopic characterization of a Nibl-bound ATR—perhaps *Lactobacillus reuteri* PduO, which was used in the computational study described in Chapter 4—would shed light on how Nibl interacts with the enzyme's active site. Our group has conducted extensive studies of PduO, including the collection of rR, MCD, and EPR spectra of the PduO/Co^{II}Cbl/ATP complex.[29,30] Collecting the equivalent spectra for PduO/Nibl/ATP would enable a study of the structural and electronic changes that occur during the Co^{II}Cbl → Co^{I}Cbl step of the ATR's catalytic cycle.

References

(1) Jensen, K. P.; Ryde, U. Theoretical Prediction of the Co–C Bond Strength in Cobalamins. *J. Phys. Chem. A* **2003**, *107* (38), 7539–7545. https://doi.org/10.1021/jp027566p.

(2) Jensen, K. P.; Ryde, U. Cobalamins Uncovered by Modern Electronic Structure Calculations. *Coordination Chemistry Reviews*. Elsevier March 1, 2009, pp 769–778. https://doi.org/10.1016/j.ccr.2008.04.015.

(3) Kornobis, K.; Kumar, N.; Wong, B. M.; Lodowski, P.; Jaworska, M.; Andruniów, T.; Ruud, K.; Kozlowski, P. M. Electronically Excited States of Vitamin B_{12}: Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated Ab Initio Methods. *J. Phys. Chem. A* **2011**, *115* (7), 1280–1292. https://doi.org/10.1021/jp110914y.

(4) Solheim, H.; Kornobis, K.; Ruud, K.; Kozlowski, P. M. Electronically Excited States of Vitamin B_{12} and Methylcobalamin: Theoretical Analysis of Absorption, CD, and MCD Data. *J. Phys. Chem. B* **2011**, *115* (4), 737–748. https://doi.org/10.1021/jp109793r.

(5) Banerjee, R.; Ragsdale, S. W. The Many Faces of Vitamin B_{12}: Catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* **2003**, *72*, 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828.

(6) Brooks, A. J.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Spectroscopic and Computational Studies on the Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase: Evaluation of Enzymatic Contributions to Co–C Bond Activation in the Co^{3+} Ground State. *J. Am. Chem. Soc.* **2004**, *126* (26), 8167–8180. https://doi.org/10.1021/ja039114b.

(7) J. Brooks, A.; Vlasie, M.; Banerjee, R.; C. Brunold, T.; Brooks, A. J.; Vlasie, M.; Banerjee, R.; Brunold, T. C.; J. Brooks, A.; Vlasie, M.; et al. Co-C Bond Activation in Methylmalonyl-CoA Mutase by Stabilization of the Post-Homolysis Product Co^{2+} Cobalamin. *J. Am. Chem. Soc.* **2005**, *127* (47), 16522–16528. https://doi.org/10.1021/ja0503736.

(8) Brooks, A. J.; Fox, C. C.; Marsh, E. N. G.; Vlasie, M.; Banerjee, R.; Brunold, T. C. Electronic Structure Studies of the Adenosylcobalamin Cofactor in Glutamate Mutase. *Biochemistry* **2005**, *44* (46), 15167–15181. https://doi.org/10.1021/bi051094y.

(9) Shibata, N.; Sueyoshi, Y.; Higuchi, Y.; Toraya, T. Direct Participation of a Peripheral Side Chain of a Corrin Ring in Coenzyme B_{12} Catalysis. *Angew. Chemie Int. Ed.* **2018**, *57* (26), 7830–7835. https://doi.org/10.1002/anie.201803591.

(10) Shibata, N.; Toraya, T. Structural Basis for the Activation of the Cobalt-Carbon Bond and Control of the Adenosyl Radical in Coenzyme B_{12} Catalysis. *ChemBioChem*. John Wiley & Sons, Ltd March 14, 2023, p e202300021. https://doi.org/10.1002/cbic.202300021.

(11) 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.

(12) Liptak, M. D.; Fleischhacker, A. S.; Matthews, R. G.; Telser, J.; Brunold, T. C. Spectroscopic and Computational Characterization of the Base-off Forms of

Cob(II)Alamin. *J. Phys. Chem. B* **2009**, *113* (15), 5245–5254. https://doi.org/10.1021/jp810136d.

(13) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Kinetic and Spectroscopic Studies of the ATP:Corrinoid Adenosyltransferase PduO from *Lactobacillus reuteri*: Substrate Specificity and Insights into the Mechanism of Co(II)Corrinoid Reduction. *Biochemistry* **2008**, *47* (34), 9007–9015. https://doi.org/10.1021/bi800419e.

(14) 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.

(15) Mera, P. E.; Maurice, M. S.; Rayment, I.; Escalante-Semerena, J. C. Residue Phe112 of the Human-Type Corrinoid Adenosyltransferase (PduO) Enzyme of *Lactobacillus reuteri* Is Critical to the Formation of the Four-Coordinate Co(II) Corrinoid Substrate and to the Activity of the Enzyme. *Biochemistry* **2009**, *48* (14), 3138–3145. https://doi.org/10.1021/bi9000134.

(16) Park, K.; Mera, P. E.; Moore, T. C.; Escalante-Semerena, J. C.; Brunold, T. C. Unprecedented Mechanism Employed by the Salmonella Enterica EutT ATP:CoIrrinoid Adenosyltransferase Precludes Adenosylation of Incomplete Co^{II}rrinoids. *Angew. Chem., Int. Ed.* **2015**, *54* (24), 7158–7161. https://doi.org/10.1002/anie.201501930.

(17) Gouda, H.; Li, Z.; Ruetz, M.; Banerjee, R. Coordination Chemistry Controls Coenzyme B_{12} Synthesis by Human Adenosine Triphosphate:Cob(I)Alamin Adenosyltransferase. *Inorg. Chem.* **2023**, *62* (32), 12630–12633. https://doi.org/10.1021/acs.inorgchem.3c02163.

(18) Petrenko, T.; Neese, F. Analysis and Prediction of Absorption Band Shapes, Fluorescence Band Shapes, Resonance Raman Intensities, and Excitation Profiles Using the Time-Dependent Theory of Electronic Spectroscopy. *J. Chem. Phys.* **2007**, *127* (16), 164319. https://doi.org/10.1063/1.2770706.

(19) Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C.; A. Stich, T.; J. Brooks, A.; R. Buan, N.; C. Brunold, T. Spectroscopic and Computational Studies of Co^{3+}-Corrinoids: Spectral and Electronic Properties of the B_{12} Cofactors and Biologically Relevant Precursors. *J. Am. Chem. Soc.* **2003**, *125* (19), 5897–5914. https://doi.org/10.1021/ja029328d.

(20) Wasinger, E. C.; De Groot, F. M. F.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. L-Edge X-Ray Absorption Spectroscopy of Non-Heme Iron Sites: Experimental Determination of Differential Orbital Covalency. *J. Am. Chem. Soc.* **2003**, *125* (42), 12894–12906. https://doi.org/10.1021/ja034634s.

(21) Toda, M. J.; Kozlowski, P. M.; Andruniów, T. Assessing Electronically Excited States of Cobalamins via Absorption Spectroscopy and Time-Dependent Density Functional Theory. In *Transition Metals in Coordination Environments: Computational Chemistry and Catalysis Viewpoints*; Broclawik, E., Borowski, T., Radoń, M., Eds.; Springer International Publishing, 2019; pp 219–258. https://doi.org/10.1007/978-3-030-11714-6_8.

(22) Andruniów, T.; Jaworska, M.; Lodowski, P.; Zgierski, M. Z.; Dreos, R.; Randaccio, L.; Kozlowski, P. M. Time-Dependent Density Functional Theory Study of Cobalt Corrinoids: Electronically Excited States of Coenzyme B_{12}. *J. Chem. Phys.* **2009**, *131* (10), 105105. https://doi.org/10.1063/1.3190326.

(23) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Resonance Raman Co−C Stretching Frequencies Reflect Bond Strength Changes in Alkyl Cobalamins, but Are Unaffected by Trans Ligand Substitution. *J. Am. Chem. Soc.* **1996**, *118* (38), 9182–9183. https://doi.org/10.1021/ja962003a.

(24) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Force Constants from Resonance Raman Spectra of Alkylcobalamins: Insensitivity to Dimethylbenzylimidazole Coordination. *Inorganica Chim. Acta* **1998**, *270* (1), 392–398. https://doi.org/https://doi.org/10.1016/S0020-1693(97)05994-X.

(25) Dong, S.; Padmakumar, R.; Maiti, N.; Banerjee, R.; Spiro, T. G. Resonance Raman Spectra Show That Coenzyme B_{12} Binding to Methylmalonyl-Coenzyme a Mutase Changes the Corrin Ring Conformation but Leaves the Co–C Bond Essentially Unaffected. *J. Am. Chem. Soc.* American Chemical Society 1998, pp 9947–9948. https://doi.org/10.1021/ja981584w.

(26) Dong, S.; Padmakumar, R.; Banerjee, R.; Spiro, T. G. Co-C Bond Activation in B_{12}-Dependent Enzymes: Cryogenic Resonance Raman Studies of Methylmalonyl-Coenzyme A Mutase. *J. Am. Chem. Soc.* **1999**, *121* (30), 7063–7070. https://doi.org/10.1021/ja982753f.

(27) Kozlowski, P. M.; Andruniow, T.; Jarzecki, A. A.; Zgierski, M. Z.; Spiro, T. G. DFT Analysis of Co-Alkyl and Co-Adenosyl Vibrational Modes in B_{12}-Cofactors. *Inorg. Chem.* **2006**, *45* (14), 5585–5590. https://doi.org/10.1021/ic052069j.

(28) Kieninger, C.; Wurst, K.; Podewitz, M.; Stanley, M.; Deery, E.; Lawrence, A. D.; Liedl, K. R.; Warren, M. J.; Kräutler, B. Replacement of the Cobalt Center of Vitamin B_{12} by Nickel: Nibalamin and Nibyric Acid Prepared from Metal-Free B_{12} Ligands Hydrogenobalamin and Hydrogenobyric Acid. *Angew. Chemie* **2020**, *132* (45), 20304–20311. https://doi.org/10.1002/ange.202008407.

(29) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Spectroscopic Characterization of Active-Site Variants of the PduO-Type ATP:Corrinoid Adenosyltransferase from Lactobacillus Reuteri: Insights into the Mechanism of Four-Coordinate Co(II)Corrinoid Formation. *Inorg. Chem.* **2012**, *51* (8), 4482–4494. https://doi.org/10.1021/ic202096x.

(30) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Resonance Raman Spectroscopic Study of the Interaction between Co(II)rrinoids and the ATP:Corrinoid Adenosyltransferase PduO from *Lactobacillus reuteri*. *J. Biol. Inorg. Chem.* **2016**, *21* (5–6), 669–681. https://doi.org/10.1007/s00775-016-1371-x.

