ELSEVIER

Biochimica et Biophysica Acta 1413 (1999) 55–61

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.elsevier.com/locate/bba

Short sequence-paper

Cloning and sequence analysis of the gene encoding *Methylophilus methylotrophus* cytochrome *c*″, a unique protein with a perpendicular orientation of the histidinyl ligands

Nicholas J. Price [a], Erik Vijgenboom [b], Graça Ribeiro [c], João V. Costa [1,c], Gerard W. Canters [b], Helena Santos [a,*]

[a] *Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2780 Oeiras, Portugal*
[b] *Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, P.O. Box 9502, 2300 RA Leiden, Netherlands*
[c] *Instituto Gulbenkian de Ciência, Apartado 14, 2780 Oeiras, Portugal*

Received 1 March 1999; received in revised form 20 July 1999; accepted 22 July 1999

Abstract

Cytochrome *c*″ from *Methylophilus methylotrophus* is an unusual monohaem protein that undergoes a major redox-linked spin-state transition: one of the two axial histidines bound to the iron in the oxidised form is detached upon reduction and a proton is taken up. A 3.5-kb DNA fragment, containing the gene encoding cytochrome *c*″ (*cycA*), has been cloned and sequenced. The cytochrome *c*″ gene codes for a pre-protein with a typical prokaryotic 20-residue signal sequence, suggesting that the protein is synthesised as a precursor which is processed during its secretion into the periplasm. The C-terminus of cytochrome *c*″ has homology with the corresponding region of an oxygen-binding haem protein (SHP) from phototrophically grown *Rhodobacter sphaeroides*. SHP is similar in size and in the location of its haem-binding site. Immediately downstream from cytochrome *c*″ a second open reading frame (ORF) codes for a 23-kDa protein with similarity to the cytochrome *b*-type subunit of Ni-Fe hydrogenase. The possibility of coordinated expression of *cycA* and this ORF is discussed.

Keywords: Cytochrome *c*″; Redox protein; Nucleotide sequencing; *Methylophilus methylotrophus*

The soluble cytochrome *c*″ isolated from *Methylophilus methylotrophus* is a small protein (15 kDa) containing a protohaem IX group covalently linked to the polypeptide backbone by thioether bridges at the conserved site Cys-X-X-Cys-His. The axial ligands are two histidine residues in the oxidised form [1], and a single histidine residue in the reduced form [2]. EPR, m.c.d., and NMR studies provided evidence for an unusual near-perpendicular (85°) orientation of the axial ligand planes in the oxidised form [1,3]. Characterisation of the haem environment by NMR suggests that the redox-linked protonation

Abbreviations: CTAB, cetyltrimethylammonium bromide; *cycA*, gene encoding cytochrome *c*″; EPR, electron paramagnetic resonance; GIDA, glucose inhibited division protein; I, inosine; m.c.d, magnetic circular dichroism; NMR, nuclear magnetic resonance; ORF(s), open reading frame(s); *oriC*, origin of replication; PCR, polymerase chain reaction; p*I*, isoelectric point; R, purine; RBS, ribosome-binding site; SHP, oxygen-binding haem protein from *Rhodobacter sphaeroides*; ThdF, thiophene and furan oxidation protein; Y, pyrimidine

* Corresponding author. Fax: +351-1-442-8766; E-mail: santos@itqb.unl.pt

[1] Present address: Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, 1300 Lisboa, Portugal.

