

## Digital dissection of arsenate reductase enzyme from an arsenic hyperaccumulating fern *Pteris vittata*

Zarrin Basharat\*<sup>1</sup>, Deeba Noreen Baig<sup>2</sup>, Azra Yasmin<sup>1</sup>

<sup>1</sup>Microbiology & Biotechnology Research Lab, Department of Environmental Sciences, Fatima Jinnah Women University, 46000 Rawalpindi, Pakistan.

<sup>2</sup>Department of Biological Sciences, Forman Christain College (A Chartered University), 54600 Lahore, Pakistan.

**\*Corresponding author**

**Email: zarrin.iiui@gmail.com**

**NOTE:** *This is version 1 of the preprint and may contain grammatical and proof reading mistakes. Errors and omissions excepted.*

## Abstract

Action of arsenate reductase is crucial for the survival of an organism in arsenic polluted area. *Pteris vittata*, also known as Chinese ladder brake, was the first identified arsenic hyperaccumulating fern with the capability to convert [As(V)] to arsenite [As(III)]. This study aims at sequence analysis of the most important protein of the arsenic reduction mechanism in this specie. Phosphorylation potential of the protein along with possible interplay of phosphorylation with *O*- $\beta$ -GlcNAcylation was predicted using neural network based webservers. Secondary and tertiary structure of arsenate reductase was then analysed. Active site region of the protein comprised a rhodanese-like domain. Cursor dynamics simulation revealed that folds remained conserved in the rhodanese main but variations were observed in the structure in other regions. This information sheds light on the various characteristics of the protein and may be useful to enzymologists working on the improvement of its traits for arsenic reduction.

**Keywords:** *Pteris vittata*, arsenate reductase, structure, phosphorylation.

## Introduction

Arsenic contamination in soil and groundwater poses significant threat to human health and cleanup of arsenic contaminated soil and water using phytoremediation approach is a cost effective strategy (Meadows, 2014). Presence of arsenic in the environment has led to the adaptation of various prokaryotic and eukaryotic species with a repertoire of enzymes for detoxification of this metalloid and hence, aid cleanup of the environment. The initial and crucial step of arsenic reduction involves conversion of arsenate [As(V)] to arsenite [As(III)]. This is accomplished by arsenate reductase enzyme which catalyzes reduction of [As(V)] to [As(III)] and vice versa (Martin *et al.*, 2001). Arsenate reductase belongs to the family of oxidoreductase enzymes, especially the ones that act with phosphorus or arsenic being donor and disulfide acting as acceptor. Similarity of arsenate to phosphate makes it perfect for ready plant uptake of arsenate as it is mistaken for plant nutrient group phosphate (Ullrich-Eberius *et al.*, 1989; Meharg and McNair, 1992; Dhankher *et al.*, 2006).

*Pteris vittata*, also known as the Chinese brake or Ladder brake, is an arsenic hyperaccumulating fern and used as a model in arsenic pollution mitigation studies (Wilkins and Salter, 2003). Root uptake of [As(V)] is followed by reduction to [As(III)]. [As(III)] is then transported to the lamina of the *Pteris vittata* frond and deposited in the cells as free arsenite (Indriolo *et al.*, 2010). The frond can hold up to 22,630 mg As kg<sup>-1</sup> (Wang, 2002). *Physcomitrella patens* is a moss species and belongs to one of the oldest species of plants that inhabited land (Heckman, 2001). Sequence, topology, structure and functional information is critical for the master regulator of arsenic hyperaccumulation in this plant as the expression, pathway and pollutant homeostasis network associated with it, that supports hyperaccumulation or tolerance can be important for environmental cleanup (Basharat and Yasmin, 2016a). This rationale of this study was to analyse arsenate reductase in this

hyperaccumulating specie, using informatics tools, which is a quick and handy method of studying protein of interest.