Appendix

Appendix

Chapter 2

Vibronic Coupling in Vitamin B_{12}: A Combined Spectroscopic and Computational Study

Contents

Table S1. Cartesian coordinates (Å) for DFT geometry-optimized model of CNCbl

Atom	x	y	z	Atom	x	y	z
C	-2.497	-4.259	2.145	H	-2.844	3.339	-3.770
C	-3.094	-3.026	4.762	Co	1.100	-0.359	-0.653
C	0.368	0.974	1.911	H	-0.385	1.641	3.836
C	-1.060	-2.278	1.608	H	-1.957	-4.712	1.290
C	-1.900	-2.923	2.521	H	-3.559	-4.137	1.839
C	-2.224	-2.300	3.772	H	-2.488	-4.988	2.978
C	-1.735	-1.020	4.076	H	-3.371	-2.378	5.613
C	-0.934	-0.367	3.124	H	-2.574	-3.923	5.162
C	-0.569	-0.997	1.907	H	-4.025	-3.386	4.282
N	-0.329	0.891	3.087	H	0.946	1.848	1.597
N	0.244	-0.126	1.181	H	-0.802	-2.775	0.672
C	-1.509	-0.259	-1.825	H	-1.995	-0.552	5.034
C	1.973	-0.601	-2.278	H	-1.859	-3.319	-3.155
N	2.510	-0.754	-3.309	H	4.017	-3.956	1.229
C	-2.223	-1.131	-2.942	H	4.667	-1.562	1.482
C	-2.032	-2.580	-2.355	H	4.613	3.142	-0.025
C	-0.708	-2.423	-1.637	H	-2.167	2.385	-1.590
C	0.157	-3.528	-1.328	H	-0.221	0.733	-3.255
C	1.286	-3.328	-0.548	H	-1.854	0.734	0.086
C	2.354	-4.362	-0.169	H	-2.746	-0.796	-0.111
C	2.954	-3.665	1.104	H	-3.277	0.680	-0.993
C	2.826	-2.234	0.668	H	-3.692	0.423	-3.177
C	3.760	-1.241	0.962	H	-4.330	-0.971	-2.306
C	3.676	0.091	0.573	H	-4.117	-1.036	-4.079
C	4.826	1.074	0.630	H	-1.586	-0.032	-4.732
C	4.062	2.427	0.617	H	-0.317	-1.014	-3.989
C	2.779	2.031	-0.106	H	-3.657	-2.363	-0.884
C	1.912	2.943	-0.687	H	-2.717	-3.851	-0.702
C	0.693	2.488	-1.311	H	0.539	-5.628	-1.773
C	-0.277	3.350	-2.138	H	-0.269	-4.734	-3.073
C	-1.442	2.330	-2.424	H	-1.186	-5.231	-1.629
C	-0.735	0.983	-2.305	H	0.925	-5.836	0.588
C	-2.406	0.110	-0.639	H	2.653	-6.284	0.799
N	-0.445	-1.175	-1.353	H	1.873	-6.443	-0.784
N	1.721	-2.084	-0.094	H	3.679	-3.156	-1.437
N	2.594	0.663	0.001	H	4.350	-4.725	-0.927
N	0.339	1.226	-1.332	H	1.078	-3.785	2.199
C	-3.670	-0.678	-3.138	H	2.344	-4.915	2.792
C	-1.393	-1.009	-4.249	H	6.404	1.617	-0.799
H	-1.572	-1.777	-5.003	H	4.870	1.073	-1.578
C	-3.157	-3.167	-1.455	H	5.953	-0.119	-0.828
H	-3.912	-3.735	-2.039	H	5.241	0.782	2.762
C	-0.204	-4.848	-1.970	H	6.535	1.637	1.847
C	1.923	-5.806	0.117	H	6.314	-0.114	1.653
C	3.429	-4.222	-1.282	H	2.837	3.762	1.779
H	3.169	-4.659	-2.242	H	3.253	2.282	2.640
C	2.160	-3.887	2.426	H	3.383	4.536	-0.560
H	2.427	-3.160	3.230	H	1.786	5.011	0.058
C	5.560	0.907	-0.727	H	2.074	4.858	-1.692
C	5.775	0.839	1.796	H	1.410	4.190	-3.310
C	3.683	3.070	1.981	H	-0.200	4.377	-4.067
H	4.434	3.646	2.567	H	0.617	2.776	-4.052
C	2.300	4.405	-0.713	H	-1.523	5.122	-2.046
C	0.434	3.698	-3.466	H	0.108	5.363	-1.468
C	-0.732	4.648	-1.431	H	-1.443	2.547	-4.589
H	-1.088	4.597	-0.379	H	-2.784	1.596	-3.964
C	-2.170	2.488	-3.760				

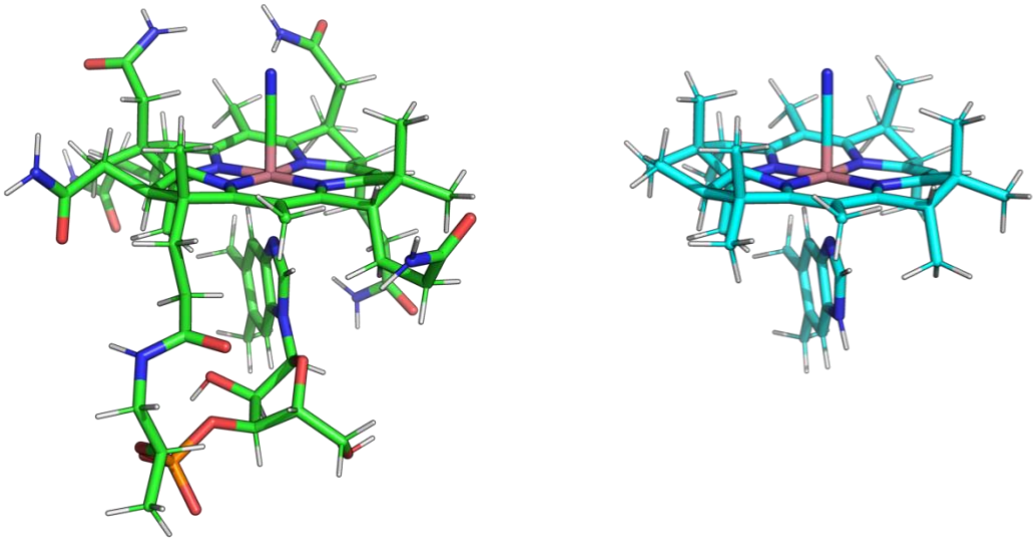


Figure S1. QM/MM-optimized structure of CNCbl (green, left) and truncated CNCbl model used for DFT and TD-DFT calculations (cyan, right).

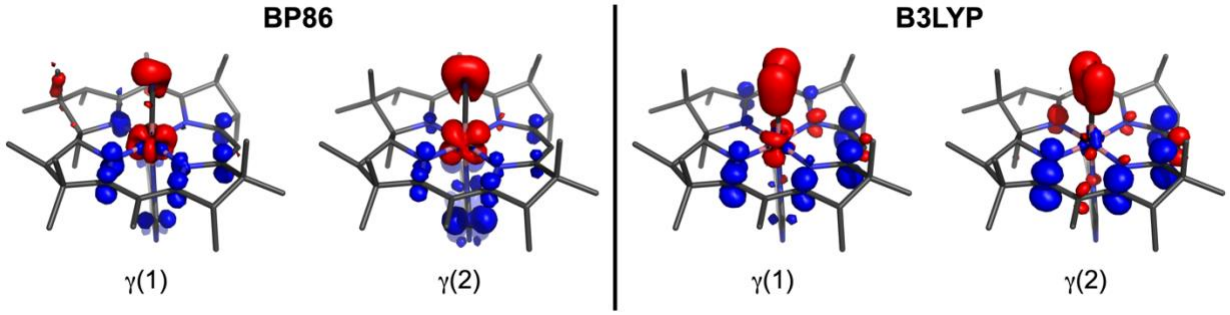


Figure S2. Electron density difference maps (EDDMs) for key electronic transitions in the γ region of TD-DFT calculated Abs spectra for CNCbl. Blue and red indicate gain and loss of electron density, respectively.

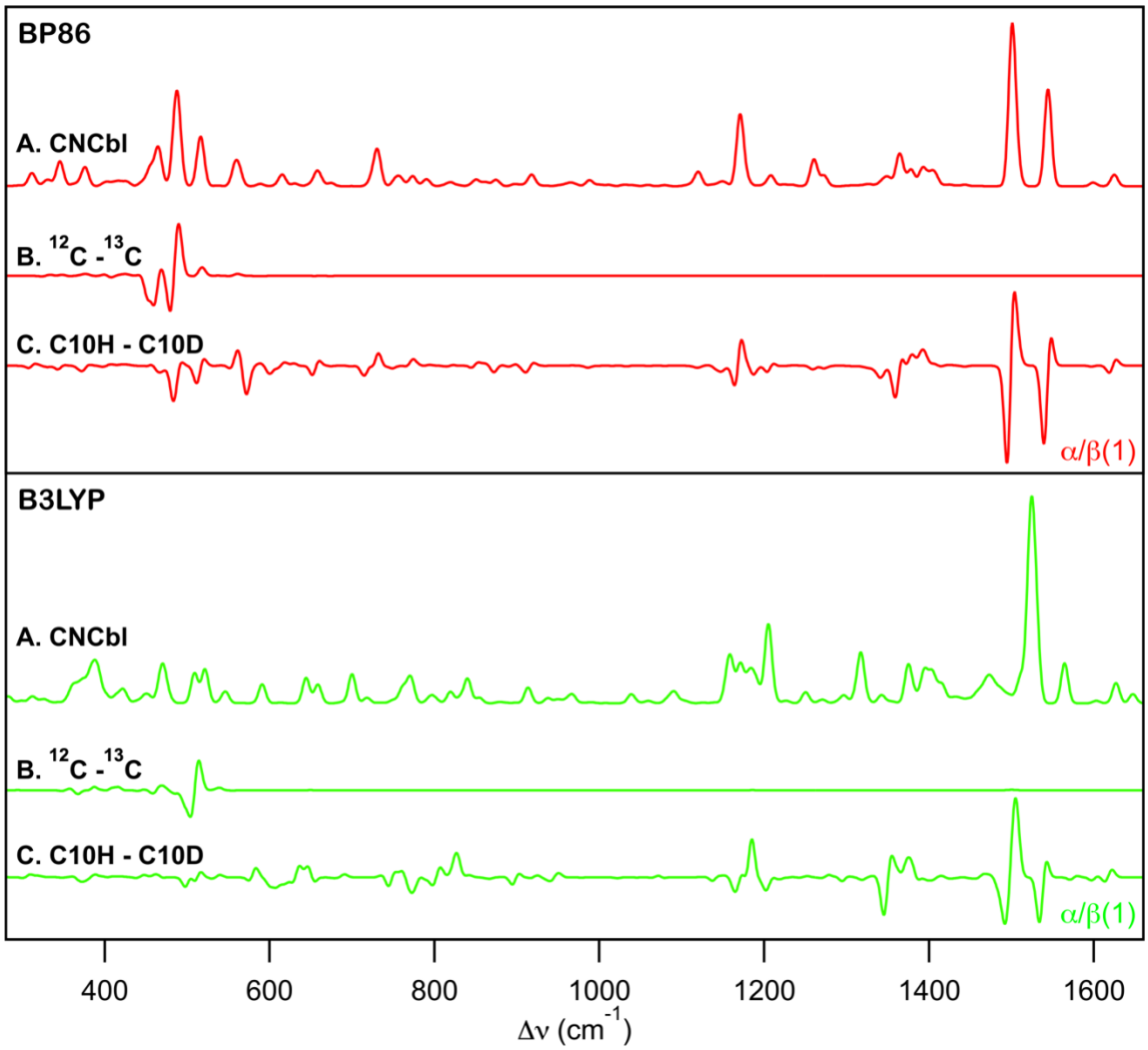


Figure S3. Computed rR spectra for CNCbl (A) for excitation in resonance with the α/β(1) transition obtained using the BP86 (top) and B3LYP (bottom) functionals, along with predicted difference spectra for $^{12}CN \rightarrow ^{13}CN$ isotopic substitution (B) and H/D exchange at the C10 position (C).

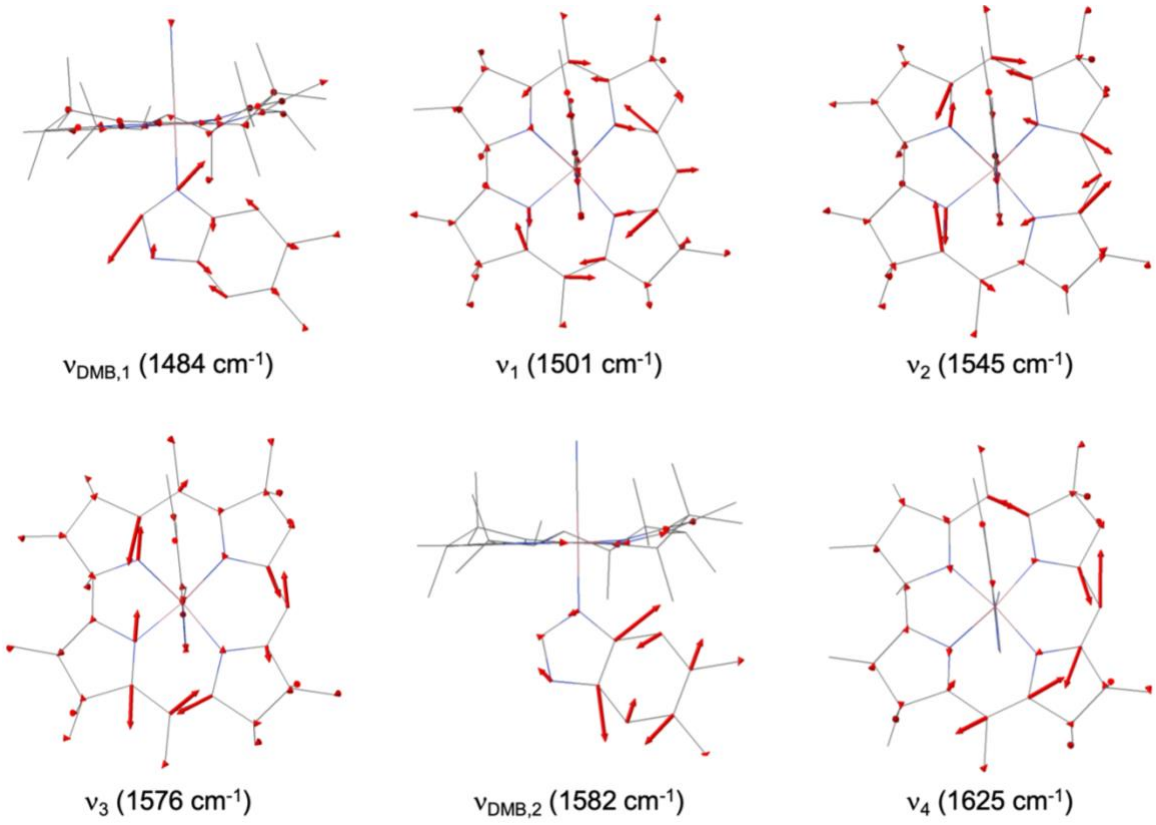


Figure S4. BP86-computed eigenvector representations of the relevant high-frequency vibrational modes included in Table 1 (DFT-computed frequencies are given in parentheses).

