Sequence and structural characterization of Trx-Grx type of monothiol glutaredoxins from *Ashbya gossypii*

Saurabh Yadav[1$], Pragati Kumari[3$] & Hemant Ritturaj Kushwaha[2]*

[1]Department of Experimental Medicine and Biotechnology, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh-160012, India; [2]Synthetic Biology and Biofuel Group; International Centre for Genetic Engineering and Biotechnology (ICGEB); New Delhi-110067, India; [3]S P. Institute of Biotechnology (Training and Research), E-734-735, Nakul Path, Lal Kothi Scheme, Jaipur, Rajasthan-302015, India; Hemant Ritturaj Kushwaha - Email: hemant@icgeb.res.in; Phone: 011-26742357; *Corresponding author
$ - Authors contributed equally

Abstract:
Glutaredoxins are enzymatic antioxidants which are small, ubiquitous, glutathione dependent and essentially classified under thioredoxin-fold superfamily. Glutaredoxins are classified into two types: dithiol and monothiol. Monothiol glutaredoxins which carry the signature "CGFS" as a redox active motif is known for its role in oxidative stress, inside the cell. In the present analysis, the 138 amino acid long monothiol glutaredoxin, AgGRX1 from *Ashbya gossypii* was identified and has been used for the analysis. The multiple sequence alignment of the AgGRX1 protein sequence revealed the characteristic motif of typical monothiol glutaredoxin as observed in various other organisms. The proposed structure of the AgGRX1 protein was used to analyze signature folds related to the thioredoxin superfamily. Further, the study highlighted the structural features pertaining to the complex mechanism of glutathione docking and interacting residues.

Keywords: Ashbya, GRX, Monothiol, glutaredoxin, Homology modelling, Docking.

Background:

Stigmatomycosis is one of the most common fungal diseases in cotton (*Gossypium hirsutum*) and various subtropical citrus fruits. The causative agent of the disease in various plant species is a fungal pathogen, *Ashbya gossypii*, first identified and characterised by Ashby and Nowell in 1926 **[1]**. *Ashbya gossypii* is a filamentous fungus which relies on heteropterous insects for its dispersal of spores or mycelia fragments. Apart from being a well known plant pathogen, *Ashbya gossypii* is also recognized for its ability to produce riboflavin (vitamin B2) which has been exploited for the commercial interest. *Ashbya gossypii* is known to be closely related to the unicellular yeast *Saccharomyces cerevisae* and share a considerable level of conservation in gene order and synteny **[2, 3]**. Earlier analysis established that over 95% proteins share homology with the *Saccharomyces cerevisiae* proteins indicating the conservation of essential cellular machinery between the two species **[3]**. Last

decade has witnessed the technological advancement in various molecular biology techniques. This improvement in technology has been widely exploited to sequence large number of genomes. With the completion and availability of genome sequence in *A. gossypii* one could look into crucial hidden mechanisms pertaining to its economic importance. *A. gossypii* genome sequence has been considered as one of the best-annotated eukaryotic genome due to the simplicity and compactness of the genome **[3]**. Reactive oxygen species (ROS) is generated as a consequence of metabolism in all aerobic organisms, in response to various environmental stress conditions such as drought, salinity etc. and also during plant-pathogen interactions **[4]**. Oxidative stress causes damage to cysteine amino acids and due to their unique ability to form intra or intermolecular disulfide bridges by oxidation, the thiol groups of cysteine residues are of physiological importance for many biological reactions such as protein folding and dynamic

regulation of protein structures under oxidative stress **[5]**. Various species have developed an elaborate mechanism comprising of antioxidants and enzymes in order to control the ROS, which can otherwise damage the cellular components **[6]**. Cellular milieu has plethora of antioxidant molecules like thioredoxins, glutaredoxins, glutathione, superoxide dismutase, catalase etc **[7]**. Glutaredoxins are small (~10-15 KDa) thermostable molecules classified as oxidoreductases of thioredoxin superfamily and conserved in both prokaryotes and eukaryotes **[8-10]**. The GRXs catalyze the reduction of oxidatively damaged proteins *via* a dithiol (2-Cys) or monothiol (1-Cys) mechanism **[11]**. The monothiol and dithiol GRXs contain a conserved Cys-X-X-Ser and Cys-X-X-Cys active site, respectively **[11]**. Although the glutaredoxin family exists in most of the fungal species, only a few fungal GRXs account for functional significance, till date.

GRXs play role in glutathionylation and de-glutathionylation reactions, thus have binding region for glutathione molecules **[12]**. These reactions regulate the activities of many enzymes and molecules which aid in signalling during normal and stress conditions **[13]**. Monothiol GRX from some fungi have been studied such as, cloning of CGFS type GRX from *Taiwanofungus camphorata* **[14]**, analysis of yeast GRX5 knockout mutant **[15]**, yeast Grx6 and Grx7 associated with the early secretory pathway **[16]** and CGFS-type GRX4 glutaredoxin from *Schizosaccharomyces pombe* **[17]** etc. In the present analysis, we have identified two monothiol glutaredoxins from *Ashbya gossypii* using *in-silico* approach and one protein having probable mitochondrial location (AgGRX1-NP_984149) was chosen for homology modelling. Both the protein sequences of *A.gossypii* were analysed and compared with respect to the other known monothiol GRX homologues. The molecular phylogeny also shed light on the evolutionary relationship of monothiol GRXs in fungi. We have also attempted to predict docking with glutathione molecule.

