The [Fe-Fe]-hydrogenase maturation protein HydF from Thermotoga maritima is a GTPase with an iron-sulfur cluster.

Xavier Brazzolotto, Jon K. Ruback, Jacques Gaillard, Serge Gambarelli, Mohamed Atta, Marc Fontecave

THE [Fe-Fe]-HYDROGENASE MATURATION PROTEIN HydF FROM *Thermotoga maritima* IS A GTPase WITH AN IRON-SULFUR CLUSTER

Xavier Brazzolotto, Jon K. Rubach, Jacques Gaillard[§], Serge Gambarelli[§], Mohamed Atta and Marc Fontecave.

Laboratoire de Chimie et Biochimie des Centres Rédox Biologiques, UMR Université Joseph Fourier / CNRS / CEA - N° 5047, Département Réponses et Dynamiques Cellulaires, [§] Laboratoire de Chimie Inorganique et Biologique, UMR-E N° 3 CEA/Université Joseph Fourier, Département de Recherche Fondamentale sur la Matière Condensée. CEA – Grenoble, 17 Rue des Martyrs, 38054, Grenoble Cedex 9, France

Running Title : HydF Iron-only Hydrogenase maturation protein.

Correspondance to: Pr Marc Fontecave, Laboratoire de Chimie et Biochimie des Centres Rédox Biologiques, UMR 5047, Département Réponses et Dynamiques Cellulaires, CEA – Grenoble, 17 Rue des Martyrs, 38054, Grenoble Cedex 9, France. Tel. (33) 4 3878 9103. Fax. (33) 4 3878 9124. E-Mail : mfontecave@cea.fr

The active site of [Fe-Fe]-hydrogenases is composed of a di-iron complex, where the two metal atoms are bridged together by a putative di(thiomethyl)amine molecule and are also ligated by di-nuclear ligands, namely carbon monoxide and cyanide. Biosynthesis of this metal site is thought to require specific protein machinery coded by the *hyd*E, *hyd*F and *hyd*G genes. The HydF protein has been cloned from the thermophilic organism *Thermotoga maritima*, purified and characterized. The enzyme possesses specific amino-acid signatures for GTP-binding and is able to hydrolyze GTP. The anaerobically reconstituted TmHydF protein binds a [4Fe-4S] cluster with peculiar EPR characteristics: an S=1/2 signal presenting a high field shifted g-value together with a S=3/2 signal, similar to those observed for [4Fe-4S] clusters ligated by only three cysteines. HYSCORE spectroscopy experiments were carried out in order to determine the nature of the cluster's fourth ligand and its exchangeability was demonstrated with the formation of a [4Fe-4S]-imidazole complex.

Hydrogenases are metalloproteins that catalyze the reversible activation of molecular hydrogen and enable an organism to either utilize H_2 as a source of reducing power or to use protons as terminal electron acceptors, thus generating H_2 gas. Based on their metal content, hydrogenases are divided into three classes, [Ni-Fe]-hydrogenases (1), [Fe-Fe]-hydrogenases (2,3) and "Iron-Sulfur cluster-free" hydrogenase (4,5,6), which do not appear to be structurally or phylogenetically related (7).

[Fe-Fe]-hydrogenases are limited to certain anaerobic bacteria and anaerobic eukaryotes and are often involved in H_2 evolution with catalytic activities up to 100 times higher than those of [Ni-Fe]-hydrogenases (8). At their active site they contain a dinuclear iron center attached to the protein by only one bond between a cysteine residue and one of the two iron atoms (Fig. 1). This cysteine also serves as a ligand for an adjacent [4Fe-4S] cluster, so there is a sulfur bridge between the two metal sites (2,3). [Fe-Fe]-hydrogenases also contain additional [2Fe-2S] and [4Fe-4S] clusters which shuttle electrons between the H_2 activating site, inside the protein, and the redox partners at the surface.

One remarkable property of the [Ni-Fe]- and [Fe-Fe]-hydrogenase active sites is the presence of carbon monoxide and cyanide ligands as clearly established by X-ray crystallography (1,2,3) and infrared spectroscopy (9). In both cases CO and CN^- are found coordinated to the iron atoms and are thought to allow stabilization of the low iron oxidation and spin states required for activity (10). Infrared spectroscopy studies have also demonstrated the presence of CO ligands in the "Iron-Sulfur cluster-free" hydrogenase (11). In the unique case of [Fe-Fe]-hydrogenases an intriguing low-molecular weight compound, still incompletely identified but often proposed to be a di(thiomethyl)amine, is bound to the di-iron site through a bridging bidentate coordination mode (Fig. 1) (12).

The presence of potentially toxic compounds such as CO and CN^- and unique ligands, such as the bridging dithiol, suggests the involvement of specific accessory proteins during the synthesis of the hydrogenases active sites. It is a recent concept that the biosynthesis of biological metal clusters and their assembly within metalloproteins are carried out and tightly controlled by specific protein machineries. For example, the ISC and SUF systems, which are involved in the production of iron-sulfur clusters and their transfer into iron-sulfur proteins, or the NIF system, which is involved in the biosynthesis of the active site of the nitrogenase protein, have been extensively studied during the last decade (13,14,15). In the case of [Ni-Fe]-hydrogenases the maturation process is rather well characterized and involves at least seven proteins as well as chemical components such as ATP, GTP and carbamoyl phosphate, which is the precursor of cyanide (16). However there is still no clear understanding of how the CO ligands are synthesized and incorporated (17)

In contrast, the maturation of [Fe-Fe]-hydrogenases is not well understood. In fact, nothing was known in terms of accessory genes involved in the biosynthesis of the di-iron site until the very recent and important discovery of two genes required for the assembly of an active [Fe-Fe]-hydrogenase from the photosynthetic green algae *Chlamydomonas reinhardtii* (18). An interesting observation is that these genes are exclusively found in [Fe-Fe]-hydrogenase-containing microorganisms. The first gene, *hyd*EF, codes for a protein that contains two independent domains homologous to distinct prokaryotic proteins, HydE and HydF. The N-terminal domain of the HydEF protein from *C. reinhardtii* and that of all HydE proteins contain amino-acid sequences characteristic for metalloenzymes belonging to the Radical-SAM enzyme superfamily (19). The second gene, *hyd*G, encodes a protein that also contains signatures characteristic for the Radical-SAM superfamily. These enzymes are characterized by the presence of a [4Fe-4S] cluster chelated by three cysteines from a conserved CX_3CX_2C sequence and the ability of this cluster to catalyze the reductive cleavage of S-adenosylmethionine (SAM), thus initiating radical reactions (20,21,22,23). Well-known examples of Radical-SAM enzymes are the activase of ribonucleotide reductase (24), the activase of pyruvate-formate lyase (21), biotin synthase (25) and the MiaB protein (26).

