Heterologous Biosynthesis and Characterization of the [2Fe-2S]-Containing N-Terminal Domain of *Clostridium pasteurianum* Hydrogenase†

Mohamed Atta,‡ Meghan E. Lafferty,§ Michael K. Johnson,§ Jacques Gaillard,|| and Jacques Meyer*,‡

Département de Biologie Moléculaire et Structurale, CEA-Grenoble, 38054 Grenoble, France, Département de Recherche Fondamentale sur la Matière Condensée, CEA-Grenoble, 38054 Grenoble, France, and Department of Chemistry and Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602

Received June 1, 1998

ABSTRACT: The primary structure of *Clostridium pasteurianum* hydrogenase I appears to be composed of modules suggesting that the various iron−sulfur clusters present in this enzyme might be segregated in structurally distinct domains. On the basis of this observation, a gene fragment encoding the 76 N-terminal residues of this enzyme has been expressed in *Escherichia coli*. The polypeptide thus produced contains a $[2Fe\text{-}2S]^{n+}$ cluster of which the oxidized level ($n = 2$) has been monitored by UV−visible absorption, circular dichroism, and resonance Raman spectroscopy. This cluster can be reduced by dithionite or electrochemically to the $n = 1$ level which has been investigated by EPR and by low-temperature magnetic circular dichroism. The redox potential of the +2 to +1 transition is −400 mV (vs the normal hydrogen electrode). The spectroscopic and redox results indicate a $[2Fe\text{-}2S]^{2+/+}$ chromophore coordinated by four cysteine ligands in a protein fold similar to that found in plant- and mammalian-type ferredoxins. Among the five cysteines present in the N-terminal hydrogenase fragment, four (in positions 34, 46, 49, and 62) are conserved in other sequences and are therefore the most likely ligands of the [2Fe-2S] site. The fifth cysteine, in position 39, can be dismissed on the grounds that the Cys39Ala mutation does not alter any of the properties of the iron−sulfur cluster. The spectroscopic signatures of this chromophore are practically identical with some of those reported for full-size hydrogenase. This confirms that *C. pasteurianum* hydrogenase I contains a [2Fe-2S] cluster and indicates that the polypeptide fold around the metal site of the N-terminal fragment is very similar, if not identical, to that occurring in the full-size protein. The N-terminal sequence of this hydrogenase is homologous to sequences of a number of proteins or protein domains, including a subunit of NADH-ubiquinone oxidoreductase of respiratory chains. From that, it can be anticipated that the structural domain isolated and described here is a building block of electron transfer complexes involved in various bioenergetic processes.

Hydrogenases, which catalyze the $H_2 \leftrightarrow 2H^+ + 2e^-$ reaction, are metalloenzymes that can be divided into two classes, the NiFe and the Fe enzymes, on the basis of metal content and sequence homology (*1*). The structure (*2*−*4*) and molecular biology (*1*, *5*) of NiFe hydrogenases are being elucidated at a fast pace. In comparison, Fe hydrogenases are still poorly understood. Only a few primary structures are known (*6*−*11*), and even the metal cluster content of these enzymes is not firmly established (*12*). The hydrogen-activating site (H cluster) of Fe hydrogenases, which may be a totally novel structure comprising up to six to eight iron atoms, has long been challenging bioinorganic chemists. The remarkable catalytic activity of Fe hydrogenases, the dihydrogen production turnover of which is 1 or 2 orders of magnitude higher than that of NiFe hydrogenases (*12*), is a further incentive for the elucidation of their active site structure. While a number of investigations have been carried out with native enzymes, the relatively low stability and in vivo concentration of Fe hydrogenases, together with the need for protein engineering, calls for the overexpression of hydrogenase genes in convenient hosts. Serious difficulties have been encountered in such attempts (*10*, *13*; our unpublished results), the limiting step appearing to be the incorporation of the hydrogen-activating site (*13*). Since the putative gene products involved in this process are still unknown, such obstacles are not expected to be easily overcome. Another significant line of investigation concerns the metal clusters other than the hydrogen-activating site, which are most likely involved in electron transfer. The identification and characterization of these accessory Fe−S clusters will provide a basis for sorting out the contributions of the various metal clusters in analytical and spectroscopic data for the whole enzyme.

A remarkable feature of Fe hydrogenase sequences, which has emerged from the first sequence determination (*6*) and has been strongly substantiated when significant comparisons became possible (*7*), is their modular structure. The large subunits of the heterodimeric enzymes (e.g., *Desulfovibrio vulgaris*; *6*) and the monomeric enzymes (e.g., *Clostridium*

† This work was supported by a grant from the National Institutes of Health (GM51962 to M.K.J.).

* To whom correspondence should be addressed: DBMS-Métalloprotéines, CEA-Grenoble, 38054 Grenoble, France. Fax: (33) 4 76 88 58 72. E-mail: jacques.meyer@cea.fr.

‡ Département de Biologie Moléculaire et Structurale, CEA-Grenoble.

§ University of Georgia.

|| Département de Recherche Fondamentale sur la Matière Condensée, CEA-Grenoble.

10.1021/bi9812928 CCC: $15.00 © 1998 American Chemical Society
Published on Web 10/16/1998

pasteurianum; *7*) are homologous and consist of a conserved C-terminal domain and of a more variable N-terminal domain (see Figure 4 in ref *7*). The former, which bears no similarity to other protein sequences, is believed to accommodate the hydrogen-activating site (H cluster). The N-terminal domains of all Fe hydrogenases consist of at least an 80−100-residue region which, according to its homology with bacterial ferredoxins (*14*), is most likely to contain two [4Fe-4S] clusters. A few hydrogenases (e.g., *C. pasteurianum*) contain an additional N-terminal extension of ca. 130 amino acids which displays no canonic cysteine residue pattern. However, some sequence similarities suggest that the first half of this extension might accommodate a [2Fe-2S] cluster, and the second half a [4Fe-4S] cluster (*1*).

