Spectroscopic and Computational Insights into the Chemistry and Biochemistry of Coenzyme B_{12}

By

Ryan L. Hall

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

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2025

Date of final oral examination: 07/28/2025

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
Danel C. Fredrickson, Professor, Inorganic Chemistry
Samuel Pazicni, Assistant Professor, Chemical Education

Spectroscopic and Computational Insights into the Chemistry and Biochemistry of Coenzyme B_{12}

Ryan L. Hall

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

Abstract

Coenzyme B_{12}, or adenosylcobalamin (AdoCbl), plays a central role in a range of biochemical transformations that often involve radical-based or organometallic reaction pathways. The research outlined in this thesis integrated spectroscopic and computational approaches to investigate the structure, reactivity, and electronic properties of AdoCbl and its biologically relevant derivatives in three distinct, but thematically related contexts. The first project focused on AdoCbl as a cofactor in the bacterial enzyme ethanolamine ammonia-lyase (EAL). Using electronic absorption (Abs) spectroscopy and density functional theory (DFT) calculations, we examined the EAL ternary (AdoCbl/EAL/substrate) complex to probe Co–C bond weakening. Our results support a model in which substrate binding induces a change in the active site H-bonding network that lowers the Co–C bond dissociation energy, thereby priming AdoCbl for homolytic cleavage. The second project was aimed at characterizing the spectral properties of halocobalamins (XCbls), a series in which halide ions (X = F, Cl, Br, I) serve as upper axial ligands of Co^{3+}Cbl. Experimental Abs and CD spectra were recorded and analyzed in the framework of time-dependent DFT calculations. Clear trends in the electronic transitions and across the halide series were observed, revealing systematic shifts attributable to halide identity and Co–X bond length. These findings enhance our

understanding of axial ligand effects on corrinoid geometric and electronic structures. In the third study, we investigated the behavior of the AdoCbl-precursor aquacobalamin (H_2OCbl^+) under reducing, acidic conditions. Reduction of H_2OCbl^+ in the presence of acid leads to the evolution of hydrogen gas. While a cobalt-hydride intermediate was initially hypothesized, Abs and circular dichroism (CD) spectroscopic studies revealed that the major steady-state product more closely resembles a Co^{1+}Cbl species rather than a Co^{3+}-hydride. This unexpected outcome challenges previous mechanistic assumptions and provides new insights into cobalamin redox chemistry in acidic environments. Together, these studies offer a multifaceted view of coenzyme B_{12} chemistry and biochemistry, spanning enzyme catalysis, molecular catalysis for $H_{2(g)}$ production, and spectroscopic changes resulting from halide coordination.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Thomas Brunold, for his mentorship and endless patience throughout my time in graduate school. Your expertise, scientific insight, and dedication to rigorous thinking have shaped the way I approach research. I am especially thankful for your ability to balance constructive criticism with encouragement, pushing me to improve while reminding me of my strengths. I am grateful for the opportunities you provided and the countless discussions that have broadened my perspective.

I would also like to sincerely thank my committee members, Judith Burstyn, Danny Fredrickson, and Sam Pazicni, for their time, thoughtful feedback, and steady support. Your diverse expertise and insightful questions challenged me to think more deeply about my work. For Sam, I couldn't imagine getting to this point without your encouragement to pursue academic research when I was a bright-eyed sophomore at UNH.

To my parents, John and Eileen, thank you for teaching me the value of perseverance, curiosity, and integrity, and for never doubting my ability to reach this goal. To my brothers Sean and Matt, thank you for celebrating my milestones, coming out to visit, and reminding me that life exists outside the lab.

To my friends, both here in Madison and from all other stages of my life, thank you for your camaraderie and humor. To my lab mates, I am grateful for the collaboration, shared frustrations, and collective problem-solving that made the ups and downs more bearable. I can't imagine going through graduate school without such a supportive and kind group.

Chapter 1

Overview of B_{12} Chemistry

1.1 Cobalt in Biological Systems

The biological requirements for Co in living systems stem from its use in cobalamin (Cbl) and related species. Cbls belong to a class of molecules referred to as corrinoids, characterized by the presence of a highly modified tetrapyrrole macrocycle termed the corrin ring, which serves as a tetradentate ligand bound to a Co ion (Figure 1.1).[1] The Co ion in this complex is redox active in biological systems, with +3 (oxidized), +2 (reduced), and +1 (supperreduced) oxidation states known to be relevant. Notably, the oxidation state of the Co ion modulates its coordination number. In their oxidized (Co^{3+}) forms, corrinoids are present with a variable ligand in the upper, or Coβ, site with the major forms encountered in biological systems containing hydroxyl- (OH), aqua- (H_2O), methyl- (CH_3), or adenosyl- (Ado) moieties. The sixth ligand, located on the lower, or Coα, face is variable and depends on the local environment. It may be the dimethylbenzimidazole (DMB) base at the terminal end of an intramolecular nucleotide loop attached to the corrin ring, a protein side chain, or solvent derived. In the reduced Co^{2+}Cbl form, only the DMB base ligates to the Co ion in the axial position. In the case of reduced cobinamide (Cbi), i.e. $Co^{2+}Cbi^{+}$, a corrinoid that lacks the terminal DMB base and nucleotide loop, a solvent derived water molecule serves as the ligand at this site. In the superreduced (Co^{1+}) state, the axial ligand sites of Cbl are vacant in aqueous solution, leaving a four-coordinate Co^{1+}Cbl species. This species is among the strongest biological nucleophiles in alkylation reactions.[2,3]

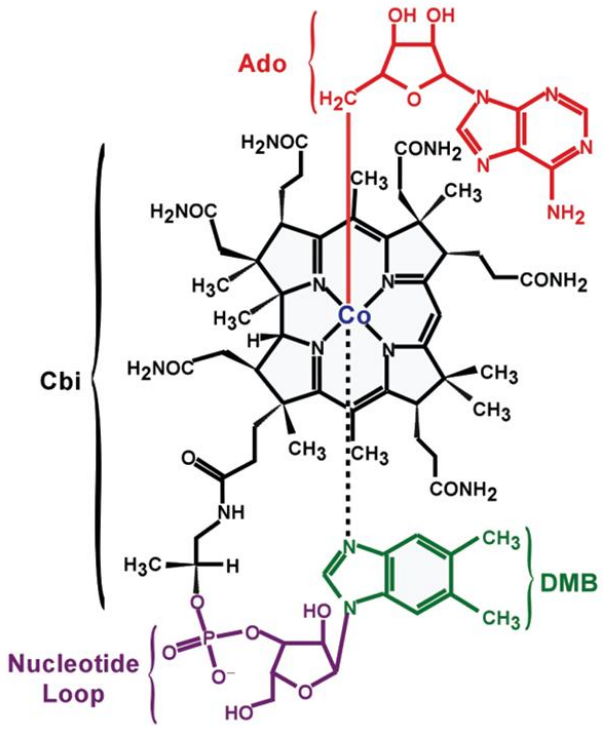

Figure 1.1. Chemical structure of adenosylcobalamin (AdoCbl), also known as coenzyme B_{12} (left). Relevant structural motifs noted in the text are highlighted in color. Abbreviations refer to Ado, adenosyl; DMB, dimethylbenzimidazole; and Cbi, cobinamide.

Two enzymatic pathways involving over 20 enzymes have been identified in bacteria and archaea devoted to the complete biosynthesis of Cbls. These convert a uroporphyrinogen-III (UroIII) precursor into AdoCbl, and differ at the point at which the Co ion is incorporated into the corrin macrocycle and whether they occur anaerobically (early insertion of Co) or aerobically (late insertion).[4,5] The elucidation of these pathways is a highlight in enzymology, as the enzymes involved carry out many unique chemical steps. These include multiple methylations of the UroIII scaffold, Co insertion, ring contraction via removal of a methylene moiety, and various side chain modifications. Addition of the axial ligands occurs via the synthesis and attachment of the nucleotide loop as well as adenosylation of the Co ion. UroIII is a common tetrapyrrole progenitor to other

biologically synthesized macrocycles including heme, chlorophyll, and the nickel-containing F430 cofactor.[4] Studies of the distribution of B_{12} across all life forms suggested that requirement of B_{12} is negatively correlated with the use of oxygen in metabolism.[5] Multicellular plants and fungi are not known to employ Cbl in their metabolism, and B_{12} auxotrophy in animals largely arises from the presence of methyl-malonyl CoA-mustase (MMCM) and methionine synthase (MetH) in their genomes. In contrast, many examples of catalytic reactions employing Cbls can be found, particularly in anaerobic bacteria and methanogens.[6,7]

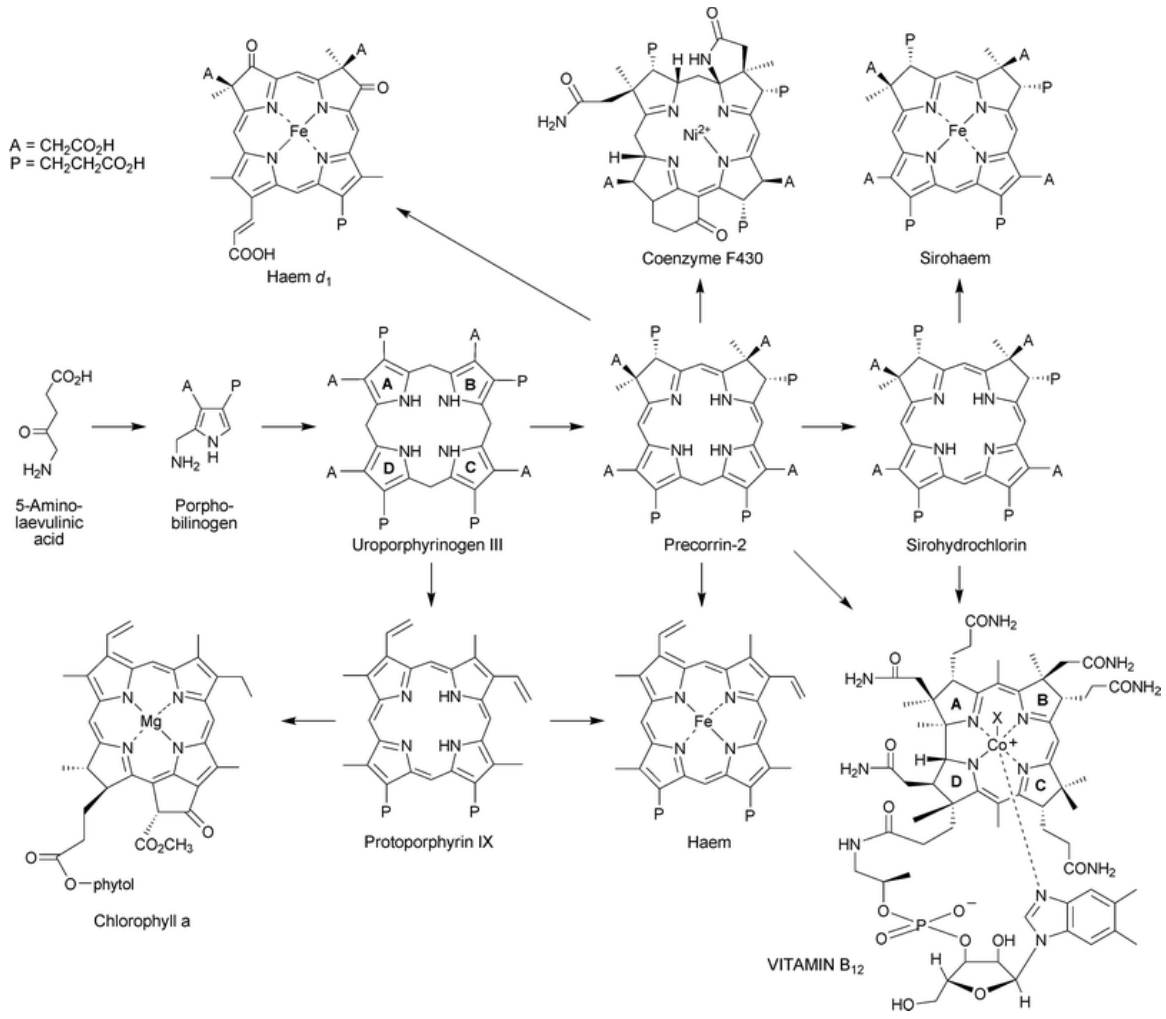

Figure 1.2. Summary of chemical intermediates observed in the biosynthesis of various tetrapyrroles. Note the common Uroporphyrinogen-III precursor in the various pathways. Adapted from ref. 4.

1.2. Biologically Active Forms of B_{12}

AdoCbl-dependent enzymes

AdoCbl-dependent enzymes are characterized by the controlled homolytic cleavage of the Co−C(Ado) bond of the bound AdoCbl cofactor in response to substrate binding. This process yields Co^{2+}Cbl and a reactive Ado-based (Ado$^{\bullet}$) radical, which is propagated to the enzyme substrate with a subsequent hydrogen abstraction to yield Ado-H. Radical mediated rearrangement of the substrate proceeds, followed by the formation of the

product after hydrogen abstraction from Ado-H to regenerate Ado•.[7,8] Recombination of Ado• with Co^{2+}Cbl to form AdoCbl completes the catalytic cycle (Figure 1.3). AdoCbl-dependent enzymes catalyze a variety of reactions, such as dehydration or deamination of substrates to yield aldehydes (diol dehydratase, glycerol dehydratase, and ethanolamine ammonia lyase (EAL)), the migration of primary amine groups (D-ornithine 4,5-aminomutase and L-leucine 2,3-aminomutase) and various carbon skeleton rearrangements (MMCM and glutamate mutase (GM)).[9] Two sub-classes of these enzymes exist, termed Class I or Class II, depending on whether or not the DMB remains ligated to the Co ion in the Co*α* site during enzyme turnover. Relative to other steps in the enzyme mechanism, particular attention has been devoted at elucidating the approach by which the enzymes facilitate the homolysis of the Co−C(Ado) bond. Notably, while the Co−C(Ado) bond of AdoCbl in aqueous solution is relatively stable (with an approximate bond dissociation energy of 31.5 kcal/mol, comparable to weak covalent bonds),[7] a significant weakening of this bond occurs in the active sites of these enzymes (by ~ 17 kcal/mol),[7,10] resulting in an acceleration of the homolytic cleavage rate of this bond by 12 orders of magnitude. The reverse reaction is essentially barrierless as it involves the recombination of two radical species to form a stable bond; thus, the overall equilibrium between bound and dissociated states depends on their relative free energies. This observation suggested two possible approaches upon which bond homolysis can be favored in the presence of the enzyme, one involving destabilization of the AdoCbl species, and another the stabilization of the dissociated radical products.[6]

In the case of MMCM and GM, crystallographic data have revealed that the AdoCbl cofactor is present in a unique conformation where the DMB is dissociated from the Co

ion, and a protein-derived His residue occupies the lower axial site instead (described as a base-off/His-on conformation and a defining feature of Class I enzymes).[11,12] An Asp residue interacts with the His moiety via a hydrogen bond. Alongside a downstream Gly moiety, these make up the characteristic $DXHX_2G$ catalytic triad of these enzymes.[6,13] Low-temperature absorption and magnetic circular dichroism (MCD) spectroscopic studies of MMCM and GM indicated that the electronic structure of AdoCbl is largely unperturbed in the presence of the enzyme loaded with substrate.[10] Furthermore, resonance Raman (rR) experiments of AdoCbl bound to MMCM with substrate did not indicate any changes consistent with a weakening of the Co−C(Ado) bond.[14] In contrast, MCD data obtained for Co^{2+}Cbl bound to MMCM and GM revealed significant changes to the electronic structure of this species relative to free Co^{2+}Cbl, attributed to the stabilization of the enzyme-cofactor complex based on computational studies.[15–17] In the context of the crystal structure, these spectroscopic results suggested that the $DXHX_2G$ triad is involved in the stabilization of Co^{2+}Cbl possibly by modulation of the Co–N(His) interaction via changes to the protonation state of the His and Glu residues.[12] Additional experimental studies of these enzymes, focusing on the interactions of the Ado$^{\bullet}$ moiety with the enzyme active site, highlighted the role of polar interactions present in the Ado-binding pocket in stabilizing the dissociated state in the presence of substrate. Thus, the overall experimental data available for Class I enzymes indicated that the stabilization of the post-homolysis products is a critical contributor to the large enzymatic acceleration of Co−C(Ado) bond homolysis.

In the case of Class II enzymes, the AdoCbl cofactor retains its DMB ligand when bound to the enzyme active site. This class of enzymes includes EAL, which catalyzes

the formation of acetaldehyde and ammonia from ethanolamine. Therefore, the mechanism of rate acceleration of Co–C bond homolysis in Class I enzymes must necessarily differ from the mechanism employed by Class I Ado-dependent eliminases due to the difference in axial ligand coordination.

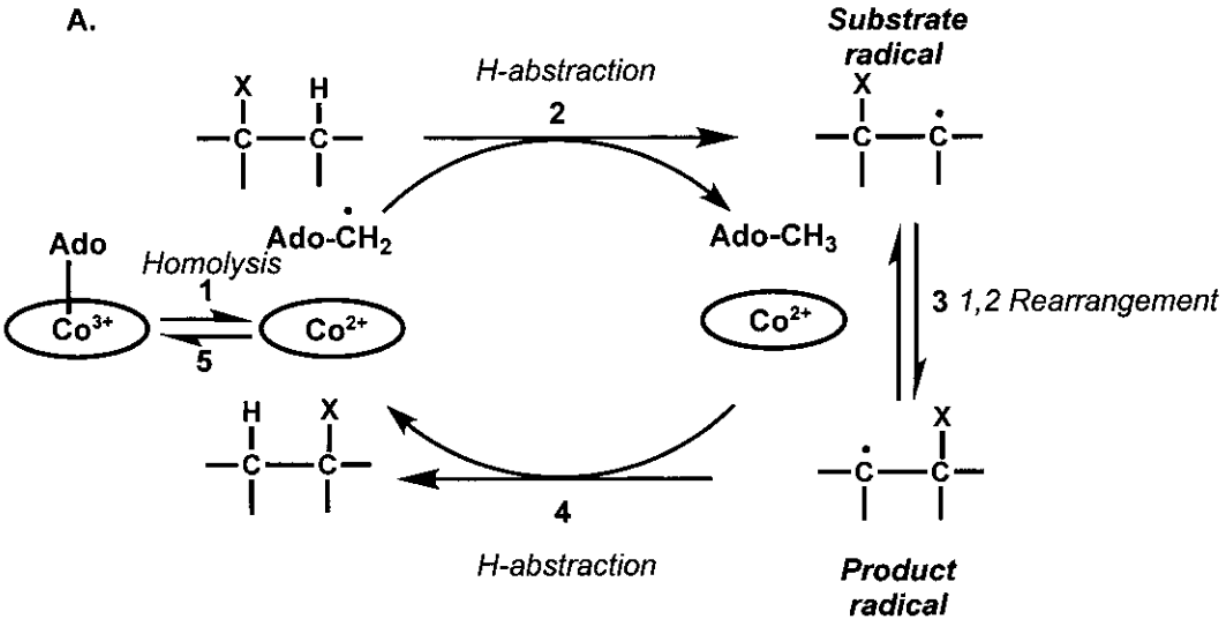

Figure 1.3. Generalized reaction cycle of AdoCbl-dependent enzymes. Various steps in the catalytic mechanism are highlighted with numbers. Although the lower axial ligand is not shown, it remains bound to the Co ion along the catalytic cycle. Note that ribonucleotide reductases follow a different overall mechanism, but share the initial AdoCbl homolysis and final radical recombination steps (steps 1 and 5). The migrating group, X, can constitute carbon, hydroxyl, or amine moieties. Adapted from ref. 7.

(ii) MeCbl-dependent enzymes: Various B_{12}-dependent enzymes involved in the transfer of methyl groups have been identified in a variety of organisms. In contrast to AdoCbl-dependent enzymes, methyltransferases cycle between the Co^{3+} and Co^{1+} oxidation states of their Cbl cofactors, concomitant with changes in the coordination environment from hexacoordinate to tetracoordinate, respectively.[8] The most extensively studied enzyme in this class is MetH, present in numerous organisms, including humans.

The catalytic cycle of this enzyme involves the reaction of methyltetrahydrofolate (CH_3-THF) with Co^{1+}Cbl to generate THF and MeCbl. Subsequent reaction of MeCbl with homocysteine to regenerate Co^{1+}Cbl yields the methionine product.[6,7] Homocysteine is required in mammals for the biosynthesis of cysteine, mediated by cystathionine β-synthase (CBS) and cystathionine γ-lyase. Methionine is particularly important in maintaining the levels of S-adenosylmethionine (SAM), an important biomolecule in a variety of biological processes including DNA methylation.[18] Furthermore, because the methyl group that MetH transfers to homocysteine originates from CH_3-THF, this enzyme is involved in the modulation of the concentration of folate derivatives in the cell. Thus, MetH can be considered a methyl group reservoir, playing a central role in coupling the folate and methionine cycles. Spectroscopic studies of a MetH variant from *Escherichia coli*, where the His residue that serves as the ligand to the Co ion in the wild-type enzyme was replaced by a Gly residue, indicated that the His ligand is critical for controlling the equilibrium among the various Cbl states encountered during turnover.[19]

References

(1) Nomenclature of Corrinoids - IUPAC-IUB Comission on Biochemical Nomenclature. *Biochem. J.* **1975**, *147* (1), 1–10.

(2) Schrauzer, G. N.; Deutsch, E. Reactions of Cobalt(I) Supernucleophiles. The Alkylation of Vitamin B_{12s}, Cobaloximes(I), and Related Compounds. *J. Am. Chem. Soc.* **1969**, *91* (12), 3341–3350.

(3) Schrauzer, G. N.; Deutsch, E.; Windgassen, R. J. The Nucleophilicity of Vitamin B_{12}s. *J. Am. Chem. Soc.* **1968**, *90* (9), 2441–2442.

(4) Warren, M. J.; Raux, E.; Schubert, H. L.; Escalante-Semerena, J. C. The Biosynthesis of Adenosylcobalamin (Vitamin B_{12}). *Nat. Prod. Rep.* **2002**, *19* (4), 390–412.

(5) Roth, J. R.; Lawrence, J. G.; Bobik, T. a. Cobalamin (Coenzyme B_{12}): Synthesis and Biological Significance. *Annual review of microbiology*. **1996**, pp 137–181.

(6) Gruber, K.; Puffer, B.; Kräutler, B. Vitamin B_{12}-Derivatives-Enzyme Cofactors and Ligands of Proteins and Nucleic Acids. *Chem. Soc. Rev.* **2011**, *40* (8), 4346–4363.

(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.

(8) Ludwig, M. L.; Matthews, R. G. Structure-Based Perspectives on B_{12} -Dependent Enzymes. **1997**, 269–313.

(9) *Enzyme-Catalyzed Electron and Radical Transfer*; Kluwer Academic Publishers: New York, 2000.

(10) 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. *Coord. Chem. Rev.* **2009**, *253* (5–6), 779–794.

(11) 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.

(12) Froese, D. S.; Kochan, G.; Muniz, J. R. C.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the Human GTPase MMAA and Vitamin B_{12}-Dependent Methylmalonyl-CoA Mutase and Insight into Their Complex Formation. *J. Biol. Chem.* **2010**, *285* (49), 38204–38213.

(13) Drennan, C.; Huang, S.; Drummond, J.; Matthews, R.; Lidwig, M. How a Protein Binds B_{12}: A 3.0 A X-Ray Structure of B_{12}-Binding Domains of Methionine Synthase. *Science.* **1994**, *266* (5191), 1669–1674.

(14) 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.

(15) 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.

(16) 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.

(17) 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.

(18) Voet, D.; Voet, J. Biochemstry.

(19) Liptak, M. D.; Fleischhacker, A. S.; Matthews, R. G.; Brunold, T. C. Probing the Role of the Histidine 759 Ligand in Cobalamin-Dependent Methionine Synthase. *Biochemistry* **2007**, *46* (27), 8024–8035.

CHAPTER 2

Spectroscopic and Computational Insights into the Mechanism of Cofactor Cobalt–Carbon Bond Homolysis by the Adenosylcobalamin-Dependent Enzyme Ethanolamine Ammonia-Lyase

This chapter has been published under the following: Elmendorf, L. D.; Hall, R. L.; Costa, F. G.; Jorge C. Escalante-Semerena, J. C.; Brunold, T. C. Spectroscopic and Computational Insights into the Mechanism of Cofactor Cobalt−Carbon Bond Homolysis by the Adenosylcobalamin-Dependent Enzyme Ethanolamine Ammonia- Lyase. *Journal of the American Chemical Society* **2025,** *147* (3), 2380-2392. https://doi.org/10.1021/jacs.4c11488

2.1 Introduction

Adenosylcobalamin (AdoCbl), also known as coenzyme B_{12}, is an organometallic cofactor vital to enzymatic isomerization and elimination reactions. 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 corrin 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 so these cofactors possess rare biological examples of an organometallic bond.

The Co–C bond lies at the heart of AdoCbl's function in B_{12}-dependent enzymes. Homolytic cleavage of this bond generates Co^{II}cobalamin (Co^{II}Cbl) and an organic radical centered on the 5'-carbon (5'C) of the adenosyl moiety (Ado•), the latter of which initiates a radical-based substrate isomerization reaction by abstracting a hydrogen atom from the substrate (see Figure 2.1).[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.[2,3] 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.

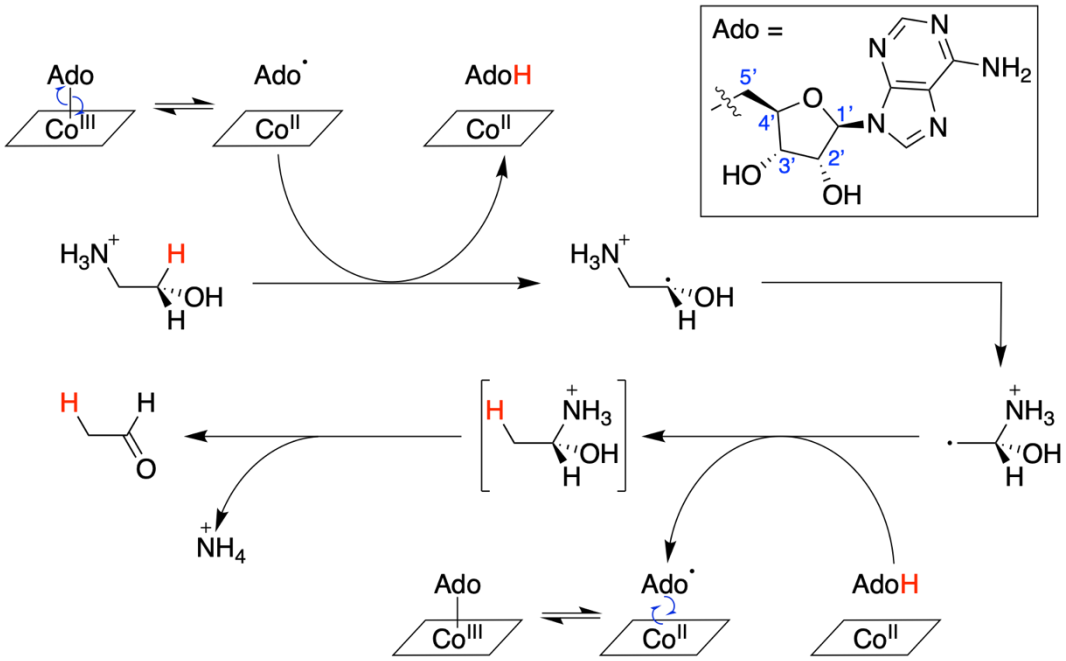

Figure 2.1. Reaction mechanism for EAL-catalyzed conversion of ethanolamine to acetaldehyde and an ammonium ion. The chemical structure and relevant atom numbering of the adenosyl (Ado) moiety is shown in the top right corner.

One such AdoCbl-dependent enzyme is ethanolamine ammonia-lyase (EAL), a Class II eliminase that catalyzes the conversion of ethanolamine to acetaldehyde and an ammonium ion (Figure 2.1).[4] As a result, EAL enables certain bacteria to use ethanolamine as the sole source of carbon and nitrogen.[5,6] 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 indicated a ~9 Å distance between Co^{II} and the substrate-like radical species in EAL,[7] compared to 6–7 Å in Class I isomerases.[8,9] EPR spectroscopy and related methods have afforded an impressively detailed understanding of the adenosyl radical (Ado•) trajectory following Co–C bond homolysis, showing that the substrate

remains largely stationary relative to the active site while the Ado• 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).[10–14] 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 5'C sufficiently close to the substrate.[15–17] 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[18–20] and X-ray crystallographic[21,22] studies. In contrast, Class I AdoCbl-dependent isomerases bind the cofactor in the so-called base-off/histidine (His)-on conformation, whereby the DMB moiety is displaced by a protein-derived His that participates in a network of conserved H-bonding residues. This conformation offers Class I isomerases increased control over the cofactor's electronic structure because the donor strength of the axially coordinated His residue can be modulated through H-bonding interactions. Based on the negligible spectral changes upon binding of AdoCbl to Class I isomerases observed using electronic absorption (Abs), magnetic circular dichroism (MCD), and resonance Raman (rR) spectroscopies, we concluded that, for this class of enzymes, the ground state of AdoCbl is largely unperturbed both in the absence and presence of substrate.[23–26] 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.[26,27] 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 must be fundamentally different. A circular dichroism (CD) spectroscopic study revealed that EAL does not undergo any major conformational changes upon binding of AdoCbl and substrate.[28] Similarly, by using time-resolved Abs spectroscopy in combination with pulsed-laser photolysis of the AdoCbl/EAL/substrate "ternary" complex, no significant perturbation of the AdoCbl ground state or stabilization of the Co^{II}Cbl/Ado$^{\bullet}$ radical pair was observed.[29] These data suggested that more subtle cofactor/enzyme/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 (with an expected Co–C bond of about 2.0 Å), likely due to the susceptibility of the Co–C bond to cleavage by free electrons generated during X-ray data collection.

Warncke and coworkers performed a detailed investigation of the formation of the Co^{II}Cbl/substrate radical pair in EAL by using time-resolved continuous-wave EPR spectroscopy in a cryosolvent system.[2] Notably, at signal-to-noise ratios of 10^3, the

Co^{II}Cbl/Ado• radical pair was not detected in these EPR spectra, indicating that the free energy of this species is at least 3.3 kcal/mol higher than that of the Co^{II}Cbl/substrate radical pair. Thermodynamic parameters of ΔH = 10.8 ± 0.8 kcal/mol and ΔS = 45 ± 3 cal/mol/K were determined for the transition from the ternary complex to the Co^{II}Cbl/substrate radical pair using Van't Hoff analysis. These data suggest that EAL compensates for AdoCbl's large, unfavorable Co–C bond dissociation energy through both enthalpic and entropic contributions. However, in a subsequent study, the same group reached the conclusion that the activation enthalpy for Co−C bond cleavage in EAL is 32 ± 1 kcal/mol, roughly equal to that of the cleavage reaction in solution.[30] Instead, they attributed EAL's acceleration of Co–C bond homolysis to a large, favorable activation entropy, at 61 ± 6 cal/(mol·K) in the enzyme compared to 7 ± 1 cal/(mol·K) in solution. More recently, Warshel and coworkers explored the origin of the entropic contributions to the Co–C bond activation mechanism employed by EAL using empirical valence bond calculations.[31] Their results suggested that the major contributions to the catalytic effect were due to the electrostatic interactions among the substrate, the ribose of the Ado group, and the active site residues. These interactions were proposed to increase in magnitude upon Co–C bond cleavage, leading to weaker interactions with the second solvation shell and thus allowing that shell to fluctuate more, which generates a larger entropic effect.

Kinetic studies in conjunction with site-directed mutagenesis have shed some light on active-site residues and enzyme/cofactor/substrate interactions crucial for EAL catalytic activity. Substitutions of an active site glutamate (E287α) shown by X-ray crystallography to interact with the substrate and Ado moiety greatly diminished the

turnover number and substrate affinity.[32] In another study,[33] a series of EAL variants with a single substitution of an active-site residue believed to hydrogen bond with the substrate (R160α, Q162α, N193α, E287α, or D362α) were used to propose a role for 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.[34] 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.[35]

In the present study, we conducted a spectroscopic and computational investigation of the mechanism by which EAL activates the AdoCbl Co–C bond in the presence of substrate. Abs, MCD, and rR spectroscopy were used to investigate changes in the AdoCbl geometric and electronic structures upon cofactor binding to the enzyme and the addition of substrate. Whole-protein models of AdoCbl/EAL both in the absence and presence of substrate were generated using molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) methods. Experimental validation of these models was achieved using published X-ray crystal structures and by comparing simulated Abs spectra obtained with time-dependent density functional theory (TD-DFT) computations to experimental spectroscopic data. A relaxed potential energy scan along the Co–C bond stretching coordinate was then performed to identify key interactions between cofactor, substrate, and active-site site residues that trigger Co–C bond homolysis and aid in the repositioning of the Ado$^\bullet$ moiety for H atom abstraction. This

scan also allowed for the calculation of Co–C bond dissociation energies (BDEs) for free AdoCbl and enzyme-bound AdoCbl in the absence and presence of substrate, with results that agree remarkably well with experimental data.