In order to study the maturation process of [Fe-Fe]-hydrogenases at the molecular level, we became interested in homologs of the *C.reinhardtii hyd*EF and *hyd*G genes. In *Thermotoga maritima,* a hyperthermophilic organism possessing only [Fe-Fe]-hydrogenases, the two domains of the HydEF protein are encoded by two separate genes, *Tmhyd*E and *Tmhyd*F, and the HydG protein is encoded by a gene localized in the same operon encoding the TmHydE protein. We have initiated a study of the three proteins, HydE, HydG and HydF, in order to understand the structural and functional properties of these interesting enzymes and their specific role in the assembly of the di-iron site of [Fe-Fe]-hydrogenases. We have recently reported a preliminary characterization of the Radical-SAM proteins, HydE and HydG (27). It was confirmed that they bind a [4Fe-4S] cluster able to reductively cleave SAM and it was also shown that they contain an additional cluster whose function is so far unknown (27). Here we report the first isolation and characterization of the third component of the assembly machinery, the HydF protein from *T.maritima*. We demonstrate that it is a [4Fe-4S] iron-sulfur protein, which binds GTP and catalyzes its hydrolysis to GDP. Possible roles for the HydF protein in the maturation process are discussed.

Materials and Methods

TmHydF cloning - The open reading frame encoding the TmHydF protein (TM0445) was PCR amplified using *T.maritima* genomic DNA, *Pwo* polymerase (Roche) and the following primers: the N-terminus primer (TmHydF-F: 5'-GAGGGTGATCATATGAGATTGCCG-3') was designed to contain a unique *Nde*I restriction site at the predicted initiation codon and the C-terminus primer (TmHydF-R: 5'-GGCCCAGATGAAAGCTTCGTCGAG-3') to hybridize in the 3'-untranslated region and to contain a unique *Hin*dIII restriction site. The PCR fragment was purified with the High Pure PCR kit (Roche), double digested with *Nde*I and *Hin*dIII (Fermentas) and gel purified before direct cloning into the pT7-7 expression vector, leading to the pTmHydF plasmid. DH5α cells were transformed with the pTmHydF plasmid and one clone containing an insert with the expected size was selected for sequencing. The

cloned *hyd*F was confirmed to be wild-type when compared to the Genbank sequence AAD35530.

TmHydF expression and purification - As *T.maritima* genes use numerous rare codons that code for arginine, isoleucine or leucine residues, *E.coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) were used for expression of the TmHydF protein. Cells transformed with the pTmHydF plasmid were grown in Miller's modified LB medium at 37 °C until the OD_{600nm} reached 1, then protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.5 mM. After 5 hours, the cells were pelleted and stored at -80 °C until use.

For purification, the cells were thawed in the presence of lysozyme (0.6 mg/ml) and 0.5% (v/v) Triton X-100, sonicated and centrifuged at 180,000 x g for 1 hour. DNA was removed from the supernatant by centrifugation at 10,000 x g after precipitation with 2% streptomycin sulfate. The cell-free extract was then loaded onto a Blue-Sepharose column (Amersham Biosciences) equilibrated with 20 mM Tris-HCl buffer pH 8.0. After extensive washing with the equilibration buffer, the protein was eluted from the column using 20 mM Tris-HCl buffer pH 8.0, containing 250 mM NaCl. Fractions containing pure protein were pooled and concentrated by ultrafiltration using a YM-30 membrane. The protein was then aliquoted, flash frozen and stored at -80 °C until use.

Protein characterization - Protein concentrations were determined with the Biorad Protein Assay, using bovine serum albumin as a standard. Aerobic UV-visible absorption spectra were recorded on a Cary 1Bio spectrophotometer (Varian). Anaerobic spectrophotometric measurements were carried out within an anaerobic glove box (Jacomex) in which the cell holder is connected to an UvikonXL spectrophotometer (BioTek Instruments) outside the box by an optical fiber. Iron and sulfur quantitations were done according to the methods of Fish (28) and Beinert (29), respectively. The oligomeric state of the protein was determined in the anaerobic glove box by analytical gel filtration on a Superdex 200 HR 10/30 column (Amersham Biosciences) equilibrated in 100 mM Tris-HCl buffer pH 8.0, 0.5 M NaCl, 2 mM dithiothreitol. Mass spectrometry and N-terminal sequencing analyses were performed by the mass spectrometry and the peptide analysis facilities of the Institut de Biologie Structurale (Grenoble, France).

In vitro reconstitution - In a anaerobic glove box, the as-purified TmHydF protein was incubated with a 10-fold molar excess of dithiothreitol (Fermentas). The protein was then incubated with a 6-fold molar excess of a ferrous salt ($NH_4FeSO_4 \cdot 6H_2O$) (Sigma Aldrich), 6-fold molar excess of cysteine (Sigma Aldrich) and 0.02 molar equivalent of the recombinant purified cysteine desulfurase IscS protein from *E.coli*, available in the laboratory. The formation of the iron-sulfur cluster was followed spectrophometrically at 410 nm. When this absorption reached a plateau the protein was treated with a 6-fold molar excess of EDTA in order to remove unspecifically bound iron and then desalted on a Sephadex-G25 column equilibrated in 20 mM Tris buffer pH 8.0, 100 mM NaCl.

GTP hydrolysis assay - Under aerobic conditions, 20 µM of TmHydF protein was incubated at 37°C with 500 µM GTP (Acros) and 1 mM $MgCl_2$ in 20 mM Tris buffer, pH 8.0, containing 200 mM NaCl. At time intervals, aliquots of the reaction mixture were collected and assayed for production of GDP. The samples were analyzed by reverse phase HPLC on a ZORBAX Eclipse XDB-C8 column, (4.6 x 150 mm, 5 µm, Agilent Technologies) eluted with an isocratic mobile phase of 50 mM sodium phosphate buffer pH 7.0, 10 mM tetrabutylammonium bromide, 10 % CH_3CN. The guanosine nucleotides were detected by their absorbance at 254 nm and under these conditions GDP eluted after 5.5 min.

Electronic Paramagnetic Resonance - X-band EPR spectra were recorded on a Brucker Instrument EMX spectrometer equipped with an Oxford Instrument ESR 900 flow cryostat. Protein samples were anaerobically reduced with 2 mM sodium dithionite before introduction in the EPR tube and freezing.