The sequence data discussed above suggest that *C. pasteurianum* hydrogenase may contain, in addition to the hydrogen-activating site, three [4Fe-4S] clusters and one [2Fe-2S] cluster. Such a cluster composition was first suggested on the basis of resonance Raman data (*15*), which were used to re-evaluate earlier analytical, EPR,[1] Mössbauer, and VTMCD data that had been interpreted in terms of four $[4Fe\text{-}4S]^{2+/+}$ clusters (*12*, *16*, *17*). Clearly, there is a pressing need for more detailed characterization of the Fe−S clusters in the enzyme, particularly those of which most, if not all, of the putative ligands occur in the 130-residue N-terminal region of the sequence. We have chosen an approach based on the premise that the modular structure of the sequence possibly reflects a composite three-dimensional structure (*18*), as observed for instance in the aldehyde oxidoreductase from *Desulfovibrio gigas* (*19*). In such case, at least some of the various Fe−S binding domains might fold and assemble their metal sites independently of the remainder of the protein. Since the main issue with regard to the accessory Fe−S sites of *C. pasteurianum* hydrogenase concerns the presence and location of the putative [2Fe-2S] cluster, we have focused our interest on a short N-terminal fragment of *C. pasteurianum* hydrogenase which is most likely to contain such a site. We here report the expression in *Escherichia coli* of the corresponding gene fragment and demonstrate that the N-terminal fragment of the enzyme does indeed contain a [2Fe-2S] cluster.

MATERIALS AND METHODS

All common DNA manipulations were like those described previously (*20*−*22*). Enzymes were purchased from Boehringer Mannheim. Oligonucleotides and culture media were from Life Technologies (Cergy-Pontoise, France).

The hydrogenase-encoding gene was lifted by polymerase chain reaction (PCR) from the 2.3 kb *Sau*3A fragment of genomic DNA initially cloned in pUC18 (pCPH31 plasmid; *7*, *23*) and subcloned in pT7-7 (*24*). The following primers have been implemented: 5′-cttaaaattaaatctaggaggctacatATGaaaac-3′ (*Nde*I site underlined, ATG codon in uppercase) hybridized to the noncoding strand at the 5′-terminus of the gene and 5′-gtattataataaagcttgataaaaatttaac-3′ (*Hin*dIII site underlined) hybridized to the coding strand ca. 30 bp downstream of the transcription termination signal (*7*). PCRs were carried out on a Perkin-Elmer 2400 machine, under the following conditions. After a 4 min denaturation at 94 °C, the *Pwo* DNA polymerase and the deoxynucleotide mix were added, and 25 cycles (1 min at 94 °C, 1 min at 50 °C, and 3 min at 72 °C) were run followed by a 7 min elongation step at 72 °C. The PCR product was digested with *Nde*I and *Hin*dIII and cloned as described previously (*21*) in pT7-7 digested with the same enzymes. The cloned gene was completely sequenced to verify that the PCR had not introduced any errors. The pT7-7-derived plasmid containing the whole hydrogenase gene (pT7CPH) was then used to amplify by PCR and clone into pT7-7 the gene fragment encoding the 76 N-terminal amino acids (plasmid pT7HN76). The primers were the one used above which hybridized to the 5′-end of the gene and another one hybridizing to the coding strand in the region of codon 76. A stop codon (taa) was introduced to replace codon 77 (aspartate), as well as a *Hin*dIII cloning site. The sequence of the latter oligonucleotide was 5′-gatttaagcttttcgttgacagcTTAggaatttg-3′ (*Hin*dIII site underlined, stop codon in uppercase). PCR amplification, digestion, cloning, and sequence verification were performed as described above. The C39A mutated pT7HN76 plasmid was prepared by using two successive rounds of PCR, as described for the variants of the [2Fe-2S] *C. pasteurianum* ferredoxin (*22*). The mutagenic oligonucleotide 5′-gttttttaaataatGCtaataatgac-3′ (mutated bases in uppercase) hybridized to the noncoding strand.

The pT7HN76 plasmid was used to transform *E. coli* K38 (HfrC λ) cells harboring the pGP1-2 plasmid (*24*) which were grown at 30 °C in Luria Broth or Terrific Broth supplemented with 100 mg/L ampicillin and 50 mg/L kanamycin. When OD_{600} reached 1, the temperature was raised to 42 °C for 1 h, and the incubation was subsequently carried out overnight at 30 °C. Cells were collected by centrifugation, resuspended in 50 mM Tris-HCl (pH 8), and stored at −70 °C. The N-terminal hydrogenase fragment (hereafter referred to as the HN76 domain or HN76 protein) was purified anaerobically at room temperature as follows. The frozen cells were thawed, broken by sonication, and centrifuged at 45 000 rpm for 1 h. Streptomycin sulfate (2% w/v) was added to the soluble cell extract to precipitate nucleic acids which were removed by centrifugation. The supernatant was loaded on a 10 cm × 2 cm anion exchange (DE 52, Whatman) column equilibrated with 50 mM Tris-HCl (pH 8). The HN76 protein was clearly visible as a reddish brown band at the top of the column. The column was washed with 5 volumes of buffer A or 5 volumes of buffer A supplemented with 0.2 M NaCl, and the HN76 protein-containing fraction was subsequently eluted with 0.4 M NaCl and 50 mM Tris-HCl (pH 8) and concentrated in an Amicon cell fitted with a YM 3 (Spectrapor) membrane. The concentrated solution (2−3 mL) was chromatographed on a 100 cm × 2.5 cm Sephadex G50 (Pharmacia) column equilibrated with 0.2 M NaCl and 50 mM Tris-HCl (pH 8). The last step, which yielded pure protein, as assessed by SDS−PAGE, consisted of anion exchange HPLC chromatography on a PL-SAX (Polymer Labs) column developed with a 0 to 1 M NaCl gradient in 50 mM Tris-HCl (pH 8).

Electrospray ionization mass spectrometry was performed as described previously (*25*, *26*).

