Heterologous expression of cyanobacterial *PCS* confers augmented arsenic and cadmium stress tolerance and higher artemisinin in *Artemisia annua* hairy roots

Neha Pandey^{1,2} · Krishna Kumar Rai² · Sanjay Kumar Rai³ · Shashi Pandey-Rai²

Received: 1 March 2021 / Revised: 1 May 2021 / Accepted: 22 May 2021 / Published online: 7 June 2021
© Korean Society for Plant Biotechnology 2021

Abstract

The present study provides the first report of heterologous expression of phytochelatin synthase from *Anabaena* PCC 7120 (*anaPCS*) into the hairy roots of *Artemisia annua*. Transformed hairy roots of *A. annua* expressing *anaPCS* gene showed better tolerance to heavy metals, viz., arsenic (As) and cadmium (Cd) owing to 143 and 191% more As- and Cd-accumulation, respectively, as compared to normal roots with a bioconcentration factor (BCF) of 9.7 and 21.1 for As and Cd, respectively. Under As and Cd stresses, transformed hairy roots possessed significantly higher amounts of phytochelatin and thiols probably due to the presence of both *AaPCS* (*Artemisia annua PCS*) and *anaPCS*. In addition, artemisinin synthesis was also induced in transformed hairy roots under heavy metals stresses. In-silico analysis revealed the presence of conserved motifs in both *AaPCS* and *anaPCS* sequences as well as structural modelling of PCS functional domain was conducted. Interaction of *AaPCS* and *anaPCS* proteins with CdCl₂ and sodium arsenate gene ontology analysis gave insights to *anaPCS* functioning in transformed hairy roots of *A. annua*. The study provides transformed hairy roots of *A. annua* as an efficient tool for effective phytoremediation with added advantages of artemisinin extraction from hairy roots used for phytoremediation.

Keywords Phytochelatin synthase · *Artemisia annua* · Hairy root · *Anabaena* PCC 7120 · Arsenic · Cadmium

Introduction

Arsenic (As) and cadmium (Cd) are among the major soil pollutants due to their high toxicity and long persistence. Both As and Cd are non-essential elements and adversely affect plant growth and development even at very low concentration in the soil.

The entry of As in plants is concentration gradient dependent. Out of four oxidation states of As [i.e., As(0), As(-III), As(III) and As(V)], only arsenate [As(V)] and arsenite [As(III)] enter into the plant roots via different transporter proteins. Among arsenate and arsenite forms of

As, arsenite is more toxic (Khalid et al. 2017). As(V) is a chemical analogue of phosphorus and hence uses Pi transporters, consisting of Pi transporter proteins (PHT), to enter into the plant (LeBlanc et al. 2013). Inside the cell, As(V) is converted into As(III) by the action of arsenate reductase (Bleeker et al. 2006). On the other hand, As(III) enters into the plant cells through Nodulin 26-like intrinsic proteins (NIPs) belonging to the class of plant aquaporins (Mukhopadhyay et al. 2014). Cd uptake takes place through several metal transporters such as ZIP (ZRT- and IRT-like proteins) family, YSL (Yellow Stripe-Like) transporters and NRAMP (natural resistance-associated macrophage protein) family of transporters. Cd is also known to enter plant symplast via passive transport or using Ca²⁺ channels present on guard cell plasma membrane (Ismael et al. 2019).

Once inside the plant, these heavy metals are immobilized, chelated and compartmentalized through metal-binding ligands which include organic acids, polypeptides or even small peptides. One such metal-binding polypeptide is phytochelatin (PC) which mitigate the harmful effects of heavy metal by chelation (Zhang et al. 2013). PCs are enzymatically synthesized from glutathione by

✉ Shashi Pandey-Rai
shashi.bhubotany@gmail.com

¹ Department of Botany, CMP PG College (A Constituent PG College of University of Allahabad), Prayagraj, India

² Department of Botany, Institute of Science, Banaras Hindu University, Varanasi, India

³ Department of Horticulture, Dr. Rajendra Prasad Agricultural University, Pusa, Samastipur, Bihar, India

phytochelatin synthase (PCS). Heavy metals significantly elevate the biosynthesis of PCs. Metal ions including both anions (for example arsenate) and cations (for example cadmium) tend to increase the level of PCs (Cobbett 2000). Cloning of *PCS* gene has been achieved in many plant species such as *Arabidopsis*, rice, and wheat, and overexpression of PCS has been linked with greater heavy metal tolerance (Vatamaniuk et al. 1999).

Plant's potential to uptake heavy metals can be easily evaluated by calculating bioconcentration factor. Plants with a value of bioconcentration factor > 1 for any heavy metal has been suggested to be categorized as bio-accumulators and could be beneficial for phytoextraction (Usman et al 2013).

Hairy roots are formed after infection with a Gram-negative bacterium *Agrobacterium rhizogenes*, which transfers its T-DNA region (transfer DNA) into the host cell, integrates with the plant genome and expresses. Cloning and expression of recombinant proteins can be achieved by using many systems such as bacterial, yeast expression system or mammalian cells. However, each of them poses some or other limitations such as risks of viral infections in bacterial systems, toxic molecules in the mammalian system or high production cost (Gutierrez-Valdes et al. 2020). For a couple of decade's plants, systems are being used for several advantages including low cost. Hairy root cultures have emerged as a potential expression system with low maintenance cost and high efficiency. As hairy roots are derived from the wounded region of plants, they possess the genetic makeup of the plant part from which they are derived and hence do not behave like a true root. This property makes hairy root cultures as an excellent expression system. Hairy roots have found their potential application for phytoremediation of heavy metals (Ago-stini et al. 2013).

Artemisia annua, popularly known for its antimalarial drug artemisinin, has been found to act as a hyperaccumulator of arsenic. In our previous studies (Kumari et al. 2017; Rai et al. 2014), *A. annua* was found to survive well even at high As-concentration and have hyperaccumulating capacity. Many heavy metal stresses including both As (Kumari et al. 2017) and Cd (Li et al. 2012) have been reported to induce biosynthesis of artemisinin. Various reports have been published for hairy root induction in *A. annua* for enhanced secondary metabolite production (Patra and Srivastava 2014). *A. annua* hairy root potential for phytoremediation of heavy metals has not been investigated. In the present investigation, we have cloned a cyanobacterial PCS gene (from *Anabaena* PCC 7120) into hairy roots of *A. annua*. We hypothesize that transformed hairy roots overexpressing PCS gene can be used as an excellent tool for phytoremediation of arsenic and cadmium.

