International Research Journal of Biological Sciences _____ ISSN 2278-3202
Vol. **1(6),** 38-42, October **(2012)** *I. Res. J. Biological Sci.*

Insights to Sequence Information of Lactoylglutathione Lyase Enzyme from Different Source Organisms

Dwivedi Vivek Dhar[1], Sharma Tanuj[1], Mishra Sarad Kumar[2] and Pandey Amit Kumar[3]
[1]Department of Bioinformatics, UCST, Dehradun, INDIA
[2]Department of Biotechnology, DDU University, Gorakhpur, INDIA
[3]Department of Forest Pathology, FRI, Dehradun, INDIA

Available online at: www.isca.in
Received 25th July 2012, revised 4th August 2012, accepted 14th August 2012

Abstract

Lactoylglutathione lyases (also known as glyoxalase I) are widely distributed enzymes among plants, fungi and bacteria. It is an enzyme that catalyzes the isomerization of hemithioacetal adducts, which are formed in a spontaneous reaction between a glutathionyl group and aldehydes such as methylglyoxal. In the present study, thirty full- length amino acid sequences of lactoylglutathione lyases from bacteria, fungi, and plants were collected and subjected to multiple sequence alignment (MSA), pattern identification, domain identification discovering individual amino acid composition, and phylogenetic tree construction. MSA revealed that one tyrosine residue were identically found in all analyzed species, two tyrosine, one arginine, one leucine, one glycine, one histidine, one phenylalanine, one proline, one aspartic acid and one glutamic acid residues were identically found in all the bacterial and fungal sources, one phenylalanine, one tyrosine, one histidine, one proline, and one glycine residues were identically found in all bacterial and plant sources while two glycine, two tyrosine, two aspartic acid and one proline residues were identically found in all plants and fungal lactoylglutathione lyases. Two major sequence clusters were constructed by phylogenetic analysis. One cluster contains ten species of fungi, five species of plant, and two species of bacteria, whereas the other one contains eight species of bacteria, four species of plant and one species of plant was outgrouped from both clusters. The amino acid composition result revealed that the average frequency of amino acid glycine is 7.86 percent that is very high in comparison to other amino acids and an average frequency of is 1.07 that is very low in all analyzed species. In addition, nine motifs which were unique for their groups were also identified.

Keywords: Lactoylglutathione lyase, phylogenetic analysis, conserved regions, motifs, Domains, amino acid composition.

Introduction

Lactoylglutathione lyase (LGL) is an enzyme involved in the detoxification of methylglyoxal, a highly toxic electrophilic glycolytic by-product that reacts with and inactivates intracellular macromolecules, including both proteins and nucleic acids[1]. Therefore, its rapid degradation is vital for cell survival. The formation of methylglyoxal occurs via enzymatic production during glycolysis from the fragmentation of triose phosphates[2-7]. LGL is involved in methylglyoxal detoxification via the formation of *S*-d-lactoylglutathione from the hemimercaptal adduct that is formed nonenzymatically between glutathione and the 2-oxoaldehyde methylglyoxal. Glyoxalase II then converts *S*-d-lactoylglutathione into reduced glutathione and d-lactate[8-11]. An examination of the *S. mutans* UA159 genome does not reveal the presence of a glyoxalase II homologue, suggesting an alternate pathway by which *S*-d-lactoylglutathione is neutralized. In *Escherichia coli*, methylglyoxal was accumulated under physiological conditions of uncontrolled carbohydrate metabolism and the concentration of methylglyoxal was seen to be greater in highly metabolically active human red blood cells, where *lgl* expression appeared to be regulated by the rate of glycolysis[12-13]. Considering the above facts, a study of amino acid sequences of lactoylglutathione lyase from different sources of organisms is quite challenging. In the present study, we performed the individual *in silico* studies of amino acid sequences obtained from bacteria, fungi, and plants, and correlated them on the basis of some common features.