[1] Abbreviations: CD, circular dichroism; *Cp*, *Clostridium pasteurianum*; EPR, electron paramagnetic resonance; Fd, ferredoxin; HN76, N-terminal (residues 1−76) domain of *C. pasteurianum* hydrogenase I; PCR, polymerase chain reaction; RR, resonance Raman; VTMCD, variable-temperature magnetic circular dichroism; WT, wild-type.

```
                 1         2         3         4         5         6         7     7
           1     0         0         0         0         0         0         0     6
HN76       MKTIIINGVQFNTDEDT.TILKFARDNNIDISALCFLNNCNNDINKCEICTVEVEGTGLVT.ACDTLIEDGMIINTNS
           ** *:::* ::  :::   :*   ::  *:*  **:*      :  * :* **::*    *  *     :** : **:
Thma       MK.IYVDGREVIINDNERNLLEALKNVGIEIPNLCYLSEASI.YGACRMCLVEINGQIT.T.SCTLKPYEGMKVKTNT
             * :**: :: ::  * *:::  : *: ** **** :    :* ** * * :**::     **::   *: *  *
Aeut       SIQITIDGKTLTTEEG.RTLVDVAAENGVYIPTLCYLKD.KPCLGTCRVCSVKVNGNVAA..ACTVRVSKGLNVEVND
            * : *:***:  :**  :::: *   *:****:***    *    *: * *  :*       *  :      :  *
Hpyl       MITMNINGKTIECQEGQ.SVLEAARSAGIYIPTICYLSGCSP.TVACKMCMVEMDGKRVY..SCNTKAKNNATILTNT
           *: : *:*: :   ** :::** **  **:***:*:     *    *::*:*:::*       ***   :   * * *
Mthe       MVNLTIDGQRVTAPEG.MTILEVARENGIHIPTLCHHPKLRP.LGYCRLCLVDIEGAAKPMTACNTPVAEGMVIRTST
```

FIGURE 1: Sequence alignments of the N-terminal domain of *C. pasteurianum* hydrogenase I (*HN76*) with some of the most similar sequences retrieved from data banks: *Thma*, N terminus of the α-subunit of the Fe hydrogenase from *Thermotoga maritima* (T. W. O'Rourke, M. F. J. M. Verhagen, and M. W. W. Adams, unpublished; GenBank accession number AF044577); *Aeut*, N terminus of the Hox U subunit of the NAD-reducing NiFe hydrogenase from *Alcaligenes eutrophus* (*29*); *Hpyl*, putative NQO3 subunit of the NADH-ubiquinone oxidoreductase from *Helicobacter pylori* (*30*); and *Mthe*, α-subunit of the formate dehydrogenase from *Moorella thermoacetica* (X.-L. Li, L. G. Ljungdahl, and D. J. Gollin, unpublished; GenBank accession number U73807). The four conserved cysteine residues are bold and underlined. Stars are for identical residues and semicolons for similar ones.

Redox titrations were carried out in an anaerobic glovebox (<1 ppm O_2). The reaction mixture (2.3 mL) contained 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 0.06 mM HN76 protein, and the following mediators: indigo disulfonic acid, safranine T, and benzyl viologen, each at a concentration of 2.5 μM. The potential was measured between the platinum electrode and the Ag/AgCl reference electrode of a combined electrode. HN76 was titrated in both the reductive and the oxidative directions, with stepwise additions of dithionite or ferricyanide, respectively. UV−visible absorption spectra in the 300−800 nm range were recorded at each step, and the absorbance at 418 nm, where the contribution of the mediators is negligible, was used for the calculations (*27*).

Electrochemical experiments were performed in a closed three-electrode cell as described previously (*28*). The three-electrode design included a saturated Ag/AgCl reference MI-401F microelectrode from Microelectrodes Inc. (Bedford, NH), a platinum wire counter electrode from EG&G Instruments, and a glassy carbon working electrode from Radiometer. The three electrodes were connected to a model 263 potentiostat controlled with model 270/250 (EG&G Instruments) software. The carbon disk tip of the working electrode was prepared as follows: immersion in 65% HNO_3 at room temperature for several hours, thorough washing with water, gentle polishing with Al_2O_3, rinsing with water, and drying. The deaerated sample (15−20 μL) was placed on the working electrode, and the other electrodes were positioned to contact the drop. Other experimental conditions are described in the legend of Figure 7.

UV−visible absorption spectra were recorded on a Hewlett-Packard 8452 diode array spectrophotometer. Resonance Raman spectra were recorded by collecting 90° scattering from the surface of a 10 μL frozen droplet of sample (20 K) on the coldfinger of an Air Products Displex DE-202 closed cycle helium refrigerator, using lines from a Coherent Innova 10W Ar^+ laser. The scattered light was analyzed using a ISA U1000 double monochromator fitted with a cooled RCA 31034 photomultiplier tube with photon counting electronics. CD spectra were recorded using a Jasco J500C or J715 spectropolarimeter interfaced to an Oxford Instruments SM3 or Spectromag 4000 superconducting magnet for VTMCD studies. The 50% (v/v) ethylene glycol that was added to the VTMCD samples to ensure the formation of an optical glass on freezing had no effect on the EPR or UV−visible absorption spectra of the samples. X-Band EPR spectra were recorded on a Bruker Instruments ESP 300D spectrometer equipped with an Oxford Instruments ESR 900 flow cryostat (4.2−300 K). The samples used for EPR and VTMCD studies were reduced anaerobically by sodium dithionite (2 mM final concentration) in a glovebox (<1 ppm O_2). Unless otherwise indicated, the samples for spectroscopic measurements were in 50 mM Tris-HCl buffer (pH 7.8).

RESULTS AND DISCUSSION

The N-terminal fragment was designed as follows. It had to include at least four cysteine residues likely to be the ligands of the expected Fe−S cluster. According to the sequence alignments shown in Figure 1, the four conserved cysteines in positions 34, 46, 49, and 62 were the best candidates. Then, the sequence had to exclude the following run of cysteines (starting at residue 98; *7*) that are the presumptive ligands of the second Fe−S cluster. Finally, according to structure predictions, two α-helices separated by a hydrophilic loop are expected to occur in segments made up of residues 60−70 and 80−90. These features suggested that serine 76 was a sensible choice for the C-terminal end of the N-terminal domain. The corresponding fragment of the hydrogenase gene was amplified and cloned in pT7-7, thus yielding the pT7HN76 plasmid, as described in Materials and Methods. Extracts of *E. coli* cells transformed with this plasmid and grown under conditions suitable for the overproduction of the encoded protein were distinctly red-brown.

