Research Article

Bioinformatics approach toward the Identification of Binding Pockets of Rice Metallothionein and its interaction with the heavy metals

DEVVRET*, PRAVEEN KUMAR, KUMUD PANT, ASHISH THAPLIYAL, NEEMA TUFCHI
Department of Biotechnology, Graphic Era University, Dehradun, Uttarakhand, India, 248002.
Email: devvret@gmail.com, (M) 9795094229

ABSTRACT
Rice is a staple crop that is consumed by half of the world's population. Paddy rice is highly prone to heavy metal toxicity. Low level of heavy metals in rice plant is beneficial for plant growth but its excess can affect the growth of the plant. Consumption of rice contaminated with heavy metal has a potential health risk to human population. Metallothionein (MT) is a metal binding proteins that comes under a class of conserved evolutionary protein molecules. It plays an essential function in the detoxification of heavy metals and homeostasis. Metals with relatively high densities, atomic weights, or atomic numbers are known as heavy metals. Cadmium (Cd), arsenic (As), chromium (Cr), lead (Pb) and zinc (Zn) are highly reactive and toxic to living cells. Plants have described two basic mechanisms such as chelation and sequestration of heavy metals by ligands to counter this problem. The two best-characterized heavy metal-binding ligands in plant cells are the phytochelatins and metallothionein. In the current study, we have generated the three dimensional structure of rice MT protein, identification of active sites of metallothionein and docking study was performed. Homology model of the metallothionein was designed through Phyre2 server. Heavy metals were docked with the metallothionein. The docking analysis and active sites were predicted through iGemDock software. The stable homology structure of the metallothionein was modelled and binding pockets were analyzed. This analysis of binding pocket will help in the *in vitro* study of metallothionein protein and understanding its role in metal detoxification.

Keywords: Metallothionein, Homology modelling, Docking, Binding Pocket.

INTRODUCTION

Rice (*Oryza sativa*) is an important staple crop that is consumed by half of the world's total population. It occupies an important place in Indian food system and is essential for food security. Rice provides approximately 75% of the average calories and 55% of protein in average daily diet of a person [1]. More than 90% of the world's rice is cultivated and consumed in Asia [2]. Approximately three billion people depend upon it as a major source of their sustenance [3].

Affect of Heavy metal contamination

Heavy metal ions, such as cadmium (Cd), arsenic (As), chromium (Cr), lead (Pb), zinc (Zn) and mercury (Hg), are highly reactive and toxic to living cells [4].In the previous studies it was found that rice is particularly efficient in arsenic uptake from paddy soil, leading to accumulation in rice grain at concentrations that may pose a health risk to people consuming large amounts of rice in their diet [5] [6].

Metallothionein

Margoshes and Vallee identified the Metallothionein from horse kidney in 1957 and suggested that metallothionein are members of cys-rich N and C termini protein, which have low molecular mass (4–8 KD).Metallothionein play an important role in heavy metal detoxification and homeostasis of intracellular metal ions in plant [7]. Heavy metals such as cadmium (Cd), arsenic (As), chromium (Cr), lead (Pb), zinc (Zn) and mercury (Hg), are toxic for living system and highly reactive [4]. Phytochelatin and metallothionein is a family of enzymatically synthesized cysteine-rich peptides that are highly characterized metal binding proteins in plant cells [7][8]. Plant metallothionein are class I proteins containing two smaller Cysteine-rich domains (CRD) and a large spacer region which lack in cysteine residues. Class I metallothionein have the capacity to bind both physiological (such as zinc, copper and selenium) and xenobiotic heavy metals (such as cadmium, mercury, silver and arsenic) through the thiol group of its cysteine residues. Expression of metallothionein gene is regulated by abiotic stress including heavy metal which plays an important role in metal detoxification and homeostasis [9][10] [11]. These functional capabilities of metallothionein allow them to play a significant role in mobilization of metal ions from senescing leaves and the sequestration of excess metal ions in trichomes.

Material and Methods

Retrieval of proteins: The protein sequences of rice metallothionein (MT) were retrieved from the Rice genome annotation project database

(URL:http://rice.plantbiology.msu.edu/, [12]. There are four types of metallothioneins present in rice plant and each of them performs different functions. Here we have only retrieved the type I MT for our study. The detailed information about the protein is mentioned in the Table 1.

Table 1: List of type I Metallothionein proteins.

