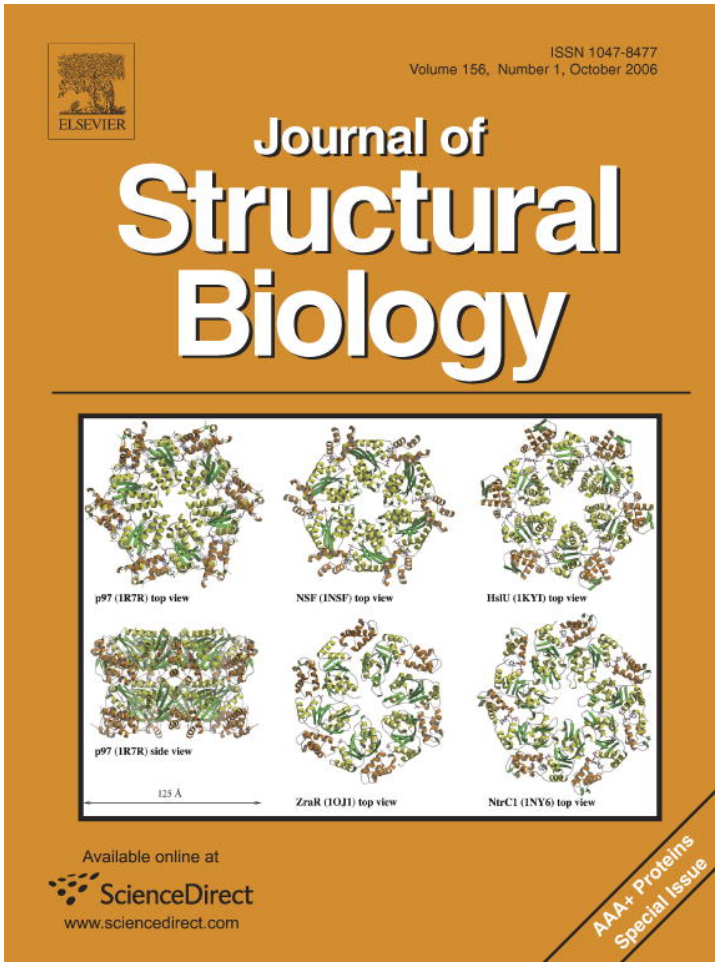

Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 156 (2006) 200–209
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

MoxR AAA+ ATPases: A novel family of molecular chaperones?

Jamie Snider, Walid A. Houry *

1 King's College Circle, Medical Sciences Building, Department of Biochemistry, University of Toronto, Toronto, Ont., Canada M5S 1A8

Received 23 December 2005; received in revised form 15 February 2006; accepted 16 February 2006
Available online 23 March 2006

Abstract

The MoxR AAA+ family is a large, diverse group of ATPases that, so far, has been poorly studied. Members of this family are found throughout the Bacteria and Archaea superkingdoms, but have not yet been detected in Eukaryota. The limited experimental data available to date suggest that members of this family might have chaperone-like activities. Here we present an extensive phylogenetic analysis which builds upon our previously published work, and reveals that the MoxR family can be divided into at least seven subfamilies, including MoxR Proper (MRP), TM0930, RavA, CGN, APE2220, PA2707, and YehL. We also include a comprehensive overview and gene context analysis for each of these subfamilies. Our data reveal distinct conserved associations of certain MoxR family members with specific genes, including further support for our previously reported observation that many members of the MoxR AAA+ family are found near Von Willebrand Factor Type A (VWA) proteins and are likely to function with them. We propose, based on bioinformatic analyses and the available literature, that the MoxR AAA+ proteins function with VWA domain-containing proteins to form a chaperone system that is important for the folding/activation of proteins and protein complexes by primarily mediating the insertion of metal cofactors into the substrate molecules.

Keywords: AAA+; Von Willebrand Factor Type A; MoxR; Chaperones; Metal insertion

1. Introduction

ATPases Associated with various cellular Activities, or AAA+ proteins, are a large superfamily of P-loop NTPases, whose members display a remarkable range of functional diversity. These proteins are involved in processes ranging from protein refolding and degradation to DNA repair and replication (Iyer et al., 2004). Regardless of their specific function, however, AAA+ proteins display a general role in molecular remodeling events, using the power of ATP hydrolysis to bring about conformational changes. AAA+ proteins are defined by the presence of the AAA+ module, a 200–250 amino acid region responsible for ATP binding and hydrolysis. These modules are comprised of two distinct structural subdomains: an N-terminal α/β-core subdomain and a smaller, C-terminal α-helical subdomain (Ogura and Wilkinson, 2001). AAA+ modules contain a variety of conserved sequence motifs responsible for ATP sensing and hydrolysis. Major motifs include the Walker A motif (GxxGxGKT), the Walker B motif (hhhhDE), Sensor I, and Sensor II (Neuwald et al., 1999). AAA+ proteins generally function as oligomers, often forming hexameric rings (Hanson and Whiteheart, 2005).

In our effort to discover novel molecular chaperones in bacteria, we were intrigued by several reports in the literature of AAA+ proteins that were required for the proper maturation of specific proteins or complexes. Upon further analysis of these AAA+ proteins, we discovered that they belonged to the MoxR family. This family contains a large number of poorly characterized AAA+ proteins. It is remarkably widespread throughout bacteria and archaea, with members being represented in all major lineages (Frickey and Lupas, 2004; Iyer et al., 2004; Neuwald et al., 1999); however, no eukaryotic members have yet been identified (Iyer et al., 2004). Though the exact functional role of MoxR proteins is unclear at present, research performed to

* Corresponding author. Fax: +1 416 978 8548.
E-mail address: walid.houry@utoronto.ca (W.A. Houry).

1047-8477/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2006.02.009

date suggests a chaperone-like role in the assembly and activation of specific protein complexes. Phylogenetic analysis of the MoxR AAA+ protein family shows that it can be divided into at least seven smaller subfamilies (Fig. 1 and Supplementary Table 1). These include MoxR Proper (MRP), TM0930 (formerly APE0892), RavA, CGN, APE2220, PA2707, and YehL. Of these, only members of the MRP, RavA, and CGN subfamilies have been studied experimentally. Here, we provide an overview of the MoxR AAA+ proteins based on bioinformatic analyses and the available literature, highlighting the novel chaperone-like activity of members of this family.

2. Materials and methods

Three MoxR AAA+ amino acid sequences were selected from each MoxR subfamily based on a previously published phylogenetic analysis (Snider et al., 2006). Each of the 18 total sequences was subjected to BlastP analysis against 275 of the microbial genomes available in the NCBI database using default parameters (Altschul et al., 1990). All homologous protein sequences matching with an Expect value $\leqslant 0.05$ were selected, compiled into a single list and filtered to eliminate redundancy. Each sequence was then compared against the Clusters of Orthologous Groups (COG) database using the COGNITOR program and all those identified as belonging to the MoxR COG0714 were selected (Tatusov et al., 2000). The sequences were subjected to preliminary alignments using the MUSCLE (Edgar, 2004) and CLUSTALW (Thompson et al., 1994) programs with default parameters, and then the sequences were shortened to include only the AAA+ modules. All sequences with incomplete AAA+ modules were removed and not included in the subsequent analysis.