The red protein obtained after the last purification step (Materials and Methods) was analyzed by automated Edman degradation which yielded a sequence identical with the N-terminal sequence of *C. pasteurianum* hydrogenase I (*7*). The initial sequencing yield indicated that the N-terminal methionine was not blocked. Electrospray ionization mass spectra recorded in the negative-ion detection mode displayed two main peaks (not shown). The $M = 8343$ Da component can be attributed to the apoprotein having the expected 76 N-terminal residues of hydrogenase (calculated $M = 8345$ Da). The second peak, at $M = 8521$ Da, corresponds to the mass of the apoprotein and two iron and two sulfur atomic masses. This indicates a significant contribution of the holoprotein in these spectra, as previously reported for other iron−sulfur proteins investigated so far (*25*, *26*), and is consistent with the presence of a [2Fe-2S] cluster (see below). The purified protein migrated as a single band on SDS−

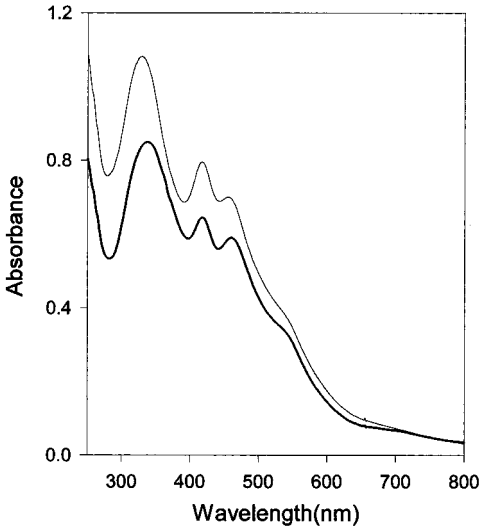

FIGURE 2: UV−visible absorption spectra of the purified HN76 wild type (thin line) and C39A (thick line). The protein concentrations were 0.5 mM, and the solvent was 0.2 M NaCl and 20 mM potassium phosphate (pH 7.0). The optical path length was 1 mm.

polyacrylamide gel electrophoresis (not shown). The apparent mass of 14 000 Da differed significantly from the one calculated from the gene sequence and determined by mass spectrometry (see above). Such anomalous behavior on polyacrylamide gels has previously been observed (*31*−*33*) with small and highly charged proteins (the calculated isoelectric point of the HN76 protein is 3.7).

The UV−visible absorption spectrum of the purified HN76 protein (Figure 2) displays maxima at 336, 418, and 460 nm, and a shoulder at ca. 540 nm, which are characteristic of $[2Fe\text{-}2S]^{2+}$ clusters with all-cysteine ligation. The absence of the 280 nm peak arising from aromatic residues is not unexpected, since the sequence contains neither tryptophan nor tyrosine, and only three phenylalanines which have a weak absorbance at 280 nm. Upon addition of dithionite, the absorption decreased over the entire 350−800 nm range (not shown), as expected upon reduction of the chromophore to the $[2Fe\text{-}2S]^{+}$ level. The four cysteine residues in positions 34, 46, 49, and 62, which are conserved in homologous domains of other proteins (Figure 1), most likely provide the ligand set required by the [2Fe-2S] chromophore. The fifth cysteine residue in position 39, an unlikely ligand of the [2Fe-2S] cluster, was altogether dismissed by the observation that all spectroscopic properties of the C39A variant were nearly identical with those of the wild-type protein (see Figure 1 and below). Furthermore, the purified C39A variant was more abundant and more stable than the wild-type protein. It is therefore anticipated to become an asset for further structural investigations of the HN76 domain.

[2Fe-2S] clusters exhibit intense visible CD compared to other biological Fe−S clusters, and CD spectra are much more sensitive than the corresponding absorption spectra to the protein folding in the vicinity of the cluster (*34*, *35*). Thus far, five distinct types of spectra have been reported for $[2Fe\text{-}2S]^{2+}$ centers in proteins. Plant- and mammalian-type Fds exhibit almost identical CD spectra (*36*, *37*) that are quite distinct from those observed for Rieske-type centers (*38*), mammalian ferrochelatase (*39*), wild-type *Cp* 2Fe Fd (*34*), or a variant form of *Cp* 2Fe Fd that has undergone a

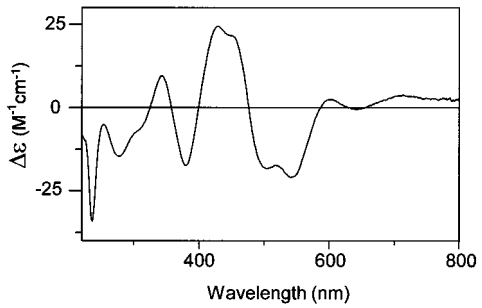

FIGURE 3: Room-temperature CD spectrum of purified wild-type HN76. The protein concentration was 0.5 mM, and the solvent was 0.2 M NaCl and 50 mM Tris-HCl (pH 7.8). The optical path length was 1 mm.

cluster-driven protein reorganization (*35*). In each case, the distinctive CD characteristics are likely to reflect differences in ligation and/or protein folding in the vicinity of the cluster. The CD spectrum of the wild-type HN76 protein as isolated (Figure 3) closely resembles those of plant- and mammalian-type Fds, and the spectrum was unperturbed in the C39A variant (not shown). This provides further support for complete cysteinyl ligation involving Cys34, Cys46, Cys49, and Cys62. Moreover, the CD data indicate that the protein folding in the vicinity of the cluster is very similar to that found in structurally characterized plant-type Fds (*40*).

