Polish Journal of Microbiology
2009, Vol. 58, No 2, 181–184

SHORT COMMUNICATION

AAN82231 Protein from Uropathogenic *E. coli* CFT073 Is a Close Paralog of DsbB Enzymes and Does Not Belong to the DsbI Family

MARCIN PAWŁOWSKI[1, 2,*], ANNA M. ŁASICA[3], ELŻBIETA K. JAGUSZTYN-KRYNICKA[3] and JANUSZ M. BUJNICKI[1,4]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, Poland
[2] Institute of Biochemistry and Biophysics PAS, Warsaw, Poland
[3] Department of Bacterial Genetics, Institute of Microbiology, University of Warsaw, Poland
[4] Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznań, Poland

Received 12 April 2009, accepted 29 April 2009

Abstract

Dsb proteins control the formation and rearrangement of disulfide bonds during the folding of membrane and exported proteins. DsbA is an oxidant that catalyzes formation of disulfide bonds in newly synthesized, and yet unfolded proteins. In order to act catalytically again, it has to be reoxidized by a transmembrane protein DsbB characterized by two pairs of disulfides. DsbB is related to another protein family DsbI, characterized by the presence of only one disulfide, and an additional C-terminal beta-propeller domain. The protein AAN82231 from *E. coli* strain CFT073 has been recently described as a new member of the DsbI family (Grimshaw *et al.*, 2008). It was found that AAN82231 forms a functional redox pair with DsbL – a newly described DsbA-like protein. Here, we report that AAN82231 shares no characteristic features with the DsbI proteins. Instead, according to phylogenetic analyses AAN82231 clearly belongs to another, previously described subfamily of DsbB paralogs. To facilitate classification of DsbB and DsbI homologs, we propose a new nomenclature system and present an updated phylogenetic analysis of the DsbB superfamily, which comprises the following families: "orthodox" DsbB, its paralogs now named DsbB2 (including AAN82231), DsbI and two groups of so far uncharacterized DsbB paralogs termed DsbB3 and DsbB4. We have also developed a web server dedicated to phylogenetic assignment of DsbB/DsbI candidate proteins that will be identified in the future.

Key words: Disulfide bonds, DsbB superfamily, DsbI family, oxidoreductase, propeller

Many periplasmic proteins of Gram-negative bacteria contain disulfide bridges that stabilize the protein structures. In *Escherichia coli* this process is facilitated by a Dsb (disulfide bond) family of the redox proteins (reviewed in: Ito and Inaba, 2008; Nakamoto and Bardwell, 2004). DsbA is the primary disulfide bond donor in the periplasmic space. DsbA is reactivated by DsbB, which oxidizes the Cys30 and the Cys33 active-site residues of DsbA as they become reduced upon substrate oxidation (Missiakas *et al.*, 1993). DsbB is integrated into the cytoplasmic membrane by its four transmembrane segments (TM1-TM4). Each of its two periplasmic regions (P1 and P2) contains a pair of cysteines, Cys41 and Cys44 in P1 and Cys104 and Cys130 in P2, that are essential for the DsbB function. Recently, we characterized a new family of proteins related to the DsbB family, and named them DsbI (Raczko *et al.*, 2005). The N-terminal domain of DsbI is related to DsbB proteins. The HHsearch (Soding, 2005) analysis carried out for DsbI members confirmed their relationship to Pfam 02600 (DsbB) (Pvalue 2.9E-32) however N-terminal domain of DsbI comprises five predicted transmembrane segments, while the C-terminal domain is predicted to locate in the periplasm and to fold into a beta-propeller structure. Another key difference between DsbI and "classical" DsbB proteins is the lack of the second pair of Cys residues in sequence of DsbI subfamily members. For convenience, we will refer to the TM domain of DsbB and DsbI (regardless

* Corresponding author: M. Pawlowski, Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, Trojdena 4, 02-109, Warsaw, Poland; phone: (+48) 22 5970750; fax: (+48) 22 5970715; **e-mail: marcinp@genesilico.pl**

of the number of the helices) as members of the DsbB superfamily.

