Cloning, sequencing and expression studies of the genes encoding amicyanin and the beta-subunit of methylamine dehydrogenase from thiobacillus-versutus

Ubbink, M.; Keef, M.A.G. van; Kleinjan, D.J.; Hoitink, C.W.G.; Huitema, F.; Beintema, J.J.; ... ; Canters, G.W.

Eur. J. Biochem. 202, 1003–1012 (1991)

Cloning, sequencing and expression studies of the genes encoding amicyanin and the β-subunit of methylamine dehydrogenase from *Thiobacillus versutus*

Marcellus UBBINK [1], Mario A. G. van KLEEF [2], Dirk-Jan KLEINJAN [1], Carla W. G. HOITINK [1], Fienke HUITEMA [2], Jaap J. BEINTEMA [3], Johannis A. DUINE [2] and Gerard W. CANTERS [1]

[1] Gorlaeus Laboratories, Leiden University, Leiden, The Netherlands
[2] Department of Microbiology and Enzymology, Delft University of Technology, Delft, The Netherlands
[3] Department of Biochemistry, Groningen University, Groningen, The Netherlands

(Received May 31, 1991) – EJB 91 0710

The genes encoding amicyanin and the β-subunit of methylamine dehydrogenase (MADH) from *Thiobacillus versutus* have been cloned and sequenced. The organization of these genes makes it likely that they are coordinately expressed and it supports earlier findings that the blue copper protein amicyanin is involved in electron transport from methylamine to oxygen. The amino acid sequence deduced from the nucleotide sequence of the amicyanin-encoding gene is in agreement with the published protein sequence. The gene codes for a pre-protein with a 25-amino-acid-long signal peptide. The amicyanin gene could be expressed efficiently in *Escherichia coli*. The protein was extracted with the periplasmic fraction, indicating that pre-amicyanin is translocated across the inner membrane of *E. coli*. Sequence studies on the purified β-subunit of MADH confirm the amino acid sequence deduced from the nucleotide sequence of the corresponding gene. The latter codes for a pre-protein with an unusually long (56 amino acids) leader peptide. The sequencing results strongly suggest that pyrroloquinoline quinone (PQQ) or pro-PQQ is not the co-factor of MADH.

When *Thiobacillus versutus* is offered methylamine as its sole source of carbon, nitrogen and energy, it synthesises a redox chain which converts the amine into the aldehyde and which transports the electrons generated in the methylamine conversion eventually to oxygen as the final electron acceptor [1]. Because the redox chain is short and the constituting proteins are relatively easy to manipulate, there is currently strong interest in the physiology, enzymology and mode of operation of this chain [1–9]. Although the composition of this chain has not yet been established with absolute certainty, there is agreement [1, 3, 5, 10, 11] that the first two proteins in the redox path are: (a) methylamine dehydrogenase (MADH), an $\alpha_2\beta_2$ dimer with a molecular mass of 123 kDa and a quinone compound as a prosthetic group [12, 13] and (b) amicyanin, a type I copper protein with a molecular mass of 11.7 kDa [14, 15]. Crystals of MADH have been analysed with X-ray diffraction techniques but the interpretation of the electron density map has been hampered up till now by a lack of amino acid sequence information [16, 17]. Especially, unequivocal interpretation of the electron density around the side chains purportedly forming the sites of attachment of the co-factor remained problematical up till now [16, 17].

A structural study of amicyanin by means of X-ray and NMR techniques is under way [18, 19]. Ongoing studies of the mechanism of the electron transfer between the various reaction partners in the chain have shown among others, that the redox activity of the amicyanin *in vitro* is regulated by pH [6]. The terminal oxidase in the redox chain is an aa_3-type cytochrome oxidase [1, 2]. The link between the oxidase and the amicyanin is presumably formed by cytochrome c_{550} [1, 20], a single heme containing class I *c*-type cytochrome with a molecular mass of 15 kDa, which recently has been characterised extensively [4].

Study of the genes coding for the above mentioned redox proteins was deemed interesting for several reasons. Their organization might confirm the supposed functionality of the redox chain. The nucleotide sequence of the gene(s) coding for the MADH subunits might give information about the nature of the sites of attachment of the co-factor in the small (β-)subunit. Knowledge of the primary structure of MADH would significantly further the interpretation of the electron-density map of MADH. Finally, with the genes of the redox proteins available, a study at the molecular level of the electron-transfer mechanism would become possible by means of site-directed mutagenesis, a technique which has already proven its potential for the study of biological electron tranfer [21, 22].

Here we report, first, on amino-acid-sequencing studies of the β-subunit of MADH. Secondly, the cloning and sequencing of the genes coding for this subunit and for

Correspondence to G. W. Canters, Gorlaeus Laboratories, Leiden University, P.O. Box 9502, NL-2300 RA Leiden, The Netherlands

Abbreviations. IPTG, isopropyl β-D-galactopyranoside; MADH, methylamine dehydrogenase; ORF, open reading frame; PQQ, pyrroloquinoline quinone.

Enzymes. Cytochrome-*c* oxidase (EC 1.9.3.1); (methyl)amine dehydrogenase (EC 1.4.99.3).

Notes. The structure of the co-factor of MADH from *M. extorquens* AM1 has recently been reported as consisting of two covalently linked tryptophan residues with a quinone moiety [56]. In view of the evidence reported in the present paper it is highly likely now that the *T. versutus* MADH contains the same di-tryptoquinone co-factor [57].

The novel nucleotide sequence data published here has been deposited with the GenBank sequence data bank and is available under accession number M58001.

amicyanin are described. The genes appear to be located in one operon, which is different from the operon in which the cytochrome-c_{550} encoding gene has been found, as will be reported elsewhere. Finally, the results of heterologous expression experiments with amicyanin and the MADH β-subunit-encoding genes are reported.

MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

T. versutus ATCC 25364^T [23] cells were obtained from J. Frank, Delft University of Technology. *E. coli* JM105 was used as a host and grown aerobically at 37°C in Luria-Bertani broth. The plasmid pUC18 was used as a cloning vector [24]. M13mp18 and mp19 phages were used for DNA sequencing.

Purification of MADH and preparation of the β-subunit

T. versutus MADH was purified as described [16]. The pure enzyme was dissociated into its subunits in 6 M guanidine/HCl in 50 mM sodium phosphate buffer, pH 7.0. The separation of the two subunits was performed on a Sephadex G-100 column equilibrated with 0.5 M guanidine/HCl in 50 mM phosphate buffer, pH 7.0. The subunits were dialyzed against a 0.1 M ammonium carbonate buffer, pH 8.0 and lyophilised. The β-subunit was subsequently pyridylethylated by treatment with 5 µl tributylphosphine plus 5 µl 4-vinylpyridine in 0.5 ml 50% pyridine in water (by vol.) for 3 h at 45°C.