2.2 Materials and Methods

Materials

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

Protein Preparation and Purification

Plasmid pEUT31 encoding *Salmonella enterica* EAL was transformed into chemically competent *E. coli* BL21-CodonPlus™ cells (Agilent) and cells were grown overnight on LB agar containing chloramphenicol (25 μg/mL) and ampicillin (100 μg/mL) at 37 °C.[36] The colonies were scraped into a homogenous mixture and 1-mL of this cell suspension was added to a 4 L Erlenmeyer flask containing 2 L of terrific broth (TB) medium (12 g/L tryptone, 24 g/L yeast extract, 0.4% v/v glycerol, 2 mM $MgSO_4$, 17 mM KH_2PO_4, 72 mM K_2HPO_4) supplemented with chloramphenicol (25 μg/mL) and ampicillin (100 μg/mL). Metal springs were placed in the flask to increase aeration. Flasks were incubated at 25 °C shaking at 150 RPM. At an optical density (OD) of 0.5 – 0.6 (600 nm), isopropyl b-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Flasks were incubated for 18 h at 25 °C and the cells were harvested by centrifugation at 6,000 ×*g* for 15 min in a refrigerated Avanti J-20 XPI equipped with a JLA 8.1 rotor. Cell pellets were stored at -80 °C until used. The cell pellet was suspended in Buffer A (Tris•HCl, 50 mM, pH 7.5 at 4 °C, Fisher Scientific; glycerol 10% v/v) containing lysozyme (1 mg/mL, Millipore Sigma), DNase (0.1 mg/mL, Millipore Sigma), phenylmethylsulfonyl

fluoride (PMSF, 1 mM, Fisher Scientific) at a 1 mL:1 g ratio of buffer:cell pellet for 30 min at 4 °C, and lysed by sonication (2 s on, 2 s off, 30 s) twice with a five minute rest. Lysed cells were centrifuged at 14,000 ×*g* for 30 min at 4 °C in an Avanti J-25I centrifuge equipped with a JA 25.50 rotor. The supernatant was decanted, and the pellet was resuspended in Buffer B (Tris•HCl, 50 mM, pH 7.5 at 4 °C) at an equal volume and incubated at 4 °C for 30 min. The cell suspension was centrifuged at 40,000 x *g* at 4 °C in an Avanti J-25I centrifuge equipped with a JA 25.50 rotor for 30 min. The supernatant was decanted into a beaker, and ground ammonium sulfate was slowly added to the supernatant, while stirring, until the solution reached a final concentration of 25% (w/v) ammonium sulfate. The cloudy suspension was centrifuged at 10,000 ×*g* for 30 min, and the supernatant was decanted. 1.5 mL of Buffer B was added to the pellet, and the pellet was incubated rocking overnight at 4 °C. The suspended pellet was transferred to dialysis tubing with a 6-8 kDa molecular weight cut-off and the protein was dialyzed in Buffer B three times against 1 L of Buffer B, for a minimum of three hours each. Protein was quantified using the Pierce BCA assay with the microtiter plate procedure and bovine serum albumin as the protein standard. The protein was stored by flash freezing pellets in liquid nitrogen and kept at -80 °C. A gel image demonstrating the purity of the protein preparation used in this study is shown in Figure S2.1.

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 prepared using 50 mM tris(hydroxymethyl)aminomethane (Tris) buffer with 1 mM TCEP (pH 7.5). MCD samples additionally contained 60% (v/v)

polyethylene glycol 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 ~100-fold molar excess of substrate.

All samples containing AdoCbl were handled under red light to prevent premature photolysis of the Co–C bond. For collecting MCD spectra of EAL containing Co^{II}Cbl, samples were prepared in an oxygen-free glovebox, 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 in AdoCbl) into N_2(l).

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.

CD and 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 aligned parallel and antiparallel to the direction of light propagation to eliminate contributions from the natural CD (and from glass strain for frozen samples).

Computational models

Whole-protein 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).[22] Molecular dynamics (MD) simulations were performed using the GROMACS 2019.6 software with the AMBER ff14SB force field,[37] supplemented by AMBER-compatible parameters developed specifically for cobalamins.[38,39] A detailed description of how each structure was prepared for MD simulations is provided in the Supporting Information. In brief, each model was solvated in a 103.5 nm^3 TIP3P[40] water box and subjected to an 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.[41] 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.[42] The QM region was defined as the corrin ring, the entirety of the α-acetamide side chain that interacts with the Ado ligand in the post-homolysis state,[22] the first carbon of all other side chains, the DMB ligand, and the entire Ado moiety, as well as the entire side chains of active site residues N193α, S247α, D362α, the polar heads and first aliphatic

carbon atoms of active site residues R160α, Q162α, E287α, relevant water molecules, and the substrate when present. These atoms were treated with the B3LYP hybrid functional[43,44] and Grimme's D3 empirical dispersion corrections,[45] using the TZVP basis set[46] for Co and its six coordinating atoms and the 6-31G* basis set[47] for all other QM atoms. All atoms in the MM region, including additional water molecules and the remainder of the EAL a/b dimer, were modeled with the AMBER force field[48] as implemented in Gaussian 16 and the AMBER-compatible cobalamin parameters described above.[38] 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.

Single Point QM/MM Energy Calculations

A series of single point (SP) QM/MM calculations were performed for every stationary point along the PECs, using the same QM region for all three models in order 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 excluded any enzyme residues, active-site water molecules, and substrate. For these SP

computations, the larger def2-TZVP basis set[49] was used for all QM atoms along with the same functional and dispersion corrections as described above. All atoms in the MM region were modeled with the AMBER force field[48] as described above. Entropy contributions were not considered in our MD simulations and QM/MM computations, as it is not possible to calculate vibrational frequencies for such large systems and because of the way by which the models were divided into QM and MM regions.

TD-DFT Computed Abs Spectra

Abs spectra for the cofactor regions of the QM/MM-optimized models of AdoCbl, holoEAL, and the ternary complex were calculated using time-dependent density functional theory (TD-DFT) within the Tamm-Dancoff approximation[50] as implemented in the ORCA 4.2.1 software package.[51] 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,[43,44] using the def2-TZVP basis set[49] for Co and its six coordinating atoms and the def2-SV(P) basis set[49] for all other atoms. The resolution of identity (RI) approximation for electron correlation[52] was applied using the def2-SVP/C auxiliary basis set.[53] The 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}.

2.3. Results and Analysis

Spectroscopic Data

Free and EAL-bound AdoCbl

Figure 2.2 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,[28,29,32,35] 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 of the Abs features in the γ region of the EAL-bound cofactor, as has been observed previously for other AdoCbl-dependent enzymes[25,26] and was attributed to changes in the cofactor's dielectric environment upon enzyme binding. Consistent with our CD and Abs spectra, the MCD spectrum of the ternary complex only shows minor differences from that of free AdoCbl. Because at low temperatures, features arising from these diamagnetic species were obscured by temperature-dependent features from a minor fraction of paramagnetic Co^{II}Cbl, only the 250 K MCD spectra are shown in Figure 2.2. The splitting of the broad, negatively-signed MCD feature at ~20,700 cm^{-1} from free AdoCbl to the AdoCbl/EAL/2AP complex can be attributed to a slight decrease in sample heterogeneity, suggesting that the Ado moiety adopts a more constrained conformation in the EAL active site than in solution. The other MCD features are largely unchanged in shape, energy, and intensity.

Differences between free and EAL-bound AdoCbl were investigated further using resonance Raman (rR) spectroscopy (Figure S2.2). 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.[23,24,54] 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 (δ_{Co-C-C}).[54] Upon AdoCbl binding to EAL, this feature displays a minor downshift by 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.[23,24] 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 substrate.

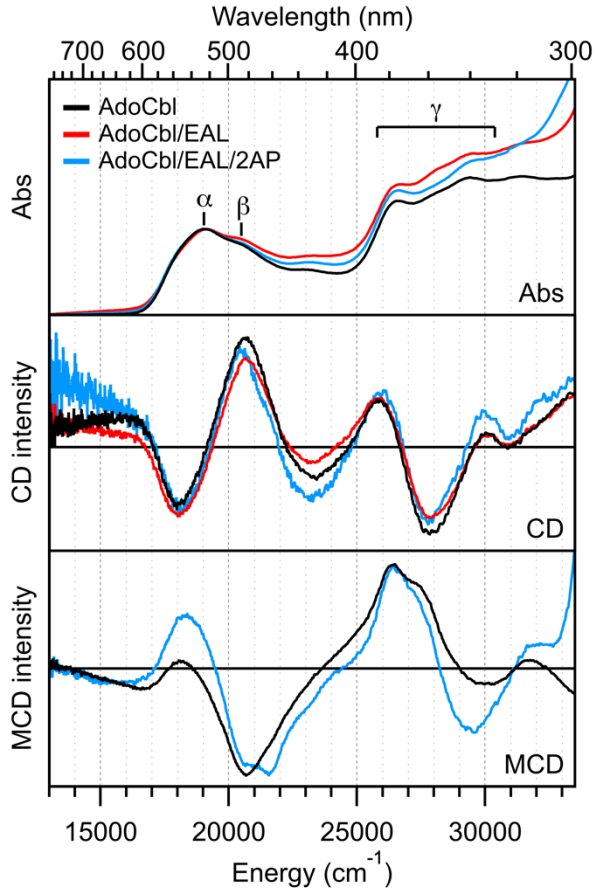

Figure 2.2. Abs (top) and CD (middle) spectra at 298 K and MCD spectra at 250 K and 7 T (bottom) of AdoCbl (black), EAL-bound AdoCbl (red), and EAL-bound AdoCbl in the presence of the slow substrate 2AP (light blue). To facilitate a comparison of the different

spectra, intensities were normalized to the most intense transition in each spectrum. Band designations are provided above the Abs spectra.

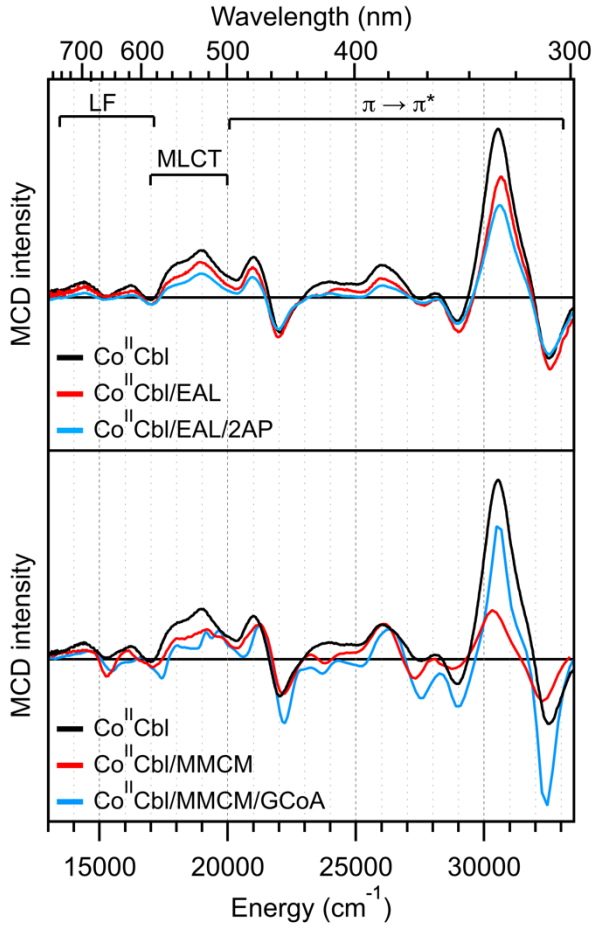

Figure 2.3. Top: MCD spectra at 4.5 K and 7 T of free Co^{II}Cbl (black), EAL-bound Co^{II}Cbl (red), and EAL-bound Co^{II}Cbl in the presence of the slow substrate 2AP (light blue). Band assignments are provided at the top. Bottom: Analogous spectra collected for free Co^{II}Cbl (black), MMCM-bound Co^{II}Cbl (red), and MMCM-bound Co^{II}Cbl in the presence of slow substrate glutaryl-CoA (GCoA) (light blue). Adapted with permission from 27. Copyright 2005 American Chemical Society.

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 2.3) and to MCD spectra of Co^{II}Cbl reported previously.[55] 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 2.3 for comparison) and GM.[26,27] 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: Co d → d [or ligand field (LF)] transitions below ~17,000 cm^{-1}, metal-to-ligand charge transfer (MLCT) transitions from 17,000 cm^{-1} to 20,000 cm^{-1}, and corrin $\pi \rightarrow \pi^*$ transitions above 20,000 cm^{-1}.[55] For MCMM and GM, Co^{II}Cbl binding to the enzyme active site and the addition of a substrate did not significantly perturb the LF and $\pi \rightarrow \pi^*$ regions of the cofactor 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 in the Co^{II} state, possibly due to proton uptake by a “catalytic triad” of residues (which includes the axial His ligand) that occurs in response to Co–C bond homolysis. This makes the His ligand an integral component of the catalytic mechanism employed by MMCM. 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 the mechanism of Co–C bond activation employed by EAL differs from that of the Class I enzymes. This finding is not surprising given that EAL binds AdoCbl in the base-on conformation and the DMB ligand is unable to participate in any H-bonding interactions.

Collectively, our data indicate that cofactor 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 turned to computational techniques to identify key interactions between the cofactor, enzyme active site, and substrate that contribute to the enormous rate acceleration of Co–C bond homolysis in the ternary complex.

Computational data

Construction of computational models

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 free 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 focused 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 a single structure of each model for subsequent QM/MM geometry optimization. The optimized models were then used for a relaxed potential energy surface scan in which the Co–C bond was elongated in 0.1 Å increments and a full QM/MM geometry optimization of all other atoms 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).

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

We first validated our combined MD + QM/MM approach using solvated free 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.[56,57] After MD simulation and QM/MM geometry optimization, the key geometric parameters of the solvated AdoCbl model agree well with those observed in the crystal structure of AdoCbl,[57] with Co–C bond lengths of 2.04 and 2.02 Å, respectively (Table 2.1 and Figure 2.4). 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 in the experimental structure.[58] To validate the computed electronic structure of solvated AdoCbl, TD-DFT was used to predict the corresponding Abs spectrum. The predicted spectrum reproduces the major experimental absorption features reasonably well (Figure 2.5), especially considering that TD-DFT fails to account for vibronic coupling that governs the Abs band shapes of B_{12} species.[59] Furthermore, the computed Co–C bond homolysis energy for this model (*vide supra*) is 27.8 kcal/mol, in excellent agreement with the experimentally-determined value of 30 ± 2 kcal/mol.[60] The fact that our DFT computations with the B3LYP functional slightly underestimate the Co–C BDE of AdoCbl is consistent with previous reports.[61,62]

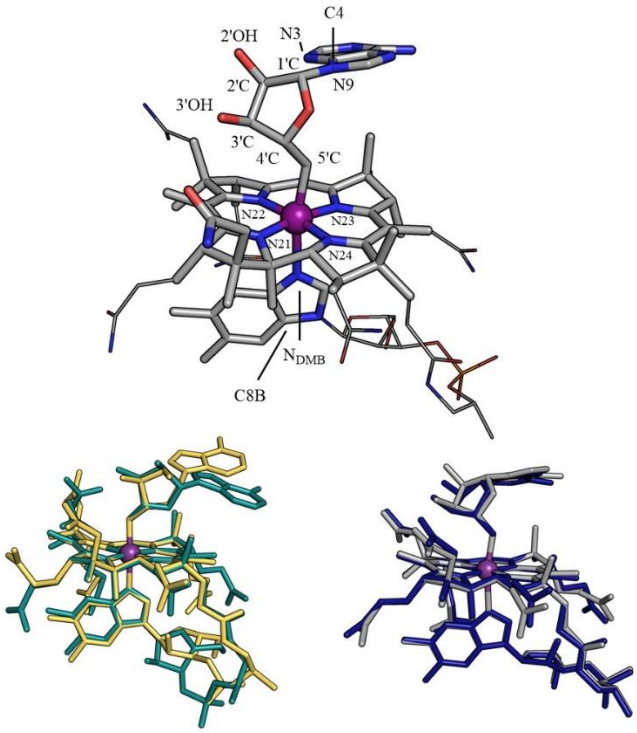

Figure 2.4. Top: AdoCbl numbering scheme with the cofactor QM region and MM region shown as thick and thin sticks, respectively. Bottom left: AdoCbl crystal structure (yellow) and QM/MM-optimized model of solvated AdoCbl (green). Bottom right: AdoCbl portions of QM/MM-optimized models of AdoCbl/EAL (blue) and AdoCbl/EAL/EA (gray).

Table 2.1. Key geometric 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.[57] [b.] Average of the two crystallographically independent AdePeCbl molecules (PDB: 3ABS).[21] [c.] To facilitate a comparison with AdoCbl, the Co-bound methylene carbon atom of AdePeCbl is referred to as the 5'C atom. [d.] DMB tilt defined as the Co–N_{DMB}–C8B bond angle.

Warranted by the close agreement between the experimental and computationally optimized AdoCbl structures, we applied the same MD + QM/MM methodology to the EAL-bound AdoCbl models. The holoEAL model, which consisted of one AdoCbl molecule bound to 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, AdoCbl binding to EAL causes a slight increase in Co–C bond length from 2.04 Å to 2.06 Å and a more substantial lengthening of the Co–N_{DMB} bond from 2.14 Å to 2.50 Å (Table 2.1 and Figure 2.4). In the absence of an EAL-bound AdoCbl crystal structure with an intact Co–C bond, it is best to compare our model of holoEAL to the crystal structure of EAL containing the cofactor derivative adeninylpentylcobalamin (AdePeCbl) with an intact

Co–C bond measuring 2.00 Å (PDB: 3ABS).[21] 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. Hence, 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 a Co–N_{DMB} lengthening are 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 substrate EA. The five substrate-binding residues identified in the crystal structure of AdePeCbl/EAL 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 the AdePeCbl/EAL/EA structure (PDB: 3ABS) and optimized ternary complex model is merely 0.55 Å (Figure S2.10). The AdoCbl structure in the optimized ternary complex model shows only minor atomic displacements from that in the optimized holoEAL model (Figure 2.4). The RMSD of the heavy atom positions of the cofactor QM region between these two AdoCbl-bound EAL models is 0.26 Å and increases to only 0.46 Å when comparing all heavy atoms of AdoCbl. Both models show a displacement of the Co^{III} ion from the corrin ring towards the Ado moiety (Figure 2.4) 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.

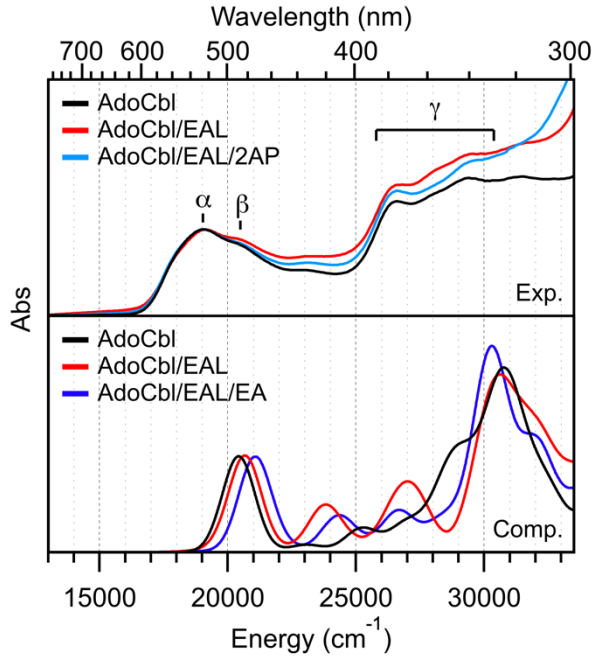

Figure 2.5. Experimental 298 K (top) and TD-DFT-computed (bottom) Abs spectra of AdoCbl (black), holoEAL (red), and the ternary complex (blue).

Further evaluation of computational models

To assess the feasibility of our computational models beyond structural comparisons, we used TD-DFT to simulate Abs spectra for the computational models of free and EAL-bound AdoCbl. Changes in the strength of the axial bonding interactions in alkylcobalamins are known to affect the position of the low-energy α/β peaks in the Abs spectrum.[56,63] These peaks correspond to the electronic origin and vibrational sideband of a corrin π (HOMO) → π* (LUMO) transition.[59] Because the HOMO also contains small contributions from Co–C σ and Co–N_{DMB} σ* bonding interactions, the energy of this transition is sensitive to axial ligand perturbations.[27,56] Thus, as the Co–C bond lengthens from solvated AdoCbl to holoEAL to the ternary complex, the Co–C σ bonding interaction

weakens and a red shift of the α/β peaks is expected.[25] Conversely, as the Co–N_{DMB} bond lengthens upon AdoCbl binding to EAL, a blue shift of the α/β peaks is expected. Consequently, on the basis of our previously established, quantitative spectro/structural correlations,[25] the simultaneous lengthening of the Co–C and Co–N_{DMB} bonds should result in relatively small 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 from free AdoCbl, by 235 cm^{-1} and 645 cm^{-1} for holoEAL and the ternary complex, respectively (Figure 2.5). 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, though other possible models cannot be ruled out solely on the basis of TD-DFT computed spectra.

Much of the current understanding of the AdoCbl Co–C bond activation mechanism employed by Class II eliminases is based on the crystal structures of EAL obtained by Toraya and coworkers.[21,22] Notably, in the holoEAL crystal structure, a Co···C distance of ~3.0 Å was observed, which was described as an intact but activated Co–C bond.[22] By performing a relaxed potential energy scan along the Co–C bond coordinate (*vide infra*) and computing the Abs spectrum for the optimized structure with a Co–C bond length of 3.0 Å using TD-DFT, drastic spectral differences from free AdoCbl are predicted for this "intact" EAL-bound AdoCbl cofactor containing such an elongated Co–C bond (Figure S2.16). The computed spectrum does not, in any way, resemble that of free AdoCbl, and is in fact much more similar to the Abs spectrum of Co^{II}Cbl,[55] with an intense band centered at ~23,000 cm^{-1} as its lowest energy feature. We therefore conclude that the X-

ray crystal structure of holoEAL published by Toraya and coworkers[22] features a cleaved Co–C bond, likely due to X-ray induced reduction during data collection. This is not unexpected, given that similarly long Co···C distances are observed in the X-ray crystal structures of other Class II, as well as Class I, AdoCbl-dependent isomerases that contain the native cofactor.[64]

PEC scans for 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 the lengthening of the AdoCbl Co–C bond to quantify the impact of EAL binding and the addition of substrate on the degree of Co–C bond activation for homolysis. The QM regions of the different models used in these PEC scans are shown in Figure 2.6. The corresponding QM energies for all stationary points during a 0.1 Å incremental Co–C bond elongation for the three AdoCbl-containing models are plotted in Figure 2.7A. Since the computational models contained a different number of atoms in the QM region (Figure 2.6), a direct comparison of the absolute QM energies between them is not possible. All QM energies in Figure 2.7A 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 the experimentally determined decrease in Co–C bond dissociation energy of ~17

kcal/mol in AdoCbl-dependent enzymes following substrate binding.[3,65] While the QM/MM computations for holoEAL failed to converge for Co–C distances beyond 3.16 Å, the PEC for this species is nearly identical to that of the solvated free cofactor. Thus, our computational data correctly reproduce the experimental finding that Co–C bond activation by EAL requires the presence of substrate.

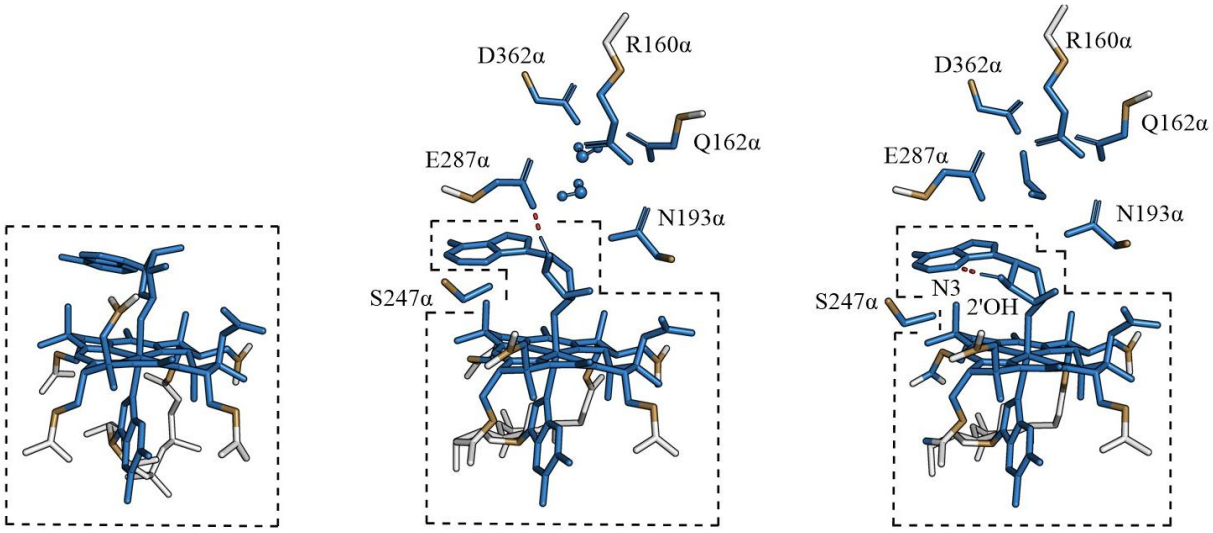

Figure 2.6. Left: solvated AdoCbl QM region. Middle: holoEAL QM region. Right: ternary complex QM region. The cofactor QM region for all three models is indicated by the dashed boxes. QM atoms are shown in blue, link atoms are shown in tan, and select MM atoms are shown in white. The H-bond involving the ribose 2'OH group is shown in red for the holoEAL and the ternary complex models. This H-bond is contained within the cofactor QM region for the ternary complex but spans the boundary of the cofactor QM region for holoEAL.

To permit a comparison of the absolute QM energies for free and EAL-bound AdoCbl, SP energies were calculated for all steps along each PEC, this time using a larger basis set (def2-TZVP) with an identical QM region for all models (hereafter referred to as the cofactor QM region, Figure 2.6). The cofactor QM energy of each QM/MM-optimized structure is plotted in Figure 2.7B, with the lowest QM energy of free AdoCbl (i.e., the QM energy of AdoCbl at the start of its PEC) subtracted as a reference. Figure 2.7B indicates that the energies of EAL-bound AdoCbl models are higher than those of

free AdoCbl, indicating that the EAL protein matrix destabilizes AdoCbl. While the destabilization of AdoCbl in the ternary complex is seemingly less pronounced than in holoEAL (4.5 and 11.7 kcal/mol, respectively), inspection of key H-bonding interactions reveals that the ternary complex contains an intramolecular H-bond within the Ado moiety, between the ribose 2'OH hydroxyl group and the adenine N3 atom, that is absent in free AdoCbl and holoEAL (Figure 2.6). 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 (Figures 2.7A and S2.17), it does not contribute to the energies of the cofactor QM regions (Figure 2.7B). To compensate for the stabilization of AdoCbl in the ternary complex due to this intramolecular H-bond, Figure 2.7B 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.

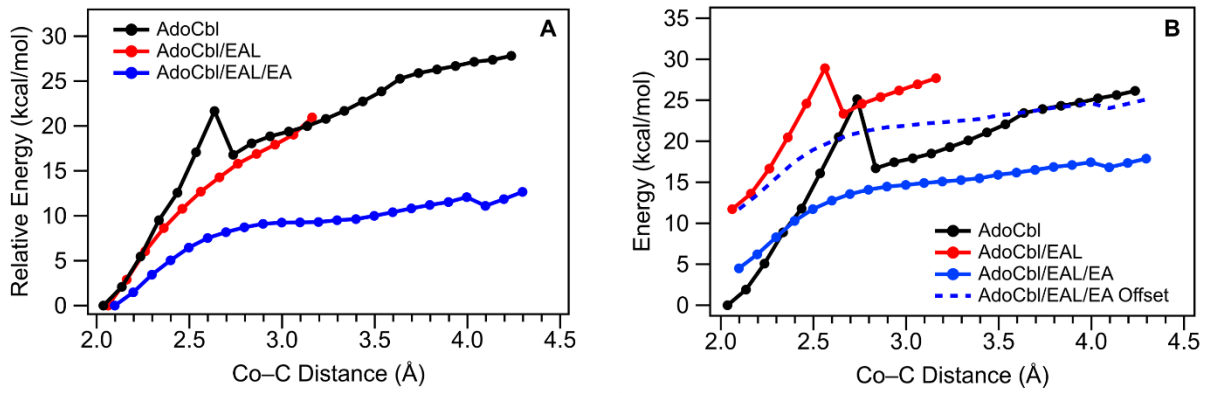

Figure 2.7. Computed PECs for Co–C bond cleavage of free and EAL-bound AdoCbl. Filled circles represent stationary points of QM/MM-optimized geometries. Discontinuities are due to a change in spin state from a closed-shell singlet at short Co–C bond distances to an open-shell singlet at larger Co–C bond distances. A: Relative energies of the entire QM regions during Co–C bond elongation obtained using the TZVP/6-31G* split basis set. B: Absolute energies of the cofactor QM regions for all stationary points in panel A using the def2-TZVP basis set for all atoms. The energy of the optimized AdoCbl model was subtracted from the energies of all other models.

While our computations predict a significant destabilization of EAL-bound AdoCbl (Figure 2.7B), 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 difference between the energies of the solvated cofactor and the cofactor in the ternary complex at a Co···C distance of >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 26.1 kcal/mol when the larger basis set is used (Figure 2.7B), still in good agreement with the experimental Co–C bond dissociation energy of 30 ± 2 kcal/mol.[60] 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 difference in cofactor QM energy increase in response to Co–C bond elongation beyond 4.30 Å between the solvated AdoCbl and ternary complex is therefore 12.7 kcal/mol, again in excellent agreement with the experimentally determined difference in Co–C bond dissociation energy of ~17 kcal/mol.[3,65] 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.

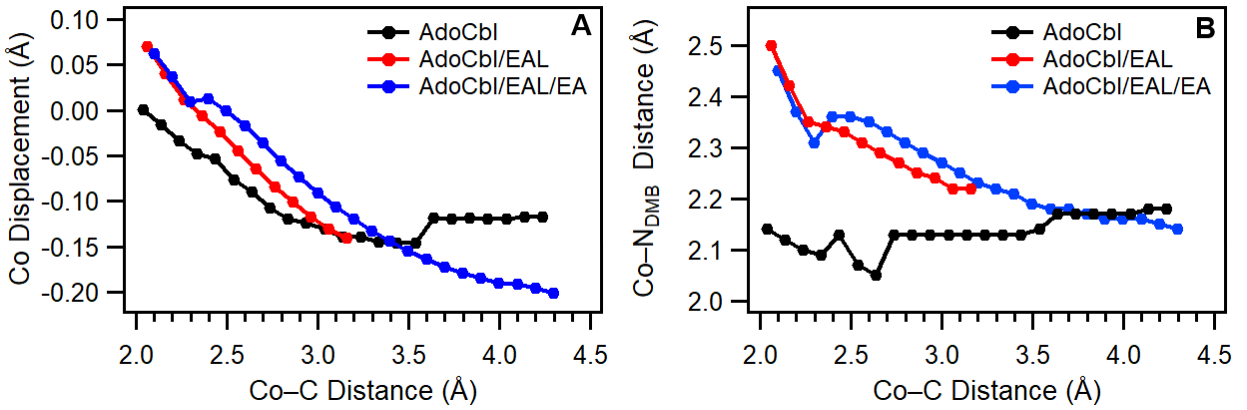

Figure 2.8. A: Co displacement from the corrin macrocycle as a function of Co–C distance. Positive values correspond to Co displacements towards the Ado moiety. B: Co–N_{DMB} bond distance as a function of Co–C distance.

To investigate the structural origin of the differences between the PECs in Figure 2.7, we first examined how key geometric parameters change as a function of Co–C bond elongation. Comparing the QM/MM-optimized models of free AdoCbl, holoEAL, and the ternary complex reveals that 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 2.1). 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 2.8A). This Co displacement is absent in the QM/MM-optimized free AdoCbl model and is therefore a consequence of EAL-induced structural changes. The observed Co displacements of 0.07 and 0.06 Å in the QM/MM-optimized holoEAL and the ternary complex models, respectively, must be energetically more favorable than an additional elongation of the Co–C bond. These results indicate that stabilizing interactions exist 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 can 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 (Figure 2.8). Alternatively, in the ternary complex, the Co–N_{DMB} distance decreases substantially, from 2.45 to 2.14 Å during incremental Co–C bond lengthening to 4.24 Å. 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.

Although 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 accelerate Co–C bond homolysis, the same trend in computed Co displacement and Co–N_{DMB} bond length is observed for holoEAL, and thus in the absence of substrate. We therefore conclude that the movement of the Co ion relative to the corrin ring to allow for the formation of a stronger Co–N_{DMB} bond after Co–C bond cleavage is not the dominant contributor 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 primarily responsible for Co–C bond activation in the ternary complex.

Substrate-induced active site changes

In the optimized holoEAL model, three H-bonding interactions exist between active site residues and the Ado moiety: N193α donates a H-bond to the ribose 3'OH, E287α accepts a H-bond from the ribose 2'OH, and S247α donates a H-bond to the adenine N3 (Figure 2.9, top). These protein/Ado interactions are maintained throughout the Co–C bond elongation process. The Ado• radical must move toward the substrate after Co–C bond homolysis, given that the substrate radical is positioned ~9 Å from the Co^{II} ion after H atom abstraction by Ado•.[7] 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.

