Dalton Transactions

Compositional and Structural Insights into the Nature of the H-Cluster Precursor on HydF

Journal:	*Dalton Transactions*
Manuscript ID	DT-ART-04-2018-001654.R1
Article Type:	Paper
Date Submitted by the Author:	27-May-2018

Complete List of Authors:	Scott, Anna; Montana State University, Chemistry & Biochemistry Szilagyi, Robert; Montana State University, Chemistry and Biochemistry Mulder, David; National Renewable Energy laboratory, Ratzloff, Michael; National Renewable Energy laboratory, Byer, Amanda; Montana State University, Chemistry & Biochemistry King, Paul; National Renewable Energy Laboratory, Broderick, William; Montana State University, Chemistry & Biochemistry Shepard, Eric; Montana State University, Chemistry & Biochemistry Broderick, Joan; Montana State University, Chemistry & Biochemistry

Journal Name

ARTICLE

Compositional and Structural Insights into the Nature of the H-Cluster Precursor on HydF

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Anna G. Scott,[a] Robert K. Szilagyi,[a] David W. Mulder,[b] Michael W. Ratzloff,[b] Amanda S. Byer,[a] Paul W. King,[b] William E. Broderick,[a] Eric M. Shepard,[a] and Joan B. Broderick*[a]

Assembly of an active [FeFe]-hydrogenase requires dedicated maturation enzymes that generate the active-site H-cluster: the radical SAM enzymes HydE and HydG synthesize the unusual non-protein ligands – carbon monoxide, cyanide, and dithiomethylamine – while the GTPase HydF serves as a scaffold for assembly of the 2Fe subcluster containing these ligands. In the current study, enzymatically cluster-loaded HydF ($[2Fe]_F$) is produced by co-expression with HydE and HydG in an *Escherichia coli* host followed by isolation and examination by FTIR and EPR spectroscopy. FTIR reveals the presence of well-defined terminal CO and CN^- ligands; however, unlike in the [FeFe]-hydrogenase, no bridging CO is observed. Exposure of this loaded HydF to exogenous CO or H_2 produces no significant changes to the FTIR spectrum, indicating that, unlike in the [FeFe]-hydrogenase, the 2Fe cluster in loaded HydF is coordinatively saturated and relatively unreactive. EPR spectroscopy reveals the presence of both [4Fe-4S] and [2Fe-2S] clusters on this loaded HydF, but provides no direct evidence for these being linked to the $[2Fe]_F$. Using the chemical reactivity and FTIR data, a large collection of computational models were evaluated. Their scaled quantum chemical vibrational spectra allowed us to score various $[2Fe]_F$ structures in terms of their ability to reproduce the diatomic stretching frequencies observed in the FTIR experimental spectra. Collectively, the results provide new insights that support the presence of a diamagnetic, but spin-polarized Fe^I-Fe^I oxidation state for the $[2Fe]_F$ precursor cluster that is coordinated by 4 CO and 2 CN^- ligands, and bridged to an adjacent iron-sulfur cluster through one of the CN^- ligands.

Introduction

Hydrogenases catalyze a stoichiometrically simple reaction, the interconversion of molecular hydrogen and protons, and play an important role in energy utilization and conversion in a variety of microorganisms.[1-4] While the compositional details vary, all hydrogenases employ organometallic species at their active sites to catalyze this reversible reaction. The [FeFe]-hydrogenase (HydA), for example, has an active site H-cluster composed of a [4Fe-4S] cubane ($[4Fe\text{-}4S]_H$) covalently bridged *via* a protein-derived cysteine thiolate ligand to a diiron subcluster ($[2Fe]_H$) that has no additional protein-derived ligands. Rather than an amino acid ligand, $[2Fe]_H$ is coordinated by three CO and two CN^- ligands, and a bridging dithiomethylamine (DTMA) ligand. The diatomic π-acid ligands are prominent in hydrogenase chemistry, as they are important in stabilizing the low spin and low oxidation states of the iron ions toward facilitating interactions with H^+, H^-, and H_2.[2-9]

The H-cluster is unique to [FeFe]-hydrogenases, and requires a dedicated biological machinery to aid its assembly and thus to generate an active enzyme.[10-13] The H-cluster assembly machinery includes three enzymes: the radical S-adenosyl-L-methionine (SAM) enzymes HydE and HydG, and the GTPase HydF.[14-17] HydG utilizes tyrosine as a substrate, cleaving it to form *p*-cresol and dehydroglycine (DHG);[18] the DHG is subsequently cleaved to generate the CN^-[19] and CO[20] that ultimately reside in the H-cluster of HydA.[21] This complicated reaction requires two Fe-S clusters on HydG,[22] a radical SAM [4Fe-4S] cluster near the N-terminus of HydG that catalyzes the reductive cleavage of SAM and the initial cleavage of tyrosine *via* a reversible H-atom abstraction event.[23] In addition, a novel [4Fe-4S]-[(κ^3-Cys)Fe] cluster was described near the C-terminus that appears to be involved in DHG conversion to CN^- and CO.[24-27] While the specific role for HydE in H-cluster biosynthesis is still unresolved,[17, 28, 29] we postulated that HydE may utilize a substrate with a thiol functionality to generate thioformaldehyde, two of which could condense with ammonia to form the DTMA bridge,[30] in a

[a] Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana 59717 United States.
[b] Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States.
† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: Additional FTIR data and EPR spectra. Details of Computational Modeling and Structure Refinement. Selection of the ωB97X-D/def2TZP and ωB97X-D/SVP levels of theory. Development of scaled quantum chemical force fields. Geometric, energetic, and vibrational analyses of **Models A-G**. Atomic positional coordinates of all the models and converged, ground state electronic structures of $[4Fe\text{-}4S]\times[2Fe]_F$ clusters in the format of Gaussian09 formatted checkpoint files are provided at DOI: 10.5281/zenodo.1052038. See DOI: 10.1039/x0xx00000x

manner analogous to the generation of DTMA in synthetic models of $[2Fe]_H$.[31]

The products of the HydE- and HydG-catalyzed reactions are thought to assemble a $[2Fe]_H$ precursor on HydF, as evidenced by the observation of CO and CN^- stretching bands on HydF purified from a system expressing all three hydrogenase maturases ($HydF^{EG}$).[32, 33] While the structural details of this $[2Fe]_H$ precursor on HydF were not further elucidated at the time, the observation that $HydF^{EG}$ activates $HydA^{\Delta EFG}$ supports the hypothesis that the CO and CN^- bound species on HydF is functionally relevant.[32-37] Berggren *et al.* have shown that HydF can also be loaded with a synthetic 2Fe subcluster containing CO and CN^- ligands, and this synthetic cluster ($[2Fe]_s$) can then be transferred to activate HydA.[9] In this semisynthetic HydF, $[2Fe]_s$ is bridged to a [4Fe-4S] cluster by a CN^- ligand.[9] Redox-active [4Fe-4S] and [2Fe-2S] clusters have been observed on HydF,[32, 34, 37-44] and a [4Fe-4S] cluster was bound to the cysteines of the conserved $CXHX_{46-53}HCXXC$ motif in a crystal structure of HydF.[45] This [4Fe-4S] is presumably the cluster bridged to $[2Fe]_s$ in semisynthetic HydF. The ability to load HydF with a synthetic 2Fe subcluster and the subsequent activation of $HydA^{\Delta EFG}$ are significant in providing insights into the function of HydF,[9, 45-47] however it does not directly address the nature of the enzymatically-loaded form of HydF produced by HydE and HydG, or the mechanism by which HydE and HydG synthesize $[2Fe]_F$.

Here we describe spectroscopic and computational studies directed at understanding the enzymatically-loaded form of HydF produced by co-expression with the radical SAM maturases HydE and HydG, with a specific focus on the [2Fe] H-cluster precursor (hereafter referred to as $[2Fe]_F$) present on purified $HydF^{EG}$. Our detailed FTIR, EPR, and density functional characterization of the $[2Fe]_F$ states of *Strep*-tag II $HydF^{EG}$ reveal unique structure and reactivity differences of $[2Fe]_F$ in comparison to the H-cluster of HydA. A wide array of computational models and their scaled quantum chemical vibrational spectra allowed us to score various $[2Fe]_F$ compositions and structures in terms of their ability to reproduce the diatomic stretching frequencies observed in the FTIR experimental spectra. Collectively, the results support the presence of a diamagnetic, spin-polarized Fe^I-Fe^I precursor cluster coordinated by 4 CO and 2 CN^- ligands that is bridged to an adjacent iron-sulfur cluster through one of the CN^- ligands.

Experimental

$HydF^{EG}$ Protein Expression and Purification. Expression and purification trials carried out at the National Renewable Energy Laboratory (NREL) were performed as follows. Using the Duet™ Vector System (Novagen), a *hydF / hydG* construct with respective genes from *Clostridium acetobutylicum* (*C.a.*) and a C-terminal *Strep*-tag II on the *hydF* gene was co-transformed with a *C.a. hydE* encoding construct into Novagen *E. coli* Rosetta™ 2(DE3) competent cells. The cells were streaked on fresh LB-Agar plates supplemented with 200 µg/mL carbenicillin and 50 µg/mL streptomycin. All media used in subsequent steps was supplemented with the same antibiotics and all growth steps were performed at 37 ˚C and 230 rpm shaking. Single colonies were used to inoculate 5 mL TB media starter cultures and grown for 15 h. In the morning, the starter cultures were used to inoculate 50 mL of TB media, which were grown to an OD_{600} of 0.4. Then 6 L of TB media containing 2 mM ferric ammonium citrate was inoculated with the second cultures and grown to an OD_{600} of 0.4. The cultures were removed from the incubator and induced with 1.5 mM IPTG and supplemented with 1 mM cysteine, 0.5% glucose, and 20 mM sodium fumarate. The flasks were sparged with Ar for 15 h. The sparged growth was transferred to an anaerobic Coy chamber (3 % H_2, 97 % N_2) and the cells were harvested by centrifugation, re-suspended in a minimal volume of Buffer A (50 mM HEPES, 300 mM KCl, 5% glycerol, pH 8.0) and frozen in airtight vials in an -80 ˚C freezer and stored there until purification. All subsequent steps were performed under anaerobic conditions.

Thawed cells were transferred to an anaerobic Coy chamber (3 % H_2, 97 % N_2) and 15 mL benzonase nuclease, 2 protease inhibitor tablets, and 1 mg lysozyme per 50 mL of re-suspended cells was added to the lysate mixture and allowed to mix by stirring for 10 minutes. The cells were lysed using 10 cycles on a micro-fluidizer. The resulting mixture was poured into gas-tight centrifuge tubes and centrifuged for 1 hour at 45,000 rpm at 4 ˚C. The centrifuged tubes were brought into an MBraun chamber (100% N_2 atmosphere) and the supernatant was loaded onto a Strep-Tactin (IBA) column pre-equilibrated with Buffer A**.** The loaded affinity column was washed with 3 column volumes of Buffer A. HydF was eluted with Buffer A supplemented with 2 mM desthiobiotin in one step. UV-Vis absorbance was used to track absorbance values at 280, 395 and 420 nm throughout the purification. The eluted HydF was concentrated first using an Amicon stirred cell under argon pressure and then Amicon spin filters using a table top centrifuge at 4 ˚C in an MBraun anaerobic chamber. Concentrated samples were either analyzed immediately by FTIR or flash frozen in liquid N_2.