Digestion and peptide mapping

17.5 mg pyridylethylated β-subunit of MADH was treated with 2% (by mass) trypsin and incubated for 4 h at 37°C. The tryptic digest was lyophilised, taken up in 0.5 ml of 10% formic acid and fractionated by HPLC chromatography on a Nucleosil 10 C_{18} column (300 × 4.6 mm) equilibrated with 0.1% trifluoroacetic acid. A linear gradient of 0–67% acetonitrile in 0.1% trifluoroacetic acid was applied at a flow rate of 1.0 ml/min over a period of 60 min. Peptide elution was monitored by measuring the absorbance at 214 nm. Of the seven fractions collected (T1–T6 and Q; Fig. S1), three (T4, T5 and T6) were subjected to another purification step by HPLC on the same column, using the same gradient in 0.1% ammonium acetate, pH 6.0. The peptide containing the chromophoric co-factor could be identified on the basis of its yellow colour (peptide Q). It was further digested by a 2-h incubation at 37°C with 2% chymotrypsin, which resulted in the isolation of four more peptides (Q1–Q4). Q3 still appeared inhomogeneous and was collected in three fractions, the first of which was subjected to further investigation.

Amino acid analysis of the unmodified β-subunit and of the purified peptides was performed on a Kontron Liquimat III Analyzer after hydrolysis of the peptide under vacuum for 24 h in 6 M HCl at 110°C. The unmodified protein and peptide Q3 were sequenced on an Applied Biosystems (Model 477A) pulse-liquid protein sequencer, equipped with an on-line Model 120 A PTH-analyzer [25]. N-terminus determination of the peptides was performed using dimethylaminobenzene 4′-isothiocyanate [26]. Tryptophan was determined quantitatively after hydrolysis with 3 M mercaptoethanesulfonic acid [27].

```
1. 3′ GTY CTR TTY TAH TG
      Gln Asp Lys Ile Thr
      1               5

2. 3′ CTY TAC TAC TGN TTY CTR GT
      Glu Met Met Thr Lys Asp Gln
      71                     76

3. 3′ ATR CTR ATR AAR ACR TG
      Tyr Asp Tyr Phe Cys Thr
      89                  94

4. 5′ AAG TGG CAG CCS CAG GAC AAC GAC ATC CAG GCS TGC GAC TAC TGG
      Lys Trp Gln Pro Gln Asp Asn Asp Ile Gln Ala Cys Asp Tyr Trp
      12                                                     26
```

Fig. 1. *Mixed oligonucleotide probes used in the hybridization experiments and their corresponding amino acid sequences.* Probes 1, 2 and 3 were derived from the amicyanin amino acid sequence [14]; probe 4 was designed on the basis of the amino acid sequence of the N terminus of the β-subunit of MADH. H: A, T, or C; N: any nucleotide: R: purine; S: G or C; Y: pyrimidine. Probes 1, 2, 3 and 4 are mixtures of 24, 32, 32 and 4 oligomers, respectively

Chemicals and reagents

All restriction enzymes, T4 DNA-ligase, T4 polynucleotide kinase, and T7 DNA-polymerase were obtained commercially. Polyclonal antibodies against amicyanin and the unmodified small subunit of MADH were raised in rabbit [28].

Preparation and analysis of DNA

Chromosomal DNA of *T. versutus* was isolated from 3 g bacterial paste as described by Chater et al. [29]. Plasmid DNA from *E. coli* strains was isolated according to Birnboim and Doly [30].

Electrophoresis of DNA and hybridization conditions

DNA restriction fragments were separated on 1–2% agarose minigels. Screening of genomic and plasmid DNA digests of *T. versutus* was performed by hybridization with labelled synthetic oligonucleotide probes. Probes 1–3 were designed on the basis of the amino acid sequence of amicyanin [14] and are shown in Fig. 1. Probe 1 is a 24-fold degenerate 14-mer. Probes 2 and 3 are 32-fold degenerate 20-mers and 17-mers, respectively. Probe 4 was designed as a fourfold degenerate 45-mer oligonucleotide on the basis of the N-terminal amino acid sequence of the β-subunit of MADH (Fig. 3). Since *T. versutus* DNA has about a 68% G+C content [23], in designing probe 4 a G or a C was chosen for the third codon in cases where the genetic code is twofold degenerate; in cases of fourfold degeneracy, G and C were chosen; ATC was chosen as the codon for Ile (Fig. 1). The oligonucleotides were labelled by phosphorylation with T4 polynucleotide kinase and an equimolar amount of [γ-^{32}P]ATP. Hybridization was performed at 45°C directly on dried denatured 1% agarose gels [31] with 1–2 mg chromosomal DNA/lane according to Canters [32]. After hybridization, gels were washed three times at room temperature with 6× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl plus 0.015 M sodium citrate) and subsequently incubated for 10 min in 6× NaCl/Cit + 0.1% SDS and 10 min in 6× NaCl/Cit, at 40°C (probe 1), 48°C (probe 2), or 43°C (probe 3). Autoradiograms were obtained after 2–72 h exposure at −70°C with two intensifying screens.

Transformation and colony hybridization were performed according to Sambrook et al. [33]. Conditions for plasmid hybridization (with any of the four probes listed in Fig. 1) were identical to those for hybridization of genomic DNA,

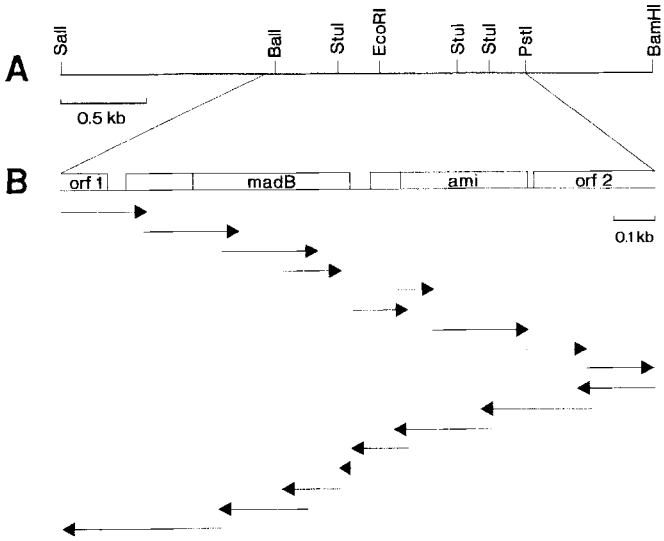

Fig. 2. *Restriction map (A) and sequencing strategy (B) for the* T. versutus mad*B and* ami *genes.* Relevant restriction sites are shown. ORFs are indicated by boxed areas. Transcription of all ORFs is from left to right. The arrows represent the templates and direction of sequencing

except that gels were washed at room temperature after hybridization.

DNA sequence analysis and strategy

M13 phage particles were purified and single-stranded DNA was extracted by the method of Sanger et al. [34]. DNA sequencing was performed by the dideoxynucleoside-triphosphate chain-termination method of Sanger et al. [35], with [α-^{32}P]dCTP as the label. Sequencing reactions were performed with T7 DNA-polymerase by following a protocol supplied by the manufacturer (Pharmacia, Uppsala, Sweden). When necessary, 7-deaza-dGTP was used instead of dGTP to relieve band compression on the gels [36]. When subcloning of specific fragments was not possible, synthetic oligonucleotide primers were used instead of the universal primer. Nucleotide sequences were analyzed with the help of the University of Wisconsin GCG software package [37].