In the ternary complex, the number of protein/Ado H-bonding interactions is reduced from three to one (Figure 2.9, bottom). 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.[22] The importance of E287α for catalysis is evidenced by the fact that the E287αX (X= D, Q, A) variants all exhibit drastically reduced rates of Co–C bond homolysis upon substrate binding.[32,33] This residue has been implicated as a dynamic "substrate trigger" for Co–C bond homolysis, moving in response to substrate binding and H-bonding to EA, thus promoting Co–C bond cleavage.[32,66] Our computational results support a slightly different mechanism of action for the substrate trigger hypothesis. Upon substrate binding, E287α no longer accepts a H-bond from the ribose 2'OH, while S247α now donates a H-bond to the *c*-acetamide side chain of the corrin ring 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 repositioning 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.[15]

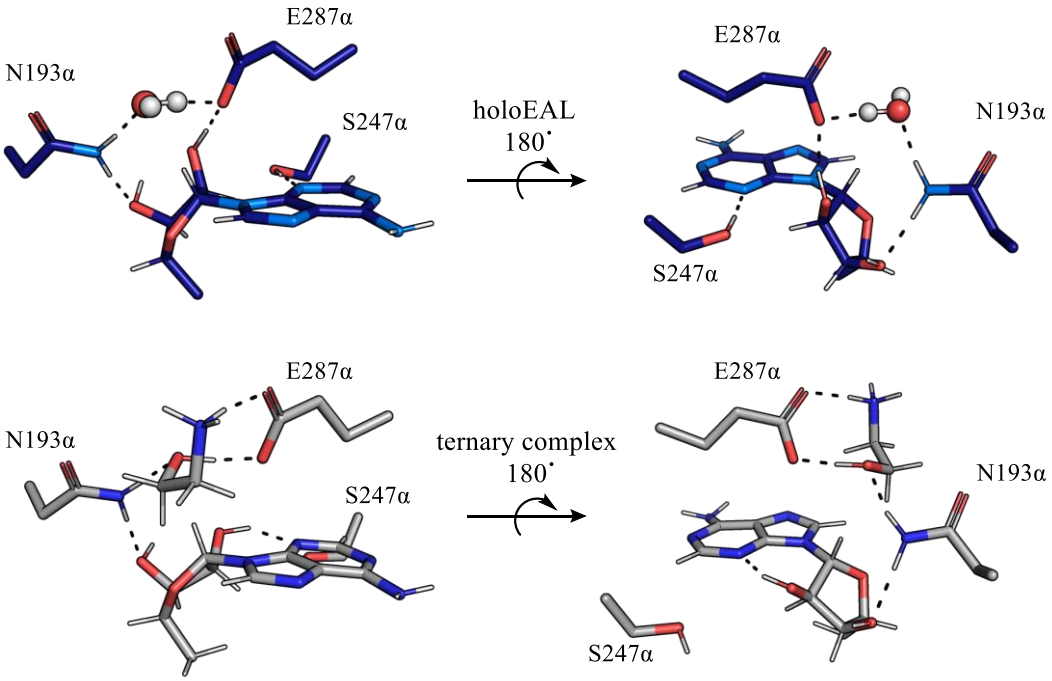

Figure 2.9. Key active site hydrogen-bonding interactions in the QM/MM-optimized models of holoEAL (top) and the ternary complex (bottom), shown from two different angles. Aliphatic hydrogen atoms on side chains are omitted for clarity. Only one of the two active site water molecules that are displaced upon substrate binding to holoEAL is shown.

Torsional strain across the N9–1'C glycosidic linkage of the Ado moiety (Figure 2.4) has been hypothesized to serve as an important contributor to the destabilization of EAL-bound AdoCbl, given that 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.[64] Our

computational models support this hypothesis, as judged on the basis of differences in the C4–N9–1'C–2'C torsional angle. While in the free AdoCbl model this angle is 149.9°, it is drastically reduced to 11.9° in our holoEAL model and to 28.4° in the model of the ternary complex (Figure 2.4), comparable to 37.1° in the AdoCbl/EAL crystal structure (PDB: 5YSN)[22]. In the crystal structure of the ternary complex with a cleaved Co–C bond (PDB: 5YSR),[22] the torsional angle is much larger, at 101.3°, than in our model of the ternary complex.

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 (Figure 2.9). Intriguingly, this H-bonding interaction is the only one expected to provide a direct link between substrate and Ado positioning, as in the ternary complex N193α additionally serves as a H-bond donor to 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 enzymatic activity while undergoing inactivation four times faster than wild-type EAL.[33]

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, S224α, was previously investigated.[67] The greatly diminished activity of the S224αA and S224αN variants led the authors to propose that H-bonding between this residue and the adenine N3 atom is important for Co–C bond activation and for preventing enzyme inactivation. This proposal is consistent with our

findings. In EAL, the H-bond donation from S247α to the 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 is expected to play a key role in the Co–C bond activation mechanism employed by EAL. Collectively, our computational models have allowed us to identify cofactor/enzyme/substrate interactions through which EAL activates the Co–C bond for homolysis upon substrate binding via destabilization of AdoCbl, with a lengthening of both axial bonds of the AdoCbl cofactor and the loss of two enzyme-cofactor H-bonds that would otherwise restrict rotation of the Ado$^{\bullet}$ radical.

2.4. 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 accelerate the rate of Co–C bond homolysis through, at least in part, stabilization of the post-homolysis products, EAL displays no appreciable stabilization of Co^{II}Cbl. Instead, our computational models of AdoCbl, holoEAL, and the ternary complex indicate that EAL destabilizes AdoCbl by elongation of the Co–axial ligand bonds and through Ado distortions, though these distortions alone are insufficient to account for the entire reduction of the barrier to Co–C bond homolysis. PECs of these structural models indicate a +12.7 kcal/mol reduction in Co–C activation barrier for enzyme-bound AdoCbl in the presence of substrate, consistent with published experimental data. Although 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 accelerate Co–C bond

homolysis, the same trend in computed Co displacement and Co–N_{DMB} bond length is observed for holoEAL, and thus in the absence of substrate. We therefore conclude that the movement of the Co ion relative to the corrin ring to allow for the formation of a stronger Co–N_{DMB} bond after Co–C bond cleavage is not the dominant contributor 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 primarily responsible for Co–C bond activation in the ternary complex. While our computational results are in good agreement with available experimental data, other factors (such as changes in activation entropy) may also play an important role in the mechanism of Co-C bond activation employed by EAL.

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.

(2) Wang, M.; Warncke, K. Kinetic and Thermodynamic Characterization of CoII-Substrate Radical Pair Formation in Coenzyme B_{12}-Dependent Ethanolamine Ammonia-Lyase in a Cryosolvent System by Using Time-Resolved, Full-Spectrum Continuous-Wave Electron Paramagnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130* (14), 4846–4858.

(3) 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.

(4) 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.

(5) Chang, G. W.; Chang, J. T. Evidence for the B_{12}-Dependent Enzyme Ethanolamine Deaminase in Salmonella. *Nature* **1975**, *254* (5496), 150–151.

(6) Roof, D. M.; Roth, J. R. Ethanolamine Utilization in Salmonella Typhimurium. *J. Bacteriol.* **1988**, *170* (9), 3855–3863.

(7) 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.

(8) 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 c Ochlearium. *Biochemistry* **1998**, *37* (12), 4105–4113.

(9) 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.

(10) 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.

(11) 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.

(12) 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.

(13) 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.

(14) Canfield, J. M.; Warncke, K. Geometry of Reactant Centers in the CoII-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.

(15) 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.

(16) Brunk, E.; Kellett, W. F.; Richards, N. G. J.; Rothlisberger, U. A Mechanochemical Switch to Control Radical Intermediates. *Biochemistry* **2014**, *53* (23), 3830–3838.

(17) 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.

(18) 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.

(19) 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,* 138-141.

(20) Ke, S. C.; Torrent, M.; Museav, D. G.; Morokuma, K.; Warncke, K. Identification of Dimethylbenzimidazole Axial Coordination and Characterization of 14N 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.

(21) 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.

(22) 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.

(23) 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**, *120*, 9947–9948.

(24) 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.

(25) 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.

(26) 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.

(27) 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.

(28) 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.

(29) 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.

(30) Wang, M.; Warncke, K. Entropic Origin of Cobalt–Carbon Bond Cleavage Catalysis in Adenosylcobalamin-Dependent Ethanolamine Ammonia-Lyase. *J. Am. Chem. Soc.* **2013**, *135* (40), 15077–15084.

(31) Schopf, P.; Mills, M. J. L.; Warshel, A. The Entropic Contributions in Vitamin B_{12} Enzymes Still Reflect the Electrostatic Paradigm. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (14), 4328–4333.

(32) 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.

(33) 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.

(34) 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.

(35) 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.

(36) Brinsmade, S. R.; Paldon, T.; Escalante-Semerena, J. C. Minimal Functions and Physiological Conditions Required for Growth of *Salmonella enterica* on Ethanolamine in the Absence of the Metabolosome. *J. Bacteriol.* **2005**, *187* (23), 8039–8046.

(37) 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.

(38) 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.

(39) 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.

(40) 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.

(41) 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.

(42) 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.

(43) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377.

(44) 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.

(45) 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).

(46) 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.

(47) 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.

(48) 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.

(49) 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.

(50) Hirata, S.; Head-Gordon, M. Time-Dependent Density Functional Theory within the Tamm-Dancoff Approximation. *Chem. Phys. Lett.* **1999**, *314* (3–4), 291–299.

(51) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22).

(52) 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.

(53) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.

(54) 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.

(55) 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.

(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.

(57) 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.

(58) 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.

(59) Elmendorf, L. D.; Brunold, T. C. Vibronic Coupling in Vitamin B_{12}: A Combined Spectroscopic and Computational Study. *Inorg. Chem.* **2023**, *62* (32), 12762–12772.

(60) Hay, B. P.; Finke, 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* (16), 4820–4829.

(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.

(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.

(63) 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. **2005**, No. 7, 16522–16528.

(64) 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* **2023**, *24* (14)

(65) 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.* **1987**, *109* (26), 8012–8018.

(66) 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.

(67) 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.

CHAPTER 3

Spectroscopic and Computational Analysis of the Halocobalamin (XCbl, X = F, Cl, Br, I) Series

3.1 Introduction

Cobalamins (Cbls) are a class of cobalt corrinoid cofactors that contain a Co ion ligated equatorially by the four nitrogens of a highly substituted macrocycle (the corrin ring) and play essential roles in metabolic pathways across diverse biological systems. The axial ligands bound to the central Co ion significantly influence the electronic structure, redox behavior, and biological activity of these complexes. While the electronic structures of the biologically active alkylcobalamins such as methylcobalamin and adenosylcobalamin have been extensively studied with both experimental and computational methods,[1,2] halocobalamins (XCbls) have comparatively been less explored,[3,4] particularly in terms of electronic structure analyses using modern computational methodology.

XCbls are readily synthetically accessible via substitution of the labile upper, or β, axial ligand of aquacobalamin in saturated aqueous halide solutions. These XCbls possess a significantly weaker σ donating upper axial ligand compared to alkylcobalamins; yet, they exhibit electronic absorption (Abs) spectra in which the dominant feature in the visible spectral region (the so-called α/β bands) is considerably red-shifted from its counterpart in the alkylcobalamin Abs spectra. This red-shift is at odds with the theoretical framework established for strongly σ donating axial ligands, according to which the α/β bands should shift to lower energy with increasing σ donor strength of the upper axial ligand.[1] Instead, for the XCbl series, the red-shift of the α/β bands appears to correlate with the halide size and not the donor strength. A prior study by Chemaly established general trends in Abs peak positions across various cobalamin derivatives (including XCbls),[4] but a quantitative interpretation of these data in terms of differences

in Cbl electronic structures was not attempted. Because axial ligand identity also impacts redox properties, ligand substitution kinetics, and cofactor reactivity, the XCbl series offers a powerful platform for refining theoretical models and probing metalloenzyme mechanisms, especially in the context of B_{12}-dependent reductive dehalogenases.[5]

In this study, we present a detailed investigation of the electronic structures of the XCbl series using electronic absorption (Abs), circular dichroism (CD), and resonance Raman (rR) spectroscopies, complemented by spectroscopically validated density functional theory (DFT) and time-dependent DFT (TD-DFT) computations. Using CD spectroscopy, we have identified the two lowest-energy, formally symmetry-forbidden Co $d \rightarrow d$ transitions and found that the energies of these transitions are highly sensitive to the identity of the bound halide. The results obtained from DFT and TD-DFT computations indicate that the length of the Co–X bond in the XCbl species is the determining factor for the spectral shifts observed in the low energy region of the Abs and CD spectra.

3.2 Methods

Synthesis

Hydroxocobalamin acetate (Tokyo Chemical Industry, >90%), sodium fluoride (Fischer Scientific, ACS grade), potassium chloride (MP Biomedicals, ACS grade), potassium bromide and potassium iodide (Fischer Scientific, ACS grade) were used as obtained. The XCbl species (X = Cl, Br, I) were synthesized by dissolving HOCbl in a freshly prepared, saturated potassium halide (XCl) solution to achieve a final XCbl concentration of 150 μM. These solutions were used for all Abs, CD, and rR spectroscopic analyses. To prepare FCbl, a saturated solution of NaF instead of KF was used, because

HOCbl does not readily dissolve in a saturated KF solution. Importantly, the Abs spectra of the FCbl species formed from dilute KF addition and the Abs spectra of FCbl formed with saturated NaF solution were identical.

Spectroscopy

Abs spectra were obtained with a Cary 5E UV/vis spectrophotometer and CD spectra were collected using a Jasco J-715 spectropolarimeter. rR spectra were obtained with excitation at 514.5 nm by a Coherent I-305 Ar^+ ion laser with ~20 mW of laser power at the sample. The scattered light was collected using a ~135° back-scattering 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 in an NMR tube placed in a quartz finger dewar filled with N_2(l). The ice peak at 228 cm^{-1} was used as an internal standard to calibrate Raman shifts.

Computational Models

All XCl computational models were derived from the high-resolution crystal structure of AdoCbl (CCDC: PAFBUV).[6] The Ado moiety was replaced by the corresponding halide and all side chains, including the nucleotide loop, were truncated and modeled as methyl groups. The dimethylbenzimidazole axial ligand was replaced with a coordinating imidazole.

Geometry optimizations were conducted using the ORCA 5.0 software package.[7,8] Calculations were performed with the PBE functional[9], the def2-TZVP basis set[10] for Co and the six coordinating atoms, and the def2-SVP basis set[10] for all other atoms. The

spin-restricted (S = 0) Kohn-Sham formalism was used for all geometry optimizations. The def2/J auxiliary basis set[11] for Coulomb integral approximation was additionally used to speed up the calculations. Atom-pairwise dispersion correction with the Beck-Johnson damping scheme (D3BJ) empirical dispersion corrections was also applied to all atoms.[12,13]

TD-DFT Computed Abs and CD Spectra

An Abs spectrum for each geometry optimized model was calculated using time-dependent density functional theory (TD-DFT) within the Tamm–Dancoff approximation as implemented in the ORCA 5.0 software package. The CAM-B3LYP functional[14] was used for all TD-DFT calculations with the same basis sets as described above. The def2-SVP/C auxiliary basis set for electron correlation was also used.[15] For all XCbl species the 60 lowest energy excited states were computed by including all single excitations between molecular orbitals (MOs) within an energy window of ±3 Hartrees with respect to the highest occupied MO (HOMO) and the lowest unoccupied MO (LUMO). The solvation model based on density (SMD)[16], implemented through the CPCM module in ORCA, was used in all TD-DFT calculations. All computed Abs and CD spectra reported here were obtained using the SMD model for water, unless otherwise noted. The TD-DFT-calculated transition energies and oscillator strengths were used to simulate Abs and CD spectra by assuming that each transition gives rise to a Gaussian band with full width at half-maximum of 2000 cm^{-1}. All TD-DFT computed spectra were uniformly red-shifted by 4500 cm^{-1} to achieve closer agreement with the experimental spectra.

The NBO7 program[17] was used to further explore the major bonding interactions for all geometry optimized species. The same parameters used for the TD-DFT

calculations were employed here. Isosurface plots of key MOs, natural bonding orbitals (NBOs), natural localized MOs (NLMOs), and electron density difference maps (EDDMs) were generated with PyMOL v3 using isodensity values of 0.03 au (for all orbital plots) and 0.003 au (for EDDMs).

Computed Raman Spectra

Vibrational frequencies and Raman scattering intensities were computed using the NumFreq keyword in ORCA 5.0 and the SMD solvent model for water. A peak width of 10 cm^{-1} was used to simulate computed Raman spectra. A scaling factor of 0.9947 was used to align the long-axis stretch of FCbl with the corresponding experimental feature at 1499 cm^{-1} and the same scaling factor was applied to all computed Raman spectra.

3.3 Results and Discussion

Experimental Electronic Absorption (Abs) and Circular Dichroism (CD) Spectra

The room temperature (RT) Abs spectra of the XCbl (X = F, Cl, Br, and I) series are shown in Figure 3.1 (top). All XCbl spectra are dominated by two features, the so-called α/β bands ($\varepsilon \approx 7000\ M^{-1}cm^{-1}$) in the visible spectral region and the γ band ($\varepsilon \approx 15000\ M^{-1}cm^{-1}$) in the near-UV. As reported previously, the α/β bands steadily red-shift from FCbl to ClCbl, BrCbl, and ICbl. These spectra are very similar in general appearance to those displayed by other Co^{3+}Cbl species, such as cyanocobalamin (CNCbl). For CNCbl, a detailed spectroscopic and computational analysis provided compelling evidence that the α/β bands arise from the same corrin $\pi \rightarrow \pi^*$ transition, with the more intense α band being the electronic origin and the β band the first member of a progression in a corrin-

based stretching mode.[18] The γ band is due to another corrin $\pi \rightarrow \pi^*$ transition that is polarized perpendicular to that responsible for the α/β bands.

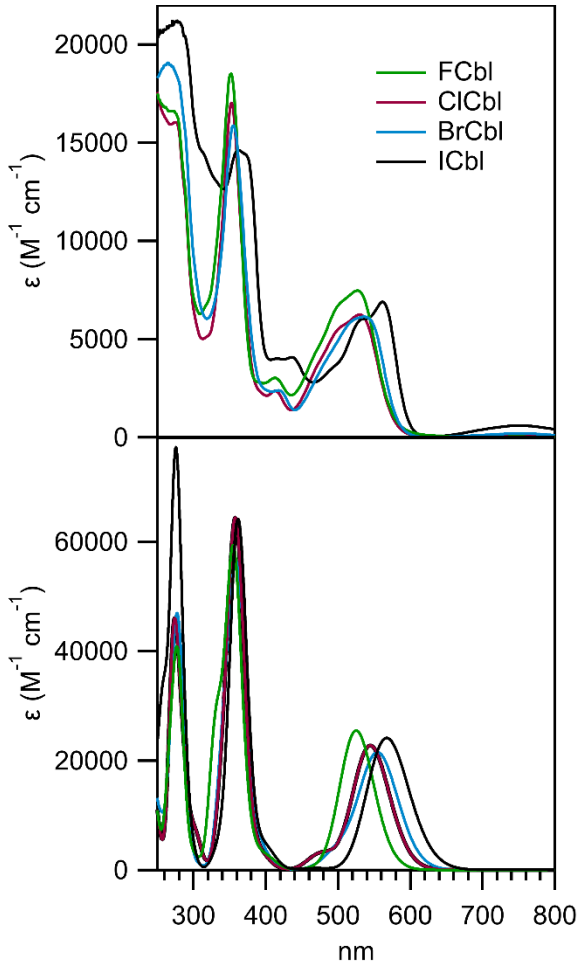

Figure 3.1. Top: Experimental RT Abs spectra of XCbl species (X = F, Cl, Br, or I). Bottom: TD-DFT computed Abs spectra. The TD-DFT computed spectra were uniformly red-shifted by 4500 cm^{-1} to facilitate a comparison with the experimental Abs spectra.

The RT CD spectra of the XCbl species are shown in Figure 3.2. While some CD peak positions and intensities were previously reported for BrCbl and ICbl,[3] the full CD spectra were not shown. In all XCbl CD spectra, a negatively signed feature can be identified that corresponds to the α/β bands in the Abs spectra. Interestingly, at least one positively signed CD feature can be identified at lower energy, the origin of which will be explored below.

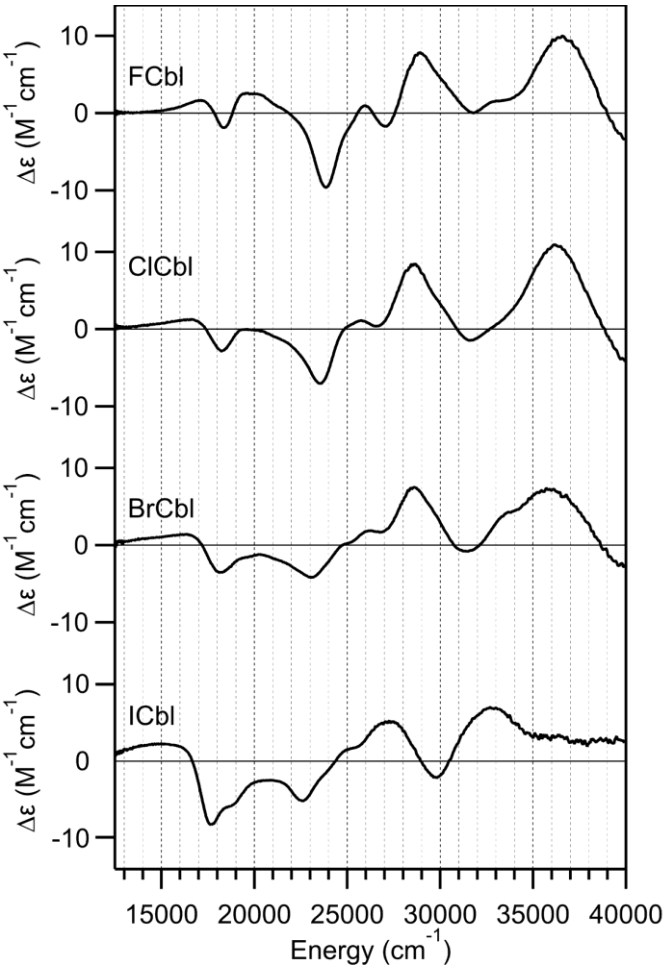

Figure 3.2. Room-temperature CD spectra of XCbl species (X = F, Cl, Br, or I).

Computational Models

To compare the geometric and electronic structures of the different XCbl species, truncated computational models were derived from the AdoCbl crystal structure and subjected to DFT geometry optimizations. Because at present ClCbl is the only XCbl species for which an X-ray crystal structure is available (CCDC: GIZYOD),[19] this structure was used to validate the geometries of the DFT optimized XCbl models.

As shown in Table 3.1, the key structural parameters of ClCbl determined experimentally and predicted computationally are in excellent agreement. Specifically, the

Co–N_{ax} and Co–Cl distances merely differ by 0.01 Å, and the average Co–N_{eq} bond distance deviates by only 0.02 Å. Additionally, the DFT predicted Cl–Co–N_{ax} bond angle is within 0.2° of the experimental value, indicating that despite of the truncation scheme applied to the computational models, these models still adequately account for the steric strain imposed by the upper and lower axial ligands. Given the excellent agreement of the key metric parameters between the experimentally determined ClCbl structure and DFT-optimized ClCbl model, an analysis of the geometric parameters of the other DFT-optimized XCbl species is warranted.

Table 3.1. Key geometric parameters for geometry optimized XCbl models and the ClCbl crystal structure

Species	**Co–X (Å)**	**Co–N_{ax} (Å)**	**X–Co–N_{ax} (°)**	**Fold Angle (°)**	**Co–N_{eq} (Å, avg.)**
FCbl (DFT)	1.85	1.96	177.8	4.3	1.92
ClCbl (DFT)	2.25	1.98	176.6	6.8	1.92
ClCbl exp.	*2.26*	*1.99*	*176.4*	*17.7*	*1.90*
BrCbl (DFT)	2.41	1.98	176.2	7.5	1.92
ICbl (DFT)	2.61	2.00	175.5	8.5	1.91

The geometry optimized XCbl models all have a similar Co–N_{ax} bond length that ranges from 1.96 Å in FCbl to 2.00 Å in ICbl. Also, the average Co–N_{eq} bond distance differs minimally (0.01 Å) among the XCbl models, indicating little change in the equatorial Co–N_{eq} bonding environment. The X–Co–N_{ax} angle shows a minor, but steady, decrease from 177.8° in FCbl to 175.5° in ICbl. A slight increase in fold angle is additionally calculated from FCbl to ICbl. This fold angle is underestimated compared to the ClCbl crystal structure, probably because of the use of a truncated ClCbl model, the exclusion of explicit solvent during gas phase geometry optimization, and the fact that crystal packing forces may distort the experimental ClCbl structure.[20]

While several geometric parameters do not change significantly between XCbls, the upper axial Co–X bond shows a drastic lengthening along the XCbl series from 1.85 Å in FCbl to 2.61 Å in ICbl. These models show that the corrin geometry, as well as the Co–N_{ax}, and Co–N_{eq} bond distances are highly similar throughout the XCbl series. The only significant structural difference among the XCbls is the length of the Co–X bond. We therefore turn to our electronic structure analysis to evaluate the impact of the lengthened Co–X bond.

Spectroscopic Validation of Computational Models

To validate the geometry optimized XCbl models for use in electronic structure analysis, we compared the TD-DDT computed Abs spectrum for each DFT-optimized XCbl model to the corresponding experimental Abs spectrum. After a uniform 4500 cm^{-1} red-shift was applied to the computed Abs spectra, excellent agreement was achieved between the computed and the experimental Abs spectra for all XCbl species (Figure 3.1). The major spectral features, including an intense band in the visible region and a more intense near-UV feature corresponding to the α/β bands and the γ peak, respectively, are present in all computed spectra with extinction coefficients that agree well with the experimental values (Figure 3.1). Also, the experimentally observed red-shift of the α/β bands from FCbl to ICbl is properly reproduced in the computed spectra.

MO Composition Analysis

Computed MO diagrams for all XCbl models are shown in Fig S3.1. To facilitate a comparison of the frontier MO energies among these species, the LUMO energy for each XCbl was set to 0 eV because this MO is almost entirely corrin π^*-based and thus nearly identical in composition across the entire XCbl series. These MO diagrams show that the splitting of the Co 3d-based MOs is similar for all XCbls. In each case, the filled Co $3d_{xz}$, $3d_{yz}$, and $3d_{x2-y2}$ -based MOs are all found ~8 to 9 eV lower than the LUMO. The formally unoccupied $3d_{xy}$-based MO is located between +1.5 and +1.6 eV above the LUMO and has ~60% Co $3d_{xy}$ character for all XCbls. Unsurprisingly, given the drastic increase in the axial Co–X bond length along the XCbl series, a large stabilization of the Co $3d_{z2}$-based MO occurs from FCbl to ICbl. As a result, this MO is the LUMO+2 in the FCbl model but becomes the LUMO+1 in the ClCbl, BrCbl, and ICbl models. The percentage Co $3d_{z2}$ orbital contribution to the LUMO+1 (LUMO+2 for FCbl) is constant at ~60% for all species, despite the variable Co–X bond length.

A minor change in the HOMO/LUMO energy gap is predicted across the XCbl series. Specifically, this energy gap decreases from 5.51 eV for FCbl to 5.42 eV for ClCbl, 5.38 eV for BrCbl, and 5.33 eV for ICbl. The small relative destabilization of the HOMO from FCbl to ICbl correlates with an increase in the total percentage of axial ligand contributions to the HOMO along this series from a total of 2% to 5%, 7%, and 9% (Figure 3.3).

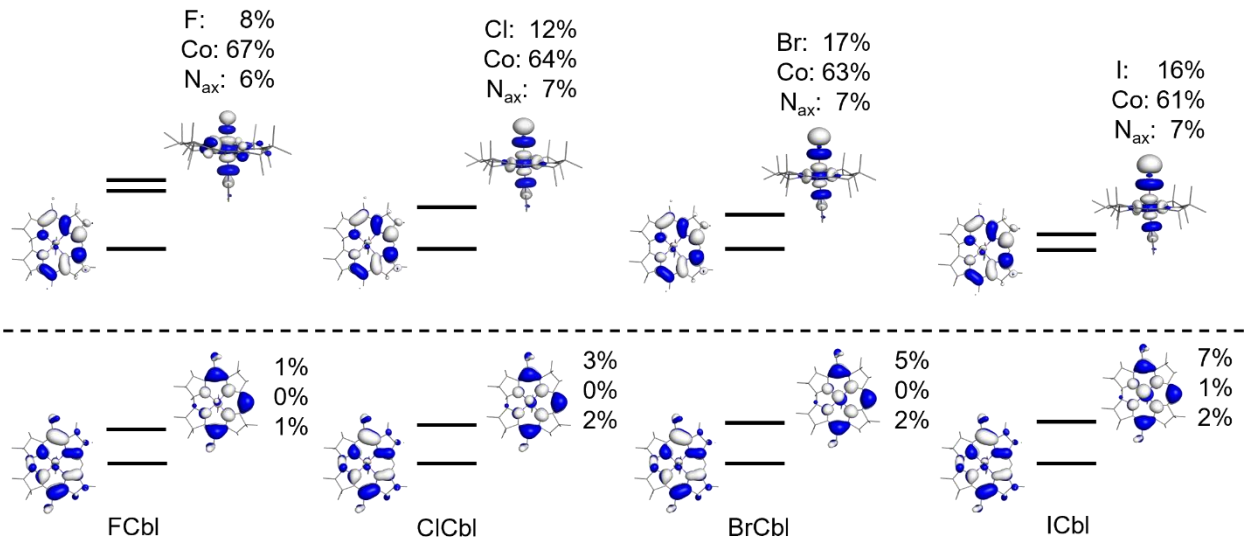

Figure 3.3. Key frontier MOs of XCbl species and percent orbital contributions from the Co atom and axial ligands for selected MOs. The dashed line separates the lower energy occupied MOs from the higher energy unoccupied MOs. Note that the HOMO/LUMO gap is not drawn to scale.

In contrast to the small variation in the HOMO/LUMO gap, there is a drastic change in the LUMO/Co $3d_{z2}$-based MO energy gap and the composition of the Co $3d_{z2}$-based MO along the XCbl series (Figure 3.3). Specifically, the computed energy difference between the LUMO and Co $3d_{z2}$-based MO is 1.30 eV for FCbl, 0.82 eV for ClCbl, 0.65 eV for BrCbl, and 0.30 eV for ICbl. The increase in Co–X bond length from FCbl to ICbl almost entirely accounts for the relative energy change of the Co d_{z2}-based MO. In fact, a nearly linear relationship exists between the LUMO/ Co $3d_{z2}$-based MO energy gap and the Co–X bond distance (Figure 3.4). Therefore, electronic transitions that have significant contributions from excitations to the Co $3d_{z2}$-based MO are expected to decrease in energy along the XCbl series.

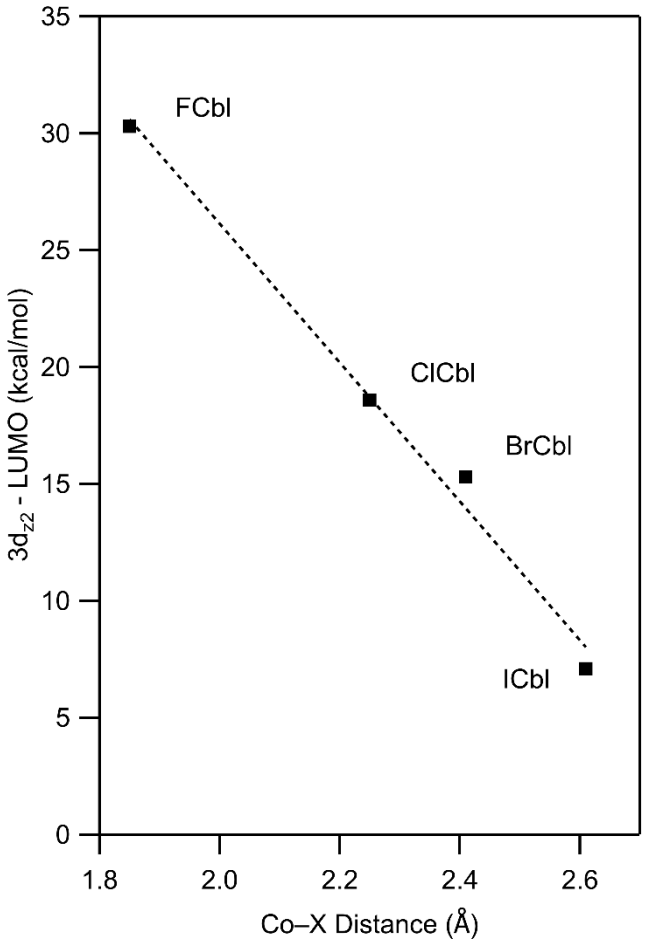

Figure 3.4. Energy gap between the unoccupied $3d_{z2}$-based MO and the LUMO for each XCbl species as a function of the Co–X bond distance.

NBO Analysis

The NBO7 program implemented in ORCA 5.0 was used to create easy to visualize, localized bonding depictions for the different XCbl species that complement the MO isosurface plots. These NBO analyses show that the compositions of the Co 3d-based orbitals are highly conserved among the different members of the XCbl series (Figure S3.2 to Figure S3.5). In each case, the three non-bonding, Co 3d-based ($3d_{xz}$, $3d_{yz}$, $3d_{x2-y2}$) NBOs and natural localized MOs (NLMOs) all have occupancies exceeding 1.960. Therefore, each XCbl is best described as a low-spin d^6, Co^{3+}Cbl species.

The NBO description of the Co–X bond shows an increase in the Co 3d orbital character from 21% in ClCbl to 24% in BrCbl and 26% in ICbl. For FCbl, the NBO analysis indicates that the dominant resonance structure constitutes an F^- ligand with four lone pairs and a Co–N_{imid} bond with 20% Co 3d character. Overall, the primary axial bonding interaction is polarized toward one of the axial ligands for all XCbls; namely, the halide in ClCbl, BrCbl, and ICbl and the coordinating imidazole in FCbl.

While the NBO/NLMO analyses are useful for producing localized bonding descriptions that are easy to visualize, they account for only one predominant resonance structure. A more complete description of the bond orders and formal number of lone pairs can be obtained by performing a natural resonance theory (NRT) analysis that potentially examines thousands of contributing resonance structures.

