Structural Studies of Metalloenzyme Complexes in Acetogenic Carbon Fixation

by

Yan Kung

B.A., Chemistry (2005)
Colby College

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
in Biological Chemistry

at the

Massachusetts Institute of Technology

June 2011

Signature of Autho

May 5, 2011

Certified by ___

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

Accepted by ___

Robert W. Field
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor JoAnne Stubbe ___

Committee Chair
Novartis Professor of Chemistry and Professor of Biology

Professor Catherine L. Drennan __

Research Supervisor
Professor of Chemistry and Biology
Howard Hughes Medical Institute Investigator and Professor

Professor Alice Y. Ting _____ __

ber
Associate Professor of Chemistry

Structural Studies of Metalloenzyme Complexes in Acetogenic Carbon Fixation

by
Yan Kung

Submitted to the Department of Chemistry
on May 5, 2011 in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Biological Chemistry

ABSTRACT

Acetogenic bacteria use the Wood-Ljungdahl carbon fixation pathway to produce cellular carbon from CO_2. This process requires several metalloenzymes that employ transition metals such as iron, nickel, and cobalt towards the production of acetyl-CoA, the final product. In one stage of the pathway, the cobalt-containing B_{12} cofactor harbored by the corrinoid iron-sulfur protein (CFeSP) transfers a methyl group from methyltetrahydrofolate (CH_3-H_4folate), which is bound by a methyltransferase enzyme (MeTr), to a nickel-containing metallocluster called the A-cluster of the downstream enzyme, acetyl-CoA synthase (ACS). Such B_{12}-dependent methyl transfer reactions require the construction of large, multimodular enzyme complexes whose three-dimensional assemblies are, at present, largely uncharacterized. X-ray crystallography was used to solve the structure of a CFeSP/MeTr complex, the first crystal structure of a B_{12}-dependent methyltransferase to depict all protein domains required for B_{12} binding, activation, protection, and catalysis. This structure, along with *in crystallo* activity data, illustrates how conformational movements, which can occur within protein crystals, enable the B_{12} cofactor to alternate between a sequestered conformation for cofactor protection and an active conformation for catalysis. Small-angle X-ray scattering (SAXS) experiments were also conducted to explore the quaternary composition of the complex in solution and revealed that multiple CFeSP/MeTr complexes can be formed. In another reaction of the Wood-Ljungdahl carbon fixation pathway, a nickel and iron containing metallocluster called the C-cluster of carbon monoxide dehydrogenase (CODH) reduces a second molecule of CO_2 to CO, an intermediate that is channeled to the ACS A-cluster. Although the structure of the C-cluster was first described a decade ago, its catalytic mechanism remained unresolved. To provide mechanistic insight into the chemistry employed at the C-cluster, crystal structures were determined with substrate and inhibitor molecules bound to the C-cluster of the CODH/ACS complex. These structures capture states of the C-cluster at key steps in the reaction and contribute to a consensus model for C-cluster chemistry. With structural descriptions for both CFeSP/MeTr and CODH/ACS complexes, this work has illuminated the molecular details for metalloenzyme complex assembly and catalysis in the acetogenic Wood-Ljungdahl carbon fixation pathway.

Thesis Supervisor: Catherine L. Drennan

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

O baffled, balked,
Bent to the very earth, here preceding what follows,
Oppressed with myself that I have dared to open my mouth,
Aware now, that amid all the blab whose echoes recoil upon me,
I have not once had the least idea who or what I am...

O I perceive I have not understood anything—not a single object—
and that no man ever can.

I perceive Nature here, in sight of the sea, is taking advantage of me,
to dart upon me, and sting me,
Because I was assuming so much,
And because I have dared to open my mouth to sing at all.

Walt Whitman

ACKNOWLEDGEMENTS

Although this thesis bears my name, scientific research is a truly collaborative endeavor, and I must thank many people who have contributed to this work. First, I thank my thesis advisor, Cathy Drennan, for giving me the great opportunity to work on such challenging and rewarding projects, for providing me with scientific and personal support, and for her time, energy, and patience over the years. Often, when things became difficult, I looked at the dedication she has to her research and her lab as motivation to keep me going.

I am also grateful to the other members of my thesis committee over the years, JoAnne Stubbe, Sarah O'Connor, and Alice Ting, for their help and suggestions as I progressed through my graduate studies. Special thanks to JoAnne, who was willing to be my thesis chair at fairly short notice.

In addition, I especially valued the interactions I had with research collaborators. Although I have never met Tzanko Doukov, his past contributions to my research years before I joined the lab are manifold, from the discovery of some initial crystallization conditions to the collection of X-ray diffraction data. Steve Ragsdale was very forthright with advice through many of the projects of my thesis, and I had the real honor of spending a great deal of time with him at many research conferences and other occasions. Güneş Bender kindly tolerated my frequent, last-minute requests for protein samples. Finally, working with Nozomi Ando was such a blast, especially with the small-angle X-ray scattering experiments.

The members of the Drennan lab are a unique bunch and have shaped my graduate student experience tremendously. I have strongly benefited from my experiences with all of the post-docs in the lab that I've known: Paul Hubbard, C. Ainsley Davis, Danny Yun, Mishtu Dey, Nozomi Ando, Daniel Dowling, Edward Brignole, and Sarah Bowman. The things I may remember most from each of you are, respectively: the picture of Paul and Angelina Jolie that I kept (and treasure– in a non-awkward way!), playing softball and volleyball in the summer with Ainsley, watching Danny shim the dongle or bash the kernel (just Google it!), frequent trips to 'Bucks with Mishtu, "Yan and Zomer! Yan and Zomer!", learning just how many structures Dan can solve in a single synchrotron trip, traumatizing Ed almost daily with my embarrassing immaturity, and surrendering my bench to Sarah and her silly NMR tubes because I felt guilty about her unfortunate desk situation. Thank you all!

The tribe of graduate students in the Drennan lab is equally unique and gifted, in more ways that I can describe. Hector Hernandez- our leader, he was like the grandpa of the lab. Thanks for always being there (no really, he was always there, blocking the glassware cabinets). Jess Vey showed me the ropes on how to maintain the X-ray equipment; I'm so glad we moved to Biology because that lab job was a lot of work! I will always look up to Kaity Ryan for her dedication- I don't think anyone set up as many crystallization trays as she did (and kept their sharp sense of humor). I'm also grateful to Cintyu Wong, who by personal example taught me the crucial lesson that is so important

to learn in research and in life: when to step back, turn around, and move confidently in a different direction. I also had the pleasure of working with Laura Phillips, who graciously offered to share her bench with a strange, new graduate student, and whose desk I promptly took over the day after she left.

I'm very fond of the newer graduate students, and because of them I'm sure that the lab will continue to be a great place to work. Thanks to Christina Stock, for being my skating buddy and for her intriguing taste in music. Thanks to Peter Goldman for periodically surprising us all with different sides to his peculiar personality and for giving us Eurotrash Thursdays. Thanks to Jeremy Setser for so perfectly replicating the way that I sigh in exasperation and for frequently talking with Peter in an odd and cryptic language. Thanks to Marcus Gibson for boldly accepting his role as our token inorganic chemist. Thanks to Michael Funk for always seeming so stressed out that it calms me down (and cracks me up). And thanks to Marco Jost for actually being an American.

I loved that I was able to work alongside Leah Blasiak in the lab. Her calm and confident balance of hard work and a rich outside life was inspiring to me, and I am truly grateful to have shared both a friendship and a research project with her.

Although Christine Phillips-Piro graduated from the lab over a year ago, she left such an indelible mark on all of us that I can't understand how we've managed this long without her. A generous and close friend, I spent countless days and nights hanging out with her and Nick; it was difficult to adjust to life in Boston without her, her wonderfully cooked meals, her manic kittens, and her uproarious, out-sized personality.

Other lab members to thank: Becky Bjork (whose data amused us with its 98% "complete mess"), Bill Hillmann, Jun Wang, and honorary lab member Nicholas Piro. Thanks as well to the undergraduate researchers over the years whom I have mentored: Sarah Whiteside, Sarah King, and Johan Nakuci. I am also very thankful to the old Building 56 fifth-floor community: Laura J., Mathew, Mary, Melva, Wan-Chen, Maria, Eugene, Alisha, Jen, and Bryan. The worst part about moving our lab was leaving you all.

Finally, I thank my family. My parents have taught me so much and provided me with such a full and privileged upbringing. They always knew that I could do more, reach higher, and push myself farther. I am also grateful to my older sister, Cee, who for years paved the road for me, clearing the way to make my life just that much easier. Last of all, I thank Casey Miller, my partner of over ten years. I made it this far with you by my side, and without your support, love, wisdom, and advice, I would surely have cracked and crumbled. I couldn't have done this without you!

TABLE OF CONTENTS

Chapter I: Acetogenesis and the Wood-Ljungdahl carbon fixation pathway

I.A. Biological carbon fixation

Life on Earth requires a biological means to fix inorganic carbon in the environment to organic carbon usable for cellular metabolism and biomass generation. Organisms capable of such a task are termed autotrophs and can either use sunlight as an energy source for carbon dioxide reduction, as with photoautotrophs, or energy can be derived from chemical reactions using inorganic substrates, as in the case of chemoautrotrophs. Autotrophic organisms provide all biological carbon to sustain life on the planet. Until now, six pathways for fixing of carbon have been discovered.

The **Calvin-Benson-Bassham cycle** (Figure I.1), often simplified as the Calvin or Calvin-Benson cycle, is utilized by photoautotrophic organisms including plants, algae, cyanobacteria, and other bacteria such as green and purple sulfur bacteria. In the Calvin cycle, CO_2 reacts with ribulose 1,5-bisphosphate (RuBP) to form two molecules of 3-phosphoglycerate (3PG) in a reaction catalyzed by RuBP carboxylase/oxygenase, also known as RuBisCO, the most abundant enzyme on Earth[1]. 3PG can then be directly used for metabolic purposes through glyceraldehyde-3-phosphate (G3P), either by glycolysis or in the production of starch for carbon storage. G3P is also used to regenerate RuBP for subsequent reactions with CO_2, wherein five G3P molecules are combined to form a series of 3-, 4-, 5-, and 7-carbon intermediates, resulting in three RuBP molecules.

The Calvin cycle, which constitutes the light-independent reactions of photosynthesis, is the predominant source of biological carbon on the planet, fixing 10^{11} tons of CO_2 per year[2] . However, it is also perhaps the most recently acquired means of carbon fixation in the evolution of life on Earth[3]. Because oxygenic photosynthesis is the

source of molecular oxygen in the atmosphere, prior to the advent of photosynthesis and the Calvin cycle other carbon fixation pathways must have provided biological carbon for the ancient, anaerobic world.

Some anaerobic bacteria and archaea use the **reductive or reverse citric acid cycle** (Figure I.2) for CO_2 fixation[4,5]. Here, as the name suggests, the oxidative citric acid cycle (also known as the tricarboxylic acid cycle or the Krebs cycle) is essentially run in the reverse direction, where CO_2 incorporation occurs during the carboxylation of succinate and α-ketoglutarate. As an anaerobic, autotrophic pathway, the reverse citric acid has been suggested to be a precursor to the oxidative citric acid cycle[3].

Three related carbon fixation pathways that perform carboxylation of acetyl-CoA have recently been uncovered (Figure I.3). In the **3-hydroxypropionate/4-hydroxybutyrate cycle**[6,7] (Figure I.3, red arrows) found in some aerobic archaea, acetyl coenzyme A (CoA) and subsequently propionyl-CoA are carboxylated using bicarbonate by the same biotin-dependent enzyme, acetyl-CoA-propionyl-CoA carboxylase, in a string of reactions that produce succinyl-CoA, via methylmalonyl-CoA. Succinyl-CoA is then converted into two acetyl-CoA molecules, through 4-hydroxybutyrate, to start the cycle again, with the net production of an additional molecule of acetyl-CoA. Tandem carboxylations of acetyl-CoA and propionyl-CoA using bicarbonate are also performed in the **3-hydroxypropionate bi-cycle**[6,8,9] (Figure I.3, blue arrows), found only in the Chloroflexus bacterial species thus far. In one half of the bi-cycle, propionyl-CoA carboxylation and subsequent rearrangement to succinyl-CoA is followed by the regeneration of one molecule of acetyl-CoA plus one molecule of glyoxylate. In the other half of the bi-cycle, propionyl-CoA may combine with glyoxylate from the first half to

produce methylmalyl-CoA, which after a series of transformations is split to form acetyl-CoA and pyruvate. The acetyl-CoA products from both halves of the bi-cycle may then begin the cycle anew. The newly discovered **dicarboxylate/4-hydroxybutyrate cycle**[6,10] (Figure I.3, black arrows) found in some anaerobic archaea also involves two carboxylation steps. The first is once again the carboxylation of acetyl-CoA, this time using CO_2 instead of bicarbonate to directly form pyruvate in a reaction catalyzed by pyruvate synthase, also called pyruvate:ferredoxin oxidoreductase (PFOR), an enzyme that utilizes the cofactor thiamine pyrophosphate (TPP) and is employed in many biological contexts. Pyruvate is then converted to phosphoenolpyruvate (PEP), which undergoes the second carboxylation to form oxaloacetate, this time using bicarbonate, in a reaction catalyzed by PEP carboxylase. Oxaloacetate then undergoes a partial reverse citric acid cycle, stopping at succinyl-CoA. This intermediate is finally split into two molecules of acetyl-CoA, via 4-hydroxybutyrate, just as in the 3-hydroxypropionate/4-hydroxybutyrate cycle.

The sixth and final pathway is the **Wood-Ljungdahl carbon fixation pathway**[11], also called the reductive acetyl-CoA pathway (Figure I.4). The metabolic hallmark of strictly anaerobic acetogenic bacteria, it is also used for carbon fixation by methanogenic archaea, which along with sulfate-reducing bacteria can also run portions of the pathway in reverse. A detailed description of the individual steps of the Wood-Ljungdahl carbon fixation pathway is given in sections I.C and I.D below.

The Wood-Ljungdahl pathway differs significantly from other carbon fixation pathways in that it does not involve carboxylation of intermediates that are subsequently rearranged to regenerate the original species prior to the next cycle. Instead, the carbon

atoms of both CO_2 substrates directly form the carbon atoms of acetyl-CoA, and thus the pathway is not cyclic. In addition, the Wood-Ljungdahl pathway is the least energetically expensive carbon fixation pathway[6,9,11], as it requires just one molecule of ATP (Figure I.4), which may be recouped if acetyl-CoA is hydrolyzed to acetate (see section I.B). In contrast, the overall synthesis of one pyruvate molecule in the Calvin-Bensen-Bassham cycle requires seven ATP molecules, the reverse citric acid cycle requires two, the 3-hydroxypropionate/4-hydroxybutyrate cycle requires nine, the 3-hydroxypropionate bi-cycle requires seven, and the dicarboxylate/4-hydroxybutyrate cycle requires five molecules of ATP. Finally, the Wood-Ljungdahl pathway is also particularly unique in its use of enzymes that harbor trace metal ions, which include nickel, iron, cobalt, and tungsten or molybdenum, some of whom form metal-carbon bonds. These metalloenzymes are some of the most oxygen sensitive enzymes known, and thus the Wood-Ljungdahl carbon fixation pathway is utilized only by obligately anaerobic organisms.

The presence of complex catalytic metallocofactors, the strictly anaerobic requirement for the pathway, and the thermophilic nature of some organisms that utilize the pathway have inspired theories that the Wood-Ljungdahl carbon fixation pathway may have been among the first autotrophic pathways in the origins of life on Earth[3]. These unique characteristics of the Wood-Ljungdahl carbon fixation pathway have attracted the interest of biochemists, microbiologists, and inorganic chemists, and the structural study of key metalloenzyme complexes of the pathway are the focus of this thesis.

I.B. The Wood-Ljungdahl carbon fixation pathway in acetogenic bacteria

An acetogenic bacterium, or "acetogen" for short, is any anaerobic bacterial species that is capable of reducing environmental CO_2 to acetyl-CoA as a carbon source and for energy conservation through the Wood-Ljungdahl carbon fixation pathway[12,13]. Bacteria that fall under this category are diverse and include gram-positive and gram-negative species, rod, coccoidal, and spiral cellular morphologies, and thermophilic, mesophilic, and psychrotrophic temperature optima[13]. Acetogens live in a variety of habitats across the biosphere, such as terrestrial soils, freshwater and marine sediments, hot springs, and the intestinal tract of animals, including humans. Acetogenesis is thus not a phylogenetic trait, but is instead a metabolic characteristic used by a wide range of microorganisms. Despite this diversity, all acetogenic bacteria are anaerobic; some enzymes of the Wood-Ljungdahl pathway are highly sensitive to oxygen, which can damage the metallocofactors required for activity.

The model acetogenic organism is *Moorella thermoacetica*, known as *Clostridium thermoaceticum* prior to reclassification in 1994[14]. Its ability to grow to high cell densities allowed relatively simple laboratory cultivation, and its mildly thermophilic nature (temperature optimum of 55 °C) allowed isolation of thermostable samples. Acetogenesis was, however, discovered in 1932 in sewage[15], and the first acetogen, *Clostridium aceticum*, was isolated four years later[16], but the culture was soon lost. Now, over 100 acetogenic species have been isolated to date, representing 22 distinct genera[13]. It is estimated that collectively, acetogenic bacteria use the Wood-Ljungdahl carbon fixation pathway to generate 10^{12} kg of acetate per year from CO_2[12], of which 10^{10} kg are produced from acetogens in the human gut alone[17,18], and 10^9 kg per day are thought to

derive solely from acetogens in the hindgut of termites, organisms that play a major role in cellulose degradation[19].

Methanogeneic archaea, or "methanogens" for short, can also grow autotrophically via acetogenesis using the Wood-Ljungdahl carbon fixation pathway[11,20]. Methanogens are strict anaerobes that derive energy by the generation of methane (CH_4) from CO_2/H_2, acetate, and sometimes from other one-carbon methyl sources, such as methanol, methylamines, and methylthiols[20,21]. If methanogenic growth is driven by reduction of CO_2 with H_2 to CH_4, methanogens can use the Wood-Ljungdahl pathway for carbon fixation. Here, a methyl group derived from reduced CO_2 may either be converted to methane, or it may become the methyl of acetyl-CoA. On the other hand, acetoclastic methanogens essentially run the pathway in reverse, generating CO_2 and methane from acetate oxidation. Methanogens produce an estimated 10^9 tons of methane per year[20,22], approximately two thirds of which derives from acetate[21,23].

Although this chapter is primarily focused on mechanisms for carbon fixation, these pathways all require varying amounts of energy derived by ATP hydrolysis, as discussed above. In the Calvin cycle, for example, the necessary ATP and reducing NADPH equivalents needed for carbon fixation are derived from sunlight energy in the light-dependent reactions of photosynthesis. However, the analogous means by which acetogenic carbon fixation is coupled to energy conservation in acetogens is still unclear; as mentioned above, one molecule of ATP is required in the Wood-Ljungdahl carbon fixation pathway to form acetyl-CoA, which is used as a metabolic intermediate in the production of biomass. Alternatively, acetyl-CoA may instead be used for ADP phosphorylation to recoup the molecule of ATP; therefore, the overall ATP requirement

for carbon fixation in acetogens is slightly less than one equivalent. In this process, phosphotransacetylase[24] uses inorganic phosphate to phosphorylate the acetyl moiety, liberating CoA. Acetate kinase then transfers the phosphate to ADP, forming ATP and acetate as a byproduct. Despite this recouping option that somewhat offsets the ATP requirement, there is still no *net* gain of ATP during carbon fixation. Therefore, a link that couples carbon fixation and a mechanism for net energy production for the cell is likely to exist, perhaps through the establishment of a transmembrane proton gradient. In methanogenesis, much of the link between CO_2 reduction and energy conservation is known[22], where CO_2 reduction is coupled to the formation of a transmembrane proton gradient that is harnessed by an ATP synthase. In *M. thermoacetica*, the chemiosmotic F_1F_0 ATP synthase has been isolated[25], and cytochrome-based electron transfer pathways have been proposed[11,26-29]. Still, the precise connection between the Wood-Ljungdahl carbon fixation pathway and a net energy conservation mechanism has not yet been identified in acetogenic bacteria.

I.C. Enzymes of the Wood-Ljungdahl pathway: the Eastern branch

The Wood-Ljungdahl carbon fixation pathway may be divided into two branches[11,30,31]: the methyl or "Eastern" branch and the CO or "Western" branch, each branch concerning one molecule of CO_2 (Figure I.4, red and blue).

In the Eastern branch (Figure I.5), one CO_2 molecule is reduced by six electrons to a methyl group. These reactions require folic acid and are by no means restricted to the Wood-Ljungdahl carbon fixation pathway. Heterotrophic organisms such as humans use several enzymes analogous to Eastern branch enzymes, but derive one-carbon (C_1) units

for entry into the folate pool not by CO_2 reduction but through the conversion of serine to glycine by serine hydroxymethyltransferase[32]. Folate-dependent C_1 transfer constitutes an important feature in nucleotide and amino acid metabolism, as well as methylation biochemistry. Methanogenic archaea also use enzymes analogous to the acetogenic Eastern branch enzymes in the reduction of CO_2 towards methane, using a unique folate analog called methanopterin[33].

The first reaction of the Eastern branch is the reduction of CO_2 to formate using NADPH and molybdopterin, catalyzed by formate dehydrogenase. *M. thermoacetica* formate dehydrogenase is a ~340 kDa $\alpha_2\beta_2$ enzyme that contains tungsten, a selenocysteine residue, and several iron-sulfur clusters[34,35]. Whereas other NADPH-utilizing formate dehydrogenases use molybdenum to coordinate molybdopterin, the *M. thermoacetica* enzyme uses tungsten instead of molybdenum for the same purpose. In the molybdenum enzyme that operates in the direction of formate oxidation, it is thought that selenocysteine that ligates a Mo(VI) ion is displaced by formate, which binds via a carboxylate oxygen[36,37]. A catalytic base removes the C-H proton, liberating CO_2 while Mo(VI) is reduced to Mo(IV). To regenerate Mo(VI), electrons are transferred through [4Fe-4S] clusters to an external electron acceptor, and selenocysteine can re-ligate the Mo(VI) ion; the molybdopterin cofactor is not released following catalysis. It is thought that the tungsten-containing formate dehydrogenase follows the same mechanism, with reducing equivalents deriving from NADPH oxidation[11].

In the second step, formate is attached to the N^{10} position of the folate cofactor in a reaction catalyzed by 10-formyl-tetrahydrofolate synthetase[38-40]. As the name of the enzyme suggests, this reaction requires the hydrolysis of ATP to first activate formate to

a formylphosphate intermediate. The N^{10} position of tetrahydrofolate (H_4folate) then attacks the formyl moiety to form the product, with phosphate as a leaving group. This step represents the only ATP-requiring step of the entire Wood-Ljungdahl carbon fixation pathway.

Third, 5,10-methenyl-H_4folate cyclohydrolase[41] catalyzes cyclization to form a five-membered ring that contains N^5, N^{10}, and the formyl carbon, releasing water and a proton. In the fourth step, catalyzed by 5,10-methylene-H_4folate dehydrogenase[42,43], NAD(P)H-dependent reduction converts the methenyl group to a methylene group. In some organisms, the cyclohydrolase and the dehydrogenase are distinct enzymes, whereas in other organisms, including *M. thermoacetica*, they operate as a bifunctional complex[11,42].

The fifth and final step of the Eastern branch is catalyzed by 5,10-methylene-H_4folate reductase, which in acetogenic bacteria is an oxygen-sensitive multimeric enzyme harboring an iron-sulfur cluster[44]. Interestingly, only the acetogenic enzyme contains this cluster, and all others, including the human enzyme, are oxygen-stable[11]. The enzyme also uses FAD as an electron mediator, with electrons deriving from ferredoxin and, ultimately, NAD(P)H. The product of the reaction is 5-methyl-H_4folate (CH_3-H_4folate), completing the reduction of the CO_2-derived carbon atom.

I.D. Enzymes of the Wood-Ljungdahl pathway: the Western branch

In the Western branch of the Wood-Ljungdahl pathway (Figure I.6), the methyl group from the Eastern branch is removed from CH_3-H_4folate through the combined action of two enzymes. In parallel, a third enzyme reduces a second molecule of CO_2 to

carbon monoxide (CO), which is an intermediate that is then combined with the methyl group and CoA to form acetyl-CoA in reactions catalyzed by a fourth enzyme. Although there are fewer enzymes in the Western branch (four, instead of five), the chemistry performed is still poorly understood in comparison to the reactions of the Eastern branch.

Removal of the methyl group from CH_3-H_4folate produced in the Eastern branch requires two of the four Western branch enzymes: the B_{12}- and Fe_4S_4-containing corrinoid iron-sulfur protein (CFeSP) and the enzyme CH_3-H_4folate:CFeSP methyltransferase (MeTr). Here, the methyl group of CH_3-H_4folate is transferred to the cobalt center of the B_{12} cofactor harbored by CFeSP. This CFeSP/MeTr-dependent methyl transfer reaction forms the subject of Chapters II and III of this thesis. Briefly, the methyl group in question is bonded to the N^5-tertiary amine of CH_3-H_4folate and is not electrophilic, and H_4folate is a poor leaving group[45]. CH_3-H_4folate thus binds MeTr, which activates the methyl group for attack by the strongly nucleophilic B_{12} cofactor of CFeSP, with cobalt in the +1 oxidation state, Co(I). The CFeSP Fe_4S_4 cluster is responsible for passing an electron to cobalt to maintain the Co(I) state should occasional oxidation yield the inactive Co(II) state[46,47]. Finally, following methyl transfer the B_{12} cobalt is converted to the CH_3-Co(III) state. This species represents one of multiple organometallic intermediates of the Wood-Ljungdahl pathway's Western branch.

All of the other organometallic intermediates involve the second molecule of CO_2, where the only enzymes recruited are carbon monoxide dehydrogenase (CODH) and acetyl-CoA synthase (ACS), both nickel- and iron-containing enzymes. Some unrelated aerobic organisms contain a copper-molybdenum CODH variant that uses FAD and molybdopterin cytosine dinucleotide cofactors in addition to several iron-sulfur clusters[48].

This Mo-Cu CODH will not be discussed in this thesis, as we will consider only the anaerobic, acetogenic CODH of the Wood-Ljungdahl carbon fixation pathway. In acetogens, CODH and ACS form a ~310 kDa bifunctional $\alpha_2\beta_2$ CODH/ACS enzyme complex[49], where CODH assembles as a homodimer (β_2) with one ACS monomer (α) bound on either side.

In each CODH monomer, a biologically unprecedented nickel-iron-sulfur cluster called the C-cluster resides at the active site and catalyzes the interconversion of CO_2 to carbon monoxide (CO). While it was previously thought that the C-cluster was composed of a Fe_4S_4 cubane, [4Fe-4S], with a nickel atom coordinated nearby[50,51], initial crystallographic studies illustrated that nickel is part of a distorted [Ni-3Fe-4S] cubane[52,53]. A fourth iron atom is coordinated nearby, bridged to a sulfide of the cubane, and is often referred to as the "unique" iron or Ferrous Component II (FCII). Therefore, the overall composition of the C-cluster is [Ni-4Fe-4S] (Figure 1.7.A). The catalytic mechanism of the C-cluster has been the subject of controversy and forms the focus of Chapters IV and V. Briefly, in the direction of CO oxidation, it is believed that CO binds nickel of the cluster, while a water molecule binds the unique iron[54]. Upon deprotonation of water, the resulting hydroxide ion attacks CO to form a nickel-bound COOH species. Further deprotonation yields Ni-COO^-, and CO_2 may then be liberated with two-electron reduction of the C-cluster. To re-oxidize the C-cluster, electrons are then passed from the C-cluster through additional Fe_4S_4 clusters harbored by the enzyme and ultimately to an electron acceptor protein.

In addition to the Wood-Ljungdahl carbon fixation pathway, in which CODH operates in the direction of CO_2 reduction to CO, CODH is also present in other systems.

A monofunctional CODH exists in microorganisms such as *Rhodospirillum rubrum*[55] and *Carboxydothermus hydrogenoformans*[56], which can catalyze CO oxidation to CO_2 as an energy and carbon source. CO oxidation yields electrons that may be coupled to ATP production, and CO_2 may then be fixed via the Wood-Ljungdahl pathway or other autotrophic pathways. Globally, CODH activity has been estimated to remove 10^8 tons of CO from the environment per year[57]. In addition, acetoclastic methanogenic archaea possess both CODH, ACS, and CFeSP as components of the large acetyl-CoA decarbonylase/synthase (ACDS) complex[58], which is involved in the conversion of acetyl-CoA to CO_2 and methane. This series of reactions may be thought of as a modified reverse Wood-Ljungdahl pathway; the acetyl-CoA methyl group is similarly transferred to a folate derivative, methanopterin. However, the methyl group is not subsequently oxidized to CO_2 but is passed to coenzyme M (CoM) to form methyl-CoM. Finally, through the action of methyl-CoM reductase, the methyl group is liberated as methane.

CO generated at the CODH C-cluster is an intermediate that is passed to the active site metallocluster of ACS termed the A-cluster, approximately 70 Å away. To travel this distance, CO does not dissociate from the CODH/ACS complex but instead travels through a remarkable tunnel within the complex itself[49,59-62]. Like the C-cluster, the A-cluster is also a Ni-Fe-S cluster, but the atoms are arranged in a completely different and novel geometry. Crystallographic and biochemical studies have revealed that the A-cluster contains a [4Fe-4S] cubane coordinated through a cysteine ligand to a binuclear Ni site[49,59,63-65]. Therefore, the overall composition of the A-cluster is [4Fe-4S]-Ni-Ni (Figure I.7.B). The A-cluster binds CO originating from CO_2 reduction by CODH, the CO_2-derived methyl group from the Eastern branch and donated from B_{12} of CFeSP,

and CoA to form acetyl-CoA. Although the mechanism of the A-cluster is currently the subject of considerable controversy, it is believed that the nickel atom closest to the [4Fe-4S] cubane (also known as the proximal nickel, Ni_p) binds both CO from CODH and the methyl group from CFeSP to form a nickel-bound acetyl group, via CO insertion[11,65]. CoA may then liberate the acetyl group to produce acetyl-CoA, completing the Wood-Ljungdahl pathway.

I.E. Primary goals of the thesis

As we have seen, the Western branch of the Wood-Ljungdahl carbon fixation pathway utilizes multiple metalloenzymes to perform chemistry on one-carbon molecules in the creation of several organometallic intermediates. CFeSP and MeTr work together to transfer a methyl group from a folate cofactor, CH_3-H_4folate, to a B_{12} cofactor harbored by CFeSP, which then delivers the methyl group to ACS. Meanwhile, the C-cluster of CODH reduces CO_2 to CO, an intermediate that is also transmitted to ACS. Thus, the methyl group and CO converge on the ACS A-cluster to form the final product, acetyl-CoA.

In this thesis, we shall explore the one-carbon transformations performed by these four enzymes from a structural perspective. As we shall see, these enzymes assemble into large enzyme complexes, namely the CFeSP/MeTr complex of B_{12}-dependent methyl transfer and the bifunctional CODH/ACS complex. Overall, the studies described in this thesis provide structural contributions towards both the description of quaternary structure and structural rearrangements (Chapters II and III), as well as the elucidation of metalloenzyme mechanisms (Chapters IV and V).

In Chapter II, we begin by studying the B_{12}-dependent methyl transfer reaction carried out by CFeSP and MeTr from *M. thermoacetica*. B_{12}-dependent methyltransferases are characterized by the use of multiple protein domains recruited to transfer a single methyl group[45,66]. Further, large conformational domain motions have long been postulated to be crucial in these enzymes; however, as knowledge of their three-dimensional domain assembly has been lacking, these domain movements have not been visualized until this work. Here, the structure of the CFeSP/MeTr enzyme complex is presented, and key objectives of this study include: (i) characterization of the three-dimensional arrangement of domains in B_{12}-dependent methyl transfer enzymes, (ii) description of structural movements that must occur during the reaction cycle, and (iii) providing structural insight into why large protein machinery is enlisted to transfer a single methyl group.

The CFeSP/MeTr structure revealed unexpected features, most notably the binding of two CFeSPs to one MeTr homodimer. To investigate the oligomeric complex in solution, small-angle X-ray scattering (SAXS) was employed. In Chapter III, we explore the complexes of CFeSP and MeTr that form under different conditions, with the objective of gaining insight into the way(s) in which these proteins may bind each other by examining their quaternary structures in solution.

In Chapter IV, we change our focus to the CODH-catalyzed reduction of CO_2 to CO in crystallographic studies of the nickel-iron-sulfur C-cluster. Although the structure of the C-cluster was first revealed in 2001[52,53], the mechanism for CO/CO_2 interconversion and the requirement for a putative sulfide bridge in the cluster remained uncertain and controversial. In this work, crystal structures of the C-cluster from the *M.*

thermoacetica bifunctional CODH/ACS complex are solved bound with substrate and inhibitor molecules. The central goal of this project is to use X-ray crystallography to help unravel the catalytic mechanism of this unprecedented metallocluster.

While work on the *M. thermoacetica* CODH/ACS C-cluster progressed, additional crystallographic studies were reported that also depicted substrates bound to the C-cluster of CODH from other organisms[67,68]. These structures, in conjunction with the structures described in Chapter IV, have helped to reconcile the seemingly conflicting literature on the CODH C-cluster and converge on a potentially unifying mechanism for both C-cluster reactivity and inhibition. In Chapter V, we compile and compare the C-cluster crystal structures to put forth a mechanism that brings together decades of study of this fascinating metallocluster.

Finally, we conclude in Chapter VI by expanding our focus to consider the broader significance of the structural work described in this thesis, including the implications both for B_{12}-utilizing methyl transfer enzymes in other systems and for C-cluster-mimicking model systems. In addition, the Wood-Ljungdahl carbon fixation pathway has attracted interest for possible industrial applications in alternative energy and carbon sequestration. We thus further consider the potential relevance of the acetogenic Wood-Ljungdahl carbon fixation pathway in the search for solutions to the global climate and energy crisis.

I.F. Figures

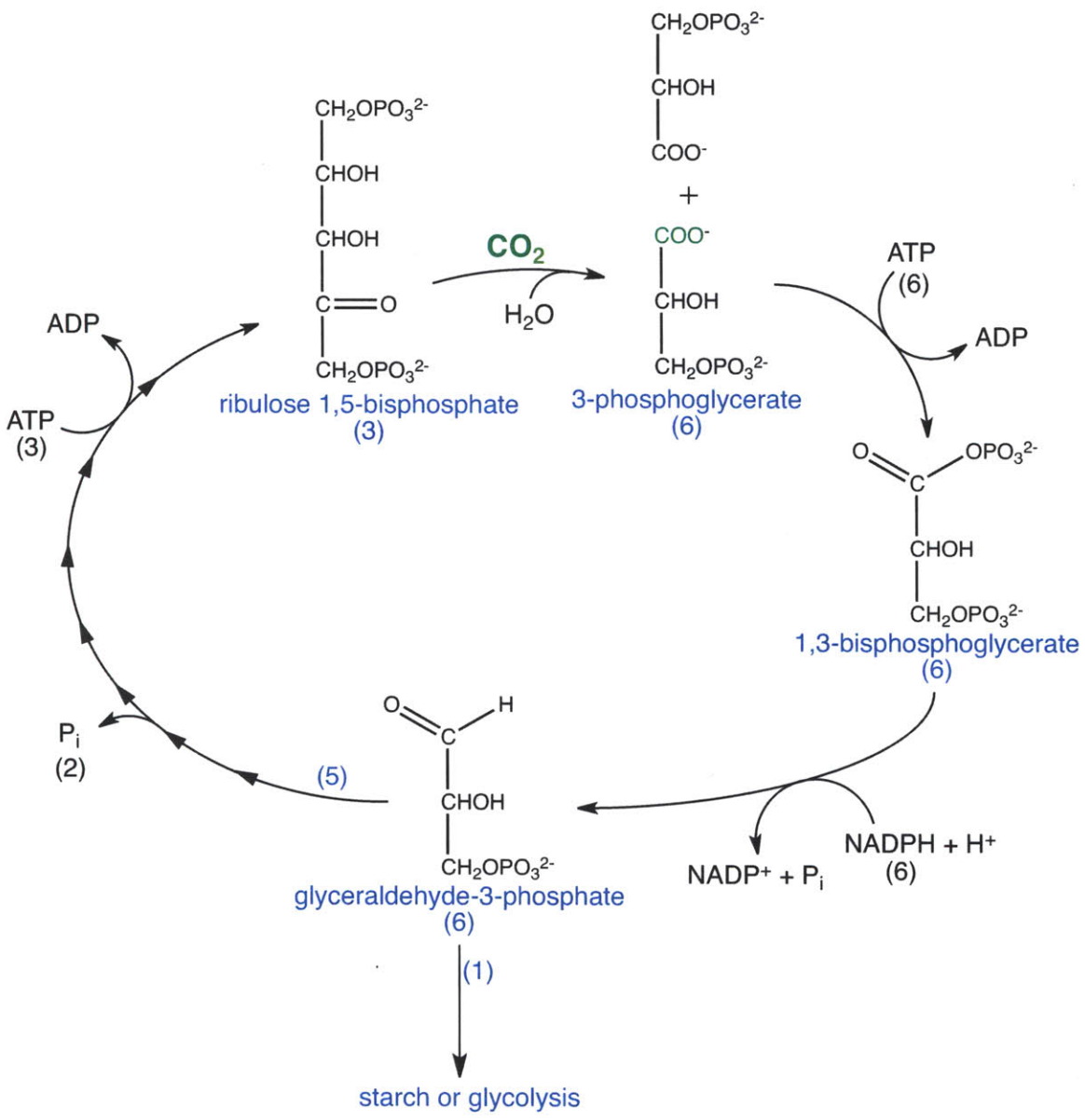


Figure I.1. The Calvin-Benson-Bassham carbon fixation cycle. Key intermediates are shown and labeled in blue, the fixed CO_2 molecule is labeled in green, and stoichiometry is indicated by numbers in parentheses. Multiple steps are indicated by arrows with multiple arrowheads.

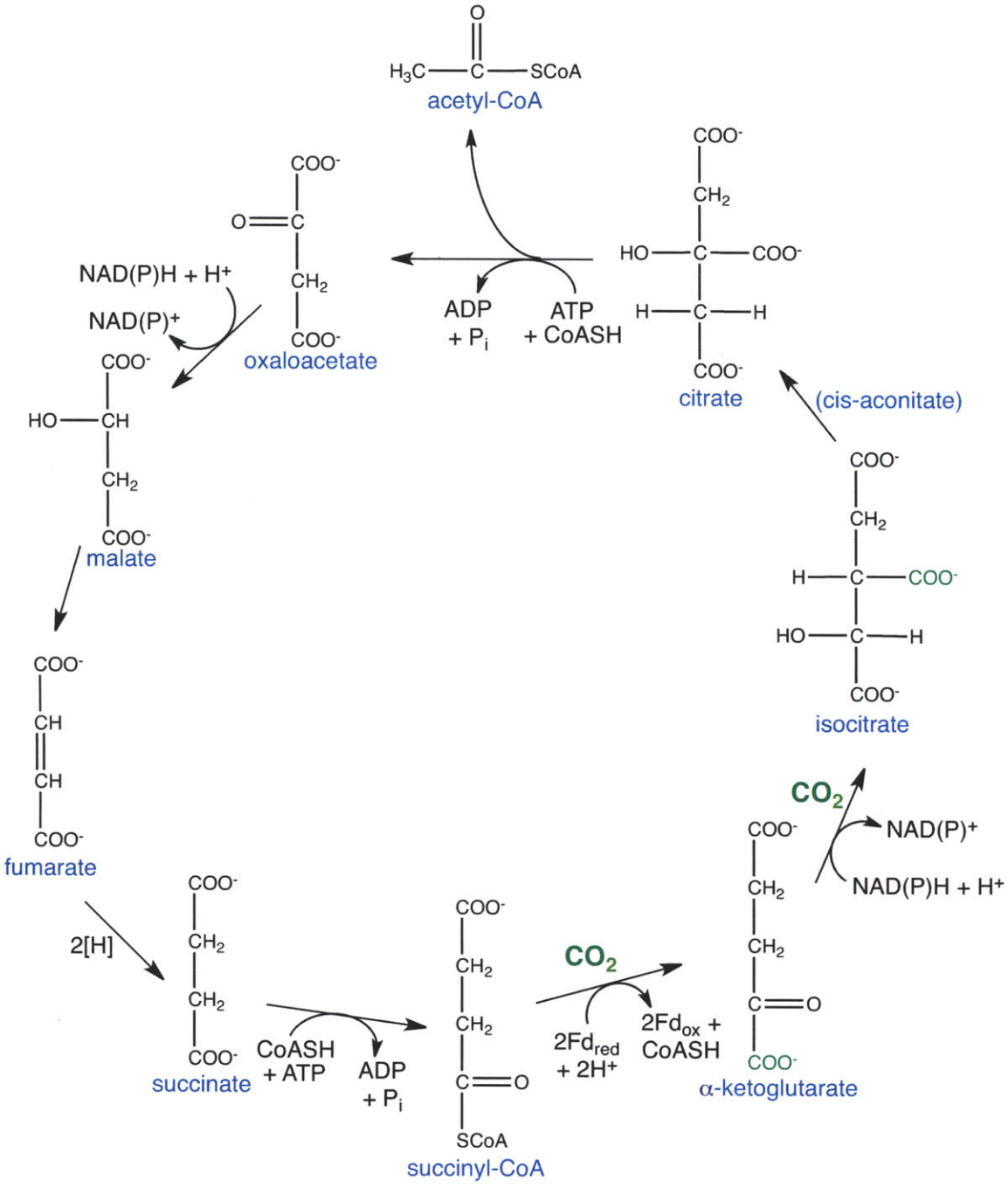


Figure I.2. The reverse citric acid cycle. Intermediates are labeled in blue, and fixed CO_2 molecules are labeled in green. "[H]" refers to a reducing equivalent derived from an unknown source.

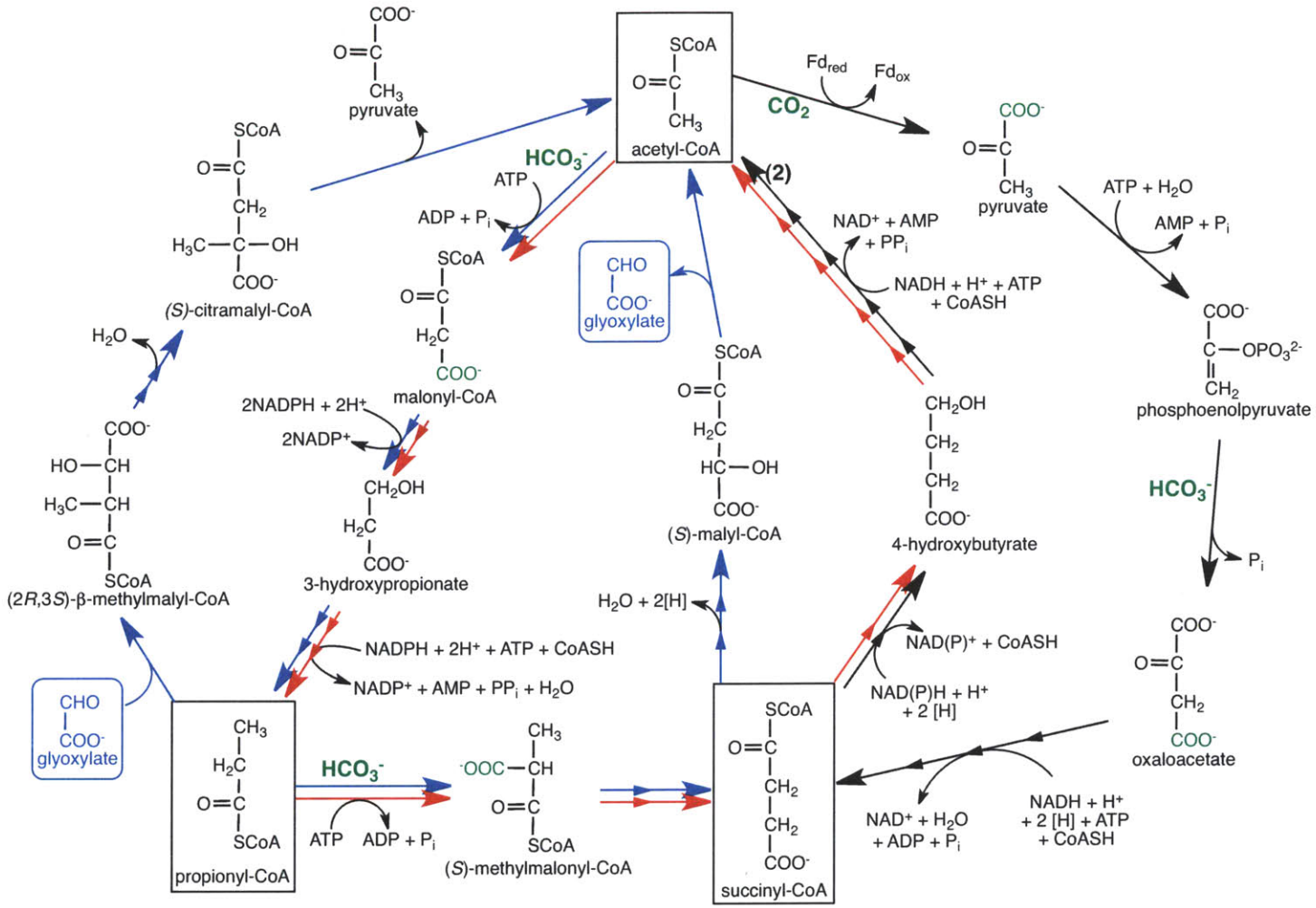


Figure I.3. Three related carbon fixation cycles. The 3-hydroxypropionate bi-cycle is indicated by the blue arrows, the 3-hydroxypropionate/4-hydroxybutyrate cycle is indicated by red arrows, and the dicarboxylate/4-hydroxybutyrate cycle is indicated by black arrows. The fixed CO_2 or HCO_3^- molecules are labeled in green. Multiple steps are indicated by arrows with multiple arrowheads. Although the three pathways share many intermediates, the enzymes used are not all the same.

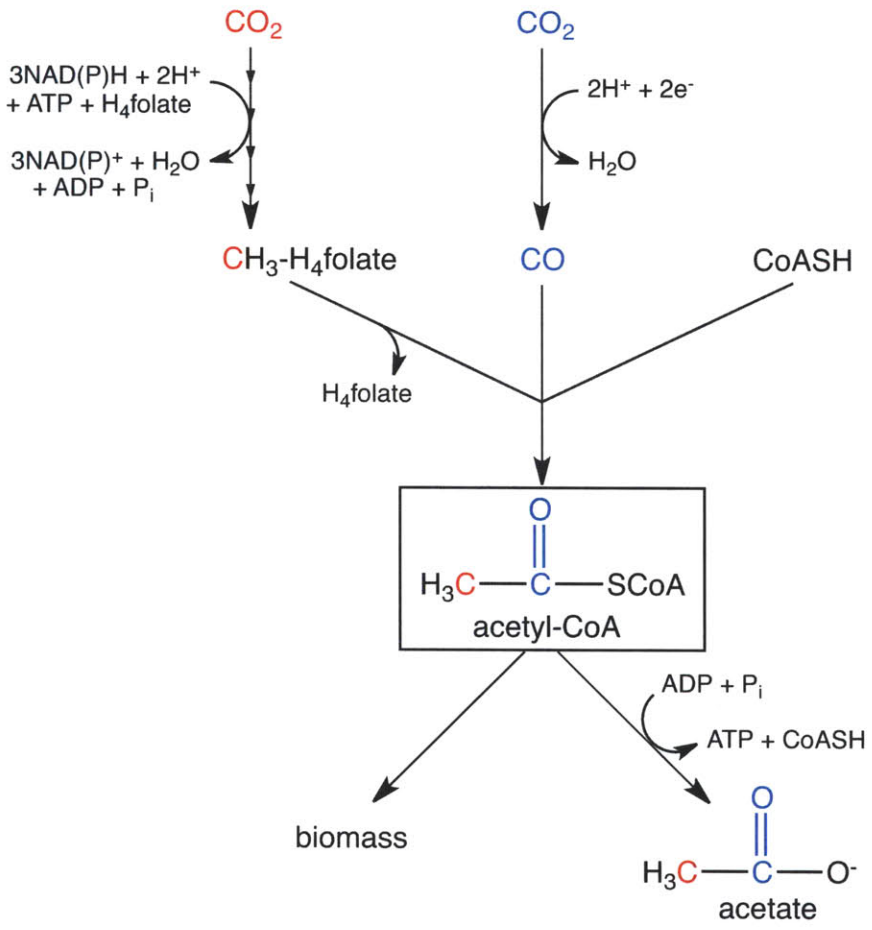


Figure I.4. The Wood-Ljungdahl carbon fixation pathway. In this non-cyclic carbon fixation pathway used by strict anaerobes, the carbon atoms from two molecules of CO_2 (blue and red) become the carbon atoms of the acetyl-CoA acetyl moiety. Multiple steps are indicated by arrows with multiple arrowheads.

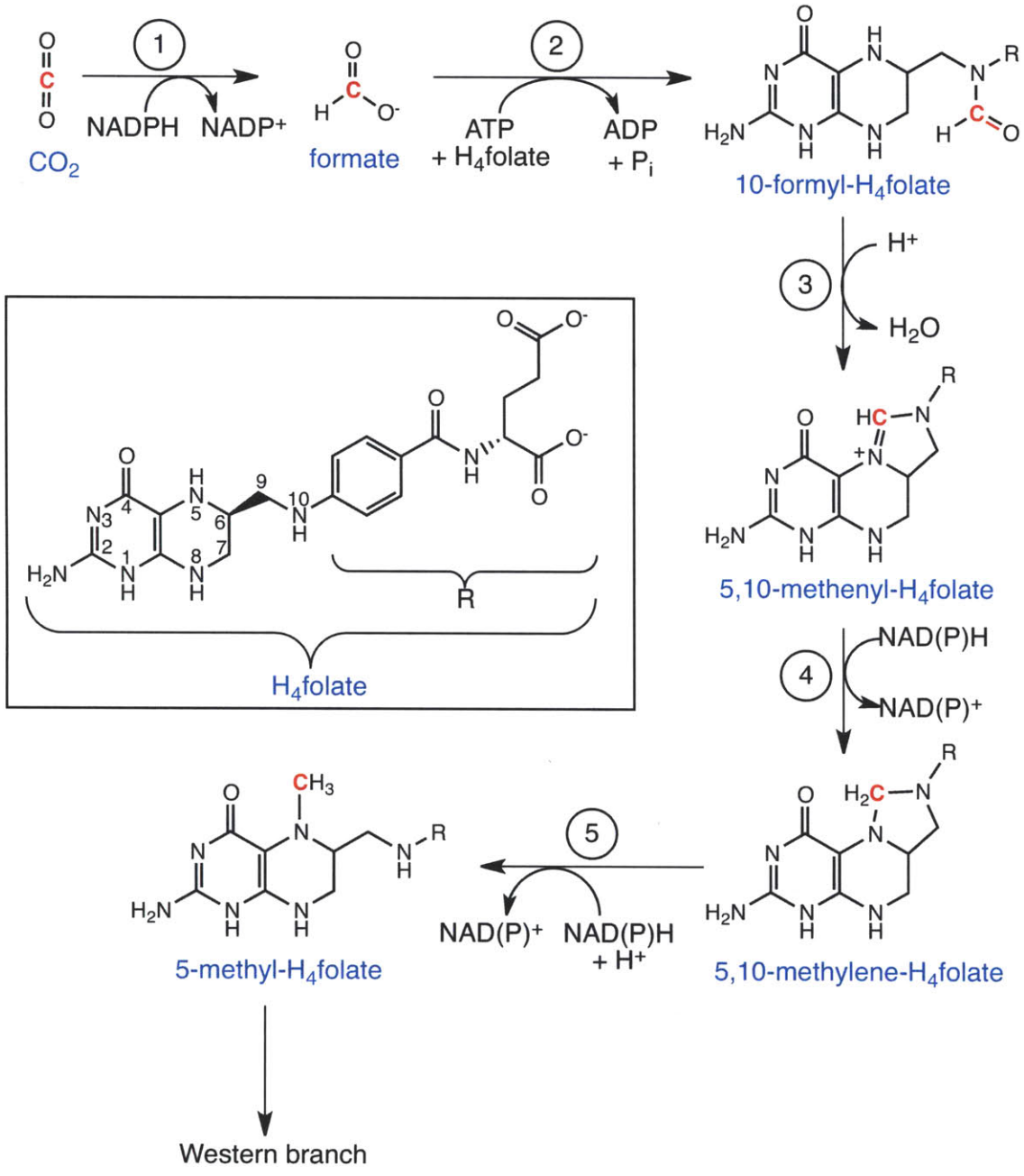


Figure I.5. The Eastern branch of the Wood-Ljungdahl carbon fixation pathway. Intermediates are labeled in blue, and the CO_2-derived carbon atom is colored in red. The full tetrahydrofolate (H_4folate) cofactor is shown inside the box. Steps are numbered 1-5 and are catalyzed by the following enzymes. 1: formate dehydrogenase, 2: 10-formyl-H_4folate synthetase, 3: 5,10-methenyl-H_4folate cyclohydrolase, 4: 5,10-methylene-H_4folate dehydrogenase, and 5: 5,10-methylene-H_4folate reductase.

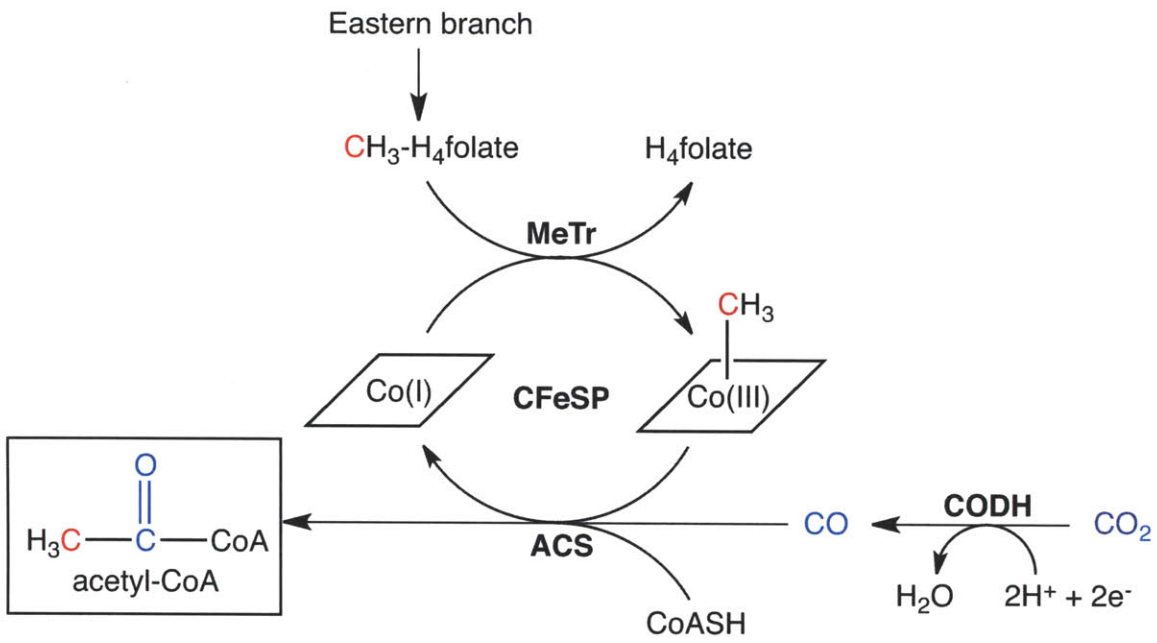


Figure I.6. The Western branch of the Wood-Ljungdahl carbon fixation pathway. The carbon atom derived from the Eastern branch is labeled in red, and the CO_2 carbon atom originating in the Western branch is labeled in blue. The four enzymes of the Western branch are labeled in bold: carbon monoxide dehydrogenase (CODH), acetyl-CoA synthase (ACS), the corrinoid iron-sulfur protein (CFeSP), and CH_3-H_4folate:CFeSP methyltransferase (MeTr). In acetogenic bacteria such as *M. thermoacetica*, CODH and ACS operate together as a bifunctional enzyme complex. The B_{12} cofactor harbored by CFeSP is shown as a rhombus with cobalt (Co) in the center, whose oxidation state is indicated in parentheses.

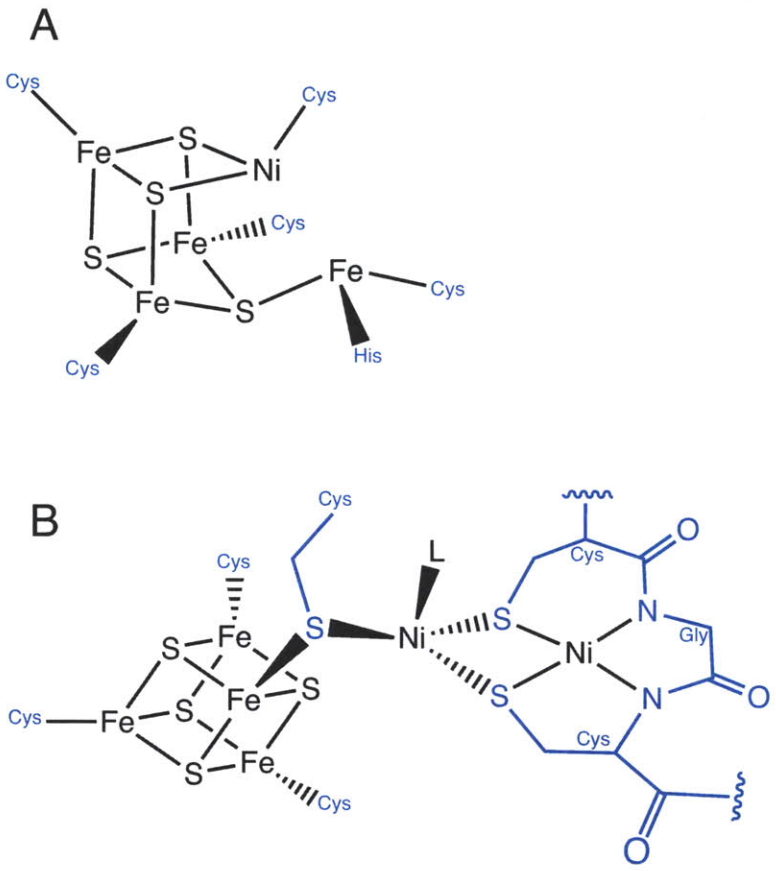


Figure I.7. The metalloclusters of CODH/ACS. (A) The [Ni-4Fe-4S] C-cluster of CODH without substrates bound. **(B)** The [4Fe-4S]-Ni-Ni A-cluster of ACS, where "L" refers to a putative substrate binding site and an unknown ligand seen in crystal structures of the A-cluster. Protein residues that ligate the metals of the metalloclusters are colored in blue.

I.G. References

1 Ellis, R. J. (1979) The Most Abundant Protein in the World. *Trends Biochem. Sci.* **44**, 241-244.

2 Siegenthaler, U. & Sarmiento, J. L. (1993) Atmospheric Carbon Dioxide and the Ocean. *Nature* **365**, 119-125.

3 Wächtershäuser, G. (1990) Evolution of the First Metabolic Cycles. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 200-204.

4 Evans, M. C. W., Buchanan, B. B. & Arnon, D. I. (1966) A New Ferredoxin-Dependent Carbon Reduction Cycle in a Photosynthetic Bacterium. *Proc. Natl. Acad. Sci. U.S.A.* **55**, 928-934.

5 Buchanan, B. B. & Arnon, D. I. (1990) A Reverse KREBS Cycle in Photosynthesis: Consensus At Last. *Photosynth. Res.* **24**, 47-52.

6 Berg, I. A., Kockelkorn, D., Ramos-Vera, W. H., Say, R. F., Zarzycki, J., Hügler, M., Alber, B. E. & Fuchs, G. (2010) Autotrophic Carbon Fixation in Archaea. *Nat. Rev. Microbiol.* **8**, 447-460.

7 Berg, I. A., Kockelkorn, D., Buckel, W. & Fuchs, G. (2007) A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea. *Science* **318**, 1782-1786.

8 Herter, S., Fuchs, G., Bacher, A. & Eisenreich, W. (2002) A Bicyclic Autotrophic CO_2 Fixation Pathway in *Chloroflexus aurantiacus*. *J. Biol. Chem.* **23**, 20277-20283.

9 Zarzycki, J., Brecht, V., Müller, M. & Fuchs, G. (2009) Identifying the Missing Steps of the Autotrophic 3-Hydroxypropionate CO_2 fixation Cycle in *Chloroflexus aurantiacus*. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21317-21322.

10 Huber, H., Gallenberger, M., Jahn, U., Eylert, E., Berg, I. A., Kockelkorn, D., Eisenreich, W. & Fuchs, G. (2008) A Dicarboxylate/4-Hydroxybutyrate Autotrophic Carbon Assimilation Cycle in the Hyperthermophile Archaeum *Ignicoccus hospitalis*. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7851-7856.

11 Ragsdale, S. W. & Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 Fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898.

12 Drake, H. L., Daniel, S. L., Matthies, C. & Küsel, K. (1994) "Acetogenesis, acetogenic bacteria, and the acetyl-CoA pathway: past and current perspectives" in *Acetogenesis*, Ed. H.L. Drake. Chapman and Hall: New York, 3-60.

13 Drake, H. L., Gößner, A. S. & Daniel, S. L. (2008) Old Acetogens, New Light. *Ann. N.Y. Acad. Sci.* **1125**, 100-128.

14 Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H. & Farrow, J. A. E. (1994) The Phylogeny and Genus *Clostridium*: Proposal of Five New Genera and Eleven New Species Combinations. *Int. J. Syst. Bacteriol.* **44**, 812-826.

15 Wieringa, K. T. (1936) Over het verdwijnen van waterstof en koolzuur onder anaerobe voorwaarden. *Antoinie Leeuwenhoek* **3**, 263-273.

16 Wieringa, K. T. (1940) The Formation of Acetic Acid from Carbon Dioxide and Hydrogen by Anaerobic Spore-Forming Bacteria. *Antoinie Leeuwenhoek* **6**, 251-262.

17 Lajoie, S. F., Bank, S., Miller, T. L. & Wolin, M. J. (1988) Acetate Production from Hydrogen and [^{13}C]Carbon Dioxide by the Microflora of Human Feces. *Appl. Environ. Microbiol.* **54**, 2723-2727.

18 Wolin, M. J. & Miller, T. L. (1994) "Acetogenesis from CO_2 in the Human Colonic Ecosystem" in *Acetogenesis*, Ed. Harold L. Drake. Chapman and Hall: New York, 365-385.

19 Breznak, J. A. & Kane, M. D. (1990) Microbial H_2/CO_2 Acetogenesis in Animal Guts: Nature and Nutritional Significance. *FEMS Microbiol. Lett.* **87**, 309-313.

20 Thauer, R. K. (1998) Biochemistry of Methanogenesis: A Tribute to Marjory Stephenson. *Microbiology.* **144**, 2377-2406.

21 Ferry, J. G. (1992) Biochemistry of Methanogenesis. *Crit. Rev. Biochem. Mol. Biol.* **26**, 473-503.

22 Thauer, R. K., Kaster, A.-K., Seedorf, H., Buckel, W. & Hedderich, R. (2008) Methanogenic Archaea: Ecologically Relevant Differences in Energy Conservation. *Nat. Rev. Microbiol.* **6**, 579-591.

23 Ferry, J. G. (1995) CO Dehydrogenase. *Annu. Rev. Biochem.* **49**, 305-333.

24 Drake, H. L., Hu, S.-I. & Wood, H. G. (1981) Purification of Five Components from *Clostridium thermoaceticum* Which Catalyze Synthesis of Acetate from Pyruvate and Methyltetrahydrofolate: Properties of Phosphotransacetylase. *J. Biol. Chem.* **256**, 11137-11144.

25 Das, A. & Ljungdahl, L. G. (1997) Composition and Primary Structure of the F_1F_0 ATP Synthase from the Obligately Anaerobic Bacterium *Clostridium thermoaceticum*. *J. Bacteriol.* **179**, 3746-3755.

26 Das, A., Hugenholtz, J., Halbeek, H. v. & Ljungdahl, L. G. (1989) Structure and Function of a Menaquinone Involved in Electron Transport in Membranes of *Clostridium thermoautotrophicum* and *Clostridium thermoaceticum*. *J. Bacteriol.* **171**, 5823-5829.

27 Hugenholtz, J. & Ljungdahl, L. G. (1989) Electron Transport and Electrochemical Proton Gradient in Membrane Vesicles of *Clostridium thermoautotrophicum*. *J. Bacteriol.* **171**, 2873-2875.

28 Hugenholtz, J., Ivey, D. M. & Ljungdahl, L. G. (1987) Carbon Monoxide-Driven Electron Transport in *Clostridium thermoautotrophicum* Membranes. *J. Bacteriol.* **169**, 5845-5847.

29 Gottwald, M., Andreesen, J. R., LeGall, J. & Ljungdahl, L. G. (1975) Presence of Cytochrome and Menaquinone in *Clostridium formiaceticum* and *Clostridium thermoaceticum*. *J. Bacteriol.* **122**, 325-328.

30 Ragsdale, S. W. (1997) The Eastern and Western Branches of the Wood/Ljungdahl Pathway: How the East and West Were Won. *Biofactors.* **6**, 3-11.

31 Ragsdale, S. W. (2008) Enzymology of the Wood-Ljungdahl Pathway of Acetogenesis. *Ann. N.Y. Acad. Sci.* **1125**, 129-136.

32 Matthews, R. G. & Drummond, J. T. (1990) Providing One-Carbon Units for Biological Methylations: Mechanistic Studies on Serine Hydroxymethyltransferase, Methylenetetrahydrofolate Reductase, and Methyltetrahydrofolate-Homocysteine Methyltransferase. *Chem. Rev.* **90**, 1275-1290.

33 DiMarco, A. A., Bobik, T. A. & Wolfe, R. S. (1990) Unusual Coenzymes of Methanogenesis. *Annu. Rev. Biochem.* **59**, 355-394.

34 Ljungdahl, L. G. & Andreesen, J. R. (1975) Tungsten, A Component of Active Formate Dehydrogenase from *Clostridium thermoaceticum*. *FEBS Lett.* **54**, 279-282.

35 Yamamoto, I., Saiki, T., Liu, S.-M. & Ljungdahl, L. G. (1983) Purification and Properties of NADP-dependent Formate Dehydrogenase from *Clostridium thermoaceticum*, a Tungsten-Selenium-Iron Protein. *J. Biol. Chem.*, 1826-1832.

36 Raaijmakers, H. C. A. & Romão, M. J. (2006) Formate-Reduced *E. coli* Formate Dehydrogenase H: The Reinterpretation of the Crystal Structure Suggests a New Reaction Mechanism. *J. Biol. Inorg. Chem.* **11**, 849-854.

37 Boyington, J. C., Gladyshev, V. N., Khangulov, S. V., Stadtman, T. C. & Sun, P. D. (1997) Crystal Structure of Formate Dehydrogenase H: Catalysis Involving Mo, Molybdopterin, Selenocysteine, and an Fe_4S_4 Cluster. *Science.* **275**, 1305-1308.

38 Mejillano, M. R., Jahansouz, H., Matsunaga, T. O., Kenyon, G. L. & Himes, R. H. (1989) Formation and Utilization of Formyl Phosphate by N^{10}-

Formytetrahydrofolate Synthetase: Evidence for Formyl Phosphate as an Intermediate in the Reaction. *Biochemistry.* **28**, 5136-5145.

39 Radfar, R., Shin, R., Sheldrick, G. M., Minor, W., Lovell, C. R., Odom, J. D., Dunlap, R. B. & Lebioda, L. (2000) The Crystal Structure of N^{10}-Formyltetrahydrofolate Synthetase from *Moorella thermoacetica. Biochemistry.* **39**, 3920-3926.

40 Lovell, C. R., Przybyla, A. & Ljungdahl, L. G. (1988) Cloning and Expression in *Escherichia coli* of the *Clostridium thermoaceticum* gene encoding thermostable formyltetrahydrofolate synthetase. *Arch. Microbiol.* **149**, 280-285.

41 Clark, J. E. & Ljungdahl, L. G. (1982) Purification and Properties of 5,10-Methenyltetrahydrofolate Cyclohydrolase from *Clostridium formicoaceticum. J. Biol. Chem.* **257**, 3833-3836.

42 O'Brien, W. E., Brewer, J. M. & Ljungdahl, L. G. (1973) Purification and Characterization of Thermostable 5,10-Methylenetetrahydrofolate Dehydrogenase from *Clostridium thermoaceticum. J. Biol. Chem.* **248**, 403-408.

43 Moore, M. R., O'Brien, W. E. & Ljungdahl, L. G. (1974) Purification and Characterization of Nicotinamide Adenine Dinucleotide-dependent Methylenetetrahydrofolate Dehydrogenase from *Clostridium formicoaceticum. J. Biol. Chem.* **249**, 5250-5253.

44 Clark, J. E. & Ljungdahl, L. G. (1984) Purification and Properties of 5,10-Methylenetetrahydrofolate Reductase, an Iron-Sulfur Flavoprotein from *Clostridium formicoaceticum. J. Biol. Chem.* **259**, 10845-10849.

45 Matthews, R. G. (2001) Cobalamin-Dependent Methyltransferases. *Acc. Chem. Res.* **34**, 681-689.

46 Menon, S. & Ragsdale, S. W. (1998) Role of the [4Fe-4S] Cluster in Reductive Activation of the Cobalt Center of the Corrinoid Iron-Sulfur Protein from *Clostridium thermoaceticum* during Acetate Biosynthesis. *Biochemistry.* **37**, 5689-5698.

47 Menon, S. & Ragsdale, S. W. (1999) The Role of an Iron-Sulfur Cluster in an Enzymatic Methylation Reaction. *J. Biol. Chem.* **274**, 11513-11518.

48 Dobbek, H., Gremer, L., Kiefersauer, R., Huber, R. & Meyer, O. (2002) Catalysis at a Dinuclear [CuSMo(=O)OH] Cluster in a CO Dehydrogenase Resolved at 1.1-Å Resolution. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15971-15976.

49 Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002) A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Science.* **298**, 567-572.

50 Hu, Z., Spangler, H. J., Andersen, M. E., Xia, J., Ludden, P. W., Lindahl, P. A. & Münck, E. (1996) Nature of the C-Cluster in Ni-Containing Carbon Monoxide Dehydrogenases. *J. Am. Chem. Soc.* **118**, 830-845.

51 Ragsdale, S. W. & Kumar, M. (1996) Nickel-Containing Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Chem. Rev.* **96**, 2515-2539.

52 Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, O. (2001) Crystal Structure of a Carbon Monoxide Dehydrogenase Reveals a [Ni-4Fe-5S] Cluster. *Science.* **293**, 1281-1285.

53 Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. & Ludden, P. W. (2001) Life on Carbon Monoxide: X-ray Structure of *Rhodospirillum rubrum* Ni-Fe-S Carbon Monoxide Dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11973-11978.

54 Kung, Y. & Drennan, C. L. (2010) A Role for Nickel-Iron Cofactors in Biological Carbon Monoxide and Carbon Dioxide Utilization. *Curr. Opin. Chem. Biol* **15**.

55 Uffen, R. L. (1976) Anaerobic Growth of a *Rhodopseudomonas* Species in the Dark with Carbon Monoxide as Sole Carbon and Energy Substrate. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3298-3302.

56 Svetlichny, V. A., Sokolova, T. G., Gerhardt, M., Ringpfeil, M., Kostrikina, N. A. & Zavarzin, G. A. (1991) *Carboxydothermus hydrogenoformans* gen. nov., sp. nov., a CO-Utilizing Thermophilic Anaerobic Bacterium from Hydrothermal Environments of Kunashir Island. *Syst. Appl. Microbiol.* **14**, 254-260.

57 Bartholomew, G. W. & Alexander, M. (1979) Microbial Metabolism of Carbon Monoxide in Culture and in Soil. *Appl. Environ. Microbiol.* **37**, 932-937.

58 Grahame, D. A. (1991) Catalysis of Acetyl-CoA Cleavage and Tetrahydrosarcinapterin Methylation by a Carbon Monoxide Dehydrogenase-Corrinoid Enzyme Complex. *J. Biol. Chem.* **266**, 22227-22233.

59 Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003) Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] Clusters in Closed and Open Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. *Nat. Struct. Biol.* **10**, 271-279.

60 Doukov, T. I., Blasiak, L. C., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2008) Xenon in and at the End of the Tunnel of Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **47**, 3474-3483, doi:10.1021/bi702386t.

61 Maynard, E. L. & Lindahl, P. A. (1999) Evidence of a Molecular Tunnel Connecting the Active Sites for CO_2 Reduction and Acetyl-CoA Synthesis in

Acetyl-CoA Synthase from *Clostridium thermoaceticum*. *J. Am. Chem. Soc.* **121**, 9221-9222.

62 Seravalli, J. & Ragsdale, S. W. (2000) Channeling of Carbon Monoxide during Anaerobic Carbon Dioxide Formation. *Biochemistry.* **39**, 1274-1277.

63 Svetlitchnyi, V., Dobbek, H., Meyer-Klaucke, W., Meins, T., Thiele, B., Römer, P., Huber, R. & Meyer, O. (2004) A Functional Ni-Ni-[4Fe-4S] Cluster in the Monomeric Acetyl-CoA Synthase from *Carboxydothermus hydrogenoformans*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 446-451.

64 Seravalli, J., Xiao, Y., Gu, W., Cramer, S. P., Antholine, W. E., Krymov, V., Gerfen, G. J. & Ragsdale, S. W. (2004) Evidence that NiNi Acetyl-CoA Synthase Is Active and That the CuNi Enzyme Is Not. *Biochemistry.* **43**, 3944-3955.

65 Hegg, E. L. (2004) Unraveling the Structure and Mechanism of Acetyl-Coenzyme A Synthase. *Acc. Chem. Res.* **37**, 775-783.

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

67 Jeoung, J.-H. & Dobbek, H. (2007) Carbon Dioxide Activation at the Ni,Fe-Cluster of Anaerobic Carbon Monoxide Dehydrogease. *Science.* **318**, 1461-1464.

68 Gong, W., Hao, B., Wei, Z., Ferguson, D. J. J., Tallant, T., Krzycki, J. A. & Chan, M. K. (2008) Structure of the $\alpha_2\varepsilon_2$ Ni-Dependent CO Dehydrogenase Component of the *Methanosarcina barkeri* Acetyl-CoA Decarbonylase/Synthase Complex. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9558-9563.

Chapter II: Crystal structure of the acetogenic CFeSP/MeTr complex

II.A. Summary

The organometallic cofactor B_{12}, which has been called "nature's most beautiful cofactor"[1], is used in methyl group transfer from one molecule to another in biological processes as diverse as methionine synthesis in humans and CO_2 fixation in acetogenic bacteria[2-4]. This seemingly straightforward reaction requires large, multimodular enzyme complexes that adopt multiple conformations to alternately activate, protect, and perform catalysis on the reactive B_{12} cofactor. Crystal structures determined thus far have provided structural information for only fragments of these complexes[5-13], inspiring speculation regarding the overall protein assembly and conformational movements inherent to activity. Here, we present the 2.38 Å X-ray crystal structure of a complete ~220 kDa complex that contains all of the enzymes responsible for B_{12}-dependent methyl transfer, namely the corrinoid iron-sulfur protein (CFeSP) and its methyltransferase (MeTr) from the model acetogen *Moorella thermoacetica*. This structure depicts for the first time the three-dimensional construction of all protein modules required for the activation, protection, and catalytic steps of B_{12}-dependent methyl transfer. Furthermore, the structure is presented alongside *in crystallo* UV-Vis spectroscopic data, which confirm enzymatic activity in crystals and demonstrate the largest known conformational movement of proteins in a crystalline state.

