Cloning, Sequencing, and Expression of the Gene Encoding the High-Molecular-Weight Cytochrome c from Desulfovibrio vulgaris Hildenborough

W. BRENT R. POLLOCK,1 MOHAMED LOUIFI,2 MIREILLE BRUSCHI,2 BARBARA J. RAPP-GILES,3 JUDY D. WALL,2 AND GERRIT VOORDOUW1*  

Division of Biochemistry, Department of Biological Sciences, The University of Calgary, Calgary, Alberta, Canada T2N 1N4; Laboratoire de Chimie Bactérienne, Centre National de la Recherche Scientifique, 13277 Marseille Cedex 9, France; and Department of Biochemistry, University of Missouri, Columbia, Missouri3

Received 26 July 1990/Accepted 16 October 1990

By using a synthetic deoxyoligonucleotide probe designed to recognize the structural gene for cytochrome cc3 from Desulfovibrio vulgaris Hildenborough, a 3.7-kb XhoI genomic DNA fragment containing the cc3 gene was isolated. The gene encodes a precursor polypeptide of 58.9 kDa, with an NH2-terminal signal sequence of 31 residues. The mature polypeptide (55.7 kDa) has 16 heme binding sites of the form C-X-X-C-H. Covalent binding of heme to these 16 sites gives a holoprotein of 65.5 kDa with properties similar to those of the high-molecular-weight cytochrome c (Hmc) isolated from the same strain by Higuchi et al. (Y. Higuchi, K. Inaka, N. Yasuoka, and T. Yagi, Biochim. Biophys. Acta 911:341–348, 1987). Since the data indicate that cytochrome cc3 and Hmc are the same protein, the gene has been named hmc. The Hmc polypeptide contains 31 histidyl residues, 16 of which are integral to heme binding sites. Thus, only 15 of the 16 hemes can have bis-histidyl coordination. A comparison of the arrangement of heme binding sites and coordinated histidines in the amino acid sequences of cytochrome cc3 and Hmc from D. vulgaris Hildenborough suggests that the latter contains three cytochrome c-like domains. Cloning of the D. vulgaris Hildenborough hmc gene into the broad-host-range vector pJRDR215 and subsequent conjugational transfer of the recombinant plasmid into D. desulfuricans G200 led to expression of a periplasmic Hmc gene product with covalently bound hemes.

C-type cytochromes are distinct from other classes of cytochrome in that they are covalently bound to their hemes by α-thioether bonds (7). These bonds arise from the Markovnikov addition of sulphydryl groups to the vinyl groups of the heme. The sulphydryl groups are derived from the two cysteiny1 residues present in the heme binding sequence C-X-X-C-H. The histidyl residue in this sequence occupies the fifth coordination position of the heme Fe atom, while the sixth coordination position is occupied by either a histidinyl or a methionyl residue. The nature of the ligand occupying this position has a direct effect on the redox potential of the heme (21).

Desulfovibrio vulgaris contains at least four different c-type cytochromes: cytochrome c553 (23, 32), cytochrome c553 (14, 35), cytochrome c553 (18), and high-molecular-weight cytochrome c (Hmc) (13, 38). Cytochrome c553 is the only Desulfovibrio cytochrome reported to be a monoheme and to have a His-Met-coordinated heme Fe atom; the other three cytochromes have more than one heme, all of which are thought to have His-His coordination. Complete amino acid sequence data and three-dimensional structural information are required to ascertain the exact number and organization of the c-type hemes in each of these cytochromes. A determination of the cellular localization of these cytochromes is also important, since Desulfovibrio spp. contain both a periplasmic and a cytoplasmic chain of redox carriers (16, 27). This information can be obtained by sequencing the structural genes for these cytochromes. At present, only the D. vulgaris subsp. vulgaris Hildenborough (hereafter referred to as D. vulgaris Hildenborough) cytochromes c553 (32) and c4 (35) structural genes have been sequenced. For the two remaining cytochromes, partial amino acid sequence data are available for cytochrome cc3 (18), and the isolation and amino acid composition for Hmc have been reported (13). This work has now been extended to D. vulgaris Hildenborough cytochrome cc3, the gene for which has been cloned and sequenced. Perhaps surprisingly, it appears that cytochrome cc3 and Hmc are the same protein, and therefore this work completes the characterization of genes for known periplasmic c-type cytochromes in D. vulgaris.

MATERIALS AND METHODS

Strains, vectors, and media. The bacterial strains and plasmids used are described in Table 1. Desulfovibrio strains were cultured for 18 h at 37°C in stoppered bottles containing 100 ml of Postgate C medium (27) made anaerobic by flushing with N2. The inoculum consisted of 3 ml of a Postgate B (27) culture that had been grown under the same conditions. Growth of Escherichia coli TG2 in TY medium was carried out as previously described (26). E. coli DH5α and Desulfovibrio desulfuricans G200, used for conjugational transfer of broad-host-range plasmids, were cultured as described elsewhere (35a).