Expression of the genes encoding amicyanin and the MADH β-subunit

A 0.9-kbp *Eco*RI-*Pst*I restriction fragment (Fig. 2) was subcloned in pUC18 (plasmid pMU13) and transformed to *E. coli* JM105 according to Sambrook et al. [33]. An overnight culture of the transformant was diluted 1 : 100 in Luria-Bertani broth and grown for 2.5 h at 37°C. Isopropyl-β-D-galactopyranoside (IPTG) was added to a final concentration of 300 μM and incubation was proceeded for another 4 h. Cells were harvested by centrifugation. Proteins from the periplasmic space were isolated using chloroform treatment [38] and separated by SDS/PAGE (20% homogeneous gels) according to Laemmli [39] on a Phast System (Pharmacia). Gels were stained with Coomassie brilliant blue. Whole cell lysates were obtained by three different methods: cells were harvested by centrifugation and the pellet was either suspended in SDS-containing electrophoresis buffer [39] and heated for 5 min at 100°C, or treated with lysozyme/Triton, or incubated with lysozyme followed by 6 freeze/thaw cycles. Gel electrophoresis was performed as before. Proteins were electrophoretically transferred from gel to nitrocellulose by using the Phast Transfer system of Pharmacia. The possible presence of amicyanin or MADH was checked for by immunochemical staining of the nitrocellulose filter or by an ELISA test with polyclonal antisera as primary antibodies and goat anti-rabbit peroxidase as a secondary antibody.

RESULTS

Amino acid sequencing of the MADH β-subunit

The amino acid composition of the MADH β-subunit was determined by total hydrolysis. The results are given in Table S1. Edman degradation of the unmodified MADH β-subunit was straightforward and revealed the identity of the first 25 N-terminal residues, except for residue 23. The results are presented in Table S2. The sequence run was terminated after cycle 31. Subsequently the amino acid composition and the N-terminal amino acid sequence of the peptides T1 – T6 (except T2) were determined. The results are presented in Table S3. By comparison with the already determined amino acid sequence of the N terminus of the β-subunit of MADH, peptides T1, T2 and T3 could be positioned in the amino acid chain. The amino acid compositions of Q1, Q2 (see Table S4) and Q4 were determined, and from this it appeared that Q4 was identical with T6 [see Table S3; notice that Q4 was obtained as a subdigest of peptide Q and that T6 had a different t_R to Q (Fig. S1) but similar to Q4]. Peptide Q3 (first fraction) was subjected to 11 cycles of Edman degradation. In each step two amino acids were liberated, indicating that Q3 consists of two peptides which are held together by covalent links to the co-factor. The sequencing results of the Q3 peptide could be grouped in two peptides Q3A and Q3B (Table S5) by analogy with the results obtained by Ishii et al. [40] for the β-subunit of MADH of *Methylobacterium extorquens* AM1 (previously called *Pseudomonas* AM1 [40a]) and also on the basis of the DNA sequencing results. The amino acid at position 6 in Q3A could not be identified and apparently represents one of the attachment sites of the co-factor. When the DNA sequencing results became available, peptides T1 – T6 and Q1 – Q3 could be unambiguously positioned in the primary structure and no further attempts at amino acid sequencing were undertaken. The final results are presented in Fig. 3.

Cloning of the amicyanin and MADH-β genes

The *T. versutus* amicyanin gene was identified by screening digests of *T. versutus* chromosomal DNA with oligonucleotide probes. Several DNA fragments hybridised specifically with probe 2: a 3.4-kbp fragment of a *Bam*HI-*Sal*I digest and a 3.8-kbp fragment of a *Bam*HI-*Pst*I digest. These fragments also hybridised with probes 1 and 3, but hybridization appeared to be less selective.

The 3.4-kbp *Bam*HI-*Sal*I fragment was cloned into plasmid pUC18. After transformation to *E. coli*, about 360 transformants were screened by colony hybridization with probe 2 for insertion of the amicyanin gene. One colony gave a strong positive hybridization signal with probe 2. Subsequent incubation of the plasmid with various restriction enzymes showed that a 0.9-kbp *Pst*I-*Eco*RI fragment produced strong hybridization signals with all three probes, indicating that the whole amicyanin gene is located on the 3.4-kbp *Bam*HI-*Sal*I fragment (Fig. 2). Starting the search for the MADH β-sub-

unit-encoding gene it was then decided to first check the *Bam*HI-*Sal*I fragment before initiating screening of chromosomal digests. It appeared that probe 4 produced strong hybridization signals with this fragment and a 1.8-kbp *Eco*RI-*Sal*I subfragment (Fig. 2), indicating that at least a part of the β-subunit of MADH was present. The 3.4-kbp *Bam*HI-*Sal*I fragment was subsequently subcloned and sequenced partly in both directions by dideoxynucleotide sequencing according to the strategy shown in Fig. 2. The results are shown in Fig. 3. The data have been entered into the GenBank data bank under accession number M58001.

Expression of amicyanin and the MADH β-subunit

The 0.9-kbp *Eco*RI-*Pst*I fragment containing the amicyanin gene was subcloned in pUC18 (pMU13) and subsequently transformed to *E. coli* JM105. The transformed strain was cultured and the periplasmic fraction was analyzed for protein composition. SDS/PAGE and immunoblotting with rabbit antiserum directed against purified amicyanin showed that an amicyanin-like protein was present in the periplasmic fraction, which co-migrated with amicyanin of *T. versutus* (Fig. 4). Apparently, amicyanin is translocated to the *E. coli* periplasm and the signal peptide is cleaved to yield mature amicyanin. Expression was strongly enhanced by the addition of IPTG to the medium, indicating that the gene is transcribed from the *lac* promoter on pUC18. Neither in the periplasmic fraction nor in whole cell lysates could a 14-kDa band corresponding with signal-peptide-containing pre-amicyanin be detected.

To test for possible expression of *mad*B in *E. coli* the *Sal*I-*Pst*I and *Bal*I-*Pst*I fragments (Fig. 2) were cloned into pUC18