A manuscript similar to this chapter has been submitted for publication as:
Kung, Y., Ando, N., Doukov, T.I., Blasiak, L.C., Bender, G., Seravalli, J., Ragsdale, S.W., and Drennan, C.L. Visualising molecular juggling within a B_{12}-dependent methyltransferase complex.

II.B. Acknowledgements

The microspectrophotometer used was constructed by Nozomi Ando, with whom the *in crystallo* data were collected. Tzanko I. Doukov obtained initial crystallization hits and collected the lower resolution X-ray diffraction data. Leah C. Blasiak processed, solved, and refined the structure using these initial data. Javier Seravalli purified the proteins used in the initial crystallization, while Güneş Bender, with the aid of Joseph E. Darty, purified the proteins used in the subsequent crystallizations that led to the final structure, working in the laboratory of Stephen W. Ragsdale at the University of Michigan.

This work was supported by NIH grants GM69857 (C.L.D.) and GM39451 (S.W.R.) and the MIT Energy Initiative (C.L.D.). C.L.D. is a Howard Hughes Medical Institute Investigator. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

II.C. Introduction

B_{12}-dependent methyl transfer lies at the heart of methylation biochemistry and is an essential reaction in human health and microbial CO_2 sequestration[3,4]. In humans, the 140 kDa enzyme methionine synthase (MetH) methylates homocysteine to form methionine to maintain cellular pools of folic acid and *S*-adenosylmethionine (AdoMet), the universal methyl donor. MetH mutation or vitamin B_{12} deficiency leads to a build-up of the MetH substrate, the folic acid derivative methyltetrahydrofolate (CH_3-H_4folate), preventing the product, tetrahydrofolate (H_4folate), from re-entering the folate pool where it may be used for other biological processes, such as nucleotide biosynthesis. This so-called "methyl trap" can thus cause serious health consequences, including megaloblastic anemia and birth abnormalities such as neural tube defects[14]. Acetogenic bacteria, including *M. thermoacetica*, use B_{12}-containing CFeSP and MeTr together to catalyze a similar methyl transfer reaction in the Wood-Ljungdahl carbon fixation pathway for growth on CO_2 as the sole carbon source[15].

For both MetH and CFeSP/MeTr methyl transfer systems, CH_3-H_4folate is the methyl donor, and protein-bound B_{12} is the methyl carrier. In acetogenic bacteria, the CH_3-H_4folate methyl group is derived from enzymatic reduction of CO_2, while in humans, the CH_3-H_4folate methyl group is derived from the β-carbon of serine by the action of serine hydroxymethyltransferase[16]. CH_3-H_4folate also is the form of folic acid that is circulated in the blood; upon entering a cell it must be demethylated by MetH first to free folate for further methionine synthesis or for other cellular purposes[17-19]. Although CH_3-H_4folate is a common methyl source, methyl removal from the N^5 tertiary amine is chemically challenging because the product, H_4folate, is a poor leaving group[2,3].

Therefore, a particularly powerful nucleophile is required, and B_{12} with cobalt in the +1 oxidation state, a Co(I) species dubbed a "supernucleophile"[20], is recruited. Such strong reactivity comes at a price: reduction of the as-isolated, inactive Co(II) state to generate active Co(I) is thermodynamically challenging, as the B_{12} Co(II/I) reduction potential is one of the lowest in biology, -504 mV in CFeSP and -526 mV in MetH[21,22]. For MetH, reduction is driven by the favorable cleavage of AdoMet and reductive methylation to CH_3-Co(III), which enters the reaction cycle. In CFeSP, an electron is first delivered from a partner protein to an Fe_4S_4 cluster harbored by an activation domain[23,24]. The electron is then passed to Co(II) to yield Co(I) (Eq. 1), which may then attack CH_3-H_4folate to form CH_3-Co(III) (Eq. 2). CFeSP then delivers the methyl group to the $Ni_2Fe_4S_4$ active site metallocluster (A-cluster) of acetyl coenzyme A synthase (ACS) where it becomes the methyl group of acetyl-CoA. MetH delivers its methyl group from CH_3-Co(III) to homocysteine to produce methionine. Upon methyl delivery, the B_{12} cobalt returns to the nucleophilic Co(I) species.

$$[4Fe\text{-}4S]^{1+} + Co(II) \rightleftharpoons [4Fe\text{-}4S]^{2+} + Co(I) \qquad \text{(Eq. 1)}$$

$$Co(I) + CH_3\text{-}H_4folate \rightleftharpoons CH_3\text{-}Co(III) + H_4folate \qquad \text{(Eq. 2)}$$

During the catalytic cycle (Figure II.1), B_{12} performs a series of molecular juggling acts, binding multiple sites in distinct domains. B_{12} bound in one domain first undergoes activation by interacting with a second domain before catalyzing two methyl transfer reactions in active sites supplied by additional domains while intermittently protecting reactive intermediates Co(I) and CH_3-Co(III) from unwanted side reactions using yet another domain. Crystal structures of a MetH B_{12}-binding fragment[10] and

CFeSP from *Carboxydothermus hydrogenoformans* (*Ch*CFeSP)[13] both depict a "resting" state where B_{12} is buried by a protective "capping" domain, shielded from unwanted chemistry but inaccessible to substrate. Because methyl transfer employs an S_N2 mechanism[25], large conformational changes must "uncap" B_{12} before chemistry can occur. B_{12} is "uncapped" in structures of MetH fragments that exhibit B_{12} activation[5,6,12], but no structure has been solved that shows B_{12}- and CH_3-H_4folate-binding domains together to illustrate methyl transfer. Although structures are available for all MetH domains[5-7,10-12], no structure exists of the intact enzyme to describe how the entire machinery is constructed. In the CFeSP/MeTr system, structures of MeTr from *M. thermoacetica* are available[8,9], but no structure of CFeSP from this organism has been determined. Furthermore, the CFeSP Fe_4S_4 domain was disordered in the *Ch*CFeSP structure[13] and has thus never been visualized. Overall, the nature of the conformational movements intrinsic to B_{12}-dependent methyl transfer has remained elusive, as the relative orientation of all domains in any system has not been observed until now.

II.D. Results

Overall structure

The structure of the CFeSP/MeTr complex from *M. thermoacetica* was solved by molecular replacement using the individual structures of MeTr[9] and *Ch*CFeSP[13] as independent search models (Table II.1). In the 2.38 Å resolution structure (Figure II.2), the individual MeTr and CFeSP components closely resemble the structures of these proteins alone. The MeTr component, a 57 kDa homodimer, is virtually identical to prior structures of both MeTr[8,9], rms deviation for C_α atoms (rmsd): 0.39 Å (Figure II.3.A), and

the analogous MetH domain that binds CH_3-H_4folate[11], rmsd: 1.03-1.08 Å. MeTr and MetH use $(\beta/\alpha)_8$ triosephosphate isomerase (TIM) barrels to bind and activate CH_3-H_4folate for nucleophilic attack. Two CFeSPs are present in the complex, each containing two subunits in a heterodimer. The small subunit is a 35 kDa TIM barrel which acts as the B_{12} "cap" in the *Ch*CFeSP structure, while the large subunit (48 kDa) has three domains joined by linkers: an N-terminal Fe_4S_4 activation domain (residues 1-57), a central TIM barrel domain (residues 93-312), and a C-terminal B_{12}-binding domain (residues 325-446). With the exception of the Fe_4S_4 and B_{12} domains, discussed below, both CFeSP copies align well to the *Ch*CFeSP structure, rmsd: 0.81-0.85 Å, with small and large subunits bound together in the same fashion (Figure II.3.B).

In the overall ~220 kDa $\alpha_2\beta_2\gamma_2$ CFeSP/MeTr assembly, the MeTr homodimer lies in the center, bound to one CFeSP heterodimer on either side (Figure II.2). Each MeTr monomer has a C-terminal α-helix (residues 255-262) protruding from the rim of the TIM barrel. Contacts between this helix and its preceding loop with a helix from the CFeSP small subunit (residues 191-204) form the primary interactions between MeTr and CFeSP (Figure II.4). Weaker interactions between MeTr and CFeSP Fe_4S_4 domains (Figure II.5) appear to have sufficiently stabilized the Fe_4S_4 domains, allowing their visualization for the first time, as described below.

Taken as a whole, the MeTr homodimer, CFeSP small subunits, and CFeSP large subunit central domains form a structural scaffold of six TIM barrels (Figure II.2). Here, the MeTr homodimer and two CFeSP small subunits, one on each side, create a set of four coplanar TIM barrels. The CFeSP large subunit central TIM barrel domains rest on top of this plane, bound to the CFeSP small subunits as described above. Only the N-

terminal Fe_4S_4 and C-terminal B_{12} domains of the CFeSP large subunit are not TIM barrels, and these are connected to the CFeSP large subunit central domain by long linkers described below.

The CFeSP Fe_4S_4 domain

Our data allow for the visualization of the N-terminal Fe_4S_4 domain of the CFeSP large subunit for the first time (Figures II.2 and II.6). The Fe_4S_4 domains are bundles of three short α-helices connected by long loops that coordinate the Fe_4S_4 cubane. The four residues that ligate the cluster are Cys17, Cys20, Cys25, and Cys42, where the first three cysteines are on the same loop and constitute a $C\text{-}X_2\text{-}C\text{-}X_4\text{-}C$ motif. Although this motif is seen in other proteins that contain Fe_4S_4 clusters involved in electron transfer, such as carbon monoxide dehydrogenase (CODH), ferredoxin, and hydrogenase, the fold of the Fe_4S_4 domain appears to be novel, as a DALI[26] search against structures in the Protein Data Bank (PDB) yielded no significant matches. The presence of the Fe_4S_4 domain is not conserved across B_{12}-dependent methyl transfer systems, and, as mentioned, the mechanism for B_{12} activation is different in MetH.

The Fe_4S_4 domain protrudes from the CFeSP large subunit central domain, placing the cluster ~20 Å away from the TIM barrel core. The Fe_4S_4 domain is connected to the large subunit TIM barrel by a 35-residue linker (Figure II.7.), a primarily unstructured coil with a conserved 14-residue β-hairpin[13]. Although the Fe_4S_4 domain was disordered in the *Ch*CFeSP structure, Fe anomalous data allowed identification of the cluster's position, giving a long Fe_4S_4-Co distance of 52 Å[13]. The Fe_4S_4 clusters in

CFeSP/MeTr are also far from B_{12} cobalt atoms: intra-CFeSP Fe_4S_4-Co distances are 54 Å and 55 Å, and inter-CFeSP Fe_4S_4-Co distances are 39 Å and 41 Å (Figure II.7).

The position and flexibility of the B_{12} domains

B_{12} domains of both MetH and CFeSP use Rossmann-like folds to bind B_{12} in the base-off conformation[10,13]. MetH binds B_{12} with a conserved axial histidine ligand coordinating the cobalt center[10] (His759 in *Escherichia coli* MetH), whereas CFeSP lacks a protein ligand to bind the B_{12} cobalt[13,22,27-30]. The lack of axial ligation by either the B_{12} base or a protein residue in CFeSP has been suggested to tune the cobalt reduction potential, stabilizing Co(I) by approximately +130 mV to allow for more facile Co(II) reduction by the Fe_4S_4 cluster[22].

In the CFeSP/MeTr complex, each B_{12} domain sits between the TIM barrels of a MeTr monomer and a CFeSP small subunit, which are adjacent and nearly perpendicular to each other (Figures II.2 and II.8-10). Both B_{12} domains in the complex lie in the same conformation, as an alignment of the CFeSPs superimposes their B_{12} cofactors (Figure II.11). However, when aligned with the "resting" structure of *Ch*CFeSP, the B_{12} domains have moved away from the "capping" CFeSP small subunit, shifting the Co centre by 6.9 Å such that the two hydrogen bonds between the corrin ring and the small subunit are replaced by one hydrogen bond to MeTr and one new hydrogen bond to the small subunit (Figures II.8.A and II.8.B).

Although the structure does not contain CH_3-H_4folate, its binding site can be identified based on an alignment with the structure of CH_3-H_4folate-bound MeTr[9] (Figure II.9). This superposition illustrates that B_{12} in the CFeSP/MeTr structure is positioned "en

route" to the active site, having moved 6.9 Å closer, but with ~18 Å to go, to be in position for methyl transfer. An unstructured 12-residue linker connects the B_{12} domain to the central TIM barrel domain (Figures II.7 and II.9) and could allow the B_{12} domain to swing from its "resting" position to a catalytic position over MeTr. Flexibility of the B_{12} domain is also evident in its high *B*-factors, a measure of thermal motion within the crystal, which are about twice those for the rest of the structure (Figure II.12 and Table II.2). In addition, electron density for the B_{12} domain is weaker than for the rest of the structure (comparison of electron density for all domains in Figures II.6 and II.13-18). The *B*-factors and electron density of the B_{12} domains indicate that they are flexible and mobile within the crystal, existing as an ensemble of conformations of which only the conformer of greatest population is observed.

In crystallo *activity of the CFeSP/MeTr complex*

Given such flexibility, the possibility that the B_{12} domain can sample the full ~18 Å to afford turnover within intact crystals was explored using anaerobic *in crystallo* UV-Vis spectroscopy to monitor absorption spectra over the course of the reaction as cobalt progresses through its oxidation states. *In crystallo* spectra were obtained in parallel to analogous solution spectra (Figure II.19), which match well-established CFeSP UV-Vis absorption features[23,24,27,31,32].

The spectrum of as-isolated CFeSP/MeTr crystals strongly resembles that of CFeSP in solution (Figure II.19, blue lines), with broad features at ~400 nm and ~470 nm arising from the Fe_4S_4 cluster and B_{12} cofactor, suggesting that no significant change occurred during crystallization. Following established protocols[23,24,28,30-32], B_{12} reduction

was achieved chemically by treatment with titanium(III) citrate, yielding a sharp 390 nm peak indicative of four-coordinate Co(I) in both solution and *in crystallo* spectra (Figure II.19, black lines). Further treatment of crystals with CH_3-H_4folate yields a decrease in absorbance at 390 nm and a new peak at 450 nm (Figure II.19, red lines), characteristic of protein-bound, five-coordinate CH_3-Co(III)[23,24,31,32], indicating that turnover has occurred.

To ensure that methyl transfer did not occur from free CH_3-H_4folate unbound to MeTr, a control reaction was performed (Figure II.19, green line) and confirms that turnover does not occur in the absence of MeTr. The spectra also demonstrate that B_{12} remains CFeSP-bound both in the crystal and in solution, as the observed CH_3-Co(III) 450 nm peak corresponds to protein-bound B_{12}, while free, unbound B_{12} has its absorption peak shifted to ~520 nm[23,27,28]. Absorption features disappear when light is passed through the solution surrounding the crystals, indicating that spectra represent protein in crystals and not protein that may have been liberated into the solution.

Taken together, the solution and *in crystallo* data confirm that the CH_3-H_4folate methyl group is enzymatically transferred to the CFeSP-bound B_{12} cofactor within the crystal, providing evidence that the B_{12} domain, which makes no lattice contacts (Figure II.20), is able to move the B_{12} cofactor at least ~18 Å within the crystal to trigger methyl transfer.

II.E. Discussion

Methylation chemistry is ubiquitous and important in all living organisms, from transcriptional regulation by DNA and histone methylation, to amino acid and nucleic

acid biosynthesis, to the ability of many microorganisms, including acetogens, methanogens, and methanotrophs, to live on CO_2 and other one-carbon compounds. While AdoMet is often associated with methyl transfer, its methyl group originally derives from CH_3-H_4folate through the action of methionine synthase. In addition, folate derivatives that are one-carbon donors in nucleic acid biosynthesis, such as 5-formyl-H_4folate and 5,10-methylene-H_4folate, are produced from H_4folate pools that must be regenerated by a methionine synthase. Therefore, methyl transfer from CH_3-H_4folate lies at the heart of methylation biochemistry.

With the first crystal structure of an enzyme complex that contains the CH_3-H_4folate domain and the B_{12} domain, in addition to the capping and activation modules, we can now consider why the transfer of a single methyl group requires such an elaborate protein assembly, and how B_{12} is alternately protected from unwanted chemistry by a capping domain, and then made available to the Fe_4S_4 and CH_3-H_4folate domains for electron transfer or catalysis. In the ensuing discussion, we examine the molecular juggling acts associated with each step of methyl transfer.

The use of the B_{12} Co(I) state by B_{12}-dependent methyltransferases presents a challenge in generating a cofactor such a low reduction potential: how can Co(I) be produced and protected? In MetH, activation of Co(II) is accomplished by reductive methylation from Co(II) to CH_3-Co(III), using an electron, derived from flavodoxin in *E. coli* or methionine synthase reductase in humans, and a methyl group from AdoMet. Because AdoMet is itself made from methionine, if it were used as the methyl donor for each turnover, no net production of methionine would be achieved, giving rise to a futile cycle. To achieve constructive methionine synthesis, the AdoMet domain must be

prevented from accessing B_{12} during the catalytic cycle, allowing methylation only by CH_3-H_4folate. The enzyme appears to accomplish this task by keeping the AdoMet domain at a distance, requiring a large conformational change for AdoMet to react with B_{12}[5,6,12]. Because CFeSP does not use AdoMet in Co(I) generation, futile cycling is not an issue, and a simpler strategy of B_{12} activation is employed. *In vivo*, an electron is first delivered to the protein-bound Fe_4S_4 cluster from a separate protein, such as CODH, ferredoxin, or pyruvate:ferredoxin oxidoreductase (PFOR)[24], and is subsequently passed to B_{12} in the Co(II) state by thermodynamically favorable electron transfer (reduction potentials of -523 mV for the $[4Fe\text{-}4S]^{2+/+}$ cluster[22] and -504 mV for Co(II/I)).

Although there is a rationale for a MetH design where the AdoMet and B_{12} domains are held far apart, there is no such logic that would require a long distance between the Fe_4S_4 and B_{12} cofactors in CFeSP, and yet both *Ch*CFeSP and CFeSP/MeTr structures show distances between the cluster and the B_{12} cobalt of >39 Å (Figure II.7), whereas <14 Å is expected for facile electron transfer[33]. It is interesting that the Fe_4S_4 cluster of one CFeSP in the CFeSP/MeTr complex is closer to the B_{12} cobalt from the other CFeSP than its own B_{12} cobalt (for example, 39 Å instead of 55 Å); however, the distance is still too great for intermolecular electron transfer. Although the linker between the Fe_4S_4 and central TIM barrel domains is certainly long enough to enable a substantial conformational change that could bring the cluster near the B_{12} cobalt, it is unclear why a highly flexible and extended Fe_4S_4 domain would be used as the immediate source of electrons—a simpler strategy could have involved docking of the electron transfer protein directly to the B_{12} domain. The explanation may lie in an undescribed, additional role for the Fe_4S_4 domain and its flexible linker. For example, binding the Fe_4S_4 domain to MeTr

or ACS may help orient CFeSP with these partner proteins. Alternatively, an extended Fe_4S_4 cluster domain could also be involved in passing electrons to the ACS active site A-cluster prior to methyl transfer.

Following Co(I) generation, we propose the following model for the methyl transfer reaction from CH_3-H_4folate to B_{12} (Figure II.21), consistent with the CFeSP/MeTr structure. First, a mechanism is needed for protecting the highly reactive Co(I) species as well as the CH_3-Co(III) state. Prior to catalysis, CFeSP exists in a "resting" state, with B_{12} "capped" and protected by the small subunit[13] ("resting" state in Figure II.21 and shown in Figure II.22.A). The previous structure of the MetH B_{12} binding fragment also illustrates a capping domain that covers and protects the active face of the corrin ring[10] (Figure II.22.B.). From this conformation, however, either the "cap" or the B_{12} domain must move to allow substrate access. The CFeSP/MeTr structure reveals that the B_{12} domain moves away from the "cap" to adopt an ensemble of "en route" conformations while still partially protected by the MeTr barrel. The high *B*-factors of the B_{12} domain and the weaker density of this region of the structure suggest that upon MeTr binding, the B_{12} domain is freed to sample multiple conformations, with the highest-occupancy position being the location seen in the crystal structure, ~7 Å away from the "resting" position in the direction of the MeTr CH_3-H_4folate binding site ("En route" state in Figure II.21). From these data, the B_{12} domain appears to be the flexible component, with the MeTr and CFeSP small subunit TIM barrels shielding B_{12} from unwanted chemistry while in motion.

It is interesting to consider why binding of CFeSP to MeTr does not simply position the B_{12} domain over the active site for catalysis ("folate-on" state in Figure

II.21). One explanation is that the structure represents an inactive complex; however, *in crystallo* results demonstrate that CFeSP/MeTr crystals are active. An alternative explanation posits that an ensemble of flexible "en route" conformations exists when CH_3-H_4folate is absent, and that CH_3-H_4folate binding would shift their conformational equilibrium to favor the "folate-on" state. While binding of CFeSP to MeTr can occur irrespective of CH_3-H_4folate binding[32], the effect of apo-MeTr on CFeSP conformation was previously unknown. This structure indicates that binding of CFeSP to apo-MeTr "loosens" the B_{12} domain from its "resting" state, preparing it for catalysis without triggering movement directly to the "folate-on" conformation, which could block CH_3-H_4folate binding. Following methyl transfer, CH_3-Co(III) must also be protected, and the small subunit can "re-cap" B_{12} ("Product-bound" state in Figure II.21). The ordered movements of the B_{12} domain proposed in this model (Figure II.21) are consistent with MetH studies where ligation, alkylation, and redox state of the B_{12} cobalt can favor/disfavor certain binding modes, alternately shifting the equilibrium of conformers for ordered domain rearrangements during the reaction cycle[6,12,34,35]. This model is thus expected to be a common theme in folate-B_{12} methyl transfer.

Further, *in crystallo* data show how large conformational movements can occur within crystals, with the B_{12} domain swinging the cofactor at least ~18 Å to afford methyl transfer. Catalysis performed by crystallized enzymes is well documented[36]; in one notable case, crystallized DNA polymerase I catalyzed elongation of a DNA template by 6 nucleotides, a ~20 Å translocation of the substrate into a cavity within the lattice recognized as the largest measured molecular movement within a crystal lattice[37]. In addition, diffusion of molecules as large as proteins through wide solvent channels in

protein and virus crystals is often observed in soaking experiments[38]. While such studies illustrate sizeable movements of molecules through a protein lattice, to our knowledge no conformational movement of this magnitude of the crystallized protein itself has been measured until now. Such dramatic B_{12} domain movement is likely facilitated by the fact that a majority of the proteins are composed of rigid TIM barrels which provide all of the lattice contacts. B_{12} domain motion is also related to the issue of the large size of B_{12}-dependent methyl transfer complexes: such large protein complexes may be needed to transfer a single methyl group because a stable scaffold is required to maintain structural integrity during the conformational juggling that enables activation, protection, and catalysis of the highly reactive B_{12} cofactor. A further discussion of this concept, focusing on the implications for MetH and other B_{12}-dependent methyl transfer systems, is provided in Chapter VI.

II.F. Materials and Methods

Protein purification and crystallization

CFeSP was expressed and purified anaerobically from *M. thermoacetica* ATCC 39073 as described[27], except for the following modifications. CFeSP was purified in an anaerobic chamber (Coy Laboratories) from cell extracts using DEAE-cellulose and high resolution Q-Sepharose anion exchange chromatography followed by phenyl-Sepharose hydrophobic interaction chromatography. Fractions containing CFeSP were concentrated and buffer exchanged using Amicon ultracentrifuge concentrators. MeTr was expressed and purified anaerobically from recombinant *E. coli* as described[8]. Concentrations of

CFeSP and MeTr protein samples were determined using the Rose-Bengal method[39] and kept in storage buffer (50 mM Tris-HCl, pH 7.6, 100 mM NaCl, 2 mM dithiothreitol).

Crystals of the CFeSP/MeTr complex were grown by hanging drop vapor diffusion in an anaerobic chamber at room temperature by adding 1 μL of precipitant solution (100 mM Bis-Tris, pH 6.5, 100 mM calcium acetate, 9% PEG 5000 monomethyl ether, 20% glycerol) to 2 μL of an equimolar mixture of CFeSP and MeTr (~250 μM monomer for each), over a 0.5 mL reservoir solution of precipitant solution. Large, brown, rod-shaped crystals appeared overnight. Crystals were looped and cryo-cooled in liquid nitrogen anaerobically prior to collection of X-ray diffraction data at 100 K.

Data collection, structure determination, and refinement

Two X-ray diffraction data sets were collected. A lower-resolution data set (3.3 Å) was collected at the Advanced Light Source (ALS) beam line 5.0.2, and a higher-resolution data set (2.38 Å) was later collected at the Advanced Photon Source (APS) beam line 24ID-C.

The initial data set to 3.3 Å resolution was processed in HKL2000 and Scalepack[40]. Crystals belonged to the space group $P2_12_12_1$ with cell dimensions (Å): a=137.42, b=159.87, and c=241.92. The structure was solved by molecular replacement in Phaser[41], using individual structures of MeTr[9] (PDB ID: 2E7F) and *Ch*CFeSP[13] (PDB ID: 2H9A) lacking its B_{12} domain as independent search models. Two CFeSP/MeTr complexes (~220 kDa each) were found in the asymmetric unit. Iterative rounds of refinement with residue-grouped *B*-factors were carried out in CNS[42] and PHENIX[43], with model building in Coot[44]. The four B_{12} domains present in the complex were kept as

a polyalanine model. Final R_{work} and R_{free} values for the model were 29.2% and 33.7%, respectively, when refinement of the structure to higher resolution began. Data collection and refinement statistics for this data set are shown in Table II.1. Ramachandran analysis was carried out in PROCHECK[45].

The data set to 2.38 Å resolution was processed in HKL2000 and Scalepack[40]. Although this crystal formed in similar conditions as the crystal which gave the 3.3 Å resolution data set, the space group was now $P2_12_12$, with unit cell dimensions (Å): a=125.71, b=242.84, and c=79.67. The structure for this crystal was thus solved by molecular replacement in Phaser[41] using the MeTr homodimer and CFeSP heterodimers lacking B_{12} domains from the previously refined model of the 3.3 Å resolution structure as independent search models. Only one CFeSP/MeTr complex was present in the asymmetric unit. Iterative rounds of refinement were carried out in CNS[42] and PHENIX[43], with model building in Coot[44]. Data collection and refinement statistics are shown in Table II.1, and average *B*-factors for each domain of the final model are given in Table II.2. Ramachandran analysis was carried out in PROCHECK[45]. The final model contains residues 1-262 (of 262) for both MeTr chains (A and B), residues 5-442 (of 446) for both CFeSP large subunit chains (C and E), and residues 1-323 (of 323) for both CFeSP small subunit chains (D and F).

Except for the Fe_4S_4 and B_{12} domains, the entire structure is composed of TIM barrels whose densities are well-defined (Figure II.13). Electron density for the Fe_4S_4 domains (Figures II.6, II.14, and II.16) is weaker, consistent with the fact that these domains exhibit higher *B*-factors (Table II.2). However, reasonable density is present for the entire main chain and most side chains, allowing us to build a model for this domain.

Even so, several side chains of the Fe_4S_4 domains (52 total residues in the final model) lacked electron density, and thus atoms were truncated past the C_β atom for the following residues (chain C: underlined, 13 total; chain E: italicized, 15 total): *Y9*, *K10*, L12, *K14*, *K15*, *N16*, *E19*, *K36*, *S38*, L39, *D40*, S41, *Y44*, V45, *S46*, *D47*, *R50*, *E51*, and *D54*.

Although density for the B_{12} domains is weak and discontinuous in many regions (Figure II.18), density for the B_{12} cofactors is unambiguous (Figure II.17), and density is also clear for some helical regions, including those near B_{12} (Figure II.15). Because the structure of a CFeSP B_{12} domain bound with B_{12} was already known[13], the clear density of the B_{12} cofactor and the resolvable helices were used to position the B_{12} domain during model building. Many side chains of the CFeSP B_{12} domains (118 total residues in the final model) lacked electron density, and thus atoms were truncated past the C_β atom for the following residues (chain C: underlined, 50 total; chain E: italicised, 60 total): K324, *L325*, *N326*, E327, *I328*, *V331*, N332, E333, *N334*, *S335*, *V337*, *V339*, *S344*, *L345*, *V350*, *I354*, *E355*, S356, *T357*, *K358*, *I359*, *S361*, *Y362*, *L363*, *L364*, *V366*, E383, *E385*, *K386*, *I387*, *V390*, *K392*, *K393*, *V394*, *D395*, *L396*, *D397*, *N398*, *K399*, *V400*, *V401*, *R402*, *H403*, *R404*, *I405*, *I406*, *V411*, *V413*, *L414*, *K415*, *K417*, *L418*, *E419*, *D420*, *L421*, *T422*, *W424*, E425, *V426*, *I427*, *V428*, *R431*, E432, *S434*, *I436*, *V437*, *F439*, *R441*.

The B_{12} cofactor in the final model contains 5,6-dimethylbenzimidazole as the lower ligand moiety, as in cobalamin. Previous studies have shown that CFeSP isolated from *M. thermoacetica* harbors an unusual B_{12} derivative that contains 5-methoxybenzimidazole as the lower ligand instead[46]. However, disorder of the B_{12} cofactor and B_{12} domain due to thermal motion of these regions in the CFeSP/MeTr crystal resulted in weak electron density for substituents of the benzimidazolyl ring

(Figures II.10 and II.17). Therefore, the presence of this unusual B_{12} derivative cannot be confirmed from these crystallographic studies, and cobalamin is the form of B_{12} modeled in the structure.

Previous spectroscopic studies[28] in addition to the crystal structure of *Ch*CFeSP[13] have indicated that a water molecule coordinates the central cobalt of B_{12} in the as-isolated CFeSP. Here, Co(II) is expected to be the major species and would thus be five-coordinate. However, because of disorder no electron density is observed to suggest a water molecule bound to cobalt (Figures II.10 and II.17). Accordingly, a water molecule has not been modeled.

All figures of the structure were prepared in PyMOL[47].

Solution and in crystallo *UV-Vis spectroscopy*

Titanium(III) citrate (100 mM in 50 mM Tris, pH 7.6) was prepared as previously described[48], and *(6S)-5*-methyl-*5,6,7,8*-tetrahydrofolate (CH_3-H_4folate) containing one glutamate tail was purchased from Schircks Laboratories. As-isolated, reduced, and methylated CFeSP samples in solution were prepared in a room temperature anaerobic chamber (MBraun) following similar procedures to those previously described[23,24,28,30-32]. Briefly, purified CFeSP (20 μM) was used for the as-isolated sample, CFeSP mixed with Ti(III) citrate (1 mM) was used for the reduced sample, and CFeSP mixed with equimolar MeTr, Ti(III) citrate (1 mM), and CH_3-H_4folate (1 mM) was used for the methylated sample. Spectra were taken using a Nanodrop 2000c (Thermo Scientific) in a quartz cuvette or on the sample stage in the anaerobic chamber directly after mixing; identical solutions lacking CFeSP were used as blanks.

To obtain *in crystallo* UV-Vis spectra, CFeSP/MeTr crystals in as-isolated, reduced, and methylated forms were prepared in a similar fashion. In a room temperature anaerobic chamber (Coy Laboratories), crystals were looped into a 2 μL drop which was placed on a cover slide and contained one of the following three solutions for as-isolated, reduced, and methylated samples, respectively: well solution, well solution with Ti(III) citrate (10 mM), and well solution with Ti(III) citrate (10 mM) and CH_3-H_4folate (1 mM). A ring of epoxy surrounding each drop was applied to the cover side, and a second cover slide was placed on top, sandwiching the drops within a uniform distance separation and sealing the crystals within an anaerobic environment. Upon curing of the epoxy, crystals were brought out of the anaerobic chamber and mounted on an XZ translation stage (Newport, UMR8.25 & SM-13) in a fiber optic coupled microspectrophotometer (Ocean Optics, Jaz) with 40 mm diameter reflective objectives (Optique Peter, France) and a deuterium-halogen lamp (DH2000-BAL, Ocean Optics), similar to the instrumental setup described previously[49,50]. Stray light was blocked with black-out material. The light focus was coarsely aligned to the crystals by visual inspection and then finely aligned by monitoring light transmission in real time. Data were acquired at room temperature with the SpectraSuite software (Ocean Optics). The background transmission was measured through the solution immediately surrounding the crystals. The dark current was measured with the light shuttered off. Sample, reference, and dark current spectra were acquired by averaging 10-50 scans with total exposure times of 90-1000 ms. All experiments were completed within 60 min of sample preparation, and crystals remained intact over the course of the experiment, as observed using a microscope following data collection.

II.G. Tables and Figures

Table II.1. Data collection and refinement statistics. Information for both the initial, lower-resolution and the final, higher-resolution data sets are given.

	Lower resolution*	Higher resolution*
Data Collection		
Beam line	Advanced Light Source, 5.0.2	Advanced Photon Source, 24ID-C
Resolution (Å)	25.00-3.30 (3.42-3.30)	50.00-2.38 (2.47-2.38)
Space group	$P2_12_12_1$	$P2_12_12$
Cell dimensions		
a, *b*, *c* (Å)	137.42, 159.87, 241.92	125.25, 243.84, 79.44
Wavelength (Å)	1.1000	0.9792
Completeness (%)	88.2 (73.8)	99.7 (100.0)
Redundancy	4.7	5.1
$I/\sigma I$	11.6 (2.6)	19.0 (2.8)
R_{sym} (%)	12.2 (46.6)	9.1 (52.5)
Wilson *B* (Å^2)	74.84	48.34
Solvent content (%)	58.6	56.4
Refinement		
Number of reflections	70,058	94,458
Resolution	25.00-3.31	50.00-2.38
R_{work}/R_{free} (%)	29.20/33.69	21.51/25.07
Ramachandran analysis (%)		
Most favored	74.0	87.3
Additionally allowed	21.2	11.5
Generously allowed	3.1	0.8
Disfavored	1.7	0.3
R.m.s. deviations		
Bond lengths (Å)	0.011	0.006
Bond angles (°)	1.484	1.150

* Highest resolution bin in parentheses

Table II.2. Average *B*-factors for each domain of the CFeSP/MeTr structure.

Protein-Chain	Domain	Residues	No. of atoms	Avg. *B*-factor (Å^2)
MeTr-A	TIM	1-262	1,993	66.78
MeTr-B	TIM	1-262	1,993	66.55
CFeSP large subunit-C	Fe_4S_4	5-57	332	104.80
CFeSP large subunit-C	Linker 1	58-92	273	57.04
CFeSP large subunit-C	TIM	94-313	1,683	62.68
CFeSP large subunit-C	Linker 2	314-324	94	87.91
CFeSP large subunit-C	B_{12}	325-442	729	127.24
CFeSP small subunit-D	TIM	1-262	2,461	40.01
CFeSP large subunit-E	Fe_4S_4	5-57	320	117.88
CFeSP large subunit-E	Linker 1	58-92	273	50.59
CFeSP large subunit-E	TIM	94-313	1,683	46.34
CFeSP large subunit-E	Linker 2	314-324	98	95.57
CFeSP large subunit-E	B_{12}	325-442	708	129.99
CFeSP small subunit-F	TIM	1-262	2,470	42.92
All atoms			15,773	63.56
All protein			15,110	63.40
Protein: B_{12} domains			1,437	128.60
Protein: non-B_{12} domains			13,673	56.55
Ligands (B_{12})			182	103.11
Fe_4S_4 cluster			16	123.98
Waters			465	51.21

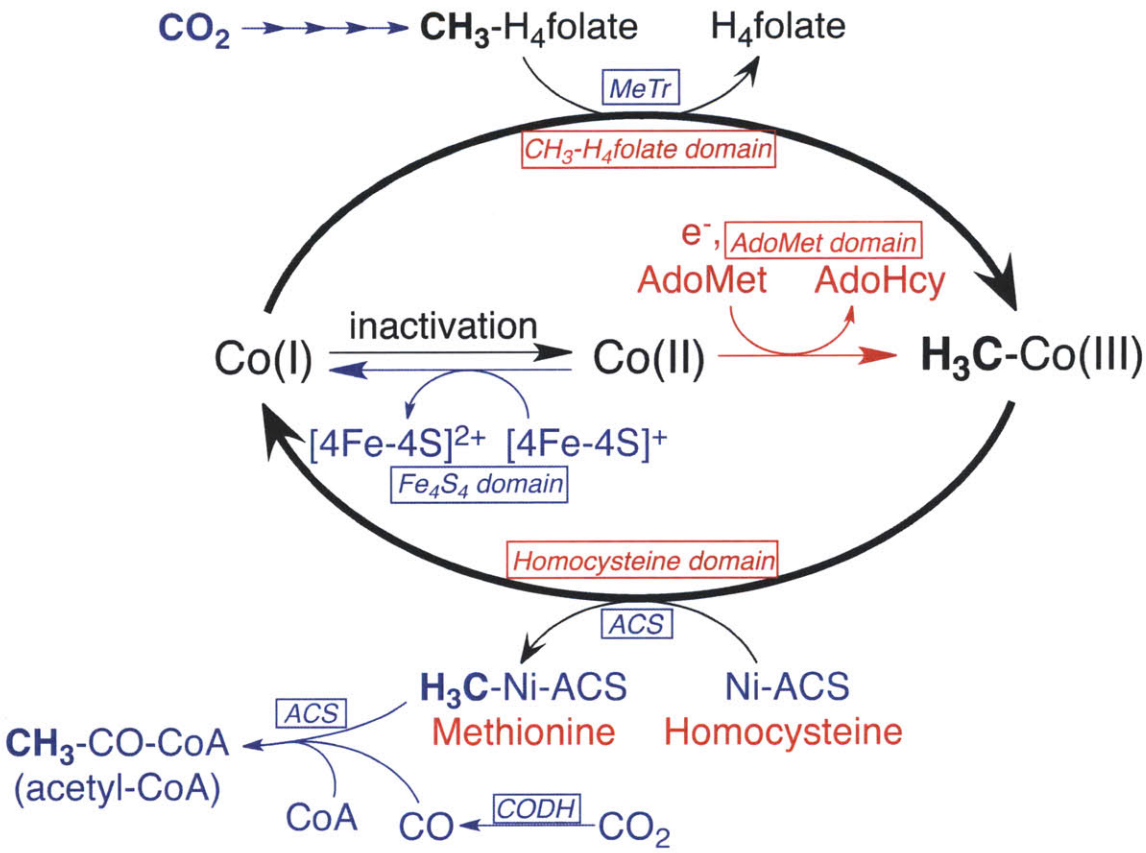


Figure II.1. Cobalt redox states in the B_{12}-dependent methyl transfer reaction cycle. Steps carried out in acetogenesis in blue, those performed by MetH in red, and those used in both systems in black. Transferred methyl group in bold. Active B_{12} in the Co(I) state attacks the methyl group of CH_3-H_4folate, giving CH_3-Co(III). In CFeSP, the methyl group is subsequently delivered to the Ni-Fe-S A-cluster of acetyl-CoA synthase (ACS), which then combines the methyl group with CoA and CO, also derived from CO_2 in a reaction catalyzed by carbon monoxide dehydrogenase (CODH), to form acetyl-CoA. In MetH, the methyl group is transferred to homocysteine to yield methionine. Intermittent oxidation of Co(I) to Co(II) necessitates reactivation. In CFeSP, an electron is passed from an electron donor protein, such as CODH, ferredoxin, or pyruvate:ferredoxin oxidoreductase[24], to the Fe_4S_4 cluster, which then reduces the B_{12} cobalt. In MetH, reductive methylation occurs via an electron from flavodoxin in *Escherichia coli* or from methionine synthase reductase in humans and a methyl group from *S*-adenosylmethionine (AdoMet), yielding *S*-adenosylhomocysteine (AdoHcy).

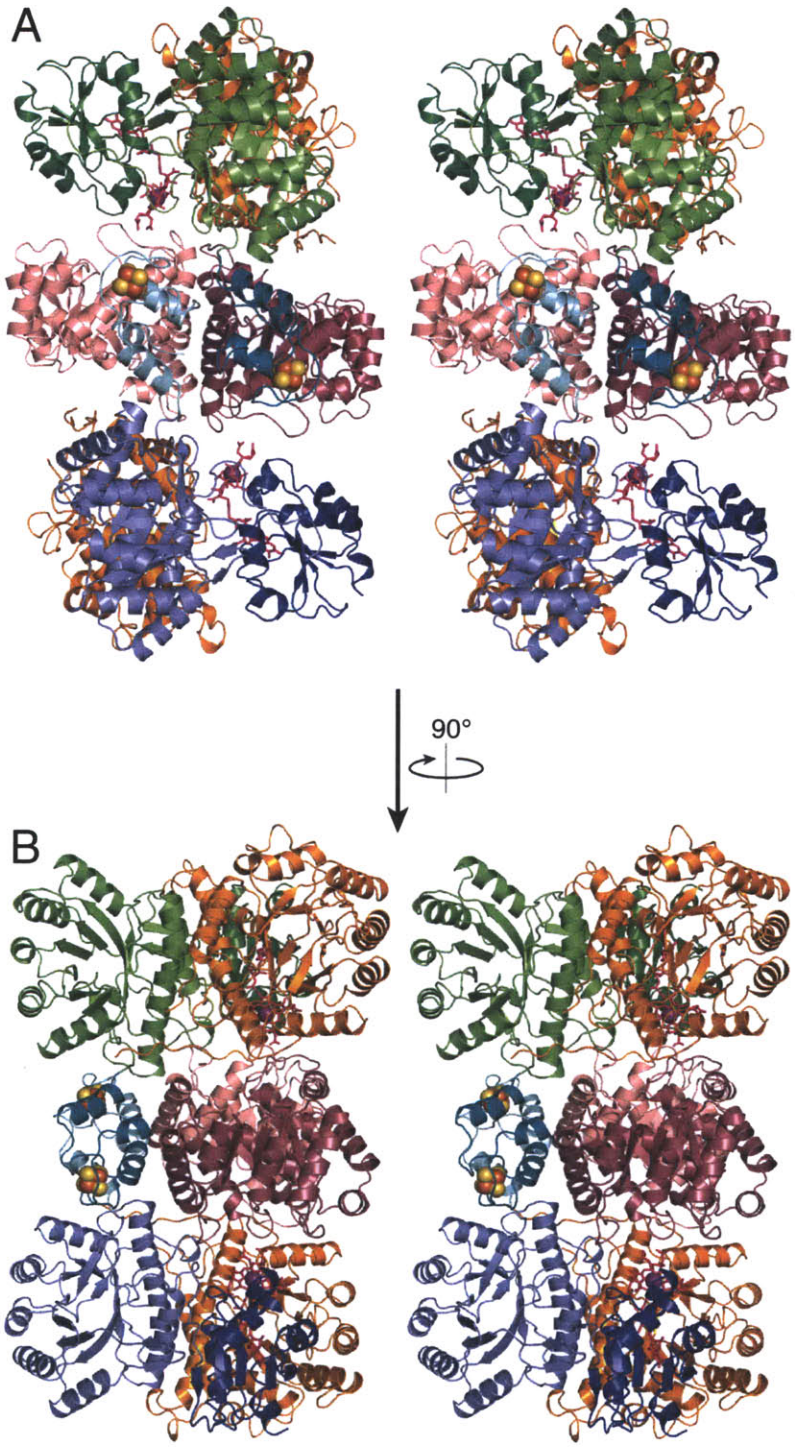


Figure II.2. The overall CFeSP/MeTr complex. A, Stereoviews of the complex (ribbons) in one orientation and **B**, at a 90° rotation about the y axis. MeTr homodimer in light and dark pink, CFeSP small subunits in orange, and CFeSP large subunit Fe_4S_4 domains in teal and cyan, TIM barrel domains in green and blue, and B_{12} domains in dark green and dark blue. B_{12} cofactors in magenta sticks with cobalt as violet spheres. Fe_4S_4 clusters in spheres with Fe in orange and S in yellow.

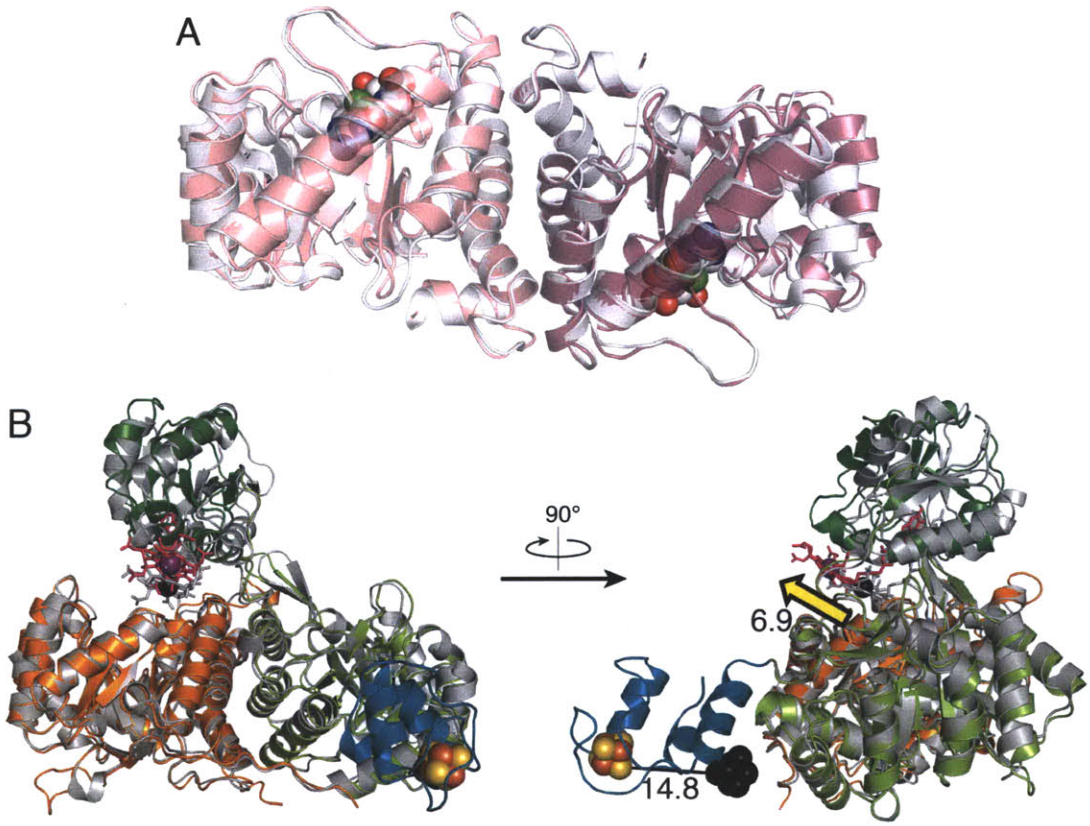


Figure II.3. Alignments of CFeSP/MeTr components with their individual, previously-solved structures. A, Superposition of the MeTr homodimer component of CFeSP/MeTr (light and dark pink semi-transparent ribbons) with the structure of MeTr[9] (PDB ID: 2E7F, grey semi-transparent ribbons) bound with CH_3-H_4folate (spheres: O in red, N in blue, and C in grey, except for the methyl group to be transferred, in green). **B**, Superposition of one CFeSP of the CFeSP/MeTr complex (ribbons; small subunit in orange; large subunit: Fe_4S_4 domain in teal, TIM barrel domain in light green, B_{12} domain in dark green; B_{12} cofactor in pink sticks, with cobalt as a violet sphere; Fe_4S_4 cluster in spheres: Fe in orange and S in yellow) with the previous structure of *Ch*CFeSP[13] (PDB ID: 2H9A, grey ribbons, B_{12} as grey sticks, with cobalt as a black sphere, Fe_4S_4 cluster in black spheres). Distances of cofactor positions from the *Ch*CFeSP structure to the CFeSP/MeTr structure given in Å. Movement of the B_{12} cofactor illustrated by the yellow arrow.

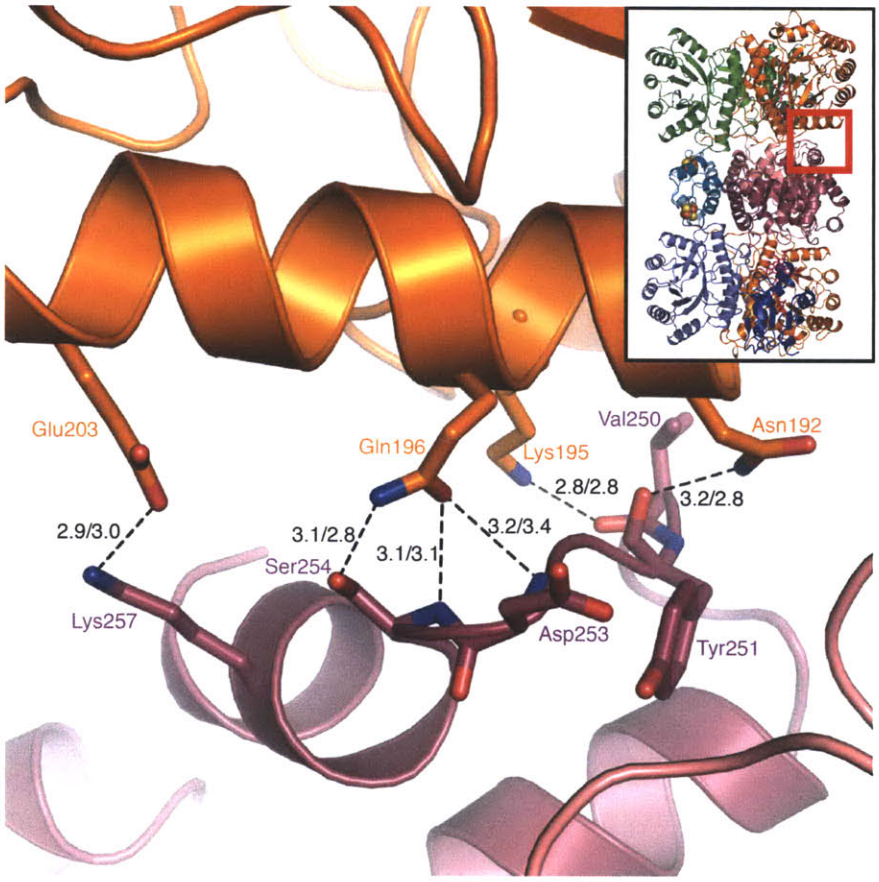


Figure II.4. Interface between MeTr and the CFeSP small subunit. MeTr and CFeSP small subunit in dark pink and orange ribbons, respectively, with selected residue side chains in sticks (oxygen in red and nitrogen in blue). Distances in Å from both MeTr and CFeSP interfaces of the CFeSP/MeTr complex. Inset figure indicates the zoomed area with a red box.

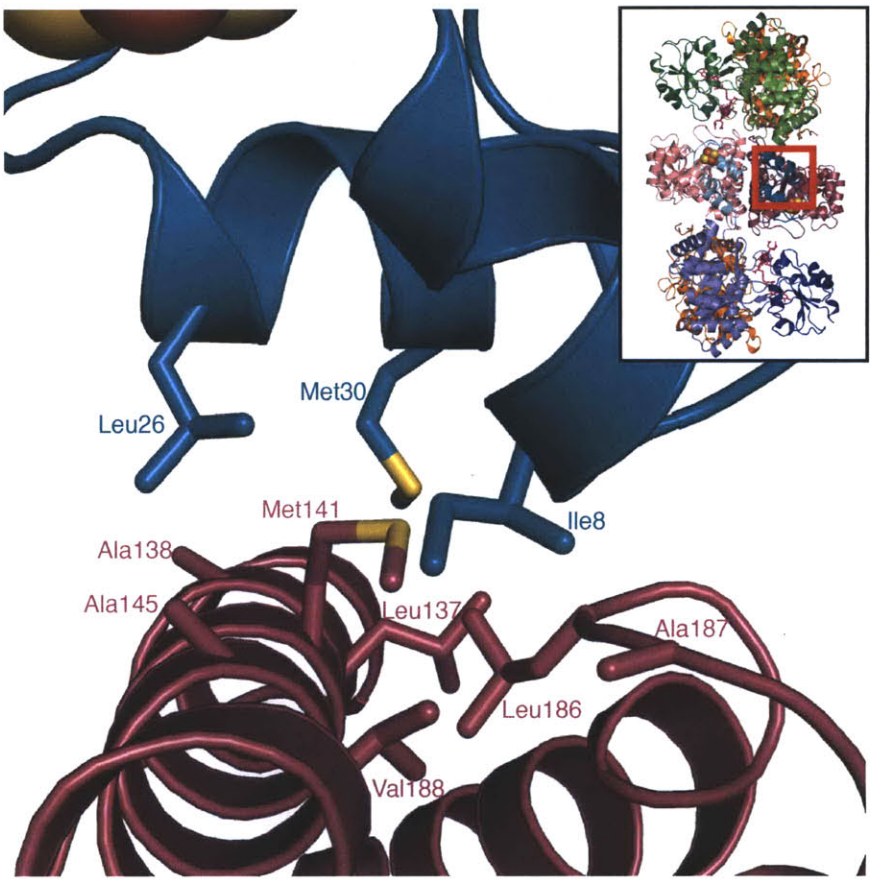


Figure II.5. The primarily hydrophobic interface between MeTr and the CFeSP large subunit Fe_4S_4 domain. MeTr and Fe_4S_4 domain in dark pink and teal ribbons, respectively, with selected residue side chains in sticks (sulfur in yellow). Inset figure indicates the zoomed area with a red box.

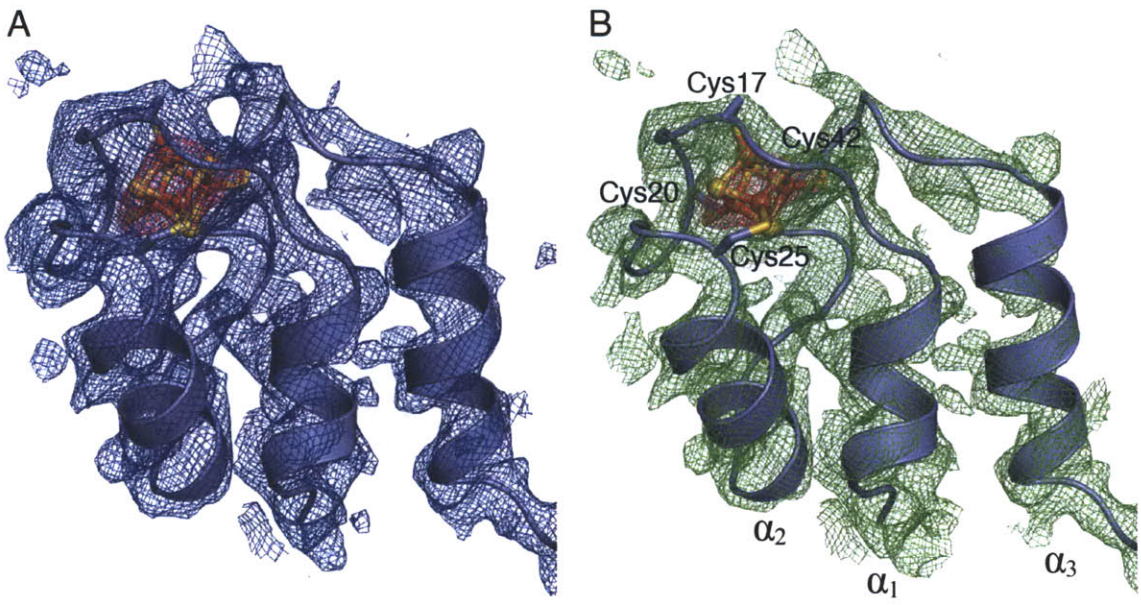


Figure II.6. Electron density for CFeSP large subunit Fe_4S_4 domain. A, $2F_o$-F_c map in blue mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the Fe_4S_4 cluster. **B,** Composite omit map in green mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the Fe_4S_4 cluster. Fe_4S_4 domain in slate ribbons, Fe_4S_4 clusters in ball-and-stick, and cysteine ligands to the cluster are labeled in **B** and shown in sticks (iron in orange, sulfur in yellow). Helices in **B** are numbered α_1-α_3 according to sequence order. Although Cys17, Cys20, and Cys25 exist in a C-X_2-C-X_4-C motif observed in other Fe_4S_4 proteins involved in electron transfer, such as ferredoxin, CODH, and hydrogenase, the fold of the CFeSP Fe_4S_4 domain appears to be novel. Additional views of electron ensity for Fe_4S_4 domain side chains and the region near the cluster are shown in Figures II.14 and II.16, respectively.

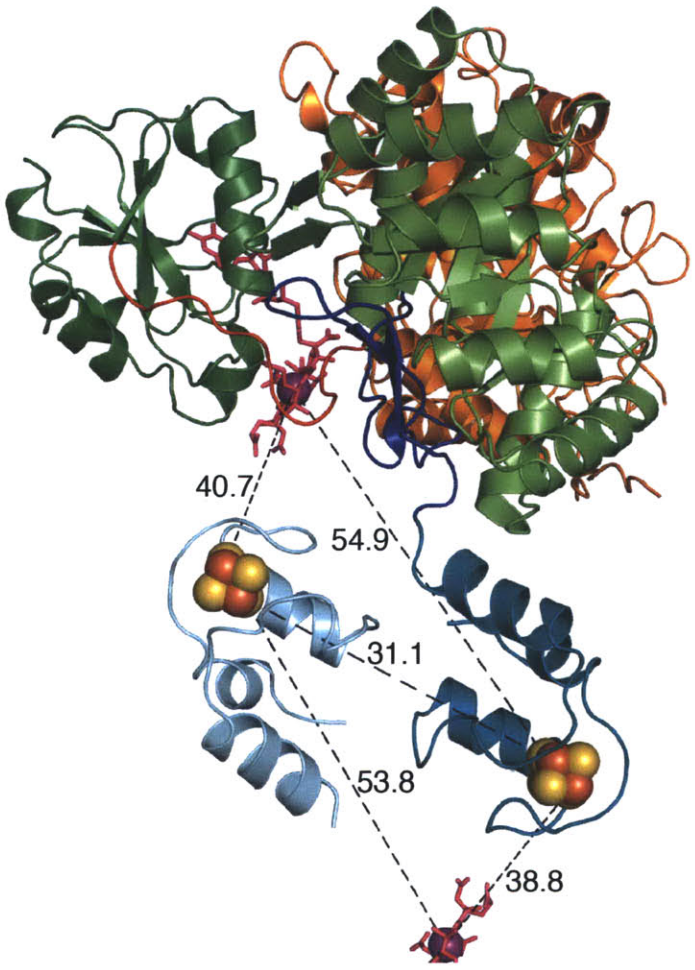


Figure II.7. The Fe_4S_4 and B_{12} cofactor distances in the CFeSP/MeTr complex. The MeTr homodimer is hidden for clarity. While one entire CFeSP is shown, only the Fe_4S_4 domain and B_{12} cofactor from the other CFeSP is displayed. Same orientation and color scheme as the left side of Figure II.2, except the linker between the CFeSP large subunit Fe_4S_4 and TIM barrel domains is dark blue, and the linker between the TIM barrel and B_{12} domains is red. Distance labels in Å.

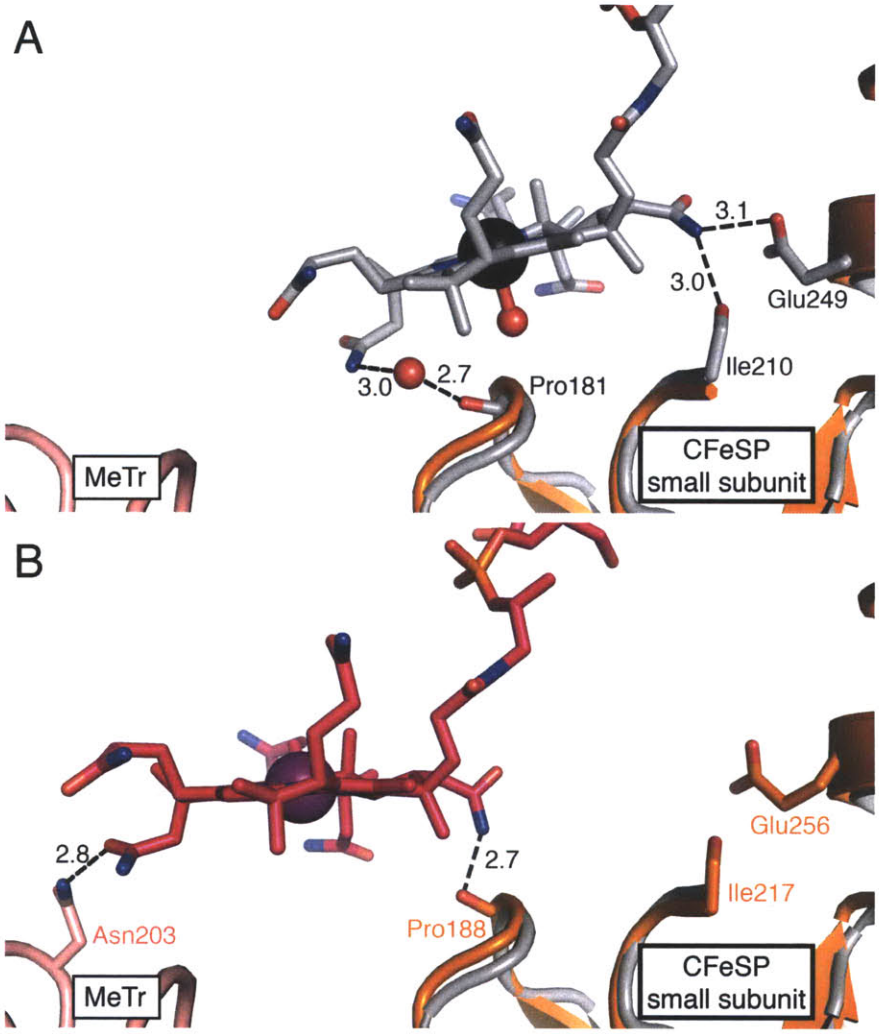


Figure II.8. Comparison of B_{12} domain positions between *Ch*CFeSP and CFeSP/MeTr. A, *Ch*CFeSP[13] (grey ribbons, PDB ID: 2H9A) and its B_{12} cofactor in sticks (C: grey, O: red, N: blue, and Co: black sphere; water: red spheres), aligned by the CFeSP small subunit with the CFeSP/MeTr structure (orange/pink ribbons). **B**, CFeSP/MeTr B_{12} cofactor in sticks (C: magenta, O: red, N: blue, Co: violet sphere). **A** and **B** are identical in orientation. Distances in Å. Electron density for B_{12} is shown in Figure II.10 and II.17.

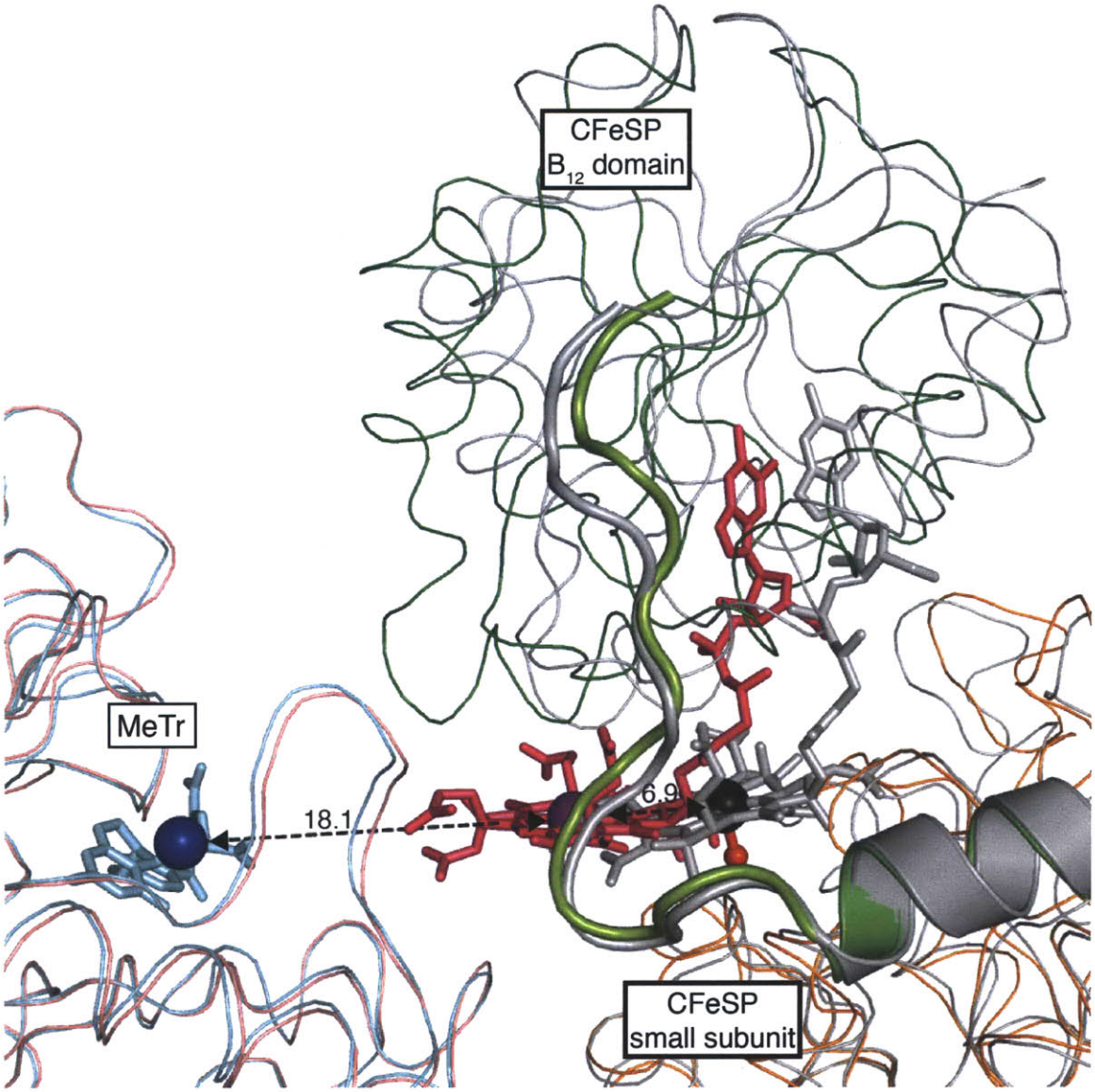


Figure II.9. Positioning of B_{12} towards the MeTr active site. Alignment of *Ch*CFeSP[13] (grey ribbons, PDB ID: 2H9A) with CFeSP/MeTr (orange/green/pink ribbons) by CFeSP small subunits. Linkers between B_{12} and TIM barrel domains plus a helix of the TIM barrel domain in wider ribbons. The remainder of the large subunit is not shown. CH_3-H_4folate-bound MeTr structure (cyan ribbons) aligned to MeTr in CFeSP/MeTr. CH_3-H_4folate in cyan sticks and the methyl group to be transferred as a blue sphere. Distances in Å.

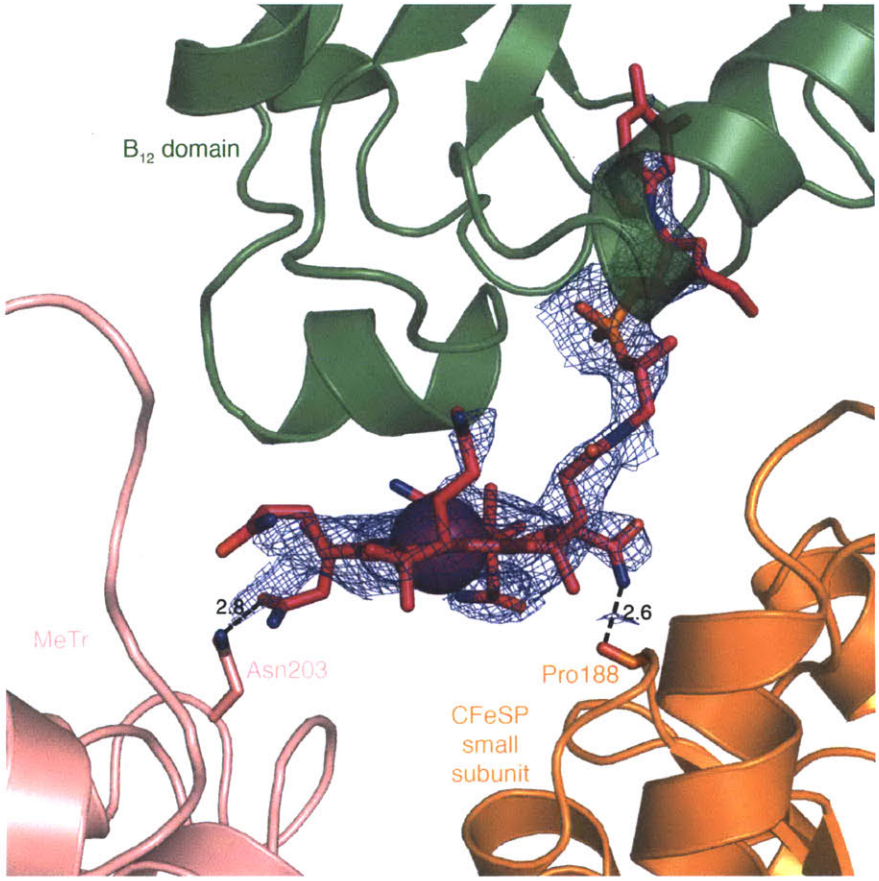


Figure II.10. The position of B_{12} the CFeSP/MeTr complex. $2F_o$-F_c electron density at 0.7 σ (blue mesh) around one B_{12} cofactor of the CFeSP/MeTr complex. MeTr, CFeSP small subunit, and CFeSP large subunit B_{12} domain in pink, orange, and green ribbons, respectively. CFeSP large subunit central TIM barrel and Fe_4S_4 domains are not shown. The B_{12} cofactor, Asn203 side chain of MeTr, and Pro188 backbone carbonyl of the CFeSP small subunit shown in sticks: C in magenta, N in blue, O in red, P in orange, with Co as a violet sphere. Distances in Å. Additional views of electron density for the B_{12} cofactor are shown in Figure II.17.

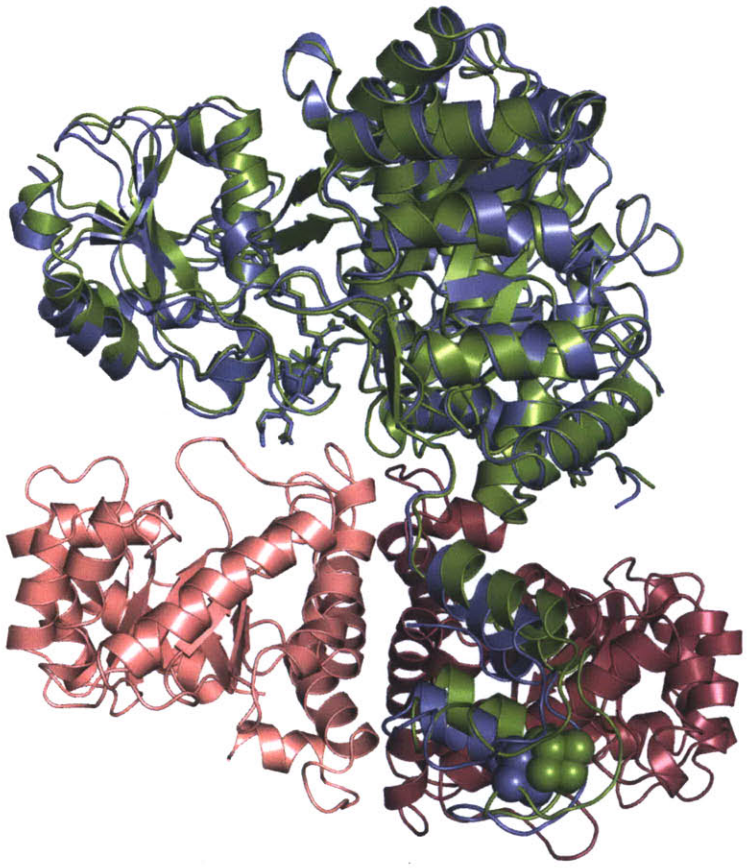


Figure II.11. Alignment of the two CFeSP heterodimers in the CFeSP/MeTr complex by their small subunits. The alignment (rmsd: 0.30 Å) superimposes their B_{12} cofactors. MeTr subunits in light and dark pink ribbons. Reference and aligned CFeSPs of the complex in green and blue, respectively: protein in ribbons, B_{12} in sticks with cobalt as spheres, and Fe_4S_4 cluster in spheres.

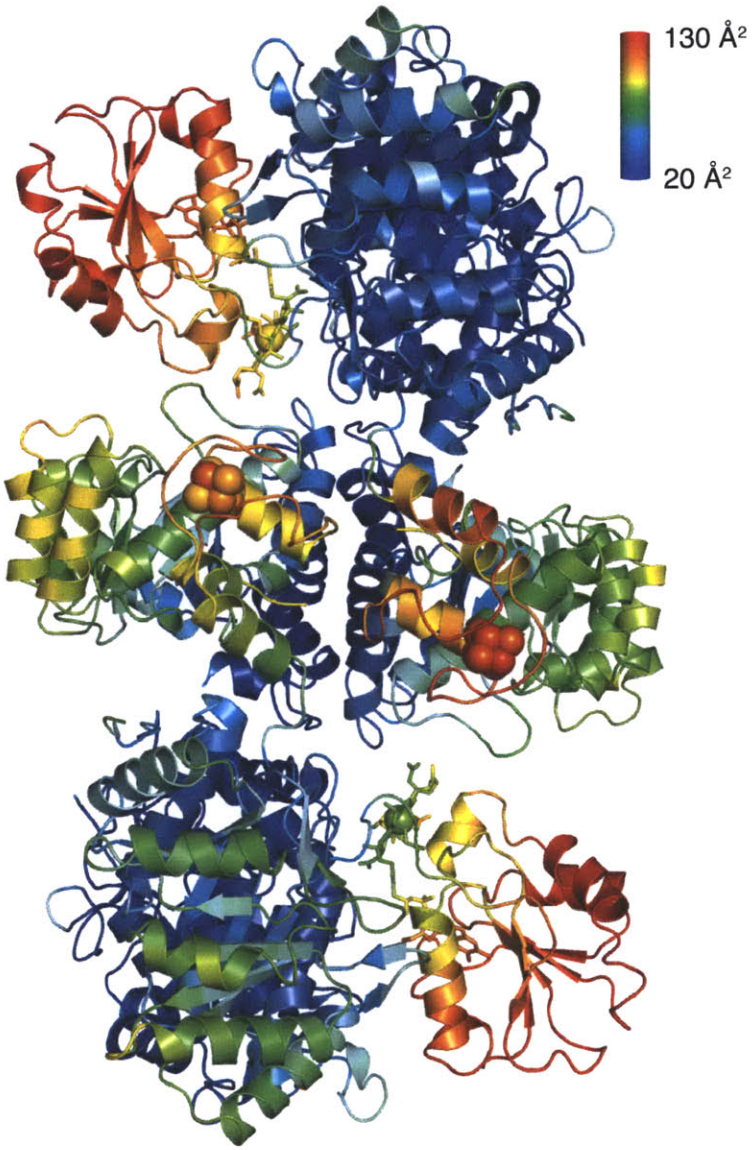


Figure II.12. *B*-factors of the CFeSP/MeTr complex. CFeSP/MeTr in ribbons, coloured by *B*-factor: lower *B*-factors in blue and green, higher *B*-factors in red and orange. B_{12} domains exhibit significantly higher *B*-factors than the rest of the structure, which indicate a greater degree of disorder in the crystal for these regions. See Table II.2 for average refined *B*-factors for each domain. Identical orientation as the left side of Figure II.2.