Biochemical reagents. All enzymes were obtained from either Pharmacia, Inc., or Boehringer Mannheim Biochemicals. The radiochemicals [α-35S]dATP (400 Ci mmol−1; 10 mCi ml−1), [α-32P]dATP (3,000 Ci mmol−1; 10 mCi ml−1), and [γ-32P]dATP (3,000 Ci mmol−1; 10 mCi ml−1) were purchased from Amersham Corp. and were used for dideoxynucleotide sequencing, nick translation, and 5' end labeling, respectively. Ficoll 400 was purchased from Pharmacia.
Polyvinylpyrrolidone (molecular weight, 40,000), bovine serum albumin (fraction V), molecular biology-grade sodium dodecyl sulfate (SDS), and salmon sperm DNA (sodium salt) were purchased from Sigma Chemical Co. Low- and high-gelling-temperature (LGT and HGT) agarose were obtained from Bethesda Research Laboratories, Inc. Nitrocellulose and Hybond-N hybridization transfer membranes were acquired from Schleicher & Schuell, Inc., and Amersham, respectively. Acrylamide and N,N'-methylenebisacrylamide were purchased from BDH Ltd. Low-molecular-weight protein electrophoresis markers were obtained from Pharmacia. Dialysis tubing (molecular size cutoff, 6 to 8 kDa) was obtained from Spectrum Medical Industries, Inc. All other reagent-grade chemicals were obtained from either Sigma or Fisher Scientific Co.

**Purification and peptide sequencing of cytochrome c₃.** Cytochrome c₃ was purified from French-pressed *D. vulgaris* Hildenborough cells essentially as described by Loutfi et al. (18), who referred to this protein as cytochrome c₃ (M., 26,000). In addition, a new purification procedure was used (not shown) in which this cytochrome was isolated from the periplasmic fraction of *D. vulgaris* Hildenborough prepared as described elsewhere (31). The yields of both procedures were approximately 1 mg/100 g of wet cells. These isolated proteins were pure as judged by SDS-polyacrylamide gel electrophoresis (PAGE) (not shown), which indicated molecular sizes of 62 and 56 kDa for the holoprotein and apoprotein, respectively. The ε₅₅₀ of the reduced holoprotein was found to be 424 M⁻¹ cm⁻³, a value similar to that of Higuchi et al. (13), who showed the presence of 16.3 hemes per polypeptide.

The carboxymethylated apoprotein was prepared as previously described (18) and treated variously with CNBr, Asp-N endopeptidase, Arg-C endopeptidase, or Glu-C endopeptidase. Automated amino acid sequencing (18) of isolated peptides yielded 47 peptides which could be organized into 12 polypeptides as indicated in the legend to Fig. 2. The stretches of Hmc amino acid sequence covered by these peptide groups are indicated in Fig. 2.