S. No.	Locus ID	MT types	Protein name	Size of protein	Chromosome no.
1.	*Os01g05585*	*OsMT-I-IIa*	*Os01g05585*	64	1
2.	*Os01g05650*	*OsMT-IIb*	*Os01g05650*	82	1
3.	*Os01g74300*	*OsMT-I-IIc*	*Os01g74300*	81	1
4.	*Os03g17870*	*OsMT-I-Ia*	*Os03g17870*	73	3
5.	*Os05g11320*	*OsMT-I-IIIa*	*Os05g11320*	66	5
6.	*Os11g47809*	*OsMT-I-Ib*	*Os11g47809*	75	11
7.	*Os12g38010*	*OsMT-I-Ic*	*Os12g38010*	79	12
8.	*Os12g38290*	*OsMT-I-Ie*	*Os12g38290*	77	12
9.	*Os12g38270*	*OsMT-I-IId*	*Os12g38270*	79	12
10.	*Os12g38300*	*OsMT-I-If*	*Os12g38300*	80	12

Multiple sequence alignment of proteins

To find out the cysteine rich residues and the evolutionary conservation of MT proteins, multiple sequence alignment of the proteins were performed by Mega7 software using ClustalW algorithm [13].

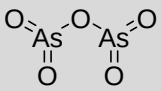
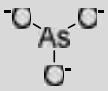
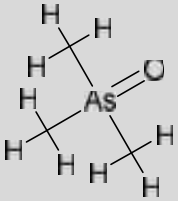
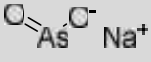
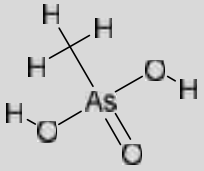
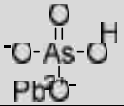
Homology modelling

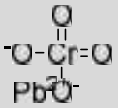
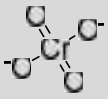
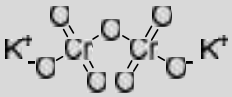
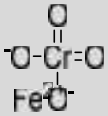
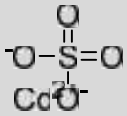
Homology modelling is used to predict the 3D-structure of an unknown protein based on the known structure of a similar protein. The three dimensional crystallographic structure of the metallothionein was not available in any of the structural databases of protein. The proteins were modelled through Phyre2 Server, which uses advanced remote homology or template detection technique to build three dimensional structure of protein [14]. The validation of structure models were obtained from the Ramachandran plot analysis using RAMPAGE server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php).

Ligand preparation

Ligands were retrieved in mol2 format from the Zinc database (http://zinc.docking.org/). These ligand molecules were then converted into .pdb format by using Open babel software which is open source chemical toolbox. The detailed information of the ligand was mentioned in Table 2.

Table 2: Detail information of the heavy metal compounds (ligand).

Heavy Metal	Compound Name	PubChemCID	Mol. Formula	Mol. Wt.(gm/mol)	2D Structure
Arsenic compound	Arsenic pentoxide	14771	As_2O_5	229.818	
	Arsenite	544	AsO_3	122.919	
	Trimethyl Arsenic oxide	104930	C_3H_9AsO	136.026	
	Sodium Arsenite	443495	$AsNaO_2$	129.909	
	Methyl Arsonic acid	8948	CH_5AsO_3	139.97	
	Lead hydrogen arsenate	24572	$PbHAsO_4$	347.126	

Lead compound	Lead chromate	24460	$PbCrO_4$	323.192	
	Chromate	24461	CrO_4^{-2}	115.992	
Chromium compound	Potassium dichromate	24502	$K_2Cr_2O_7$	294.182	
	Chromite	18541957	$CrFeO_4$	171.837	
Cadmium compound	Cadmium sulphate	24962	$Cd(SO_4)_2$	208.47	
Zinc compound	Zinc oxide	14806	ZnO	81.379	

Docking and Binding Site analysis:

In-silico molecular docking is a technique of structural biology that is used to study the chemical interactions between two molecules. Strength of association or binding affinity of the heavy metals with the metallothionein proteins were predicted through the iGEMDOCK software [15].

