ISSN 0974 - 9373

Vol. 15(2011) Special Issue 2 **Journal of International Academy of Physical Sciences** pp. 501-506

In Silico Genome Analysis of Gammaproteobacteria with Reference to Metal Binding Sites*

Navaneet Chaturvedi
Center of Bioinformatics, University of Allahabad, India
Email: bioinfonavneet@gmail.com

P. N. Pandey
Department of Mathematics, University of Allahabad, India
Email: pnpiaps@gmail.com

(Received June 14, 2011)

Abstract: Identification of functional residues in proteins and their functional annotation has been addressed by several workers. However, proteome and genome analysis of proteobacteria involved in metal resistance is not yet explored. Genome analysis and study of metal binding proteins from Gamma-proteobacteria give us opportunities to understand the role of Gamma-proteobacteria in Bioremediation. In this study we analyze the genes containing metal-binding proteins and peptides pattern using appropriate tools. We have found the protein family of *C. sakazakii* BAA-894 (Class:Gammaproteobacteria) chromosomal genome encoded proteins and involvement of these families in metal interaction by specific pattern search.

Keyword: Metal Resistance, Gamma-proteobacteria, *Cronobacter sakazakii*

1. Introduction

Recently it is proved that many heavy metals constitute a global environmental hazard. Microorganisms especially bacteria could be used to clean up metal contaminations by removing metals from contaminated water and waste streams, sequestering metals to facilitate their extraction. Basically bacteria from class gammaproteobacteria and higher organisms have developed resistance mechanisms to toxic metals to make them innocuous. *Cronobacter sakazakii* (previously known as *Enterobacter sakazakii*) is gram negative proteobacteria, rods shaped from *Enterobacteriaceae* family. *Cronobacter* is widespread in both the environment and in plant material, including water, soil and a variety of processed foods and fresh produce[1, 2]. *Cronobacter sakazakii BAA-894* is

*Paper presented in CONIAPS XIII at UPES, Dehradun during June 14-16, 2011.

well known as food pathogen and for neonatal infection but recent research has been reported about its some sort of metal resistive nature and the signal of the identification of metal binding nature[2]. Earlier many workers had studied reports on the significant importance of cadmium (Cd) and arsenic (As) heavy metals. Even though Cd is frequently recruited by many agricultural crops and fruits and because of its presence in many food substances, the cadmium metal was mainly concerned with bioremediation and phytoremediation studies[3]**.** Moreover, the metabolism of heavy metals has been reported too and the direction of transport of heavy metals in bacteria has been suggested, like efflux direction, uptake direction and both side direction[4]. Therefore several bacteria have exhibited mechanism to tolerate the toxicity of heavy metal ions. One of the study on direction of transport include efflux of metal ions outside the cell, uptake or both phenomenon of the metal ions inside the cell and reduction of heavy metal ions to a less toxic state[5]. For the proper bioremediation, the metal resistant bacteria have been shown to have great potential to resist the toxicity. The study of plasmid metal resistance operand flanked by mobile genetic elements and the identification of these plasmid genomic islands has been reported as well[6]**.** Most resistance systems are based on the energy-dependent efflux of toxic metal ions[7].

Previously there is couple of study have addressed by several workers about the metal resistive capacity of *Enterobacterales* family[8]. In this work it was found that the several protein classes of *Cronobacter sakazakii* BAA-894 strain have involved in metal binding. The chromosomal genome *Cronobacter sakazakii BAA-894* was taken as the study of interest.

2. Material and Method

The chromosomal genome sequences of *C. sakazakii* BAA-894 were downloaded from Genbank genome database of NCBI[9] and in addition the protein table was considered for classifying the similar protein class by text search. Mostly proteins sequences are characterized as hypothetical proteins. BLASTp[10] was applied to know the similar protein sequence for selected hypothetical protein sequences of corresponding genes. Metal binding residues were obtained from metal mine database[11]. The entire binding pattern was identified by pattern search method. *C. sakazakii* BAA-894 protein family have identified by the searching of the specific residue pattern for corresponding metals. The related protein crystal structure (pdb-id) of corresponding metal was also taken into account. 17 different protein families were identified as a metal binding interaction.