HYSCORE spectroscopy - HYSCORE experiments were performed on a Bruker Elexsys E-580 X-band pulsed spectrometer with a Bruker ER4118X dielectric resonator and a continuous flow He cryostat (Oxford Instrument CF935) controlled by an Oxford Instrument temperature controller ITC 503. Experiments were performed at 10 K using the standard 4-pulse sequence ($\pi/2$-τ-$\pi/2$-t1-π-t2-$\pi/2$-echo) with a nominal pulse width of 16 ns for $\pi/2$ and π pulses, a τ value of 120 ns and a pulse repetition rate of 2 kHz. Unwanted echoes were removed by four-step phase cycling. A 128x128 dataset was recorded with times t1 and t2 incremented in 24 ns steps from an initial value of 200 ns. This dataset was processed using Xepr

software (Bruker). The background decay in both dimensions was subtracted using a linear fit followed by apodization with a Hamming window and zero-filling to 512 points in each dimension. Then the 2D Fourier Transform magnitude spectrum was calculated and presented as a contour plot.

Results

Purification of the recombinant TmHydF protein - The purification protocol described in the Materials and Methods section resulted in the production of a pure protein that migrates as a single band on a denaturating polyacrylamide gel (data not shown) and which was confirmed as being the TmHydF protein by N-terminal sequencing. The molecular weight was determined by mass spectrometry to be at 45110±3 Da, in good agreement with the mass calculated from the amino acid sequence (45091 Da). The protein behaved as a monomer in solution as shown by gel filtration on a calibrated Superdex 200 column (data not shown).

The TmHydF protein is a GTPase - Primary sequence analysis of the N-terminal domain of HydF proteins from various [Fe-Fe]-hydrogenase-containing microorganisms revealed the presence of several conserved consensus sequences, similar to those involved in guanine nucleotide binding in Small-G proteins (Fig. 2) (30,31). The first motif is the (G/A)X_4GK(T/S) sequence responsible for the binding of α- and β-phosphate groups of the nucleotide (the P loop), which is localized at residues 16-23 in the TmHydF protein. One of the three threonine residues, T43, T44 or T45, might correspond to the conserved threonine residue of the G2 loop involved in Mg^{2+} binding. The third sequence is DX_2G, residues 69-72, which provides the G3 loop involved in the interaction with the γ-phosphate and Mg^{2+}. The G4 loop, (N/T)(K/Q)XD, which is supposed to interact with the nucleotide, is present between residues 123 and 126. Finally, the G5 loop, for recognition of the guanine base, (T/G/C)(C/S)A, is more difficult to find. It could involve residues V150, S151 and A152, where the replacement of a threonine residue by a valine could be considered as an isosteric change.

GTP binding to the TmHydF protein was demonstrated by fluorescence spectroscopy. Upon excitation at 285 nm, the TmHydF protein displays an intense fluorescence band at 334 nm. Addition of GTP resulted in a quenching of the fluorescence of the protein indicating formation of a complex. The Kd value for the dissociation of the TmHydF protein-GTP complex could thus be determined by titration of the TmHydF protein with increasing concentrations of GTP, monitored by fluorescence spectroscopy. The data were best fit to the standard binding equation assuming one saturation site and allowed the determination of a value of approximately 3 µM for Kd (data not shown). GTP hydrolysis was assayed by incubating the as-purified TmHydF protein with an excess of GTP and Mg^{2+} at pH 8.0. A slow production of GDP (0.03 min^{-1}) was observed as monitored by HPLC (Fig. 3). When the protein was omitted from the reaction mixture, no production of GDP could be detected. Furthermore no ATPase activity could be observed. These data clearly demonstrate that the TmHydF protein is a GTPase, with a rather low activity as observed in the case of many Small-G proteins (30). It is important to note that this very weak activity offered the possibility to determine the Kd value for GTP as described above.

The TmHydF protein is an Iron-Sulfur protein - Concentration of solutions containing the as-purified protein resulted in the appearance of a light-brown color. In addition, analysis of an alignment of HydF protein primary sequences from different [Fe-Fe]-hydrogenase-containing microorganisms showed the complete conservation of three cysteine residues located in the C-terminal region of the protein, C302, C353 and C356 for the TmHydF protein (Fig. 4). Cysteine residues, especially within the CX_2C motif, are often observed as ligands of Fe-S clusters. The as-isolated protein was found to contain substoichiometric amounts of iron and sulfur atoms (<0.2 of each atom per equivalent protein). This is possibly due to the loss of the oxygen-sensitive cluster during aerobic purification, as is often observed with Fe-S proteins. Finally, we determined that TmHydF protein could bind up to four iron and four sulfur atoms per polypeptide chain, after anaerobic treatment with an excess of ferrous iron and sulfide and desalting (data not shown).

The UV-visible spectrum of the reconstituted TmHydF protein (Fig. 5*A*) displays a broad absorption band centered around 410 nm which can be assigned to sulfur-to-iron charge transfer transitions characteristic of a [4Fe-4S] cluster and responsible for the brown color of the protein. The presence of a [4Fe-4S] cluster in the TmHydF protein was further demonstrated by EPR spectroscopic analysis after anaerobic reduction with

an excess of dithionite. The EPR spectrum displayed several signals (Fig. 5*B* and *C*). An axial signal is seen in the g=2 region (g=2.045, g=1.904) with microwave power saturation and temperature saturation dependence properties characteristic for a reduced S=1/2 $[4Fe\text{-}4S]^{1+}$ center (Fig. 5*B*). Additional resonances were observed around g=5.2 and g=1.75, which belong to the broad spectrum of the $[4Fe\text{-}4S]^{1+}$ center in the S=3/2 spin state (32) (Fig. 5*C*). Quantitation of the g=2 signal (0.1-0.2 spin/protein) revealed that less than 20% of the $[4Fe\text{-}4S]^{1+}$ cluster was in the S=1/2 reduced state. The amount of reduced cluster in the S=3/2 state is much more difficult to appreciate considering the broadness of the signal and the difficulty to obtain the zero-fill splitting parameter. However, there is no doubt that there was a significant production of S=3/2 clusters during anaerobic treatment of the TmHydF protein with dithionite.

These results unambiguously show that the TmHydF protein is a [4Fe-4S] cluster-containing protein. So far we can not find any correlation between the presence of the cluster and the GTPase activity. Reconstituted and as-isolated TmHydF protein preparations have similar GTPase specific activities and furthermore no modification of the EPR spectrum could be observed upon addition of GTP or GDP.