Expression and purification trials were carried out at Montana State University (MSU) and mirrored those performed at the National Renewable Energy Laboratory (NREL) with the following alterations. Transformations were done using BL21-(DE3)-RIL competent cells. Ampicillin was used instead of carbenicillin during growth steps and sparging of cell growths was done with N_2. All lysis and purification steps were performed in an anaerobic Coy chamber (4 % H_2, 96 % N_2, 4 ˚C). Lysis was carried out *via* sonication (Model 505 sonic dismembrator, Fisher Scientific) on cells re-suspended in Buffer A supplemented with 9 mg lysozyme, 1 mg DNAse, 1 mg RNAse, 1 mL Triton™ X-100, 100 mg $MgCl_2$, and 9 mg PMSF per 50 mL of re-suspended cells.

FTIR Sample Preparation and Data Collection. FTIR spectra (512 scans, 2 cm^{-1} resolution) were collected on a Nicolet 6700 FTIR spectrometer using a high D*MCT detector and custom-built cryogen-free cryostat (ColdEdge Technologies, Inc.). Small aliquots of sample (15 µL) were loaded into a sample cell consisting of a set of CaF_2 windows with a 15 µm spacer. Spectra were manually baseline corrected using the OMNIC software program (Thermo Fisher Scientific).

EPR Sample Preparation and Data Collection. EPR samples were prepared in an MBraun chamber at ≤ 1 ppm O_2 levels. Aliquots of $HydF^{EG}$ were loaded into EPR tubes (Wilmad LabGlass, 4 mm OD, NJ, USA), capped with a rubber septum, and immediately flash frozen outside the box in liquid N_2. Dithionite reduced samples were prepared by supplementing enzyme with 2 mM sodium dithionite in 50 mM HEPES, 300 mM KCl, 5% glycerol, pH 8.0 buffer. Samples were incubated with dithionite for 10 minutes prior to flash freezing in the aforementioned manner. All samples were stored in a liquid N_2 Dewar until data collection occurred.

Low temperature (≤ 75 K) EPR measurements were made using a Bruker (Billerica, MA) EMX X-band (9.4 GHz) spectrometer equipped with a Bruker/Cold Edge (Sumitomo Cryogenics) 10 K waveguide cryogen free system with an Oxford MercuryiTC controller unit and helium Stinger recirculating unit (Sumitomo Cryogenics, ColdEdge Technologies, Allentown, PA). Unless otherwise noted, typical EPR parameters were: 9.38 GHz microwave frequency, 500 µW microwave power, 100 kHz modulation frequency, 10 G modulation amplitude, 327.68 ms time constant, and spectra were averaged over 4 scans. OriginPro 8.5 was utilized to baseline correct and plot all experimental spectra. Simulations of experimental data were carried out using the EasySpin software platform[48] and generated the reported g-values presented in the text and supporting information.

Computational Models and Methods. Based on the crystal structure of the $(Et_4N)_2[Fe_2(DTMA)(CO)_4(CN)_2]$ biomimetic complex,[31] computational models were constructed for systematically mapping the variability of CO/CN^- stretching frequencies as a function of composition, ligand coordination isomerism, oxidation and spin states of the 2Fe subcluster, protein ligand environment, and the presence of a covalently linked iron-sulfur cluster. For the latter 6Fe-containing computational models, as also discussed by Berggren *et al.*,[9] a site differentiated [4Fe-4S] cluster environment was taken from the highest resolution HydA structure (PDB ID 3C8Y)[49] containing three cysteine ligands with adjacent amide functional groups. During the preparation of the manuscript, the crystal structure of the non-activated HydF metalloenzyme from *Thermosipho melanesiensis* (*Tm*HydF, PDB ID 5KH0) with a site differentiated [4Fe-4S] cluster was published.[45] Given this structure had lower resolution (2.80 Å) than the near-atomic-resolution structure of 3C8Y (1.39 Å), it was necessary to refine the crystallographic atomic positional coordinates of the [4Fe-4S] cluster in *Tm*HydF and its inner-sphere ligand

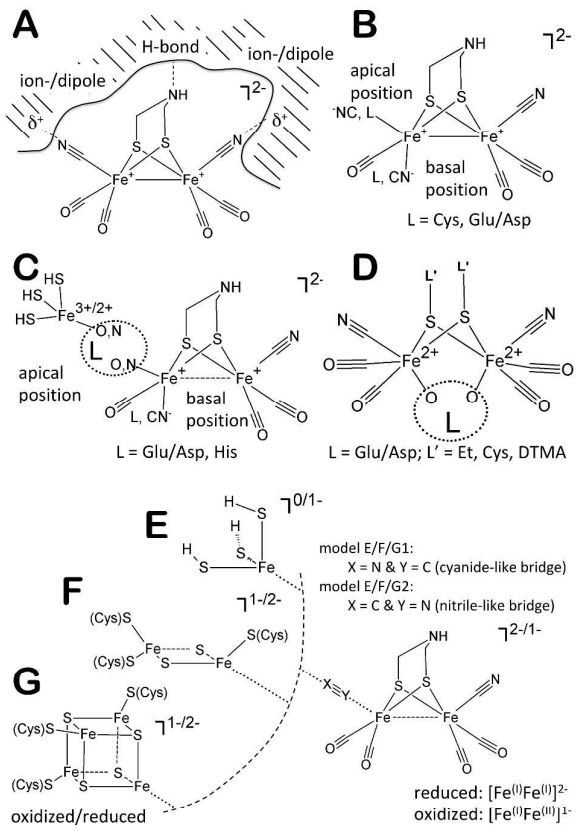

Figure 1. Overview of computational models for the $[2Fe]_F$ subcluster in $HydF^{EG}$ considered in the given work.

environment in a stepwise manner. The final structural optimizations were carried out with constrained backbone atoms in a model polarizable continuum for the electrostatic influence of protein environment. Details of the refinement are provided as Section 3 of Supporting Information (Figures S3.1-S3.4 and Table S3.1).

Figure 1 summarizes all models considered for $[2Fe]_F$ composition and structure in this study. These include a non-specific binding mode (**Model A**), one that is suggested by the similarity of the spectral features to synthetic 2Fe-clusters (**Model B**), and those that have already been proposed in the literature (**Model C**,[10, 42, 45] **Model D**,[50] **Model G**[9]). **Model F** was inspired by our spectroscopic evidence for [2Fe-2S] clusters on HydF,[37, 38, 51] and **Model E** completes the spectrum of possible iron species bridged to the $[2Fe]_F$. The variation in cluster oxidation states that were taken into account is indicated by the total molecular charges in Figure 1.

Given the approximate nature of density functional theory, the selection of the most reasonable set of exchange and correlation functionals for reproducing experimental CO/CN^- stretching frequencies is not straightforward, in part due to the requirements for accurate description of the location of stationary points and the curvature of the potential energy

surface. Prior published works[52-55] provided examples for how to select a reasonable level of theory and treat environmental effects induced by the presence of counter ions and inclusion of solvation for $[2Fe]_F$ cluster models. We adopted these approaches with a more critical evaluation of computational accuracy, in which we emphasized achieving the highest possible accuracy in reproducing molecular structures, relative coordination isomer energies, and vibrational signatures using the same set of functionals. The structural accuracy was described by the root mean square (r.m.s.) deviations of inner-sphere metal-ligand bond lengths and interatomic distances for radial information, as well as atomic positional coordinates for both radial and angular information as a function of the employed exchange and correlation functionals. In most calculations, we used the def2TZVP[56] all electron, triple-ζ quality basis set with a single set of polarization functions. Exceptions were made for the [2Fe-2S] and [4Fe-4S] cluster containing models (**Models F and G**, respectively; Figure 1) with two or three $(CH_3NH_3)^+$ counter-ions due to the large number of total electrons, where a smaller (def2SVP[57]) basis set was used. We confirmed (see Tables S4.1 and S4.2) that this smaller basis set also yields theoretically converged results for molecular geometry and vibrational frequencies in comparison to those obtained by the larger def2TZVP basis set. Another significant deviation from the earlier computational studies was the way we scored the agreement between the experimental and calculated Fe-CO/CN^- stretching frequencies. Instead of defining a set of rules[54] for assigning and correlating calculated and experimental values, we used separate average frequencies for the CO and CN^- manifolds of stretching frequencies. Notably, negligible difference was observed when we used individual CO/CN^- stretching frequencies in comparison to their manifolds. All details for the FTIR-based spectroscopic calibration of density functionals are provided in Section 4 of Supporting Information, Figures S4.1-4.3, Tables S4.1 and S4.2.

In brief, we have selected a range-separated hybrid DFT theory, ωB97X-D[58] with dispersion correction, since this functional gives the smallest r.m.s. deviations (0.02 and 0.08 Å) for interatomic distances and Cartesian coordinates of all atomic positions, for the $[Fe_2(DTMA)(CO)_4(CN)_2]^{2-}$ complex from a comprehensive range of GGA, metaGGG, and hybrid functionals. The corresponding r.m.s. values are worse for the BP86 (0.03 Å and 0.20 Å), B3LYP (0.05 Å and 0.12 Å), M06x (0.15 Å and 0.14 Å), and double hybrid B2PLYPD3 (0.06 Å and 0.27 Å) functionals. It is important to highlight that improvement in molecular structure counter intuitively paralleled a significant deviation from the experimental diatomic stretching frequencies. This can be mitigated by using a scaled quantum chemical force field approach[59-61] by empirically adjusting the curvature of the potential energy at a stationary point as has been done earlier for other functionals.[52-55] For the highest levels of theory considered in this study (ωB97X-D/def2TZVP), we developed an empirical shift (+102 cm^{-1}) along with a scaling factor (0.890). All reported stretching frequencies for the computational models discussed herein are adjusted in this manner. Inclusion of a polarizable continuum model[62, 63] does not notably perturb the correction factors (slope = 0.879 and intercept = +135 cm^{-1}). The performance of the (ωB97xD/def2TZVP) level of theory was further validated for various coordination isomers of the $[Fe_2(DTMA)(CO)_4(CN)_2]^{2-}$ complex with alternative ligand compositions and oxidation states (see Section 5 of Supporting Information, Figures S5.1 and S5.2). The crystallographically characterized structures[31, 64] correspond to the two lowest energy conformers at ωB97X-D/def2TZVP level, while their relative energies vary for other functionals. From a systematic comparison of experimental and calculated CO/CN^- stretching frequency manifolds, the lower limit for a significant difference between two structures was determined to be approximately 15±5 cm^{-1}. This value is important for the given study as we used it to define our confidence interval for meaningfully differentiating between two different calculated spectroscopic signatures. The largest computational models examined herein (**Models F and G**, Figure 1) using a simulated protein environment (*in proteo* modeling as provided by COSMO PCM[62, 63]) with water solvent parameters (except the dielectric constant was ε=10) and neutralizing counter ions (methylammonium cations) required the employment of a different set of correction factors (slope = 1.038; intercept = -218 cm^{-1}). These shifting and scaling factors for the more extended models are significantly different from the earlier ones due to the cumulative effects of using a smaller basis set and the presence of explicit counter ions in addition to the polarization continuum in order to consider protein environment effects in our calculations.