Materials and methods

Plant material and hairy root induction

Uniform seeds of *Artemisia annua* were surface sterilized in sodium hypochlorite (10% w/v) solution for 15–20 min and washed three-to-four times in sterilized water. Seeds were then treated with ethanol (70% v/v) for 30 s and were then allowed to germinate on Murashige and Skoog (MS) medium containing sucrose (3% w/v) at pH 5.8. Germination initiated under the controlled condition of light and temperature (16/8 h light/dark regime and 25 ± 2 °C). The in-vitro raised plantlets were then used for induction of hairy roots by infecting the leaves/stem with *Agrobacterium rhizogenes* ATCC 15834 maintained in liquid YMB medium with O.D. at 600 nm ranged between 0.9 and 1.0 (Hooykass et al. 1977). The infected plantlets were then cultured on full-strength MS medium for 2 days and then transferred to half-strength MS medium containing cefotaxime (250 mg l^{-1}) under the same growth conditions. The plantlets were further sub-cultured by transferring them to new antibiotic supplemented medium for 2 weeks, and finally, the bacterial-free hairy roots were cultured on liquid MS medium. Hairy root cultures were incubated in dark inside the rotary shaker with constant rotation (100 rpm) and were sub-cultured twice a month into the fresh medium.

For experiments, three tissue samples were taken: normal roots (which were derived from plantlets), hairy roots (derived from leaves by *A. rhizogenes* infection) and transformed hairy roots (also derived from leaves by *A. rhizogenes* infection and expressing *anaPCS* gene).

Cyanobacterial strain and plasmids

Anabaena sp. PCC 7120 was photoautotrophically grown in BG-11 medium (without any source of combined nitrogen) prepared in Tris buffer with a photoperiod of 14:10 h under light (day light fluorescent tubes) intensity of $72 \mu\text{mol m}^{-2} \text{ s}^{-1}$ PAR (photosynthetically active radiation) at 24 ± 2 °C (Rippka et al. 1979). The *E. coli* strains DH5 α was used for cloning and BL21 (Novagen) for overexpression. The cells possessing recombinant plasmid were transferred on fresh Luria–Bertani (LB) medium supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin (Sambrook and Russell 2001). For cloning, pGEX-5X-2 was used as a vector (Chaurasia et al. 2008).

Cloning and confirmation of *anaPCS* gene and transformation in *A. annua* hairy roots

Total genomic DNA from *Anabaena* PCC7120 was isolated as per the standard CTAB method (Srivastava et al. 2007) and the region encoding phytochelatin synthase (*alr0975*, denoted as *anaPCS* in the manuscript) was amplified using gene-specific primers having restriction sites for *EcoRI* and *NotI* (Chaurasia et al. 2008). The desired gene amplicon was then purified by the QIAquick gel extraction kit (Qiagen), and cloned in pGEX-5X-2 vector (Novagen) containing *EcoRI* and *NotI* restriction sites, and the resulting construct pGEX-5X-2-*anaPCS* was used to confirm *anaPCS* nucleotide sequence verification and expression analysis according to Chaurasia et al. (2008). For heterologous expression analysis, *anaPCS* gene was amplified using PF 5′-GCTCTAGAC GATAGTTATGAACTCTTTA-3′, and PR 5′-GCAGAG CTCCTAATCTTGTGTTTACTTAC-3′, containing *XbaI* and *SacI* restriction sites and inserted into a pBI121 (Clontech, USA) under the control of promoter CMV35S and as binary vector introduced into the *A. rhizogenes* ATCC 15834 by electroporation for further transformation of *A. annua* hairy roots and expression studies. The confirmation of the resulting recombinant vector pBI121-*anaPCS* was done by performing double digestion of the plasmid with *XbaI* and *SacI* enzymes and its sequencing.

Plant stress treatments and measurement of arsenic and cadmium content

To assess the role of *anaPCS* in conferring heavy metal tolerance, *A. annua* hairy roots transformed with pBI121-*anaPCS*, non-transformed hairy roots (empty vector), and normal roots were subjected to arsenic (100 μ M) and cadmium stress (50 μ M) using sodium arsenate ($\text{Na}_2\text{HAsO}_4 \cdot 7\text{H}_2\text{O}$) and cadmium chloride (CdCl_2), respectively, for 5 days. The treatments were divided in to three sets, set 1 with no treatment (control), set 2 with 100 μ M of arsenate ($\text{Na}_2\text{HAsO}_4 \cdot 7\text{H}_2\text{O}$) and set 3 with CdCl_2 . Each treatment set contains three groups of samples, untransformed hairy roots, transformed hairy roots and normal roots of the plant.

All the experiments were performed in triplicates. The accumulation of arsenic content in treated/non-treated hairy roots and normal roots was determined by the SEM-EDS method and was estimated by atomic absorption spectrometry (Perkin-Elmer A Analyst 600) as described by Kumari et al (2018).

Estimation of phytochelatin synthase, reduced glutathione and thiol

Phytochelatin content was determined as described earlier (Rai et al. 2011). For glutathione estimation, 0.5 g tissue was homogenized in an extraction buffer containing sulphosalicylic acid (5% w/v). The absorbance of the reaction mixture containing potassium phosphate buffer (100 mM, pH 7.0), DTNB (5′5′-dithiobis-2-nitrobenzoic acid) and tissue extract was read at 412 nm (Anderson 1985). For total thiol estimation, 0.5 g tissue was homogenized in extraction buffer containing ascorbate buffer (20 mM) and EDTA (20 mM), and after centrifugation at 12,000 $\times$ g for 15 min (4 $^{\circ}$ C), 0.5 ml supernatant was mixed with reaction mixture that contained Tris-HCl buffer (200 mM) and DTNB (10 mM). Reaction mixture developed yellow color after incubation for 20 min at 30 $^{\circ}$ C and absorbance was read at 412 nm (Nagalakshmi and Prasad 2001).