Material and Methods

The full-length amino acid sequences of lactoylglutathione lyase from bacteria, fungi and plants were searched and retrieved from Entrez protein database available at NCBI. The sequences were arranged in bacterial, fungal, and plant profile, respectively. The multiple sequence alignment of the individual profiles was performed using MUSCLE at the European Bioinformatics Institute. Motifs were discovered in profiles using the expectation maximization approach[14] implemented in multiple EM for motif elicitation server. Further, the discovered motifs were used to search their protein family using Pfam at the Sanger Institute. The UPGMA approach implemented in the Mega program was employed for constructing phylogenetic relationships among sequences. The statistical reliability of the phylogenetic tree was tested by bootstrap analyses with 500 replications. Mega program is also used for discovering individual amino acid composition.

Results and Discussion

The accession number of retrieved sequences along with the species name and origin is listed in table-1. For MSA four profiles were created. One is from bacterial and fungal origin, second is from bacterial and plants origin, third is from plants and fungal origin, and fourth is from all bacterial, fungal, and plants origin. MSA showed the presence of some conserved regions in all the sequences from different profiles, while others were restricted only to their groups. one tyrosine residue were identically found in all analyzed species, two tyrosine, one arginine, one leucine, one glycine, one histidine, one phenylalanine, one proline, one aspartic acid and one glutamic acid residues were identically found in all the bacterial and fungal sources, one phenylalanine, one tyrosine, one histidine, one proline, and one glycine residues were identically found in all bacterial and plant sources while two glycine, two tyrosine, two aspartic acid and one proline residues were identically found in all plant and fungal lactoylglutathione lyase.

For phylogenetic analysis four major profiles were created. Profile one was of bacterial origin, second was of fungal origin, third was of plants origin, and fourth was joint profile of bacterial, fungal, and plants, origin.

Phylogenetic analysis[15] of bacterial profile showed two major clusters (figure-1). Cluster I consist of seven species which was further divided into two subclusters. Subcluster I contains four species (*Ferrimonas balearica, Mesorhizobium amorphae, Francisella noatunensis,* and *Mesorhizobium opportunistum*). Subcluster II contains three species (*Marinobacter adhaerens, Cardiobacterium valvarum, and Halomonas sp.*). Cluster II consist of three species namely *Agrobacterium tumefaciens, Clostridium perfringens, and Streptococcus pyogenes.*

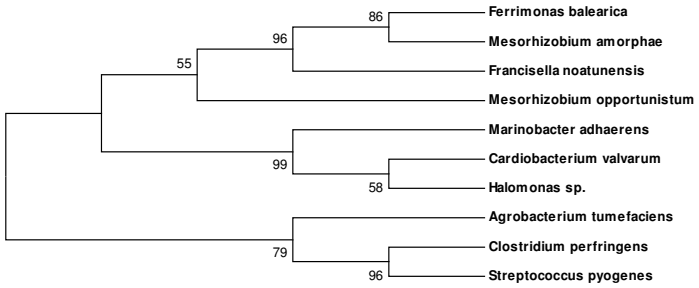

Figure-1
Phylogenetic analysis of bacterial profile using UPGMA method

Phylogenetic analysis[16] of fungal profile showed two major clusters (figure-2). Cluster I consist of four species which was further divided into two subclusters. Subcluster I contains two species (*Metarhizium acridum, and Magnaporthe oryzae).* Subcluster II also contains two species (*Neurospora tetrasperma, and Neurospora crassa*). Cluster II consist of two species namely *Colletotrichum higginsianum, and Glomerella graminicola*. *Verticillium dahliae, Exophiala dermatitidis, Paracoccidioides brasiliensis* and *Metarhizium anisopliae* were ditantantly related to each other and not included in any cluster.