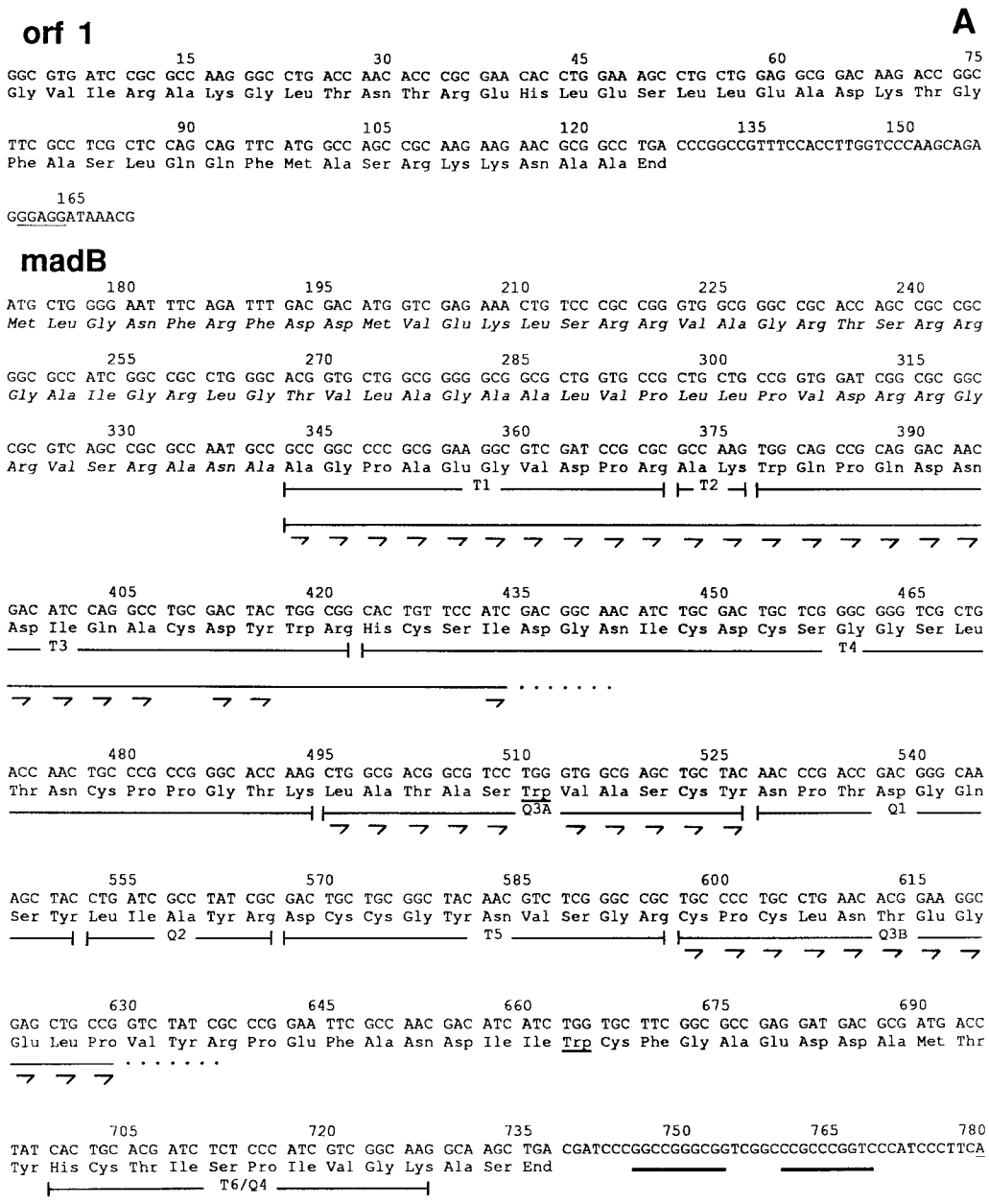

Fig. 3A

ami **B**

```
  795                 810                 825                 840                 855
ATG ATT TCT GCA AAG ACG CTT CGG CCG GCC ATC GCG GCC ATC GCC CTT TTC GCC ATC GGC GCG ACC GGA GCA TGG
Met Ile Ser Ala Lys Thr Leu Arg Pro Ala Ile Ala Ala Ile Ala Leu Phe Ala Ile Gly Ala Thr Gly Ala Trp

  870                 885                 900                 915                 930
GCA CAG GAC AAG ATC ACC GTC ACC AGC GAA AAG CCG GTC GCG GCC GCC GAT GTG CCC GCC GAT GCC GTC GTC GTC
Ala Gln Asp Lys Ile Thr Val Thr Ser Glu Lys Pro Val Ala Ala Ala Asp Val Pro Ala Asp Ala Val Val Val

  945                 960                 975                 990                1005
GGC ATC GAG AAG ATG AAA TAC CTG ACC CCG GAA GTG ACC ATC AAG GCG GGC GAG ACC GTC TAT TGG GTC AAC GGC
Gly Ile Glu Lys Met Lys Tyr Leu Thr Pro Glu Val Thr Ile Lys Ala Gly Glu Thr Val Tyr Trp Val Asn Gly

 1020                1035                1050                1065                1080
GAG GTG ATG CCC CAC AAC GTC GCC TTC AAG AAG GGC ATC GTG GGC GAG GAC GCG TTC CGC GGC GAG ATG ATG ACC
Glu Val Met Pro His Asn Val Ala Phe Lys Lys Gly Ile Val Gly Glu Asp Ala Phe Arg Gly Glu Met Met Thr

 1095                1110                1125                1140                1155
AAG GAC CAG GCC TAT GCG ATC ACC TTC AAC GAG GCG GGA AGC TAT GAC TAT TTC TGC ACG CCC CAT CCC TTC ATG
Lys Asp Gln Ala Tyr Ala Ile Thr Phe Asn Glu Ala Gly Ser Tyr Asp Tyr Phe Cys Thr Pro His Pro Phe Met

 1170                1185
CGG GGC AAG GTC ATC GTC GAA TAG GTGCCGCC
Arg Gly Lys Val Ile Val Glu End
```

orf 2

```
1200                1215                1230                1245                1260
ATG TGG ATT CCC TAT GAC CTG CGC GCC TCG CTG AAG GGC GAA ACC GAC CGG GAA ACG CTG CGC CTG TCG CGG GCC
Met Trp Ile Pro Tyr Asp Leu Arg Ala Ser Leu Lys Gly Glu Thr Asp Arg Glu Thr Leu Arg Leu Ser Arg Ala

1275                1290                1305                1320                1335
GAC GGG CAG CCG CGC CAG TTC CTG GTG GGG TTC TTC CTG CGC AAC CCG GTC ACG CAG GCC TGG GAA CTG GAC CTG
Asp Gly Gln Pro Arg Gln Phe Leu Val Gly Phe Phe Leu Arg Asn Pro Val Thr Gln Ala Trp Glu Leu Asp Leu

1350                1365                1380                1395                1410
GCG GCC GAG GAC GGC CTG CCG GAA CTG CCT GCC GTG CCG GCG GGG GCG AGT TTC AGC ATC TGC CCG AAC GAG GCC
Ala Ala Glu Asp Gly Leu Pro Glu Leu Pro Ala Val Pro Ala Gly Ala Ser Phe Ser Ile Cys Pro Asn Glu Ala

1425                1440                1455                1470                1485
GGC AAG CTG GCC GAG GTG ATC TAT CGC CTG CCC GCG CGT TCG GCC ACC GAG GCG CTG GAG CTT GCC CAT GCC GAC
Gly Lys Leu Ala Glu Val Ile Tyr Arg Leu Pro Ala Arg Ser Ala Thr Glu Ala Leu Glu Leu Ala His Ala Asp

1500
CTG CAG
Leu Gln
```

Fig. 3. *DNA sequence of the* T. versutus mad*B (A) and* ami *(B) genes and inferred protein sequences.* Putative ribosome-binding sites are underscored. Leader peptides are shown in italics. A region of dyad symmetry is underlined in bold. Tryptic (T1 – T6) or chymotryptic (Q1 – Q4) peptides of the MADH β-subunit for which sequences or amino acid composition were determined are indicated by horizontal lines. Arrows indicate the results of amino acid sequencing work. Trp57 and Trp108, involved in the formation of the co-factor of the β-subunit of MADH, are doubly underscored