0005-2728/99/$ – see front matter
PII: S0005-2728(99)00085-7

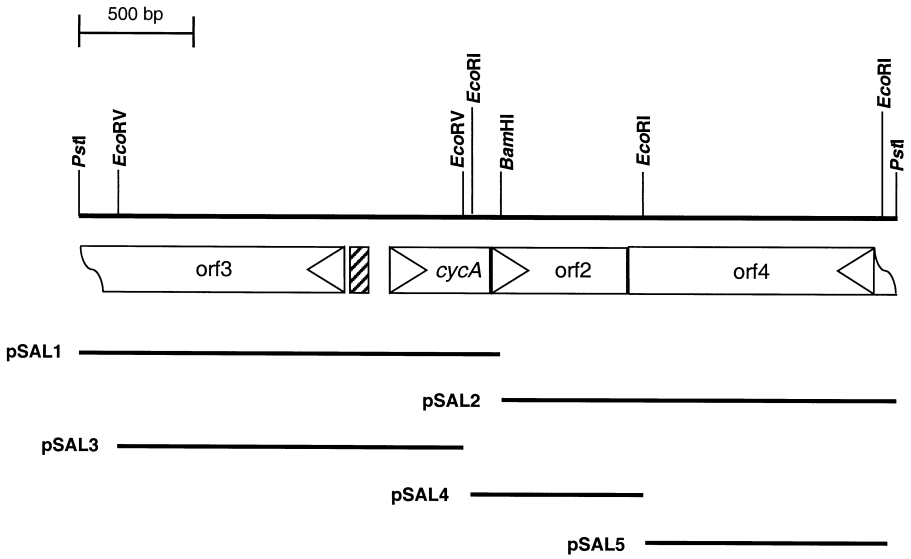

Fig. 1. Genetic and physical map of the *M. methylotrophus* DNA containing the cytochrome *c″* gene, *cycA*. The orientation and localisation of the ORFs are indicated by four triangles. The solid lines represent the extent of DNA sequenced. Subclones were constructed in pUC18 or pBluescriptII KS$^+$ (Stratagene). pSAL1 and pSAL2 were constructed using the unique *Bam*HI site and *Pst*I cloned into pUC18 and pUC19. The *Pst*I fragment was digested with *Eco*RV giving rise to pSAL3. pSAL4 and pSAL5 were constructed by cloning the *Eco*RI fragments of the *Pst*I fragment into *Eco*RI-digested pUC18. The relevant restriction sites are shown. The putative origin of replication is indicated (hatched box).

occurs via a channel running through the cleft on the haem face opposite to that containing the histidinyl ligand that detaches upon reduction of the iron atom [4]. It has been shown that the midpoint redox potential of this haem protein has a strong pH dependence (redox-Bohr effect) in the physiological pH range [5]. Therefore, cytochrome *c″* provides an interesting example of a soluble protein capable of coupling electron and proton transfer, in vitro, but its physiological role has not been elucidated. Recently, amino acid sequence homology has been found between cytochrome *c″* and an oxygen-binding protein (SHP) from *Rhodobacter sphaeroides* [6]. Here we report on the nucleotide sequence of the cytochrome *c″* gene and show that the gene appears to be located in an operon that includes, at least, one other redox protein.

M. methylotrophus (NCIMB 11585) was grown on a methanol containing synthetic medium [7], harvested and genomic DNA isolated by the CTAB method adapted from the protocol by Murray and Thompson [8]. Degenerate oligonucleotides 5′ CTC GAA TTC AAY CCI ATG TAY GAR GCI CC 3′ and 5′ CTC GGA TCC GC IGG RTT RTT IGT RTG RCA 3′ were designed based on the published partial amino acid sequence corresponding to *Eco*RI-NPMYEAP and *Bam*HI-APNNTHC, respectively [6]. Following Touchdown PCR [9], a 132-bp product was cloned into pBluescriptII KS$^+$ (Stratagene), sequenced and used as a homologous probe to screen against digested chromosomal DNA. Positive signals were recuperated from an identical agarose gel, shotgun cloned into pBluescriptII KS$^+$ and transformed into *Escherichia coli* XL1-Blue. One transformant was selected by colony hybridisation using the same probe. Two complete and another two incomplete ORFs were identified in the 3.5-kb *Pst*I fragment as shown in Fig. 1. The complete nucleotide sequences and the derived amino acid sequences are shown in Figs. 2 and 3. The first ORF, *cycA* (438 bp, position 1353–1790 in Fig. 2) encodes for cytochrome *c″*. A ribosome-binding site [10] (AGGAG, double underlined in Fig. 2) occurs 8–12 bp upstream from the start codon. Preceding the start site are two hexamers (underlined in Fig. 2) with considerable homology to the −10 (**TA**TAA**T**) and −35 (**TTG**ACA) consensus sequences of *E. coli* [11]. These putative sequences are separated by 18 bp, within the observed range of 15–19 bp [12]. The mature cytochrome *c″* is preceded by a 20 amino acids long sig-

