The *napEDABC* gene cluster encoding the periplasmic nitrate reductase system of *Thiosphaera pantotropha*

Ben C. BERKS,*† David J. RICHARDSON,*§ Ann REILLY,* Anthony C. WILLIS‡ and Stuart J. FERGUSON†

*Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, University Plain, Norwich NR4 7TJ, and †Department of Biochemistry and ‡MRC Immunochemistry Unit, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

The *napEDABC* locus coding for the periplasmic nitrate reductase of *Thiosphaera pantotropha* has been cloned and sequenced. The large and small subunits of the enzyme are coded by *napA* and *napB*. The sequence of NapA indicates that this protein binds the GMP-conjugated form of the molybdopterin cofactor. Cysteine-181 is proposed to ligate the molybdenum atom. It is inferred that the active site of the periplasmic nitrate reductase is structurally related to those of the molybdenum-dependent formate dehydrogenases and bacterial assimilatory nitrate reductases, but is distinct from that of the membrane-bound respiratory nitrate reductases. A four-cysteine motif at the N-terminus of NapA binds a [4Fe-4S] cluster. The DNA- and protein-derived primary sequence of NapB confirm that this protein is a dihaem *c*-type cytochrome and, together with spectroscopic data, indicate that both NapB haems have bis-histidine ligation. *napC* is predicted to code for a membrane-anchored tetrahaem *c*-type cytochrome that shows sequence similarity to the NirT cytochrome *c* family. NapC may be the direct electron donor to the NapAB complex. *napD* is predicted to encode a soluble cytoplasmic protein and *napE* a monotopic integral membrane protein. *napDABC* genes can be discerned at the *aeg-46.5* locus of *Escherichia coli* K-12, suggesting that this operon encodes a periplasmic nitrate reductase system, while *napD* and *napC* are identified adjacent to the *napAB* genes of *Alcaligenes eutrophus* H16.

INTRODUCTION

Thiosphaera pantotropha possesses a soluble periplasmically located respiratory nitrate reductase [1,2]. This periplasmic nitrate reductase is a heterodimer of a 93 kDa (NapA) subunit and a 16 kDa cytochrome c_{552} (NapB) [3]. A combination of spectroscopic and MS evidence suggests that NapB binds two haem groups [3,4]. The periplasmic nitrate reductase also contains molybdenum and a $[4Fe\text{-}4S]^{2+,1+}$ centre [3,4]. It has been suggested that both these cofactors are bound by NapA and that the molybdenum is present in the form of a molybdopterin cofactor [3–5]. The molybdenum atom is thought to be involved in transfer of an oxygen atom from nitrate to water [5,6].

T. pantotropha periplasmic nitrate reductase is normally only expressed at high levels during aerobic growth [1,2,7]. Expression is maximal when cells are grown on highly reduced carbon sources, but is unaffected by the presence or absence of nitrate. The function of the periplasmic nitrate reductase system in *T. pantotropha* is unclear. A rôle in disposing of excess reductant has been suggested [7].

Periplasmic nitrate reductases resembling the *T. pantotropha* enzyme have been described for other bacteria [8–13], and a sequence has recently been reported for the two structural genes (*napAB*) of a periplasmic nitrate reductase coded on an *Alcaligenes eutrophus* megaplasmid [12]. We now report the sequence of the complete periplasmic nitrate reductase (*nap*) locus of *T. pantotropha*.

It has recently been proposed that *T. pantotropha* be reclassified as a strain of *Paracoccus denitrificans* ([14]; see also [15]).

EXPERIMENTAL

Protein sequencing

T. pantotropha periplasmic nitrate reductase was purified as previously described [3]. In initial experiments the two enzyme subunits were separated by SDS/10%-PAGE. Before electrophoresis Nap (150 μg) was denatured by heating at 100 °C for 3 min in the presence of 3% SDS and 5% 2-mercaptoethanol. The resolved polypeptides were electroblotted on to a ProBlott membrane (Applied Biosystems Ltd., Warrington, Cheshire, U.K.) and NapA detected by staining with Coomassie Blue [16]. NapA isolated by this method was found to be N-terminally blocked. A membrane strip bearing the 93 kDa NapA subunit was excised and digested *in situ* with trypsin (sequencing grade from Promega Ltd., Southampton, U.K.) using an estimated enzyme/substrate ratio of 1:20 (w/w), essentially by the method of Aebersold et al. [17]. Released peptides were resolved by reverse-phase HPLC in 0.1% trifluoroacetic acid (TFA) and a 0–100% acetonitrile gradient using a Brownlee Aquapore RP-300 column (100 mm × 2 mm) (Applied Biosystems Ltd.) connected to a Severn Analytical HPLC system (HPLC technology Ltd., Macclesfield, Cheshire, U.K.). Peptides were collected manually and N-terminally sequenced (see Table 1 below) on an Applied Biosystems 470A protein sequencer with on-line phenylthiohydantoin-amino acid analyser.

As NapB is not stained by Coomassie Blue [3], this subunit was prepared by subjecting purified periplasmic nitrate reductase (100 μg) to reverse-phase HPLC as described above, except that a Vydac C4 column (150 mm × 2 mm; Hichrom Ltd., Reading,

Abbreviations used: MGD, GMP derivative of the molybdopterin cofactor; TFA, trifluoroacetic acid.

§ To whom correspondence should be addressed.

The nucleotide sequence presented here has been deposited in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number Z36773.

U.K.) was used. Two major peaks were observed and were assigned to NapA (eluted at 58% acetonitrile) and NapB (eluted at 44% acetonitrile) on the basis of their relative absorbances at 215 nm. Both NapA and NapB prepared by this method were found to be N-terminally blocked. Fractions containing NapB were freeze-dried, then subjected to trypsin digestion, separation of peptide fragments by reverse-phase HPLC and N-terminal sequencing of selected peaks as described above and in the legend to Figure 1 (below).

Bacterial strains and genetic techniques

T. pantotropha was grown as previously described [2]. *Escherichia coli* strains JM89, DH5a and SURE (Stratagene cloning systems) and the plasmids pUC18 and pBSIIKS+ (Stratagene) were used for routine DNA manipulations. Standard cloning methods were as described [18]. The degenerate primers used to amplify the *napA* fragment were designed taking into account the codon bias of known *P. denitrificans* genes and had the sequences GA(G/A)GA(G/A)GC(G/C)TTCGA(C/T)GA(C/T)ATGG (from peptide 9235H) and AGTC(G/C)GCGAA(C/T)TC(C/T)-TCGAA(G/C)GT (from peptide 9216H). The amplification reaction used *Taq* polymerase, and the product was cloned into plasmid pCR™II (Invitrogen Corporation) to give pTA1. Probes used in Southern blotting were labelled using the DIG system (Boehringer Mannheim). A *T. pantotropha* chromosomal library was constructed by cloning chromosomal DNA partially digested with *Mbo*I into the *Bam*HI site of SuperCos1 (Stratagene cloning systems). Clones of interest were selected by Southern blotting, using the pTA1 insert as a probe. Double-stranded DNA was sequenced with a Pharmacia automated laser fluorescence ('A.L.F.') system using T7 polymerase according to the manufacturer's protocol. DNA sequences were determined from a combination of subclones and custom primers. All DNA sequence in the region shown in Figure 2 (below) was determined a minimum of three times in both the forward and reverse directions.

Computer-assisted sequence analysis

The University of Wisconsin Genetic Computer Group software package release 7.3 [19] with its GCGNCBI and GCGEMBL extensions was employed. Database searches used the program BLAST [19a]. Determination of open reading frames included the use of the program GenMark [19b]. Membrane-spanning transmembrane helices were assigned for aligned multiple sequences by using the statistical analysis described by Persson and Argos [20]. These assignments were in broad agreement with analysis of individual sequences using the statistical/topology analysis by Jones et al. [21] and with hydropathy analysis using the physicochemical scale of Engelman et al. [22]. The topology of transmembrane sequences with respect to the two sides of the cytoplasmic membrane was assigned for each sequence by the algorithm of Jones et al. [21]. In all cases these predictions agreed with simple assignment of topology by the 'inside positive' rule of von Heijne [23]. Where indicated the possible presence of signal sequences was assessed using the method of von Heijne [24].