From FCbl to ICbl, the NRT analysis indicates that the number of lone pair electrons in Co 3d orbitals increases very slightly from 3.05 to 3.09 electrons (Table 3.2). More drastic changes are observed for the computed covalency of the Co–X bond, which is predicted to increase sharply from 50% in FCbl to ~70% in ClCbl, BrCbl, and ICbl. As a result, the computed Co–X bond order is substantially smaller in FCbl than in the other XCbl species. Nevertheless, the NRT analysis agrees with the NBO results that in all XCbls the Co ion has ~6 non-bonding d electrons (6.10 in FCbl to 6.18 in ICbl).

Table 3.2. Results obtained from natural population analyses and natural resonance theory analyses for XCbls

	Natural Charge			Natural Bond Order		Fold Angle
Species	Co	X	N_{His}	Co–X	Co–N_{His}	
FCbl	1.07	-0.71	-0.54	0.51	0.49	4.31
ClCbl	0.94	-0.57	-0.54	0.70	0.24	6.81
BrCbl	0.92	-0.52	-0.55	0.71	0.22	7.45
ICbl	0.91	-0.49	-0.56	0.72	0.22	8.47

Abs Spectral Assignments

As mentioned above, the TD-DFT computed Abs spectra for all XCbl species exhibit an intense band in the visible region that corresponds to the α/β bands in the experimental Abs spectra. Because of the red-shift of the α/β bands from FCbl to ICbl, the computed transition responsible for these bands is number three for FCbl, ClCbl, and BrCbl, but number four for ICbl. As a result, the TD-DFT computed Abs spectrum for ICbl contains an additional, less intense feature at slightly lower energy that originates from a transition primarily involving HOMO $\rightarrow$ LUMO+1 one-electron excitation (Figure 3.5).

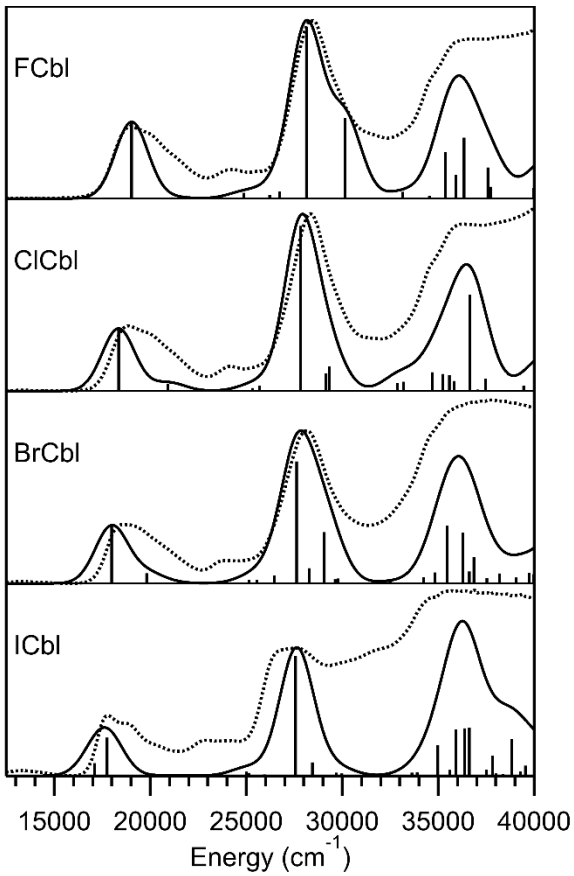

Figure 3.5. TD-DFT computed Abs spectra (solid traces, solid sticks) plotted against the experimental Abs spectra (dotted traces) of the XCbl series. The intensities of the experimental and computed spectra were normalized to the γ region peak.

The position of the α/β bands in the experimental Abs spectra of Co^{3+}Cbl species has been shown to be highly sensitive to the identity of the upper axial ligand, decreasing in energy with increasing σ donor strength of this ligand.[1] This trend has been attributed to, primarily, an increase in the Co–N_{ax} antibonding interaction in the HOMO with increasing σ donor strength of the upper axial ligand. However, in the case of the XCbl series, the σ donor strength of the axial ligand cannot be the primary reason for the observed red-shift of the α/β bands from FCbl to ICbl, as these bands are observed at a

lower energy for ICbl than for MeCbl or CNCbl,[1] despite I^- being a much weaker σ donor than CH_3^- and CN^-.

The red-shift of the α/β bands across the XCbl series is instead attributed to the small decrease in the HOMO/LUMO gap and, for ICbl, the relative stabilization of the Co d_{z2}-based MO. The transition responsible for the α/β bands is primarily HOMO $\rightarrow$ LUMO in character, with this one-electron excitation contributing 90% in the case of FCbl, 90% for ClCbl, 91% for BrCbl, and 83% for ICbl. Interestingly, for ICbl only, this transition is predicted to have an additional 9% contribution from the HOMO $\rightarrow$ LUMO+1 excitation. The drastically red-shifted α/β bands observed for ICbl can therefore be attributed to i) the fact that this species has the smallest HOMO/LUMO energy gap of any XCbl and ii) a sizable contribution from the HOMO $\rightarrow$ LUMO+1 transition that is not present for the other XCbls.

Upon closer inspection of the low-energy region of the TD-DFT computed Abs spectra for all XCbl species, two very weak features can clearly be seen on the low-energy side of the intense band corresponding to the α/β bands in the experimental Abs spectra (Figure 3.6). Like the α/β bands, these two features are predicted to red-shift from FCbl to ICbl.

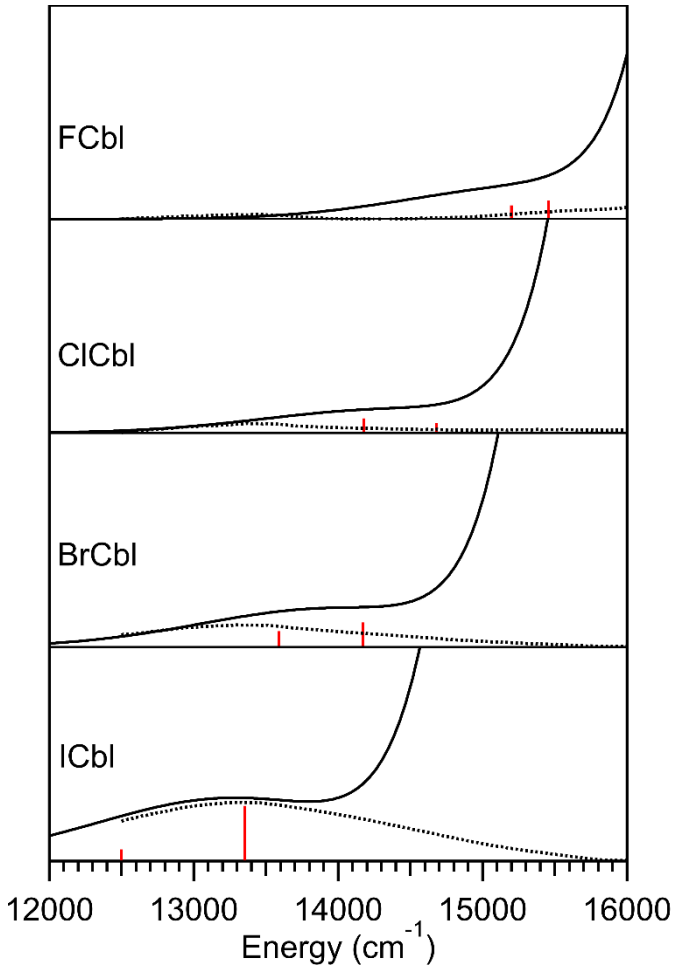

Figure 3.6. Low-energy region of the TD-DFT computed Abs spectra (solid traces, solid red sticks) plotted against the experimental Abs spectra (dotted traces) of the XCbl series. The computed spectra were scaled by a factor of x100 from the scaling used in Figure 3.5.

Previous TD-DFT computations for H_2OCbl indicated that in this species, for which a weak Abs feature is also observed on the low-energy side of the α/β bands, the lowest energy transition primarily corresponds to a corrin $\pi \rightarrow$ Co $3d_{z2}$ (LMCT) excitation.[1] By analogy, the lowest-energy, weak Abs features of the XCbl series were also tentatively assigned to corrin $\pi \rightarrow$ Co $3d_{z2}$ transitions.[4] TD-DFT computed electron density difference maps (EDDMs) for the two low energy transitions of the XCbl species (Figure 3.7),

however, clearly show that these transitions are not corrin π-based, but instead possess significant Co d → d character. Specifically, the principal contributors to the lowest and second lowest energy transitions are the Co $3d_{yz} \rightarrow 3d_{z2}$ and Co $3d_{xz} \rightarrow 3d_{z2}$ excitations, respectively. The computed red shifts of these transitions from FCbl to ICbl can be attributed to the stabilization of the Co $3d_{z2}$-based acceptor MO along this series.

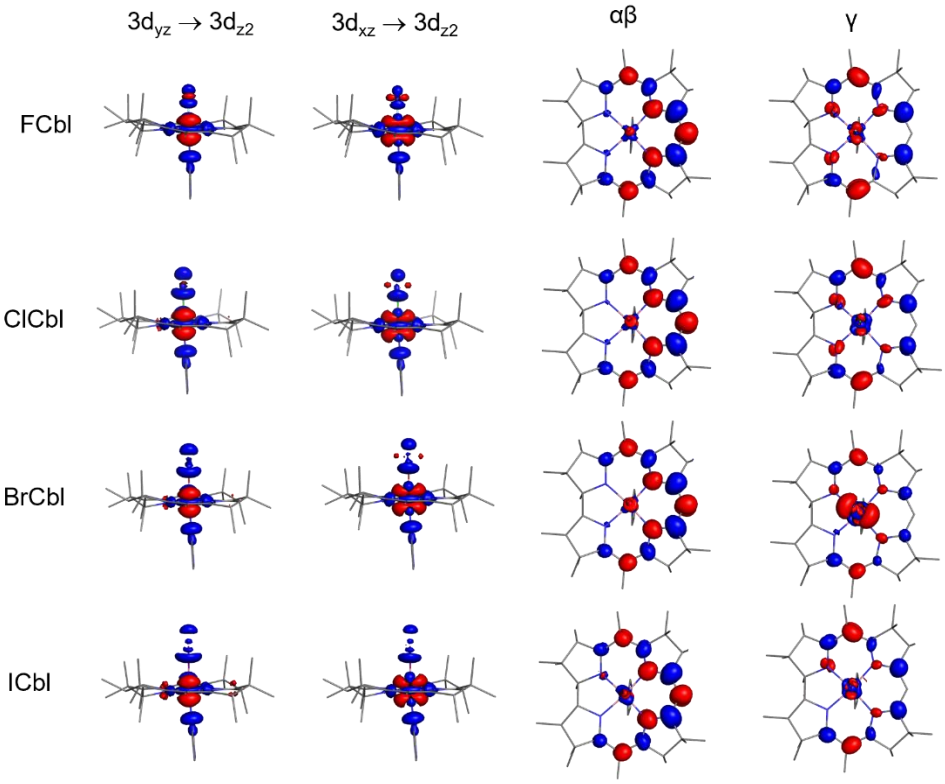

Figure 3.7. EDDMs for selected d → d transitions and the dominant transitions contributing to the αβ region and the γ region of XCbls. Red and blue show what???

CD Spectral Analysis

In addition to reproducing the experimental Abs data previously reported for all XCbls, we have also collected the first CD spectra of the entire XCbl series between 800

and 250 nm (Figure 3.2). Some CD peak positions and intensities were previously reported for BrCbl and ICbl,[3] but the full CD spectra were not shown. To interpret these spectra, TD-DFT was also used to compute CD spectra for the XCbl series, which are shown in Figure 3.7. While the computed CD spectra do not match the experimental data quite as well in general appearance as do the computed Abs data, this discrepancy can be attributed to the fact that the CD spectra of Cbls are highly sensitive to temperature and corrin ring distortions on a scale of a few kcal/mol, and our truncated models and gas phase optimizations may not capture subtle conformational effects that are present in solution at room temperature.[3] However, the computed CD spectra still provide useful information on the nature of the low energy electronic transitions.

In the low energy region of all computed XCbl CD spectra (Figure 3.7, right), the two lowest energy Co d $\rightarrow$ d transitions are positively signed. Consistent with this prediction, all experimental XCbl CD spectra show a positively signed feature in the low energy region. Therefore, we attribute the lowest energy, positively signed feature in the experimental CD spectra to at least one Co d $\rightarrow$ d transition. Next, higher in energy, a negatively signed feature corresponding to the α/β bands in the Abs spectra is observed in both the computed and experimental CD spectra. While the energies of this negatively signed feature in the experimental CD spectra do not exactly match the energies of the α/β bands in the corresponding Abs spectra, this apparent discrepancy can be attributed to the overlap of multiple positively and negatively signed CD transitions in this region.

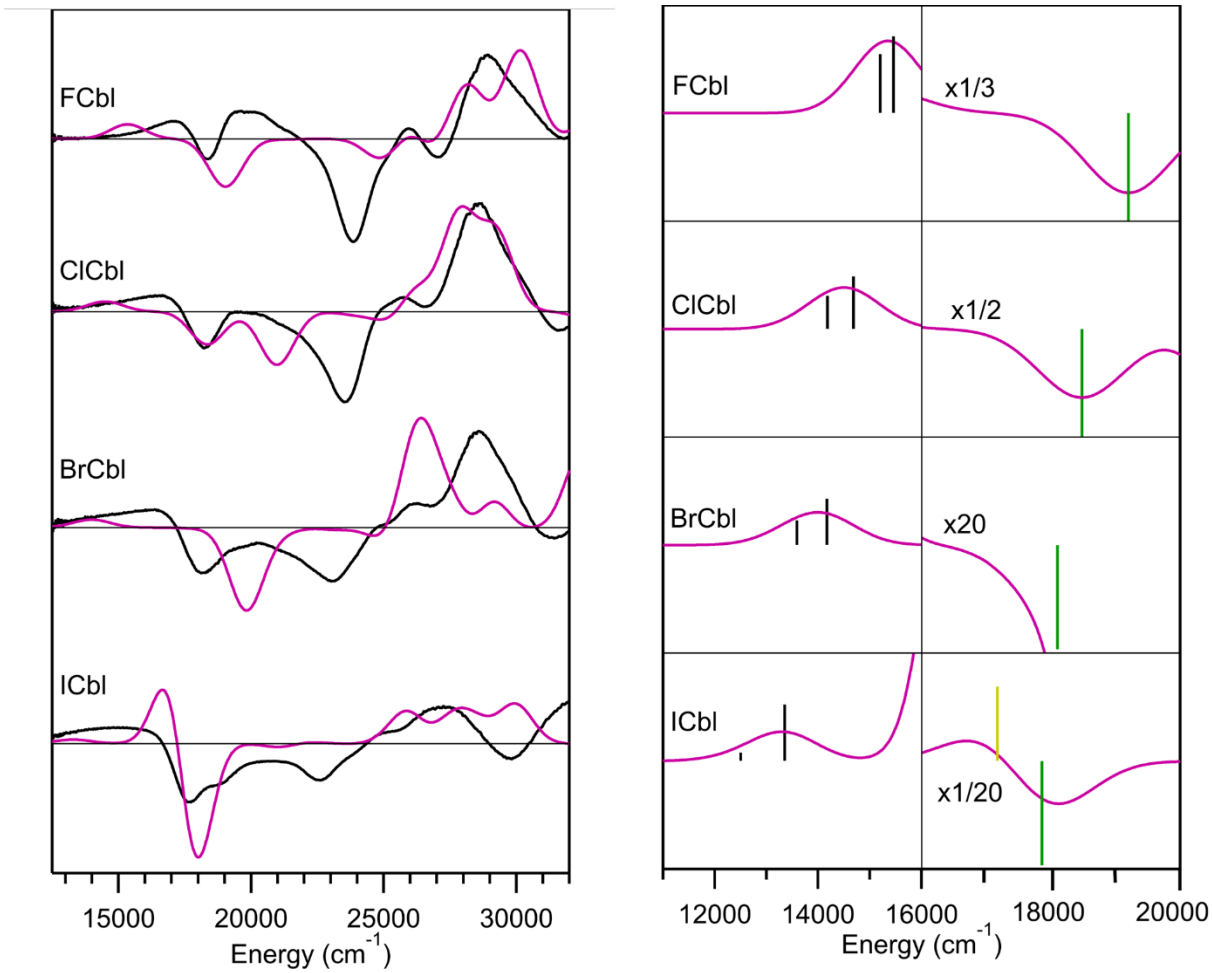

Figure 3.8. Left: Experimental RT CD spectra (black traces) and TD-DFT computed CD spectra (purple traces) of XCbl. Right: Low energy region of the computed CD spectra (purple traces) with the individual transitions shown as black sticks (Co d $\rightarrow$ d transitions), green sticks (transition responsible for $\alpha\beta$ peaks), and a yellow stick (HOMO $\rightarrow$ $3d_{z2}$ transition of ICbl, above 20000 cm^{-1} for other XCbl species).

In the experimental FCbl CD spectrum, an additional positively signed feature spanning from ~19400 to 20100 cm^{-1} (515 to 498 nm) is observed directly toward higher energy of the negatively signed feature associated with the α/β bands in the corresponding Abs spectrum. This feature is absent from the experimental CD spectra of all other XCbls. Based on our TD-DFT computations for FCbl, this positively signed feature can be assigned to the Co $3d_{xz} \rightarrow 3d_{z2}$ transition. Because the Co $3d_{z2}$-based MO becomes relatively stabilized from FCbl to ICbl, the corresponding transition gradually

red-shifts along this series and eventually gives rise to a well-resolved, positively signed low-energy feature in both the experimental and computed CD spectra of ICbl.

The features in the higher energy region of the experimental CD spectra show a minor red shift down the XCbl series but otherwise the spectra are very similar in general appearance. The transition responsible for these features are primarily corrin $\pi \rightarrow \pi^*$ in character, and so their energies are likely not drastically affected by the axial Co–X bond.

Resonance Raman Data

For all XCbl species, rR spectra were obtained with λ_{ex} = 514.5 nm, roughly corresponding to the energy of the transition responsible for the α/β bands in the corresponding Abs spectra.[1] The corrin short-axis polarized stretch, found at ~1540 cm^{-1} for Co^{3+}Cbls, is sensitive to the Co oxidation state and shifts to lower energy with decreasing Co partial charge.[2] This peak is located at 1544, 1542, and 1543 cm^{-1} in the rR spectra of FCbl, ClCbl, and BrCbl, respectively (Figures S3.6 to S3.9). In the ICbl rR spectrum, this peak could not be discerned, likely because the transition responsible for the α/β bands of this species is much lower in energy than the laser excitation wavelength used to collect rR data. However, based on the Abs and CD spectra, as well as the computational data obtained in this study, it is obvious that ICbl can also be described as a typical Co^{3+}Cbl species. Consistent with this conclusion, the computed Raman spectra for the XCbl series are nearly identical to one another (Figure S3.10).

4.4 Conclusions

This study provides a comprehensive spectroscopic and computational analysis of the halocobalamin (XCbl, X = F, Cl, Br, I) series, offering new insight into the relationship between axial ligand identity and electronic structure in Co(III)-corrinoid complexes.

Experimental Abs, CD, and rR spectra disclose systematic trends regarding the α/β band positions and d $\rightarrow$ d transitions energies across the series, with notable red-shifts correlating with increasing halide size rather than σ donor strength. Through spectroscopically validated DFT and TD-DFT calculations, we established that the red-shift of the α/β bands is predominantly driven by changes in the Co–X bond length, which modulates the energy of the Co $3d_{z2}$-based molecular orbital. Importantly, these computations reproduce the experimental spectra with high fidelity and reveal the key orbital contributions to the observed electronic transitions. NBO and NRT analyses provide further insight into the electronic structure differences across the series of XCbl species, confirming the expected low-spin Co d^6 configuration of the metal center. Collectively, these experimental and computational findings highlight the unique electronic landscape of halocobalamins and refine our understanding of axial ligand effects in cobalamin chemistry.

References

(1) 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.

(2) Park, K.; C. Brunold, T. 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.

(3) 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* **1967**, *6* (7), 2178–2189. https://doi.org/10.1021/bi00859a040.

(4) Chemaly, S. M. Cobalamins and the Spectrochemical Series. *Dalt. Trans.* **2008**, No. 42, 5766–5773. https://doi.org/10.1039/b806927a.

(5) Jugder, B. E.; Ertan, H.; Lee, M.; Manefield, M.; Marquis, C. P. Reductive Dehalogenases Come of Age in Biological Destruction of Organohalides. *Trends Biotechnol.* **2015**, *33* (10), 595–610. https://doi.org/10.1016/j.tibtech.2015.07.004.

(6) Mebs, S.; Henn, J.; Dittrich, B.; Paulmann, C.; Luger, P. Electron Densities of Three B12 Vitamins. *J. Phys. Chem. A* **2009**, *113* (29), 8366–8378. https://doi.org/10.1021/jp902433x.

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

(8) Neese, F. Software Update: The ORCA Program System—Version 5.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2022**, *12* (5), 1–15. https://doi.org/10.1002/wcms.1606.

(9) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. https://doi.org/10.1021/acsami.8b15784.

(10) 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.

(11) 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.

(12) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. https://doi.org/10.1002/jcc.

(13) 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). https://doi.org/10.1063/1.3382344.

(14) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1–3), 51–57. https://doi.org/10.1016/j.cplett.2004.06.011.

(15) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor. Chem. Acc.* **2007**, *117* (4), 587–597. https://doi.org/10.1007/s00214-007-0250-5.

(16) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396. https://doi.org/10.1021/jp810292n.

(17) NBO 7.0. E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou, C. R. Landis, and F. Weinhold, Theoretical Chemistry Institute, University of Wisconsin, Madison, WI (2018).

(18) Elmendorf, L. D.; Brunold, T. C. Vibronic Coupling in Vitamin B_{12}: A Combined Spectroscopic and Computational Study. *Inorg. Chem.* **2023**, *62* (32), 12762–12772. https://doi.org/10.1021/acs.inorgchem.3c01305.

(19) Randaccio, L.; Furlan, M.; Geremia, S.; Slouf, M. Preparation and X-Ray Analysis of Crystals of Azido- and Chlorocobalamin Containing LiCl: A Structural Model for the Interactions of the Corrin Ring with Ionic Species. *Inorg. Chem.* **1998**, *37* (20), 5390–5393. https://doi.org/10.1021/ic980181u.

(20) Rovira, C.; M. Kozlowski, P. 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 4

Spectroscopic and Computational Analysis of Hydridocobalamin (HCbl)

4.1. Introduction

Cobalamins (Cbls) are a class of cofactors composed of a tetrapyrrole macrocycle, termed the corrin ring, featuring a central cobalt atom with four nitrogen atoms in a pseudo-equatorial plane. Additionally, a 5',6'-dimethylbenzimidazole (DMB) moiety is tethered to the corrin ring through a nucleotide loop that binds to the Co atom on the lower, or α, face. The upper axial ligand varies across the Cbl family and differentiates their functions. Adenosylcobalamin and methylcobalamin, containing a 4'5'-deoxyadenosyl moiety and a methyl group, respectively, are the biologically active forms in human metabolism. Additional forms, such as aquacobalamin (H_2OCbl^+) and hydroxocobalamin (HOCbl), are readily converted to other Cbls. The Co atom in the corrin ring is known to access three oxidation states, Co^{3+}, Co^{2+}, and Co^{1+}, depending on reaction conditions.

Hydridocobalamin (HCbl) represents a proposed, but elusive, intermediate in the reactivity of cobalamin cofactors, with particular relevance to hydrogen evolution chemistry and the mechanistic understanding of B_{12}-dependent catalysis and of B_{12} as a photosensor in the transcription factor CarH.[1] In nature, cobalamins mediate a wide variety of chemical transformations, but the direct formation of metal-hydride species in the biological context remains speculative. Model systems such as cobaloximes have established that cobalt-hydride intermediates can facilitate proton reduction to H_2, providing functional parallels to hydrogenase enzymes.[2] Extending these insights to corrinoids is important for assessing whether B_{12} scaffolds can participate in similar processes and for evaluating their potential as bioinspired hydrogen-evolution catalysts. The historical assignment of HCbl as a hydrido–Co(III) species, based largely on limited

electronic absorption (Abs) data, left unresolved questions regarding its actual electronic structure, stability, and relevance to catalysis. Understanding whether HCbl is truly formed has broader implications for the design of corrinoid analogues with tailored reactivity, for example in synthetic biology applications where B_{12} cofactors are repurposed for abiological transformations.

Chemaly and Pratt investigated the formation of $H_{2(g)}$ by Cbl under reducing conditions in acidic solution.[3] They reported the Abs spectrum of “HCbl”, formed by Cbl reduction with Zn dust in glacial HOAc. $H_{2(g)}$ gas was readily formed during the reaction. The Abs spectrum of “HCbl” showed close resemblance to the Abs spectrum of base-off ethylcobalamin (EtCbl). However, they stated that the peaks at 470 nm and 384 nm were variable in intensity throughout the Abs measurements, a result they attributed to the presence of variable amounts of Co^{2+}Cbl (which exhibits a strong absorption feature near 470 nm) and Co^{1+}Cbl (large absorption near 384 nm). They still, however, concluded their Abs spectrum corresponded to HCbl and not just a mixture of other Cbls. However, the lack of other spectroscopic data from their study made their conclusion tenuous, and their classification of HCbl resembling a base-off alkylCbl serendipitous. They attributed the Abs spectral changes observed when H_2OCbl^+ was reduced in acidic solution to formation of a protonated Co^{1+}Cbl species. Whether this species was protonated at Co or somewhere on the corrin ring was not determined. To answer this question, Kozlowski and coworkers performed a computational analysis of HCbl and reported that the electronic structure of HCbl was best described as a protonated Co^{1+}Cbl species.[4] This conclusion was based on HCbl models that were spectroscopically validated using the inconclusive Abs data from Chemaly and Pratt.

We replicated the experimental conditions for producing “protonated Co^{1+}Cbl” here and sought to determine i) whether the originally reported Abs spectrum did, in fact, correspond to an HCbl species and ii) whether the electronic structure of this species is best described as a protonated Co^{1+}Cbl or a hydride-bound Co^{3+}Cbl. We further explored the mechanism for catalytic $H_{2(g)}$ production of Cbl by performing intrinsic reaction coordinate (IRC) computations to determine the energies associated with the formation of base-off HCbl, conversion to base-off H_2Cbl, and the release of H_2 from H_2Cbl. Beyond mechanistic chemistry, this work connects to renewable-energy catalysis by clarifying how axial ligand identity, oxidation state, and protonation pathways influence dihydrogen release. By probing HCbl formation and reactivity under controlled conditions, this study addresses a long-standing uncertainty in B_{12} chemistry, refining our understanding of how corrinoids can catalyze the formation of $H_{2(g)}$ and informing future efforts to exploit these cofactors in both biological and technological hydrogen-production systems.

4.2. Methods

Synthesis

Hydroxocobalamin acetate (Tokyo Chemical Industry, >90%) glacial acetic acid (Sigma Aldrich, >99.7%), and zinc dust (Sigma Aldrich, <10 μm, >98%) were used as obtained.

Spectroscopy

Abs spectra were obtained with a Cary 5E UV/vis spectrophotometer and CD spectra were collected using a Jasco J-715 spectropolarimeter.

Computational models

Computational models were derived from the high-resolution crystal structure of AdoCbl (CCDC: PAFBUV).[5] The Ado moiety was replaced by the corresponding methyl, hydrogen, or dihydrogen ligand and all side chains, including the nucleotide loop, were truncated and modeled as methyl groups. The dimethylbenzimidazole axial ligand was replaced with a coordinating water molecule.

Geometry optimizations were conducted using the ORCA 5.0 software package.[6,7] Calculations were performed with the PBE functional,[8] the def2-TZVP basis set[9] for Co and the six coordinating atoms, and the def2-SVP basis set[9] for all other atoms. The spin-restricted (S = 0) Kohn-Sham formalism was used for all geometry optimizations. The def2/J auxiliary basis set[10] for Coulomb integral approximation was additionally used to speed up the calculations. Atom-pairwise dispersion correction with the Beck-Johnson damping scheme (D3BJ) empirical dispersion corrections was also applied to all atoms.[11,12]

IRCs for the elongation of the Co–H bond in the base-off HCbl and base-off H_2Cbl geometry-optimized models were generated by performing relaxed potential energy scans where a single Co–H bond was elongated in 0.1 Å increments with a geometry optimization step performed after each 0.1 Å elongation. The dispersion-corrected SCF energy values at each stationary point were then plotted as a function of Co–H distance. An analogous procedure was used to generate the IRC of Co–C elongation in base-off MeCbl. For the IRC associated with the formation H_2Cbl, an H^+ cation was initially placed 2.00 Å away from the upper axial hydride of base-off HCbl, and the H–H distance was

shortened in 0.1 Å increments with a geometry optimization step performed after each 0.1 Å decrease.

TD-DFT computed Abs spectra

An Abs spectrum for each geometry optimized model was calculated using time-dependent density functional theory (TD-DFT) within the Tamm–Dancoff approximation as implemented in the ORCA 5.0 software package. The CAM-B3LYP functional[13] was used for all TD-DFT calculations with the same basis sets as described above. The def2-SVP/C auxiliary basis set for electron correlation was also used.[14] The 60 lowest energy excited states were computed by including all single excitations between molecular orbitals (MOs) within an energy window of ±3 Hartrees with respect to the highest occupied MO (HOMO) and the lowest unoccupied MO (LUMO). The solvation model based on density (SMD), [15] implemented through the CPCM module in ORCA, was used in all TD-DFT calculations. All computed Abs spectra reported here were obtained using the SMD solvent model for acetic acid, unless noted otherwise. The 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 2000 cm^{-1}. All computed spectra were uniformly red-shifted by 3000 cm^{-1} to achieve closer agreement with the experimental spectra.

The NBO7 program[16] with the same parameters employed for the TD-DFT calculations was used to further explore the major bonding interactions for the geometry optimized species. Isosurface plots of natural bonding orbitals (NBOs) and natural localized MOs (NLMOs) were generated with PyMOL v3 using an isodensity value of 0.03 au.

4.3. Results and Discussion

Abs and CD spectra of reduced Co^{3+}Cbl

The RT Abs spectrum of H_2OCbl^+ reduced with Zn dust in glacial HOAc, hereafter referred to as the reduced Cbl species in HOAc, shows clearly discernable features at 17600 cm^{-1} (569 nm) and 21500 cm^{-1} (466 nm), along with a prominent peak at 25400 cm^{-1} (393 nm) (Figure 4.1). This spectrum bears striking similarity to that of aqueous Co^{1+}Cbl formed by reduction of H_2OCbl^+ with Zn dust in water in the presence of an NH_4Cl electrolyte,[17] with the aqueous Co^{1+}Cbl peaks found slightly higher in energy at 17800 cm^{-1} (563 nm), 21600 cm^{-1} (463 nm), and 25800 cm^{-1} (388 nm) (Figure 4.1). Besides the small shifts of all Abs features, the peak intensity at ~21500 cm^{-1} appears ~4× greater for the reduced Cbl species in HOAc than for aqueous Co^{1+}Cbl. Based on the experimental Abs data presented here, the product of H_2OCbl^+ reduction in glacial HOAc most closely resembles that of aqueous Co^{1+}Cbl and does not correspond to a base-off, alkyl-Co^{3+}Cbl species as originally reported by Chemaly and Pratt.[3]

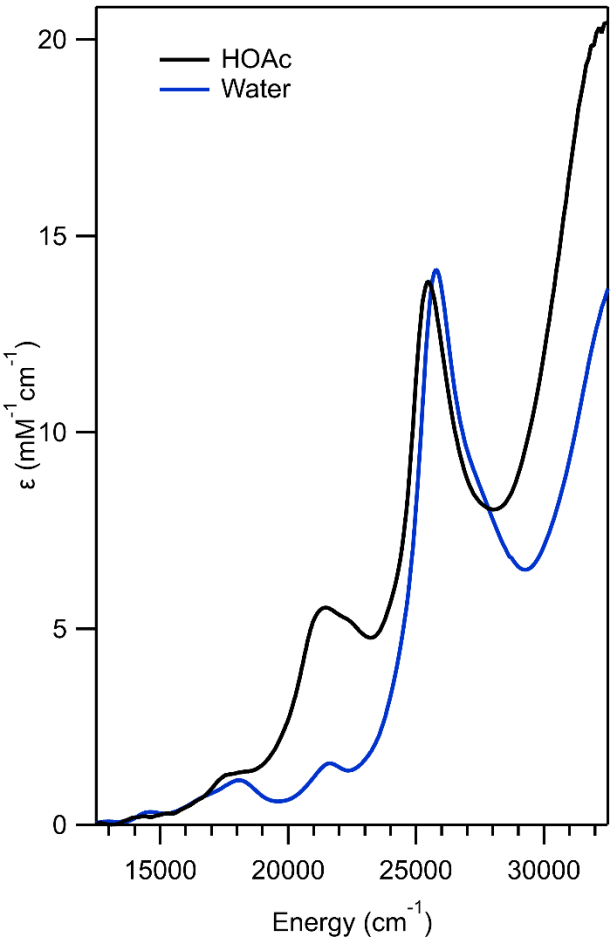

Figure 4.1. Experimental RT Abs spectra of the products formed during the reaction of H_2OCbl^+ with Zn dust in HOAc (black trace) and with Zn dust in water (blue trace).

RT CD spectra were acquired to further experimentally characterize the electronic structure of the reduced Cbl species in HOAc (Figure 4.2). The CD spectra show that the aqueous Co^{1+}Cbl species and the reduced Cbl species formed in HOAc are even more similar than is evident from the Abs spectra. The red-shift of the electronic transitions of the reduced Cbl species in HOAc is clearly reproduced in the CD spectra. Additionally, the molar CD ($\Delta\varepsilon$) values of all peaks for the reduced Cbl species in HOAc and for aqueous Co^{1+}Cbl are extremely similar.