## Material and Methods

Sequence comprising of 435 amino acid residues of arsenate reductase for *Pteris vittata* was retrieved from NCBI database with accession no: ABD64881. ProtParam tool (<http://web.expasy.org/protparam>) was used for physiochemical property analysis of the protein, while subcellular localization was calculated using CELLO v.2.5 (<http://cello.life.nctu.edu.tw/>) as previously described (Basharat, 2015). Conserved domains were studied using Conserved domain database (Marchler-Bauer *et al.*, 2012). Position of phosphorylation sites in arsenate reductase of both species was studied using neural network based tool NetPhos 2.0 (Blom *et al.*, 1999; Basharat and Yasmin, 2016b; Basharat and Yasmin, 2016c). It makes predictions for possible phosphorylation potential at serine, threonine and tyrosine sites in the eukaryotic proteins. Kinases responsible for phosphorylation were predicted using another neural network based tool NetPhosK (Blom *et al.*, 2004). Possible serine or threonine sites that could act as substrates for phosphate as well as showed potential for *O*- $\beta$ -GlcNAcylation were predicted using YinOYang server (Gupta and Brunak, 2001).

1D structure was studied using Protter (<http://wlab.ethz.ch/protter>). Secondary structure motifs, wiring diagram of the protein was obtained by PROMOTIF program (Hutchison and Thronton, 1996) implemented at PDBSum server web interface (Laskowski, 2001; Laskowski *et al.*, 1997; Laskowski, 2004; Laskowski, 2009). The tertiary structure of arsenate reductase was modeled using I-TASSER (Zhang, 2004; Roy *et al.*, 2010; Yang *et al.*, 2014). Models were built on the template structures of proteins from PDB library, identified by LOMETS (Wu and Zhang, 2007) and ranked as best template by threading

programs in I-TASSER. Final model was chosen based on confidence or C-score, centred on the comparative clustering of structural density as well as consensus significance score of multiple threading templates. Structure quality assessment was done by Ramachandran plot analysis (Ramachandran *et al.*, 1963; Morris *et al.*, 1992; Lovell *et al.*, 2003). Percentage of residues in favourable regions was calculated. ERRAT webserver (Colovos and Yeates, 1993) was also used to assess quality of structure. Model was rendered in Pymol v1.4.1. Missing hydrogens were added and energy minimization was performed with Molecular operating Environment (MOE). Dynamics simulation was carried out for 500 ps using MOE. GROMOS forcefield was applied, with NVT statistical ensemble for conformation generation and Nose-Poincare-Anderson algorithm (Bond, 1999) to solve the equation of motion. Timestep of 2 fs was selected for discretizing the equation of motion. Temperature (300K) and pressure (101 KPa) remained constant.

## Results and Discussion

In plants, few detailed studies have been conducted on arsenate reductase enzyme. Arsenate reductase has been reported in *Arabidopsis thaliana* (Chao *et al.*, 2014; Sánchez-Bermejo *et al.*, 2014), *Oryza sativa* (Duan *et al.*, 2007), green alga *Chlamydomonas reinhardtii* (Yin *et al.*, 2011) and the arsenic hyperaccumulator *Pteris vittata* (Duan *et al.*, 2011; Ellis *et al.*, 2006; Cesaro *et al.*, 2015) till to date. Here, *in silico* analysis has been attempted for this enzyme at sequence and structure level. Molecular weight of the protein was determined as 45271.7 daltons, theoretical pI: 9.31, total number of negatively charged residues (Asp + Glu): 37, total number of positively charged residues (Arg + Lys): 46, instability index (II): 29.98 (which classified the protein as stable). Subcellular localization analysis based on support vector machine technique revealed that the protein resided in the chloroplast region.