The remaining 596 sequences were aligned once again using the MUSCLE program. The phylogenetic tree was constructed from this alignment via the neighbor-joining method using the PROTDIST and NEIGHBOR programs available in the PHYLIP package (Felsenstein, 1996). A PMB model of substitution (Veerassamy et al., 2003) and a coefficient of variation of 0.7 were used. For each of the 596 sequences included in the analysis, the protein sequences encoded by the 10 genes surrounding each MoxR (5 on both sides, where available) were extracted from the NCBI database. A total of 6366 sequences were identified and compared against the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2005), available through the NCBI website, to detect putative domains. Sequences from each of the major subfamilies identified in our phylogenetic tree were grouped, and the total occurrence of individual domains in a given subfamily was examined. Domains occurring in high frequency were selected, and the proteins containing those domains were manually examined. This allowed the identification of proteins whose genes occur in the neighborhood of MoxR genes with high frequency.

3. Results and discussion

To obtain a global view of the MoxR AAA+ family, we performed a large scale phylogenetic analysis, building and expanding upon our previously published work (Snider et al., 2006). 596 members of the MoxR COG0714 were identified by BlastP analysis with 18 representative MoxR sequences selected from our previous analysis (Snider et al., 2006), against 275 microbial genomes available in the NCBI database. These sequences were aligned and used in the construction of a phylogenetic tree as described in Section 2 (Fig. 1 and Supplementary Table 1). In our previous work, in which 156 MoxR proteins from 94 organisms were analyzed, we detected only six distinct subfamilies. Our current analysis is more comprehensive and covers a much broader spectrum of organisms resulting in the detection of an additional major MoxR subfamily (APE2220 branched from previously assigned PA2707 subfamily). This brings the total subfamily count to seven (Fig. 1): MoxR Proper (MRP), TM0930, RavA, CGN, APE2220, PA2707, and YehL. The signature sequences characterizing each subfamily are given in Supplementary Fig. 1.

In an effort to learn more about the function and role of the various MoxR AAA+ subfamily members, we also conducted an analysis of the genes surrounding each MoxR gene. By identifying genes which occur in close proximity to MoxR genes, it is hoped that potential functional partners can be identified. Where sequencing information was available, the protein products of five genes on either side of a *moxR* gene were examined for putative domains using the Conserved Domain Database (CDD). The sequences were grouped based upon the individual MoxR subfamilies identified in the phylogenetic tree, and then analyzed to identify

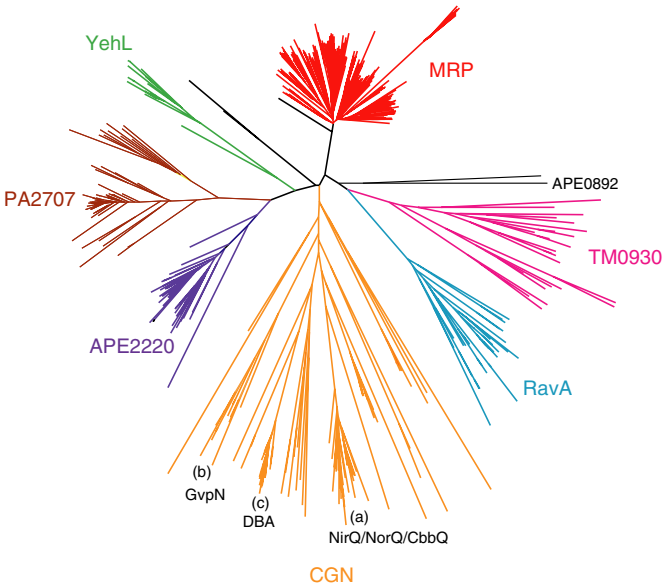

Fig. 1. Phylogenetic tree of the MoxR AAA+ family. The tree was constructed from 596 MoxR AAA+ (COG0714) sequences as described in the Materials and Methods. Each subfamily is represented by a distinct color. Small subgroups within the CGN branch are labeled.

neighboring genes. We selected this particular approach because it is highly specific and allowed us to examine the different MoxR subfamilies in detail with a minimum of background noise.

3.1. MoxR proper subfamily

Of all of the MoxR subfamilies identified, MoxR proper (MRP) appears to be the largest and most diverse. From the 596 MoxR AAA+ ATPases identified in our study, 287 (48.2%) belong to the MRP subfamily (Fig. 1 and Supplementary Table 1). These are distributed across 155 of the 275 organisms included in our analysis. Most of these organisms contain multiple MRP proteins. *Rhodopirellula baltica SH 1* is the most extreme example, with a genome encoding an astounding 15 MRP subfamily members (Supplementary Table 1). The range of organisms is remarkably diverse, including members of all major subdivisions of the Proteobacteria, as well as the Acidobacteria, Actinobacteria, Aquificae, Bacteroidetes, Chlamydiae, Chlorobi, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Firmicutes, Planctomycetes, Spirochaetes, and Thermotogae phyla. Representatives are also found in the Archaea superkingdom (Fig. 2). Notably, no members of the MRP family are present in *Escherichia coli* K12 or any of the other enterobacteria included in our analysis.

Our analysis of MRP encoding genes shows that they are closely associated with genes encoding proteins of unknown function belonging to the COG1721 (Fig. 3A). Of the 287 MRP members, 258 (90%) are close to genes encoding COG1721 proteins. Two hundred and fifty of these (97%) are immediately adjacent to COG1721 proteins. In addition, three MRP genes are close to two copies of COG1721 encoding genes. Most of the COG1721 proteins (94%) also contain a PFAM DUF58 domain (Domain of Unknown Function 58). A small subset of the COG1721 proteins (~13%) also contain a Von Willebrand Factor Type A (VWA) domain. The functions of COG1721 and DUF58 proteins is currently unknown. Based upon the remarkably high conservation of position between these genes and the MRP genes, however, it seems reasonable to assume that they likely function together.

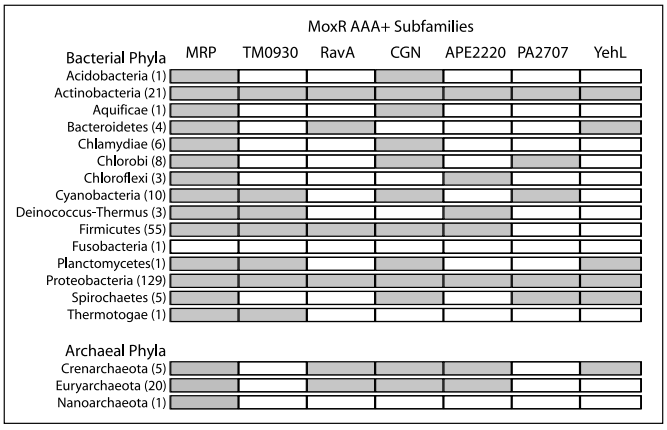

Fig. 2. Distribution of MoxR AAA+ proteins across organisms. Organisms are sorted according to phyla. A filled box indicates that a particular subfamily is present in at least one member of a given phylum. The number of organisms from each phylum included in our analysis is shown in parenthesis.