RESULTS AND DISCUSSION

Analysis of NapB haem-binding polypeptides

We have previously suggested, on the basis of MS, redox potentiometry and haem quantifications, that NapB contains two covalently bound (*c*-type) haem groups [3,4]. This was confirmed by the experiments shown in Figure 1. HPLC-purified NapB was subjected to trypsin digestion, and the resulting polypeptides resolved by reverse-phase HPLC (Figure 1). By monitoring the separation at a wavelength (410 nm) close to the Soret-band maximum of *c*-type cytochromes, two peptides with covalently bound haem were identified (Figure 1). The amino acid sequences of these peptides were determined (Table 1). Each peptide sequence ends in an arginine residue. As trypsin cleaves on the C-terminal side of basic amino acids, this indicated that the complete sequence of each peptide had been obtained. Both peptide sequences contain a blank-Xaa$_2$-blank-His motif in which the blank cycles presumably contain a modified amino acid. Haem in *c*-type cytochromes is usually covalently bound to cysteine residues in a Cys-Xaa$_2$-Cys-His motif. Our interpretation of the protein sequence is thus that each peptide binds a single haem group, leading to the conclusion that, in agreement with earlier results [3,4], NapB is a dihaem *c*-type cytochrome.

Nucleotide sequence of the *nap* locus

The *nap* locus was cloned using NapA peptide sequence data as described in the Experimental section. The locus contains five open reading frames, *napEDABC*, transcribed in the same direction (Figure 2). All five open reading frames have a codon usage and GC bias at the third position similar to that found in *P. denitrificans* genes (results not shown). The stop codon of *napD* overlaps the start codon of *napA*. *napA* and *napB* overlap by 28 bp, including the stop codon. The stop codon of *napB* and the start codon of *napC* are separated by 1 bp, while 9 bp separate the stop and start codons of *napE* and *napD*.

A sixth partial open reading frame transcribed in the same direction as the *nap* genes is found approx. 400 bp before the start of *napE* (Figure 2). The protein coded by this gene shows high sequence similarity (40% identity in the region sequenced) to one of the integral membrane subunits (DctM) of a dicarboxylate transporter from the closely related bacterium *Rhodobacter capsulatus* ([25]; D. J. Kelly, personal communication) as well as to related proteins in *E. coli* ([26]; 31% identity in the region sequenced) and *Bordetella pertussis* [27]. As a consequence we provisionally designate this open reading frame *dctM*. The presence of 400 bp of apparently non-coding DNA between *dctM* and *napE*, as well as the probable assignment of *dctM* to a function that is not obviously related to periplasmic nitrate reduction, all suggest that *dctM* is not part of the *nap* operon. Analysis of the 500 bp 3′ to *napC* did not identify any possible open reading frame in this region. It thus appears likely that the periplasmic nitrate reductase operon of *T. pantotropha* comprises the five genes *napEDABC*.

NapA shows highest sequence similarity to the NapA protein of the *A. eutrophus* periplasmic nitrate reductase ([12]; 71% identity of the precursor proteins) and to a predicted gene product of the *E. coli aeg-46.5* operon ([28]; reading frame *yojC*; we use the translation that appears in the Swissprot database as P33937; 66% identity comparing from residue 41 of *T. pantotropha* NapA onwards only). NapA also has strong sequence similarity (results not shown) to other polypeptides that bind the GMP derivative of the molybdopterin cofactor (MGD) indicating that, as biochemically demonstrated for the periplasmic nitrate reductase of *Rhodobacter sphaeroides* f. sp. *denitrificans* [29], *T. pantotropha* periplasmic nitrate reductase binds the MGD form of the cofactor. The *T. pantotropha* NapA precursor is a protein of 832 amino acids with a molecular mass of 92617 Da and a calculated pI of 5.9. Because we were unable to determine the N-terminal protein sequence of NapA we used the weight matrix method of von Heijne [24] to infer signal peptidase processing after amino acid 31. Processing at this

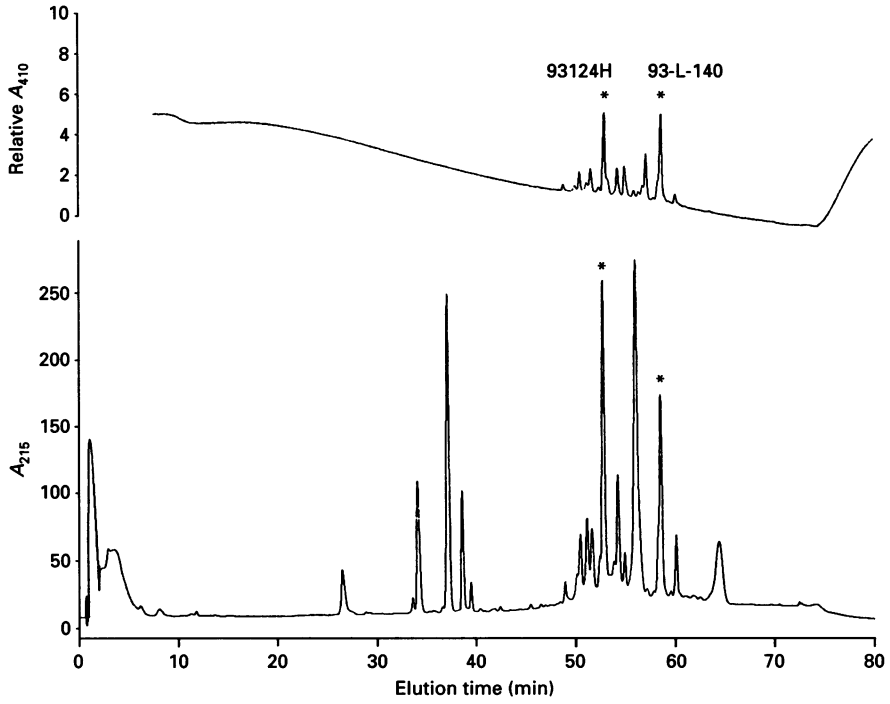

Figure 1 Reverse-phase HPLC separation of tryptic peptides of NapB by a 0–100% acetonitrile gradient in 0.1% TFA

Identification of haem-containing fractions (*) was by their absorbance at 410 nm (haem Soret band) relative to that at 215 nm (peptide bond). Absorbances are in arbitrary units.

Table 1 Sequences of tryptic peptides derived from *T. pantotropha* periplasmic nitrate reductase

(a) NapA peptide fragment	Sequence	Position in translated *napA* sequence
9211H	GQSLFDVLFR	591–600
9212H	DKGPTALGMFGSGQWTIFE	139–157
9213H	GHGHDLAPYDTYHEVR	642–657
9216H	TAEDPAATTPSTFEEFAELVSEY	317–339
9225H	IPHGVIPEQPGLHAVAQDR	451–469
9235H	DGEFTPVSWEEAFDDMA	114–130
9236H	EGLDPYVEPGAGVQFYGNPD	675–649
92-L-009	IIAVPYEPPAEPPDEEY	699–715

(b) NapB haem-containing peptide	Sequence	Position in translated *napB* sequence
93124H	-MD-HKPQFTEGSGAPMISVTHFQDR	93–118
93-L-140	YF-TA-HVQQTDVQPLVPNQFR	131–152

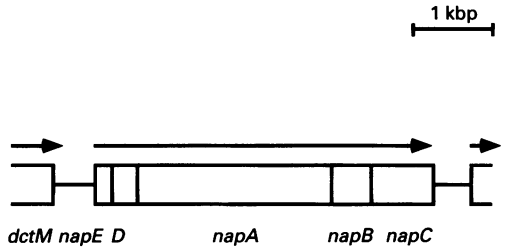

Figure 2 Schematic diagram of the *T. pantotropha nap* locus showing the area for which sequence was obtained

position corresponds closely to the N-terminus of the mature NapA protein of *A. eutrophus* (amino acid 30 of the precursor protein; [12]) and would result in a mature apoprotein of molecular mass 89496 Da with a calculated pI of 5.8. The N-terminal residue would be alanine, a residue that is often N-acetylated when found at polypeptide N-termini, possibly explaining the failure of N-terminal sequencing of the NapA subunit. The relatively long signal peptide predicted for NapA, as well as the presence near the N-terminus of a sequence with similarity to a (Ser/Thr)-Arg-Arg-Xaa-Phe-Xaa-Lys consensus, appear to be characteristic of periplasmic proteins binding MGD and some other complex redox cofactors (Figure 3; [30,31]; B. C. Berks, unpublished work).