All unconstrained optimizations were repeatedly restarted until the frequency calculations showed no imaginary normal mode vibrations. The larger protein-embedded models were restarted at least three times upon randomly applying ±0.1 Å atomic displacements to avoid getting trapped at a local minimum with high potential energy due to the conformational flexibility. The appropriate spin coupling schemes for the [4Fe-4S] constraining models were obtained by merging well-defined ionic fragments of sulphide, ferric/ferrous ions, thiolate S(Cys) ligands, and the 2Fe subcluster with the appropriate arrangements of positive or negative spins.[65]

Results and Discussion

HydF has the remarkable ability to harbor a CO and CN^--ligated iron species when co-expressed with the maturases HydE and HydG[32, 33] or when an appropriate synthetic 2Fe cluster is supplied.[9, 45] The CO and CN^--ligated 2Fe subcluster can ultimately be transferred to $HydA^{\Delta EFG}$ to generate an active hydrogenase.[32-34] Thus, HydF is a central species along the biosynthetic pathway by which the [FeFe]-hydrogenase is matured.[10, 51] Although the CO- and CN^--ligated 2Fe species on $HydF^{EG}$ ($[2Fe]_F$) has been reported from two different labs,[32, 33] the structure of $[2Fe]_F$ and the nature of its attachment to the protein matrix are not well understood. In [FeFe]-hydrogenase

itself, the $[2Fe]_H$ subcluster has no protein coordination other than the cysteine that bridges this subcluster to the [4Fe-4S] cluster;[5] whether the coordination is similar in naturally loaded HydF is unknown. HydF does contain conserved cluster-binding ligands and both [4Fe-4S] and [2Fe-2S] clusters have been observed on $HydF^{\Delta EG}$.[37, 38, 51] The [4Fe-4S] cluster of $HydF^{\Delta EG}$ is coordinated by the three conserved cysteines as well as transiently by the conserved glutamate residue.[45] Although it is unclear where the [2Fe-2S] cluster of $HydF^{\Delta EG}$ binds, electron spin relaxation studies indicate that it is separated from the [4Fe-4S] cluster by at least 25 Å.[37, 51] A recent X-ray structure of HydF shows the two cluster binding sites in dimeric HydF are approximately 35 Å apart,[45] leading to the hypothesis that the observed [2Fe-2S] clusters may be bound to the same three conserved cysteines on the other site of the HydF dimer in at least a subset of protein molecules.[37, 51] In $HydF^{EG}$, however, the coordination of $[2Fe]_F$ is not clear; thus this cluster could interact with either the [4Fe-4S] or [2Fe-2S], or with both of them. While no detailed molecular picture was provided, XAS studies did provide support for the assembly of a [4Fe-4S]×$[2Fe]_F$ unit on $HydF^{EG}$.[66] Experimental observation of loading $HydF^{\Delta EG}$ with synthetic 2Fe subclusters ($[2Fe]_S$) was reported, where spectroscopic evidence suggests $[2Fe]_S$ is bridged to a [4Fe-4S] cluster through a CN^- ligand.[9] The recent X-ray crystal structure of HydF reveals a bound [4Fe-4S] cluster without the presence of $[2Fe]_F$, although a positively-charged pocket exists adjacent to the [4Fe-4S] cluster that could harbor the $[2Fe]_F$.[45] Here we provide FTIR and EPR spectroscopic characterization of the biologically relevant loaded state of $HydF^{EG}$, as well as companion computational analysis that provides insights into the nature of the clusters in this enzymatically "loaded" HydF.

Characterization of the $[2Fe]_F$ Cluster on $HydF^{EG}$ by FTIR. The purified *C. a.* $HydF^{EG}$ contains iron loading ranging between 2.1 and 4.4 Fe/protein, representing an improvement over prior work.[32, 34] While these new preparations still exhibit variable and substoichiometric iron loading (assuming 6 Fe/protein is full loading, although this is not known), it should be noted that the *in vivo* role of this protein is in the *transient* scaffolding of $[2Fe]_F$. It is thus not surprising that achieving high iron loading in as-purified protein has proven to be difficult, especially in cases where HydE, HydF, and HydG are all being overexpressed simultaneously. Regardless, the as-purified $HydF^{EG}$ exhibits well-resolved vibrational bands attributable to Fe-CO and Fe-CN^- stretching modes (Fig. 2). Bands are observed in regions characteristic of metal-bound terminal CO (1850-2015 cm^{-1}) and CN^- (2025-2110 cm^{-1}) ligands, with frequencies similar to those of holo-HydA (Figure 2), corroborating the presence of a $[2Fe]_H$-like precursor ($[2Fe]_F$) on $HydF^{EG}$. Specifically, when compared to the HydA1 [FeFe]-hydrogenase from *Chlamydomonas reinhardtii* (*Cr*HydA1) stretching modes, the FTIR features of $HydF^{EG}$ are indicative of at least two CN^- and at least four CO ligand modes; a similar vibrational stoichiometry was previously observed in $HydF^{EG}$ resulting from the homologous overexpression in its native host *Clostridium acetobutylicum*.[33]

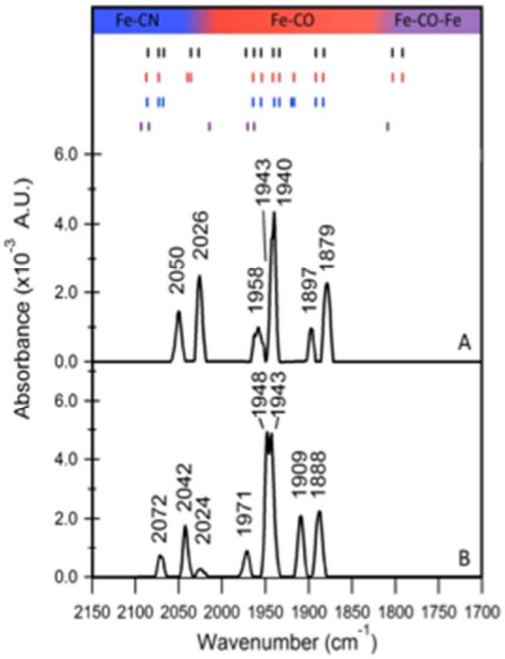

Figure 2: FTIR spectra of as-purified and DT-treated $HydF^{EG}$ along with reference FTIR bands for *Cr*HydA1 [FeFe]-hydrogenase (top). **A**. As-purified $HydF^{EG}$ (2.3 mM). **B**. DT-treated $HydF^{EG}$ (2.1 mM). Spectra in both **A** and **B** were recorded at 13 K. Peak shifts are likely due to redox changes at or near the $[2Fe]_F$ cluster, since HydF preparations with either spectral features present show no reactivity with CO or H_2. Reference FTIR bands for CrHydA1 are given for as-purified (black), H_2 (red), DT (blue), and CO (purple) treated enzymes.[1]

While $[2Fe]_F$ has similarities to $[2Fe]_H$, the FTIR data points to an important structural difference: the lack of a bridging CO ligand (μ-CO), as evidenced by the definitive absence of features in the μ-CO stretching region (1770-1850 cm^{-1}) in both the as-purified and reduced forms of $HydF^{EG}$ (Fig. 2). In addition, the terminal CO and CN^- vibrational bands of $[2Fe]_F$ occur at slightly different energies than for $[2Fe]_H$ in holo-HydA and as such may reflect differences from an altered active site environment, solvent exposure, hydrogen bonding, oxidation state, and/or composition of $[2Fe]_F$. The clear absence of any Fe-CO-Fe vibrational bands in current preparations of as-purified $HydF^{EG}$ (Figures 2 and S1.1) provides an important distinction to prior work. Happe and coworkers previously observed considerable speciation in a preparation of $HydF^{EG}$ in the absence of dithionite; one of the bands persisted at 1811 cm^{-1} and led to the conclusion that at least one oxidized enzyme state contained a Fe-CO-Fe moiety.[33] This observation in combination with EPR spectroscopic analysis indicated that the reduced form of $HydF^{EG}$ was EPR silent and did not harbour a Fe-CO-Fe species, while the oxidized form was EPR active and contained a bridging CO ligand.[33] Current results with $HydF^{EG}$ under as-purified conditions (either with or without

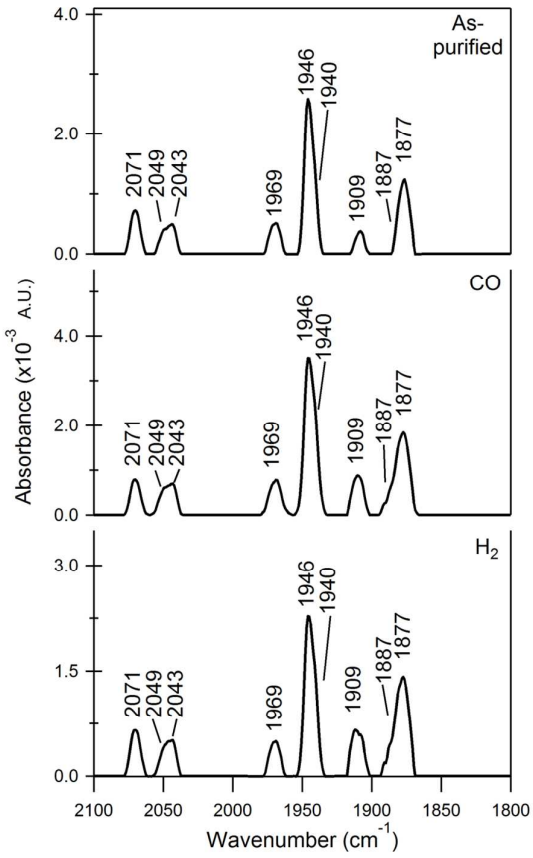

Figure 3: FTIR spectra of as-purified, CO-treated and H_2-treated $HydF^{EG}$ samples. The lack of spectral changes upon treatment is an indication for the presence of coordinatively saturated Fe-sites (sample concentration: 2.1 mM, 13 K).

dithionite), combined with the lack of reactivity with either H_2 or CO (see discussion below), together indicate that $[2Fe]_F$ does not harbour a Fe-CO-Fe moiety.

Reduction of $HydF^{EG}$ with dithionite results in shifting of some of the νCN^- and νCO modes to higher vibrational frequencies (Figure 2b), which is counterintuitive. Cluster reduction is expected to shift bands to lower energies as the π-acid ligands attract the extra electron density. Upshifts generally come from either oxidation or the redistribution of spin from the Fe atoms onto ligands or to a nearby cluster due to direct covalent interaction. Spectral differences of less than 22 and 13 cm^{-1} for CN^- and CO ligands, respectively, can be attributed to redox changes within or near $[2Fe]_F$. As such, they are not thought to represent significant perturbations in its structure given the delocalized nature of the electronic structure. Treatment with dithionite may also induce speciation, as evidenced by the appearance of an additional feature at 2072 cm^{-1} (Figure 2b).