Of the entire MRP subfamily, 105 of the 287 family members (~37%) were found near VWA encoding genes. The VWA domain is a metal binding domain often involved in mediating protein–protein interactions (Whittaker and Hynes, 2002). Binding of metal, which is typically magnesium, occurs through a non-contiguous metal ion dependant adhesion site (MIDAS) that is important for binding to protein ligands (Xiong et al., 2002). Although these proteins have been well-studied in eukaryotic organisms, where they are known to be involved in a wide-range of processes including cell-adhesion, transport, the complement system, proteolysis, transcription, DNA repair, and ribosome biogenesis, their roles in bacteria and archaea are much less understood. Research has implicated them, however, in bacterial surface adhesion, fibrinogen binding, serum opacity, and metal insertion (Kachlany et al., 2000; Katerov et al., 2000; Willows, 2003). In 68 of the 287 MRP family members, the MRP genes are close to more than one VWA encoding gene (generally two, see Fig. 3A). Virtually, all of the MRP genes associated with VWA encoding genes are also near COG1721 genes, and, interestingly, 21% of all of the VWA genes are fused to a COG1721. Close to half of these MRP genes are also closely associated with genes encoding proteins containing tetratricopeptide repeat (TPR) domains, which are known to be important in mediating protein–protein interactions (Blatch and Lassle, 1999). In some instances, the TPR and VWA domains are on a single polypeptide chain.

Another 26% of MRP genes are in close genomic proximity to genes encoding proteins with a putative transglutaminase domain (CDD 22804) (Fig. 3A). None of these MRP genes are associated with VWA genes, although virtually all of them are associated with COG1721 genes. This relationship makes it tempting to speculate that these MRP proteins may interact with the transglutaminases, perhaps utilizing them as substrates.

Limited research has been conducted on MRP genes located near methanol dehydrogenase gene clusters in several organisms. The first *moxR* gene described, and from which the entire family derives its name, was an MRP subfamily member from *Paracoccus denitrificans*. *P. denitrificans* is gram-negative soil bacterium capable of growing methylotrophically on one-carbon compounds such as methanol or methylamine (Van Spanning et al., 1991). Oxidation of methanol is carried out by the enzyme methanol dehydrogenase (MDH), a tetrameric complex consisting of two identical large (α) and two identical small (β) subunits. MDH utilizes pyrroloquinoline quinone (PQQ, Fig. 4A) as a cofactor and cytochrome *c* as an electron acceptor. Activity of MDH is also dependent upon bound Ca^{2+} ions which are believed to be important for PQQ coordination and activation (Anthony, 2004). Studies in *P. denitrificans*

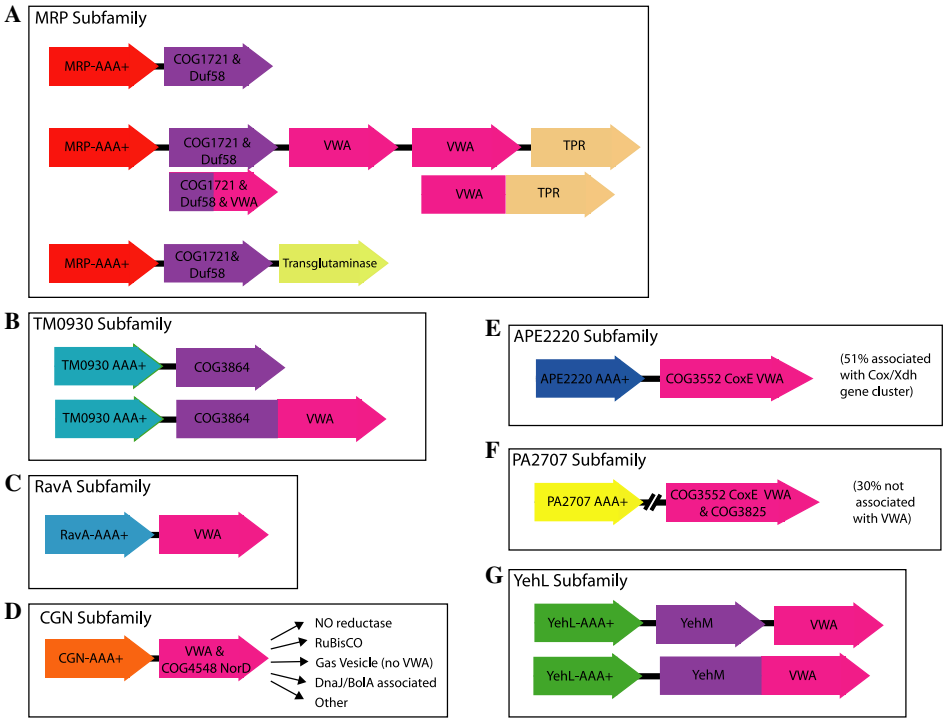

Fig. 3. General gene structure for each MoxR AAA+ subfamily. (A–G) Show the major conserved gene organization present around the MoxR genes for each subfamily. Note that the gene structure for MRP subfamily members associated with VWA proteins is quite diverse [(A), line 2]. The gene structure shown is the most frequently occurring (> 42% of cases). In (F), the double slash separating the genes indicates that the positions of the VWA genes with respect to the PA2707 AAA+ genes are highly variable.

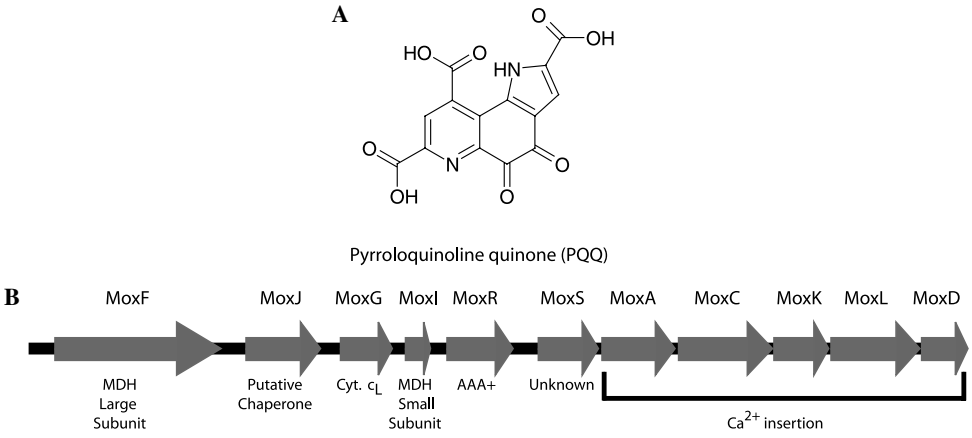

Fig. 4. (A) Structure of Pyrroloquinoline quinone (PQQ) (Anthony, 2004). (B) Shown is the organization of the *Methylobacterium extorquens* methanol dehydrogenase (MDH) gene region.

showed that, when cells contained an insertion mutant of the *moxR* gene, they were unable to grow on minimal media using methanol as a carbon source. Growth on methylamine or succinate, however, was unaffected. Examination of the cells showed that MDH, PQQ, and cytochrome *c* levels were all normal, however, the expressed MDH was inactive. Thus, the authors concluded that the *moxR* gene product appears to be important for activation of methanol dehydrogenase, possibly via modification and activation of one of the subunits of MDH, or via regulation of genes encoding the MDH activators (Van Spanning et al., 1991).