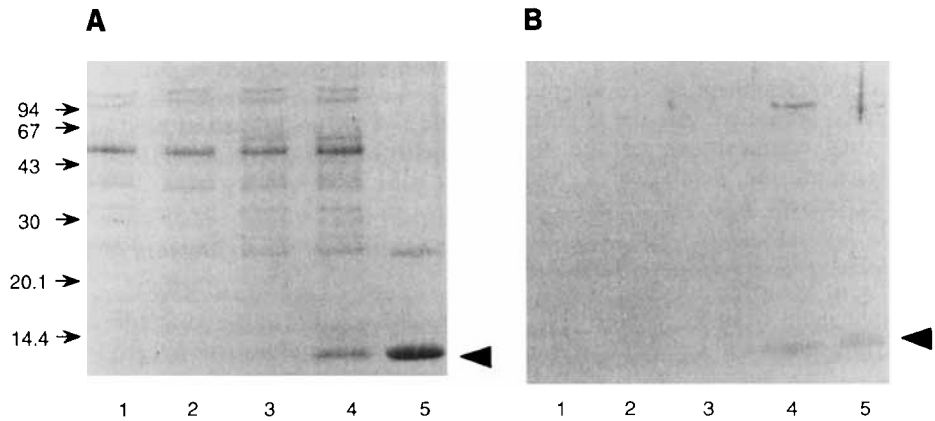

Fig. 4. *SDS/PAGE and immunoblotting of periplasmic proteins isolated (see Materials and Methods for details) from* E. coli *strain JM105 in which the* ami *gene is expressed under control of the* lac *promoter.* (A) SDS/polyacrylamide gel, in which proteins were stained with Coomassie brilliant blue. (B) Immunoblot of a gel identical to that in (A) which was probed with rabbit antiserum directed against purified amicyanin. Lanes 1, *E. coli* JM105/pUC18; 2, as 1 but with IPTG induction; 3, *E. coli* JM105/pUC18 containing the *Eco*RI-*Pst*I fragment (pMU13); 4, as 3 but with IPTG induction; 5 partially purified amicyanin from *T. versutus*. The sizes (kDa) of the molecular mass markers are indicated on the left. The arrows at the right-hand side of the gels indicate the position of the amicyanin band [12(±1) kDa]

(resulting in plasmids pMK1 and pMK2, respectively) and transformed to *E. coli* JM101. The transformed strains were cultured in the presence and absence of 300 μM IPTG. Whole cell lysates, obtained as described in Materials and Methods, were separated by electrophoresis and the gels were coloured with Coomassie brilliant blue. No extra bands at 14.2 kDa or 20.4 kDa corresponding with the processed or unprocessed MADH β-subunit could be detected in the gels, while amicyanin bands were present in samples from cultures induced by IPTG. An ELISA test for the MADH β-subunit was

negative. It is concluded that the amount of MADH β-subunit if present at all, had to be less than 0.1% of total protein content.

DISCUSSION

Amino acid sequence of the small MADH β-subunit

Although the final amino acid sequence of the MADH β-subunit was established from the corresponding DNA sequence, the amino acid sequencing results were necessary for two reasons. Firstly, primary structure information was needed to be able to later unambiguously identify the gene encoding the MADH β-subunit. Secondly, information about the nature and the position of the attachment of the co-factor in the polypeptide chain can only be obtained from studying the amino acid sequence directly. The interpretation of the amino acid sequencing experiments initially profited strongly from the sequencing work of Ishii et al. [40] on the MADH β-subunit of *M. extorquens* AM1. (Eventually a large homology was found between the MADH β-subunits from the two organisms). The N-terminal amino acid sequence appears in full agreement with the results deduced later from the DNA sequencing work (Fig. 3). The attachment site for the co-factor in peptide Q3A after alignment with the full amino acid sequence, appears to be residue 57 (nucleotides 510–512; Fig. 3). The second attachment site (namely, in peptide Q3B) is assigned on the basis of the strong similarity with the *M. extorquens* AM1 MADH [40] as position 108 (nucleotides 663–665; Fig. 3). After DNA sequencing, peptides T1–T6 and Q1 and Q2 could unambiguously be positioned in the amino acid chain. The observed cleavage positions by trypsin and chymotrypsin are in agreement with the specificities of the two enzymes, except for a chymotryptic cleavage after Tyr119 with trypsin (peptide T6).

Nucleotide sequence and analysis of the T. versutus *amicyanin and MADH genes*

Within the sequenced DNA fragment two complete open reading frames (ORF) were identified. All the determined amino acid sequences and compositions of the tryptic and chymotryptic peptides of the β-subunit of MADH (Tables S1–S5) are in agreement with the amino acid sequence deduced from the first complete ORF (Fig. 3). The amino acid sequence derived from the DNA sequence of the second complete ORF is in complete agreement with the published sequence of amicyanin from *T. versutus* [14]. The two genes will be denoted by *mad*B and *ami* in the rest of the paper.

The ATG startcodon (nucleotide 791) of the *ami* gene is located 78 bp upstream of the codon which corresponds to the first residue of mature amicyanin (nucleotide 869). A consensus ribosome-binding site [41] (AGGAGG, underlined in Fig. 3) occurs 6–11 bp upstream from the start codon. The mature amicyanin is preceded by a 25-amino-acid-long signal peptide (not counting the ATG start codon), apparently serving the purpose of translocating the protein across the periplasmic membrane. The signal sequence shows the typical signal peptide characteristics [42]: positively charged amino acids near the N-terminus, a central hydrophobic region (amino acids −17 to −8; nucleotides 818–847) and two small uncharged residues (alanine in both cases) in positions −3 and −1. Similar findings have been reported for other bacterial blue copper protein genes [32, 43, 44]. The mature amicyanin has a M_r of 11690 according to the amino acid sequence, in agreement with literature data [15].

The *mad*B gene is located upstream of the *ami* gene. A Shine-Dalgarno sequence [41] (GGAGG, underlined in Fig. 3) occurs 8–12 bp in front of the ATG start codon. Comparison of the amino acid and DNA sequencing results reveals that residues 1–25 of the mature protein as identified by amino acid sequencing are identical to amino acids 58–82 as derived from the nucleotide sequence (nucleotides 342–416) of the *mad*B gene. This indicates that the first 171 nucleotides of the *mad*B gene code for a putative 56-amino-acid-long leader peptide (not counting the ATG start codon). Cleavage of the signal sequence results in a protein with a calculated relative molecular mass of M_r = 14198, in agreement with the molecular-mass determination by gel electrophoresis [17]. The leader peptide is unusually long and contains positively as well as negatively charged residues at both the N-terminal and C-terminal regions. It, therefore, does not conform to the consensus structure of simple prokaryotic signal peptides from Gram-negative micro-organisms. There is, however, some resemblance to the hydrogenases, which are heterologous dimers of a large α-subunit and a small β-subunit [45], the latter preceded by a long leader peptide. It has been suggested that this peptide is involved in translocation of both subunits across the inner membrane [46].