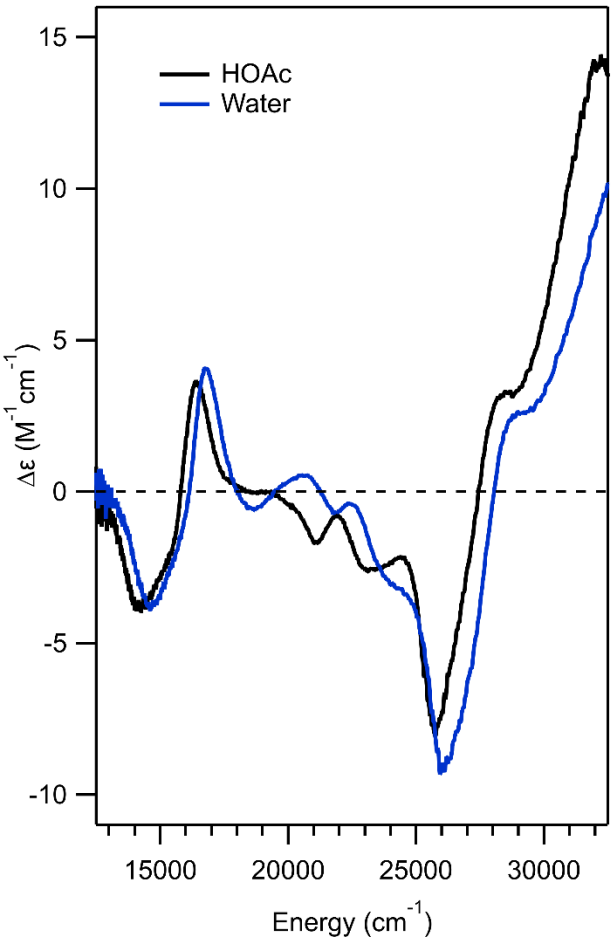

Figure 4.2. Experimental RT CD spectra of the products formed during the reaction of H_2OCbl^+ with Zn dust in HOAc (black trace) or in water (blue trace).

The experimental Abs data, complemented with the CD data here, show that the steady-state species formed during H_2OCbl^+ reduction with Zn dust in glacial HOAc most closely resembles that of the previously characterized, aqueous $Co^{1+}Cbl$. Based on these spectroscopic results, the electronic structure of the reduced Cbl species formed in HOAc is best described as a perturbed $Co^{1+}Cbl$ and bears no resemblance to a $Co^{3+}Cbl$.[18]

Computational Results

To evaluate whether the perturbed Co^{1+}Cbl species formed during H_2OCbl^+ reduction with Zn dust in glacial HOAc corresponds to the proposed HCbl intermediate, a DFT geometry-optimized model of base-off HCbl was generated (Figure 4.3). For comparison, the same computational approach was used to generate a model of base-off MeCbl, for which Abs data can readily be obtained experimentally. The key geometric bonding parameters of the central Co atom in base-off HCbl and MeCbl are shown in Table 4.1.

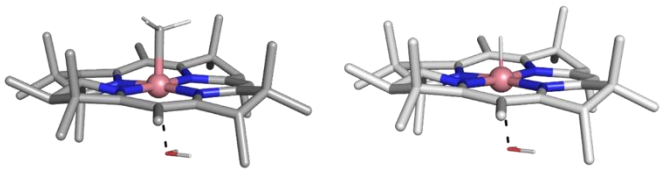

Figure 4.3. The DFT geometry optimized base-off MeCbl model (left) and the base-off HCbl model (right). Only the H atoms of the axial ligands are shown for clarity.

Table 4.1. Key geometric parameters for geometry optimized base-off MeCbl and base-off HCbl models

Species	Co–C / Co–H (Å)	Co–OH_2 (Å)	Co–N_{eq} (Å, avg.)
base-off MeCbl	1.96	2.49	1.90
base-off HCbl	1.44	2.46	1.90

The TD-DFT computed Abs spectrum for the base-off HCbl model (Figure 4.4) does not, in any way, resemble the experimental Abs spectrum of the perturbed Co^{1+}Cbl species observed during the reaction of H_2OCbl^+ with Zn dust in HOAc. Instead, the computed base-off HCbl Abs spectrum closely matches those reported for base-off alkyl-Cbls.[18] To further affirm the similarity of the Abs spectra of base-off HCbl and base-off alkyl-Cbls, TD-DFT was used to compute the Abs spectrum for the base-off MeCbl model.

As expected, the TD-DFT computed Abs spectra for base-off HCbl and base-off MeCbl are nearly identical (Figure 4.4). To validate the TD-DFT computed Abs spectrum for the base-off MeCbl model, we collected a RT Abs spectrum of MeCbl dissolved in glacial HOAc (Figure 4.4). This spectrum is indistinguishable from previously reported spectra of aqueous alkyl-cobinamides, which lack the nucleotide loop and DMB base.[18] Therefore, we conclude that the axial DMB of MeCbl is protonated in glacial HOAc to form a base-off MeCbl species. Importantly, the experimental and TD-DFT computed Abs spectra of base-off MeCbl are in excellent agreement. Thus, the lack of any resemblance between the Abs spectrum of the reduced Cbl species in HOAc and the TD-DFT computed Abs spectrum for HCbl provides further evidence that the major species formed during the reaction of H_2OCbl^+ with Zn dust in HOAc is a weakly perturbed $Co^{1+}Cbl$ rather than HCbl.

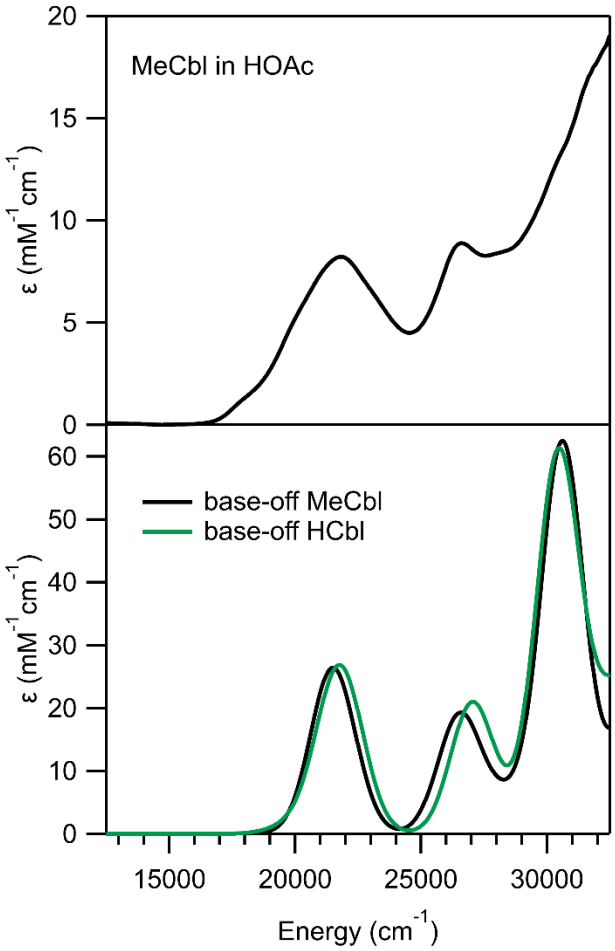

Figure 4.4. Experimental RT Abs spectrum of MeCbl dissolved in HOAc (top) and the TD-DFT computed Abs spectra for base-off MeCbl and base-off HCbl (bottom). The TD-DFT computed spectra were uniformly red-shifted by 3000 cm^{-1} to facilitate a comparison with the experimental Abs spectrum.

base-off HCbl, Natural bonding orbitals (NBOs) and natural localized MOs (NLMOs) were computed for the base-off MeCbl (Figure 4.5) and base-off HCbl (Figure 4.6) models to further investigate the electronic structures of these species. The NBO and NLMO results are consistent with a Co^{3+}Cbl description of base-off HCbl, since three Co d-orbitals with high lone pair electron occupancies can clearly be identified. The NBO analysis of the base-off HCbl model additionally indicates a highly covalent Co–H bond similar in character to the Co–C bond in base-off MeCbl. Therefore, we conclude that

base-off HCbl is best described as a Co^{3+}Cbl species with an axial hydride ligand, bearing striking similarity to alkylcobinamides and base-off alkylcobalamins.

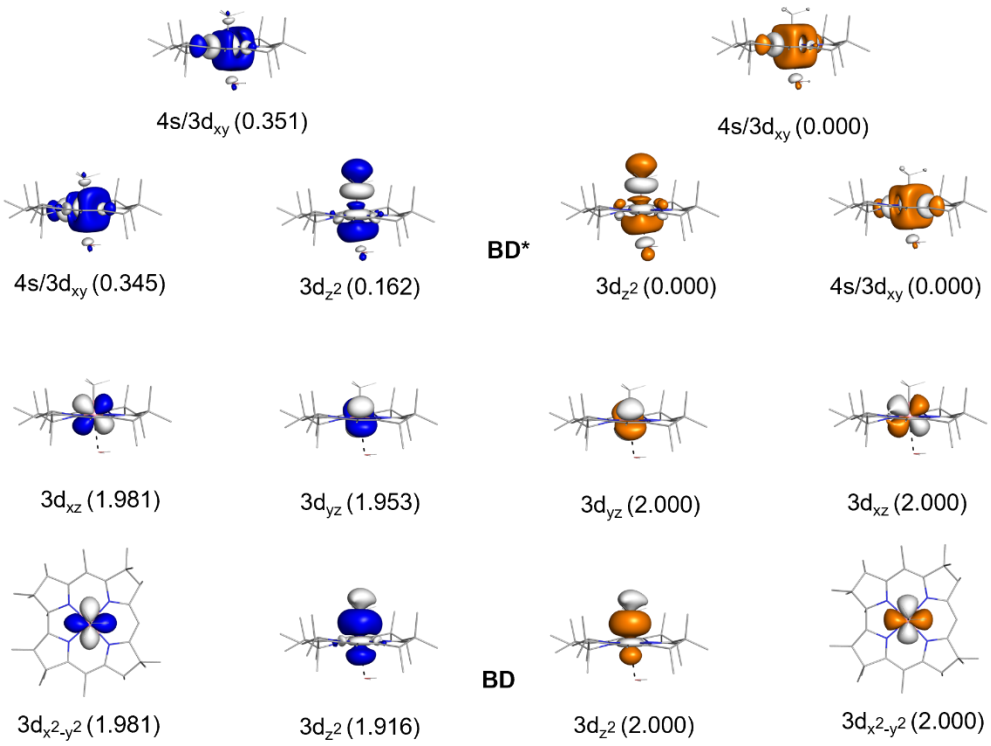

Figure 4.5. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of base-off MeCbl.

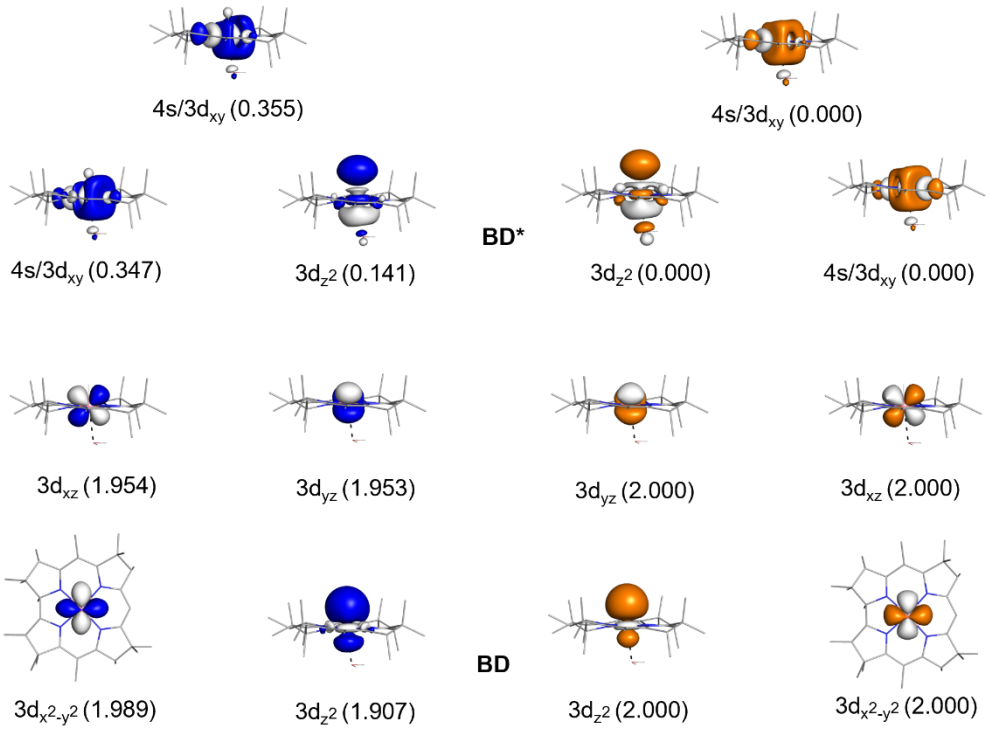

Figure 4.6. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of base-off HCbl.

Mechanism of $H_{2(g)}$ production

Though we determined that the reduced Cbl species in HOAc is best described as a perturbed Co^{1+}Cbl, whose electronic structure bears no resemblance to that predicted for a hydrido-Co^{3+}Cbl species, we investigated the mechanism of $H_{2(g)}$ production with the assumption that an HCbl species is formed during the reduction of H_2OCbl^+ in glacial HOAc but rapidly reacts to produce $H_{2(g)}$. In this scenario, the perturbed Co^{1+}Cbl observed experimentally is first protonated by HOAc to form a hydride-bound base-off HCbl. Base-off HCbl is then rapidly protonated by HOAc to form a base-off H_2Cbl species that instantaneously releases $H_{2(g)}$ to form a five-coordinate Co^{3+}Cbl species. The resulting

five-coordinate Co^{3+}Cbl species is then reduced by Zn dust to reform the perturbed Co^{1+}Cbl. This mechanistic proposal is summarized in Figure 4.7.

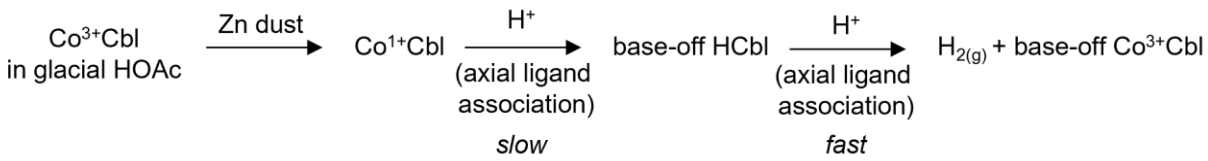

Figure 4.7. Proposed reaction mechanism for the catalytic $H_{2(g)}$ generation by Co^{1+}Cbl..

A relaxed potential energy scan for base-off HCbl was performed by elongating the Co–H bond in 0.1 Å increments to determine the strength of the Co–H bond (Figure 4.8). The computed energy cost for Co–H bond elongation reaches 70.1 kcal/mol from an initial Co–H distance of 1.44 Å to a final Co–H distance of 3.24 Å (net displacement of 1.80 Å). For comparison, the computed Co–C bond elongation energy for base-off MeCbl is lower, at 57 kcal/mol, from the initial distance of 1.97 Å to a final Co–C distance of 3.77 Å (net displacement of 1.80 Å). Therefore, H^+ association to the perturbed Co^{1+}Cbl should result in a strong, highly covalent Co–H bond.

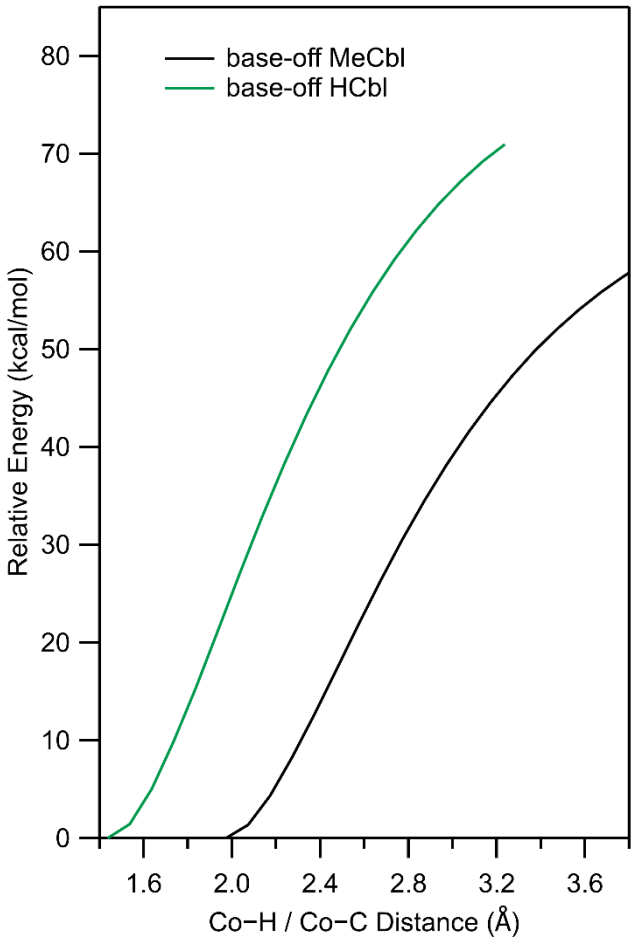

Figure 4.8. Relaxed potential energy scans for the elongation of the Co–C bond in base-off MeCbl (black trace) and the elongation of the Co–H bond in base-off HCbl (green trace) in 0.1 Å increments. For both species, the relative energies of the first 18 stationary points of these scans are shown.

To assess the ease of dihydrogen release from the putative base-off HCbl intermediate, an additional relaxed potential energy scan was performed starting with the geometry-optimized base-off HCbl model and an additional H^+ placed 2.00 Å away from the axial hydride. The H–H distance was decreased in 0.1 Å increments and a geometry optimization was performed after each step. The resulting IRC shows that protonation of base-off HCbl to form H_2-bound base-off H_2Cbl has an energy barrier of 21.3 kcal/mol (Figure 4.9). The resulting H_2Cbl species contains H_2 bound to the Co atom, with a slightly

elongated H–H bond distance of 0.85 Å and the electronic structure analysis (by means of computed Abs spectrum) of this species being consistent with a Co^{3+}Cbl description (Figure S4.3). Therefore, H^+ association to HCbl is significantly more enthalpically favorable than H^+ or H^- dissociation from HCbl.

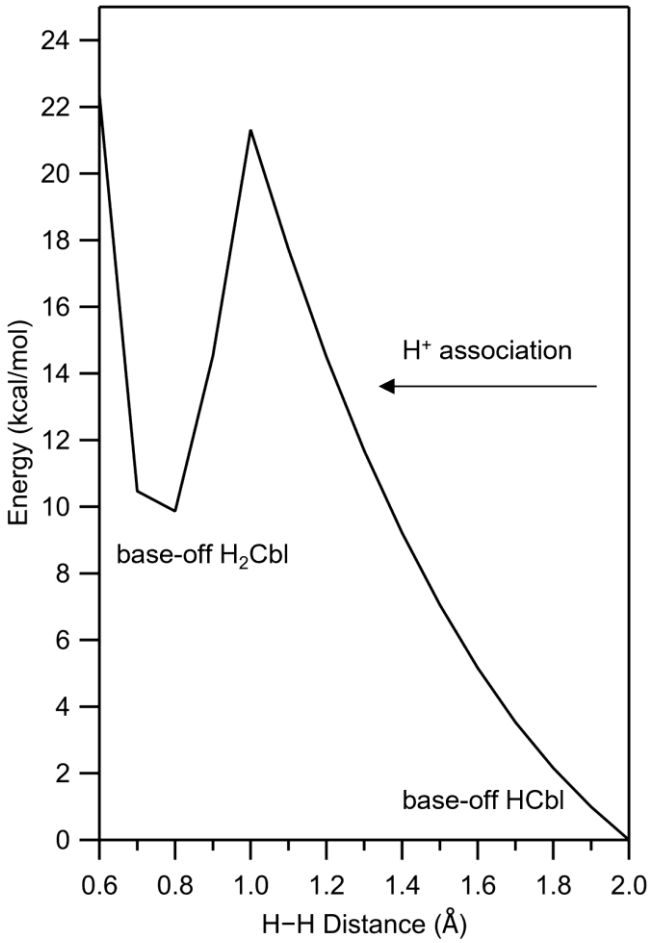

Figure 4.9. Relaxed potential energy scan for the protonation of HCbl starting with the geometry optimized base-off HCbl model and an H^+ placed 2.00 Å from the bound hydride to form the putative six-coordinate, base-off H_2Cbl species. The H–H distance was shortened in 0.1 Å increments.

To explore the IRC of the reverse process described above (i.e. H^+ dissociation from base-off H_2Cbl) a geometry-optimized base-off H_2Cbl model was generated and one of the Co–H coordinates was elongated incrementally by 0.1 Å with a geometry

optimization performed after each step (Figure 4.10). Notably, an extremely low energy barrier of 4.0 kcal/mol was computed for the elongation of this Co–H coordinate. Instead of an H^+ dissociating from base-off H_2Cbl to form base-off HCbl, we found that when a single Co–H coordinate is elongated starting from base-off H_2Cbl, the other Co–H bondconcurrently elongates to form H_2 and a five-coordinate Co^{3+}Cbl species. Therefore, the calculated energy barrier for the release of H_2 from base-off H_2Cbl is ~5× lower than the 21.3 kcal/mol needed for releasing H^+ from base-off H_2Cbl. Remarkably, the energy barrier for $H_{2(g)}$ dissociation from base-off H_2Cbl is over an order of magnitude lower than the energy barrier calculated for H^+ dissociation from base-off HCbl (70.1 kcal/mol). Entropy effects that are not accounted for in these computations are expected to even further increase the favorability of $H_{2(g)}$ dissociation.

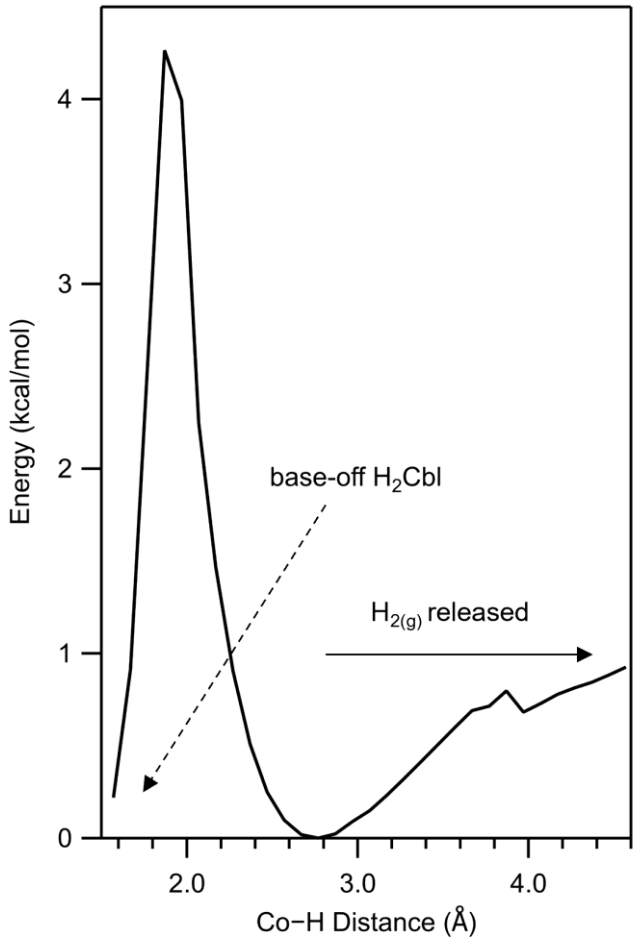

Figure 4.10. Relaxed potential energy scan for the dissociation of $H_{2(g)}$ starting with the geometry optimized base-off H_2Cbl model obtained by elongating one of the Co–H bonds in 0.1 Å increments.

4.4. Conclusions

Catalytic production of $H_{2(g)}$ from Cbl in acidic solution is shown here to operate from a steady state concentration of a perturbed Co^{1+}Cbl species. The perturbed Co^{1+}Cbl observed experimentally bears no electronic resemblance to the computed electronic structure of a protonated Co^{1+}Cbl, since our computations support that a protonated Co^{1+}Cbl is actually best characterized as a hydride-bound Co^{3+}Cbl. The exact nature of the perturbation to Co^{1+}Cbl in HOAc leading to the red-shifted Abs and CD spectra is

unknown but could result from i) a change in the dielectric environment and/or ii) the harsh reaction conditions including the presence of Zn^{2+} cations and OAc^- anions, all of which may interact with Co^{1+}Cbl non-covalently to perturb this species. Such perturbations have been observed resulting from Li^+ cations and Cl^- anions interacting with Cbl in the solid state.[19] Ultimately, base-off HCbl and base-off H_2Cbl are not stable enough to be observed under our experimental conditions, a conclusion supported by the relaxed potential energy scans performed for the formation of base-off HCbl, formation of base-off H_2Cbl, and the release of $H_{2(g)}$ from base-off H_2Cbl. Carrying out the reduction of Cbl with a strong reducing agent (Zn dust) in acidic solution at low temperatures, large $H_{2(g)}$ pressures, and/or ultrafast spectroscopic methods are likely needed to capture and characterize the Co^{3+}Cbl-like HCbl and H_2Cbl reactive intermediates.

References

(1) Jost, M.; H. Simpson, J.; L. Drennan, C. The Transcription Factor CarH Safeguards Use of Adenosylcobalamin as a Light Sensor by Altering the Photolysis Products. *Biochemistry* **2015**, *54* (21), 3231–3234. https://doi.org/10.1021/acs.biochem.5b00416.

(2) 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.

(3) Chemaly, S. M.; Pratt, J. M. The Chemistry of Vitamin B_{12}. Part 24. Evidence for Hydride Complexes of Cobalt(III) Corrinoids. *J. Chem. Soc. Dalt. Trans.* **1984**, 595–599.

(4) Toda, M. J.; Lodowski, P.; Mamun, A. Al; Kozlowski, P. M. Electronic and Photolytic Properties of Hydridocobalamin. *J. Photochem. Photobiol. B Biol.* **2021**, *224* (August), 112295. https://doi.org/10.1016/j.jphotobiol.2021.112295.

(5) 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.

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

(7) Neese, F. Software Update: The ORCA Program System—Version 5.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2022**, *12* (5), 1–15. https://doi.org/10.1002/wcms.1606.

(8) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. https://doi.org/10.1021/acsami.8b15784.

(9) 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.

(10) 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.

(11) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. https://doi.org/10.1002/jcc.

(12) 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).

(13) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1–3), 51–57. https://doi.org/10.1016/j.cplett.2004.06.011.

(14) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor. Chem. Acc.* **2007**, *117* (4), 587–597. https://doi.org/10.1007/s00214-007-0250-5.

(15) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396. https://doi.org/10.1021/jp810292n.

(16) NBO 7.0. E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou, C. R. Landis, and F. Weinhold, Theoretical Chemistry Institute, University of Wisconsin, Madison, WI (2018).

(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. **2006**, No. 15, 9144–9156. https://doi.org/10.1021/ja061433q.

(18) 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.

(19) Randaccio, L.; Furlan, M.; Geremia, S.; Slouf, M. Preparation and X-Ray Analysis of Crystals of Azido- and Chlorocobalamin Containing LiCl: A Structural Model for the Interactions of the Corrin Ring with Ionic Species. *Inorg. Chem.* **1998**, *37* (20), 5390–5393. https://doi.org/10.1021/ic980181u.

Appendix

Appendix

Chapter 2

Spectroscopic and Computational Insights into the Mechanism of Cofactor Cobalt–Carbon Bond Homolysis by the Adenosylcobalamin-Dependent Enzyme Ethanolamine Ammonia-Lyase

Supporting Information

Appendix 1.

Detailed computational workflow used for building structural models of solvated AdoCbl, holoEAL, and the ternary complex prior to QM/MM optimization.

Solvated AdoCbl

Initial atomic coordinates for the solvated AdoCbl model were extracted from the corresponding X-ray crystal structure (CCDC ID: PAFBUV).[1] Published AMBER MM parameters for AdoCbl were used[2] and units were converted for GROMACS compatibility. AdoCbl partial charges and Lennard-Jones parameters were taken from the literature.[3] AdoCbl was solvated in a 103.5 × 103.5 × 103.5 nm^3 TIP3P water box using GROMACS. An energy minimization was carried out using the steepest descent method. A constant volume equilibration at 300 K was performed for 100 ps with a time step of 2 ps using 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 AdoCbl of 10 kJ mol^{-1}. After these equilibrations, a 5 ns MD run was completed at 300 K with a time step of 2 fs and long-range interactions were treated with the particle mesh Ewald method. Following the 5 ns MD run, the heavy atoms of the AdoCbl model (Figure S2.3) were used for GROMACS clustering with an RMSD cutoff of 0.08 nm to create 18 clusters (Figure S2.4). The first cluster's average occupancy corresponded to timepoint 420 ps. Therefore, the AdoCbl coordinates and all waters within 29 nm of AdoCbl at timepoint 420 ps were selected for subsequent QM/MM optimization.

HoloEAL

Initial atomic coordinates for the holoEAL model were extracted from the crystal structure of EAL complexed with AdoCbl (PDB: 5YSN).[4] One α/β unit of holoEAL was protonated at pH 7.0 using the PDB2PQR server.[5] Published AMBER MM parameters for AdoCbl were used[2] and units were converted for GROMACS compatibility. AdoCbl partial charges and Lennard-Jones parameters were taken from the literature.[3] The active site was visually inspected to confirm appropriate protonation of charged side chains at pH 7.0, namely that E287a, D362a were deprotonated and R160a was protonated. The protein was modeled with the AMBER14SB forcefield.[6] The holoEAL model was solvated in 103.5 × 103.5 × 103.5 nm^3 TIP3P water box and 17 Na^+ ions were added to neutralize the overall charge. Waters added by GROMACS within 10 Å of the Co atom were removed to ensure the holoEAL active site contained only crystallographic waters. These steps constitute Step A in Figure S2.2. An energy minimization was performed on the solvated holoEAL model using the steepest descent method in GROMACS to reform the

Co–C bond of AdoCbl within the EAL active site, followed by a constant volume equilibration at 300 K for 100 ps with a time step of 2 ps using 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^{-1}. After these equilibrations, a 5 ns MD run was completed at 300 K with a time step of 2 fs and long-range interactions were treated with particle mesh Ewald method. Following the 5 ns MD run, the heavy atoms of the AdoCbl/EAL model substructure (Figure S2.6) were used for GROMACS clustering with an RMSD cutoff of 0.041 nm to create 15 clusters. The first cluster's average occupancy corresponded to timepoint 3190 ps. Therefore, the holoEAL coordinates and all waters within 5 Å of the protein at timepoint 3190 ps were selected for subsequent QM/MM optimization. These steps constitute Step B in Figure S2.2.

Ternary Complex

Initial atomic coordinates for the ternary complex model were extracted from the crystal structure of EAL complexed with AdoCbl (PDB: 5YSN).[4] One α/β unit of holoEAL was protonated at pH 7.0 using the PDB2PQR server.[5] The active site was visually inspected to confirm appropriate protonation of charged side chains at pH 7.0, namely that E287a, D362a were deprotonated and R160a was protonated. The protein was modeled with the AMBER14SB forcefield.[6] Published AMBER MM parameters for AdoCbl were used[2] and converted for GROMACS compatibility. AdoCbl partial charges and Lennard-Jones parameters were taken from the literature.[3] To model EA binding, heavy atom coordinates of EA were extracted from the crystal structure of the AdePeCbl/EAL/EA complex (PDB: 3ABS).[7] H atoms were added to EA using Pymol, and EA was parameterized using the General AMBER forcefield[8] and AM1-BCC charge method[9] within AmberTools. GROMACS-compatible units were generated using the amb2gro program. The two active site waters found closest to Eα287 were removed and replaced with EA. The AdoCbl/EAL/EA model was solvated in 103.5 × 103.5 × 103.5 nm^3 TIP3P water box and 16 Na^+ counterions were added to neutralize the overall charge. Waters added by GROMACS 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 reform the Co–C bond within the EAL active site, followed by a constant volume equilibration at 300 K for 100 ps with a time step of 2 ps using position restraints on the heavy atoms of the protein, AdoCbl, and EA 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, AdoCbl, and EA of 10 kJ mol^{-1}. After these equilibrations, a 5 ns MD run was completed at 300 K with a time step of 2 fs and long-range interactions were treated with particle mesh Ewald method. Following the 5 ns MD run, heavy atoms of the AdoCbl/EAL/EA model substructure (Figure S2.9) were used for GROMACS clustering with an RMSD cutoff of 0.042 nm. The first cluster's average occupancy corresponded to time point 2060 ps. Therefore, the AdoCbl/EAL/EA coordinates and all waters within 5 Å of the protein at timepoint 2060 ps were selected for subsequent QM/MM optimization.

Appendix 2.

Sample input files for holoEAL energy minimization, NVT equilibration, NPT equilibration, and 5 ns MD simulation.