An exchangeable site at the [4Fe-4S] cluster of the TmHydF protein - The EPR spectrum of the TmHydF protein is quite unusual. Not only does it show the presence of both S=1/2 and S=3/2 clusters but the high-field feature of the S=1/2 signal is centered at g=1.904, which is significantly high-field shifted with regard to standard g values for reduced [4Fe-4S] clusters, generally measured at g=1.94 (33). The observation of a protein-bound $[4Fe\text{-}4S]^{1+}$ cluster presenting a S=3/2 signal together with low g values for the S=1/2 signal (g=2.10, g=1.87, g=1.80) has been reported by Adams and coworkers for the ferredoxin from *Pyrococcus furiosus* (34). This ferredoxin is one of the rare examples of a spectroscopically well-characterized protein with a [4Fe-4S] cluster that has incomplete cysteinyl coordination, with three cysteine ligands and an aspartate as the fourth ligand (34). This raised the possibility of the presence of a non-cysteinyl ligand at the [4Fe-4S] cluster of the TmHydF protein.

To further study the ligands of the cluster, HYSCORE spectroscopy, a two-dimensional pulsed EPR technique, has been used to characterize the reconstituted protein after anaerobic reduction with dithionite. HYSCORE spectroscopy is the method of choice to detect hyperfine coupling of nuclei with low gyromagnetic moments in non-oriented systems. One of the main advantages of HYSCORE is the ability to sort three types of nuclei: the strongly ($|a|/2 > \nu_n$) and weakly ($|a|/2 < \nu_n$) coupled ones and the "distant" nuclei characterized by very low hyperfine constants. In the latter case the corresponding peaks lie on the diagonal of the (+,+) quadrant, whereas the strongly and weakly coupled nuclei appear off the diagonal in the (-,+) and (+,+) quadrants, respectively. Moreover, with a strongly coupled I=1 nucleus such as ^{14}N, it is possible to observe a characteristic pattern in the (-,+) quadrant with so-called double quanta-double quanta (dqdq) correlation features (35,36). They correspond to $\Delta I=2$ nuclear transitions and are, to the first order, insensitive to orientation broadening. It is possible to obtain a rough estimate of the hyperfine coupling constant *a* (supposed to be mainly isotropic) and quadrupolar coupling constant *K* from the frequency position of the dqdq correlation peaks using equation 1 (η being the asymmetry parameter) (37). Thus, a nitrogen atom ligated to a Fe-S cluster is easily detectable due to the strong coupling of the nuclear spin (I=1) with the unpaired electron.

$$(1) \quad \nu_{dqdq}^2 = 4\left[\left(\nu_I \pm \frac{a}{2}\right)^2 + K^2\left(3+\eta^2\right)\right]$$

The HYSCORE spectrum of the anaerobically reduced protein presented no cross correlation peaks except those corresponding to distant ^{13}C carbon atoms present in natural abundance and ^{1}H atoms (Fig. 6*A*), thus excluding a nitrogen atom in the coordination sphere of the cluster. In addition, HYSCORE spectra of the protein in D_2O did not show any strongly coupled (exchangeable) ^{2}H nucleus (data not shown) as would have been expected if the fourth ligand was a water molecule. Further experiments are required to identify the fourth ligand.

The exchangeability of the fourth ligand was assayed by incubating the reduced protein with 5 mM imidazole. The EPR spectrum of the reaction mixture did not show any significant

change. On the other hand, the HYSCORE spectrum was drastically modified (Fig. 6*B*). The most striking change was the appearance of a complex pattern of peaks in the (-,+) quadrant, characteristic of a strongly coupled nitrogen atom. In particular, it was possible to observe a pair of dqdq correlation peaks at (-2.4, 7.0) and (-7.0, 2.4) MHz, respectively. Using equation (1), 4.8 MHz was obtained for the hyperfine coupling constant *a* and 0.53 MHz for the quadrupolar coupling constant *K*. These values are in the range of those obtained for Rieske proteins, where the N_δ histidine nitrogen is ligated to the [2Fe-2S] cluster (38,39). This suggests that imidazole is N-ligated to the fourth iron of the cluster in a very similar fashion to that observed for the histidine ligand in the Rieske [2Fe-2S] cluster. These results also show that the [4Fe-4S] cluster is solvent accessible, as its fourth ligand is easily exchangeable with imidazole.

Discussion

The results presented here provide the first detailed characterization of the HydF protein, which is involved in the maturation of [Fe-Fe]-hydrogenases. We have clearly established that the TmHydF protein is an iron-sulfur protein carrying a [4Fe-4S] cluster, on the basis of iron and sulfur quantitation, UV-visible and EPR spectroscopy. It is very likely that this cluster is ligated by only three cysteines and that these cysteines are the three conserved ones, C302, C353 and C356, the latter two belonging to a CX_2C motif, which is frequently involved in cluster chelation. The presence of a non-cysteinyl ligand is also indicated by the significantly low g value (g=1.904) in the EPR spectum of the S=1/2 reduced cluster, as compared to the spectra of clusters with four cysteine ligands. Also, the presence of a significant amount of reduced clusters in the S=3/2 spin state suggests an unusual coordination. There are few examples of [4Fe-4S] clusters ligated by a protein site different from a cysteine. An interesting one is provided by hydrogenases themselves. The three-dimensional structures of [Fe-Fe]-hydrogenase and [Ni-Fe]-hydrogenase reveal the presence of a [4Fe-4S] cluster coordinated by a histidine residue, through the N_ε atom in the first case and the N_δ atom in the second case (2,1). Hydrogenases are difficult to investigate spectroscopically due to the large number of Fe atoms and clusters within the same enzyme and, unfortunately, these histidine-ligated clusters have never been studied in the reduced state and thus nothing is known about their EPR characteristics. The second example is the cluster found in the ferredoxin from *P.furiosus*, see results section, which has a cluster that is ligated by a carboxylate from an aspartate residue (34). From the HYSCORE results reported we definitively exclude a nitrogen-based ligand from the protein. Thus an O-coordination provided by the protein would be a possibility. Also, there are several examples of [4Fe-4S] clusters ligated by three cysteines and an exogenous ligand. The most extensively studied of such clusters are those found in aconitase, with the citrate substrate bound to the fourth iron, or in Radical-SAM enzymes, with S-adenosylmethionine coordinated to one iron through its aminocarboxyl moiety (40,41). It is thus also possible that in the purified protein studied here, the fourth coordination site is occupied by a low-molecular weight compound, a cofactor or a substrate, not yet identified. Also based on HYSCORE experiments, we exclude the possibility that the fourth ligand is a water molecule (see above).