Protein phosphorylation represents the addition of a phosphate group to the amino acid side chain of an amino acid by a protein kinase (Tan, 2011). A phosphorylated amino acid may be dephosphorylated by a phosphatase or replaced by some other post translational modification as *O*- $\beta$ -GlcNAcylation (Basharat and Yasmin, 2015). Phosphorylation can modulate the enzymatic activity, so it is necessary to study phosphorylation profile for gaining insights into protein function. A total of 34 phosphosites (19 serine, 10 threonine, 5 tyrosine) were predicted in *Pteris vittata* arsenate reductase while interplay of phosphosites with *O*- $\beta$ -glycosylation was observed at six sites (Table 1). Conserved domain analysis revealed that a rhodanese-like homology domain was present in the sequence at position 262-376 with an E-value threshold of  $1.00 \times 10^{-08}$ . Rhodanese-like domain is an alpha beta fold domain with a cysteine residue comprising active site, characteristic of stress proteins as cyanide and arsenate resistance proteins (Bordo, 2002).

1D plot revealed that the protein has a transmembrane region (residue 224-249) (Fig. 1). Secondary structure analysis disclosed presence of 1 sheet, 1 beta hairpin, 3 strands, 26 helices, 53 helix-helix interactions, 43 beta turns and 9 gamma turns (Fig. 2). 3D structure (Fig. 3) quality analysis using Ramachandran plot analysis revealed that 85.0% residues existed in the favoured, 10.2% in allowed regions while 4.8% as outliers. ERRAT displayed a value of 94.84 (~95%). This indicated the good quality of structure as good resolution structures have an ERRAT value of around 95% or higher. Short timescale dynamics simulation was carried out to study initial changes in protein structure conformation. Simulation analysis showed C $\alpha$  (protein backbone) fluctuations and large variation in protein conformation was observed (Fig. 4). Larger timescale dynamics simulation is further required to study protein behaviour and map fluctuations and reoccurrence of native state accurately.

This study is a useful addition to arsenate reductase knowledgebase and has shows that bioinformatics analysis of a protein sequence is a quick method for inferring functional

properties. It also paves way for investing future research efforts focusing on comparative analysis of close and distant homologs as a wealth of information for this enzyme lies unexplored in other plants that have recently been sequenced. Evolutionary dynamics of protein phosphorylation occurring at the residue level should also be attempted. Analysis of protein-protein interactions with arsenate reductase helper enzymes is currently being conducted by the authors, which can further enhance the understanding of evolution and its impact in this enzyme system in closely related species. Experimental research in this domain is proposed and would definitely lead to new post translational modification prospects in arsenic reduction phenomenon in hyperaccumulators, as well as other plants.

## Figures/Table Legends

Table 1. Phosphorylation and *O*- $\beta$ -GlcNAcylation potential of arsenate reductases in *Pteris vittata*. – indicates a complete absence or lack of information for specified residues. + indicates presence of a particular trait in the enzyme at mentioned site.

Fig. 1. Topology of the arsenate reductase protein. Red colour shows the signal peptide region. N-glyco motif is shown in green colour.

Fig. 2. Schematic secondary structure wiring diagram of arsenate reductase for *Pteris vittata*.


Helices  are labelled as H1, H2, ... and strands by their sheets A and B. Motifs  $\beta$  and  $\gamma$  turns are also labelled.

Fig. 3. Modeled 3D structure of *Pteris vittata*. C-terminal is depicted by red and N-terminus by blue colour. Phosphorylated sites are depicted by spheres while rhodanese-like domain is encircled.

Fig. 4. (A) Correlation plot of energy of the simulated protein versus time. (B)  $C\alpha$  fluctuation of native (green) and simulated (green) protein state for 500 ps. Beta sheets (residue 55-58 and 61-65) conform to loops. Helices in the rhodanes domain remained stable whereas helix at the C-terminal lost conformation to form loop (encircled).