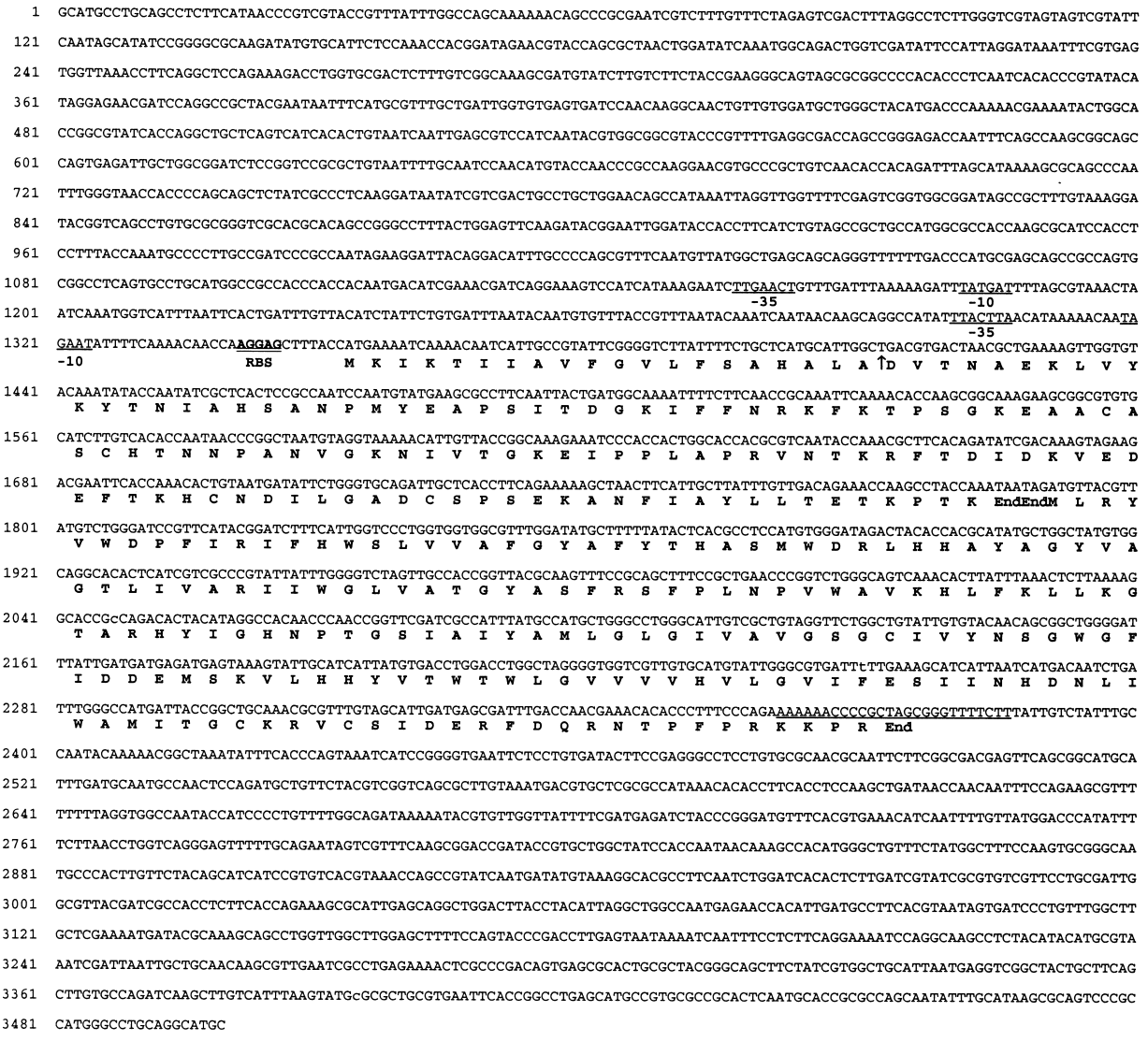

Fig. 2. Sense strand nucleotide sequence of a 3.5-kb fragment of *M. methylotrophus* DNA. The deduced amino acid sequences of the cytochrome *c*″ gene and orf2 are indicated using the standard single-letter code. Putative ribosomal binding site is underlined (bold). An arrow marks the signal peptide cleavage site. A region with dyad-symmetry is underlined (nucleotides 2361–2386). The GenBank accession number for the *M. methylotrophus cycA* nucleotide sequence is AF119838.

nal peptide (including the ATG start codon) apparently serving the purpose of directing the protein to the periplasm. The sequence of the signal peptide shows high similarity to the sequences of signal peptides of other prokaryotes [13]. The amino acid sequence of cytochrome *c*″ has been completed by protein sequencing and is identical to that deduced from the nucleotide sequence [14].

Cysteine residues 49 and 52 together with histidine 53 constitute the typical haem-binding site of the mature protein as found in most *c*-type cytochromes. However, unlike most cytochromes *c*, the haem is located in the middle of the protein. Equally unusual for monohaem cytochromes, cytochrome *c*″ contains a further two cysteine residues not required to bind the haem, at positions 96 and 104 of the mature protein. The distal histidine is adjacent at position 95. A similar occurrence has been found for cytochrome c_5 from *Azotobacter vinelandii* [15], *Pseudomonas mendocina* [16] and more recently for SHP [5]. Using Chou and Fasman prediction [17], the polypeptide chain of cytochrome *c*″ is organised into five