The observation of a cluster-imidazole complex during incubation of the TmHydF protein with an excess of imidazole shows that the fourth site is easily accessible by exogenous ligands and readily exchangeable. This [4Fe-4S]-imidazole complex is an interesting novel species since it would provide the first opportunity to spectrospically characterize an interaction between a [4Fe-4S] cluster and a nitrogen ligand such as imidazole or histidine, as found in some enzymes such as hydrogenases (see above) and to identify specific signatures for these entities that differentiate them from standard clusters. This is currently under investigation.

The second important observation reported here is the ability of the HydF protein to bind GTP (Kd=3 µM) and to catalyze GTP hydrolysis. This GTPase activity requires Mg^{2+} and is consistent with the presence, in the N-terminal part of HydF proteins, of several conserved consensus sequences found in Small-G proteins. So far we have been unable to find any interdependency between the presence of the [4Fe-4S] cluster and the GTPase activity. In particular, similar enzyme activities were observed using preparations of the TmHydF protein with different cluster contents. Iron-sulfur proteins with the ability to bind and hydrolyze nucleoside

triphosphates (ATP or GTP) are rare. One fascinating example is the bacterial nitrogenase which couples the energy of ATP binding and hydrolysis to electron-transfer reactions (42). In this case the Fe protein of nitrogenase (NifH) binds the nucleotide and a single [4Fe-4S] cluster, and the electron transfer from that cluster to the MoFe protein where N_2 is reduced is driven by the hydrolysis of two molecules of ATP. Future experiments will aim at finding whether the cluster in HydF proteins has a redox function and whether the latter is controlled by GTP binding and hydrolysis.

Finally, an important issue is how the HydF protein participates to the maturation of [Fe-Fe]-hydrogenases. These enzymes harbor a very peculiar active site with CO, CN ligands to the iron atoms as well as a putative di(thiomethyl)amine bridging the two irons (Fig. 1). Furthermore, this di-iron complex shares with an adjacent [4Fe-4S] cluster a sulfur ligand provided by one bridging cysteine. Activation of hydrogenase thus requires synthesis of these components and assembly within the active site. It is too early to identify where the HydF protein participates in this process. However, it is interesting to note that a protein, with GTPase activity and Kd values for GTP comparable to those displayed by the TmHydF protein, is involved in the maturation of [Ni-Fe]-hydrogenase, which also contains a dinuclear metal site (consisting of a Ni atom and an Fe atom with CO and CN ligands). It has been shown that this protein, HypB, participates in the GTP-dependent insertion of the nickel atom into an hydrogenase form only containing the $Fe(CO)(CN)_2$ motif (16). It is thus tempting to speculate that the HydF protein is also involved in the GTPase-driven insertion of the diiron complex, or part of it or of the [4Fe-4S] cluster to which the diiron complex binds, into the [Fe-Fe]-hydrogenase active site. It seems to be a common theme that a nucleotidase activity is required in metallosite biosynthesis. In addition to the HydF and HypB proteins, one can mention the CooC and UreG proteins, involved in nickel insertion during maturation of carbon monoxide dehydrogenase (43) and urease (44), respectively, as well as NifH which also functions in the ATP-dependent insertion of the FeMo-Co into aponitrogenase (45).

Together with the purification and the characterization of the HydE and HydG proteins that we recently reported (27), this work on the HydF protein provides a strong basis to design an *in vitro* system for maturation of [Fe-Fe]-hydrogenases and to understand the mechanism of the assembly of the diiron active site. It is remarkable that this process depends so much on Fe-S clusters, two from HydE, two from HydG and one from HydF.

REFERENCES

1. Volbeda, A., Charon, M.H., Piras, C., Hatchikian, E.C., Frey, M., and Fontecilla-Camps, J.C. (1995) *Nature* **373**, 580-587
2. Peters, J.W., Lanzilotta, W.N., Lemon, B.J.., and Seefeldt, L.C. (1998) *Science* **282**, 1853-1858
3. Nicolet, Y., Piras, C., Legrand, P., Hatchikian, E.C., and Fontecilla-Camps, J.C. (1999) *Struct. Fold. Des.* **7**, 13-23
4. Lyon, E.J., Shima, S., Buurman, G., Chowdhuri, S.., Batschauer, A., Steinbach, K., and Thauer, R.K. (2004) *Eur. J. Biochem.* **271**, 195-204
5. Shima, S., Lyon E.J., Thauer, R.K., Mienert, B., and Bill E. (2005) *J. Am. Chem. Soc.* **127**, 10430-10435
6. Shima, S., Lyon, E.J., Sordel-klippert, M., Kauβ, M., Kahnt, J., Thauer, R..K., Steinbach, K., Xie, X., Verdier, L., and Griersinger, C. (2004) *Angew. Chem. Int. Ed.* **43**, 2547-2551
7. Vignais, P.M., Billoud, B., and Meyer, J. (2001) *FEMS Microbiol. Rev.* **25**, 455-501
8. Adams, M.W. (1990) *Biochim. Biophys. Acta.* **1020**, 115-145