## References

- Basharat, Z., Yasmin, A. 2015. *In silico* assessment of phosphorylation and O- $\beta$ -GlcNAcylation sites in human NPC1 protein critical for Ebola virus entry. *Infect. Genet. Evol.* 34, 326-338.
- Basharat, Z. 2015. Computational sequence analysis and *in silico* modeling of a stripe rust resistance protein encoded by wheat TaHSC70 gene. *J. Comp. Bio. Bioinfo. Res.* 7, 1-7.
- Basharat, Z., Yasmin, A., 2016a. Genome editing weds CRISPR: what is in it for phytoremediation?. *PeerJ Preprints*, e2034v1.
- Basharat, Z. and Yasmin, A., 2016b. Understanding properties of the master effector of phage shock operon in *Mycobacterium tuberculosis* via bioinformatics approach. *bioRxiv*, 050047.
- Basharat, Z., Yasmin, A., 2016c. January. Bioinformatic analysis of human *Gaq* Q209 missense mutations associated with uveal melanoma. In 13th Int. Bhurban Conf. on Applied Sci. Tech. (IBCAST). IEEE. 106-111.
- Blom, N., Gammeltoft, S., Brunak, S. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 294, 1351-1362.
- Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S. 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics.* 4, 1633-1649.
- Bond, S.D., Benedict, J.L., Laird, B.B. 1999. The Nosé-Poincaré method for constant temperature molecular dynamics. *J. Comp. Phys.* 151, 114–134.
- Bordo, D. 2002. The rhodanese/Cdc25 phosphatase superfamily: Sequence-structure-function relations. *EMBO Rep.* 3, 741-746.
- Cesaro, P., Cattaneo, C., Bona, E., Berta, G., Cavaletto, M. 2015. The arsenic hyperaccumulating *Pteris vittata* expresses two arsenate reductases. *Sci Rep.* 5, 14525.
- Chao, D.Y., Chen, Y., Chen, J., Shi, S., Chen, Z., Wang, C., Danku, J.M., Zhao, F.J., Salt, D.E. 2014. Genome-wide association mapping identifies a new arsenate reductase enzyme critical for limiting arsenic accumulation in plants. *PLoS Biol.* 12, e1002009.
- Colovos, C., Yeates, T.O. 1993. Verification of protein structures: patterns of nonbonded atomic interactions. *Prot Sci.* 2, 1511-9.
- Dhankher, O., Rosen, B., McKinney, E., Meagher, R. 2006. Hyperaccumulation of arsenic in the shoots of *Arabidopsis* silenced for arsenate reductase (ACR2). *Proc. Nat. Acad. Sci.* 103, 5413-5418.

- Duan, G., Hu, Y., Liu, W., Kneer, R., Zhao, F., Zhu, Y. 2011. Evidence for a role of phytochelatins in regulating arsenic accumulation in rice grain. *Environ. Exp. Botany*. 71, 416-21.
- Duan, G., Zhou, Y., Tong, Y., Mukhopadhyay, R., Rosen, B., Zhu, Y. 2007. A CDC25 homologue from rice functions as an arsenate reductase. *New Phytologist*. 174, 311-321.
- Ellis, D.R., Gumaelius, L., Indriolo, E., Pickering, I.J., Banks, J.A., Salt, D.E. 2006. A novel arsenate reductase from the arsenic hyperaccumulating fern *Pteris vittata*. *Plant Physiol*. 141, 1544-54.
- Gupta, R., Brunak, S. 2001. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* 2002, 310-322.
- Heckman D. 2001. Molecular evidence for the early colonization of land by fungi and plants. *Science*. 293(5532):1129-1133.
- Hutchinson E, Thornton J. 1996. PROMOTIF-A program to identify and analyze structural motifs in proteins. *Prot Sci*. 5(2):212-220. doi: 10.1002/pro.5560050204.
- Indriolo E, Na G, Ellis D, Salt D, Banks J. 2010. A vacuolar arsenite transporter necessary for arsenic tolerance in the arsenic hyperaccumulating fern *Pteris vittata* is missing in flowering plants. *Plant Cell*. 22(6):2045-2057.
- Laskowski R, Hutchinson E, Michie A, Wallace A, Jones M, Thornton J. 1997. PDBsum: a web-based database of summaries and analyses of all PDB structures. *Trends in Biochem Sci*. 22(12):488-490.
- Laskowski R. 2001. PDBsum: summaries and analyses of PDB structures. *Nuc Acids Res*. 29(1):221-222.
- Laskowski R. 2004. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nuc Acids Res*. 33(Database issue):D266-D268.
- Laskowski R. 2009. PDBsum new things. *Nuc Acids Res*. 37(Database):D355-D359.
- Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. 2003. Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins: Struct Func Bioinformatics*. 50(3):437-50.
- Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire M, Geer L, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F. 2012. CDD: conserved domains and protein three-dimensional structure. *Nuc Acids Res*. 41(D1):D348-D352.
- Martin P, DeMel S, Shi J, Gladysheva T, Gatti D, Rosen B, Edwards BF. 2001. Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme. *Structure*. 9(11):1071-1081.