Energy Minimization

```
title                = Minimization
integrator           = steep
emtol                = 1000.0
emstep          = 0.01
nsteps               = 50000
nstlist              = 1
cutoff-scheme    = Verlet
ns_type                    = grid
rlist                = 1.2
coulombtype      = PME
rcoulomb           = 1.2
vdwtype          = cutoff
vdw-modifier     = force-switch
rvdw-switch      = 1.0
rvdw                 = 1.2
pbc            = xyz
DispCorr         = no
```

NVT Equilibration

```
title                = HoloEAL NVT equilibration
define               = -DPOSRES
integrator           = md
nsteps               = 50000
dt                   = 0.002
nstenergy            = 500
nstlog               = 500
nstxout-compressed       = 500
continuation             = no
constraint_algorithm     = lincs
constraints              = h-bonds
lincs_iter               = 1
lincs_order              = 4
cutoff-scheme            = Verlet
ns_type                  = grid
nstlist                  = 20
```

```
rlist                = 1.2
vdwtype                = cutoff
vdw-modifier           = force-switch
rvdw-switch            = 1.0
rvdw                 = 1.2
coulombtype            = PME
rcoulomb             = 1.2
pme_order              = 4
fourierspacing         = 0.16
tcoupl               = V-rescale
tc-grps              = Protein_5D1_B21_3CO1 Water_and_ions
tau_t                = 0.1   0.1
ref_t                = 300   300
pcoupl                 = no
pbc                  = xyz
DispCorr               = no
gen_vel                = yes
gen_temp               = 300
gen_seed               = -1
```

NPT Equilibration

```
title                = HoloEAL NPT equilibration
define                 = -DPOSRES
integrator             = md
nsteps               = 50000
dt                   = 0.002
nstenergy              = 500
nstlog               = 500
nstxout-compressed     = 500
continuation           = yes
constraint_algorithm   = lincs
constraints            = h-bonds
lincs_iter           = 1
lincs_order            = 4
cutoff-scheme          = Verlet
ns_type              = grid
nstlist              = 20
rlist                = 1.2
vdwtype                = cutoff
```

```
vdw-modifier            = force-switch
rvdw-switch             = 1.0
rvdw                    = 1.2
coulombtype             = PME
rcoulomb                = 1.2
pme_order               = 4
fourierspacing          = 0.16
tcoupl                  = V-rescale
tc-grps                 = Protein_5D1_B21_3CO1 Water_and_ions
tau_t                   = 0.1   0.1
ref_t                   = 300   300
pcoupl                  = Berendsen
pcoupltype              = isotropic
tau_p                   = 2.0
ref_p                   = 1.0
compressibility         = 4.5e-5
refcoord_scaling        = com
pbc                     = xyz
DispCorr                = no
gen_vel                 = no
```

MD Simulation

```
title                   = HoloEAL MD simulation
integrator              = md
nsteps                  = 2500000
dt                      = 0.002
nstenergy               = 5000
nstlog                  = 5000
nstxout-compressed      = 5000
continuation            = yes
constraint_algorithm    = lincs
constraints             = h-bonds
lincs_iter              = 1
lincs_order             = 4
cutoff-scheme           = Verlet
ns_type                 = grid
nstlist                 = 20
rlist                   = 1.2
```

```
vdwtype                 = cutoff
vdw-modifier            = force-switch
rvdw-switch             = 1.0
rvdw                    = 1.2
coulombtype             = PME
rcoulomb                = 1.2
pme_order               = 4
fourierspacing          = 0.16
tcoupl                  = V-rescale
tc-grps                 = Protein_5D1_B21_3CO1 Water_and_ions
tau_t                   = 0.1   0.1
ref_t                   = 300   300
pcoupl                  = Parrinello-Rahman
pcoupltype              = isotropic
tau_p                   = 2.0
ref_p                   = 1.0
compressibility         = 4.5e-5
pbc                     = xyz
DispCorr                = no
gen_vel                 = no
```

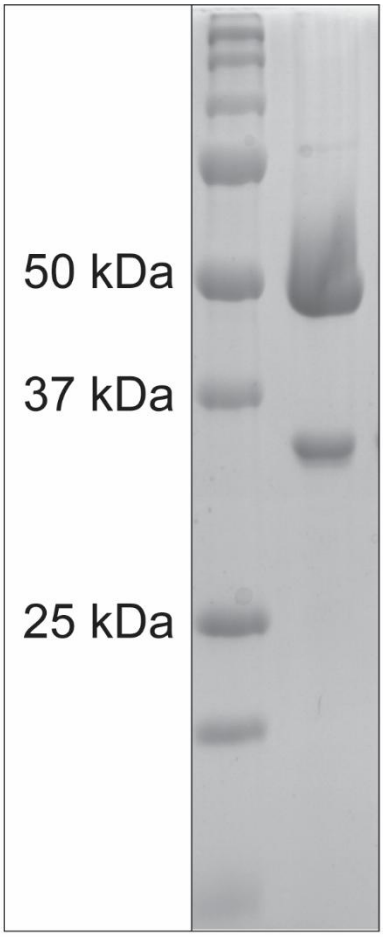

Figure S2.1. SDS-PAGE gel image demonstrating the purity of the EAL preparation used in this study. Purified EAL (12 mg) was loaded onto a 15% SDS-PAGE gel. A constant current of 200 V was applied to the gel for ~50 min. The molecular masses of the EutB and EutC subunits are 50 kDa and 30 kDa, respectively.

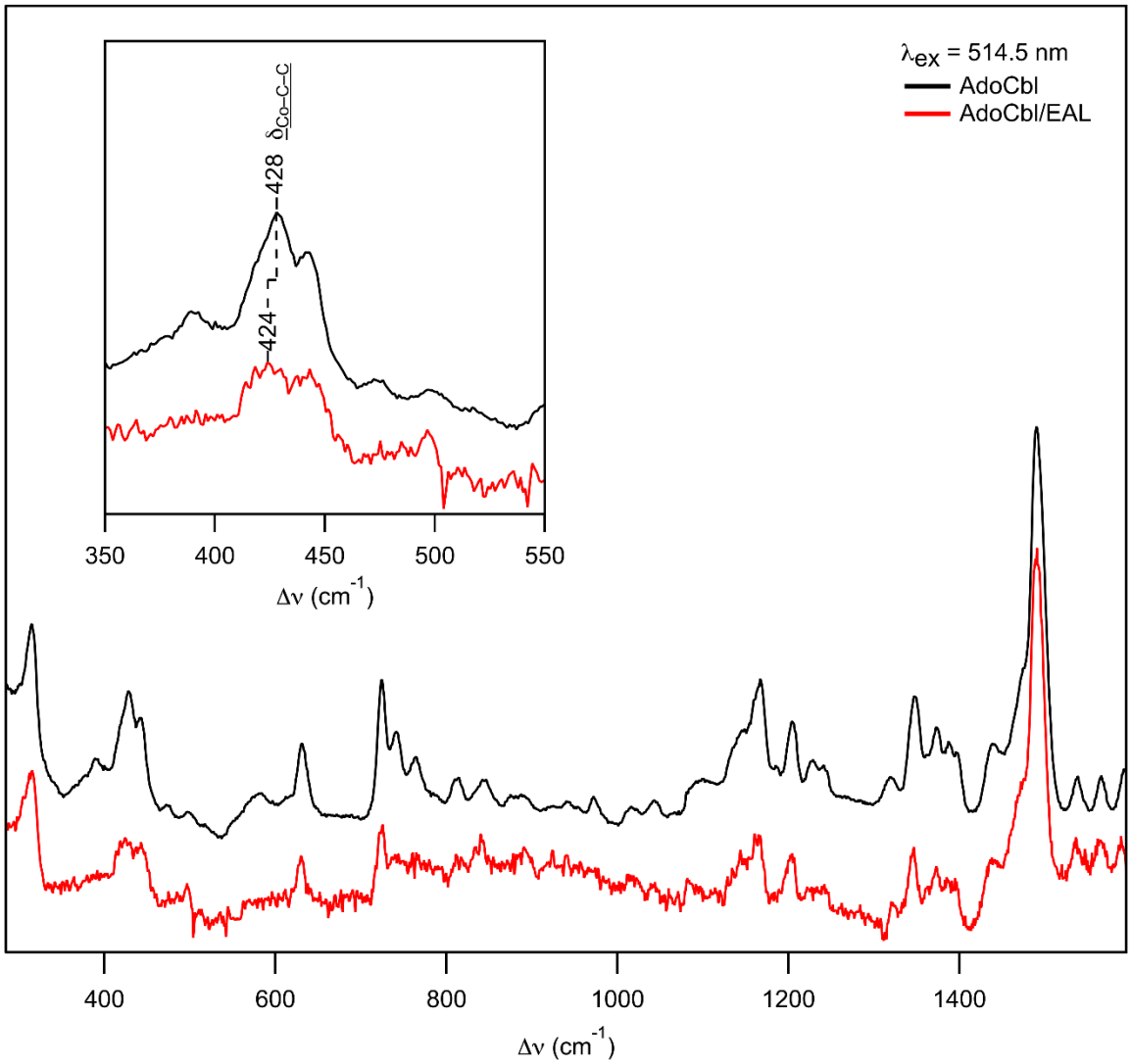

Figure S2.2. Resonance Raman spectra of free (black) and EAL-bound (red) AdoCbl, obtained with 514.5 nm laser excitation. Inset: Magnification of the region containing a peak associated with a Co–C–C bending mode (δ_{Co-C-C}).

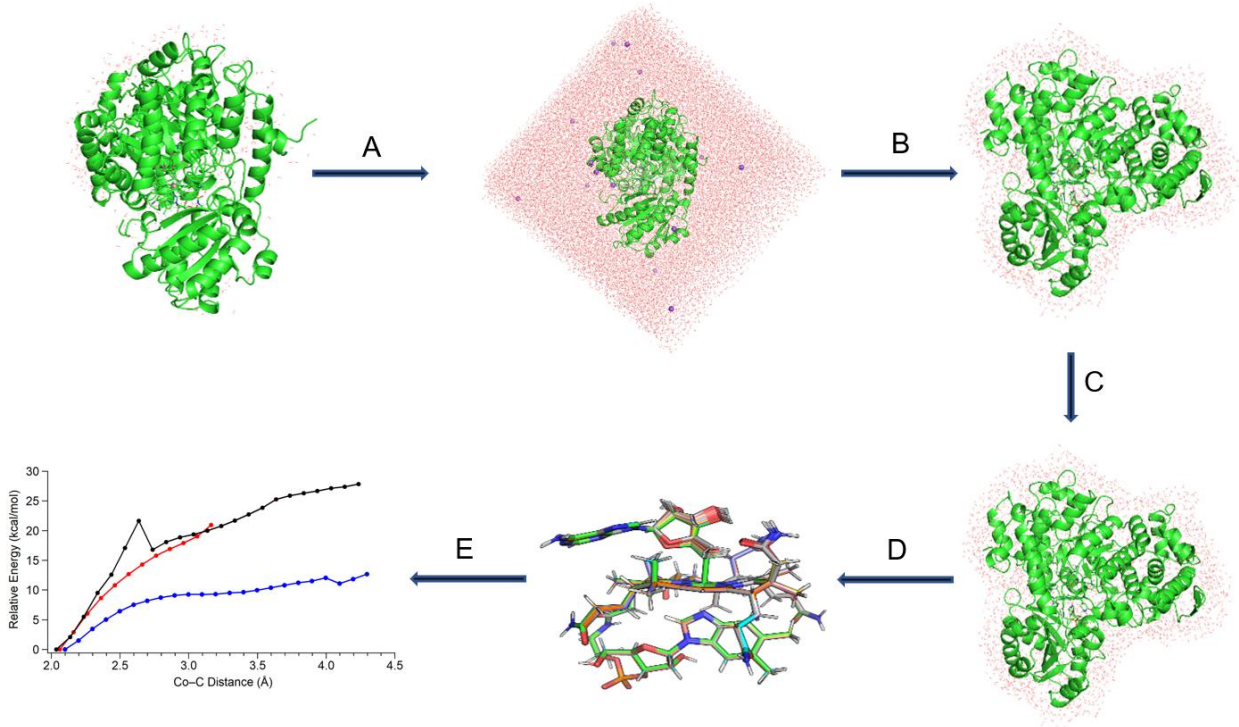

Figure S2.3. MD + QM/MM computational workflow used for generating the computational models described in the text. Steps A and B are described in the holoEAL section of Appendix 1. Step C represents the QM/MM geometry optimization after a single structure was selected from the clustering algorithm. Step D represents Co–C bond elongation performed in 0.1 Å increments with a subsequent QM/MM geometry optimization. Step E represents the construction of QM energy plots for every stationary point along the Co–C bond elongation coordinate.

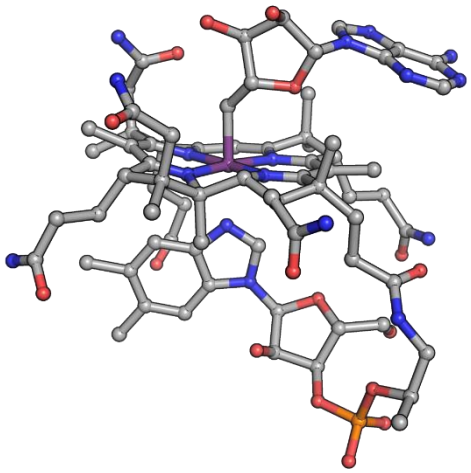

Figure S2.4. Heavy atoms of the AdoCbl model used for the RMSD-based clustering calculation. All non-hydrogen QM atoms are contained within this structure.

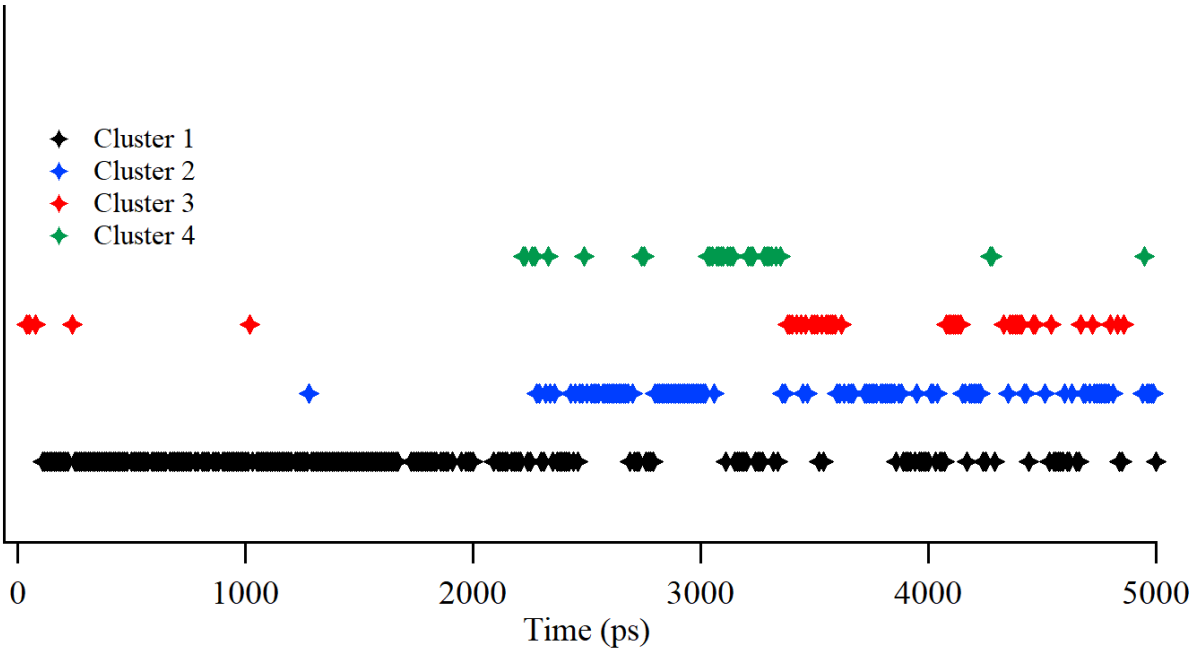

Figure S2.5. Cluster occupancy for the heavy atoms of the AdoCbl model (Figure S2.3) obtained using an RMSD cutoff of 0.08 nm. Number of structures in each cluster out of 501: Cluster 1 = 246 ; Cluster 2 = 111 ; Cluster 3 = 43 ; Cluster 4 = 30 ; Clusters 5 through 18 (not shown) = 71.

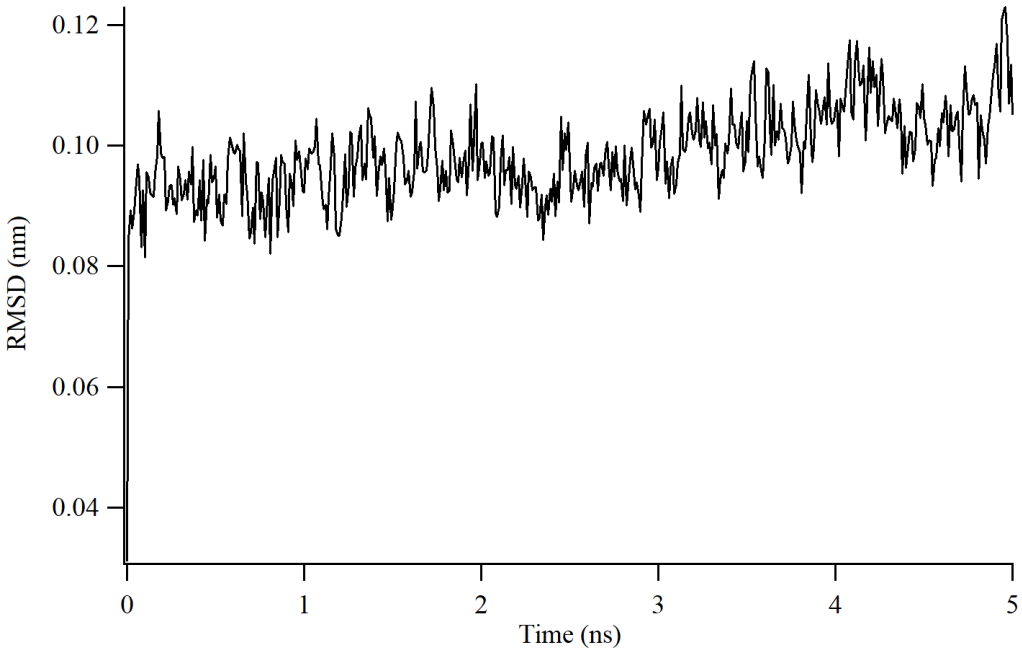

Figure S2.6. RMSD distances of the AdoCbl/EAL model protein backbone calculated for the 5 ns MD run in reference to the crystal structure PDB: 5YSN before energy minimization, NVT, and NPT equilibration.

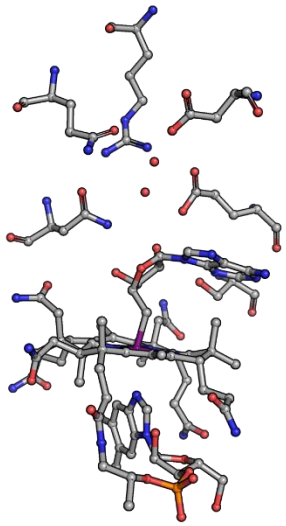

Figure S2.7. Heavy atoms of the AdoCbl/EAL model used for RMSD-based clustering calculation. All non-hydrogen QM atoms are contained within this structure.

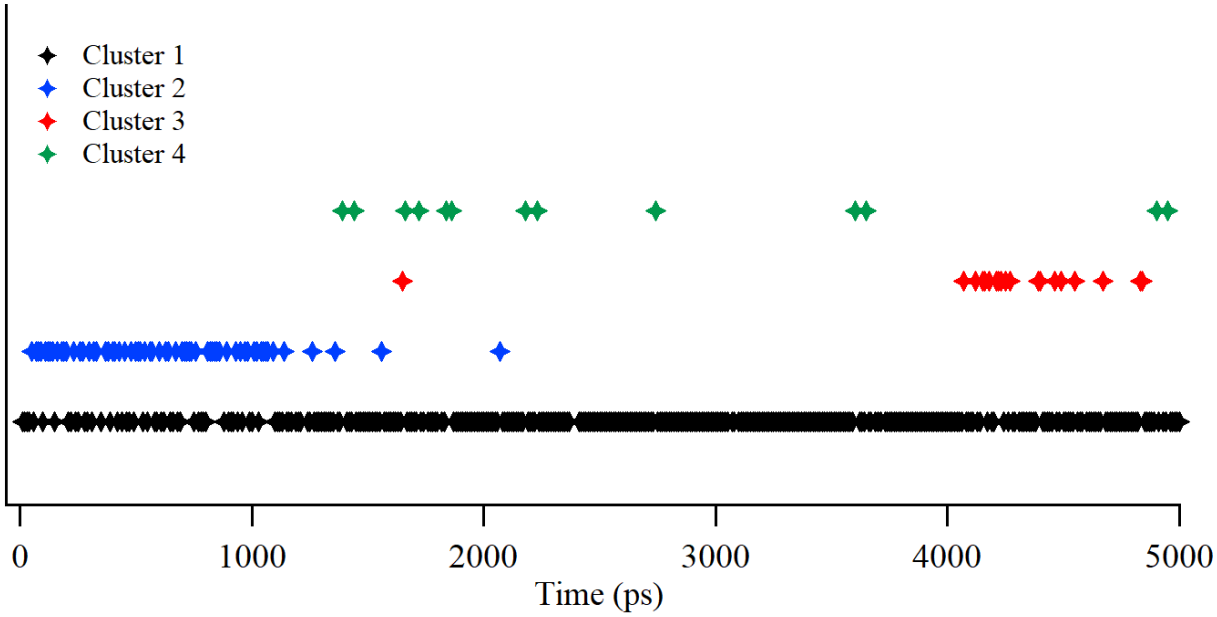

Figure S2.8. Cluster distribution for the heavy atoms of the AdoCbl/EAL model substructure (Figure S2.6) obtained using an RMSD cutoff of 0.041 nm. Number of structures in each cluster out of 501: Cluster 1 = 381 ; Cluster 2 = 64 ; Cluster 3 = 19 ; Cluster 4 = 13 ; Clusters 5 through 15 (not shown) = 24.

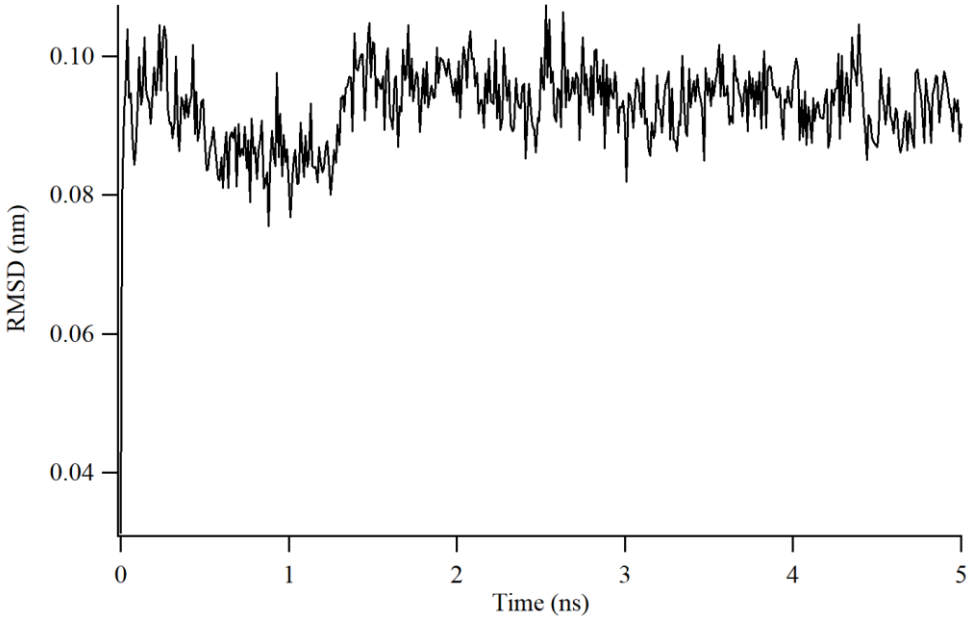

Figure S2.9. RMSD distances of the AdoCbl/EAL/EA model protein backbone calculated for the 5 ns MD run in reference to the crystal structure PDB: 5YSN before energy minimization, NVT, and NPT equilibration.

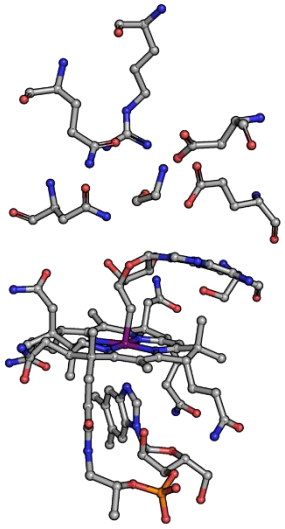

Figure S2.10. Heavy atoms of the AdoCbl/EAL/EA model used for RMSD-based clustering calculation. All non-hydrogen QM atoms are contained within this structure.

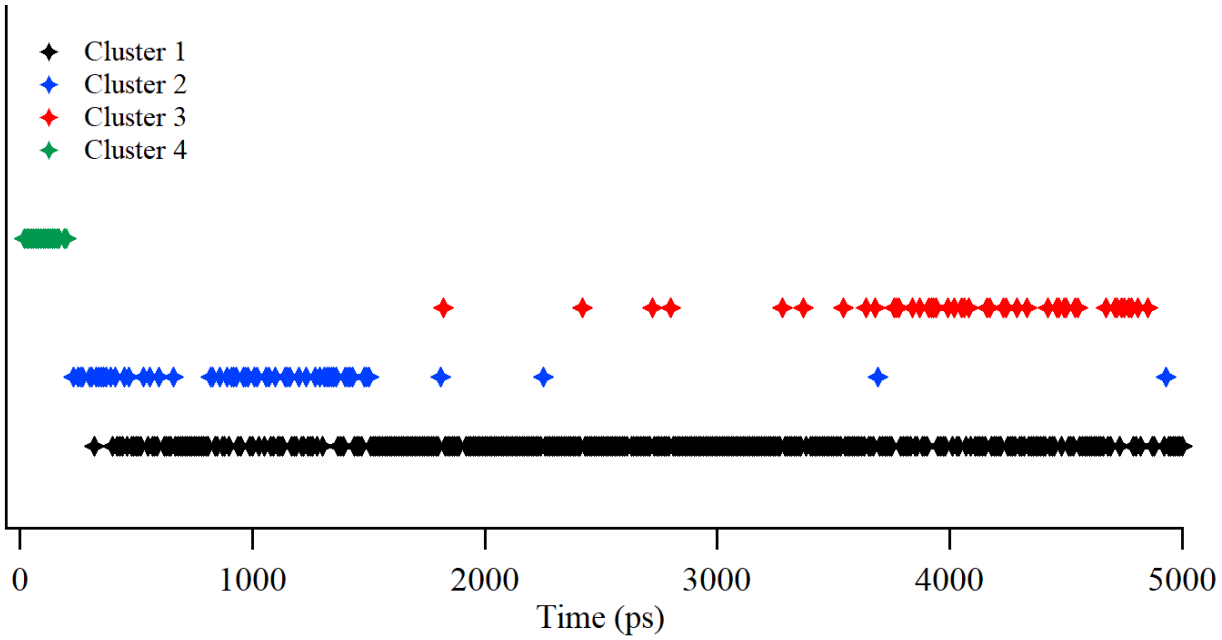

Figure S2.11. Cluster distribution for the heavy atoms of the AdoCbl/EAL/EA model substructure (Figure S2.9) using an RMSD cutoff of 0.042 nm. Number of structures in each cluster out of 501: Cluster 1 = 354 ; Cluster 2 = 58 ; Cluster 3 = 45; Cluster 4 = 18 ; Clusters 5 through 14 (not shown) = 26.

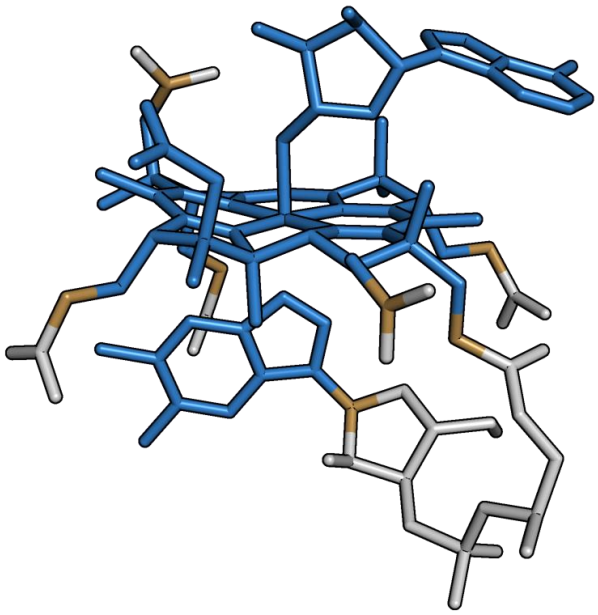

Figure S2.12. QM Region of model of solvated AdoCbl. QM atoms are shown as blue sticks, link atoms in tan, and MM atoms as white sticks.

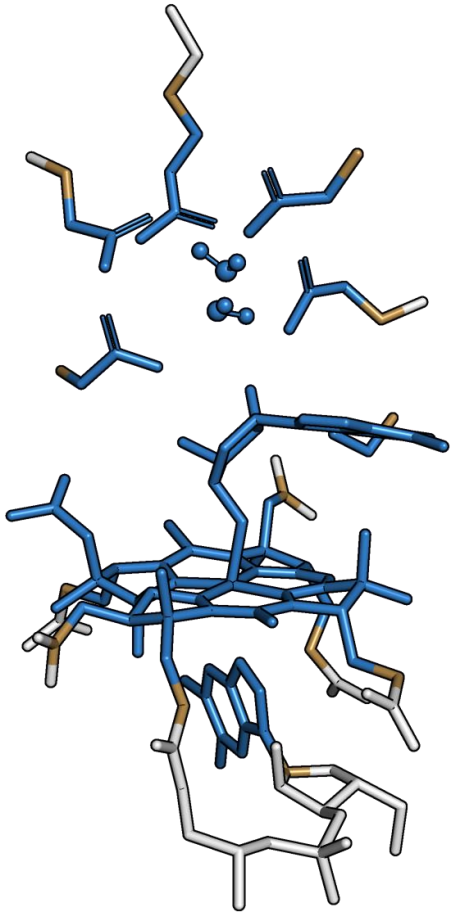

Figure S2.13. QM Region of model of holoEAL. QM atoms are shown as blue sticks, link atoms in tan, and MM atoms as white sticks.

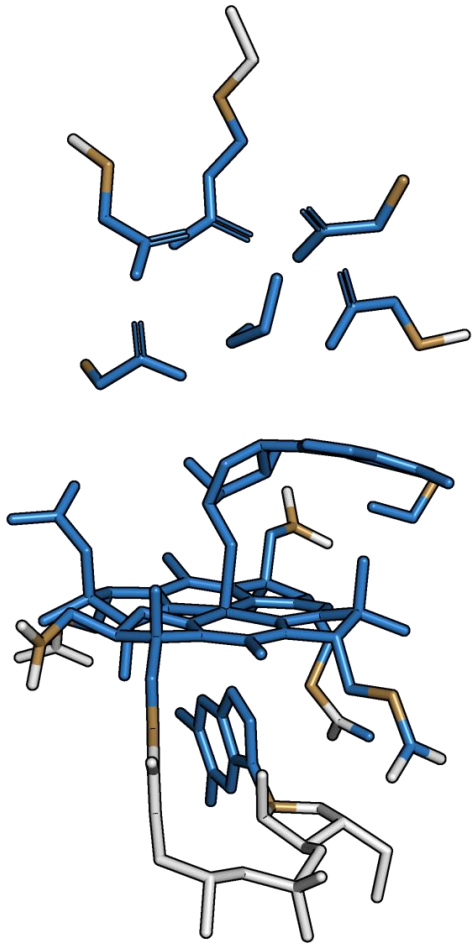

Figure S2.14. Ternary Complex QM Region. QM atoms are shown as blue sticks, link atoms in tan, and MM atoms as white sticks.

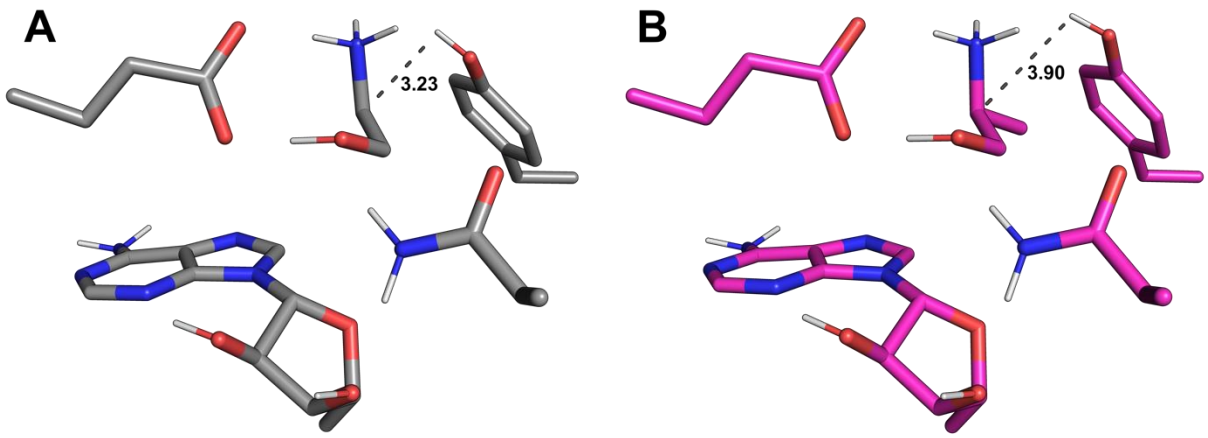

Figure S2.15. Proximity of Y404a to substrate in QM/MM-optimized models of AdoCbl/EAL/EA (gray) and AdoCbl/EAL/2AP (pink).