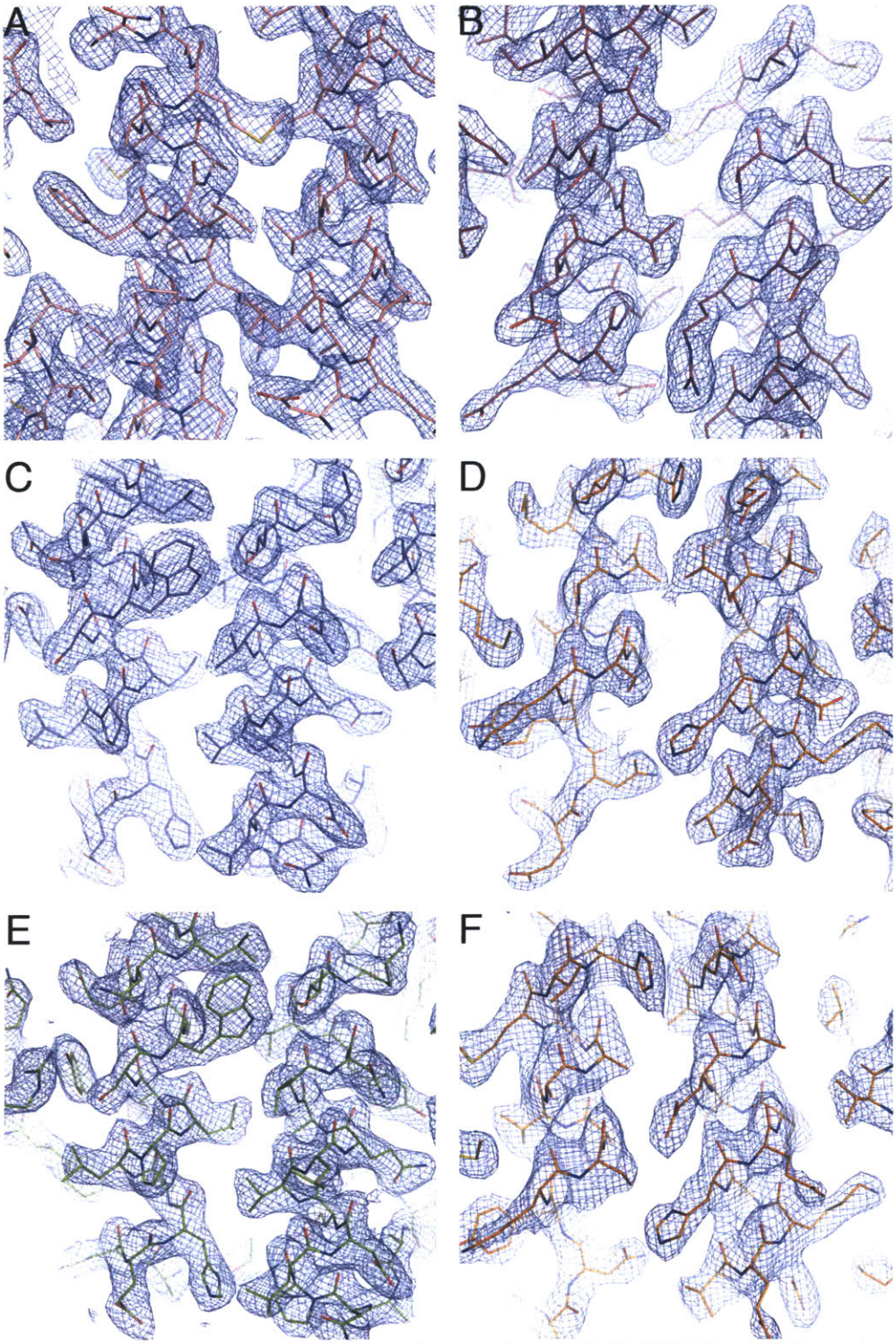


Figure II.13. Electron density for helical regions of all TIM barrel domains (sticks). **A**, MeTr, chain A (carbon in light pink). **B**, MeTr, chain B (carbon in dark pink). **C**, CFeSP large subunit, chain C (carbon in slate). **D**, CFeSP large subunit, chain E (carbon in green). **E**, CFeSP small subunit, chain D (carbon in orange). **F**, CFeSP small subunit (carbon in orange). Oxygen in red, nitrogen in blue, sulfur in yellow. $2F_o$-F_c maps in blue mesh (1.0 σ).

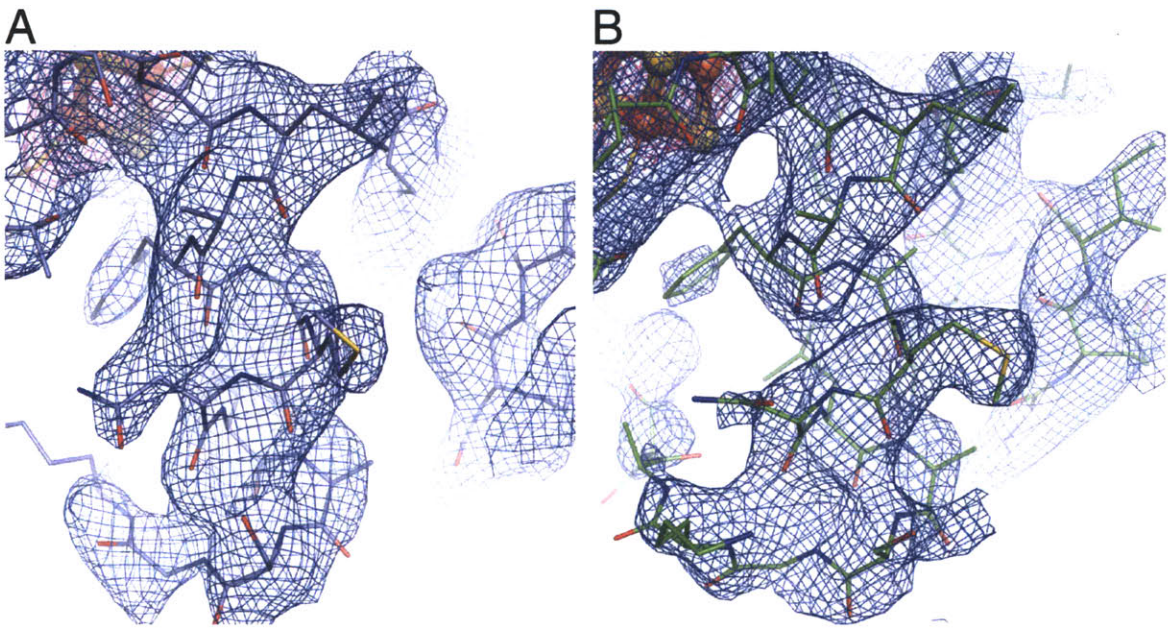


Figure II.14. Electron density for helical region of the CFeSP large subunit Fe_4S_4 domains (sticks). **A**, CFeSP Fe_4S_4 domain, chain C (carbon in slate). **B**, CFeSP Fe_4S_4 domain, chain E (carbon in green). Fe_4S_4 clusters in ball-and-stick. Oxygen in red, nitrogen in blue, sulfur in yellow, and iron in orange. $2F_o$-F_c maps in blue mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the Fe_4S_4 cluster. In the final model, residues of the Fe_4S_4 domains whose side chains lacked electron density were truncated past the C_β atom.

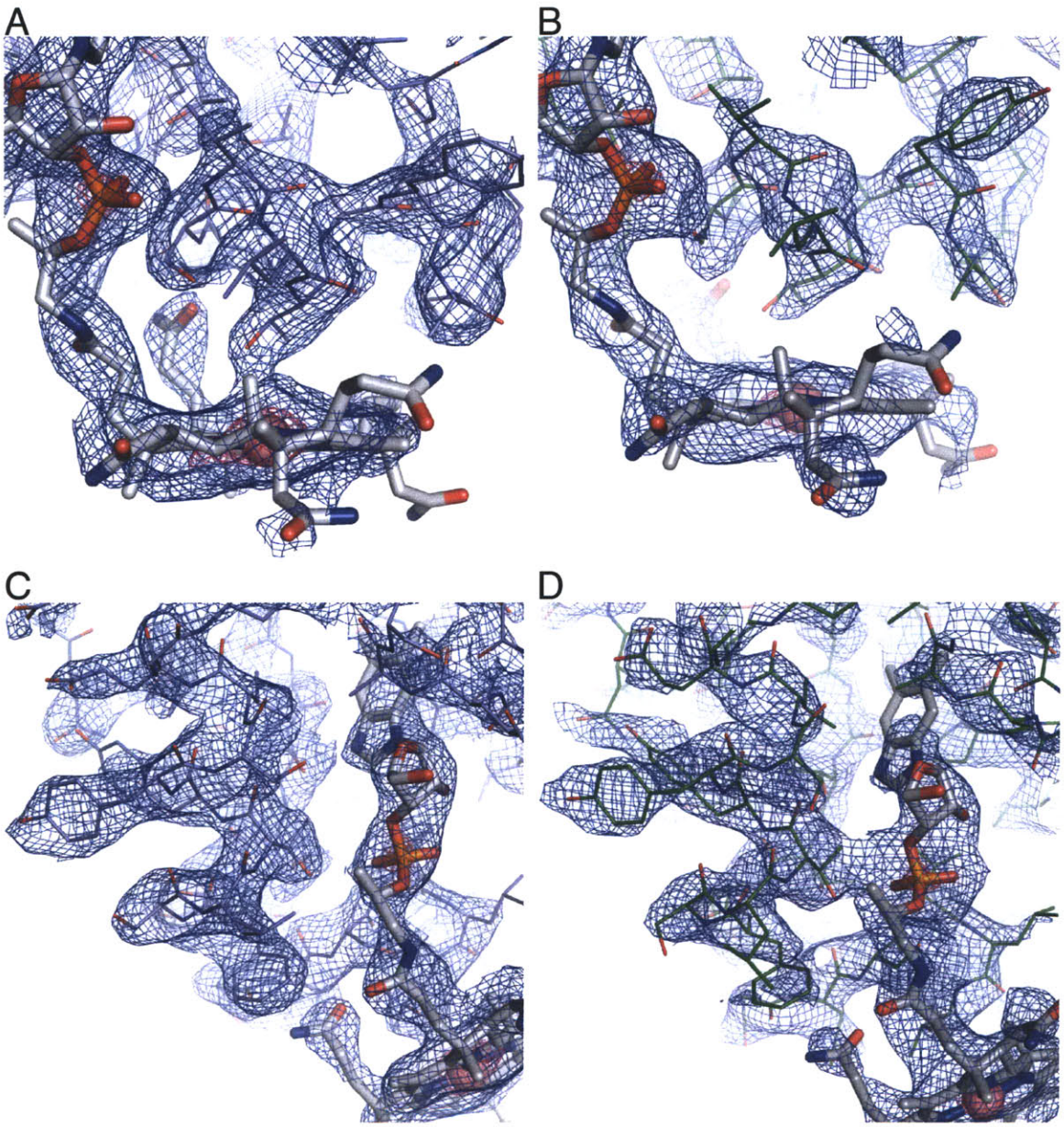


Figure II.15. Electron density for helical regions (sticks) in the vicinity of the B_{12} cofactor of the CFeSP large subunit B_{12} domains. A and **C**, CFeSP B_{12} domain, chain C (carbon in slate). **B** and **D**, CFeSP B_{12} domain, chain E (carbon in green). B_{12} in wider sticks: carbon in grey and cobalt as a pink sphere. Oxygen in red, nitrogen in blue, and phosphorus in orange. $2F_o$-F_c maps in blue mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the cobalt and phosphorus positions. In the final model, residues of the B_{12} domains whose side chains lacked electron density were truncated past the C_β atom.

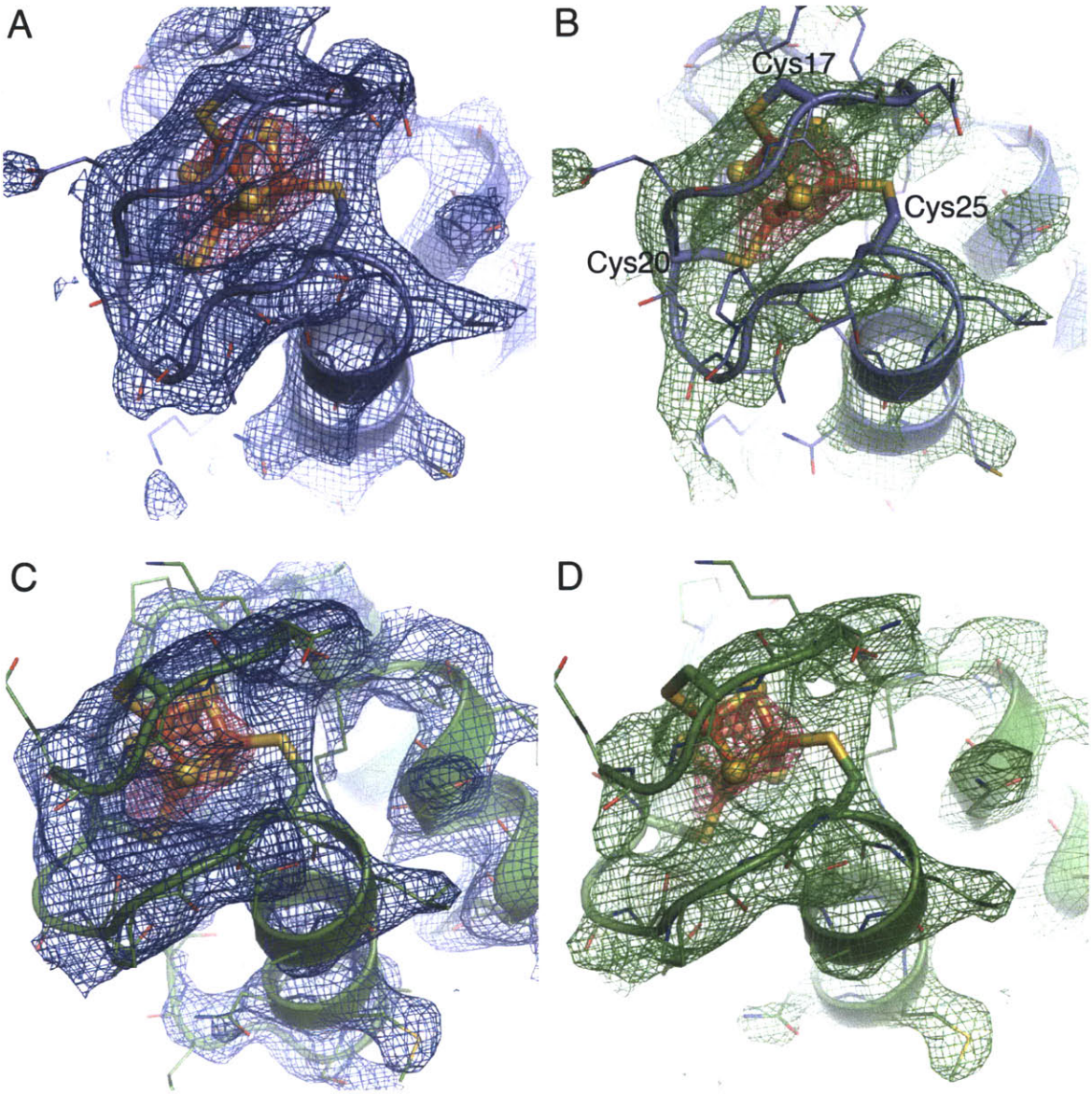


Figure II.16. Electron density for the Fe_4S_4 clusters and their vicinity (ribbons and sticks). **A-B**, CFeSP large subunit Fe_4S_4 domain, chain C (carbon in slate). **C-D**, CFeSP large subunit Fe_4S_4 domain, chain E (carbon in green). Fe_4S_4 clusters in ball-and-stick. Oxygen in red, nitrogen in blue, sulfur in yellow, and iron in orange. The Cys-X_2-Cys-X_4-Cys loop that coordinates the cluster is shown in the foreground and labeled in **B**. **A** and **C** show $2F_o$-F_c maps in blue mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the Fe_4S_4 cluster. **B** and **D** show composite omit maps in green mesh (0.7 σ) and pink mesh (3.0 σ) to highlight the Fe_4S_4 cluster. In the final model, residues of the Fe_4S_4 domains whose side chains lacked electron density were truncated past the C_β atom.

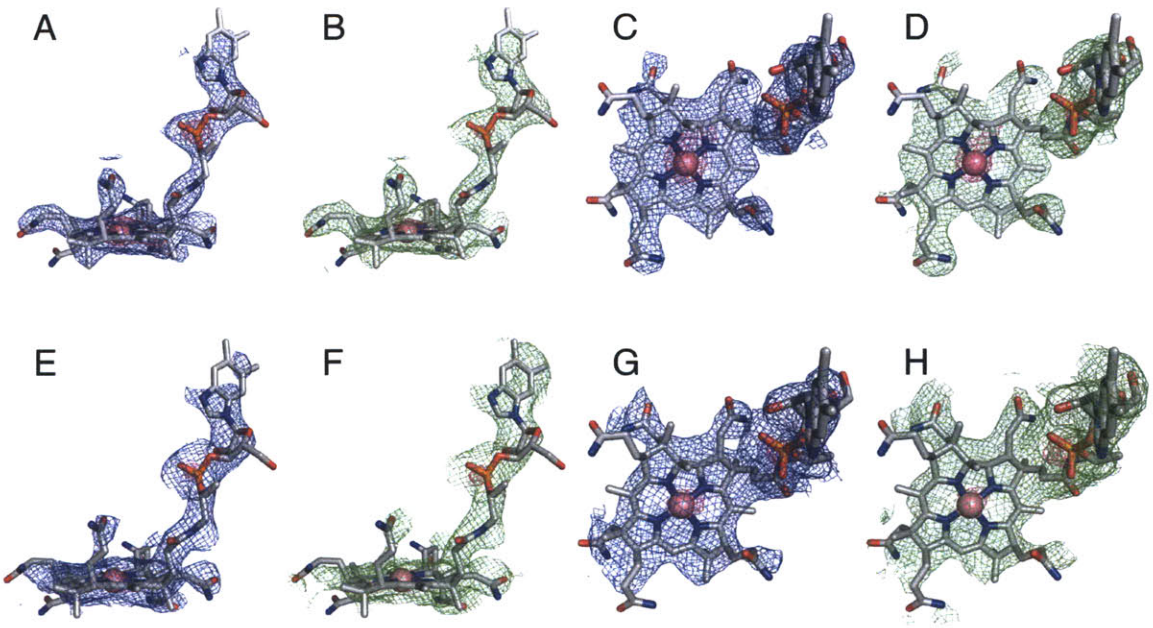


Figure II.17. Electron density for B_{12} cofactors (sticks). A-D, CFeSP large subunit B_{12}, chain C. **E-H**, CFeSP large subunit B_{12}, chain E. Carbon in grey, oxygen in red, nitrogen in blue, phosphorus in orange, and cobalt as a pink sphere. **A-B** and **E-F** show side views, while **C-D** and **G-H** show top views. **A**, **C**, **E**, and **G**, $2F_o$-F_c map in blue mesh (0.7 σ) and pink mesh (3.0 σ) to highlight cobalt and phosphorus locations. **B**, **D**, **F**, and **H**, composite omit map in green mesh (0.7 σ) and pink mesh (3.0 σ) to highlight cobalt and phosphorus locations.

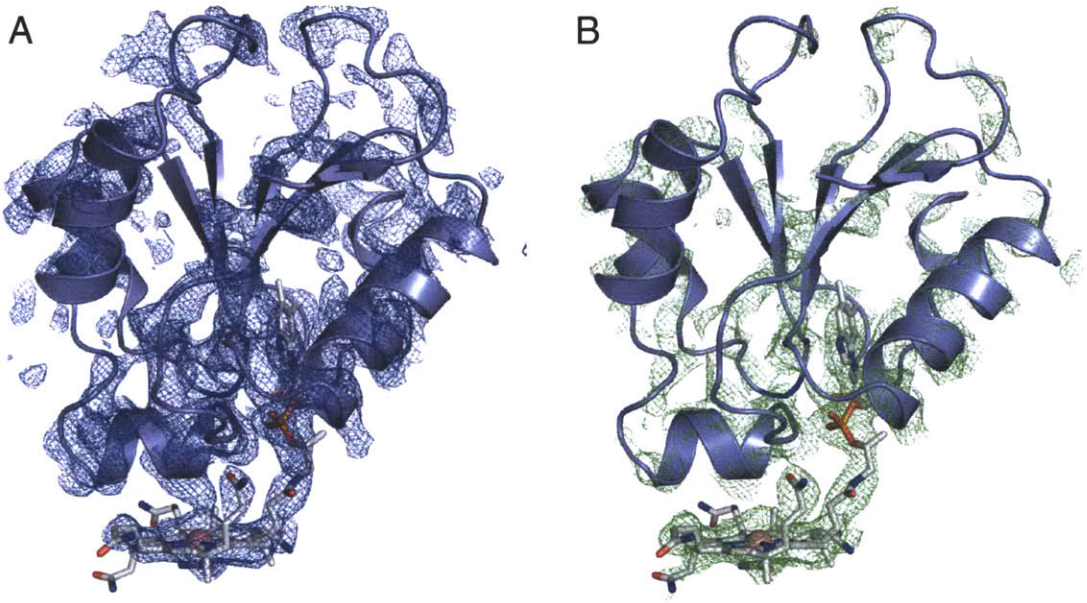


Figure II.18. Electron density for the CFeSP large subunit B_{12} domain (slate ribbons). A, $2F_o$-F_c map in blue mesh (0.7 σ). **B**, composite omit map in green mesh (0.7 σ). B_{12} in sticks, with carbon in grey, oxygen in red, nitrogen in blue, phosphorus in orange, and cobalt as a pink sphere. Although density for the B_{12} domain is weak due to movement within the crystal lattice, density for the B_{12} cofactor itself is clear (Figure II.17), and many regions are resolvable (Figure II.15). Because the structure of a CFeSP B_{12} domain with B_{12} bound is known (PDB ID: 2H9A), the positions of B_{12} and the resolvable regions in this structure were used to orient the placement of the B_{12} domain while model building. In the final model, residues of the B_{12} domains whose side chains lacked electron density were truncated past the C_β atom.

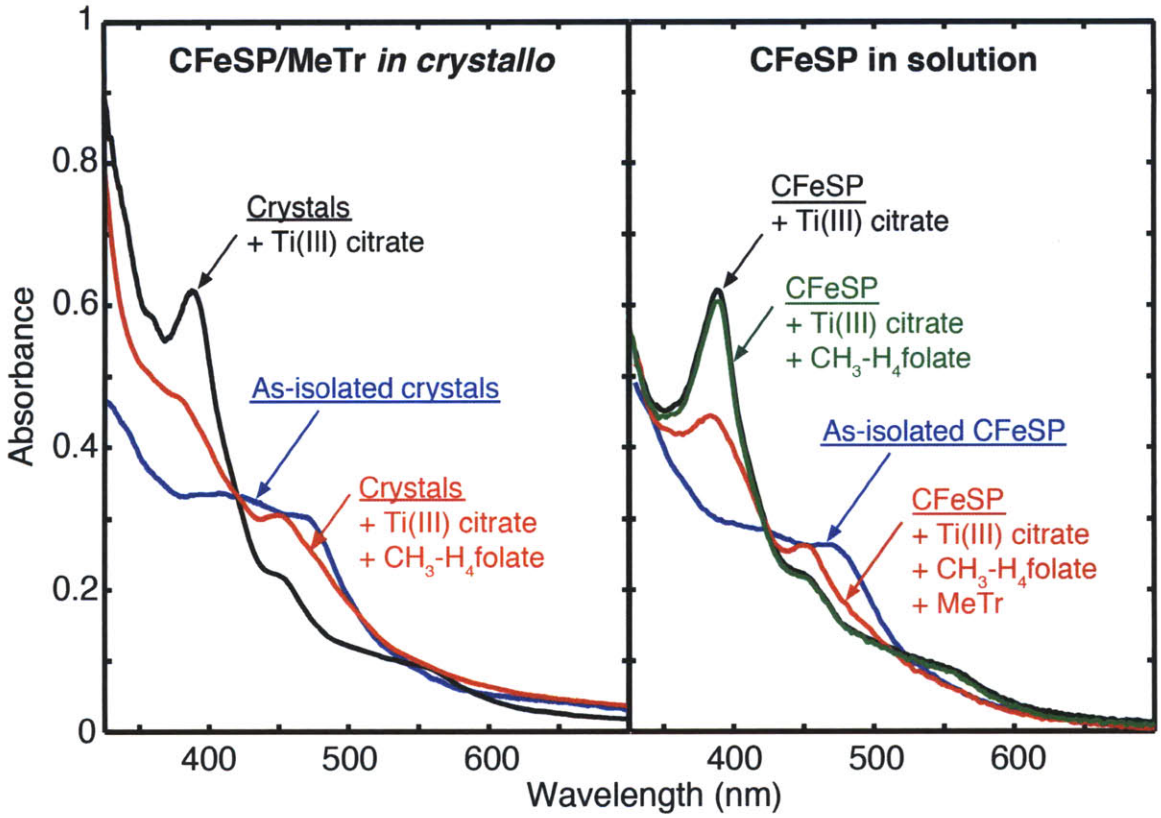


Figure II.19. Methyl transfer activity of CFeSP/MeTr crystals. UV-Vis absorption spectra for as-isolated (blue), reduced (black), and methylated (red) samples of CFeSP/MeTr crystals (left pane) and CFeSP in solution (right pane). In a control (green, right pane), the reaction in the absence of MeTr shows no turnover. Peaks at 390 nm and 450 nm are indicative of CFeSP-bound B_{12} in Co(I) and CH_3-Co(III) states, respectively. Absorbance spectra are scaled relative to each other to account for variable crystal sizes and pathlengths, where absolute peak absorbances did not exceed 1 absorbance unit.

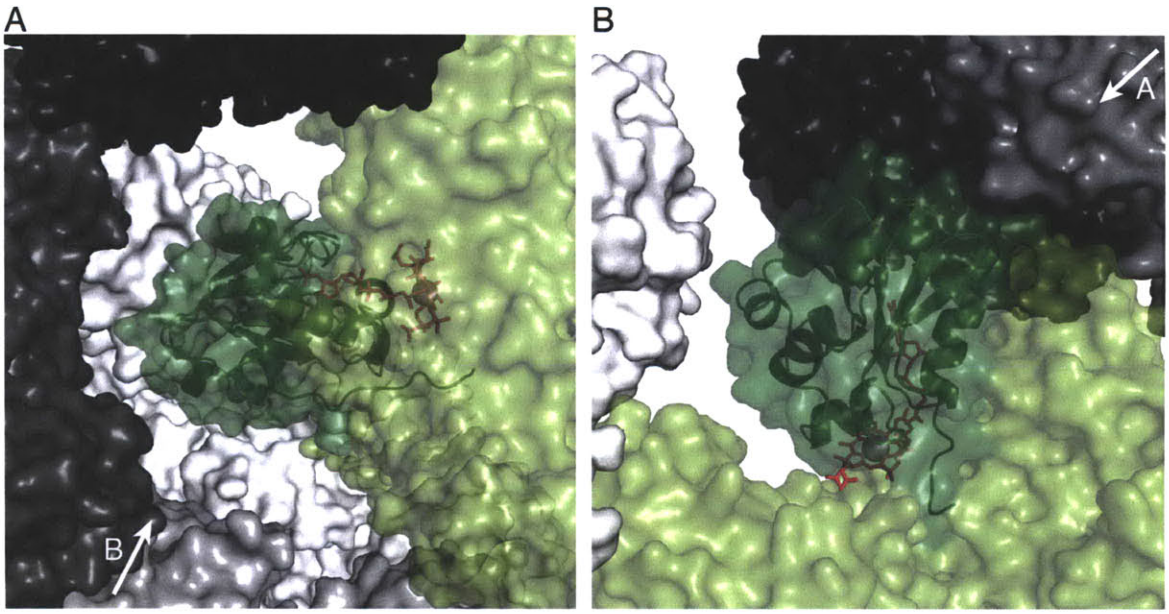


Figure II.20. Lack of lattice contacts between the B_{12} domain and symmetry-related neighboring protein molecules. The CFeSP/MeTr complex present in the asymmetric unit shown as a semi-transparent green surface, with B_{12} domain in dark green and in ribbons and B_{12} cofactor in magenta sticks, with cobalt as a sphere. Five neighboring CFeSP/MeTr complexes are located in the vicinity and are shown in varying shades of grey surface, as seen in two orientations, **A** and **B**, and as indicated by the labeled arrows.

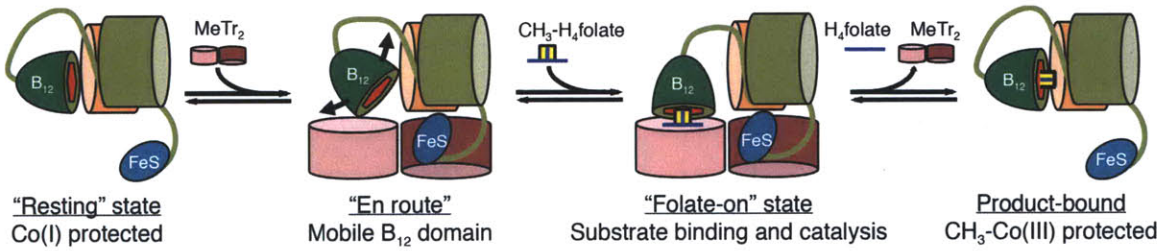


Figure II.21. Cartoon model of B_{12}-dependent methyl transfer in CFeSP/MeTr. For simplicity, only one of the two CFeSP heterodimers in the full CFeSP/MeTr complex is shown. Same color scheme as Figure II.2. Loops represent linker regions. Red hexagons indicate the B_{12} cofactor, and the transferred methyl group is highlighted with a yellow box. Curved arrows denote the flexibility the "en route" B_{12} domain prior to CH_3-H_4folate binding.

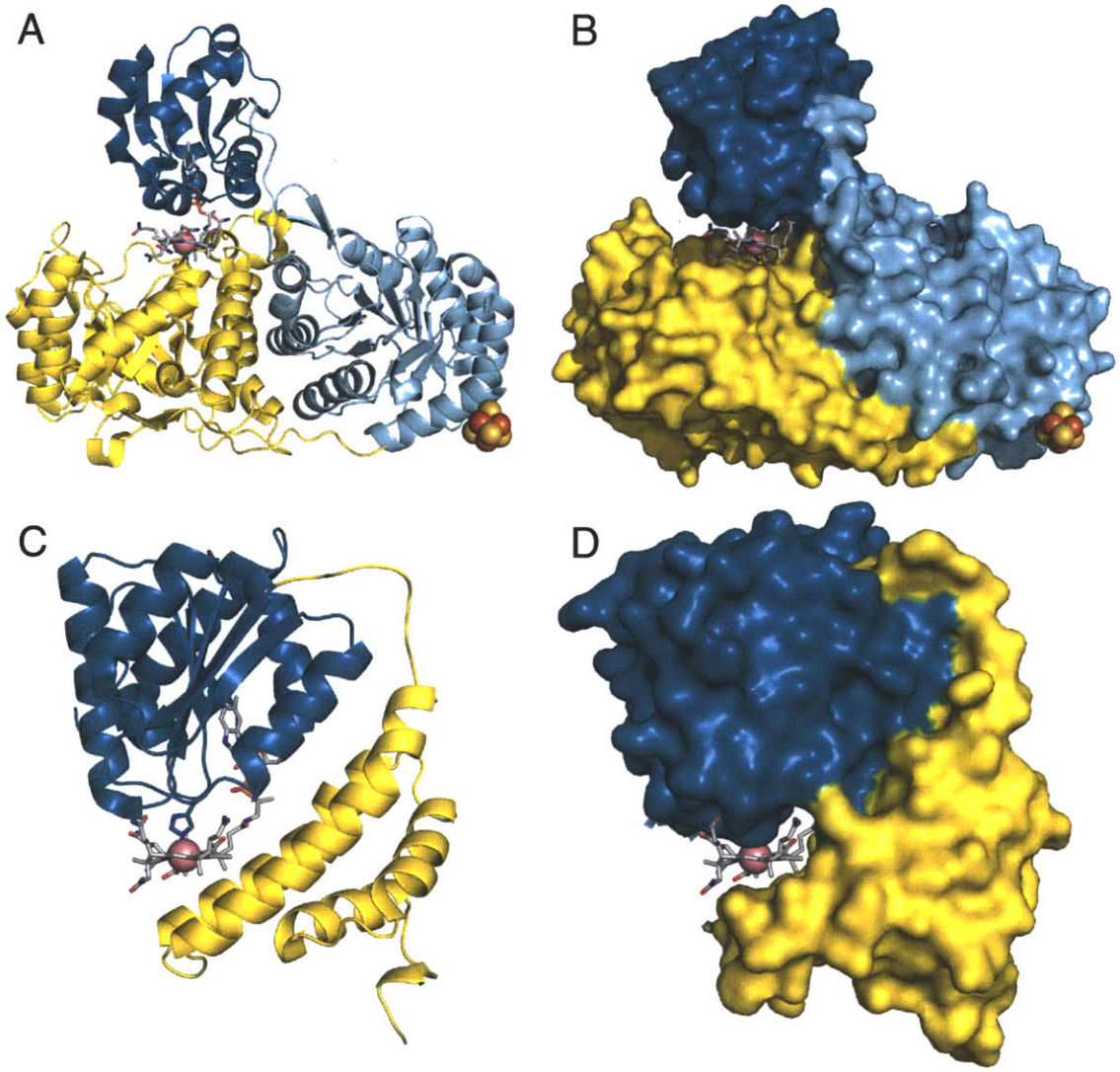


Figure II.22. The "capping" of the B_{12} cofactor. A-B, *Ch*CFeSP[13] (PDB ID: 2H9A) small subunit (yellow) caps the bottom face of B_{12}. *Ch*CFeSP large subunit B_{12} domain in dark cyan, central TIM in light cyan, and Fe_4S_4 cluster in spheres: iron in orange, sulfur in yellow. *Ch*CFeSP in **A** is shown in ribbons, while *Ch*CFeSP in **B** is shown in surface representation. **C-D**, MetH B_{12} domain[10] (PDB ID: 1BMT) "capping" subdomain (yellow) caps the bottom face of B_{12}. B_{12} binding domain in dark blue. MetH B_{12} domain in **C** is shown in ribbons, while MetH B_{12} domain in **D** is shown in surface representation. In **C**, His759, the axial ligand of the B_{12} cobalt, is shown as sticks (carbon in dark cyan, nitrogen in blue). B_{12} cofactors in sticks: carbon in grey, oxygen in red, nitrogen in blue, phosphorus in orange, and cobalt as a pink sphere.

II.H. References

1 Stubbe, J. (1994) Binding Site Revealed of Nature's Most Beautiful Cofactor. *Science.* **266**, 1663-1664.

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

3 Matthews, R. G. (2001) Cobalamin-Dependent Methyltransferases. *Acc. Chem. Res.* **34**, 681-689.

4 Matthews, R. G., Koutmos, M. & Datta, S. (2008) Cobalamin-Dependent and Cobamide-Dependent Methyltransferases. *Curr. Opin. Struct. Biol.* **18**, 658-666.

5 Bandarian, V., Pattridge, K. A., Lennon, B. W., Huddler, D. P., Matthews, R. G. & Ludwig, M. L. (2002) Domain Alternation Switches B_{12}-Dependent Methionine Synthase to the Active Conformation. *Nat. Struct. Biol.* **9**, 53-56.

6 Datta, S., Koutmos, M., Pattridge, K. A., Ludwig, M. L. & Matthews, R. G. (2008) A Disulfide-Stabilized Conformer of Methionine Synthase Reveals an Unexpected Role for the Histidine Ligand of the Cobalamin Cofactor. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4115-4120.

7 Dixon, M. M., Huang, S., Matthews, R. G. & Ludwig, M. L. (1996) The Structure of the C-terminal Domain of Methionine Synthase: Presenting *S*-adenosylmethionine for Reductive Methylation of B_{12}. *Structure.* **4**, 1263-1275.

8 Doukov, T., Seravalli, J., Stezowski, J. J. & Ragsdale, S. W. (2000) Crystal Structure of a Methyltetrahydrofolate- and Corrinoid-Dependent Methyltransferase. *Structure.* **8**, 817-830.

9 Doukov, T. I., Hemmi, H., Drennan, C. L. & Ragsdale, S. W. (2007) Structural and Kinetic Evidence for an Extended Hydrogen-bonding Network in Catalysis of Methyl Group Transfer: Role of an Active Site Asparagine Residue in Activation of Methyl Transfer by Methyltransferases. *J. Biol. Chem.* **282**, 6609-6618.

10 Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G. & Ludwig, M. L. (1994) How a Protein Binds B_{12}: A 3.0 Å X-ray Structure of B_{12}-Binding Domains of Methionine Sythase. *Science.* **266**, 1669-1674.

11 Evans, J. C., Huddler, D. P., Hilgers, M. T., Romanchuk, G., Matthews, R. G. & Ludwig, M. L. (2004) Structures of the N-terminal Module Imply Large Domain Motions During Catalysis by Methionine Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3729-3736.

12 Koutmos, M., Datta, S., Pattridge, K. A., Smith, J. L. & Matthews, R. G. (2009) Insights into the Reactivation of Cobalamin-Dependent Methionine Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18527-18532.

13 Svetlitchnaia, T., Svetlitchnyi, V., Meyer, O. & Dobbek, H. (2006) Structural Insights into Methyltransfer Reactions of a Corrinoid Iron-Sulfur Protein Involved in Acetyl-CoA Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14331-14336.

14 Banerjee, R. B. & Matthews, R. G. (1990) Cobalamin-Dependent Methionine Synthase. *FASEB J.* **4**, 1450-1459.

15 Ragsdale, S. W. & Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 Fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898.

16 Matthews, R. G. & Drummond, J. T. (1990) Providing One-Carbon Units for Biological Methylations: Mechanistic Studies on Serine Hydroxymethyltransferase, Methylenetetrahydrofolate Reductase, and Methyltetrahydrofolate-Homocysteine Methyltransferase. *Chem. Rev.* **90**, 1275-1290.

17 Tibbetts, A. S. & Appling, D. R. (2010) Compartmentalization of Mammalian Folate-Mediated One-Carbon Metabolism. *Annu. Rev. Nutr.* **30**, 57-81.

18 Scott, J. M. (1999) Folate and Vitamin B_{12}. *Proc. Nutr. Soc.* **58**, 441-448.

19 Lucock, M. (2000) Folic Acid: Nutritional Biochemistry, Molecular Biology, and Role in Disease Processes. *Mol. Genet. Metab.* **71**, 121-138.

20 Schrauzer, G. N. & Deutsch, E. (1969) Reactions of Cobalt(I) Supernucleophiles: The Alkylations of Vitamin B_{12s}, Cobaloximes(I), and Related Compounds. *J. Am. Chem. Soc.* **91**, 3341-3350.

21 Banerjee, R. B., Harder, S. R., Ragsdale, S. W. & Matthews, R. G. (1990) Mechanism of Reductive Activation of Cobalamin-Dependent Methionine Synthase: An Electron Paramagnetic Resonance Spectroelectrochemical Study. *Biochemistry.* **29**, 1129-1135.

22 Harder, S. R., Lu, W.-P., Feinberg, B. A. & Ragsdale, S. W. (1989) Spectroelectrochemical Studies of the Corrinoid/Iron-Sulfur Protein Involved in Acetyl Coenzyme A Synthesis by *Clostridium thermoaceticum*. *Biochemistry.* **28**, 9080-9087.

23 Menon, S. & Ragsdale, S. W. (1998) Role of the [4Fe-4S] Cluster in Reductive Activation of the Cobalt Center of the Corrinoid Iron-Sulfur Protein from *Clostridium thermoaceticum* during Acetate Biosynthesis. *Biochemistry.* **37**, 5689-5698.

24 Menon, S. & Ragsdale, S. W. (1999) The Role of an Iron-Sulfur Cluster in an Enzymatic Methylation Reaction. *J. Biol. Chem.* **274**, 11513-11518.

25 Zydowsky, T. M., Courtney, L. F., Frasca, V., Kobayashi, K., Shimizu, H., Yuen, L. D., Matthews, R. G., Benkovic, S. J. & Floss, H. G. (1986) Stereochemical

Analysis of the Methyl Transfer Catalyzed by Cobalamin-Dependent Methionine Synthase from *Escherichia coli* B. *J. Am. Chem. Soc.* **108**, 3152-3153.

26 Holm, L. & Rosenström, P. (2010) Dali Server: Conservation Mapping in 3D. *Nucleic Acids Res.* **38**, W545-W549.

27 Ragsdale, S. W., Lindahl, P. A. & Münck, E. (1987) Mössbauer, EPR, and Optical Studies of the Corrinoid/Iron-Sulfur Protein Involved in the Synthesis of Acetyl Coenzyme A by *Clostridium thermoaceticum*. *J. Biol. Chem.* **262**, 14289-14297.

28 Stich, T. A., Seravalli, J., Venkateshrao, S., Spiro, T. G., Ragsdale, S. W. & Brunold, T. C. (2006) Spectroscopic Studies of the Corrinoid/Iron-Sulfur Protein from *Moorella thermoacetica*. *J. Am. Chem. Soc.* **128**, 5010-5020.

29 Wirt, M. D., Kumar, M., Ragsdale, S. W. & Chance, M. R. (1993) X-ray Absorption Spectroscopy of the Corrinoid/Iron-Sulfur Protein in Acetyl Coenzyme A Synthesis by *Clostridium thermoaceticum*. *J. Am. Chem. Soc.* **115**, 2146-2150.

30 Wirt, M. D., Kumar, M., Wu, J.-j., Scheuring, E. M., Ragsdale, S. W. & Chance, M. R. (1995) Structural and Electronic Factors in Heterolytic Cleavage: Formation of the Co(I) Intermediate in the Corrinoid/Iron-Sulfur Protein from *Clostridium thermoaceticum*. *Biochemistry.* **34**, 5269-5273.

31 Zhao, S., Roberts, D. L. & Ragsdale, S. W. (1995) Mechanistic Studies of the Methyltransferase from *Clostridium thermoaceticum*: Origin of the pH Dependence of the Methyl Group Transfer from Methyltetrahydrofolate to the Corrinoid/Iron-Sulfur Protein. *Biochemistry.* **34**, 15075-15083.

32 Seravalli, J., Zhao, S. & Ragsdale, S. W. (1999) Mechanism of Transfer of the Methyl Group from (*6S*)-Methytetrahydrofolate to the Corrinoid/Iron-Sulfur Protein Catalyzed by the Methyltransferase from *Clostridium thermoaceticum*: A Key Step in the Wood-Ljungdahl Pathway of Acetyl-CoA Synthesis. *Biochemistry.* **38**, 5728-5735.

33 Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. (1999) Natural Engineering Principles of Electron Tunnelling in Biological Oxidation-Reduction. *Nature.* **402**, 47-52.

34 Bandarian, V., Ludwig, M. L. & Matthews, R. G. (2003) Factors Modulating Conformational Equilibria in Large Modular Proteins: A Case Study with Cobalamin-Dependent Methionine Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8156-8163.

35 Jarrett, J. T., Amaratunga, M., Drennan, C. L., Scholten, J. D., Sands, R. H., Ludwig, M. L. & Matthews, R. G. (1996) Mutations in the B_{12}-Binding Region of

Methionine Synthase: How the Protein Controls Methylcobalamin Reactivity. *Biochemistry.* **35**, 2464-2475.

36 Bourgeois, D. & Royant, A. (2005) Advances in Kinetic Protein Crystallography. *Curr. Opin. Struct. Biol.* **15**, 538-547.

37 Johnson, S. J., Taylor, J. S. & Beese, L. S. (2003) Processive DNA Synthesis Observed in a Polymerase Crystal Suggests a Mechanism for the Prevention of Frameshift Mutations. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3895-3900.

38 Lin, T., Chen, Z., Usha, R., Stauffacher, C. V., Dai, J.-B., Schmidt, T. & Johnson, J. E. (1999) The Refined Crystal Structure of Cowpea Mosaic Virus at 2.8 Å Resolution. *Virology.* **265**, 20-34.

39 Elliott, J. I. & Brewer, J. M. (1978) The Inactivation of Yeast Enolase by 2,3-butanedione. *Arch. Biochem. Biophys.* **190**, 351-357.

40 Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307-326.

41 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007) *Phaser* Crystallographic Software. *J. Appl. Cryst.* **40**, 658-674.

42 Brünger, A. T., Adams, P. D., Clore, M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998) *Crystallography & NMR System*: A New Software Suite for Marcromolecular Structure Determination. *Acta Cryst.* **D54**, 905-921.

43 Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010) *PHENIX*: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Cryst.* **D66**.

44 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010) Features and Development of Coot. *Acta Cryst.* **D66**, 486-501.

45 Laskowski, R. A., MacArthur, M. W. & Thornton, J. M. (1993) PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Cryst.* **26**, 283-291.

46 Ljungdahl, L. G., LeGall, J. & Lee, J.-P. (1973) Isolation of a Protein Containing Tightly Bound 5-Methoxybenzimidazolylcobamide (Factor IIIm) from *Clostridium thermoaceticum*. *Biochemistry.* **12**, 1802-1808.

47 DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA (http://www.pymol.org).

48 Zehnder, A. J. B. & Wuhrmann, K. (1976) Titanium(III) Citrate as a Nontoxic Oxidation-Reduction Buffering System for the Culture of Obligate Anaerobes. *Science.* **194**, 1165-1166.

49 Royant, A., Carpentier, P., Ohana, J., McGeehan, J., Paetzold, B., Noirclerc-Savoye, M., Vernède, X., Adam, V. & Bourgeois, D. (2007) Advances in Spectroscopic Methods for Biological Crystals. 1. Fluorescence Lifetime Measurements. *J. Appl. Cryst.* **40**, 1105-1112.

50 Barstow, B., Ando, N., Kim, C. U. & Gruner, S. M. (2008) Alteration of Citrine Structure by Hydrostatic Pressure Explains the Accompanying Spectral Shift. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13362-13366.

Chapter III: Structural insight into CFeSP/MeTr complexes in solution

III.A. Summary

Enzymes responsible for B_{12}-dependent methyl transfer adopt large complexes composed of multiple domains that are each responsible for discrete functions of the catalytic cycle. The Wood-Ljungdahl carbon fixation pathway contains one such B_{12} methyltransferase complex, where the corrinoid iron-sulfur protein (CFeSP) harbors the B_{12} cofactor. B_{12} in this system carries a methyl group from methyltetrahydrofolate (CH_3-H_4folate) bound in the active site of CH_3-H_4folate:CFeSP methyltransferase (MeTr) to the downstream enzyme acetyl coenzyme A synthase (ACS). This process involves the formation of a CFeSP/MeTr complex, whose crystal structure is currently the only available structure depicting all domains of a B_{12} methyltransferase complex. Because a single snapshot is not sufficient to describe the various subunit constructions that must occur in the cell, further structural characterization of the quaternary assembly of components that make up the complex is necessary. Here, CFeSP and MeTr alone and in complex have been characterized in solution by anaerobic small-angle X-ray scattering (SAXS). These studies show that the structures of free CFeSP and MeTr in solution both match previously determined crystal structures of these proteins alone. However, two different CFeSP/MeTr complexes with different oligomeric compositions, one of which resembles the CFeSP/MeTr crystal structure, are observed to form depending upon the solution conditions. These surprising results have led to a new model for this B_{12}-dependent methyltransferase where multiple active complexes with modular assemblies can form to allow an inherent adaptability to various intracellular conditions.

A manuscript similar to this chapter is currently in preparation for publication:
Ando, N., Kung, Y., Bender, G., Ragsdale, S.W., and Drennan, C.L. (2011)

III.B. Acknowledgements

Experimental design, sample preparation, and SAXS data collection were carried out collaboratively with Nozomi Ando, who performed data manipulation and analyses. Proteins were purified by Güneş Bender in the laboratory of Stephen W. Ragsdale at the University of Michigan. Sample preparations were performed in part in the laboratories of Sol M. Gruner and Brian R. Crane at Cornell University.

III.C. Introduction

Derivatives of vitamin B_{12}, a cobalt-containing organometallic cofactor, are used as methyl group carriers in a variety of biological contexts, from methionine biosynthesis and regeneration by methionine synthase (MetH) to anaerobic CO_2 fixation by acetogenic bacteria, where the B_{12}-containing corrinoid iron-sulfur protein (CFeSP) together with its methyltransferase (MeTr) play a key role. Enzymes and enzyme complexes that are responsible for B_{12}-dependent methyl transfer are large and multimodular, consisting of numerous domains. CFeSP and MeTr work together in a CFeSP/MeTr complex to catalyze two methyl transfer reactions (Figure III.1), first from methyltetrahydrofolate (CH_3-H_4folate), a folic acid (vitamin B_9) derivative that binds MeTr, to the B_{12} cofactor bound to CFeSP, and then to the Ni-Fe-S active site metallocluster of acetyl coenzyme A synthase (ACS). At the A-cluster, the methyl group is combined with CO and coenzyme A (CoA) to form the methyl moiety of acetyl-CoA in the final steps of the acetogenic Wood-Ljungdahl carbon fixation pathway. In this way, half of the carbon atoms fixed when acetogens grow autotrophically must transit through the CFeSP/MeTr B_{12}-dependent methyltransferase system.

CFeSP contains two polypeptide chains: a small subunit (35 kDa) that adopts a (β/α) triosephosphate isomerase (TIM) barrel fold, and a large subunit (48 kDa) that contains three domains connected by linkers: an N-terminal domain that harbors a Fe_4S_4 cluster responsible for activation of the B_{12} cofactor, a central TIM barrel domain, and a C-terminal B_{12}-binding domain. A crystal structure of CFeSP from *Carboxydothermus hydrogenoformans* has been determined that shows a heterodimer of small and large subunits[1]. The Fe_4S_4 domain was disordered and not modeled in this structure. MeTr is a

57 kDa homodimer of TIM barrel monomers, each of which may bind one molecule of CH_3-H_4folate. Crystal structures of MeTr from *Moorella thermoacetica*, the model acetogenic organism, have been obtained and show a homodimeric species[2,3]. In addition, a crystal structure has been solved containing a complex of CFeSP/MeTr from *M. thermoacetica* (Chapter II), where two CFeSP heterodimers of small and large subunits are bound on either side of one central MeTr homodimer. This ~220 kDa complex thus contains 2:1 stoichiometry of CFeSP and the MeTr homodimer.

Such large enzyme complexes are required for the B_{12}-dependent transfer of a single methyl group because of the many steps performed during the reaction cycle, each occurring in separate protein domains. In CFeSP, B_{12} is bound in the Rossmann-like fold of the C-terminal B_{12} domain of the large subunit, while the substrate CH_3-H_4folate is bound to the MeTr active site. During successive methyl transfer reactions between CH_3-H_4folate, B_{12}, and the ACS A-cluster, the cobalt ion of B_{12} cycles between the powerfully nucleophilic Co(I) species and the methylated CH_3-Co(III) species. However, both of these B_{12} states are reactive and must be protected to avoid side-reactions. Protection is achieved by a "capping" domain, which in the case of CFeSP is the small subunit TIM barrel itself, that sequesters the reactive face of the B_{12} macrocycle before catalysis[1]. Despite this protective mechanism, B_{12} with Co(I) is sensitive to oxidation, which occasionally inactivates B_{12} to the Co(II) state[4]. CFeSP thus harbors yet another domain to allow the cofactor to re-enter the catalytic cycle. Here, an electron from a partner protein, such as ferredoxin, carbon monoxide dehydrogenase (CODH), and hydrogenase[4,5], is first passed to an Fe_4S_4 cluster harbored by the CFeSP large subunit's

N-terminal Fe_4S_4 domain and then to Co(II) of the B_{12} cofactor, reducing it back to the active, nucleophilic Co(I) state.

Given that B_{12} binding, activation, protection, and catalytic functions are performed by separate domains in B_{12}-dependent methyl transfer systems, B_{12}-dependent methyltransferases are often described as adopting a "beads on a string" topology in their primary sequences[6]. However, structural characterization of the quaternary organization of these domains has been challenging. For example, structures of individual MetH fragments are available, but no structural model yet exists of the full-length MetH enzyme. A structure of the CFeSP/MeTr complex mentioned above that exhibits 2:1 stoichiometry of CFeSP and the MeTr homodimer is the only crystal structure containing all domains in any B_{12}-dependent methyl transfer system (Chapter II).

The CFeSP/MeTr crystal structure provides an insightful snapshot of an assembled B_{12}-dependent methyltransferase complex; however, studies that investigate the relationship between the quaternary structures of individual and complexed components in solution for any system have been lacking. In this study, we used anaerobic small-angle X-ray scattering (SAXS) to explore the quaternary structure of CFeSP/MeTr components under different solution conditions. Our results show that under all concentrations tested, MeTr maintains homodimeric association, while CFeSP remains a heterodimer of large and small subunits. However, when the CFeSP heterodimer CFeSP and the MeTr homodimer are mixed, they unexpectedly form two different CFeSP/MeTr complexes with different subunit compositions, depending upon the solution conditions. The ability to form complexes of different quaternary structures

may reflect an inherent flexibility and modularity of the acetogenic CFeSP/MeTr methyltransferase system.

III.D. An overview of the mechanics and utility of small-angle X-ray scattering

X-ray crystallography of biological macromolecules has led an explosion of atomic-resolution structural information over the last few decades. Crystal structures can present molecular details of complex biological processes, offering structural and mechanistic insight into protein function. However, other structural techniques can also be employed to address biochemical questions that are less or not at all accessible by X-ray crystallography. In this chapter, one such technique, small-angle X-ray scattering (SAXS), is exploited to gain additional knowledge into the CFeSP/MeTr B_{12}-dependent methyl transfer system that is not otherwise available using crystallography alone.

The use of SAXS has a long history in biological experimentation[7], but a more recent and rapid surge in the development of molecular biology, instrumentation, and computation technologies has given rise to a large growth in the application and relative ease of this technique[8]. Although SAXS does not offer atomic resolution, it has many advantages that make it an attractive tool in a structural biologist's toolkit. SAXS is a solution-based technique, and so crystallization of the protein of interest is not required. The solution conditions can also be varied, allowing for the structural dependence on components in the solution to be easily probed, provided that these do not lead to aggregation. Information regarding the size and shape of species in solution may be obtained, and in many cases electron density may also be determined *ab initio*, yielding a shape reconstruction that can be presented as a surface envelope around the particle in

question. Fitting shape reconstructions with known crystal structures can be particularly informative and may reveal changes in size, shape, or oligomerization. In addition, as we shall discover, deconvolution analysis of SAXS data can provide information about the distribution of multiple species that may be present simultaneously in solution. Inherent to all of these applications are assumptions and limitations that must be addressed to justify the validity of the employed course of analysis.

Like X-ray crystallography, SAXS is an X-ray technique where the sample is placed in a transmission setup in the path of a highly collimated X-ray beam (Figure III.2); monochromatic X-rays derived from synchrotron radiation are most often employed. As X-rays hit the sample, they are scattered by electrons within each molecule as they pass through. Downstream of the sample, the scattered X-rays continue to travel through a vacuumed flight tube, typically ~1-2 m in length, to minimize additional scattering from air before reaching a two-dimensional detector, typically a charge-coupled device (CCD), which records the intensity of the scattered X-rays. Because particles in the sample tumble freely in solution as the X-rays pass, the X-ray scattering is radially symmetric about the center of the beam, unlike with X-ray crystallography where diffraction reflections or "spots" are observed due to the regularly ordered arrangement of particles in the crystal lattice. Thus, the recorded scattering appears as a diffuse pattern that is most intense at the center and becomes weaker farther from the center. A beam stop is used to absorb the direct X-ray beam and to protect the detector from damage, and the size and position of the beam stop can be important factors to recording the lowest-angle scattering of the sample. The beam stop is often equipped with a photodiode to measure the transmission intensity, which is used for normalizing data before background

subtraction. The scattering angles in SAXS are relatively small, giving rise to the name of the technique, with most observed scattering intensities residing within 6° of the direct beam.

The scattering intensities recorded on the detector are then azimuthally integrated and background-subtracted to generate a one-dimensional SAXS or scattering profile, a plot of the measured scattering intensity (I) as a function of the magnitude of the scattering vector, q:

$$q = \frac{4\,\pi\,\sin\theta}{\lambda}$$

where θ is half the scattering angle and λ is the wavelength of the incident X-rays; thus q has units of $Å^{-1}$. Because λ is fixed and θ is small (<3°), the scattering profile is an $I(q)$ plot that can be thought of as the scattering intensity as a function of the scattering angle. In this way, the scattering profile, $I(q)$, is essentially a one-dimensional plot of the measured scattering intensity versus the distance from the beam center in the detector image.

Scattering intensities decrease with increased distance from the beam center, and thus the $I(q)$ scattering profile appears roughly Gaussian in shape at low q. The precise shape of the scattering profile is directly related to the shape of the species in solution. Overall, larger species contain more electron density and lead to higher scattering intensities at low q, giving rise to a tighter and taller Gaussian (Figure III.2). A higher concentration of molecules will thus increase the *height* of the profile, but the *shape* of the curve will remain constant if the size and shape of the species in solution is also constant. Aside from such crude observations, the finer details inherent to the scattering

profile are not immediately intuitive, and further analysis must be performed to extract more detailed information.

First, it is useful to carry out the Fourier transform of the scattering profile, which gives the inter-electron distance distribution function, $P(r)$, of the scattering molecule. In crystallography, such a Fourier transform of the diffraction amplitudes gives the Patterson function, and in the same way, $P(r)$ is analogous to a one-dimensional Patterson. Thus the $P(r)$ versus r plot shows the frequency of all inter-electron distances within the scattering molecule, and can provide valuable information on the shape of the species in solution.

Second, the Guinier approximation can be applied to the scattering profile to compute additional parameters: the forward or zero-angle scattering intensity, $I(0)$, and the radius of gyration, R_g. The Guinier approximation simply states that at low scattering angles and thus low q, the decrease in scattering intensity is roughly Gaussian. Mathematically, this can be expressed as:

$$I(q) \approx I(0)\, e^{\frac{-q^2 R_g^2}{3}}$$

$I(0)$, which is the scattering intensity at the zero scattering angle, cannot be experimentally determined, not only because the beam stop blocks all X-rays at the zero scattering angle, but also because the *scattering* contribution to the intensity at zero angle cannot be separated from the direct transmitted X-ray beam. As noted in the above expression, R_g of the species in solution can also be computed and is related to the width of the Gaussian (Figure III.2).

The Guinier approximation also allows the construction of a Guinier plot, which is $\ln(I)$ versus q^2. Because the Guinier approximation assumes a Gaussian relationship

between I and q, the Guinier plot should thus be linear; in fact, linearity of the Guinier plot is used as validation of sample quality, confirming that sample aggregation has not occurred and that inter-particle interactions are minor contributions to the scattering, relative to the electrons of the molecules in the sample themselves.

Finally, the *ab initio* shape reconstruction of species in solution can also be determined from the scattering profile. Here, small virtual beads or "dummy" residues are assembled together computationally using various physical restraints to form a hypothetical protein shape whose theoretical scattering profile is computed. Using iterative rounds of simulated annealing, these "dummy" beads or residues are rearranged, minimizing differences between the calculated theoretical scattering profile with the experimentally derived scattering profile, until an adequate match is obtained. Many such runs, typically around ten, are performed and averaged to give the final shape reconstruction, which represents the most probable model for the scattering species. The quality of the average model is represented in the normalized spatial discrepancy (NSD) value: 0 indicates identical objects, while values >1 are considered non-unique and are thus unstable model solutions[9].

III.E. Results

For all anaerobic SAXS experiments described below, proteins samples were of the same as-isolated form as previously used in crystallization of the CFeSP/MeTr complex, with the CFeSP B_{12} cobalt atom primarily in the Co(II) oxidation state. The CH_3-H_4folate substrate was not added during any of these experiments, and thus turnover did not occur. Anaerobic SAXS has been previously performed using sealed, enclosed

flow-cells[10], epoxy-sealed capillaries[11,12], and SAXS cells with flat mica windows that were sealed either with clay[13,14] or with O-rings connected to syringes that were filled with protein samples[11,15,16]. In this study, both acrylic SAXS cells with flat Kapton windows and glass X-ray capillaries were filled with samples in the anaerobic chamber and sealed with epoxy to maintain anaerobic conditions prior to data collection (see Methods for details). Epoxy-sealed capillaries have been traditionally used for in-vacuum SAXS setups and is a simple and reliable method for anaerobic containment. The oxygen permeability of the epoxy-sealed cells with Kapton windows was tested spectroscopically with methylviologen and was found to be insignificant over the time course of the experiments (~5 min of air exposure).

Structural behavior of CFeSP and MeTr alone in solution

Prior to the structural characterization of *M. thermoacetica* CFeSP/MeTr complexes in solution using anaerobic SAXS, the solution behavior of the individual subunits must first be well understood. To this end, free MeTr and CFeSP were examined in the standard *in vitro* assay buffer (50 mM Tris, pH 7.6, 100 mM NaCl, 2 mM dithiothreitol) at varying concentrations below and above the presumed physiological concentration of ~50 μM[17]. The background-subtracted scattering profiles, $I(q)$, for both proteins are shown in Figure III.3.A-B. Guinier analysis was used to assess data quality and determine the overall radius of gyration, R_g, which is a function of spatial size[18]. The Guinier plots, $\ln(I)$ versus q^2, display linearity at low q for both proteins at all concentrations (Figure III.4.A-B), indicating an absence of non-specific aggregation caused by radiation damage.

R_g values are useful to investigate the oligomerization states of both CFeSP and MeTr in solution. R_g determined from linear fits over the q range satisfying the Guinier approximation ($q \cdot R_g < 1.3$) show minimal concentration dependence, with R_g decreasing slightly with increasing protein concentration (Figure III.4.C-D). This small decrease is attributed to small inter-particle repulsions, such as excluded volume interactions. The linear extrapolation to zero protein concentration eliminates these repulsive effects and gives R_g values for MeTr and CFeSP of 27.6 ± 0.4 Å and 31.1 ± 0.7 Å, respectively. These experimental R_g values agree well with the theoretical values of 26.8 Å calculated from the *M. thermoacetica* homodimeric MeTr crystal structure[3] and 29.8 Å from heterodimeric CFeSP composed of large and small subunits from the CFeSP/MeTr crystal structure, respectively. It should be noted that the theoretical R_g of the CFeSP structure from *C. hydrogenoformans*[1], which is the only crystal structure available of a CFeSP alone, is too small at 28.0 Å, when compared with the experimental R_g determined here, as this crystal structure lacks the Fe_4S_4 domain. Modeling the Fe_4S_4 domain from the *M. thermoacetica* CFeSP/MeTr structure into the *C. hydrogenoformans* structure by aligning their CFeSP large subunit TIM barrel domains (root mean square deviation for C_α atoms: 0.60 Å) leads to a larger theoretical R_g of 29.6 Å, which is in better agreement with the experimental R_g and likely arises from an extended Fe_4S_4 domain conformation in solution. This model is also the most realistic model for a full-length CFeSP free in solution: the *C. hydrogenoformans* CFeSP structure is the only crystal structure available of a CFeSP alone, but it lacks the Fe_4S_4 domain. By simply appending this domain from the *M. thermoacetica* CFeSP/MeTr structure, a model for a full-length CFeSP alone in solution can be obtained. Combining the known structural

elements of CFeSP from both organisms for the purpose of SAXS analysis is further valid because both CFeSPs are of similar size (~83 kDa) and shape (Figure II.3.B). This model for CFeSP thus represents a "hybrid" of structures that will be referred to as the "hybrid" model and will be used in the rest of the SAXS analysis.

The molecular weights of the individual subunits in solution can be quantified by several SAXS analytical methods to further investigate the oligomerization states of free CFeSP and MeTr. Here, a Porod invariant method, which relies on the area under the curve of $I \cdot q^2$ versus q, was employed[19]. Using data collected at 100 μM protein concentration up to a maximum q of 0.2 $Å^{-1}$, the molecular weights were calculated to be 53 and 83 kDa for MeTr and CFeSP, respectively. These values are in excellent agreement to the actual molecular weights of homodimeric MeTr and heterodimeric CFeSP of 57 and 83 kDa, respectively. Because the overall sizes as measured by R_g do not change significantly with protein concentration over the investigated range (Figures III.4.C-D), the oligomerization states also do not change for either protein with increasing concentration. Therefore, MeTr remains a homodimer and CFeSP remains a heterodimer of small and large subunits that both match their respective crystal structures at all concentrations tested.

The solution conformations of MeTr and CFeSP can be further compared to their crystal structures by fitting the theoretical scattering profiles calculated from the atomic coordinates to the experimental data[20]. The calculated scattering profile of the MeTr homodimeric crystal structure shows an excellent fit to the data collected at 25 μM protein concentration, while that of only a MeTr monomer generates a poor fit (Figure III.3.A). This result agrees with the molecular weight estimations discussed above as well

as previously reported gel filtration results, which suggested that MeTr is a homodimer in solution[21]. Consistent with the R_g analysis, the experimental scattering profile of CFeSP also agrees well with the theoretical scattering profile calculated from the "hybrid" CFeSP model (Figure III.3.B).

Ab initio shape reconstructions of the solution conformations were then determined for direct comparison with the crystal structures in real space (Table III.1). For MeTr, ten bead models were generated from scattering data collected at 100 μM homodimer concentration in SAXS sample cells with fixed path lengths in the program GASBOR[22], using *P*2 symmetry and a total of 562 dummy residues. The averaged model is shown in Figure III.5.A and exhibits a normalized spatial discrepancy (NSD) of 0.925 (var = 0.011). The crystal structure of homodimeric MeTr[3] fits well into the reconstructed density for this protein. Similarly, ten bead models of CFeSP were generated from data collected at 100 μM using the program DAMMIF[23], and averaged by the program DAMAVER[24] to generate the most probable model, NSD: 0.971 (var = 0.197) (Figure III.5.B). The reconstructed density shows excellent alignment with the core of the CFeSP "hybrid" model, composed of the small and large subunit TIM barrels and the B_{12} domain. However, the Fe_4S_4 domain of the "hybrid" model protrudes from the density in this alignment. This fit could be interpreted to indicate that the Fe_4S_4 domain in solution is closely packed to the CFeSP core. However, in the R_g analysis described above, the contribution from an extended Fe_4S_4 domain was important to explaining the magnitude of the experimental R_g. Therefore, the absence of density in the shape reconstruction is more likely due to flexibility of the Fe_4S_4 domain in solution, which samples many conformations whose positions are averaged out. This interpretation is consistent with the

crystal structure of CFeSP alone from *C. hydrogenoformans* that also contains a disordered Fe_4S_4 domain. Such domain flexibility is also manifested in the relatively high NSD variance of the CFeSP reconstructions, while that of MeTr is in a typical range.

Having established that the solution conformations of MeTr and CFeSP are well-described by their respective crystal structures, we compared their forward scattering intensities, $I(0)$, to confirm their relative protein concentrations with the spectroscopically determined concentrations measured by both Bradford[25] and Rose-Bengal[26] methods. For this analysis, $I(0)$ was determined from scattering data collected over a range of protein concentrations using SAXS sample cells with fixed path-lengths, as described in the Methods below, in order to minimize any effect on the scattering intensity due to varying path-lengths instead of varying concentrations between samples. $I(0)$ was then normalized by the square of the molecular weight (MW^2), to give values that are proportional to the molar protein concentrations (Figure 6). The linear relationship in the plot of $I(0)/MW^2$ versus concentration confirms the previously described observation that the oligomerization states for both MeTr and CFeSP, reflected in the MW term, remain constant with increased protein concentration. More importantly, however, the molecular weight-normalized data points for both MeTr and CFeSP plots are co-linear, indicating that the SAXS-derived relative concentrations of both protein samples are accurate to their previously measured concentrations. Confirmation of the relative concentrations is of crucial importance for subsequent experiments involving complexes between CFeSP and MeTr, where the two proteins are mixed.

With clear validation that MeTr assembles as a homodimer and CFeSP as a heterodimer of large and small subunits that closely resemble their crystal structures over

all concentrations tested, we may now proceed to experiments where the two protein samples are mixed together to investigate CFeSP/MeTr complexes.

Titration of CFeSP into MeTr

To determine the subunit composition and stoichiometry of the CFeSP/MeTr complex in solution, CFeSP was titrated into 50 μM MeTr homodimer under the standard *in vitro* assay buffer conditions, such that the molar ratio of the two proteins varied across a wide range. The total protein concentration in these samples spanned 2.9-15.4 mg/mL, which is within the typical range studied by SAXS. The resultant scattering profiles (Figure III.7.A) display good linearity at low q when represented as Guinier plots (Figure III.8.A), indicating that inter-particle interactions are small and the solutions can be considered dilute. This validation thus allows the scattering profiles to be described as the sum of scattering contributions for individual species in solution. The titration mixtures were therefore decomposed and deconvoluted into dissociated and associated species by fitting linear combinations of scattering from individual species to the experimental profiles[18].

In this deconvolution of the scattering profiles, four possible species were considered. First, uncomplexed MeTr or CFeSP may be present. Due to the experiments described above investigating these proteins alone in solution, uncomplexed MeTr and CFeSP are composed of the MeTr homodimer and the CFeSP heterodimer of small and large subunits, respectively, over the concentration range studied. When CFeSP and MeTr samples are mixed, a third possible species representing the complex obtained in the CFeSP/MeTr crystal structure may also form. The crystal structure depicts two

CFeSPs bound to one MeTr homodimer; this possible species in solution can be described as having 2:1 stoichiometry (2CFeSP:1MeTr homodimer), and will be referred to as the 2:1 complex. Finally, because the crystal structure shows that each CFeSP binds to distinct and analogous sites of independent MeTr monomers, it is conceivable that upon titration of CFeSP into MeTr a fourth species could exist that contains just one CFeSP per MeTr homodimer. This 1CFeSP:1MeTr homodimer species thus exhibits 1:1 stoichiometry and will be referred to as the 1:1 complex.

In total, linear combinations of the theoretical scattering profiles for each of these four species were fit to the experimental scattering profiles for each sample of the titration to determine the individual contributions in solution for CFeSP, MeTr homodimer, 1:1 complex, and 2:1 complex. Theoretical scattering profiles were calculated in CRYSOL[20] and linear combinations were fit to the data in OLIGOMER[27] over the q range 0.0179-0.1634 $Å^{-1}$. Higher q data was not included, as they are less accurate due to errors in background subtraction. Scattering profile fits are shown in Figure III.7.A, and the resulting, deconvoluted volume fractions of each species over the course of the titration are shown in Figure III.7.C. The deconvolution analysis illustrates a clear trend: as CFeSP is added, free MeTr homodimer is consumed to form the 1:1 complex. This complex is maximally formed at equimolar concentrations of CFeSP and MeTr homodimer, which is consistent with the equimolar subunit stoichiometry. Interestingly, the 2:1 complex observed in the crystal structure does not form in appreciable amounts, even in the presence of excess CFeSP. Instead, unbound CFeSP alone is observed to build up. As the standard assay buffer was used in these experiments, the 1:1 complex formed is active and able to perform B_{12}-dependent methyl transfer.

However, because the 2:1 complex is observed in the crystal structure, it must also form under some solution condition. A previous study had shown that addition of the crystallization precipitant solution can stabilize a larger oligomer observed in the crystal structure that does not form when the crystallization solution is absent[28]. Therefore, we repeated the titration experiment in a solution that precisely mimics the crystallization condition, a 2:1 mixture of assay buffer and the crystallization precipitant (100 mM Bis-Tris, pH 6.5, 100 mM calcium acetate, 9% polyethylene glycol monomethylether 5000, and 20% glycerol). The MeTr homodimer concentration was again fixed at 50 μM, and increasing amounts of CFeSP were titrated. Scattering profiles were obtained (Figure III.7.B) and gave linear Guinier plots (Figure III.8.B), which verified that the contributions to the scattering from inter-particle interactions were minimal. Scattering profiles were then linearly deconvoluted, providing good fits (Figure III.7.B). The resulting volume fractions of each species were again plotted against the increasing CFeSP concentrations over the course of the titration experiment to give clear trends (Figure III.7.D). As in the previous titration, free MeTr homodimer is consumed to form complexes as more CFeSP is added. However, in the crystallization condition, both 1:1 and 2:1 complexes were formed. At low ratios of CFeSP:MeTr homodimer concentrations, the smaller 1:1 complex is favored, while the larger 2:1 complex becomes favored in the presence of excess CFeSP.

Although the scattering profiles of the 1:1 and 2:1 CFeSP/MeTr complexes are difficult to differentiate visually, by converting them into the real space function, $P(r)$, a more intuitive comparison can be determined. As described above, $P(r)$ shows the distribution of electron-pair distances in the protein electron density. The longest

distance, D_{max}, thus corresponds to the maximum protein dimension. Because the precision of D_{max} is dependent on the lowest q data, which is sensitive to sample quality, small differences in D_{max} values (such as 10 Å) cannot be distinguished. However, D_{max} values can be used to distinguish 40 Å differences, which is the change in the maximum dimensions between the 1:1 and 2:1 complexes based on the CFeSP/MeTr crystal structure. Scattering profiles from various titration experiments described above were converted into $P(r)$ by the indirect Fourier transform method implemented in the program GNOM[29] (Figure III.9). Consistent with the formation of a 1:1 complex, the observed D_{max} for experiments in assay buffer did not exceed 100 Å, the maximum dimension of the 1:1 complex, even in the presence of excess CFeSP (Figure III.9). However, in the crystallization condition, D_{max} increased to 140 Å, the maximum length in the 2:1 complex. In addition, the predominant peak remains at a similar radius, indicative of a lengthening of a particle that maintains the same cross-sectional radius. These results strongly support the deconvolution analyses, which indicate that the 1:1 complex is the dominant form in solution under the standard assay condition, while the 2:1 complex can be formed in the crystallization condition.

While the crystal structure of the CFeSP/MeTr complex gave a detailed depiction of the 2:1 complex, the solution experiments in assay buffer are the first to structurally demonstrate a 1:1 complex of CFeSP and MeTr homodimer. Therefore, an additional experiment was performed to further characterize the 1:1 complex in solution. Mixtures of CFeSP and MeTr homodimer with the molar ratio fixed at 1:1 were examined over a wide range of concentrations (25-150 μM). The resultant R_g values determined from Guinier analysis show a slight and linear decrease with increasing total protein

concentration (Figure III.10), which supports the fact that higher order complexes (such as a 2:1 complex) are not favored even at high protein concentrations. If a larger species did form, the line would instead transition upwards. A linear extrapolation to zero protein concentration to remove the repulsive inter-particle effects yields R_g of 35.5 ± 0.5 Å, which is in excellent agreement with the theoretical value of 34.8 Å for a 1:1 complex extracted from CFeSP/MeTr crystal structure, where one CFeSP is simply removed. In contrast, the larger 2:1 complex with theoretical R_g of 43.2 Å is too large and does not agree with the experimental R_g value. Further, the data-point that represents the sample at lowest protein concentration (25 μM) also lies on this line, indicating that the same 1:1 complex is still present without dissociation into individual CFeSP and MeTr homodimer protein components. This observation is consistent with a previously estimated CFeSP/MeTr dissociation constant in the low micromolar range[30].

In addition, an *ab initio* shape reconstruction of the 1:1 complex was obtained using a 100 μM sample in assay buffer that contained a 1:1 molar ratio of CFeSP and MeTr homodimer. Ten bead models were generated from the scattering data using the program DAMMIF[23] and averaged in DAMAVER[24] to obtain the most probable model, NSD: 0.811 (var = 0.048) (Figure III.11). A model structure of the 1:1 complex, in which one CFeSP was simply removed from the CFeSP/MeTr crystal structure, fits well into the reconstructed density. Only the B_{12} and Fe_4S_4 domains protrude slightly from the reconstruction density in this alignment, again consistent with the CFeSP/MeTr crystal structure, where these domains are disordered, exhibit higher *B*-factors, and contain weaker electron density.