**Gene cloning.** DNA manipulations were carried out as described by Maniatis et al. (19). Samples (10 μg) of *D. vulgaris* Hildenborough genomic DNA were digested with various restriction endonucleases, electrophoresed on a 0.7% (wt/vol) HGT agarose gel as described elsewhere (36), and blotted onto a nitrocellulose membrane. Following prehybridization for 15 min at 68°C in 0.2% SDS-10X Denhardt solution (19) and 6X SSC (19), the blot was incubated with the 5'-end-labeled probe P43 for 16 h at 68°C in 6X SSC. The sequence of the deoxyoligonucleotide probe P43, a 47-mer, was 5'-AAGATCGA(AG)AAACCGCGACACCG GC(G)TGCGT(CG)GACTGCCACAAAGGA. This probe was designed to recognize the COOH-terminal region of the structural gene encoding the sequence K-I-E-K-P-A-N-T-A-V-D-C-H-K-E. The bias toward G or C in the third position of each codon was chosen on the basis of the known codon usage in the nucleotide sequences of the *hydA*B (34) and *cyc* (35) genes from *D. vulgaris* Hildenborough. After incubation, the blot was washed thrice with 2X SSC, dried, and autoradiographed. Autoradiography (not shown) revealed a 3.7-kb *XhoI* fragment of chromosomal DNA to be the one most suitable for cloning. This was accomplished by size fractionating chromosomal DNA digested with *XhoI* on 1% (wt/vol) LGT agarose and ligating the 3- to 4.5-kb fraction into the vector pUC8, which had been digested with calf alkaline phosphatase and the restriction endonuclease *Sall*. *E. coli* TG2 cells, made competent by CaCl₂ treatment, were transformed with the ligation mixture and spread onto TY plates containing isopropylidene-β-D-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), and ampicillin. Southern blots of plasmids isolated from white recombinant colonies and digested with both EcoRI and HindIII were probed with P43 as described above. A restriction map of the 3.7-kb insert of the only positive plasmid obtained (pP6A) is shown in Fig. 1.
Shotgun nucleotide sequencing. The procedure followed for shotgun nucleotide sequencing was essentially that outlined by Bankier and Barrell (1). Plasmid pP6A (1 μg ml⁻¹) was sonicated, and the resulting fragments were end reaped and size fractionated as described above. The 0.3- to 0.6-kb fraction was isolated from the gel and ligated into the replicative form of M13mp8 digested with Smal and calf alkaline phosphatase. Competent E. coli TG2 cells were transfected with the ligation mixture and spread onto TY plates with top agar containing IPTG and X-Gal. Single-stranded DNA was isolated from white recombinant plaques by using a 1.5-ml minipreparation procedure (1). Purified phage DNAs (2 μl) were dot blotted on a Hybond-N hybridization transfer membrane and fixed by exposure to UV light. Following digestion with EcoRI and PstI, the 3.7-kb insert of plasmid pP6A was isolated by 1% (wt/vol) LGT agarose gel electrophoresis, radiolabeled by nick translation (19), and denatured by boiling. The filter was prehybridized in 6x SSC–10x Denhardt solution-0.5% (wt/vol) SDS for 1 h at 68°C. The 3.7-kb DNA probe was then added to the prehybridization solution, and hybridization was continued for 16 h at 68°C. After hybridization, the filter was washed twice in 2x SSC for 1 min at room temperature, followed by a 1-h wash at 68°C in 6x SSC–0.5% (wt/vol) SDS. The dried filter was autoradiographed, and positive clones were sequenced by the method of Sanger et al. (28) as detailed by Bankier and Barrell (1). The sequencing data were processed and analyzed by using the programs of Staden and McClellan (29, 30).

Broad-host-range plasmid construction and bacterial conjugation. Plasmid pP6A was digested with Smal and size fractionated as described above. The 2.2-kb fragment was isolated from the gel and ligated to the vector pJR2125, which had been digested with Stal and calf alkaline phosphatase. E. coli TG2 was transformed with the ligation mixture and spread onto TY-kanamycin plates. Recombinant plasmids were purified and analyzed by restriction digestion with BamHI, resulting in excision of the hmc gene as a 2.6-kb BamHI fragment (not shown). The recombinant plasmid and the vector alone were each transformed into E. coli DH5α and subsequently transferred to D. desulfuricans G200 as described elsewhere (35a).

Total (chromosomal and plasmid) DNA of the recombinant exconjugant was prepared as described by Marmur (20), digested with BamHI, and subjected to Southern analysis (not shown). The 2.2-kb Smal fragment derived from plasmid pP6A was radiolabeled by nick translation and used as the probe to verify the presence of plasmid pBPHMC-1 within the recombinant exconjugant D. desulfuricans G200 (pBPHMC-1).

SDS-PAGE and fluorophotography. Desulfovibrio periplasmic proteins were isolated essentially as described by van der Westen et al. (31). Cultures of D. desulfuricans G200(pJR2125, D. desulfuricans G200(pBPHMC-1), and D. vulgaris Hildenborough grown in 100 ml of Postgate C medium were harvested by centrifugation at 10,000 × g for 10 min at 4°C. Each pellet was resuspended in 2 ml of H₂O and transferred to three separate 15-ml glass centrifuge tubes, to each of which was added 2 ml of 0.1 M Tris hydrochloride–0.1 M EDTA (pH 9.0). These suspensions were heated to 30°C in a 60°C water bath and then shaken for 15 min at room temperature, after which they were centrifuged for 10 min at 10,000 × g. The supernatant, which contained the periplasmic proteins, was dialyzed for 20 h against two 1,000-ml volumes of cold deionized H₂O.

For SDS-PAGE, 1-ml aliquots of each dialysate were placed into microfuge tubes and dried by vacuum centrifugation, resuspended in 100 μl of H₂O, 100 μl of SDS incubation buffer (26), and 20 μl of 2-mercaptoethanol, and placed in a boiling water bath for 10 min. SDS-PAGE and fluorophotography were carried out as previously described (26) except that the acrylamide gradient was 10 to 15% and the slab gel was photographed by using Polaroid type 55 film and an exposure time of 10 min.

Nucleotide sequence accession number. The sequence data presented in this paper have been submitted to GenBank and assigned accession number M34607.