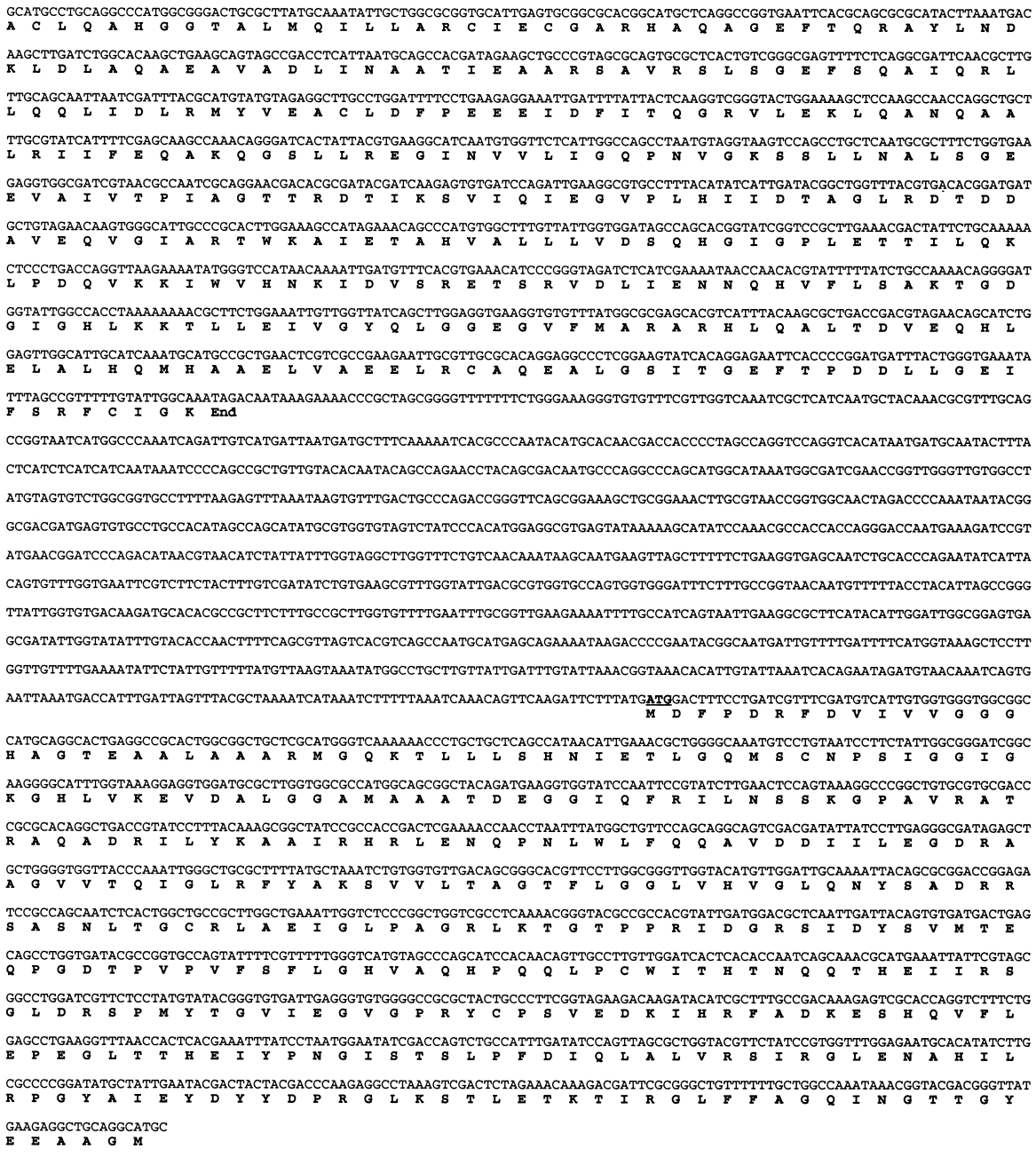

Fig. 3. Antisense strand nucleotide sequence of a 3.5-kb fragment of *M. methylotrophus* DNA. The deduced amino acid sequences of the two incomplete ORFs, orf3 and orf4, are indicated. Putative start site for orf4 is underlined (bold).

α-helical segments totalling ~47% of the amino acids with very little β-strand structure (data not shown).

The second ORF (orf2 in Fig. 1, 585 bp, position 1791–2375) is located immediately downstream of the two stop codons of the *cycA* gene. Although there is no obvious Shine–Dalgarno sequence [10] upstream from the start codon, it is conceivable that the two genes are translated by the same ribosome. The gene ends with a single termination codon and codes for a 194-residue peptide with a calculated charge of +8 resulting in a p*I* of 10.4. This gene product may be membrane-bound containing four hydrophobic transmembrane regions (bracketed re-