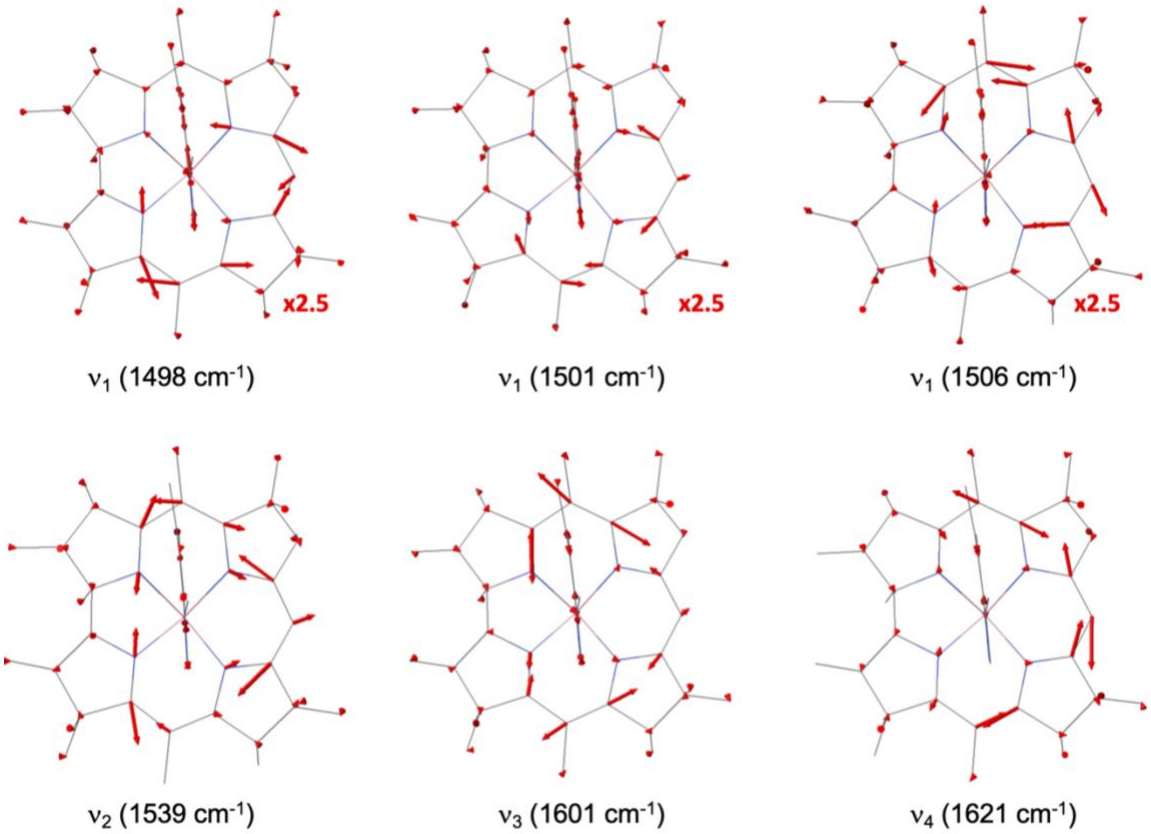


Figure S5. B3LYP-computed eigenvector representations of the relevant high-frequency vibrational modes included in Table 1 (DFT-computed frequencies are given in parentheses). For clarity, the magnitude of vectors is scaled up by 2.5-fold for select modes (labeled ×2.5). Note that ν_1 consists of three closely spaced vibrational modes.

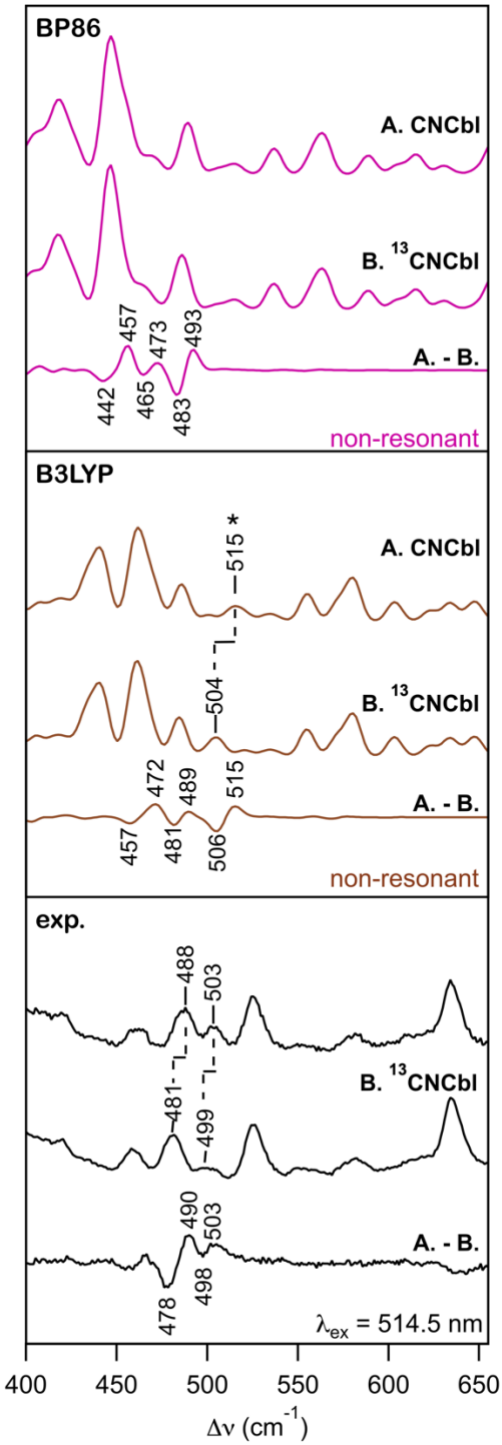


Figure S6. Low-frequency regions of computed (top and center) off-resonance Raman spectra of ^{12}CNCbl and ^{13}CNCbl, and their difference spectra, shown along with the experimental (bottom) rR spectra of these species collected with 514.5 nm excitation. The peak indicated by an asterisk in the B3LYP-predicted spectrum for ^{12}CNCbl contains contributions from the Co–C–N bending modes at 513 cm^{-1} and 518 cm^{-1}, depicted in Figures 8 and S4 respectively. The energies of rR features that display a visible shift in response to isotopic labeling are indicated.

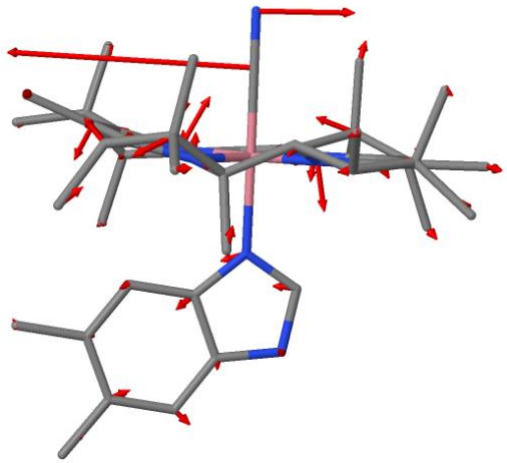


Figure S7. B3LYP-computed eigenvector representation of the mode at 518 cm^{-1} that downshifts by 10 cm^{-1} upon ^{12}CNCbl → ^{13}CNCbl isotopic substitution (marked by an asterisk in Figure S6). Unlike the vibrational modes depicted in Figure 8, this mode displays virtually no resonance enhancement in the computed rR spectra for excitation in the α/β region.

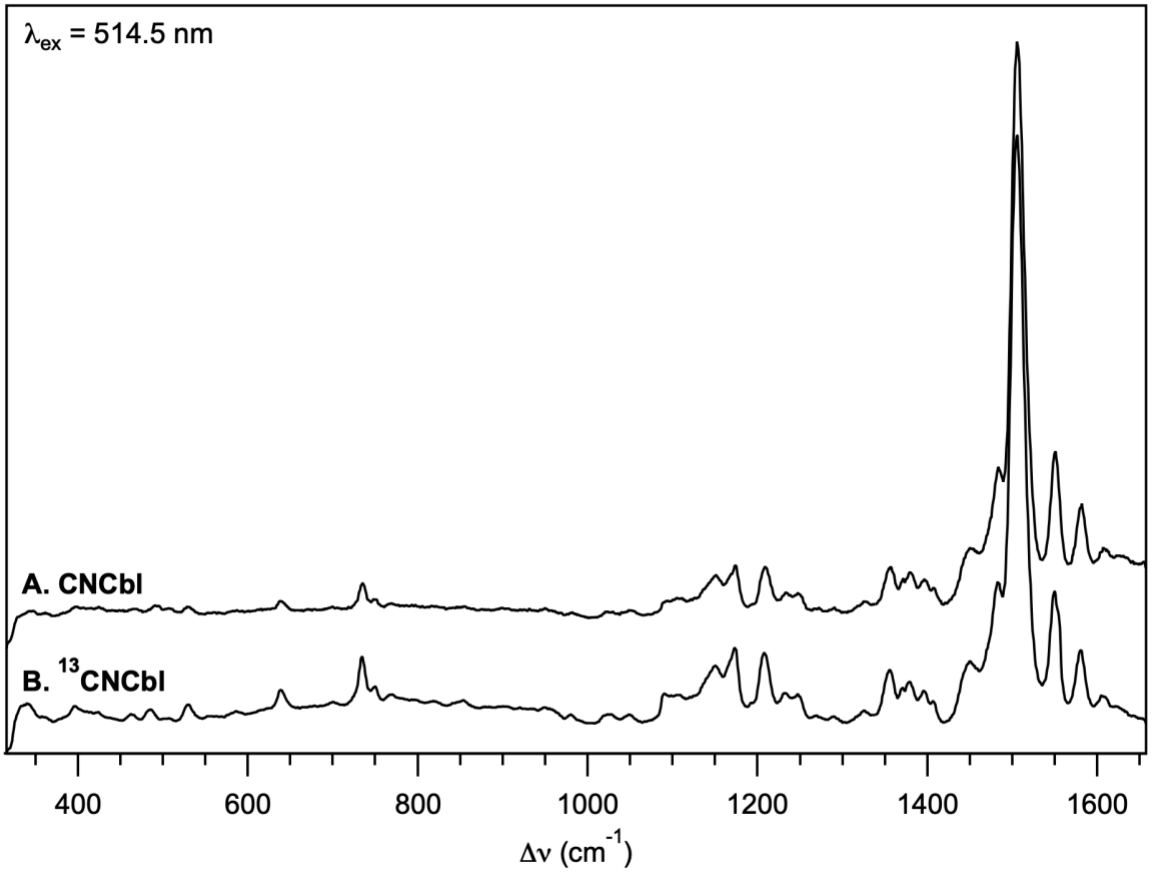


Figure S8. Full experimental rR spectra of ^{12}CNCbl (A) and ^{13}CNCbl (B) obtained at 77 K with 514.5 nm excitation.