Table -1
Retrieved sequence from NCBI/Entrez and their accession number

Source	Species	Accession no.
Bacteria	Ferrimonas balearica	YP_003913279.1
Bacteria	Marinobacter adhaerens	ADP98654.1
Bacteria	Mesorhizobium opportunistum	AEH89136.1
Bacteria	Francisella noatunensis	AFJ43975.1
Bacteria	Clostridium perfringens	EIA18179.1
Bacteria	Streptococcus pyogenes	NP_268789.1
Bacteria	Cardiobacterium valvarum	ZP_09445317.1
Bacteria	Mesorhizobium amorphae	ZP_09089707.1
Bacteria	Agrobacterium tumefaciens	EHH07353.1
Bacteria	Halomonas sp.	ZP_08960109.1
Fungi	Neurospora tetrasperma	EGZ68442.1
Fungi	Metarhizium anisopliae	EFY94694.1
Fungi	Metarhizium acridum	EFY91247.1
Fungi	Magnaporthe oryzae	EHA48833.1
Fungi	Colletotrichum higginsianum	CCF47148.1
Fungi	Exophiala dermatitidis	EHY60225.1
Fungi	Verticillium dahliae	EGY14417.1
Fungi	Paracoccidioides brasiliensis	EEH40065.1
Fungi	Glomerella graminicola	EFQ30294.1
Fungi	Neurospora crassa	XP_960441.1
Plant	Zea mays	NP_001146873.1
Plant	Arabidopsis thaliana	AEE28246.1
Plant	Medicago truncatula	AES92643.1
Plant	Vitis vinifera	XP_002273346.2
Plant	Oryza sativa	AEK99333.1
Plant	Cucumis melo	ADN34057.1
Plant	Gossypium hirsutum	ACJ11750.1
Plant	Glycine max	NP_001236152.1
Plant	Solanum lycopersicum	NP_001234447.1
Plant	Arachis hypogaea	ACF74334.1

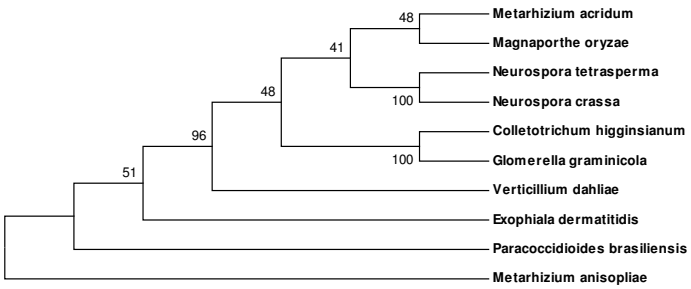

Figure-2
Phylogenetic analysis of fungal profile using UPGMA method

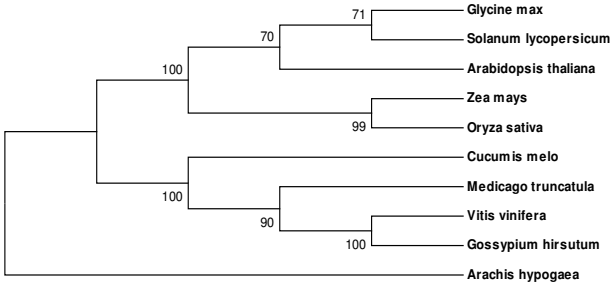

Figure-3
Phylogenetic analysis of plants profile using UPGMA method

Phylogenetic analysis of plant profile showed two major clusters (figure-3). Cluster I consist of five species which was further divided into two subclusters. Subcluster I contains three species (*Glycine max, Solanum lycopersicum,* and *Arabidopsis thaliana).* Subcluster II also contains two species (*Zea mays, and Oryza sativa*). Cluster II consist of four species namely *Cucumis melo, Medicago truncatula, Vitis vinifera and Gossypium hirsutum. Arachis hypogaea* was distantly related and not included in any cluster.

When joint profile of bacteria, fungi, and plant, sequences were taken for phylogenetic analysis, two major clusters were formed (figure-4). One cluster contains ten species of fungi, five species of plant, and two species of bacteria, whereas the other one contains eight species of bacteria, four species of plant and one species of plant was outgrouped from both clusters. This suggests that fungi were closely related to plants in comparison to bacteria in terms of lactoylglutathione lyases. Nine motifs are identified with their Pfam analysis which was unique for their groups given in table-2. The amino acid composition result revealed that the average frequency of amino acid glycine is 7.86 percent that is very high in comparison to other amino acids and an average frequency of is 1.07 that is very low in all analyzed species (figure-4).