In agreement with the peptide mapping experiments, the DNA-derived NapB sequence shows two *c*-type cytochrome consensus haem-binding motifs. NapB shows clear sequence similarity to the NapB subunit of *A. eutrophus* (45% identity of the precursor proteins) as well as the gene product of *E. coli aeg46.5* locus reading frame *yejY* (41% identity of the precursor proteins) (Figure 4) but shows no significant sequence similarity to other *c*-type cytochromes. The NapB precursor protein has a mass of 17789 Da and a calculated pI of 8.2. Although the N-terminal sequence of the mature NapB protein could not be obtained, the site of signal peptidase processing can be predicted using previously obtained MS data [15]. The molecular mass of the mature holoprotein (possessing two haems) was determined

Species	Peptide	Sequence	Length (no. of amino acids)	Evidence
T. pantotropha	NapA	MTISRRDLLKAQAAGIAAMAANIPLSSQAPA	31	VH
A. eutrophus	NapA	MKISRRDFIKQTAITATASVAGVTLPAGA	29	EX
E. coli	NapA	MKLSRRSFMKANxxxxxxxxxxxxxxxxxx?		
E. coli	FdoG	MQVSRRQFFKICAGGMAGTTAAALGFAPSVALA	33	VH
E. coli	FdnG	MDVSRRQFFKICAGGMAGTTVAALGFAPKQALA	33	VH
W. succinogenes	FdhA	MSEALSGRGNDRRKFLKMSALAGVAGVSQAVG	32	EX
E. coli	TorA	MNNNDLFQASRRRFLAQLGGLTVAGMLGPSLLTPRRATAAQA	42	EX
W. succinogenes	PsrA	METTMTRRDFLKSAGAAGAAGLVWSQTIPGTLGAL	35	EX

Figure 3 Comparison of the unusually long signal peptides of periplasmically located MGD-binding proteins

'EX' signifies that the signal-peptidase cleavage site has been determined experimentally, whereas 'VH' indicates that it was predicted using the weight-matrix method of von Heijne [24]. The sequences are those of periplasmic nitrate reductase NapA subunits of *T. pantotropha, A. eutrophus* H-16 [12] and *E. coli* (poorly sequenced in this area; the translation of reading frame *yojC* used is that found in the Swissprot database [P33937]; [28]); the catalytic subunits of the *E. coli* membrane-bound formate dehydrogenases -O (FdoG; [64]) and -N (FdnG; [34]), *Wolinella succinogenes* membrane-bound formate dehydrogenase (FdhA; [30]); *E. coli* periplasmic trimethylamine *N*-oxide reductase (TorA [50]); *W. succinogenes* polysulphide reductase catalytic subunit (PsrA; [31]).

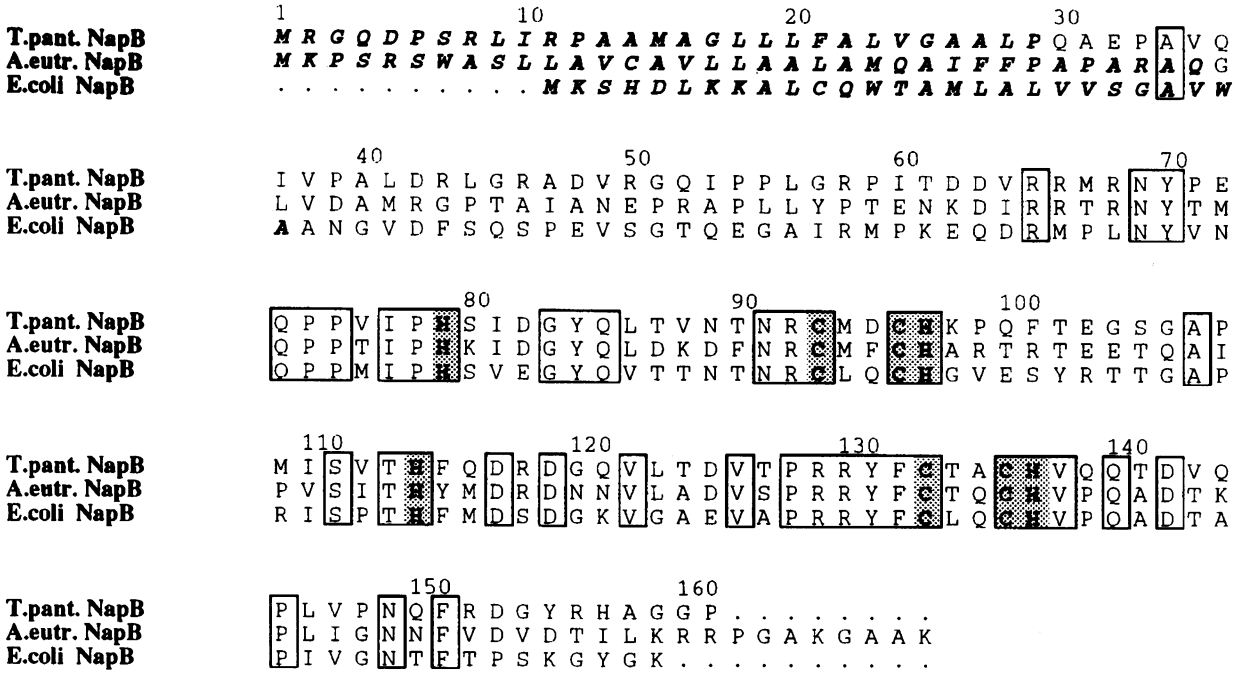

Figure 4 Alignment of the amino acid sequences of the NapB precursor proteins of *T. pantotropha, A. eutrophus* [12] and *E. coli* (product of reading frame *yejY* at the *aeg-46.5* locus [28])

A frameshift at amino acid 127 of *E. coli* NapB (base 25402 of U00008 [28]) has been corrected; also an N-terminal methionine residue, eight amino acids after that originally suggested [28], has been used as the *E. coli* NapB start site. The two *c*-type haem-binding motifs and the putative haem iron axial ligands are shaded. Amino acids that are identical in all three sequences are boxed. Signal peptides, assigned for *T. pantotropha* NapB as discussed in the text, experimentally determined for *A. eutrophus* NapB and predicted by the von Heijne method [24] for *E. coli* NapB, are printed in ***bold italics***. Note that the alanine at position 34 is predicted to be in the mature sequence of *T. pantotropha* but in the signal peptides of *A. eutrophus* and *E. coli* NapB.

to be 15991 Da and that of two samples of the chemically prepared apoprotein to be 14766 and 14795 Da. The average molecular mass of the mature apoprotein is thus 14774 Da. This value is most consistent with cleavage after amino acid 29, which would give a 14800 Da mature apoprotein. Processing at this site would result in a glutamine at the mature protein N-terminus. This would be expected to cyclize, forming pyroglutamate, which is resistant to Edman degradation, and would result in a mature apoprotein molecular mass of 14783 Da. All three NapB proteins appear to have unusually long signal peptides (Figure 4), but they do not have the conserved sequence motif found in the NapA signal peptides (Figure 3).