Methodology:

In-silico isolation of monothiol GRXs from *Ashbya gossypii*
In order to identify monothiol GRX protein sequences from *Ashbya gossypii*, BLAST **[18]** searches were made in non redundant database of NCBI using yeast monothiol GRXs as query. The sequences obtained were aligned using MUSCLE multiple alignment software **[19]**. The specific sequences of monothiol glutaredoxin proteins were identified using the conserved signature of monothiol glutaredoxin proteins. Further, in order to have an exhaustive overview, pfam domain of glutaredoxin (PF00462) was used to search monothiol GRXs in *A. gossypii* genome. All the sequences obtained were aligned and checked for conserved monothiol glutaredoxin signature "CGFS".

Sequence analysis and multiple alignments
The sequence analysis of two proteins (NP_984149 and NP_986777) was performed by Compute *pI*/Mw, using ExPASy server (http://www.expasy.org/tools/). The homologues were searched using BLASTp program available at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The protein sequences of full-length monothiol glutaredoxin from various organisms were aligned with MUSCLE software (v3.8) **[19]**. The putative localization of monothiol glutaredoxin in *Ashbya*

gossypii, was predicted using TargetP (v1.01) **[20]** and PSORT **[21]** web based software.

Phylogenetic analysis
The full-length protein sequences obtained from NCBI were used to perform phylogenetic analysis using Mega (ver 5.05) software **[22]**. The tree was calculated using maximum likelihood approach using 500 replicas. The final tree was plotted in Mega (ver 5.05) software and bootstrap values were mentioned on the tree.

Secondary structure analysis and fold recognition
Secondary structures of monothiol glutaredoxin in *Ashbya gossypii* were predicted using JNET **[23]**, SABLE **[24]**, STRIDE **[25]**, PSIPRED **[26]** and SAM-T08 **[27]** softwares. Fold-recognition analysis was carried out using mGENETHREADER **[28]**, and 3DPSSM **[29]** softwares. The architectural motifs and the topology of proteins with known 3-D structure were analysed according to SCOP **[30]** and CATH **[31]** classifications. Topology of the modelled monothiol glutaredoxin protein in *Ashbya gossypii* was analysed using PDBSum (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum).

Comparative modelling and analysis
The comparative modelling of 3-D structure of monothiol glutaredoxin protein in *Ashbya gossypii* was performed in a stepwise procedure and templates were searched from PDB database using BLASTp program. The crystal structure of human glutaredoxin with bound glutathione in FeS cluster (PDB entry: 2WUL) and Arabidopsis monothiol glutaredoxin (PDB entry: 3IPZ) were identified as the template for modelling monothiol glutaredoxin domain in *Ashbya gossypii*. The template structures were downloaded and aligned using STAMP, structure alignment software, STAMP **[32]**. These aligned structures were used as a profile for aligning the target sequence using ClustalX **[33]**. The automated comparative protein modelling program MODELLER9v10 **[34]** was used to generate a 100 all-atom model by alignment of the target sequence with the selected template sequences in an alignment file. Molecular visualisation and analysis of the final modelled structure were carried out with Visual Molecular Dynamics (VMD) (http://www.ks.uiuc.edu/Research/vmd/) **[35]**.

Validation of AgGRX1 structure
The best structure model was chosen on the basis of the stereochemistry quality report generated using PROCHECK (used for inspection of ψ/ϕ Ramachandran plot) **[36]** as shown in **Table 1 (see supplementary material)**. The spatial restraints and the energy minimisation steps were performed within Modeller using the CHARMM22 force field for proper stereochemistry of proteins. Further, PROSA-web test was applied on final model to check for energy criteria in comparison with the potential of mean force derived from a large set of known protein structures **[37]**.

Docking analysis
The glutathione (GSH) molecule was docked into the active site of the modelled protein using PatchDock software **[38]**. The software works on the principle of finding the suitable docking transformations that yields molecular shape complementarity using geometry based docking algorithm. The docking was performed with the default parameters.

Discussion:

Sequence analysis of AgGRX1 monothiol glutaredoxin

The protein sequences of monothiol glutaredoxin in *Ashbya gossypii* has been identified using Blast and hmmer software (see Methodology) using whole genome protein sequence of *Ashbya gossypii*. The search resulted in the identification of two sequences with accession number NP_984149.1 and NP_986777.1 which are subsequently named as AgGRX1 and AgGRX2 respectively. The identified protein sequences were further validated as monothiol glutaredoxins using their signatures "CGFS". The protein sequences of AgGRX1 and AgGRX2 were classified as monothiol glutaredoxins where cysteine is conserved only at the first position of redox active motif. However, cysteines at the first and fourth position of the redox active motif like 'CXXC', is conserved in dithiol GRXs. The proteins sequence of AgGRX1 and AgGRX2 were found to be of 138 and 237 amino acids respectively. The AgGRX1 showed 68% identity with 'CGFS' type glutaredoxin of *S.cerevisiae* i.e. ScGRX5 while AgGRX2 showed 65% identity with both ScGRX4 and ScGRX3 respectively. AgGRX1 and AgGRX2 have a molecular weight of 15.0 kDa and 26.0 kDa respectively. Their theoretical pI's are 4.76 and 4.52 respectively. The AgGRX1 possesses a 32 amino acid transit peptide for mitochondrial localization while predicted localization of AgGRX2 is cytoplasmic **Table 2 (see supplementary material).** Analysis of the GRX module present in AgGRX1 and AgGRX2 showed single instance of their occurrence on the protein sequence. Recently, members of monothiol glutaredoxins having 'CXYS' motif in the active site has been identified namely, ScGrx6 (CSYS) and ScGrx7 (CPYS) **[39].** The genome of *S. cerevisiae* and *S. pombe* contains approximately 5 and 3 monothiol GRXs, respectively **[40, 41].** In *S. cerevisiae*, ScGRX6, ScGRX7 and SpGRX3 has single GRX module of 'CS/PYS' type. Similar single GRX module of 'CGFS' type were earlier found to be present in *S. cerevisiae* (ScGRX3, ScGRX4, ScGRX5) and *S. pombe* (SpGRX4 and SpGRX5) **[42].** Some cases of convergent evolution in domain architecture have been identified, where two independent recombination events of a TRX domain to a GRX domain have occurred **[43].** The TRX domain in hybrid molecules helps in the localization of the protein. Earlier reports

shows that ScGRX3, ScGRX4 and SpGRX4 belongs to the TRX-GRX class and Posses a WAxPC motif **[5].** In ScGRX3, the TRX domain seems to be required for the nuclear targeting of the molecule **[44].** Further analysis showed the presence of WAxPC motif at the N-terminal region of AgGRX2 while in AgGRX1 the motif was found missing in the protein sequence. In order to analyze the conservation in AgGRX1 and AgGRX2, the protein sequences were aligned with the representative monothiol glutaredoxin protein sequences from prokaryotes (*E.coli, Thiobacillus denitrificans and Pseudomonas aeruginosa*), fungi (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*), animals (*Xenopus tropicalis, Drosophila melanogaster and Homo sapiens*), algae (*Chlamydomonas reinhardtii, Volvox carteri, Micromonas pusilla and Ostreococcus lucimarinus*) and plants (*Arabidopsis thaliana, Zea mays, Pteris vittata, Populus trichocarpa, Ricinus communis, Oryza sativa, Glycine max and Physcomitrella patens*). Their protein entry code, protein length and putative localization have been shown in **Table 2**.

The resulting multiple sequence alignment showed that 'CGFS' being the redox active motif was highly conserved in AgGRX1 and AgGRX2 **(Figure 1)**. Glutaredoxins contain a putative glutathione binding site which consists of a glycine pair at C-terminal. The putative glutathione binding site (V/QGG) was present in both AgGRX1 and AgGRX2. 'IGG' was reported as a putative glutathione binding site in Grx2 of human, mouse and rat **[45].** In fungal glutaredoxins, towards C-terminal region, two consecutive glycine residues assist in the formation of the GSH cleft of the GRX molecule **[5].** In all the sequences analysed, tripeptide 'PAK' is present only in AgGRX1 and yeast GRX5 but was absent in AgGRX2. A conserved stretch of '**WPT**[I/F]**PQL**' was observed in both the sequences and was not found in the sequences of *H.sapiens, D. melanogaster, X. tropicalis, C. reinhardtii, V. carteri, O. sativa, G. max and M. pusilla.* Similar stretch was also reported in *Synechocystis*, which assists during glutathione binding **[46].** The study of the fungal monothiol GRXs showed diversity at sequence level, but the redox active motif and invariable residues involved in the glutathione binding were conserved.
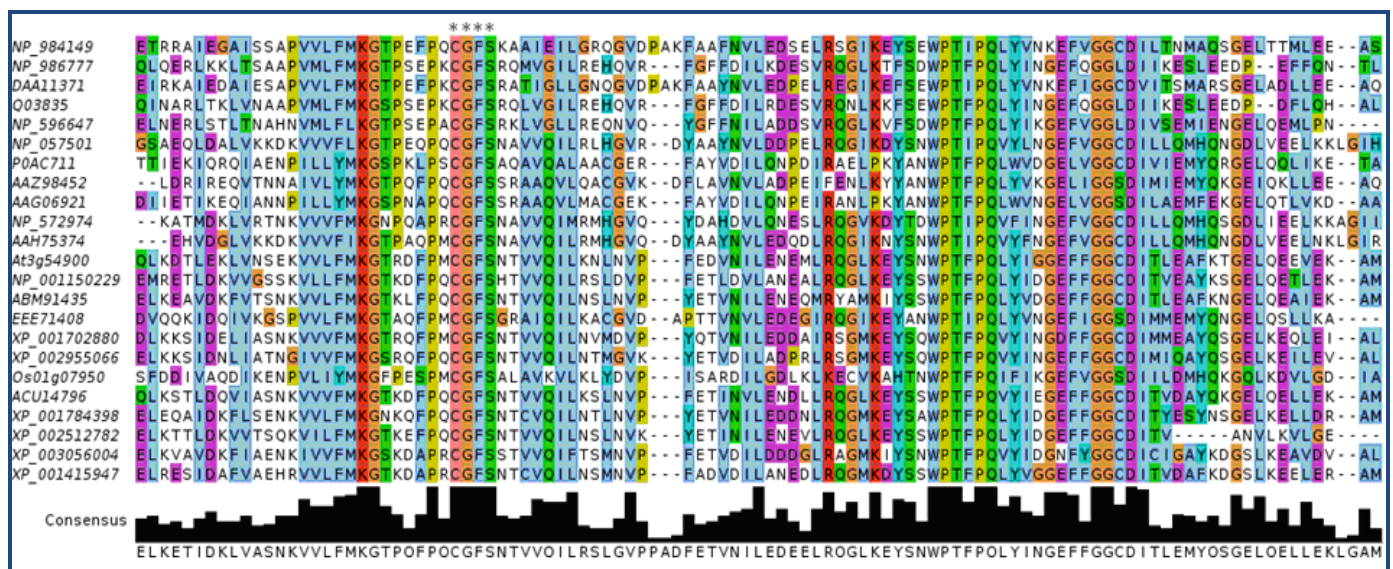