Appendix 2.1

Representative input files for DFT and TD-DFT computations performed in Orca version 4.2.1 are provided below. For brevity, the coordinate block is omitted in each example, but the coordinates are available in Table S1 above. It should be noted that, in addition to providing the cartesian coordinates, the coordinate block also specifies the use of the TZVP basis set for the Co atom and its six coordinated atoms. This is achieved by including “newgto "TZVP" end” after the coordinates for these atoms.

Note that two files are provided for the advancd spectral analysis portion; the initial calculation produces a second file that can be edited and run through the orca_asa program. Again, some parts of the file (in this case, the lengthy $vib_freq_gs and $sdnc blocks) are omitted for brevity.

TD-DFT

```
! RKS B3LYP/G def2-SVP def2-SVP/C Grid4 TightSCF NormalPrint

%scf  maxiter 300 damp fac 0.85 erroff 0.005 end
        end

%pal nprocs 32
end

%tddft NRoots 60
        MaxDim 5
        EWin -4,4
        TDA true
        Etol -1
        RTol 1e-4
        end

%pointcharges “charge_CNCbl_tddft.txt"

* xyz +1 1
#add coordinates here
*
```

Vibrational frequency calculation

```
! RKS B3LYP/G def2-SVP def2-SVP/C Grid4 TightSCF NormalPrint MORead

%moinp "CNCbl_tddft_b3lyp.gbw" # the gbw file produced in the initial TD-DFT calculation

%scf  maxiter 300 damp fac 0.85 erroff 0.005 end
        end
```

```
%pal nprocs 32
end

!NumFreq

%freq CentralDiff true Increment 0.01
        end

%elprop Polar 1
        end

%pointcharges “charge_CNCbl_tddft.txt"

* xyz +1 1
#add coordinates here
*
```

Advanced spectral analysis (initial input file)

```
! RKS B3LYP def2-SVP def2-SVP/C Grid4 TightSCF NormalPrint MORead NMGrad

%moinp "CNCbl_tddft_b3lyp.gbw" # the gbw file produced in the initial TD-DFT calculation

%scf  maxiter 3000 damp fac 0.85 erroff 0.005 end
        end

%pal nprocs 32
        end

%tddft NRoots 60
        MaxDim 5
        EWin -4,4
        TDA true
        Etol -1
        RTol 1e-4
        end

%pointcharges "charge_CNCbl_tddft.txt"

%rr States 1, 3, 4, 10, 14, 16
        HessName "CNCbl_freq_b3lyp.hess" # hess file from frequency calculation
        ASAInput true
        end

* xyz +1 1
#add coordinates here
*
```

Advanced spectral analysis (input file run through orca_asa)

```
%sim
      model IMDHO
      method Heller
      AbsRange     10000.0,  45000.0
      NAbsPoints    3501

      RRPRange     10000.0,  45000.0
      NRRPPoints    3501

      RRSRange       10,    1700
      NRRSPoints 1691

            RRSE 20814, 23371, 24577, 29201, 31890, 32192
      RRS_FWHM 10.0

      AbsScaleMode Ext
      RamanOrder  1

      EnInput E0

      CAR   0.800
end

%heller
      IntMethod NumInt
      NumIntMethod GL
      IntConvTol      1.000000e-03
      NumIntTimeTol   1.000000e-03
      RRIntConvTol    1.000000e-03
      RRNumIntTimeTol 1.000000e-03
end

$el_states
6
  1 20814.36   160.00    0.00    0.5379   1.3669   0.0439
  2 23371.46   160.00    0.00   -0.3726   0.0881  -0.2478
  3 24577.11   160.00    0.00   -0.2806   0.3758  -0.2912
  4 29201.01   160.00    0.00   -0.7399   0.2487  -0.3641
  5 31889.65   160.00    0.00   -0.8382   0.2728  -0.3065
  6 32191.71   160.00    0.00   -1.0141   0.3433  -0.7063

#add the $vib_freq_gs block here

#add the $sdnc block here
```

Appendix 2.2

A normal coordinate analysis for the Co–C–N unit of CNCbl was performed using Wilson's method of F and G matrices (Wilson, E. B. J.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955), which states that:

$$|FG - E\lambda| = 0$$

The G matrix:

inverse masses: $\mu_{Co} = \frac{1}{58.933}$ $\mu_C = \frac{1}{12.011}$ $\mu_N = \frac{1}{14.007}$

bond lengths and angles: r_{CoC} = 1.86 Å r_{CN} = 1.17 Å ϕ = 179.3°

$$G = \begin{bmatrix} \mu_{Co} + \mu_C & \mu_C \cdot \cos(\phi) & \frac{\mu_C \cdot \sin(\phi)}{r_{CN}} \\ \mu_C \cdot \cos(\phi) & \mu_C + \mu_N & \frac{\mu_C \cdot \sin(\phi)}{r_{CoC}} \\ \frac{\mu_C \cdot \sin(\phi)}{r_{CN}} & \frac{\mu_C \cdot \sin(\phi)}{r_{CoC}} & \frac{\mu_{Co}}{{r_{CoC}}^2} + \frac{\mu_N}{{r_{CN}}^2} + \mu_C \cdot \left[\left(\frac{1}{{r_{CoC}}^2}\right) + \left(\frac{1}{{r_{CN}}^2}\right) - \frac{2 \cdot \cos(\phi)}{r_{CoC} r_{CN}}\right] \end{bmatrix}$$

The F matrix:

$$F = \begin{bmatrix} k_{Co-C} & 0 & 0 \\ 0 & k_{C-N} & 0 \\ 0 & 0 & k_{Co-C-N} \end{bmatrix}$$

Fitting to experimental data:

MATLAB was used to fit the unknowns k_{Co-C}, k_{C-N}, and k_{Co-C-N} to the experimental rR frequencies:

$$\nu\ (cm^{-1}) = \begin{bmatrix} 488 \\ 503 \\ 2127 \end{bmatrix}$$

$$\lambda = eig(F \cdot G)$$

$$\nu = \frac{\sqrt{\lambda \times N_A \times 10^5}}{2 \times \pi \times c}$$

The best fit was achieved for: k_{Co-C} = 2.67 mdyn/Å, k_{C-N} = 16.4 mdyn/Å, and k_{Co-C-N} = 0.677 mdyn Å/rad^2.

Chapter 3

Spectroscopic and Computational Investigation of the AdoCbl-Dependent Enzyme Ethanolamine Ammonia-Lyase: Destabilization of Protein-Bound AdoCbl Contributes to Accelerated Co–C Bond Homolysis

Supporting Information

Appendix 1.

The details of the computational workflow for building structural models are described below.

Solvated AdoCbl

Atomic coordinates for all atoms of AdoCbl were taken from the crystal structure of AdoCbl (CCDC ID: PAFBUV). AdoCbl AMBER MM parameters were taken from literature and units were converted for GROMACS compatibility. AdoCbl partial charges and Lennard-Jones parameters were taken from the literature. AdoCbl was solvated in 103.5 x 103.5 x 103.5 nm^3 TIP3P water box using GROMACS. An energy minimization was performed using the steepest descent method to form the Co–C bond. A constant volume equilibration at 300 K was performed for 100 ps with a time step of 2 ps with position restraints on the heavy atoms of the protein and AdoCbl of 10 kJ mol^{-1}. A constant pressure equilibration was then performed at 300 K with position restraints on the heavy atoms of the protein and AdoCbl of 10 kJ/mol. After these equilibrations, a 5 ns MD run was performed 300 K with a time step of 2 fs and long-range interactions with particle mesh Ewald. After the 5 ns MD run, the heavy atoms (Figure S3) of AdoCbl were used for GROMACS clustering with an RMSD cutoff of 0.08 nm to create 18 clusters. The first cluster's average occupancy corresponded to timepoint 420 ps. The AdoCbl coordinates and all waters within 29 nm of AdoCbl at timepoint 420 ps were selected for subsequent QM/MM optimization.

HoloEAL

Atomic coordinates were taken from the crystal structure of EAL complexed with AdoCbl (PDB: 5YSN). One α/β unit of HoloEAL was extracted from the crystal structure at protonated at pH 7.0 using PDB2PQR server. AdoCbl AMBER MM parameters were taken from literature and units were converted for GROMACS compatibility. AdoCbl partial charges and Lennard-Jones parameters were taken from the literature. The active site was visually inspected to confirm appropriate protonation, namely Eα287, Dα362 carry a negative charge and Rα160 has a positive charge. The protein was modeled with the AMBER14SB forcefield. The HoloEAL system was solvated in 103.5 x 103.5 x 103.5 nm^3 TIP3P water box using GROMACS. 17 Na^+ counterions were added to neutralize the system. Waters added within 10 Å of the Co atom were removed to ensure the active site contained only crystallographic waters. An energy minimization was performed using the steepest descent method to form the Co–C bond. A constant volume equilibration at 300 K was performed for 100 ps with a time step of 2 ps with position restraints on the heavy atoms of the protein and AdoCbl of 10 kJ mol^{-1}. A constant pressure equilibration was then performed at 300 K with position restrains on the heavy atoms of the protein and AdoCbl

of 10 kJ/mol. After these equilibrations, a 5 ns MD run was performed 300 K with a time step of 2 fs and long-range interactions with particle mesh Ewald.. After the 5 ns MD run, selected heavy atoms (Figure S6) of the HoloEAL active site were used for GROMACS clustering with an RMSD cutoff of 0.041 nm. Cluster 1 represented 381 timeframes. The first cluster's average occupancy corresponded to timepoint 3190 ps. The AdoCbl/HoloEAL system and waters within 5 Å of the protein at timepoint 3190 ps were selected for subsequent QM/MM optimization.

Ternary Complex

Atomic coordinates were taken from the crystal structure of EAL complexed with AdoCbl (PDB: 5YSN). One α/β unit of HoloEAL was extracted from the crystal structure at protonated at pH 7.0 using PDB2PQR server. The active site was visually inspected to confirm appropriate protonation, namely Eα287, Dα362 carried a negative charge and Rα160 was protonated. The two active site waters were replaced by ethanolamine (EA), and the coordinates for the heavy atoms of EA were taken from the 3ABS crystal structure. AdoCbl AMBER MM parameters were taken from literature and units were converted for GROMACS compatibility (Marques,). AdoCbl partial charges and Lennard-Jones parameters were taken from the literature. (CHARMM). The protein was modeled with the AMBER14 forcefield. EA was geometry optimized and parameterized using AmberTools and the SQM program with AM1 before addition to the AdoCbl/EAL system. AdoCbl/EAL/EA was solvated in 103.5 x 103.5 x 103.5 nm^3 TIP3P water box using GROMACS. Waters added by GROMACS within 10 Å of the Co atom were removed to ensure the active site contained only crystallographic waters. 16 Na^+ counterions were added by GROMACS to neutralize the system. An energy minimization was performed using the steepest descent method to form the Co–C bond. A constant volume equilibration at 300 K was performed for 100 ps with a time step of 2 ps with position restraints on the heavy atoms of the protein, EA, and AdoCbl of 10 kJ mol^{-1}. A constant pressure equilibration was then performed at 300 K with position restrains on the heavy atoms of the protein and AdoCbl of 10 kJ/mol. After these equilibrations, a 5 ns MD run was performed 300 K with a time step of 2 fs and long-range interactions with particle mesh Ewald. After the 5 ns MD run, select heavy atoms of the AdoCbl/EAL/EA active site (Figure S9) were used for GROMACS clustering with an RMSD cutoff of 0.042 nm. See SI for more information on clustering procedure. Cluster 1 contained 354 timepoints with the average coordinates represented by timepoint 2060 ps. The AdoCbl/HoloEAL system and waters within 5 Å of the protein at timepoint 2060 ps were selected for subsequent QM/MM optimization.