DsbI has been found to be important for pathogenicity of *Helicobacter pylori* (Godlewska *et al.*, 2006). However, thus far no DsbA-like partner has been identified for DsbI. Recently, protein AAN82231 from a highly virulent *E. coli* strain CFT073 has been found to form a functional redox pair with a newly described protein DsbL, the most oxidizing thioredoxin-like protein known to date. AAN82231 is a homolog of DsbB and has been classified into the DsbI by the Glockshuber group (Grimshaw *et al.*, 2008). However, we found that this assignment is clearly incorrect and in fact, AAN82231 belongs to the DsbB gamma 2 subfamily presented in our earlier analysis (Raczko *et al.*, 2005). Firstly, AAN82231 and its closest homologs contain two pairs of Cys (P1 and P2), like DsbB and not just one like DsbI. Moreover, according to several methods for prediction of TM helices, *e.g.* those run *via* the ConPredII metaserver (Arai *et al.*, 2004), AAN82231 contains four TM segments rather than five expected for a member of the DsbI family.

Phylogenetic relationships within the DsbB superfamily. In order to accurately determine the phylogenetic relationship between "classical" DsbB, AAN82231, DsbI and other homologs of DsbB, we have carried out database searches with PSI-BLAST (Altschul *et al.*, 1997) to identify all homologs of DsbB and DsbI proteins. Like previously reported (Kimball *et al.*, 2003; Raczko *et al.*, 2005), most members of the DsbB superfamily are found in *Proteobacteria*, suggesting that the DsbB family evolved in an ancient *Proteobacterium* and that other organisms acquired copies of DsbB-like genes by horizontal gene transfers. Other taxons include *Firmicutes*, *Chlamydiae* and even *Archaea*, *e.g. Halorubrum lacusprofundi* ATCC 49239, *Haloarcula marismortui* ATCC 43049, and *Natronomonas pharaonis* DSM 2160. To our knowledge it is the first report of archaeal members of the DsbB superfamily. We have constructed a multiple sequence alignment for representative sequences related by <30% identity (data not shown). Figure 1 presents the Maximum Likelihood phylogenic tree of the DsbB superfamily built using PHYML (Guindon and Gascuel, 2003), with the WAG model of amino acid substitution. The significance of individual branches has been assessed using the parametric approximate likelihood ratio test (aLRT). This tree is similar to the tree shown in our original publication describing DsbI (Raczko *et al.*, 2005) however, it contains more sequences and its branches have better statistical support, thus enabling definite classification of DsbB-like proteins into well-defined families.

The topology of the DsbB-like protein tree reveals several well-resolved branches. The branch containing the classical DsbB protein from *E. coli* and its gammaproteobacterial orthologs, as well as single branches from *Alpha-* and *Betaproteobacteria* conform to the established organismal phylogeny, suggesting that they have been derived from regular vertical descent from the ancient *Proteobacterium*. The gammaproteobacterial lineage containing the classical DsbB has a sister lineage with members from partially overlapping set of species. We have previously dubbed these branches gamma 1 and gamma 2 (Raczko *et al.*, 2005). The topology of both branches agrees with the organismal phylogeny, with the only deviation resulting from certain species missing from one branch or the other. For instance, *E. coli* K12 contains only the classical DsbB, while *E. coli* CFT073 contains members of both lineages, namely an ortholog of the classical DsbB and AAN82231. This strongly suggests that these two lineages are products of a gene duplication that took place in an ancestor of *Gammaproteobacteria*. In all other branches, including DsbI, the topology of protein tree does not agree with the species tree, suggesting that they all have been created by horizontal gene transfers and further duplications at the later stage of evolution of the DsbB superfamily. A corollary of this analysis is that the protein AAN82231 from *E. coli* CFT073 belongs to the gamma 2 branch of DsbB proteins, and not to the DsbI family as suggested by the Glockshuber and coworkers (Grimshaw *et al.*, 2008). The name DsbI, as we originally proposed, should be reserved to members of the DsbB superfamily that possess five TM helices in the catalytic domain and a b-propeller domain in the C-terminus, and to their clear orthologs, as determined by phylogenetic analysis. According to our analysis there are no members of the DsbI family in any strain of *E. coli*. Thus, there is no evidence to support the statement contained in the title of the article by the Glockshuber and coworkers (Grimshaw *et al.*, 2008) *i.e.* that "DsbL and DsbI form a specific dithiol oxidase system for periplasmic arylsulfate sulfotransferase in uropathogenic *Escherichia coli*". DsbL is not a partner of DsbI, but of a close paralog of DsbB from a different family.