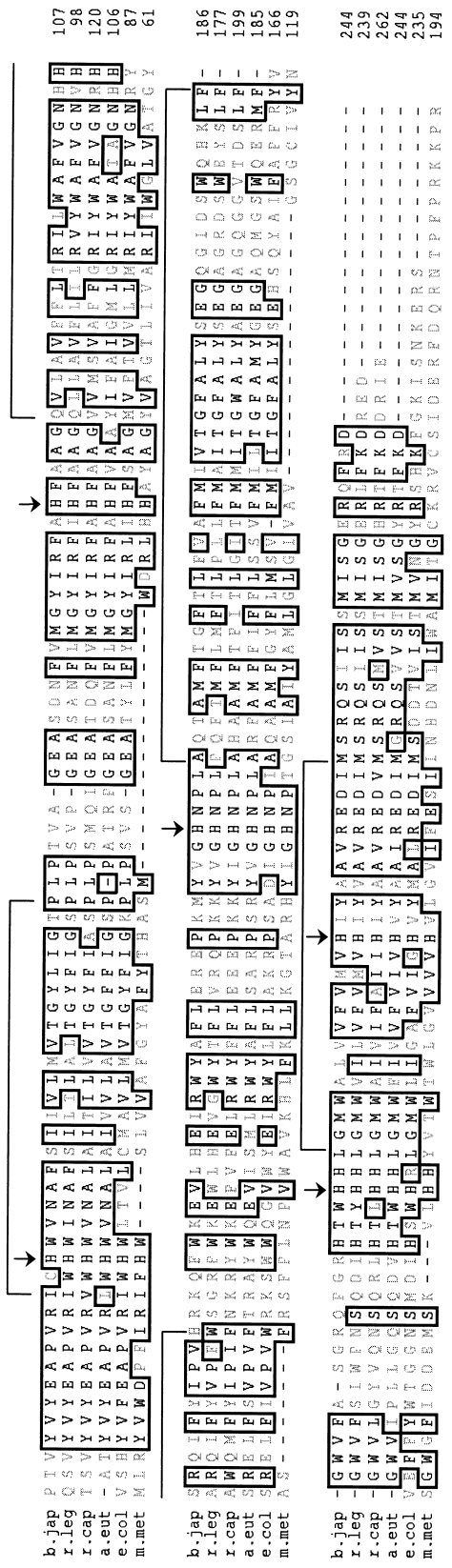

Fig. 4. Alignment of orf2 (m.met) with amino acid sequences predicted by various *HydC* genes. Vertical bars designate the hydrophobic stretches. The bacterial species are labelled as follows: b.jap, *Bradyrhizobium japonicum* [25]; r.leg, *Rhizobium leguminosarum* [26]; r.cap, *Rhodobacter capsulatus* [27]; a.eut, *Alcaligenes eutrophus* [28]; e.col, *E. coli* [29]; m.met, *M. methylotrophus* (this study). Fully conserved histidines are designated by arrows.
←

gions in Fig. 4). A striking similarity was found with the *b*-type cytochrome subunit of several Ni-Fe hydrogenases usually referred to as the gene *hupC* ([18,19] and references therein), with identity in the range of 32–66% for membrane spanning regions, and no specific homology with the rest of the sequence (see Fig. 4). *HupC* gene products have been shown to accept electrons from hydrogenase in *Rhodobacter capsulatus* [19]. Dross and co-workers [20] have suggested that the putative *b*-type cytochrome subunits may function as electron donors to the respiratory quinones. However, orf2 is not homologous with any other known cytochrome *b*. Five conserved histidine residues representing potential haem ligands were found, two of which were in the first and last transmembrane spanning regions in the hydrogenase subunit as seen in other Ni-Fe hydrogenases. Haem *b* of cytochrome *b* is generally considered to be ligated within hydrophobic regions via two histidine residues [21].

Upstream of the *cycA* gene, up to 75% identity was found with the *E. coli* replication origin (*oriC*). On the opposite strand (orf3 in Fig. 3, position 2353–3499), 77% identity with the *E. coli gidA* gene coding for glucose inhibited division protein (GIDA) reported to be involved in cell division [22]. In all bacteria studied, the *gidA* gene is found