Figure 1: Multiple sequence alignment of AgGRX1 (NP_984149), AgGRX2 (NP_986777) with- *Saccharomyces cerevisiae* (Q03835, DAA11371), *Schizosaccharomyces pombe* (NP596647), *Homo sapiens* (NP057501), *E.coli* (P0AC71), *Thiobacillus denitrificans* (AAZ98452),

BIOINFORMATION

Pseudomonas aeruginosa (AAG06921), *Drosophila melanogaster* (NP572974), *Xenopus tropicalis* (AAH75374), *Arabidopsis thaliana* (At3g54900), *Zea mays* (NP001150229), *Pteris vittata* (ABM91435), *Populus trichocarpa* (EEE71408), *Chlamydomonas reinhardtii* (XP001702880), *Volvox carteri* (XP002955066), *Oryza sativa* (Os01g07950), *Glycine max* (ACU14796), *Physcomitrella patens* (XP001784398), *Ricinus communis* (XP002512782), *Micromonas pusilla* (XP003056004), *Ostreococcus lucimarinus* (XP001415947). The redox active motif 'CGFS' is marked by asterisks*. The conservation and consensus sequence is depicted below the alignment.
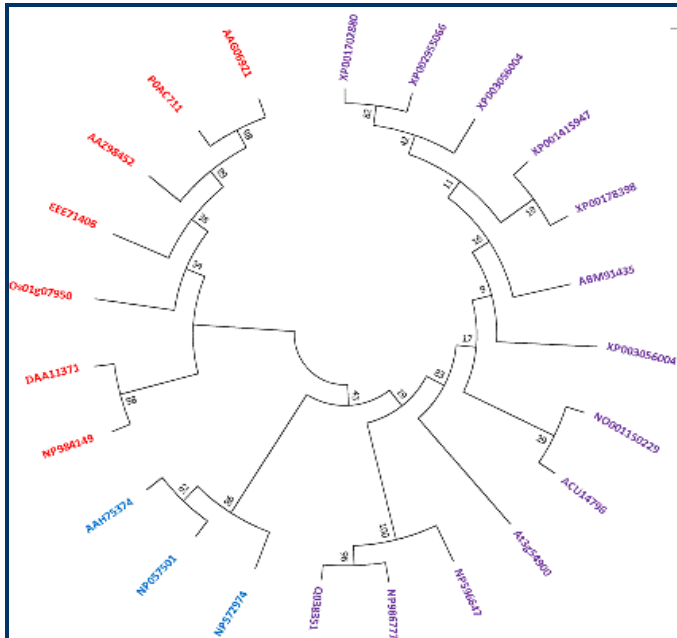

Figure 2: Phylogenetic tree of monothiol glutaredoxins from different organisms showing various clades. Clade A: Purple colour label, Clade B: Blue colour labels, Clade C: Red colour labels

Phylogenetic analysis of 'CGFS' GRX proteins

Phylogenetic and evolutionary relationships among monothiol glutaredoxins of *Ashbya gossypii* and its homologues were investigated using the full-length protein sequences to construct an unrooted phylogenetic tree. The tree was sub-divided into three distinct clades: 'A', 'B', and 'C' **(Figure 2)**. The largest clade 'A' (purple color labels) encompasses AgGRX2 along with algae (*Chlamydomonas reinhardtii, Volvox carteri, Micromonas pusilla and Ostreococcus lucimarinus*), plants (*Arabidopsis thaliana, Zea mays, Pteris vittata, Ricinus communis, Glycine max and Physcomitrella patens*) and fungi (ScGRX3 and SpGRX4). In the clade 'B' (Blue color labels) - animals (*Xenopus tropicalis, Drosophila melanogaster and Homo sapiens*), were clustered together. Clade 'C' (red color labels) had AgGRX1, prokaryotes (*E.coli, Thiobacillus denitrificans and Pseudomonas aeruginosa*), Yeast GRX5, plants (*Populus trichocarpa and Oryza sativa*). AgGRX1 and AgGRX2 were significantly divergent and grouped under different clades. Similarly, Yeast GRX3 and GRX5 were placed in different clades. AgGRX2, ScGRX3 and SpGRX4 contain 'WaxPC' motif and belong to TRX-GRX hybrid class and all were placed in clade 'A'. AgGRX1, Yeast GRX5 and *Oryza sativa* were having mitochondrial localization signal and clustered together in clade 'C'.