The effect of polyethylene glycol

Because the addition of crystallization buffer was necessary to form the same 2:1 complex in solution that was observed in the crystal structure, we additionally investigated which component of the precipitant contributed to this shift in oligomerization state from the 1:1 complex to the 2:1 complex. Of all crystallization solution components, a major effect on the formation of the 2:1 complex was only observed with polyethylene glycol monomethyl ether (PEG MME) 5000. As Figure III.12.A shows, an increasing concentration of PEG MME 5000 leads to the formation of a species that exhibits D_{max} approaching 140 Å, which is consistent with the 2:1 complex in the crystal structure. At the highest concentration of PEG MME 5000 tested (12%), even higher-order protein association is observed. For comparison, the crystallization condition contains 3% PEG MME 5000, as the precipitant solution with 9% PEG MME 5000 was diluted three-fold into the assay buffer for each crystallization drop. Therefore, the addition of PEG MME 5000 is sufficient and necessary for the formation of complexes larger than the 1:1 complex.

R_g analysis, however, shows that the 2:1 complex does not form exclusively (Figure III.12.B). As stated above, the calculated theoretical R_g values for the 1:1 and 2:1 complex based on the CFeSP/MeTr crystal structure are 34.8 Å and 43.2 Å, respectively. It is clear that with the addition of PEG MME 5000 the R_g values are greater than 34.8 Å, indicating that complexes larger than the 1:1 complex have formed. However, R_g does not reach 43.2 Å at PEG concentrations of 9% or below, suggesting that a mixture of 1:1 and 2:1 complexes is present. At 12% PEG MME 5000, R_g increases past 43.2 Å,

signifying that higher-order protein associations than the 2:1 complex are formed, consistent with the $P(r)$ results described above.

III.F. Discussion

B_{12}-dependent methyl transfer reactions play important roles in biological systems, from methionine biosynthesis and regeneration to the microbial metabolism of one-carbon compounds. The enzymes responsible for these processes adopt large assemblies of multiple subunits to carry out a sequence of reactions in the catalytic cycle. Crystal structures containing individual fragments for several of these enzyme assemblies have been determined and are described in Chapter VI; however, the overall structural arrangements of B_{12}-dependent methyltransferase components remain largely undetermined. The crystal structure of a CFeSP/MeTr complex, the subject of Chapter II, is the only structure that depicts the domains that bind, protect, activate, and perform catalysis on the B_{12} cofactor together to illustrate their relative orientations. Nonetheless, this snapshot cannot alone explain the multitude of structural configurations that must occur in the cell.

In this study, anaerobic SAXS experiments of MeTr and CFeSP in solution were performed to gain insight into the quaternary structures of these species alone and together in complex. Multiple analytical approaches demonstrate that at the full range of protein concentrations explored, CFeSP assembles as a heterodimer of large and small subunits and MeTr as a homodimer. Further, *ab initio* shape reconstructions of these species match exceptionally well with their respective crystal structures (Figure III.5).

With the quaternary structures for both proteins well established, they could be mixed to investigate complex formation.

Upon mixing of CFeSP and MeTr samples, however, surprising results emerged. It was expected that CFeSP and MeTr in solution would form a complex similar to the crystal structure, a 2:1 complex containing two CFeSPs bound on either side of one central MeTr homodimer (2CFeSP:1MeTr homodimer). However, under assay solution conditions, CFeSP and MeTr instead form an unexpected complex with 1:1 stoichiometry (1CFeSP:1MeTr homodimer). Despite the addition of excess CFeSP, a 2:1 complex still does not form appreciably and a build-up of free CFeSP is observed instead (Figure III.7.C). Therefore, only one out of two available MeTr active sites is capable of methyl transfer activity at any time, regardless of CFeSP concentration. Because prior biochemical studies of MeTr were performed in the same or similar solution conditions[4,5,17,30-32], the exclusive formation of the 1:1 complex observed here indicates that the effective MeTr concentrations in those experiments were actually only half of the presumed MeTr concentrations. The results of the SAXS experiments under assay conditions thus underscore the importance of structural information in the interpretation of biochemical results.

Before CFeSP binding, the MeTr homodimer is composed of identical monomers that are bound together and related by two-fold symmetry[2,3]; therefore, the binding affinities of the first CFeSP to either of the two MeTr monomers are equivalent. In the assay buffer, binding of one CFeSP to the MeTr homodimer dramatically lowers the binding affinity for the second CFeSP to such an extent that binding of a second CFeSP was not detected. Under these conditions, CFeSP and the MeTr homodimer thus exhibit a

negative cooperativity binding mode, where binding of CFeSP to one of two equivalent MeTr binding sites strongly decreases the binding affinity of another CFeSP to the second MeTr site.

A possible structural basis for the observed negative cooperativity binding behavior could be rationalized if binding of one CFeSP to the MeTr homodimer blocks the binding site for a second CFeSP. Although a crystal structure of the 1:1 complex does not yet exist that may show such a binding mode, the *ab initio* shape reconstruction density of the 1:1 complex fits well with the CFeSP/MeTr crystal structure of the 2:1 complex where one CFeSP is simply removed (Figure III.11). Here, the MeTr binding site for the CFeSP small subunit (Figure II.4), where all direct CFeSP/MeTr interactions are located, is on the opposite side of the MeTr homodimer to its equivalent binding site for the second CFeSP. This orientation suggests that binding of one CFeSP should not impact the binding ability of the second CFeSP. However, a second binding interface also exists, where the CFeSP Fe_4S_4 domain makes hydrophobic interactions with MeTr (Figure II.5). The crystal structure of the 2:1 complex shows that the Fe_4S_4 domains from both CFeSPs lie in close proximity (Figure II.2), binding to the same sites of adjacent MeTr monomers. Therefore, it is possible that the Fe_4S_4 domain of CFeSP in the 1:1 complex is flexible and may block the binding site for the Fe_4S_4 domain of a second CFeSP, potentially providing a structural basis for negative cooperativity. Indeed, the shape reconstruction density for the Fe_4S_4 domain in the 1:1 complex appears weak, indicating that the domain is flexible in solution (Figure III.11). In addition, due to the required proximity of Fe_4S_4 domains in the 2:1 complex, it is possible that the Fe_4S_4 domains of two CFeSPs repel each other in solution, which would also disfavor the

assembly of the 2:1 complex and cause the observed negative cooperativity binding mode.

However, in the presence of the crystallization precipitant, the 2:1 complex can be observed (Figure III.7.D). The precipitant was added to precisely match the crystallization condition, where two volumes of assay buffer are mixed with one volume of the precipitant solution. Under these conditions, when the concentration of CFeSP is less than the concentration of the MeTr homodimer, the 1:1 complex composes a majority of complexes, with few 2:1 complexes formed. At the equimolar CFeSP:MeTr homodimer point, a maximal amount of the 1:1 complex is reached, and addition of excess CFeSP promotes the formation of the 2:1 complex observed in the crystal structure. From these titration experiments, CFeSP and MeTr can be described as acting like building blocks, where CFeSP can bind to the MeTr homodimer first at a 1:1 ratio, and up to a maximum of two CFeSPs per MeTr homodimer can bind when sufficient CFeSP is present.

The unexpected results described above illustrate the existence of two possible CFeSP/MeTr complexes in different *in vitro* solution conditions. Currently available biochemical data show that the 1:1 complex formed in the assay buffer[4,5,17,30-32] and the 2:1 complex in the CFeSP/MeTr crystals (Chapter II) are both catalytically active. However, the solution and *in crystallo* activities cannot be kinetically compared, and solution conditions that produce a homogeneous sample of the 2:1 complex have not yet been identified. Further kinetic studies must therefore be performed in parallel with structural techniques to investigate if activity is enhanced upon formation of the 2:1 complex. Because the physiological relevance for either complex cannot yet be

established solely on the basis of their activities, both must be taken into account when considering the possible *in vivo* composition of CFeSP/MeTr complexes. In particular, three models should be considered.

In a first model (Model A), only the 1:1 complex observed in the assay buffer is physiological to the exclusion of the 2:1 complex, and CFeSP and the MeTr homodimer bind together with negative cooperativity. However, with this model, the formation of a MeTr homodimer would appear ineffective, as only one active site can be used at a time. In addition, the symmetric 2:1 complex has been observed in the crystal structure and in the solution SAXS experiments that contain PEG MME 5000. Therefore, the observed formation of the 2:1 complex under certain conditions is not consistent with Model A.

In a second model (Model B), only the 2:1 complex of the crystal structure is physiological to the exclusion of the 1:1 complex, where CFeSP and the MeTr homodimer bind together in a positively cooperative binding mode and binding of one CFeSP enhances the affinity for the second CFeSP such that the 2:1 complex is favored. However, a solution that homogeneously forms such a 2:1 complex has not been identified, and the only conditions that can exclusively form the 2:1 complex are the crystals of CFeSP/MeTr themselves. Yet when CFeSP and the MeTr homodimer are mixed in a solution identical to the crystallization condition, the 1:1 complex does form, and addition of CFeSP past the 1:1 molar ratio is required to favor the 2:1 complex.

Therefore, it is possible that both the 1:1 and 2:1 complexes of CFeSP and MeTr homodimer can exist *in vivo*. This third model (Model C), in which both complexes are physiologically relevant, is consistent with the experiments conducted in the presence of both the crystallization precipitant and one of its components, PEG MME 5000; only the

experiments performed in the standard assay buffer, which lacks PEG, exhibit the formation of only one of the two possible CFeSP/MeTr complexes.

PEG is often used *in vitro* to simulate intracellular macromolecular crowding by contributing to the excluded volume effect, which dramatically increases the effective concentration of species in solution[33-35]. Here, PEG repels protein particles to enhance protein-protein interactions, allowing proteins that may otherwise repel each other *in vitro* to form complexes. For example, in studies of large multicomponent macromolecules such as the 70 S ribosome[36], RuBisCO activase and its complex with RuBisCO[37], the bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS)[38], the T4 DNA replication complex[39], α- and β-synuclein[40,41], and cytoskeletal elements such as actin[42], spectrin[43], and microtubules[44], PEG has been shown to markedly enhance subunit association, leading to substantial increases in binding affinity, protein stability, and/or activity. However, because PEG can also engage in attractive interactions with the protein of interest, its use as an intracellular crowding mimic is imperfect and may lead to non-physiological behavior[35]. The caveats of PEG as an ideal crowding agent notwithstanding, the presence of PEG is still sufficient to promote the formation of the 2:1 complex (Figures III.7.D and III.12.A-B), a feat not possible by simply increasing the protein concentration (Figure III.10). Therefore, the repulsion between PEG and the protein particles is able to overcome repulsive protein-protein forces and favor 2:1 complex formation, perhaps mimicking analogous *in vivo* crowding effects. Such protein-protein repulsive forces that PEG may counteract could arise from the CFeSP Fe_4S_4 domains, which appear close together in the crystal structure of the 2:1 complex as previously mentioned.

In Model C, both the 1:1 and 2:1 complexes are physiologically relevant. Such a scenario would give rise to a flexibility and modularity of the acetogenic B_{12}-dependent methyl transfer system, allowing one or two CFeSPs to bind at independent MeTr sites of the MeTr homodimer while maintaining catalytic activity. In the cell, this behavior could be advantageous. CFeSP interacts with its electron donor protein and with ACS, the downstream enzyme[45], while MeTr may also interact with the upstream enzyme 5,10-methylene-H_4folate reductase, which supplies MeTr with CH_3-H_4folate. It is thus possible that the MeTr homodimer and CFeSP must dissociate from each other in order to allow these additional interactions. Should one CFeSP dissociate from the MeTr homodimer in this manner, it would be advantageous to retain activity of the other MeTr monomer in the resulting asymmetric 1:1 complex. If sufficient CFeSP then becomes available, a maximum of two CFeSPs can bind one MeTr homodimer to allow for enhanced activity in a fully formed 2:1 complex. Although crystals of the 2:1 complex are active (Figure II.19), it is currently unknown whether both catalytic sites are active at the same time; however, both active sites and B_{12} conformations in the CFeSP/MeTr crystal structure are highly similar, and there is currently no structural feature that implies half-sites reactivity. In this way, activity of the CFeSP/MeTr system would be modular and flexible to accommodate additional protein interactions.

Furthermore, CFeSP contains two cofactors that are oxygen sensitive, and CFeSP is prone to repeated inactivation[4,31]. If the integrity of the full 2:1 complex in the cell is compromised due to CFeSP instability or inactivation, the 1:1 complex can still be active. When acetogenic organisms such as *M. thermoacetica* grow autotrophically, the cellular supply of every carbon atom is dependent upon CFeSP and MeTr. Therefore, if lower

overall levels of active CFeSP are present in the cell relative to MeTr homodimer levels, it would be advantageous to continue to perform carbon fixation even though half of the methyl transfer activity would be lost in the 1:1 complex compared to the 2:1 complex.

In contrast to Model C, Model A does not permit increased activity at higher CFeSP levels, while Model B does not allow for any activity should active CFeSP levels decrease. Instead, Model C, in which 1:1 and 2:1 complexes can be formed and are active *in vivo*, enables flexibility of the CFeSP/MeTr system so that both complexes are independently capable of B_{12}-dependent methyl transfer. This model would accommodate interactions with other proteins and permit adaptations to fluctuations in relative concentrations of CFeSP and MeTr homodimer in the cell. The SAXS experiments described here thus raise the previously unanticipated possibility that two different CFeSP/MeTr complexes may form and be active *in vivo*. Further studies may now be performed to test this model and could help unravel the complexity regarding the physiological composition(s) of CFeSP/MeTr complexes.

III.G. Conclusions

The quaternary structures of the acetogenic B_{12}-dependent methyltransferase components CFeSP and MeTr in solution have been probed alone and in complex by anaerobic SAXS. Although the composition of individual CFeSP and MeTr proteins match well with prior structural studies, the observed assembly of two possible CFeSP/MeTr complexes is unanticipated and necessitates a more nuanced understanding of quaternary structures in B_{12} methyltransferase systems. The known crystal structure of a CFeSP/MeTr complex contains two CFeSPs bound on either side of one MeTr

homodimer, a complex with 2:1 stoichiometry. However, under currently-used enzyme activity assay conditions a complex containing only one CFeSP bound to one MeTr homodimer was formed, a species with 1:1 stoichiometry that leaves an unused MeTr active site, even with excess CFeSP present. Interestingly, under solution conditions that mimic the crystallization condition, the 1:1 complex is formed at low ratios of CFeSP to the MeTr homodimer, but the 2:1 complex is favored when excess CFeSP is supplied. The presence of macromolecular crowding agent PEG in the crystallization condition is attributed in part to causing this effect and could represent the crowded conditions within the cell. This surprising behavior demonstrates that two CFeSP/MeTr complexes of different stoichiometries can exist in solution, and suggests that the assembly of this B_{12}-dependent methyltransferase complex exhibits a modular architecture where one or two CFeSPs can bind to one MeTr homodimer. Such behavior would provide an inherent flexibility of the system, allowing the MeTr homodimer to adapt to fluctuations in levels of active CFeSP or permitting CFeSP and one MeTr monomer to dissociate for binding with other protein partners while the other MeTr monomer remains active.

III.H. Materials and Methods

Protein expression and purification

CFeSP was expressed and purified anaerobically from *M. thermoacetica* ATCC 39073 as described[5], except for the following modifications. CFeSP was purified from cell extracts using DEAE-cellulose and high resolution Q-Sepharose anion exchange chromatography followed by phenyl-Sepharose hydrophobic interaction chromatography. Fractions containing CFeSP were concentrated and buffer exchanged using Amicon

ultracentrifuge concentrators in an anaerobic chamber (Coy Laboratories). MeTr was expressed and purified anaerobically from recombinant *Escherichia coli* as described[2]. Concentrations of CFeSP and MeTr protein samples were determined using the Bradford[25] and Rose-Bengal[26] methods and stored in assay buffer (50 mM Tris-HCl, pH 7.6, 100 mM NaCl, 2 mM dithiothreitol).

Small-angle X-ray scattering (SAXS)

SAXS was performed at the Cornell High Energy Synchrotron Source (CHESS) G1 station using a 10.5 keV 250 μm square X-ray beam with a flux of several 10^{12} photons/s. Data were collected at room temperature on a custom 1024×1024 pixel CCD detector[46] with a sample-to-detector distance of ~1.1 m. All samples were prepared in an anaerobic chamber (Coy Laboratories). For *ab initio* shape reconstructions of MeTr and CFeSP and for determination of relative protein concentrations, the protein solutions and matching buffers were contained in 2 mm path-length acrylic cells (ALine Inc.)[47] with 7.5 μm Kapton windows (Chemplex). For all other experiments, glass and boron-rich glass X-ray capillaries were used instead. Prepared cells and capillaries were sealed with epoxy within the anaerobic chamber to maintain anaerobic conditions prior to data collection. To prevent potential oxygen permeation through the Kapton windows, which occurs long timescales (>15 min), each acrylic cell was stored in an individual airtight container, which was not opened until immediately before data collection. Several 1 and 2 s exposures were taken separated by 10 s pauses, where the entire sequence did not exceed 5 min per sample. Exposures that did not display apparent radiation-induced changes were averaged after previously described image correction procedures[47]. The

corrected scattering images were integrated about the beam center and normalized by the transmitted intensities measured by a PIN diode beamstop[47]. Background scattering was subtracted from the protein solution scattering to produce the one-dimensional protein scattering profile, $I(q)$, as a function of q, where $q = 4\pi/\lambda \sin\theta$; 2θ is the scattering angle, and λ is the X-ray wavelength.

A Guinier approximation was applied to the low q region of the scattering profile:

$$I(q) \approx I(0)e^{-R_g^2 q^2/3}$$

where the radius of gyration, R_g, and the forward scattering intensity, $I(0)$, were determined from a linear fit to the Guinier plot, $\ln(I)$ vs. q^2, for the q range that satisfies the $q \cdot R_g < 1.3$ condition[48].

The pair distance distribution function, $P(r)$, was calculated from the experimental $I(q)$ with the indirect Fourier transform method[18] implemented in the program GNOM[29]. The maximum electron pair distance (i.e. maximum protein dimension), D_{max}, was chosen where $P(r)$ naturally approaches zero[18].

Low-resolution models of protein structures were generated from the GNOM outputs with a high-resolution limit of $q \cdot R_g \sim 8$ using the *ab initio* reconstruction programs DAMMIF[23] and GASBOR[22]. The program, DAMAVER[24] was used to align *ab initio* models, reject outliers, and perform averaging to produce the most probable models. Averaged models were aligned to crystal structures in the program SUPCOMB20[9].

Theoretical scattering curves were calculated from atomic coordinates using CRYSOL[20]. For experiments with MeTr and CFeSP alone, theoretical scattering curves were fit to experimental scattering curves in MATLAB (The MathWorks), whereas linear

combinations of theoretical scattering curves were fitted to titration data using a non-negative least-squares fitting algorithm implemented in the program OLIGOMER[27]:

$$I_{fit}(q) = \sum_{i}^{N} v_i I_i(q)$$

where $I_i(q)$ and v_i are the theoretical scattering curve and volume fraction for the i^{th} species, and N is the number of species.

III.I. Tables and Figures

Table III.1. Shape reconstruction statistics.

	MeTr	CFeSP	1:1 complex
Figure	III.5.A	III.5.B	III.11
Protein concentration (μM)	100	100	150
q range ($Å^{-1}$)	0.015-0.229	0.017-0.221	0.031-0.238
Real space range (Å)	0-88	0-90	0-100
GNOM total estimate	0.633	0.672	0.742
Shape reconstruction	gasbor22i	dammif 1.1.0	dammif 1.1.0
Symmetry	*P2**	*P1*	*P1*
Search space	sphere	sphere	sphere
$\sqrt{(\chi^2)}$	2.35-2.56	2.133-2.136	1.185-1.186
# of models averaged/total	10/10	10/10	10/10
Damavar NSD (var)	0.925 (0.011)	0.971 (0.197)	0.811 (0.048)

* 262 x 2 total dummy residues

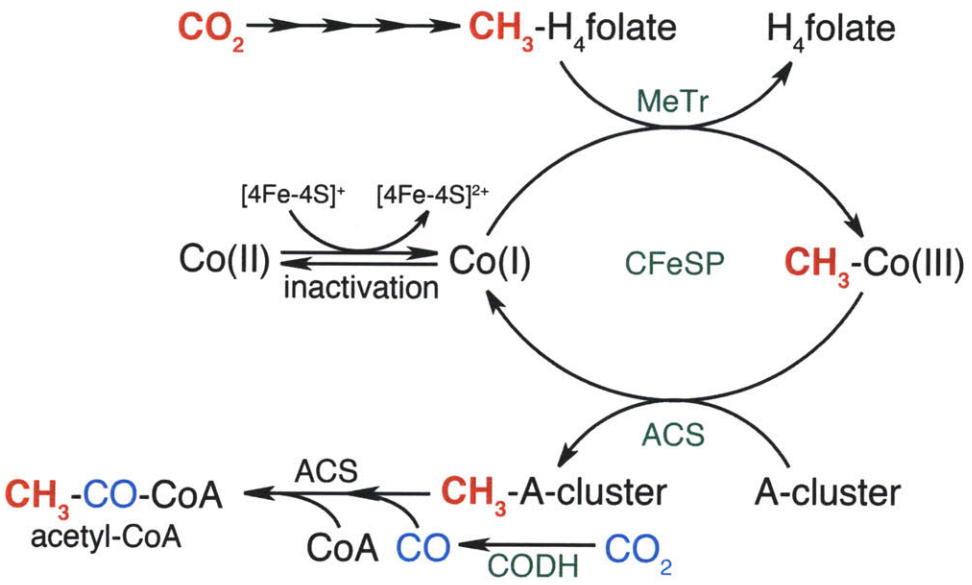


Figure III.1. Overall scheme of CFeSP and MeTr in the Wood-Ljungdahl carbon fixation pathway. One molecule of CO_2 (red, bold) is reduced to a methyl group in a series of folate-dependent reactions to produce CH_3-H_4folate, the substrate of MeTr. MeTr and B_{12}-containing CFeSP form a complex to transfer the methyl group to the Co(I) center of B_{12} cofactor, forming a CH_3-Co(III) intermediate. The methyl group is then delivered to a Ni-Fe-S cluster of acetyl-CoA synthase (ACS) called the A-cluster, reducing the B_{12} cobalt back to the Co(I) state. ACS subsequently catalyzes formation of acetyl-CoA by combining he methyl group with CoA and CO, itself derived from a second molecule of CO_2 (blue) by the action of carbon monoxide dehydrogenase (CODH). Periodic oxidation of the reactive Co(I) center of B_{12} causes inactivation to the Co(II) state. CFeSP can be reactivated by an electron that is transferred from the CFeSP [4Fe-4S] cluster and is originally derived from a partner protein, such as CODH, ferredoxin, or pyruvate:ferredoxin oxidoreductase (PFOR)[4]. Enzyme labels are colored in green.

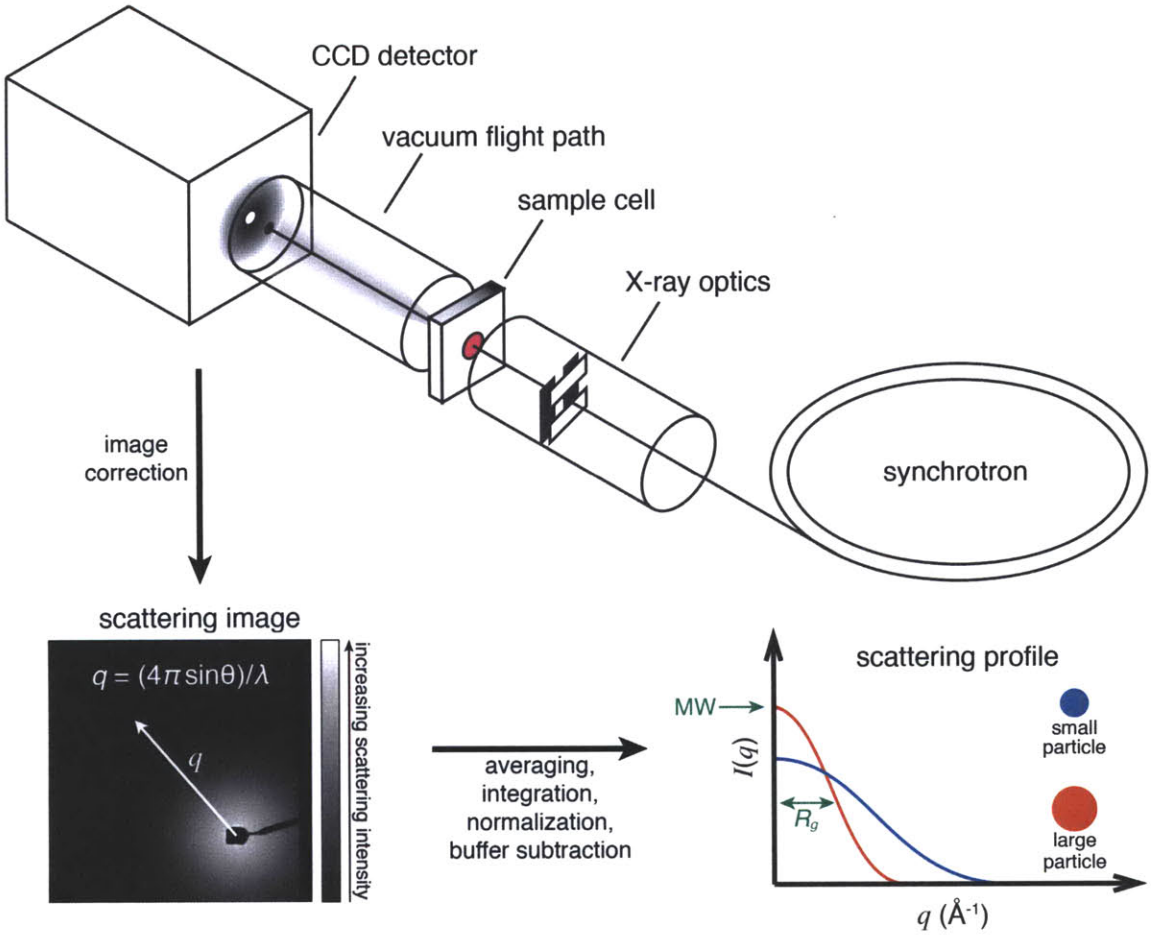


Figure III.2. Overview of a small-angle X-ray scattering (SAXS) experiment. X-ray synchrotron radiation is focused using X-ray optics and passed through the sample (pink) in the sample cell. X-rays scatter through the sample and continue to travel through a vacuum flight path (typically ~1-2 m in length) until they hit the CCD detector. The image is collected and the intensities, I, are corrected. The intensities are averaged about the scattering vector, q, integrated, normalized, and buffer-subtracted to give a scattering profile, $I(q)$ versus q, which can give size and shape information. Small particles (blue scattering profile) scatter to wider angles, resulting in a wider Gaussian shape, while scattering from larger particles (red scattering profile) lead to greater overall intensity and a tighter Gaussian width. Figure modified from images created previously[49], with permission of the author.

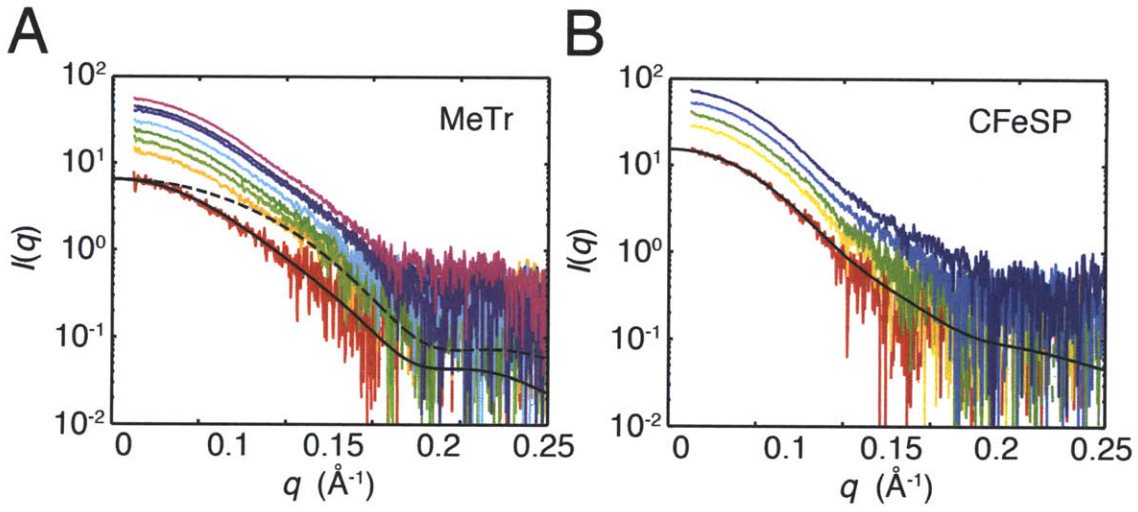


Figure III.3. Scattering profiles for MeTr and CFeSP alone in assay buffer. A, Scattering profiles for MeTr alone in assay buffer at varying concentrations (50-400 μM dimer concentration). The theoretical scattering profiles for the MeTr homodimer (solid black line) and MeTr monomer (dashed black line) were obtained from the MeTr crystal structure (PDB ID: 2E7F). **B**, Scattering profiles for CFeSP alone in assay buffer at varying concentrations (25-150 μM). The theoretical scattering profile for the CFeSP heterodimer (solid black line) was obtained from the "hybrid" CFeSP model (see text for explanation of the "hybrid" model). Scattering profiles are depicted in colors that range from red to violet with increasing protein concentration. No change in the shape of the scattering curves is observed with increasing protein concentration, indicating that there is also no change in the shape of the species in solution.

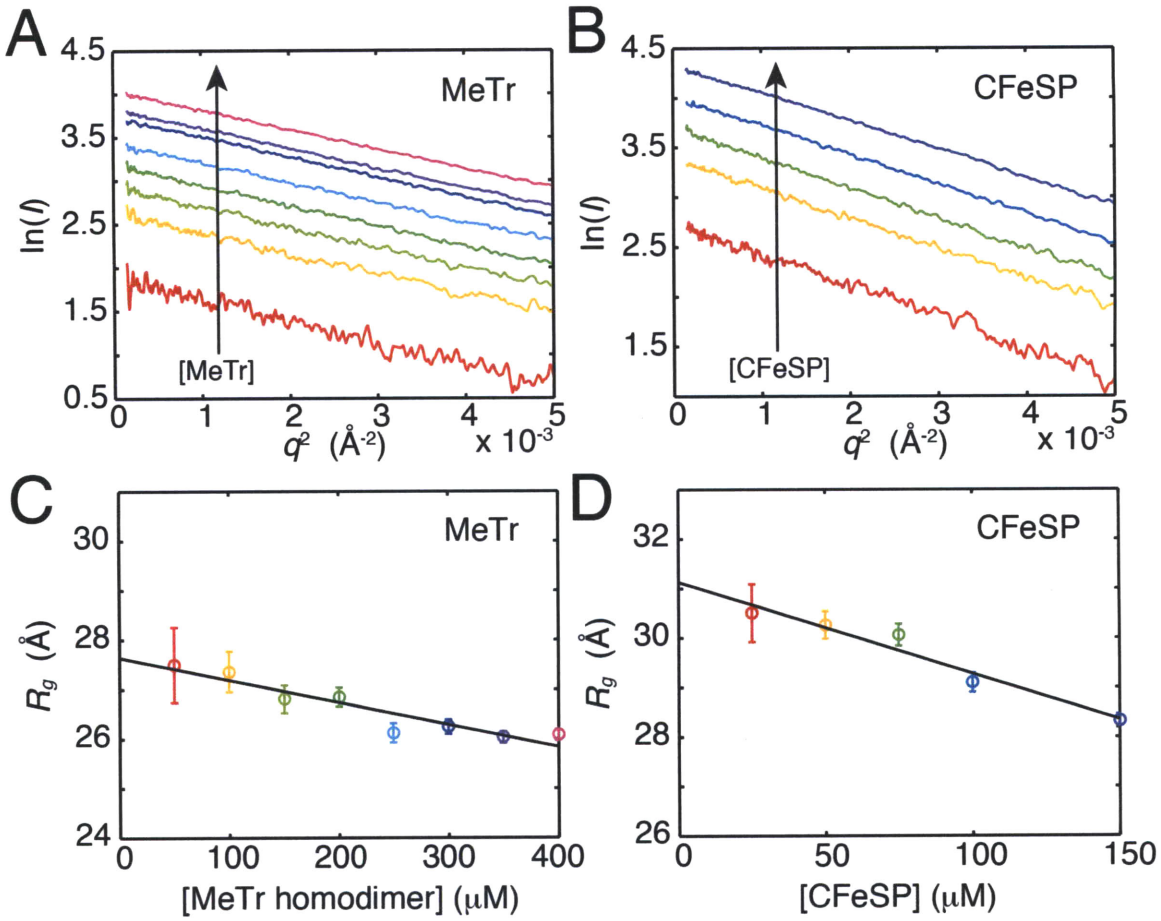


Figure III.4. Guinier plots and radii of gyration determination for MeTr and CFeSP alone in assay buffer. A, Guinier plots for MeTr in assay buffer, and **B**, Guinier plots for CFeSP in assay buffer. **C**, Radius of gyration (R_g) as a function of MeTr homodimer concentration, showing that MeTr exists as a homodimer at all concentrations measured. R_g extrapolated to zero protein concentration is 27.6 ± 0.4 Å, compared to the theoretical R_g of 26.8 Å based on the MeTr homodimeric crystal structure (PDB ID: 2E7F). **D**, R_g as a function of CFeSP concentration, showing that CFeSP exists as a heterodimer of small and large subunits at all concentrations measured. The slight downward slopes of both lines are indicative of minor repulsive inter-particle effects that reduce the apparent R_g as the concentration of particles is increased. Linear extrapolation to zero concentration eliminates these effects and gives R_g of 31.1 ± 0.7 Å, which matches well to the theoretical value of 29.6 Å based on the crystal structure of the "hybrid" model (see text for explanation of the "hybrid" model). Identical colors for **A**/**C** pairs and for **B**/**D** pairs indicate scattering profiles/R_g values for the same sample.

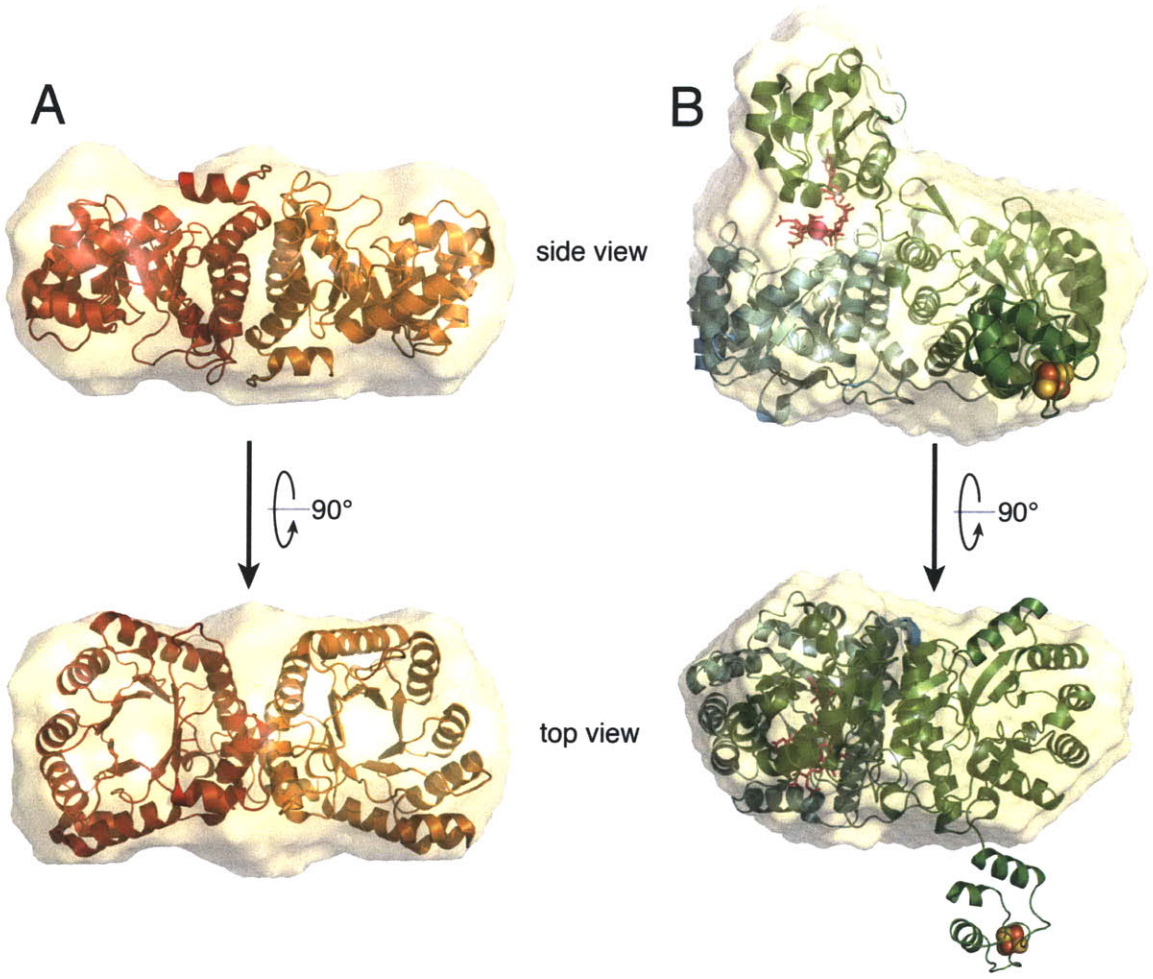


Figure III.5. *Ab initio* shape reconstructions for the MeTr and CFeSP in assay buffer. A, Shape reconstruction (surface) for MeTr determined from the sample at 100 μM (dimer concentration), depicting the homodimeric assembly of MeTr (NSD: 0.925, var: 0.011). The crystal structure of the MeTr homodimer (PDB ID: 2E7F) is aligned and shown in red and orange ribbons. **B**, Shape reconstruction for CFeSP (surface) determined from the sample at 100 μM, depicting the heterodimeric assembly of small and large subunits (NSD: 0.971, var: 0.197). The structure of CFeSP alone is the "hybrid" model and is shown in cyan and green ribbons for the small and large subunits, respectively, with the Fe_4S_4 shown in spheres (Fe in orange and S in yellow) and B_{12} in magenta sticks, with the central cobalt as a magenta sphere. Figure prepared in PyMOL[50].

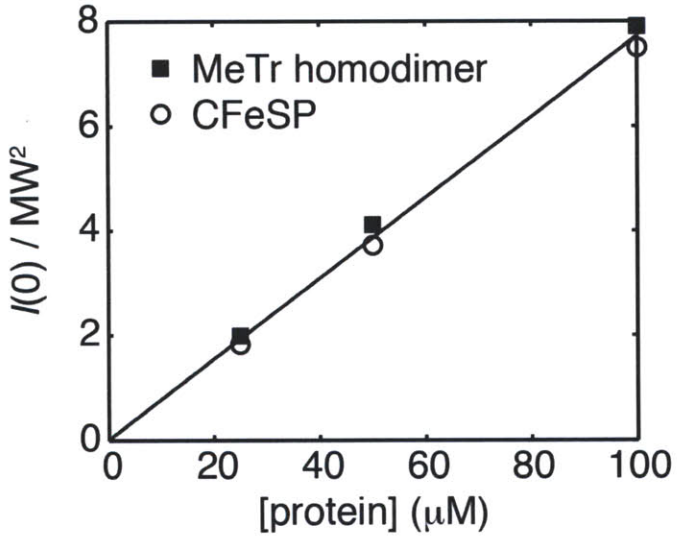


Figure III.6. Confirmation of relative protein concentrations. Plots of the size-adjusted forward scattering intensities, $I(0)$, against protein concentration for samples of both MeTr homodimer (closed squares) and CFeSP (open circles). Points for both samples are nearly colinear, indicating that their experimental and spectroscopically determined relative concentrations are in agreement. Concentrations used for both MeTr homodimer and CFeSP were 25, 50, and 100 μM, in samples prepared in acrylic cells with Kapton windows.

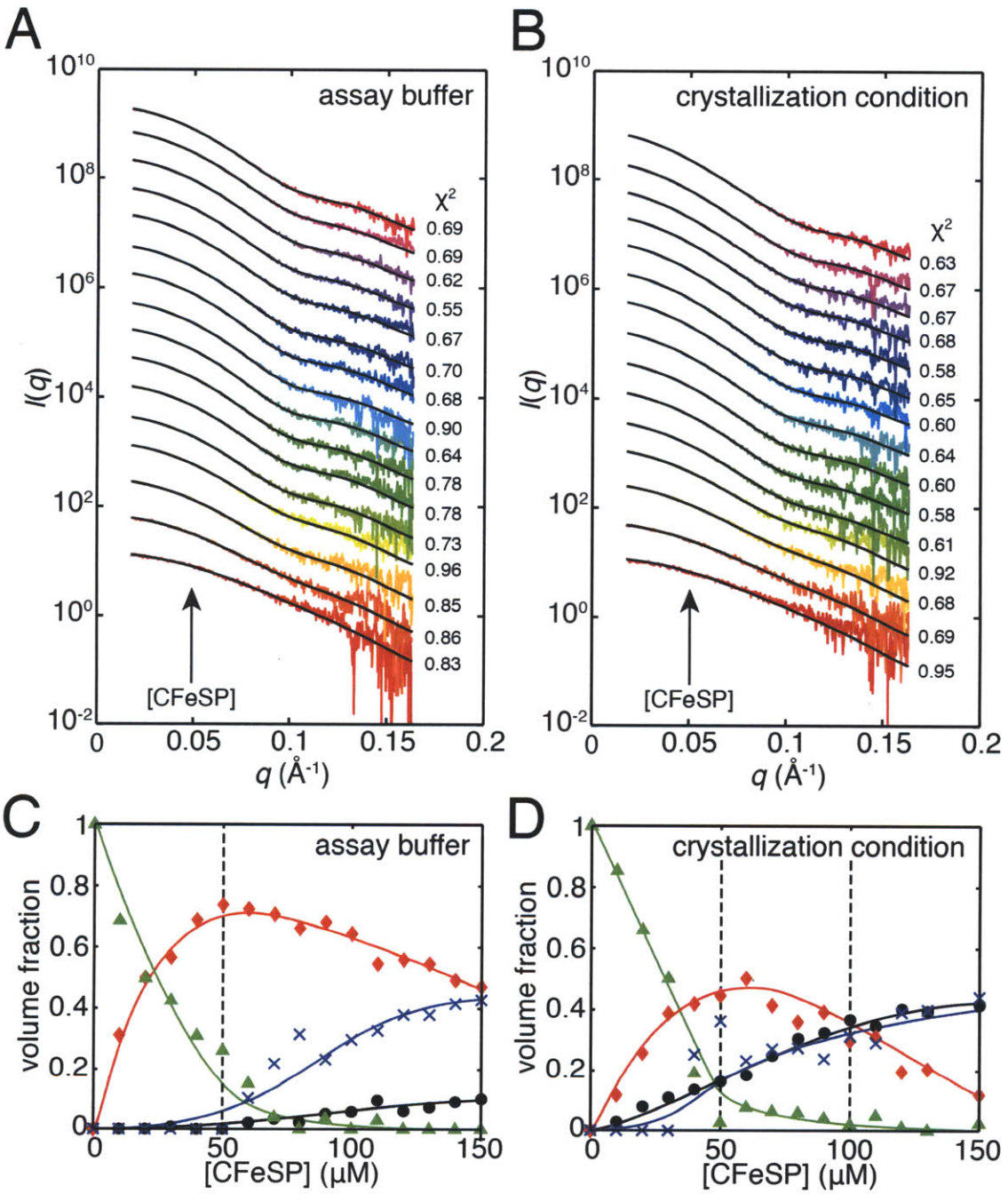


Figure III.7. Scattering profiles and deconvolutions of CFeSP titration experiments. **A**, Scattering profiles for CFeSP titration into MeTr homodimer (50 μM) in assay buffer, and **B**, the crystallization condition. Profile colors range from red to violet and indicate increasing CFeSP concentrations. Deconvoluted scattering fits are depicted with a solid black line in each scattering profile. **C**, Plot of deconvolution results (q range: 0.0179-0.1634 Å^{-1}) for CFeSP titration performed in assay buffer, and **D**, in the crystallization condition. Green triangles represent free MeTr homodimer, blue crosses represent free CFeSP, red diamonds represent the 1:1 complex, and black circles represent the 2:1 complex. Colored lines are visual guides, and dashed black vertical lines denote 50 and 100 μM CFeSP concentration levels.

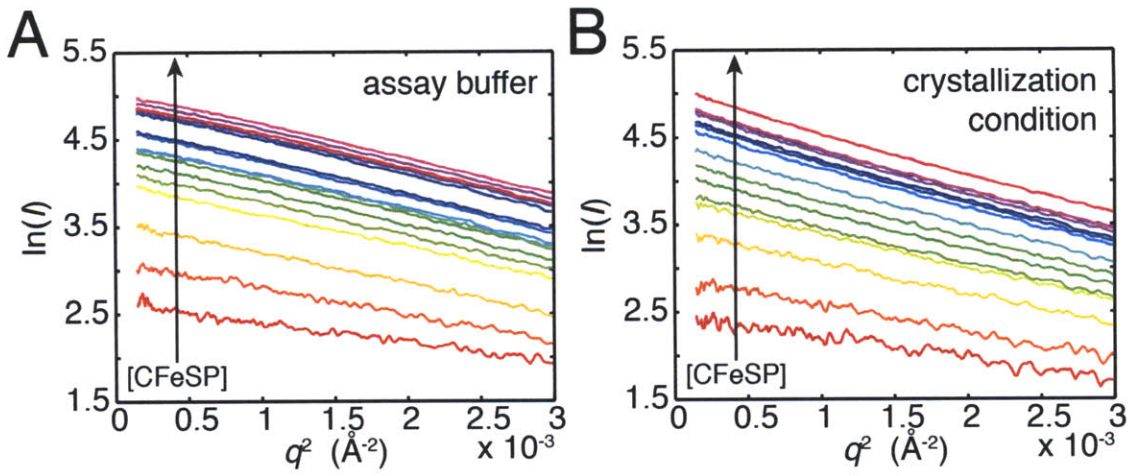


Figure III.8. Guinier plots for CFeSP titration experiments. A, Guinier plot for the CFeSP titration (0-150 μM) into MeTr homodimer (fixed at 50 μM) in assay buffer, and **B** in the crystallization condition.

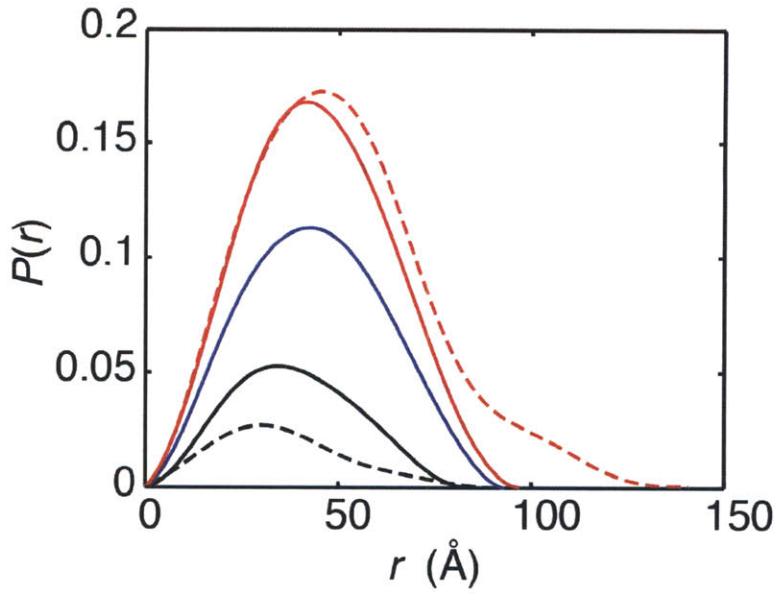


Figure III.9. Electron-pair distance distributions, *P*(*r*), for samples containing CFeSP, MeTr, or both. For samples in assay buffer containing 50 μM MeTr homodimer (dashed black line), 50 μM CFeSP (solid black line), 50 μM MeTr homodimer and 50 μM CFeSP (solid blue line), and 50 μM MeTr homodimer and 100 μM CFeSP (solid red line), the longest electron-pair distance, D_{max}, does not exceed 100 Å, consistent with the 1:1 complex. However, for 50 μM homodimer and 100 μM CFeSP in the crystallization condition (dashed red line), D_{max} is extended to ~140 Å, consistent with the longest dimension in the CFeSP/MeTr crystal structure, which depicts the 2:1 complex.

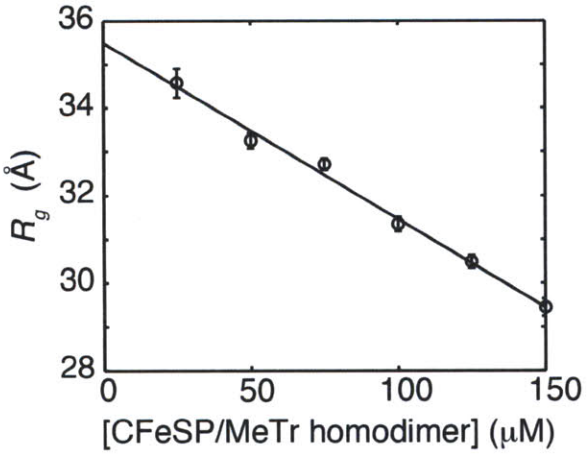


Figure III.10. Radius of gyration (R_g) for the 1:1 complex of CFeSP and MeTr homodimer. As total protein concentration is increased, the 1:1 stoichiometry of CFeSP and MeTr homodimer is maintained, as R_g does not increase, indicating that a larger complex does not form. The slightly decreasing slope is symptomatic of the increasing protein concentration causing proportionally increasing inter-particle repulsive effects, that induce a decrease in apparent R_g. Linear extrapolation to zero concentration eliminates these effects and gives R_g of 35.5 ± 0.5 Å, which matches well with the theoretical value of 34.8 Å for the 1:1 complex extracted from the CFeSP/MeTr crystal structure that depicts the 2:1 complex.

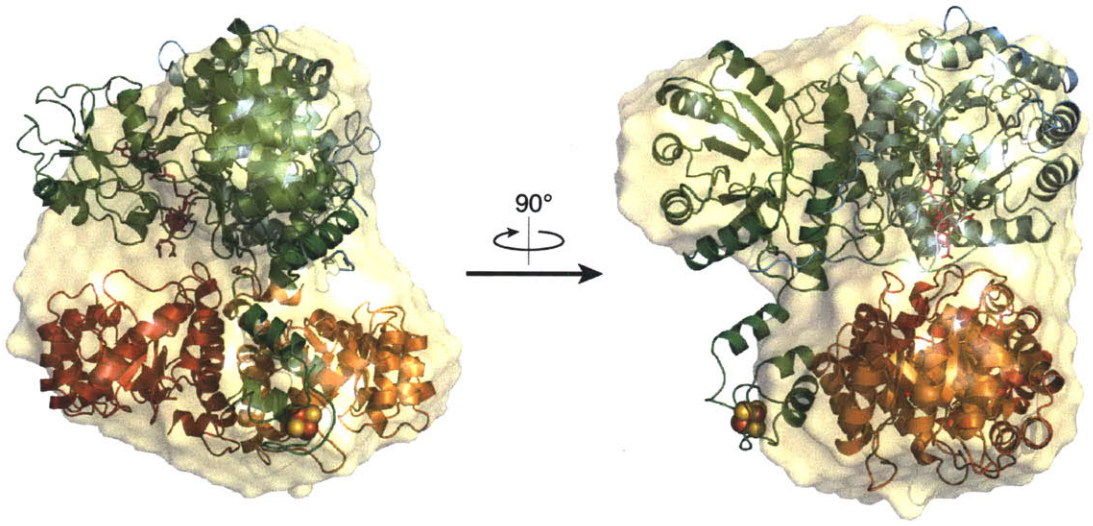


Figure III.11. *Ab initio* shape reconstruction of the 1:1 complex. The shape reconstruction (surface) was obtained from a 100 μM solution of equimolar CFeSP and MeTr homodimer in assay buffer (NSD: 0.811, var: 0.048). The 1:1 complex (ribbons) was extracted from the CFeSP/MeTr crystal structure, which depicted the 2:1 complex, after removing one CFeSP and then aligned to shape reconstruction. MeTr monomers are shown in red and orange, while CFeSP is shown in cyan and green for the small and large subunits, respectively, with Fe_4S_4 cluster in spheres (Fe in orange and S in yellow) and B_{12} in magenta sticks, with the cobalt as a magenta sphere. Figure prepared in PyMOL[50].

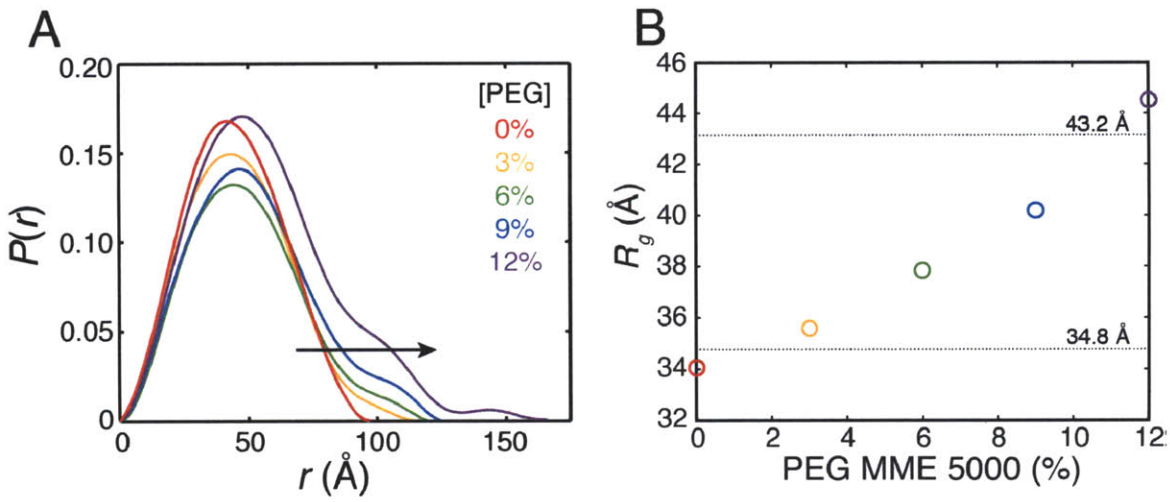


Figure III.12. The effect of increasing PEG concentration on CFeSP/MeTr complexes. A, Electron pair-distance distribution, $P(r)$, as a function of the radius with increasing PEG MME 5000 concentrations, as indicated by the black arrow. Without PEG, the maximum electron pair distance (D_{max}) is ~100 Å, but is shifted to approach ~140 Å of the 2:1 complex as PEG concentration is increased. At the maximum PEG concentration tested, 12%, higher ordered oligomerization is detected, with D_{max} greater than 150 Å. For comparison, the crystallization precipitant contained 9% PEG MME 5000, yielding a final concentration of 3% PEG MME 5000 in the crystallization condition, as crystallization drops contained two volumes of assay buffer and one volume of precipitant. **B**, The radius of gyration, R_g, with increasing PEG MME 5000 concentrations (open circles, whose colors follow the coloring in **A**). Dotted lines indicate the theoretical R_g values for the 1:1 and 2:1 complexes of 34.8 and 43.2 Å, respectively. All samples contained 50 μM MeTr homodimer and 100 μM CFeSP in assay buffer plus PEG MME 5000 in the final concentrations indicated.

III.J. References.

1 Svetlitchnaia, T., Svetlitchnyi, V., Meyer, O. & Dobbek, H. (2006) Structural Insights into Methyltransfer Reactions of a Corrinoid Iron-Sulfur Protein Involved in Acetyl-CoA Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14331-14336.

2 Doukov, T., Seravalli, J., Stezowski, J. J. & Ragsdale, S. W. (2000) Crystal Structure of a Methyltetrahydrofolate- and Corrinoid-Dependent Methyltransferase. *Structure.* **8**, 817-830.

3 Doukov, T. I., Hemmi, H., Drennan, C. L. & Ragsdale, S. W. (2007) Structural and Kinetic Evidence for an Extended Hydrogen-bonding Network in Catalysis of Methyl Group Transfer: Role of an Active Site Asparagine Residue in Activation of Methyl Transfer by Methyltransferases. *J. Biol. Chem.* **282**, 6609-6618.

4 Menon, S. & Ragsdale, S. W. (1999) The Role of an Iron-Sulfur Cluster in an Enzymatic Methylation Reaction. *J. Biol. Chem.* **274**, 11513-11518.

5 Ragsdale, S. W., Lindahl, P. A. & Münck, E. (1987) Mössbauer, EPR, and Optical Studies of the Corrinoid/Iron-Sulfur Protein Involved in the Synthesis of Acetyl Coenzyme A by *Clostridium thermoaceticum*. *J. Biol. Chem.* **262**, 14289-14297.

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

7 Guinier, A. (1939) La Diffraction des Rayons X aux Très Petits Angles: Application a L'Étude de Phénomènes Ultramicroscopiques. *Ann. Phys. Onzieme Serie.* **12**, 161-237.

8 Jacques, D. A. & Trewhella, J. (2010) Small-Angle Scattering for Structural Biology – Expanding the Frontier While Avoiding the Pitfalls. *Protein Sci.* **19**, 642-657.

9 Kozin, M. B. & Svergun, D. I. (2001) Automated Matching of High- and Low-Resolution Structural Models. *J. Appl. Cryst.* **34**, 33-41.

10 Sarma, R., Mulder, D. W., Brecht, E., Szilagyi, R. K., Seefeldt, L. C., Tsuruta, H. & Peters, J. W. (2007) Probing the MgATP-Bound Conformation of the Nitrogenase Fe Protein by Solution Small-Angle X-ray Scattering. *Biochemistry.* **46**, 14058-14066.

11 Chen, L., Gavini, N., Tsuruta, H., Eliezer, D., Burgess, B. K., Donaich, S. & Hodgson, K. O. (1994) MgATP-induced Conformational Changes in the Iron Protein from *Azotobacter vinelandii*, as Studied by Small-Angle X-ray Scattering. *J. Biol. Chem.* **269**, 3290-3294.

12 Eliezer, D., Frank, P., Gillis, N., Newton, W. E., Doniach, S. & Hodgson, K. O. (1993) Small-Angle X-ray Scattering Studies of the Iron-Molybdenum Cofactor from *Azotobacter vinelandii* Nitrogenase. *J. Biol. Chem.* **268**, 20953-20957.

13 Grossman, J. G., Hasnain, S. S., Yousafzai, F. K., Smith, B. E. & Eady, R. R. (1997) The First Glimpse of a Complex of Nitrogenase Component Proteins by Solution X-ray Scattering: Conformation of the Electron Transfer Transition State Complex of *Klebsiella pneumoniae* Nitrogenase. *J. Mol. Biol.* **266**, 642-648.

14 Corbett, M. C., Hu, Y., Fay, A. W., Tsuruta, H., Ribbe, M. W., Hodgson, K. O. & Hedman, B. (2007) Conformational Differences between *Azotobacter vinelandii* Nitrogenase MoFe Proteins As Studied by Small-Angle X-ray Scattering. *Biochemistry.* **46**, 8066-8074.

15 Akiyama, S., Fujisawa, T., Ishimori, K., Morishima, I. & Aono, S. (2004) Activation Mechanisms of Transcriptional Regulator CooA Revealed by Small-Angle X-ray Scattering. *J. Mol. Biol.* **341**, 651-668.

16 Ellis, J., Gutierrez, A., Barsukov, I. L., Huang, W.-C., Grossman, J. G. & Roberts, G. C. K. (2009) Domain Motion in Cytochrome P450 Reductase: Conformational Equilibria Revealed by NMR and Small-Angle X-ray Scattering. *J. Biol. Chem.* **284**, 36628-36637.

17 Zhao, S., Roberts, D. L. & Ragsdale, S. W. (1995) Mechanistic Studies of the Methyltransferase from *Clostridium thermoaceticum*: Origin of the pH Dependence of the Methyl Group Transfer from Methyltetrahydrofolate to the Corrinoid/Iron-Sulfur Protein. *Biochemistry.* **34**, 15075-15083.

18 Glatter, O. & Kratky, O. (1982) "Data Treatment" in *Small Angle X-ray Scattering*, Ed. O. Glatter. Academic Press: London.

19 Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. & Craievich, A. F. (2010) Determination of the Molecular Weight of Proteins in Solution from a Single Small-Angle X-ray Scattering Measurement on a Relative Scale. *J. Appl. Cryst.* **143**, 101-109.

20 Svergun, D., Barberato, C. & Koch, M. H. J. (1995) CRYSOL-A Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.* **28**, 768-773.

21 Drake, H. L., Hu, S.-I. & Wood, H. G. (1981) Purification of Five Components from *Clostridium thermoaceticum* Which Catalyze Synthesis of Acetate from Pyruvate and Methyltetrahydrofolate: Properties of Phosphotransacetylase. *J. Biol. Chem.* **256**, 11137-11144.

22 Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001) Determination of Domain Structure of Proteins from X-ray Solution Scattering. *Biophys. J.* **80**, 2946-2953.

23 Franke, D. & Svergun, D. I. (2009) DAMMIF, A Program for Rapid *Ab-Initio* Shape Determination in Small-Angle Scattering. *J. Appl. Cryst.* **42**, 342-346.

24 Volkov, V. V. & Svergun, D. I. (2003) Uniqueness of *Ab-Initio* Shape Determination in Small-Angle Scattering. *J. Appl. Cryst.* **36**, 860-864.

25 Bradford, M. M. (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248-254.

26 Elliott, J. I. & Brewer, J. M. (1978) The Inactivation of Yeast Enolase by 2,3-butanedione. *Arch. Biochem. Biophys.* **190**, 351-357.

27 Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003) *PRIMUS*: A Windows PC-Based System for Small-Angle Scattering Data Analysis. *J. Appl. Cryst.* **36**, 1277-1282.

28 Hamiaux, C., Pérez, J., Prangé, T., Veesler, S., Riès-Kautt, M. & Vachette, P. (2000) The BPTI Decamer Observed in Acidic pH Crystal Forms Pre-Exists as a Stable Species in Solution. *J. Mol. Biol.* **297**, 697-712.

29 Svergun, D. I. (1992) Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Cryst.* **25**, 495-503.

30 Seravalli, J., Zhao, S. & Ragsdale, S. W. (1999) Mechanism of Transfer of the Methyl Group from (*6S*)-Methytetrahydrofolate to the Corrinoid/Iron-Sulfur Protein Catalyzed by the Methyltransferase from *Clostridium thermoaceticum*: A Key Step in the Wood-Ljungdahl Pathway of Acetyl-CoA Synthesis. *Biochemistry.* **38**, 5728-5735.

31 Menon, S. & Ragsdale, S. W. (1998) Role of the [4Fe-4S] Cluster in Reductive Activation of the Cobalt Center of the Corrinoid Iron-Sulfur Protein from *Clostridium thermoaceticum* during Acetate Biosynthesis. *Biochemistry.* **37**, 5689-5698.

32 Stich, T. A., Seravalli, J., Venkateshrao, S., Spiro, T. G., Ragsdale, S. W. & Brunold, T. C. (2006) Spectroscopic Studies of the Corrinoid/Iron-Sulfur Protein from *Moorella thermoacetica*. *J. Am. Chem. Soc.* **128**, 5010-5020.

33 Ellis, R. J. (2001) Macromolecular Crowding: Obvious but Underappreciated. *Trends Biochem. Sci.* **26**, 597-604.

34 Minton, A. (2000) Implications of Macromolecular Crowding for Protein Assembly. *Curr. Opin. Struct. Biol.* **10**, 34-39.

35 Zhou, H.-X., Rivas, G. & Minton, A. P. (2008) Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences. *Annu. Rev. Biochem.* **37**, 375-397.

36 Zimmerman, S. B. & Trach, S. O. (1988) Effects of Macromolecular Crowding on the Association of *E. coli* Ribosomal Particles. *Nucleic Acids Res.* **16**, 6309-6326.

37 Salvucci, M. E. (1992) Subunit Interactions of Rubisco Activase: Polyethylene Glycol Promotes Self-Association, Stimulates ATPase and Activation Activities, and Enhances Interactions with Rubisco. *Arch. Biochem. Biophys.* **298**, 688-696.

38 Rohwer, J. M., Postma, P. W., Kholodenko, B. N. & Westerhoff, H. V. (1998) Implications of Macromolecular Crowding for Signal Transduction and Metabolite Channeling. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10547-10552.

39 Jarvis, T. C., Ring, D. M., Daube, S. S. & von Hippel, P. H. (1990) "Macromolecular Crowding": Thermodynamic Consequences for Protein-Protein Interactions within the T4 DNA Replication Complex. *J. Biol. Chem.* **265**, 15160-15167.

40 Munishkina, L. A., Cooper, E. M., Uversky, V. N. & Fink, A. L. (2004) The Effect of Macromolecular Crowding on Protein Aggregation and Amyloid Fibril Formation. *J. Mol. Recognit.* **17**, 456-464.

41 Yamin, G., Munishkina, L. A., Karymov, M. A., Lyubchenko, Y. L., Uversky, V. N. & Fink, A. L. (2005) Forcing Nonamyloidogenic β-Synuclein To Fibrillate. *Biochemistry.* **44**.

42 Tellam, R. L., Sculley, M. J., Nichol, L. W. & Willis, P. J. (1983) The Influence of Poly(ethylene glycol) 6000 on the Properties of Skeletal-Muscle Actin. *Biochem. J.* **213**, 651-659.

43 Cole, N. & Ralston, G. B. (1994) Enhancement of Self-Association of Human Spectrin by Polyethylene Glycol. *Int. J. Biochem.* **26**, 799-804.

44 Walsh, J. L., Keith, T. J. & Knull, H. R. (1989) Glycolytic Enzyme Interactions with Tubulin and Microtubules. *Biochim. Biophys. Acta* **999**, 64-70.

45 Tan, X. S., Sewell, C. & Lindahl, P. A. (2002) Stopped-Flow Kinetics of Methyl Group Transfer between the Corrinoid-Iron-Sulfur Protein and Acetyl-Coenzmye A Synthase from *Clostridium thermoaceticum. J. Am. Chem. Soc.* **124**.

46 Tate, M. W., Eikenberry, E. F., Barna, S. L., Wall, M. E., Lowrance, J. L. & Gruner, S. M. (1995) A Large-Format High-Resolution Area X-ray Detector Based on a Fiber-Optically Bonded Charge-Coupled Device (CCD). *J. Appl. Cryst.* **28**, 196-205.

47 Ando, N., Chenevier, P., Novak, M., Tate, M. W. & Gruner, S. M. (2008) High Hydrostatic Pressure Small-Angle X-ray Scattering Cell for Protein Solution Studies Featuring Diamond Windows and Disposable Sample Cells. *J. Appl. Cryst.* **41**, 167-175.

48 Svergun, D. I. & Koch, M. H. J. (2003) Small-Angle Scattering Studies of Biological Macromolecules in Solution. *Rep. Prog. Phys.* **66**, 1735-1782.

49 Ando, N. (2009) *Biomacromolecules Under High Hydrostatic Pressure.* Ph.D. Dissertation, Cornell University.

50 DeLano, W. L. (2010) The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. (http://www.pymol.org).

Chapter IV: Crystal structures of substrate- and inhibitor-bound CODH C-clusters

IV.A. Summary

Nickel-containing carbon monoxide dehydrogenases (CODHs) reversibly catalyze the oxidation of carbon monoxide to carbon dioxide and are of vital importance in the global carbon cycle. The unusual, catalytic CODH C-cluster has been crystallographically characterized as either a $NiFe_4S_4$ or a $NiFe_4S_5$ metal center, the latter containing a fifth, additional sulfide that bridges Ni and a unique Fe site. To determine whether this bridging sulfide is catalytically relevant and to further explore the mechanism of the C-cluster, crystal structures of the ~310 kDa bifunctional CODH/acetyl coenzyme A synthase complex from *Moorella thermoacetica* were solved bound with a substrate H_2O/OH^- molecule and with a cyanide inhibitor. X-ray diffraction data were collected from native crystals and from identical crystals soaked in a solution containing potassium cyanide. In both structures, the substrate H_2O/OH^- molecule binds to the unique Fe site of the C-cluster. We also observe cyanide binding in a bent conformation to Ni of the C-cluster, adjacent to the substrate H_2O/OH^- molecule. Importantly, the bridging sulfide is not present in either structure. As these forms of the C-cluster represent the coordination environment immediately before the reaction takes place, our findings do not support a fifth, bridging sulfide playing a catalytic role in the enzyme mechanism. The crystal structures presented here, along with recent structures of CODHs from other organisms, have led us towards a unified mechanism for CO oxidation by the C-cluster, the catalytic center of an environmentally important enzyme.

A manuscript similar to this chapter was published as:
Kung, Y., Doukov, T.I., Seravalli J., Ragsdale, S.W., and Drennan, C.L. (2009) Crystallographic Snapshots of Cyanide- and Water-Bound C-Clusters from Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry* **48**, 7432-7440.

IV.B. Acknowledgements

Javier Seravalli purified the proteins in the laboratory of Stephen W. Ragsdale, and Tzanko I. Doukov performed crystallization, collected X-ray diffraction data, and processed the native data set.

This work was supported by NIH Grants GM69857 (C.L.D.) and GM39451 (S.W.R) and the MIT Energy Initiative. C.L.D. is a Howard Hughes Medical Institute Investigator. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

IV.C. Introduction

Carbon monoxide dehydrogenases (CODHs) are key enzymes in the global carbon cycle and catalyze the reversible conversion of CO to CO_2. In some anaerobic bacteria, including the phototroph *Rhodospirillum rubrum* and the thermophile *Carboxydothermus hydrogenoformans*, monofunctional Ni-containing CODHs allow these organisms to use CO as their sole carbon and energy source[1,2]. CODH activity accounts for the removal of ~10^8 tons of CO from the environment every year[3].

Acetogenic bacteria, including well-characterized *Moorella thermoacetica*, couple CODH-catalyzed CO_2 reduction with acetyl coenzyme A (CoA) synthesis in the bifunctional enzyme complex CODH/acetyl-CoA synthase (ACS) as part of the Wood-Ljungdahl carbon fixation pathway (Figures I.4-I.6)[4-7]. Briefly, in the "Eastern" branch of the pathway, one molecule of CO_2 is reduced to a methyl group in a series of folate-dependent steps. In the "Western" branch, the methyl group is then transferred from methyltetrahydrofolate (CH_3-H_4folate) to the corrinoid iron-sulfur protein (CFeSP) by CH_3-H_4folate:CFeSP methyltransferase (MeTr). Ni-containing CODH/ACS is also a principal player in the Western branch, where a second molecule of CO_2 is reduced to a CO intermediate by the CODH active site C-cluster. CO then travels ~70 Å through a remarkable tunnel within the enzyme to the ACS active site A-cluster[8-12], where it is combined with the CFeSP-derived methyl group and coenzyme A (CoA) to form acetyl-CoA. Acetyl-CoA can encounter one of two cellular fates: it can either be converted to biomass, or the high energy of the thioester bond can be harnessed first by phosphotransacetylase to produce acetyl phosphate and then by acetate kinase, which transfers the phosphate group to adenosine diphosphate (ADP), yielding adenosine

triphosphate (ATP) and acetate. In this manner, it is estimated that anaerobic acetogens are responsible for roughly 10% of the total global biological output of ~10^{11} U.S. tons of acetate per year[13]. Methanogenic archea also harbor a CODH/ACS variant, referred to as acetyl-CoA decarbonylase/synthase (ACDS), which operates in the reverse direction by degrading acetyl-CoA to the greenhouse gases methane and CO_2[14,15].

All CODHs characterized thus far are homodimeric, can catalyze both forward and reverse reactions, and contain five metalloclusters: two buried active site Ni-Fe-S C-clusters, two Fe_4S_4 B-clusters, and one Fe_4S_4 D-cluster located at the dimeric interface near the surface of the enzyme. The CODH component of methanogenic ACDS also harbors two E- and two F-clusters, which are both Fe_4S_4 centers. The B- and D-clusters are typical cubane Fe_4S_4 clusters and are involved in electron transfer. The unusual C-cluster is the site of CO_2 reduction/CO oxidation, whose structure and catalytic mechanism have been the subject of considerable debate. Initial crystal structures of CODHs from *R. rubrum* (*Rr*CODH)[16], *C. hydrogenoformans* (*Ch*CODH)[17], and the bifunctional ~310 kDa CODH/ACS from *M. thermoacetica* (*Mt*CODH/ACS)[8,10] all show the same overall C-cluster geometry: Ni is part of a distorted cubane-like $NiFe_3S_4$ cluster, and an additional, unique Fe, often referred to as ferrous component II (FCII), is connected to a sulfide of the cubane (Figure IV.1). However, one major difference between these structures is apparent: the *Ch*CODH structure contains a fifth sulfide that bridges the unique Fe and the Ni of the cubane (Figure IV.1.A), while no such sulfide bridge is observed in the *Rr*CODH and *Mt*CODH/ACS structures (Figure IV.1.B). As the bridging sulfide occupies the putative substrate binding sites, its presence or absence is of considerable catalytic significance.

Since 2001, the catalytic relevance and mechanistic role of this bridging sulfide has been hotly contested. For *Ch*CODH, the observation that the bridging sulfide is displaced and the enzyme becomes inactivated by long exposures to CO (>20 h) prompted the suggestion that *Mt*CODH/ACS and *Rr*CODH structures lack the bridging sulfide due to CO inhibition[18]. However, the same CO treatment that led to inactivation of *Ch*CODH did not affect the activity of *Mt*CODH/ACS, even after >600 h of exposure[19]. Furthermore, when sodium sulfide was added to active samples of *Rr*CODH and *Mt*CODH/ACS in an attempt to generate a sulfide bridge, these enzymes became inactivated[19]. As the crystals of *Rr*CODH and dissolved crystals of *Mt*CODH/ACS are active[10,16], and their structures do not contain the sulfide bridge, it has also been suggested that sulfide does not play a catalytic role[20]. Instead, the sulfide-containing *Ch*CODH structures may represent a pre-catalytic state, and the bridging sulfide could be displaced upon substrate binding.

To investigate where substrates bind in the well-characterized *Mt*CODH/ACS, we have solved two ligand-bound structures. First, to identify the binding site for a substrate water molecule on the C-cluster of *Mt*CODH/ACS, we grew crystals in the complete absence of CO to prevent turnover. This lack of CO has also allowed us to explore the relationship between the bridging sulfide and CO exposure. The resulting native structure shows H_2O/OH^- bound to the unique iron of the C-cluster and no bridging sulfide, despite the absence of CO. Second, we wished to identify the binding site for substrate CO on the C-cluster of *Mt*CODH/ACS. However, infrared spectroscopy has indicated that CO does not bind at a single, stable location that would be amenable to crystallographic observation[21]. CO was also recently observed to exchange rapidly to form CO_2 in

*Ch*CODH even in the absence of external electron acceptors[22]. Therefore, to identify the CO binding site, we used cyanide, a known competitive inhibitor of monofunctional CODHs[23]. This cyanide-liganded structure, also obtained in the absence of CO, shows cyanide bound to the Ni of the C-cluster. Despite the fact that *Mt*CODH/ACS is the best characterized CODH regarding substrate binding, no corroborating crystal structures with substrates or analogues bound have been captured until now.

As we were carrying out these crystallographic studies, a third set of *Ch*CODH structures were reported which no longer contained the bridging sulfide and instead depicted CO_2 bound to the C-cluster in its place[24]. Until this point, no CODH structures had substrates bound to the C-cluster, and thus the mechanism of the C-cluster and the role of sulfide in catalysis remained elusive. Even more recently, a structure of the CODH component of methanogenic ACDS from *Methanosarcina barkeri* (*Mb*CODH) bound with a putative CO molecule and H_2O/OH^- also did not contain the bridging sulfide[25]. With these CODH structures and the new *Mt*CODH/ACS structures presented here, we are now in a unique position to compare the C-clusters bound with different ligands, across the three types of CODHs: monofunctional CODH, bifunctional CODH/ACS, and CODH-containing methanogenic ACDS. This new crop of structures has led the field towards a unified mechanistic view of the CODH C-cluster.