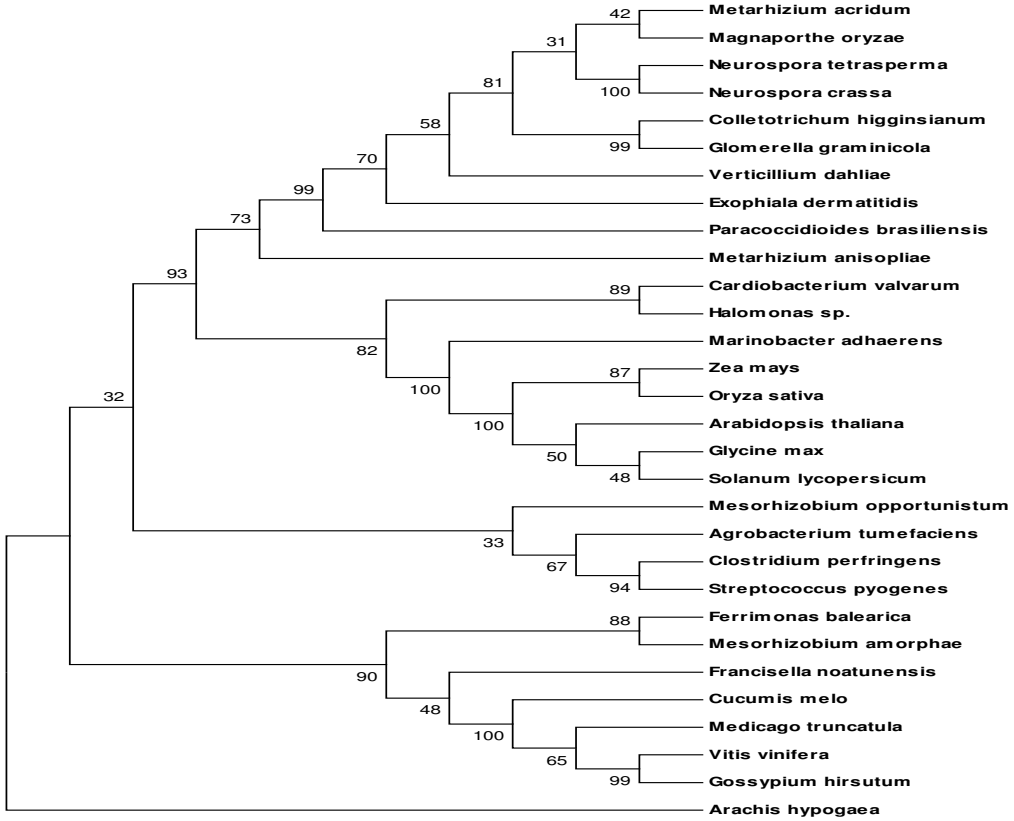

Figure-4
Phylogenetic analysis of joint profile of bacterial, fungal, and plants sequences using UPGMA method

Table-2
Motifs searched with meme program

Source	Motif width	Motif present in number of sequences	Motif	Pfam
Bacteria	36	9	MKFLHTMLRVKDLDRSLDFYTNAFGMTEVRRLDFEE	Glyoxalase family
Bacteria	16	9	IAFIKDPDGYKIEVIQ	Glyoxalase 2 family
Bacteria	21	9	YDEGNGFGHIAVGVEDIYAAC	Pfam entry not found
Fungi	50	8	PQGFGHICVSVDDIDAACARFEALKVNWKKRLTDGRMKNVAFLLDPDNYW	Glyoxalase 2 family
Fungi	50	8	IELTHNYGTENDPSYTVNNGNTEPHRGFGHTCISVDNIQAACQRLEDAGY	Glyoxalase 4 family
Fungi	41	8	HSMIRVKDPKASVKFYELLGMSVIKKLEFPEAKFDLYFLAY	Glyoxalase family
Plant	41	10	QQTMLRVKDPKRSLDFYSKVLGMSLLKRLDFPEMKFSLYFM	Glyoxalase family
Plant	21	10	RGFGHIGVTVDDVYKACERFE	Glyoxalase family
Plant	41	10	APVDDLDRTIWTFTQCATMELTHNWGTEEDPEFKGYHNGNS	Pfam entry not found

