

## Short sequence-paper

Cloning and sequence analysis of the dimethylsulfoxide reductase structural gene from *Rhodobacter capsulatus*<sup>1</sup>Anthony L. Shaw<sup>a</sup>, Graeme R. Hanson<sup>b</sup>, Alastair G. McEwan<sup>a, \*</sup><sup>a</sup> Department of Microbiology, The University of Queensland, Brisbane 4072, Australia<sup>b</sup> Centre for Magnetic Resonance, The University of Queensland, Brisbane 4072, Australia

Received 6 May 1996; revised 2 July 1996; accepted 17 July 1996

---

Abstract

---

The dimethylsulfoxide reductase structural gene (*dorA*) of *Rhodobacter capsulatus* was cloned from a  $\lambda$  expression library. The nucleotide sequence of the *dorA* gene was determined and it was found to encode a protein of 825 amino acids. Comparison of the deduced amino-acid sequence of DorA with N-terminal sequence of purified dimethylsulfoxide reductase from *Rhodobacter capsulatus* showed that the pre-protein possesses a 41-amino-acid N-terminal signal polypeptide. All of the conserved segments which have been described in bacterial enzymes which bind molybdopterin guanine dinucleotide (Berks, B.C., Ferguson, S.J., Moir, J.W.B. and Richardson, D.J. (1995) Biochim. Biophys. Acta 1232, 97–173) were identified in *Rhodobacter capsulatus* dimethylsulfoxide reductase.

**Keywords:** Anaerobic respiration; Dimethylsulfoxide; Oxidoreductase; Molybdenum oxotransferase; Primary structure

---

Many facultative anaerobic bacteria are able to use dimethylsulfoxide (DMSO) and trimethylamine-*N*-oxide (TMAO) as electron acceptors in anaerobic respiration [1,2] and a number of terminal reductases which catalyse DMSO and/or TMAO reduction have been described. In *Escherichia coli* a periplasmic TMAO reductase and a membrane-associated DMSO reductase have been characterised [2] and the operons encoding these enzymes have been sequenced [3,4]. Purple phototrophic bacteria, such as *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*, possess a periplasmic DMSO reductase which has been purified and extensively characterised (reviewed in Ref. [5]). The unifying feature of DMSO/TMAO reductases is that they all contain a pterin molybdenum cofactor at their active site [2,6]. These enzymes are members of the molybdenum oxotransferase family and they catalyse the transfer of an oxo group to or from a substrate in a

reaction which is coupled to a two electron transfer process [6,7]. It has been shown, using *R. sphaeroides* DMSO reductase, that the oxo group from DMSO is found in the product, water [8] and this consistent with previous observations of oxo group transfer in xanthine dehydrogenase [9]. The discovery of dinucleotide forms of the pterin molybdenum cofactor (reviewed in Ref. [7]) and the recent publication of crystal structures of the aldehyde:ferredoxin oxidoreductase and xanthine oxidase-related aldehyde oxidoreductase has revealed an unexpected diversity in the structure of pterin molybdenum cofactors [10,11]. The DMSO reductase from *Rhodobacter* is an attractive model enzyme for investigating the molecular properties of molybdenum oxotransferases since it contains molybdopterin guanine dinucleotide (MGD) as its only prosthetic group [12,13]. Recently, the DNA sequence of the DMSO reductase structural gene from *R. sphaeroides* was reported [14,15] and in this paper we report the cloning and DNA sequence analysis of the DMSO reductase structural gene from *R. capsulatus*.

Genomic DNA from *R. capsulatus* strain 37b4 [16] was isolated as described in [17] and used to construct a  $\lambda$  expression library. DNA fragments in the range 4–9 kb, produced by partial digestion of chromosomal DNA with *Sau3AI*, were purified from an agarose gel and ligated into

---

Abbreviations: MGD, molybdopterin guanine dinucleotide; DMSO, dimethylsulfoxide; TMAO, trimethylamine-*N*-oxide.

\* Corresponding author. Fax: +61 7 33654620; e-mail: mcewan@bio-sci.uq.oz.au.

<sup>1</sup> The nucleotide sequence data presented in this paper have been deposited in the EMBL/GenBank Data Libraries under accession number: U49506.

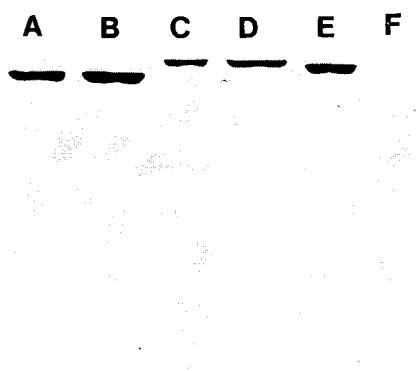


Fig. 1. Western blot of proteins expressed from pALS1 and pALS2. Crude cell extracts of *E. coli* XL1-Blue MRF' (Stratagene) were separated by SDS-PAGE and blotted onto Hybond N<sup>+</sup> membrane then probed with an anti DMSO reductase antibody (*R. capsulatus*) as in Ref. [18]. Lanes A and B contain extracts from cells containing pALS2, lanes C and D contain extracts from cells containing pALS1. Lane E contains purified *R. capsulatus* DMSO reductase, lane F contains an extract from cells containing a pBK-CMV phagemid with no insert.