Similar results were obtained in *Methylobacterium extorquens*, where cells containing an insertion mutant of an MRP gene located near the MDH gene cluster also lost the ability to grow on methanol (Toyama et al., 1998). The *moxR* gene is located adjacent to another gene, *moxS*, which encodes a protein of unknown function (Fig. 4B). Interestingly, the *moxRS* genes studied in *M. extorquens* are located between two MDH gene clusters: the first (*moxFJGI*) encodes the MDH structural genes and the second encodes proteins known to be required for Ca^{2+} insertion (*moxACKLD*) (Amaratunga et al., 1997; Richardson and Anthony, 1992) (Fig. 4B). Disruption of genes in the latter region was

observed to result in MDH lacking Ca^{2+} and containing an abnormally bound PQQ cofactor (Richardson and Anthony, 1992). Intriguingly, the *moxC* and *moxL* genes encode proteins containing VWA domains. Such an arrangement is reminiscent of the large subset of MRP genes found in close proximity to VWA containing proteins. This, in conjunction with the tendency to find members of other MoxR subfamily genes in association with VWA encoding genes, may suggest an association between the *moxR* gene product and those of the *moxACKLD*, thereby implicating MoxR in metal insertion into MDH. Such an involvement in metal insertion is reminiscent of the metal chelatase enzymes, to which the MoxR family is closely related (Iyer et al., 2004). These enzymes utilize both AAA+ and VWA domains to mediate insertion of metal ions into porphyrin rings as part of the synthesis of cobalamin or (bacterio)chlorophyll. In the case of Mg chelatase, one of the subunits, BchD, actually contains a AAA+ and VWA domain fused together on a single polypeptide (Fodje et al., 2001).

It is noteworthy that the *moxR* gene in *M. extorquens* does not appear to be in the vicinity of a COG1721 encoding gene, while in *P. denitrificans*, which has a similar gene arrangement in this area, the *moxR* gene is adjacent to a COG1721 gene. This may suggest that the COG1721 protein is not absolutely required for MoxR function in MDH activation. Alternatively, the COG1721 gene in *M. extorquens* may simply have not yet been properly identified or may exist elsewhere in the genome.

Since only a very small subset of the MRP subfamily members are found in organisms that contain the methanol dehydrogenase gene cluster, it is clear that the function of these proteins is more general in nature. It is tempting to speculate that the MRP, and indeed MoxR proteins as a whole, may be involved in general metal insertion processes, possibly working as molecular chaperones helping to mediate proper insertion of metal ions/metal cofactors into substrate proteins.

3.2. TM0930 subfamily

The TM0930 subfamily contains 24 of the 596 MoxR AAA+ proteins analyzed (Fig. 1 and Supplementary Table 1). TM0930 genes are present in 21 different organisms, including members of the Actinobacteria, Cyanobacteria, Deincoccus-Thermus, Firmicutes, Planctomycetes, Proteobacteria (Beta, Delta/Epsilon, and Alpha subdivisions), and Thermotogae phyla (Fig. 2). No archaeal members were identified. TM0930 genes are all found near (typically immediately adjacent to) genes encoding proteins belonging to COG3864 (Fig. 3B). This COG represents a group of bacterial proteins of unknown function. A subset of these COG3864 gene products (37.5%) also contain a C-terminal VWA domain. Members of this subfamily have not been studied experimentally, and, thus, nothing is known about their specific function. The conservation of gene position between the TM0930 and COG3864/VWA encoding genes suggests that they may function together.

It should be noted that our current analysis suggests that APE0892, which was previously grouped with some members of the now expanded TM0930 branch (Snider et al., 2006), is actually distinct and seems to represent a minor branch (Fig. 1 and Supplementary Table 1). This conclusion is further supported by our gene neighborhood analysis since APE0892 is not found in the neighborhood of a COG3864 encoding gene.

3.3. RavA subfamily

The RavA subfamily consists of AAA+ proteins comprised of an N-terminal MoxR-type AAA+ module, followed by a poorly conserved C-terminal region of variant length. Of the 596 MoxR AAA+ proteins we identified, 39 belong to the RavA subfamily (Fig. 1 and Supplementary Table 1). These are distributed across 35 of the 275 organisms analyzed (Supplementary Table 1). RavA proteins are found in a range of bacteria belonging to the Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria (Beta, Delta/Epsilon, and Gamma subdivisions) phyla, as well as in members of Crenarchaeota and Euryarchaeota phyla of the Archaea (Fig. 2). Most organisms contain only a single copy of the *ravA* gene, however, certain Archaea, particularly *Methanocaldococcus jannaschii* DSM2611 and the three members of the Methanosarcina subdivision, each contain two copies (Supplementary Table 1).

The genes encoding RavA proteins are almost always found adjacent to genes encoding proteins containing a VWA domain (Snider et al., 2006) (Fig. 3C). *Cytophaga hutchinsonii* and *Chromobacterium violaceum* RavA's are exceptions, with no nearby VWA-containing gene being detected. Our analysis did not detect any other genes co-occurring with members of the RavA subfamily as a whole.

Of the numerous RavA-type MoxR AAA+ proteins identified, only *E. coli* K12 MG1655 RavA (YieN_ECOLI) has been characterized in detail (Snider et al., 2006). Studies on WT MG1655 cells grown in rich and minimal media at 37 °C have shown that induction of RavA is maximal towards late log and early stationary phase. The *ravA* promoter region contains a nearly perfect σ^S consensus sequence, and induction has been shown to be eliminated in strains deficient in *rpoS*, suggesting that this stationary phase sigma factor is responsible, at least in part, for *ravA* regulation. Northern blotting and RT-PCR studies have demonstrated that the *ravA* gene comprises an operon with the VWA-protein encoding gene adjacent to it, *viaA* (YieM_ECOLI). Considering the invariant adjacency of *ravA* and *viaA* genes across the various organisms studied (Fig. 3C), it is likely this operon arrangement is conserved. The RavA protein itself has also been purified and characterized. Studies have shown that it is a functional ATPase which forms hexameric rings in the presence of nucleotide. The enzyme also hydrolyzes GTP, although not as effectively as ATP. *E. coli* K12 RavA consists of three discrete fragments, including the AAA+ core and α-helical subdomains, as well as a poorly conserved C-terminal domain of

unknown function. *ravA* deletion mutants do not possess any obvious phenotype under a wide range of conditions, so the exact function of the protein in the cell is still elusive.

The RavA protein was found to interact strongly with the inducible lysine decarboxylase enzyme, LdcI, encoded by the *cadA* gene (Snider et al., 2006). LdcI is a pyridoxal phosphate (PLP, a vitamin B6 derivative) dependent decameric enzyme (Sabo et al., 1974; Sabo and Fischer, 1974) that plays a major role in the acid stress response in bacteria (Merrell and Camilli, 1999; Park et al., 1996; Soksawatmaekhin et al., 2004). The RavA–LdcI complex has been visualized by negative stain electron microscopy, and has been shown to form a remarkable 'cage-like' structure comprised of two LdcI decamers linked by up to five RavA oligomers. RavA does not appear to be important for the expression, biogenesis, or function of the LdcI enzyme, so the exact role of the RavA–LdcI complex is still unclear. Interestingly, the RavA protein does not interact with the constitutive lysine decarboxylase, LdcC, even though LdcC is 69% identical and 84% similar to LdcI. This suggests that the interaction between RavA and LdcI is highly specific. It has been proposed that the RavA–LdcI complex may be involved in the regulation of RavA activity under low pH conditions. The complex may be important, for instance, in sequestering RavA from, or enhancing RavA activity towards its substrates. Alternatively, complex formation may direct RavA towards an entirely new set of substrates. The exact role of RavA in the cell is under investigation.