near to the *oriC* [23]. Although the frame is incomplete, between 40 and 73% homology to GIDA from several bacteria was found when compared against protein databases. Downstream from orf2 on the antisense strand, another incomplete ORF (orf4, position 1–1107 in Fig. 3) was identified. This ORF is similar to the carboxy termini of the thiophene and furan oxidation proteins (ThdF) from *E. coli* (∼51% identical residues in the sequenced portion), *Haemophilus influenzae* (∼51%), *Buchnera aphidicola* (∼41%) and *Pseudomonas putida* (∼53%). Although *E. coli* ThdF was identified because of its involvement in thiophene

and furan oxidation [24], the exact function of this protein remains unknown.

Although cytochrome *c*″ and the possible *b*-type cytochrome subunit of hydrogenase are quite probably co-expressed, it is not known if they are functionally related. In other organisms, the organisation of the hydrogenase genes appears to be polycistronic while being very conserved ([21] and references therein). However, homology between orf2 and several known sequences of these cytochromes *b* is considerably lower than expected. Although similar accessory genes generally follow the *b*-type cytochrome gene, in this study, no ORFs were found to be proximal other than the high similarity with a ThdF protein on the opposite strand. Cytochrome *c*″ is not known to be involved in hydrogen metabolism, suggesting that the gene product of orf2 is probably not a hydrogenase subunit. Also the analysis of the sequence data suggests that there is no reason to believe that orf2 is a hydrogenase subunit. Therefore it seems reasonable to propose that orf2 could be a *b*-type cytochrome. The sequence data suggest that *cycA* and orf2 are in the same operon and are translationally coupled which supports the idea that the two have a physiological relation.