The reactivity of $[2Fe]_F$ in as-purified $HydF^{EG}$ can provide further insights into this cluster's composition and structure; thus, in studies paralleling those of the 2Fe subcluster of the H-cluster, we examined whether $[2Fe]_F$ could interact with CO or H_2 in a manner similar to the 2Fe subcluster of the H-cluster. Previous studies have shown that the treatment of holo-HydA with exogenous CO and H_2 is a powerful probe of redox chemistry and structural changes of $[2Fe]_H$ through its catalytic cycle.[1, 7, 67, 68] For example, addition of exogenous CO leads to a CO inhibited form of holo-HydA with a simplified IR spectrum (Fig. 2 top, purple bands). [1, 7, 69] The spectral signature is consistent with an increase of ligand symmetry due to CO binding at the apical, open coordination site on the distal Fe of $[2Fe]_H$, where H_2 is also thought to bind.[1, 7, 67, 70] Exposure of the H-cluster to exogenous H_2, on the other hand, leads to a more complex IR spectrum reflecting a mixture of reduced and oxidized states that are in equilibrium as a result of H_2 binding and activation at the distal Fe (Fig. 2 top, red bands).[1, 68] In contrast to holo-HydA, treatments of $HydF^{EG}$ with exogenous CO and H_2 result in no significant changes in the spectral features (Fig. 3). Although general sample handling and slight variations in the methodology of enzyme preparations was shown to cause some minor spectral differences (Fig. S1), no significant changes were observed in the overall spectral composition among different enzyme preparations, whether or not they were treated with CO or H_2. These results show that $HydF^{EG}$ is unreactive towards CO and H_2, which require an open site on the $[2Fe]_H$ subsite for binding, thereby suggesting the presence of coordinatively saturated Fe ions in the $[2Fe]_F$ subsite.

Characterization of Cluster Species on $HydF^{EG}$ by EPR. EPR spectroscopy is complementary to FTIR in characterizing the paramagnetic cluster states of the $[2Fe]_F$ species on $HydF^{EG}$. We have previously reported detailed EPR spectroscopic characterization of $HydF^{\Delta EG}$, demonstrating the presence of redox-active [2Fe-2S] and [4Fe-4S] clusters in the protein.[38] These two different clusters on $HydF^{\Delta EG}$ are bound to sites that are at least 25 Å apart, suggesting the possibility that each occupies a cluster-binding site in separate monomers of HydF.[37, 45] In order to determine whether both types of cluster are present on naturally loaded $HydF^{EG}$, we employed EPR spectroscopy on $HydF^{EG}$ under as-isolated and reducing conditions. Our results show similarly to the case for $HydF^{\Delta EG}$,[37, 38] that $HydF^{EG}$ contains both [4Fe-4S] and [2Fe-2S] clusters in EPR-active states under certain conditions (Figure 4); this observation also appears to coalesce with the previously reported results using homologously overexpressed $HydF^{EG}$.[33] Our prior characterization of His-tagged $HydF^{EG}$ resulting from heterologous overexpression in *E.coli* showed only nominal $[2Fe\text{-}2S]^+$ cluster signals in as-purified enzyme samples, although at the time this small signal was attributed to $[3Fe\text{-}4S]^+$ clusters.[32] Regardless, our current data reveal that the $[2Fe\text{-}2S]^+$ cluster is observed in as-isolated $HydF^{EG}$, as is the case for $HydF^{\Delta EG}$ (Figure S2.1 and Table S2.1).[37, 38] The assignment is based on its EPR parameters being identical to typical $[2Fe\text{-}2S]^+$ clusters with primarily sulfur ligation. We

propose that these clusters are bound to the conserved cysteine residues of the $CXHX_{46-53}HCXXC$ motif.[71] While the role of the [2Fe-2S] cluster is currently unresolved, it is clear that it is observed in preparations of HydF from a multitude of sources.[33, 37-39, 42, 44, 66] Prior work has suggested that $[2Fe\text{-}2S]^+$ and $[4Fe\text{-}4S]^+$ clusters are coordinated to the dimeric form of HydF, which is also the quaternary state which activates $HydA^{\Delta EFG}$.[37] Thus it is tempting to speculate that the [2Fe-2S] cluster may be involved in electron transfer steps during $[2Fe]_F$ assembly, or possibly as a placeholder for HydG-derived $Fe(CO)_2(CN)$-cysteine synthon units.[24, 51, 72]

The $[4Fe\text{-}4S]^+$ cluster EPR signal appears upon reduction with similar g-values to what has been reported for $HydF^{\Delta EG}$, [37, 38] (see Supporting Information, Table S2.1). This [4Fe-4S] cluster is most likely bound to the protein by the three conserved cysteine residues of the $CXHX_{46-53}HCXXC$ motif, as observed in the HydF crystal structure.[45] Together, our results point to some of the three-cysteine cluster-binding sites being occupied by [4Fe-4S] clusters, while others are occupied by [2Fe-2S] clusters. None of these EPR signals are thought to arise from the $[2Fe]_F$ observed by FTIR, as $[2Fe]_F$ is expected to be EPR silent due to the presence of low-spin Fe^{II} or coupled low-spin Fe^{I} in this cluster as a result of the strong π-acid diatomic ligands. If the $[2Fe]_F$ were bridged to a [4Fe-4S] cluster, we might expect perturbations in the EPR spectrum due to the influence of the nearby subcluster, or even an altered spectrum resulting from a [6Fe] cluster as is observed in the H-cluster of HydA. However, as indicated above, our S=1/2 $[4Fe\text{-}4S]^+$ EPR spectra are very similar to those for the unloaded $HydF^{\Delta EG}$, giving little indication of the interaction between $[2Fe]_F$ and these EPR-active clusters (Figure S2.2 and Table S2.1). It is interesting to note that $[2Fe]_S$ loading in $HydF^{\Delta EG}$ similarly shows no substantial perturbations to the S=1/2 $[4Fe\text{-}4S]^+$ EPR signal.[45]

Computational Modeling of $[2Fe]_F$ Composition and Structure. The appearance of well-resolved, iron-bound CO and CN^- stretching bands in the experimental FTIR spectra (Figure 2) and the unambiguous absence of a μ-CO ligand provide a foundation for the structural modeling of the $[2Fe]_F$ cluster. The observed reactivity patterns, including the inability to generate a μ-CO state upon chemical treatment with exogenous CO gas and the lack of reactivity with H_2 (Figure 3), sets up strict boundaries for modeling the coordination environment of Fe ions in $[2Fe]_F$. Each of the seven models represented in Figure 1 has undergone extensive evaluations that allowed us to conclude that **Models A-D** do not meet a number of criteria, most importantly the significant disagreements (greater than 15±5 cm^{-1}) between the experimental and calculated CO/CN^- stretching manifolds. For brevity, here we report only key structural features or complex stability for evaluating each model for the $[2Fe]_F$, and refer the reader to Sections 6-12 of Supporting Information for further details.

The non-specific binding model (**Model A,** Figure 1) assumes a $2Fe^I/4CO/2CN^-$ composition for the cluster nested in the protein pocket without any covalent coordination by a residue. As a model for an arbitrary protein pocket, we considered two $CH_3NH_3^+$ molecules in close proximity to the CN^- groups and a CH_3OH moiety H-bonded to the DTMA bridgehead group. This model provided good agreement for the CO manifold (within 2 cm^{-1}); however, it deviates considerably (up to 23 cm^{-1}) for the experimental CN^- stretching frequencies (Figure S6.1).

Model B (Figure 1) invokes displacement of a CO ligand by a protein derived ligand in either an apical or a basal position of the 2Fe cluster that results in a $2Fe^I/3CO/2CN^-/L$ cluster composition, where L is a HydF-derived residue (Cys or Glu/Asp) modeled by ethylthiolate or acetate in this study. Structural optimizations of this model repeatedly resulted in one of the remaining CO ligands gradually migrating into a bridging position (Figure S7.1 for Cys; Figure S7.2 for Glu/Asp). The presence of a μ-CO ligand correlates with the absence of the Fe^I–Fe^I bond that with the above ligand composition creates a coordinatively unsaturated, trigonal bipyramidal Fe site. It was remarkable to observe that if there is a CN^- ligand in one of the apical positions that is proximal (and thus H-bonded) to the secondary amine bridgehead, the Fe

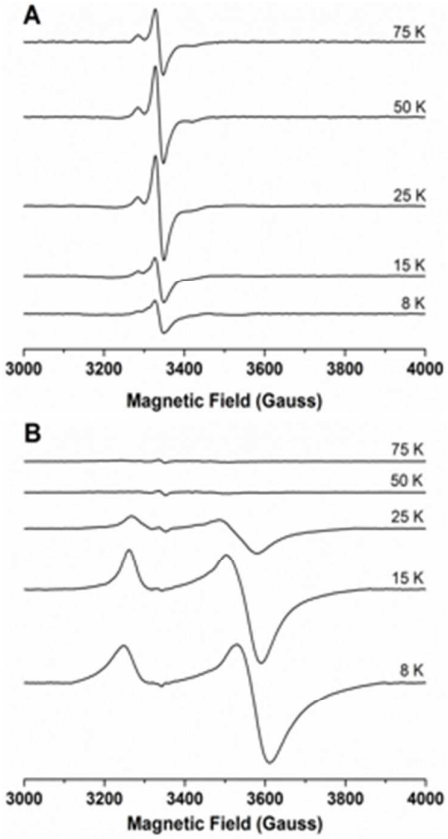

Figure 4: Low temperature CW X-band EPR spectra for strep-tagged $HydF^{EG}$. **A**. As-purified $HydF^{EG}$ (335 μM protein, 3.63 ± 0.46 Fe/dimer). **B**. Dithionite reduced (2 mM final) protein (335 μM protein, 3.63 ± 0.46 Fe/dimer).

Table 1. Comparison of the difference between calculated and experimental manifolds of CO and CN^- stretching frequencies ($\Delta\nu$CO and $\Delta\nu CN^-$, respectively in cm^{-1}) and energies differences between the two CN^- bridge isomers (ΔE^{SCF}(**F2-F1**) in kJ mol^{-1}) for **Model F**.

		vicinal, *cis*-Cys_2 constrained		vicinal, *cis*-Cys_2 relaxed		geminal-Cys_2 relaxed		Cys_3 relaxed	
		$\Delta\nu$CO	$\Delta\nu CN^-$	$\Delta\nu$CO	$\Delta\nu CN^-$	$\Delta\nu$CO	$\Delta\nu CN^-$	$\Delta\nu$CO	$\Delta\nu CN^-$
Model F1	oxidized	10	10	27	10	22	16	17	27
(cyanide-like)	reduced	24	23	24	33	19	26	16	41
Model F2	oxidized	26	68	26	70	23	79	21	72
(nitrile-like)	reduced	25	63	22	65	15	65	18	51
ΔE^{SCF}(**F2-F1**)	oxidized	+23		+58		+73		+66	
	reduced	+15		-18		+36		-17	

coordination environment remains square pyramidal with only a slightly perturbed Fe^I–Fe^I bonding interaction. The loss of a CO ligand in going from **Model A** to **Model B** will considerably alter the stretching frequencies, since CO dissociation causes the Fe^I-sites to donate more electron density to the remaining CO ligands. This in turn results in red shifting the CO manifold by 79 – 91 cm^{-1} (Cys model) and 67 – 87 cm^{-1} (Glu/Asp model), making the agreement with experimental FTIR data significantly worse than in **Model A**.