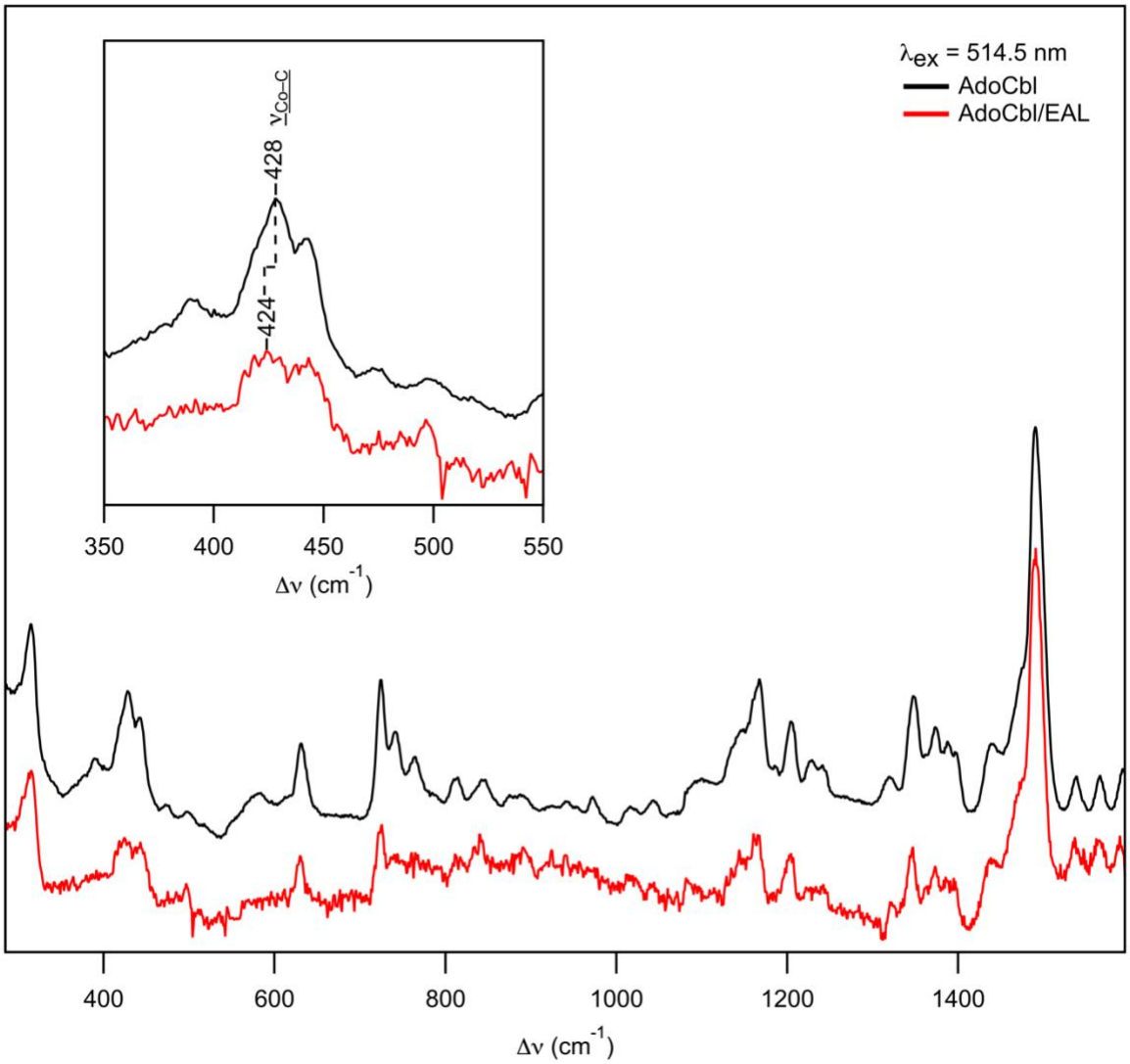


Figure S1. rRaman spectra of free (black) and EAL-bound (red) AdoCbl, obtained with 514.5 nm laser excitation. Inset: A closer look at the ν_{Co-C} stretching peaks in these spectra.

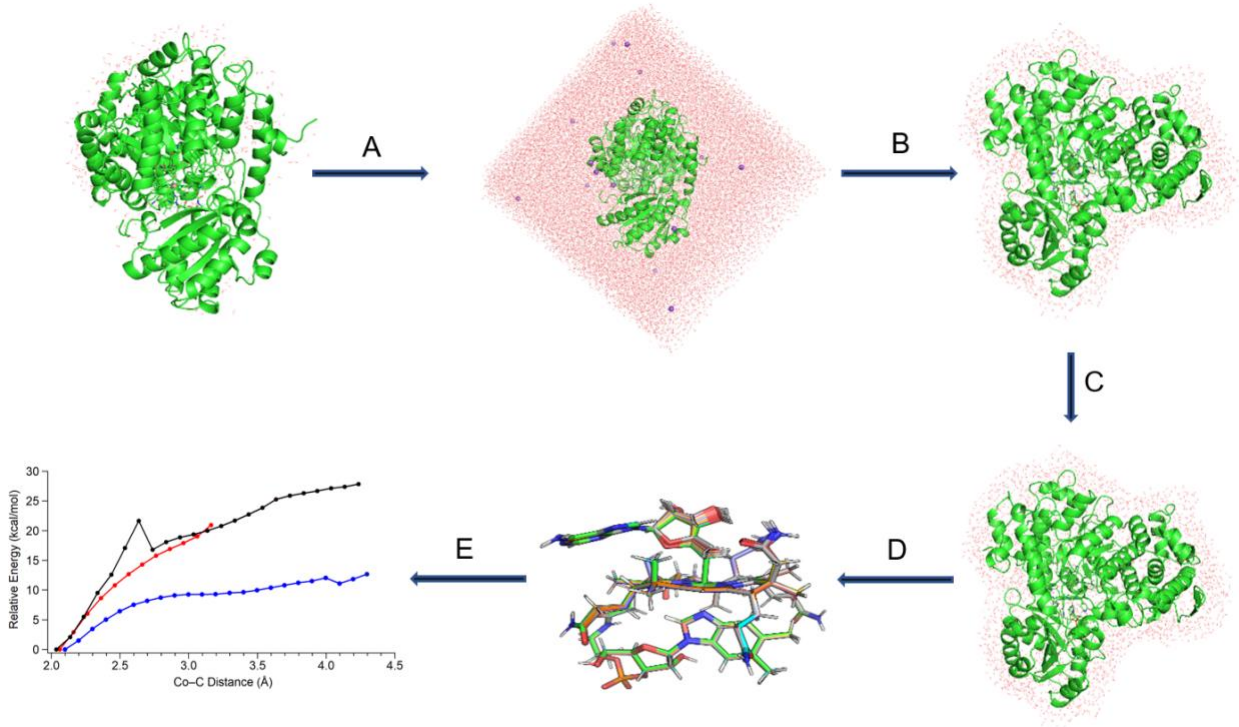


Figure S2. MD + QM/MM computational workflow. Additional information for each step is described above.

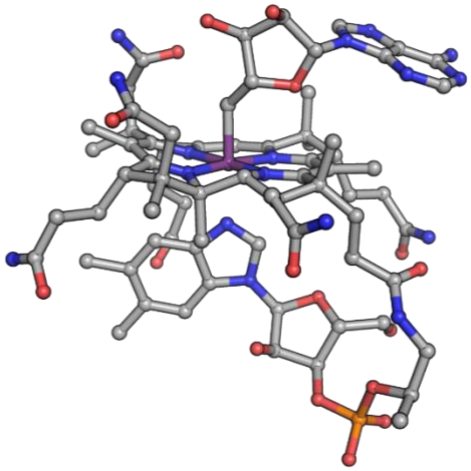


Figure S3. The heavy atoms of AdoCbl used for the RMSD-based clustering algorithm. All non-hydrogen QM atoms are contained within this structure.

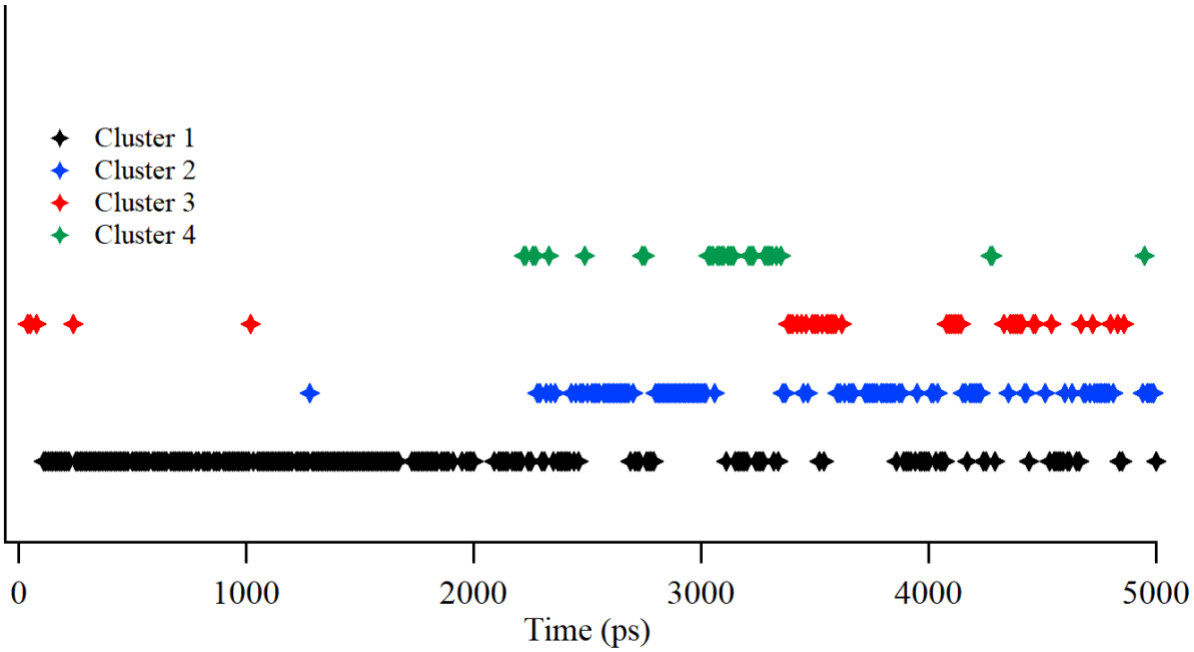


Figure S4. Cluster occupancy for the heavy atoms of AdoCbl (Figure S3) using an RMSD cutoff of 0.08 nm. Number of structures in each cluster out of 501: Cluster 1 = 246 ; Cluster 2 = 111 ; Cluster 3 = 43 ; Cluster 4 = 30 ; Clusters 5 through 18 (not shown) = 71.

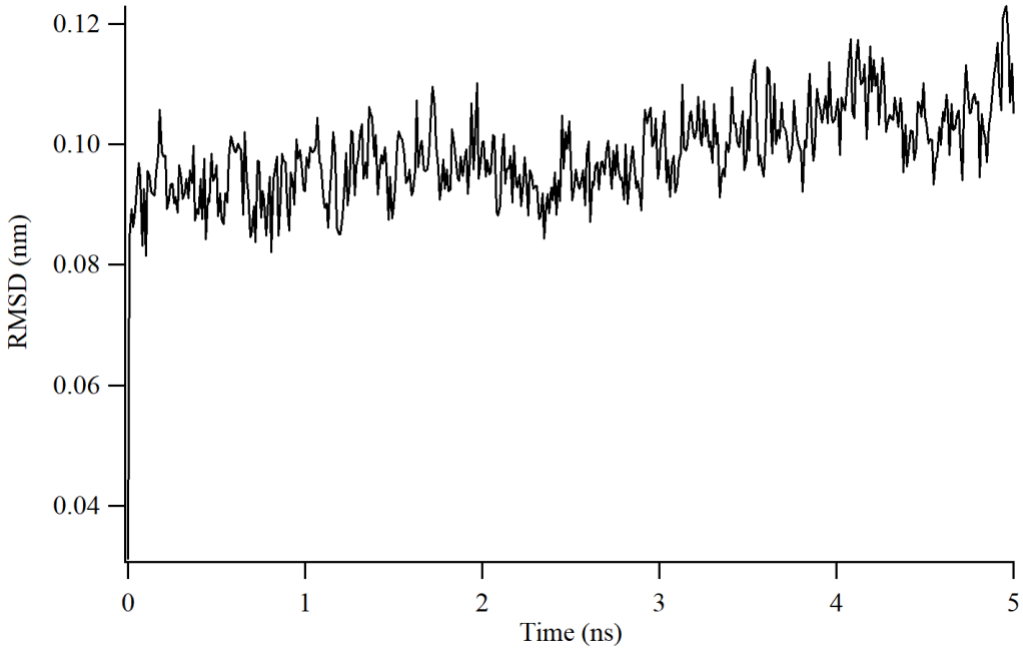


Figure S5. Root-mean-square deviation of the AdoCbl/EAL system protein backbone. RMSD distances calculated for the 5 ns MD run in reference to the crystal structure (PDB: 5YSN) protein backbone before energy minimization, NVT, and NPT equilibration.