IV.D. Results

The native C-cluster

Following the initial CODH structure determinations, a primary point of contention regarding the geometry and mechanism of the C-cluster involves the presence

or absence of a fifth sulfide bridging Ni and the unique Fe. The structures of *Rr*CODH[16], *Mt*CODH/ACS[8-10], and *Mb*CODH[25] lack this bridging sulfide (Figure IV.1.B); however, it is present in some structures of the *Ch*CODH C-cluster[17,18] (Figure IV.1.A). When explaining the absence of the bridging sulfide in the *Rr*CODH and *Mt*CODH/ACS structures, it was proposed that these proteins could have been inhibited by the CO to which they were exposed, leading to an inactivated C-cluster[18]. Here two structures of *Mt*CODH/ACS are presented (Table IV.1) from proteins purified and crystallized in the absence of CO, one containing a native C-cluster (Figure IV.2) and one containing a C-cluster bound by cyanide (Figure IV.3). Both C-clusters have atomic positions consistent with those seen in all prior CODH structures, and neither contains density for the bridging sulfide (Figure IV.4).

In the native structure, a positive F_o-F_c electron density peak greater than 4 σ is present, indicating a ligand bound to the unique Fe of the C-cluster (Figure IV.2.A). As a monatomic ligand best fit the density, we placed a H_2O/OH^- molecule in this location. After refinement, no F_o-F_c difference density was present in this region, with H_2O/OH^- positioned 2.03 Å from the unique Fe and 2.60 Å to Lys587 (Figure IV.2.B). All reported distances are derived from chain B of the final model, as this chain has the lowest *B*-factors of the four CODH chains present in the asymmetric unit. To confirm that this density does not represent a sulfide, we also modeled a sulfur atom into this site; after refinement, negative F_o-F_c difference density appeared near the sulfur, indicating that sulfur is an incorrect assignment (Figure IV.5).

The occupancies of the bound H_2O/OH^- ligands in all chains were then refined. Occupancies and *B*-factors of atoms are interrelated and thus cannot be refined

simultaneously; when the occupancies for the H_2O/OH^- ligands were set to 1.0, as in the final model, their *B*-factors were generally somewhat higher than the average *B*-factor of the C-cluster atoms. Correspondingly, when *B*-factors of the H_2O/OH^- ligands were set to the average *B*-factor of their respective C-cluster atoms, occupancies generally fell to an average value of 0.9, denoting 90% occupancy of the oxo ligand. The results of these occupancy and *B*-factor refinements are shown in Table IV.2.

Previous structures of the C-cluster without substrates or the bridging sulfide indicate that an open coordination site on the unique Fe is present before substrate binding[8-10,16]. When the C-cluster binds the H_2O/OH^- substrate as seen in our native structure, the coordination environment of the unique Fe converts from trigonal pyramidal to a distorted tetrahedral geometry. The positions of the cluster sulfide, Cys317, and His283 ligands to the unique Fe remain largely unchanged, as do the positions of the other C-cluster atoms.

The cyanide-bound C-cluster

The structure of cyanide-treated *Mt*CODH/ACS crystals also shows density consistent with a H_2O/OH^- molecule (Figure IV.3.A and Figure IV.6) bound in an analogous position as in the native structure, with a 2.05 Å Fe-oxo bond. However, additional electron density for a larger ligand bound to Ni of the C-cluster is present, represented in both the $2F_o$-F_c and the F_o-F_c electron density maps (Figure IV.3.A). Because the only difference between the native and cyanide data sets was that, with the latter, crystals were soaked in a solution containing cyanide before data collection, we considered if the electron density represents a cyanide ligand by refining molecules of

different sizes into this density. Assignment of this new site as a H_2O/OH^- molecule proved to be a poor fit to the density, as refinement moved the molecule to a distance of 2.47 Å from Ni, with some positive F_o-F_c density still remaining (Figure IV.6.E-F). Refinement was also carried out with a sulfur atom in this location; however, strong negative F_o-F_c density peaks were observed, signifying that sulfur was also an incorrect assignment (Figure IV.6.G-H). F_o-F_c density was eliminated only when a cyanide ligand was modeled into this region (Figures IV.3.B and IV.6.A-B).

Occupancies of the cyanide and H_2O/OH^- ligands were also refined. As with the native structure, when occupancies of the cyanide and H_2O/OH^- atoms were set to 1.0, their *B*-factors were somewhat higher than the average *B*-factor of the C-cluster. When *B*-factors of the cyanide and H_2O/OH^- atoms were set to the average *B*-factor of their respective C-cluster atoms, occupancies generally fell to an average value of 0.7, denoting 70% occupancy. The results of these occupancy and *B*-factor refinements are shown in Table IV.3.

Cyanide binds to Ni of the C-cluster in a bent conformation (Figure IV.3.B), with a C-Ni distance of 1.99 Å, a C-H_2O/OH^- distance of 3.54 Å, and an N-H_2O/OH^- distance of 2.94 Å The Ni-C-N bond angle is nonlinear and is not uniform across the four CODH subunits: chain A, 115°; chain B, 111°; chain C, 101°; and chain D, 130°; for an average of 114°. These discrepancies may be due to differences in the relative order and disorder of the four chains in the asymmetric unit, as chains C and D have higher *B*-factors, indicating more disorder. Three refinement programs, REFMAC[26], CNS[27], and PHENIX[28], consistently produced a bent Ni-CN coordination; when cyanide was fixed linearly to Ni and the cyanide carbon assigned as *sp* hybridized, refinement nonetheless

routinely moved the cyanide ligand back into its favored bent conformation. Ile591 is found in close proximity to the cyanide nitrogen, presenting a potential steric clash that prevents linear binding of cyanide (Figure IV.7).

Previous structures of the C-cluster without substrate ligands or the bridging sulfide indicated possible empty coordination sites on Ni of the C-cluster before substrate binding. In the cyanide-bound structure presented here, binding of the substrate mimic results in a distorted tetrahedral geometry of Ni. This geometry is in contrast to the structures of *Ch*CODH that contain the bridging sulfide, which depicts square planar-like geometry. The positions of the other atoms and ligands to the C-cluster do not change significantly as a result of cyanide binding.

IV.E. Discussion

The CODH C-cluster is responsible for the catalysis of one of the most fundamental chemical reactions in biology: the reversible interconversion of CO and CO_2. Despite the importance of this metallocluster, there has been significant disagreement regarding the structural elements required for activity. One point of contention has been whether the fifth (bridging) sulfide observed in structures of *Ch*CODH[17,18], but not in other CODHs[8-10,16,25], is catalytically relevant in all CODH enzymes, represents a pre-catalytic state in all CODH enzymes, or has different roles in different CODH variants. Interspecies variations could explain the presence of the bridging sulfide in *Ch*CODH and the lack of the sulfide in CODHs from other organisms. However, given the CODH sequence homology, it is unlikely that C-clusters performing the same chemistry would employ different mechanisms with different substrate binding

features, and there is a lack of biochemical or biophysical evidence supporting this explanation. The suggestion that a sulfide bridge could be important for maintaining the structural integrity of the C-cluster in a pre-catalytic state, while not being directly involved in catalysis, has precedence in the proposed mechanism of some [NiFe] hydrogenases. Here, a sulfide ligand bridging Ni and Fe has been suggested to exist in a pre-catalytic, inactive state to protect the metals in the active site prior to catalysis[29,30]. The purpose of the fifth sulfide in the catalytic cycle of the C-cluster, however, had been difficult to evaluate until structural data on substrate and analogue binding was acquired.

In the native structure of *Mt*CODH/ACS, crystallized in the absence of CO, no bridging sulfide is observed. The C-cluster instead depicts a H_2O/OH^- molecule bound to the unique Fe. Electron density maps following refinement of this ligand as sulfur and not H_2O/OH^- indicate that assignment as sulfur does not fit the density (Figures IV.5.C-D). The native structure thus represents the first time that the substrate H_2O/OH^- molecule has been observed in *Mt*CODH/ACS, the best characterized CODH. Although it is difficult to unambiguously prove that a ligand is a H_2O/OH^- molecule using crystallography, our assignment is consistent with electron-nuclear double resonance (ENDOR) studies of *Mt*CODH[31], which indicated that H_2O/OH^- binds to the unique Fe. Structural comparisons are also valuable in assigning H_2O/OH^- as the ligand; recent *Ch*CODH and *Mb*CODH structures also contain H_2O/OH^- bound in the same position on Fe (Figure IV.8), a site that is unavailable in structures that contain the bridging sulfide. H_2O/OH^- in *Ch*CODH was modeled as the ligand partly based on spectroscopic work performed on *Mt*CODH/ACS. Previous *Mt*CODH/ACS structures[8-10], however, did not contain density for H_2O/OH^-. Of the three stable oxidation states of the C-cluster (C_{ox},

C_{red1}, C_{red2}), it has been shown that both CO and H_2O/OH^- substrates bind only the C_{red1} state, with oxidation of CO to CO_2 further reducing the cluster to the C_{red2} state, the only state that binds CO_2[31-34]. It is possible that in the previous *Mt*CODH/ACS structures heterogeneity in the electronic states of the clusters prevented full occupancy of bound H_2O/OH^-, where some oxidation to the C_{ox} state and some turnover to the C_{red2} state due to CO exposure may have occurred. In the structures presented here, the purification and crystallization protocol appears to have maintained the C_{red1} state for H_2O/OH^- binding at ~90% or greater occupancy. The H_2O/OH^- substrate molecule is bound 2.03 Å from the unique Fe, compared to 1.94 and 2.14 Å in the *Ch*CODH and *Mb*CODH structures, respectively.

H_2O/OH^- is also 2.60 Å from a conserved Lys587, whose role in the reaction mechanism of the C-cluster is intriguing. In general, hydrogen atoms cannot be observed crystallographically due to the limits in resolution; therefore, we cannot assign with confidence the protonation state of H_2O/OH^- or Lys587. As Fe-bound H_2O must be deprotonated before nucleophilic attack on CO, the proximity of Lys587 to H_2O/OH^- in our structure may lead one to propose that Lys587 acts as the general base. Due to pKa considerations, it would be atypical for lysine to accept protons from water; however, the proximity of both Lys587 and the unique Fe of the C-cluster could lower the pKa of H_2O. It is also possible that Lys587 stabilizes the OH^- form, serving to activate hydroxide attack on Ni-bound CO. Indeed, mutation of Lys587 has been shown to reduce the activity of *Mt*CODH[35].

Because infrared spectroscopy has shown that *Mt*CODH/ACS does not bind CO in a single, stable location[21], we have used cyanide, which is isosteric and isoelectronic to

CO, to probe the CO substrate binding site. The literature is inconsistent regarding the location of the cyanide binding site on the C-cluster of CODH, with some spectroscopic studies suggesting cyanide binds to iron and others suggesting multiple, perhaps bridging, binding sites[22,23,31,32,36-39]. However, recent inhibition studies and extended X-ray absorption fine structure (EXAFS) data on *Ch*CODH show cyanide as a competitive inhibitor of C-cluster activity, binding to the same location as CO[23]. In the cyanide-bound *Mt*CODH/ACS structure presented here, cyanide is bound to Ni at a distance of 1.99 Å. For *Ch*CODH, EXAFS data on cyanide binding are also consistent with cyanide binding to Ni, with a Ni to C distance in the range of 1.81-1.84 Å[23]. Therefore, while some of the spectroscopy on the cyanide-bound CODH is inconsistent with our structure, the recent EXAFS results on *Ch*CODH are consistent. A further discussion of cyanide inhibition in CODH is provided in Chapter V.

In the structure presented here, cyanide is bound to Ni of the C-cluster in an analogous conformation as the proposed CO ligand modeled in the *Mb*CODH structure[25]. That ligand's identification as CO was only inferred from the electron density, and the authors indicated that the ligand could also be a formyl group. If assignment as CO was correct, it is puzzling that turnover did not occur; the authors attribute this to the low pH of the crystallization conditions which may have disfavored the deprotonation necessary for activity. In contrast, we can more confidently report our ligand's identity as cyanide, as the sole difference between our native and cyanide data was that, with the latter, crystals were soaked in a solution containing cyanide before data collection. As mentioned, no bridging sulfide is seen in this structure, and such a bridge would not be possible with CN/CO bound to Ni in this position. Interestingly, we observe cyanide

bound to Ni in a bent conformation, with an average angle of 114°. A diatomic ligand seen in the *Mb*CODH structure, modeled as CO, also exhibits a similar bent conformation (Figure IV.8.A), with a Ni-C-O bond angle of 103°. In the *Mb*CODH structure, when modeled CO was fixed in a linear fashion with respect to Ni, a steric clash was observed between CO and an isoleucine residue near the C-cluster. This isoleucine was seen in close proximity to the oxygen of CO, preventing linear binding of a diatomic molecule. Here, we similarly observe Ile591 impart a potential steric clash that cannot be overcome during refinement (Figure IV.7). There is precedence for a bent cyanide metal ligand in several heme-containing metalloproteins, including catalases[40], peroxidases[41,42], sulfite reductase hemoprotein[43], hemoglobin and myoglobin[44], cytochrome cd_1 nitrite reductase[45], and cytochrome P450 enzymes[46,47]. A study of the non-heme mononuclear iron enzyme superoxide reductase also identifies bent CN geometry, where the Fe-C-N bond angle was observed as 123° and 133° for two different substrates[48]. In all cases, bent geometry was attributed to steric constraints and hydrogen bonding upon adopting a bent conformation. In our structure, not only does Ile591 similarly impart a steric hindrance to linear binding, the bent conformation could also allow the cyanide ligand to form a hydrogen bond with the Fe-bound water molecule at 2.94 Å away or perhaps with His113, which is at a farther distance of 3.28 Å (Figure IV.7).

Although cyanide binds Ni in a bent fashion in our structure, this does not require that CO in this bent arrangement represents the conformation needed for hydroxide attack on the CO carbon. Indeed, Gong *et al.* propose that before nucleophilic attack a change in CO binding occurs[25]. Starting from the bent arrangement seen in their *Mb*CODH structure, the authors propose that the CO carbon moves toward the H_2O/OH^- ligand,

causing Ni to switch from the tetrahedral-like geometry observed in our structure to a more square planar geometry, while the CO oxygen does not move. Such a "carbon shift" causes CO to be bent in the opposite direction, allowing the carbon to be closer to the attacking hydroxide, without movement of the CO oxygen. This proposal stemmed from the authors' observation that when compared with the *Ch*CODH CO_2-bound structure, discussed below, the position of the CO oxygen remains essentially constant after conversion to CO_2, while the CO_2 carbon was shifted, altering the Ni coordination geometry.

Further, Gong *et al.* propose that deprotonation of water to hydroxide may control the suggested "carbon shift" and thus the direction of CO bending. Here, CO is initially bent toward H_2O, allowing for hydrogen bonding. Presumably, as H_2O is not the catalytic substrate, CO cannot undergo the proposed "carbon shift", as this would place CO and water, two nonreacting species, too close in proximity. Only after water is deprotonated to the active hydroxide species may the "carbon shift" take place, positioning the carbon closer to the hydroxide for chemistry to occur. Pointing to the fact that their CO-bound structure was obtained at pH 4.6, the authors suggest that their observed bent conformation may be attributed to the acidic conditions making deprotonation of water, and catalysis, more unfavorable and that bending in the opposite direction may be observed at neutral pH. Gong *et al.* thus suspect that their protein was inactive and did not turn over to form CO_2 because of these acidic conditions.

Although the cyanide-bound structure presented here was obtained from samples crystallized at pH 7.6, we nonetheless clearly observe the same bent conformation as was seen with the *Mb*CODH structure in the presence of CO. Unlike CO, where a partial

positive charge exists on carbon when bound to a metal ion, cyanide is not subject to attack by hydroxide and functions as an inhibitor, binding to the same site as CO without initiating catalysis. As such, had cyanide undergone the "carbon shift" to be bent in the opposite direction, the distance between the cyanide carbon and H_2O/OH^- would again be too close, approximately the length of a covalent bond, which is sterically unfavorable between two nonreacting species. Perhaps, because cyanide does not react with H_2O/OH^-, it cannot undergo the "carbon shift" due to either steric or electronic constraints and thus does not exhibit bending in the opposite direction.

A recent *Ch*CODH crystal structure illustrates a CO_2-bound C-cluster[24]. Here, *Ch*CODH crystals soaked in a solution containing sodium bicarbonate, buffered at pH 8.0, and with redox potential of -600 mV vs. standard hydrogen electrode (SHE) yielded a C-cluster with electron density modeled as CO_2 between Ni and the unique Fe. The bridging sulfide is not observed in this structure, as the modeled CO_2 molecule now occupies the region where the *Ch*CODH bridging sulfide was previously seen[17,18]. The redox potential of -600 mV was attained using titanium(III) citrate with the intention of obtaining the C_{red2} state for CO_2 binding; however, the authors did not indicate why CO_2 bound in the C_{red2} state did not induce turnover to CO and water. In this structure, CO_2 is modeled with the carbon bound 1.96 Å to Ni, one oxygen is bound 2.05 Å to the unique Fe, and the second oxygen is located 2.89 Å from His93 (His113 in *Mt*CODH/ACS). In comparing our cyanide structure with the CO_2-bound C-cluster, the locations of both Fe-bound oxygens are essentially the same, and the cyanide nitrogen is in the same position as the second CO_2 oxygen. Similar to the *Mb*CODH structure, the only significant difference is once again the location of the carbon, suggestive of the proposed "carbon

shift". To illustrate, a superposition of all three structures is shown in Figure IV.8.B. The substrate- and inhibitor-bound C-clusters of all structures are also shown in Figure IV.9.

As the cyanide inhibitor is not subject to nucleophilic attack by hydroxide, we have captured a state of the enzyme immediately before the reaction or any "carbon shift" takes place, with both substrates lined up for catalysis. In this crucial stage, we observe no evidence of a bridging sulfide. The same scenario was observed with the putative CO-bound *Mb*CODH. As the recent CO_2-bound *Ch*CODH structure also does not contain the bridging sulfide, and as it represents the cluster environment immediately after hydroxide attack on CO, this structure provides further support for the position that a bridging sulfide must not be catalytically relevant. Our structures of *Mt*CODH/ACS together with other recent CODH structures[25,49] and biochemical data[19] thus put to rest the proposition that a bridging sulfide is catalytically relevant and make possible a unified mechanistic view of C-cluster chemistry.

In our proposed reaction mechanism for CO oxidation, shown in Figure IV.10, a water molecule is bound to the unique Fe of the C-cluster in the C_{red1} redox state, as observed in the native structure of *Mt*CODH/ACS shown here. This state is referred to in Figure IV.10 as state I. Then, CO is proposed to bind Ni in a bent conformation (state II), as observed in the cyanide-bound crystal structure described here, with the carbon too distant from the Fe-bound water to undergo a reaction. In the next step, deprotonation of water to the reactive hydroxide species may promote a "carbon shift" that would place the carbon closer to the Fe-bound hydroxide, in a position that would promote the reaction between CO and OH^- (state III). Following nucleophilic attack of hydroxide onto Ni-bound CO, the Ni-COOH intermediate (state IV) must also be deprotonated to give a

Ni-COO^- species (state V). CO_2 is then released and the C-cluster becomes reduced to the C_{red2} state (state VI). Finally, two electrons are transferred to the B-cluster, the D-cluster, and ultimately an electron acceptor protein, ferredoxin, hydrogenase, or pyruvate:ferredoxin oxidoreductase (PFOR)[50,51]. This electron transfer reaction re-oxidizes the C-cluster to the C_{red1} state (state I), and the catalytic cycle continues.

Two deprotonations are carried out during the course of the reaction, but identifying the catalytic base(s) responsible is difficult. For both deprotonations, Lys587 is most optimally positioned, but as discussed, pK_a considerations may limit the lysine to playing a role in stabilizing the negatively charged hydroxide and the Ni-bound carboxylate intermediate. If this is the case, His113 could instead be the basic residue (Figures IV.7 and IV.11), as previously suggested[16], and the recent CO_2-bound *Ch*CODH shows the water-derived oxygen located 3.93 Å away from the corresponding histidine, His93. Further, a base may not be necessary for the first deprotonation, as the pK_a of the water molecule may be dramatically lowered in its distinctive environment, bound to the unique Fe of the C-cluster.

Protons generated by the reaction must be delivered from the C-cluster to the exterior of the protein to access the bulk solvent, and the first CODH structures revealed a channel that emanates from Lys587, contains water molecules, and is lined with several conserved histidine residues. At that time, it was suggested that this channel could serve as a proton transfer network[16]. Mutational studies have since shown that several of the residues that make up the channel are important for activity[35]. The structures presented here contain substrates that mark the site of catalysis and support the involvement of these residues in proton transfer. Figure IV.11 shows the positions of cyanide and

H_2O/OH^- relative to the conserved residues that make up the proton transfer network, Lys587, His113, His116, His119, His122, and Asn284, the latter of which would not be directly involved in proton transfer but contributes hydrogen bonds that may be structurally important.

IV.F. Conclusions

The structures presented here of native and cyanide-bound C-clusters of the well-characterized *Mt*CODH/ACS have confirmed the initial proposal that a sulfide bridge does not play a catalytic role in CO oxidation and CO_2 reduction in *M. thermoacetica*. Recent crystal structures of *Ch*CODH have further indicated that a bridging sulfide is also not catalytic in *C. hydrogenoformans*. Whether the bridging sulfide plays a physiological role in a pre-catalytic state of the C-cluster remains to be determined, but the catalytic C-cluster should be described as a $NiFe_4S_4$ and not a $NiFe_4S_5$ cluster. With the first structure of the H_2O/OH^--bound C-cluster of *Mt*CODH/ACS, we have also linked spectroscopic work that was carried out using *Mt*CODH/ACS, which pointed to Fe-bound H_2O/OH^-, with crystallographic evidence. Our data also provide the first three-dimensional structural view of the inhibitor cyanide bound to Ni of the C-cluster. After years of considerable debate, these structures have contributed towards a unified view of the chemistry performed by the CODH C-cluster, as for the first time there is consistency between the structures of ligand-bound CODHs, with crystallographic snapshots of H_2O/OH^--, CO-, CN-, and CO_2-bound C-clusters available to provide a portrait of CODH reactivity.

Due to the clear importance of CODHs in the global carbon cycle and in light of the current crisis of global climate change, the details surrounding CO_2/CO interconversion in nature have been eagerly anticipated. Our structures of the CODH C-cluster and the mechanistic insights they offer may have broader applications in the conception and development of model systems that can be used for the removal of environmental CO_2 and CO.

IV.G. Materials and Methods

Protein purification and crystallization

M. thermoacetica (formerly *Clostridium thermoaceticum* strain ATCC 39073) was cultured on glucose in the absence of CO as described[52]. Therefore, any CO present would only be due to CO produced as an intermediate during the growth of the organism. CO production in growth under glucose has been found to be small but not insignificant, with a maximum of 53 ppm of CO in the gas phase[53]. CODH/ACS was purified as previously described[54,55] under strictly anaerobic conditions, with the final sample exhibiting specific activity for CO oxidation to indicate that no CO inactivation had occurred (366 units/mg). Crystals were grown anaerobically as described[10] in an anaerobic chamber (Coy Laboratories) in the absence of CO with 0.25 μL of a microcrystal solution to aid in crystal growth[9]. Native crystals were flash-cooled in liquid nitrogen using 20% glycerol as a cryoprotectant. Cyanide-bound crystals were obtained by soaking native crystals in the precipitant solution that was supplemented with 100 μM KCN for 1 h at room temperature under anaerobic conditions before flash-cooling for X-ray diffraction data collection.

Data collection and refinement

X-ray diffraction data of native and cyanide crystals were collected at Stanford Synchrotron Radiation Laboratory (SSRL) beam line 11-1 and at Advanced Light Source (ALS) beam line 5.0.2, respectively. The asymmetric unit was composed of two ~310 kDa $\alpha_2\beta_2$ tetramers. Native and cyanide data were processed using MOSFLM/SCALA[56,57] and DENZO/SCALEPACK[58], respectively, with good quality diffraction data to 2.15 Å for both structures (Table IV.1).

The first published *Mt*CODH/ACS structure[10] (PDB ID: 1MJG) was used for an initial rigid body refinement, which resulted in *R*-factors below 25% for both the native and the cyanide-bound structures. To obtain test set reflections that were of higher resolution than the published *Mt*CODH/ACS structure, 5% of the reflections from this region of the data were randomly generated. Subsequent rounds of refinement were carried out in REFMAC[26] from the CCP4 Program Suite[56] using noncrystallographic symmetry (NCS) restraints, with iterative model building in COOT[59]. Occupancy refinements of water and cyanide ligands to the C-cluster were carried out in PHENIX[60]. The final model contains residues 2-674 (of 674) for the four CODH chains A-D and residues 2-729 (of 729) for the four ACS chains M-P. Ramachandran analysis of the final model was calculated in PROCHECK[61], and the final data collection and refinement statistics are shown in Table IV.1.

All figures of the structures were made in PyMOL[62]. The *Mt*CODH/ACS native and cyanide-bound C-cluster structures have been deposited into the Protein Data Bank with accession codes 3I01 and 3I04, respectively.

IV.H. Tables and Figures

Table IV.1. Data collection and refinement statistics.

	Native*	Cyanide-bound*
Data Collection		
Space group	*P*1	*P*1
Cell Dimensions		
a, b, c (Å)	99.65, 136.87, 140.86	99.83, 136.77, 141.61
α, β, γ (°)	101.26, 109.11, 104.08	101.23, 109.18, 103.87
Wavelength (Å)	0.95469	1.0000
Resolution (Å)	36.96-2.15 (2.21-2.15)	50.0-2.15 (2.23-2.15)
R_{sym} (%)	6.7 (40.6)	5.1 (16.7)
Mean $I/\sigma I$	12.2 (2.3)	25.9 (6.8)
Completeness (%)	92.3 (90.4)	96.3 (86.6)
Redundancy	2.1	3.5
Refinement		
Resolution (Å)	36.96-2.15	48.39-2.15
R_{work}/R_{free} (%)	18.6/24.2	17.2/22.1
number of atoms		
protein	43,732	43,619
heteroatoms (Fe, S, Ni, Cu)	124	124
ligands	16	24
waters	1,413	2,153
other ions	28	28
average B-factor, all atoms (Å^2)	36.9	35.6
protein	37.1	35.7
C-clusters	28.2	25.2
C-cluster ligands	31.7	32.1
heteroatoms (Fe, S, Ni, Cu)	31.1	28.4
ligands	56.5	55.1
waters	31.2	33.8
other ions	41.0	39.1
rms deviations		
bond lengths (Å)	0.021	0.019
bond angles (°)	1.871	1.650
Ramachandran plot (%)		
most favored	88.1	89.8
additionally allowed	10.5	8.9
generously allowed	0.9	0.7
disallowed	0.6	0.5

* Highest resolution shell is shown in parentheses

Table IV.2. Occupancy and *B*-factor refinement trials of H_2O/OH^- ligands. Native *Mt*CODH/ACS structure.

Chain	Average *B*-factor (Å^2)		Oxygen *B*-factor (Å^2) (occupancy set to 1.0)	Oxygen occupancy, (*B*-factor set to C-cluster average)
	Chain	C-cluster		
A	27.03	25.16	29.51	0.9
B	24.40	21.04	20.73	1.0
C	29.44	29.35	34.90	0.9
D	32.26	37.21	41.75	0.8
Average	28.28	28.19	31.72	0.9

Table IV.3. Occupancy and *B*-factor refinement trials of the H_2O/OH^- and cyanide ligands. Cyanide-bound *Mt*CODH/ACS structure.

Chain	Average *B*-factor (Å^2)		Atom *B*-factor (Å^2) (occupancy set to 1.0)		Atom occupancy (*B*-factor set to C-cluster average)
	Chain	C-cluster			
A	26.13	23.40	O	24.74	1.0
			C	31.25	0.6
			N	32.22	0.7
B	24.21	20.80	O	24.71	0.8
			C	25.48	0.8
			N	25.33	0.8
C	27.92	25.39	O	31.97	0.8
			C	33.80	0.5
			N	31.85	0.9
D	30.00	31.28	O	37.69	0.8
			C	42.58	0.4
			N	43.73	0.7
Average	27.06	25.22		32.11	0.7

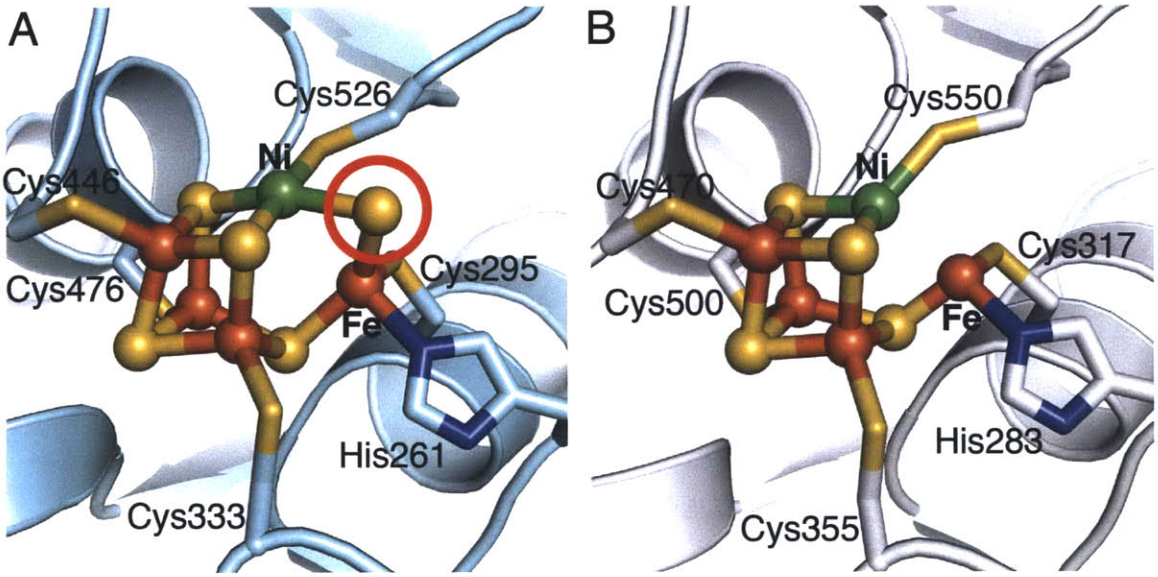


Figure IV.1. Representative structures of CODH C-clusters. A, The C-cluster from *Ch*CODH[17] (cyan ribbons, PDB ID: 1SU8) with the sulfide bridge (red circle). Residue numbering follows the *Ch*CODH sequence **B**, The C-cluster from *Mt*CODH/ACS[10] (grey ribbons, PDB ID: 1MJG) without the sulfide bridge. Residue numbering follows the *Mt*CODH/ACS sequence. C-clusters are depicted in ball-and-stick, while ligands are shown in sticks: Ni in green, Fe in orange, S in yellow, and Ni in blue. Ni and the unique Fe of the C-cluster are labeled Ni and Fe, respectively.

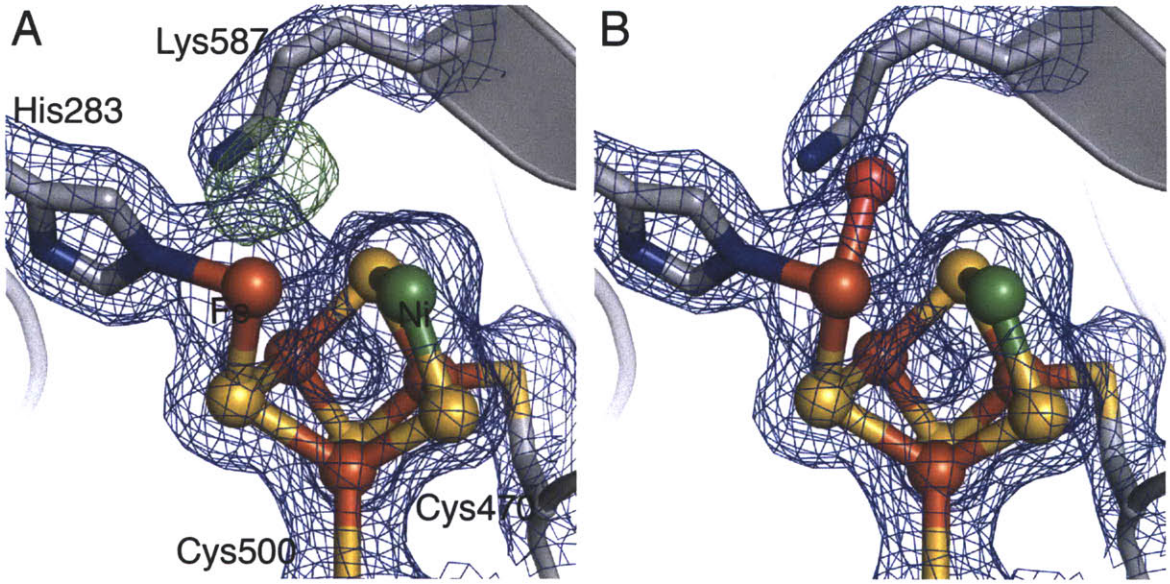


Figure IV.2. Electron density of the native *Mt*CODH/ACS C-cluster. A, Native C-cluster before and **B**, after H_2O/OH^- is modeled. H_2O/OH^- is bound 2.03 Å to the unique Fe and 2.60 Å to Lys587. Protein chain shown in grey ribbons, C-clusters in ball-and-stick, and ligands are shown as sticks: Ni in green, Fe in orange, S in yellow, N in blue. H_2O/OH^- is depicted as a red sphere. $2F_o$-F_c (1.5 σ) and F_o-F_c (4.0 σ) electron density maps are shown in blue and green mesh, respectively. Ni and the unique Fe are labeled in **A** as Ni and Fe, respectively, with residues following *Mt*CODH/ACS numbering. For clarity, not all protein ligands to the cluster are shown.

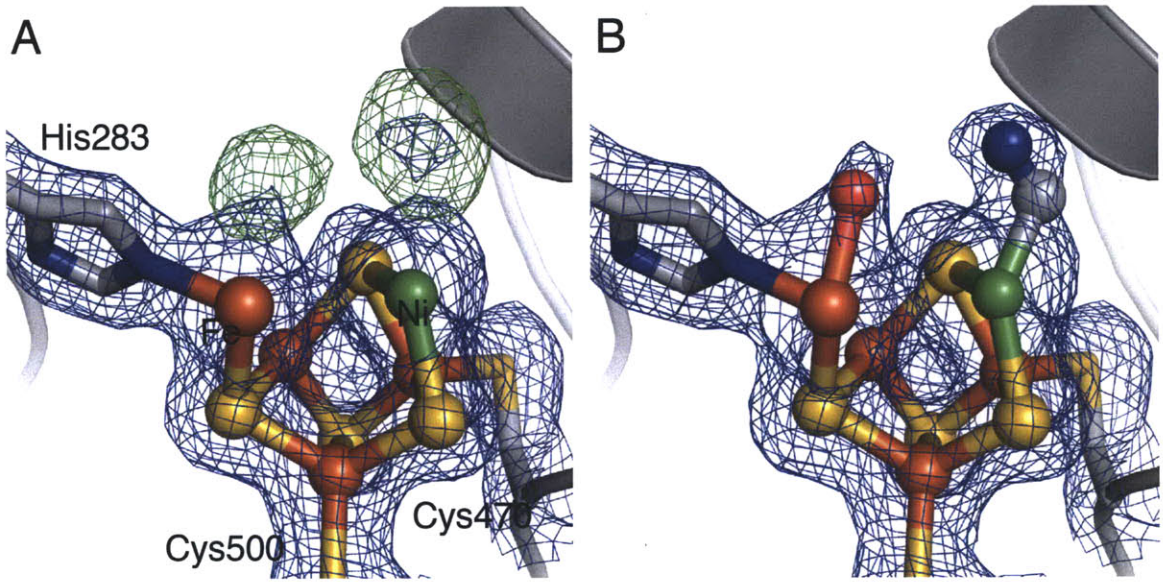


Figure IV.3. Electron density of the *Mt*CODH/ACS cyanide-bound C-cluster. A, Cyanide-bound C-cluster before and **D**, after H_2O/OH^- and cyanide are modeled. H_2O/OH^- is bound 2.05 Å to the unique Fe, and cyanide is bound 1.99 Å to Ni, with an average Ni-C-N angle of 114°. Protein chain shown in grey ribbons, C-clusters in ball-and-stick, and ligands are shown as sticks: Ni in green, Fe in orange, S in yellow, N in blue. H_2O/OH^- is depicted as a red sphere. $2F_o$-F_c (1.5 σ) and F_o-F_c (4.0 σ) electron density maps are shown in blue and green mesh, respectively. Ni and the unique Fe are labeled in **A** as Ni and Fe, respectively, with residues following *Mt*CODH/ACS numbering. For clarity, not all protein ligands to the cluster are shown, and Lys587 shown in Figure IV.2 is hidden.

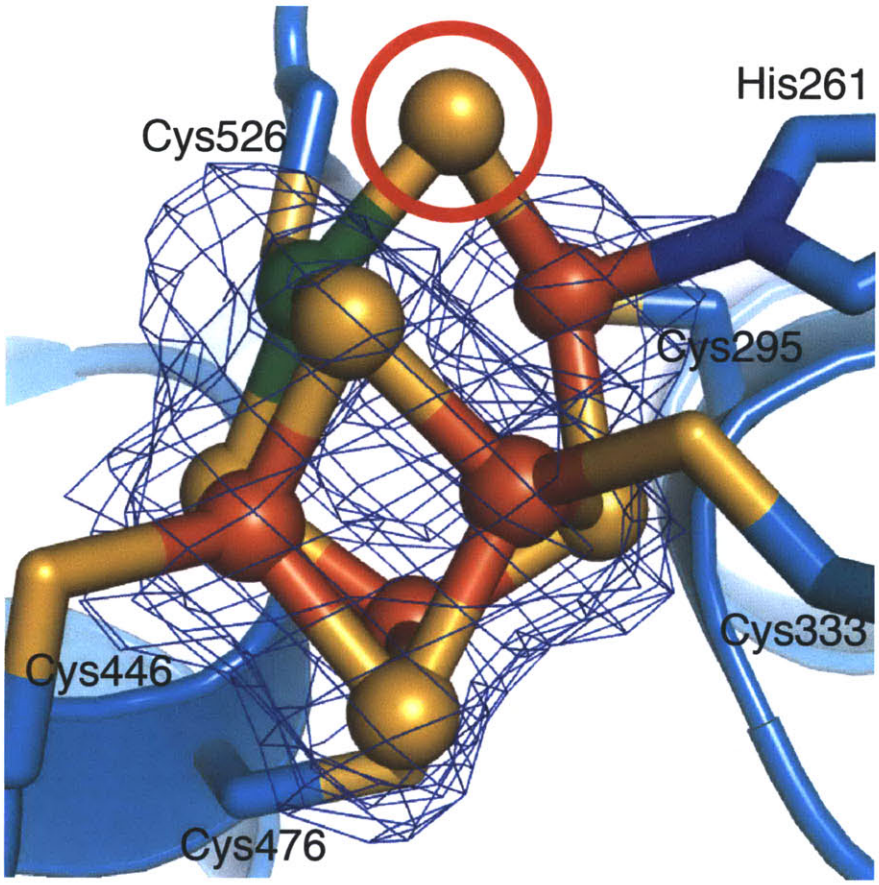


Figure IV.4. Lack of density to support a bridging cyanide. C-cluster model of *Ch*CODH with the bridging sulfide (cyan ribbons, PDB ID: 1SU8) superimposed with the $2F_o$-F_c density (blue mesh, 2.0 σ) of the cyanide-bound C-cluster of *Mt*CODH/ACS presented here. Residue numbering following the *Ch*CODH sequence. The bridging sulfide protrudes from the density at the top of the figure, circled in red.

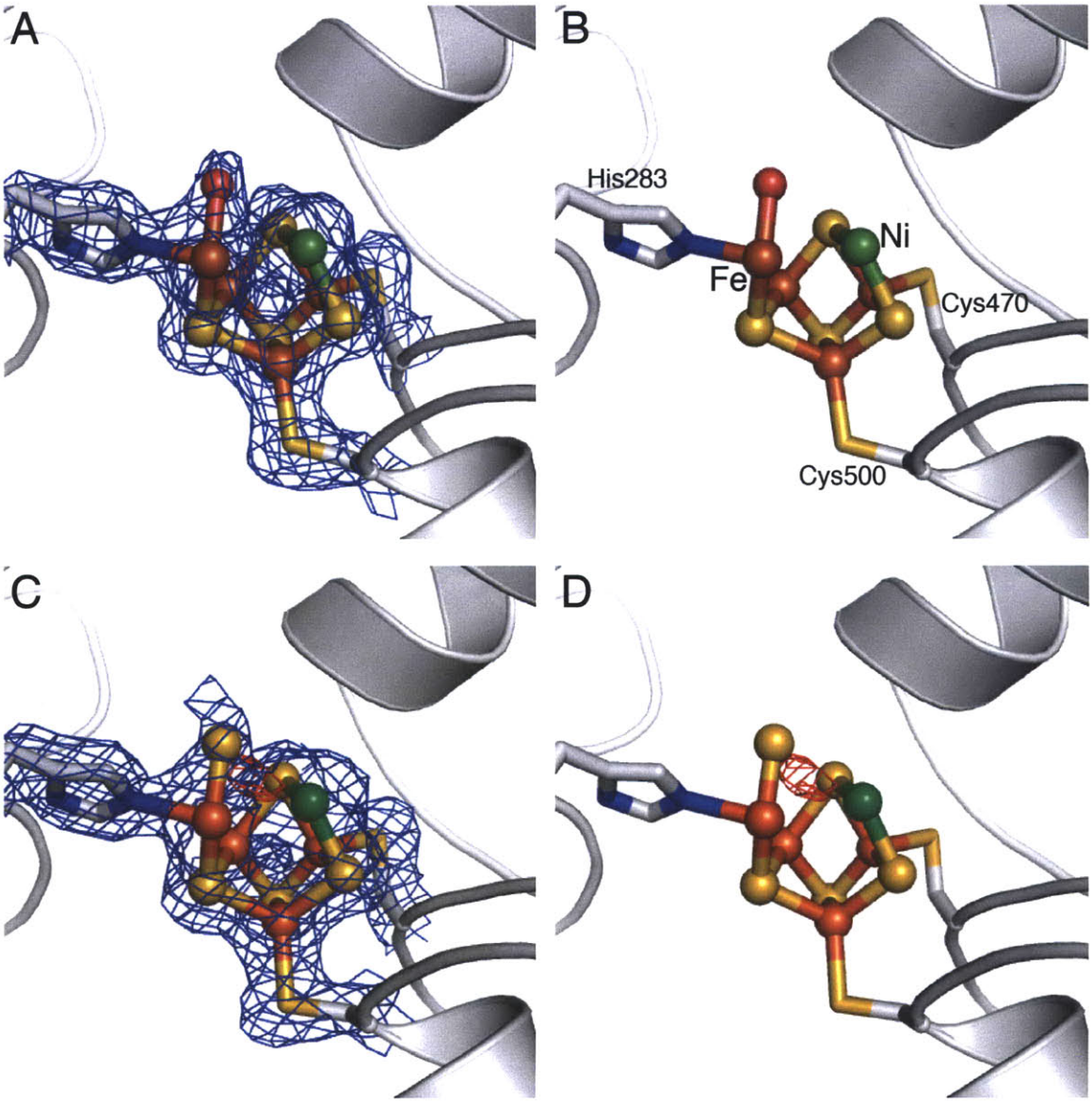


Figure IV.5. Refinement trials for the native *Mt*CODH/ACS structure. Refinement results after **A-B**, H_2O/OH^- and **C-D**, sulfur was modeled bound to the unique Fe, labeled as Fe in **B**. In both **A** and **C**, $2F_o$-F_c electron density maps are contoured at 1.5 σ (blue mesh) and F_o-F_c maps are contoured at -3.0 σ (red mesh). No positive F_o-F_c density is present at +3.0 σ. **B** and **D** show only the F_o-F_c difference density. *Mt*CODH/ACS is shown in grey ribbons, the C-cluster and H_2O/OH^- and sulfur ligands are shown in ball-and-stick, and protein ligands to the cluster are shown in sticks. C in grey, N in blue, O in red, Fe in orange, Ni in green, and S in yellow. For clarity, not all protein ligands to the cluster are shown.

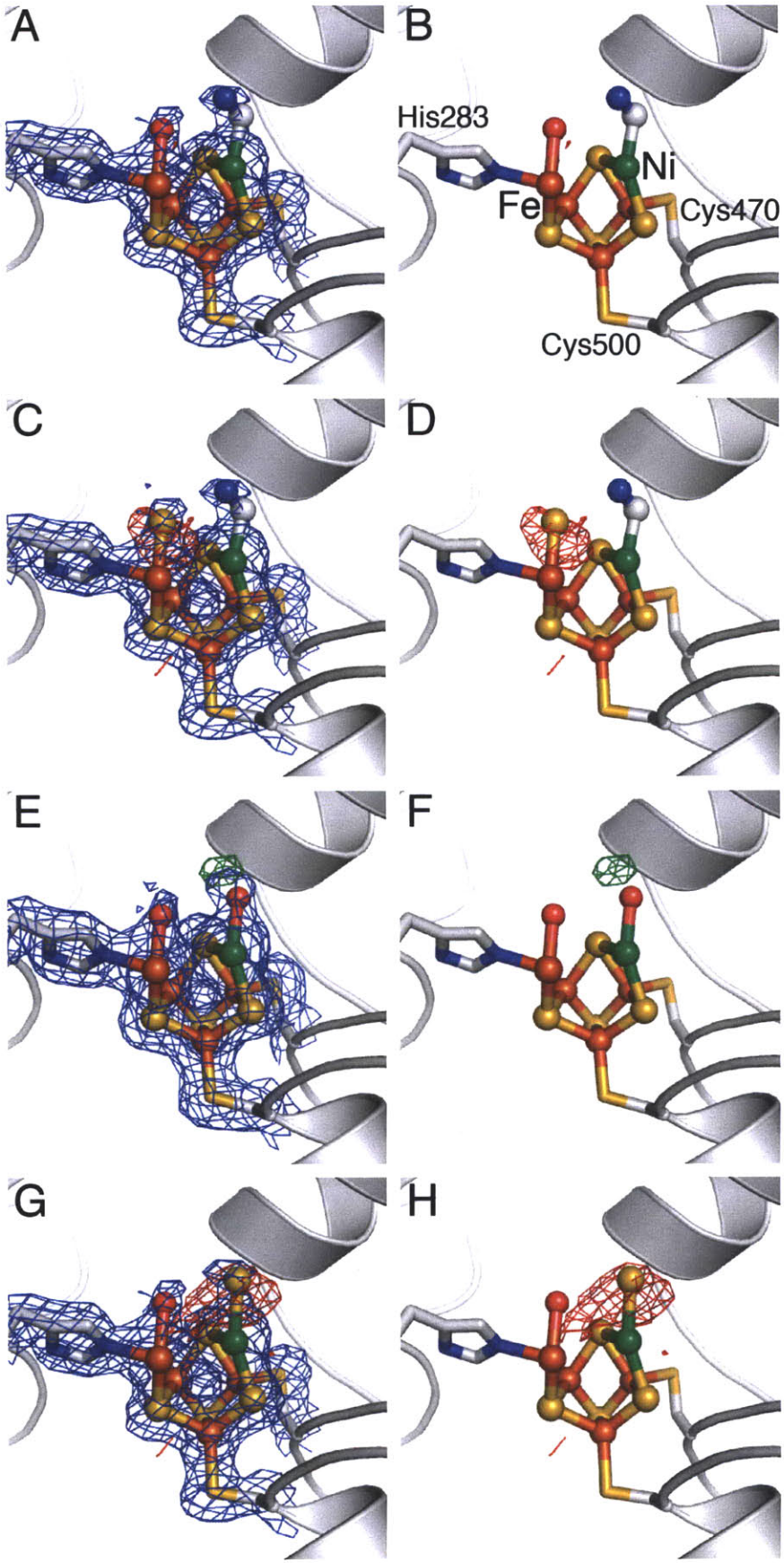

A
B
His283
Ni
Fe
Cys470
Cys500
C
D
E
F
G
H

Figure IV.6. Refinement trials for the cyanide-bound *Mt*CODH/ACS structure. Refinement results after the following were modeled: **A-B**, H_2O/OH^- bound to the unique Fe and cyanide bound to Ni, labeled in **B** as Fe and Ni, respectively; **C-D**, sulfur bound to Fe and cyanide bound to Ni; **E-F**, H_2O/OH^- bound to both Fe and Ni; **G-H**, H_2O/OH^- bound to Fe and sulfur bound to Ni. **A**, **C**, **E**, and **G** show $2F_o$-F_c electron density maps contoured at 1.5 σ (blue mesh) and F_o-F_c difference density maps contoured at -3.5 σ (red mesh) and +3.5 (green mesh). **B**, **D**, **F**, and **H** show only the F_o-F_c difference density maps. *Mt*CODH/ACS is shown in grey ribbons, the C-cluster and H_2O/OH^-, cyanide, and sulfur ligands are shown in ball-and-stick, and protein ligands to the cluster are shown in sticks. C in grey, N in blue, O in red, Fe in orange, Ni in green, and S in yellow. For clarity, not all protein ligands to the cluster are shown.

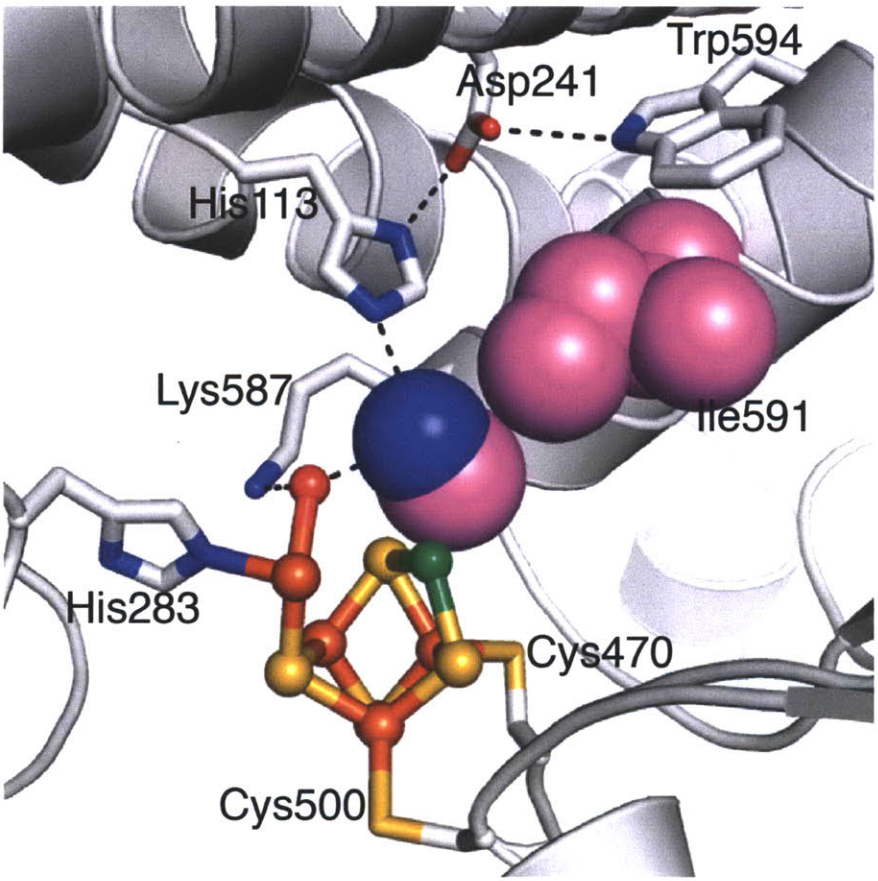


Figure IV.7. Residues surrounding the cyanide binding site of *Mt*CODH/ACS. Ile591 (*Mt*CODH/ACS numbering) hinders linear Ni-cyanide binding at the C-cluster. Cyanide and Ile591 are shown in spheres with van der Waals radii and with carbon atoms in violet. *Mt*CODH/ACS in grey ribbons, C-cluster in ball-and-stick, and protein residues shown as sticks. Ni in green, Fe in orange, S in yellow, N in blue, and O in red. For clarity, not all protein ligands to the cluster are shown.

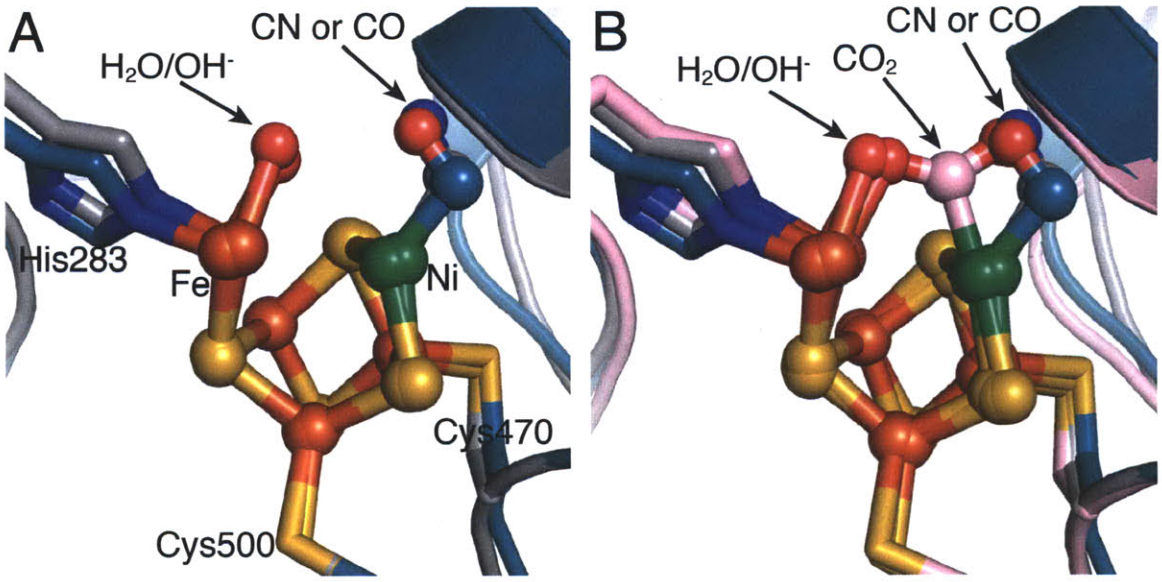


Figure IV.8. Superposition of cyanide-bound *Mt*CODH/ACS structure with other CODH C-cluster structures with bound substrates. A, C-clusters from cyanide-bound *Mt*CODH/ACS (grey ribbons) and putative CO-bound *Mb*CODH (teal ribbons), with residue numbering following the *Mt*CODH/ACS. Ni and the unique Fe of the C-cluster are labeled Ni and Fe, respectively. **B**, C-clusters from **A** plus that from *Ch*CODH (pink ribbons). C-clusters are depicted in ball-and-stick and protein ligands are shown in sticks. Ni in green, Fe in orange, S in yellow, N in blue, and O in red. For clarity, not all protein ligands to the cluster are shown.

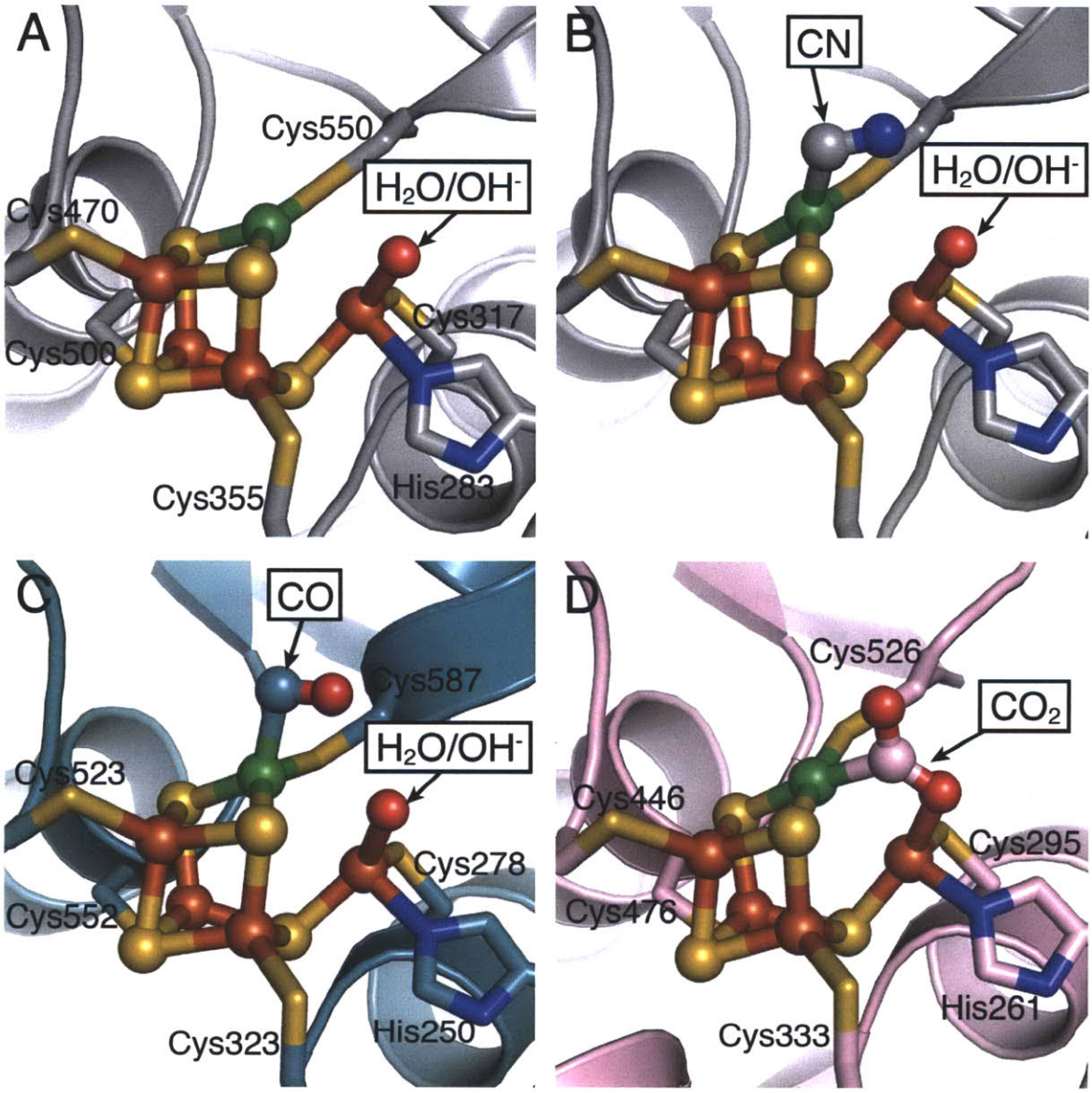


Figure IV.9. Structures of substrates and/or inhibitors bound to the C-clusters from various CODHs. A, Native *Mt*CODH/ACS structure (grey ribbons) with H_2O/OH^- bound. **B**, Cyanide-bound *Mt*CODH/ACS structure (grey ribbons) with cyanide and H_2O/OH^- bound. Residue numbering for *Mt*CODH/ACS shown in **A** only. **C**, CO and H_2O/OH^- bound to *Mb*CODH (teal ribbons), with residue numbering following the *Mb*CODH sequence. **D**, CO_2 bound to *Ch*CODH (pink ribbons), with residue numbering following the *Ch*CODH sequence. C-cluster are depicted in ball-and-stick while protein ligands are shown in sticks. Ni in green, Fe in orange, S in yellow, N in blue, and O in red.

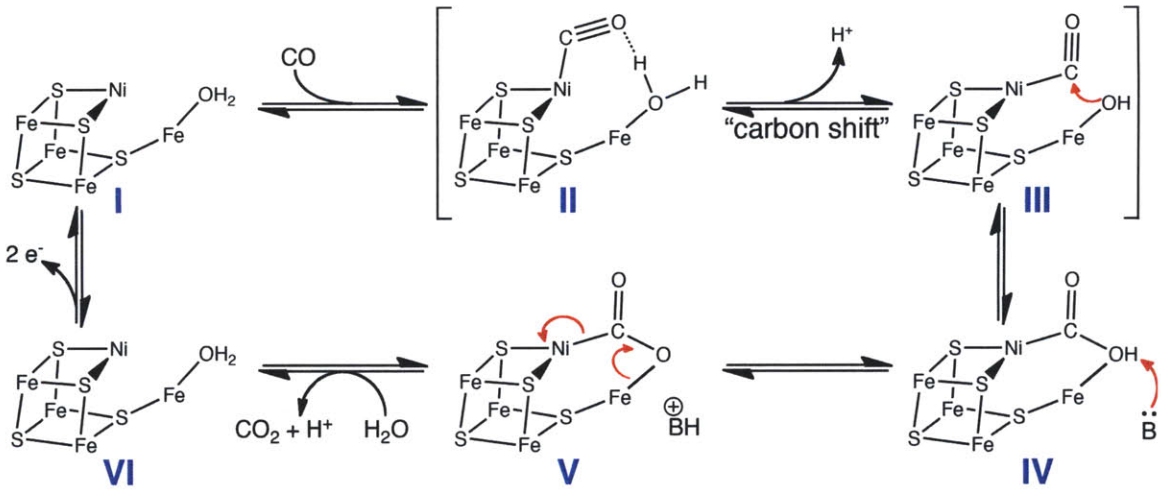


Figure IV.10. Proposed mechanism of the C-cluster. The native *Mt*CODH/ACS structure presented here represent state I of this mechanism, while the cyanide-bound *Mt*CODH/ACS structure represents state II, as does the putative CO-bound *Mb*CODH structure. The CO_2-bound *Ch*CODH structure represents state IV or V, which differ only in the protonation state of the intermediate and would thus be difficult to distinguish using X-ray crystallography. States I and VI are structurally identical and differ only the redox state of the cluster. "B" refers to a basic residue involved in deprotonation. The identity of this base has not been resolved, and candidates include His113 and Lys587.

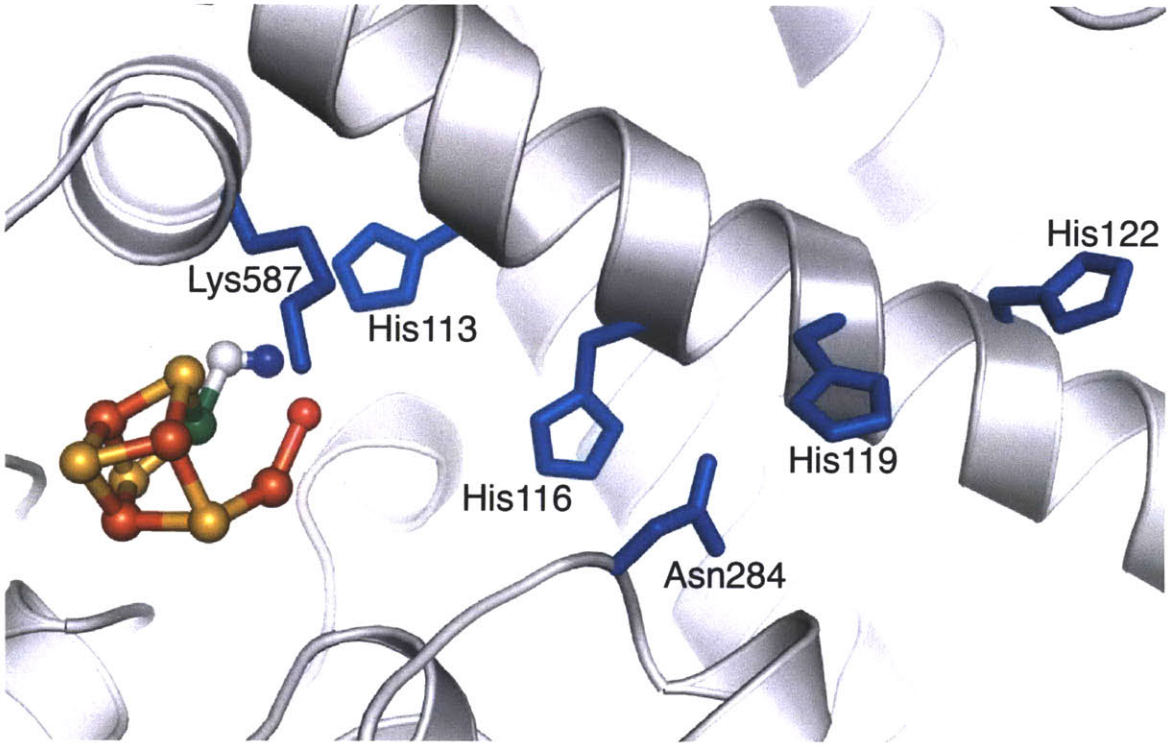


Figure IV.11. Proposed proton transfer network. *Mt*CODH/ACS is depicted in grey ribbons, residues involved in the proton transfer network are shown in slate sticks, and the C-cluster atoms with cyanide and H_2O/OH^- ligands are shown in ball-and-stick: Ni in green, Fe in orange, S in yellow, O in red, and N in blue.

IV.I. References

1 Svetlichny, V. A., Sokolova, T. G., Gerhardt, M., Ringpfeil, M., Kostrikina, N. A. & Zavarzin, G. A. (1991) *Carboxydothermus hydrogenoformans* gen. nov., sp. nov., a CO-Utilizing Thermophilic Anaerobic Bacterium from Hydrothermal Environments of Kunashir Island. *Syst. Appl. Microbiol.* **14**, 254-260.

2 Uffen, R. L. (1976) Anaerobic Growth of a *Rhodopseudomonas* Species in the Dark with Carbon Monoxide as Sole Carbon and Energy Substrate. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3298-3302.

3 Bartholomew, G. W. & Alexander, M. (1979) Microbial Metabolism of Carbon Monoxide in Culture and in Soil. *Appl. Environ. Microbiol.* **37**, 932-937.

4 Lindahl, P. A. (2002) The Ni-Containing Carbon Monoxide Dehydrogenase Family: Light at the End of the Tunnel? *Biochemistry.* **41**, 2097-2105.

5 Müller, V., Imkamp, F., Rauwolf, A., Küsel, K. & Drake, H. L. (2004) "Molecular and Cellular Biology of Acetogenic Bacteria" in *Strict and Facultative Anaerobes: Medical and Environmental Aspects*, eds Michiko M. Nakano & Peter Alan Zuber. Horizon Bioscience: Wymondham, UK.

6 Ragsdale, S. W. (1997) The Eastern and Western Branches of the Wood/Ljungdahl Pathway: How the East and West Were Won. *Biofactors.* **6**, 3-11.

7 Ragsdale, S. W. & Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 Fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898.

8 Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003) Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] Clusters in Closed and Open Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. *Nat. Struct. Biol.* **10**, 271-279.

9 Doukov, T. I., Blasiak, L. C., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2008) Xenon in and at the End of the Tunnel of Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **47**, 3474-3483, doi:10.1021/bi702386t.

10 Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002) A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Science.* **298**, 567-572.

11 Maynard, E. L. & Lindahl, P. A. (1999) Evidence of a Molecular Tunnel Connecting the Active Sites for CO_2 Reduction and Acetyl-CoA Synthesis in Acetyl-CoA Synthase from *Clostridium thermoaceticum. J. Am. Chem. Soc.* **121**, 9221-9222.

12 Seravalli, J. & Ragsdale, S. W. (2000) Channeling of Carbon Monoxide during Anaerobic Carbon Dioxide Formation. *Biochemistry.* **39**, 1274-1277.

13 Drake, H. L., Daniel, S. L., Matthies, C. & Küsel, K. (1994) "Acetogenesis, acetogenic bacteria, and the acetyl-CoA pathway: past and current perspectives" in *Acetogenesis*, Ed. H.L. Drake. Chapman and Hall: New York, 3-60.

14 Grahame, D. A. (1991) Catalysis of Acetyl-CoA Cleavage and Tetrahydrosarcinapterin Methylation by a Carbon Monoxide Dehydrogenase-Corrinoid Enzyme Complex. *J. Biol. Chem.* **266**, 22227-22233.

15 Terlesky, T. C., Nelson, M. J. K. & Ferry, J. G. (1986) Isolation of an Enzyme Complex with Carbon Monoxide Dehydrogenase Activity Containing Corrinoid and Nickel from Acetate-Grown *Methanosarcina thermophila. J. Bacteriol.* **168**, 1053-1058.

16 Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. & Ludden, P. W. (2001) Life on Carbon Monoxide: X-ray Structure of *Rhodospirillum rubrum* Ni-Fe-S Carbon Monoxide Dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11973-11978.

17 Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, O. (2001) Crystal Structure of a Carbon Monoxide Dehydrogenase Reveals a [Ni-4Fe-5S] Cluster. *Science.* **293**, 1281-1285.

18 Dobbek, H., Svetlitchnyi, V., Liss, J. & Meyer, O. (2004) Carbon Monoxide Induced Decomposition of the Active Site [Ni-4Fe-5S] Cluster of CO Dehydrogenase. *J. Am. Chem. Soc.* **126**, 5382-5387.

19 Feng, J. & Lindahl, P. A. (2004) Effect of Sodium Sulfide on Ni-Containing Carbon Monoxide Dehydrogenases. *J. Am. Chem. Soc.* **126**, 9094-9100.

20 Drennan, C. L. & Peters, J. W. (2003) Surprising Cofactors in Metalloenzymes. *Curr. Opin. Struct. Biol.* **13**, 220-226.

21 Chen, J., Huang, S., Seravalli, J., Gutzma, H. J., Swartz, D. J., Ragsdale, S. W. & Bagley, K. A. (2003) Infrared Studies of Carbon Monoxide Binding to Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase from *Moorella thermoacetica. Biochemistry.* **42**, 14822-14830.

22 Seravalli, J. & Ragsdale, S. W. (2008) ^{13}C NMR Characterization of an Exchange Reaction between CO and CO_2 Catalyzed by Carbon Monoxide Dehydrogenase. *Biochemistry.* **47**, 6770-6781.

23 Ha, S.-W., Korbas, K., Klepsch, M., Meyer-Klaucke, W., Meyer, O. & Svetlitchnyi, V. (2007) Interaction of Potassium Cyanide with the [Ni-4Fe-5S] Active Site Cluster of CO Dehydrogenase from *Carboxydothermus hydrogenoformans. J. Biol. Chem.* **282**, 10639-10646.

24 Jeoung, J.-H. & Dobbek, H. (2007) Carbon Dioxide Activation at the Ni,Fe-Cluster of Anaerobic Carbon Monoxide Dehydrogease. *Science.* **318**, 1461-1464.

25 Gong, W., Hao, B., Wei, Z., Ferguson, D. J. J., Tallant, T., Krzycki, J. A. & Chan, M. K. (2008) Structure of the $\alpha_2\varepsilon_2$ Ni-Dependent CO Dehydrogenase Component of the *Methanosarcina barkeri* Acetyl-CoA Decarbonylase/Synthase Complex. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9558-9563.

26 Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **D53**, 240-255.

27 Brünger, A. T., Adams, P. D., Clore, M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998) *Crystallography & NMR System*: A New Software Suite for Marcromolecular Structure Determination. *Acta Cryst.* **D54**, 905-921.

28 Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010) *PHENIX*: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Cryst.* **D66**.

29 Higuchi, Y., Ogata, H., Miki, N., Yosuoka, N. & Yagi, T. (1999) Removal of the Bridging Ligand Atom at the Ni-Fe Active Aite of [NiFe] Hydrogenase Upon Reduction with H_2, As Revealed by X-ray Structure Analysis at 1.4 Å Resolution. *Structure.* **7**, 549-556.

30 Higuchi, Y. & Yagi, T. (1999) Liberation of Hydrogen Sulfide During the Catalytic Action of *Desulfovibrio* Hydrogenase Under the Atmosphere of Hydrogen. *Biochem. Biophys. Res. Commun.* **255**, 295-299.

31 DeRose, V. J., Telser, J., Andersen, M. E., Lindahl, P. A. & Hoffman, B. M. (1998) A Multinuclear ENDOR Study of the C-Cluster in CO Dehydrogenase from *Clostridium thermoaceticum*: Evidence for H_xO and Histidine Coordination to the $[Fe_4S_4]$ Center. *J. Am. Chem. Soc.* **120**, 8767-8776.

32 Andersen, M. E. & Lindahl, P. A. (1994) Organization of Clusters and Internal Electron Pathways in CO Dehydrogenase from *Clostridium thermoaceticum*: Relevance to the Mechanism of Catalysis and Cyanide Inhibition. *Biochemistry.* **33**, 8702-8711.

33 Andersen, M. E. & Lindahl, P. A. (1996) Spectroscopic States of the CO Oxidation/CO_2 Reduction Active Site of Carbon Monoxide Dehydrogenase and Mechanistic Implications. *Biochemistry.* **35**, 8371-8390.

34 Seravalli, J., Kumar, M., Lu, W.-P. & Ragsdale, S. W. (1997) Mechanism of Carbon Monoxide Oxidation by the Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase from *Clostridium thermoaceticum*: Kinetic Characterization of the Intermediates. *Biochemistry.* **36**, 11241-11251.

35 Kim, E. J., Feng, J., Bramlett, M. R. & Lindahl, P. A. (2004) Evidence for a Proton Transfer Network and a Required Persulfide-Bond-Forming Cysteine Residue in Ni-Containing Carbon Monoxide Dehydrogenases. *Biochemistry.* **43**, 5728-5734.

36 Andersen, M. E., DeRose, V. J., Hoffman, B. M. & Lindahl, P. A. (1993) Identification of a Cyanide Binding Site in CO Dehydrogenase from *Clostridium thermoaceticum* Using EPR and ENDOR Spectroscopies. *J. Am. Chem. Soc.* **115**, 12204-12205.

37 Ensign, S. A., Bonam, D. & Ludden, P. W. (1989) Nickel is Required for the Transfer of Electrons from Carbon Monoxide to the Iron-Sulfur Center(s) of Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum*. *Biochemistry.* **28**, 4968-4973.

38 Ensign, S. A., Hyman, M. R. & Ludden, P. W. (1989) Nickel-Specific, Slow-Binding Inhibition of Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum* by Cyanide. *Biochemistry.* **28**, 4973-4979.

39 Hu, Z., Spangler, H. J., Andersen, M. E., Xia, J., Ludden, P. W., Lindahl, P. A. & Münck, E. (1996) Nature of the C-Cluster in Ni-Containing Carbon Monoxide Dehydrogenases. *J. Am. Chem. Soc.* **118**, 830-845.

40 Al-Mustafa, J., Sykora, M. & Kincaid, J. R. (1995) Resonance Raman Investigation of Cyanide Ligated Beef Liver and *Aspergillus niger* Catalases. *J. Biol. Chem.* **270**, 10449-10460.

41 Al-Mustafa, J. & Kincaid, J. R. (1994) Resonance Raman Study of Cyanide-Ligated Horseradish Peroxidase. Detection of Two Binding Geometries and Direct Evidence for the "Push-Pull" Effect. *Biochemistry.* **33**, 2191-2197.

42 Blair-Johnson, M., Fiedler, T. & Fenna, R. (2001) Human Myeloperoxidase: Structure of a Cyanide Complex and Its Interaction with Bromide and Thiocyanate Substrates at 1.9 Å Resolution. *Biochemistry.* **40**, 13990-13997.

43 Han, S., Madden, J. F., Siegel, L. M. & Spiro, T. G. (1989) Resonance Raman Studies of *Escherichia coli* Sulfite Reductase Hemoprotein. 3. Bound Ligand Vibrational Modes. *Biochemistry.* **28**, 5477-5485.