3.4. CGN subfamily

The CGN subfamily, (CbbQ/GvpN/NorQ), refers to a large, highly diversified branch of the MoxR AAA+ phylogenetic tree. From our analysis, 114 of 596 MoxR AAA+ proteins belong to this subfamily (Fig. 1 and supplementary Table 1). These are found in 81 of the organisms included in our study, including members of all major subdivisions of the Proteobacteria, as well as of the Acidobacteria, Actinobacteria, Aquificae, Chlamydiae, Chlorobi, Cyanobacteria, Firmicutes, Planctomycetes, and Spirochaetes phyla. CGN proteins are also found in members of the Euryarchaeota and Crenarchaeota phyla of the Archaea superkingdom (Fig. 2).

A global analysis of the family shows that 83 of the CGN genes co-occur with VWA encoding genes (73%) (Fig. 3D). All of these genes co-occur with a single VWA gene, with the exception of the CGN found in *Shewanella denitrificans* OS217, which occurs in proximity to two VWA genes. Approximately 80% of the CGN genes are immediately adjacent to the VWA gene, and, in the case of one CGN found in *Desulfitobacterium hafniense* DCB-2, the VWA and AAA+ are encoded by the same gene. Of these VWA proteins, 86% belong to COG4548 (NorD) (Fig. 3D). Members of this COG are similar to the NorD protein which is associated with the activation of nitric oxide reductase enzyme (see below).

The extreme diversity of this subfamily makes a general gene-structure analysis difficult, and there do not appear to be any genes, other than the VWA, that are strongly associated with members of this branch as a whole. The CGN subfamily is perhaps the best experimentally studied of all of the MoxR subfamilies, however, and some functional information is available. Experimental work has been performed on the NirQ/NorQ-, CbbQ- and GvpN-type members.

3.4.1. NirQ/NorQ-type members

Our analysis detects 22 CGN genes associated with nitric-oxide reductase encoding genes (Fig. 1). Almost all of these genes also appear to co-occur with VWA-encoding genes. This group includes the NirQ/NorQ enzymes, which have been studied experimentally in a variety of organisms. Mutagenesis experiments have been performed in *Pseudomonas stutzeri*, *P. denitrificans*, *Rhodobacter sphaeroides 2.4.3*, and *Pseudomonas aeruginosa*. These genes are found in association with the nitrite reductase/nitric oxide reductase gene clusters and appear to play an important role in bacterial denitrification. Denitrification is a microbial respiratory process involving the use of oxidized nitrogen compounds as alternative electron acceptors. The entire pathway consists of four reduction reaction steps, leading from nitrate to dinitrogen. The four enzymes involved in this process are nitrate reductase, nitrite reductase, nitric oxide reductase, and nitrous oxide reductase (Philippot, 2002; Zumft, 2005) (Fig. 5). Research to date suggests that the NirQ/NorQ proteins may play a role in the activation of the respiratory, short-chain nitric oxide reductase, which acts at the third step in this pathway.

Enzyme activity studies on *P. aeruginosa* and *P. stutzeri* mutants deleted of *nirQ*/*norQ* genes revealed the loss of

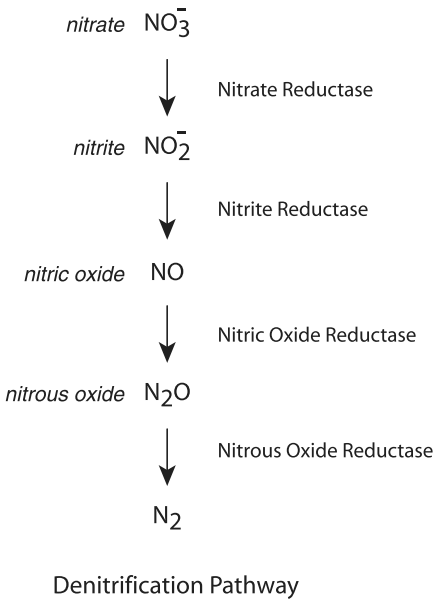

Fig. 5. Pathway of microbial denitrification.

nitric oxide reductase activity, although the induction from the nitric oxide reductase promoter was not compromised (Arai et al., 1999; Jungst and Zumft, 1992). Thus, it appears that the *nirQ/norQ* gene products may be important for regulating the activity of nitric oxide reductase at the enzyme level (Jungst and Zumft, 1992). Deletion or insertion mutagenesis of the *nirQ/norQ* genes in *P. aeruginosa*, *P. denitrificans*, and *R. sphaeroides 2.4.3* resulted in strains incapable of anaerobic growth on minimal medium using nitrate or nitrite as a sole electron acceptor (Arai et al., 1998; Bartnikas et al., 1997; de Boer et al., 1996). In addition, in *P. denitrificans* and *R. sphaeroides*, insertion mutagenesis of the neighboring *norD* genes, which encode VWA proteins, also produced a similar phenotype (Bartnikas et al., 1997; de Boer et al., 1996). Mutant phenotypes could be partially complemented by introduction of plasmids containing the *nirQ/norQ* or *norD* genes (Arai et al., 1999; Bartnikas et al., 1997; de Boer et al., 1996; Jungst and Zumft, 1992).

The *nirQ/norQ* genes are often found closely associated with *nirO/norE* and *nirP/norF* genes, and, in *P. aeruginosa*, the three genes have been shown to comprise an operon (*nirQOP*) (Arai et al., 1994). Our analysis detects *nirO/norE* genes near 16 of the 22 *nirQ/norQ* type CGN genes identified. *nirQ/norE* and *nirP/norF* encode small transmembrane proteins whose exact function is unknown, although NirO is homologs to subunit III of bacterial and mitochondrial cytochrome oxidases (Arai et al., 1994). The exact role of these proteins is unclear, and there have been conflicting reports as to whether or not deletion of these genes produces an effect on nitric oxide reductase activity (Arai et al., 1998; de Boer et al., 1996).

3.4.2. CbbQ-type members

The CbbQ-type proteins are closely related to the NirQ/NorQ proteins. Our study detected 8 CbbQ-type CGN genes (Fig. 1) co-occurring with genes encoding bacterial 1,5-bisphosphate carboxylase/oxygenase (RuBisCO); all but one also co-occur with a VWA gene. The *cbbQ* genes have been shown to be important in the activation of RuBisCO. This enzyme is responsible for the fixation of CO_2, catalyzing the carboxylation of ribulose-1,5-bisphosphate (RuBP) to form two molecules of 3-phosphoglycerate. Two major forms of RuBisCO are found in chemoautotrophic bacteria: form I is a hexadecameric enzyme composed of 8 large and 8 small subunits, and form II is a dimeric, tetrameric, or octameric complex composed of only large subunits (Shively et al., 1998). CbbQ genes have been found in association with both types of RuBisCO genes.