The putative NapC apoprotein has a molecular mass of 27236 Da, a calculated pI of 8.2 and contains four consensus *c*-

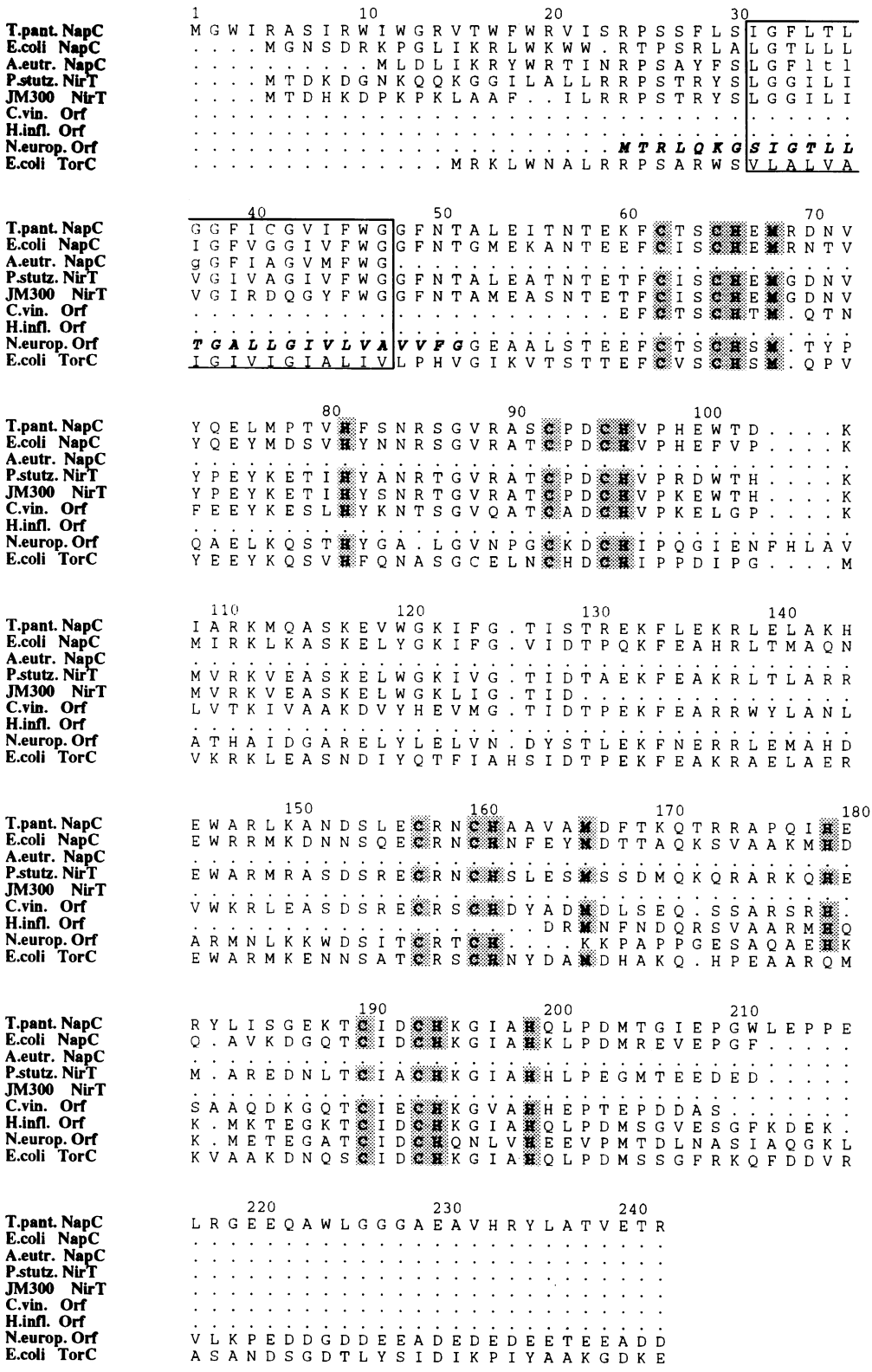

Figure 5 Amino-acid-sequence comparison of the NapC proteins of *T. pantotropha*, *E. coli (yejX* gene product; [28]) and *A. eutrophus* H-16 (partial sequence from [12] with two frameshift corrections) with structurally similar cytochromes

Shown are the sequences of the NirT proteins of *Pseudomonas stutzeri* strains ZoBell and JM300 [65,66]; product of a *Chromatium vinosum* open reading frame (partial sequence; [51]); product of a *Haemophilus influenzae* open reading frame (partial sequence; [67]); *E. coli* periplasmic trimethylamine *N*-oxide reductase operon TorC protein [50]; product of an open reading frame associated with a *Nitrosomonas europaea* hydroxylamine oxidoreductase gene cluster [68]. Note that in this Figure the *N. europaea* protein and *E. coli* TorA proteins are truncated at their C-terminii by 22 and 168 amino acids respectively. The four *c*-type haem binding motifs and possible haem iron axial ligands are shaded. Boxing delineates the proposed transmembrane helices. The proposed signal peptide of the *N. europaea* protein is printed in ***bold italics***.

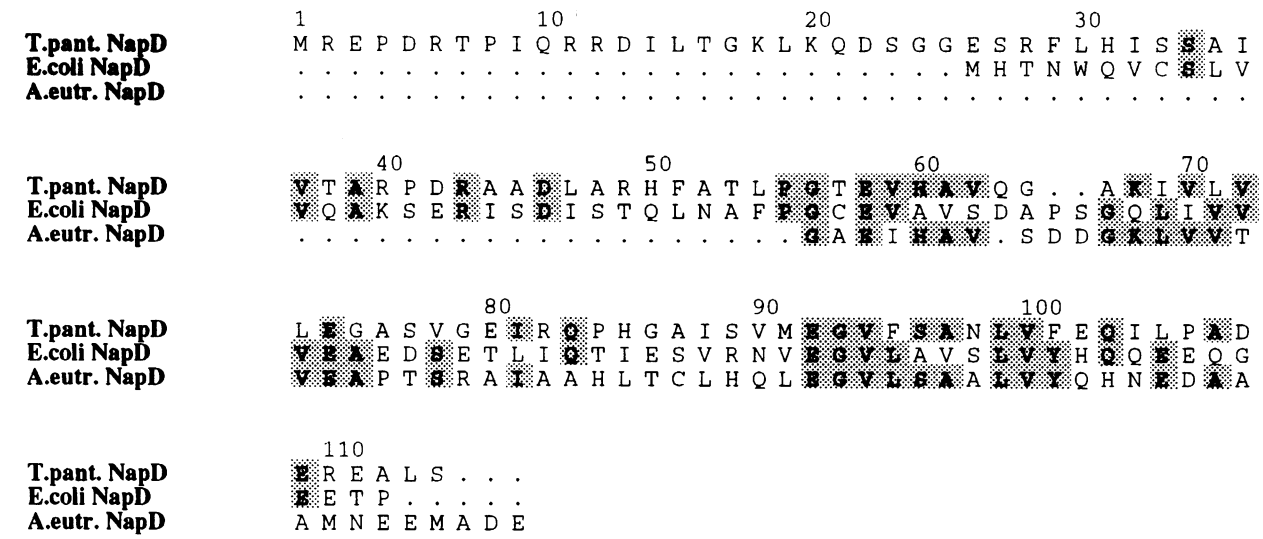

Figure 6 Amino-acid-sequence comparison of the NapD proteins of *T. pantotropha*, *E. coli* (*yojF* gene product;[28]), and *A. eutrophus* H-16 (partial C-terminal sequence from [12])

Shaded amino acids are identical in at least two of the three sequences.

Figure 7 Amino-acid-sequence comparison of *T. pantotropha* NapE and the product of a possible open reading frame divergently transcribed from the structural gene (*nirU*) of the copper nitrite reductase of *Pseudomonas* sp. strain G-179 [33]

Note that the C-terminus of the *Pseudomonas* protein is incomplete. Shaded amino acids are identical in both sequences. Boxing delineates the proposed transmembrane helix.

type cytochrome haem-binding sites. A gene (*yejX*) coding for a protein with high sequence similarity to NapC is found directly following the *napB* (*yejY*) gene at the *E. coli aeg46.5* locus (55% protein identity; Figure 5). Part of a gene coding for a NapC-like protein is also found directly following *A. eutrophus napB* (Figure 5). Hydropathy analysis suggests that NapC has a membrane-spanning helix near its N-terminus. No signal peptidase site is predicted by the von Heijne [24] method. Topological analysis, as well as the expectation that *c*-type haem groups are only found in the periplasmic compartment [32], suggests that NapC is a monotopic integral membrane protein in which the N-terminus is located in the cytoplasm and the bulk of the protein containing the *c*-type cytochrome redox centres is in the periplasm.