44 Bolognesi, M., Rosana, C., Losso, R., Borassi, A., Rizzi, M., Wittenberg, J. B., Boffi, A. & Ascenzi, P. (1999) Cyanide Binding to *Lucina pectinata* Hemoglobin I and to Sperm Whale Myoglobin: An X-Ray Crystallographic Study. *Biophys. J.* **77**, 1093-1099.

45 Jafferji, A., Allen, J. W. A., Ferguson, S. J. & Fülöp, V. (2000) X-ray Crystallographic Study of Cyanide Binding Provides Insights into the Structure-Function Relationship for Cytochrome cd_1 Nitrite Reductase from *Paracoccus pantotrophus*. *J. Biol. Chem.* **275**, 25089-25094.

46 Deng, T., Macdonald, I. D. G., Simianu, M. C., Sykora, M., Kincaid, J. R. & Sligar, S. G. (2001) Hydrogen-Bonding Interactions in the Active SItes of Cytochrome P450cam and Its Site-Directed Mutants. *J. Am. Chem. Soc.* **123**, 269-278.

47 Simianu, M. C. & Kincaid, J. R. (1995) Resonance Raman Spectroscopic Detection of Both Linear and Bent Fe-CN Fragments for the Cyanide Adducts of Cytochrome P-450 Camphor and Its Substrate-Bound Forms. Relevance to the "Charge Relay" Mechanism. *J. Am. Chem. Soc.* **117**, 4628-4636.

48 Clay, M. D., Yang, T.-C., Jenney, F. E. J., Kung, I. Y., Cosper, C. A., Krishnan, R., Kurtz, D. M. J., Adams, M. W. W., Hoffman, B. M. & Johnson, M. K. (2006) Geometric and Electronic Structures of Cyanide Adducts of the Non-Heme Iron Active Site of Superoxide Reductases: Vibrational and ENDOR Studies. *Biochemistry.* **45**, 427-438.

49 Jeoung, J.-H. & Dobbek, H. (2009) Structural Basis of Cyanide Inhibition of Ni, Fe-Containing Carbon Monoxide Dehydrogenase. *J. Am. Chem. Soc.* **131**, 9922-9923.

50 Menon, S. & Ragsdale, S. W. (1996) Unleashing Hydrogenase Activity in Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase and Pyruvate:Ferredoxin Oxidoreductase. *Biochemistry.* **35**, 15814-15821.

51 Ragsdale, S. W., Clark, J. E., Ljungdahl, L. G., Lundie, L. L. & Drake, H. L. (1983) Properties of Purified Carbon Monoxide Dehydrogenase from *Clostridium thermoaceticum*, a Nickel, Iron-Sulfur Protein. *J. Biol. Chem.* **258**, 2364-2369.

52 Ljungdahl, L. G. & Andreesen, J. R. (1978) Formate Dehydrogenase, a Selenium-Tungsten Enzyme from *Clostridium thermoaceticum*. *Methods Enzymol.* **53**, 360-372.

53 Diekert, G., Hansch, M. & Conrad, R. (1984) Acetate Synthesis from 2 CO_2 in Acetogenic Bacteria: Is Carbon Monoxide an Intermediate? *Arch. Microbiol.* **138**, 224-228.

54 Andreesen, J. R., Schaupp, A., Neurauter, C., Brown, A. & Ljungdahl, L. G. (1973) Fermentation of Glucose, Fructose, and Xylose by *Clostridium thermoaceticum*: Effect of Metals on Growth Yield, Enzymes, and the Synthesis of Acetate from CO_2. *J. Bacteriol.* **114**, 743-751.

55 Ragsdale, S. W., Ljungdahl, L. G. & DerVartanian, D. V. (1983) Isolation of Carbon Monoxide Dehydrogenase from *Acetobacterium woodii* and Comparison

of its Properties with Those of the *Clostridium thermoaceticum* Enzyme. *J. Bacteriol.* **155**, 1224-1237.

56 Collaborative Computational Project, N. (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* **D50**, 760-763.

57 Leslie, A. G. W. (1992) Recent Changes to the MOSFLM Package for Processing Film and Image Plate Data. *Joint CCPE + ESF-EAMCB Newsletter on Protein Crystallography* **No. 26**.

58 Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307-326.

59 Emsley, P. & Cowtan, K. (2004) Coot: Model-Building Tools for Molecular Graphics. *Acta Cryst.* **D60**, 2126-2132.

60 Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002) PHENIX: Building New Software for Automated Crystallographic Structure Determination. *Acta Cryst.* **D58**, 1948-1954.

61 Laskowski, R. A., MacArthur, M. W. & Thornton, J. M. (1993) PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Cryst.* **26**, 283-291.

62 DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA (http://www.pymol.org).

Chapter V: A structural convergence on CODH C-cluster catalysis and inhibition

V.A. Summary

Ni-Fe containing enzymes are involved in the biological utilization of carbon monoxide, carbon dioxide, and hydrogen. Interest in these enzymes has increased in recent years due to hydrogen fuel initiatives and concerns over the development of new methods for CO_2 sequestration. One Ni-Fe enzyme, carbon monoxide dehydrogenase (CODH), is a key player in the global carbon cycle and carries out the interconversion of the environmental pollutant CO and the greenhouse gas CO_2. The Ni-Fe center responsible for this important chemistry, the C-cluster, has been the source of much controversy, but several recent structural studies have helped to direct the field toward a unifying mechanism. Here, a summary of the current state of understanding of this fascinating metallocluster is presented.

A manuscript similar to this chapter has been published as:
Kung, Y. and Drennan, C.L. (2011) A Role for Nickel-Iron Cofactors in Biological Carbon Monoxide and Carbon Dioxide Utilization. *Curr. Opin. Chem. Biol.* **15**, 276-283.

V.B. Introduction

Although enzymes that utilize iron-containing active sites catalyze a wide range of well-known chemical transformations, three remarkable enzymes combine iron and nickel into complex metalloclusters that extend Nature's biochemical toolkit and lie at the heart of fundamental biological processes involving microbial hydrogen utilization and carbon fixation. [NiFe]-hydrogenase can catalyze both H_2 oxidation and evolution in anaerobic microbes to consume or produce protons and electrons, the biological equivalent of the hydrogen fuel cell anode[1,2]. Involved in carbon fixation, the enzyme acetyl coenzyme A synthase (ACS) contains a Ni-Fe-S active site metal center called the A-cluster that combines carbon monoxide (CO) with a methyl group and coenzyme A (CoA) to form acetyl-CoA, generating a source of carbon for a variety of microbes. CO is often provided to ACS by another Ni-Fe enzyme called carbon monoxide dehydrogenase (CODH), a dimeric enzyme which contains a distinctive Ni-Fe-S metal center termed the C-cluster that carries out the reversible reduction of carbon dioxide (CO_2) to CO, the biological equivalent of the water-gas shift reaction and a mechanism for CO_2 utilization. Interest in all three enzymes has increased dramatically in recent years due to renewed attention to the development of hydrogen fuel cells and the design of CO_2 sequestration technologies.

While the first X-ray crystal structures of [NiFe]-hydrogenase[3], ACS[4], and CODH[5,6] revealed the overall architecture of these complex metallocenters, recent work has been aimed at developing an understanding of the mechanisms by which these clusters catalyze their respective reactions. Major advances have been made on the CODH C-cluster within the last three years, with the description of several new crystal

structures. Although the CODH literature has had its share of controversies, the CODH community is now converging on a consensus mechanism, a timely achievement toward understanding one of Nature's solutions for CO_2 utilization.

V.C. CODH and the global carbon cycle

CODH plays a central role in the global carbon cycle in anaerobic microorganisms (Figure V.1). Some microbes, such as *Rhodospirillum rubrum* and *Carboxydothermus hydrogenoformans* depend upon a monofunctional CODH in their ability to use CO oxidation as a sole carbon and energy source[7,8]. It is estimated that CODH activity accounts for the annual removal of 10^8 tons of CO from the environment[9]. Acetogenic bacteria, such as *Moorella thermoacetica*, couple CODH-catalyzed reduction of CO_2 to CO with the synthesis of acetyl-CoA in a bifunctional CODH/ACS complex[10]. Here, CO produced from CO_2 at the C-cluster is an intermediate, which then binds the active site metallocluster of ACS, the A-cluster, where it finally becomes the carbonyl of acetyl-CoA. Acetyl-CoA is then either converted into cellular biomass, or its high-energy thioester bond can be cleaved to drive phosphorylation of adenosine diphosphate (ADP) to adenosine triphosphate (ATP), producing acetate as a waste product. It is estimated that 10^{10} tons of acetate are produced globally from CO_2 through this process every year by anaerobic acetogens[11]. Additionally, CODH and ACS components are present in the acetyl-CoA decarbonylase/synthase (ACDS) complex, a multienzyme machine that is a major route towards methane production in methanogenic archaea, which generate an estimated 10^9 tons of methane per year[12,13].

V.D. Initial structural and mechanistic studies of CODH

Although it had been well established that CODH harbored a Ni-Fe-S active site[14], it was not until the initial X-ray crystal structures of CODH from *R. rubrum* (*Rr*CODH) (Figure V.2)[6], *C. hydrogenoformans* (*Ch*CODH)[5], and *M. thermoacetica* CODH/ACS (*Mt*CODH/ACS) (Figure V.3.A)[4] that the arrangement and geometry of the metals were determined. Early spectroscopic studies had suggested that the C-cluster was composed of a [4Fe-4S] cubane with a unique Ni site nearby[14,15]. However, all of these CODH structures revealed an unprecedented metallocluster that can be described as a distorted [Ni-3Fe-4S] cubane coordinated to a unique Fe site, also called ferrous component II (FCII).

Despite exhibiting the same arrangement of metals in the C-cluster, these initial structures possessed key differences that hindered full mechanistic understanding. Perhaps most importantly, the *Ch*CODH C-cluster contained an additional sulfide ligand in a position bridging Ni of the distorted cubane and the unique Fe (Figure V.4.A), a feature absent in the *Rr*CODH and *Mt*CODH/ACS structures (Figure V.4.B). This inconsistency led to controversy over the correct composition of the cluster, and further experiments were conducted which argued either for[16] or against[17,18] a catalytic role for the sulfide bridge.

Related to the issue of the sulfide bridge is the crucial question of where substrates bind to the C-cluster for the interconversion of CO and CO_2. In the direction of CO oxidation, CO and H_2O must bind the C-cluster, H_2O is deprotonated, and CO_2 is formed, generating two protons and two electrons that reduce the cluster. Although the pathway by which protons exchange with the bulk solvent is not firmly established, a

network of histidine residues that link the buried C-cluster with the solvent exterior has been suggested as a possible route[6,19]. Electrons are passed from the C-cluster to the surface of the protein through additional [4Fe-4S] clusters that form a wire seen in all CODH structures (Figure V.2). Ferredoxin, pyruvate:ferredoxin oxidoreductase (PFOR), and hydrogenase have been proposed as ultimate electron acceptors[20,21]; intermolecular electron transfer has been suggested to be rate-limiting[22], with specific activity depending upon the electron acceptor employed[23,24]. In the direction of CO_2 reduction, CO_2 must bind the two-electron reduced C-cluster[25,26], and with the addition of two protons, H_2O and CO are formed. CO generated at the C-cluster is a gaseous intermediate that travels approximately 70 Å through an extraordinary hydrophobic tunnel within the enzyme complex[4,27-30] to the ACS A-cluster (Figure V.3.B).

Although there remained uncertainties surrounding the transfer of electrons and protons to the buried C-cluster, the most attention in the literature has been paid to binding sites for CO and H_2O on the metals of the C-cluster. Studies conducted before the *Ch*CODH and *Rr*CODH structure determinations had indicated that Ni and Fe are involved in binding the substrate CO and water molecules, respectively[15,31]. In the *Ch*CODH structure, however, the sulfide bridge fills coordination sites to complete square planar geometry around Ni and distorted tetrahedral geometry around the unique Fe, making the substrate binding locations unclear. Although *Rr*CODH and two independent *Mt*CODH/ACS structures contain empty coordination sites in place of the sulfide bridge, electron density for an unassigned ligand apical to Ni was present in the *Rr*CODH structure and one *Mt*CODH/ACS structure (Figure V.4.B)[6,27]. Without a clear

identification of substrate binding sites, the mechanism of the CODH C-cluster remained enigmatic.

V.E. Substrate- and inhibitor-bound C-cluster structures identify the active site

Over the past several years, many crystal structures have been solved that depict substrates bound to the C-cluster. Structures of *Ch*CODH[32] and *Mt*CODH/ACS[33] show the substrate water molecule bound to the unique Fe site in an identical fashion, completing a distorted tetrahedral geometry (Figure V.5.A). These observations are consistent with previous studies[31] which also indicated that water binds the unique Fe. None of these structures contain the sulfide bridge, as the water molecule occupies the sulfide coordination site on Fe.

A structure of the CODH component of the ACDS complex from *Methanosarcina barkeri* (*Mb*CODH) depicts a putative CO molecule bound to Ni of the C-cluster in a position adjacent the water molecule, which remains bound to the unique Fe (Figure V.5.B)[34]. With both substrates bound to the cluster, it was hypothesized that the low pH of the crystallization condition (4.6) prevented turnover by disfavoring deprotonation of water to the active hydroxide nucleophile, allowing the capture of the C-cluster state immediately before catalysis. CO is bound to Ni in an unexpected bent conformation, with a Ni-C-O bond angle of 103°, completing a distorted tetrahedral geometry. Interestingly, a structure of cyanide, a CODH inhibitor, bound to the *Mt*CODH/ACS C-cluster illustrates analogous bent geometry[33] (Figure V.5.B), with a Ni-C-N bond angle of 114°. The substrate water molecule in this cyanide-bound structure also remains bound to the unique Fe. A conserved isoleucine residue is seen in both CO- and CN-bound

structures to sterically block linear binding of these ligands to Ni (Figure V.5.B). Such bent coordination is not likely to be stable; indeed, infrared spectroscopy had suggested that there is no single, stable site for CO binding to the C-cluster[35]. Thus, it is possible that binding of CO to give bent geometry is mechanistically important, as the enhanced stability of a linear binding mode may impede turnover. In such a model, isoleucine would contribute to ground state destabilization, lowering the activation barrier to catalysis by preventing linear substrate binding.

The structures of CO and cyanide bound to the C-cluster place the carbon atom and water molecule too far apart for catalysis; a shift in coordination must occur during the reaction. The crystal structure of the product CO_2 bound to the *Ch*CODH C-cluster[32] provides a unique perspective on how such a shift may occur. Here, the CO_2 carbon is bound to Ni, while one CO_2 oxygen is bound to the unique Fe. A superposition of CO_2-bound *Ch*CODH with CO-bound *Mb*CODH and cyanide-bound *Mt*CODH/ACS structures exhibits a nearly identical position of all substrate atoms except for the carbon atom (Figure V.5.C). While the oxygen/nitrogen atoms remain stationary across the structures, the carbon atom has shifted closer to the water molecule in the CO_2-bound structure. This "carbon shift" has been proposed[33] to alter the Ni coordination geometry from distorted tetrahedral in the CO-bound and CN-bound forms to square planar in the CO_2-bound form.

Taken together, these structures reveal the active site on the C-cluster of CODH. In the direction of CO oxidation, crystal structures from *Ch*CODH, *Mb*CODH, and *Mt*CODH/ACS depict water bound to the unique Fe, while structures of *Mb*CODH and *Mt*CODH/ACS show CO or an inhibitor CN^- bound to Ni, respectively. Meanwhile, one

structure of *Ch*CODH shows CO_2, the substrate in the direction of CO_2 reduction, bound to the cluster. In all structures, regardless of the organism, no sulfide bridge is present when a substrate molecule is bound. Consistent with all of the structures of substrate- and inhibitor-bound C-clusters, a catalytic mechanism is proposed (Figure V.6).

V.F. The complexity of cyanide inhibition

A distinguishing feature of all Ni-containing CODHs is potent inhibition by cyanide[21,23,36-40]. Although direct binding to the C-cluster has been implicated as the root of cyanide inhibition, years of study have not led to a clear inhibitory mechanism. On the one hand, cyanide has been described as a competitive inhibitor that binds Ni: Ni-deficient *Rr*CODH does not bind cyanide[37,38], and X-ray absorption spectroscopy (XAS) indicated that cyanide shares with CO a binding site on Ni with a Ni-C distance of 1.81-1.84 Å[40]. In contrast, other studies have suggested that cyanide binds Fe: electron-nuclear double resonance (ENDOR) spectroscopy indicated that cyanide displaces the Fe-bound water molecule[31], and Mössbauer spectroscopy showed a change in quadrupole splitting (ΔE_Q) of the unique Fe signal upon cyanide treatment[15]. It has also been suggested that cyanide may bind at multiple sites[22].

However, these data appear to be in conflict only if it is assumed that cyanide must adopt a single binding mode. Indeed, two crystal structures of cyanide bound to the C-cluster show that the same inhibitor can actually adopt multiple binding modes. As mentioned above, the cyanide-bound *Mt*CODH/ACS structure[33] shows cyanide bound to Ni in a bent conformation, in an analogous fashion as CO in the *Mb*CODH structure, with water still bound to Fe (Figure V.5.B). The cyanide carbon completes distorted

tetrahedral geometry around Ni and mimics how CO binds to the active C-cluster. On the other hand, a cyanide-bound *Ch*CODH structure[41] shows CN^- bound to Ni with effectively linear geometry (Ni-C-N bond angle of 175°), conferring square planar geometry around Ni. Notably, the substrate water molecule is absent in this structure, which represents an inhibited form where neither substrate is bound to the cluster. A superposition of the two cyanide-bound C-clusters is shown in Figure V.5.D.

These dissimilar crystal structures depicting cyanide binding mirror the seemingly contradictory spectroscopic results described above. It has recently been suggested that a rapid, reversible cyanide binding step is followed by a slow rearrangement step to achieve tighter binding[22,41], explaining how cyanide could be both a rapid, reversible inhibitor under some conditions and a slow-binding inhibitor under others[38]. Here, it may be reasoned that the *Mt*CODH/ACS cyanide structure illustrates an "easily reversible" cyanide binding mode, where cyanide binds Ni in the same bent manner as CO. This bent, Ni-bound cyanide structure is consistent with studies indicating that cyanide is a competitive inhibitor of CO and, like CO, binds Ni. In a subsequent, slow rearrangement step, the substrate water molecule bound to the unique Fe may then be displaced, freeing space to allow cyanide to relax into a more favored linear binding mode, represented by the *Ch*CODH cyanide structure. One would expect this structure to represent cyanide in a "tight binding" mode. The displacement of the Fe-bound water molecule upon cyanide treatment seen in the *Ch*CODH structure is consistent with the ENDOR experiments[31] and further clarifies that water displacement is not a result of direct binding of cyanide to Fe, an assumption made in the ENDOR study. Water displacement and linear cyanide binding to Ni can also explain why Mössbauer spectroscopy showed a change in ΔE_Q of

the unique Fe signal upon cyanide treatment[15]. Therefore, while neither cyanide-bound C-cluster structure can alone reconcile all of the spectroscopic studies, both structures together can provide an explanation for the seemingly inconsistent data on CODH cyanide inhibition, a 20-year mystery in the field.

In asking why the *Ch*CODH and *Mt*CODH/ACS structures revealed different cyanide binding modes, the most likely answer lies in the dissimilar cyanide crystal soaking protocols. *Ch*CODH crystals were soaked for 30 min in 70 mM KCN[41], while *Mt*CODH/ACS crystals were soaked for 1 hour in just 100 μM KCN[33]. Because *Ch*CODH crystals were exposed to cyanide concentrations orders of magnitude higher than *Mt*CODH/ACS crystals, it is possible that the equilibrium was shifted toward displacement of Fe-bound water and linear cyanide binding. Regardless, these structures have offered a clearer picture of the possible mechanism for cyanide inhibition (Figure V.6).

V.G. Conclusions

The truly distinctive $NiFe_4S_4$ CODH C-cluster has inspired decades of biochemical research; however, the literature has been fraught with controversy and contradictions regarding the C-cluster's structure, mechanism, and mode of inhibition. Crystal structures of CODHs from several organisms have brought the C-cluster into focus and allowed us to rationalize the abundance of seemingly inconsistent biochemical data. From these structural studies, a unified view of the C-cluster has emerged, presenting key insights into the function of this remarkable and environmentally important metallocluster.

V.H. Figures

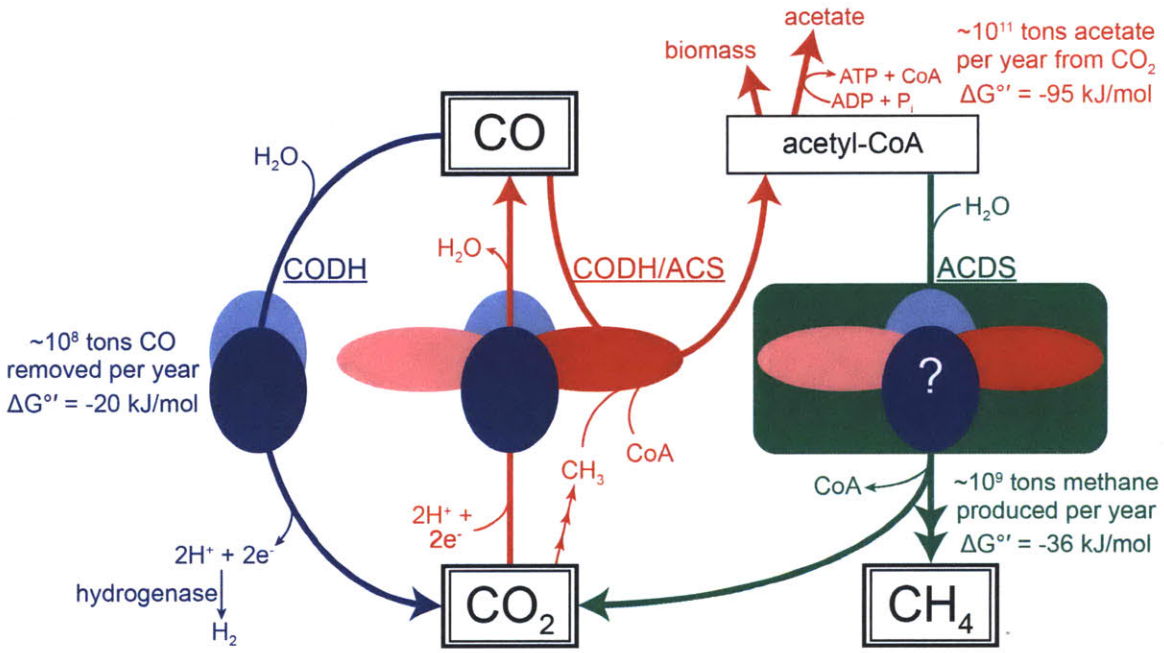


Figure V.1. Schematic of NiFe-containing CODH in the microbial carbon cycle and its contributions to CO, CO_2, and methane production and consumption. CODH is shown as a dimer in light and dark blue ovals, ACS as light and dark red ovals, and the methanogenic ACDS complex as a green rectangle containing both CODH and ACS components in an unknown arrangement. Blue arrows indicate monofunctional CODH activity, red arrows indicate CODH/ACS bifunctional activity in the acetogenic pathway, and green arrows indicate methanogic CODH activity related to the ACDS complex.

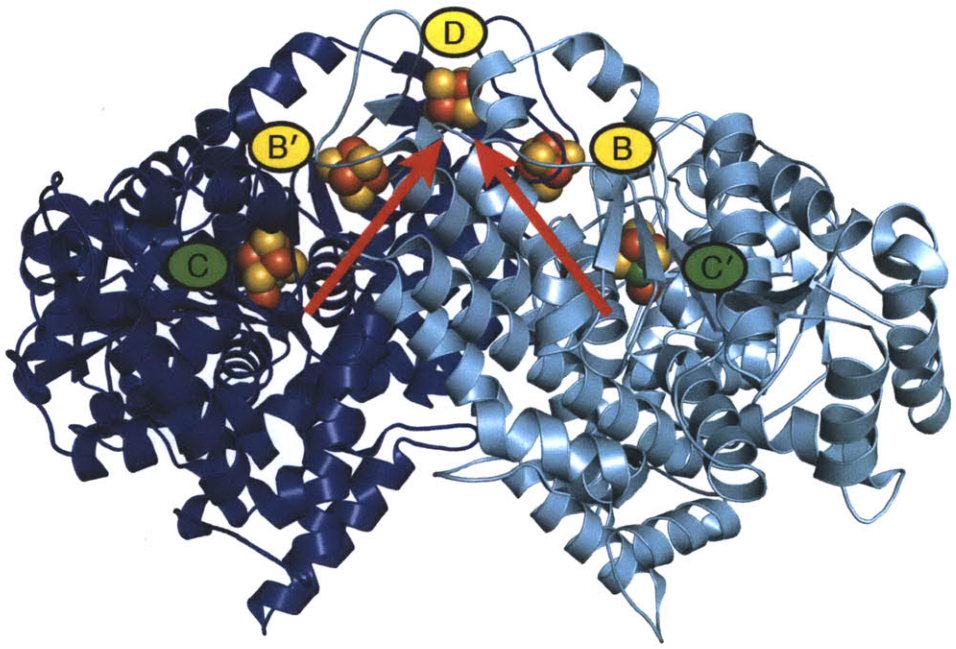


Figure V.2. Monofunctional CODH and its metalloclusters. The *Rr*CODH homodimer[6] (blue and cyan ribbons, PDB ID: 1JQK). Metalloclusters in spheres: Ni in green, Fe in orange, and S in yellow. Active site C-clusters are labeled in green, while B-clusters and D-cluster are labeled in yellow. Red arrows in **A** indicate the direction of electron transfer during CO oxidation. All figures prepared in PyMOL[42].

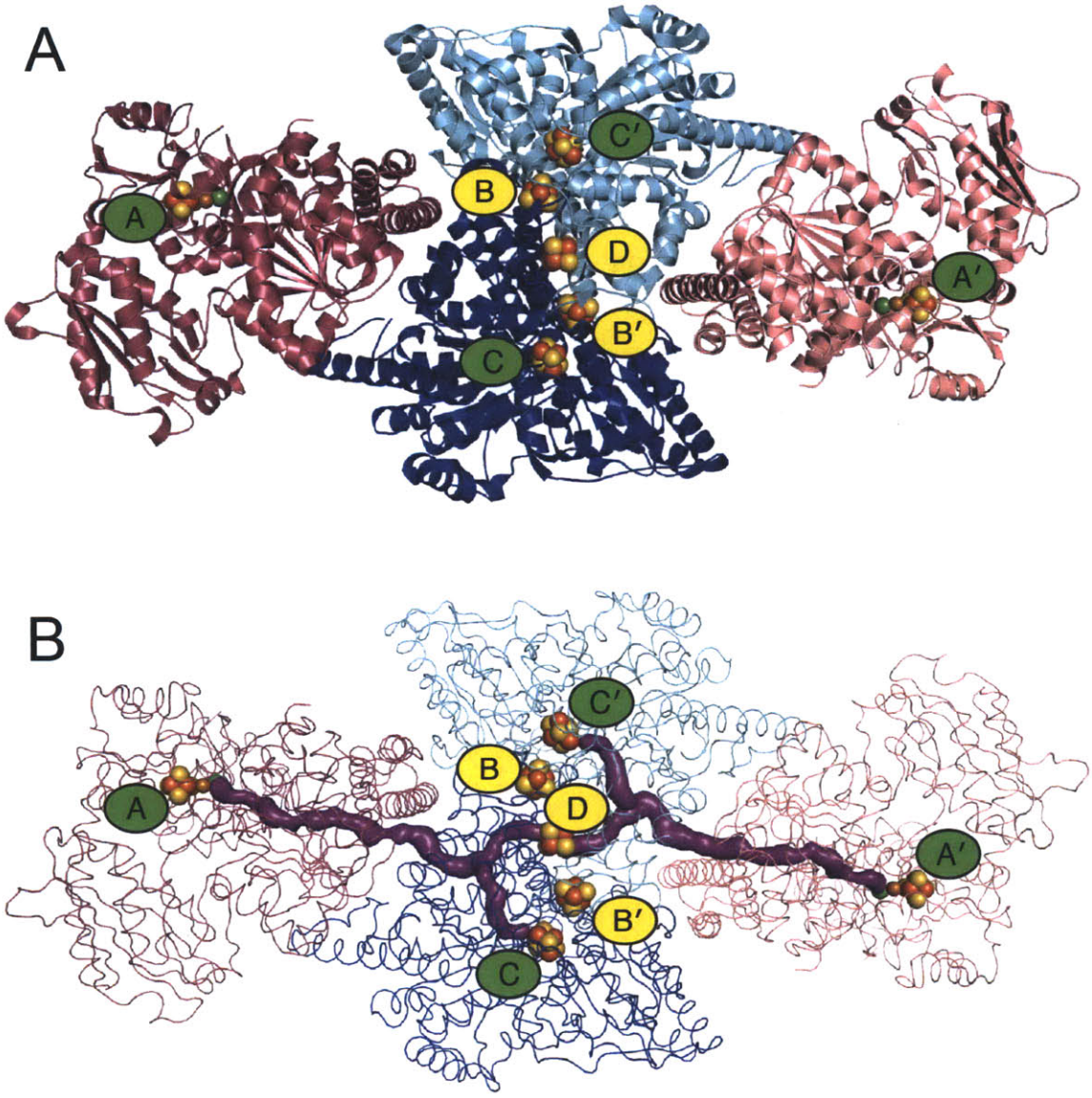


Figure V.3. The acetogenic bifunctional CODH/ACS complex. A, The *Mt*CODH/ACS complex[4] (CODH: blue and cyan ribbons, ACS: dark and light link ribbons, PDB ID: 1MJG). **B**, *Mt*CODH/ACS, emphasizing the extended tunnel within the complex (violet surface, prepared in using the CAVER[43] plugin for PyMOL[42]), identified from Xe-pressurized crystals[28] (PDB ID: 2Z8Y). From A-cluster to A-cluster, the tunnel is 138 Å long. **A** and **B** share the same colors and orientation, except thinner ribbons are depicted in **B**. Active site C-clusters and A-clusters are labeled in green, while B-clusters and the D-cluster are labeled in yellow. All metalloclusters shown in spheres: Ni in green, Cu in brown, Fe in orange, and S in yellow. Note that several *Mt*CODH/ACS structures including those depicted in this figure contain a Cu atom in the A-cluster[4,28,33] which is replaced by Ni in active ACS[44-46].

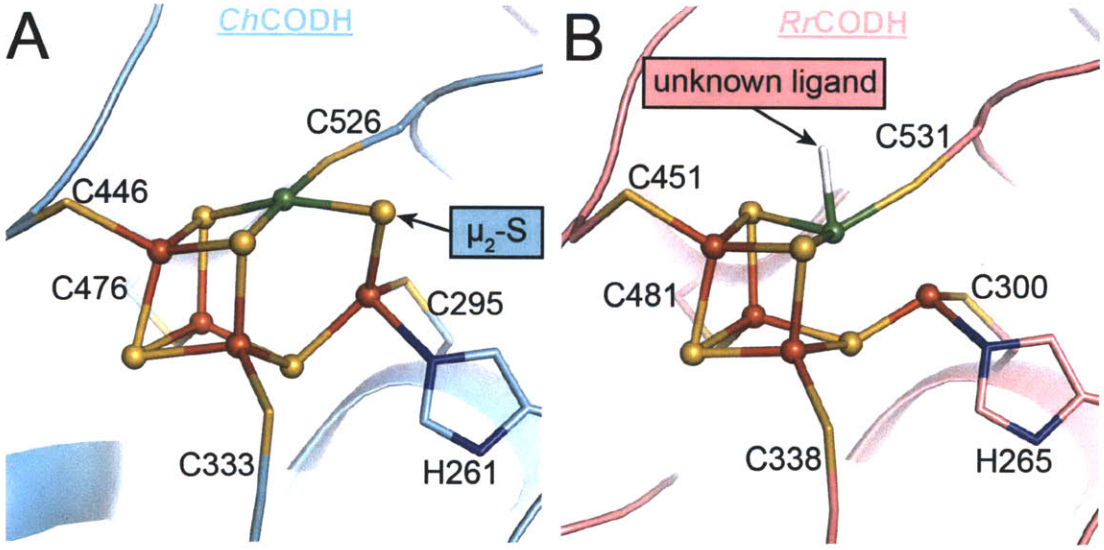


Figure V.4. Initial structures of CODH C-clusters. A, The C-cluster of *Ch*CODH[5] (cyan ribbons, PDB ID: 1SU8), which contains a sulfido ligand (labeled μ_2-S) that bridges Ni and the unique Fe. **B**, The C-cluster of *Mt*CODH/ACS[4] (pink ribbons, PDB ID: 1MGJ), which does not contain the bridging sulfide. C-clusters in ball-and-stick: Ni in green, Fe in orange, and S in yellow. Protein ligands in sticks: N in blue, S in yellow, and C following protein ribbon coloring. Residue numbers follow the respective protein sequence.

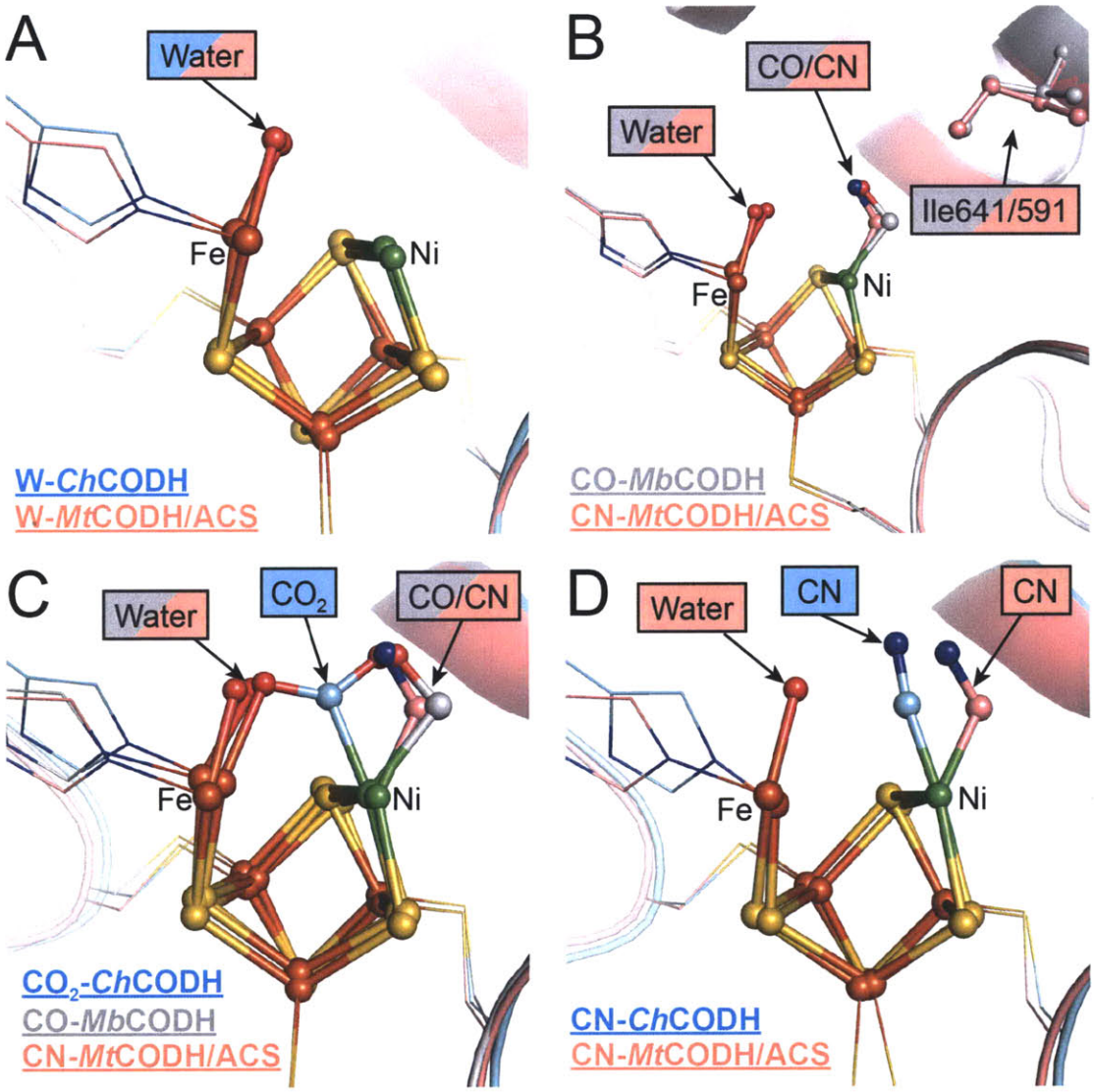


Figure V.5. Substrate-bound and inhibitor-bound C-cluster structures. A, Superposition of water-bound C-clusters from *Ch*CODH[32] (PDB ID: 3B51) and *Mt*CODH/ACS[33] (PDB ID: 3I01). **B**, Superposition of CO/water-bound and cyanide/water-bound C-clusters *Mb*CODH[34] (PDB ID: 3CF4) and *Mt*CODH/ACS[33] (PDB ID: 3I04), respectively. A conserved isoleucine residue (Ile641 in *Mb*CODH and Ile591 in *Mt*CODH/ACS) that is believed to sterically impede linear CO/CN-binding in shown in ball-and-stick. **C**, Superposition of the structures in **B** plus the CO_2-bound CO_2-bound C-cluster from *Ch*CODH[32] (PDB ID: 3B52). **D**, Superposition of the cyanide-bound C-clusters from *Mt*CODH/ACS[33] (PDB ID: 3I04) and *Ch*CODH[41] (PDB ID: 3I39). Protein chains in ribbons: *Ch*CODH in cyan, *Mt*CODH/ACS in pink, and *Mb*CODH in grey. C-clusters in ball-and-stick and protein ligands to the cluster in sticks: Ni in green, Fe in orange, S in yellow, Ni in blue, O in red, and C following protein ribbon coloring. The Ni and unique Fe sites are labeled "Ni" and "Fe", respectively. For clarity, not all protein ligands to the cluster are shown.

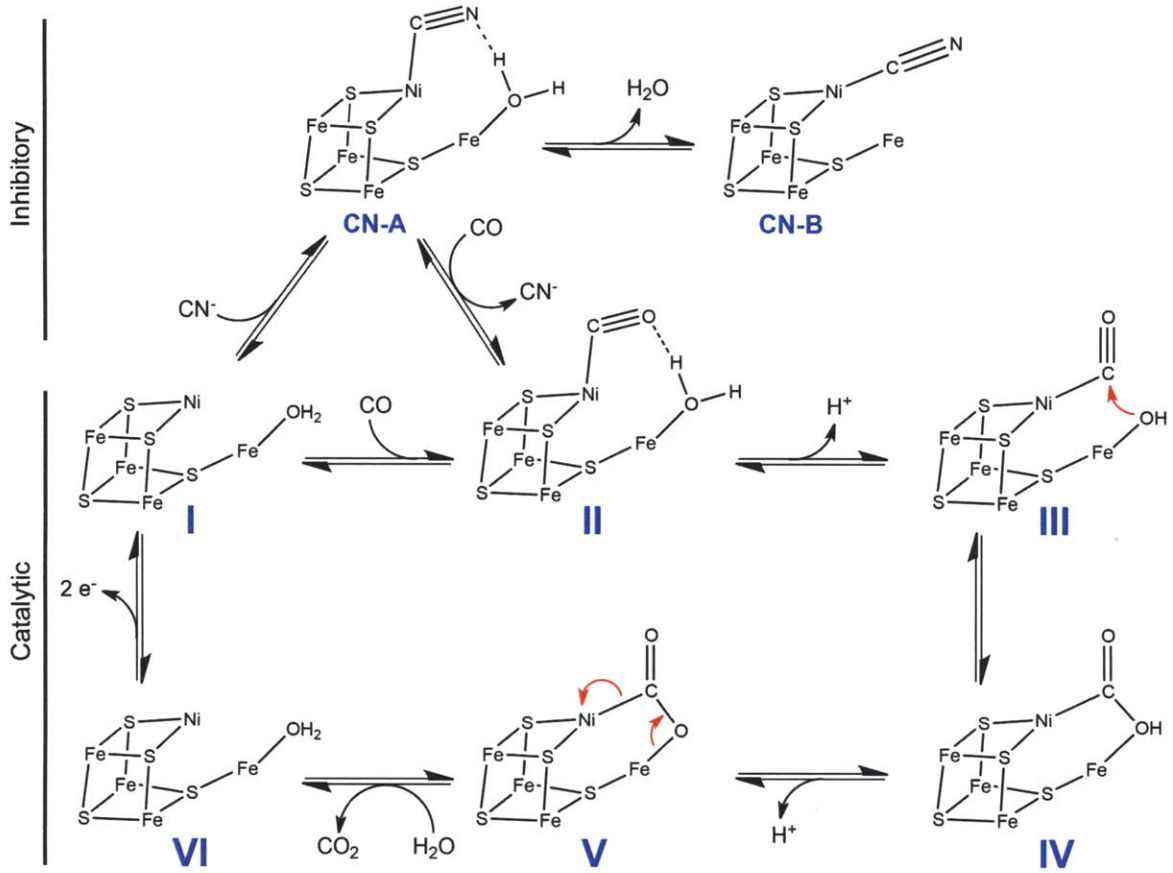


Figure V.6. Proposed catalytic and inhibitory mechanisms of the CODH C-cluster, consistent with crystal structures containing bound substrate and inhibitor molecules. To begin the reaction, a water molecule binds the unique Fe of the C-cluster (state I), as seen in *Ch*CODH[32] and *Mt*CODH/ACS[33] structures. CO binds Ni with bent geometry (state II), as seen in CO-bound *Mb*CODH[34] and CN-bound *Mt*CODH/ACS[33]. Upon deprotonation of water, a "carbon shift" may occur to position the carbon in proximity to the resulting hydroxide (state III), promoting nucleophilic attack to yield a Ni-COOH intermediate (state IV). A second deprotonation gives a Ni-COO^- species (state V). The crystal structure of CO_2-bound *Ch*CODH[32] resembles states IV and V, which differ only in protonation state. CO_2 may then be released, with the C-cluster becoming reduced by two electrons (state VI). Electrons are then passed to external electron acceptor proteins through a wire of [4Fe-4S] clusters within the CODH dimer, as seen in Figure V.2, completing the catalytic cycle. Protons generated during CO oxidation may access the external solvent through a histidine-lined channel. In the proposed inhibitory mechanism, cyanide competes with CO for binding to Ni of the C-cluster with bent geometry, as seen in the structure of cyanide-bound *Mt*CODH/ACS[33] (state CN-A). Following displacement of Fe-bound water, cyanide may relax to a tighter linear binding mode, as seen in the structure of cyanide-bound *Ch*CODH[41] (state CN-B).

V.I. References

1 Fontecilla-Camps, J. C., Volbeda, A., Cavazza, C. & Nicolet, Y. (2007) Structure/Function Relationships of [NiFe]- and [FeFe]-Hydrogenases. *Chem. Rev.* **107**, 4273-4303.

2 Vignais, P. M. & Billoud, B. (2007) Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. *Chem. Rev.* **107**, 4206-4272.

3 Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M. & Fontecilla-Camps, J. C. (1995) Crystal Structure of the Nickel-Iron Hydrogenase from *Desulfovibrio gigas*. *Nature.* **373**, 580-587.

4 Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002) A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Science.* **298**, 567-572.

5 Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, O. (2001) Crystal Structure of a Carbon Monoxide Dehydrogenase Reveals a [Ni-4Fe-5S] Cluster. *Science.* **293**, 1281-1285.

6 Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. & Ludden, P. W. (2001) Life on Carbon Monoxide: X-ray Structure of *Rhodospirillum rubrum* Ni-Fe-S Carbon Monoxide Dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11973-11978.

7 Svetlichny, V. A., Sokolova, T. G., Gerhardt, M., Ringpfeil, M., Kostrikina, N. A. & Zavarzin, G. A. (1991) *Carboxydothermus hydrogenoformans* gen. nov., sp. nov., a CO-Utilizing Thermophilic Anaerobic Bacterium from Hydrothermal Environments of Kunashir Island. *Syst. Appl. Microbiol.* **14**, 254-260.

8 Uffen, R. L. (1976) Anaerobic Growth of a *Rhodopseudomonas* Species in the Dark with Carbon Monoxide as Sole Carbon and Energy Substrate. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3298-3302.

9 Bartholomew, G. W. & Alexander, M. (1979) Microbial Metabolism of Carbon Monoxide in Culture and in Soil. *Appl. Environ. Microbiol.* **37**, 932-937.

10 Ragsdale, S. W. & Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 Fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898.

11 Drake, H. L., Daniel, S. L., Matthies, C. & Küsel, K. (1994) "Acetogenesis, Acetogenic Bacteria, and the Acetyl-CoA Pathway: Past and Current Perspectives" in *Acetogenesis*, Ed. H.L. Drake. Chapman and Hall: New York, 3-60.

12 Grahame, D. A. (1991) Catalysis of Acetyl-CoA Cleavage and Tetrahydrosarcinapterin Methylation by a Carbon Monoxide Dehydrogenase-Corrinoid Enzyme Complex. *J. Biol. Chem.* **266**, 22227-22233.

13 Thauer, R. K. (1998) Biochemistry of Methanogenesis: A Tribute to Marjory Stephenson. *Microbiology.* **144**, 2377-2406.

14 Ragsdale, S. W. & Kumar, M. (1996) Nickel-Containing Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Chem. Rev.* **96**, 2515-2539.

15 Hu, Z., Spangler, H. J., Andersen, M. E., Xia, J., Ludden, P. W., Lindahl, P. A. & Münck, E. (1996) Nature of the C-Cluster in Ni-Containing Carbon Monoxide Dehydrogenases. *J. Am. Chem. Soc.* **118**, 830-845.

16 Dobbek, H., Svetlitchnyi, V., Liss, J. & Meyer, O. (2004) Carbon Monoxide Induced Decomposition of the Active Site [Ni-4Fe-5S] Cluster of CO Dehydrogenase. *J. Am. Chem. Soc.* **126**, 5382-5387.

17 Feng, J. & Lindahl, P. A. (2004) Effect of Sodium Sulfide on Ni-Containing Carbon Monoxide Dehydrogenases. *J. Am. Chem. Soc.* **126**, 9094-9100.

18 Gu, W., Seravalli, J., Ragsdale, S. W. & Cramer, S. P. (2004) CO-Induced Structural Rearragement of the C Cluster in *Carboxydothermus hydrogenoformans* CO Dehydrogenase-Evidence from Ni K-Edge X-ray Absorption Spectroscopy. *Biochemistry.* **43**, 9029-2035.

19 Kim, E. J., Feng, J., Bramlett, M. R. & Lindahl, P. A. (2004) Evidence for a Proton Transfer Network and a Required Persulfide-Bond-Forming Cysteine Residue in Ni-Containing Carbon Monoxide Dehydrogenases. *Biochemistry.* **43**, 5728-5734.

20 Menon, S. & Ragsdale, S. W. (1996) Unleashing Hydrogenase Activity in Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase and Pyruvate:Ferredoxin Oxidoreductase. *Biochemistry.* **35**, 15814-15821.

21 Ragsdale, S. W., Clark, J. E., Ljungdahl, L. G., Lundie, L. L. & Drake, H. L. (1983) Properties of Purified Carbon Monoxide Dehydrogenase from *Clostridium thermoaceticum*, a Nickel, Iron-Sulfur Protein. *J. Biol. Chem.* **258**, 2364-2369.

22 Seravalli, J. & Ragsdale, S. W. (2008) ^{13}C NMR Characterization of an Exchange Reaction between CO and CO_2 Catalyzed by Carbon Monoxide Dehydrogenase. *Biochemistry.* **47**, 6770-6781.

23 Andersen, M. E., DeRose, V. J., Hoffman, B. M. & Lindahl, P. A. (1993) Identification of a Cyanide Binding Site in CO Dehydrogenase from *Clostridium thermoaceticum* Using EPR and ENDOR Spectroscopies. *J. Am. Chem. Soc.* **115**, 12204-12205.

24 Svetlitchnyi, V., Peschel, C., Acker, G. & Meyer, O. (2001) Two Membrane-Associated NiFeS-Carbon Monoxide Dehydrogenases from the Anaerobic Carbon-Monoxide-Utilizing Eubacterium *Carboxydothermus hydrogenoformans*. *J. Bacteriol.* **183**, 5134-5144.

25 Andersen, M. E. & Lindahl, P. A. (1996) Spectroscopic States of the CO Oxidation/CO_2 Reduction Active Site of Carbon Monoxide Dehydrogenase and Mechanistic Implications. *Biochemistry.* **35**, 8371-8390.

26 Seravalli, J., Kumar, M., Lu, W.-P. & Ragsdale, S. W. (1997) Mechanism of Carbon Monoxide Oxidation by the Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase from *Clostridium thermoaceticum*: Kinetic Characterization of the Intermediates. *Biochemistry.* **36**, 11241-11251.

27 Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003) Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] Clusters in Closed and Open Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. *Nat. Struct. Biol.* **10**, 271-279.

28 Doukov, T. I., Blasiak, L. C., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2008) Xenon in and at the End of the Tunnel of Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **47**, 3474-3483, doi:10.1021/bi702386t.

29 Maynard, E. L. & Lindahl, P. A. (1999) Evidence of a Molecular Tunnel Connecting the Active Sites for CO_2 Reduction and Acetyl-CoA Synthesis in Acetyl-CoA Synthase from *Clostridium thermoaceticum*. *J. Am. Chem. Soc.* **121**, 9221-9222.

30 Seravalli, J. & Ragsdale, S. W. (2000) Channeling of Carbon Monoxide during Anaerobic Carbon Dioxide Formation. *Biochemistry.* **39**, 1274-1277.

31 DeRose, V. J., Telser, J., Andersen, M. E., Lindahl, P. A. & Hoffman, B. M. (1998) A Multinuclear ENDOR Study of the C-Cluster in CO Dehydrogenase from *Clostridium thermoaceticum*: Evidence for H_xO and Histidine Coordination to the [Fe_4S_4] Center. *J. Am. Chem. Soc.* **120**, 8767-8776.

32 Jeoung, J.-H. & Dobbek, H. (2007) Carbon Dioxide Activation at the Ni,Fe-Cluster of Anaerobic Carbon Monoxide Dehydrogease. *Science.* **318**, 1461-1464.

33 Kung, Y., Doukov, T. I., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2009) Crystallographic Snapshots of Cyanide- and Water-Bound C-Clusters from Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **48**, 7432-7440.

34 Gong, W., Hao, B., Wei, Z., Ferguson, D. J. J., Tallant, T., Krzycki, J. A. & Chan, M. K. (2008) Structure of the $\alpha_2\varepsilon_2$ Ni-Dependent CO Dehydrogenase Component

of the *Methanosarcina barkeri* Acetyl-CoA Decarbonylase/Synthase Complex. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9558-9563.

35 Chen, J., Huang, S., Seravalli, J., Gutzma, H. J., Swartz, D. J., Ragsdale, S. W. & Bagley, K. A. (2003) Infrared Studies of Carbon Monoxide Binding to Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase from *Moorella thermoacetica*. *Biochemistry.* **42**, 14822-14830.

36 Andersen, M. E. & Lindahl, P. A. (1994) Organization of Clusters and Internal Electron Pathways in CO Dehydrogenase from *Clostridium thermoaceticum*: Relevance to the Mechanism of Catalysis and Cyanide Inhibition. *Biochemistry.* **33**, 8702-8711.

37 Ensign, S. A., Bonam, D. & Ludden, P. W. (1989) Nickel is Required for the Transfer of Electrons from Carbon Monoxide to the Iron-Sulfur Center(s) of Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum*. *Biochemistry.* **28**, 4968-4973.

38 Ensign, S. A., Hyman, M. R. & Ludden, P. W. (1989) Nickel-Specific, Slow-Binding Inhibition of Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum* by Cyanide. *Biochemistry.* **28**, 4973-4979.

39 Grahame, D. A. & Stadtman, T. C. (1987) Carbon Monoxide Dehydrogeanse from *Methanosarcina barkeri*: Disaggregation, Purification, and Physicochemical Properties of the Enzyme. *J. Biol. Chem.* **262**, 3706-3712.

40 Ha, S.-W., Korbas, K., Klepsch, M., Meyer-Klaucke, W., Meyer, O. & Svetlitchnyi, V. (2007) Interaction of Potassium Cyanide with the [Ni-4Fe-5S] Active Site Cluster of CO Dehydrogenase from *Carboxydothermus hydrogenoformans*. *J. Biol. Chem.* **282**, 10639-10646.

41 Jeoung, J.-H. & Dobbek, H. (2009) Structural Basis of Cyanide Inhibition of Ni, Fe-Containing Carbon Monoxide Dehydrogenase. *J. Am. Chem. Soc.* **131**, 9922-9923.

42 DeLano, W. L. (2010) The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. (http://www.pymol.org).

43 Beneš, P., Chovancová, E., Kozlíková, B., Pavelka, A., Strnad, O., Brezovský, J., Šustr, V., Klvaňa, M., Szabó, T., Gora, A., Zamborský, M., Biedermannová, L., Medek, P., Damborský, J. & Sochor, J. (2010) CAVER 2.1. (http://ww.caver.cz).

44 Hegg, E. L. (2004) Unraveling the Structure and Mechanism of Acetyl-Coenzyme A Synthase. *Acc. Chem. Res.* **37**, 775-783.

45 Seravalli, J., Xiao, Y., Gu, W., Cramer, S. P., Antholine, W. E., Krymov, V., Gerfen, G. J. & Ragsdale, S. W. (2004) Evidence that NiNi Acetyl-CoA Synthase Is Active and That the CuNi Enzyme Is Not. *Biochemistry.* **43**, 3944-3955.

46 Svetlitchnyi, V., Dobbek, H., Meyer-Klaucke, W., Meins, T., Thiele, B., Römer, P., Huber, R. & Meyer, O. (2004) A Functional Ni-Ni-[4Fe-4S] Cluster in the Monomeric Acetyl-CoA Synthase from *Carboxydothermus hydrogenoformans*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 446-451.

Chapter VI: Beyond the Wood-Ljungdahl carbon fixation pathway

VI.A. Summary

Earlier chapters of this thesis describe the structural characterization of metalloenzyme complexes of the Wood-Ljungdahl carbon fixation pathway in *Moorella thermoacetica*. As discussed, the structural conclusions gained from these studies have made a significant impact on our understanding of the Wood-Ljungdahl pathway, but they can offer additional insight into other systems as well. For example, B_{12}-dependent methyl transfer also occurs outside of the Wood-Ljungdahl pathway. Methionine synthase (MetH) from *Escherichia coli* is the most studied enzyme of this class, and as we shall see, the structure of the B_{12}-dependent methyltransferase complex explored in Chapters II and III provides new structural details that can inform MetH and other B_{12}-mediated methyl transfer systems. On the other hand, the C-cluster of carbon monoxide dehydrogenase (CODH), the focus of Chapters IV and V, is unprecedented and does not exist in any other known enzyme. However, interest in the environmental impact of CODHs in nature has inspired the ongoing development of biomimetic model complexes that aim to imitate C-cluster structure and catalysis. Therefore, our enhanced structural insight into the CODH C-cluster may also have outside implications for the development of biomimetic compounds, and thus we shall survey the current state of C-cluster model complex construction. Overall, the substantial structural insights into the metalloenzymes studied in this work have a considerable impact on research into both biological and synthetic systems outside of the Wood-Ljungdahl carbon fixation pathway.

VI.B. Structural similarities and divergences in B_{12}-dependent methyltransferases

Introduction

Chapters II and III of this thesis represent the first structural descriptions of a B_{12}-dependent methyl transfer complex where the proteinaceous components responsible for B_{12} binding, activation, protection, and methyl transfer are assembled together. The crystal structure of the *Moorella thermoacetica* corrinoid iron-sulfur protein (CFeSP) in complex with methyltetrahydrofolate:CFeSP methyltransferase (MeTr) discussed in Chapter II shows how protein domains responsible for disparate but coordinated functions are organized in three dimensions. These functions necessitate significant domain rearrangements over the course of the full reaction cycle. Because these structural data now identify the relative orientation of domains, the required domain rearrangements can now be recognized and rationalized. For example, by modeling the substrate methyltetrahydrofolate (CH_3-H_4folate) into this structure of the complex, the position of the B_{12} cofactor in relation to the site of catalysis can now be considered. Additionally, the previous structure of CFeSP alone[1] may also be compared to the structure of the CFeSP/MeTr complex to understand the movement of B_{12} from its "resting" and protected position to a catalytic position. In this way, the CFeSP/MeTr structure sets the stage for understanding the molecular motions required for the many steps along the B_{12}-dependent methyl transfer reaction cycle.

In vitro activity of CFeSP/MeTr in solution studies has been established[2-7], exploiting the B_{12} cofactor for as a handle for UV-Vis spectroscopy. However the domain motions that facilitate methyl transfer activity in solution have remained invisible. In Chapter II, UV-Vis spectra obtained from crystals of this complex were described and

appear nearly identical to the solution spectra (Figure II.19). These data not only support the physiological relevance of the CFeSP/MeTr assembly obtained through crystallography, they also further provide a glimpse into the domain movements that enable catalysis. In addition, anaerobic small-angle X-ray scattering (SAXS) experiments described in Chapter III probed the structure of the CFeSP/MeTr complex in solution. These data demonstrate that the complex observed in the crystal structure can also form in certain solution conditions, where CFeSP and the MeTr homodimer may either bind each other in 1:1 or 2:1 stoichiometry. Using SAXS, it was determined that CFeSP/MeTr assembles with 1:1 stoichiometry under the assay conditions used in prior activity studies[2-7], whereas the 2:1 complex observed in the crystal structure occurs under conditions that include the crystallization component polyethylene glycol (PEG). Although the precise physiological oligomeric composition of the complex remains speculative, it is clear from SAXS studies that CFeSP and a MeTr homodimer act as separate particles that can dock together in 1:1 or 2:1 ratios in solution.

Although these structural investigations of the ~220 kDa CFeSP/MeTr complex are the first such characterizations of fully assembled B_{12}-dependent methyltransferases, they offer considerable insight into previous structural work on these systems. While no comparable structure exists of the 136 kDa *Escherichia coli* B_{12} enzyme methionine synthase (MetH), by far the most characterized B_{12}-dependent methyltransferase, many crystal structures have been solved of MetH fragments[8-15]. CFeSP/MeTr and MetH share several, but not all, protein components, and the detailed structural comparison between the CFeSP/MeTr complex and MetH fragments that follows gives further insight into the ways in which these two systems are at once similar as well as distinct. Afterward,

attention will be placed on a discussion of the overall size of these protein complexes and how the CFeSP/MeTr structure may suggest a possible rationale for the large amount of protein Nature has recruited to perform B_{12}-dependent methyl transfer.

B_{12} nomenclature and reactivity

Vitamin B_{12} (Figure VI.1) is often loosely used as an umbrella term to refer to the cobalamin-related cofactors that contain a cobalt atom coordinated equatorially by four nitrogen ligands of a tetrapyrrolic macrocycle called a corrin, which is connected via an aminopropanol linker to the unusual ribonucleoside base 5,6-dimethylbenzimidazole (DMB) that may fill an axial coordination site of the cobalt atom. Strictly speaking, however, vitamin B_{12} formally refers to the form of cobalamin called cyanocobalamin, where cyanide fills the remaining axial position on cobalt to complete octahedral geometry. Commonly described additional cobalamin derivatives include the following forms, which differ in the identity of the axial ligand that substitutes for cyanide: coenzyme B_{12} or adenosylcobalamin (AdoCbl) contains a 5′-deoxy-5′-adenosine axial ligand, methylcobalamin contains a methyl group bound to cobalt, and hydroxocobalamin contains a hydroxyl ligand. It should be noted that some B_{12} derivatives in biology use ribonucleoside bases that are not DMB, and such is the case for CFeSP, where 5-methoxybenzimidazole replaces DMB[16]. The role of this unusual base (if any) is unclear. Because DMB is not present, these B_{12} forms are not called cobalamins, and the more general term "cobamide" is used instead, with the CFeSP B_{12} derivative referred to as 5-methoxybenzimidazolyl cobamide.

The subclass of AdoCbl B_{12} enzymes invoke radical-based chemistry[17] and include methylmalonyl-CoA mutase, glutamate mutase, lysine 5,6-aminomutase, diol dehydratase, ethanolamine ammonia lyase, and class II ribonucleotide reductase. In all of these enzymes, the AdoCbl Co(III)-C bond is cleaved homolytically to generate cob(II)alamin and a 5′-deoxyadenosine radical. This radical species is identical to the radical formed in the *S*-adenosylmethionine (SAM, AdoMet) radical superfamily of enzymes[18-21] and is used to initiate radical-mediated transformations of the substrate.

A second group of B_{12} enzymes is the reductive dehalogenase subclass[17] present in anaerobic microorganisms that are capable of detoxifying halogenated aromatic or aliphatic compounds. Very little is known about these enzymes, which include 3-chloro-4-hydroxybenzoate dehalogenase and haloalkane dehalogenases. The overall mechanism of these B_{12} enzymes is poorly understood, and even the role of B_{12} in these systems is unknown. In one proposal, the halide is eliminated upon formation of an organometallic aryl- or alkyl-Co(III) species; alternatively, B_{12} has been proposed to serve as an electron donor during the reaction[17].

Methylcobalamin-containing enzymes perform methyl transfer reactions[17], where a methyl group is transferred from a methyl donor molecule to a methyl acceptor, and methylcobalamin is the intermediary methyl carrier. B_{12}-dependent methionine synthase (MetH) is the best characterized enzyme of this subclass and uses methyltetrahydrofolate (CH_3-H_4folate) as the methyl donor and homocysteine (Hcy) as the methyl acceptor. Several other B_{12} methyltransferases have been characterized as well, including additional enzymes that also use CH_3-H_4folate as the methyl donor, as with CFeSP, in addition to other enzymes that use methyl donors such as methyl-folate analogs

(including methyltetrahydromethanopterin), methylamines, methylthiols, methoxylated aromatics, methyl halides, and methanol[17]. Organisms that use these substrates are obligate anaerobes, including methanogenic and methylotrophic microbes that rely on their B_{12}-dependent methyltransferases to convert these compounds to usable carbon sources. Because CFeSP belongs to the methyltransferase subclass of B_{12} enzymes typified by MetH, we continue our discussion of the structural features of B_{12} methyltransferases in the sections that follow.

Overview of MetH and other B_{12}-dependent methyltransferase structures

Besides CFeSP/MeTr, nearly all crystal structures of enzymes related to B_{12}-dependent methyl transfer systems are structures of MetH fragments[8-15], with the exceptions being a structure of a complex of enzymes responsible for methyl transfer from methanol[22] and a handful of unpublished structures that have been deposited into the PDB but not described in the literature. Therefore, comparisons may now be made between the CFeSP/MeTr complex structure and other systems that may inform the MetH reaction cycle and the B_{12} methyl transfer process. Before delving into details and comparisons between these crystal structures, the following provides an overview of the structures that exist, tallying all of the structural components that are currently available.

In regards to the CFeSP/MeTr system, a structure is available of *Carboxydothermus hydrogenoformans* CFeSP alone, containing both small and large subunits[1]. The N-terminal Fe_4S_4 domain of the large subunit was disordered and left unmodeled; however, the location of the Fe_4S_4 cluster was identified by Fe anomalous

diffraction methods. Also, structures of *M. thermoacetica* MeTr alone are available, with and without CH_3-H_4folate bound[23,24], in addition to a MeTr N199A mutant.

MetH is a large, multicomponent enzyme containing four domains connected by linker regions, forming a primary sequence topology that has been described as "beads on a string"[17]. Each domain binds a different substrate or cofactor species (from N- to C-terminus): homocysteine (Hcy), methyltetrahydrofolate (CH_3-H_4folate), cobalamin (B_{12}), and *S*-adenosylmethionine (AdoMet) (Figure VI.2). Overall, the following structures of MetH fragments are currently available, listed from N- to C-terminus.

Several structures have been solved of an N-terminal di-domain fragment containing the Hcy and CH_3-H_4folate domains, with and without zinc, cadmium (a zinc mimic), Hcy, and CH_3-H_4folate bound[11,14]. The constructs used in these studies were of a *Thermotoga maratima* MetH fragment containing residues 1-566, corresponding to residues 1-626 in *E. coli* MetH, which shares 25% sequence identity. No structure yet exists of the *E. coli* Hcy or CH_3-H_4folate domains.

One structure has been solved of the *E. coli* MetH B_{12} domain alone[8] (residues 651-896). Structures are also available for the C-terminal AdoMet domain alone for the *E. coli* enzyme (residues 901-1227) with AdoMet bound[9] and for the human enzyme[12] (residues 926-1264, corresponding to residues 892-1227 of *E. coli* MetH, with 49% sequence identity). In addition, many structures of a C-terminal di-domain fragment that contains both B_{12} and AdoMet domains have been determined (residues 649-1227); the initial structures contain a H759G mutation[10], while later structures retain the histidine residue but contain an engineered disulfide linkage that was used to help stabilize the complex in the conformation observed in the initial structures[13,15].

A crystal structure of protein components from a B_{12}-dependent methyl transfer system present in some methanogenic archaea[22], including *Methanosarcina barkeri*, has been determined. Prior to the work described in this thesis, this structure represented the only view of a B_{12} domain in complex with a domain that binds a methyl donor. The methyl donor here is not CH_3-H_4folate but methanol, and the enzymes involved are MtaA-C. MtaC is a 27 kDa B_{12}-binding protein analogous in function to the B_{12} domains of MetH and CFeSP. MtaB is a zinc protein that does not exhibit sequence similarity to any protein of known function and binds and activates methanol for methyl transfer to the B_{12} of MtaC. MtaA is also a zinc protein and is thought to bind coenzyme M (CoM) for methyl delivery from MtaC to form methyl-CoM, the central intermediate toward production of methane by methyl-CoM reductase (MCR) or oxidation to CO_2. The crystal structure contains MtaB and MtaC in complex, suggesting how a methyl group may be transferred from methanol to B_{12}.

Three structures of proteins likely involved in B_{12}-dependent methyl transfer have been deposited into the PDB, but at the time of writing have not been published or otherwise described in the literature. The first is a structure of a B_{12}-binding protein of unknown function from *M. thermoacetica*. Because this protein exhibits 27% sequence identity to the *E. coli* MetH B_{12} domain, it could represent a standalone B_{12} domain protein involved in methionine synthesis. The second structure depicts a *M. barkeri* B_{12}-binding protein called MtmC that is similar to MtaC and is involved in the monomethylamine methyltransferase (MMAMT) system, where a methyl group is transferred from monomethylamine to CoM. The third structure depicts a protein that is

annotated as the CH_3-H_4folate domain of MetH from *Bacteroides thetaiotaomicron* and contains CH_3-H_4folate bound.

A compilation of all available structures relating to B_{12}-dependent methyl transfer, including their respective Protein Data Bank (PDB) accession codes, is given in Table VI.1, and a schematic is provided in Figure VI.3. Here, we explore similarities and divergences in all currently available crystal structures for each modular component involved in B_{12}-dependent methyl transfer.

The B_{12}-binding domains

In 1994, the 3.0 Å resolution structure of B_{12} bound to a B_{12}-binding fragment of *E. coli* MetH was solved (27 kDa, residues 651-896) and represented the first crystal structure of a B_{12} cofactor bound to any protein[8] (Figure VI.4). Here, B_{12} (cobalamin) is bound in the enzyme's B_{12} domain, a Rossmann-like α/β fold with five parallel β-strands and six α-helices, similar to the traditional Rossmann fold that contains six parallel β-strands and four α-helices. B_{12} in the structure exists in the "base-off/His-on" conformation, where axial DMB coordination is replaced by a conserved histidine ligand from the protein, His759, and the DMB nucleotide "tail" is extended away from the corrin ring to become buried within the Rossmann-like fold. The more recent structures of MtaC from the MtaBC complex structure[22], in addition to the two unpublished structures of the putative MetH-like B_{12} protein from *M. thermoacetica* and the *M. barkeri* MMAMT B_{12} protein MtmC, are all virtually superimposable with the *E. coli* MetH B_{12} domain (Figure VI.5.A) and retain the same histidine ligand.

A similar, but slightly different structure is observed in CFeSP from both *M. thermoacetica* (Chapter II) and *C. hydrogenoformans*[1] (Figure VI.5.B). B_{12} is also bound in a Rossmann-like fold, but the domain contains one fewer β-strand and exhibits a different secondary structural topology and connectivity (Figure VI.5.C-D). The *M. thermoacetica* and *C. hydrogenoformans* B_{12} nucleotide tails are similarly buried within the Rossmann-like fold; however, no protein ligand coordinates the B_{12} cobalt in CFeSP. In fact, the region abutting the corrin ring is entirely different between MetH and CFeSP. In MetH, the His759 ligand resides on a loop which is replaced in CFeSP by a short helix (Figure VI.5.B). Curiously, the hydroxyl group of a conserved threonine residue from this helix sits close to the B_{12} cobalt, but it is still too far (~3.6 Å) to be a cobalt ligand[1]. However, current structural data do not rule out the possibility that the threonine residue may ligate the cobalt in another protein conformation. Indeed, structures of MetH His759 bound[8,15] and unbound[10,13,15] to the B_{12} cobalt have been obtained, and it has been proposed that this alternation is important for tuning the structural and catalytic properties of B_{12} during different stages of catalysis[15]. For CFeSP, spectroscopic and computational studies support the presence of some hydroxyl ligand to the B_{12} cobalt in the Co(II) and Co(III) forms[6], though this has been interpreted as a water ligand and not a protein ligand, as threonine coordination is unprecedented. In support of a water molecule instead of threonine ligation, electron density for water was observed on the opposite face of the corrin ring in the *C. hydrogenoformans* CFeSP structure (Co-O distance of 2.5 Å)[1]. Poor electron density due to B_{12} domain mobility in the *M. thermoacetica* CFeSP/MeTr structure prevented observation of the same putative water molecule.

The B_{12} "cap": schemes employed for B_{12} protection

The initial MetH B_{12} domain structure revealed an N-terminal subdomain composed of a four-helix bundle (residues 658-738), with an eight-residue long unstructured linker connecting it to the Rossmann-like fold responsible for binding of the B_{12} cofactor[8]. This subdomain has been referred to as the "capping" subdomain, as it appears to sterically block the active face of the corrin ring, covering the side of the cofactor with which the transferred methyl group must react (Figures II.22.C-D and VI.4). The corrin ring is thus sandwiched between the Rossmann-like domain and the "capping" subdomain. In this way, it is thought that the "capping" helices sequester the cofactor from the solvent, protecting reactive B_{12} intermediates from unwanted side reactions in a conformation that has been referred to as a "resting" state when the B_{12} domain is not positioned for catalysis. Mutagenesis experiments support a protective role for the "capping" subdomain, where Phe→Ala and Leu→Ala mutations in "capping" residues near the B_{12} cobalt atom expose the cofactor and significantly enhance the rate of methylcobalamin photolysis[25].

Interestingly, the unpublished structure of the MMAMT MtmC B_{12} protein from *M. barkeri* also contains the same "capping" subdomain that superimposes with the MetH "capping" subdomain (Figure VI.6). In the structure of the MtaBC complex, B_{12}-binding MtaC also contains the same "capping" subdomain, as well as an additional 25-residue N-terminal extension. The putative MetH-like B_{12} protein from *M. thermoacetica* may also contain a "capping" subdomain, but its unpublished structure lacks the N-terminal 84 residues. Nevertheless, the four-helix bundle of the "capping" subdomain appears to be a

common strategy for B_{12} protection in a "resting" state from systems as diverse as methionine biosynthesis and anaerobic one-carbon utilization.

However, CFeSP does not contain the "capping" subdomain. As discussed extensively in Chapter II, the CFeSP small subunit TIM barrel is instead thought to serve the same protective function[1], sequestering the reactive face of the B_{12} cofactor in the "resting" state prior to binding of MeTr (Figure II.22). Therefore, in a concept similar to MetH, the corrin ring of B_{12} is seen sandwiched between the Rossmann-like B_{12} domain and the CFeSP small subunit TIM barrel, instead of a four-helix bundle. Despite the fact that MetH and CFeSP utilize the same methyl group donor, CH_3-H_4folate, they appear to use completely different protein architectures for protecting the B_{12} cofactor, a possible example of convergent evolution.

Dissimilar routes to B_{12} activation

The mechanism for B_{12} activation is not shared between MetH and CFeSP. As discussed in Chapter II, inactive MetH with B_{12} in the Co(II) state is activated to methyl-Co(III) by reductive methylation, carried out in the AdoMet domain using an electron from flavodoxin and a methyl group from AdoMet. In contrast, for CFeSP a low-potential electron from a redox partner protein is injected into the cobalt center via the CFeSP large subunit Fe_4S_4 cluster to give the nucleophilic Co(I) state of B_{12}.

The CFeSP Fe_4S_4 activation domain was visualized for the first time in the CFeSP/MeTr structure described in Chapter II. The Fe_4S_4 domain appears as a bundle of short helices with long loops that coordinate the cluster, positioning the cluster at a significant distance (>20 Å) away from the rest of the CFeSP large subunit. The

conformation of the CFeSP/MeTr complex also places the Fe_4S_4 clusters far from B_{12} cobalt atoms (>38 Å); therefore, the structure does not represent the conformation necessary for electron transfer. With CFeSP bound to MeTr, the structure instead offers significant insight into the methyl transfer reaction, and formation of the B_{12} activation conformation may require dissociation of MeTr and/or binding of the redox partner protein. In any case, the structure of the CFeSP activation conformation remains elusive, and it is clear that large domain rearrangements from the known CFeSP structures must occur to afford B_{12} reduction by the Fe_4S_4 cluster.

On the other hand, multiple structures of *E. coli* and human MetH fragments have been solved that contain the C-terminal AdoMet activation domain alone[9,22]. The first such structure consisted of the *E. coli* AdoMet domain alone[9] (38 kDa, residues 901-1227) and showed a novel, crescent-shaped protein fold not seen in any other enzyme that resembled a curved hand, where AdoMet binds in the "palm" of the hand (Figure VI.7.A). The same fold was subsequently observed in the analogous human MetH AdoMet domain structure[22].

In later studies, *E. coli* MetH AdoMet domain structures depicted a larger (65 kDa) MetH fragment that additionally incorporated the entire B_{12} domain (residues 649-1227) and showed for the first time how B_{12} may interact with another domain (Figures VI.7.B-D)[10,13,15]. The first of these C-terminal di-domain structures[10] revealed a 25 Å movement of the "capping" subdomain discussed above, involving a ~63° rotation to allow B_{12} and AdoMet domains to interact (Figure VI.7.D). Soaking of these crystals with AdoMet yielded weak density for AdoMet bound to its domain that placed the methyl group ~6 Å away from the B_{12} cobalt, a distance too long for methyl transfer. A

tyrosine residue, Tyr1139, is further located in the path between the methyl group and cobalt (Figure VI.7.C), and thus it was concluded that additional, undefined rearrangements beyond mere displacement of the "capping" subdomain must take place to bring AdoMet and B_{12} closer together. Critically, a H759G MetH mutant was used in initial studies, where the histidine that ligates the B_{12} cobalt is absent, preventing axial coordination of the cobalamin; this mutant was used because previous studies had indicated that such a base-off species favors the AdoMet activation conformation[26,27].

More recently, numerous structures of the same C-terminal MetH fragment (residues 649-1227) incorporate an engineered disulfide linkage: in a double mutant (I690C/G743C), cysteines were placed in the "capping" subdomain and the linker that connects it with the B_{12} domain to covalently tether the "cap" away from the B_{12} cofactor[13]. The disulfide appeared to hold the "capping" subdomain in the same location as observed in the earlier B_{12}-AdoMet di-domain structure. Importantly, this construct also retained His759, which was observed to dissociate from the B_{12} cobalt and form a new inter-domain interaction with Asp1093 of the AdoMet domain[13] (Figure VI.7.C). Additional structures of the same construct further illustrate how differences in the ligation and oxidation state of the B_{12} cobalt may affect the cofactor's positioning and local environment[15]. These slight changes were proposed to tune the reactivity of B_{12} and allow for the AdoMet activation component of the reaction cycle to progress in an ordered fashion[15].