Comparative modelling of AgGRX1 protein

The structure model of monothiol glutaredoxin protein in *Ashbya gossypii*, BLAST searches were performed against the PDB for proteins with similar sequence and known 3D structures using the protein sequence of AgGRX1. The search identified crystal structure of human glutaredoxin with bound

glutathione in FeS cluster (PDB entry: 2WUL) and Arabidopsis monothiol glutaredoxin (PDB entry: 3IPZ) were identified as the template for modelling monothiol glutaredoxin domain in *Ashbya gossypii* (see Materials and Methods). The Pfam analysis of AgGRX1 showed the presence of glutaredoxin domain (89 to 153 amino acids). Due to the lack of template structure for the N-terminal region (1-88 amino acids), only glutaredoxin domain of the protein has been modelled. The Ramachandran plot for the modelled domain of AgGRX1 showed 93.5% residues in most favourable regions with the remaining 6.5% of residues occurring in allowed regions while none of the residues were found to be in generously allowed and disallowed region **(Figure 3A)**. The PROCHECK result summary showed none out of 109 residues labelled. The torsion angles of the side chain designated by $\chi1$-$\chi2$ plots showed only 2 labelled residues out of 60 in the modelled structure of AgGRX1. The observed G-factor scores of the modelled AgGRX1 glutaredoxin domain were found to be 0.07 for dihedral bonds, -0.26 for covalent bonds and -0.05 overall. The distribution of the main chain bond lengths and bond angles were 98.7% within limits for the modelled AgGRX1 protein structure. The PROSA-web energy plots for modelled glutaredoxin domain in AgGRX1 protein showed a z-score for pair, surface and combined energy which was found to be -7.39 **(Figure 3B & Figure 3C)**.
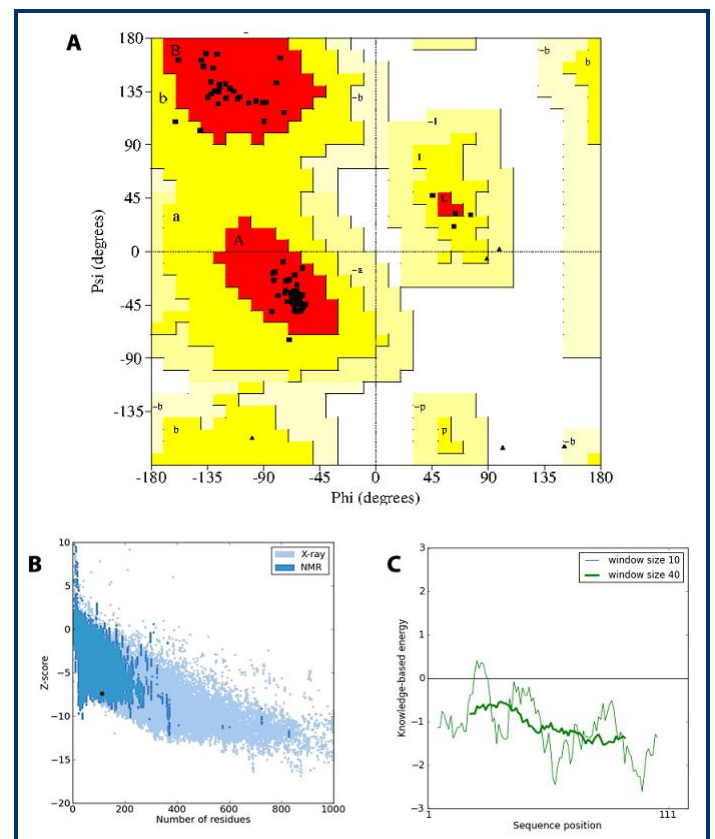