Measurement of antioxidative enzymes

Fresh tissues (500 mg) were homogenized with 5 ml Tris buffer (50 mM, pH = 7.0) that contains EDTA (2 mM), PMSF (1 mM), and PVP (1%). After centrifugation of homogenate at 12,000 $\times$ g for 15 min, immediately the supernatant was used for the assay.

Antioxidative enzymes catalase, ascorbate peroxidase, glutathione reductase and peroxidases were measured as described earlier (Kumari et al. 2017; Pandey and Pandey-Rai 2014). All the spectrophotometric analysis was carried out using Hitachi U-2910 spectrophotometer.

Expression analysis of *rolB* and *PCS* gene and artemisinin biosynthetic pathway genes

Total RNA was isolated using Trizol reagent (Invitrogen) as per the manufacturer's recommendation. One microgram of total RNA was used as template for cDNA synthesis with cDNA synthesis kit (Bio-Rad Laboratories, USA) following the manufacturer's instructions.

RT-PCR analysis using primers listed in supplementary table 1 was performed by following the method described earlier (Pandey and Pandey-Rai 2014). All reactions were run in triplicate along with a negative control. Primer accuracy was tested using gDNA as positive control. PCR products were run on agarose gel and band intensities were measured using Quantity One software (Bio-Rad) installed with Gel Doc 2000 system (Bio-Rad).

PCR was performed in reaction mixture (50 μ l final volume) containing 50 ng cDNA, 2.5 μ l 10 $\times$ PCR buffer with 15 mM MgCl_2 , 200 μ M dNTPs, 10 pmol of each primer and 0.2 U *Taq* DNA polymerase (Bangalore Genei, India) in an iCycler (Bio-Rad, USA). The thermal amplification protocol

was as follows: 1 cycle at 95 °C for 5 min, and 30 cycles at 95 °C for 30 s, 59 °C (value varied for different primer sets) for 45 s and 72 °C for 45 s.

Artemisinin estimation

Artemisinin was extracted and prepared according to Zhao and Zeng (1985) and Pu et al. (2009) with minor modifications. Approximately 100 mg dry tissue was extracted in 20 ml petroleum ether by ultrasonication (35 min). The extract was evaporated and dissolved in 1 ml ethanol. After 1 min centrifugation at 10,000×g, 0.1 ml of the supernatant was treated with 0.4 ml of 0.2% (w/v) NaOH at 45 °C for 35 min and cooled to room temperature. The solution was acidified with 0.5 ml of 0.05 M acetic acid and filtered through a Millipore filter (0.45 µm) before HPLC analysis. HPLC analysis conditions: Waters 600E HPLC system equipped with 150×9 4.6 mm Hypersil BDS C8 column; injected samples (10 µl) were eluted with methanol/0.01 M Na₂HPO₄–NaH₂PO₄ buffer (pH 7) (2:3, v/v) at 1 ml min^{−1} and monitored at 260 nm. The concentration of artemisinin was measured using the standard curve.

In-silico analysis for comparative phylogeny, structural and functional analysis of *anaPCS* and *AaPCS* gene.

The sequence of phytochelatin synthase (*PCS*) gene in *A. annua* (*AaPCS*) and *Anabaena* PCC7120 (*anaPCS*) was obtained using BLAST-P (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) by taking homology with the *Arabidopsis* phytochelatin synthase (*AtPCS*) gene (Locus: AAD41794.1) identified from TAIR (<https://www.arabidopsis.org/>). The sequences with maximum query coverage and percentage identity (> 90%) were selected for the phylogenetic tree construction and sequential characterization. The sequences were aligned using Bio-Edit tool (Hall 1999) and phylogenetic tree was constructed using MEGA 7 suite (Tamura et al. 2013). The result of BLAST-P search was further subjected to PDB database to evaluate the similarity indices of the sequences and confirmed by circos visualization tool (Krzywinski et al. 2009). All the identified probable PCS protein sequences were subjected to multiple sequence alignment analysis using CLC bio workbench and with the help of Clustal W (<http://www.ebi.ac.uk/tools/msa/clustalw2/>) for their sequential classification (Krzywinski et al. 2009). The functional domain regions occupied by PCS proteins were searched using InteProScan <http://www.ebi.ac.uk/Tools/pfa/iprscan> (Jones et al. 2014), NCBI CDD server <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml> (Marchler-Bauer et al. 2011) and ExPASy-Prosite scan <http://prosite.expasy.org/scanprosite> (de Castro et al. 2006). The occurrence of potential motifs

in PCS protein in all the members was identified using MEME Suite 4.1.1.2 (Multiple EM for Motif Elicitation) (Bailey et al. 2006) by keeping the selection parameter to any number of repeats with motif width of 10 and 30 residues and by keeping a maximum number of motifs to 40.

Structural modelling and model validation

The 3D homology modelling of identified protein sequences for PCS protein in *A. annua* and *Anabaena* PCC 7120 (*AaPCS* and *anaPCS*) was done using template proteins available in protein database which were searched using BLAST-P programme of Protein Data Bank (www.rcsb.org/pdb/) (Berman et al. 2000). The result of BLAST-P analysis showed most significant and accurate templates for the modelling of functional domain structures of *AaPCS* and *anaPCS* proteins. Three most closely related proteins were selected for the homology modelling of *AaPCS* and *anaPCS* proteins using MODELLER module of Discovery Studio 3.0 (accelrys.com; Shahi et al. 2013). Five models were generated by the MODELLER using Crystal structure of *anaPCS* (template 1; PDB ID: 2BTW), Crystal structure of a putative C39-like peptidase from *Bacillus anthracis* (template 2; PDB ID: 3ERV) and Crystal structure of the N-terminal peptidase C39-like domain of the toxin secretion ATP-binding protein from *Vibrio parahaemolyticus* (template 3; PDB ID: 3B79). The model with lowest Discrete Optimized Protein Energy (DOPE) score was further refined for Cα traces by ModRefiner using two-step atomic level energy minimization module (Xu and Zhang 2011). The modelled proteins were then allowed to superimpose with each template using Superpose version 1.0 (Maiti et al. 2004) to access the topological details of the modelled proteins.