Both Form I and Form II RuBisCO enzymes can be stimulated by CbbQ and its associated VWA domain containing protein, CbbO (Hayashi et al., 1997; Hayashi et al., 1999). Studies have shown that overexpression of *Hydrogenophilus thermoluteolus* CbbQ and/or CbbO with form I RuBisCO from the same organism in an *E. coli* system results in the production of a more active form of RuBisCO than when the enzyme is overexpressed alone. RuBisCO purified from the cells upon co-expression of CbbQ and/or CbbO possessed a V_{max} almost twofold greater than enzyme expressed without CbbQ/O. The K_m of the enzyme for RuBP and Mg^{2+} was also elevated, while the K_m for CO_2 was unchanged. The conformational state of the co-expressed RuBisCO was also observed to be different than RuBisCO expressed alone, as determined using both ANS staining and circular dichroism spectroscopy (Hayashi et al., 1997). Form II RuBisCO from *Hydrogenovibrio marinus* overexpressed with its CbbQ was also observed to undergo activation and a change in conformation (Hayashi et al., 1999).

Although the mechanism of action of CbbQ on RuBisCO enzymes is unclear, the ability of CbbQ to activate both forms of RuBisCO suggests that it interacts with the large, rather than the small subunit (which is not present in form II) (Hayashi et al., 1999). Analysis of purified recombinant CbbQ showed that it is a functional ATPase with activity dependent upon the presence of Mg^{2+} ions (Hayashi and Igarashi, 2002).

3.4.3. GvpN-type members

The GvpN-type CGN MoxR proteins (Fig. 1) have also been studied experimentally and have been shown to have an important role in gas vesicle formation. Our analysis detected 11 CGN protein-encoding genes located in close proximity to gas-vesicle biosynthesis genes (Fig. 3D). None of these are associated with VWA protein-encoding genes. Gas vesicles are intracellular, hollow, gas-filled structures comprised solely of protein. They are found in aquatic microorganisms where they provide buoyancy, allowing cells to control their depth in water, thus optimizing exposure to oxygen and light. The major component of gas vesicles is the small, hydrophobic GvpA protein, which forms the ribs of the main structure. GvpC, a larger hydrophilic protein, is a minor component located on the outer surface of the gas vesicles where it serves to stabilize the entire structure (Walsby, 1994). Mutations in *gvpN* produce varying effects on gas vesicle formation. Insertion mutagenesis of *Halobacterium salinarium gvpN* resulted in cells which formed a large number of gas vesicles of unusually small size compared with WT cells. Removal of the bulk of the insertion cassette did not affect the phenotype, suggesting that it is specific to GvpN and not a result of a polar effect (DasSarma et al., 1994). Different results were obtained in another study using *Haloferax volcanii* cells. When the *gvp* gene cluster was deleted in these cells and the cells were transformed with the *H. salinarium gvp* gene cluster containing a mutant form of *gvpN*, the cells formed a very low number of gas vesicles compared to cells transformed with the WT gene cluster. Northern blotting revealed that the other three genes in the same operon as *gvpN* were expressed at normal levels, indicating that the phenotype was not a result of a polar effect and was not due to alterations in the expression level of these genes (Offner et al., 1996).

Thus, the exact role of GvpN is unclear and may differ slightly between different organisms. Based upon the studies on other MoxR AAA+ proteins, however, it is tempting

to speculate that GvpN may have a general, chaperone-like function, possibly acting on GvpA and/or GvpC to ensure that they are properly folded and successfully assembled into the gas vesicle structure.

3.4.4. DnaJ/BolA associated members

A small subset of CGN genes are found near genes encoding a DnaJ-type molecular chaperone and a BolA-type protein (Fig. 1). These CGN genes, which we have called the DnaJ/BolA associated (DBA) subgroup, are also associated with a VWA-encoding gene. DnaJ has been well-studied in *E. coli*, where it works with the DnaK/Hsp70 molecular chaperone and the nucleotide exchange factor GrpE to promote protein folding and the oligomerization/dissociation of protein complexes (Houry, 2001). No Hsp70 or GrpE encoding genes were detected in any of the 10 genes surrounding CGN. BolA is a transcription factor and has also been studied in *E. coli*, where it has been shown to play a role in cell morphology, cell division, and cell stress response (Santos et al., 2002). Whether the co-occurence of these CGN genes with *dnaJ* and *bolA* genes has any functional significance is unclear, and will need to be determined experimentally. It is interesting to postulate, for example, that the CGN and VWA proteins may work with DnaJ in order to promote the proper folding of substrate proteins.

With the lack of experimental work on any of the remaining CGN members, and the absence of any larger-scale gene conservation, no conclusions as to the function of the remaining members can yet be drawn.

3.5. APE2220 subfamily

The APE2220 subfamily contains 49 of the 596 MoxR proteins included in our analysis (Fig. 1 and Supplementary Table 1). These are found distributed across 31 organisms, including members of the Actinobacteria, Chloroflexi, Deinococcus-Thermus, Firmicutes, and Proteobacteria (Alpha, Beta, and Delta/Epsilon subphyla) phyla, as well as two members of the Archaea superkingdom (Fig. 2).

Examination of the gene structure of this family reveals that 46 of the 49 APE2220 genes are in close proximity (typically adjacent) to genes encoding VWA proteins belonging to the CoxE COG3552 (Fig. 3E). Members of this COG are of unknown function, but their genes are often found as part of a carbon monoxide dehydrogenase (Cox) gene cluster. Indeed, 25 of the 49 APE2220 genes found near these VWA proteins appear to be part of a carbon monoxide/xanthine dehydrogenase-type gene cluster, with the specific gene arrangement and composition varying between organisms. The remaining 21 APE2220 members have a varied genomic environment.

Although no experimental work has been done on members of the APE2220 subfamily, it seems reasonable to assume that those proteins whose genes are found in association with the dehydrogenase cluster are likely to be involved in the biogenesis and/or activation of these enzymes in a manner analogous to that of other MoxR AAA+ proteins. The carbon monoxide and xanthine dehydrogenase enzymes are both members of the molybdenum hydroxylase family, and are dependent upon metal cofactors for their activity (Hille, 2005). It is tempting to speculate that the APE2220 and their associated VWA proteins may be important for proper insertion of these cofactors into these enzymes.

3.6. PA2707 subfamily

Of the 596 MoxR AAA+ proteins used in our analysis, 64 belong to the PA2707 subfamily (Fig. 1 and Supplementary Table 1). These can be found in a total of 52 of the 275 organisms studied, including members of the Actinobacteria, Chlorobi, Cyanobacteria, Proteobacteria (all major subphyla), and Spirochaetes phyla. No archaeal members were detected (Fig. 2).

Forty-five of the 64 PA2707 genes (70%) are located in close proximity to genes encoding VWA proteins (Fig. 3F). The position of the VWA genes with respect to the PA2707 genes varies between organisms, although they are adjacent in about 60% of the cases. No PA2707 gene is in proximity to more than a single VWA encoding gene. All of the VWA proteins are identified as CoxE type, with all but 3 belonging to COG3552 CoxE VWA, similar to the VWA proteins near APE2220 subfamily members. All of the VWA proteins also belong to COG3825, an uncharacterized class of bacterial proteins whose function is not known. Interestingly, and in contrast to members of the APE2220 subfamily, these genes do not appear to be part of a *cox* gene cluster, although a small subset (~23%) do appear to be close to genes encoding a cytochrome *c*. It is interesting to note that *Trichodesmium erythraeum* IMS101, a filamentous cyanobacteria found in nutrient poor tropical and subtropical ocean waters, contains 9 PA2707 subfamily genes with two adjacent to VWA-encoding genes. No research has been performed to date on any of the members of the PA2707 subfamily.