Figure 3: Ramachandran plot of the modelled AgGRX1 protein (**A**) Z-plot (**B**) and e-plot (**C**).
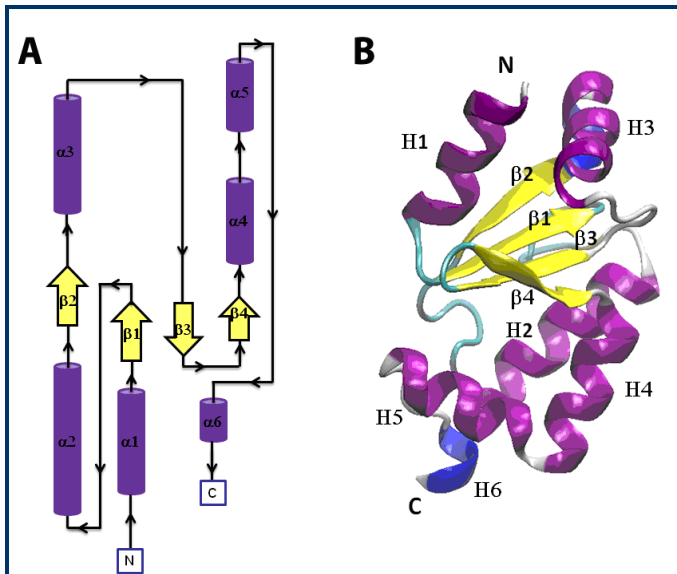
Figure 4: Secondary structure topology of AgGRX1 glutaredoxin protein. **A)** Topology diagram of AgGRX1 protein, the four sheets and six helices are represented in purple and yellow colour respectively; **B)** Cartoon representation of the modelled AgGRX1protein with the labelled secondary structures.
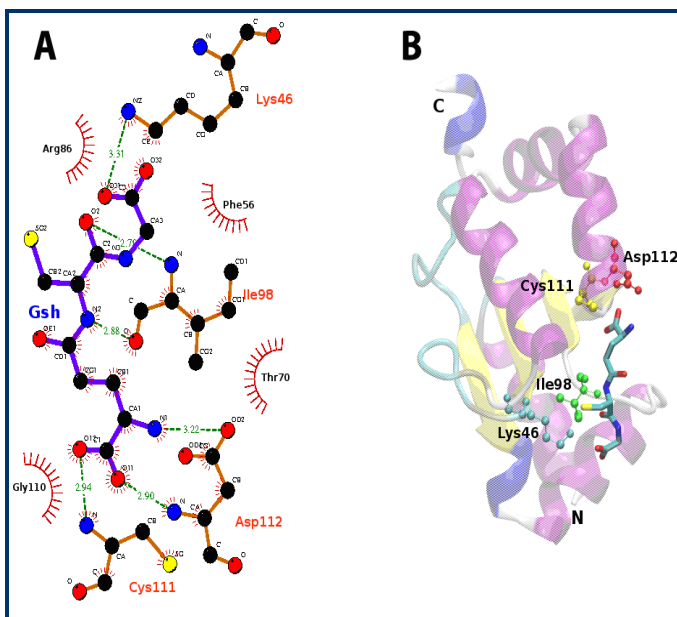

Figure 5: **A)** Glutathione binding to the modelled AgGRX1 domain showing hydrogen bonds and other non-bonded interactions; **B)** Cartoon representation of the modelled AgGRX1protein docked with glutathione molecule showing residues forming hydrogen bonds with glutathione.

Analysis of modelled AgGRX1 protein

AgGRX1 model shows the presence of four stranded β-sheet flanked by six α-helices **(Figure 4)**. In AgGRX1, redox active motif 'CGFS' was present at the start of α2, putative glutathione binding sites '**WPT**[I/F]**PQL**' was between α3 and β3 and 'VGG' was located between α4 and β4 structures. Tripeptide 'PAK' which was unique to AgGRX1 was located between α2 and β2 structures **(Figure 1)**. In AgGRX1 model there is a core of four mixed β-strands out of which, three (β1, β2 and β4) are

parallel and one (β3) remains antiparallel **(Figure 4A & Figure 4B)**. The AgGRX1 model of glutaredoxin domain was similar to other known monothiol GRXs. The overall fold of the AgGRX1 glutaredoxin domain remains close to previously known thioredoxins with (αβαβαββαα) topology, thus having a signature fold of thioredoxin family with central mixed β-sheet sandwiched between α-helices **[47]**

Docking of AgGRX1 with glutathione (GSH) molecule

In order to understand the mechanism of monothiol glutaredoxins in maintaining the cellular redox homeostasis, the AgGRX1 protein was docked with the glutathione molecule using PatchDock software (See materials and Methods). Docking of glutathione molecule into the active site of AgGRX1 revealed the hydrogen bond interaction of Lys46, Ile98, Cys111 and Asp112 **(Figure 5A & Figure 5B)**. Similar interaction of the glutathione with the active site cysteine was observed in human Grx2 proteins **[48]**. Other residues such as Phe56, Thr97, Arg86 and Gly110 were found to have non-bonded interaction with the glutathione molecule **(Figure 5A)**. All the residues were found to be well conserved in the multiple sequence alignment of the monothiol glutaredoxins **(Figure 1)**. Earlier, residues such as Lys23, Thr71 and Asp86 were found to be located in the GRX cleft and had role in glutathione binding in *Synechocystis* SyGrx3 protein **[46]**.