The predicted models were further evaluated qualitatively in terms of their geometric analysis, stereochemical orientation and backbone conformation patterns of nonbonded atomic interactions. For the backbone conformations, the predicted models were subjected to Ramachandran plot analysis for evaluating the backbone dihedral phi (φ) and psi (Ψ) angles using PROCHECK module of PDBSum server <http://www.ebi.ac.uk/pdbsum/> (Laskowski et al. 2005) which was further attested by RAMPAGE server (Lovell et al. 2003). The models were also evaluated by a single model method using ProSA (Wiederstein and Sippl 2007), Qmean (Benkert et al. 2009), RESPROX (Resolution by Proxy), ERRAT (Colovos and Yeates 1993). For the quantitative evaluation, the models were subjected to VADAR (Volume, Area, Dihedral Angle Reporter) analysis (Willard et al. 2003). The protein models generated for the functional domain were further submitted to an online repository of PMDB (Castrignano et al. 2006) to obtain the accession identities.

Ligand modelling and protein docking

The molecular docking studies between ligands (Cadmium chloride; CdCl_2 and sodium arsenate; $\text{Na}_2\text{HAsO}_4 \cdot 7\text{H}_2\text{O}$) and the AaPCS and anaPCS protein sequences were carried out using Hex 8.0 molecular docking server (Macindoe et al. 2010) to reveal the probable binding modes for ligand compounds with in the active site of PCS protein. The ligands, viz., CdCl_2 and sodium arsenate, were drawn in 2D structure using Chemdraw software (ChemDraw, CambridgeSoft Corporation, 875 Massachusetts Avenue, Cambridge, MA 02139, USA). The 3D structures of both AaPCS and anaPCS proteins were generated using MODELLER module of Discovery studio 3.0 and were further explored for molecular interaction studies with ligands using Hex 8.0 molecular docking server. The protein and ligand interaction studies were performed following correlation criterion, i.e., DARS + Shape + Electro, FFT Mode-3D fast lite and grid range of 0.6 with Receptor:Ligand:Twist:Distance range of 180:180:360:40. Several conformations of docked complexes were obtained; however, the docked complex having lowest binding energy values was analyzed by DS Studio 3.0 for identifying the key residues involved in the interaction with CdCl_2 and sodium arsenate.

Protein–protein interaction, structural and functional annotation

The functional protein–protein interaction complex involved in signaling and metabolic pathways at both cellular, biological and molecular levels in AaPCS and AnaPCS proteins was searched through STRING server (Search Tool for the Retrieval of Interacting Genes/Proteins) database version 10.0 <http://string-db.org/> (Szklarczyk et al. 2015). The functional interactive sources were identified based on seven core shells such as co-expression, co-occurrence, neighbor, gene fusion, text-mining and experiments by setting confident score (0.40) to default. The identified proteins were then sequentially analyzed for their structural and functional organization based on gene ontological (GO) terms using CATH server <http://www.cathdb.info/> (Sillitoe et al. 2015) and FunFHMME <http://www.cathdb.info/search/by-funfhmmr> (Das et al. 2015). The GO enrichment data obtained were further visualized using hypergeometric distribution test analysis of REVIGO web server <http://revigo.irb.hr/> (Supek et al. 2011). The identified GO terms were further fetched to CELLO2GO web server <http://cello.life.nctu.edu.tw/cello2go/> (Yu et al. 2014) for their probable subcellular localization.

Active site prediction

The prediction of transition metal-binding sites and identification of potential residues involved in the making of the active sites with AaPCS and anaPCS was done using metapocket server <http://metapocket.embl.org> (Huang 2009).

Statistical analysis

All the experiments were performed in three biological replications and subjected to analysis of variance (ANOVA). The mean differences were compared by performing Duncan's multiple range test (DMRT) using SPSS software (SPSS Inc., Version 20.0) and values at $P \leq 0.05$ were considered significant.

Results

Induction and establishment of normal roots, hairy roots and transformed hairy root

Plantlets of *A. annua* were successfully established under *in-vitro* conditions (supplementary Fig. 1) and produced visible roots after 15 days of inoculation onto the MS medium. Infection by *A. rhizogenes* carrying binary vector (pBI121-*anaPCS*) and co-culturing of infected *A. annua* leaves has resulted in transformed hairy root (containing *anaPCS* gene) induction after 3 week of inoculation at a relative transformation frequency of 71%. Normal hairy roots not carrying *anaPCS* gene were also obtained in 15 days by infecting *A. annua* leaves with *A. rhizogenes* and co-culturing leaf discs. The presence of *rolB* gene in normal hairy roots and transformed hairy roots (with *anaPCS*) confirmed their origin after *A. rhizogenes* infection. The *rolB* gene was absent in normal roots (Fig. 1a, b).

Transformed hairy roots expressing *anaPCS* show enhanced tolerance (accumulation) to heavy metal stresses

In *A. annua*, hairy roots accumulated very high amount of As and Cd as compared to normal roots under As- and Cd-stress, respectively (Fig. 2). However, hairy roots expressing *anaPCS* gene accumulated even higher amount of both heavy metals as compared to non-transformed hairy roots. Under As-stress, non-transformed hairy roots and transformed hairy roots accumulated 91.3% and 143.4% higher amount of As, respectively, as compared to normal roots of *A. annua*. Among hairy roots, 72.7% greater As accumulation was observed in transformed hairy roots expressing *anaPCS* gene. Similar trend was observed for Cd-accumulation, where non-transformed hairy roots and transformed hairy roots accumulated 36% and

Fig. 1 **a** Normalized gene expression analysis. **b** RT-PCR for expression analysis of *rolB* gene in hairy root, normal root and transformed hairy roots expressing *anaPCS* gene

191% more Cd as compared to normal roots. Further among hairy roots, transformed hairy roots showed 55.4% greater Cd-accumulation as compared to non-transformed hairy roots.

The bioconcentration factor (BCF) for As was found to be 3.71, 5.71 and 9.7 for normal roots, non-transformed hairy roots and transformed hairy roots, respectively (Fig. 3). Similarly, for Cd, bioconcentration factor was 9.4, 12.7 and 21.1 for normal roots, non-transformed hairy roots and transformed hairy roots, respectively. The results showed that bioconcentration of Cd is higher than As and hairy roots expressing *anaPCS* gene have greater BCF for both As and Cd as compared to non-transformed hairy roots.