9. Van der Spek, T.M., Arendsen, A.F., Happe, R.P., Yun, S., Bagley, K.A., Stufkens, D..J., Hagen, W.R., and Albracht, S.P. (1996) *Eur. J. Biochem.* **237**, 629-634
10. Pierik, A.J., Hulstein, M., Hagen, W.R., and Albracht, S.P. (1998) *Eur. J. Biochem.* **258**, 572-578
11. Lyon, E.J., Shima, S., Boecker, R., Thauer, R.K., Grevels, F.-W., Bill, E., Roseboom, W., and Albracht, S.P. (2004) *J. Am. Chem. Soc.* **126**, 14239-14248
12. Nicolet, Y., Cavazza, C., and Fontecilla-Camps, J.C. (2002) *J. Inorg. Biochem.* **91**, 1-8
13. Frazzon, J., and Dean, D.R. (2003) *Curr. Opin. Chem. Biol.* **7**, 166-173
14. Johnson, D.C., Dean, D.R., Smith, A.D., and Johnson, M.K. (2005) *Annu. Rev. Biochem.* **74**, 247-281
15. Rubio, L.M., and Ludden, P.W. (2005) *J. Bacteriol.* **187**, 405-414
16. Blokesch, M., Paschos, A., Theodoratou, E., Bauer, A., Hube, M.., Huth, S., and Böck, A. (2002) *Biochem. Soc. Trans.* **30**, 674-680
17. Roseboom, W., Blokesch, M., Böck, A., and Albracht, S.P. (2005) *FEBS Lett.* **579**, 469-472
18. Posewitz, M.C., King, P.W., Smolinski, S.L., Zhang, L., Seibert, M., and Ghirardi, M.L. (2004) *J. Biol. Chem.* **279**, 25711-25720
19. Sofia, H.J., Chen, G., Hetzler, B.G., Reyes-Spindola, J.F., and Miller, N.E. (2001) *Nucleic Acids Res.* **29**, 1097-1106
20. Fontecave, M., Mulliez, E., and Ollagnier-de-Choudens, S. (2001) *Curr. Opin. Chem. Biol.* **5**, 506-511
21. Cheek, J., and Broderick, J.B. (2001) *J. Biol. Inorg. Chem.* **6**, 209-226
22. Frey, P., and Magnusson, O.T. (2003) *Chem. Rev.* **103**, 2129-2148
23. Jarret, J.T. (2003) *Curr. Opin. Chem. Biol.* **7**, 174-182
24. Fontecave, M., Mulliez, E., and Logan, D.T. (2002) *Prog. Nucleic Acids Res. Mol. Biol.* **72**, 95-127
25. Fontecave, M., Ollagnier-de-Choudens, S., and Mulliez, E. (2003) *Chem. Rev.* **103**, 2149-2166
26. Pierrel, F., Douki, T., Fontecave, M., and Atta, M. (2004) *J. Biol. Chem.* **279**, 47555-47563
27. Rubach, J.K., Brazzolotto, X., Gaillard, J. and Fontecave, M. (2005) *FEBS Lett.* **579**, 5055-5060
28. Fish, W.W. (1988) *Methods Enzymol.* **158**, 357-364
29. Beinert, H. (1983) *Anal. Biochem.* **131**, 373-378
30. Paduch, M., Jelen, F., and Otlewski, J. (2001) *Acta Biochim. Pol.* **48**, 829-850
31. Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2005) *J. Mol. Biol.* **317**, 41-72
32. Flint, D.H., Emptage, M.H., Finnegan, M.G., Fu, W., and Johnson, M.K. (1993) *J. Biol. Chem.* **268**, 14732-14742
33. Guigliarelli, B. and Bertrand, P. (1999) *Adv. Inorg. Chem.* **47**, 421-497
34. Conover, R.C., Kowal, A.T., Fu, W., Park, J.B., Aono, S., Adams, M.W., Johnson, M.K. (1990) *J. Biol. Chem.* **265**, 8533-8541
35. Schweiger, A., and Jeschke, G. (2001) *Principles of pulse electron paramagnetic resonance*, Oxford University Press.
36. Maryasov, A.G., and Bowman, M.K. (2004) *J. Phys. Chem. B* **108**, 9412-9420
37. Tyryshkin, A.M., Dikanov, S.A., Reijerse, E.J., Burgard, C., and Hüttermann, J. (1999) *J. Am. Chem. Soc.* **121**, 3396-3406
38. Dikanov, S.A., Xun, L., Karpiel, A.B., Tyryshkin, A.M., and Bowman, M.K. (1996) *J. Am. Chem. Soc.* **118**, 8408-8416
39. Samoilova, R.I., Kolling, D., Uzawa, T., Iwasaki, T., Crofts, A.R., and Dikanov. S.A. (2002) *J. Biol. Chem.* **277**, 4605-4608.
40. Layer, G., Moser, J., Heinz, D.W., Jahn., D., and Schubert, W.D. (2003) *EMBO J.* **22**, 6214-6224
41. Hanzelmann, P., and Schindelin, H. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 12870-12875
42. Rees, D.C., and Howard, J.B. (2000) *Curr. Opin. Chem. Biol.* **4**, 559-566
43. Jeon, W.B., Cheng, J., and Ludden, P.W. (2001) *J. Biol. Chem.* **276**, 38602-38609
44. Moncrief, M.B.C., and Hausinger, R.P. (1997) *J. Bacteriol.* **179**, 4081-4086
45. Rangaraj, P., Ryle, M.J., Lanzilotta, W.M., Ludden, P.W., and Shah, V.K. (1999) *J. Biol. Chem.* **274**, 19778-19784

FOOTNOTES

XB was supported by a post-doctoral fellowship from the Centre National de la Recherche Scientifique and JKR was supported by a post-doctoral grant from the National Science Foundation. The authors would like to thank Pr R. Huber (Regensburg University, Germany) for the gift of *Thermotoga maritima* genomic DNA, and all the members of the Fontecave's Lab for helpful discussions.

[1]The abbreviations used are: PCR, polymerase chain reaction; EDTA, ethylenediaminetetraacetic acid.

FIGURE LEGENDS

Figure 1. Schematic representation of the [Fe-Fe]-hydrogenase active site complex.

Figure 2. Multiple sequence alignment of the N-terminal GTP-binding motifs of HydF proteins from different [Fe-Fe]-hydrogenase containing organisms. Consensus sequences for tri-phosphate nucleotide binding are shaded in gray.

Figure 3. GTP hydrolysis catalyzed by the TmHydF protein. 20 µM of as-purified (*filled circles*) or reconstituted (*open circles*) TmHydF protein was incubated at 37°C with 500 µM GTP and 1 mM $MgCl_2$ in 20 mM Tris buffer, pH 8.0, containing 200 mM NaCl.

Figure 4. Multiple sequence alignment of a single region located in the C-terminal end of HydF proteins from different [Fe-Fe]-hydrogenase containing organisms. Conserved cysteine residues are shaded in black.

Figure 5. Spectroscopic characterization of the anaerobically reconstituted TmHydF protein. *A*, UV-visible spectrum; *B and C*, X-band EPR spectra of the anaerobically reduced protein. Amplitude modulation 1 mT (*B*, temperature 10K, microwave power 1 mW; *C*, temperature 4K, microwave power 4 mW, the central peaks (*) correspond to the S=1/2 cluster).

Figure 6. HYSCORE spectra of the anaerobically reduced TmHydF protein, recorded at g=1.904. *A*, before; *B*, after addition of 5 mM imidazole.