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
Arachis hypogaea	3.00	2.00	6.50	11.00	4.00	7.00	3.50	6.00	7.00	8.50	3.00	8.50	5.50	5.00	4.50	4.00	3.50	5.00	0.50	2.00	200
Agrobacterium tumefaciens	8.40	0.76	6.11	10.69	7.63	8.40	2.29	6.11	6.11	8.40	1.53	3.05	3.05	3.05	6.11	3.82	1.53	7.63	0.00	5.34	131
Clostridium perfringens	4.76	0.79	8.73	10.32	3.17	7.14	2.38	4.76	8.73	9.52	3.17	1.59	3.17	0.79	3.97	5.56	4.76	7.14	0.00	9.52	126
Streptococcus pyogenes	7.20	0.80	9.60	7.20	6.40	7.20	4.80	4.80	7.20	8.80	0.80	4.00	3.20	3.20	3.20	4.80	4.00	4.00	0.00	8.80	125
Mesorhizobium opportunistum	8.90	1.37	7.53	7.53	2.74	7.53	1.37	4.79	4.79	8.22	3.42	5.48	4.79	2.74	6.85	5.48	4.11	4.79	2.05	5.48	146
Metarhizium anisopliae	4.26	1.22	5.78	6.08	5.17	9.12	3.04	3.65	4.56	10.03	1.82	5.47	4.26	3.95	5.78	8.21	5.17	6.99	1.52	3.95	329
Paracoccidioides brasiliensis	6.90	1.25	6.58	6.90	5.33	6.58	3.45	4.08	5.96	7.52	2.19	6.90	4.39	3.13	5.33	5.96	5.33	5.64	2.19	4.39	319
Exophiala dermatitidis	4.45	1.06	6.14	6.14	5.08	6.78	3.18	4.24	8.05	8.90	2.54	5.30	4.03	4.45	5.08	4.87	7.42	5.93	2.33	4.03	472
Metarhizium acridum	5.98	1.42	6.55	6.84	4.56	6.84	3.13	4.27	7.41	7.12	3.13	6.84	3.70	3.99	5.41	4.84	6.27	5.70	0.85	5.13	351
Verticillium dahliae	7.67	1.60	8.31	6.71	5.75	7.35	2.88	4.15	7.35	7.03	2.56	7.03	2.24	1.92	4.79	4.15	6.71	6.07	1.60	4.15	313
Neurospora tetrasperma	6.03	1.27	6.67	8.57	5.40	8.57	2.86	3.17	8.25	8.57	2.86	6.98	2.86	3.17	4.44	3.49	4.76	5.71	1.59	4.76	315
Neurospora crassa	6.03	1.27	6.67	8.57	5.40	8.57	2.86	3.17	8.25	8.57	2.86	6.98	2.86	3.17	4.44	3.49	4.76	5.71	1.59	4.76	315
Magnaporthe oryzae	5.88	1.12	6.44	7.00	5.04	6.72	3.36	4.48	7.28	7.28	3.08	6.72	3.92	3.92	5.60	6.44	5.60	5.32	1.12	3.64	357
Colletotrichum higginsianum	5.80	1.45	6.96	6.67	5.51	5.80	2.61	4.93	7.83	7.54	2.61	7.25	2.90	5.22	4.93	4.35	7.25	5.22	1.16	4.06	345
Glomerella graminicola	5.83	1.62	7.12	7.44	5.18	6.47	2.91	5.18	8.41	7.12	2.59	6.47	3.56	4.53	4.21	4.21	6.15	5.50	1.29	4.21	309
Cucumis melo	5.10	0.68	6.46	7.82	3.40	8.16	1.70	5.78	6.80	9.86	3.40	3.74	4.08	4.42	4.42	4.76	6.12	7.48	0.68	5.10	294
Medicago truncatula	7.12	0.36	7.83	7.83	4.27	8.90	2.14	5.34	6.76	8.90	0.71	3.56	3.91	2.85	4.27	2.14	7.12	9.61	0.71	5.69	281
Vitis vinifera	5.48	0.68	6.51	9.25	4.45	8.56	1.37	5.48	9.25	7.53	2.74	2.74	5.48	1.71	4.45	2.40	7.19	8.22	0.68	5.82	292
Gossypium hirsutum	5.54	0.69	6.23	8.65	4.50	9.00	1.73	5.19	8.65	7.27	2.42	2.42	5.19	1.38	4.50	4.50	6.57	9.00	0.69	5.88	289
Francisella noatunensis	5.60	1.60	10.40	4.80	4.00	9.60	4.00	7.20	8.00	8.00	4.00	3.20	1.60	3.20	2.40	1.60	5.60	8.00	0.80	6.40	125
Ferrimonas balearica	7.30	0.73	7.30	8.03	3.65	13.14	2.92	6.57	2.92	8.03	2.92	2.92	2.19	3.65	4.38	3.65	6.57	6.57	0.73	5.84	137
Mesorhizobium amorphae	7.30	0.73	7.30	8.03	4.38	11.68	2.92	5.84	5.11	9.49	4.38	2.19	5.11	0.00	5.84	1.46	6.57	5.84	0.73	5.11	137
Cardiobacterium valvarum	5.98	1.09	6.52	7.07	5.43	8.15	3.26	7.07	5.43	4.89	3.26	4.35	4.89	2.72	7.07	4.89	6.52	4.35	2.72	4.35	184
Halomonas sp.	8.43	0.56	8.43	7.30	7.87	6.74	2.81	2.25	4.49	6.74	3.93	3.37	5.06	6.18	4.49	4.49	5.06	7.30	2.25	2.25	178
Marinobacter adhaerens	4.40	0.55	9.89	8.79	6.59	8.79	2.75	2.75	6.59	7.14	5.49	2.75	6.04	4.40	5.49	0.55	5.49	5.49	1.10	4.95	182
Zea mays	6.76	0.90	7.66	5.86	6.76	7.66	2.70	4.05	7.21	8.11	2.70	2.70	5.86	2.25	5.86	6.76	6.76	5.41	1.35	2.70	222
Oryza sativa	5.82	0.53	7.41	6.88	6.88	8.99	2.12	3.17	7.94	5.82	3.70	3.70	6.88	2.12	4.76	5.82	5.82	6.35	1.59	3.70	189
Arabidopsis thaliana	6.38	1.70	5.96	6.38	7.23	6.81	1.70	5.96	6.38	5.53	2.98	4.26	5.53	1.70	5.53	8.51	8.94	3.83	1.28	3.40	235
Glycine max	5.41	0.54	6.49	7.57	6.49	8.11	2.16	5.41	8.65	4.86	3.78	4.86	6.49	2.70	3.24	5.41	7.57	4.86	1.62	3.78	185
Solanum lycopersicum	5.95	0.54	7.57	7.03	7.03	8.11	2.16	4.32	8.65	6.49	3.24	3.24	5.95	1.62	2.70	9.19	6.49	4.32	1.62	3.78	185
Avg	5.97	1.07	7.03	7.50	5.27	7.86	2.70	4.66	7.17	7.77	2.83	5.01	4.22	3.26	4.86	4.76	6.04	6.14	1.29	4.58	242.3