the *Bam*HI-digested and dephosphorylated arms of  $\lambda$ ZAP-Express (Stratagene). The recombinant library was propagated according to the manufacturer's instructions on *E. coli* XL1-Blue MRF' grown on LB medium. Plaques were transferred to Hybond N<sup>+</sup> filters (Amersham) according to Stratagene protocols. Filters were probed with an anti-DMSO reductase (*R. capsulatus*) antibody [18] and developed using an alkaline phosphatase-conjugated secondary antibody and chromogenic substrates as described by Stratagene. Screening of the  $\lambda$  library with the anti-DMSO reductase antibody identified two putative immunopositive clones. The two clones were excised as derivatives of pBK-CMV phagemids using Stratagene protocols. The phagemids, pALS1 and pALS2, harboured in *E. coli*, were induced with IPTG. Fig. 1 shows that the products of IPTG-induced expression from both phagemids reacted with the anti-DMSO reductase antibody (lanes A–D). A comparison of the expressed proteins with purified DMSO reductase from *R. capsulatus* (lane E) shows that the product of pALS1 was slightly larger while the polypeptide produced from pALS2 was of a slightly lower molecular mass. Although these data provide strong evidence that both clones contained the *R. capsulatus* DMSO reductase structural gene (*dorA*) no enzyme activity was detected in periplasmic or cytoplasmic extracts. Further Western blotting experiments showed that the expressed proteins were located in the cytoplasm (data not shown).

The *R. capsulatus* DNA inserts in pALS1 and pALS2 was sequenced using the dye terminator method. Cycle sequencing reactions were performed on a Perkin Elmer Cetus DNA thermal cycler using Applied Biosystems or Perkin Elmer sequencing premix. Sequences were analysed using an Applied Biosystems 373A automated sequencer. DNA was sequenced in both directions starting from the T3 and T7 promoters in the vector then using primers

designed using new sequence information. The nucleotide sequence of the *R. capsulatus* DNA insert in pALS1 indicated that, as expected, this clone was an in-frame fusion of a sequence with homology to DMSO reductases to the N terminus of  $\beta$ -galactosidase encoded by the vector. Further sequence analysis revealed that this fusion was at the *Sau*3AI site (nucleotide 452) in the complete *dorA* sequence. Sequencing of this clone generated about 90% of the *dorA* sequence and a stop codon was identified at nucleotide 2530 in the complete sequence (Fig. 2). The complete sequence of *dorA* was obtained by sequencing a part of the second clone pALS2 by the primer walking method described above. This 2478 nucleotide sequence is preceded by a Shine–Dalgarno sequence (Fig. 2) and further sequencing revealed that the entire *dorA* gene was present.

*R. capsulatus* DorA clearly is most similar to *R. sphaeroides* DmsA [15], *E. coli* TorA [4] and BisC from *E. coli* and *R. sphaeroides* [21,22], and noticeably less similar to *E. coli* DmsA [3] (Table 1). All of the above enzymes form a sub-class of S- and N-oxide reductases in the molybdenum oxotransferase family which is distinct from the formate dehydrogenase/nitrate reductase class on the basis of primary sequence. The similarity between the DMSO reductases of the two *Rhodobacter* species is not surprising in view of their almost identical molecular properties. The close relationship between biotin sulfoxide reductase and DMSO reductase in *R. sphaeroides* has also been noted [14,22] while it has also been found that the periplasmic *E. coli* TMAO reductase (TorA) cross-reacts with the anti-DMSO reductase (*R. capsulatus*) antibody [23].

The *dorA* gene of *R. capsulatus* encodes a protein of deduced molecular mass 87 302 Da. The sequence of the N-terminus of purified mature DMSO reductase was determined by Alta Bioscience (University of Birmingham) using an ABI 473A automatic protein sequencer. The N-terminal sequence was EALANGTVMMSGSHWGVF. This was in agreement with the deduced amino-acid sequence except that there were 41 additional amino acids encoded between the initiator ATG, which is located 3' of the putative ribosome binding site and the N-terminus of the mature enzyme. The presence of an amino terminal signal sequence which is removed during secretion is consistent with the periplasmic location of the enzyme. The deduced molecular mass of the mature protein is 82 812 Da; this compares favourably with 82 000 Da as determined by SDS-PAGE [13].