New nomenclature and a new web server dedicated to classification of DsbB/DsbI proteins. To avoid nomenclatural problems in the future, we suggest to rename members of the "gamma 2" branch of DsbB (including AAN82231) as DsbB2. We also suggest to use (at least provisionally) the name DsbB3 for the uncharacterized family of DsbB paralogs previously dubbed "gamma 3", whose characteristic feature is the presence of the 2nd Cys-Cys pair in the C-terminal tail rather than in the periplasmic loop. Finally, the remaining family of DsbB homologs that appear related to DsbI but lack the C-terminal b-propeller domain and possess the 2nd Cys-Cys pair in the periplasmic loop, may be provisionally referred to as

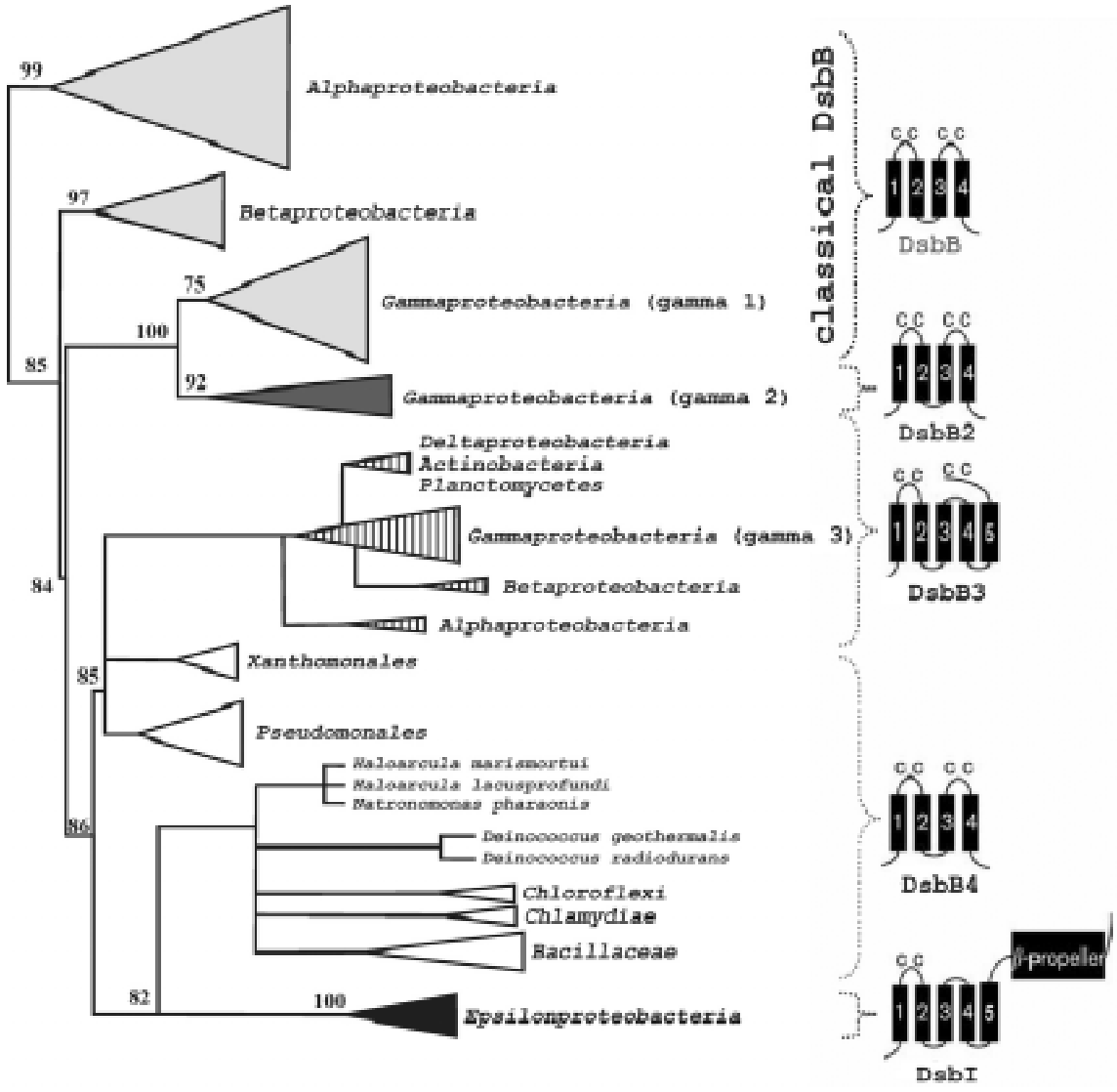

Fig. 1. Phylogenetic tree of the DsbB superfamily, with families shown in different patterns. The localization of AAN82231 from *E. coli* CFT073 is included in the dark grey triangle representing DsbB2 – gamma 2 lineage. "Compressed" branches are shown as triangles with the names of taxa indicated. Numbers at the nodes describe the statistical support (in percent) according to the aLRT test. Nodes with support <70% are regarded as unresolved and were grouped together with their sister lineages. The right panel illustrates the typical topology characteristic for each lineage, including the number of TM segments and the position of periplasmic Cys residues. Noteworthy, majority of DsbB2 proteins contains 4 TM segments however, DsbB2 members from *Campylobacter* have 5 TM segments.

DsbB4, at least until their function is determined experimentally. To facilitate the classification of DsbB and DsbI homologs in the future, we developed a simple web server available at the URL: http://iimcb.genesilico.pl/mp/DsbBI/index.py. The query sequence submitted by the user is compared to hidden Markov model profiles of DsbB/DsbI subfamilies with the HHsearch method (Soding, 2005). A ranking of "hits" is displayed, and the most likely localization of the sequence is indicated on the phylogenetic tree. This website provides also the multiple sequence alignment and the phylogenetic tree of the DsbB superfamily. We hope this resource will be useful for the community of researchers involved in analyses of proteins from the Dsb pathway and it will stimulate experimental characterization of the currently "hypothetical" members and identification of redox partners for members of branches other than DsbB (DsbA) and DsbB2 (DsbL).