Result

Metallothioneins are cysteine rich, low molecular weight proteins that have many attributed functions. It plays an important role in metal detoxification and homeostasis. It is also involved in the regulation of metal ions which can affect the plant physiology and provide protection against oxidative stress [16]. This cysteine-rich proteins acts as a chelator for heavy metals, because metals have high affinity towards the thiol group of the cysteine (Cys) residues and thereby it protect plants from the heavy metal toxicity[17]. In the present study, we have retrieved ten sequences of type I metallothionein protein in fasta format. Locus ID, metallothionein types, protein name, size of protein and location on the chromosome is mentioned in Table 1. All the sequences were then aligned through ClustalW algorithm using Mega7 software. The analysis showed that these sequences were evolutionary conserved and rich in cysteine residues (Figure 1 and 2).

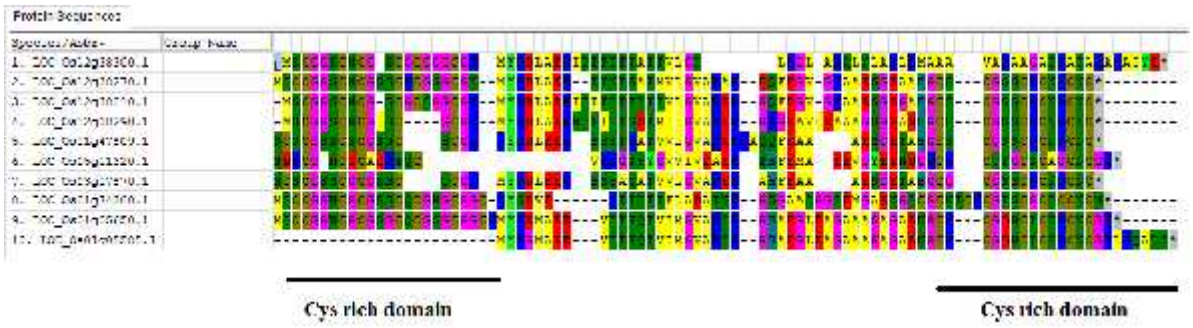


Figure 1: The multiple sequence alignment of Metallothionein protein showing the cysteine-rich region.

Docking

Docking is a tool of structural biology that is used to study about the interaction between two molecules. Here in the present study, we have performed the docking analysis of the metallothionein protein with several heavy metals. Metallothionein plays an important role in metal detoxification, so to the check binding affinity and the binding site of the metal compounds, docking was performed. The docking analysis suggested that the binding affinity of Lead nitrate is more than any other metal compounds selected for the study. Total free energy scored for inorganic Lead nitrate was found lowest -173.16, amongst all the heavy metal compounds. This suggests that paddy rice is highly prone to Lead nitrate contamination. Higher concentration of lead in soil can affect the germination of rice seeds and may reduce the growth of seedling [18].Arsenic pentaoxide also showed binding affinity close to lead nitrate. This suggests that metallothionein play an important role in metal detoxification by binding these heavy metal compounds and thus helping in

metal chelation. The interaction property of heavy metals was determined using iGemDock interaction analysis tool, in case of lead nitrate out of twenty two residues that are playing role in lead-metallothionein interaction, fifteen serine residues showed better interaction values. In the whole study, only lead nitrate is showing the interaction with maximum serein residues, and others are showing the interaction with the cysteine residues. In the case of arsenic pentaoxide, out of twenty five residues, thirteen cysteine residues, in the site 3, 7, 9, 12, 14 possess the arsenic pentaoxide absorbance property. This determines that cysteine residues play important role in functioning of metallothionein protein. The absorbance property of metallothionein with different heavy metals is given in Table 4.