Prokaryotic signal sequences are typically about 20 amino acids long [24]. Periplasmic proteins that contain MGD have been reported to possess a conserved motif (RRXF) in an unusually long signal sequence of greater than 30 amino acids [4,15,25]. *R. capsulatus* DorA also contains this sequence as do other periplasmic N- and S-oxide reductases (Fig. 3). The *Haemophilus influenzae* BisC also has a signal sequence (Fig. 3) and this is

|                                                             |      |                                                             |      |
|-------------------------------------------------------------|------|-------------------------------------------------------------|------|
| TGGCGGCGCTTGATCAGCAGCACTCATCCCCAACCGAAGAAAGAGCAGTCAATGAC    | 60   | GCTGGCCTCGATGCTGGGTGAGATCGGGCTCGCGGGCGCGGCTTCGGGCTGTCTATCA  | 1280 |
| <i>M T</i>                                                  | 2    | <i>L A S M L G Q I G L R G G G F G L S Y H</i>              | 402  |
| GAAGTTTTCGGAACAGAGCTGCGCGCAGAGCTTTACCGCGCGCTTTCCTCAGTACTC   | 120  | CTATTGGGGCGGTGGCAGCCCTCGAGCAGCGGTCCGGCGCTTCGGGCATCACCAGTGG  | 1340 |
| <i>K F S G N E L R A E L Y R R A F L S Y S</i>              | 22   | <i>Y S G G G T P S S S G P A L S G I T D G</i>              | 422  |
| GGTTCACCGGCGCGCTGGGATGTTGGGCGGTGCTTTCGGCAAGGCGCGCGCGC       | 180  | CGGGCGACGAAGGGCGGGAATGGCTGGCGGCGAGCGGCGCTTCGGGTATCCGGTGGC   | 1400 |
| <i>V A P G A L G M F G R S L L A K G A R A</i>              | 42   | <i>G R R R G R N G W R R A A L R C I P V A</i>              | 442  |
| CGAGGCGCTGGCCAATGGCAGGTGATGTCGGGCGAGCCACTGGGGCGTCTTACCGCGAC | 240  | GCGCGTGGTGCAGATGCTGGAACACCGCGCGCGGAATTCGACTTCAACGGTACGCGGTC | 1460 |
| <i>E A L A N G T V M S G S H W G V F T A T</i>              | 62   | <i>R V V D M L E N P G A E F D F N G T R S</i>              | 462  |
| GGTCGAAACGGCGCGCGCCACCGCTTACCCCTGGGAAAGACCGCATCCGACGCC      | 300  | GAAATCCCGGATGTGAAGATGGCTATTGGGTGGCGGAACCCCTTCGTGTACCATCA    | 1520 |
| <i>V E N G R A T A F T F F E K D H P T P</i>                | 82   | <i>K F P D V K M A Y W V G G T P S C H H Q</i>              | 482  |
| GATGCTGGAAGCGTGTGGATCGATCTATTCGCGACGCGGATCAAATATCCGATGGT    | 360  | GGACCGCAACCGCATGGTCAAGGCTGGGAAACCTGGAAACCTTCATCGTGCATGACTT  | 1580 |
| <i>M L E G V L D S I Y S P T R I K Y P M V</i>              | 102  | <i>D R N R M V K A W E K L E T F I V H D F</i>              | 502  |
| CGCGCGCAATTCCTGAAAGGCGTGAATGCTGATCGCTCCACCGCGGCAACCGCA      | 420  | CCAGTGGACGCCCGCGCGGCGCATGCGCATCTGCTGCCCGGACGACGAGTATGA      | 1640 |
| <i>R R E F L E K G V N A D R S T R G N G D</i>              | 122  | <i>Q W T P T A R H A D I V L P A T T S Y E</i>              | 522  |
| TTTTCGTCCTCGTGGGATCAGGCGCTCGATCTGATGGTTCGGGCGAGGTCAAAG      | 480  | ACGCAACGACATCGAGCAGATCGGCGATTATTGCAACACCGGCATCTGGCGATGAAGAA | 1700 |
| <i>F R P V S W D Q A L D L H G R G E V K R</i>              | 142  | <i>R N D I E T I G D Y S N T G I L A M K K</i>              | 542  |
| GGTCGAAGGAGACCTACGCGCGCGCGCTTTCGGCGGCTCTATGGCTGAAAGGCC      | 540  | GATCGTCAGCGCGCTTACGAAGCGCGCGGATACGACATCTTCGCCGCGGTGCCGA     | 1760 |
| <i>V E G D L R P A G V F G G S Y G W K S P</i>              | 162  | <i>I V E P L Y E A R S D Y D I F A A V A E</i>              | 562  |
| CGGGCGGCTGCACAAATGCACACGCTTCTGCGCGGATGCTGACGCTGGCGGCGGCTA   | 600  | ACGGCTGGGCAAGGCAAGGATTCACCGAAGGCAAGGACGAGATGGGCTGGATCAAGTC  | 1820 |
| <i>G R L H N C T T L L R M L T L A G G Y</i>                | 182  | <i>R L G K G K E F T E G K D E M G W I K S</i>              | 582  |