3.7. YehL subfamily

The YehL subfamily is the smallest MoxR subfamily detected in our study. Of the 596 MoxR AAA+ ATPases we identified only 14 (2.3%) as belonging to YehL subfamily (Fig. 1 and Supplementary Table 1). These are distributed across 12 of the 275 organisms in our analysis (Fig. 2). These organisms include members of the Actinobacteria and Proteobacteria (Beta and Gamma subdivisions), as well as single members from the Bacteroidetes, Planctomycetes, and Spirochaetes phyla. A lone archaeal YehL was also identified in *Aeropyrum pernix* K1. The Actinobacteria *Streptomyces avermitilis* MA-4680 and *Streptomyces coelicolor* A3(2) contain two different *yehL* genes.

Examination of *yehL* gene structure (Fig. 3G) reveals that these genes are also found in close proximity to genes encoding VWA proteins, once again suggesting the

possibility that these genes function together. Of the 14 *yehL* genes, only the representative from *Kineococcus radiotolerans* is not in close proximity to a VWA gene. Since the sequencing around the *yehL* gene in this organism is incomplete, however, the possibility that an as yet undetected VWA encoding gene may be nearby cannot be dismissed.

Twelve of the 14 *yehL* genes are also adjacent to a moderately conserved gene encoding a long protein of unknown function. The product of this gene, which we will refer to as *yehM* after the *E. coli* K12 homologue, contains no recognizable domains or motifs. The *yehM* gene is not detected in *A. pernix* and *K. radiotolerans*, although once again the latter may be due to incomplete sequencing of the gene region. In *R. baltica SH1*, the *yehM* and VWA genes are fused together, encoding a predicted product of ~1300 amino acids in length with a C-terminal VWA domain. The same is observed for one of the two *yehM* genes detected in *S. coelicolor* A3(2). In all but one of the remaining cases, the *yehM* gene is located between the *yehL* and VWA genes. In all of these cases it encodes a protein of ~700–800 amino acids in length.

The conserved association of the *yehL*, *yehM*, and VWA encoding genes strongly suggests that they may function together. To date, no experimental work has been performed on members of the YehL MoxR subfamily, so information about their roles in the cell is not available. The extremely limited distribution of these proteins, however, may suggest a highly specialized role, perhaps acting as molecular chaperones for a single or very narrow range of substrate proteins.

4. Conclusions

The large size and remarkable diversity of the MoxR AAA+ family emphasizes the importance of this particular class of proteins. Surprisingly, however, relatively little is known about these enzymes and the specific functions of the MoxR AAA+ proteins remain elusive. Our phylogenetic analysis has identified at least seven distinct MoxR AAA+ subfamilies. Analysis of adjacent genes reveals distinctly different patterns of neighboring genes for each of these subfamilies, although one major commonality between all of the subfamilies is a tendency to be near genes encoding proteins containing VWA domains.

Analysis of the limited experimental research available on MoxR AAA+ suggests a role in the assembly and activation of protein complexes. Exactly, how these proteins function is not clear, although a role in metal insertion seems a reasonable hypothesis. The similarity between the MoxR AAA+ proteins and the metal chelatase enzymes is notable, particularly in light of the fact that the latter enzymes utilize a VWA in conjunction with a AAA+ module to carry out the insertion of Co^{2+} and Mg^{2+} into porphyrin rings (Fodje et al., 2001). MoxR AAA+ proteins in conjunction with VWA domain-containing proteins might insert metals directly into proteins rather than into porphyrin rings or other related structures. This is supported by the experimental data demonstrating that the activities of MDH, NO reductases, and RuBisCO are dependent on the presence of metal cofactors, and that these activities are disrupted or reduced upon deletion of the respective MoxR AAA+ and/or VWA associated proteins. Also, many members of the APE2220 AAA+ subfamily appear to associate with carbon monoxide dehydrogenase/xanthine dehydrogenase-type gene clusters, which encode enzymes dependent upon metal cofactors for their function.

The possibility that some MoxR proteins may have different chaperone-like functions not involving metal insertion must also be considered. Gas vesicles, whose assembly appears to be dependent upon the CGN-type MoxR AAA+ protein, GvpN, are not reported to incorporate metal ions into their structure. Notably, none of the GvpN-encoding genes appears to occur in proximity to VWA-encoding gene. Thus, it is possible that MoxR-type proteins not found in association with VWA proteins may not play a role in metal insertion events, but rather have different chaperone type functions.

Considering the size and diversity of the MoxR AAA+ protein family, it is remarkable that so little is known about it. Continued research in this area is essential and is sure to provide fascinating insights into the function of what appears to be a novel class of molecular chaperones.

Acknowledgments

J.S. is the recipient of a graduate fellowship from the Natural Sciences and Engineering Research Council of Canada. W.A.H. is Canadian Institutes of Health Research New Investigator. This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada to W.A.H.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2006.02.009.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Amaratunga, K., Goodwin, P.M., O'Connor, C.D., Anthony, C., 1997. The methanol oxidation genes mxaFJGIR (S) ACKLD in *Methylobacterium extorquens*. FEMS Microbiol. Lett. 146, 31–38.

Anthony, C., 2004. The quinoprotein dehydrogenases for methanol and glucose. Arch. Biochem. Biophys. 428, 2–9.

Arai, H., Igarashi, Y., Kodama, T., 1994. Structure and ANR-dependent transcription of the nir genes for denitrification from *Pseudomonas aeruginosa*. Biosci. Biotechnol. Biochem. 58, 1286–1291.

Arai, H., Kodama, T., Igarashi, Y., 1998. The role of the nirQOP genes in energy conservation during anaerobic growth of *Pseudomonas aeruginosa*. Biosci. Biotechnol. Biochem. 62, 1995–1999.

Arai, H., Kodama, T., Igarashi, Y., 1999. Effect of nitrogen oxides on expression of the nir and nor genes for denitrification in *Pseudomonas aeruginosa*. FEMS Microbiol. Lett. 170, 19–24.

Bartnikas, T.B., Tosques, I.E., Laratta, W.P., Shi, J., Shapleigh, J.P., 1997. Characterization of the nitric oxide reductase-encoding region in *Rhodobacter sphaeroides 2.4.3*. J. Bacteriol. 179, 3534–3540.

Blatch, G.L., Lassle, M., 1999. The tetratricopeptide repeat: a structural motif mediating protein–protein interactions. Bioessays 21, 932–939.

DasSarma, S., Arora, P., Lin, F., Molinari, E., Yin, L.R., 1994. Wild-type gas vesicle formation requires at least ten genes in the gvp gene cluster of *Halobacterium halobium* plasmid pNRC100. J. Bacteriol. 176, 7646–7652.

de Boer, A.P., van der Oost, J., Reijnders, W.N., Westerhoff, H.V., Stouthamer, A.H., van Spanning, R.J., 1996. Mutational analysis of the nor gene cluster which encodes nitric-oxide reductase from *Paracoccus denitrificans*. Eur. J. Biochem. 242, 592–600.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Felsenstein, J., 1996. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods Enzymol. 266, 418–427.