RESULTS

Cloning. Deoxyxylonucleotide P43, which was designed to recognize the COOH-terminal sequence of cytochrome cc₃, was used to screen purified recombinant plasmids for the presence of this gene. Screening of approximately 140 plasmids yielded one positive plasmid, pP6A. A restriction map for the 3.7-kb XhoI insert of pP6A is shown in Fig. 1. hmc coding region. The sequence of the first 2,160 nucleotides (nt) of the pP6A insert is shown in Fig. 2. The region containing the structural gene for the cytochrome (nt 472 to 2106) was obtained on both strands and has been translated into protein. The NH₂-terminal sequence found for cytochrome cc₃ (18) by protein sequencing (peptide P₂; KALP...) is present at nt 565. It is clear from the nucleotide sequence (Fig. 2) that this sequence is preceded by an NH₂-terminal signal sequence of 31 residues which initiates at nt 472. The gene-translated protein sequence confirmed the directly determined sequences for the 12 polypeptides, which together span 85% of the mature amino acid sequence (Fig. 2).

The region of the structural gene corresponding to the mature cytochrome sequence (nt 565 to 2106) encodes an apoprotein of 55.7 kDa. This sequence includes 16 C-X-X-C-H heme binding sites, and the holoprotein, with 16 covalently bound hemes, has a calculated molecular size of 65.5 kDa. As discussed in detail below, the properties of the cytochrome defined by the nucleotide sequence depicted in Fig. 2 are reminiscent of those of the Hmc first correctly described by Higuchi et al. (13). Therefore, this gene will be referred to as the hmc gene, and its product will be referred to as Hmc.

The Hmc polypeptide contains 31 His residues, 16 of which are integral to the 16 C-X-X-C-H heme binding sites. Thus, only 13 of the 16 hemes can have bis-histidinyl coordination. The heme binding sites and histidyl residues of the Hmc polypeptide are spatially arranged into three complete cytochrome c₃-like domains (II, III, and IV) and one incomplete domain (I) that lacks both the second heme binding site and its corresponding sixth-coordinate-position histidine (Fig. 3). Visible spectroscopic data on Hmc (13) do not disallow a Met residue at this position. The ε₉₅₃ value for His-Met-coordinated c-type cytochromes is typically on the order of 0.9 mM⁻¹ cm⁻¹, as determined for cytochromes c (9) and ε₉₅₃ (23, 38). Higuchi et al. (13) determined an ε₉₅₃ value of 428 mM⁻¹ cm⁻¹ for Hmc based on a molecular size of 75 kDa, which corresponds to an ε₉₅₃ value of 374 mM⁻¹ cm⁻¹ when adjusted to the correct molecular size of 65.5 kDa. Applying this adjusted ε₉₅₃ value to the visible spectrum of D. vulgaris Hildenborough Hmc shown by Higuchi et al. (A₅₅₃ = 0.14; concentration of Hmc, 370 nM) gives an A₉₅₃ on the order of 0.0003. Thus, if there is a His-Met-coordinated heme in Hmc, it has remained undetected because the A₉₅₃ is only 0.2% of the heme A₅₅₃ in Hmc.
FIG. 2. Nucleic acid sequence of the first 2,160 nt of the 3.7-kb XhoI fragment of _D. vulgaris_ Hildenborough chromosomal DNA depicted in Fig. 1. The −35 and −10 regions of a putative promoter sequence are indicated, as is the ribosome binding site (rbs). The _hmc_ structural gene spans the region from nt 472 (initiator methionine residue) to nt 2109 (termination codon). A vertical arrow marks the signal peptide cleavage site. The overlined regions represent 436 of the 514 amino acid residues of the mature Hmc polypeptide which have been sequenced at the peptide level in polypeptides P1 to P8. The corresponding positions of these polypeptides in the gene-translated sequence are P1, nt 565 to 921; P2, nt 1375 to 1530; P3, nt 955 to 1026; P5, nt 1162 to 1224; P8, nt 1315 to 1374; P9, nt 922 to 955; P10, nt 1069 to 1104; P11, nt 1105 to 1119; P12, nt 1255 to 1293; P13, nt 1594 to 1635; and P14, nt 2008 to 2106. The region corresponding to the deoxyoligonucleotide probe P43 is underlined, and the asterisk in that region denotes the one nucleotide that is mismatched in the P43 sequence. Restriction sites for enzymes XhoI, _Hind_III, and _KpnI_ are also noted.
of the heme Fe atoms in Hmc is coordinated to a Met residue, this would not be the first report of a c-type cytochrome containing both His-His- and His-Met-coordinated hemes; the tetraheme cytochrome c of the Rhodopseudomonas viridiss photosynthetic reaction center has one His-His- and three His-Met-coordinated heme Fe atoms (6). The hmc gene does not encode any transmembrane sequences other than the NH2-terminal signal sequence, and its product is therefore expected to reside in the periplasmic space.