The low-temperature resonance Raman spectrum (457.9 nm excitation) of wild-type HN76 as isolated (Figure 4) displays a band pattern, in the 250−450 cm^{-1} Fe−S stretching region, characteristic of all-cysteine ligated $[2Fe\text{-}2S]^{2+}$ clusters (*22*, *41*). In particular, the intense 290 cm^{-1} band is unique, among Fe−S active sites, to the binuclear clusters. The frequencies and enhancement profiles of individual bands are very similar to those of putidaredoxin and adrenodoxin, archetypal mammalian-type Fds for which detailed Fe−S stretching mode assignments are available on the basis of isotope shifts and normal mode calculations (*41*, *42*). Moreover, the RR spectra of the HN76 domain and of thionine-oxidized *C. pasteurianum* hydrogenase I (*15*) are remarkably similar. The only significant differences reside in the relative intensity and frequencies of the bands at 337 and 356 cm^{-1} in hydrogenase I. However, these differences are readily interpreted in terms of contributions from the Fe−S stretching modes of the multiple $[4Fe\text{-}4S]^{2+}$ clusters that are present in the enzyme. Resonance enhancement of the Fe−S stretching modes of $[4Fe\text{-}4S]^{2+}$ clusters with 457.9 nm excitation is almost 1 order of magnitude less than for $[2Fe\text{-}2S]^{2+}$ clusters, and on the basis of the RR spectra of *D. vulgaris* hydrogenase (*15*), the dominant bands from the $[4Fe\text{-}4S]^{2+}$ clusters in the enzyme are expected near 340 and 356 cm^{-1}. Given the exquisite sensitivity of RR spectra to Fe−S chromophore structure, the observation of near-identical RR spectra for the [2Fe-2S] clusters in the HN76 fragment and full-size hydrogenase I attests to similar, if not identical, polypeptide chain folding in the vicinity of the cluster.

As purified, the HN76 protein showed no detectable EPR signals at 10 K in the $g \sim 2$ region. In contrast, the dithionite-reduced protein displayed a strong, nearly axial, signal with apparent g values of 2.047, 1.954, and 1.911 (Figure 5). This spectrum was not detectably broadened up to at least 60 K, which indicated relaxation properties

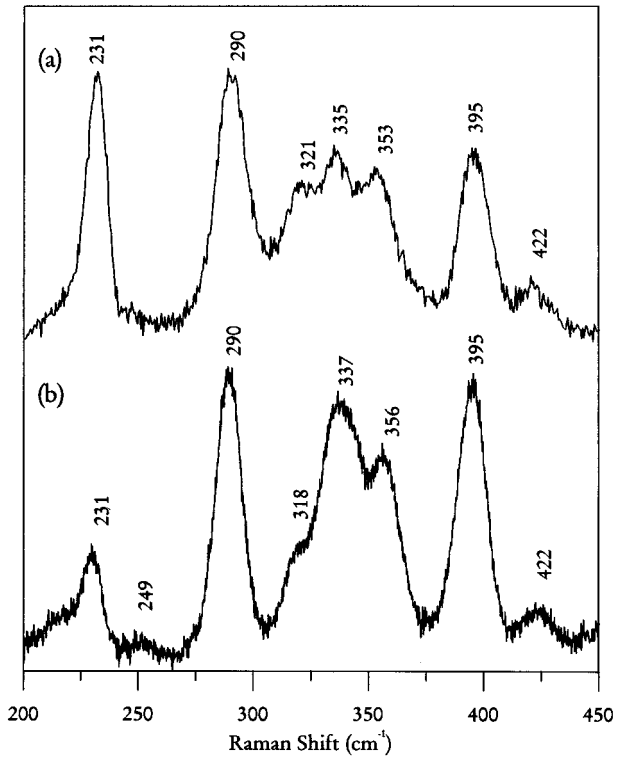

FIGURE 4: Low-temperature resonance Raman spectra of *C. pasteurianum* hydrogenase I: (a) oxidized HN76 fragment and (b) thionine-oxidized hydrogenase I. Both spectra were collected with 457.9 nm laser excitation. The sample concentration for spectrum a was approximately 1.5 mM, and the solvent was 50 mM Tris-HCl buffer (pH 8.0) with 0.2 M NaCl. The sample temperature was 30 K. The spectrum is the sum of 32 scans with each scan involving photon counting for 1 s every 0.5 cm^{-1} with 6 cm^{-1} spectral resolution. The sample concentration for spectrum b was approximately 1 mM, and the medium was 100 mM Tris-HCl buffer (pH 7.8). The sample temperature was 19 K. The spectrum is the sum of 36 scans, with each scan involving photon counting for 1 s every 0.2 cm^{-1} with 6 cm^{-1} spectral resolution (data taken from ref *15*). The 231 cm^{-1} bands arise from frozen water.

characteristic of the $S = {}^1/_2$ ground state of $[2Fe\text{-}2S]^+$ clusters. The power saturation behavior of the EPR spectra of wild-type and C39A HN76 proteins has been investigated by varying the microwave power in the 0.001−20 mW range at fixed temperature (10 K) and was found to be identical to that of plant- and mammalian-type [2Fe-2S] ferredoxins (not shown). Unfortunately, comparison with the EPR spectrum of reduced hydrogenase is not very informative. The EPR spectrum of *Cp* hydrogenase I reduced with dithionite under a H_2 atmosphere comprises a broad, complex resonance centered at around $g = 1.95$ that cannot be simulated as the sum of individual noninteracting components (*12*). EPR studies over the temperature range of 3.8−100 K with microwave powers in the range of 0.001−100 mW using three different frequencies (X-, Q-, and S-band) confirm that this resonance arises from weak intercluster spin−spin interactions (A. T. Kowal, M. K. Johnson, and M. W. W. Adams, unpublished observations). These weak interactions prevent observation of an EPR resonance from a magnetically isolated $S = {}^1/_2$ $[2Fe\text{-}2S]^+$ cluster in hydrogenase I.