In **Model C** (Figure 1), we considered the HydF-derived ligand to be bridging between the unique Fe-site of a [4Fe-4S] cluster and one of the Fe-sites of the $[2Fe]_F$. Structural optimizations showed that the 1,1- or the 1,3-carboxylate bridged structures do not give stable stationary structures as the mononuclear iron thiolate and the $[2Fe]_F$ cluster model spontaneously dissociate (Figure S8.1). An imidazole bridge between the two Fe sites can provide a stable structure; however, the CO stretching frequencies remain grossly underestimated by 33 – 63 cm^{-1} in the $[Fe^IFe^I]$ state (Figure S8.2). One-electron oxidation of the $[2Fe]_F$ model reduces the electron density at the Fe-sites and thus significantly reduces the Fe→CO back-donation. This results in CO stretches being now overestimated by 19 – 27 cm^{-1}, in addition to the CN^- stretches being shifted by 36 – 42 cm^{-1} in the $[Fe^IFe^{II}]$ cluster (Figure S8.2) relative to the $[Fe^IFe^I]$ state. None of the models based on **Model C** with a 3CO/2CN^- diatomic ligand composition gave any reasonable agreement with the experimental boundary conditions used for scoring the possible structure and composition of the $[2Fe]_F$.

Model D provided us with the opportunity to evaluate the possibility of a thiolate S(Cys) or a carboxylate O(Glu/Asp) ligand occupying the bridging position between the two Fe-sites, opposite to the DTMA ligand, as a clear explanation for lack of reactivity and lack of any μ-CO band in FTIR measurements. Coordinatively saturated, hexacoordinate Fe sites were obtained only for the diferrous state; however, the CO and CN^- manifolds were again 99 and 48 cm^{-1} blue-shifted relative to the experimental values for HydF (Figures S9.1-S9.3) for similar reasons as described above for the one-electron oxidized form of **Model C**. The consideration of a mixed valence, $[Fe^IFe^{II}]$ paramagnetic state resulted in a coordinatively vacant Fe^I site, which is primed for exogenous CO/H_2 coordination. This model composition and oxidation state are clearly inconsistent with the observed lack of reaction of $[2Fe]_F$ with either CO or H_2. Due to the unsaturated coordination environments, we can therefore eliminate this as a viable model, although the deviation between experimental and calculated IR stretching frequencies approaches the threshold limit of 15±5 cm^{-1}. In our set of seven models and ligand conformational and configurational variations, **Model G** is the only one to date that provides good agreement between the calculated and experimental IR frequencies and meets all of the experimental boundary conditions; thus, this model is described below in detail along with its derivatives of **Models E** and **F**.

Virtual Chemical Models for the $[2Fe]_F$ Bridged Structures (Models E-G). The earlier work examining the *in vitro* maturation of $HydA^{\Delta EFG}$ *via* a synthetically loaded $[Fe_2(DTMA)(CO)_4(CN)_2]$-HydF complex proposed that the $[2Fe]_S$ subcluster bridges to the site-differentiated Fe-site of the [4Fe-4S] cluster coordinated to HydF[9] while the possibility for CN^- ligand isomerism was also discussed. Thus, we also considered CN^- forming an Fe-C bond with either the $[2Fe]_F$ (**Models E1, F1, and G1**, cyanide-like coordination to $[2Fe]_F$) or with the [4Fe-4S] cluster (**Models E2, F2, and G2**, nitrile-like coordination to $[2Fe]_F$). Since our previous models (**Models A – D**) exhibited unambiguous energetic preference for the apical CN^- position, we only focused on the apical isomers as shown in Figure 1, **Model E-G**. This assumption was further supported by screening of fully optimized, stationary structures containing an $[Fe(SH)_3]$ moiety (**Model E**) as a minimalist model for a site-differentiated [4Fe-4S] cluster. **Model** E structures revealed (see Figures S10.1 and S10.2) that when the $[Fe(SH)_3]$ moiety is coordinated to the basal CN^- ligand, structural rearrangement occurs at the distal site. Similarly to **Model B**, the distal Fe site in **Model E** becomes coordinatively unsaturated as one of the CO ligands shifts towards a semi-bridging position. This structural change creates a vacant site at the distal Fe in a trigonal bipyramidal coordination environment, which is not in agreement with the experimental observation in the given study of lack of reactivity toward H_2 and coordination of exogenous CO in $[2Fe]_F$.

As an intermediate cluster composition, we also evaluated the possibility for a [2Fe-2S] rhomb linked $[2Fe]_F$ structure (**Model F**). Given the larger structural variability of a site-differentiated rhomb *versus* cubane, we considered three different ligand compositions that entailed coordination by two cysteine residues in vicinal, *cis* (Figures S11.1-S11.4) and geminal (Figures S11.5 and S11.6) arrangement, as well as the three cysteine coordinated (Figures S11.7 and S11.8) clusters. These structures were directly created from the HydA structure without altering the atomic positions of the two cysteine ligands. Table 1 summarizes the deviations between the calculated and experimental manifolds of CO and CN^- stretching frequencies, and relative energies of cyanide and nitrile linkage isomers; Figures S11.1 – S11.8 in the supporting information provide more details about each structure.

Considering the selected confidence interval of 15±5 cm-1 in $\Delta\nu CO$ and $\Delta\nu CN^-$ values, there is only one structure (oxidized, vicinal, *cis*-Cys_2 **Model F1** with constrained Cys ligand positions from HydA) that fully meets the requirement for agreement with the experimental stretching manifolds. The other structural models fall near this threshold assignment, but it is clear that in order to consider them as operative models the local protein environment would have to influence the CO and CN^- stretching manifolds. However, it is evident from Table 1 that none of the **Model F2** structures give acceptable agreement with the experimental enzyme data, since the CN^- manifold energy (average of two stretching modes) is blue shifted by at least 50 cm^{-1}. These results and the dominantly lower stability of **Models F2** versus **F1** (with two exceptions for the reduced vicinal, *cis*-Cys_2 and Cys_3 complexes) can rule out the possibility for bridging CN^- ligand isomerism.

Turning to the models containing a [4Fe-4S] cluster, the unconstrained optimization of both apical **Models G** resulted in unreasonable structures as the three cysteine ligands of the [4Fe-4S] cluster folded onto the $[2Fe]_F$ cluster to participate in the extensive H-bonding network near $[2Fe]_F$ due to the lack of any environmental constraints placed upon the *in vacuo* models (see Figure S12.1). These large-scale structural differences between the **Model G1** and **G2** coordination isomers upon optimization were mitigated by constraining the backbone atoms of the anchoring Cys ligands to their crystallographic positions.

Figure 5 compares the structure, energetics, stretching frequencies, and atomic spin densities for the optimized structures of **Models G1** and **G2** with constrained backbone atoms (see also Figure S12.2). As found earlier by others at a different level of theory[9] than we used here, the energy of **Model G1** is lower than that calculated for **G2**. Furthermore, constraining the atomic positions introduces approximately 40 kJ mol^{-1} of strain, which is a considerable energy cost from the protein environment for keeping the cysteine thiolate ligands in the observed position. When the [4Fe-4S] cluster is coordinated at the C-end (**Model G2**), the bridging CN^- ligand's stretching frequency shifts above 2100 cm^{-1}; thus the CN^-

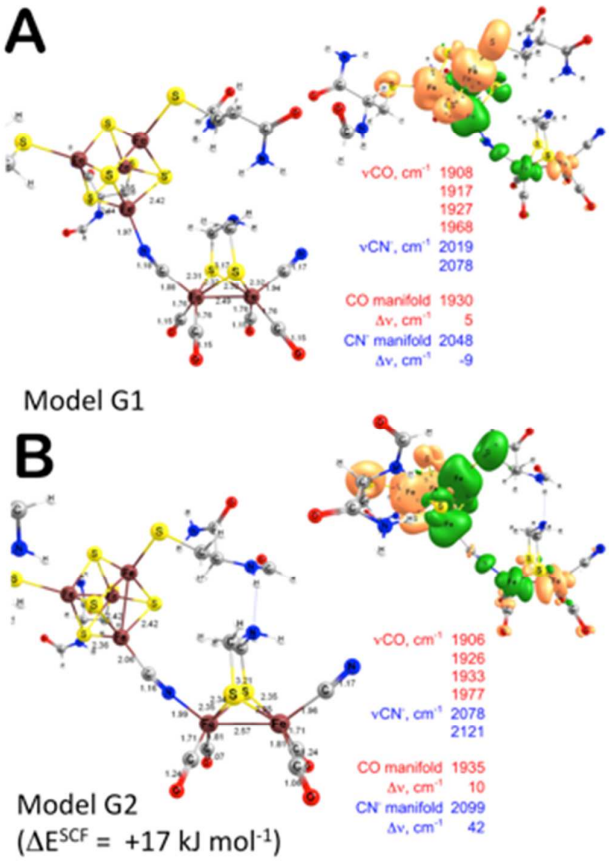

Figure 5: *In vacuo* molecular structures with selected bond lengths, relative energies, diatomic stretching frequencies, and atomic spin density contours (at 0.003 $e^-/Å^3$ level) for **Model G1** (**A**) and **Model G2** (**B**) as cyanide coordination isomers between the oxidized $[4Fe\text{-}4S]^{2+}$ cluster and fully-reduced $[2Fe^I]$-subcluster. The αC, amide N, and CO backbone atoms of the three cysteine ligands were kept frozen during the structural optimizations at their crystallographic positions.

manifold shifts by about 42 cm^{-1} away from experimental $HydF^{EG}$ values. On the contrary, the comparison of the scaled CO and CN^- stretching frequencies and the experimental values for $HydF^{EG}$ exhibit a remarkable agreement with **Model G1** (Figure 5A). The differences between the calculated frequency manifold and the experimental data are within ±10 cm^{-1}, which is below our confidence interval of 15±5 cm^{-1}.

During the preparation of the manuscript, the [4Fe-4S] cluster-loaded HydF structure (PDB ID: 5KH0) from *Thermosipho melanesiensis* (*Tm*HydF) was published.[45] We have therefore repeated the above calculation for an excised [4Fe-4S] cluster with one glutamate and three cysteine ligands (Figure S3.2). The atomic positions of the cluster centers were refined stepwise with three different magnetic coupling schemes (Figure S3.3) due to the structure's modest resolution (2.80 Å). In the refinement, the *in proteo* solvated, fixed backbone

stationary structure underwent about 610 kJ mol^{-1} of structural relaxation (Table S3.1) relative to the crystal structure. Upon full structural optimization of the Fe and S ions, there is an additional 183 kJ mol^{-1} relaxation, which is again a considerable protein constraint that is imposed on the [4Fe-4S] clusters in *Tm*HydF. Another unexpected result was the negative dissociation energy (-83 kJ/mol) of the glutamate ligand to form a [4Fe-4S] cluster with a three coordinate, unique Fe-site and a free glutamate anion under *in proteo* modeling conditions. This observation may be rationalized by protein strain effects that result in tight packing of the [4Fe-4S] cluster into the nest of *Tm*HydF. Using the 'native' protein environment, it was remarkable to find that the $[2Fe]_F$ model of $[Fe^{I}_2(DTMA)(CO)_4(CN)_2]^{2-}$ with apical CN^- ligands coordinates spontaneously (ΔG^{rxn} = -29 kJ mol^{-1}) to the site-differentiated cluster *via* one of the N-end of a bridging CN^- ligand (**Model G1**), while the coordination isomer formation (**Model G2**) is non-spontaneous (ΔG^{rxn} = +6 kJ mol^{-1}). The binding enthalpy (ΔH^{rxn}) between the two clusters is 34 kJ mol^{-1} favored for **Model G1** relative to **G2**. More importantly, **Model G1** in the *Tm*HydF environment reproduces within 8 and 6 cm^{-1} of the experimental CO and CN^- stretching modes, while **Model G2** overestimates them by 37 cm^{-1} (see Figure S12.8 for details).