Nucleotide sequences up-and downstream from hmc. An E. coli gal^70 consensus promoter is present at nt 209 (Fig. 2). A similar promoter sequence was found in the upstream region of the D. vulgaris Hildenborough cya gene encoding cytochrome c_3 (35). The −35 sequences of the two promoters are identical (TTGACA), as are four of the six nucleotides of the −10 sequences (TACCAT for cya). The ribosome binding site at nt 460 contains the core AGGA sequence which is typical of all D. vulgaris Hildenborough genes sequenced to date except the rbo gene (2, 4, 32, 34, 35, 36).

No hairpin loop transcription termination sequences have been detected downstream of the hmc gene, and sequencing of the downstream region, although incomplete, indicates that the hmc gene may be the first cistron of a redox operon, which is composed of at least three open reading frames (not shown). The metabolic role of Hmc has not been firmly established, and the sequencing of this operon may help to uncover the redox partners of Hmc.

Probing of Desulfovibrio genomes with D. vulgaris Hildenborough hmc. Samples of total DNA from several Desulfovibrio species were digested with EcoRI, subjected to Southern blotting, and probed with the 2.2-kb Smal fragment of plasmid p63A (not shown). The region spanning nt 1 to 2109 of the sequence illustrated in Fig. 2 represents 94% of this 2.2-kb Smal fragment. Therefore, this fragment is mainly representative of the hmc gene and its upstream region.

Autoradiography showed single bands for the following Desulfovibrio species: D. vulgaris Brockhurst Hill, Hildenborough, Monticello, and Wandle; D. desulfuricans Berre Sol, Canet 41, and Teddington R; and D. africanus Bhengazi and Walvis Bay. These results suggest that the hmc gene may be present in these Desulfovibrio species. No hybridizing bands were observed for D. gigas, D. multispirans, D. salseigens, and D. desulfuricans El Agheila Z and Norway, which suggests that hmc is either not present or not sufficiently homologous for hybridization in these bacteria. None of the samples tested displayed multiple bands, which indicates that the hmc gene is present in single copy in the positive chromosomes.

Expression of Hmc in D. desulfuricans G200. Electrophoresis in the presence of a reductive chelator such as 2-mercaptoethanol or dithiothreitol leads to the removal of iron from heme to yield the protoheme, which upon excitation with UV light fluoresces in the visible range (10, 15). Therefore, following electrophoresis under denaturing conditions in the presence of 2-mercaptoethanol, c-type cytochromes, which are covalently bound to their hemes, can be visualized under UV light, whereas other heme-containing proteins cannot since their hemes are not retained upon denaturation.

The periplasmic protein fractions of D. vulgaris Hildenborough (Fig. 4, lanes 3) and the two exconjugants D. desulfuricans G200(pPDR215) (lanes 1) and D. desulfuricans G200(pPBP-HMC-1) (lanes 2) were subjected to SDS-PAGE. Directly upon completion of electrophoresis, the gel was analyzed for postelectrophoretic protoheme fluorescence (Fig. 4A). Comparison of lanes 1 and 2 shows that D. desulfuricans G200(pPBP-HMC-1) expressed the D. vulgaris Hildenborough hmc gene as a fluorescent protein of 66 kDa. Lanes 1 to 3 in Fig. 4A all share a fluorescent band corresponding to 15 kDa that is due to the presence of the cytochromes c_3 native to both D. desulfuricans G200 (lanes 1 and 2; 35a) and D. vulgaris Hildenborough (lane 3; 26). The periplasmic fraction of D. desulfuricans G200 also contains a fluorescent protein of 45 kDa (lanes 1 and 2).

The apparent absence of a fluorescent band corresponding to 66 kDa (mature Hmc has a molecular size of 65.5 kDa; see above) in lane 3 is due to the low concentration of Hmc in D. vulgaris Hildenborough. The clear presence of this band in lane 2 illustrates that the exconjugant was able to express the hmc gene product at higher levels than did the parent organism. Figure 4B, which shows the gel illustrated in Fig. 4A after staining with Coomassie blue, indicates that the
VOL.

binant

exconjugant

are

the

SDS-PAGE

protoheme fluorescence analysis of

D.

vulgaris

A

of

4A

encoding

cytochrome

higher

level

(not

pP6A

fied by

D.

exconjugant

amino

terminal

cytoplasmic

proteins

is marked

with their

Hildenborough.

All three

Hildenborough

2i, .;

these

peptides.