Figure 1

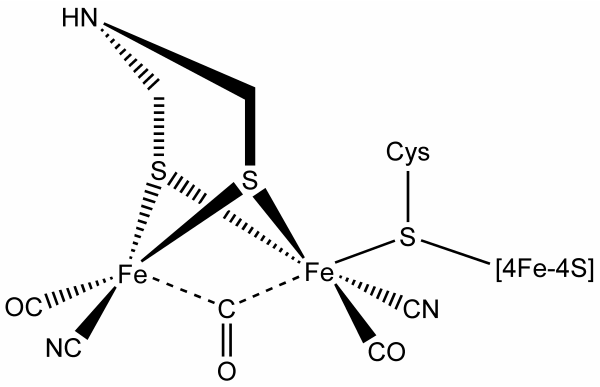


Figure 2

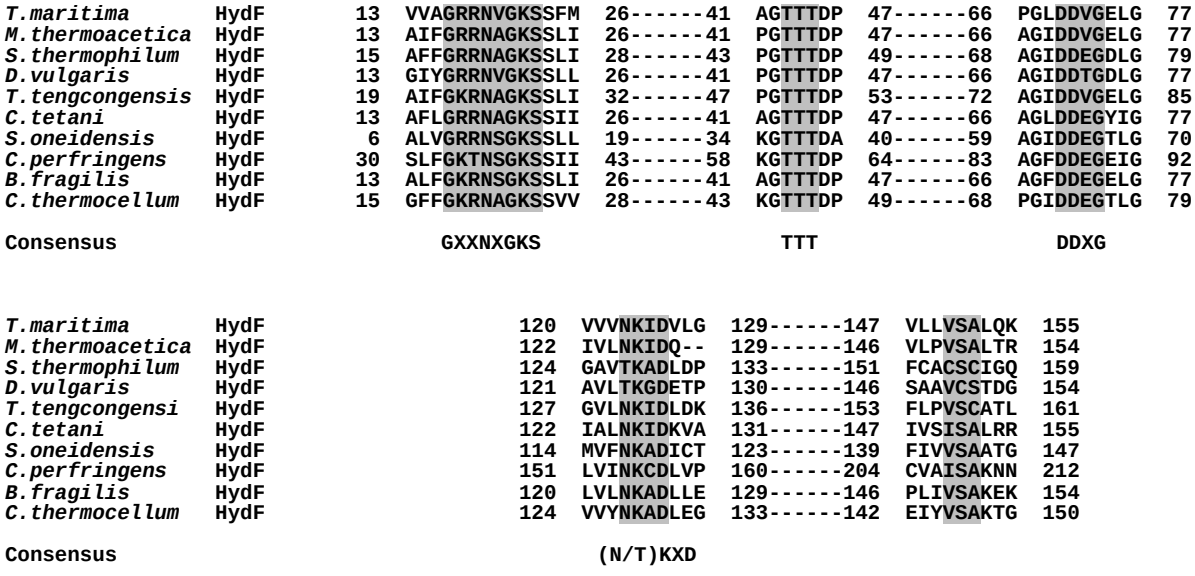

T.maritima HydF 13 VVAGRRNVGKSSFM 26------41 AGTTTDP 47------66 PGLDDVGELG 77
M.thermoacetica HydF 13 AIFGRRNAGKSSLI 26------41 PGTTTDP 47------66 AGIDDVGELG 77
S.thermophilum HydF 15 AFFGRRNAGKSSLI 28------43 PGTTTDP 49------68 AGIDDEGDLG 79
D.vulgaris HydF 13 GIYGRRNVGKSSLL 26------41 PGTTTDP 47------66 AGIDDTGDLG 77
T.tengcongensis HydF 19 AIFGKRNAGKSSLI 32------47 PGTTTDP 53------72 AGIDDVGELG 85
C.tetani HydF 13 AFLGRRNAGKSSII 26------41 AGTTTDP 47------66 AGLDDEGYIG 77
S.oneidensis HydF 6 ALVGRRNSGKSSLL 19------34 KGTTTDA 40------59 AGIDDEGTLG 70
C.perfringens HydF 30 SLFGKTNSGKSSII 43------58 KGTTTDP 64------83 AGFDDEGEIG 92
B.fragilis HydF 13 ALFGKRNSGKSSLI 26------41 AGTTTDP 47------66 AGFDDEGELG 77
C.thermocellum HydF 15 GFFGKRNAGKSSVV 28------43 KGTTTDP 49------68 PGIDDEGTLG 79
Consensus GXXNXGKS TTT DDXG
T.maritima HydF 120 VVVNKIDVLG 129------147 VLLVSALQK 155
M.thermoacetica HydF 122 IVLNKIDQ-- 129------146 VLPVSALTR 154
S.thermophilum HydF 124 GAVTKADLDP 133------151 FCACSCIGQ 159
D.vulgaris HydF 121 AVLTKGDETP 130------146 SAAVCSTDG 154
T.tengcongensi HydF 127 GVLNKIDLDK 136------153 FLPVSCATL 161
C.tetani HydF 122 IALNKIDKVA 131------147 IVSISALRR 155
S.oneidensis HydF 114 MVFNKADICT 123------139 FIVVSAATG 147
C.perfringens HydF 151 LVINKCDLVP 160------204 CVAISAKNN 212
B.fragilis HydF 120 LVLNKADLLE 129------146 PLIVSAKEK 154
C.thermocellum HydF 124 VVYNKADLEG 133------142 EIYVSAKTG 150
Consensus (N/T)KXD

Figure 3

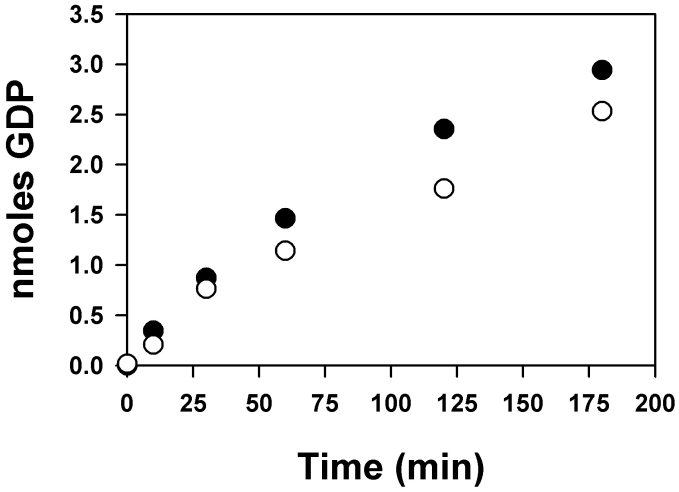

nmoles GDP
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
25
50
75
100
125
150
175
200
Time (min)