It is important to highlight that the antiferro- and ferromagnetic coupling interactions among the Fe ions in the [4Fe-4S] cluster trigger a strong spin polarization also in the formally diamagnetic $[2Fe]_F$ cluster by effectively breaking the covalent Fe–Fe bond. This localizes the unpaired electrons on each of the low-spin $3d^7$ Fe^{I} sites, which is well illustrated by the green and orange lobes of spin density contours at the Fe sites of $[2Fe]_F$ (Figure 5) with shapes of $3d_{z^2}$ orbitals. The paramagnetic Fe centres provide a kinetic barrier to reactivity with diamagnetic substrates such as H_2 or exogenous CO. The two S=1/2 Fe^{I} sites are antiferromagnetically coupled through the bridging thiolate sulfurs. This is facilitated by the hyperfine exchange interactions with the sulfurs, which is based on covalent Fe–S bonds. The covalency and the spin-polarization of the entire [6Fe] cluster in biomimetic synthons has been shown from S K-edge XAS measurements.[73, 74] Another important feature of the spin density plots is the well-defined contour dominantly located at the proximal atom of the

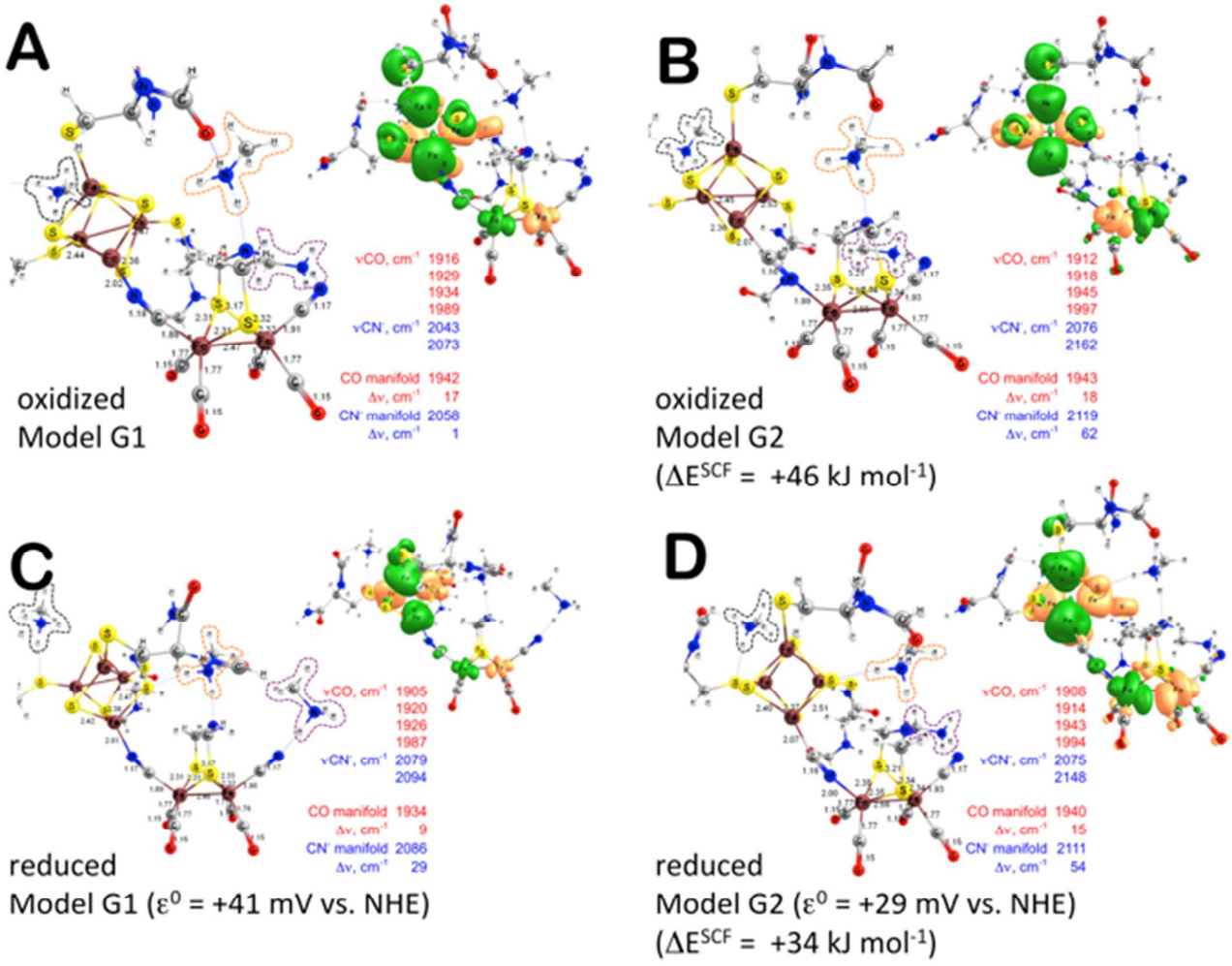

Figure 6: Solvated and unconstrained molecular structures with selected bond lengths, relative energies, diatomic stretching frequencies, and atomic spin density contours (at 0.003 e$^-$/Å^3 level) for **Model G1** (**A,C**) and **Model G2** (**B,D**) as cyanide coordination isomers between the oxidized $[4Fe\text{-}4S]^{2+}$ (**A,B**) and reduced $[4Fe\text{-}4S]^{+}$ (**C,D**) cluster and fully-reduced $[2Fe^{I}]$-subcluster (**A-D**).

bridging diatomic ligand between the two clusters. The spin density contour's phase matches the phase of the conjoint rhomb of the [4Fe-4S] subcluster and the phase of the proximal Fe site of the [2Fe] subcluster. This is the same for either the cyanide-like (on the C center) and the nitrile-like (on the N center) bridge between the two subclusters.

The **Model G** structures in Figure 5 did not take into account any electrostatic or electronic effects from either the protein environment or the presence of charged residues that screen the overall molecular charge of -3 for the $[4Fe\text{-}4S]\times[2Fe]_F$ model. Therefore, we carried out an additional series of calculations using an ε=10 polarizable continuum with three explicit methylammonium counter ions as an extreme electrostatic and H-bonding effect from the protein environment on the structure, stability, and properties of the $[4Fe\text{-}4S]\times[2Fe]_F$ construct. Figure 6 displays our most complete virtual chemical models for the protein bound $[2Fe]_F$ with different bridging cyanide positions and for the oxidized (A and B panels) and reduced (C and D panels) [4Fe-4S] cluster states. Figures S12.3 and S12.7 provide further details for the *in proteo* computational models in oxidized and reduced states, respectively. The structural optimizations using constrained backbone atoms show that one of the counter ions remains in the vicinity of $[2Fe]_F$ (dashed orange line) due to H-bonding interactions with the bridging amine group and the amide group of the closest thiolate S(Cys) ligand. The other cation near $[2Fe]_F$ (dashed purple line) forms an ion/dipole interaction with the terminal CN^- ligand at the apical position of the distal Fe site. The third methylammonium cation (dashed black line) is H-bonded to the sulfide anion opposite to the unique Fe-site of the [4Fe-4S] cluster. In the oxidized model, the energy gap increased between **Models G1** and **G2** up to +46 kJ mol^{-1}, in comparison to +17 kJ mol^{-1} for the *in vacuo* structures, while **Model G1** remains energetically favored. The spin density contour plots remain similar to the *in vacuo* structures, and to each other as well. It is noticeable that the Fe…Fe distance is considerably longer (+0.08 Å) in the oxidized **Model G2** structure, in addition to the elongation of the Fe-S bonds relative to **Model G1**. These changes correspond to a more open geometric and a more spin polarized electronic structure in **Model G2** as can be seen from the small but noticeable lobes of spin density at the S and the CO ligands in Figure 6B. Unexpectedly, the involvement of the CN^- ligands in ion/dipole or H-bonding interaction, as well as the presence of a neutralizing polarizable continuum, only slightly shifts the CN^- stretching frequencies, but considerably influences the CO stretching modes. The +17 cm^{-1} deviation in CO manifold energy in **Model G1** is within the considered confidence interval of 15±5 cm^{-1}. At the current level of theory, the agreement in the experimental and computational stretching frequencies of diatomic ligands is remarkable given the simplified treatment of protein environmental effects. In respect to the experimental reactivity data, the structural optimizations maintain the coordinatively saturated environments on the Fe centers of $[2Fe]_F$. On the other hand, the predicted diatomic stretching frequencies for **Model G2** do not provide good agreement with the experimental values, especially for the CN^- stretching modes.

In going from the oxidized (Figure 6A) to the reduced counterpart of **Model G1** in Figure 6C, the relative position of the $[Fe_4S_4(SCys)_3]^{2-}$ and $[2Fe]^{2-}$ clusters changes, as the network of weak interactions becomes significantly altered. We estimate the reduction potential *versus* normal hydrogen electrode (NHE, +4.43 eV) to be +41 mV, which is within a reasonable biologically achievable range. The spin density contour plots (top right corner insets) for the total of 19 unpaired electrons do not highlight straightforwardly the location of the additional unpaired electron as reduction influences the entire $[4Fe\text{-}4S]\times[2Fe]_H$ construct, due to extensive delocalization of molecular orbitals among the two clusters.[74] This is also reflected in the changes of CO and CN^- manifolds. The CN^- manifold shifts outside the acceptance limit (15±5 cm^{-1}). It is important to emphasize that the difference between the oxidized (Figure S12.5) and reduced forms (Figures S12.6 and S12.7) are in a similar range of 12-18 cm^{-1} of the experimentally detected values (see Figures 2 and 3, and Figure S1). The calculated CO band positions shift to lower energy (-8 cm^{-1}) or remain practically unchanged (-3 cm^{-1}) in **Models G1** and **G2**; respectively, while the CN^- bands shift up considerably as observed also in our FTIR experiments (Figure 2, $\Delta\nu CN^-$ = +16 – +22 cm^{-1}). This can be rationalized by the superimposed effect of the delocalization of the additional electron into a [6Fe] cluster and the overall larger electron density that will reduce the strength of the Fe-CN^- interactions. Experimentally, the CO stretching frequencies shift to higher energies (Figure 2, $\Delta\nu CO$ = +3 – +16 cm^{-1}), which can be used to estimate the significance of explicit protein environmental effects (presence of nucleophilic dipoles, H-bonding residues) that are only treated as a polarizable continuum in our computational study. Furthermore, it is important to consider that **Model G** is a delocalized, covalent 6Fe-cluster unit. Thus, the locus of the reduction can shift from one of the subclusters to the other, or the entire 6Fe-cluster can be affected. In addition, the presence of an additional unpaired electron is also expected to perturb the spin/spin-coupling within the entire cluster. This complicates the straightforward delineation of the observed changes in diatomic stretching frequencies as a function of redox state. However, taking the experimental differences, we can propose that as the C-O bonds in CO become stronger (positive $\Delta\nu CO$), the Fe–CO bonds become weaker, and the locus of reduction is on the [4Fe-4S]-subcluster. The latter can be induced by a stronger set of H-bonding to the sulphides and thiolates. Similarly, the $[2Fe]_F$ can also be perturbed by stronger dipole/H-bonding interactions that pull the CO ligands away from the [2Fe]-subcluster. Addition of an electron makes the entire [6Fe] cluster more nucleophilic, which attracts the H-bonding donors and the positive ends of the dipoles/cationic side chains toward the cluster. This may result in an electron density shift away from $[2Fe]_F$, which would weaken Fe-CO bonds. Moreover, the trends are the same for reduction potentials, relative energies, structural changes, and CO/CN^-

stretching frequencies in reduced **Model G2** relative to oxidized **Model G2**, or the reduced **Model G1**. However, none of the models with the coordination isomer of the bridging cyanide ligand (specifically, C-end to [4Fe-4S], N-end to [2Fe]), as proposed by Berggren *et al.*[9] give acceptable agreement with the current set of experimental results.

contribute to poor agreement with the experimental observables.