Overall, in all of the many MetH C-terminal di-domain structures, either mutagenesis of the B_{12}-coordinating His759 or tethering of the "capping" subdomain away from the B_{12} cofactor using a disulfide crosslink were used as chemical tricks to

trap the fragment in a conformation that enabled the capture of the B_{12} activation complex. It is possible that similarly shrewd tactics may be necessary to capture the CFeSP activation conformation in the future.

Domains that bind methyl group donors and acceptors

Eight structures of the N-terminal MetH di-domain fragment from *T. maritima* (25% sequence identity to *E. coli* MetH) have been solved in two different space groups, depicting the Hcy and CH_3-H_4folate domains[11,14]. Both domains fold as $(\beta/\alpha)_8$ triosephosphate isomerase (TIM) barrels and are connected by a 19-residue linker region (residues 292-314) that contains a β-hairpin (Figure VI.8.A). The TIM barrels are seen bound together, and the interface is partially formed by sandwiching the linker region in between the barrels (Figure VI.8.B). Due to the large buried surface of the interface, the authors proposed that the Hcy and CH_3-H_4folate domains would not disassemble and reassemble during the catalytic cycle[11].

TIM barrels appear to be a common fold for binding methyl group donors and acceptors: in addition to the Hcy and CH_3-H_4folate domains of MetH, MeTr also adopts a TIM barrel fold to bind CH_3-H_4folate. The methanol-binding MtaB protein also assembles as a TIM barrel and contains additional C-terminal helices that continue to encircle the barrel. The unpublished structure of the CH_3-H_4folate domain from *B. thetaiotaomicron* MetH (22% and 63% sequence identity to *T. maritima* and *E. coli* MetH, respectively) also shows a TIM barrel, with two additional α-helices at the C-terminus.

Regarding substrate binding, CH_3-H_4folate and Hcy in the *T. maritima* MetH N-terminal di-domain structures bind to the upper loops of their respective TIM barrel domains to mark the methyl donating and methyl accepting active sites, respectively. Curiously, in all MetH di-domain structures the two active sites are located ~50 Å apart and are positioned facing away from each other (Figure VI.8.A). Because *E. coli* MetH is thought to be monomeric[28], an issue discussed in further detail below, the large distance between the Hcy and CH_3-H_4folate active sites of the *T. maritima* MetH implied exceptionally large domain movements during the course of methyl transfer. However, no structural information has been obtained that could shed light into how the MetH B_{12} domain may access either CH_3-H_4folate or Hcy substrates, and there is no structural information that shows any connection between the Hcy or CH_3-H_4folate domains with the rest of MetH.

In the CFeSP/MeTr B_{12}-dependent methyl transfer system, MeTr serves an analogous in function to the MetH CH_3-H_4folate domain. Although the *T. maritima* MetH CH_3-H_4folate domain and *M. thermoacetica* MeTr maintain only 25% sequence identity, they share a conserved TIM barrel fold that is highly superimposable (Figure VI.8.C) with root-mean-squared deviation for C_α atoms (rmsd) of 1.09 Å, and their active sites are virtually identical, as noted previously[24].

In the active site of both MeTr and the MetH CH_3-H_4folate domain, an invariant asparagine residue interacts with the N^5 position of the CH_3-H_4folate substrate (Figure VI.9.A), which must be protonated during the reaction cycle to allow for methyl transfer to B_{12}, with H_4folate as a leaving group. Because asparagine is not a general acid, it is puzzling that this conserved residue lies at the presumed protonation site in all

CH_3-H_4folate domains. Such a situation in which no clear proton donor is positioned at a protonation site is also observed in purine nucleoside phosphorylase (PNP), where protonation of the purine N^7 position must accompany phosphorolysis of the purine-ribose linkage to yield the purine base and ribose-1-phosphate[29-31]. In PNP, a conserved asparagine residue similarly occupies the presumed ideal site for a catalytic acid. Therefore, the current model based on PNP studies and adapted for MeTr and MetH by structural comparisons[24] proposes that the asparagine side chain, uniquely capable of playing a dual role as a hydrogen bond donor and acceptor, is ideally situated to hydrogen bond with both unprotonated and protonated intermediates of the reaction. The source of the required proton is less clear and is thought to be from any of several residues involved in a hydrogen-bonding network within the substrate binding pocket to which the asparagine belongs (Figure VI.9.A). Following protonation of the N^5 position, the asparagine side chain could then aid in the stabilization of the protonated substrate[24], facilitating methyl transfer to the B_{12} cobalt.

Although the CH_3-H_4folate binding modules and active sites for CFeSP/MeTr and MetH systems are strikingly similar, the methyl group acceptor modules of CFeSP and MetH systems are entirely different, as the methyl group is passed to different downstream partners. In MetH, the methyl group is delivered to Hcy bound to the N-terminal zinc-containing Hcy domain to yield methionine, whereas CFeSP delivers the methyl group to an active site metallocluster called the A-cluster of acetyl coenzyme A synthase (ACS), a separate enzyme. Therefore, the ability to make structural comparisons between all methyl accepting modules in B_{12}-dependent methyl transfer is limited, especially as structures of the MetH Hcy domain[11,14] and ACS[32-36] are the only known

structures of B_{12}-related methyl accepting domains. The MetH Hcy domain contains a zinc ion coordinated by three cysteine residues and Hcy[11,14] (Figure VI.9.B). Here, zinc acts as a Lewis acid to activate nucleophilic attack of the Hcy thiolate onto the methyl group of methylcobalamin. In ACS, the catalytic mechanism for methyl transfer is controversial but may involve an A-cluster nickel atom, whose oxidation state, geometry, and composition are unknown, acting as a nucleophile to remove the methyl group from B_{12}, forming a Ni-C bond[37,38]. Following the delivery of the methyl group, B_{12} in both MetH and CFeSP systems reverts to the active Co(I) state and can continue to perform the next round of methyl transfer.

Possible dimerization of B_{12} methyltransferases at the CH_3-H_4folate domain

As discussed in Chapters II and III, CFeSP heterodimers bind to one MeTr homodimer to form the CFeSP/MeTr complex. In this way, a stable MeTr homodimer exists as a central protein that binds the CH_3-H_4folate methyl donor, to which CFeSP B_{12} domains must interact to afford methyl transfer. The homodimeric assembly of MeTr has long been noted: gel filtration indicates that MeTr is a homodimer[39], and all crystal structures of MeTr indeed illustrate such a homodimer in the asymmetric unit[23,24], including the CFeSP/MeTr complex discussed in Chapter II. In addition, the small-angle X-ray scattering (SAXS) solution studies described in Chapter III support the assembly of homodimeric MeTr. In the homodimer, the MeTr monomers are related by two-fold symmetry, where their TIM barrels sit side-by-side with their active sites pointing in opposite directions[23,24].

With MetH, however, the *E. coli* protein is assumed to be monomeric, a notion that is often mentioned in the literature but is rarely referenced; it appears that the monomeric composition of *E. coli* MetH is based upon early gel filtration studies[28]. However, all crystal structures of the MetH Hcy/CH_3-H_4folate di-domain from *T. maritima* (25% identity to the *E. coli* MetH CH_3-H_4folate domain) depict a homodimer of the CH_3-H_4folate domain related by non-crystallographic symmetry (ncs) within the asymmetric unit[11,14] (Figure VI.10.A). The CH_3-H_4folate domain ncs homodimer is a curious observation that has not been discussed in the literature, either because it has remained unnoticed or because a clear explanation for its assembly has been elusive. It is possible that the homodimer is simply a result of crystal packing; however, eight crystal structures from two space groups have shown the same homodimer in the asymmetric unit (Figure VI.10.B).

The crystal structure of the *M. thermoacetica* MeTr ncs homodimer superimposes surprisingly well with that of the *T. maritima* MetH CH_3-H_4folate domain ncs homodimer (Figure VI.11.A-B), with rmsd of 3.32 Å, despite sharing only 25% sequence identity. Thus, the individual monomers of both MeTr and MetH CH_3-H_4folate domain structures are related to each other by two-fold symmetry, with side-by-side TIM barrels whose active sites point in opposite directions. It is thus conceivable that MetH also dimerizes at the CH_3-H_4folate domain in an analogous fashion as MeTr, at least in the case of *T. maritima* MetH. Therefore, it may be necessary to conduct additional experiments to more firmly establish the MetH oligomeric state in solution.

The unpublished structure of the *B. thetaiotaomicron* MetH CH_3-H_4folate domain alone also shows a homodimer (Figure VI.11.C-D), with side-by-side crystallographic

TIM barrel monomers. Each *B. thetaiotaomicron* MetH CH_3-H_4folate domain monomer assembles as a TIM barrel that aligns well with the TIM barrels of the other CH_3-H_4folate domains. However, this homodimer is different than the *M. thermoacetica* MeTr and *T. maritima* MetH CH_3-H_4folate domain ncs homodimers. Here, the second monomer that completes the homodimer binds the opposite side of the TIM barrel (Figure VI.11.D). It is unclear where this different arrangement would place the Hcy domains, which are not part of the *B. thetaiotaomicron* MetH CH_3-H_4folate domain structure. Interestingly, the two additional C-terminal helices of the *B. thetaiotaomicron* MetH CH_3-H_4folate domain mentioned above appear to block what would have been the second monomer's binding site (Figure VI.11.C). Therefore, two different CH_3-H_4folate domain homodimers are possible. Because the *B. thetaiotaomicron* and *E. coli* MetH CH_3-H_4folate domains share 63% sequence identity, which extends to the additional C-terminal helices, it is possible that *E. coli* MetH may also dimerize in the same fashion as observed with *B. thetaiotaomicron*. In any case, it is clear that the potential for dimerization of *E. coli* and *T. maritima* MetH at the CH_3-H_4folate domain should be more carefully investigated in solution.

Although CH_3-H_4folate binding homodimers of MeTr in CFeSP/MeTr and *T. maritima* MetH CH_3-H_4folate domains superimpose well, their alignment does not superimpose the MetH Hcy domains onto any protein component of CFeSP/MeTr (Figure VI.12). This observation is consistent with the fact that the CFeSP/MeTr system does not contain a protein with an analogous function to the Hcy domain that could conceivably be required to fall in the same region. Instead, the methyl group in CFeSP is finally delivered to the active site A-cluster of ACS, a species wholly distinct from Hcy.

It is possible that ACS could bind to CFeSP/MeTr at a site similar to the Hcy domain location in MetH, but no structural information is available regarding the ACS interaction with CFeSP.

Structural insight into the methyl group transfer reaction

As stated, there is no structural information available to describe how the B_{12} domain of MetH interacts with the methyl donor (CH_3-H_4folate) or acceptor (Hcy) domains. The CFeSP/MeTr crystal structure, however, does provide a structural depiction of B_{12} domains together with a methyl donor domain, shedding much light onto how a mobile B_{12} domain may navigate from its "resting state" towards the MeTr active site for methyl transfer. As discussed in Chapter II, the CFeSP/MeTr structure represents an "en route" conformation that, while highly relevant to the inherent conformational flexibility of B_{12}-dependent methyltransferases, does not depict B_{12} within methyl transferring distance to the methyl donor. The 2.5 Å resolution crystal structure of a complex between MtaB and MtaC (Figure VI.13.A) represents the only structure currently available of B_{12} bound to a domain from which it accepts a methyl group, where the cofactor is near the active site[22]. Unlike MetH and CFeSP/MeTr, however, the MtaA-C system does not utilize CH_3-H_4folate or any other pterin derivative as a methyl donor.

In the MtaBC structure, MtaB exists as a TIM barrel, with additional C-terminal helices that continue to wrap around the barrel, and binds methanol using an active site zinc ion. As mentioned above, MtaC contains a Rossmann-like fold that binds the B_{12} cofactor. MtaC also harbors an N-terminal "capping" subdomain that forms the same four-helix bundle as observed in the MetH B_{12} domain structure[8]. In the MtaBC

assembly, however, the "capping" subdomain is swung away from the corrin ring, strongly reminiscent of the MetH B_{12}-AdoMet di-domain structures[10]. In fact, upon aligning the MtaBC and MetH B_{12}-AdoMet di-domain structures by their Rossmann-like B_{12} binding domains, the "caps" superimpose (Figure VI.13.B), a feature not mentioned in the prior description of the MtaBC structure[22]. In fact, the "capping" subdomains are located in the same position across the MtaBC and all MetH B_{12}-AdoMet structures[10,13,15], including those that do and do not contain the disulfide crosslink that holds the "cap" apart from the B_{12} domain. This comparison indicates that the "uncapped" form of B_{12} is the same in both MetH and MtaBC systems, despite the fact that methanol-binding MtaB, a TIM barrel, and the MetH AdoMet domain, a crescent shaped domain, serve completely different purposes and adopt completely different folds.

The MtaB active site zinc ion and the MtaC B_{12} cobalt are located 7.7 Å apart in the MtaBC structure. A methanol molecule bound to zinc was modeled into the structure, placing the methyl group and the B_{12} cobalt ~4.5 Å apart, too long for van der Waals contact[22]. Also, few contacts exist between the Rossmann domain of MtaC and MtaB, and the *B*-factors of the B_{12}-binding MtaC molecules are significantly higher than the MtaB molecules (Figure VI.13.C). This lack of contacts, the high *B*-factors, and the long Zn-Co distance led the authors to conclude that the active site captured represents an "open" state, and that side chain rearrangements and further conformational movements may afford tighter binding and a more closed active site to induce methyl transfer. Therefore, a tighter complex between B_{12} and methyl donor or acceptor domains that illustrates methyl transfer still awaits structural characterization. Given the inherently flexible nature of B_{12} domains, discussed in Chapter II, it is possible that more advanced

tactics than simple co-crystallization of domains, such as chemical crosslinking or enzymatic trapping, may be necessary to more adequately capture B_{12} methyl transfer.

The large size of B_{12} methyltransferases and the role of TIM barrels

In *E. coli*, MetH is a 136 kDa protein responsible for transferring a single methyl group of just 15 Da in size; as suggested above, MetH may even adopt larger oligomers. The structure of the CFeSP/MeTr complex from *M. thermoacetica* is over 220 kDa in size and performs the same, seemingly simple task. Both enzyme systems contain multiple TIM barrels that contribute greatly to the overall protein bulk, with the CFeSP/MeTr complex containing six TIM barrels comprising 1,610 residues out of 2,060, or 78% of the total complex. The TIM barrel content is also substantial, though somewhat less prominent, in *E. coli* MetH, where the Hcy and CH_3-H_4folate domains are also TIM barrels and encompass 638 residues of 1,227, or 52% of the total enzyme.

The TIM barrel domains of CFeSP do not bind substrates and serve no direct role in catalysis. As discussed, the small subunit TIM barrel is thought to fulfill a protective, "capping" function in the B_{12} domain "resting" state[1]. The large subunit TIM barrel, however, does not appear to have any clear function. It is possible that the heavy TIM barrel content of CFeSP serves as a structural scaffold to maintain rigidity of the CFeSP/MeTr complex, a concept alluded to briefly in Chapter II. Indeed, the large interface between the CFeSP TIM barrels suggests a stable surface, where large and small subunits are unlikely to dissociate from each other during the reaction sequence[1]. In contrast, the remaining CFeSP domains that are not TIM barrels, the Fe_4S_4 and B_{12}

domains, are thought to be considerably flexible and are tethered through linkers to the TIM barrel core.

In addition, the TIM barrels could be important to provide a surface to which the methyl donor and acceptor proteins may interact. Consistent with this notion, the CFeSP/MeTr structure shows that MeTr binds to CFeSP through helices of the CFeSP small subunit TIM barrel. From this TIM barrel scaffold, the mobile B_{12} domain can simply "swing" to meet MeTr and perform the first methyl transfer of the reaction cycle. In the second methyl transfer reaction from B_{12} to the ACS A-cluster, it is possible that ACS could also bind to a CFeSP TIM barrel, instead of directly to the B_{12} domain, though no structural information for the CFeSP/ACS interaction yet exists. In this way, the CFeSP TIM barrels would form a surface onto which upstream and downstream protein partners bind, where flexible B_{12} and Fe_4S_4 domains can then swing to reach the active sites presented. When no protein partner is bound to CFeSP, the B_{12} domain can then swing back to "re-cap" the B_{12} cofactor against the small subunit TIM barrel.

Unlike CFeSP, whose TIM barrel domains do not bind substrates, all MetH domains serve catalytic roles, either harboring an active site that performs chemistry on B_{12} (the Hcy, CH_3-H_4folate, and AdoMet domains) or carrying the B_{12} cofactor itself (the B_{12} domain). If a function of the CFeSP TIM barrel domains is simply to provide a stable structural scaffold to tether the mobile B_{12} and Fe_4S_4 domains and to interact with upstream and downstream protein partners, it is intriguing that these analogous domains do not exist in MetH. However, such a structural scaffold may not be entirely necessary for MetH. Unlike CFeSP, the upstream and downstream protein partners in MetH, the CH_3-H_4folate and Hcy domains, are already covalently tethered as part of the MetH

polypeptide. Therefore, the requirement of stable TIM barrels to tether the flexible B_{12} domain could already be provided by the CH_3-H_4folate and Hcy domains without requiring the additional TIM barrels in CFeSP, whose methyl donor and acceptor domains reside on separate polypeptides. However, the way in which the MetH B_{12} domain interacts with the CH_3-H_4folate and Hcy domains remains elusive.

Because structural information on B_{12}-dependent methyl transfer proteins outside of MetH and CFeSP are lacking, it is perhaps premature to generalize the importance of TIM barrels as structural scaffolds in these systems. The only other structure currently available to extend the TIM barrel scaffolding model, aside from the unpublished standalone MetH domain structures, is that of MtaB (Table VI.1), the methanol-activating protein from *M. barkeri* discussed previously. MtaB also adopts a TIM barrel fold, despite a lack of sequence similarity to other methyl donor or acceptor domains or, in fact, to any other protein of known function[22]. MtaB further contains seven additional C-terminal helices comprising 115 residues that continue to encircle the TIM barrel, increasing its thickness (Figure VI.13.A). Notably, this helical layer contacts another MtaB monomer in the asymmetric unit as well as the N-terminal extension of its MtaC partner, which may help to displace the MtaC "capping" subdomain and allow for the MtaB-MtaC interaction. Further, Hagemeier *et al.* also suggest that the MtaB TIM barrel could interact with MtaA[22], though this proposal awaits structural evidence. Regardless, the MtaBC structure indicates that the MtaB TIM barrel may also play a role as a structural scaffold, consistent with the MetH and CFeSP structures. The heavy presence and protein interactions of TIM barrels in all B_{12}-dependent methyl transfer enzymes characterized thus far are consistent with a structural role for TIM barrels as scaffolding

elements, and it will be fascinating to see if this continues to be a recurring theme as future crystal structures of proteins related to B_{12}-dependent methyl transfer are surely determined.

C. Mimicking the CODH C-cluster in model complexes

The water-gas shift reaction and the C-cluster

Although Nature has long used the nickel-containing C-cluster of carbon monoxide dehydrogenase (CODH) to interconvert CO_2 and CO, humans have also developed industrial processes that perform an analogous transformation. In the industrial water-gas shift (WGS) reaction, CO and water vapor are converted to CO_2 and dihydrogen in the presence of a heterogeneous catalyst, according the following equation:

$$CO_{(g)} + H_2O_{(v)} \rightleftharpoons CO_{2(g)} + H_{2(g)}$$

The reaction is exothermic (ΔH_{rxn} = -41 kJ mol^{-1}), and increasing the temperature shifts the equilibrium towards the reactants. Thus the reaction is typically performed at relatively low temperatures[40], around 100 °C.

A major use of the WGS reaction is as a step in anthropogenic nitrogen fixation through the Haber-Bosch process, a fundamental industrial process responsible for fixing 10^{15} kg of nitrogen per year to ammonia for use as a raw material for production processes as diverse as nylon, plastics, resin, feed, fertilizer, and explosives[41]. Due to the heavy use of fertilizers in the agricultural industry, it is estimated that roughly 40-50% of nitrogen in human proteins originally derives from the Haber-Bosch process, with the remaining nitrogen obtained from microbial nitrogen fixation[41]. In the Haber-Bosch process, N_2 and three equivalents of H_2 form two equivalents of ammonia, and the WGS

reaction is used a source of dihydrogen. CO for the WGS reaction is provided from fossil fuels. Another major industrial use of the WGS reaction is the production of synthesis gas (syngas), a mixture of CO and H_2 that is used for the production of synthetic natural gas and petroleum fuels from coal.

Various metal-based catalysts are used for the WGS reaction, such as iron, copper, gold, platinum, rhodium, ruthenium, palladium, and zinc compounds, as well as oxides of the lanthanide cerium[40-44]. In most applications, a mixture of catalysts are often used. The precise mechanism of catalysis is somewhat controversial, and different catalysts may employ different mechanisms. One mechanistic proposal termed the "carbonyl mechanism" involves a metal-carboxyl (M-COOH) intermediate, while others involve heterogeneous redox chemistry[44].

Despite the current widespread use of the WGS reaction, other industrial methods for CO/CO_2 interconversion have been investigated, as discussed below. Motivations for research in this area include the search for catalysts that are cheaper, involve more abundant elements, are easier to synthesize, or those that react in more ambient or environmentally friendly conditions. Because Nature uses the CODH C-cluster to perform a similar process as the WGS reaction, a deeper understanding of the C-cluster mechanism, the subject of Chapters IV and V, may help guide the development of a biomimetic catalyst.

Model complexes geometrically similar to the C-cluster

The geometry of the metals in the C-cluster came as a shock to the bioinorganic chemistry community when crystal structures of CODH were first revealed in 2001[45,46].

Prior to the publication of these structures, it was assumed that nickel was bridged in some fashion to a more traditional [4Fe-4S] cubane commonly observed in biology. Biomimetic systems had thus been prepared that bridged nickel(II) to [4Fe-4S] cubanes. One such assembly contained a central [4Fe-4S] cubane with two flanking Ni-N_2S_2 moieties, bound through bridging sulfides[47] (Figure VI.14.A). An additional compound contained one Ni linked to a [4Fe-4S] cluster[48]. Further attempts were also made to construct binuclear and trinuclear iron-nickel complexes that mimicked fragments of the presumed cluster geometry and bound CO[49,50] (Figure VI.14.B).

The structure of the active site C-cluster was markedly different from the presumed models. In the CODH C-cluster, nickel occupies a cuboidal-like [Ni-3Fe-4S] cluster, and a unique iron is also coordinated nearby via a μ_3-sulfido linkage with a sulfur of the cubane. Such an assembly had no precedent in biology or chemistry[51]. Several similar cuboidal [Ni-3Fe-4S] species had been previously synthesized, prior to the knowledge of the CODH C-cluster structure[52,53]. These have involved either the reaction of a linear trinuclear iron system and $Ni(PPh_3)_4$ to give a [Ni-3Fe-4S] cubane[52] (Figure VI.15.A) or reaction with an incomplete [3Fe-4S] cluster with $NiCl(PPh_3)_3$ to produce the same structural core (Figure VI.15.B)[53]. After the structure of the C-cluster was known, additional complexes containing a [Ni-3Fe-4S] cubane were also prepared and characterized that represent models that most closely resemble the C-cluster[54]. One such complex that strongly resembles that of the C-cluster is composed of a distorted cubane, where nickel and a sulfide are pulled apart to break the Ni-(μ_3-S) bond, with low-spin Ni(II) possessing square planar geometry (Figure VI.15.C). However, to more accurately reflect the C-cluster an additional *exo*-iron to the cubane must also be present, bound

through the cubane sulfide that is separated from nickel, but such a complex has yet to be synthesized.

Although the Ni-Fe heterometallic cubanes described have been constructed to resemble the overall geometry of the C-cluster, these complexes are often incapable of binding CO, water/hydroxide, or cyanide, and testing for reactivity by treatment with CO can lead to degradation. A recent study illustrates a series of mononuclear Ni complexes that attempt to mimic substrate binding to the C-cluster, where Ni is coordinated by three nitrogens of a pincer ligand. Here, Ni has been shown to bind hydroxide, CO_2, formate, and cyanide[55] (Figure VI.16.A). Ni-Fe binuclear complexes using similar ligands were also constructed and could bind CO_2 and cyanide[55] (Figure VI.16.B). However, these complexes could not perform CO oxidation or CO_2 reduction, and more reactive complexes have been sought.

Catalytically active CODH-inspired nickel complexes

Because C-cluster-like Ni-Fe cuboidal complexes that have been synthesized are unreactive towards CO oxidation, some research has been directed at producing functional models that may not as accurately resemble the geometry of the CODH C-cluster but still make use of nickel. One example is a binuclear Ni(II) complex with thiolate ligands (Figure VI.17.A) that has been shown to perform CO oxidation with nucleophilic attack by water at room temperature in aqueous solution, yielding CO_2 and Ni(I)[56]. Electron transfer to methylviologen can reoxidize the system to Ni(II). Interestingly, cyanide in this system is inhibitory, a feature that is shared with the CODH C-cluster (Chapters IV and V).

Meanwhile, square-planar Ni(II) complexes[57,58] and Ni-phthalocyanines[59] can reduce CO_2 to CO. A phosphine-Ni(0) complex with bound CO_2 can be reduced by thiols and other proton donors give CO and water, where protonation of Ni-CO_2 is key to the mechanism[60,61]. Interestingly, the mechanism of reduction of this complex in 1988 was suggested at the time to be relevant to the understanding of the CODH C-cluster mechanism. Indeed, the current consensus mechanism outlined in Chapter V does involve protonation of a Ni-COO^- intermediate in the direction of CO_2 reduction.

Most recently, a square planar nickel complex utilizing a similar nitrogen-coordinating pincer ligand as described above was created that was capable of rapidly converting CO_2 to bicarbonate[62] (Figure VI.17.B). Here, a terminal Ni-hydroxide complex (Ni-O_aH) is reacted with CO_2 (O_b=C=O_c), which approaches out-of-plane to insert an oxygen into the Ni-O_a bond, forming Ni-O_b-C. Although the product of the reaction, bicarbonate (HCO_3^-), is not the same as that of the CODH C-cluster, CO, this finding does represent a possible Ni-inspired route towards carbon fixation. However, this mechanism differs considerably from the C-cluster, as hydroxide is bound to nickel and not iron, as has been observed in the CODH crystal structures discussed in Chapters IV and V. A binuclear Ni-Fe complex that can catalyze CO_2 to CO interconversion and resembles the CODH C-cluster has yet to be constructed.

VI.D. The Wood-Ljungdahl pathway in CO_2 sequestration and alternative energy

The modeling studies described above represent the possibility of synthetically mimicking the C-cluster, an intriguing route towards CO_2 fixation and sequestration. Ultimately, it may be possible to use such model complexes as scrubbers for CO_2

removal from flue gas produced by fossil fuel burning power plants and manufacturing facilities. Carbon sequestration from air is also an active area of research[63,64], but the significantly lower CO_2 concentrations compared to flue gas make this method more difficult and costly.

In an alternative approach, acetogenic and methanogenic microorganisms that use the Wood-Ljungdahl carbon fixation pathway could theoretically be cultured in mass quantities for CO_2 removal from the atmosphere. The major barrier is the extreme oxygen sensitivity of the metalloenzymes in this pathway. However, oxygen tolerance of these enzymes has not been studied in detail, and it is possible that engineered enzymes could be created that exhibit increased oxygen tolerance. This concept is already an active area of research in relation to hydrogenases[65-67] and similar studies may conceivably be extended to CODH, ACS, and CFeSP.

Acetogenic organisms could also be used for the maintenance of microbial fuel cells (MFCs), electrochemical cells where bacteria are cultured to pass electrons to or from external metal electrodes to generate a current. The bacteria employed in MFCs require a fuel source, and sucrose and acetate are examples are currently used carbon sources. Theoretically, acetogens could be co-cultured in MFCs to provide acetate from CO_2, in much the same way as they already do in Nature for termites and for other microbes that live symbiotically with acetogens in the mammalian intestinal rumen. MFCs are often maintained under anaerobic environments to preferentially allow the electrode and not dioxygen to act as the terminal electron acceptor. Therefore, the issue of oxygen sensitivity in obligately anaerobic acetogenic bacteria would already be satisfied in MFCs.

The Wood-Ljungdahl carbon fixation pathway plays a crucial role in the global carbon cycle and allows various microorganisms to use CO_2 and CO as carbon and energy sources. It thus seems feasible to exploit the insight gained from research into the enzymology of this pathway towards industrial CO_2 sequestration or energy generation. As the problem of global climate change intensifies, perhaps Nature's approach towards CO_2 utilization can inspire the development of novel solutions to the current carbon and energy crisis.

VI.E. Future directions

Chapter II describes the crystal structure of the CFeSP/MeTr complex where the B_{12} domain is observed "en route" to the MeTr active site, but there are many other forms of the enzyme complex that should be characterized and compared. Chief among these is the form of the complex bound with the CH_3-H_4folate substrate to MeTr. Chapter II proposes that the B_{12} domain is flexible upon CFeSP binding to MeTr, existing as an ensemble of conformers. The equilibrium of B_{12} domain positions is thought to be influenced by the binding of substrate, the redox and ligation state of the B_{12} cobalt, and the presence of upstream and downstream protein partners. Therefore, it will be interesting to determine whether the binding of substrate induces additional conformational movements of the B_{12} domain, providing support for the proposal that the B_{12} domain equilibrium may be influenced in this way. Further complexes that could be studied crystallographically include the methylated form of CFeSP, where B_{12} is converted to CH_3-Co(III), as well as CFeSP containing different redox states of B_{12}. Ultimately, a better structural understanding of each step of the reaction cycle would

contribute significant insight into the CFeSP/MeTr and other related B_{12}-dependent methyltransferase systems.

X-ray crystallography has made major contributions to our understanding of the CODH C-cluster mechanism and has been a significant driving force behind past and future attempts to create biomimetic model complexes for synthetic C-cluster chemistry. As described above, an intriguing, new area of research may be initiated that investigates the oxygen sensitivity of CODH. Because the D-cluster resides at the very surface of the protein, it is likely that it is first damaged by oxidation to lead to an inactive enzyme. It is possible that other means of electron transfer could be explored that bypass this cluster; adsorption of CODH to an electrode surface has already been performed and has led to successful CO/CO_2 interconversion[68,69]. In this way, it may be feasible to create an enzymatic fuel cell with electrons deriving from H_2 oxidation by hydrogenase being used for CO_2 reduction by CODH[70]. Moreover, there are some reports in the literature that some CODHs may exhibit inherent hydrogenase activity[71], with direct electron transfer to connect CO/CO_2 and H_2/H^+ redox half-cells. Further investigation into the prospect for understanding hydrogenase activity in CODH enzymes may thus represent an interesting avenue for future research into the C-cluster.

The work described in this thesis concerns the structures of metalloenzymes that contribute one-carbon groups toward the synthesis of acetyl-CoA in the Wood-Ljungdahl carbon fixation pathway. CFeSP and MeTr, the subjects of Chapters II and III, are involved in the delivery of a methyl group for acetyl-CoA, while CODH, the focus of Chapters IV and V, is responsible for providing the carbonyl moiety. However, the way in which these two one-carbon groups converge and are combined at the ACS A-cluster

in the final step of the pathway is relatively unknown and highly controversial. The ACS A-cluster has come to represent a sizeable hole in our fundamental understanding of the Wood-Ljungdahl pathway, appearing ever larger as mechanistic details of the upstream reactions have become progressively clearer. As illustrated in Chapter V, considerable advances towards understanding the mechanism of the C-cluster have come from crystallographic studies of substrates and inhibitors bound to the cluster. Perhaps a similar approach can be exploited to help reveal further catalytic insight into the enigmatic yet environmentally important ACS A-cluster. Such discoveries may one day bring to light the mechanisms employed by all of the metalloenzymes complexes of the Wood-Ljungdahl carbon fixation pathway.

Table VI.1. Currently available crystal structures of protein components for B_{12}-dependent methyl transfer.

PDB ID	Space Group	Enzyme(s)	Organism	Domain(s)	Residues	Mutations(s)	Bound species	Reference
(TBD)	$P2_12_12$	CFeSP/MeTr	*M. thermoacetica*	CFeSP-large	5-442	WT	B_{12}, Fe_4S_4	Chapter II
				CFeSP-small	1-323	WT		
				MeTr	1-262	WT		
2H9A	$C2$	CFeSP	*C. hydrogenoformans*	Large	59-441	WT	B_{12}, Fe_4S_4*	1
				Small	3-309	WT		
1F6Y	$P2_12_12_1$	MeTr	*M. thermoacetica*	MeTr	1-262	WT		23
2E7F	$P2_12_12_1$	MeTr	*M. thermoacetica*	MeTr	1-262	WT	CH_3-H_4folate	24
2OGY	$P2_12_12_1$	MeTr	*M. thermoacetica*	MeTr	1-262	N199A	CH_3-H_4folate	24
1Q7Z	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-559	WT	Cd^{2+}	11
1Q85	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-559	WT (SeMet)	Cd^{2+}	11
1Q8A	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-559	WT (SeMet)	Cd^{2+}, Hcy	11
1Q8J	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-559	WT	Cd^{2+}, Hcy, CH_3-H_4folate	11
1Q7M	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-559	WT		11
1Q7Q	$P2_12_12_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-565	WT		11
3BOL	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-560	WT	Zn^{2+}, Hcy	14
3BOF	$P2_1$	MetH	*T. maritima*	Hcy, CH_3-H_4folate	1-560	WT	Zn^{2+}, Hcy	14
1BMT	$P2_12_12_1$	MetH	*E. coli*	B_{12}	651-896	WT	B_{12}	8
1MSK	$P2_12_12_1$	MetH	*E. coli*	AdoMet	901-1227	WT	AdoMet	9
1K7Y	$P4_32_12$	MetH	*E. coli*	B_{12}- AdoMet	649-1227	H759G	B_{12}	10
1K98	$P4_32_12$	MetH	*E. coli*	B_{12}- AdoMet	649-1227	H759G	B_{12}, AdoMet	10
3BUL	$P4_32_12$	MetH	*E. coli*	B_{12}- AdoMet	651-1227	I690C/G743C	B_{12}	13
3IVA	$P4_32_12$	MetH	*E. coli*	B_{12}- AdoMet	651-1226	I690C/G743C	B_{12}, AdoHcy	15
3IV9	$P4_32_12$	MetH	*E. coli*	B_{12}- AdoMet	651-1227	I690C/G743C	B_{12}	15
2O2K	$P2_12_12_1$	MetH	Human	AdoMet	926-1264	WT		12
2I2X	$P2_1$	MtaBC	*M. barkeri*	MtaB	3-461	WT	Zn^{2+}	22
				MtaC	1-258	WT	B_{12}	
1Y80	$P2_12_12$	MetH?	*M. thermoacetica*	B_{12}	85-209	WT	B_{12}	Unpublished
3EZX	$C222_1$	MMAMT: MtmC	*M. barkeri*	B_{12}	2-216	WT	B_{12}	Unpublished
3K13	$C2$	MetH	*B. thetaiotaomicron*	CH_3-H_4folate	350-647	WT (SeMet)	CH_3-H_4folate	Unpublished

* Fe_4S_4 cluster was located by Fe anomalous diffraction, but the Fe_4S_4 domain was disordered

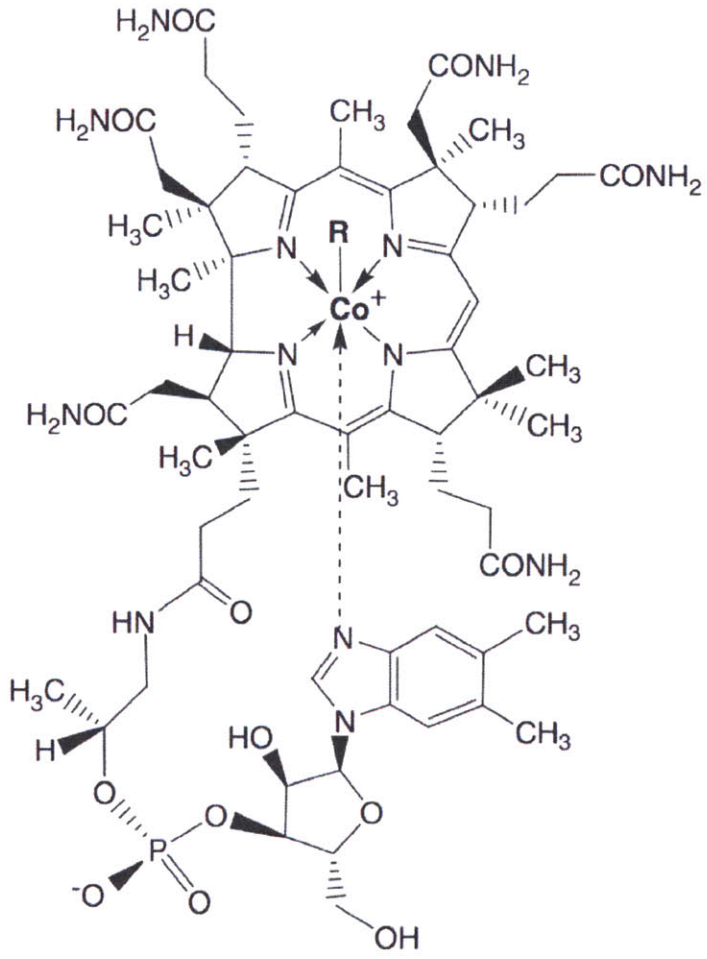


Figure VI.1. General structure of the cobalamin B_{12} derivative. R = CN, CH_3, 5′-deoxy-5′-adenosine, H_2O, and OH in cyanocobalamin, methylcobalamin, adenosylcobalamin, aquocobalamin, and hydroxocobalamin, respectively. Cobalt in all of these forms exists in the +3 oxidation state. Arrows from the equatorial nitrogen ligands of the tetrapyrrole indicate dative bonds to the cobalt center. The dashed arrow indicating a dative bond between the 5,6-dimethylbenzimidazole (DMB) base and the cobalt center represents the fact that some enzymes bind cobalamin in the "base-off" form, where DMB does not coordinate cobalt. Image modified from a previously published figure[72].

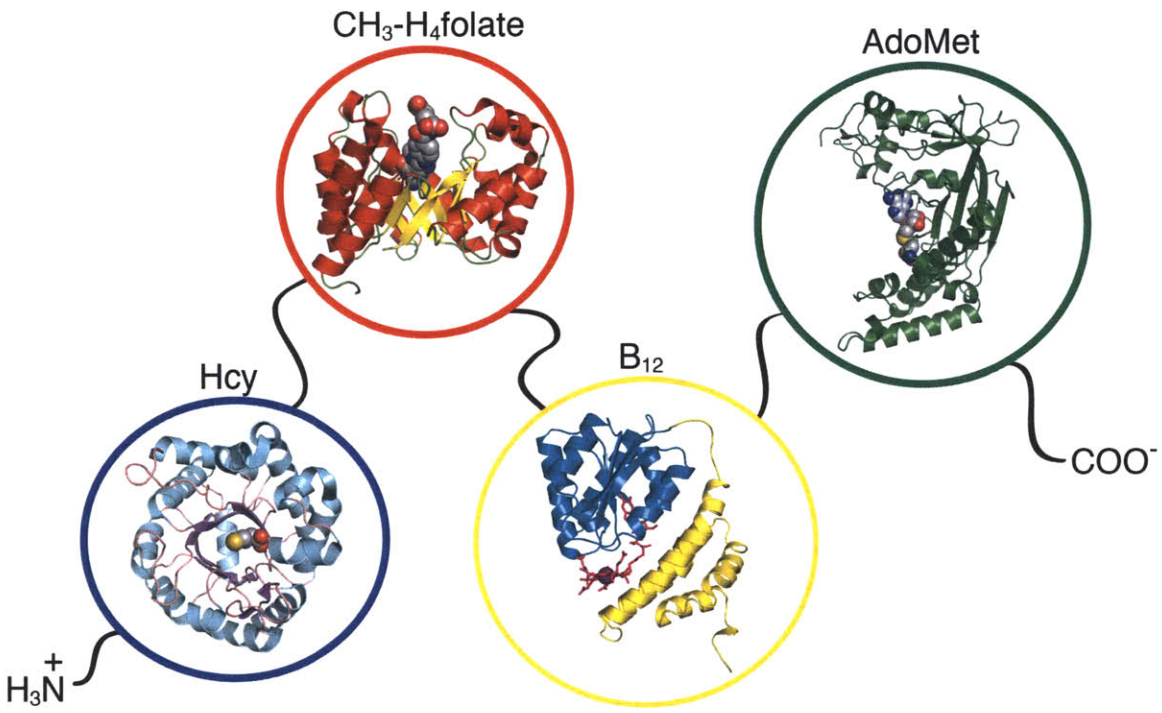
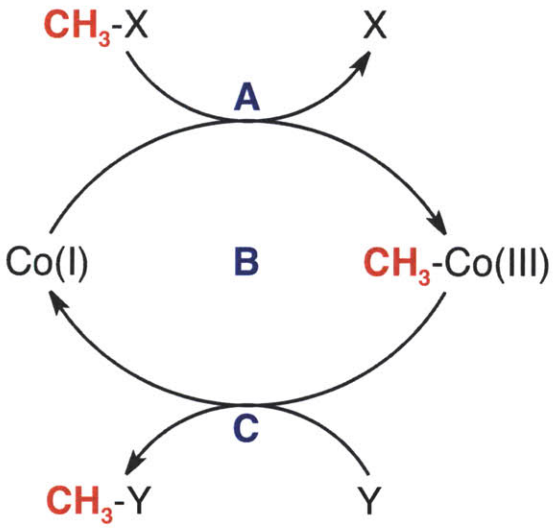


Figure VI.2. The "beads-on-a-string" modular topology of domains in MetH. Crystal structures of each domain exist and are shown "strung" together from N- to C-terminus: homocysteine (Hcy), methyltetrahydrofolate (CH_3-H_4folate), cobalamin (B_{12}), and *S*-adenosylmethionine (AdoMet). Structures are shown in ribbons of varying colors, and small molecule substrates bound are shown in spheres: C in grey, O in red, N in blue, and S in yellow. B_{12} bound to the B_{12} domain is show in pink sticks, with the central cobalt as a violet sphere.



system	A	B	C	CH_3-X	Y
CFeSP/MeTr	MeTr	CFeSP (B_{12})	ACS	CH_3-H_4folate	A-cluster
MetH	MetH (CH_3-H_4folate)	MetH (B_{12})	MetH (Hcy)	CH_3-H_4folate	Hcy
MtaA-C	MtaB	MtaC	MtaA	CH_3-OH	CoM
MMAMT	MtmB	MtmC	MtbA	CH_3-NH_3^+	CoM

Figure VI.3. Schematic of B_{12}-dependent methyltransferase systems where crystal structures are known for some protein component. In the scheme, the transferred methyl group is colored in red, and proteins components are labeled in blue (A-C). In the table, individual protein domains are in parentheses, and the crystal structures of protein components colored in blue have been solved, while those colored black have yet to be determined. Only protein components involved in methyl transfer activity are included, while those responsible for other functions, such as B_{12} "capping" and activation, are not shown.

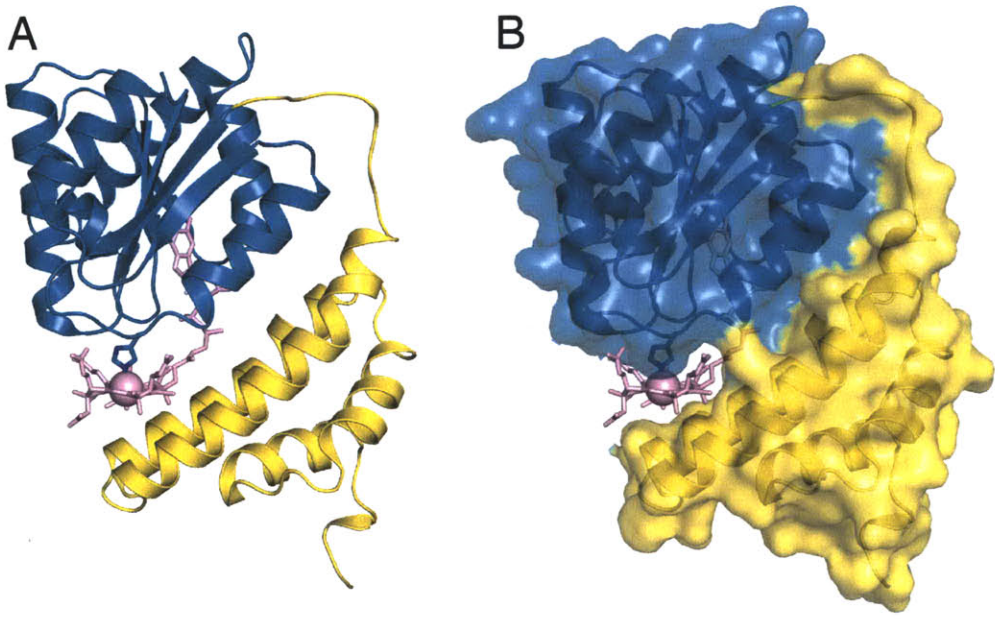


Figure VI.4. The MetH B_{12} domain. A, *E. coli* MetH (PDB ID: 3BMT) B_{12}-binding Rossmann-like fold in teal ribbons, "capping" subdomain in yellow ribbons, His759 in sticks (C in teal and N in blue), and B_{12} in pink sticks, with cobalt as a pink sphere. **B**, same as **A** with a semi-transparent protein surface shown to illustrate B_{12} sequestration and protection by the "capping" subdomain.

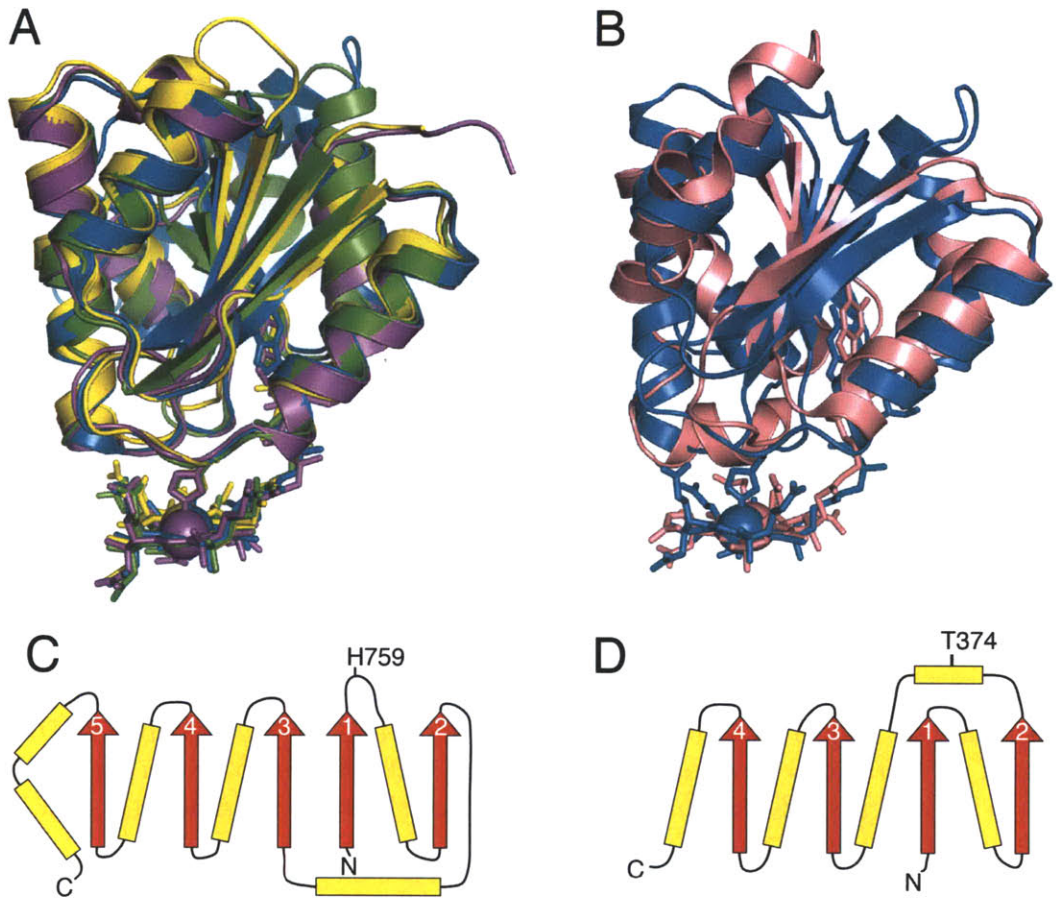


Figure VI.5. The B_{12} domains of B_{12}-dependent methyltransferases. A, Structural alignment of putative methyltransferase B_{12}-binding domains with *E. coli* MetH in teal (PDB ID: 1BMT): *M. barkeri* MtaC in green (PDB ID: 2I2X, rmsd: 0.69 Å), unpublished *M. barkeri* MtmC in yellow (PDB ID: 3EZX, rmsd: 0.98 Å), unpublished *M. thermoacetica* putative MetH in magenta (PDB ID: 1Y80, rmsd: 0.68 Å). Histidines that coordinate the B_{12} cobalt shown in sticks and B_{12} in sticks, with cobalt as a sphere. **B**, Manual structural alignment of the *C. hydrogenoformans* CFeSP B_{12}-binding domain in pink (PDB ID: 2E9A) to that of *E. coli* MetH in teal (PDB ID: 1BMT). Histidine of MetH (His759) that coordinates the B_{12} cobalt is shown in sticks and B_{12} in sticks, with cobalt as a sphere. **C**, Topology diagrams of *E. coli* MetH and **D**, *C. hydrogenoformans* CFeSP. The MetH β-strands numbered 5, 4, 3, and 1 superimpose with the CFeSP β-strands numbered 4, 3, 1, and 2, respectively. The position of His759 that coordinates the B_{12} cobalt in MetH is labeled in **C**. In the CFeSP structure, a short helix replaces the His759-containing loop of MetH, and Thr374 labeled in **D** is instead the closest residue to the B_{12} cobalt.

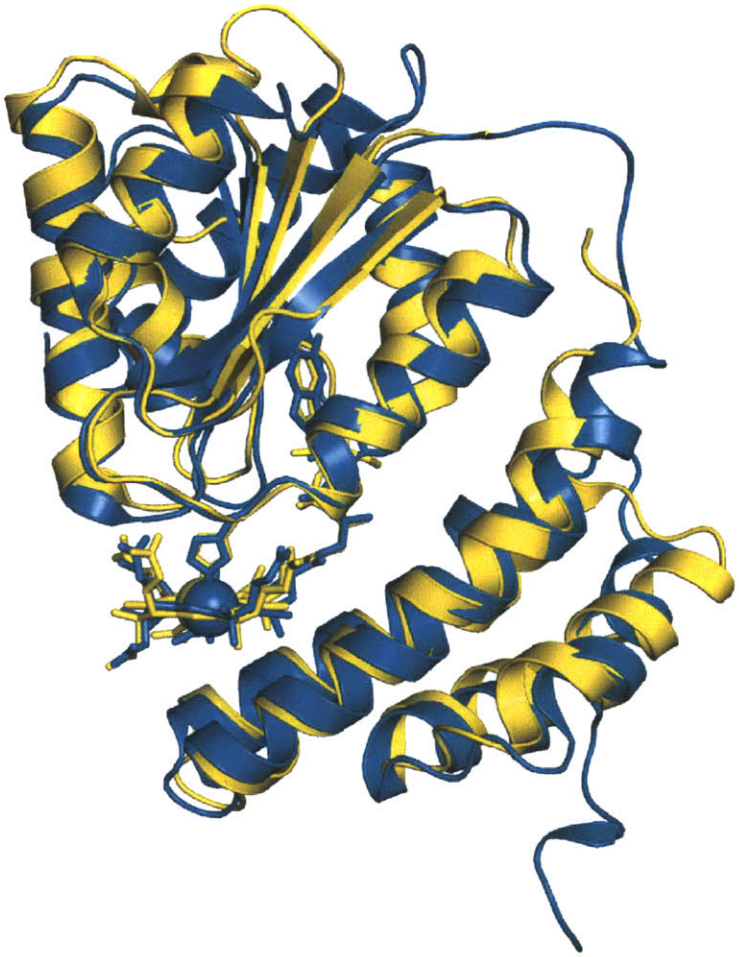


Figure VI.6. Alignment of B_{12}-binding domains superimposes the "capping" subdomain. The *E. coli* MetH B_{12} domain is shown in teal (PDB ID: 1BMT), and the unpublished *M. barkeri* MMAMT B_{12} binding MtmC protein is shown in yellow (PDB ID: 3EZX, rmsd: 0.98 Å). A chain break is present in the latter structure between the B_{12}-binding Rossmann fold and the "capping" subdomain. Histidines that coordinate the B_{12} cobalt are depicted in sticks, and B_{12} is shown in sticks with the cobalt as a sphere.

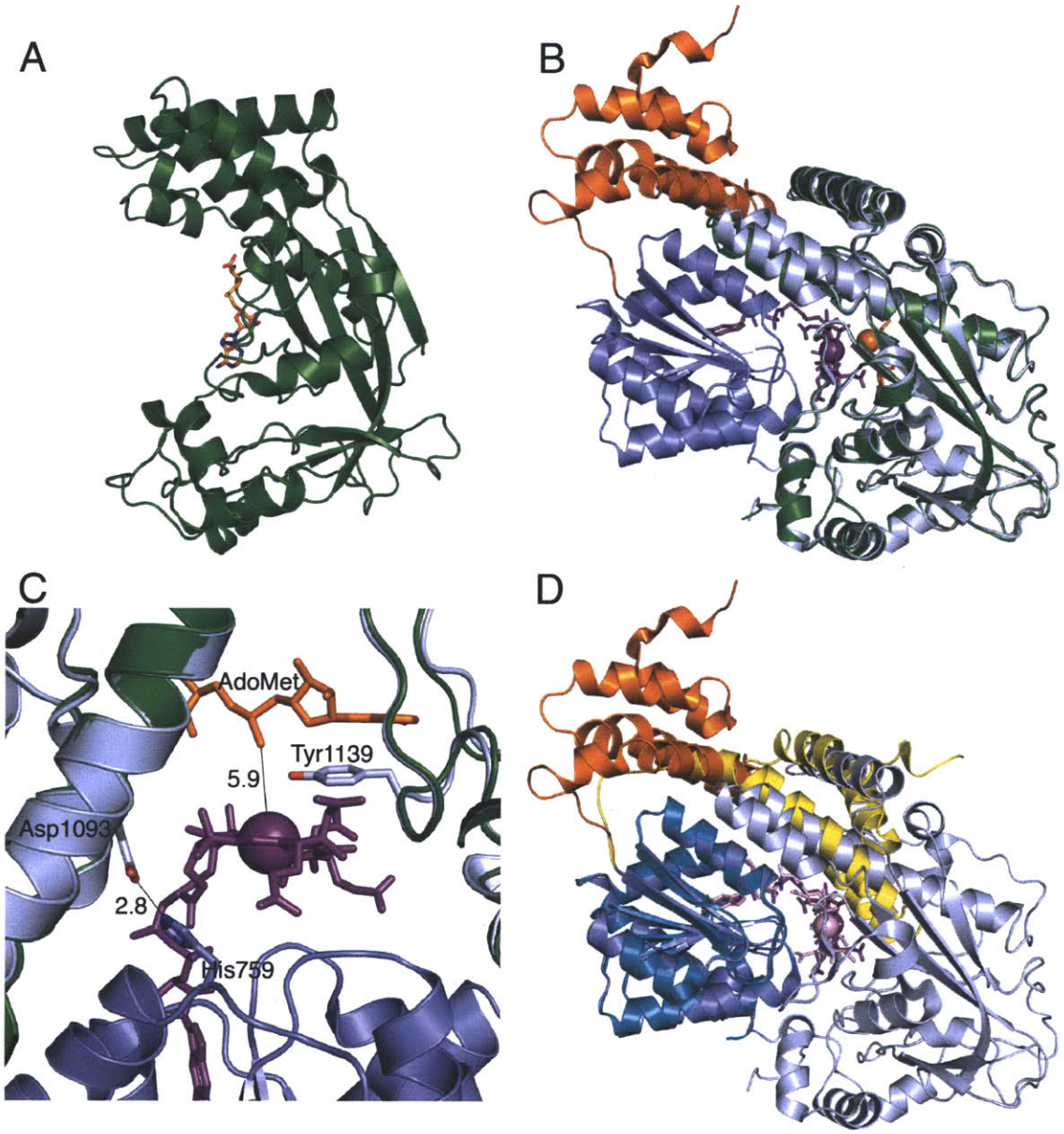


Figure VI.7. Crystal structures informing B_{12} activation in *E. coli* MetH. A, Structure of the AdoMet domain (PDB ID: 1MSK) in green ribbons, with AdoMet in sticks (C in orange, O in red, N in blue, and S in yellow). **B**, The structure in **A** (green ribbons, AdoMet in orange sticks, with the methyl group to be transferred as a sphere) aligned with the C-terminal B_{12}/AdoMet di-domain (PDB ID: 1K98, contains the H759G mutation): "capping" subdomain in orange, B_{12}-binding domain in slate, AdoMet domain in light blue, and B_{12} in violet sticks with cobalt as a sphere. **C**, Close up of the region surrounding the B_{12} cofactor in the B_{12}/AdoMet di-domain (PDB ID: 3BUL, with H759). Same coloring as **B**. His759, Tyr1139, and Asp1093 are labeled and shown in sticks, with C in light blue, O in red, and N in blue. Distances in Å. **D**, Movement of the B_{12} "cap". Same depiction as **B**, but without the AdoMet domains structure, and with the structure of the *E. coli* MetH B_{12} domain alone (PDB ID: 1BMT): "capping" subdomain in yellow, B_{12}-binding domain in teal, and B_{12} in pink sticks, with cobalt as a sphere.

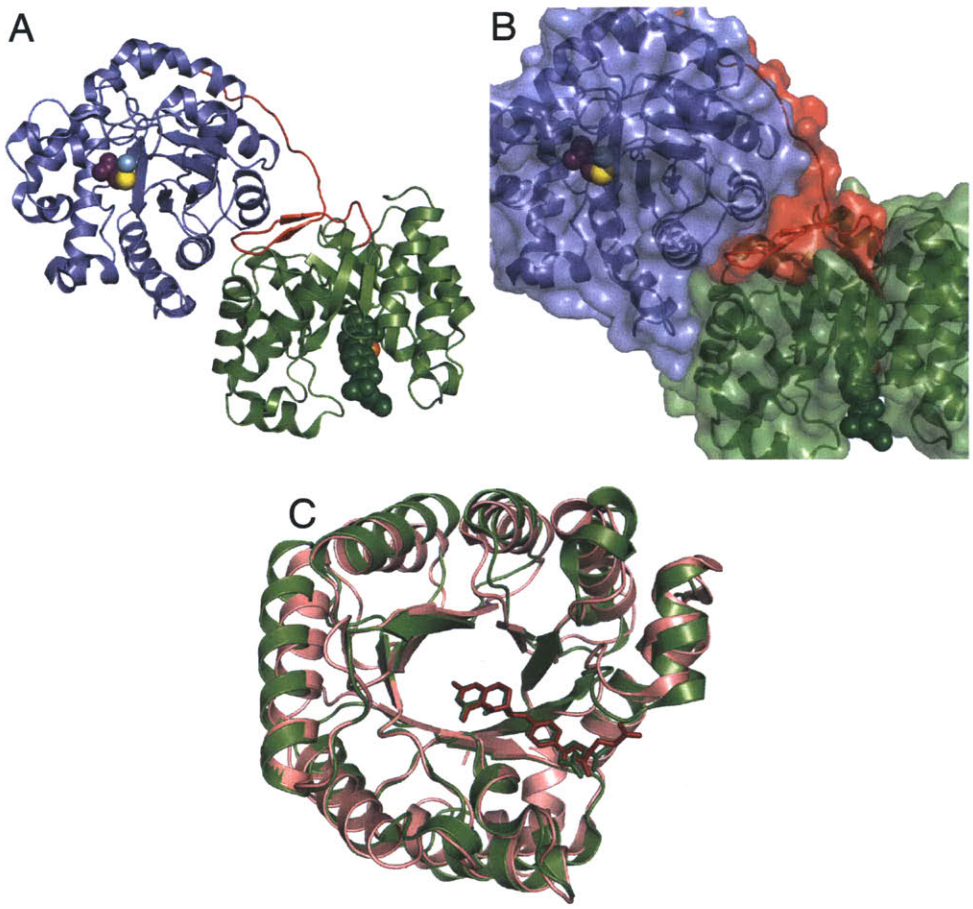


Figure VI.8. The N-terminal domains of MetH. A, Crystal structure of the N-terminal Hcy/CH_3-H_4folate di-domain of *T. maritima* MetH (PDB ID: 1Q8J). Hcy domain in slate, CH_3-H_4folate domain in green, and inter-domain linker in red ribbons. Hcy in violet spheres, except the sulfur, which accepts the transferred methyl group, is yellow. Zinc shown as a cyan sphere. CH_3-H_4folate in dark green spheres, except the methyl group to be transferred, in orange. **B**, Close up of the large domain interface, which is partly formed by the linker. Same depiction and orientation as **A**, except protein surfaces are shown. **C**, Superposition of the *T. maritima* MetH CH_3-H_4folate domain (green ribbons, CH_3-H_4folate in green sticks) with a *M. thermoacetica* MeTr monomer (PDB ID: 2E7F, rmsd: 1.09 Å) in pink ribbons with CH_3-H_4folate in red sticks.

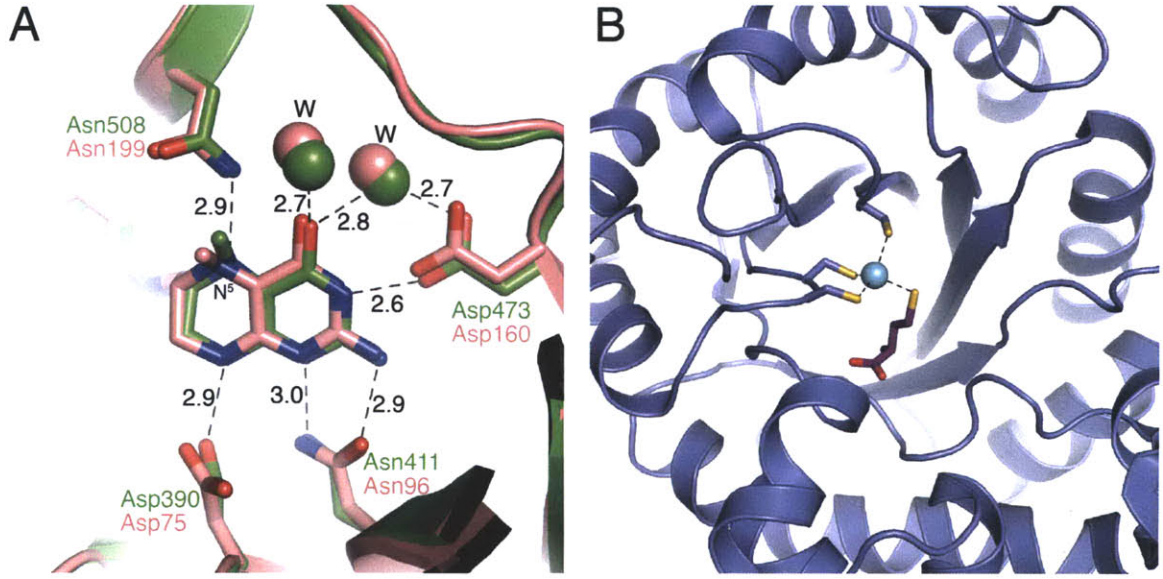


Figure VI.9. The CH_3-H_4folate and Hcy active sites for methyl transfer. A, The conserved methyl-donating CH_3-H_4folate active site. *E. coli* MetH CH_3-H_4folate domain (from PDB ID: 1Q8J) in green ribbons, and *M. thermoacetica* MeTr monomer (PDB ID: 2E7F) in pink ribbons. CH_3-H_4folate and residues surrounding the pterin ring are shown in sticks (N in blue, O in red, and C following the protein coloring). Residues are labeled in green and pink for MetH and MeTr sequences, respectively. Conserved water molecules ("W") are shown in spheres. Distances in Å derived from the MetH structure. **B**, The methyl-accepting Hcy active site from *E. coli* MetH in slate ribbons (from PDB ID: 1Q8J). Hcy in sticks (C in violet, S in yellow, O in red, N in blue), zinc as a cyan sphere, and cysteines that coordinate the zinc ion in sticks (C in slate and S in yellow).

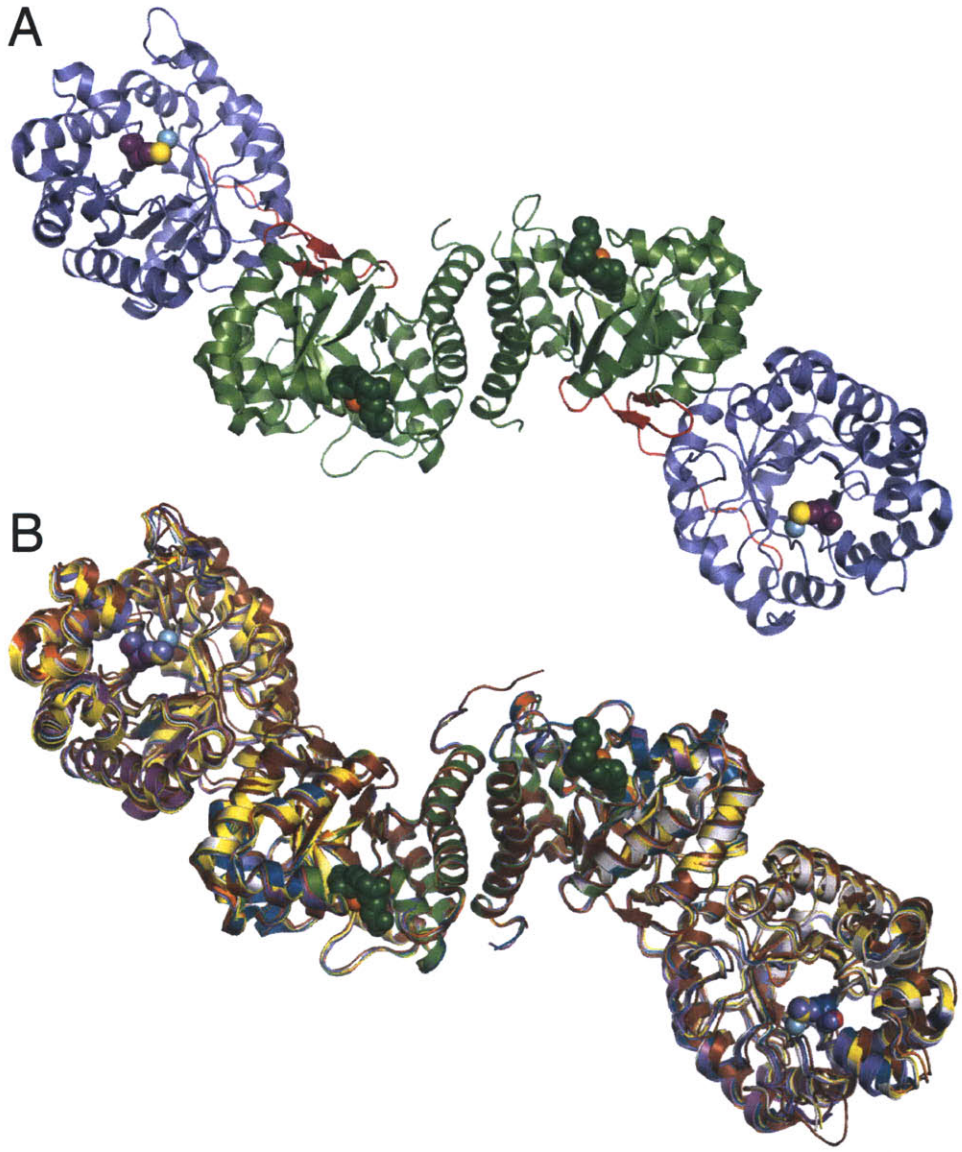


Figure VI.10. *T. maritima* MetH N-terminal Hcy/CH_3-H_4folate di-domain assembly. **A**, Structural components of one asymmetric unit within a *T. maritima* Hcy/CH_3-H_4folate di-domain (PDB ID: 1Q8J). Same coloring as Figure VI.7.A. **B**, Superimposable alignment of all *T. maritima* Hcy/CH_3-H_4folate di-domains to the structure shown in **A**. By PDB ID: 1Q85 in orange (rmsd: 0.22 Å), 1Q7Z in grey (rmsd: 0.28 Å), 1Q8A in slate with Hcy in slate spheres (rmsd: 0.23 Å), 1Q7M in yellow (rmsd: 0.33 Å), 1Q7Q in brown (rmsd: 0.66 Å), 3BOL in teal with Hcy in teal spheres (rmsd: 0.31 Å), 3BOF in magenta with Hcy in magenta spheres (rmsd: 0.25 Å).

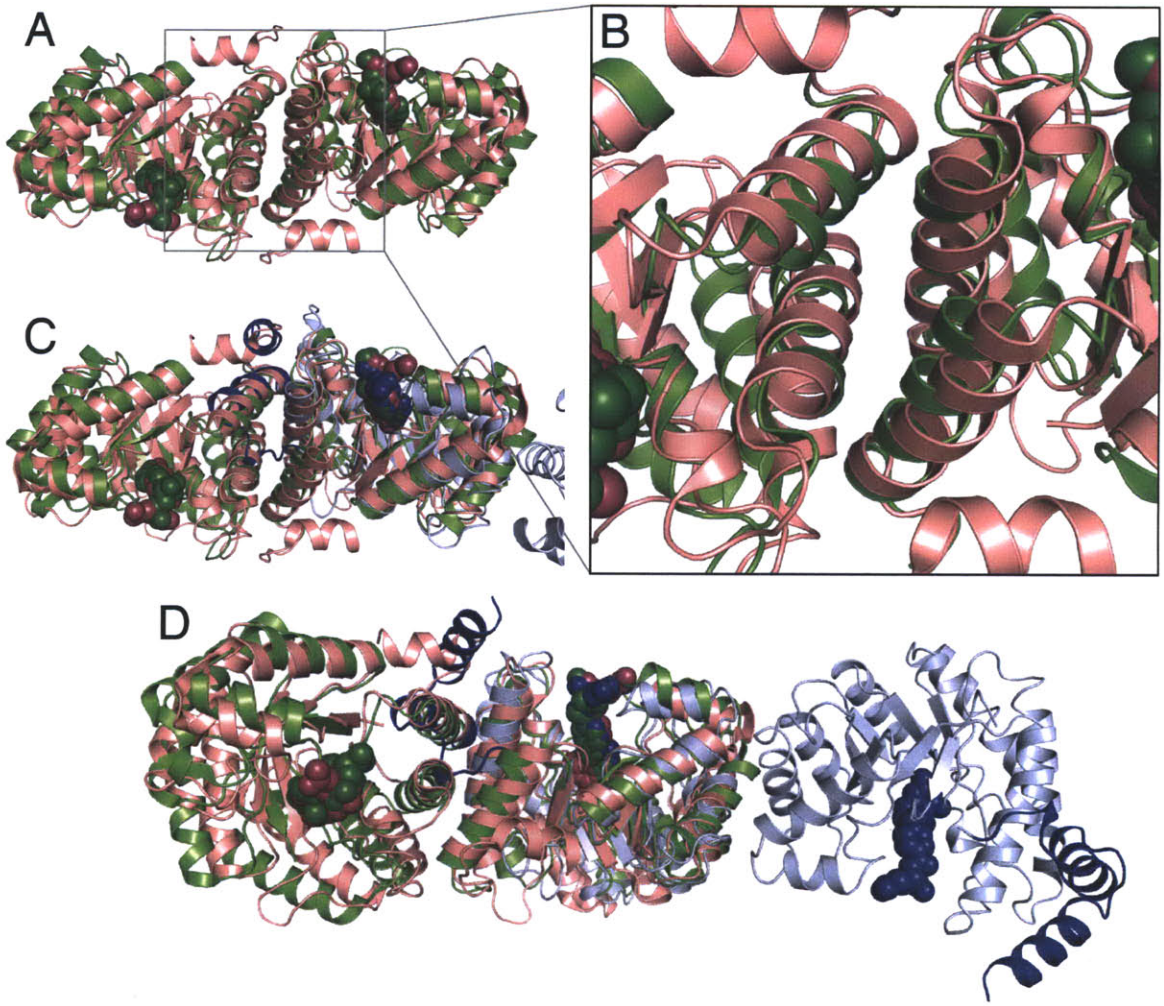


Figure VI.11. Structural comparisons of CH_3-H_4folate domain dimers. **A**, Superposition of the CH_3-H_4folate-binding domain ncs homodimers of *T. maritima* MetH (PDB ID: 1Q8J, green ribbons, CH_3-H_4folate as dark green spheres) and *M. thermoacetica* MeTr (PDB ID: 2E7F, light pink ribbons, CH_3-H_4folate as dark pink spheres), with rmsd: 3.32 Å. The grey square indicates the close-up image in **B**, a zoom-in of the MeTr and MetH CH_3-H_4folate domain homodimeric interfaces. **C**, Same orientation and coloring as **A**, plus the structure of the CH_3-H_4folate-binding domain of *B. thetaiotaomicron* MetH (PDB ID: 3K13, rmsd 1.60 Å), shown in light blue ribbons, with the additional C-terminal helices in dark blue ribbons and CH_3-H_4folate as dark blue spheres. **D**, Same structures and coloring as **C**, expanded to show the second *B. thetaiotamicron* MetH CH_3-H_4folate domain monomer, and in a slightly altered orientation.

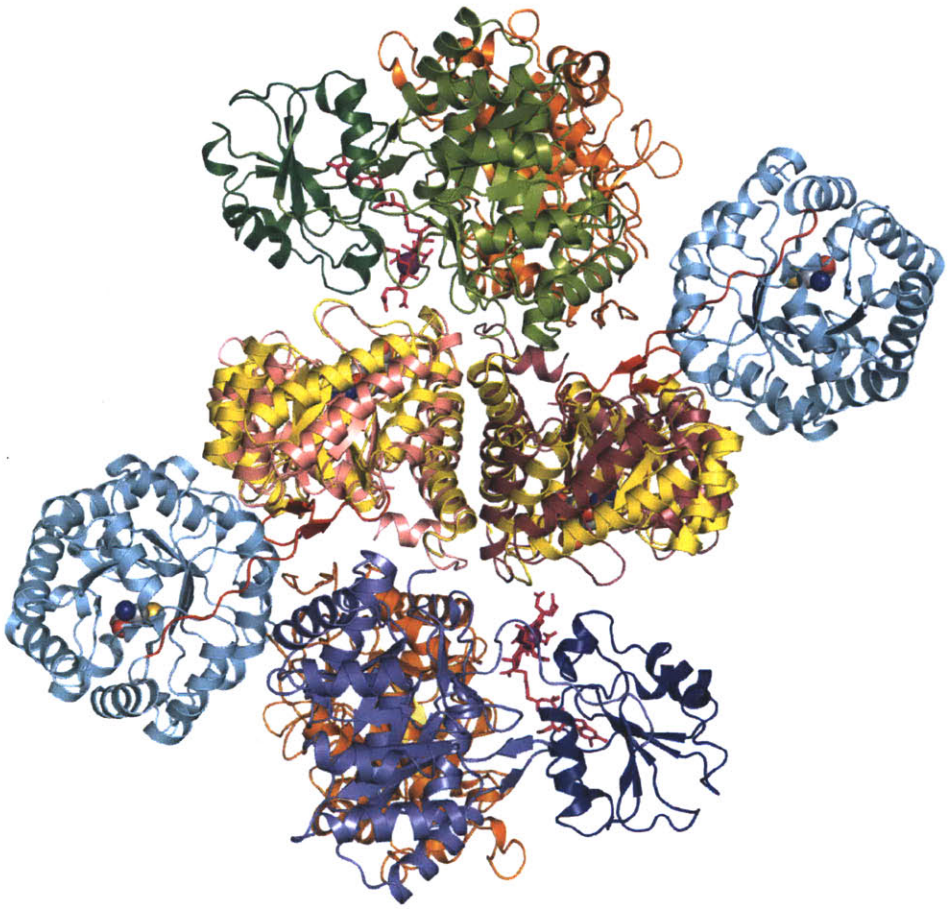


Figure VI.12. Alignment of CFeSP/MeTr with the *T. maritima* MetH N-terminal di-domain. Structure of the CFeSP/MeTr complex aligned to the *T. maritima* MetH Hcy/CH_3-H_4folate N-terminal di-domain structure (PDB ID: 1Q8J) by their CH_3-H_4folate binding domains (rmsd 3.38 Å). For CFeSP/MeTr, same orientation and coloring as Figure II.2, with Fe_4S_4 domains hidden for clarity. For MetH, CH_3-H_4folate domains in yellow ribbons, CH_3-H_4folate domains in cyan ribbons, and linkers in red ribbons. The Hcy and CH_3-H_4folate are shown in spheres: C in grey, N in blue, O in red, S in yellow. The alignment does not place MetH Hcy domains in the same region as any components of the CFeSP/MeTr structure.