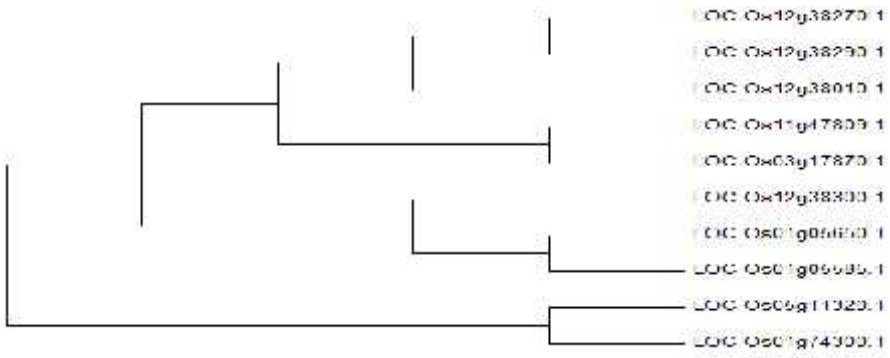
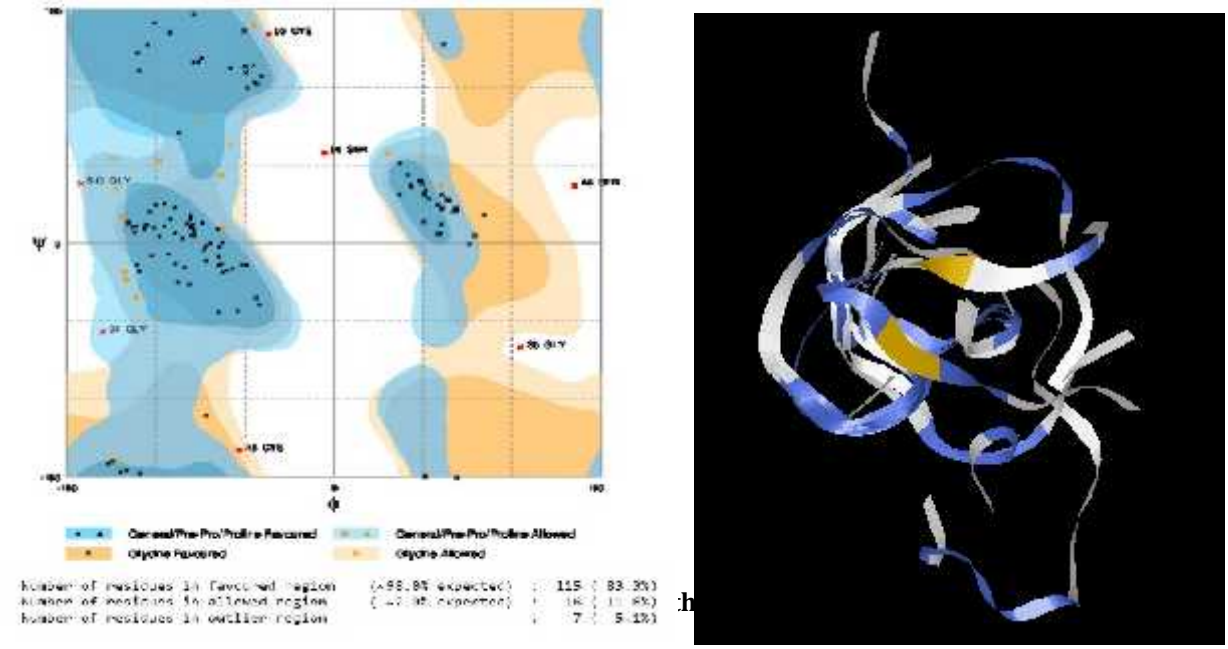


Figure 2:Maximum parsimony tree showing the evolutionary relationship of the MT proteins.

The crystallographic structures were not available in the database so we have modelled each protein sequence through Phyre2**,** which is an online server that uses the advanced template detection and homology modelling technique. It uses multiple template for homology modelling of a protein. The structure of the proteins was then validated through Ramachandran Plot analysis by using RAMPAGE server. Only the protein with ID Os12g38300 showed the better result than other proteins. Out of 138 amino acid residues, 115 residues fall in the most favoured region, 16 in the most allowed region and 7 amino acid residues in the outlier region of the plot. Ramachandran plot analysis suggests that the protein is stable. The multiple template for modelling the proteins were c2mrtA_ (Chain 'A'), d1e9ma_ (Chain 'B'), d1xlqa1 (Chain 'C'), d1l5pa_ (Chain 'D'), c3ah7A_ (Chain 'E'), d1i7ha_ (Chain 'F'), c2wlbB_ (Chain 'G'), c3lxfC_ (Chain 'H'), c3huiA_ (Chain 'I'), d1b9ra_ (Chain 'N'), c2msqA_ (Chain 'R'), c1m0jA_ (Chain 'S'). Ligands were retrieved from the Pubchem and Zinc Database in .sdf format and converted into three dimensional structure or in .pdb format, required for docking analysis. The details like molecular weight, valency state etc. of the ligands were mentioned in Table 2.