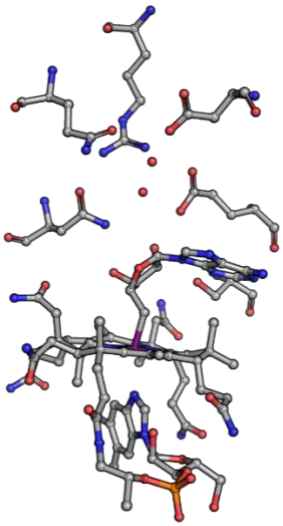


Figure S6. The heavy atoms of AdoCbl/EAL system used for RMSD clustering calculation. All non-hydrogen QM atoms are contained within this structure.

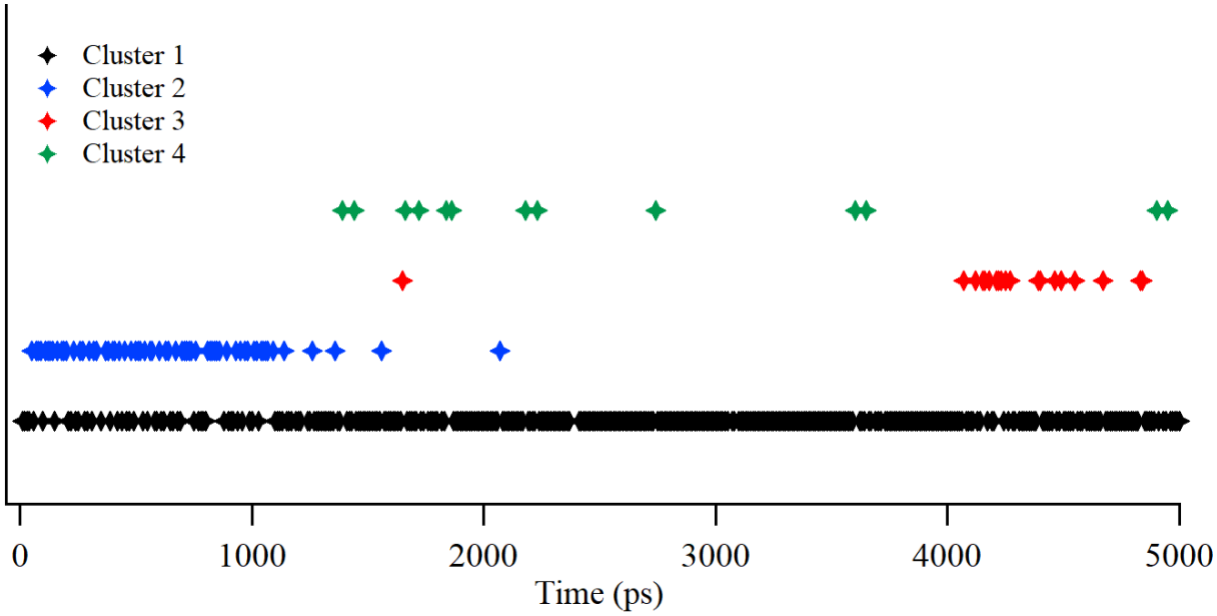


Figure S7. Cluster distribution for the heavy atoms of AdoCbl/EAL substructure (Figure S6) using an RMSD cutoff of 0.041 nm. Number of structures in each cluster out of 501: Cluster 1 = 381 ; Cluster 2 = 64 ; Cluster 3 = 19 ; Cluster 4 = 13 ; Clusters 5 through 15 (not shown) = 24.

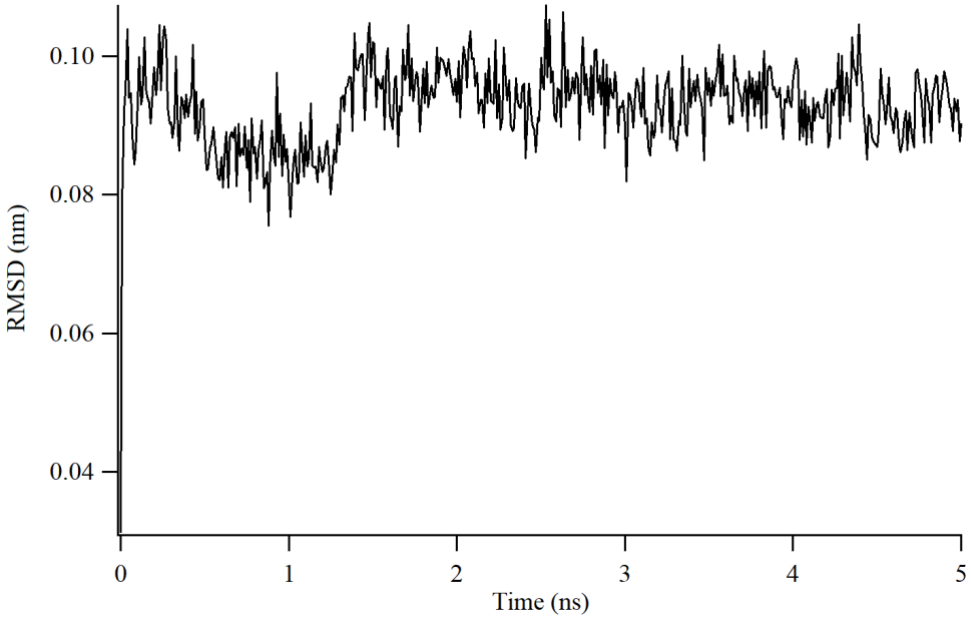


Figure S8. Root-mean-square deviation of the AdoCbl/EAL/EA system protein backbone. RMSD distances calculated for the 5 ns MD run in reference to the crystal structure (PDB: 5YSN) protein backbone before energy minimization, NVT, and NPT equilibration.

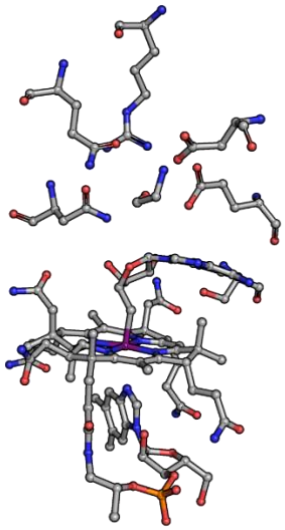


Figure S9. The heavy atoms of AdoCbl/EAL/EA system used for RMSD calculation for clustering. All non-hydrogen QM atoms are contained within this structure.

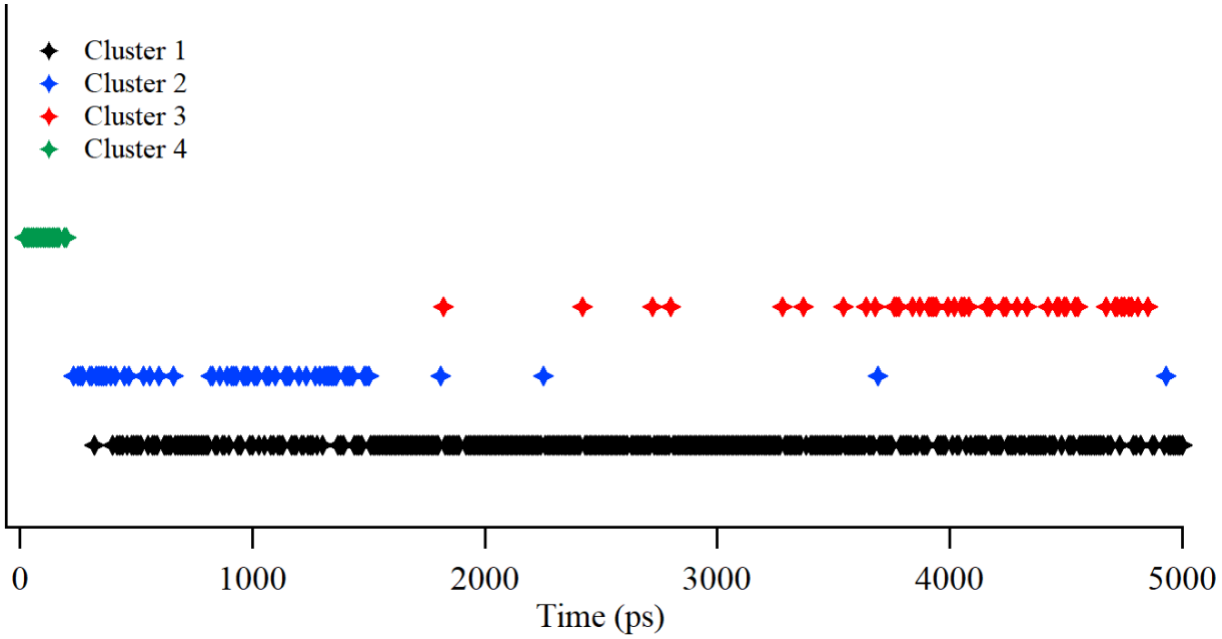


Figure S10. Cluster distribution for the heavy atoms of AdoCbl/EAL/EA substructure (Figure S9) using an RMSD cutoff of 0.042 nm. Number of structures in each cluster out of 501: Cluster 1 = 354 ; Cluster 2 = 58 ; Cluster 3 = 45; Cluster 4 = 18 ; Clusters 5 through 14 (not shown) = 26.

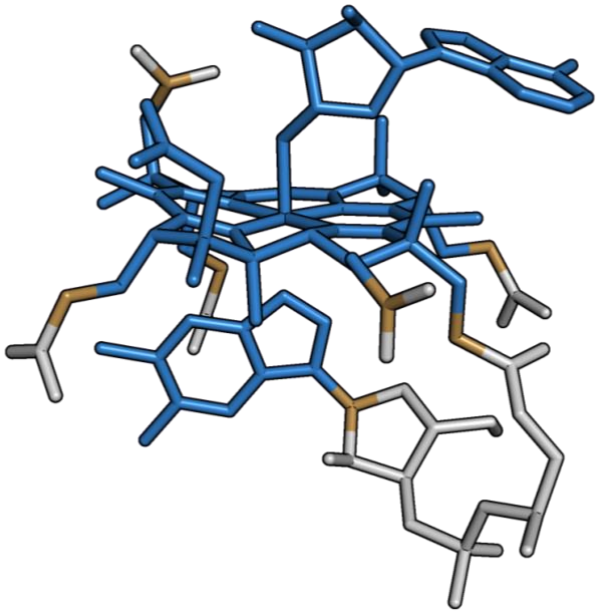


Figure S11. Cofactor QM Region. QM atoms shown as blue sticks, link atoms shown in tan, and MM atoms are white sticks.

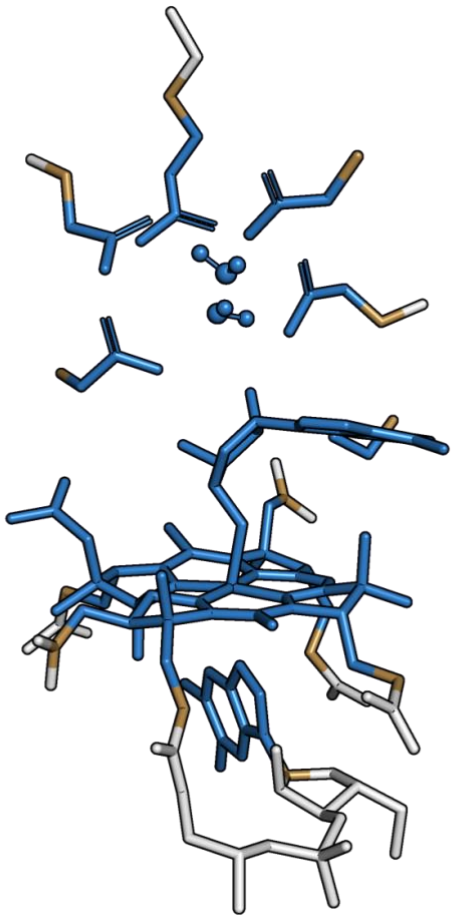


Figure S12. HoloEAL QM Region. QM atoms shown as blue sticks, link atoms shown in tan, and MM atoms are white sticks.