Figure-5
Amino acid composition of lactoylglutathione lyase (given in percent) in different source organisms

Conclusion

In silico analysis of the sequences showed sequence based similarities depending on their source organism. one tyrosine residue were identically found in all analyzed species, two tyrosine, one arginine, one leucine, one glycine, one histidine, one phenylalanine, one proline, one aspartic acid and one glutamic acid residues were identically found in all the bacterial and fungal sources, one phenylalanine, one tyrosine, one histidine, one proline, and one glycine residues were identically found in all bacterial and plant sources while two glycine, two tyrosine, two aspartic acid and one proline residues were identically found in all plant and fungal lactoylglutathione lyase. This suggests that these conserved amino acid residues have an important function in lactoylglutathione lyase sequences and in its evolution from lower organisms (bacteria) to higher organisms (plants). Some motifs which were unique for their group were also identified. In all species of bacteria, fungi, and plants an average frequency of amino acid glycine is 7.86 percent that is very high in comparison to other amino acids. This suggests that the amino acid glycine play a very important role in the composition lactoylglutathione lyases. Two major sequence clusters were constructed by phylogenetic analysis One cluster contains ten species of fungi, five species of plant, and two species of bacteria, whereas the other one contains eight species of bacteria, four species of plant and one species of plant was outgrouped from both clusters. This classification can significantly contribute in the understanding of the evolutionary relations between the species at molecular level. However, owing to the considerable importance of lactoylglutathione lyase, more contribution is warranted for the detailed investigation of the activity and functional analysis of enzymes.