Normal roots, hairy roots (with only *rol* gene) and transformed hairy roots (with *rol* and *anaPCS* gene) have morphological differences in terms of average weight and average weight increase and are listed in supplementary table 6. As visible in the Supplementary Fig. 10, branching pattern is also different among normal, hairy and transformed hairy roots. Transformed hairy roots (with *anaPCS* gene) had more branching as compared to normal and hairy roots (without *anaPCS* gene).

Phytochelatins and phytochelatin synthase expression

Heavy metal stresses enhanced the phytochelatin content in all the tissue samples; however, the concentration was

significantly more in hairy roots rather than normal roots (Fig. 4a). As-stress boosted phytochelatin content up to 28.3%, 44.7% and 14.7% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to their respective control. Cd-stress enhanced phytochelatin level up to 25%, 35.7% and 9.52% in normal, hairy roots and transformed hairy roots, respectively, in comparison to their respective control. Although the stress-induced percent increase of phytochelatin was comparatively less prominent in transformed hairy roots, the amount of phytochelatin was more. Under conditions of As- and Cd-stress, transformed hairy roots expressing *anaPCS* gene was shown to contain 35.3% and 37.7% more phytochelatin as compared to non-transformed hairy roots.

The expression of phytochelatin synthase was measured through RT-PCR. Expression of *anaPCS* was observed only in transformed hairy roots, and because normal roots and hairy roots were not carrying *anaPCS* gene copy, no expression was observed in them (Fig. 4b). Expression of *anaPCS* was reported to be induced up to 2.1- and 2.3-fold under As- and Cd-treatment, respectively, as compared to control. Expression of *AaPCS* (*Artemisia annua PCS*) was observed in all the tissue samples. *AaPCS* expression was significantly induced under As- and Cd-stress in hairy roots and transformed hairy roots; however, no significant elevation in *AaPCS* was observed in normal roots under heavy metal stress conditions.

Fig. 2 Heavy metal accumulation in normal roots, hairy roots and transformed hairy roots (expressing *anaPCS* gene). **a** Arsenic content and **b** cadmium content. Vertical bars denote mean $\pm$ SD and different alphabets show statistical significance among treatments (control/As-treatment/Cd-treatment) at $p < 0.05$

Fig. 3 Bioconcentration factor (BCF) of arsenic and cadmium-stressed normal roots, hairy roots and transformed hairy roots (expressing *anaPCS* gene). Vertical bars denote mean $\pm$ SD and different alphabets show statistical significance among treatments (control/As-treatment/Cd-treatment) at $p < 0.05$

Thiol and GSH level in normal roots, hairy roots and transformed hairy roots under As- and Cd-stress

In control samples (without heavy metal stress) of normal roots, hairy roots and transformed hairy roots, thiol content was found to 11.2, 12.2 and 21.1 $\mu\text{M g}^{-1}$ FW (Fig. 5a). Under condition of As-stress, thiol content was elevated up to 43.75%, 59.83% and 87.2% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to their respective controls. However, under Cd-stress, the percentage increase was 39.2%, 38.5% and 52.6% for normal roots, hairy roots and transformed hairy roots, respectively, as compared to their respective control. Among hairy roots, thiol content was increased up to 102.5% under As-stress and 90.5% under Cd-stress in transformed hairy roots as compared to non-transformed hairy roots.

In contrast to thiol content, GSH level (Fig. 5b) was reduced under both heavy metal stresses in all tissue samples. Under As-stress 6.8%, 34.4% and 50.6% reduction in GSH level was observed in normal roots, hairy roots and transformed hairy roots, respectively, as compared to their respective control. Compared with the control, Cd-stress dropped GSH level up to 9.8%, 32.2% and 45.7% in normal roots, hairy roots and transformed hairy roots, respectively. As compared to non-transformed hairy roots, transformed hairy roots expressing *anaPCS* gene contained 30.5% and 26.2% lower levels of GSH under As- and Cd-stress, respectively.

Stress-induced antioxidative enzymes in normal roots, hairy roots and transformed hairy roots

Variable activities of four antioxidative enzymes were measured under heavy metal stress in all tissue samples (Fig. 6a–d). Under As-stress, catalase activity was reported to be induced up to 20%, 12.5% and 20% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to control. Similarly, Cd-stress increased catalase activity up to 26.6%, 12.5% and 13.3% in normal roots, hairy roots and transformed hairy roots as compared to control. Although catalase activity under stresses was higher than control, their activity was similar among all three tissues. Transformed hairy roots did not show any significant change in catalase activity under heavy metal stresses in *A. annua*.

Under As-stress, APX activity was enhanced up to 21.9%, 67.6% and 34.7% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to control. Under Cd-stress, APX activity was reported to be 17.4%, 72.3% and 5.37% in normal roots, hairy roots and transformed hairy roots, respectively, compared with the control. In comparison to non-transformed hairy roots, transformed hairy roots expressing *anaPCS* gene showed 31.9% greater APX activity under As-stress; however, Cd-stressed

Fig. 4 Phytochelatin content and phytochelatin synthase expression in normal roots (NR), hairy roots (HR) and transformed hairy roots (THR) (expressing *anaPCS* gene) under arsenic and cadmium stress.

a Phytochelatin content; **b** RT-PCR for expression analysis of *Anabaena* PCC 7120 PCS gene (*anaPCS*); **c** RT-PCR for expression analysis of *Artemisia annua* PCS (*AaPCS*)

Fig. 5 **a** Thiol content and **b** GSH content in normal roots, hairy roots and transformed hairy roots (expressing *anaPCS* gene) under arsenic and cadmium stress. Vertical bars denote mean $\pm$ SD and different alphabets show statistical significance among treatments (control/As-treatment/Cd-treatment) at $p < 0.05$

transformed hairy root did not show any significant increment in APX activity as compared to non-transformed hairy roots.

Peroxidase activity was also enhanced up to 150%, 139.4% and 101.3% in As-stressed normal roots, hairy roots and transformed hairy roots, respectively, as compared to their respective control. Increment of 130.7%, 153.5% and 90.5% in peroxidase activity was reported in Cd-stressed normal roots, hairy roots and transformed hairy roots, respectively, as compared to control.