| TGTGAACGCGCGCGGATTTATTCGACCGCGCGCGCGGATGATCATGCCGATGTGGT    | 660  | CTTCTACGACGATGCGCGCAAGCAGGCAAGCGGGGTTCGAGATGCCCGGCTTCGACGC  | 1880 |
| <i>V N G A G D Y S T G A A Q V I M P H V V</i>              | 202  | <i>F Y D D A A K Q A K R G S R C P A F D A</i>              | 602  |
| CGGCACGCTGGAAGTCTATGAACAGCAGACCGCTGGCGGTGCTGGCGGAAACACCGA   | 720  | CTTCTGGGCGAAGGATCGTGAATTCGGGTCACCGACGCGCGGACTTCGTGCGCTA     | 1940 |
| <i>G T L E V Y E Q T A W P V L A E N T E</i>                | 222  | <i>F W A E G I V E F P V T D G A D F V R Y</i>              | 622  |
| AGTCATGGTGTCTGGCGCGCGATCCGATCAAGACAGCAGATATCGGCTGGGTGTATCC  | 800  | TGCCAGCTTCGGGAAGATCCGCTGCTCAACCCGCTGGGCGCGCGCGGCTGATCGA     | 2020 |
| <i>V M V F W A A D P I K T A D I G W V Y P</i>              | 242  | <i>A S F R E D P L L N P L G T P T G L I E</i>              | 642  |
| CGAATGCGCGCTATCCGGGACTGAGCGCTCAAGGCCAAGGCGACCAAGGTATCGT     | 860  | GATCTACTCGAAGAACATCGAGAAGATGGGCTATGACGATGCCCGCGCATCCGACCTG  | 2080 |
| <i>E H G A Y P G T E A L K A K G T K V I V</i>              | 262  | <i>I Y S K N I E K M G Y D D C P A H P T W</i>              | 662  |
| CATCGATCGGTCGCGACCAAGCGGTGAATTCCTGGCGCGGATCAAGTCACGCCGAA    | 920  | GATGGAACCGCTGGAACGCTCGACGGCGCGGGCGGAATATCCGCTGCACATCGCGGC   | 2140 |
| <i>I D P V R T K T A E F F G A D H V P T P K</i>            | 282  | <i>M E P L E R L D G P G A K Y P L H I A A</i>              | 682  |
| ACCGCAGACCGATGTCGCGATGCTGGGATGCGCGCATACGCTGGTGGCGAAGACCT    | 980  | TGCGACCGGTCGAACCGGTGTACTCGCACCGCTTACCGGCTCAACGGCACGGTGTGCG  | 2200 |
| <i>P Q T D V A I M L G M A H T L V A E D L</i>              | 302  | <i>R T R S T G V L A P V H R L N G T V L R</i>              | 702  |
| GTATGTAAGGACTTCATCGCAACTACACCTCGGGCTTCGACAAGTTCCTGCCCTATCT  | 1040 | CGAAGGCTATGCGGTGCGGGGCGAGCGCTCGCTGATGCACCCGACGACGCGCGCGC    | 2260 |
| <i>Y V K D F I A N Y T S G F D K F L P Y L</i>              | 322  | <i>E G Y A V Q G H E P C L M H P D D A A A</i>              | 722  |
| GATGGCGGAGACCGACAGCAGCGCGAAGACCGCGAATGGCGCTCGGATATCAGCGCGT  | 1100 | GCGCGGATCGCGGATGGGACGCTGGTGGGTCGACATGATCGCGGTGAGATCTGAC     | 2320 |
| <i>M G E T D S T P K T A E W A S D I S G V</i>              | 342  | <i>R G I A D G D V V R V H N D R G Q I L T</i>              | 742  |
| TCCGGCGGAGACGATCAAGGAATGGCGCGGCTGTTCAAATCGAAACGACGATGCTGGC  | 1160 | CGGGGTCAAGGTGACCGATGCGGTGATGAAGGGGTATCCAGATCTACGAAGGGGGCTG  | 2380 |
| <i>P A E T I K E L A R L F K S K R T M L A</i>              | 362  | <i>G V K V T D A V M K G V I Q I Y E G G W</i>              | 762  |
| GGCGGGCTGGTGCATGACGGATGATCAGCGGAGCGCGGATTCGATGCTGGTGAC      | 1220 | GTATGATCCCTCGGAGTGCAGGCGGGGCGGCTCGACAATACGGCGACGTTAACGT     | 2440 |
| <i>A G W S M Q R M H H G E Q A H W M L V T</i>              | 382  | <i>Y D P S D V T E A G T L D K Y G D V N V</i>              | 782  |
|                                                             |      | GCTGTCGGCGGATATCGGATGTCGAACTGGCGGAGGCGCACTGTGGTGCAGACCGTGT  | 2500 |
|                                                             |      | <i>L S A D I G M S K L A Q G N C G Q T V L</i>              | 802  |
|                                                             |      | GGCGGAGTTCGAGAATACACCGCGCGCGCTGACCGCTTGGTTCGCGCGAA          | 2560 |
|                                                             |      | <i>A E V E K Y T G P A V T L T G F G R A K</i>              | 822  |
|                                                             |      | GGCGGTGGAATAGCCCTTGAGCCTGTACGGAGGGTCTCTCC                   | 2602 |
|                                                             |      | <i>A V E *</i>                                              | 825  |