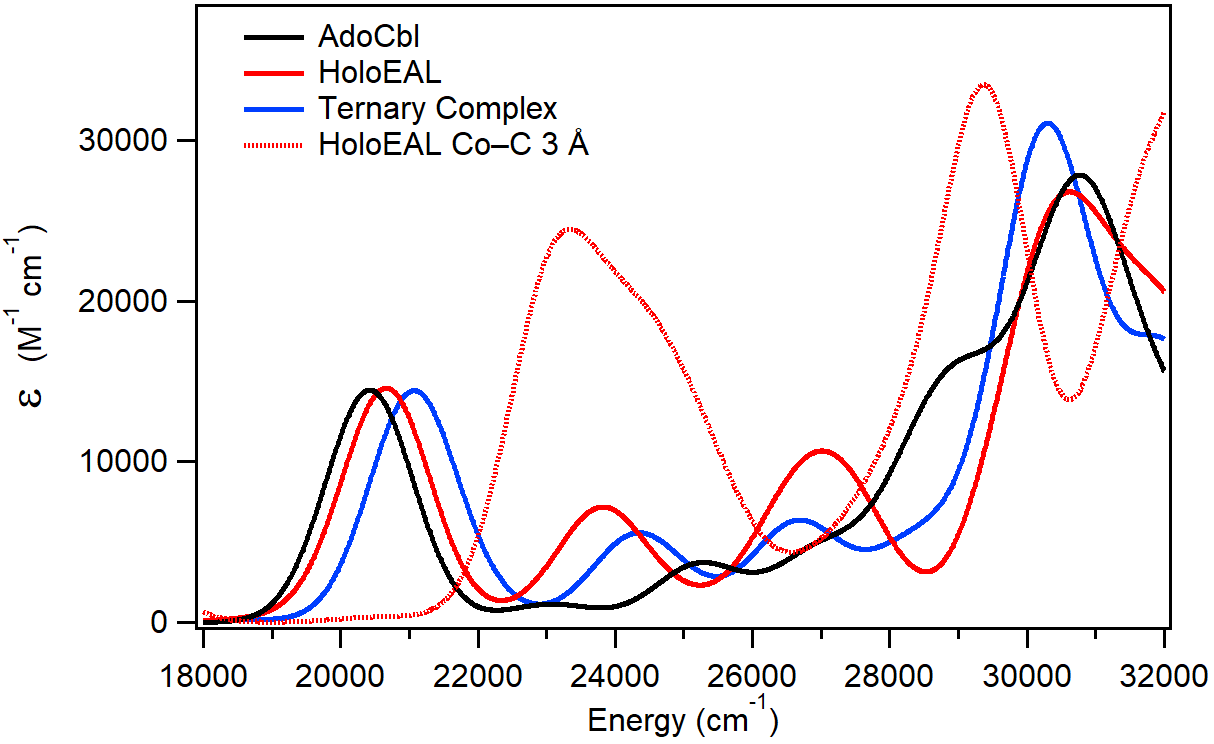

Figure S2.16. TD-DFT computed absorption spectra for the QM/MM-optimized models of free and EAL-bound AdoCbl (solid traces) and the QM/MM-optimized holoEAL model with a fixed Co–C bond distance of 3 Å as observed in the crystal structure PDB: 5YSN.

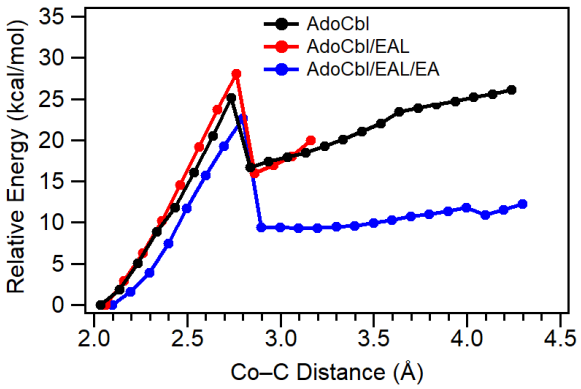

Figure S2.17. SP QM energies for all stationary points during Co–C bond elongation using the full QM region of each model and the larger def2-TZVP basis set for all QM atoms. All energies are in reference to each model's ground state QM energy prior to Co–C bond elongation.

Appendix

Chapter 3

Spectroscopic and Computational Analysis of the Halocobalamin (XCbl, X = F, Cl, Br, I) Series

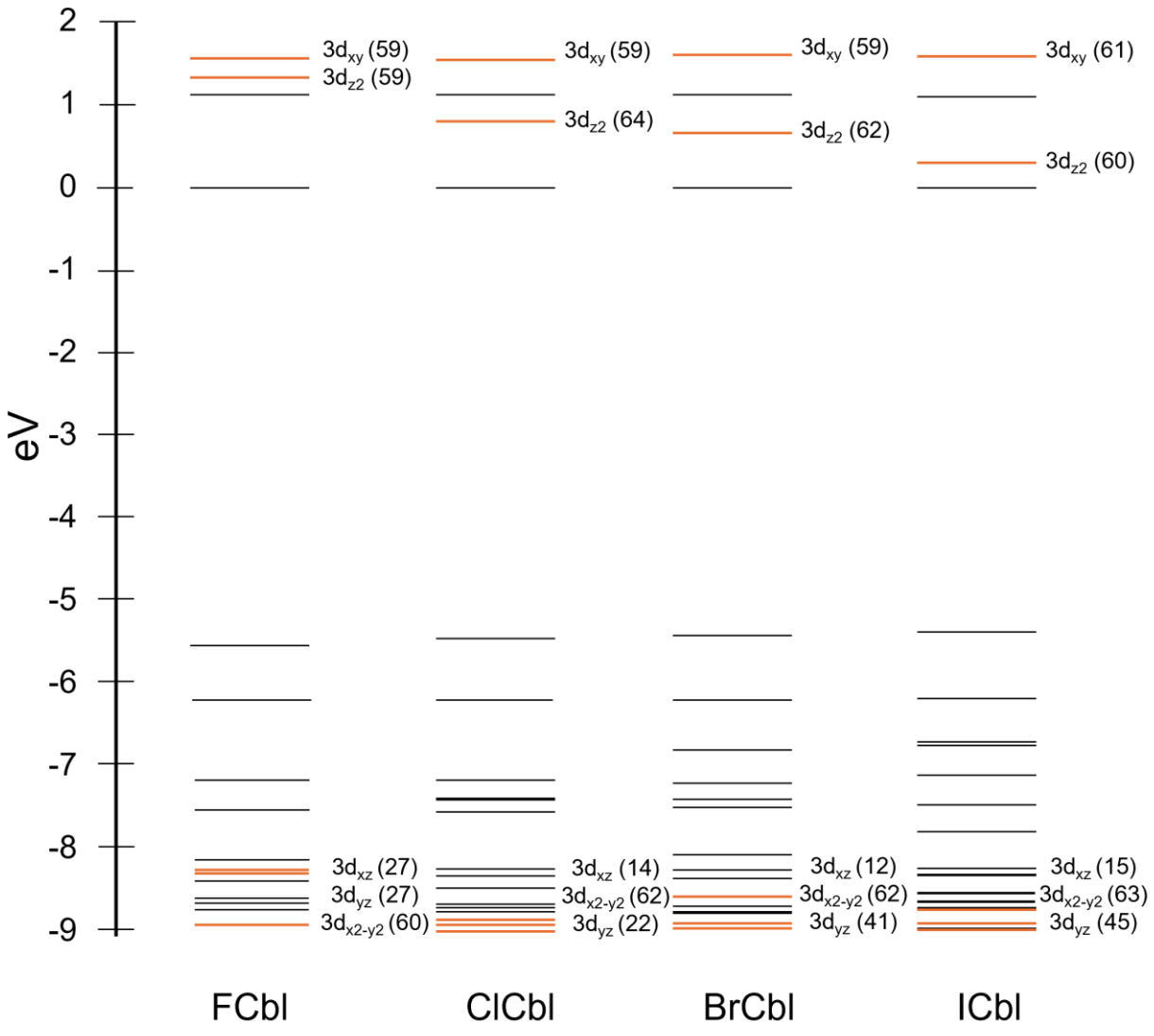

Figure S3.1. To-scale MO diagrams covering the d orbitals. The LUMO was shifted to 0 eV for all species to facilitate a comparison of the relative MO energies.

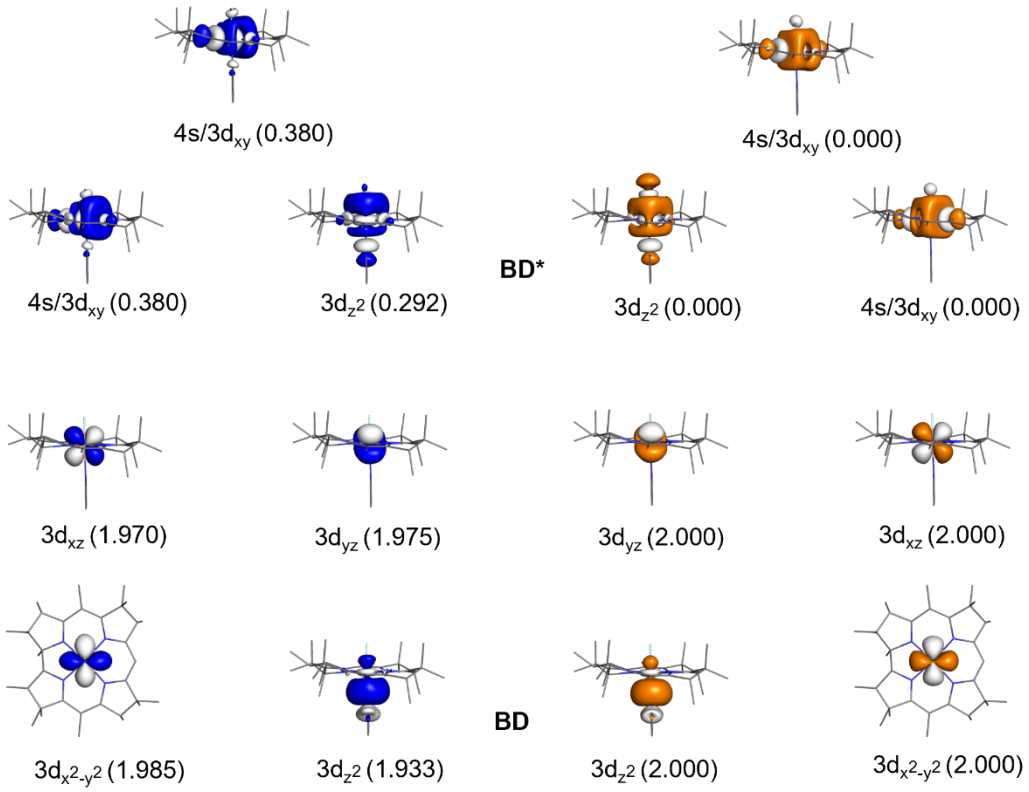

Figure S3.2. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of FCbl.

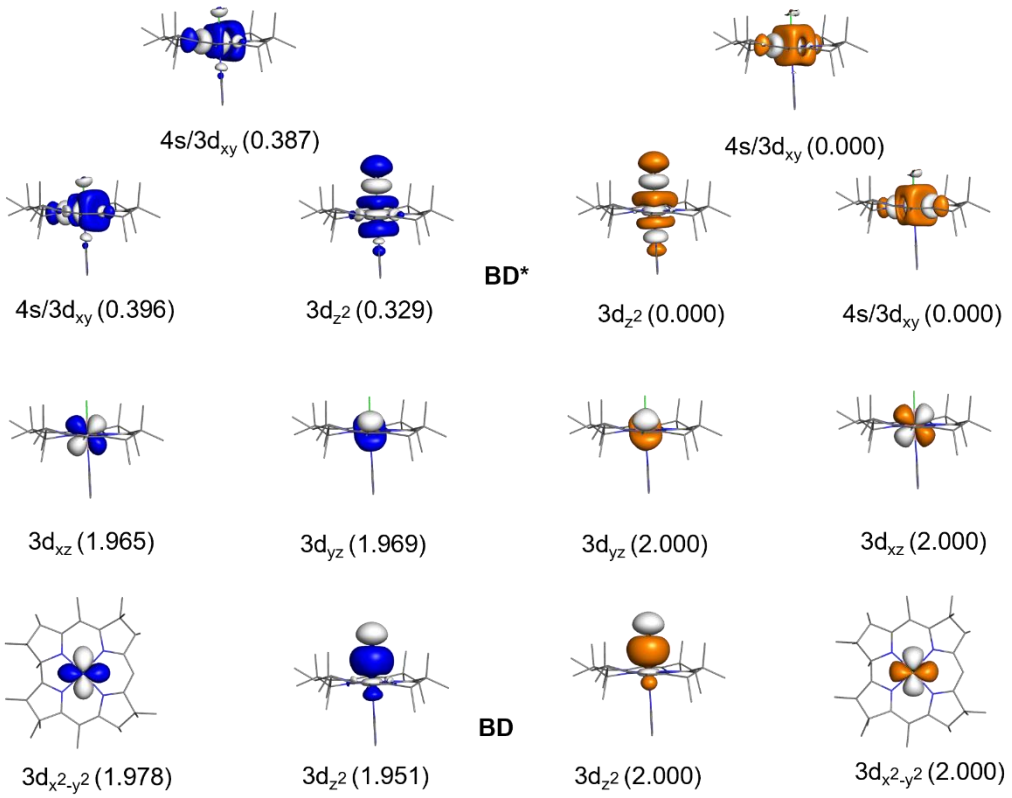

Figure S3.3. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of ClCbl.

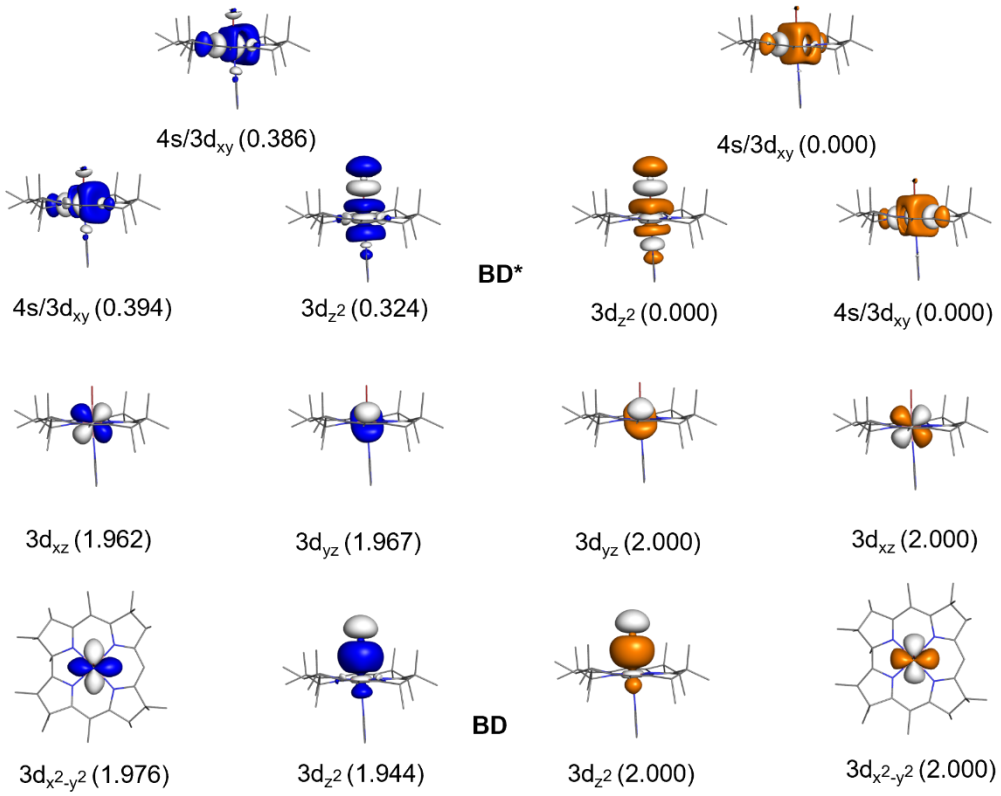

Figure S3.4. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of BrCbl.

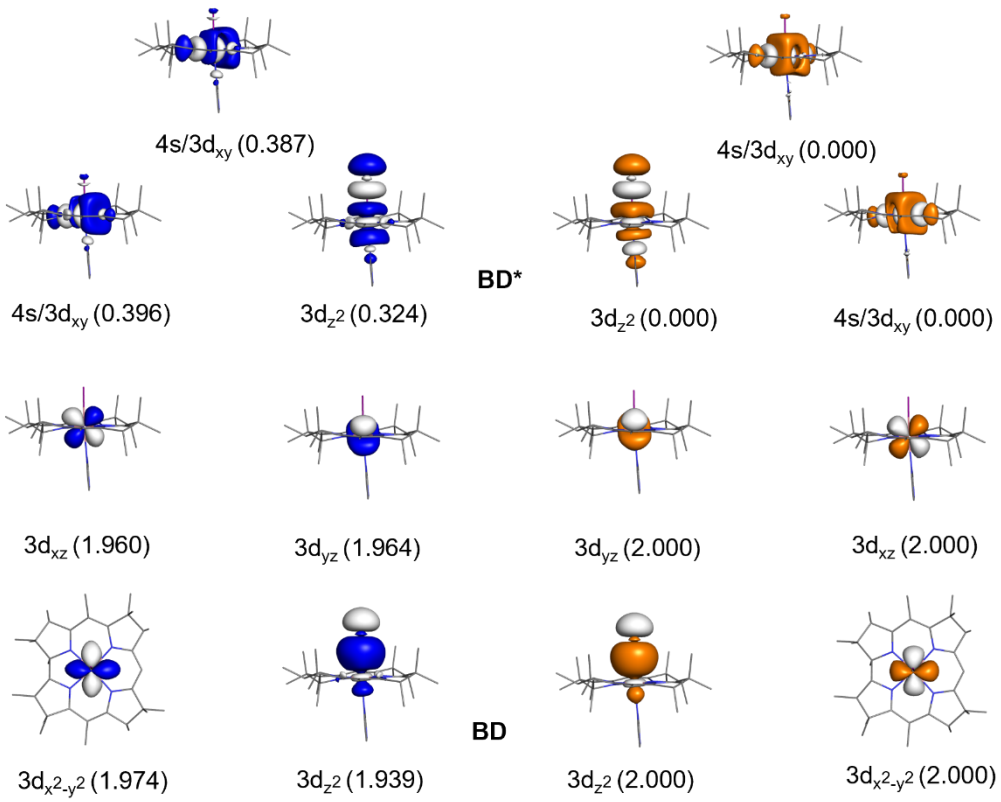

Figure S3.5. Isosurface plots and occupancies of the metal-based NBOs (blue and white) and corresponding NLMOs (orange and white) of ICbl.

Table S3.1. Computed shifts (cm^{-1}) for the two low energy d→d transitions and the αβ transition from methanol to water as predicted by the SMD solvation model

Species	$3d_{yz} \rightarrow 3d_{z2}$	$3d_{xz} \rightarrow 3d_{z2}$	αβ
FCbl	+0	+1	-2
ClCbl	+4	+5	-2
BrCbl	+4	+4	-3
ICbl	+22	+17	-9

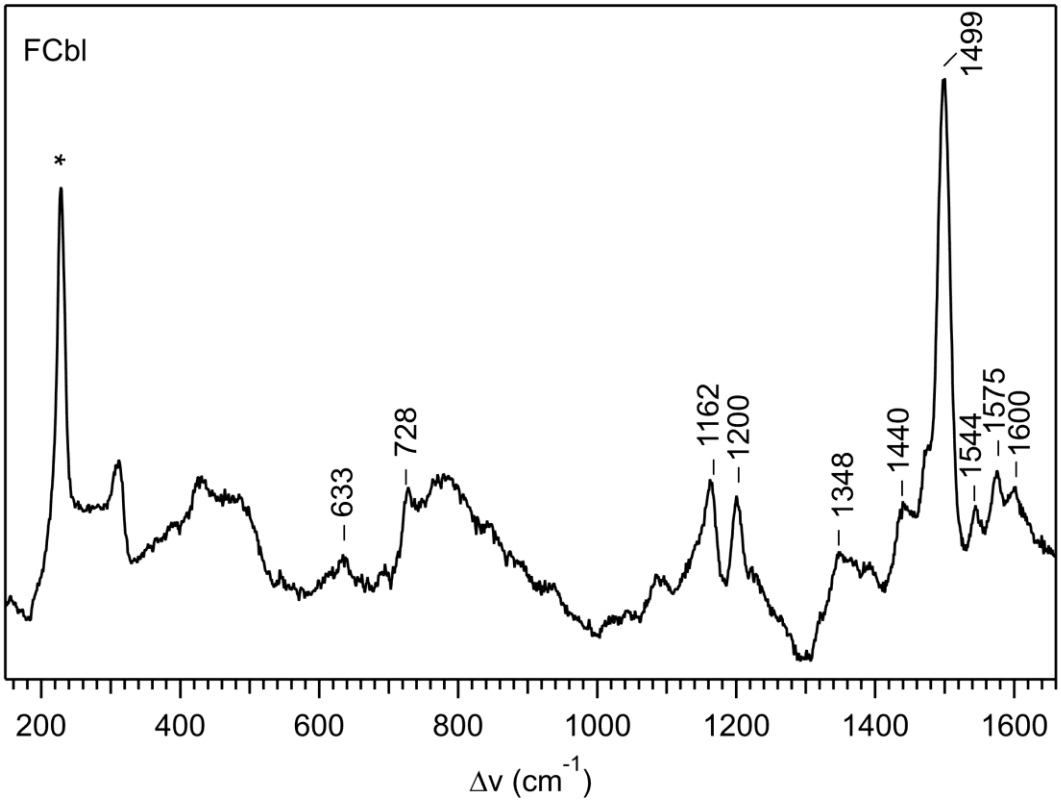

Figure S3.6. rR spectrum at 77 K of FCbl collected with λ_{ex} = 514.5 nm. The ice peak at 228 cm^{-1} is marked with an asterisk.

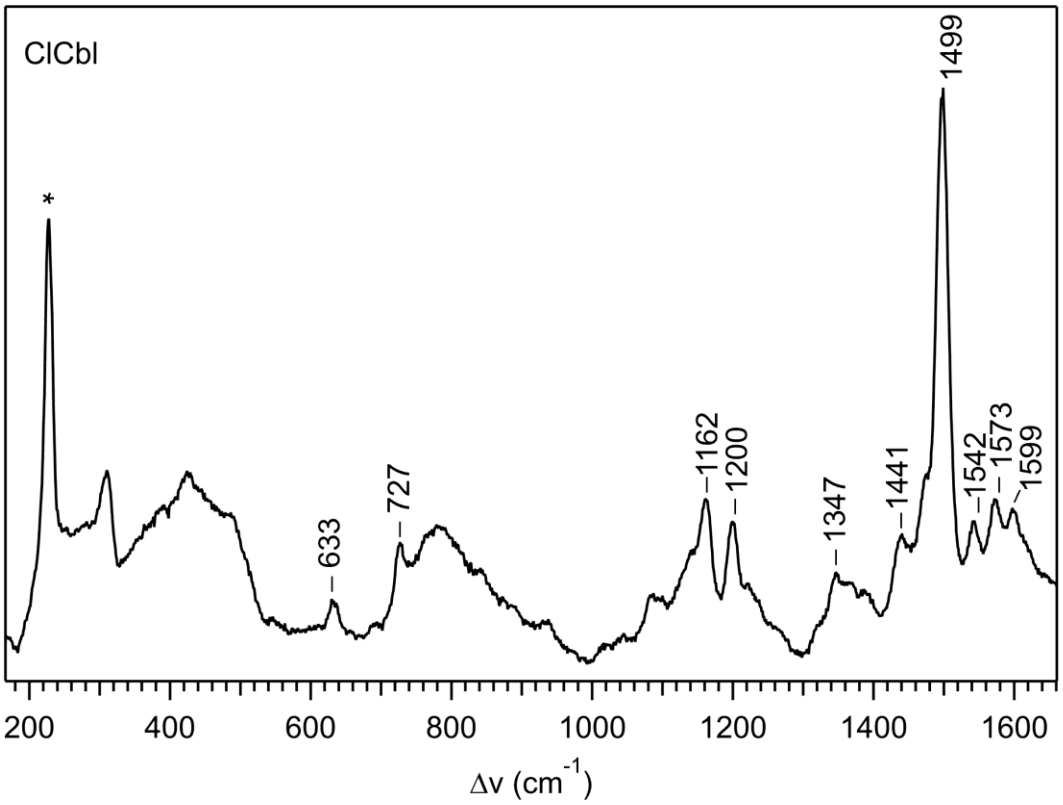

Figure S3.7. rR spectrum at 77 K of ClCbl collected with λ_{ex} = 514.5 nm. The ice peak at 228 cm^{-1} is marked with an asterisk.

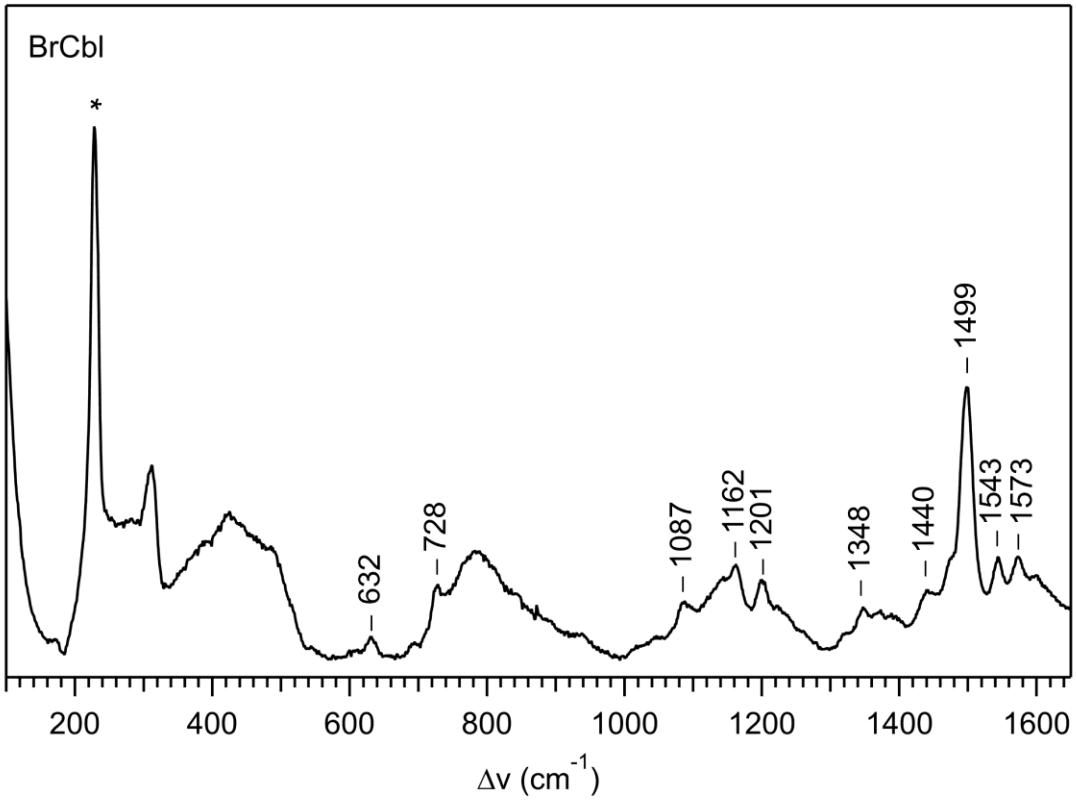

Figure S3.8. rR spectrum at 77 K of BrCbl collected with λ_{ex} = 514.5 nm. The ice peak at 228 cm^{-1} is marked with an asterisk.

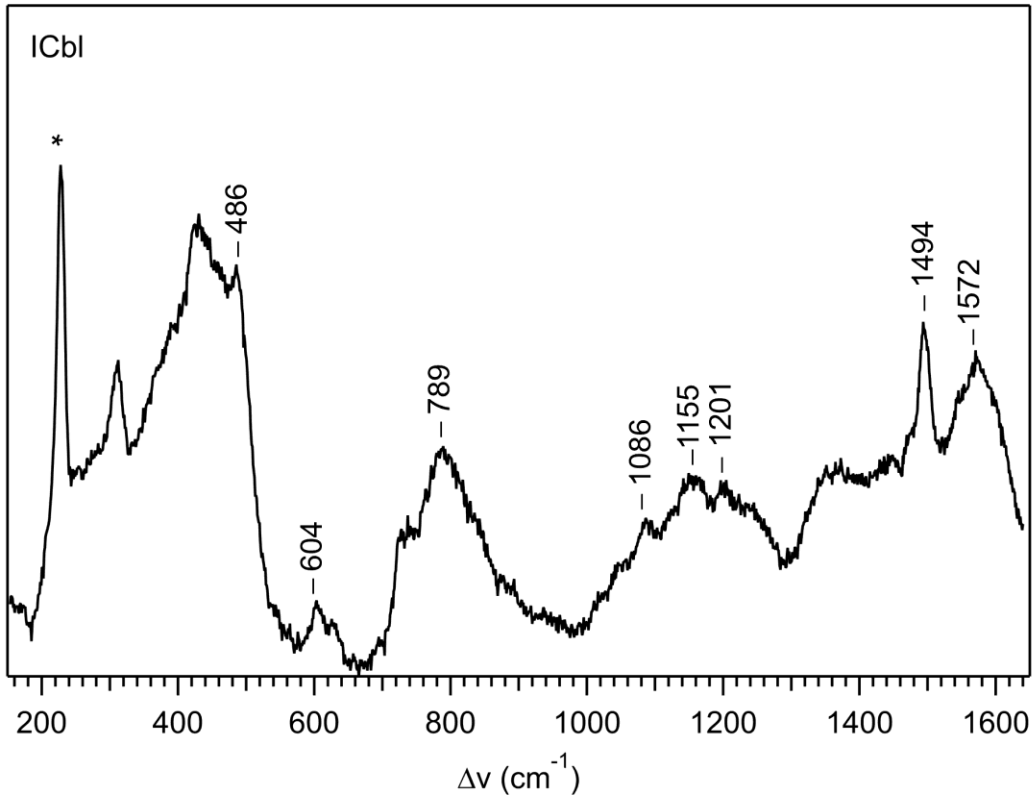

Figure S3.9. rR spectrum at 77 K of BrCbl collected with λ_{ex} = 514.5 nm. The ice peak at 228 cm^{-1} is marked with an asterisk.

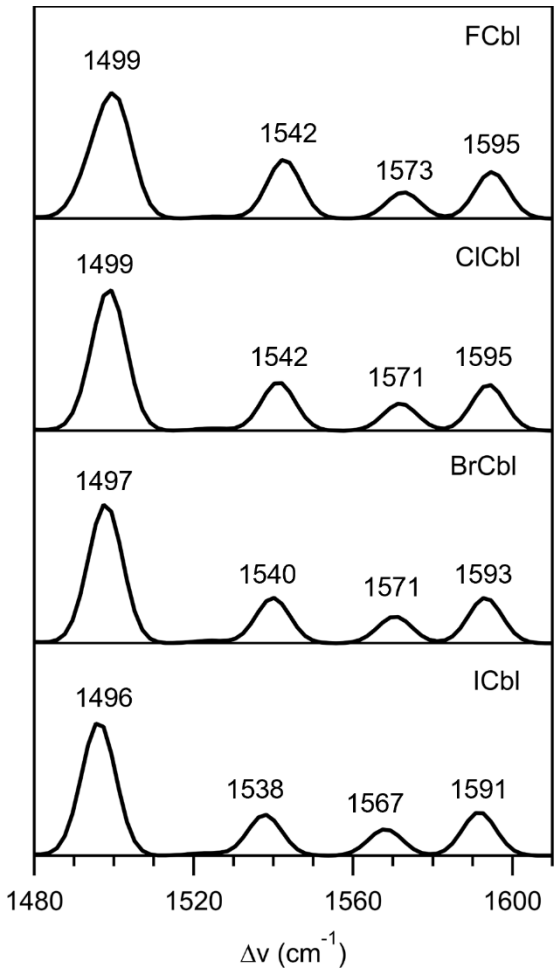

Figure S3.10. High energy region of the computed Raman spectra for XCbl species.