The

presence

Hildenborough

sequence

given

was

Comparison

of

D. vulgaris

Hildenborough

sequence

generated

by

D.

desulfuricans G200(pBPHMC-1)

was

D.

exconjugant

was

transformed

with plasmid

pBPHMC-1.

The presence of plasmid pBPHMC-1 in the recombinant exconjugant D. desulfuricans G200(pBPHMC-1) was verified by probing total DNA with the 2.2-kb SmaI fragment of pP6A (not shown).

Periplasmic localization of Hmc. According to a predictive method, which involves a comparison of the mature NH2-terminal amino acid residues of both known periplasmic and cytoplasmic proteins from Desulfovibrio spp. (16), Hmc would be classified as a cytoplasmic protein. The genetic and biochemical work presented here indicates that this assignment is incorrect.

The hmc gene encodes an additional 31-amino-acid signal sequence (Fig. 2). Comparison of this sequence with the signal peptides of the periplasmic D. vulgaris Hildenborough cytochromes c3 and c553 (Fig. 5) shows that it is also a signal peptide with the same characteristic positive and hydrophobic regions as in the other two sequences. However, the Hmc signal sequence has a greater overall length, has three NH2-terminal Arg residues that are spaced out rather than adjacent, and has a signal peptide cleavage site that is not of the form A↓A as is seen in the signal sequences of the other two percytochromes. Hmc can be purified from the periplasm of D. vulgaris Hildenborough in quantities similar to those obtained by using whole cells (see Materials and Methods). This protein was also found to be localized in the periplasm when it was overexpressed in D. desulfuricans G200 cells containing plasmid pBPHMC-1 (Fig. 4). The periplasmic localization of Hmc is in accord with an empirical rule which states that maturation of c-type cytochromes requires a membrane translocation step. This rule applies to both proaryotic and eucaryotic cytochromes, e.g., the mitochondrial, monoheme cytochromes c and c1 (24, 25), for which it is known that the heme is covalently inserted by the enzyme heme lyase (8) following or during translocation of the apocytochrome.

**DISCUSSION**

The Hmc class of proteins. Cytochrome c3, which has also been referred to as cytochrome c3 (M, 26,000) (18) and octaheme cytochrome c3 (17, 22), from D. vulgaris Hildenborough was reported to be an octaheme dimer composed of two identical tetraheme subunits, each with a molecular size of 20 kDa (Table 2). Partial amino acid sequence data obtained from this same cytochrome was used for the formulation of deoxyoligonucleotide probe P43 (see Materials and Methods), which was used to isolate the recombinant plasmid pP6A. Surprisingly, nucleotide sequencing (Fig. 2)

![Figure 4](image1.png)

**FIG. 4. Expression of D. vulgaris Hildenborough Hmc by D. desulfuricans G200 transformed with plasmid pBPHMC-1. (A) Post-SDS-PAGE protoheme fluorescence analysis of the periplasmic fractions from D. desulfuricans G200(pIRD215) (lane 1), the recombinant exconjugant D. desulfuricans G200(pBPHMC-1) (lane 2), and D. vulgaris Hildenborough (lane 3). (B) Coomassie blue staining of the same gel. Positions of the size standards are marked on the left and are given in kilodaltons. Approximately 30 μg of protein was loaded in each lane.**

![Figure 5](image2.png)

**FIG. 5. (A) Comparison of the NH2-terminal signal peptide sequences for three periplasmic, c-type cytochromes from D. vulgaris Hildenborough. All three sequences have been determined by nucleic acid sequencing of their respective structural genes. The cleavage site is marked by the vertical line, and +1 marks the first amino acid residue of the mature polypeptide. Positively charged residues occurring in the signal sequence are underlined. (B and C) Comparison of the L (B) and S (C) peptide sequences determined for D. vulgaris Miyazaki Hmc with their homologous regions in D. vulgaris Hildenborough Hmc. Numbers indicate positions of the amino acid residues in the Hildenborough sequence. The underlined Lys residues have not been sequenced but are assumed from the activity of lysyl-endopeptidase on these peptides. Direct homology is indicated by an asterisk.**
of the *D. vulgaris* Hildenborough chromosomal DNA insert carried by this plasmid indicated the presence of an open reading frame for a mature hexadecahemocytocrome c of 65.5 kDa rather than the expected tetrahemocytocrome c of 20 kDa. In addition to the nucleotide sequence data, the following evidence indicates that in *D. vulgaris* Hildenborough, cytochrome *cc*$_3$ is a 65.5-kDa protein. (i) The amino acid sequence data obtained by direct peptide sequencing (polypeptides P$_1$ to P$_{15}$; Fig. 2) preclude the protein from being a 43.3-kDa, octaheme homodimer. Rather than the 4 heme binding sequences (C-X-X-C-H) expected in this model (18), polypeptides P$_1$ to P$_{15}$ contain 14 separate sequences C-X-X-C-H and cover 85% of the mature polypeptide sequence encoded by the hmc gene (Fig. 2). (ii) SDS-PAGE of preparations of holocytochrome *cc*$_3$ indicates a major band at 70 kDa (18), whereas the apoprotein gives a major band at 56 kDa (not shown). These values correspond to the 65.5-kDa holoprotein and 55.7-kDa processed apoprotein encoded by the *hmc* gene. It is likely that the lower-molecular-weight bands observed in preparations of cytochrome *cc*$_3$ from *D. vulgaris* Hildenborough are the result of limited proteolysis between different domains (Fig. 3). (iii) As discussed also by Yagi and Ogata (38), the amino acid composition and physical properties of *D. vulgaris* Hildenborough cytochrome *cc*$_3$ (18) are similar to those of the high-molecular-weight cytochrome c from the same strain (13). These data are summarized in Table 2, which also contains the amino acid composition derived from the gene sequence. Hmc and cytochrome *cc*$_3$ are identical with respect to pl in the oxidized state, $M_r$ by SDS-PAGE, and number of hemes per polypeptide and are found to have very similar amino acid compositions (Table 2).