The stop codon of the *ami* gene (nucleotide 1187) overlaps with a putative ribosome-binding site (AGG, underlined in Fig. 3) of a third ORF (marked by orf 2 in Fig. 3), which is followed by an ATG initiator codon (nucleotide 1198). A codon usage analysis (results not shown) supports the presence of a potential ORF. The nucleotide sequence from this ORF was compared with known sequences of the GenBank, but no obvious homology could be found. Similarly, upstream of the *mad*B gene the 3′ region of a fourth unidentified ORF was found (indicated by orf 1 in Fig. 3).

Neither *E. coli* [47] nor *Paracoccus denitrificans* promotor-like sequences ([20] and references cited therein), nor potential ϱ-independent transcription terminators could be found within the sequenced region, although a region of dyad symmetry is present between the *madB* and *ami* genes (nucleotides 745–769, underlined in bold in Fig. 3). The same situation exists in *P. denitrificans* [11].

The nature of the co-factor in MADH

Several suggestions have been made in the past with respect to the structure of the co-factor: (a) based on spectral evidence and sensitivity for carbonyl group reagents, a pyridoxal-phosphate-like compound was suggested [48]; (b) mass spectroscopy of the semicarbazide-derivatised co-factor suggested a PQQ-like structure without carboxylic acid groups [49]; (c) derivatization with phenylhydrazine provided a product considered to be the phenylhydrazone of PQQ [13]; (d) X-ray diffraction studies of crystals of MADH suggested pro-PQQ [50] or a variant thereof [17] as the co-factor. Amino acid sequencing of the β-subunit of MADH from *M. extorquens* AM1 indicated that the co-factor is covalently bound to unidentified residues at positions 55 and 106 (57 and 108 using the numbering in *T. versutus*) [40]. Recently it appeared that these positions concern tryptophans, as deduced from the nucleotide sequence of the gene [51]. The results presented here for the β-subunit of MADH from *T. versutus* are in full agreement with those for *M. extorquens*.

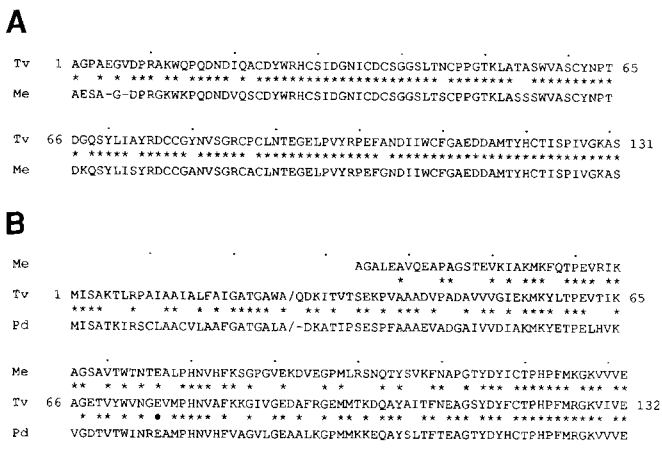

Fig. 5. *Comparison of amino acid sequences of the β-subunits of MADH (A) and of amicyanins (B) from different organisms.* Tv, deduced amino acid sequence from *T. versutus*; Pd, deduced amino acid sequence of amicyanin from *P. denitrificans* [11]; Me, amino acid sequences from *M. extorquens* AM1 [40, 51, 53]

X-ray analysis gave an electron-density map which could be fitted with pro-PQQ, a structure comprising amino acid residues 57 and 108 (erroneously considered to be glutamic acid and a modified tyrosine, respectively) [50]. Since the residues appear to be tryptophans, the volume predicted from the electron-density map does not allow PQQ or pro-PQQ to bridge these two tryptophans, the latter forming the body of the co-factor. Work is in progress now to elucidate the precise structure of the quinonoid tryptophan dimer.

Comparison with other sequences

The amino acid sequence of the β-subunit of MADH from *T. versutus* shows a very high similarity with the protein sequence of *M. extorquens* [40, 51]. Of the 131 amino acids, 115 are identical (Fig. 5A). This is consistent with immunological data, indicating that the structural features of the β-subunit of MADH are conserved in taxonomically diverse bacteria [52]. Of the amino acid sequence of *T. versutus* amicyanin 62% and 51% is identical with the sequences of the amicyanins found in *P. denitrificans* and *M. extorquens*, respectively [11, 53] (Fig. 5B). An overall DNA similarity of 72% was found for the *ami* genes from *T. versutus* and *P. denitrificans*. The findings confirm that *T. versutus* is strongly interrelated with *P. denitrificans* as concluded previously from a phylogenetic analysis of the genus Thiobacillus [54].

The sequenced fragment has a G+C content of 66.9%, which agrees with an estimated total G+C content of 68% for *T. versutus* [23] and the observed strong bias toward the use of G and C in the first and third positions.

Concluding remarks

The results reported here show that the genes coding for amicyanin and the β-subunit of MADH are found on adjacent locations on the bacterial chromosome in *T. versutus*, most likely as part of one operon. Although the MADH α-subunit-encoding gene has not yet been located, the shortness of the stretch between the *ami* and the *mad*B genes and the lack of clear terminator or promoter sequences in this region make it likely that amicyanin and MADH are coordinately expressed. From a functional point of view this would make sense since it is known that the *T. versutus* amicyanin can operate as an electron acceptor of the MADH *in vitro* [55] while for the closely related *P. denitrificans* the same has been shown to apply *in vivo* [11].

Remarkably, the *ami*, but not the *mad*B gene could be expressed in *E. coli*. The successful expression of the *ami* gene opens the way to site-directed mutagenesis of amicyanin. In expression studies of *mad*B, the *ami* gene functioned as an internal control. The presence of amicyanin in cells containing pMK1 and pMK2 implies that transcription of the *mad*B gene must have taken place, since no promotor sequences can be found between the *ami* and *mad*B gene. Furthermore, amicyanin production was dependent on the presence of IPTG in the culture medium. Therefore, the absence of the MADH β-subunit must be due either to lack of translation or rapid degradation of the protein. It is conceivable that enzymes are necessary for co-factor synthesis and that these enzymes are not present in *E. coli*. It is also possible that the translated polypeptide cannot be translocated because of the unusual leader sequence and that the presence of the α-subunit is required for successful transport across the inner membrane. (The latter question will be addressed when the identification of the α-subunit-encoding gene, presently underway, has been completed.) Finally, the combined information available from the amino acid sequencing work, the DNA sequencing work, the enzymological [13] and the crystallographic studies [50] of MADH demonstrates that the MADH co-factor is located in the β-subunit, contains an ortho-quinone function and that its *in vivo* synthesis involves two tryptophan side chains. Further work is needed to establish the precise nature of the co-factor.

The authors thank Dr L. F. Oltmann and Dr R. van Spanning for providing them with information concerning the studies on *P. denitrificans*. They also thank P. A. Jekel and P. Wietzes for technical assistance. Finally, they thank M. van de Kamp and T. den Blaauwen for fruitful discussions. Part of this work was carried out under a contract of the EEC Biotechnology Action Programme.