Acknowledgements

We are thankful to department of biotechnology DDU Gorakhpur University, Gorakhpur for providing laboratory facilities & encouragement and to director of UCST, Dehradun for their cooperation during the study. We are grateful to head of Department of Forest Pathology Division, FRI, Dehradun for his kind support and necessary suggestions whenever I needed.

References

1. Ackerman R.S., Cozzarelli N.R. and Epstein W., Accumulation of toxic concentrations of methylglyoxal by wild-type *Escherichia coli* K-12, *J. Bacteriol*, **119,** 357-362 **(1974)**

2. Allen R.E., Lo T.W. and Thornalley P.J., Purification and characterisation of glyoxalase II from human red blood cells, *Eur. J. Biochem*., **213**, 1261-1267 **(1993)**

3. Banas J.A., Virulence properties of *Streptococcus mutans*, *Front. Biosci*., **9,** 1267-1277 **(2004)**

4. Cooper R.A., Metabolism of methylglyoxal in microorganisms, *Annu. Rev. Microbiol*., **38,** 49-68 **(1984)**

5. Eriksson S., Lucchini S., Thompson A., Rhen M., and Hinton J.C., Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*, *Mol. Microbiol***, 47,** 103-118 **(2003)**

6. Freedberg W.B., Kistler W.S. and Lin E.C., Lethal synthesis of methylglyoxal by *Escherichia coli* during unregulated glycerol metabolism, *J. Bacteriol*, **108,** 137-144 **(1971)**

7. Frickel E.M., Jemth P., Widersten M. and Mannervik B., Yeast glyoxalase I is a monomeric enzyme with two active sites, *J. Biol. Chem*., **276,** 1845-1849 **(2001)**

8. Hanna M.N., Ferguson R.J., Li Y.H. and Cvitkovitch D.G., *uvrA* is an acid-inducible gene involved in the adaptive response to low pH in *Streptococcus mutans*, *J. Bacteriol*., **183,** 5964-5973 **(2001)**

9. Inoue Y. and Kimura A., Methylglyoxal and regulation of its metabolism in microorganisms, *Adv. Microb. Physiol*., **37,** 177-227 **(1995)**

10. Kalapos M.P., Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications, *Toxicol. Lett*., **110,** 145-175 **(1999)**

11. Khalichi P., Cvitkovitch D.G. and Santerre J.P., Effect of composite resin biodegradation products on oral streptococcal growth, *Biomaterials,* **25,** 5467-5472 **(2004)**

12. Kim I., Kim E., Yoo S., Shin D., Min B., Song J. and Park C., Ribose utilization with an excess of mutarotase causes cell death due to accumulation of methylglyoxal, *J. Bacteriol*., **186,** 7229-7235 **(2004)**

13. Ko J., Kim I., Yoo S., Min B., Kim K. and Park C., Conversion of methylglyoxal to acetol by *Escherichia coli* aldo-keto reductases, *J. Bacteriol*., **187,** 5782-5789 **(2005)**

14. Bailey T.L. and Elkan C., Expectation maximization pproach, Mach Learn, **21(1–2),** 51–80 **(1995)**

15. Dwivedi V.D. and Mishra S.K., Amino acid sequence analysis of glutamate dehydrogenase from different source organisms, *Online J Bioinform*., **13(2),** 184-191 **(2012)**

16. Bhatt T.K., Phylogenetic Studies on tRNA Dependent Amidotransferase from Plasmodium Falciparum, *ISCA J. Biological Sci.,* **1(3),** 20-24 **(2012)**