Table S3.2. Cartesian coordinates of FCbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.803293	0.661741	-0.060755
C	-3.669748	1.772931	0.650433
C	-2.846741	3.067835	0.327829
C	-1.429320	2.529422	0.230560
C	-0.247161	3.343945	0.368747
C	1.026056	2.795380	0.257706
C	2.357804	3.542864	0.484915
C	3.371913	2.539749	-0.140372
C	2.626294	1.237927	-0.014746
C	3.264040	0.000000	0.000000
C	2.635414	-1.242037	0.052507
C	3.373536	-2.562559	0.220709
C	2.326114	-3.532695	-0.398286
C	1.022890	-2.788173	-0.142846
C	-0.231421	-3.380368	-0.115199
C	-1.408125	-2.583035	0.146081
C	-2.778991	-3.112492	0.611331
C	-3.658843	-1.820065	0.492981
C	-2.638159	-0.691553	0.668424
N	-1.429188	1.238539	0.013582
N	1.292679	1.449979	0.048247
N	1.300679	-1.436066	-0.010985
N	-1.386400	-1.274692	0.152970
C	-3.252487	3.892047	-0.905391
H	-2.580797	4.762893	-1.032443
H	-4.281406	4.284955	-0.795724
H	-3.211788	3.304361	-1.842493
C	-3.190341	0.424942	-1.529658
H	-2.547854	-0.352425	-1.982703
H	-3.081237	1.345807	-2.130128
H	-4.240583	0.089662	-1.612978
C	-3.641374	1.601774	2.182722
H	-4.119705	2.479601	2.661102
H	-2.604524	1.524914	2.568256
H	-4.199864	0.705797	2.512174
C	-5.125229	1.855921	0.179639
H	-5.225755	2.033414	-0.906560
H	-5.644405	2.684206	0.702498
H	-5.675922	0.926247	0.422097
H	-2.919495	3.742162	1.205913
H	4.361493	0.011202	-0.000986
C	-0.438140	4.808094	0.710062

H	0.269628	5.141894	1.489548
H	-1.448671	4.999758	1.109731
H	-0.303361	5.479389	-0.162761
C	2.601223	3.587778	2.013892
H	2.586846	2.569795	2.451630
H	1.826495	4.185785	2.531296
H	3.587078	4.048997	2.225302
C	2.496197	4.947317	-0.121999
H	1.984179	5.720050	0.475724
H	2.110045	5.004128	-1.158159
H	3.568566	5.227054	-0.146107
H	4.310002	2.517282	0.451244
C	3.727736	2.774105	-1.619503
H	4.276274	3.724043	-1.763831
H	2.816928	2.798897	-2.252375
H	4.370501	1.952371	-1.991208
C	3.490423	-2.837297	1.740058
H	3.972257	-3.821269	1.910221
H	2.493893	-2.843074	2.224293
H	4.104120	-2.059267	2.234600
C	4.761603	-2.600517	-0.423832
H	4.749987	-2.274659	-1.481399
H	5.176634	-3.627425	-0.381504
H	5.468041	-1.947250	0.125650
H	2.346291	-4.501595	0.138882
C	2.506743	-3.787514	-1.904347
H	3.453050	-4.323256	-2.111109
H	2.519141	-2.835559	-2.473104
H	1.679106	-4.407379	-2.300851
C	-0.361404	-4.874280	-0.310176
H	-0.887927	-5.358211	0.535511
H	0.622367	-5.365380	-0.392388
H	-0.930813	-5.127444	-1.227285
C	-2.613321	-3.573091	2.076857
H	-3.576263	-3.942093	2.480739
H	-2.249118	-2.753244	2.727239
H	-1.882864	-4.401506	2.149499
C	-3.406343	-4.224297	-0.245398
H	-2.976025	-5.222208	-0.050948
H	-3.314714	-4.005975	-1.328410
H	-4.488101	-4.291504	-0.009922
C	-4.872465	-1.777089	1.415493
H	-5.525616	-0.915320	1.180439
H	-4.583671	-1.697022	2.481712

H	-5.492758	-2.688651	1.301008
H	-4.012285	-1.796023	-0.559198
H	-2.451591	-0.485961	1.740705
F	0.018124	0.031425	1.853405
C	-0.019896	-0.570328	-4.118824
C	-0.032118	-1.050498	-2.823956
N	0.048095	0.803894	-4.008237
C	0.075788	1.133614	-2.689552
H	-0.050553	-1.081741	-5.085330
H	0.080255	1.466479	-4.784349
H	-0.083205	-2.079111	-2.459956
H	0.118802	2.153971	-2.298791
N	0.029192	0.022348	-1.955113

Table S3.3. Cartesian coordinates of ClCbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.800555	0.684523	-0.224200
C	-3.704441	1.800648	0.433250
C	-2.845178	3.088531	0.186795
C	-1.430808	2.537177	0.175279
C	-0.251720	3.340898	0.384016
C	1.021767	2.791993	0.289096
C	2.351128	3.526884	0.558955
C	3.369721	2.542942	-0.090275
C	2.624039	1.238489	-0.003239
C	3.260419	0.000000	0.000000
C	2.631637	-1.241911	0.057714
C	3.369137	-2.557202	0.259367
C	2.314237	-3.549277	-0.308392
C	1.012680	-2.789956	-0.088989
C	-0.245041	-3.376505	-0.067017
C	-1.432465	-2.571889	0.104098
C	-2.838953	-3.091257	0.463822
C	-3.701537	-1.802749	0.235114
C	-2.700771	-0.675834	0.500722
N	-1.427261	1.247216	-0.047305
N	1.290158	1.450139	0.052800
N	1.297083	-1.436525	-0.002062
N	-1.408704	-1.262016	0.094549
C	-3.155023	3.936819	-1.058256
H	-2.464522	4.799678	-1.123499
H	-4.183808	4.342368	-1.013441
H	-3.057052	3.362380	-1.999609
C	-3.092543	0.463927	-1.716884
H	-2.438892	-0.323652	-2.133811

H	-2.929166	1.386713	-2.302271
H	-4.141460	0.148824	-1.867725
C	-3.799526	1.612601	1.960606
H	-4.300938	2.492818	2.410042
H	-2.799232	1.512955	2.429318
H	-4.397882	0.722530	2.231070
C	-5.119508	1.912186	-0.144828
H	-5.137914	2.099481	-1.233895
H	-5.660340	2.746472	0.345460
H	-5.704639	0.992418	0.049519
H	-2.969081	3.751189	1.068007
H	4.357888	0.010557	0.004908
C	-0.448443	4.793286	0.768079
H	0.235543	5.095764	1.581038
H	-1.470478	4.974872	1.142686
H	-0.283313	5.493874	-0.075859
C	2.571310	3.511964	2.092182
H	2.530013	2.479238	2.492450
H	1.797426	4.103981	2.617586
H	3.561076	3.948430	2.335451
C	2.502741	4.951365	0.005207
H	1.984130	5.704134	0.622533
H	2.132310	5.045671	-1.033953
H	3.576297	5.227689	0.006856
H	4.303827	2.503511	0.506707
C	3.735149	2.821871	-1.559186
H	4.284090	3.775970	-1.670510
H	2.828831	2.865733	-2.197312
H	4.380910	2.012024	-1.951272
C	3.507666	-2.774107	1.786345
H	3.978747	-3.757719	1.986470
H	2.520802	-2.744695	2.290013
H	4.141279	-1.986662	2.239130
C	4.748154	-2.618742	-0.402928
H	4.721395	-2.336144	-1.472640
H	5.166455	-3.642222	-0.325035
H	5.461142	-1.943042	0.109639
H	2.331468	-4.489653	0.277286
C	2.487734	-3.883909	-1.799764
H	3.429722	-4.436164	-1.980913
H	2.504979	-2.963155	-2.417566
H	1.654099	-4.517788	-2.159530
C	-0.368397	-4.878947	-0.191775
H	-0.960143	-5.313331	0.636958
H	0.616095	-5.374713	-0.169439
H	-0.866637	-5.182812	-1.134536
C	-2.803354	-3.510245	1.950746
H	-3.800605	-3.861685	2.280214
H	-2.487856	-2.674898	2.606978
H	-2.089141	-4.340803	2.109764
C	-3.394355	-4.224266	-0.414424

H	-2.990139	-5.218512	-0.156224
H	-3.206716	-4.036249	-1.490788
H	-4.493004	-4.280870	-0.273659
C	-5.001003	-1.740806	1.030566
H	-5.623863	-0.883405	0.712422
H	-4.818230	-1.637728	2.117908
H	-5.609600	-2.654007	0.874775
H	-3.945094	-1.794668	-0.848185
H	-2.615479	-0.481952	1.588227
Cl	0.022763	0.028363	2.250644
C	0.114495	-0.597190	-4.131724
C	0.052904	-1.066080	-2.833881
N	0.186586	0.777655	-4.030245
C	0.168915	1.118124	-2.714031
H	0.116314	-1.116307	-5.094687
H	0.251466	1.433028	-4.810577
H	-0.017698	-2.091210	-2.463175
H	0.203988	2.141326	-2.330202
N	0.088780	0.014001	-1.973733

Table S3.4. Cartesian coordinates of BrCbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.796048	0.689550	-0.262862
C	-3.710462	1.806168	0.380055
C	-2.842974	3.092699	0.155590
C	-1.430092	2.538710	0.165298
C	-0.252398	3.339922	0.389667
C	1.021270	2.791294	0.297168
C	2.350406	3.526374	0.565978
C	3.368319	2.543565	-0.086694
C	2.623776	1.238583	0.000055
C	3.259789	0.000000	0.000000
C	2.631074	-1.241869	0.059020
C	3.369419	-2.556103	0.262616
C	2.311215	-3.551897	-0.291313
C	1.010630	-2.789850	-0.076442
C	-0.247695	-3.375562	-0.055622
C	-1.436810	-2.569857	0.093145
C	-2.851509	-3.087174	0.422092
C	-3.707239	-1.799416	0.166581
C	-2.713625	-0.673233	0.458335
N	-1.424376	1.248249	-0.057177
N	1.289536	1.448828	0.060512
N	1.295684	-1.435952	0.003658
N	-1.411059	-1.259258	0.082898
C	-3.128635	3.949063	-1.089871
H	-2.434983	4.810497	-1.136857
H	-4.157060	4.356885	-1.061294
H	-3.014694	3.379980	-2.032736
C	-3.061515	0.475120	-1.761333

H	-2.406294	-0.315963	-2.168523
H	-2.881512	1.398404	-2.341056
H	-4.109641	0.167504	-1.931721
C	-3.839827	1.612947	1.904110
H	-4.349197	2.492667	2.345471
H	-2.850938	1.509084	2.395918
H	-4.446298	0.723218	2.157284
C	-5.112948	1.924921	-0.226808
H	-5.109016	2.116566	-1.315268
H	-5.660280	2.759500	0.255690
H	-5.705942	1.007091	-0.047722
H	-2.979897	3.751092	1.038127
H	4.357275	0.010205	0.002377
C	-0.451421	4.789563	0.782396
H	0.233626	5.088590	1.595633
H	-1.473235	4.965856	1.160463
H	-0.289964	5.495708	-0.057625
C	2.575206	3.511005	2.098405
H	2.538634	2.477914	2.498435
H	1.800952	4.099833	2.626820
H	3.564333	3.950600	2.338524
C	2.498714	4.951385	0.012633
H	1.980003	5.702896	0.631442
H	2.125881	5.045516	-1.025702
H	3.571776	5.229484	0.012397
H	4.303382	2.503697	0.508788
C	3.731893	2.825073	-1.555590
H	4.280364	3.779579	-1.665725
H	2.825060	2.869814	-2.192837
H	4.377544	2.016188	-1.949829
C	3.519973	-2.764149	1.789693
H	3.993805	-3.746143	1.991370
H	2.537232	-2.733390	2.301591
H	4.156114	-1.973441	2.233133
C	4.743674	-2.620589	-0.409666
H	4.708851	-2.343278	-1.480533
H	5.163092	-3.643451	-0.329909
H	5.460299	-1.942213	0.094187
H	2.329307	-4.486064	0.304176
C	2.478809	-3.904050	-1.779436
H	3.419250	-4.459963	-1.957302
H	2.495444	-2.990635	-2.407997
H	1.642820	-4.540710	-2.128519
C	-0.368996	-4.879587	-0.163328
H	-0.991488	-5.300235	0.649338
H	0.613204	-5.376007	-0.095790
H	-0.832140	-5.198206	-1.119277
C	-2.853134	-3.498185	1.911639
H	-3.859356	-3.844274	2.218771
H	-2.550493	-2.660138	2.570643
H	-2.146004	-4.330216	2.093126

C	-3.386180	-4.223631	-0.464621
H	-2.989819	-5.217210	-0.192131
H	-3.171340	-4.040906	-1.536816
H	-4.488061	-4.278537	-0.351261
C	-5.028183	-1.734139	0.925487
H	-5.641294	-0.877825	0.586167
H	-4.875753	-1.626351	2.017028
H	-5.632546	-2.647824	0.756768
H	-3.920933	-1.793384	-0.923055
H	-2.656468	-0.482976	1.548669
Br	0.023416	0.018595	2.407805
C	0.154089	-0.597396	-4.137916
C	0.079328	-1.065419	-2.840342
N	0.226005	0.777560	-4.036444
C	0.195583	1.118456	-2.720443
H	0.165128	-1.116857	-5.100662
H	0.299428	1.432345	-4.816566
H	0.003978	-2.090297	-2.469703
H	0.227670	2.141829	-2.336697
N	0.107233	0.015023	-1.980834

Table S3.5. Cartesian coordinates of ICbl model in Å

C	-2.788700	0.697661	-0.323130
C	-3.717470	1.815530	0.297467
C	-2.838486	3.099723	0.105241
C	-1.428545	2.541116	0.147058
C	-0.253219	3.338854	0.395414
C	1.020976	2.791019	0.307811
C	2.349499	3.526803	0.576200
C	3.367855	2.544465	-0.077708
C	2.623950	1.239004	0.006599
C	3.259116	0.000000	0.000000
C	2.630667	-1.242043	0.057087
C	3.370106	-2.555039	0.262741
C	2.306201	-3.556274	-0.269498
C	1.007494	-2.789559	-0.060839
C	-0.252109	-3.373620	-0.039389
C	-1.443408	-2.566211	0.075454
C	-2.869425	-3.080411	0.357285
C	-3.713017	-1.792877	0.063503
C	-2.730356	-0.668185	0.392973
N	-1.419876	1.250116	-0.075633
N	1.289263	1.447460	0.072170
N	1.294115	-1.435250	0.008190
N	-1.414410	-1.254586	0.062745
Co	0.000000	0.000000	0.000000
C	-3.088434	3.967132	-1.140545
H	-2.391379	4.826720	-1.160960
H	-4.116065	4.377809	-1.136373

H	-2.950115	3.405343	-2.084571
C	-3.013602	0.491499	-1.829189
H	-2.354668	-0.303730	-2.221282
H	-2.808248	1.415529	-2.399227
H	-4.059390	0.194601	-2.030050
C	-3.896274	1.615734	1.815447
H	-4.417875	2.494331	2.244537
H	-2.924152	1.508116	2.339351
H	-4.511782	0.725635	2.044728
C	-5.100388	1.944468	-0.351076
H	-5.063060	2.141106	-1.437998
H	-5.657166	2.780060	0.118667
H	-5.704126	1.029590	-0.193996
H	-2.995137	3.752612	0.988687
H	4.356664	0.009661	0.000071
C	-0.456035	4.785103	0.798220
H	0.227739	5.079703	1.614060
H	-1.478840	4.955287	1.176819
H	-0.295959	5.497772	-0.036521
C	2.577819	3.513574	2.107894
H	2.541858	2.481031	2.509731
H	1.804875	4.102781	2.637850
H	3.567648	3.953197	2.344919
C	2.494339	4.951746	0.021750
H	1.973369	5.702121	0.640102
H	2.121299	5.044123	-1.016692
H	3.566598	5.232743	0.021569
H	4.302198	2.503607	0.518943
C	3.734369	2.828988	-1.545285
H	4.283492	3.783483	-1.652200
H	2.829148	2.875318	-2.184536
H	4.380578	2.020675	-1.939792
C	3.538219	-2.748565	1.789900
H	4.007863	-3.731620	1.996219
H	2.561647	-2.705544	2.312993
H	4.184484	-1.957361	2.217578
C	4.736622	-2.625054	-0.425123
H	4.688513	-2.356920	-1.497779
H	5.157894	-3.646828	-0.341626
H	5.458746	-1.941895	0.064160
H	2.325434	-4.480547	0.341306
C	2.465020	-3.936051	-1.751886
H	3.403410	-4.497096	-1.924196
H	2.480144	-3.034540	-2.397348
H	1.626030	-4.577438	-2.084458
C	-0.370737	-4.879710	-0.118158
H	-1.042428	-5.277982	0.665232
H	0.604357	-5.375061	0.023639
H	-0.774882	-5.223799	-1.092306
C	-2.929297	-3.482220	1.847974
H	-3.948400	-3.820909	2.118886

H	-2.647862	-2.641276	2.512919
H	-2.234289	-4.316528	2.062895
C	-3.372463	-4.220381	-0.543670
H	-2.988446	-5.213400	-0.252135
H	-3.116429	-4.043119	-1.607714
H	-4.477939	-4.273154	-0.472106
C	-5.064070	-1.723217	0.766672
H	-5.661315	-0.867656	0.398299
H	-4.958031	-1.611190	1.863300
H	-5.661762	-2.637026	0.576353
H	-3.882012	-1.789307	-1.033995
H	-2.713800	-0.482682	1.486157
I	0.020931	-0.005324	2.605182
C	0.221068	-0.595370	-4.155658
C	0.122038	-1.062249	-2.859061
N	0.295976	0.779462	-4.053140
C	0.244118	1.119818	-2.737270
H	0.246700	-1.114606	-5.118313
H	0.385490	1.433552	-4.832241
H	0.035510	-2.086775	-2.489421
H	0.272691	2.143319	-2.353254
N	0.139079	0.017673	-1.999233

Appendix

Chapter 3

Spectroscopic and Computational Analysis of Hydridocobalamin (HCbl)

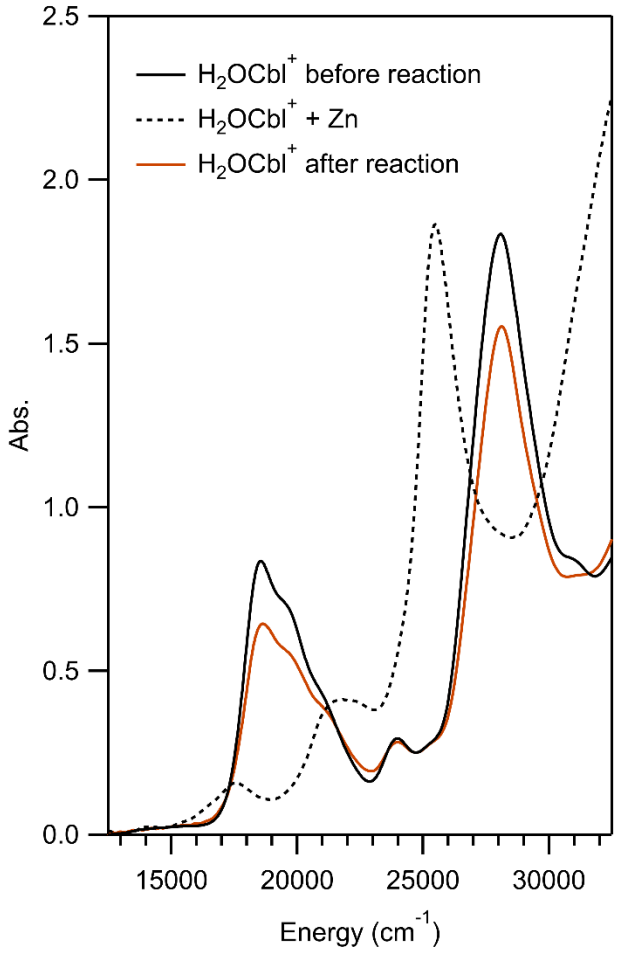

Figure S4.1. Experimental RT Abs spectra of H_2OCbl^+ in HOAc (black trace) before addition of Zn dust to the cuvette, the Abs spectrum 20 minutes after Zn dust was added (dotted trace), and the Abs spectrum after the reaction solution was exposed to air for two hours (red trace).

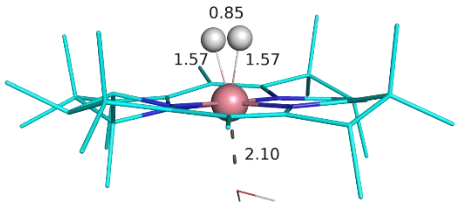

Figure S4.2. Geometry optimized base-off H_2Cbl with relevant bond distances shown in Å. Only the H atoms located on the axial ligands are shown for clarity.

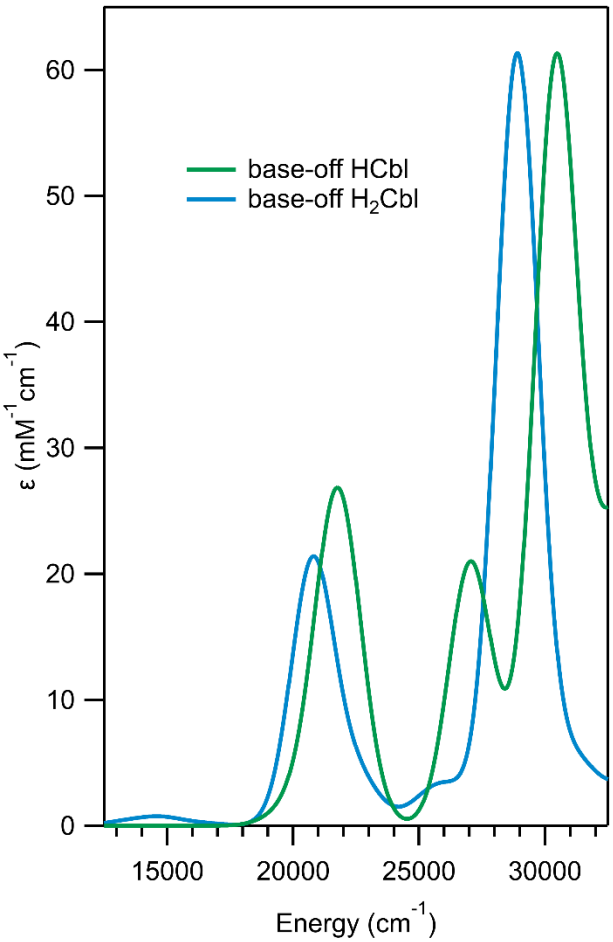

Figure S4.3. TD-DFT computed Abs spectra for base-off HCbl (green trace) and base-off H_2Cbl (light blue trace). The TD-DFT computed spectra were uniformly red-shifted by 3000 cm^{-1}like the TD-DFT computed Abs spectra presented in the main text.

Table S4.1. Cartesian coordinates of DFT-optimized base-off MeCbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.743224	0.697862	-0.504862
C	-3.711948	1.825424	0.025121
C	-2.815594	3.105680	-0.115258
C	-1.414792	2.541658	0.029083
C	-0.250315	3.336110	0.325742
C	1.024004	2.791271	0.241956
C	2.352216	3.532525	0.496683
C	3.367516	2.545168	-0.155419
C	2.628746	1.239939	-0.039755
C	3.264969	0.000000	0.000000
C	2.634520	-1.241188	0.054721
C	3.361947	-2.558427	0.274251
C	2.311210	-3.549221	-0.300934
C	1.008929	-2.785233	-0.107851
C	-0.250113	-3.367357	-0.105886
C	-1.441352	-2.558766	-0.022515
C	-2.882724	-3.075430	0.152987
C	-3.702815	-1.779361	-0.165859
C	-2.740383	-0.658638	0.231603
N	-1.387865	1.243213	-0.169893
N	1.285997	1.439825	0.008235
N	1.289272	-1.423240	-0.029780
N	-1.400172	-1.243084	-0.002194
C	-2.978192	3.964048	-1.381616
H	-2.280966	4.823430	-1.361260
H	-4.003301	4.375359	-1.449562
H	-2.774790	3.396068	-2.309351
C	-2.829161	0.477862	-2.022500
H	-2.079842	-0.266615	-2.347278
H	-2.624442	1.412870	-2.574115
H	-3.834601	0.124316	-2.316716
C	-4.015831	1.653016	1.526565
H	-4.580523	2.533120	1.893293
H	-3.090251	1.574750	2.132652
H	-4.636456	0.759516	1.727357
C	-5.038218	1.939138	-0.735197
H	-4.908564	2.112495	-1.818500
H	-5.634584	2.781636	-0.331159

H	-5.648712	1.023690	-0.611062
H	-3.030833	3.768265	0.749010
H	4.362768	0.010038	0.017332
C	-0.467148	4.778573	0.731372
H	0.235537	5.089367	1.523914
H	-1.483849	4.928055	1.136433
H	-0.346921	5.490010	-0.111287
C	2.602220	3.544608	2.024545
H	2.598632	2.517543	2.442240
H	1.829655	4.129417	2.560153
H	3.588177	4.002273	2.243401
C	2.484447	4.951672	-0.077073
H	1.966257	5.707223	0.537642
H	2.098216	5.028555	-1.111380
H	3.554906	5.238781	-0.093347
H	4.310799	2.518938	0.427974
C	3.711612	2.814023	-1.631641
H	4.280385	3.755230	-1.753116
H	2.796489	2.886122	-2.254841
H	4.334195	1.991515	-2.035094
C	3.474440	-2.769938	1.804607
H	3.943336	-3.751474	2.019664
H	2.475606	-2.750251	2.285123
H	4.095917	-1.979261	2.268918
C	4.751631	-2.631157	-0.364583
H	4.741694	-2.352644	-1.435531
H	5.163738	-3.656304	-0.277635
H	5.460562	-1.956763	0.155466
H	2.317802	-4.485335	0.291786
C	2.502753	-3.903169	-1.785702
H	3.447398	-4.456434	-1.948564
H	2.529464	-2.993113	-2.420494
H	1.673124	-4.538628	-2.149861
C	-0.366307	-4.872223	-0.203173
H	-1.076135	-5.278764	0.541118
H	0.600912	-5.370541	-0.022364
H	-0.723214	-5.200866	-1.200463
C	-3.052667	-3.522292	1.622344
H	-4.082566	-3.887661	1.802454
H	-2.849805	-2.696515	2.333456

H	-2.359842	-4.349242	1.869647
C	-3.325173	-4.189892	-0.810548
H	-2.969667	-5.192594	-0.515757
H	-2.989767	-3.987639	-1.847319
H	-4.433172	-4.236126	-0.821400
C	-5.093064	-1.715886	0.457411
H	-5.664315	-0.849929	0.072550
H	-5.052801	-1.627176	1.560884
H	-5.682984	-2.621806	0.213173
H	-3.805725	-1.756453	-1.270489
H	-2.816020	-0.465395	1.321692
C	0.007258	0.042512	1.963881
H	1.058352	0.064556	2.279994
H	-0.489989	-0.859994	2.341520
H	-0.518085	0.951763	2.283421
O	0.329485	-0.087980	-2.463806
H	0.757540	0.779996	-2.596889
H	1.082773	-0.706485	-2.390100

Table S4.2. Cartesian coordinates of DFT-optimized base-off of base-off HCbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.757311	0.686399	-0.401260
C	-3.712186	1.815305	0.148946
C	-2.825057	3.096671	-0.031597
C	-1.417938	2.539562	0.077473
C	-0.250184	3.338404	0.344079
C	1.023278	2.790997	0.262550
C	2.352358	3.535301	0.503958
C	3.365159	2.543559	-0.143565
C	2.628472	1.238208	-0.019313
C	3.268020	0.000000	0.000000
C	2.638687	-1.241428	0.054194
C	3.370387	-2.561085	0.242007
C	2.310471	-3.543871	-0.329925
C	1.011083	-2.784529	-0.102860
C	-0.245791	-3.369648	-0.083038
C	-1.434913	-2.564096	0.045952
C	-2.867696	-3.082809	0.275575

C	-3.702415	-1.787721	-0.008824
C	-2.727619	-0.663982	0.347380
N	-1.392913	1.239348	-0.116730
N	1.285630	1.436415	0.041509
N	1.292154	-1.422301	-0.011464
N	-1.394999	-1.248720	0.067676
C	-3.028485	3.940613	-1.301787
H	-2.333945	4.802381	-1.312844
H	-4.056631	4.347962	-1.342906
H	-2.852028	3.363052	-2.229070
C	-2.888861	0.451764	-1.913934
H	-2.148569	-0.295391	-2.254206
H	-2.703312	1.382374	-2.479654
H	-3.901562	0.093564	-2.175649
C	-3.966113	1.656643	1.661456
H	-4.525574	2.536671	2.036352
H	-3.020698	1.592751	2.237403
H	-4.573372	0.761236	1.892427
C	-5.062419	1.915809	-0.569625
H	-4.968003	2.081384	-1.657795
H	-5.650922	2.757887	-0.153383
H	-5.662691	0.997467	-0.419223
H	-3.017604	3.767670	0.831472
H	4.365862	0.011644	0.001269
C	-0.462039	4.787783	0.727921
H	0.241511	5.108554	1.515836
H	-1.477787	4.947791	1.130925
H	-0.338776	5.485407	-0.125679
C	2.609193	3.562494	2.030687
H	2.610314	2.539415	2.457578
H	1.837331	4.150557	2.563847
H	3.594748	4.025412	2.240149
C	2.482672	4.949380	-0.083218
H	1.969951	5.711141	0.528293
H	2.090253	5.018040	-1.115765
H	3.553495	5.234377	-0.108696
H	4.310329	2.521458	0.437056
C	3.705536	2.801813	-1.622822
H	4.271653	3.743449	-1.752869
H	2.788731	2.865812	-2.244338

H	4.329289	1.977665	-2.020998
C	3.513448	-2.798041	1.765987
H	3.992310	-3.780279	1.954169
H	2.524495	-2.792998	2.266364
H	4.139529	-2.011914	2.231745
C	4.747996	-2.621749	-0.424172
H	4.718106	-2.325056	-1.489871
H	5.161758	-3.648045	-0.362244
H	5.466365	-1.955971	0.094001
H	2.327889	-4.488703	0.248483
C	2.476009	-3.875337	-1.822982
H	3.418277	-4.424642	-2.011073
H	2.490110	-2.955903	-2.444425
H	1.640473	-4.506094	-2.181810
C	-0.363756	-4.872828	-0.200544
H	-1.038878	-5.293323	0.568217
H	0.610967	-5.372691	-0.072681
H	-0.765400	-5.184371	-1.186093
C	-2.978820	-3.535380	1.748829
H	-4.000040	-3.904100	1.967415
H	-2.748892	-2.712267	2.454450
H	-2.275000	-4.361952	1.964535
C	-3.345659	-4.195453	-0.673163
H	-2.978408	-5.198337	-0.394019
H	-3.050638	-3.990926	-1.721695
H	-4.453231	-4.242332	-0.641332
C	-5.066159	-1.727948	0.671030
H	-5.655028	-0.862614	0.312036
H	-4.981535	-1.642015	1.772215
H	-5.663271	-2.634787	0.448620
H	-3.851604	-1.766011	-1.108207
H	-2.760920	-0.460122	1.437282
H	-0.042905	0.037517	1.436372
O	0.289737	-0.100386	-2.442514
H	0.615170	0.797036	-2.650908
H	1.103669	-0.640821	-2.407299

Table S4.3. Cartesian coordinates of DFT-optimized base-off H_2Cbl model in Å

Co	0.000000	0.000000	0.000000
C	-2.783964	0.690518	-0.454566
C	-3.723637	1.827745	0.115226
C	-2.831503	3.106607	-0.075283
C	-1.425104	2.552383	0.030978
C	-0.257021	3.347198	0.318996
C	1.022734	2.803978	0.244499
C	2.350463	3.544150	0.495699
C	3.371505	2.546567	-0.136679
C	2.634144	1.243109	-0.028597
C	3.265324	0.000000	0.000000
C	2.640502	-1.246207	0.048127
C	3.368614	-2.564900	0.236438
C	2.309982	-3.552209	-0.338687
C	1.005592	-2.801806	-0.123109
C	-0.253695	-3.380174	-0.113356
C	-1.447129	-2.570959	0.007116
C	-2.882508	-3.084165	0.236864
C	-3.720214	-1.787643	-0.047839
C	-2.745219	-0.657702	0.300281
N	-1.415083	1.256206	-0.194266
N	1.292257	1.455399	0.009265
N	1.298437	-1.439170	-0.039020
N	-1.422129	-1.258743	0.015655
C	-3.032038	3.949336	-1.347802
H	-2.338726	4.811490	-1.363111
H	-4.059318	4.358155	-1.378511
H	-2.866749	3.371113	-2.276555
C	-2.952200	0.453429	-1.962335
H	-2.264764	-0.333624	-2.321498
H	-2.753113	1.374861	-2.538622
H	-3.986111	0.137529	-2.191236
C	-3.958440	1.665311	1.630234
H	-4.512030	2.544897	2.013206
H	-3.011443	1.601925	2.205914
H	-4.568058	0.773125	1.866631
C	-5.082607	1.941728	-0.584591
H	-5.006053	2.115954	-1.672521

H	-5.656164	2.785596	-0.152782
H	-5.690339	1.029812	-0.428194
H	-3.011699	3.780552	0.787678
H	4.363028	0.007946	0.014582
C	-0.470663	4.793690	0.704827
H	0.231620	5.114118	1.492950
H	-1.487827	4.958199	1.100692
H	-0.340406	5.485509	-0.151987
C	2.586002	3.572872	2.027980
H	2.589117	2.552426	2.462321
H	1.812353	4.165446	2.552490
H	3.568879	4.037675	2.242311
C	2.486513	4.956250	-0.094343
H	1.971160	5.720657	0.510879
H	2.111116	5.024651	-1.132929
H	3.558031	5.237523	-0.104065
H	4.306418	2.517270	0.458858
C	3.738459	2.800123	-1.612302
H	4.305276	3.743084	-1.723153
H	2.836394	2.865472	-2.254045
H	4.377861	1.982240	-1.997464
C	3.500967	-2.790962	1.765732
H	3.980024	-3.772409	1.953820
H	2.511237	-2.792244	2.265792
H	4.129137	-2.007620	2.232441
C	4.750104	-2.623300	-0.421883
H	4.729610	-2.327974	-1.487935
H	5.158047	-3.651222	-0.357585
H	5.468818	-1.965079	0.104653
H	2.322437	-4.494111	0.243668
C	2.482690	-3.887507	-1.831194
H	3.426751	-4.437586	-2.004261
H	2.505313	-2.972218	-2.458462
H	1.653703	-4.524399	-2.193152
C	-0.375551	-4.881369	-0.230195
H	-1.048832	-5.298646	0.541165
H	0.597395	-5.385135	-0.109649
H	-0.784944	-5.185451	-1.214226
C	-2.979300	-3.528849	1.714751
H	-4.001627	-3.892163	1.934773

H	-2.748990	-2.705085	2.419789
H	-2.282248	-4.360210	1.931827
C	-3.359152	-4.200641	-0.707365
H	-2.999706	-5.203543	-0.420365
H	-3.068108	-4.004205	-1.758318
H	-4.466395	-4.241236	-0.673073
C	-5.079642	-1.721736	0.639541
H	-5.668274	-0.857421	0.278755
H	-4.991913	-1.636126	1.740179
H	-5.679470	-2.626277	0.418314
H	-3.873141	-1.768929	-1.146431
H	-2.780205	-0.453188	1.391595
H	-0.345737	0.344028	1.491855
H	0.245407	-0.268878	1.529356
O	0.137964	-0.101465	-2.089935
H	0.259851	0.803761	-2.447064
H	0.975130	-0.573138	-2.288829