The putative *napD* gene product would be a 113-amino-acid protein with a molecular mass of 12123 Da and a calculated pI of 6.5. It shows sequence similarity (23% amino acid identity) to the product of *E. coli* reading frame *yojF* which immediately precedes the *E. coli napA*-like gene (Figure 6). A NapD-like protein may also be coded directly 5′ to the *napA* gene of *A. eutrophus* (Figure 6). On the basis of hydropathy analysis and lack of a predicted signal peptide it is probable that NapD is a soluble cytoplasmic protein.

The *napE* gene product would be a 60-amino-acid protein with a molecular mass of 6618 Da and a calculated pI of 9.4. NapE is predicted to be a monotopic integral membrane protein with the polypeptide N-terminus in the cytoplasmic compartment (Figure 7). NapE shows sequence similarity (39% amino acid identity) to the product of a potential open reading frame divergently transcribed from the structural gene (*nirU*) of the copper nitrite reductase of *Pseudomonas* sp. strain G-179 ([33]; Figure 7). The weak hybridization of the *Pseudomonas* copper nitrite reductase gene probe to *P. denitrificans* DNA [33], an organism that does not possess a copper nitrite reductase, may be a consequence of the presence of this *napE*-like gene fragment on the probe DNA.

NapA analysis

Analysis (results not shown) of aligned MGD-binding protein sequences, including *T. pantotropha* and *E. coli* NapA, reveals two polypeptide regions that are highly conserved between enzymes of similar substrate specificity, but which are much less conserved between enzymes of differing substrate specificities. These regions of the polypeptide are likely to be near the substrate binding/active site of the enzymes.

The first region (Figure 8) has previously been suggested to be involved in substrate specificity by other workers [12,30,34]. We extend these analyses by suggesting that at least two distinct specificity classes can be distinguished in this region (Figure 8).

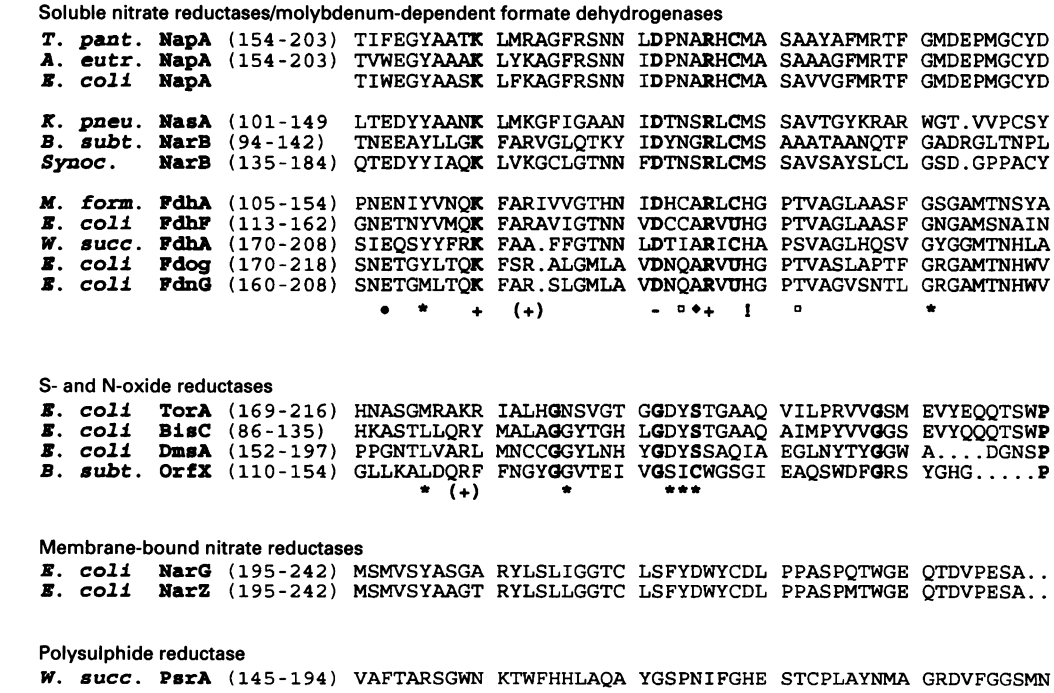

Figure 8 Comparison of the proposed specificity-determining regions of MGD-binding proteins

The sequences are those of the proteins from Figure 3 and additionally the assimilatory nitrate reductases of *Klebsiella pneumoniae* (NasA; [70]), *Bacillus subtilis* (NarB; [71]) and *Synechococcus* PCC7942 (NarB; [72]); *Methylobacterium formicicum* F_{420}-dependent formate dehydrogenase catalytic subunit (FdhA; [73]); *E. coli* formate hydrogenlyase system formate dehydrogenase (FdhF; [74]); *E. coli* D-biotin sulphoxide reductase (BisC; [26]); *E. coli* membrane-bound dimethyl sulphoxide reductase catalytic subunit (DmsA; [75]); product of a *Bacillus subtilis* open reading frame [76]; catalytic subunits of the *E. coli* major (NarG; [77,78]) and minor (NarZ; [79]) membrane-bound nitrate reductases. Residues in **bold** are identical. '*' indicates residues that are conserved in all but one sequence. Strict conservation of amino acids with charged side chains is indicated by '(+)' (Arg/Lys) and ' − ' (Asp/Glu). For the putative 'formate site', '!' indicates a residue that is conserved in the formate dehydrogenases, but is a different, but conserved, amino acid in the soluble nitrate reductases; ●, an amino acid conserved in the formate dehydrogenases and assimilatory nitrate reductases but not the periplasmic nitrate reductases; □, an amino acid conserved amongst the soluble nitrate reductases but not found in the formate dehydrogenases; ◆, an amino acid conserved between the formate dehydrogenases and periplamic nitrate reductases. A '*' in the putative 'S-oxide site' indicates the position in our alignment corresponding to the molybdenum serine ligand proposed for two of the enzymes by Gladyshev and co-workers [37].

One comprises the molybdenum-dependent formate dehydrogenases and soluble nitrate reductases (which contain a putative 'formate site'), the other the S- and N-oxide reductases (which contain a putative 'S-oxide site'). The sequence of MGD-binding proteins in this region may thus be used as a fingerprint for substrate class. Support for the idea that this polypeptide region forms part of the active site comes from the observation that the 'conserved' cysteine in the 'formate site' is replaced in the *E. coli* formate dehydrogenases by a translationally inserted selenocysteine (Figure 8; [35]). The elaborate machinery required to insert the selenocysteine and the deleterious effects of a selenocysteine to cysteine change [36] argue for the importance of this residue in catalysis, while recent EPR evidence strongly suggests that the selenocysteine is a ligand of the cofactor molybdenum atom [37]. It is likely that the equivalent cysteine in the soluble nitrate reductases (cysteine-181 of the *T. pantotropha* NapA precursor protein) is also a molybdenum ligand and that the functional rationale for this ligation will be similar in all enzymes with the 'formate site'. Consistent with the absence of a conserved cysteine in the 'S-oxide site' proposed here, variable-temperature magnetic CD and resonance Raman spectroscopies have provided no evidence for molybdenum cysteine co-ordination in the MGD-dependent dimethyl-sulphoxide reductase of *R. capsulatus* [38–40]. In addition we have argued from Mo(V) EPR data that the molybdenum atom of *T. pantotropha* periplasmic nitrate reductase has an additional -SR ligand relative to that of *R. capsulatus* dimethyl-sulphoxide reductase [5,41]. Very recently Gladyshev and co-workers [37] have made the reasonable suggestion that, in certain S-oxide reductases, the molybdenum atom may have an O- rather than an S- or Se- ligand. The amino acid that these workers speculate provides the O-ligation corresponds to a serine or cysteine in our proposed 'S-oxide' site (Figure 8). The similarity of the proposed active-site region of soluble nitrate reductase and formate dehydrogenases may be a reflection of the similar size, shape and charge of the two substrates, which are both approximately trigonal planar univalent anions. One function of the conserved basic residues in the 'formate site' might, therefore, be in the formation of salt bridges to the anionic substrate. In contrast, the 'S-oxide site', which has to bind uncharged substrates, has only a single conserved charged residue. Further, the 'S-oxide site' has a number of conserved glycine residues that might permit the entry of bulkier substrate molecules into the active site.