Fig. 2. *Rhodobacter capsulatus* dimethylsulfoxide reductase nucleotide and deduced amino-acid sequence. Complete *R. capsulatus* DMSO reductase sequence. Bold type indicates a Shine–Dalgarno sequence, the underlined amino-acid sequence is from N-terminal amino-acid sequencing matched with the sequence derived from DNA sequence data. The italicised amino-acid sequence indicates the signal sequence, a potential cleavage site is also predicted between amino acids 42 and 43 by the method of Von Heijne [19].

Table 1

Percent identity and similarity of amino-acid sequences of some *S*- and *N*-oxide reductases to *R. capsulatus* DMSO reductase

| Species                | Protein                          | Identity (%) | Similarity (%) |
|------------------------|----------------------------------|--------------|----------------|
| <i>R. sphaeroides</i>  | DmsA, DMSO reductase             | 73           | 83             |
| <i>H. influenzae</i>   | BisC, Biotin sulfoxide reductase | 47           | 66             |
| <i>E. coli</i>         | BisC, Biotin sulfoxide reductase | 45           | 64             |
| <i>E. coli</i>         | TorA, TMAO reductase             | 45           | 64             |
| <i>R. sphaeroides</i>  | BisC, Biotin sulfoxide reductase | 40           | 59             |
| <i>E. coli</i>         | DmsA, DMSO reductase             | 29           | 51             |
| <i>W. succinogenes</i> | PsrA, Polysulfide reductase      | 23           | 50             |

Comparisons were made using the method of Needleman and Wunsch [20] with each protein being compared individually to DmsA.

| Species                 | Peptide | Sequence                                                       | Length (no. of amino acids) | Evidence |
|-------------------------|---------|----------------------------------------------------------------|-----------------------------|----------|
| <i>R. capsulatus</i>    | DorA    | .....MTKFSGNELRAELYSRRFLSYSVAPGAL.MFGRSLL...AKGARA.....        | 41                          | EX       |
| <i>R. sphaeroides</i>   | DmsA    | .....MTKLSGQELHAELSRRFLSYTAAVGALGLCGTSL...AQGARA.....          | 42                          | VH       |
| <i>E. coli</i>          | TorA    | .....MNNDLFQASRRRFLAQLGLTVAGMLGFSLLTPRRATAAQA.....             | 42                          | EX       |
| <i>H. influenzae</i>    | DmsA    | .....MSNFNQISRRDFVKASSAG.AALAVSN..LTLPFNVMA.....               | 35                          | VH       |
| <i>H. influenzae</i>    | BisC    | .....MKNNVNEQRDRFLKKTSLGVAGSALSGMGVGVSKSAVA.....               | 40                          | VH       |
| <i>W. succinogenes</i>  | Psra    | .....METTMTTRDFLKSAGAAGAAGLVWSQTFPGTLGAL.....                  | 35                          | EX       |
| <i>E. coli</i>          | FdoG    | .....MQVSRROFFKICAGGMAGTTAAA..LGFAPSVALA.....                  | 33                          | VH       |
| <i>E. coli</i>          | FdnG    | .....MDVSRROFFKICAGGMAGTTVAA..LGFAPKQALA.....                  | 33                          | VH       |
| <i>W. succinogenes</i>  | FdhA    | .....MSEALSGRGN.....DRRFLKMSALAGVAGVQAVG.....                  | 32                          | EX       |
| <i>A. eutrophus</i>     | NosZ    | .....MSKEKASIGNPGGGIGRRQLGLTAALAGLAGVACTDKGAAPAAAAGAPASGAHG    | 56                          | VH       |
| <i>T. pantotropha</i>   | NapA    | .....MTISRDLKKAQAAG.IAAMAAN..IPLSSQAPA.....                    | 31                          | VH       |
| <i>A. eutrophus</i>     | NapA    | .....MKISRDRDFIKQTAIT.ATASVAG..VTLPAGA.....                    | 29                          | EX       |
| <i>P. denitrificans</i> | MauA    | MLGNFRFDDMVEKLSRRVAGQTSRRSVIGKLGATMLGIGLVPLLPVDRRGRVSRANA..... | 56                          | VH       |
| <i>R. capsulatus</i>    | FbcF    | .....MSHAEDNAGTRDRFLYHATAATGVVVTGAAG.....                      | 31                          | VH       |