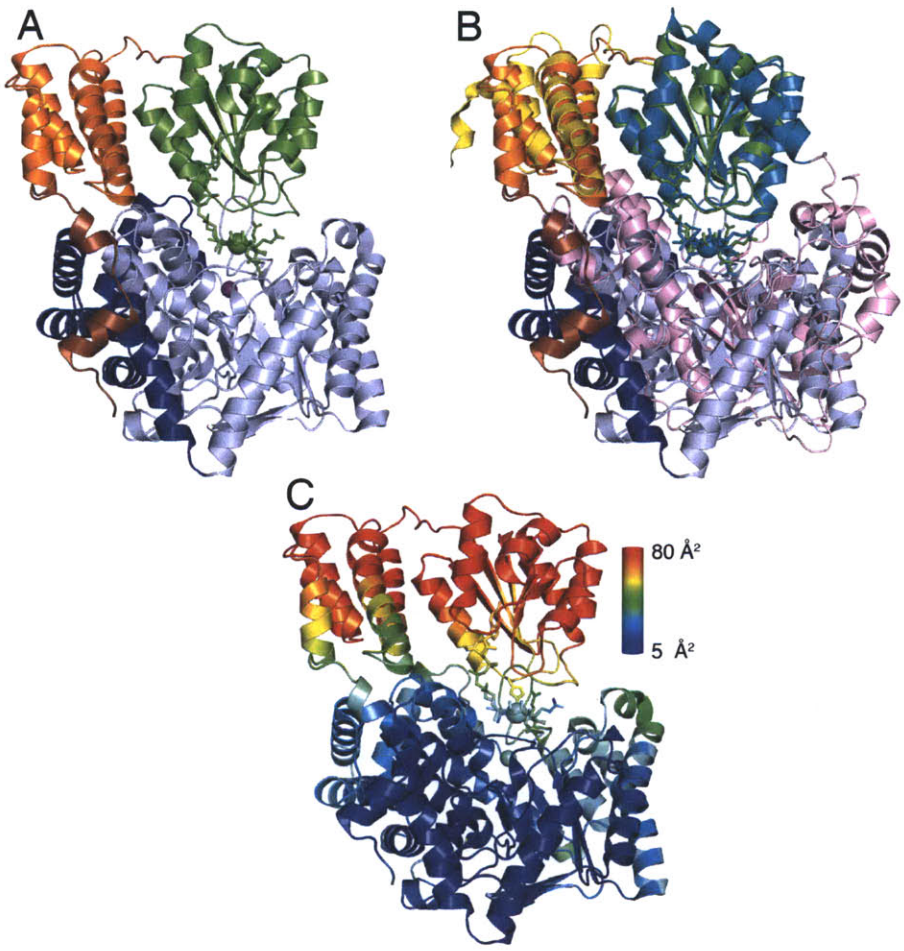


Figure VI.13. Structure of the MtaBC complex involved in methanol metabolism. A, Crystal structure of *M. barkeri* MtaBC in ribbons (PDB ID: 2I2X). MtaB: TIM barrel in light blue, C-terminal helices that wrap the barrel in dark blue, and zinc as a violet sphere. MtaC: N-terminal extension in brown, "capping" subdomain in orange, B_{12}-binding domain in green, and B_{12} domain in green, with B_{12} in green sticks with cobalt as a green sphere. **B**, Alignment of **A** with the *E. coli* MetH C-terminal B_{12}/AdoMet di-domain in ribbons (PDB ID: 1K98) by their B_{12} domains superimposes their "capping" subdomains. Same orientation and MtaBC coloring as **A**. MetH coloring: "capping" subdomain in yellow, B_{12}-binding domain in teal, and AdoMet domain in pink ribbons, with B_{12} in teal sticks and cobalt as a teal sphere. **C**, Crystal structure of MtaBC shown in ribbons with the same orientation as **A** and colored by *B*-factor, as labeled.

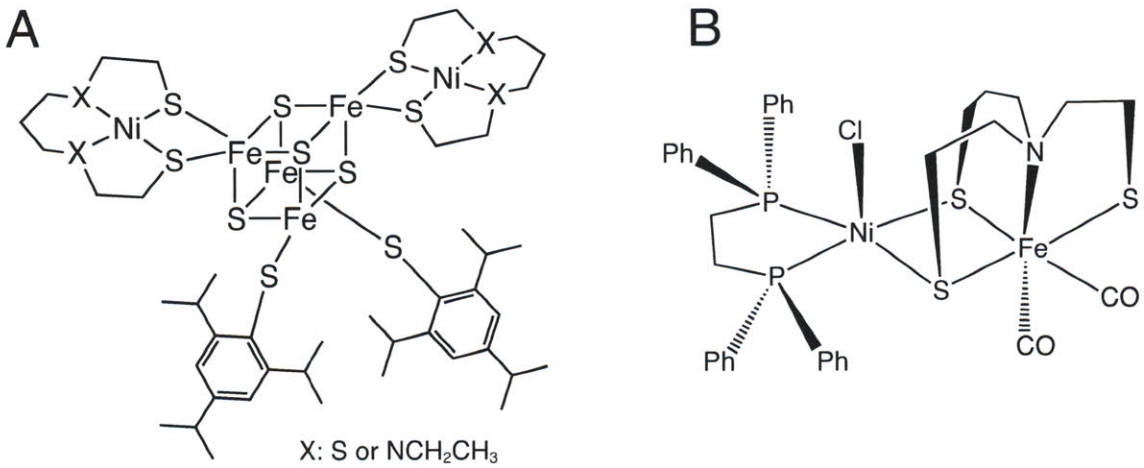


Figure VI.14. Examples of C-cluster model complexes prior to the CODH structural determination. **A**, Structure of a [4Fe-4S] cubane bridged to two nickel sites. **B**, structure of a binuclear iron-nickel complex that could bind CO.

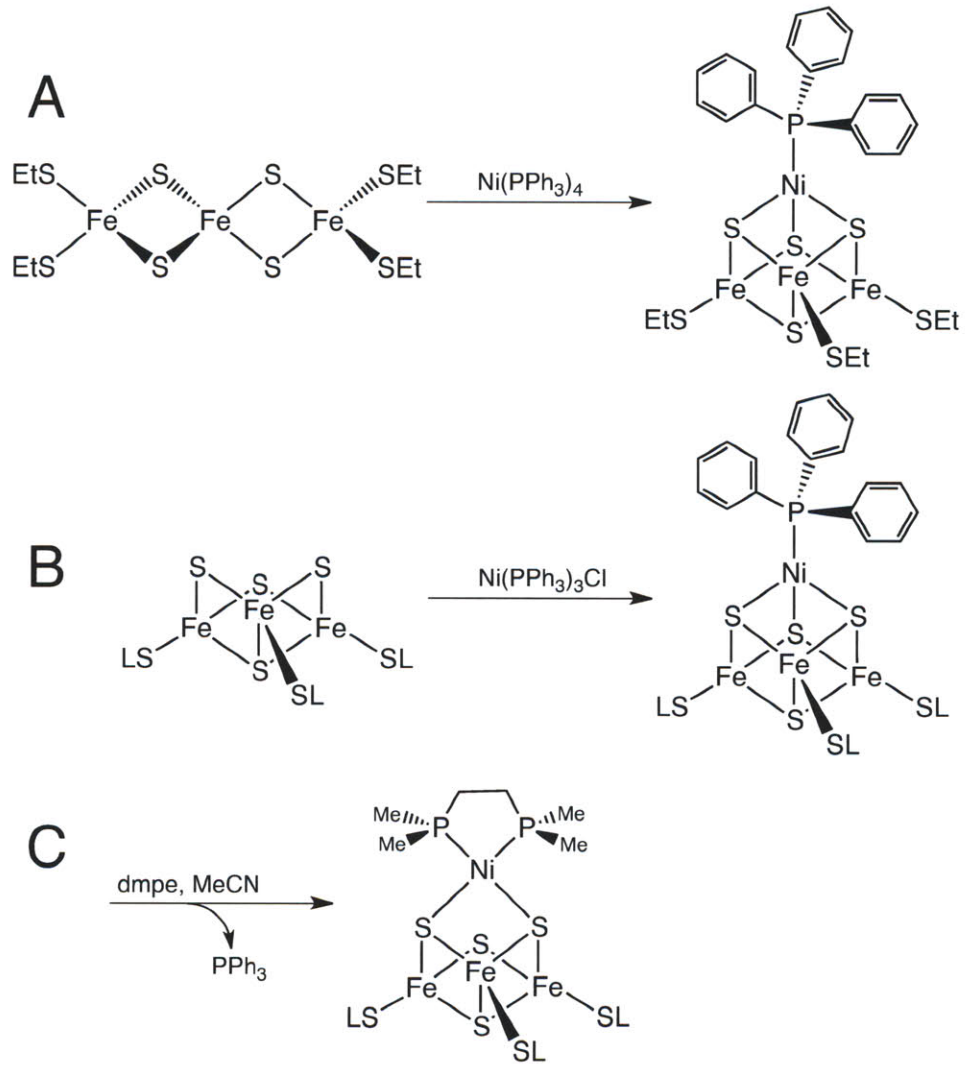


Figure VI.15. Examples of C-cluster model complexes that incorporated nickel into the cubane. A, Construction of a [Ni-3Fe-4S] cubane from a linear triiron species or **B**, from a preassembled [3Fe-4S] cubane, prior to the knowledge of the C-cluster structure. **C**, Continuation of **B** to produce a square planar nickel site and a distorted cubane, where nickel has moved further from one of the cubane sulfides. In **B** and **C**, "L" represents 1,3,5-tris((4,6-dimethyl-3-mercaptophenyl)thio)-2,4,6-tris(p-tolylthio)benzene, a ligand that coordinates the cluster in a tridentate binding mode. In **C**, dmpe represents 1,2-bis(dimethylphosphino)ethane, a bidentate ligand.

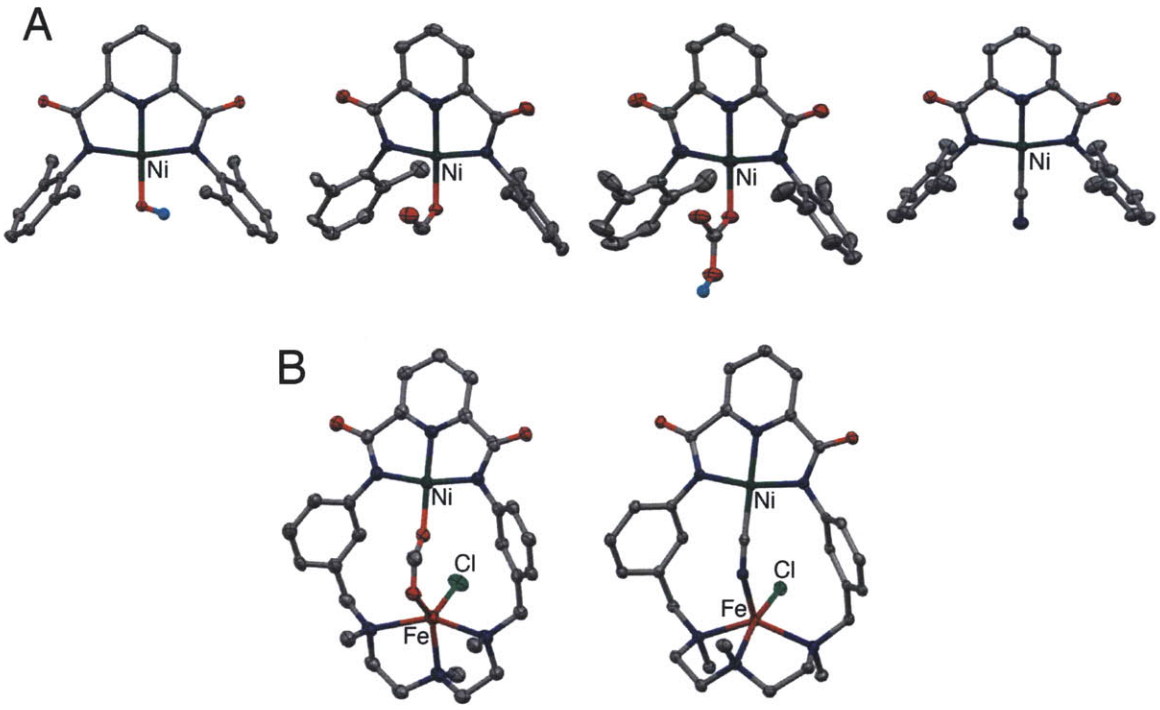


Figure VI.16. C-cluster model complexes that bind CODH substrates and inhibitors. **A**, Crystal structures of mononuclear nickel complexes bound with hydroxide, CO_2, formate, and cyanide. **B**, Crystal structures of binuclear iron-nickel complexes bound with CO_2 and cyanide. Ni, Fe, and Cl are labeled. Figures modified from previously published images[55].

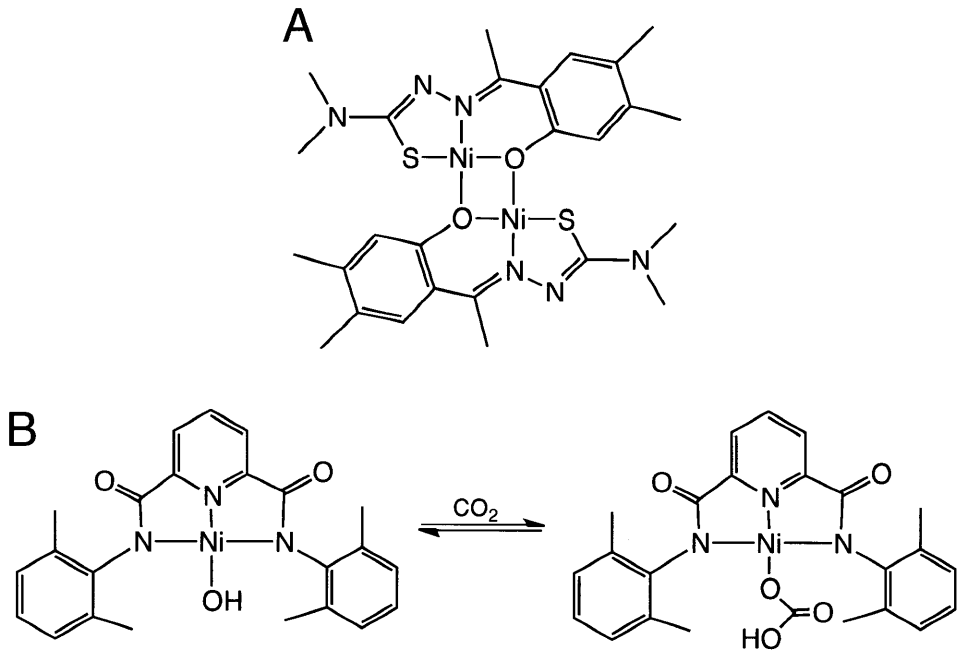


Figure VI.17. Catalytically active model complexes inspired by the CODH C-cluster. **A**, A binuclear Ni-Ni complex that can perform CO oxidation using water at room temperature in aqueous solution. **B**, Reaction of a mononuclear nickel complex that can convert CO_2 to bicarbonate.

VI.G. References

1 Svetlitchnaia, T., Svetlitchnyi, V., Meyer, O. & Dobbek, H. (2006) Structural Insights into Methyltransfer Reactions of a Corrinoid Iron-Sulfur Protein Involved in Acetyl-CoA Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14331-14336.

2 Menon, S. & Ragsdale, S. W. (1998) Role of the [4Fe-4S] Cluster in Reductive Activation of the Cobalt Center of the Corrinoid Iron-Sulfur Protein from *Clostridium thermoaceticum* during Acetate Biosynthesis. *Biochemistry.* **37**, 5689-5698.

3 Menon, S. & Ragsdale, S. W. (1999) The Role of an Iron-Sulfur Cluster in an Enzymatic Methylation Reaction. *J. Biol. Chem.* **274**, 11513-11518.

4 Ragsdale, S. W., Lindahl, P. A. & Münck, E. (1987) Mössbauer, EPR, and Optical Studies of the Corrinoid/Iron-Sulfur Protein Involved in the Synthesis of Acetyl Coenzyme A by *Clostridium thermoaceticum*. *J. Biol. Chem.* **262**, 14289-14297.

5 Seravalli, J., Zhao, S. & Ragsdale, S. W. (1999) Mechanism of Transfer of the Methyl Group from (*6S*)-Methytetrahydrofolate to the Corrinoid/Iron-Sulfur Protein Catalyzed by the Methyltransferase from *Clostridium thermoaceticum*: A Key Step in the Wood-Ljungdahl Pathway of Acetyl-CoA Synthesis. *Biochemistry.* **38**, 5728-5735.

6 Stich, T. A., Seravalli, J., Venkateshrao, S., Spiro, T. G., Ragsdale, S. W. & Brunold, T. C. (2006) Spectroscopic Studies of the Corrinoid/Iron-Sulfur Protein from *Moorella thermoacetica*. *J. Am. Chem. Soc.* **128**, 5010-5020.

7 Zhao, S., Roberts, D. L. & Ragsdale, S. W. (1995) Mechanistic Studies of the Methyltransferase from *Clostridium thermoaceticum*: Origin of the pH Dependence of the Methyl Group Transfer from Methyltetrahydrofolate to the Corrinoid/Iron-Sulfur Protein. *Biochemistry.* **34**, 15075-15083.

8 Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G. & Ludwig, M. L. (1994) How a Protein Binds B_{12}: A 3.0 Å X-ray Structure of B_{12}-Binding Domains of Methionine Sythase. *Science.* **266**, 1669-1674.

9 Dixon, M. M., Huang, S., Matthews, R. G. & Ludwig, M. L. (1996) The Structure of the C-terminal Domain of Methionine Synthase: Presenting *S*-adenosylmethionine for Reductive Methylation of B_{12}. *Structure.* **4**, 1263-1275.

10 Bandarian, V., Pattridge, K. A., Lennon, B. W., Huddler, D. P., Matthews, R. G. & Ludwig, M. L. (2002) Domain Alternation Switches B_{12}-Dependent Methionine Synthase to the Active Conformation. *Nat. Struct. Biol.* **9**, 53-56.

11 Evans, J. C., Huddler, D. P., Hilgers, M. T., Romanchuk, G., Matthews, R. G. & Ludwig, M. L. (2004) Structures of the N-terminal Module Imply Large Domain

Motions During Catalysis by Methionine Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3729-3736.

12 Wolthers, K. R., Toogood, H. S., Jowitt, T. A., Marshall, K. R., Leys, D. & Scrutton, N. S. (2007) Crystal Structure and Solution Characterization of the Activation Domain of Human Methionine Synthase. *FEBS J.* **274**, 738-750.

13 Datta, S., Koutmos, M., Pattridge, K. A., Ludwig, M. L. & Matthews, R. G. (2008) A Disulfide-Stabilized Conformer of Methionine Synthase Reveals an Unexpected Role for the Histidine Ligand of the Cobalamin Cofactor. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4115-4120.

14 Koutmos, M., Pejchal, R., Bomer, T. M., Matthews, R. G., Smith, J. L. & Ludwig, M. L. (2008) Metal Active Site Elasticity Linked to Activation of Homocysteine in Methionine Synthases. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3286-3291.

15 Koutmos, M., Datta, S., Pattridge, K. A., Smith, J. L. & Matthews, R. G. (2009) Insights into the Reactivation of Cobalamin-Dependent Methionine Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18527-18532.

16 Ljungdahl, L. G., LeGall, J. & Lee, J.-P. (1973) Isolation of a Protein Containing Tightly Bound 5-Methoxybenzimidazolylcobamide (Factor IIIm) from *Clostridium thermoaceticum*. *Biochemistry.* **12**, 1802-1808.

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

18 Wang, S. C. & Frey, P. A. (2007) *S*-Adenosylmethionine as an Oxidant: the Radical SAM Superfamily. *Trends Biochem. Sci.* **32**, 101-110.

19 Frey, P. A., Hegeman, A. D. & Ruzicka, F. J. (2008) The Radican SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol.* **43**, 63-88.

20 Frey, P. A. & Magnusson, O. T. (2003) *S*-Adenosylmethionine: A Wolf in Sheep's Clothing, or a Rich Man's Adenosylcobalamin? *Chem. Rev.* **103**, 2129-2148.

21 Vey, J. L. & Drennan, C. L. (2011) Structural Insights into Radical Generation by the Radical SAM Superfamily. *Chem. Rev.*, doi:10.1021/cr9002616.

22 Hagemeier, C. H., Krüer, M., Thauer, R. K., Warkentin, E. & Ermler, U. (2006) Insight into the Mechanism of Biological Methanol Activation Based on the Crystal Structure of the Methanol-Cobalamin Methyltransferase Complex. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18917-18922.

23 Doukov, T., Seravalli, J., Stezowski, J. J. & Ragsdale, S. W. (2000) Crystal Structure of a Methyltetrahydrofolate- and Corrinoid-Dependent Methyltransferase. *Structure.* **8**, 817-830.

24 Doukov, T. I., Hemmi, H., Drennan, C. L. & Ragsdale, S. W. (2007) Structural and Kinetic Evidence for an Extended Hydrogen-bonding Network in Catalysis of Methyl Group Transfer: Role of an Active Site Asparagine Residue in Activation of Methyl Transfer by Methyltransferases. *J. Biol. Chem.* **282**, 6609-6618.

25 Jarrett, J. T., Drennan, C. L., Amaratunga, M., Scholten, J. D., Ludwig, M. L. & Matthews, R. G. (1996) A Protein Radical Cage Slows Photolysis of Methylcobalamin in Methionine Synthase from *Escherichia coli*. *Bioorgan. Med. Chem.* **4**, 1237-1246.

26 Jarrett, J. T., Amaratunga, M., Drennan, C. L., Scholten, J. D., Sands, R. H., Ludwig, M. L. & Matthews, R. G. (1996) Mutations in the B_{12}-Binding Region of Methionine Synthase: How the Protein Controls Methylcobalamin Reactivity. *Biochemistry.* **35**, 2464-2475.

27 Hoover, D. M., Jarrett, J. T., Sands, R. H., Dunham, W. R., Ludwig, M. L. & Matthews, R. G. (1997) Interaction of *Escherichia coli* Cobalamin-Dependent Methionine Synthase and Its Physiological Partner Flavodoxin: Binding of Flavodoxin Leads to Axial Ligand Dissocaition from the Cobalamin Cofactor. *Biochemistry.* **36**, 127-138.

28 Frasca, V., Banerjee, R. V., Dunham, W. R., Sands, R. H. & Matthews, R. G. (1988) Cobalamin-Dependent Methionine Synthase from *Escherichia coli* B: Electron Paramagnetic Resonance Spectra of the Inactive Form and the Active Methylated From of the Enzyme. *Biochemistry.* **27**, 8458-8465.

29 Lewandowicz, A., Ringia, E. A. T., Ting, L.-M., Kim, K., Tyler, P. C., Evans, G. B., Zubkova, O. V., Mee, S., Painter, G. F., Lenz, D. H., Furneaux, R. H. & Schramm, V. L. (2005) Energetic Mapping of Transition State Analogue Interactions with Human and *Plasmodium falciparum* Purine Nucleoside Phosphorylases. *J. Biol. Chem.* **280**, 30320-30328.

30 Kicska, G. A., Tyler, P. C., Evans, G. B., Furneaux, R. H., Shi, W., Fedorov, A., Lewandowicz, A., Cahill, S. M., Almo, S. C. & Schramm, V. L. (2002) Atomic Dissection of the Hydrogen Bond Network for Transition-State Analogue Binding to Purine Nucleoside Phosphorylase. *Biochemistry.* **41**, 14489-14498.

31 Fedorov, A., Shi, W., Kicska, G., Tyler, P. C., Furneaux, R. H., Hanson, J. C., Gainsford, G. J., Larese, J. Z., Schramm, V. L. & Almo, S. C. (2001) Transition State Structure of Purine Nucleoside Phosphorylase and Principles of Atomic Motion in Enzymatic Catalysis. *Biochemistry.* **40**, 853-860.

32 Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003) Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] Clusters in Closed and Open Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. *Nat. Struct. Biol.* **10**, 271-279.

33 Doukov, T. I., Blasiak, L. C., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2008) Xenon in and at the End of the Tunnel of Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **47**, 3474-3483, doi:10.1021/bi702386t.

34 Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002) A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Science.* **298**, 567-572.

35 Kung, Y., Doukov, T. I., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2009) Crystallographic Snapshots of Cyanide- and Water-Bound C-Clusters from Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **48**, 7432-7440.

36 Svetlitchnyi, V., Dobbek, H., Meyer-Klaucke, W., Meins, T., Thiele, B., Römer, P., Huber, R. & Meyer, O. (2004) A Functional Ni-Ni-[4Fe-4S] Cluster in the Monomeric Acetyl-CoA Synthase from *Carboxydothermus hydrogenoformans*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 446-451.

37 Ragsdale, S. W. (2008) Enzymology of the Wood-Ljungdahl Pathway of Acetogenesis. *Ann. N.Y. Acad. Sci.* **1125**, 129-136.

38 Ragsdale, S. W. & Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 Fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898.

39 Drake, H. L., Hu, S.-I. & Wood, H. G. (1981) Purification of Five Components from *Clostridium thermoaceticum* Which Catalyze Synthesis of Acetate from Pyruvate and Methyltetrahydrofolate: Properties of Phosphotransacetylase. *J. Biol. Chem.* **256**, 11137-11144.

40 Zhai, Y., Pierre, D., Si, R., Deng, W., Ferrin, P., Nilekar, A. U., Peng, G., Herron, J. A., Bell, D. C., Saltsburg, H., Mavrikakis, M. & Flytzani-Stephanopoulos, M. (2010) Alkali-Stabilized Pt-OH_x Species Catalyze Low-Temperature Water-Gas Shift Reactions. *Science.* **329**, 1633-1636.

41 Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P. & Sutton, M. A. (2008) Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science.* **320**, 889-892.

42 Fu, Q., Saltzburg, H. & Flytzani-Stephanopoulos, M. (2003) Active Nonmetallic Au and Pt Species on Ceria-Based Water-Gas Shift Catalysts. *Science.* **301**, 935-938.

43 Burch, R. (2006) Gold Catalysts for Pure Hydrogen Production in the Water-Gas Shift Reaction: Activity, Structure and Reaction Mechanism. *Phys. Chem. Chem. Phys.* **8**, 5483-5500.

44 Gokhale, A. A., Dumesic, J. A. & Mavrikakis, M. (2008) On the Mechanism of Low-Temperature Water Gas Shift Reaction on Copper. *J. Am. Chem. Soc.* **130**, 1402-1414.

45 Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, O. (2001) Crystal Structure of a Carbon Monoxide Dehydrogenase Reveals a [Ni-4Fe-5S] Cluster. *Science.* **293**, 1281-1285.

46 Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. & Ludden, P. W. (2001) Life on Carbon Monoxide: X-ray Structure of *Rhodospirillum rubrum* Ni-Fe-S Carbon Monoxide Dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11973-11978.

47 Osterloh, F., Saak, W. & Pohl, S. (1997) Unidentate and Bidentate Binding of Nickel(II) Complexes to an Fe_4S_4 Cluster *via* Bridging Thiols: Synthesis, Crystal Structures, and Electrochemical Properties of Model Compounds for Active Sites of Nickel Containing CO Dehydrogenase/Acetyl-CoA Synthase. *J. Am. Chem. Soc.* **119**, 5648-5656.

48 Osterloh, F., Saak, W., Haase, D. & Pohl, S. (1996) Nickel(II) Complexes Bound to an $[Fe_4S_4]$ Cluster *via* Bridging Thiolates: Synthesis and Crystal Structures of Model Compounds for the Active Site of Nickel CO Dehydrogenase. *Chem. Commun.*, 777-778.

49 Smith, M. C., Longhurst, S., Barclay, J. E., Cramer, S. P., Davies, S. C., Hughes, D. L., Gu, W.-W. & Evans, D. J. (2001) Nickel(II) in an Almost Regular Tetrahedral Thiolate Environment: A First Generation Synthetic Analogue of the Active Site of Nickel CO-Dehydrogenase. *J. Chem. Soc., Dalton Trans.*, 1387-1388.

50 Smith, M. C., Barclay, J. E., Cramer, S. P., Davies, S. C., Gu, W.-W., Hughes, D. L., Longhurst, S. & Evans, D. J. (2002) Nickel-Iron-Sulfur Complexes: Approaching Structural Analogues of the Acvite Sites of NiFe-Hydrogenase and Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *J. Chem. Soc., Dalton Trans.*

51 Evans, D. J. (2005) Chemistry Relating to the Nickel Enzymes CODH and ACS. *Coordin. Chem. Rev.* **249**, 1582-1595.

52 Ciurli, S., Ross, P. K., Scott, M. J., Yu, S.-B. & Holm, R. H. (1992) Synthetic Nickel-Containing Heterometal Cubane-Type Clusters with $NiFe_3Q_4$ Cores (Q = S, Se). *J. Am. Chem. Soc.* **114**, 5415-5423.

53 Zhou, J., Raebiger, J. W., Crawford, C. A. & Holm, R. H. (1997) Metal Ion Incorproation Reactions of the Cluster $[Fe_3S_4(LS_3)]^{3-}$, Containing the Cuboidal $[Fe_3S_4]^0$ Core. *J. Am. Chem. Soc.* **119**, 6242-6250.

54 Panda, R., Zhang, Y., McLauchlan, C. C., Rao, P. V., Oliveira, F. A. T. d., Münck, E. & Holm, R. H. (2004) Initial Structure Modification of Tetrahedral to Planar Nickel(II) in a Nickel-Iron-Sulfur Cluster Related to the C-Cluster of Carbon Monoxide Dehydrogenase. *J. Am. Chem. Soc.* **126**, 6448-6459.

55 Huang, D. & Holm, R. H. (2010) Reactions of the Terminal Ni^{II}-OH Group in Substitution and Electrophilic Reactions with Carbon Dioxide and Other Substrates: Structural Definition of Binding Modes in an Intramolecular Ni^{II}---Fe^{II} Bridged Site. *J. Am. Chem. Soc.* **132**, 4693-4701.

56 Lu, Z. & Crabtree, R. H. (1995) A Functional Modeling Study of the CO Oxidation Site of Nickel CO Dehydrogenase. *J. Am. Chem. Soc.* **117**, 3994-3998.

57 Beley, M., Collin, J.-P., Ruppert, R. & Sauvage, J.-P. (1984) Nickel(II)-Cyclam: An Extremely Selective Electrocatalyst for Reduction of CO_2 in Water. *J. Chem. Soc., Chem. Commun.*, 1315-1316.

58 Beley, M., Collin, J.-P., Ruppert, R. & Sauvage, J.-P. (1986) Electrocatalytic Reduction of CO_2 by Ni $Cyclam^{2+}$ in Water: Study of the Factors Affecting the Efficiency and the Selectivity of the Process. *J. Am. Chem. Soc.* **108**, 7461-7467.

59 Zagal, J. H. (1992) Metallophthalocyanines as Catalysts in Electrochemical Reactions. *Coordin. Chem. Rev.* **119**, 89-136.

60 Tommasi, I., Aresta, M., Giannoccaro, P., Quaranta, E. & Fragale, C. (1998) Bioinorganic Chemistry of Nickel and Carbon Dioxide: An Ni Complex Behaving as a Model System for Carbon Monoxide Dehydrogenase Enzyme. *Inorg. Chim. Acta.* **272**, 38-42.

61 Aresta, M., Quaranta, E. & Tommasi, I. (1988) Reduction of Co-ordinated Carbon Dioxide to Carbon Monoxide *via* Protonation by Thiols and Other Brønsted Acids Promoted by Ni-Systems: A Contribution to the Understanding of the Mode of Action of the Enzyme Carbon Monoxide Dehydrogenase. *J. Chem. Soc., Chem. Commun.*, 450-452.

62 Huang, D., Makhlynets, O. V., Tan, L. L., Lee, S. C., Rybak-Akimova, E. V. & Holm, R. H. (2011) Kinetics and Mechanistic Analysis of an Extremely Rapid Carbon Dioxide Fixation Reaction. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1222-1227.

63 Lackner, K. S. (2003) A Guide to CO_2 Sequestration. *Science.* **300**, 1677-1678.

64 Keith, D. W. (2009) Why Capture CO_2 from the Atmosphere? *Science.* **325**, 1654-1655.

65 Goris, T., Wait, A. F., Saggu, M., Fritsch, J., Heidary, N., Stein, M., Zebger, I., Lendzian, F., Armstrong, F. A., Friedrich, B. & Lenz, O. (2011) A Unique Iron-Sulfur Cluster is Crucial for Oxygen Tolerance of a [NiFe]-Hydrogenase. *Nat. Chem. Biol.*, doi:10.1038/nchembio.555.

66 Liebgott, P.-P., Leroux, F., Burlat, B., Dementin, S., Baffert, C., Lautier, T., Fourmond, V., Ceccaldi, P., Cavazza, C., Meynial-Salles, I., Soucaille, P., Fontecilla-Camps, J. C., Guigliarelli, B., Bertrand, P., Rousset, M. & Léger, C. (2009) Relating Diffusion Along the Substrate Tunnel and Oxygen Sensitivity in Hydrogenase. *Nat. Chem. Biol.* **6**, 63-70.

67 Buhrke, T., Lenz, O., Krauss, N. & Friedrich, B. (2005) Oxygen Tolerance of the H_2-sensing [NiFe] Hydrogenase from *Ralstonia eutropha* H16 Is Based on Limited Access of Oxygen to the Active Site. *J. Biol. Chem.* **280**, 23791-23786.

68 Lazarus, O., Woolerton, T. W., Parkin, A., Lukey, M. J., Reisner, E., Seravalli, J., Pierce, E., Ragsdale, S. W., Sargent, F. & Armstrong, F. A. (2009) Water-Gas Shift Reaction Catalyzed by Redox Enzymes on Conducting Graphite Platelets. *J. Am. Chem. Soc.* **131**, 14154-14155.

69 Parkin, A., Seravalli, J., Vincent, K. A., Ragsdale, S. W. & Armstrong, F. A. (2007) Rapid and Efficient Electrocatalytic CO_2/CO Interconversions by *Carboxydothermus hydrogenoformans* CO Dehydrogenase I on an Electrode. *J. Am. Chem. Soc.* **129**, 10328-10329.

70 Cracknell, J. A., Vincent, K. A. & Armstrong, F. A. (2008) Enzymes as Working or Inspirational Electrocatalysts for Fuel Cells and Electrolysis. *Chem. Rev.* **108**, 2439-2461.

71 Menon, S. & Ragsdale, S. W. (1996) Unleashing Hydrogenase Activity in Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase and Pyruvate:Ferredoxin Oxidoreductase. *Biochemistry.* **35**, 15814-15821.

72 Kräutler, B. (2009) Organometallic Chemistry of B_{12} Coenzymes. *Met. Ions Life Sci.* **6**, 1-51.

Afoot and light-hearted I take to the open road!
Healthy, free, the world before me!
The long brown path before me, leading wherever I choose!

Walt Whitman

Appendix I. A putative sodium binding site in acetyl coenzyme A synthase

AI.A. Summary

A novel metal binding site was identified during the refinement and model building stages of the native and cyanide-bound crystal structures of carbon monoxide dehydrogenase/acetyl coenzyme A synthase (CODH/ACS) described in Chapter IV. Previously determined crystal structures of ACS also appear to contain this site, but because a water molecule was often modeled instead, it is likely that this site has remained unnoticed until now. The site exhibits octahedral geometry, with four carbonyl backbone ligands, one water ligand, and one carboxylate ligand from the side chain of a conserved glutamate residue. Refinement trials with several metals at the site indicate that sodium best fits the crystallographic data, based on the electron density and resulting bond lengths. Preliminary studies of mutant ACS enzymes, where the coordinating glutamate residue has been substituted for aspartate or alanine, suggest that the site is important for catalysis. Additional details regarding the specific role for the putative ACS sodium site await further biochemical experimentation.

AI.B. Acknowledgements

Mutagenesis, protein purification, and exchange assays were performed by Elizabeth Pierce in the laboratory of Stephen W. Ragsdale at the University of Michigan.

AI.C. Results and Discussion

Structural features of the putative sodium binding site

Acetyl coenzyme A synthase (ACS) is composed of three domains that are numbered 1-3 from N-terminus to C-terminus (Figure AI.1.A). A large cavity between the three domains has been hypothesized as the binding site for coenzyme A (CoA)[1], due to the presence of Trp418 and Trp427, which point towards the cavity and induce fluorescence changes upon CoA binding[2], and several arginine residues, whose modification by phenylglyoxal, methylglyoxal, or butanedione inhibits binding of CoA[3]. The active site A-cluster is housed in Domain 3 and is located near the central ACS cavity (Figure AI.1.A).

During the structural refinement of the *Moorella thermoacetica* bifunctional carbon monoxide dehydrogenase/ACS (CODH/ACS) complexes described in Chapter IV, a previously unnoticed putative metal binding site was observed in Domain 2 of all four ACS subunits present in the asymmetric unit. This metal site is located at a distance of 23 Å from the A-cluster on the opposite side of the ACS cavity (Figure AI.1.B). The metal is octahedrally coordinated by four backbone carbonyl ligands (from Phe328, Asn412, Gly414, and Leu416), a water molecule, and the carboxylate side chain of a conserved glutamate residue, Glu331 (Figure AI.2). The putative metal site had been modeled in prior structures of ACS as a water molecule; however, the presence of these six ligands indicates that assignment as a water molecule is incorrect, as water can form only a maximum of four hydrogen bonds, with two hydrogen bond donors and two hydrogen bond acceptors.

To identify the possible metal at this site, refinement trials with different metals was performed in Refmac[4], using metals that may have been present in the crystallization solution. Calcium acetate at 200 mM was included in the precipitant, but refinement of with calcium gave significant negative F_o-F_c difference electron density peaks (Figure AI.3.A) and ligand-metal distances that were slightly longer than average distances for calcium[5-7] (Table AI.1). These features indicated that calcium is a poor fit. Although refinement with magnesium, a smaller divalent Group 2 metal, eliminated F_o-F_c difference density around the metal (Figure AI.3.B), the refined ligand-metal distances were much longer than typical magnesium distances[5-7] (Table AI.1). When refinement with potassium in the site was conducted, strong negative F_o-F_c difference electron density peaks were produced (Figure AI.3.C), and ligand-metal distances were slightly longer than is typical with potassium[5-7] (Table AI.1). Refinement with the smaller monovalent Group 1 metal sodium eliminated difference electron density (Figure AI.3.D) and produced metal-ligand distances that were near the typical values for sodium[5-7] (Table AI.1). Based on these refinement trials, sodium, which was present during protein purification, produces the best fit to the data.

A search through previously solved crystal structures of ACS revealed that these structures also exhibit features consistent with the putative sodium site (Figure AI.4). All *M. thermoacetica* CODH/ACS structures contain the site (Figures AI.4.A-C), which has been modeled as a water molecule. These structures include both the "closed" and "open" ACS conformations, which are thought to represent states of the enzyme required for channeling of CO from CODH and binding of the corrinoid iron-sulfur protein (CFeSP), respectively[8,9]. Because the sodium site is present in both states, it appears that the site is

not significant for conformational switching between "closed" and "open" states. In addition, the putative sodium site is also present in the crystal structure of a truncated *M. thermoacetica* ACS (Figure AI.4.D), where Domain 1 has been completely removed. The drastically altered ACS construct used here exhibits a hexameric structure with six-fold symmetry, and the sodium sites are situated towards the axis of symmetry. Therefore, it would be difficult to hypothesize a function for the sodium site based on this structure of an ACS fragment that adopts such a non-physiological quaternary organization. The 2.20 Å resolution crystal structure of a monofunctional ACS from *Carboxydothermus hydrogenoformas* is currently the only available structure of ACS that is not from *M. thermoacetica*. Unlike all other ACS structures, this structure does not contain a water molecule modeled in the putative sodium site; however, the arrangement of backbone carbonyls and Glu334 (analogous to Glu331 in ACS from *M. thermoacetica*) is consistent with a metal binding site (Figure AI.4.E). Here, the carboxylate of Glu334 need only be rotated slightly to yield the required geometry.

Preliminary biochemical characterization

To test the effect of the putative sodium site on the function of ACS, mutants of the *M. thermoacetica* enzyme were prepared by site-directed mutagenesis, expressed recombinantly in *Escherichia coli*, and purified as previously described[10,11]. Of all the protein ligands to the metal site, only one is a side chain (Glu331), and thus this residue was mutated. Two mutants were prepared, E331D and E331A, and their activities were compared to that of the wild-type enzyme using the CO/acetyl-CoA exchange assay[3,12-14]. Briefly, radiolabeled [1-^{14}C]acetyl-CoA is incubated with the enzyme and unlabeled CO

(1 atm) in the anaerobic chamber. In the active enzyme, exchange of the hot carbonyl from [1-^{14}C]acetyl-CoA with cold CO can occur. Aliquots are removed at given time intervals and quenched with 0.1 M HCl. Following the addition of a liquid scintillation cocktail, remaining radioactivity is measured by scintillation counting and represents unreacted [1-^{14}C]acetyl-CoA.

Figure AI.5 shows the results of the preliminary CO/acetyl-CoA exchange assay. Specific activities for wild-type, E331D, and E331A mutants were 1.09, 0.44, and 0.06 U/mg, respectively. Because the specific activity is strongly affected by mutation of the side chain that coordinates the sodium site, these results suggest that the sodium site is important for some step of ACS catalysis. However, the precise role of the site cannot be identified by this assay alone.

Possible functions of the sodium site

All currently available crystal structures of ACS appear to contain the putative sodium site, yet a clear function for the site remains uncertain. It is too far from the active site A-cluster (>20 Å) to be directly involved in catalysis, and sodium may therefore simply play a structural role in ACS folding. However, the location of the sodium site directly adjacent to the putative CoA binding site suggests that sodium has a role in CoA binding. In fact, Trp418, a residue mentioned earlier that is thought to be important for binding of CoA, resides on the same loop as three residues whose backbone carbonyls are ligands to the metal: Asn412, Gly414, and Leu416. It is thus possible that the sodium ion is important for positioning tryptophan for binding with CoA.

The preliminary studies described above indicate that the activities of Glu331 ACS mutants are significantly reduced, where mutation to aspartate gives an over two-fold decrease in specific activity, and mutation to alanine more strongly compromises activity by over an order of magnitude. Additional mutations, such as a more conservative glutamine substitution, could be performed to further test the importance of the metal for activity. More importantly, however, binding studies must be performed to identify whether the observed changes in activity can be specifically attributed to CoA binding. In addition, the precise CoA binding site and the residues that contact CoA must be identified, and a crystal structure of ACS or CODH/ACS bound with CoA would be invaluable towards this goal.

Table AI.1. Refinement trials with different metals at the putative ACS sodium site. PDB and CSD refer to the Protein Data Bank and Cambridge Structural Database, respectively. Average distances were taken from previously compiled values[5-7].

	Calcium			Magnesium			Potassium			Sodium		
	PDB	CSD	Refined	PDB	CSD	Refined	PDB	CSD	Refined	PDB	CSD	Refined
Avg. distance to C=O (Å)	2.36	2.39	2.50	2.26	N/A	2.45	2.80	2.67	2.50	2.46	2.37	2.46
Avg. distance to Glu (Å)	2.33	2.38	2.55	2.08	2.05	2.53	2.80	2.82	2.63	2.30	2.41	2.54
Avg. distance to water (Å)	2.40	2.39	2.55	2.09	2.07	2.50	2.82	2.80	2.51	2.42	2.41	2.50
Negative F_o-F_c density?		yes			no			yes			no	

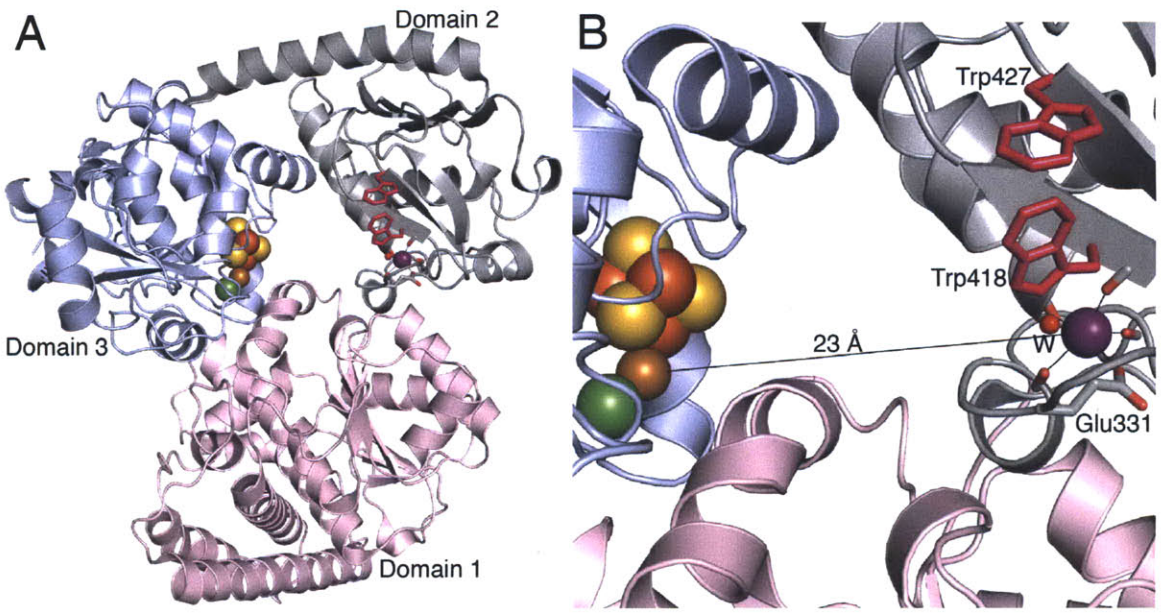


Figure AI.1. Overall structure of ACS. A, Domain arrangement enclosing a large central cavity, with Domains 1, 2, and 3 labeled and in light pink, grey, and light blue ribbons, respectively. The A-cluster is shown in spheres (Ni in green, Cu in brown, Fe in orange, and S in yellow). The putative sodium site opposite the cavity from the A-cluster is shown as a violet sphere with ligands in sticks (C in grey and O in red) and the water ligand as a red sphere. Tryptophan residues thought to be important for CoA binding are shown in dark pink sticks. **B**, Zoom in of the A-cluster and sodium site near the central ACS cavity. Same orientation and coloring as **A**.

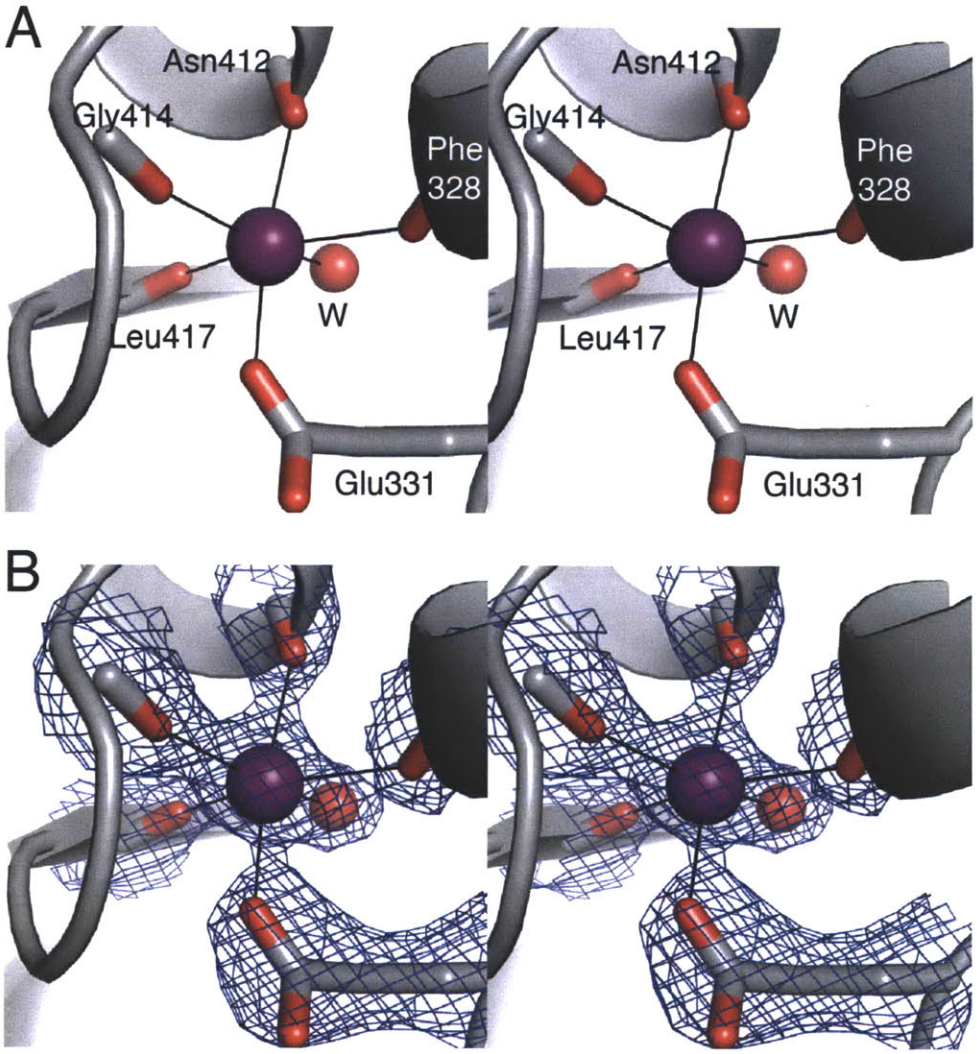


Figure AI.2. The putative sodium binding site. A, Stereoview of the putative sodium site from native CODH/ACS in grey ribbons (PDB ID: 3I01), with labeled residues in sticks (C in grey and O in red). The sodium ion is depicted as a violet sphere and the coordinating water molecule ("W") is shown as a red sphere. **B**, Stereoview of the same orientation and coloring as **A**, with the $2F_o$-F_c electron density map shown in blue mesh contoured at 1.5 σ.

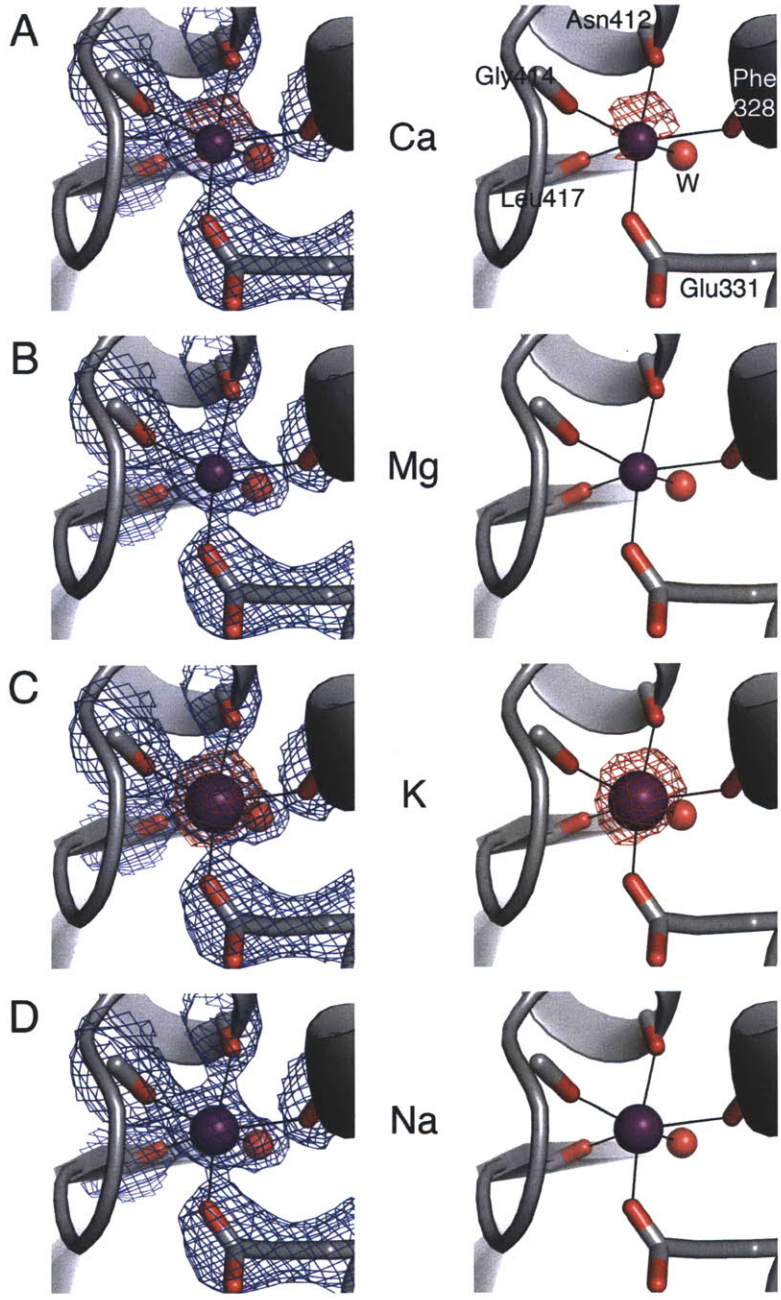


Figure AI.3. Refinement trials with different metals modeled. A, Calcium, **B**, magnesium, **C**, potassium, and **D**, sodium modeled into the new metal site. Same orientation and coloring as Figure AI.2. Left panels show $2F_o$-F_c density in blue mesh contoured at 1.5 and F_o-F_c difference density in red mesh at -3.0 σ, while right panels show only the F_o-F_c difference density. No positive F_o-F_c difference density was present for any structure, and negative F_o-F_c difference density was not observable in magnesium and sodium refinement trials.

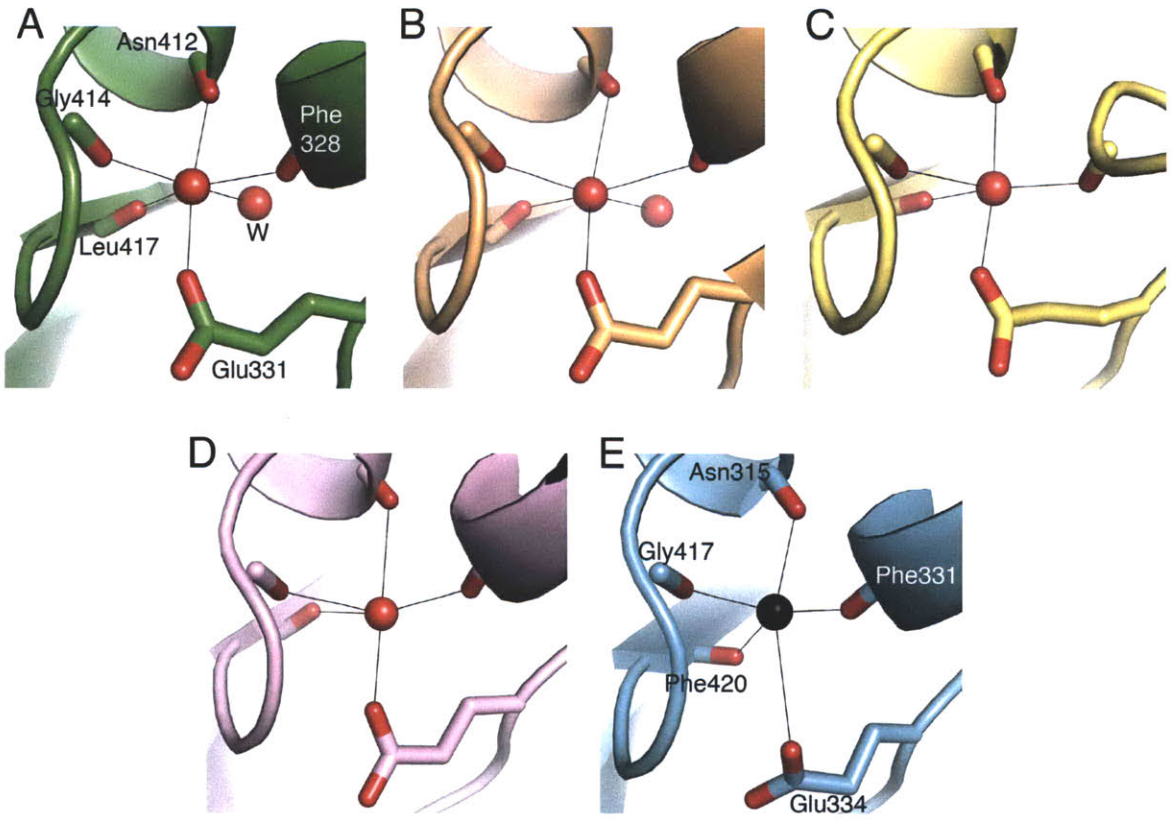


Figure AI.4. Putative sodium sites from previously solved ACS crystal structures, modeled as a water molecule. A, The first CODH/ACS structure (PDB ID: 1MJG) in green ribbons, where ACS exists in a "closed" conformation. **B**, A separate CODH/ACS structure (PDB ID: 1OAO) in orange ribbons, where the ACS subunit shown is in an "open" conformation. **C**, The CODH/ACS structure (PDB ID: 2Z8Y) in yellow ribbons, where crystals were pressurized with xenon prior to data collection to identify tunnels within the enzyme complex. The water molecule which completes octahedral geometry was not present in this model. **D**, The truncated ACS structure (PDB ID: 3GIT) in pink ribbons, where Domain 1 of ACS was removed. The water molecule which completes octahedral geometry is also absent in this model. Structures shown in **A-D** are of ACS from *M. thermoacetica*, whose sequence is labeled in **A**. **E**, The structure of monomeric ACS from *C. hydrogenoformans* (PDB ID: 1RU3) in cyan ribbons. The water molecule shown as a black sphere was not present in the final model and depicts where the putative sodium site would be located. In all structures, ligands to the putative sodium site are shown in sticks (O in red and C following the protein ribbon coloring), and the coordinating water molecule is shown as a red sphere.

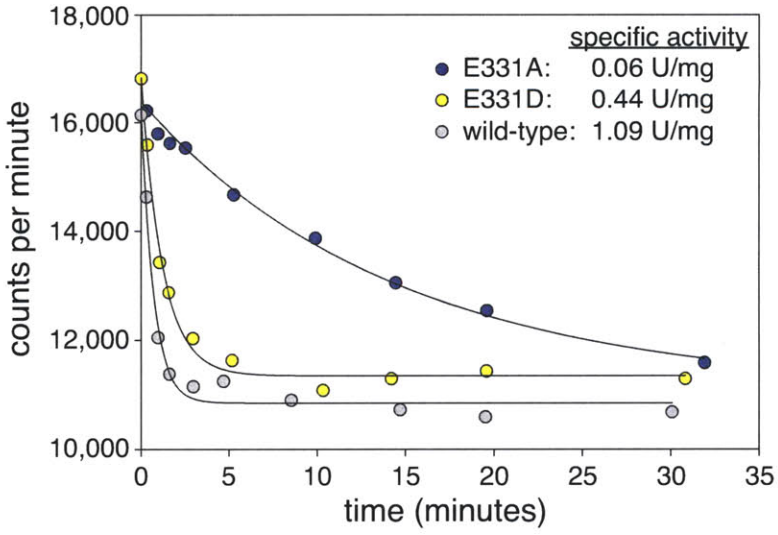


Figure AI.5. CO/[1-^{14}C]acetyl-CoA exchange assay. Mutations of Glu331, which coordinates the putative sodium site, affect the specific activity of ACS as shown.

AI.E. References

1 Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002) A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Science.* **298**, 567-572.

2 Shanmugasundaram, T., Kumar, G. K. & Wood, H. G. (1988) Involvement of Tryptophan Residues at the Coenzyme A Binding Site of Carbon Monoxide Dehydrogenase from *Clostridium thermoaceticum*. *Biochemistry.* **27**, 6499-6503.

3 Ragsdale, S. W. & Wood, H. G. (1985) Acetate Biosynthesis by Acetogenic Bacteria: Evidence that Carbon Monoxide Is the Condensing Enzyme that Catalyzes the Final Steps of the Synthesis. *J. Biol. Chem.* **260**, 3970-3977.

4 Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **D53**, 240-255.

5 Harding, M. M. (1999) The Geometry of Metal-Ligand Interactions Relevant to Proteins. *Acta Cryst.* **D55**, 1432-1443.

6 Harding, M. M. (2001) Geometry of Metal-Ligand Interactions in Proteins. *Acta Cryst.* **D57**, 401-411.

7 Harding, M. M. (2002) Metal-Ligand Geometry Relevant to Proteins and In Proteins: Sodium and Potassium. *Acta Cryst.* **D58**, 872-874.

8 Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003) Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] Clusters in Closed and Open Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. *Nat. Struct. Biol.* **10**, 271-279.

9 Doukov, T. I., Blasiak, L. C., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2008) Xenon in and at the End of the Tunnel of Bifunctional Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *Biochemistry.* **47**, 3474-3483, doi:10.1021/bi702386t.

10 Bender, G., Stitch, T. A., Yan, L., Britt, R. D., Cramer, S. P. & Ragsdale, S. W. (2010) Infrared and EPR Spectroscopic Characterization of a Ni(I) Species Formed by Photolysis of a Catalytically Competent Ni(I)-CO Intermediate in the Acetyl-CoA Synthase Reaction. *Biochemistry.* **49**, 7516-7523.

11 Seravalli, J. & Ragsdale, S. W. (2008) Pulse-Chase Studies of the Synthesis of Acetyl-CoA by Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase: Evidence for a Random Mechanism of Methyl and Carbonyl Addition. *J. Biol. Chem.* **283**, 8384-8394.

12 Raybuck, S. A., Ramer, S. E., Abbanat, D. B., Peters, J. W., Orme-Johnson, W. H., Ferry, J. G. & Walsh, C. T. (1991) Demonstration of Carbon-Carbon Bond Cleavage of Acetyl Coenzyme A by Using Isotopic Exchange Catalyzed by the CO Dehydrogenase Complex from Acetate-Grown *Methanosarcina thermophila*. *J. Bacteriol.* **173**, 929-932.

13 Seravalli, J., Kumar, M. & Ragsdale, S. W. (2002) Rapid Kinetic Studies of Acetyl-CoA Synthesis: Evidence Supporting the Catalytic Intermediacy of a Paramagnetic NiFeC Species in the Autotrophic Wood-Ljungdahl Pathway. *Biochemistry.* **41**, 1807-1819.

14 Seravalli, J., Xiao, Y., Gu, W., Cramer, S. P., Antholine, W. E., Krymov, V., Gerfen, G. J. & Ragsdale, S. W. (2004) Evidence that NiNi Acetyl-CoA Synthase Is Active and That the CuNi Enzyme Is Not. *Biochemistry.* **43**, 3944-3955.

Yan Kung

Department of Chemistry, Room 68-688
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139

Tel: *(617) 894-8812*
Fax: *(617) 258-7847*
E-mail: *ykung@mit.edu*

EDUCATION

2005-Present ***Massachusetts Institute of Technology**, Cambridge, MA*
Ph.D., Biological Chemistry
- **Dissertation:** Structural studies of metalloenzyme complexes in acetogenic carbon fixation

2001-2005 ***Colby College**, Waterville, ME*
B.A. with Honors, Chemistry
- *Summa cum laude*, Phi Beta Kappa

2003-2004 ***St. Peter's College, Oxford University**, Oxford, United Kingdom*
- Visiting Student in Biochemistry

PUBLICATIONS

Kung, Y. and Drennan, C.L. (2011) A role for nickel-iron cofactors in biological carbon monoxide and carbon dioxide utilization. *Curr. Opin. Chem. Biol.* 15, 276-283.

Kung, Y., Doukov, T.I., Seravalli, J., Ragsdale, S.W., and Drennan, C.L. (2009) Crystallographic snapshots of cyanide- and water-bound C-clusters from bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Biochemistry*. 48, 7432-7440.

Kung, Y., Ando, N., Doukov, T.I., Blasiak, L.C., Bender G., Seravalli, J, Ragsdale, S.W., and Drennan, C.L. Visualising molecular juggling in a B_{12}-dependent methyltransferase complex. *In Review.*

Kung, Y., Ando, N., Bender, G., Ragsdale, S.W., and Drennan, C.L. Global conformational movements within a B_{12}-dependent methyltransferase complex involved in acetogenic CO_2 fixation. *In Preparation.*

Ando, N., Kung, Y., Bender, G., Ragsdale, S.W., Drennan, and C.L. Structural studies of protein complex formation in solution during acetogenic B_{12}-dependent methyl transfer. *In Preparation.*

ORAL PRESENTATIONS

Feb 4, 2011 ***Gordon Research Conference (GRC): Bioinorganic Chemistry Graduate Research Seminar, Ventura, CA***
"Structural investigations of a metalloenzyme complex enabling microbial growth on CO_2"

Oct 20, 2010 ***MIT Microbiology Seminar Series***
"Structural investigations of a metalloenzyme complex in bacterial acetogenesis"

Aug 24, 2010 ***American Chemical Society Meeting, Boston, MA***
"Crystal structure of a metalloenzyme complex involved in bacterial acetogenesis"

Apr 5, 2010 ***MIT Biological Chemistry Seminar Series***
"Using CO_2 for energy: metalloenzyme complexes for carbon fixation in acetogenic bacteria"

Feb 6, 2010 ***GRC: Bioinorganic Chemistry Graduate Research Seminar, Ventura, CA***
"Crystallographic studies of CO to CO_2 interconversion at a Ni-Fe-S cluster"

Aug 6, 2009 ***GRC: Vitamin B_{12} and Corphins, Magdalen College, Oxford University, UK***
"Crystal structure of a protein complex involved in B_{12}-dependent methyl transfer"

Apr 22, 2008 ***MIT Biology/Chemistry Interdepartmental Seminar Series***
"Carbon fixation in acetogens: from CO_2 to cellular carbon in one enzyme"

POSTER PRESENTATIONS

Aug 1, 2010 ***GRC: Molecular Basis of Microbial One-Carbon Metabolism, Bates College, Lewiston, ME***
"Structural studies of protein complexes in acetogenic B_{12}-dependent methyl transfer"

Jan 31, 2010 ***GRC: Metals in Biology, Ventura, CA***
"Crystallographic studies of CO to CO_2 interconversion at a Ni-Fe-S cluster"

Jan 14, 2010 ***MIT Energy Initiative (MITEI) Energy Futures Poster Session***
"Enzymatic regulation of pollutants and greenhouse gases"

May 27, 2009 ***MIT Center for Environmental Health Sciences Annual Poster Session***
"Crystal structures of the substrate- and inhibitor-bound $NiFe_4S_4$ cluster of carbon monoxide dehydrogenase/acetyl-CoA synthase"

Jul 20, 2008 ***GRC: Molecular Basis of Microbial One-Carbon Metabolism, Bates College, Lewiston, ME***
"Crystal structures of the substrate- and inhibitor-bound $NiFe_4S_4$ cluster of carbon monoxide dehydrogenase/acetyl-CoA synthase"

Feb 19, 2005 ***AAAS Annual Meeting, Washington, D.C.***
"Production of a mutant stylar-ribonuclease that deactivates self-incompatibility in the Solanaceae family"

HONORS AND AWARDS

2009	Morse Travel Grant for GRC: Vitamin B_{12} and Corphins
2008	Morse Travel Grant for GRC: Molecular Basis of Microbial One-Carbon Metabolism
2005	*Summa cum laude*
2005	Phi Beta Kappa
2005	American Institute of Chemists (AIC) Student Award in Biochemistry
2005	Merck-AAAS Scholarship for undergraduate research and travel grant
2004	Certificate of Excellence, Department of Pharmacology, Oxford University
2001-2005	Dean's List, every semester at Colby College
2001-2005	Shelby MC Davis United World College (UWC) Scholarship to attend Colby College

RESEARCH EXPERIENCE

2006-Present ***Research Assistant**, Dept. of Chemistry, MIT*
Advisor: Catherine L. Drennan, Professor of Chemistry and Biology, HHMI

- **Graduate thesis work:** Structural studies of metalloenzyme complexes in the anaerobic acetogenic CO_2 fixation pathway (collaboration with Stephen W. Ragsdale, University of Michigan): 1) X-ray crystallographic, small-angle X-ray scattering (SAXS), and *in crystallo* spectrophotometric studies of a ~220 kDa metalloenzyme complex of the B_{12}-containing corrinoid iron-sulfur protein and its methyltransferase enzyme, which depicted all protein components required for B_{12}-dependent methyl transfer together for the first time, identifying the modular arrangement of protein subunits and the substantial conformational movements inherent to B_{12} activation, protection, and catalysis. 2) X-ray crystallography of bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase

(CODH/ACS), a single ~310 kDa metalloenzyme complex which directly converts CO_2 to acetyl-CoA through a CO intermediate; solved crystal structures of the CODH/ACS [Ni-Fe_3-S_4]-Fe active site "C-cluster" bound with substrate and inhibitor molecules in the mechanistic elucidation of biological CO/CO_2 interconversion.

- Additional enzymes studied by X-ray crystallography include: 1) epothilone synthetase B (EpoB), a nonribosomal peptide synthetase (NRPS) module within a hybrid NRPS/polyketide synthase (PKS) biosynthetic pathway that produces epothilone, an anticancer agent (collaboration with Christopher T. Walsh, Harvard Medical School). 2) Several Fe_4S_4-containing *S*-adenosylmethionine (SAM, AdoMet) radical enzymes such as biotin synthase, which catalyzes the final, sulfur-insertion reaction in biotin (vitamin B_7) biosynthesis (collaboration with Joseph T. Jarrett, University of Hawaii), and spore photoproduct lyase, which catalyzes the repair of UV-damaged DNA thymine dimers from bacterial spores (collaboration with Joan B. Broderick, Montana State University).

2004-2005 ***Undergraduate Researcher**, Dept. of Chemistry, Colby College*
Advisor: Stephen U. Dunham, Assistant Professor of Chemistry
- Conducted undergraduate honors thesis research investigating the impact of mutant stylar ribonucleases on the inhibition of self-fertilization in select populations of the Solanaceae family of plants by molecular cloning of wild type and mutant enzymes.

2003-2004 ***Undergraduate Researcher**, Dept. of Pharmacology, Oxford University*
Advisor: Edith Sim, Professor of Pharmacology
- Studied mechanisms of tuberculosis resistance to isoniazid drug treatments through the construction of a *Mycobacterium tuberculosis* model in zebrafish by molecular cloning of arylamine *N*-acetyltransferase (NAT) genes from various mycobacterial species.

Jan 2003 ***Research Intern**, Dept. of Microbiology, Mount Sinai Hospital, Toronto, Canada*
Advisor: Allison McGeer, M.D.
- Investigated methicillin-resistant *Staphylococcus aureus* (MRSA) outbreaks in Toronto's hospital network and identified isolates by genotyping with pulsed-field gel electrophoresis.

May-Aug 2002 ***Summer Research Intern**, Infection Control, Mount Sinai Hospital, Toronto, Canada*
Advisor: *Barbara M. Willey*
- Designed and conducted a research study investigating outbreaks within Toronto's long-term care facilities of multidrug resistant bacteria with extended-spectrum β-lactamase activity: screened patient samples for microbial culture and antibiotic resistance, identified bacterial strains, discovered new cases, and postulated methods of disease propagation.

TEACHING AND MENTORING EXPERIENCE

Spring 2010 ***Teaching Assistant**: Massachusetts Institute of Technology (MIT)*
5.78/7.71J – Biophysical Chemistry Techniques
- Prepared class material, led laboratory exercises, maintained class website, and graded homework and exams in this graduate-level course.

Spring 2008 ***Howard Hughes Medical Institute (HHMI) Teaching Fellow:** MIT*
5.111 – Principles of Chemical Science
- Taught a recitation class of 16 students in twice-weekly sections for this undergraduate course in general chemistry: prepared class material, lectured, and led the section, and graded homework and examinations for the course of about 100 students.

Summer 2007 ***HHMI-MIT Mentoring Training Course**: MIT*

- Enrolled in an HHMI-affiliated training course at MIT to learn and refine effective undergraduate mentoring strategies in laboratory research.

2007-2010 ***Laboratory Mentor**: Drennan Laboratory, MIT*
- Trained and mentored a total of three undergraduate students conducting research in the Drennan lab over four years.

Spring 2006 ***Teaching Assistant**: Massachusetts Institute of Technology*
5.32 – Intermediate Chemical Experimentation
- Taught the undergraduate biochemistry laboratory course at MIT: prepared laboratories, gave course lectures, graded reports, and administered final oral examinations.

Fall 2005 ***Teaching Assistant**: Massachusetts Institute of Technology*
5.111 – Principles of Chemical Science
- Taught a recitation class of 25 students in twice-weekly sections for this undergraduate course in general chemistry: prepared class material, lectured, and led the section, and graded homework and examinations for the course of about 250 students.

RELATED ACTIVITIES

2006-Present **President** (2008-2009), **Treasurer** (2007-2008), and **Member** (2009-2011) of the MIT Chemistry Graduate Student Council: held regular events for graduate students, funded student activities, and represented student interests in departmental affairs.

2004-2005 **Editor-in-Chief** of the monthly campus magazine *The Community Voice* at Colby College: wrote, reported, published, and distributed student and faculty articles on intellectual, cultural, and political issues affecting both the college and the world at large.

2004-2005 **Student appointment** on the Colby College Committee on Race and Racism: investigated incidences of racial harassment, suggested changes to harassment policies, conducted analysis of minority attrition rates, and developed proposals to improve minority retention

REFERENCES

Graduate Advisor
Catherine L. Drennan
Professor of Chemistry and Biology
HHMI Investigator and Professor
Massachusetts Institute of Technology
Room 68-680, 77 Massachusetts Avenue
Cambridge, MA 02139
Tel: 617-253-5622
Fax: 617-258-7847
E-mail: cdrennan@mit.edu

Thesis Committee Chair
Professor JoAnne Stubbe
Novartis Professor of Chemistry
Professor of Biology
Massachusetts Institute of Technology
Room 18-598, 77 Massachusetts Avenue
Cambridge, MA 02139
Tel: 617-253-1814
Fax: 617-324-0505
E-mail: stubbe@mit.edu

Research Collaborator
Stephen W. Ragsdale
Professor of Biological Chemistry
University of Michigan
5220D MSRB3, Box 5606
Ann Arbor, MI 48109
Tel: 734-615-4621
Fax: 734-763-4581
E-mail: sragsdal@umich.edu