Conclusions

The combined spectroscopic and computational studies support a model in which a [2Fe] H-cluster precursor is

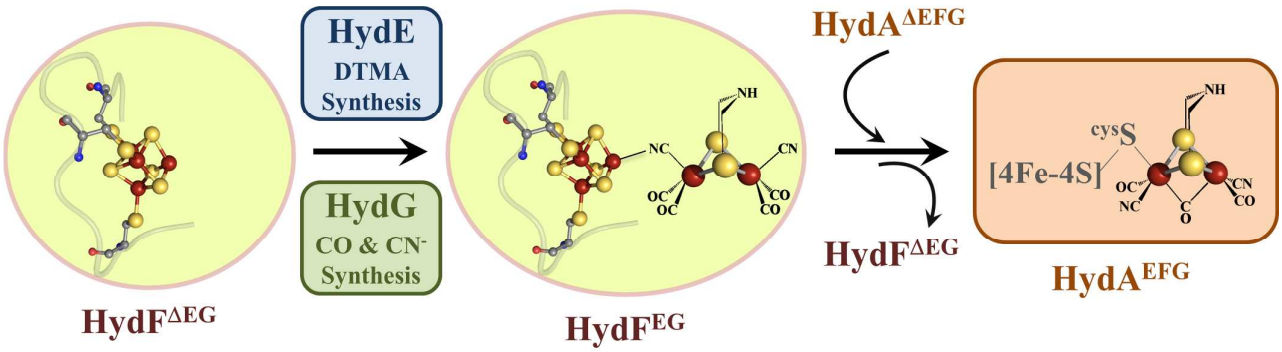

Figure 7. Schematic representation of the maturation of [FeFe]-hydrogenase (HydA) by the maturases HydE, HydG, and HydF.

The manual composition of the electronic structure for the entire [4Fe-4S]×$[2Fe]_F$ model from well-defined ionic fragments allows for evaluating the intramolecular electron transfer from $[2Fe]_F$ into the [4Fe-4S] cluster. This process would result in a paramagnetic, reduced $[4Fe\text{-}4S]^+$ cluster (S_1 = ½) that is linked through a CN^- ligand to a paramagnetic, mixed valence [Fe^I-Fe^{II}] cluster (S_2 = ½). Both ferro- (+½,+½) and antiferromagnetically (+½,-½) coupled states were considered. The energetic difference between the S_t = 1 and 0 states was calculated to be approximately 27 cm^{-1}. The magnitude of this coupling is indicative of non-negligible magnetic interaction between the two sub-clusters, which agrees with advanced EPR results for the $[2Fe]_H$ by Lubitz *et al.*[75, 76] Relevant to the observed reactivity of $[2Fe]_F$, this model completely rearranges upon intramolecular one-electron transfer as one of the terminal CO ligands at the proximal Fe-site shifts towards a bridging CO position. Spin density contour plots (Figure S12.4) reveal that the proximal Fe site will localize the unpaired electron of the $[2Fe]_F$, as the distal Fe^{II} site completes its coordination environment with a semi-bridging μ-CO. Thus, the structural changes upon intramolecular oxidation of the $[2Fe]_F$ are already detrimental to the agreement with the $HydF^{EG}$ experimental FTIR band positions, as the CO and CN^- stretching manifolds blue shift by 54–62 cm^{-1}. The peculiarity of these structures is that a vacant coordination site emerges at the proximal Fe site by the formation of a semi-bridging CO (indicated by the dashed lines in Figure S12.4). The structural changes for the full one-electron oxidized [4Fe-4S]×$[2Fe]_H$ model (Figure S12.5) with the locus of oxidation at $[2Fe]_H$ are even more exaggerated than those seen for the intramolecular electron transfer. The presence of an [Fe^IFe^{II}] cluster in $[2Fe]_F$ is unlikely due to the predicted formation of the μ-CO, formation of a vacant coordination site, and the concomitant shift in the CO/CN^- stretching frequencies that will all assembled on HydF through the enzymatic actions of HydE and HydG (Figure 7). This [2Fe] cluster on HydF is coordinated by terminal CO and CN^- ligands, with no bridging CO and no open sites for coordination of exogenous ligands, such as CO or H_2. The saturated $[2Fe]_F$ described herein would constrain HydF as a hydrogenase scaffold, as an open coordination site would be required for it to function as a catalyst. Further, imposing this specific geometric and electronic configuration is likely central to HydF's functional role in assembling and transferring the $[2Fe]_F$ cluster to $HydA^{\Delta EFG}$. The systematic evaluation of seven compositionally significantly different computational models for the $[2Fe]_F$ in both reduced Fe^IFe^I and oxidized Fe^IFe^{II} states using a FTIR spectroscopy-validated scaled quantum chemical force field unambiguously supported the assignment of the observed diatomic stretching frequencies.

We propose that the [2Fe]-cluster on HydF is most likely linked to a redox active [4Fe-4S] cluster through a cyanide ligand in an apical position, with the cyanide C bound to the $[2Fe]_F$ cluster and the N linked to the [4Fe-4S] cluster as represented by **Model G1**. The presence of a specific isomer (**Model G2**) was proposed earlier; however, the higher energy of **G2** vs. **G1**, and the significant blue shift (> 20 cm^{-1}) of the cyanide stretching frequencies for the nitrile-like bridge in **G2** in comparison to the cyanide-like bridge in **G1,** do not support the previous assignment. Furthermore, translocation of the $[2Fe]_F$ cluster in the **G2** model to $HydA^{\Delta EFG}$ would require breaking the strong NC-[4Fe-4S] bond in addition to the weak CN…$[2Fe]_F$ interaction, followed by isomerism to NC-$[2Fe]_F$. This could be a thermodynamically exigent process under physiological conditions.

In **Model G1**, the cyanide bridge electronically links the two subclusters and extends the covalent interactions and magnetic coupling throughout the entire [6Fe] construct. One-

electron reduction does not significantly perturb the diatomic stretching frequencies given that the additional electron is greatly delocalized over the entire [4Fe-4S]×[2Fe] cluster. In fact, multiple electron redox chemistry of the conjoined clusters could be responsible for the activation of the $[2Fe]_F$ subcluster for translocation into $HydA^{\Delta EFG}$, perhaps concomitant with rotation about the diiron bond to generate a catalytically competent, μ-CO bridged $[2Fe]_H$. Interestingly, redox chemistry is invoked in the building of the H-cluster from $[2Fe]_S$ as well.[77]

It seems likely that the process of maturation of $HydA^{\Delta EFG}$ will involve similar steps, regardless of whether the [2Fe] unit comes from a synthetic cluster in solution, or from chemically or enzymatically-loaded HydF. In vitro maturation using $[2Fe]_S$ has been examined in some detail in recent studies, and includes binding and translocation of the synthetic cluster into the active site, ligand exchange and loss of CO, and closing of the delivery channel.[77] In the case of maturation using the enzymatically-loaded HydF protein, a transient protein-protein complex would likely be involved that would allow direct transfer of $[2Fe]_F$ from HydF to HydA, perhaps via the intermediate P observed during maturation with $[2Fe]_S$.[77] Transfer of $[2Fe]_F$ would require breaking the cyanide bridge to the [4Fe-4S] cluster on HydF, translocation of the $[2Fe]_F$ from HydF to HydA, forming the cysteinate bridge to the [4Fe-4S] on HydA, and loss of a CO ligand concomitant with formation of the bridging CO.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been supported by the U.S. Department of Energy, Office of Basic Energy Sciences grant DE-SC0005404 (to J.B.B. and E.M.S.) and the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Lab. Some aspects of the computational modeling were supported by the U.S. National Science Foundation grant CLP 1609557 (to E.M.S., R.K.S., and J.B.B.). Computations were carried out on the Hyalite High-Performance Computing System, operated and supported by University Information Technology Research Cyberinfrastructure at Montana State University.

Notes and references

1. D. W. Mulder, M. W. Ratzloff, E. M. Shepard, A. S. Byer, S. M. Noone, J. W. Peters, J. B. Broderick and P. W. King, *J. Am. Chem. Soc.*, 2013, **135**, 6921-6929.
2. W. Lubitz, H. Ogata, O. Rudiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081-4148.
3. J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King and M. W. W. Adams, *BBA - Mol. Cell Res.*, 2015, **1853**, 1350-1369.
4. D. Schilter, J. M. Camara, M. T. Huynh, S. Hammes-Schiffer and T. B. Rauchfuss, *Chem. Rev.*, 2016, **116**, 8693-8749.
5. J. W. Peters, W. N. Lanzilotta, B. J. Lemon and L. C. Seefeldt, *Science*, 1998, **282**, 1853-1858.
6. Y. Nicolet, C. Piras, P. Legrand, C. E. Hatchikian and J. C. Fontecilla-Camps, *Structure*, 1999, **7**, 13-23.
7. Z. Chen, B. J. Lemon, S. Huang, D. J. Swartz, J. W. Peters and K. A. Bagley, *Biochemistry*, 2002, **41**, 2036-2043.
8. A. Silakov, C. Kamp, E. Reijerse, T. Happe and W. Lubitz, *Biochemistry*, 2009, **48**, 7780-7786.
9. G. Berggren, A. Adamska, C. Lambertz, T. R. Simmons, J. Esselborn, M. Atta, S. Gambarelli, J. M. Mouesca, E. Reijerse, W. Lubitz, T. Happe, V. Artero and M. Fontecave, *Nature*, 2013, **499**, 66-69.
10. E. M. Shepard, F. Mus, J. Betz, A. Byer, B. R. Duffus, J. W. Peters and J. B. Broderick, *Biochemistry*, 2014, **53**, 4090-4104.
11. A. S. Byer, E. M. Shepard, J. W. Peters and J. B. Broderick, *J. Biol. Chem.*, 2015, **290**, 3987-3994.
12. J. B. Broderick, A. S. Byer, K. S. Duschene, B. R. Duffus, J. N. Betz, E. M. Shepard and J. W. Peters, *J. Biol. Inorg. Chem.*, 2014, **19**, 747-757.
13. P. Dinis, B. M. Wieckowski and P. L. Roach, *Curr. Op. Struct. Biol.*, 2016, **41**, 90-97.
14. M. C. Posewitz, P. W. King, S. L. Smolinski, L. Zhang, M. Seibert and M. L. Ghirardi, *J. Biol. Chem.*, 2004, **279**, 25711-25720.
15. P. W. King, M. C. Posewitz, M. L. Ghirardi and M. Seibert, *J. Bacteriol.*, 2006, **188**, 2163-2172.
16. S. E. McGlynn, S. S. Ruebush, A. Naumov, L. E. Nagy, A. Dubini, P. W. King, J. B. Broderick, M. C. Posewitz and J. W. Peters, *J. Biol. Inorg. Chem.*, 2007, **12**, 443-447.
17. J. K. Rubach, X. Brazzolotto, J. Gaillard and M. Fontecave, *FEBS Lett.*, 2005, **579**, 5055-5060.
18. E. Pilet, Y. Nicolet, C. Mathevon, T. Douki, J. C. Fontecilla-Camps and M. Fontecave, *FEBS Lett.*, 2009, **583**, 506-511.
19. R. C. Driesener, M. R. Challand, S. E. McGlynn, E. M. Shepard, E. S. Boyd, J. B. Broderick, J. W. Peters and P. L. Roach, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 1687-1690.
20. E. M. Shepard, B. R. Duffus, S. E. McGlynn, M. R. Challand, K. D. Swanson, P. L. Roach, J. W. Peters and J. B. Broderick, *J. Am. Chem. Soc.*, 2010, **132**, 9247-9249.
21. J. M. Kuchenreuther, S. J. George, C. S. Grady-Smith, S. P. Cramer and J. R. Swartz, *PLoS ONE*, 2011, **6**, e20346.
22. R. C. Driesener, B. R. Duffus, E. M. Shepard, I. R. Bruzas, K. S. Duschene, N. J.-R. Coleman, A. P. G. Marrison, E. Salvadori, C. W. M. Kay, J. W. Peters, J. B. Broderick and P. L. Roach, *Biochemistry*, 2013, **52**, 8696-8707.
23. B. R. Duffus, S. Ghose, J. W. Peters and J. B. Broderick, *J. Am. Chem. Soc.*, 2014, **136**, 13086-13089.
24. J. M. Kuchenreuther, W. K. Myers, D. L. M. Suess, T. A. Stich, V. Pelmenschikov, S. A. Shiigi, S. P. Cramer, J. R. Swartz, R. D. Britt and S. J. George, *Science*, 2014, **343**, 424-427.
25. P. Dinis, D. L. M. Suess, S. J. Fox, J. E. Harmer, R. C. Driesener, L. De La Paz, J. R. Swartz, J. W. Essex, R. D. Britt and P. L. Roach, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 1362-1367.
26. D. L. M. Suess, C. C. Pham, I. Bürstel, J. R. Swartz, S. P. Cramer and R. D. Britt, *J. Am .Chem. Soc.*, 2016, **138**, 1146-1149.