The excited state electronic properties of Fe−S clusters, as revealed by VTMCD spectra, are not sensitive to weak intercluster spin−spin interactions. Hence, VTMCD spectroscopy provides a direct means of comparing the electronic

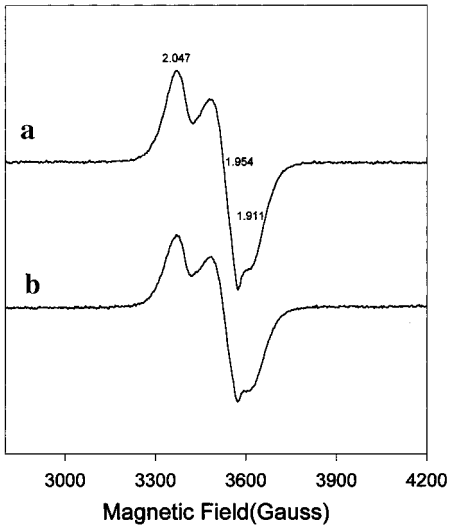

FIGURE 5: EPR spectra of dithionite-reduced HN76: (a) wild type and (b) C39A. Both spectra were recorded under the same conditions, with a temperature of 10 K, a microwave frequency of 9.655 GHz, a microwave power of 0.01 mW, and a modulation amplitude of 10 G. The protein concentrations were 0.5−0.7 mM, and the solvent was 0.2 M NaCl and 20 mM potassium phosphate (pH 7.0).

properties of the reduced $[2Fe\text{-}2S]^+$ clusters in the HN76 fragment and the full-size enzyme (Figure 6). Identical VTMCD spectra were observed for both the wild type and C39A variant, and magnetization data collected at discrete wavelengths (data not shown) confirmed that all the transitions arise exclusively from the EPR detectable $S = {}^1/_2$ ground state. The VTMCD of the dithionite-reduced HN76 fragment is very similar to those reported for reduced adrenodoxin and putidaredoxin, particularly in the charge transfer region below 600 nm (*41*). Moreover, comparison of the spectra shown in Figure 6 reveals that the VTMCD bands of the $[2Fe\text{-}2S]^+$ fragment are clearly apparent in the spectrum of reduced hydrogenase I. Indeed, quantitative subtraction of the reduced HN76 VTMCD spectrum from that of reduced hydrogenase I under identical conditions yields the characteristic VTMCD spectrum of $S = {}^1/_2$ $[4Fe\text{-}4S]^+$ clusters such as those in reduced *Cp* 2[4Fe-4S] Fd (*43*) (data not shown). These results demonstrate that detailed electronic properties of the $[2Fe\text{-}2S]^+$ cluster in the full-size enzyme are completely preserved in the HN76 fragment.

The redox potentials of wild-type and C39A HN76 proteins were measured by square wave voltammetry (Figure 7). The electrochemical responses of both proteins were reversible, and their redox potentials were found to be practically identical, with values (vs the normal hydrogen electrode) of -404 ± 10 (WT) and -396 ± 10 mV (C39A). The cathodic and anodic waves both had peak widths at half-height of 128 ± 2 mV, and the peak currents were proportional to the square root of the scan rate up to 60 mV/s, hence demonstrating the involvement of a fully reversible one-electron process. The reduction potential of HN76 has also been measured by spectrophotometric titration. Reductive titration with dithionite yielded a value of −390 mV, in agreement with the electrochemical measurements. Oxidative back-titration with ferricyanide, however, was too sluggish, maybe as a result of limited protein stability under these conditions, and did not allow reliable redox potential measurements. In any case, the redox potential of the HN76

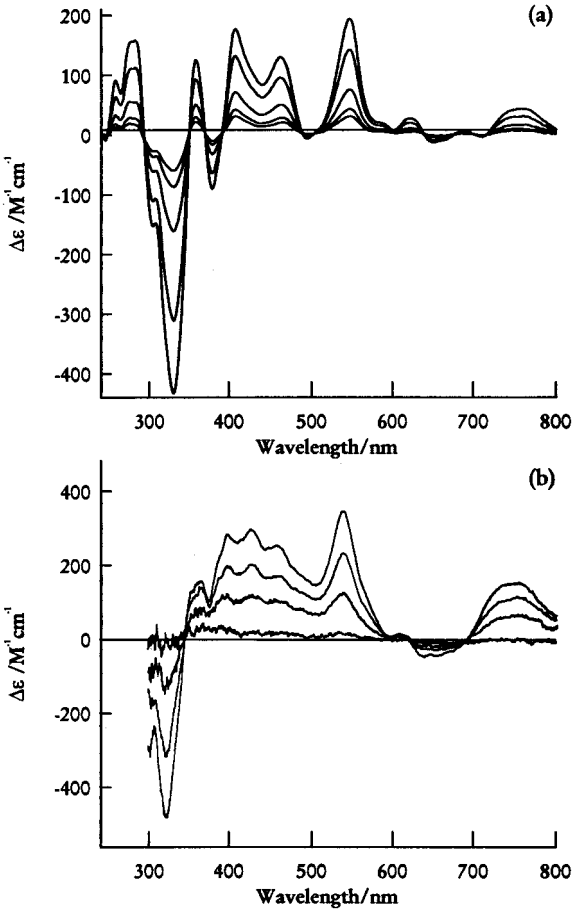

FIGURE 6: VTMCD spectra of dithionite-reduced *C. pasteurianum* hydrogenase I: (a) HN76 fragment and (b) full-size hydrogenase I. The sample whose spectra are shown in panel a was 0.56 mM in 50 mM Tris-HCl buffer (pH 8), 0.2 M NaCl, 50% (v/v) ethylene glycol, and 0.5 mM sodium dithionite. MCD spectra were recorded in a 1 mm cuvette with a magnetic field of 6 T at 1.7, 4.3, 10.2, 19, and 29 K. The sample whose spectra are shown in panel b was 0.136 mM in 50 mM Tris-HCl buffer (pH 7.8), 50% (v/v) ethylene glycol, and 1 mM sodium dithionite under an atmosphere of hydrogen. MCD spectra were recorded in a 1.6 mm cuvette with a magnetic field of 4.5 T at 1.62, 4.22, 8.9, and 73 K (data taken from ref *17*). In both spectra, the MCD intensity of all bands increases with decreasing temperature.

fragment as determined here is consistent with a previous report stating that in *C. pasteurianum* hydrogenase I all iron−sulfur clusters other than the hydrogen-activating site had redox potentials of about −400 mV (*44*).