Meadows R. 2014. How Plants Control Arsenic Accumulation. PLoS Biol. 12(12):e1002008.

Meharg A, Macnair M. 1992. Suppression of the high affinity phosphate uptake system: A mechanism of arsenate tolerance in *Holcus lanatus* L. J Exp Bot. 43(4):519-524.

Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. 1992. Stereochemical quality of protein structure coordinates. Proteins: Struct Func Bioinformatics. 12(4):345-64.

Ramachandran GN, Ramakrishnan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. J Mol Biol. 7(1):95-9.

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 5(4):725-738.

Sánchez-Bermejo E, Castrillo G, Del Llano B, Navarro C, Zarco-Fernández S, Martínez-Herrera DJ, Leo-del Puerto Y, Muñoz R, Cámara C, Paz-Ares J, Alonso-Blanco C. 2014. Natural variation in arsenate tolerance identifies an arsenate reductase in *Arabidopsis thaliana*. Nature Comm. 5.

Tan C. 2011. Sequence, structure, and network evolution of protein phosphorylation. Science Signal. 4(182):mr6-mr6.

Ullrich-Eberius C, Sanz A, Novacky A. 1989. Evaluation of arsenate and vanadate associated changes of electrical Membrane Potential and Phosphate Transport in *Lemna gibba* G1. J Exp Bot. 40(1):119-128.

Wang J. 2002. Mechanisms of arsenic hyperaccumulation in *Pteris vittata*: Uptake kinetics, interactions with phosphate and arsenic speciation. Plant Physiol. 130(3):1552-1561.

Wilkins C, Salter L. 2003. Arsenic hyperaccumulation in ferns: a review. Bulletin Royal Soc Chem. 8-10.

Wu S, Zhang Y. 2007. LOMETS: A local meta-threading-server for protein structure prediction. Nuc Acids Res. 7;35(10):3375-3382.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2014. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 12(1):7-8.

Yin X, Wang L, Duan G, Sun G. 2011. Characterization of arsenate transformation and identification of arsenate reductase in a green alga *Chlamydomonas reinhardtii*. J Env Sci. 23(7):1186-1193.

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 9(1):40.