Stress-induced increment in GR activity was more pronounced in transformed hairy roots than non-transformed hairy roots. As-stress induced GR activity up to 5.5%, 18.6% and 151.7% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to control. Similarly, Cd-stress enhanced GR activity up to 13.4%, 15.2% and 93.1% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to control. Transformed hairy roots expressing *anaPCS* gene showed 108.5% and 64.7% greater GR activity under As- and Cd-stress, respectively.

Artemisinin content and expression of artemisinin biosynthetic pathway genes

Artemisinin level was found to be more in hairy roots than in normal roots in both normal and stressed conditions (Fig. 7a). Under As-stress, artemisinin level was enhanced up to 50%, 33.3% and 16.6% in normal roots, hairy roots and transformed hairy roots, respectively. On the other hand, Cd-stress boosted artemisinin level up to 45.41%, 25% and 22.2% in normal roots, hairy roots and transformed hairy roots, respectively, as compared to control. Figure 7b shows the expression pattern of four key artemisinin biosynthetic pathway genes, viz., *ADS*, *CYP71AV1*, *DBR2* and *ALDH1*, in all tissue samples. All the four biosynthetic pathway genes were found to

Fig. 6 Antioxidative enzyme activities of **a** Catalase, **b** ascorbate peroxidase, **c** glutathione reductase, and **d** peroxidases, in normal roots, hairy roots and transformed hairy roots (expressing *anaPCS* gene)

under arsenic and cadmium stress. Vertical bars denote mean $\pm$ SD and different alphabets show statistical significance among treatments (control/As-treatment/Cd-treatment) at $p < 0.05$

Fig. 7 **a** Artemisinin content and **b** expression analysis of artemisinin biosynthetic enzymes (*ADS*, *CYP71AV1*, *DBR2* and *ALDH1*) through RT-PCR, in normal roots, hairy roots and transformed hairy roots (expressing *anaPCS* gene) under arsenic and cadmium stress. In **a**

vertical bars denote mean $\pm$ SD and different alphabets show statistical significance among treatments (control/As-treatment/Cd-treatment) at $p < 0.05$

be significantly overexpressed in transformed hairy roots as compared to normal roots and hairy roots (non-transformed). Under As-stress, *ADS* expression was 7- and 6.4-fold more than normal roots and hairy roots, respectively. Similarly, in transformed hairy roots, Cd-stress enhanced

ADS expression up to 2.8 and 3.2-folds as compared to normal roots and hairy roots, respectively. Similar over-expression patterns were observed for *CYP71AV1*, *DBR2* and *ALDH1*, as well.

Database mining and phylogenomic analysis of *AaPCS* and *anaPCS*

To find *A. annua* and *Anabaena* PCC 7120 *PCS* homolog, BLAST-P search was performed using the *Arabidopsis* *PCS* gene (AAD41794.1) which led to the identification of six orthologs with high homology. The sequences (PWA96196.1) comprehending the highest query coverages (100%) and identity (73%) with the target sequence were selected for further studies. Comparative phylogenomic analysis of *AaPCS* genes performed by constructing a phylogenetic tree based on UPGMA (Fig. 8a) and maximum parsimony methods (Fig. 8b) revealed monophyletic origin with *S. tuberosum* (CAD68110.1) and with *O. sativa* (AAO13349.2), respectively. In contrasts, other members, viz., *B. juncea* (CAC37692.1) and *Anabaena* PCC 7120 (BAD10973.1), were observed to form a separate clade. The result of phylogenetic analysis among different plant species revealed homology of *AaPCS* with other homolog and ortholog members, established the evolutionary relationship between them and predicted that *AaPCS* is the closest homologue of *StPCS*.

Functional sites identification and domain analysis

The functional signature sequences underlying *AaPCS* and *anaPCS* were identified by PROSITE analysis which validated and confirmed that the query sequences of both *AaPCS* and *anaPCS* belongs to Papain-like cysteine peptidase superfamily. The position of functional signature sequences constituting phytochelatin catalytic domain from N-terminal end (NTD) was also retrieved. The results of InteProScan analysis further confirm the presence of phytochelatin catalytic domain from N-terminal domain in each *AaPCS* and *anaPCS*. In *AaPCS*, the N-terminal phytochelatin catalytic domain (NTD) prevailed from 7 to 234 amino acid residue where core signature sequences were occupied between Tyr⁷ and Arg²³⁴ (Supplementary Fig. 2), whereas in *anaPCS*, the N-terminal phytochelatin catalytic domain (NTD) existed between 25 to 237 amino acid residues consisting functional core residues in between Leu²⁵ and Lys²³⁷ (Supplementary Fig. 3). The multiple sequence alignment (Supplementary Fig. 4) results depicted conserved residues within the N-terminal phytochelatin catalytic domain (NTD) of the protein along with some insertion and deletion of the amino acid residues at certain positions that may have led to the sequence divergence in other homologs and orthologs,

Fig. 8 Phylogenetic tree: **a** UPGMA method and **b** maximum parsimony method showing origin, evolutionary relationship and functional homology of phytochelatin synthase protein in *A. annua*,

S. tuberosum, *A. thaliana*, *B. juncea*, *O. sativa* and *Anabaena* PCC 7120. The trees were generated using 1000 boot replication values based on maximum-likelihood methods

Fig. 9 Comparative analysis of similarities and differences in phytochelatin synthase (PCS) protein among *A. annua*, *S. tuberosum*, *A. thaliana*, *B. juncea*, *O. sativa* and *Anabaena* PCC 7120. The Circos software was used to generate circular map based on similarity percentage matrices using Clustal W algorithm

thus delineating their ancestral evolutionary origin. Furthermore, circos visualization tools confirmed the ancestral evolutionary similarity and differences across the members of phytochelatin synthase family (Fig. 9). At 50% cut-off values, *AaPCS* displayed close relation with *S. tuberosum* and *A. thaliana*, whereas *anaPCS* did not found in relation with any of the other homolog or ortholog members.