CONCLUSIONS

The analytical and spectroscopic data presented here demonstrate that a short N-terminal fragment (76 residues) of *C. pasteurianum* hydrogenase I stabilizes a $[2Fe\text{-}2S]^{2+/+}$ iron−sulfur cluster, thus bearing out predictions based on the apparently modular primary structure of the enzyme (*7*). Most of the spectroscopic signatures of this cluster are very similar to, if not identical with, those previously reported for the binuclear iron−sulfur cluster present in the full-size enzyme (*15*, *17*). This is particularly true regarding the RR spectra, which are exquisitely sensitive to the polypeptide environment of Fe−S clusters in general. From that, it is inferred that the polypeptide folds around the [2Fe-2S] cluster in full-size hydrogenase and in its N-terminal fragment are at least very similar. This result validates the gene fragmentation approach for the structural analysis of hydrogenase and suggests its application to other domains of this enzyme.

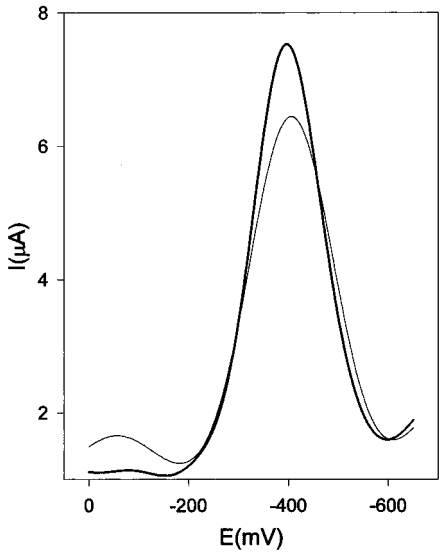

FIGURE 7: Square-wave voltammograms of wild-type HN76 (thin line) and C39A (thick line). Data were recorded at the planar glassy carbon electrode. The protein concentrations were 1−1.2 mM, and the solvent was 0.2 M NaCl and 20 mM potassium phosphate buffer (pH 7.0). The pulse frequency was 5 Hz, the scan increment 2 mV, and the pulse height amplitude 25 mV. The potentials are vs the normal hydrogen electrode.

The confirmed presence of a [2Fe-2S] cluster in *Cp* hydrogenase I, together with the analysis of the sequence of this protein, suggests that it contains, in addition to the H (hydrogen-activating) site, one $[2Fe\text{-}2S]^{2+/+}$ and three $[4Fe\text{-}4S]^{2+/+}$ clusters. This is in accord with the re-evaluation of the cluster content of the enzyme necessitated by RR studies (*15*), and supersedes the original suggestion of four $[4Fe\text{-}4S]^{2+/+}$ clusters and one H cluster that was inferred from analytical, EPR, VTMCD, and Mössbauer data (*12*). When it is considered that the enzyme contains at least 20 iron atoms per molecule (*44*), the present data suggest that the H cluster is composed of at least six iron atoms. This would be consistent with most data on various Fe-only hydrogenases (*1*) and narrows the scope of hypothetical or synthetic models of this site.

All properties of the $[2Fe\text{-}2S]^{2+/+}$ cluster present in the HN76 domain are indicative of all-cysteinyl ligation. The most likely ligands, on the basis of sequence comparisons and site-directed mutagenesis, are the cysteine residues in positions 34, 46, 49, and 62. This cysteine pattern is reminiscent of, though not identical with, that observed in plant- and mammalian-type ferredoxins (*45*). Moreover, in light of the close similarity in the CD spectra, it seems likely that these proteins and the N-terminal hydrogenase domain assume similar folds around their [2Fe-2S] chromophores. So far, only two different protein folds have been found to accommodate [2Fe-2S] clusters with four cysteine ligands; one of these is represented by plant-type ferredoxins (*40*, *45*), and the other one has so far occurred only once, as one of the two [2Fe-2S] clusters of the aldehyde oxidoreductase from *D. gigas* (*19*).

As the N-terminal, [2Fe-2S]-containing fragment of *C. pasteurianum* hydrogenase I has now been shown to be a structurally self-sufficient domain, it is particularly interesting to point out its occurrence in a number of proteins involved

in a variety of redox reactions: subunits of Fe and NiFe hydrogenases, of complex I of the respiratory chain, and of formate dehydrogenase (Figure 1). This implies that any structural information that has been (in this work) or will be (by NMR or X-ray crystallography) obtained with the hydrogenase fragment may be transposed to all similar domains present in other proteins. The evolutionary bearings of these similarities are no less interesting. Indeed, the presence of homologous protein structural modules in anaerobic hydrogen-producing bacteria, in hydrogenosomes of amitochondrial protists, and in respiratory chains of aerobic bacteria and of mitochondria may be taken as an illustration, and possibly as supportive, of a recently proposed hypothesis for the origin of eukaryotic cells through the symbiosis of a hydrogen-producing bacterium and of a hydrogen-dependent archaeon (*46*). Further genome analyses and protein structural studies will shed more light on these questions.

ACKNOWLEDGMENT

We thank J. Gagnon and J.-P. Andrieu for N-terminal sequence determination, Y. Pétillot for recording mass spectra, J.-P. Issartel for the synthesis of oligonucleotides, and J.-M. Moulis for help in setting up the electrochemical apparatus and for critical reading of the manuscript.