Although a cytochrome with the characteristics of Hmc (16-c-type heme; $M_r$, 65.5 ± 10) has not been isolated from other *Desulfovibrio* species, single hybridizing bands were observed when a Southern blot of EcoRI-digested chromosomal DNA from several *Desulfovibrio* species was probed with the nick-translated, 2.2-kb Smal fragment of pP6A (see Results). This finding suggests that Hmc is not restricted to *D. vulgaris* Hildenborough in the genus *Desulfovibrio*.

The data presented above clearly indicate that the model of *D. vulgaris* Hildenborough cytochrome *cc*$_3$ as an octaheme homodimer of 43.3 kDa is incorrect. It is possible that this is also the case for the cytochromes *cc*$_3$ identified in *D. desulfuricans* Norway (11) and *D. gigas* (3). However, these proteins differ from *D. vulgaris* Hmc in that both have never been reported to display a molecular size greater than 26 kDa. Additionally, neither *D. desulfuricans* Norway nor *D. gigas* chromosomal DNA was seen to hybridize to the 2.2-kb Smal fragment of plasmid pP6A (see Results). Both of these observations suggest that the *D. desulfuricans* Norway and *D. gigas* proteins may be quite different from Hmc. A complete description of the amino acid sequence of these proteins will be required to settle this question.

Cytochromes *cc*$_3$ and *cc*$_{555}$ from *D. vulgaris* Miyazaki are homologous to the corresponding proteins of the Hildenborough strain, and the Miyazaki strain also contains a cytochrome similar to Hmc (30a). Although this cytochrome is reported to have only 11 hemes per polypeptide (38), it does display similarity to the Hmc from the Hildenborough strain in two respects: (i) the holoform of this protein is also a monomer of 67 kDa by SDS-PAGE, and (ii) there is homology with Hmc in the limited amino acid sequence obtained for the Miyazaki protein both in the NH$_2$-terminal peptide (L; 30a) and in an internal peptide fragment (S; Fig. 5). The Miyazaki and Hildenborough Hmcs are further compared below.

### **Hmc protein structure**

Cytochromes *cc*$_3$ display a distinct linear arrangement of heme binding sites and coordinating histidinyl residues (14). In *D. vulgaris* cytochrome *cc*$_3$, the elements of this pattern (Fig. 3) span approximately 100 amino acids and are, proceeding from NH$_2$ to COOH termini: (i) a H-X-X-H sequence in which the first His residue coordinates to heme 1 and the second His residue coordinates to heme 3, (ii) a C-X-X-C-H heme binding sequence for heme 1, (iii) a single His residue that coordinates to heme 2, (iv) a C-X-X-X-C-H heme binding sequence for heme 2, (v) a single His residue that coordinates to heme 4, (vi) a C-X-X-C-H heme binding sequence for heme 3, and (vii) a C-X-X-X-X-C-H heme binding sequence for heme 4. The 15 non-heme binding site His residues of Hmc and the 16 heme binding sites are arranged relative to one another in a pattern that resembles a quadruplication of the same components as they are arrayed in cytochrome *cc*$_3$ (Fig. 3). There are some irregularities of possible consequence in this motif. The first domain (I; hemes and His 1 to 3 as in Fig. 3) is composed of only six of the eight elements present in the cytochrome *cc*$_3$ pattern. According to this pattern, the second heme binding site and its corresponding sixth-position His residue are absent. The other irregularity is related to the first in that there is an extra heme (number 12; Fig. 3) immediately following the third domain, and it is the only one that does not have a His residue, which may be related to it by the