Conserved motif analysis

To comprehend conservation and diversification of PCS proteins in *A. annua*, putative functionally conserved network elements (motifs) of *AaPCS* were predicted using MEME software which predicted seven different motifs (Fig. 10). As anticipated, MEME motif analysis revealed that all the members in phylogenetic tree shared conjoint motif distribution pattern, signifying that PCS proteins within the same subclades might be involved in similar phytochelatin catalytic function. In our results, the motif scan analysis revealed significant motifs that comprised phytochelatin catalytic domain with in PCS proteins and statistical significance of the motifs was evaluated in terms of their *E*-value that speculate their adequate patterns and *p* value that conjectures how substantially each pattern correlate with the motif. The N-terminal catalytic domain (NTD) in *AaPCS* was represented by motif 3 (MAMASLYRRLPSPPAIDF; *p* value 1.3×10^{-23}) and propensity of the amino acids constituting

this motif was revealed based on size of the alphabets that correspond to amino acids. The motif distribution pattern revealed via motif scan analysis unveil commonalities in *A. annua*, *A. thaliana*, *S. tuberosum* and *B. juncea* (motif 3: MAMASLYRRLPSPPAIDF; *p* value 1.7×10^{-65}). Furthermore, the motif results showed the presence of additional 1 motif in *O. sativa* (RRLFRCAQAQ; *p* value 2.9×10^{-11}) and additional 2 motifs in *Anabaena* PCC 7120 (QVARQGLTLD; *p* value 1.0×10^{-12} and PNLIGFNENE; *p* value 1.1×10^{-9}) may have led to sequence divergence which could have been reason behind separate clustering of both *OsPCS* and *anaPCS* in the phylogenetic tree constructed in the present study.

Structural modelling of PCS functional domain

In the present study, the N-terminal catalytic domain of both *AaPCS* and *anaPCS* was modelled using relevant templates chosen based on crystal resolution, sequence similarity and completeness. A total of five models each for *AaPCS* (Fig. 11a) and *anaPCS* (Fig. 11b) were generated by MODELLER module of DS Modeller (Supplementary Table 2). In the present study, the models retaining lowest RMSD value corresponding to C α atoms within the template crystal structures and minimum electrostatic energy (calculated) were considered as good stable models which were selected for further interactions with CdCl₂ and Na₃AsO₄ (sodium arsenate) ligands modelled using ChemDraw tool. The predicted models were submitted to PMDB database and assigned PMDB IDs for our predicted structures of *AaPCS* (PM0083228) and *anaPCS* (PM0083229).

Model evaluation and validation

The three-dimensional models both for *AaPCS* and *anaPCS* were generated by satisfying all spatial restraint using Modeller Discovery Studio Client 3.0. Five sets of models each for *AaPCS* and *anaPCS* were arranged based on their Discrete Optimized Protein Energy (DOPE) and molecular Probability Density Functions (PDFs). The models (both for *A. annua* and *Anabaena* PCC 7120) exhibiting minimum DOPE scores and optimum energy values were selected for further validation using various statistics. The model proteins were subjected to Ramachandran plot analysis using RAMPAGE and PDBSum servers to assess the behavior of amino acid residues. The RAMPAGE statistics shows the presence of 95.4% residues in most favored region with 3.2% in additionally allowed region and 1.4% in outlier region for *AaPCS*, whereas for *anaPCS*, the statistics revealed the occurrence of 98.2% residues in most favored region with 1.8% in additionally allowed region and 0.0% in outlier region. Furthermore, VADAR analysis revealed that both the modelled proteins, viz., *AaPCS* and *anaPCS* existed as helix type structure 68 and 80 (30 and 36%) with 57 and 59

Fig. 10 The motif scan analysis showing distribution and presence/absence of common and uncommon motifs found in *A. annua*, *S. tuberosum*, *A. thaliana*, *B. juncea*, *O. sativa* and *Anabaena PCC 7120* discovered through MEME and MAST analysis. **a** The block diagram showing the sequence of discovered motifs for AaPCS. The

red arrows indicate the presence of uncommon motif in *O. sativa* and black arrows indicate the presence of uncommon motifs in *Anabaena PCC 7120* which are absent in other members. **b** The sequential logo of the motif 1 showing consensus sequences present in all the representatives' members

Fig. 11 Predicted structures of functional domain of **a** *A. annua* PCS (*AaPCS*) and **b** *Anabaena PCC 7120* PCS (*anaPCS*) generated and visualized through MODELLER module of Discovery Studio 3.0

(25 and 26%) and 96 and 80 (43 and 37%) coil type configuration having mean H-bond energy (-1.6 ($SD=1.1$) against the expected values (-2.0 ($SD=0.8$)). Furthermore, PROCHECK evaluation (Supplementary Fig. 6) revealed the occurrence of 87.6% (AaPCS) and 94.5% (anaPCS) residues in most favored region (A, B, L), with 8.8% (AaPCS) and 5.0% (anaPCS) in additionally allowed region (a, b, l, p). A protein model is generally considered as a good quality model if it has 90% of the residues in most favored region which corroborates the results of the present study. Additionally, ProSA results also validated the results of RAMPAGE and PROCHECK analysis, and confirmed that both the modelled proteins were precisely similar to template proteins. Moreover, RESPROX and QMEAN analysis confirmed the reliability of our predicted models by generating a composite score based on the template's proteins (Supplementary Table 3).

Protein–ligand interaction

To further explore the core amino acid residues of AaPCS and anaPCS involved in the binding with the ligand, molecular docking studies were performed for the possible investigation of binding mode of ligand, viz., cadmium chloride ($CdCl_2$) to the given receptors. The ligand, viz., $CdCl_2$

was modelled using the ChemDraw tool and was docked with the modelled proteins through HexDock server. In the present study, most reliable protein models of AaPCS and anaPCS were docked and stability of the docked complex was measured in terms of binding energy. For AaPCS, the most stable docked complex has the binding energy of $-134.58 \text{ kcal mol}^{-1}$, whereas in case of anaPCS, the most stable complex was docked with the binding energy of $-132.45 \text{ kcal mol}^{-1}$. The probable amino acid residues which are potentially involved in the interaction with the ligand are depicted in Fig. 12a. The docking studies for AaPCS revealed that the key residues, viz., Met²⁶, Met³⁸, Glu³⁹, Phe⁴², Val⁶⁵ and Leu⁶⁶ actively participated in the interaction with $CdCl_2$ (Fig. 12b). In contrast, the residues involved in the interaction of anaPCS with $CdCl_2$ were Thr¹⁰⁸, Leu¹⁰⁹, Leu¹¹², Ile¹⁵², Val¹⁵³, Asn¹⁵⁴, Ile¹⁶⁸ and Val²¹⁹ (Fig. 12d). Key amino acid residues involved during sodium arsenate interaction with AaPCS were Phe⁸⁴, Met⁸⁸, Asp⁹⁹, Lue¹¹¹ (Fig. 12c). However, sodium arsenate and anaPCS interaction involved Thr¹⁰⁸, Leu¹⁰⁹, Leu¹¹², Ile¹⁵², Val¹⁵³, Asn¹⁵⁴, Ile¹⁶⁸, Gly²¹⁷, Val²¹⁹ (Fig. 12e).