This work was supported by E.C. Project contract numbers CHRX-CT94-0540 and FMRX-CT98-0218. We thank Dr. D.L. Turner for helpful discussions.

References

[1] M.J. Berry, S.J. George, A.J. Thomson, H. Santos, D.L. Turner, Cytochrome *c*″ isolated from *Methylophilus methylotrophus*. An example of bis-histidine coordinated Fe^{3+} haem with near-perpendicular orientation of the ligands, Biochem. J. 270 (1990) 413–417.

[2] H. Santos, D.L. Turner, Characterization and NMR studies of a novel cytochrome *c* isolated from *Methylophilus methylotrophus* which shows a redox-linked change of spin-state, Biochim. Biophys. Acta 954 (1988) 277–286.

[3] H.S. Costa, H. Santos, D.L. Turner, Ligand orientation and haem electronic structure in ferricytochrome *c*″ from *Methylophilus methylotrophus* studied by ^{13}C NMR, Eur. Biophys. J. 25 (1996) 19–24.

[4] H.S. Costa, H. Santos, D.L. Turner, Characterization of the haem environment in *Methylophilus methylotrophus* ferricytochrome *c*″ by ^{1}H-NMR, Eur. J. Biochem. 215 (1993) 817–824.

[5] H.S. Costa, H. Santos, D.L. Turner, A.V. Xavier, Involvement of a labile axial histidine in coupling electron and proton transfer in *Methylophilus methylotrophus* cytochrome *c*″, Eur. J. Biochem. 208 (1992) 427–433.

[6] K. Klarskov, G. Van Driessche, K. Backers, C. Dumortier, T.E. Meyer, G. Tollin, M.A. Cusanovich, J.J. Van Beeumen, Ligand binding and covalent structure of an oxygen-binding heme protein from *Rhodobacter sphaeroides*, a representative of a new structural family of *c*-type cytochromes, Biochemistry 37 (1998) 5995–6002.

[7] J.D. Windass, M.J. Worsey, E.M. Pioli, D. Pioli, P.T. Barth, K.T. Atherton, E.C. Dart, D. Byrom, K. Powell, P.J. Senior, Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*, Nature 287 (1980) 296–401.

[8] M.G. Murray, W.F. Thompson, Rapid isolation of high-molecular-weight plant DNA, Nucleic Acids Res. 8 (1980) 4321–4325.

[9] R.H. Don, P.T. Cox, B.J. Wainwright, J.S. Mattick, 'Touchdown' PCR to circumvent spurious priming during gene amplification, Nucleic Acids Res. 19 (1991) 4008.

[10] J. Shine, L. Dalgarno, Determinant of cistron specificity in bacterial ribosomes, Nature 254 (1980) 34–38.

[11] B.C. Hoopes, W.R. McClure, Strategies in regulation of transcription initiation, in: F.C. Neidhardt, J.L. Ingraham, K. Brooks-Low, B. Magasanik, M. Schaechter, H.E. Umbarger (Eds.), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, Vol. 2, American Society for Microbiology, Washington, DC, 1987, pp. 1231–1240.

[12] O. Raibaud, M. Schwartz, Positive control of transcription initiation in bacteria, Annu. Rev. Genet. 18 (1984) 173–206.

[13] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, Protein Eng. 10 (1997) 1–6.

[14] K. Klarskov, D. Leys, K. Backers, G. Van Driessche, H.S. Costa, H. Santos, Y. Guisez, J.J. Van Beeumen, Cytochrome *c*″ from the obligate methylotroph *Methylophilus methylotrophus*, an unexpected homologue of SHP from the phototroph *Rhodobacter sphaeroides*, Biochim. Biophys. Acta 1412 (1999) 47–55.