Fig. 3. Alignment of unusually long signal peptides of MGD-binding proteins and other periplasmic redox proteins. Signal peptidase cleavage sites have been determined experimentally (EX) or using the weight matrix method of Von Heijne (VH) [19]. The sequences are those from *R. capsulatus* DorA (this work) and Rieske Fe-S protein, FbcF [26], *R. sphaeroides* DMSO reductase, DmsA [15], *Haemophilus influenzae* DMSO reductase, DmsA and biotin sulfoxide reductase, BisC [27], *E. coli* trimethylamine-N-oxide reductase, TorA [4], formate dehydrogenase, FdoG [28] and nitrate inducible formate dehydrogenase, FdnG [29], *Wolinella succinogenes* polysulfide reductase, PsrA [30] and formate dehydrogenase, FdhA [31], *Alcaligenes eutrophus* nitrous oxide reductase, NosZ [32] and periplasmic nitrate reductase, NapA [33], *Thiosphaera pantotropha* periplasmic nitrate reductase, NapA [37] and *Paracoccus denitrificans* methylamine dehydrogenase, MauA [34].

inconsistent with the cytoplasmic location of biotin sulfoxide reductases. In view of this observation we suggest that the *bisC* gene of *H. influenzae* is more likely to encode a periplasmic TMAO reductase. The signal sequences in the Cu-containing nitrous oxide reductase (NosZ) and methylamine dehydrogenase (MauA) which contains a tryptophan tryptophylquinone (TTQ) cofactor are 56 amino acids long [35]. The RRXF motif is also found in the signal sequences of these proteins as well as the Rieske Fe-S protein of the cytochrome *bc<sub>1</sub>* complex [26] (Fig. 3). The long signal sequences may be necessary for the process which couples prosthetic group insertion to secretion and folding.

Detailed analyses of MGD-binding proteins have identified a number of highly conserved sequences [2,6]. Recently Berks et al. [36] have identified eight conserved

segments in MGD-binding proteins which account for about 350 amino acids. This comprises approximately 40% of the smallest proteins in this class, such as DMSO/TMAO reductase. These conserved segments are punctuated by sequences of variable length. Segment I is conserved in nitrate reductase (apart from NarG and NarZ), formate dehydrogenases and DmsA from *E. coli* and its main feature is a putative 4Fe-4S binding motif identified by conserved cysteine residues. The presence of this Fe-S cluster has been confirmed experimentally in the periplasmic nitrate reductase from *Thiosphaera pantotropha* [25]. The iron-sulfur binding motif of segment I is absent from *E. coli* TorA [4] and from *R. capsulatus* DorA (Fig. 2) and this is consistent with the observations that periplasmic DMSO reductases do not contain prosthetic groups other than MGD.

#### Segment III

|                 |                |                                                                   |
|-----------------|----------------|-------------------------------------------------------------------|
| <i>R. caps.</i> | DorA (116–173) | SYGWSKSEGRHLNC.TLIRRLTLIAGGYVNGAGLYSTGAAQVI.PHVVGTLLEVYEQQT.AWP   |
| <i>E. coli</i>  | DmsA (142–197) | TLGGIMTRSWPPGNTLVARLMNCCGGYLNHYGDYSSAQIAEGLNYTY.GGWADGNPSD..      |
| <i>H. infl.</i> | DmsA (127–184) | TLGGTMAKSWPPASTMIARFMNCCGGYLNHYGDYSTAQIAVGLDYTYGGGWALGNMGAD..     |
| <i>R. sph.</i>  | DmsA (110–172) | SYGWSKSEGRHLNCQVLMRRALNLAGGEVFNSSGELYSTAAQAIIIMPHVMGTLEVYEQQT.AWP |
| <i>H. infl.</i> | BisC (113–170) | SYGWFCGSLHASRTLLQRYMNATGGFVCHKGDYSTGAAQVIMPHVLTIEVYEQQT.SWE       |
| <i>E. coli</i>  | BisC (76–135)  | SYGWSRNGVLHKASTLLQRYMALAGGYTGHLGEYSTGAAQAIMPVVVGSEVYEQQT.SWP      |
| <i>E. coli</i>  | TorA (158–216) | S.GWQSTGMFHNASGMRAKRIALHGNVSGTGGDYSTGAAQVILPRVVGSMVEVYEQQT.SWP    |
| <i>R. sph.</i>  | BisC (87–147)  | SYGWTSCGRFHASTLLKRLNLVGGFTGHVDYTSIAGGFVILRHTLGDGRACGGQANTLD       |
| <i>W. suc.</i>  | Psra (153–196) | ....WNKTWFHH.....LAQAYGSPNIFGHESTCP....LAYNMAGRDVFGG...SMN        |