Support for the idea that the formate dehydrogenases and soluble nitrate reductases have related substrate binding sites comes from observations that nitrate is both a competitive inhibitor and stabilizer of molybdenum-dependent formate dehydrogenases [42–46]. A difference in the structure of the catalytic site of the soluble and membrane-bound nitrate reductases, inferred here from sequence comparisons (Figure 8), is supported by Mo(V) EPR studies of the three enzymes [5]. These considerations imply that the catalytic mechanism of the

```
T. pant. NapA (40-89)   TWSKAP C.RFCGTGCG VMVGVKE... ....GRVVAT HGDLLAEVNR GLNCVKGYFL
A. eutr. NapA (40-89)   KWSKAP C.RFCGTGCG VTVAVKD... ....NKVVAT QGDPQAEVNK GLNCVKGYFL
E. coli  NapA           KWDKAP C.RFCGTGCG VLVGTQQ... ....GRVVAC QGDPDAPVNR GLNCIKGYFL
K. pneu. NasA (2-49)    TETRTT C.PYCGVGCG VIA..... SRAPHGQVSV RGDEQHPANF GRLCVKGAAL
B. subt. NarB (-14-35)  KTYDTQ C.PFCSMQCK MQLV......EQTIVTRKKYT AIGIDNPTTQ GRLCIKGMNA
Synoc.   NarB (17-75)   DTAKTL C.PYCGVGCG LEAV....(17).....IWQI RGDRQHPSSQ GMVCVKGATT
M. form. FdhA (4-51)    KYVPTI C.PYCGVGCG MNLVVKD... ....EKVVGV EPWKRHPVNE GKLCPKGNFC
E. coli  FdhF (2-48)    KKVVTV C.PYCASGCK INLVVDN... ....GKIVRA EA.AQGKTNQ GTLCLKGYYG
W. succ. FdhA (55-102)  KKVKTI C.TYCSVGCG IIAEVVD... ....GVWVRQ EVAQDHPISQ GGHCCKGADM
E. coli  FdoG (44-98)   RETRNT C.TYCSVGCG LLMYSLGDGA KNAKASIFHI EGDPDHPVNR GALCPKGAGL
E. coli  FdnG (44-98)   KEIRNT C.TYCSVGCG LLMYSLGDGA KNAREAIYHI EGDPDHPVSR GALCPKGPGL
E. coli  NarG (44-99)   KIVRST HGVNCTGSCS WKIYVKNGLV TWETQQTDYP RTRPDLPNHE PRGCPRGASY
E. coli  NarZ (44-99)   KIVRST HGVNCTGSCS WKIYVKNGLV TWEIQQTDYP RTRPDLPNHE PRGCPRGASY
E. coli  DmsA (28-81)   KVIWSA CTVNCGSRCP LRMHVVDG.. EIKYVETDNT GDDNYDGLHQ VRACLRGRSM
B. subt. OrfX (3-51)    KVHQSA CPLNCWDSCG FLVTVDDG.. .....KVTKV DGDPNHPITE GKICGRGRML
W. succ. PsrA (45-92)   KFVPSI C.EMCTSSCT IEARVEG... ....DKGVFI RGNPKDKSRG GKVCARGGSG
                           ↑  ↑  ↑                                          ↑
```

Figure 9 Alignment of the N-terminal regions of MGD-binding polypeptides showing a conserved [4Fe-4S]$^{2+,1+}$-binding motif

'↑' indicates amino acids proposed to bind an [Fe-S] cluster. The sequence references are given in the legends to Figures 3 and 8. Note that we have inferred an N-terminal extension of the *B. subtilis* NarB sequence relative to that presented in [71].

soluble nitrate reductases is distinct from that of the membrane-bound nitrate reductase, but may have unexpected similarity to the mechanism by which formate is hydroxylated to bicarbonate by the molybdenum-dependent formate dehydrogenases. The sequence relationships further suggest that the soluble nitrate reductases have a close evolutionary relationship with the formate dehydrogenases, but not with the membrane-bound nitrate reductases.

The second proposed specificity-linked region is found around amino acids 406 to 431 (*T. pantotropha* NapA preprotein numbering). This region appears less useful than the first as a fingerprint region, because there is only strong sequence conservation between members of the formate dehydrogenase/soluble nitrate reductase specificity group. This second region was previously suggested to be conserved among formate dehydrogenases [30].

A four-cysteine motif that is found near the N-terminus of most MGD-binding polypeptides, and which has been suggested to bind a [4Fe-4S] cluster in these proteins ([4] and references cited therein), is also present in *T. pantotropha* NapA (Figure 9). *T. pantotropha* periplasmic nitrate reductase is the first MGD-dependent enzyme in which a [4Fe-4S] cluster identified by EPR [4] can be unambiguously assigned to this conserved cysteine motif. It should be emphasized that no other possible [Fe-S]-cluster-binding motifs are present in NapA or NapB. Recently Trieber and co-workers [47] have carried out site-directed mutagenesis of the conserved cysteine residues in the MGD-binding subunit of the membrane-bound dimethyl-sulphoxide reductase of *E. coli*. In the majority of cases the mutant phenotypes suggested a rôle for electron transfer to the molybdenum cofactor from redox groups in other enzyme subunits and excluded a predominantly structural rôle. Trieber and co-workers [47] did not suggest that the conserved cysteine residues bound an [Fe-S] cluster, but the reported mutant phenotypes would be consistent with such a function.

NapB analysis

The axial ligands of the haem iron atoms of *c*-type cytochromes are usually the histidine from the Cys-Xaa$_2$-Cys-His haem-binding motif and a distal methionine or second histidine. Outside the two haem-binding motifs the NapB sequences contain two conserved histidine residues (histidine-78 and -113; *T. pantotropha* precursor protein numbering), but no conserved methionine residues (Figure 4). This suggests that both the *c*-haems are bis-histidine ligated. The EPR spectra of the *T. pantotropha* NapB ferricytochromes have $g_z = 2.92$, $g_y = 2.26$ and $g_x = 1.5$ [4]. These values are consistent with bis-histidine ligation of the haems [48]. The reduction potentials (at pH 7.0) of the NapB haems are -15 mV and $+80$ mV [15]. The $+80$ mV potential is the highest that we are aware of for a *c*-type cytochrome with probable bis-histidine ligation, but is quite consistent with calculated and experimentally determined potentials of bis-histidine-ligated *b*-cytochromes (containing non-covalently bound haem; [49]) and model compound studies of bis-histidine versus histidine-methionine-co-ordinated iron porphyrins [48].

The NapB polypeptide is highly unusual in that it does not bind the dye Coomassie Brilliant Blue and cannot be detected by electrospray MS [3]. Both Coomassie Blue staining and normal-mode electrospray MS of proteins is largely dependent on the protein being positively charged. As 14 of the 132 amino acids in mature NapB are basic (Figure 4), the basis of the unusual staining and electrospray behaviour of the protein would not, therefore, appear to be a lack of potentially positively charged residues.

NapC analysis

NapC shows strong sequence similarity to a family of multi-haem *c*-type cytochromes, the first member to be described being the NirT protein of *Pseudomonas stutzeri* ZoBell (Figure 5). Like NapC, these proteins are likely to have a periplasmic domain(s) containing four *c*-type haem groups (exceptionally *E. coli* TorC has an additional polypeptide region binding a fifth *c*-type haem) and an N-terminal membrane-spanning region (exceptionally the *N. europaea* protein may have a signal peptide). Experimental support for the proposed membrane localization has been presented for *E. coli* TorC [50]. Examination of the NirT family alignments shown in Figure 5 allows identification of five residues that are candidates for the four distal haem iron ligands: absolutely conserved methionine-68, histidine-81 and histidine-194 (*T. pantotropha* NapC numbering), as well as methionine-161 and histidine-174, which are found in all but one of the sequences. That only three residues are absolutely conserved indicates that at least two haem ligation patterns are found in the NirT family of cytochromes. Dolata and co-workers [51] have argued for methionine-68, histidine-81, methionine-161 and

histidine-194 as haem iron axial ligands for the *Chromatium vinosum* NirT family cytochrome (Figure 5) because of what they claim is an obvious duplication of the two halves of the molecule. While the relatively symmetrical model of the cytochrome structure they suggest is appealing, we see no strong evidence from our sequence analysis for such a duplication. Methionine-68 is probably too close to haem-binding-site motif 1 to be a possible ligand for the haem bound at that site.