REFERENCES

1. Van Wielink, J. E., Frank Jzn., J. & Duine, J. A. (1988) in *Proceedings of the first international symposium on PQQ and quinoproteins* (Jongejan, J. A. & Duine, J. A., eds) pp. 269–279, Kluwer Academic Publishers, Dordrecht, The Netherlands.
2. Large, P. J. (1990) *FEMS Microbiol. Rev. 87*, 235–240.
3. Fukumori, Y. & Yamanaka, T. (1987) *J. Biochem. (Tokyo) 101*, 441–445.
4. Lommen, A., Ratsma, A., Bijlsma, N., Canters, G. W., van Wielink, J. E., Frank, J. & Van Beeumen, J. (1990) *Eur. J. Biochem. 192*, 653–661.
5. Gray, K. A., Davidson, V. L. & Knaff, D. B. (1988) *J. Biol. Chem. 263*, 13987–13990.
6. Lommen, A. & Canters, G. W. (1990) *J. Biol. Chem. 265*, 2768–2774.
7. Chandrasekar, R. & Klapper, M. H. (1986) *J. Biol. Chem. 261*, 3616–3619.
8. Gray, K. A., Knaff, D. B., Husain, M. & Davidson, V. L. (1986) *FEBS Lett. 207*, 239–242.
9. Tobari, J. & Harada, Y. (1981) *Biochem. Biophys. Res. Commun. 101*, 502–508.
10. Kumar, M. A. & Davidson, V. L. (1990) *Biochemistry 29*, 5299–5304.
11. Van Spanning, R. J. M., Wansell, C. W., Reijnders, W. N. M., Oltmann, L. F. & Stouthamer, A. H. (1990) *FEBS Lett. 275*, 217–220.
12. De Beer, R., Duine, J. A., Frank Jzn., J. & Large, P. J. (1980) *Biochem. Biophys. Acta 622*, 370–374.

13. Van der Meer, R. A., Jongejan, J. A. & Duine, J. A. (1987) *FEBS Lett. 221*, 299–304.
14. Van Beeumen, J., Van Bun, S., Canters, G. W., Lommen, A. & Chothia, C. (1991) *J. Biol. Chem. 266*, 4869–4877.
15. Van Houwelingen, T., Canters, G. W., Stobbelaar, G., Duine, J. A., Frank Jzn., J. & Tsugita, A. (1985) *Eur. J. Biochem. 153*, 75–80.
16. Vellieux, F. M. D., Frank Jzn., J., Swarte, M. B. A., Groendijk, H., Duine, J. A., Drenth, J. & Hol, W. G. J. (1986) *Eur. J. Biochem. 154*, 383–386.
17. Vellieux, F. M. D. & Hol, W. G. J. (1989) *FEBS Lett. 255*, 460–464.
18. Petratos, K., Dauter, Z., Wilson, K. S., Lommen, A., Van Beeumen, J. & Canters, G. W. (1988) *J. Mol. Biol. 199*, 545–546.
19. Lommen, A., Wijmenga, S., Hilbers, C. W. & Canters, G. W. (1991) *Eur. J. Biochem.*, in the press.
20. Van Spanning, R. J. M., Wansell, C., Harms, N., Oltmann, L. F. & Stouthamer, A. H. (1990) *J. Bacteriol. 172*, 986–996.
21. Van de Kamp, M., Floris, R., Hali, F. C. & Canters, G. W. (1990) *J. Am. Chem. Soc. 112*, 907–908.
22. Van de Kamp, M., Silvestrini, M. C., Brunori, M., Van Beeumen, J., Hali, F. C. & Canters, G. W. (1990) *Eur. J. Biochem. 194*, 109–118.
23. Harrison, A. P. (1983) *Int. J. Syst. Bacteriol. 33*, 211–217.
24. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene 33*, 103–119.
25. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990–7997.
26. Chang, J. Y., Bramer, D. & Wittmann-Liebold, B. (1978) *FEBS Lett. 93*, 205–214.
27. Penke, B., Ferenczi, R. & Kovacs, K. (1974) *Anal. Biochem. 60*, 45–50.
28. Harlow, E. & Lane, D. (1988) *Antibodies, a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, USA.
29. Chater, K. F., Hopwood, D. A., Kieser, T. & Thompson, C. J. (1982) *Curr. Top. Microbiol. Immunol. 96*, 69–95.
30. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.
31. Meinkoth, J. & Wahl, G. (1984) *Anal. Biochem. 138*, 267–284.
32. Canters, G. W. (1987) *FEBS Lett. 212*, 168–172.
33. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular cloning, a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.
34. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.
35. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl Acad. Sci. USA 74*, 5463–5467.
36. Mizusawa, S., Nishimura, S. & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319–1324.
37. Gribskov, M., Devereux, J. & Burgess, R. R. (1984) *Nucleic Acids Res. 12*, 539–549.
38. Ferro-Luzze Ames, G., Prody, C. & Kustu, S. (1984) *J. Bacteriol. 160*, 1181–1183.
39. Laemmli, U. K. (1970) *Nature 227*, 680–685.
40. Ishii, Y., Hase, T., Fukumori, Y., Matsubara, H. & Tobari, J. (1983) *J. Biochem. (Tokyo) 93*, 107–119.
40a. Green, P. N., Bousfield, I. J. & Hood, D. (1988) *Int. J. Syst. Bacteriol. 38*, 124–127.
41. Shine, J. & Dalgarno, L. (1980) *Nature 254*, 34–38.
42. Duffaud, G. D., Lehnhardt, S. K., March, P. E. & Inouye, M. (1985) *Curr. Top. Membr. Transp. 24*, 65–104.
43. Yamamoto, K., Uozumi, T. & Beppu, T. (1987) *J. Bacteriol. 169*, 5648–5652.
44. Hoitink, C. W. G., Woudt, L. P., Turenhout, J. C. M., van de Kamp, M. & Canters, G. W. (1990) *Gene 90*, 15–20.
45. Voordouw, G., Strang, J. D. & Wilson, F. R. (1989) *J. Bacteriol. 171*, 3881–3889.
46. Van Dongen, W., Hagen, W., van den Berg, W. & Veeger, C. (1988) *FEMS Microbiol. Lett. 50*, 5–9.
47. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res. 11*, 2237–2255.
48. Eady, R. R. & Large, P. J. (1971) *Biochem. J. 123*, 757–771.
49. McIntire, W. S. & Stults, J. T. (1986) *Biochem. Biophys. Res. Commun. 141*, 562–568.
50. Vellieux, F. M. D., Huitema, F., Groendijk, H., Kalk, K. H., Frank Jzn., J., Jongejan, J. A., Duine, J. A., Petratos, K., Drenth, J. & Hol, W. G. J. (1989) *EMBO J. 8*, 2171–2178.
51. Chistoserdov, A. Y., Tsygankov, Y. D. & Lidstrom, M. E. (1990) *Biochem. Biophys. Res. Commun. 172*, 211–216.
52. Davidson, V. L. & Neher, J. W. (1987) *FEMS Microbiol. Lett. 44*, 121–124.
53. Ambler, R. P. & Tobari, J. (1985) *Biochem. J. 232*, 451–457.
54. Lane, D. J., Stahl, D. A., Olsen, G. J., Heller, D. J. & Pace, N. R. (1985) *J. Bacteriol. 163*, 75–81.
55. Van Wielink, J. E. & Duine, J. A. (1990) *Trends Biochem. Sci. 15*, 136–137.
56. McIntyre, W. S., Wemmer, D. E., Chistoserdov, A. & Lidstrom, M. E. (1991) *Science 252*, 817–824.
57. Chen, L., Mathews, F. S., Davidson, V. L., Huizinga, E. G., Vellieux, F. M. D., Duine, J. A. & Hol, W. G. J. (1991) *FEBS Lett. 287*, 163–166.