Figure 4


T.maritima 298.. IMEGCTHRPLTEDIGRVKIPRWLVNHTGAQLNFKVIAGKDFPDLEEIEGAKLIIHCGGCV..357
M.thermoacetica 297.. IAEACTHHRQADDIGKVQIPRWLRQYVGGDLQFSWYSGSGFP--EDVASYKVIIHCGGCM..354
S.thermophilum 303.. IAETCTHHQQKDDIGTVKIPRWLRERVGGELQFEWAHGGEMP--EDLSPYRLIVHCGGCM..360
D.psychrophila 305.. IAEACTHHATEDDIGRVKIPRWLRKYSGKKVEIDVFSGRDFP--ENLEDYKLIIHCGSCM..362
D.vulgaris 299.. VAEACTHHPLADDIGRVKIPRWLRQYAGGNIHIDTFAGKDMP--ADLSGYRLIIHCGGCV..356
B.thetaiotaomicron 309.. IAEACTHAPLSEDIGRVKLPHLLRKRIGEKLSIDIVAGTDFP--QDLTPYSLVIHCGACM..366
T.tengcongensis 305.. IAEACTHHRQSDDIGTVKIPRWIRQIAGFDINFEWVSGYGYK--KDLSKYALIVHCGGCM..362
C.tetani 298.. ISEACTHHSLKGDIAKEKIPNLLKKKIGGEVNIDFSSGEDFT--KNIEKYKLIIHCGGCM..355
S.oneidensis 290.. ISEACSHNVQEDDIGRVKLPRWINSYTGKQLEFVVTSGHDFP--NDLEQYALVIHCGACM..357
C.perfringens 356.. IAEACTHHQLKGDIAREKLPTWLEETC-PGIIVHNCSGKDFP--KNLNEYALVIHCGGCM..412
B.fragilis 298.. IAEACTHAPLSEDIGRVKLPRMLRKRIGEKLHIDIVSGNDFP--KDLNCYDLIIHCGACM..355
C.thermocellum 294.. ISEGCTHHRQCDDIGTVKLPRWINNYTKKNLNFEFTSGTEFP--EDLTRYKLIVHCGGCM..351
C.acetobutylicum 300.. IAEACTHHRQSDDIGKVKIPRWLRQKTGKKLEFDFSSGFSFP--PNIEDYALIVHCAGCM..357

Figure 5

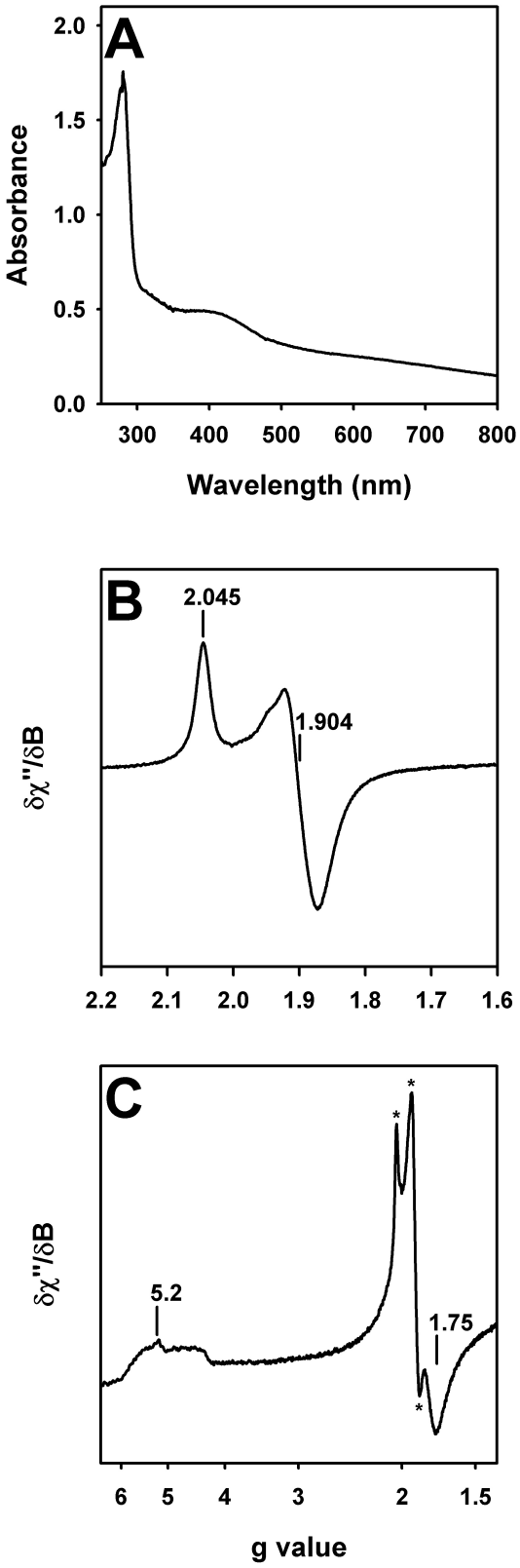
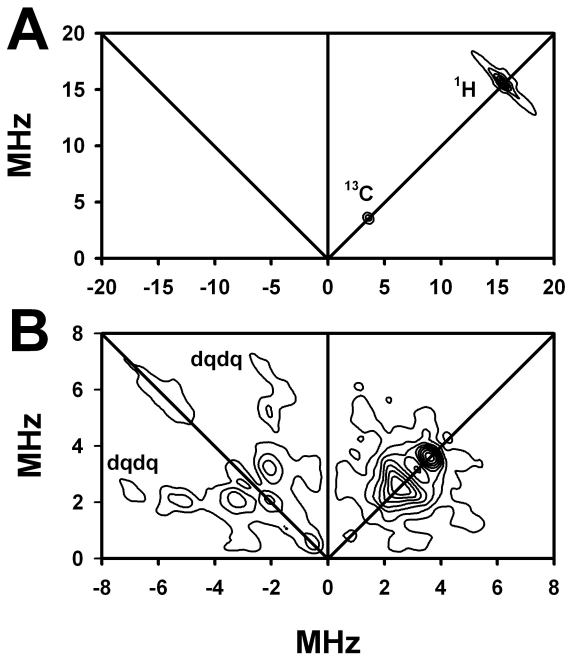

A
Absorbance
2.0
1.5
1.0
0.5
0.0
300
400
500
600
700
800
Wavelength (nm)
B
2.045
1.904
δχ''/δB
2.2
2.1
2.0
1.9
1.8
1.7
1.6
C
*
*
*
5.2
1.75
δχ''/δB
6
5
4
3
2
1.5
g value

Figure 6


A
B
MHz
MHz
MHz
^{1}H
^{13}C
dqdq
dqdq
20
15
10
5
0
-20
-15
-10
-5
0
5
10
15
20
8
6
4
2
0
-8
-6
-4
-2
0
2
4
6
8