Conclusion:

Ashbya gossypii is one of the filamentous fungi and known for its commercial importance as well as for causing the disease in various crop plants. With the genome sequence available for the fungi, various important processes can be understood in the fungus which may help in harnessing it for better commercial success. Oxidative stress is common for all the aerobic species and various organisms develop intrinsic mechanisms in order to survive under this stress condition. Glutaredoxins are one amongst the antioxidant system which deals with oxidative stress and is well conserved in prokaryotes and eukaryotes. The analysis of "CGFS" monothiol glutaredoxins of *Ashbya gossypii* can shed light on its antioxidant mechanism. The protein sequence of the AgGRX1 was used for modelling and analysis of monothiol glutaredoxins in *Ashbya gossypii*. The sequence alignment of the protein sequence revealed the conserved residues of monothiol glutaredoxins in AgGRX1. The protein sequence of AgGRX1 was found close to the *Saccharomyces cerevisiae* monothiol glutaredoxin. The structure model of AgGRX1 protein showed important features about the structural aspect of the protein and the topology resembles with thioredoxin molecules. Further, the docking analysis of glutathione molecule in the active site of GRXs revealed the probable interaction of the molecules with various residues. The cysteine which is considered as relevant for the glutathione binding as well as protein dimerization, was also found to be conserved in AgGRX1. This study will help in the understanding about the structural account of fungal monothiol glutaredoxins, its interaction with glutathione molecule and add valuable information about its biochemical function and interaction properties in molecular detail.

BIOINFORMATION

References:
[1] Ashby SF & Nowell W, *Ann Bot.* 1926 **1**: 69
[2] Wendland J & Walther A, *Nat Rev Microbiol.* 2005 **3**: 421 [PMID: 15821727]
[3] Dietrich FS *et al. Science.* 2004 **304**: 304 [PMID: 15001715]
[4] Storz G & Imlay JA, *Curr Opin Microbiol.* 1999 **2**: 188 [PMID: 10322176]
[5] Herrero E & de la Torre-Ruiz MA, *Cell Mol Life Sci.* 2007 **64**: 1518 [PMID: 17415523]
[6] Foyer CH *et al. Plant Cell Environ* 1994 **17**: 507
[7] Neill SJ *et al. J Exp Bot.* 2002 **53**: 1237 [PMID: 11997372]
[8] Rouhier N *et al. Cell Mol Life Sci.* 2004 **61**: 1266 [PMID: 15170506]
[9] Rouhier N *et al. Annu Rev Plant Biol.* 2008 **59**: 143 [PMID: 18444899]
[10] Couturier J *et al. Cell Mol Life Sci.* 2009 **66**: 2539 [PMID: 19506802]
[11] Lillig CH *et al. Biochim Biophys Acta.* 2008 **1780**: 1304 [PMID: 18621099]
[12] Nikkola M *et al. J Biol Chem.* 1991 **266**: 16105 [PMID: 1874748]
[13] Ghezzi P, *Biochem Soc Trans.* 2005 **33**: 1378 [PMID: 16246123]
[14] Ken CF *et al. J Agric Food Chem.* 2011 **59**: 3828 [PMID: 21395221]
[15] Mühlenhoff U *et al. EMBO J.* 2003 **22**: 4815 [PMID: 12970193]
[16] Izquierdo A *et al. Eukaryot Cell.* 2008 **7**: 1415 [PMID: 18503006]
[17] Kim KD *et al. Biochem Biophys Res Commun.* 2011 **408**: 609 [PMID: 21531205]
[18] Altschul SF *et al. J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]
[19] Edgar RC, *Nucleic Acids Res.* 2004 **32**: 1792 [PMCID: PMC390337]
[20] Emanuelsson O *et al. J Mol Biol.* 2000 **300**: 1005 [PMID: 10891285]
[21] Nakai K & Horton P, *Trends Biochem Sci.* 1999 **24**: 34 [PMID: 10087920]
[22] Tamura K *et al. Mol Biol Evol.* 2011 **28**: 2731 [PMID: 21546353]
[23] Cuff JA & Barton GJ, *Proteins.* 1999 **40**: 502 [PMID: 10861942]
[24] Adamczak R *et al. Proteins.* 2004 **56**: 753 [PMID: 15281128]
[25] Heinig M & Frishman D, *Nucleic Acids Res.* 2004 **32**: W500 [PMID: 15215436]
[26] McGuffin LJ *et al. Bioinformatics.* 2000 **16**: 404 [PMID: 10869041]
[27] Karplus K, *Nucleic Acids Res.* 2009 **37**: W492 [PMCID: PMC2703928]
[28] Jones DT, *J Mol Biol.* 1999 **287**: 797 [PMID: 10191147]
[29] Kelley LA *et al. J Mol Biol.* 2000 **299**: 499 [PMID: 10860755]
[30] Murzin AG *et al. J Mol Biol.* 1995 **247**: 536 [PMID: 7723011]
[31] Orengo CA *et al. Structure.* 1997 **5**: 1093 [PMID: 9309224]
[32] Russell RB & Barton GJ, *Proteins.* 1992 **14**: 309 [PMID: 1409577]
[33] Larkin MA *et al. Bioinformatics.* 2007 **23**: 2947 [PMID: 17846036]
[34] Martí-Renom MA *et al. Annu Rev Biophys Biomol Struct.* 2000 **29**: 291 [PMID: 10940251]
[35] Humphrey W *et al. J Mol Graph.* 1996 **14**: 33 [PMID: 8744570]
[36] Laskowski RA *et al. J. Appl. Cryst.* **26**: 283
[37] Wiederstein M & Sippl MJ, *Nucleic Acids Res.* 2007 **35**: W407 [PMID: 17517781]
[38] Schneidman-Duhovny D *et al. Nucleic Acids Res.* 2005 **33**: W363 [PMID: 15980490]
[39] Luo M *et al. J Mol Biol.* 2010 **398**: 614 [PMID: 20347849]
[40] Mesecke N *et al. Biochemistry.* 2008 **47**: 1452 [PMID: 18171082]
[41] Chung WH *et al. Biochem Biophys Res Commun.* 2005 **330**: 604 [PMID: 15796926]
[42] Rodríguez-Manzaneque MT *et al. Mol Cell Biol.* 1999 **19**: 8180 [PMID: 10567543]
[43] Alves R *et al. BMC Evol Biol.* 2009 **9**: 66 [PMID: 19321008]
[44] Molina MM *J Biol Chem.* 2004 **279**: 51923 [PMID: 15456753]
[45] Lundberg M *et al. J Biol Chem.* 2001 **276**: 26269 [PMID: 11297543]
[46] Picciocchi A *et al. Biochemistr.* 2007 **46**: 15018 [PMID: 18044966]
[47] Qi Y & Grishin NV, *Proteins.* 2005 **58**: 376 [PMID: 15558583]
[48] Johansson C *et al. J Biol Chem.* 2007 **282**: 3077 [PMID: 17121859]