3. Results and Discussion

The protein table (Accession number NC_009778) was recruited as chromosomal genome encoded proteome and on the basis of text search 17 protein families were identified. Table1 shows the name of several protein classes of *C. sakazakii* BAA-894. Additionally these families were contained different protein. The metalmine is a database for metal ion binding proteins, was helpful for obtaining amino acids binding pattern for different metal ions. The table 2 shows name of metals and its interaction with corresponding protein. The binding pattern was considered from metalmine databases.

Table 1: Identified Protein classes of chromosomal genome of C. sakazakii BAA-894

Serial no.	**Proteins Family/Class**	**No. of Proteins**
1	50S ribosomal proteins	27
2	30S ribosomal proteins	16
3	hypothetical proteins	71
4	Translocase proteins	11
5	Repressor proteins	18
6	Transcription Regulator proteins	52
7	Glutathion proteins	12
8	RNA polymerase proteins	11
9	Membrane proteins	18
10	Transcription factor proteins	09
11	Synthetase proteins	46
12	Oxidoreductase proteins	08
13	Reductase proteins	42
14	Dehydrogenase proteins	15
15	transporting ATPase proteins	06
16	Binding proteins	33
17	Regulatory proteins	10

The metal binding pattern residues of specific metal, shown in table2, were taken from metalmine database. Computational pattern search was applied to know the particular identified protein family and scripting on program on MATLAB[12] was also applied to search these specific patterns for a particular metal. It was made to know the location of these patterns in sequences of *C. sakazakii* BAA-894 for every identified classes of protein (Table 1). The table 2 was drawn to show the detail of particular metals have potentially able to bind the sequence of specific identified family of chromosomal genome of *C. sakazakii* BAA-894. The related crystal

structure of proteins was also used to validate the study. Metals ion, like Co, Ni, Mn, Fe, Mo and Cu were involved to interact with sequences. The identified protein families like, reductase, regulatory, repressor and dehydrogenase, were mostly involved in metal binding.

Table 2: Identified *C. sakazakii* BAA-894 protein family having pattern for corresponding metal

S.No.	Metal Name	Binding Pattern	Proteins name	Family	Realated pdb-id
1	Co	QHH	Transcriptional repressor MprA	Repressor	1h41
		QHH	transcriptional regulator MalT	Regulatory	1kej
		DHH	peptidoglycan synthetase	Synthetase	1rqb
		DHH	Ribonucleotide-diphosphate reductase subunit alpha	Reductase	1rqb
2	Ni	HEHH	Glutamyl-tRNA synthetase	Synthetase	1j5y
		HHC	Transcriptional repressor MprA	Repressor	2bj8
		HHHHD	hydroxyacylglutathione hydrolase	Hydrolase	1a5n
		CHH	NADH dehydrogenase subunit L	Dehydrogenase	1bs6
3	Mn	HHDH	hydroxyacylglutathione hydrolase	Glutathion	1ap5
		EHS	phenylalanyl-tRNA synthetase subunit beta	Synthetase	1cpo
		EHS	putative transcriptional regulator	Regulatory	-
		AEEE	phosphoadenosine phosphosulfate reductase	Reductase	1mqw
		AEEE	succinyl-CoA synthetase subunit alpha	Synthetase	1mqw
		QEMN	hypothetical protein ESA_00230 (putative methyl-accepting chemotaxis protein II)	Hypothetical	1a9x
		GDDD	dihydrofolate reductase	Reductase	1a6q
			hypothetical protein ESA_03724 (Sec-independent protein translocase protein tatA)	Hypothetical	
4	Fe	HH	30S ribosomal protein S15	30S ribosomal	1a00
			RNA-binding protein Hfq	Binding protein	1a00
			GTP-binding protein Era	Binding protein	2gtl
			ESA_00380 (putative quinol monooxygenase ygiN)	Hypothetical	1ew6
			dihydrodipicolinate reductase	Reductase	1d8u
		HM	aspartate-semialdehyde dehydrogenase	Dehydrogenase	1dw1
			aspartate-semialdehyde dehydrogenase	Dehydrogenase	1etp
			putative inner membrane protein	Membranous	1dw1
			outer membrane protein	Membranous	1jju
			ribonucleotide-diphosphate reductase subunit beta	Reductase	1bcc
			N-ethylmaleimide reductase	Reductase	1kv9
			DNA-binding transcriptional repressor FabR	Repressor	1h31
5	Mo	HC	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase	Dehydrogenase	1g20
			hypothetical protein ESA_00345 (Putative zeaxanthin glucosyl transferase CRTX)	Hypothetical	1g20