#### Segment VI

|                 |                |                                                                                |
|-----------------|----------------|--------------------------------------------------------------------------------|
| <i>R. caps.</i> | DorA (452–525) | .KLETFIVHDFQWTPT.RHADIVLPATTSYERNDIETIGDYSN.TILAM.KKIVEEPLYEARSDDYDIFAAVAER.GK |
| <i>R. sph.</i>  | DmsA (448–525) | .KLETFIVQDFQWTATARHADIVLPATTSYERNDIESVGDYSNRILLAM.KKVVDPLYEARSDDYDIFAAERLKG    |
| <i>H. infl.</i> | BisC (452–523) | .KPDWVIVNEVNWPTARMADIVLPATTSYERNDLTMAGDYSMMSVYPM.KQVVPQFEAKNDYDIFVELAKR..A     |
| <i>E. coli</i>  | BisC (409–482) | .KPELVVISECFWTPAAKHADIVLPATTSYERNDLTMTGDYSNCHLVEM.KQVVPQFEARNDFDFAELSERWEK     |
| <i>H. infl.</i> | DmsA (463–533) | TQCEMIITIDNHMTSTAKYSEILLPDCTTSEQMDFALDAFVSNNMAYVIFADQVIKESFECRPIYDMLSDIAEKM..  |
| <i>E. coli</i>  | TorA (503–576) | .NVETVIADNQWTSCTREFALIVLPATTQFERNDLDQYGNHNSNRGIAM.KQVVPQFEARNDFDIFRELRRFNR     |
| <i>E. coli</i>  | DmsA (475–548) | KKCELIVVIDCHMTSSAKYALILLPDCTASEQMDFALDASCNNMSYVIFNDQVIKPRFECKTYIEMTSELAKRL..   |
| <i>R. sph.</i>  | BisC (415–485) | .RPEETIIVQDEMFETATAKRALIVLPASTSIERNDL..AGNKRSEFILAM.GQAIAPLGEARSDFDIFNALSGLK.. |
| <i>W. suc.</i>  | Psra (466–540) | .SMDLVVCVDIQVSTDAWFAADVLPDPTTYLERDEEFTAGGKNFSFGIGRQKQVLEPLGDAKPGWKIAKELSEKMG   |

Fig. 4. Segments of amino-acid similarity in MGD-dependent enzymes. The conserved region in segment III (◇) is repeated, with a lesser degree of homology in a subset of the sequences in segment VI (↓).

A number of groups have suggested that Segment III of MGD-binding proteins contained the substrate binding site [36,38]. In a refinement of this analysis [37] two specificity classes of Segment III were identified. The first class was the formate dehydrogenase/soluble nitrate reductase class which possess a conserved cysteine or selenocysteine which may be ligands for the Mo [36,38]. There is strong evidence for this in the case of the selenocysteine residue in formate dehydrogenase [38]. The second specificity class is the *S*- and *N*-oxide reductases and they lack the conserved cysteine described above. Instead, a conserved serine has been identified [36] and postulated to be a ligand to Mo [38]. This has now been confirmed following the publication of the crystal structure of the DMSO reductase from *R. sphaeroides* [39]. The sequence GDYS of segment III is conserved in *E. coli* TorA, BisC and DmsA, *R. sphaeroides* DmsA and DsrA [14,39] and in *R. capsulatus* DorA (Fig. 4). The possible roles of Segments IV–VIII in MGD-binding proteins have been discussed recently [36] and will become clear as the structure of the high resolution crystal structure of DMSO reductase is analysed. An unexpected outcome from the work of Rees and co-workers [39] was the presence of a bis(MGD) molybdenum cofactor in *R. sphaeroides* DMSO reductase. Given the close sequence similarity between DMSO reductases from the two species of *Rhodobacter* it is expected that this novel form of the molybdenum cofactor will also be present in the enzyme from *R. capsulatus*.

This work was supported by grant A29531856 to A.G.M. and G.R.H. from the Australian Research Council. A.L.S. is the recipient of an Australian Postgraduate Award.