BIOINFORMATION

Supplementary material:

Table 1: Ramachandran plot analysis of different models by PROCHECK.

Models	Favourable (%)	Allowed (%)	Generously Allowed (%)	Disallowed (%)	G-factor overall
Model 1	93.5	6.5	0.0	0.0	-0.05
Model 2	93.5	5.4	1.1	0.0	-0.05
Model 3	92.5	6.5	1.1	0.0	-0.05
Model 4	91.4	6.5	2.2	0.0	-0.09
Model 5	94.6	5.4	0.0	0.0	-0.12
Model 6	91.4	6.5	2.2	0.0	-0.16
Model 7	91.4	7.5	1.1	0.0	-0.17
Model 8	89.2	7.5	1.1	2.2	-0.15
Model 9	92.5	7.5	0.0	0.0	-0.06
Model 10	92.5	6.5	0.0	1.1	-0.05

Table 2: The classification of 'CGFS' monothiol glutaredoxins from various fung.

Protein code	Organism name	Protein length	Putative localization	Active site sequence	Kingdom (Class)
NP984149	*Ashbya gossypii*	138	Mitochondria	CGFS	Fungi (Saccharomycetes)
NP986777	*Ashbya gossypii*	237	Cytoplasm	CGFS	Fungi (Saccharomycetes)
Q038351	*Saccharomyces cerevisiae*	285	Nucleus	CGFS	Fungi (Saccharomycetes)
DAA11371	*Saccharomyces cerevisiae*	150	Mitochondria	CGFS	Fungi (Saccharomycetes)
NP596647	*Schizosaccharomyces pombe*	244	Nucleus	CGFS	Fungi (Schizosaccharomycetes)
NP057501	*Homo sapiens*	157	Cytoplasm	CGFS	Animalia (Mammalia)
P0AC711	*E.coli*	115	Cytoplasm	CGFS	Eubacteria (Gammaproteobacteria)
AAZ98452	*Thiobacillus denitrificans*	105	Cytoplasm	CGFS	Bacteria (Betaproteobacteria)
AAG06921	*Pseudomonas aeruginosa*	108	Cytoplasm	CGFS	Bacteria (Gammaproteobacteria)
NP572974	*Drosophila melanogaster*	159	Mitochondria	CGFS	Animalia (Insecta)
AAH75374	*Xenopus tropicalis*	154	Mitochondria	CGFS	Animalia (Amphibia)
At3g54900	*Arabidopsis thaliana*	173	Chloroplast	CGFS	Plantae (Magnoliopsida)
NP001150229	*Zea mays*	172	Chloroplast	CGFS	Plantae (Liliopsida)
ABM91435	*Pteris vittata*	184	Chloroplast	CGFS	Plantae (Pteridopsida)
EEE71408	*Populus trichocarpa*	103	Cytoplasm	CGFS	Plantae (Magnoliopsida)
XP001702880	*Chlamydomonas reinhardtii*	148	Mitochondria	CGFS	Protista (Chlorophyceae)
XP002955066	*Volvox carteri*	107	Chloroplast	CGFS	Plantae (Chlorophyceae)
Os01g07950	*Oryza sativa*	185	Mitochondria	CGFS	Plantae (Liliopsida)
ACU14796	*Glycine max*	177	Chloroplast	CGFS	Plantae (Magnoliopsida)
XP001784398	*Physcomitrella patens*	116	Nucleus	CGFS	Plantae (Bryopsida)
XP002512782	*Ricinus communis*	159	Chloroplast	CGFS	Plantae (Magnoliopsida)
XP003056004	*Micromonas pusilla*	158	Mitochondria	CGFS	Viridiplantae (Prasinophyceae)
XP001415947	*Ostreococcus lucimarinus*	107	Cytoplasm	CGFS	Plantae (Prasinophyceae)