			hypothetical protein ESA_00507 (Puataive L-serine deaminase II)	Hypothetical	-
			fumarate reductase flavoprotein subunit	Reductase	-
			chemotaxis regulatory protein CheY	Regulatory	-
			PTS system transporter subunit IIA-like nitrogen-regulatory protein PtsN	Regulatory	
	MoSO2		transcriptional repressor MprA	Repressor	1t3q
6	Cu	HCHM	glutamine synthetase	Synthetase	2hx7
		HDH	hydroxyacylglutathione hydrolase	Glutathion	1s4c
			zinc uptake transcriptional repressor	Repressor	

4. Conclusion

C. sakazakii is a well known food pathogen bacteria but most of the protein family were taken part in metal binding. All metal binding sequences were accessed from protein table of Genbank database, indicated that they are all specific to the protein family. Specific heavy metal binding interaction with protein sequences is helpful for protein annotation on the functional basis. Therefore we have concluded the insight about the *C. sakazakii* bacteria that it could also be advantageous for the environmental care and for bioremediation strategy. Moreover the detailed metal binding affinities with given proteins will also be beneficial for the further study and the predicted putative metals are yet to be confirmed at a laboratory scale.

References

1. Drudy, D., Mullane, N.R., Quinn, T., Wall, P.G. and Fanning, S. Enterobacter sakazakii: an emerging pathogen in powdered infant formula. Clinical Infectious Diseases, **42** (2006) 996–1002
2. Kucerova E, Clifton SW, Xia X-Q, Long F, Porwollik S, et al. (2010) Genome Sequence of Cronobacter sakazakii BAA-894 and Comparative Genomic Hybridization Analysis with Other Cronobacter Species. *PLoS ONE* 5**(3)**.
3. Mejáre M, Bülow L., Metal-binding proteins and peptides in bioremediation and phytoremediation of heavy metals, *Trends Biotechnol.* 2001 Feb;19 **(2)**:67-73.
4. Thilakaraj R, Raghunathan K, Anishetty S, Pennathur G. , In silico identification of putative metal binding motifs, *Oxford University Press,* **(23)**3 2007, 267–271.
5. Nies DH., Microbial heavy-metal resistance, *Appl Microbiol Biotechnol* (1999) 51: 730-750.
6. Mergea M., Plasmids pMOL28 and pMOL30 of Cupriavidus metallidurans Are Specialized in the Maximal Viable Response to Heavy Metals, (2007), *Journal of Bacteriology*, 189 **(20)**: 7417-7425,

7. Silver (1996), Bacterial resistance to toxic metal ions-a review**,** *Gene,* **(179)** 9-19.
8. Garbisu C, Alkorta I, Carlson DE, Leighton T, Buchanan BB. (1997) Selenite bioremediation potential of indigenous microorganisms from industrial activated sludge. *Microbiologia* 13(**4**):437-44
9. National Center for Bioitechnology Information, http://www.ncbi.nlm.nih.gov
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215 **(3)**: 403–410
11. Metalmine database http://metalmine.naist.jp/
12. "MATLAB Online Documentation". www.mathworks.com , Retrieved 2010-06-07.