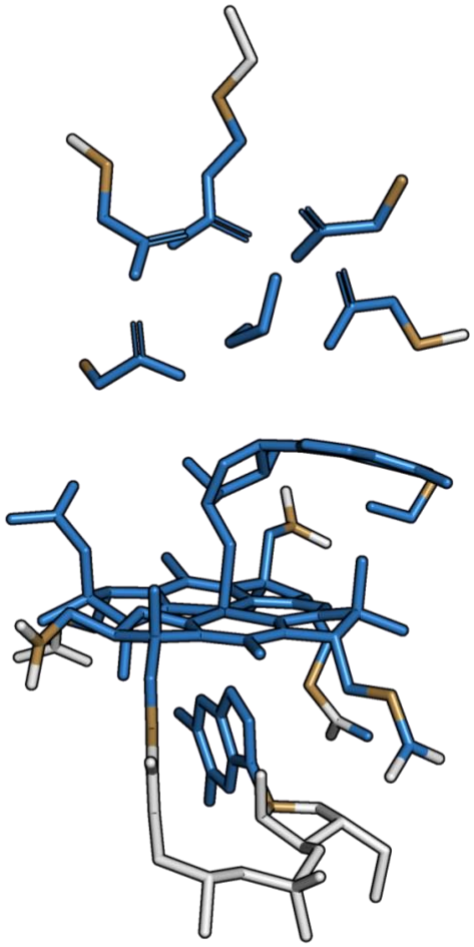


Figure S13. Ternary Complex QM Region. QM atoms shown as blue sticks, link atoms shown in tan, and MM atoms are white sticks.

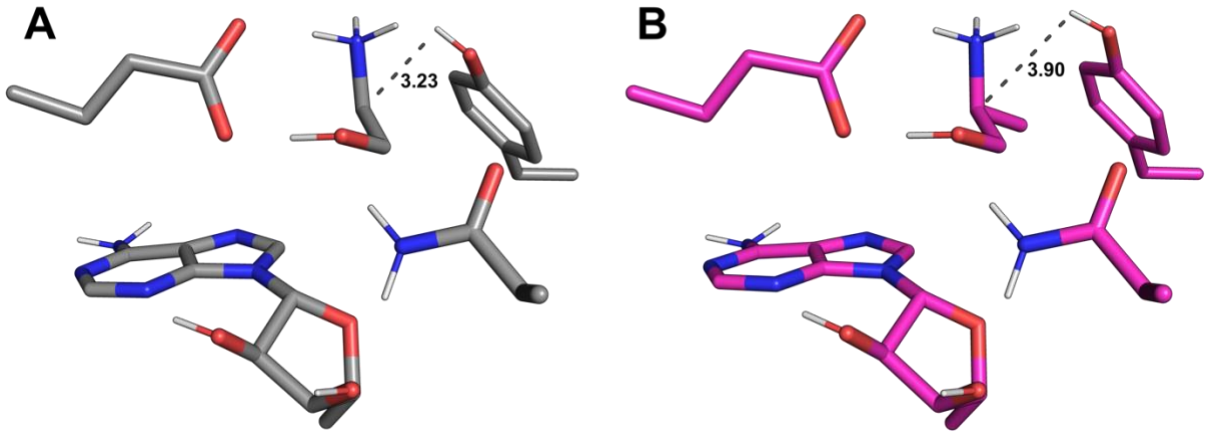


Figure S14. Proximity of Tyrα404 to substrate in QM/MM-optimized models of AdoCbl/EAL/EA (gray) and AdoCbl/EAL/2AP (pink).

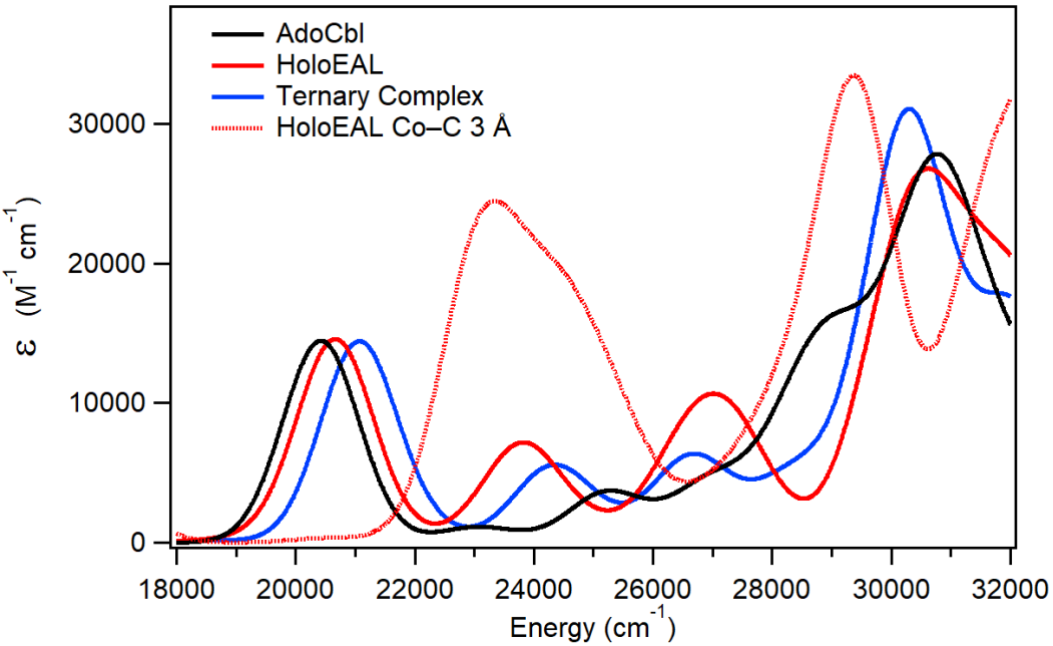


Figure S15. TD-DFT computed absorption spectra of the three ground state AdoCbl systems (solid traces) and the computed absorption spectra of the optimized HoloEAL system at a Co–C bond distance of 3 Å. The 3 Å bond distance corresponds to the CoC distance seen in the crystal structure (PDB 5YSN) used for initial MD + QM/MM.

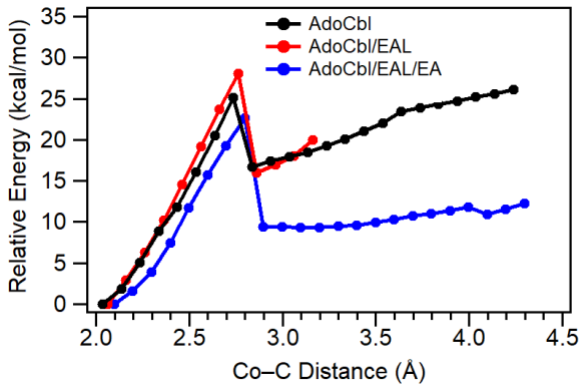


Figure S16. SP QM energies for all stationary points during Co – C elongation using the full QM region of each system and the def2-TZVP basis set on all QM atoms.

Chapter 4

Molecular-Level Insight into the Catalytic Cycle of a PduO-type Adenosyltransferase

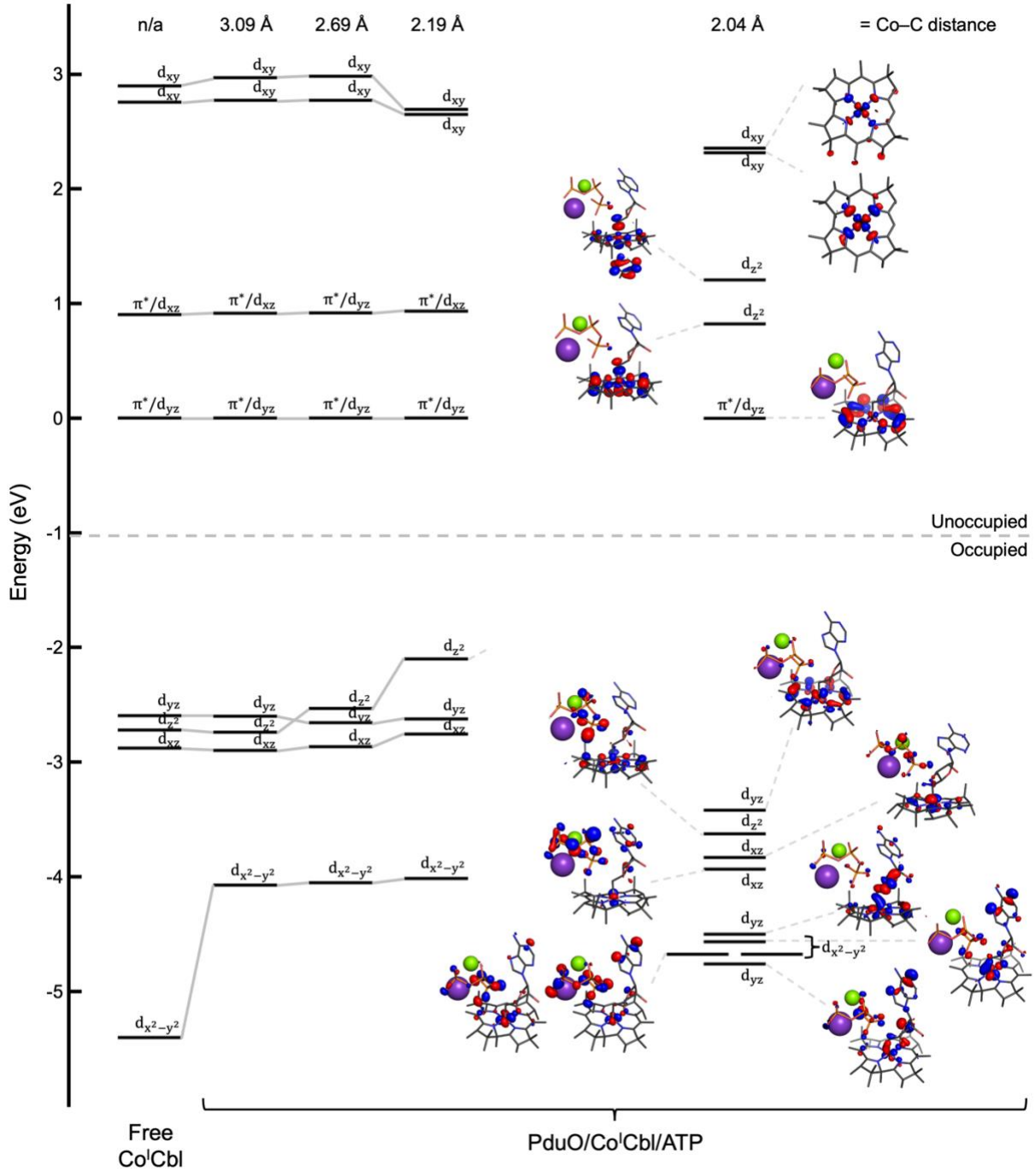


Figure S1. Relative energies and isosurface plots of the key MOs of Co^{I}Cbl, PduO/Co^{I}Cbl/ATP, and three structures (steps 8, 18, and 21) along the Co–C bond formation PES. For clarity, ATP, F112, and the two cations are hidden unless the MO contains a sizable contribution from those species. Energies are relative to that of the corresponding lowest energy unoccupied corrin-based MO.