Fodje, M.N., Hansson, A., Hansson, M., Olsen, J.G., Gough, S., Willows, R.D., Al-Karadaghi, S., 2001. Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. J. Mol. Biol. 311, 111–122.

Frickey, T., Lupas, A.N., 2004. Phylogenetic analysis of AAA proteins. J. Struct. Biol. 146, 2–10.

Hanson, P.I., Whiteheart, S.W., 2005. AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. 6, 519–529.

Hayashi, N.R., Arai, H., Kodama, T., Igarashi, Y., 1997. The novel genes, cbbQ and cbbO, located downstream from the RubisCO genes of *Pseudomonas hydrogenothermophila*, affect the conformational states and activity of RubisCO. Biochem. Biophys. Res. Commun. 241, 565–569.

Hayashi, N.R., Arai, H., Kodama, T., Igarashi, Y., 1999. The cbbQ genes, located downstream of the form I and form II RubisCO genes, affect the activity of both RubisCOs. Biochem. Biophys. Res. Commun. 265, 177–183.

Hayashi, N.R., Igarashi, Y., 2002. ATP binding and hydrolysis and autophosphorylation of CbbQ encoded by the gene located downstream of RubisCO genes. Biochem. Biophys. Res. Commun. 290, 1434–1440.

Hille, R., 2005. Molybdenum-containing hydroxylases. Arch. Biochem. Biophys. 433, 107–116.

Houry, W.A., 2001. Chaperone-assisted protein folding in the cell cytoplasm. Curr. Protein Pept. Sci. 2, 227–244.

Iyer, L.M., Leipe, D.D., Koonin, E.V., Aravind, L., 2004. Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. 146, 11–31.

Jungst, A., Zumft, W.G., 1992. Interdependence of respiratory NO reduction and nitrite reduction revealed by mutagenesis of nirQ, a novel gene in the denitrification gene cluster of *Pseudomonas stutzeri*. FEBS Lett. 314, 308–314.

Kachlany, S.C., Planet, P.J., Bhattacharjee, M.K., Kollia, E., DeSalle, R., Fine, D.H., Figurski, D.H., 2000. Nonspecific adherence by Actinobacillus actinomycetemcomitans requires genes widespread in bacteria and archaea. J. Bacteriol. 182, 6169–6176.

Katerov, V., Lindgren, P.E., Totolian, A.A., Schalen, C., 2000. Streptococcal opacity factor: a family of bifunctional proteins with lipoproteinase and fibronectin-binding activities. Curr. Microbiol. 40, 149–156.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Marchler, G.H., Mullokandov, M., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Yamashita, R.A., Yin, J.J., Zhang, D., Bryant, S.H., 2005. CDD: a Conserved Domain Database for protein classification. Nucleic Acids Res. 33, D192–D196.

Merrell, D.S., Camilli, A., 1999. The cadA gene of *Vibrio cholerae* is induced during infection and plays a role in acid tolerance. Mol. Microbiol. 34, 836–849.

Neuwald, A.F., Aravind, L., Spouge, J.L., Koonin, E.V., 1999. AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. 9, 27–43.

Offner, S., Wanner, G., Pfeifer, F., 1996. Functional studies of the gvpACNO operon of *Halobacterium salinarium* reveal that the GvpC protein shapes gas vesicles. J. Bacteriol. 178, 2071–2078.

Ogura, T., Wilkinson, A.J., 2001. AAA+ superfamily ATPases: common structure–diverse function. Genes Cells 6, 575–597.

Park, Y.K., Bearson, B., Bang, S.H., Bang, I.S., Foster, J.W., 1996. Internal pH crisis, lysine decarboxylase and the acid tolerance response of *Salmonella typhimurium*. Mol. Microbiol. 20, 605–611.

Philippot, L., 2002. Denitrifying genes in bacterial and Archaeal genomes. Biochim. Biophys. Acta 1577, 355–376.

Richardson, I.W., Anthony, C., 1992. Characterization of mutant forms of the quinoprotein methanol dehydrogenase lacking an essential calcium ion. Biochem. J. 287 (Pt. 3), 709–715.

Sabo, D.L., Boeker, E.A., Byers, B., Waron, H., Fischer, E.H., 1974. Purification and physical properties of inducible *Escherichia coli* lysine decarboxylase. Biochemistry 13, 662–670.

Sabo, D.L., Fischer, E.H., 1974. Chemical properties of *Escherichia coli* lysine decarboxylase including a segment of its pyridoxal 5′-phosphate binding site. Biochemistry 13, 670–676.

Santos, J.M., Lobo, M., Matos, A.P., De Pedro, M.A., Arraiano, C.M., 2002. The gene bolA regulates dacA (PBP5), dacC (PBP6), and ampC (AmpC), promoting normal morphology in *Escherichia coli*. Mol. Microbiol. 45, 1729–1740.

Shively, J.M., van Keulen, G., Meijer, W.G., 1998. Something from almost nothing: carbon dioxide fixation in chemoautotrophs. Annu. Rev. Microbiol. 52, 191–230.

Snider, J., Gutsche, I., Lin, M., Baby, S., Cox, B., Butland, G., Greenblatt, J., Emili, A., Houry, W.A., 2006. Formation of a distinctive complex between the inducible bacterial lysine decarboxylase and a novel AAA+ ATPase. J. Biol. Chem. 281, 1532–1546.

Soksawatmaekhin, W., Kuraishi, A., Sakata, K., Kashiwagi, K., Igarashi, K., 2004. Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. Mol. Microbiol. 51, 1401–1412.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28, 33–36.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Toyama, H., Anthony, C., Lidstrom, M.E., 1998. Construction of insertion and deletion mxa mutants of *Methylobacterium extorquens* AM1 by electroporation. FEMS Microbiol. Lett. 166, 1–7.

Van Spanning, R.J., Wansell, C.W., De Boer, T., Hazelaar, M.J., Anazawa, H., Harms, N., Oltmann, L.F., Stouthamer, A.H., 1991. Isolation and characterization of the moxJ, moxG, moxI, and moxR genes of *Paracoccus denitrificans*: inactivation of moxJ, moxG, and moxR and the resultant effect on methylotrophic growth. J. Bacteriol. 173, 6948–6961.

Veerassamy, S., Smith, A., Tillier, E.R., 2003. A transition probability model for amino acid substitutions from blocks. J. Comput. Biol. 10, 997–1010.

Walsby, A.E., 1994. Gas vesicles. Microbiol. Rev. 58, 94–144.

Whittaker, C.A., Hynes, R.O., 2002. Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. Mol. Biol. Cell 13, 3369–3387.

Willows, R.D., 2003. Biosynthesis of chlorophylls from protoporphyrin IX. Nat. Prod. Rep. 20, 327–341.

Xiong, J.P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S.L., Arnaout, M.A., 2002. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. Science 296, 151–155.

Zumft, W.G., 2005. Nitric oxide reductases of prokaryotes with emphasis on the respiratory, heme-copper oxidase type. J. Inorg. Biochem. 99, 194–215.