Acknowledgements

This work was suppored by Polish Ministry of Science and Higher Education grant (N301 10632/3600) and by Iowa NIH grant (430-46-22).

Literature

Altschul S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.

Nucleic Acids Res. 25: 3389–3402.

Arai M., H. Mitsuke, M. Ikeda, J.X. Xia, T. Kikuchi, M. Satake and T. Shimizu. 2004. ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Res.* 32: W390–393.

Godlewska R., A. Dzwonek, M. Mikula, J. Ostrowski, M. Pawlowski, J.M. Bujnicki and E.K. Jagusztyn-Krynicka. 2006. *Helicobacter pylori* protein oxidation influences the colonization process. *Int. J. Med. Microbiol.* 296: 321–324.

Grimshaw J.P., C.U. Stirnimann, M.S. Brozzo, G. Malojcic, M.G. Grutter, G. Capitani and R. Glockshuber. 2008. DsbL and DsbI form a specific dithiol oxidase system for periplasmic arylsulfate sulfotransferase in uropathogenic *Escherichia coli*. *J. Mol. Biol.* 380: 667–680.

Guindon S. and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52: 696–704.

Ito K. and K. Inaba. 2008. The disulfide bond formation (Dsb) system. *Curr. Opin. Struct. Biol.* 18: 450–458.

Kimball R.A., L. Martin and M.H. Saier, Jr. 2003. Reversing transmembrane electron flow: the DsbD and DsbB protein families. *J. Mol. Microbiol. Biotechnol.* 5: 133–149.

Missiakas D., C. Georgopoulos and S. Raina. 1993. Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds *in vivo*. *Proc. Natl. Acad. Sci. U S A* 90: 7084–7088.

Nakamoto H. and J.C. Bardwell. 2004. Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim. Biophys. Acta* 1694: 111–119.

Raczko A.M., J.M. Bujnicki, M. Pawlowski, R. Godlewska, M. Lewandowska and E.K. Jagusztyn-Krynicka. 2005. Characterization of new DsbB-like thiol-oxidoreductases of *Campylobacter jejuni* and *Helicobacter pylori* and classification of the DsbB family based on phylogenomic, structural and functional criteria. *Microbiology* 151: 219–231.

Soding J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21: 951–960.