Fi **h**

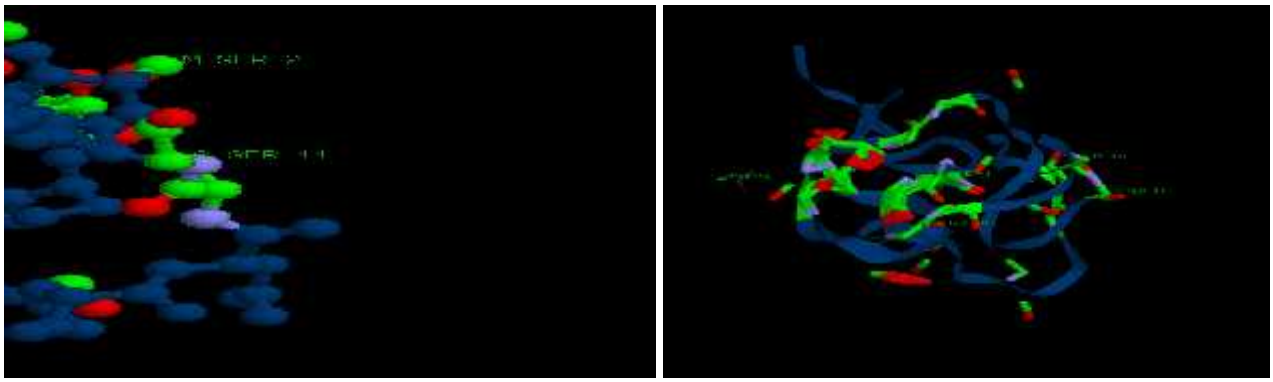


Figure 4: Docked poses of lead nitrate with metallothionein protein.

Table 3. Energy values obtained through iGemDock during docking analysis of heavy metals with metallothionein protein.

Compound	Energy	VDW	HBond
RiceMT-Lead nitrate	-175.9	-80.08	-95.82
RiceMT-Arsenic pentoxide	-173.159	-46.7324	-126.427
RiceMT-Arsenite	-135.827	-10.0062	-125.821
RiceMT-Cadmium sulphate	-135.581	-97.2889	-38.2923
RiceMT-Chromate	-120.61	-86.4751	-34.1349
RiceMT-Chromite	-130.661	-90.4421	-40.2188
RiceMT-Dimethyl arsenic acid	-125.016	-83.7157	-41.3001
RiceMT-Lead chromate	-125.966	-90.5906	-35.3755
RiceMT-Lead hydrogen arsenate	-134.64	-64.2181	-70.4217
RiceMT-Methyl arsenic Acid	-118.688	-85.5423	-33.1457
RiceMT-potassium dichromate	-171.378	-89.7862	-81.5919
RiceMT-Sodium arsenite	-137.146	-40.7387	-96.4074
RiceMT-Tri chloroarsenic	-152.514	-91.1269	-61.3866
RiceMT-Trimethylarsenic oxide	-118.622	-85.8524	-32.7693
RiceMT-Zinc oxide	-88.97	-17.84	-17.13

Table 4: Data showing interaction analysis of amino acids with metallothionein. H-S signifies hydrogen bond with side chain; H-M signifies hydrogen bond with the main chain. Number of amino acids is shown below its respective amino acid by iGemDock.

Heavy Metals	Interaction with MT	Heavy Metals	Interaction with MT	Heavy Metals	Interaction with MT
Lead nitrate	H-S SER 6 -2.5 H-M CYS 9 -3.5 H-M SER 2 -3.5 H-S SER 2 -6.5 H-M GLY 10 -3.5 H-S SER 11 -2.7 H-S SER 2 -7.8 H-M GLY 10 -3.5 H-S SER 11 -2.9 H-M SER 2 -3.5 H-S SER 2 -4.7 H-M GLY 10 -3.5 H-S SER 11 -2.5 H-M MET 1 -3.5 H-M SER 2 -3.5 H-S SER 2 -8.0 H-M GLY 10 -3.5 H-S SER 11 -2.5 H-M SER 2 -3.4	Arsenic pentoxide	H-M GLY 13 -7.0 H-M GLY 13 -4.1 H-S CYS 12 -2.5 H-M GLY 13 -4.4 H-M GLY 13 -5.0 H-S CYS 12 -2.5 H-M GLY 13 -4.8 H-M CYS 14 -3.3 H-S CYS 12 -2.5 H-M GLY 13 -5.3 H-M CYS 14 -2.8 H-S CYS 3 -2.5 H-M CYS 7 -6.7 H-S CYS 7 -2.5 H-M ASN 8 3.5 H-M CYS 9 -3.5	Arsenite	H-M GLY 13 -7 H-S CYS 12 -2.5 H-M GLY 13 -5.8 H-S CYS 12 -2.5 H-M GLY 13 -6.0 H-S CYS 12 -2.5 H-M GLY 13 -6.9 H-S CYS 12 -2.5 H-M GLY 13 -6.9 H-S CYS 12 -2.5 H-M GLY 13 -7.0 H-M CYS 7 -3.1 H-M ASN 8 -5.7 H-M CYS 9 -3.5