REFERENCES

1. Albracht, S. P. J. (1994) *Biochim. Biophys. Acta 1188*, 167−204.
2. Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, C. E., Frey, M., and Fontecilla-Camps, J. C. (1995) *Nature 373*, 580−587.
3. Volbeda, A., Garcin, E., Piras, C., de Lacey, A. I., Fernandez, V. M., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1996) *J. Am. Chem. Soc. 118*, 12989−12996.
4. de Lacey, A. I., Hatchikian, E. C., Volbeda, A., Frey, M., Fontecilla-Camps, J. C., and Fernandez, V. M. (1997) *J. Am. Chem. Soc. 119*, 7181−7189.
5. Vignais, P. M., and Toussaint, B. (1994) *Arch. Microbiol. 161*, 1−10.
6. Voordouw, G., and Brenner, S. (1985) *Eur. J. Biochem. 148*, 515−520.
7. Meyer, J., and Gagnon, J. (1991) *Biochemistry 30*, 9697−9704.
8. Santangelo, J. D., Dürre, P., and Woods, D. R. (1995) *Microbiology 141*, 171−180.
9. Malki, S., Saimmaime, I., De Luca, G., Rousset, M., Dermoun, Z., and Belaich, J.-P. (1995) *J. Bacteriol. 177*, 2628−2636.
10. Gorwa, M.-F., Croux, C., and Soucaille, P. (1996) *J. Bacteriol. 178*, 2668−2675.
11. Bui, E. T. N., and Johnson, P. J. (1996) *Mol. Biochem. Parasitol. 76*, 305−310.
12. Adams, M. W. W. (1990) *Biochim. Biophys. Acta 1020*, 115−145.
13. Voordouw, G., Hagen, W. R., Krüse-Wolters, M., van Berkel-Arts, A., and Veeger, C. (1987) *Eur. J. Biochem. 162*, 31−36.
14. Moulis, J.-M., Sieker, L. C., Wilson, K. W., and Dauter, Z. (1996) *Protein Sci. 5*, 1765−1775.
15. Fu, W., Drozdzewski, P. M., Morgan, T. V., Mortenson, L. E., Juszczak, A., Adams, M. W. W., He, S.-H., Peck, H. D., Jr., DerVartanian, D. V., LeGall, J., and Johnson, M. K. (1993) *Biochemistry 32*, 4813−4819.
16. Wang, G., Benecky, M. J., Huynh, B. H., Cline, J. F., Adams, M. W. W., Mortenson, L. E., Hoffman, B. M., and Münck, E. (1984) *J. Biol. Chem. 259*, 14328−14331.
17. Zambrano, I. C., Kowal, A. T., Mortenson, L. E., Adams, M. W. W., and Johnson, M. K. (1989) *J. Biol. Chem. 264*, 20974−20983.
18. Doolittle, R. F. (1995) *Annu. Rev. Biochem. 64*, 287−314.
19. Romão, M. J., Archer, M., Moura, I., Moura, J. J. G., LeGall, J., Engh, R., Schneider, M., Hof, P., and Huber, R. (1995) *Science 270*, 1170−1176.
20. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1998) *Current Protocols in Molecular Biology*, Wiley-Interscience, New York.
21. Fujinaga, J., and Meyer, J. (1993) *Biochem. Biophys. Res. Commun. 192*, 1115−1122.
22. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry 33*, 13642−13650.
23. Meyer, J. (1995) *Anaerobe 1*, 169−174.
24. Tabor, S. (1990) Expression using the T7 RNA Polymerase/Promoter System, in *Current Protocols in Molecular Biology* (Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds.) pp 16.2.1−16.2.11, Greene Publishing and Wiley-Interscience, New York.
25. Pétillot, Y., Forest, E., Meyer, J., and Moulis, J.-M. (1995) *Anal. Biochem. 228*, 56−63.
26. Meyer, J., Gagnon, J., Gaillard, J., Lutz, M., Achim, C., Münck, E., Pétillot, Y., Colangelo, C. M., and Scott, R. A. (1997) *Biochemistry 36*, 13374−13380.
27. Quinkal, I., Davasse, V., Gaillard, J., and Moulis, J.-M. (1994) *Protein Eng. 7*, 681−687.
28. Hagen, W. R. (1989) *Eur. J. Biochem. 182*, 523−530.
29. Tran-Betcke, A., Warnecke, U., Böcker, C., Zaborosch, C., and Friedrich, B. (1990) *J. Bacteriol. 172*, 2920−2929.
30. Tomb, J.-F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., et al. (1997) *Nature 388*, 539−547.
31. Meyer, J., Moulis, J.-M., and Lutz, M. (1984) *Biochem. Biophys. Res. Commun. 119*, 828−835.
32. Vidakovic, M. S., Fraczkiewicz, G., and Germanas, J. P. (1996) *J. Biol. Chem. 271*, 14734−14739.
33. Gao-Sheridan, H. S., Pershad, H. R., Armstrong, F. A., and Burgess, B. K. (1998) *J. Biol. Chem. 273*, 5514−5519.
34. Crouse, B. R., Yano, T., Finnegan, M. G., Yagi, T., and Johnson, M. K. (1994) *J. Biol. Chem. 269*, 21030−21036.
35. Golinelli, M.-P., Chatelet, C., Duin, E. C., Johnson, M. K., and Meyer, J. (1998) *Biochemistry 37*, 10429−10437.
36. Stephens, P. J., Thomson, A. J., Dunn, J. B. R., Keiderling, T. A., Rawlings, J., Rao, K. K., and Hall, D. O. (1978) *Biochemistry 17*, 4770−4778.
37. Ta, D. T., and Vickery, L. E. (1992) *J. Biol. Chem. 267*, 11120−11125.
38. Fee, J. A., Findling, K. L., Yoshida, T., Hille, R., Tarr, G. E., Hearshen, D. O., Dunham, W. R., Day, E. P., Kent, T. A., and Münck, E. (1984) *J. Biol. Chem. 239*, 124−133.
39. Dailey, H. A., Finnegan, M. G., and Johnson, M. K. (1994) *Biochemistry 33*, 403−407.
40. Ikemizu, S., Bando, M., Sato, T., Morimoto, Y., Tsukihara, T., and Fukuyama, K. (1994) *Acta Crystallogr. D50*, 167−174.
41. Fu, W., Drozdzewski, P. M., Davies, M. D., Sligar, S. G., and Johnson, M. K. (1992) *J. Biol. Chem. 267*, 15502−15510.
42. Han, S., Czernuszewicz, R. S., Kimura, T., Adams, M. W. W., and Spiro, T. G. (1989) *J. Am. Chem. Soc. 111*, 3505−3511.
43. Johnson, M. K., Thomson, A. J., Robinson, A. E., Rao, K. K., and Hall, D. O. (1981) *Biochim. Biophys. Acta 667*, 433−451.
44. Adams, M. W. W., Eccleston, E., and Howard, J. B. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4932−4936.
45. Matsubara, H., and Saeki, K. (1992) *Adv. Inorg. Chem. 38*, 223−280.
46. Martin, W., and Müller, M. (1998) *Nature 392*, 37−41.

BI9812928