27. A. Pagnier, L. Martin, L. Zeppieri, Y. Nicolet and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 104-109.
28. Y. Nicolet, J. K. Rubach, M. C. Posewitz, P. Amara, C. Mathevon, M. Atta, M. Fontecave and J. C. Fontecilla-Camps, *J. Biol. Chem.*, 2008, **283**, 18861-18872.
29. R. Rohac, P. Amara, A. Benjdia, L. Martin, P. Ruffié, A. Favier, O. Berteau, J.-M. Mouesca, J. C. Fontecilla-Camps and Y. Nicolet, *Nat. Chem.*, 2016, **8**, 491-500.
30. J. N. Betz, N. W. Boswell, C. J. Fugate, G. L. Holliday, E. Akiva, A. G. Scott, P. C. Babbitt, J. W. Peters, E. M. Shepard and J. B. Broderick, *Biochemistry*, 2015, **54**, 1807-1818.
31. H. Li and T. B. Rauchfuss, *J. Am. Chem. Soc.*, 2002, **124**, 726-727.
32. E. M. Shepard, S. E. McGlynn, A. L. Bueling, C. Grady-Smith, S. J. George, M. A. Winslow, S. P. Cramer, J. W. Peters and J. B. Broderick, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 10448-10453.
33. I. Czech, A. Silakov, W. Lubitz and T. Happe, *FEBS Lett.*, 2010, **584**, 638-642.
34. S. E. McGlynn, E. M. Shepard, M. A. Winslow, A. V. Naumov, K. S. Duschene, M. C. Posewitz, W. E. Broderick, J. B. Broderick and J. W. Peters, *FEBS Lett.*, 2008, **582**, 2183-2187.
35. D. W. Mulder, D. O. Ortillo, D. J. Gardenghi, A. Naumov, S. S. Ruebush, R. K. Szilagyi, B. H. Huynh, J. B. Broderick and J. W. Peters, *Biochemistry*, 2009, **48**, 6240-6248.
36. D. W. Mulder, E. S. Boyd, R. Sarma, R. K. Lange, J. A. Endrizzi, J. B. Broderick and J. W. Peters, *Nature*, 2010, **465**, 248-251.
37. E. M. Shepard, A. S. Byer, P. Aggarwal, J. N. Betz, A. G. Scott, K. A. Shisler, R. J. Usselman, G. R. Eaton, S. S. Eaton and J. B. Broderick, *Biochemistry*, 2017, **56**, 3234-3247.
38. E. M. Shepard, A. S. Byer, J. N. Betz, J. W. Peters and J. B. Broderick, *Biochemistry*, 2016, **55**, 3514-3527.
39. P. Berto, M. De Valentin, L. Cendron, F. Vallese, M. Albertini, E. Salvadori, G. M. Giacometti, D. Carbonera and P. Costantini, *Biochim. Biophys. Acta*, 2012, **1817**, 2149-2157.
40. G. Berggren, R. Garcia-Serres, X. Brazzolotto, M. Clemancey, S. Gambarelli, M. Atta, J.-M. Latour, H. L. Hernández, S. Subramanian, M. K. Johnson and M. Fontecave, *J. Biol. Inorg. Chem.*, 2014, **19**, 75-84.
41. M. Albertini, F. Vallese, M. Di Valentin, P. Berto, G. M. Giacometti, P. Costantini and D. Carbonera, *Intl. J. Hydrogen Energy*, 2014, **39**, 18574-18582.
42. M. Albertini, P. Berto, F. Vallese, M. Di Valentin, P. Costantini and D. Carbonera, *J. Phys. Chem. B*, 2015, **119**, 13680-13689.
43. M. Albertini, L. Galazzo, L. Maso, F. Vallese, P. Berto, E. De Rosa, M. Di Valentin, P. Costantini and D. Carbonera, *Top. Catal.*, 2015, **58**, 708-718.
44. J. M. Kuchenreuther, R. D. Britt and J. R. Swartz, *PLoS ONE*, 2012, **7**, e45850.
45. G. Caserta, L. Pecqueur, A. Adamska-Venkatesh, C. Papini, S. Roy, V. Artero, M. Atta, E. Reijerse, W. Lubitz and M. Fontecave, *Nat. Chem. Biol.*, 2017, **13**, 779-784.
46. J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons, G. Berggren, J. Noth, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607-609.
47. V. Artero, G. Berggren, M. Atta, G. Caserta, S. Roy, L. Pecqueur and M. Fontecave, *Acc. Chem. Res.*, 2015, **48**, 2380-2387.
48. S. Stoll and A. Schweiger, *J. Magn. Res.*, 2006, **178**, 42-55.
49. A. S. Pandey, T. V. Harris, L. J. Giles, J. W. Peters and R. K. Szilagyi, *J. Am. Chem. Soc.*, 2008, **130**, 4533-4540.
50. W.-F. Liaw, W.-T. Tsai, H.-B. Gau, C.-M. Lee, S.-Y. Chou, W.-Y. Chen and G.-H. Lee, *Inorg. Chem.*, 2003, **42**, 2783-2788.
51. E. M. Shepard, A. S. Byer and J. B. Broderick, *Biochemistry*, 2017, **56**, 4733-4734.
52. J. W. Tye, M. Y. Darensbourg and M. B. Hall, *J. Comp. Chem.*, 2006, **27**, 1454-1462.
53. J. W. Tye, M. Y. Darensbourg and M. B. Hall, *Inorg. Chem.*, 2008, **47**, 2380-2388.
54. L. Yu, C. Greco, M. Bruschi, U. Ryde, L. De Gioia and M. Reiher, *Inorg. Chem.*, 2011, **50**, 3888-3900.
55. M. Senger, S. Mebs, J. Duan, F. Wittkamp, U.-P. Apfel, J. Heberle, M. Haumann and S. T. Stripp, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 8454-8459.
56. F. Weigend and R. Ahlrichs, *J. Chem. Phys.*, 2005, **7**, 3297–3305.
57. A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem Phys.*, 1994, **100**, 5829-5835.
58. J. D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615-6620.
59. P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs and A. Vargha, *J. Am. Chem. Soc.*, 1983, **105**, 7037-7047.
60. G. Fogarasi and P. Pulay, in *Vibrational Spectra and Structure*, ed. J. R. Durig, Elsevier, Amsterdam, 1985, vol. 14, p. 125.
61. V. G. Ivanov and S. Quillard, in *Spectroscopy of Emerging Materials*, Kluwer Academic Publisher, Amsterdam, The Netherlands, 2004, vol. NAII, 165, pp. 339-350.
62. A. Klamt, *Wiley Interdisciplinary Reviews-Comp. Mol.*, 2011, **1**, 699-709.
63. A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans.*, 1993, **2**, 799-805.
64. M. Schmidt, S. M. Contakes and T. B. Rauchfuss, *J. Am. Chem. Soc.*, 1999, **121**, 9736-9737.
65. R. K. Szilagyi and M. A. Winslow, *J. Comput. Chem.*, 2006, **27**, 1385-1397.
66. I. Czech, S. Stripp, O. Sanganas, N. Leidel, T. Happe and M. Haumann, *FEBS Lett.*, 2011, **585**, 225-230.
67. A. J. Pierik, W. Roseboom, R. P. Happe, K. A. Bagley and S. P. J. Albracht, *J. Biol. Chem.*, 1999, **274**, 3331-3337.
68. A. Adamska, A. Silakov, C. Lambertz, O. Rüdiger, T. Happe, E. Reijerse and W. Lubitz, *Angew. Chem. Int. Ed.* , 2012, **51**, 11458-11462.
69. A. J. Pierik, M. Hulstein, W. R. Hagen and S. P. J. Albracht, *Eur. J. Biochem.*, 1998, **258**, 572-578.
70. B. J. Lemon and J. W. Peters, *Biochemistry*, 1999, **38**, 12969-12973.
71. N. Joshi, E. M. Shepard, A. S. Byer, K. D. Swanson, J. B. Broderick and J. W. Peters, *Febs Lett*, 2012, **586**, 3939-3943.
72. D. L. M. Suess, I. Bürstel, L. De La Paz, J. M. Kuchenreuther, C. C. Pham, S. P. Cramer, J. R. Swartz and R. D. Britt, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 11455-11460.
73. L. J. Giles, A. Grigoropoulos and R. K. Szilagyi, *J. Phys. Chem. A*, 2012, **116**, 12280-12298.

74. D. E. Schwab, C. Tard, E. Brecht, J. W. Peters, C. J. Pickett and R. K. Szilagyi, *Chem. Commun.*, 2006, DOI: 10.1039/B604994J, 3696-3698.
75. A. Silakov, E. J. Reijerse, S. P. J. Albracht, E. C. Hatchikian and W. Lubitz, *J. Am. Chem. Soc.*, 2007, **129**, 11447-11458.
76. A. Silakov, J. L. Shaw, E. J. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2010, **132**, 17578-17587.
77. C. F. Megarity, J. Esselborn, S. V. Hexter, F. Wittkamp, U.-P. Apfel, T. Happe and F. A. Armstrong, *J. Am .Chem. Soc.*, 2016, **138**, 15227-15233.