	H-S SER 2 -6.0 H-M GLY 10 -3.5 H-S SER 11 -2.5		H-M GLY 10 -6.2 H-S CYS 3 -2.5 H-M ASN 8 -3.5 H-M CYS 9 -3.5 H-M GLY 10 -4.9 H-S CYS 3 -2.5 H-M ASN 8 -3.5 H-M CYS 9 -3.5 H-M GLY 10 -5.1		H-M GLY 10 -6.9 H-M ASN 8 -4.3 H-M CYS 9 -5.2 H-M GLY 10 -5.5 H-S CYS 12 -2.5 H-S CYS 3 -2.5 H-M ASN 8 -4.1 H-M CYS 9 -5.1 H-M GLY 10 -5.8 H-S CYS 12 -2.5
Potassium dichromate	H-M GLY 10 -2.6 H-M GLY 13 -3.5 H-M CYS 7 -2.6 H-M CYS 14 -3.2 H-M ASN 8 -3.5 H-M GLY 10 -5.1 H-S CYS 7 -2.7 H-M ASN 8 -3.5 H-M CYS 9 -2.8 H-M GLY 10 -6.2 H-S CYS 7 -8 -5 H-M ASN -3.4 H-M CYS 9 -2.7 H-M GLY 10 -6.4 H-S CYS 12 -2.7	Trichloro arsenic	H-M GLY 13 -3.5 H-M CYS 7 -7.0 H-M GLY 13 -3.0 H-M GLY 13 -3.0 H-M GLY 13 -3.5 H-M GLY 13 -3.5 H-M GLY 13 -3.5 H-M ASN 8 -3.2 H-M GLY 10-5.8 H-M GLY 10-5.4 H-M GLY 10-5.6	Lead hydrogenarsenate	H-M ASN 8 -3.5 H-M GLY 10 -3.5 H-S CYS 7 -2.5 H-M CYS 9 -2.6 H-M GLY 10 -3.5 H-S CYS 7 -2.9 H-M CYS 9 -2.6 H-M GLY 10 -3.5
Trimethyl arsenic oxide	H-M GLY 13 -3.5 H-M CYS 7 -3.5 H-M GLY 13 -2.6 H-M GLY 13 -2.6 H-M GLY 13 -3.4 H-M GLY 13 -3.3 H-M GLY 13 -3.4 H-M GLY 10 -3.5 H-M GLY 10 -3.4 H-M GLY 10 -3.5	Zinc oxide	H-M GLY 13 -3.5 H-S CYS 12 -2.5 H-M GLY 13 -2.6 H-M GLY 13 -2.9 H-S CYS 12 -2.5 H-M GLY 13 -3.3 H-S CYS 12 -2.5 H-M GLY 13 -3.3 H-M GLY 13 -3.4 H-M CYS 7 -3.5 H-M ASN 8 -3.5 H-M CYS 9 -3.5 H-M GLY 10 -3.2 H-M ASN 8 -3.5 H-M CYS 9 -3.5 H-M ASN 8 -3.5 H-M CYS 9 -3.5	Methyl arsenic acid	H-M GLY 13 -3.5 H-M CYS 7 -3.5 H-M GLY 13 -2.6 H-M GLY 13 -2.6 H-M GLY 13 -3.4 H-M GLY 13 -3.3 H-M GLY 13 -3.5 H-M GLY 10 -3.5 H-M GLY 10 -3.4 H-M GLY 10 -3.5
Chromite ore	H-M CYS 7 -3.4 H-M ASN 8 -3.5 H-M GLY 10 -5.1 H-M GLY 10 -6.0 H-S CYS 7 -2.5 H-M GLY 10 -6.4	Dimethyl arsenic acid	H-M GLY 13 -3.1 H-M CYS 7 -3.0 H-M ASN 8 -3.5 H-M GLY 10 -4.7 H-M GLY 10 -5.4 H-M GLY 10 -5.9	Lead chromate	H-M CYS 7 -3.4 H-M ASN 8 -3.5 H-M GLY 10 -5.1 H-M GLY 10 -5.9 H-S CYS 7 -2.5 H-M GLY 10 -6.3
Chromate	H-M CYS 7 -3.4 H-M ASN 8 -3.5		H-M CYS 7 -3.3 H-M ASN 8 -3.5		