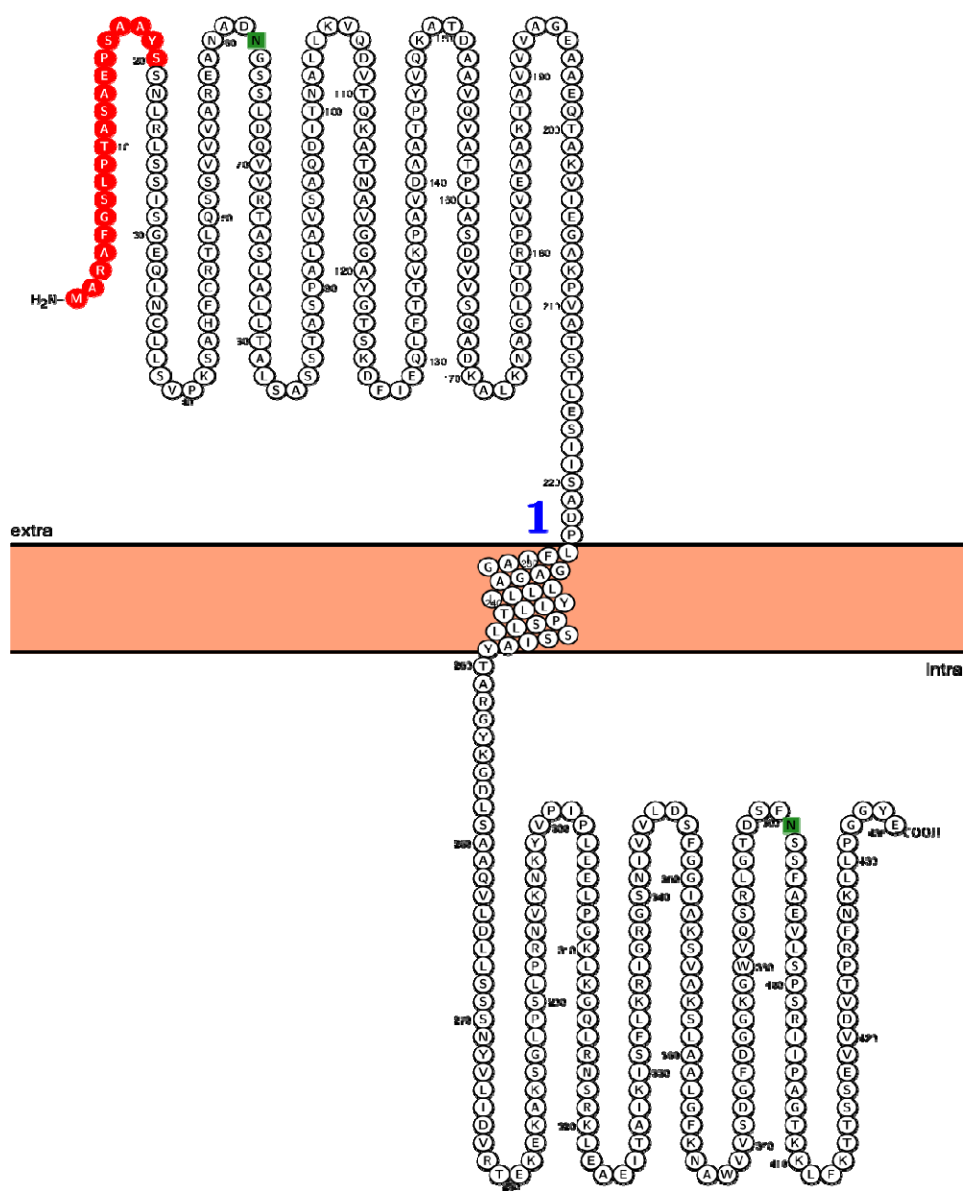


Fig. 1.