Chapter 6

The Crystal Structure of Cysteamine Dioxygenase Reveals the Origin of the Large Substrate Scope of This Vital Mammalian Enzyme

Contents

Table S1. Data collection and refinement statistics

Wavelength (Å)	1.127	
Space group	P 21 21 21	
Unit cell	a=54.296 Å, b=139.525 Å, c=142.007 Å, α=β=γ=90°	
Data collection statistics		
	overall	last shell
Resolution range (Å)	39.17 – 1.89	1.958 – 1.89
Total reflections	1144403	108025
Unique reflections	86883	4059
Multiplicity	13.2	12.6
Completeness (%)	87.21	47.39
Mean I/σ(I)	12.34	0.99
Wilson B-factor	34.34	
R-merge	0.1844	1.657
R-meas	0.1919	1.726
R-pim	0.05248	0.4784
CC(1/2)	0.994	0.474
CC*	0.998	0.802
Refinement statistics		
Reflections used in refinement	76076	4059
Reflections used for R-free	2519	161
R-work	0.1911	0.2749
R-free	0.2275	0.3125
CC(work)	0.954	0.753
CC(free)	0.945	0.789
Number of non-hydrogen atoms	8168	
macromolecules	7597	
ligands	10	
solvent	561	
Protein residues	965	
RMS bonds (Å)	0.01	
RMS angles (°)	0.88	
Ramachandran favored (%)	98.09	
Ramachandran allowed (%)	1.49	
Ramachandran outliers (%)	0.42	
Rotamer outliers (%)	0.36	
Clashscore	1.79	
Average B-factor (Å^2)	46.47	
Macromolecules (Å^2)	46.74	
Ligands (Å^2)	42.78	
Solvent (Å^2)	42.91	
Number of TLS groups	16	

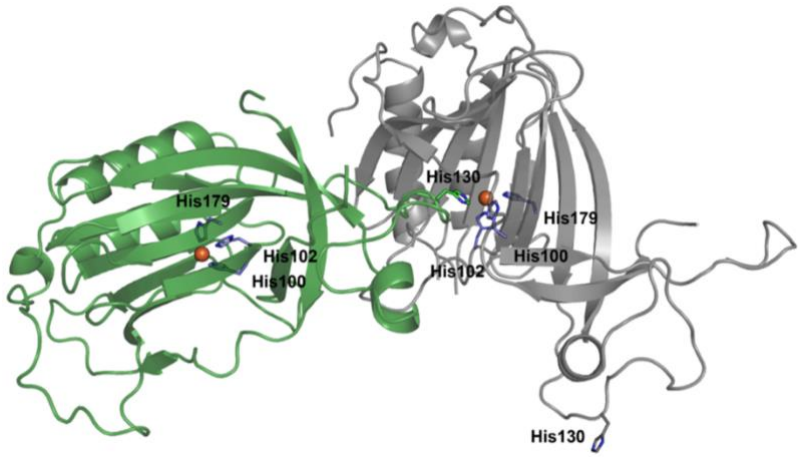


Figure S1. His130 residue of unit A is present in the active site of symmetry-related unit D.

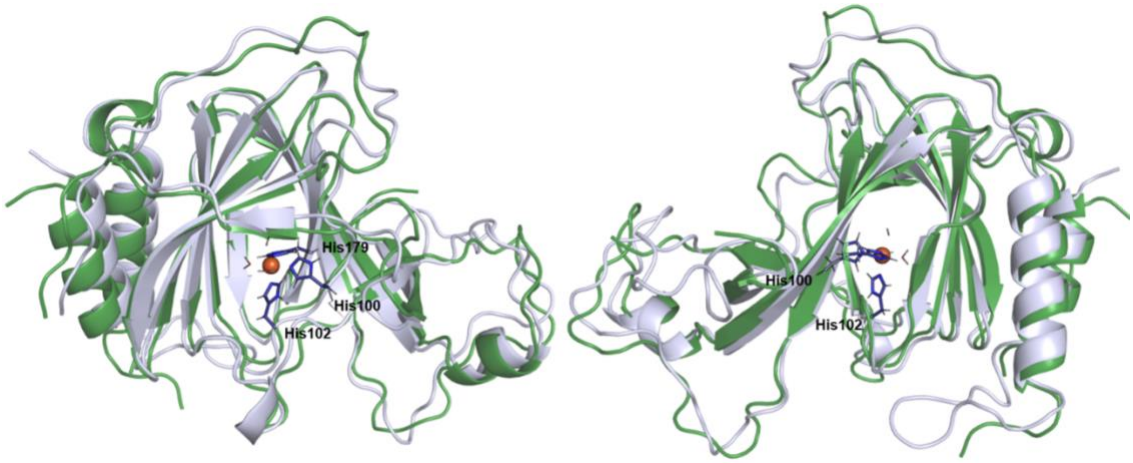


Figure S2. Comparison of the ADO X-ray crystal structure with a QM/MM optimized structure of ADO. The X-ray structure is shown in green and the QM/MM optimized model in gray.

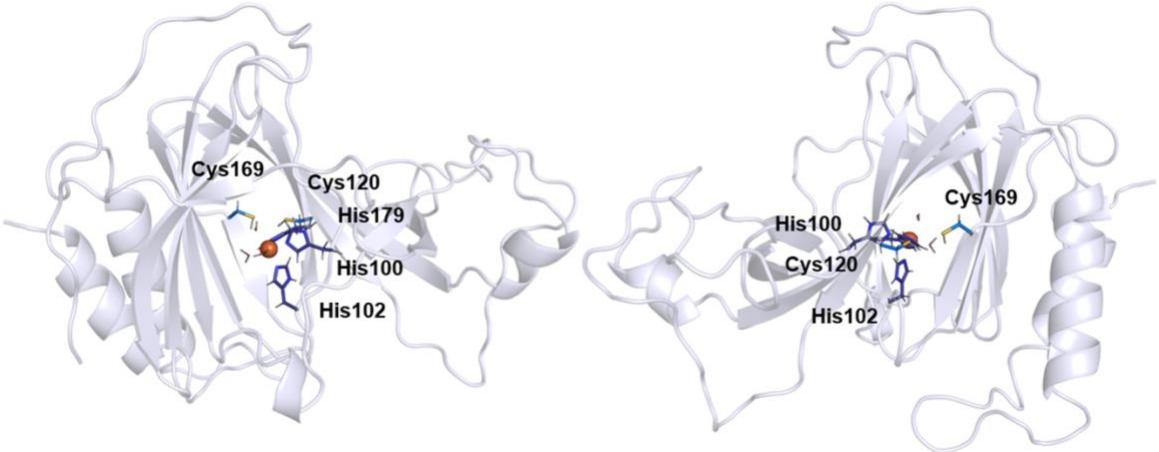


Figure S3. Front and back of the QM/MM optimized structure of ADO.

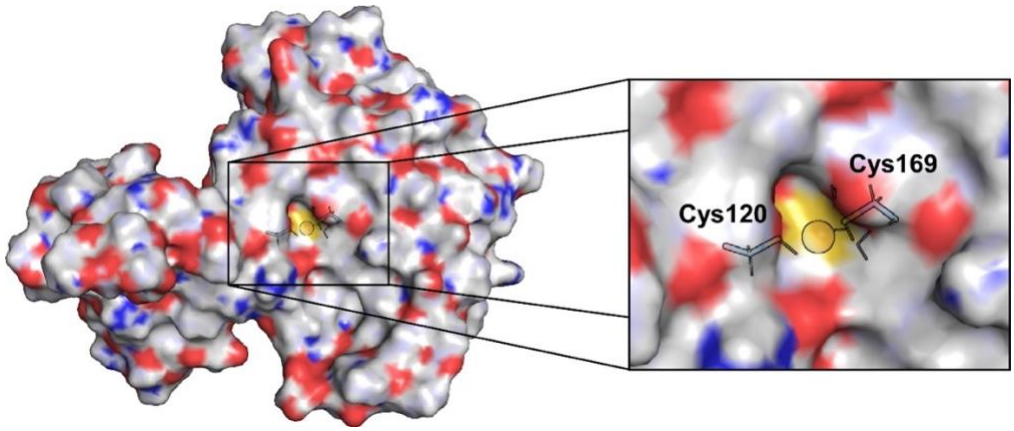


Figure S4. Surface representation of the backside of QM/MM optimized structure of ADO highlighting the co-substrate tunnel. The RBG coloring is as follows: Fe, orange; carbon, green; hydrogen white, oxygen, red; nitrogen, blue; and sulfur, yellow.

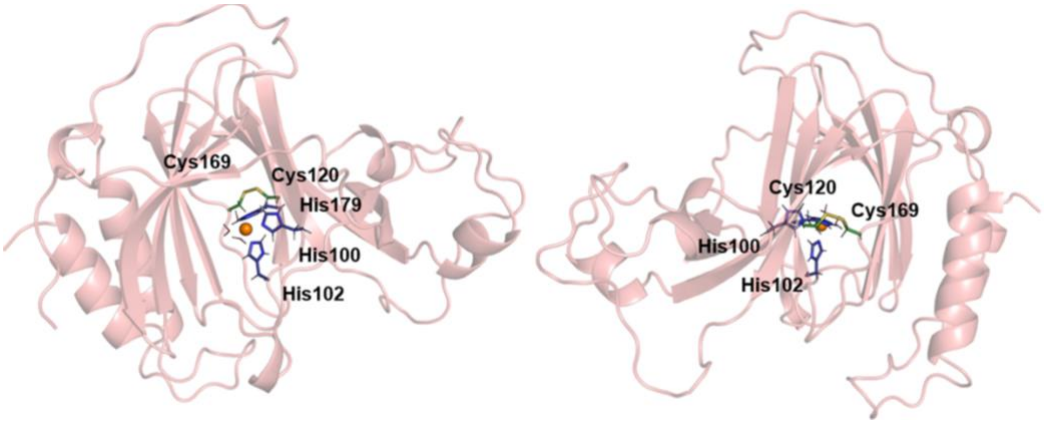


Figure S5. Front and back of the QM/MM optimized structure of a putative form of oxidized ADO. This structure has a disulfide bond between Cys120 and Cys169 (disulfide bond in yellow).

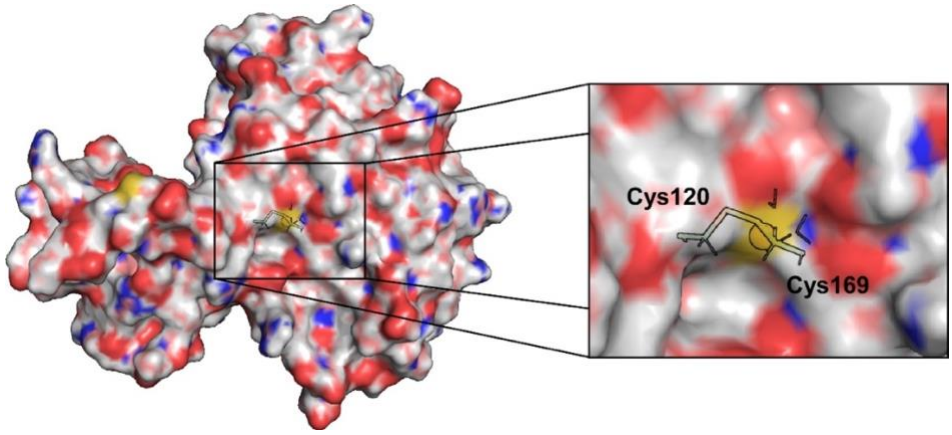


Figure S6. Surface representation of the backside of QM/MM optimized structure of oxidized ADO. A disulfide bond between Cys120 and Cys169 closes the back co-substrate tunnel. The RBG coloring is as follows: Fe, orange; carbon, green; hydrogen white, oxygen, red; nitrogen, blue; and sulfur, yellow.

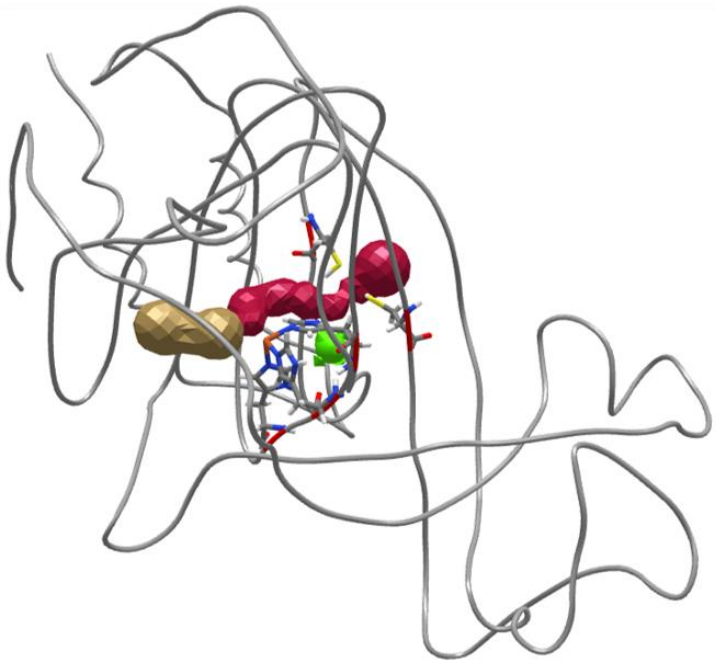


Figure S7. Tunnels calculated by MOLE2.5 into the active site of the QM/MM optimized model of peptide-bound ADO. To generate these tunnels the following parameters were used: internal threshold of 0.9 Å, bottleneck radius of 0.49 Å, and a cutoff ratio of 0.5.