	H-M GLY 10-4.2 H-M GLY 10-4.9 H-S CYS 7 -2.5 H-M GLY 10-5.4	Cadmium sulphate	H-M GLY 10 -5.4 H-M GLY 10 -6.4 H-S CYS 7 -2.5 H-M GLY 10 -6.6		

Conclusion

Rice plants show reduction in growth when they are grown in heavy metal polluted soil. Metallothionein, a cysteine rich protein plays important role in heavy metal detoxification. The homology model of the rice Metallothionein and the identified sites of interaction will provide additional help to analyze the crystal structure and to understand the role of Metallothionein protein.

Reference

1. Orrok BG, Alcoholics Anonymous: The Story of How Many Thousands of Men and Women Have Recovered From Alcoholism, Journal of the American Medical Association, 1989; 261(22): 3315-3316.
2. Tyagi AK, Khurana JP, Khurana P, et al. Structural and functional analysis of rice genome, Journal of genetics, 2004; 83(1):79-99.
3. Cantrell RP, Reeves TG. The cereal of the world's poor takes center stage, Science, 2002; 296(5565):53-53
4. Hall JL. Cellular mechanisms for heavy metal detoxification and tolerance, Journal of experimental botany, 2002; 53(366): 1-1.
5. Williams PN, Villada A, Deacon C, et al. Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley, Environmental Science & Technology, 2007; 41(19):6854-9.
6. Zhu YG, Williams PN, Meharg AA, Exposure to inorganic arsenic from rice: a global health issue?, Environmental pollution, 2008; 154(2):169-71.
7. Margoshes M, Vallee BL, A cadmium protein from equine kidney cortex, Journal of the American Chemical Society, 1957; 79(17):4813-4.
8. Yu LH, Umeda M, Liu JY, et al. A novel MT gene of rice plants is strongly expressed in the node portion of the stem, Gene, 1998; 206(1):29-35.
9. Sharma R, Kishore J, Kher SA, Differential expression of two metallothionein encoding genes during heavy metal stress in the mangrove species, Avicennia marina (Forsk.), Current Science, 2007; 93:1215-1219.
10. Huang GY, Wang YS, Expression analysis of type 2 metallothionein gene in mangrove species (Bruguieragymnorrhiza) under heavy metal stress, Chemosphere, 2009; 77(7):1026-9.
11. Singh RK, Anandhan S, Singh S, et al. Metallothionein-like gene from Cicermicrophyllum is regulated by multiple abiotic stresses, Protoplasma, 2011; 248(4):839-47.
12. Kawahara Y, Bastide M, Hamilton JP, et al. Improvement of the OryzasativaNipponbare reference genome using next generation sequence and optical map data, Rice, 2013; 6(1):4.
13. Kumar S, Stecher G, Tamura K, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets, Molecular biology and evolution, 2016; msw 054.
14. Kelley LA, Mezulis S, Yates CM, et al. The Phyre2 web portal for protein modeling, prediction and analysis, Nature protocols, 2015; 10(6):845-58.
15. Hsu KC, Chen YF, Lin SR, et al. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis, BioMedCentral Bioinformatics, 2011; 12(1):S33.
16. Jin S, Sun D, Wang J, et al. Expression of the rgMT gene, encoding for a rice metallothionein-like protein in Saccharomyces cerevisiae and Arabidopsis thaliana, Journal of genetics, 2014; 93(3):709-18.
17. Thirumoorthy N, Kumar KM et al. Citation of This Article Corresponding Author of This Article Keywords for Related Literature of This Article Article-Type of This Article Open-Access Policy of This Article, World J Gastroenterol, 2007; 13(7): 993-996.
18. Verma S, Dubey RS. Lead toxicity induces lipid peroxidation and alters the activities of antioxidant enzymes in growing rice plants, Plant Science, 2003; 164(4):645-55.