## References

- [1] Barret, E.L. and Kwan, H.S. (1985) *Annu. Rev. Microbiol.* 39, 131–149.
- [2] Weiner, J.H., Rothery, R.A., Sambasivarao, D. and Trieber, C.A. (1992) *Biochim. Biophys. Acta* 1102, 1–18.
- [3] Bilous, P.T., Cole, S.T., Anderson, W.F. and Weiner, J.H. (1988) *Mol. Microbiol.* 2, 785–795.
- [4] Mejean, V., Iobbi-Nivol, C., Lepelletier, G., Giordano, G., Chipaux, M. and Pascal, M.C. (1994) *Mol. Microbiol.* 11, 1169–1179.
- [5] McEwan, A.G. (1994) *Antonie van Leeuwenhoek* 66, 151–164.
- [6] Wootton, J.C., Nicholson, R.E., Cock, J.M., Walters, D.E., Burke, J.F., Doyle, W.A. and Bray, R.C. (1991) *Biochim. Biophys. Acta* 1057, 157–185.
- [7] Rajagopalan, K.V. (1991) *Adv. Enzymol. Rel. Areas Mol. Biol.* 64, 215–290.
- [8] Schultz, B.E., Hille, R. and Holm, R.H. (1995) *J. Am. Chem. Soc.* 117, 827–828.
- [9] Hille, R. and Sprecher, H. (1987) *J. Biol. Chem.* 262, 10914–10917.
- [10] Chan, M.K., Mukund, S., Kleitzen, A., Adams, M.W.W. and Rees, D.C. (1995) *Science* 267, 1463–1469.
- [11] Romao, M.J., Archer, M., Moura, L., Moura, J.J.G., LeGall, J., Engh, R., Schneider, M., Hof, P. and Huber, R. (1995) *Science* 270, 1170–1176.
- [12] Johnson, J.L., Bastian, N.R. and Rajagopalan, K.V. (1991) *Proc. Natl. Acad. Sci. USA* 87, 3190–3194.
- [13] McEwan, A.G., Ferguson, S.J. and Jackson, J.B. (1991) *Biochem. J.* 274, 305–307.
- [14] Barber, M.J., Van Valkenburgh, H., Trimboli, A.J., Pollock, V.V., Neame, P.J. and Bastian, N.R. (1995) *Arch. Biochem. Biophys.* 320, 266–275.
- [15] Yamamoto, I., Wada, N., Ujiye, T., Tacibana, M., Matsuzaki, M., Kajiwar, H., Watanabe, Y., Hirano, H., Okubo, A., Satoh, T. and Yamazaki, S. (1995) *Biosci. Biotech. Biochem.* 59, 1850–1855.
- [16] Kaufmann, N., Hudig, H. and Drews, G. (1985) *Mol. Gen. Genet.* 198, 153–158.
- [17] Donohue, T.J., McEwan, A.G. and Kaplan, S. (1986) *J. Bacteriol.* 168, 962–972.
- [18] Hatton, A.D., Malin, G. and McEwan, A.G. (1994) *Arch. Microbiol.* 162, 148–150.
- [19] Von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683–4691.
- [20] Needleman, S.B. and Wunsch, C.D. (1970) *J. Mol. Biol.* 48, 443–453.
- [21] Pierson, D.E. and Campbell, A.J. (1990) *J. Bacteriol.* 172, 2194–2198.
- [22] Pollock, V.V. and Barber, M.J. (1995) *Arch. Biochem. Biophys.* 318, 322–332.
- [23] Bonnett, T.C. (1995) PhD thesis, University of East Anglia, UK.
- [24] Benson, S.A., Hall, M.N. and Silhavy, T.J. (1985) *Annu. Rev. Biochem.* 54, 101–134.
- [25] Breton, J., Berks, B.C., Reilly, A., Thomson, A.J., Ferguson, S.J. and Richardson, D.J. (1994) *FEBS Lett.* 345, 76–80.
- [26] Davidson, E. and Daldal, F. (1987) *J. Mol. Biol.* 195, 25–29.
- [27] Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Uitterback, T.R., Hanna, M.C., Nguyen, D.C., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Goughagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C. (1995) *Science* 269, 496–512.
- [28] Abaiou, H., Pommier, J., Giordano, G. and Mandrand-Bertheholt, M.A. unpublished sequence. GenBank accession number X87583.
- [29] Berg, B.L., Li, J., Heider, J. and Stewart, V. (1991) *J. Biol. Chem.* 266, 22380–22385.
- [30] Krafft, T., Bokranz, M., Klimmek, O., Schroder, I., Fahrenholz, F., Kojro, E. and Kroger, A. (1992) *Eur. J. Biochem.* 206, 503–510.
- [31] Bokranz, M., Gutmann, M., Kortner, C., Kojro, E., Fahrenholz, F., Lauterbach, F. and Kroger, A. (1991) *Arch. Microbiol.* 152, 119–128.
- [32] Zumft, W.G., Dreusch, A., Lochelt, S., Cuypers, H., Freidrich, B. and Schneider, B. (1992) *Eur. J. Biochem.* 208, 31–40.
- [33] Siddiqui, R.A., Warnecke-Eberz, U., Hergsberger, A., Schneider, B., Kostka, S. and Freidrich, B. (1993) *J. Bacteriol.* 175, 5867–5876.
- [34] Chistoserdov, A., Boyd, C., Mathews, F.S. and Lidstrom, M.E. (1992) *Biochem. Biophys. Res. Commun.* 184, 1181–1189.
- [35] Hoeren, F.U., Berks, B.C., Ferguson, S.J. and McCarthy, J.E.G. (1993) *Eur. J. Biochem.* 218, 49–57.
- [36] Berks, B.C., Ferguson, S.J., Moir, J.W.B. and Richardson, D.J. (1995) *Biochim. Biophys. Acta* 1232, 97–173.
- [37] Berks, B.C., Richardson, D.J., Reilly, A., Willis, A.C. and Ferguson, S.J. (1995) *Biochem. J.* 309, 983–992.
- [38] Gladyshev, V.N., Khangulov, S.V., Axley, M.J. and Stadtman, T.C. (1994) *Proc. Natl. Acad. Sci. USA* 91, 7708–7711.
- [39] Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K.V. and Rees, D.C. (1996) *Science* 272, 1615–1621.