We suggest that NapC functions in electron transfer between the ubiquinol pool and the periplasmic nitrate reductase and that NapC is the direct electron donor to the NapAB complex. Whether NapC is itself a quinol oxidase is less easy to assess. The proposed membrane-spanning helix shows a degree of sequence conservation within the NirT family (Figure 5) that would not be expected of a simple membrane-anchoring polypeptide, suggesting that the putative transmembrane helix of NapC either makes specific interactions with another protein and/or participates in the formation of a catalytic (quinol oxidase) site. The quinol oxidase associated with the periplasmic nitrate reductase pathway is believed to oxidize quinol at the periplasmic side of the cytoplasmic membrane [2] and so is not required to translocate charge across the membrane bilayer. Therefore, despite having only one transmembrane helix, it is conceivable that NapC could be a quinol oxidase. Alternatively, NapC may receive its electrons from a separate quinol oxidase or other intermediary electron-transport protein.

We have not so far been able to detect NapC biochemically in either wild-type *T. pantotropha* or Nap-overexpressing mutant strain M-6 [2] cell fractions. Membrane cytochrome(s) with α-band maxima in the reduced-minus-oxidized difference spectra at 559 nm have been implicated in the transfer of electrons between the ubiquinol pool and the periplasmic nitrate reductase in *R. capsulatus* N22DNAR$^+$ [10]. The position of the α-band close to 560 nm was taken to indicate that a *b*-type cytochrome (α-band of the ferrocytochrome usually around 560 nm; [52]) rather than a *c*-type cytochrome (α-band of the ferrocytochrome usually around 550 nm; [52]) was involved. Electronic interactions between haem groups, as could occur in tetrahaem NapC, can cause substantial shifts in the position of the α-band maximum. It is therefore possible that the cytochrome detected spectroscopically in *R. capsulatus* might be NapC.

In *A. eutrophus* a gene coding for NapC is apparently present 3′ to NapB (Figure 5). It has been reported that transposon insertions in this area do not affect periplasmic nitrate reductase activity assayed with the non-physiological electron donor Benzyl Viologen [12]. As Viologens are direct electron donors to the soluble NapAB complex (e.g. [3]) and as physiological electron transport to the periplasmic nitrate reductase was not measured, our suggestion of a role for NapC in electron transfer to NapAB is not contradicted by these experiments.

NapD and NapE analysis

The function of NapD is unknown. The overlap of *napD* and *napA* in all three *nap* operons is suggestive of translational coupling and stoichiometric synthesis of NapD with NapA. NapD is predicted to have a cytoplasmic location and lacks conserved residues likely to ligand redox cofactors (Figure 6). Given these observations, a rôle for NapD in formation of the mature periplasmic NapAB enzyme is a possibility.

The function of NapE is also unknown. The relatively high degree of sequence conservation in the proposed transmembrane region between NapE and a possible *Pseudomonas* protein (Figure 7) suggests that these proteins interact with another integral membrane protein. NapE may be involved in mediating interactions between NapC and a quinol oxidase. In support of this idea, other proteins that are essentially single transmembrane helices have been found in apparent association with integral membrane redox complexes e.g. *Sulfolobus acidocaldarius* quinol oxidase SoxD subunit [53], the product of *orf4* of the *Hmc* operon of *Desulfovibrio vulgaris* [54], several subunits of bovine NADH:ubiquinone oxidoreductase [55] and subunit 9 of *Saccharomyces cerevisiae* cytochrome bc_1 complex [56].

The *aeg-46.5* locus of *E. coli*

The *aeg-46.5* locus of *E. coli* is anaerobically inducible by nitrate or nitrite [57–59]. It has recently been sequenced as part of the *E. coli* genome project [28]. One of the gene products (YojC) is an MGD-binding protein containing the periplasmic nitrate reductase specificity-determining region (Figure 8) and a possible signal peptide. The locus also codes for NapB- (YejY) and NapD-like (YojF) proteins, polypeptides unique to periplasmic nitrate reductase operons (Figures 4 and 6), and a NapC-like protein (YejX; Figure 5). These sequence similarities, together with the positive regulation of transcription by nitrate, enable us to identify the *aeg46.5* locus as coding for a periplasmic nitrate reductase system. The 16 and 24 kDa *c*-type cytochromes synthesized by *E. coli* grown anaerobically with nitrite as electron acceptor [60] may correspond to NapB (YejY) and NapC (YejX). Other biochemical evidence for a periplasmic nitrate reductase system in *E. coli* is lacking.

The *E. coli napDABC* gene order is the same as in *T. pantotropha* and *A. eutrophus*, although three other genes coding for iron–sulphur proteins are also present at the *E. coli* locus i.e. the gene order is *yojG/napD/napA/yojA/yejZ/napB/napC*, whereas *napE* is not found. It is possible that the putative [Fe-S]-cluster-binding proteins are a quinol oxidase system. YojA and the predicted integral membrane protein YejZ have high sequence similarity to proteins (MauM and MauN) of unknown function that are found in the methylamine dehydrogenase gene clusters of several methylotrophic organisms [61–63]. Several clusters of genes that are probably involved in periplasmic attachment of haem to apo *c*-type cytochromes follow, and are transcribed in the same orientation as, the *aeg-46.5 nap* genes.

This work was supported through a grant from the Agricultural and Food Research Council/Science and Engineering Research Council (GR/H/12225) to S.J.F. and D.J.R. We thank Dr. Deepa Nath and Dr. Dudley Page for many helpful suggestions, and Dr. Akber Siddiqui, Dr. David Kelly and Professor Bärbel Friedrich for making data available to us prior to publication.

REFERENCES

1 Bell, L. C., Richardson, D. J., and Ferguson, S. J. (1990) FEBS Lett. **265**, 85–87
2 Bell, L. C., Page, M. D., Berks, B. C., Richardson, D. J. and Ferguson, S. J. (1993) J. Gen. Microbiol. **139**, 3205–3214
3 Berks, B. C., Richardson, D. J., Robinson, C., Reilly, A., Aplin, R. T. and Ferguson, S. J. (1994) Eur. J. Biochem. **220**, 117–124
4 Breton, J., Berks, B. C., Reilly, A., Thomson, A. J., Ferguson, S. J. and Richardson, D. J. (1994) FEBS Lett. **345**, 76–80
5 Bennett, B., Berks, B. C., Ferguson, S. J., Thomson, A. J. and Richardson, D. J. (1994) Eur. J. Biochem. **226**, 189–198
6 Hille, R. (1994) Biochim. Biophys. Acta **1184**, 143–169
7 Richardson, D. J. and Ferguson, S. J. (1992) Arch. Microbiol. **157**, 535–537
8 Alef, K. and Klemme, J.-H. (1979) Z. Naturforsch. **34c**, 33–37
9 McEwan, A. G., Wetzstein, H. G., Meyer, O., Jackson, J. B. and Ferguson, S. J. (1987) Arch. Microbiol. **147**, 340–345
10 Richardson, D. J., McEwan, A. G., Page, M. L. D., Jackson, J. B. and Ferguson, S. J. (1990) Eur. J. Biochem. **194**, 263–270
11 Satoh, T. (1981) Plant Cell Physiol. **22**, 443–452
12 Siddiqui, R. A., Warnecke-Eberz, U., Hengsberger, A., Schneider, B., Kostka, S. and Friedrich, B. (1993) J. Bacteriol. **175**, 5867–5876