[15] D.C. Carter, K.A. Melis, S.E. O'Donnell, B.K. Burgess, W.F. Furey Jr., B.C. Wang, C.D. Stout, Crystal structure of *Azotobacter* cytochrome c_5 at 2.5 Å resolution, J. Mol. Biol. 184 (1985) 279–295.

[16] R.P. Ambler, E. Taylor, Amino acid sequence of cytochrome c_5 from *Pseudomonas mendocina*, Biochem. Soc. Trans. 1 (1973) 166–168.

[17] P.Y. Chou, G.D. Fasman, Prediction of the secondary structure of proteins from their amino acid sequence, Adv. Enzymol. Related Areas Mol. Biol. 47 (1978) 45–148.

[18] J.N. Reeve, G.S. Beckler, D.S. Cram, P.T. Hamilton, J.W. Brown, J.A. Krzychi, A.F. Kolodziej, L. Alex, W.H. Orme-Johnson, C.T. Walsh, A hydrogenase-linked gene in *Methanobacterium thermoautotrophicum* strain ΔH encodes a polyferredoxin, Proc. Natl. Acad. Sci. USA 86 (1989) 3031–3035.

[19] P.M. Vignais, B. Toussaint, Molecular biology of membrane H_2 uptake hydrogenases, Arch. Microbiol. 161 (1994) 1–10.

[20] F. Dross, V. Geisler, R. Lenger, F. Theis, T. Krafft, F. Fahrenholz, The quinone-reactive Ni/Fe-hydrogenase of *Wolinella succinogenes*, Eur. J. Biochem. 206 (1992) 93–102.
[21] G. Hauska, W. Nitschke, R.G. Herrmann, Amino acid identities in the three redox center-carrying polypeptides of cytochrome bc_1/b_6f complexes, J. Bioenerg. Biomembr. 20 (1988) 211–228.
[22] K. von Meyerburg, B.B. Jørgensen, J. Nielsen, F.G. Hanson, Promoters of the *atp* operon coding for the membrane-bound ATP synthase of *Escherichia coli* mapped by Tn10 insertion mutations, Mol. Gen. Genet. 188 (1982) 240–248.
[23] N. Ogasawara, H. Yoshikawa, Genes and their organization in the replication origin region of the bacterial chromosome, Mol. Microbiol. 6 (1992) 629–634.
[24] K.Y. Alam, D.P. Clark, Molecular cloning and sequence of the *thdF* gene which is involved in thiophene and furan oxidation by *Escherichia coli*, J. Bacteriol. 173 (1991) 6018–6024.
[25] L.A. Sayavedra-Soto, G.K. Powell, H.J. Evans, R.O. Morris, Nucleotide sequence of the genetic loci encoding the subunits of *Bradyrhizobium japonicum* uptake hydrogenase, Proc. Natl. Acad. Sci. USA 85 (1988) 8395–8399.
[26] E. Hidalgo, J.M. Palacios, J. Murillo, T. Ruiz-Argüeso, Nucleotide sequence and characterization of four additional genes of the *hydrogenase structural operon from Rhizobium leguminosarum* bv. Viciae, J. Bacteriol. 174 (1992) 4130–4139.
[27] P. Richaud, P.M. Vignais, A. Colbeau, R.L. Uffen, B. Cauvin, Molecular biology studies of the uptake hydrogenase of *Rhodobacter capsulatas and Rhodocyclus gelatinosus*, FEMS Microbiol. 87 (1990) 413–418.
[28] C. Kortlüke, K. Horstmann, E. Schwartz, M. Rohde, R. Binsack, B. Friedrich, A gene complex coding for the membrane-bound hydrogenase of *Alcaligenes eutrophus* H16, J. Bacteriol. 174 (1992) 6277–6289.
[29] N.K. Menon, J. Robbins, H.D. Peck Jr., C.Y. Chatelus, E.-S. Choi, A.E. Przybyla, Cloning and sequencing of a putative *Escherichia coli* [NiFe] hydrogenase-1 operon containing six open reading frames, J. Bacteriol. 172 (1990) 1969–1977.