<table>
<thead>
<tr>
<th>Property</th>
<th>This work</th>
<th>Higuchi et al. (13)</th>
<th>Loutfi et al. (18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_r$ of holoform (10$^3$)</td>
<td>65.5</td>
<td>69$^a$</td>
<td>70$^a$</td>
</tr>
<tr>
<td>Subunits</td>
<td>1</td>
<td>1</td>
<td>2 × 20 kDa$^d$</td>
</tr>
<tr>
<td>Heme content/subunit</td>
<td>16</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>pl of ferriform</td>
<td>ND$^e$</td>
<td>9.2</td>
<td>9.2</td>
</tr>
<tr>
<td>No. of residues</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>54</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>His</td>
<td>31</td>
<td>31</td>
<td>26</td>
</tr>
<tr>
<td>Arg</td>
<td>23</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>Asx</td>
<td>47</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>Thr</td>
<td>24</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Ser</td>
<td>24</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Glx</td>
<td>44</td>
<td>51</td>
<td>48</td>
</tr>
<tr>
<td>Pro</td>
<td>33</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>Gly</td>
<td>37</td>
<td>41</td>
<td>45</td>
</tr>
<tr>
<td>Ala</td>
<td>63</td>
<td>68</td>
<td>58</td>
</tr>
<tr>
<td>Cys</td>
<td>32</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>Val</td>
<td>27</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>Met</td>
<td>12</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>Ile</td>
<td>18</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Leu</td>
<td>25</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>Tyr</td>
<td>5</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Phe</td>
<td>13</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Trp</td>
<td>2</td>
<td>ND$^f$</td>
<td>ND$^f$</td>
</tr>
<tr>
<td>Total</td>
<td>514</td>
<td>512$^f$</td>
<td>512$^f$</td>
</tr>
</tbody>
</table>

---

$a$: SDS-PAGE of holoform.
$b$: Gel filtration of holoform.
$c$: Analytical ultracentrifugation of holoform.
$d$: SDS-PAGE of apoprotein.
$e$: ND, Not determined.
$f$: Adjusted to 512 residues.

---

**Table 2.** Comparison of some physical properties and amino acid composition of cytochrome *cc*$_3$ (18) and Hmc (13) with data obtained from the sequence of the gene for *D. vulgaris* Hildenborough Hmc (this study)
cytochrome c₃ motif. By this reasoning, heme 12 is the one most likely not to have bis-histidinyl coordination.

X-ray crystallographic data obtained for *D. vulgaris* Hildenborough Hmc have been presented (13) but are not of sufficient resolution for structural information. The pattern in Fig. 3 indicates a threefold symmetry for Hmc, since the 15 bis-histidyl-coordinated hemes are arranged into three complete cytochrome c₃-like domains (II to IV; Fig. 3), each of which comprises four hemes, and an incomplete domain (I; Fig. 3) with only three hemes. The single Met-His-coordinated heme may function as the electron exit point, interacting with a higher-potential redox protein. This exit heme could obtain its electrons from the three hemes in domain I (Fig. 3), which could each, in turn, receive electrons from one of the three remaining, tetraheme domains (II, III, and IV; Fig. 3). Domains II, III, and IV could each interact with a redox protein in addition to their interaction with heme 12 via domain I.

NH₂-terminal sequence data for the L and S peptides from *D. vulgaris* Miyazaki Hmc have been obtained (see above). The partial amino acid sequence of the S peptide (19.1 kDa) depicted in Fig. 5C is homologous to a strain of amino acids located between the third and fourth cytochrome c₃-like domains of the Hmc from the Hildenborough strain. This Miyazaki peptide, cleaved from the holoprotein, has been reported to be free of heme, and all of the heme was to be found on the NH₂-terminal L peptide (33.1 kDa). The discrepancy in heme content between the two Hmcs (11 and 16 hemes per 67 kDa for the Miyazaki and Hildenborough holoproteins, respectively) can be accounted for by a complete absence of heme binding sites from domain IV (Fig. 3), and this suggests that the Hmc class of cytochromes may display a high degree of variability in domain IV.

ACKNOWLEDGMENTS

W.B.R.P. acknowledges the support of a Graduate Scholarship from the Alberta Heritage Foundation for Medical Research. This work was supported by an operating grant from the National Sciences and Engineering Research Council of Canada to G.V. and by grant DE-FG02-87ER13713 from the Basic Energy Research Program of the U.S. Department of Energy awarded to J.D.W. We thank P. Coddin for assistance with the interpretation of X-ray crystallographic data and J. Boncel for help with the automated protein sequence analysis.

REFERENCES