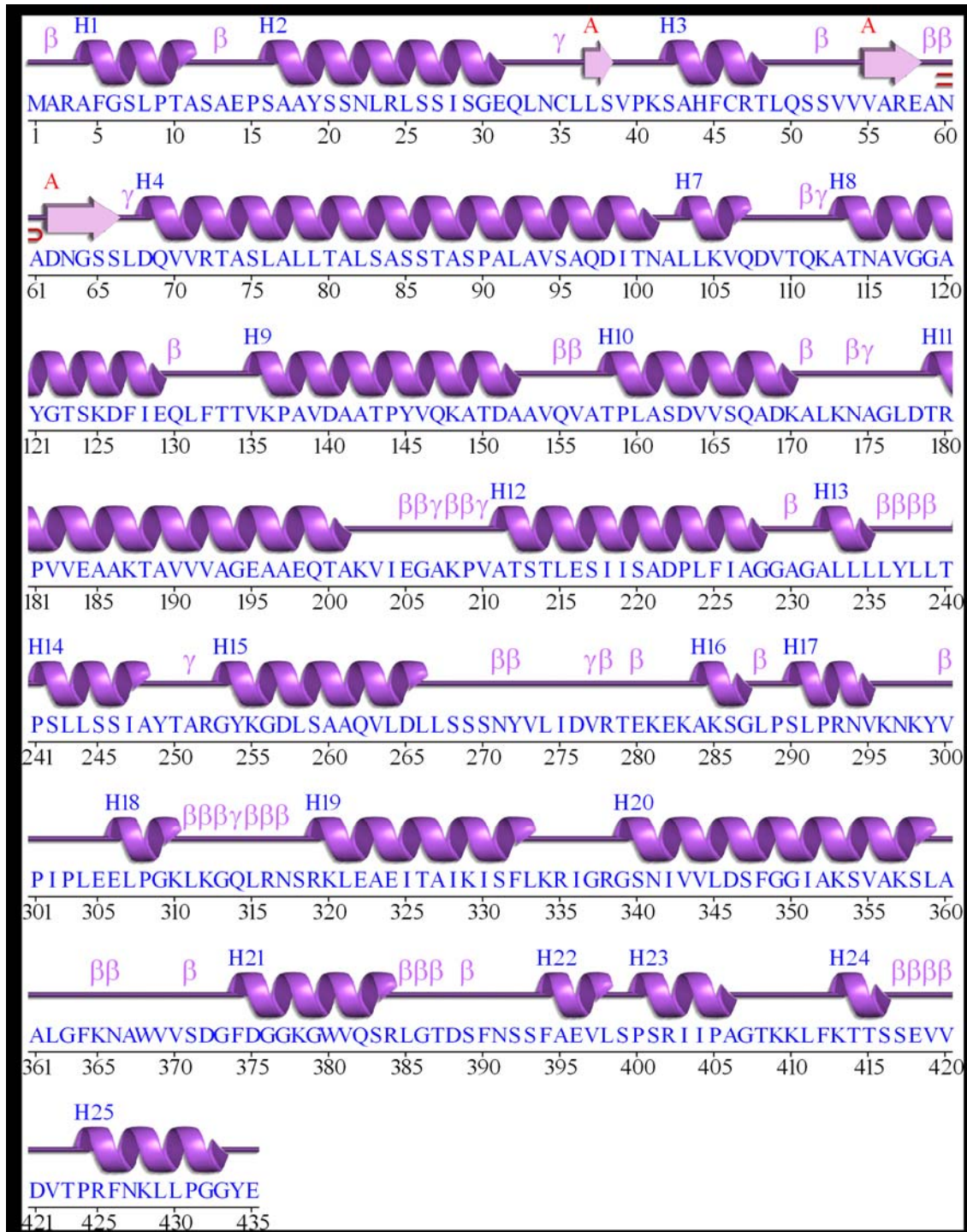
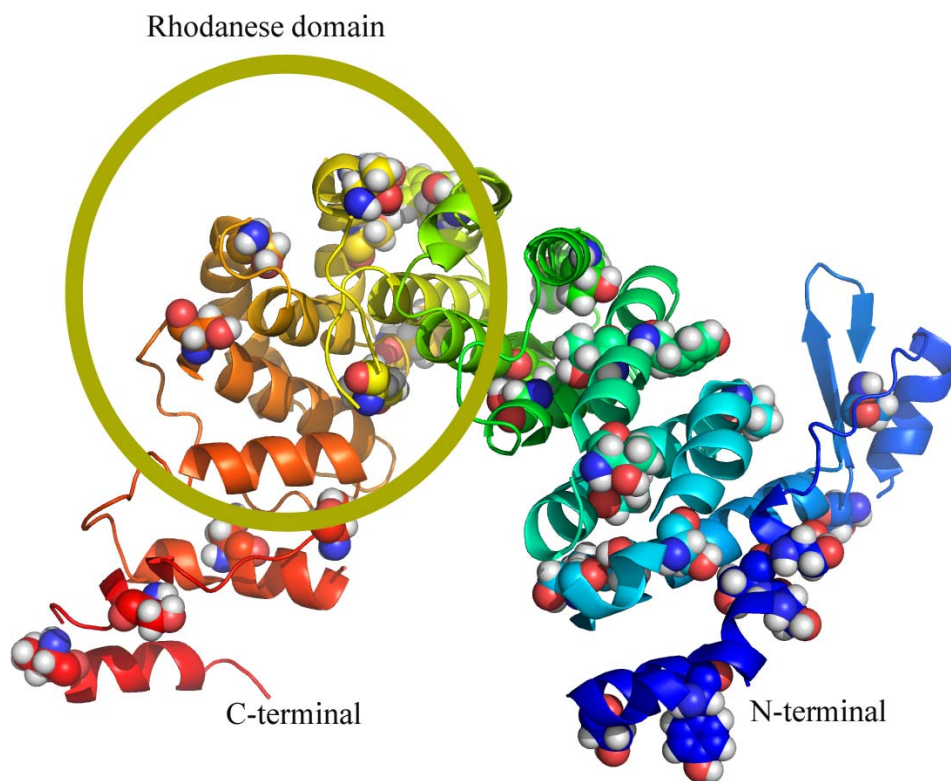
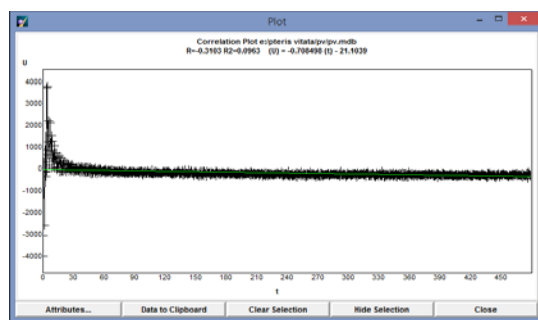