13 Bursakov, S., Liu, M.-Y., Payne, W. J., LeGall, J., Moura, I. and Moura, J. J. G. (1995) Anaerobe **1**, 55–60
14 Ludwig, W., Mittenhuber, G. and Friedrich, C. G. (1993) Int. J. System. Bacteriol. **43**, 363–367
15 Samyn, B., Berks, B. C., Page, M. D., Ferguson, S. J. and van Beeumen, J. J. (1994) Eur. J. Biochem. **219**, 585–594
16 Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
17 Aebersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E. and Kent, S. B. H. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 6970–6974
18 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecuar Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19 Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
19a Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) J. Mol. Biol. **215**, 403–410
19b Borodovsky, M. and McInininch, J. (1993) Comput. Chem. **17**, 123–133
20 Persson, B. and Argos, P. (1994) J. Mol. Biol. **237**, 182–192
21 Jones, D. T., Taylor, W. R. and Thornton, J. M. (1994) Biochemistry **33**, 3038–3049
22 Engelman, D. M., Steitz, T. A. and Goldman, A. (1986) Annu. Rev. Biophys. Biophys. Chem. **15**, 321–353
23 von Heijne, G. (1992) J. Mol. Biol. **225**, 487–494
24 von Heijne, G. (1987) Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit ? Academic Press, London and New York
25 Kelly, D. J. (1993) Accession number S19739, PIR Database
26 Sofia, H. J., Burland, V., Daniels, D. L., Plunkett, G., III and Blattner, F. R. (1994) Nucleic Acids Res. **22**, 2576–2586
27 Williams, R. J. L. (1993) Accession number X64876, EMBL Nucleotide Sequence Database, Heidelberg
28 Richterich, P., Lakey, N., Gryan, G. et al. (1993) Accession number U00008, EMBL Nucleotide Sequence Database, Heidelberg
29 Rajagopalan, K. V. and Johnson, J. L. (1992) J. Biol. Chem. **267**, 10199–10202
30 Bokranz, M., Gutmann, M., Körtner, C. et al. (1991) Arch. Microbiol. **156**, 119–128
31 Krafft, T., Bokranz, M., Klimmek, O. et al. (1992) Eur. J. Biochem. **206**, 503–510
32 Wood, P. (1983) FEBS Lett. **164**, 223–226
33 Ye, R. W., Fries, M. R., Bezborodnikov, S. G., Averill, B. A. and Tiedje, J. M. (1993) Appl. Environ. Microbiol. **59**, 250–254
34 Berg, B. L., Li, J., Heider, J. and Stewart, V. (1991) J. Biol. Chem. **266**, 22380–22385
35 Heider, J. and Böck, A. (1993) Adv. Microbiol. Physiol. **35**, 71–109
36 Axley, M. J., Böck, A. and Stadtman, T. C. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 8450–8454
37 Gladyshev, V. N., Khangulov, S. V., Axley, M. J. and Stadtman, T. C. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 7708–7711
38 Benson, N., Farrar, J. A., McEwan, A. G. and Thomson, A. J. (1992) FEBS Lett. **307**, 169–172
39 Finnegan, M. G., Hilton, J., Rajagopalan, K. V. and Johnson, M. K. (1993) Inorg. Chem. **32**, 2616–2617
40 Kilpatrick, L., Rajagopalan, K. V., Hilton, J. et al. (1995) Biochemistry **34**, 3032–3039
41 Bennett, B., Benson, N., McEwan, A. G. and Bray, R. C. (1994) Eur. J. Biochem. **225**, 321–331
42 Axley, M. J., Grahame, D. A. and Stadtman, T. C. (1990) J. Biol. Chem. **265**, 18213–18218
43 Friedebold, J. and Bowien, B. (1993) J. Bacteriol. **175**, 4719–4728
44 Karzanov, V. V., Bogatsky, Y. A., Tischkov, V. I. and Egorov, A. M. (1989) FEMS Microbiol. Lett. **60**, 197–200
45 Kröger,A., Winkler, E., Innerhofer, A., Hackenberg, H. and Schägger, H. (1979) Eur. J. Biochem. **94**, 465–475
46 Schauer, N. L., and Ferry, J. G. (1986) J. Bacteriol. **165**, 405–411
47 Trieber, C. A., Rothery, R. A. and Weiner, J. H. (1994) J. Biol. Chem. **269**, 7103–7109
48 Moore, G. R. and Pettigrew, G. W. (1990) Cytochromes *c*: Evolutionary, Structural and Physicochemical Aspects, Springer-Verlag, Berlin
49 Krishtalik, L. I., Tae, G.-S., Cherepanov, D. A. and Cramer, W. A. (1993) Biophys. J. **65**, 184–195
50 Méjèan, V., Iobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M. and Pascal, M.-C. (1994) Mol. Microbiol. **11**, 1169–1179
51 Dolata, M. M., van Beeumen, J. J., Ambler, R. P., Meyer, T. E. and Cusanovich, M. A. (1993) J. Biol. Chem. **268**, 14426–14431
52 Barsch, R. G. (1971) Methods Enzymol. **23**, 344–363
53 Lübben, M., Warne, A., Albracht, S. P. J. and Saraste, M. (1994) Mol. Microbiol. **13**, 327–335
54 Rossi, M., Pollock, W. B. R., Reij, M. W., Keon, R. G., Fu, R. and Voordouw, G. (1993) J. Bacteriol. **175**, 4699–4711
55 Walker, J.E, Arizmendi, J. M., Dupuis et al. (1992) J. Mol. Biol. **226**, 051–1072
56 Philips, J. D., Graham, L. A. and Trumpower, B. L. (1993) J. Biol. Chem. **268**, 11727–11736
57 Choe, M. and Reznikoff, W. S. (1991) J. Bacteriol. **173**, 6139–6146
58 Choe, M. and Reznikoff, W. S. (1993) J. Bacteriol. **175**, 1165–1172
59 Rabin, R. S. and Stewart, V. (1993) J. Bacteriol. **175**, 3259–3268
60 Iobbi-Nivol, C., Crooke, H., Griffiths, L. et al. (1994) FEMS Microbiol. Lett. **119**, 89–94
61 Christoserdov, A. Y., Christoserdova, L. V., McIntire, W. S. and Lidstrom, M. E. (1994) J. Bacteriol. **176**, 4052–4065
62 Christoserdov, A. Y., McIntire, W. S., Mathews, F. S. and Lidstrom, M. E. (1994) J. Bacteriol. **176**, 4073–4080
63 van der Palen, C. J. N. M. (1994) Accession number U15028, EMBL Nucleotide Sequence Database, Heidelberg
64 Blattner, F. R., Burland,V. D., Plunkett G., III, Sofia, H. J. and Daniels, D. L. (1993) Nucleic Acids Res. **21**, 5408–5417
65 Jüngst, A., Wakabayashi, S., Matsubara, H. and Zumft, W. G. (1991) FEBS Lett. **279**, 205–209
66 Smith, G. B. and Tiedje, J. M. (1992) Appl. Env. Microbiol. **58**, 376–384
67 Maskell, D. J., Szabo, M. J., Butler, P. D., Williams, A. E. and Moxon, E. R. (1991) Mol. Microbiol. **5**, 1013–1022
68 Bergmann, D. J., Arciero, D. M. and Hooper, A. B. (1994) J. Bacteriol. **176**, 3148–3153
69 Reference deleted
70 Lin, J. T., Goldman, B. S. and Stewart, V. 1993) J. Bacteriol. **175**, 2370–2378
71 Ogawa, K.-I., Nakano, M. M., Zubar, P. and Yamane, K. (1995) Accession number D30689, EMBL Nucleotide Sequence Database, Heidelberg
72 Andriesse, A. J. and Bakker, H. (1993) Accession number X74597, EMBL Nucleotide Sequence Database, Heidelberg
73 Shuber, A. P., Orr, E. C., Recny, M. A. et al. (1986) J. Biol. Chem. **261**, 12942–12947
74 Zinoni, F., Birkmann, A., Stadtman, T. C. and Böck, A. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 4650–4654
75 Bilous, P. T., Cole, S. T., Anderson, W. F. and Weiner, J. H. (1988) Mol. Microbiol. **2**, 785–795
76 Ogasawara, N., Nakai, S. and Yoshikawa, H. (1994) DNA Res. **1**, 1–14
77 Blasco, F., Iobbi, C., Giordano, G., Chippaux, M. and Bonnefoy, V. (1989) Mol. Gen. Genet. **218**, 249–256
78 Wootton, J. C., Nicolson, R. E., Cock, J. M. et al. (1991) Biochim. Biophys. Acta **1057**, 157–185
79 Blasco, F., Iobbi, C., Ratouchniak, J., Bonnefoy, V. and Chippaux, M. (1990) Mol. Gen. Genet. **222**, 104–111

Received 3 February 1995/6 April 1995; accepted 7 April 1995