Supplementary material to:

Cloning, sequencing and expression studies of the genes encoding amicyanin and the β-subunit of methylamine dehydrogenase from *Thiobacillus versutus*

Marcellus UBBINK, Mario A. G. van KLEEF, Dirk-Jan KLEINJAN, Carla W. G. HOITINK, Fienke HUITEMA, Jaap J. BEINTEMA, Johannis A. DUINE and Gerard W. CANTERS

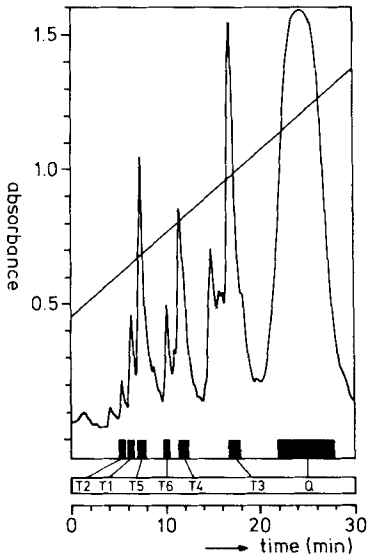

Fig. S1. *Elution pattern of the tryptic digest of the β-subunit of MADH by HPLC.* Absorbance was measured at 214 nm

Table S1. *Amino acid composition of the MADH β-subunit*
Numbers in parentheses indicate the number of times an amino acid occurs in the protein sequence as derived from the DNA sequence. Total number of residues 131

Amino acid	Occurrence
	mol/mol
Asx	19.2 (18)
Thr	7.9 (7)
Ser	10.3 (9)
Glx	9.0 (9)
Pro	10.4 (10)
Gly	13.0 (12)
Ala	10.3 (12)
Cys	11.5 (12)
Val	3.2 (5)
Met	0.7 (1)
Ile	4.8 (8)
Leu	4.9 (5)
Tyr	6.5 (7)
Phe	1.2 (2)
Lys	3.0 (3)
His	2.5 (2)
Arg	5.7 (5)
Trp	2.1 (4)[a]

[a] Two Trp residues are probably linked through a non-hydrolizable bond and are modified as part of the cofactor (see text).

Table S2. *Sequence analysis of N-terminus of the unmodified β-subunit of MADH*
Pth, phenylthiohydantoin. Cys was determined as pyridylethylcysteine

Cycle	Amino acid	Pth-Xaa content
		pmol
1	Ala	208
2	Gly	134
3	Pro	153
4	Ala	154
5	Glu	94
6	Gly	112
7	Val	132
8	Asp	42
9	Pro	54
10	Arg	18
11	Ala	95
12	Lys	51
13	Trp	26
14	Gln	30
15	Pro	52
16	Gln	42
17	Asp	14
18	Asn	25
19	Asp	41
20	Ile	29
21	Gln	8
22	Ala	24
23	(Cys)	—
24	Asp	6
25	Tyr	16
26	—	—
27	—	—
28	—	—
29	—	—
30	—	—
31	Ile	3

Table S3. *Amino acid composition of peptides T1 and T3 – T6*
The peptides are derived from a tryptic digest of the β-subunit of MADH. Numbers in parentheses indicate the number of times an amino acid occurs in the peptide chain as derived from the DNA sequence. n.d., not determined. N termini were Ala, Trp, His, Asp and His for peptides T1, T3, T4, T5 and T6/Q4, respectively

Amino acid	Content in peptide				
	T1	T3	T4	T5	T6/Q4
	mol/mol				
Asx	1.3 (1)	3.7 (4)	3.8 (4)	1.9 (2)	0.1 (0)
Thr			2.0 (2)		1.0 (1)
Ser	0.2 (0)		3.0 (3)	0.9 (1)	1.1 (1)
Glx	0.9 (1)	3.1 (3)			
Pro	1.5 (2)	1.3 (1)	2.1 (2)		1.1 (1)
Gly	2.1 (2)		4.0 (4)	2.0 (2)	1.1 (1)
Ala	1.9 (2)	1.0 (1)			
Cys		+[a](1)	+[a](4)	+[a](2)	+[a](1)
Val	0.8 (1)			1.1 (1)	0.6 (1)
Met					
Ile	0.2 (0)	1.0 (1)	1.9 (2)		1.8 (2)
Leu			1.1 (1)		
Tyr		1.2 (1)		1.1 (1)	
Phe					
His			1.0 (1)		0.6 (1)
Lys	0.1 (0)		1.1 (1)		0.5 (1)
Arg	1.0 (1)	1.0 (1)		1.0 (1)	
Trp	n.d. (0)	n.d. (2)	n.d. (0)	n.d. (0)	n.d. (0)
Total	(10)	(15)	(24)	(10)	(10)

[a] Present, but not determined quantitatively.

Table S4. *Amino acid composition of peptides Q1 and Q2 of the chymotryptic digest of tryptic peptide Q*
Numbers in parentheses indicate the number of times an amino acid occurs in the peptide chain as derived from the DNA sequence. n.d., not determined

Amino acid	Content in peptide	
	Q1	Q2
	mol/mol	
Asx	1.5 (2)	0.1 (0)
Thr	0.7 (1)	0.2 (0)
Ser	0.8 (1)	0.4 (0)
Glx	1.0 (1)	0.3 (0)
Pro	0.0 (1)	
Gly	1.1 (1)	0.3 (0)
Ala	0.5 (0)	1.0 (1)
Cys		
Val	0.3 (0)	
Met		
Ile		1.0 (1)
Leu		1.0 (1)
Tyr	1.0 (1)	0.8 (1)
Phe		
Lys		
His		
Arg		1.0 (1)
Trp	n.d.	n.d.
Total	(8)	(5)

Table S5. *Sequence analysis of peptide Q3 generated by chymotrypsin treatment of peptide Q*
In each cycle two amino acids were liberated which have been grouped into two sequences, Q3A and Q3B. Increases of the amounts of Pth-Xaa are in pmol

Cycle	Peptide			
	Q3A		Q3B	
		pmol		pmol
1	Leu	132	Cys	70
2	Ala	158	Pro	206
3	Thr	132	Cys[c]	119
4	Ala	125	Leu	183
5	Ser	52	Asn	84
6	Xaa	–[a]	Thr	116
7	Val	67	Glu	87
8	Ala	68	Gly	124
9	Ser	41	Glu	70
10	(Cys)	–[b]	Leu	110
11	Tyr	35	Pro	77

[a] Amino acid at this position could not be identified. Only glycine (< 85 pmol) could be detected.
[b] Determined as pyridylethylcysteine, but not quantitatively.
[c] Determined as pyridylethylcysteine.