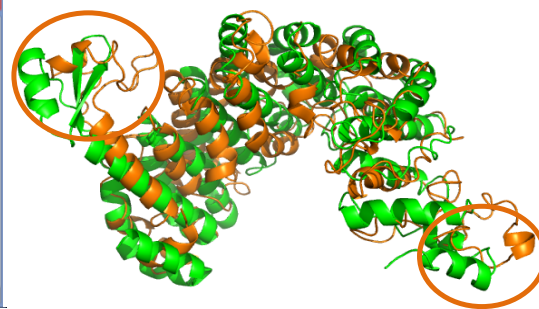
Fig. 2.



**Fig. 3.**



A



B

**Fig. 4.**

**Table 1.** Phosphorylation and *O*- $\beta$ -GlcNAcylation potential of arsenate reductases in *Pteris vittata*. – indicates a complete absence or lack of information for specified residues. + indicates presence of a particular trait in the enzyme at mentioned site.

<b>Residue</b>	<b>No.</b>	<b>Kinases responsible for Phosphorylation</b>	<b><i>O</i>-<math>\beta</math>-GlcNAcylation</b>
Ser	12	-	+
	26	-	-
	27	PKA	-
	29	-	-
	52	-	-
	65	-	-
	85	-	+
	86	-	+
Thr	89	p38MAPK, cdk5	-
	95	-	+
	166	-	+
	213	CKII	+
	217	CKI	-
	245	-	-
	286	-	-
	290	PKC	-
	318	-	-



	358	-	-
	399	-	-
	110	-	-
	123	-	-
	134	-	-
	143	-	-
	188	-	-
	200	-	-
	279	-	-
	408	-	-
	415	-	-
	423	-	-
Tyr	19	-	-
	145	-	-
	249	INSR	-
	272	-	-
	299	INSR	-