The result of molecular docking study revealed different amino acid residues were involved in the interaction of AaPCS with the ligand as those observed for anaPCS, thus clearly portraying that members of PCS family specifically

Fig. 12 **a** Comparative analysis of docked complex with experimentally resolved X-ray diffraction structures of functional domain of AaPCS and anaPCS with cadmium chloride ($CdCl_2$) and sodium

arsenate. **b, c** Structure of docked complexes in *A. annua* as visualized in Discovery Studio 3.0. **d, e** Structure of docked complex in *Anabaena* PCC 7120 as visualized in Discovery Studio 3.0

binds to ligand with varying amino acid residues that discloses a common interacting pattern.

Protein–protein interaction

The functional protein–protein interactive network of both AaPCS and anaPCS was predicted using STRING server at medium confidence level, and the results were further validated by comparing protein–protein interaction results by accessing PTIR database. At medium confidence level, AaPCS found to have strong interaction (Fig. 13a) with chloroplastic glutamatecysteine ligase (PGSC0003DMT400031827; score value 0.973 which regulates salicylic acid and phytoalexin biosynthesis, thereby regulating tolerance response through GSH-dependent developmental pathway), chloroplastic glutathione synthetase (PGSC0003DMT400073627; score value 0.955, which is exclusively involved in the glutathione biosynthesis and also regulates metal binding with in the plants) and lastly with gamma-glutamyl transpeptidases (PGSC0003DMT400022803; score value 0.955, which is mainly involved in the catabolism of glutathione, regulation of post-translational processing and cellular localization). On the other hand, anaPCS actively interacted (Fig. 13b) with ABC transporter ATP-binding protein (anaPCS; score value 0.999, which is involved in nutrients uptake including metals by utilizing energy from ATP binding), ABC

transporter permease (anaPCS; score value 0.996, regulates metal transport) and binding protein of ABC transporter component (anaPCS; score value 0.995, which is also involved in the import and export of various substrates including metal). From the results of the present study, it is evident that all the functional interactive partners of both AaPCS and anaPCS proteins belong to transmembrane protein families exhibiting metal transport activity thus signifying the coordinated role of PCS proteins in the transportation of heavy metals. The complete list of the interacting partners from both AaPCS and anaPCS have been shown in supporting information (supplementary table 4 and 5).

Superimposition results

Several studies have reported that protein structures are conserved to a greater extent compared to sequences. Keeping these facts in minds, the predicted structures of both AaPCS and anaPCS were superimposed over each other to predict their topological properties (Supplementary Fig. 7). The superimposition was performed using BLOSUM 62 matrix with gap penalty of 10 and extended penalty of 0.5. The superimposition results indicated that a total of 143 amino acid residues are substituted in between AaPCS and anaPCS with covering overall sequence similarity of 46.1% which was also further confirmed by the RMSD values of alpha carbon (2.06 Å) and backbone atoms (2.07 Å). In

Fig. 13 Functional interactive network of **a** AaPCS and **b** anaPCS with other protein family members as found on STRING server where the colored nodes describe query proteins from first shell interactors

and white nodes form second shell interactors. The large node size represents characterized proteins and smaller nodes for uncharacterized proteins

contrast, when AaPCS was allowed to be superimposed with the template 2BTW, the calculated RMSD values were 1.03 Å° for alpha carbon and 1.09 Å° for backbone atoms. Similarly, superimposition of anaPCS over template 2BTW showed RMSD values of 1.41 Å° both for alpha carbon and backbone atoms. The superimposition results confirmed the conservation of PCS functional domain structures as well as sequence across the divergent PCS members. However, structural alignment of functional domain of both AaPCS and anaPCS revealed some identical and non-identical (light blue) substitutions that extricated AaPCS domain from anaPCS. The amino acid substitutions were observed at position such as Ser⁵, Val¹⁰, Pro¹², Glu³⁴, Asn⁴⁷ and Gly⁷⁵ in AaPCS when compared to anaPCS, where Gln⁵, Thr¹², Asp²⁴¹ and Phe²⁴² were present.

Gene ontology enrichment analysis

The available protein sequences, viz., AaPCS and anaPCS, were subjected to CATH-Gene3D analysis to predict their subcellular localization and their functional annotation using gene ontology terms. Gene ontological (GO) terms serve as structural and functional descriptors characterizing genes and their products on the basis of their cellular, biological and molecular processes. The ReviGO analysis measured the biological function (Supplementary Fig. 8) of AaPCS protein revealed that 44% of the AaPCS protein is involved in the phytochelatin biosynthetic pathway (GO:0046938) and is potentially involved in the detoxification of metals such as arsenic, cadmium and copper (GO:0010273, GO:0046686, GO:0046685). The molecular function measured in terms of gene ontology revealed that the predicted proteins (56%) are mostly engaged in glutathione gamma-glutamyl cysteinyl transferase activity (GO:0016756) and respond well under metal stress condition (GO:0005507, GO:0046870). On the other hand, CELLO2GO analysis (Supplementary Fig. 9) revealed that 56.8% of the proteins are localized into plasma membrane (score value 2.839) involved in phytochelatin biosynthetic process (40.6%) and metal ion binding (35.7%). For anaPCS, 17.8, 11.1, 4.4, 8.9 and 2.2% of proteins are involved in phytochelatin biosynthetic process, detoxification of copper, response to arsenic, detoxification of cadmium ion and arsenic transport, respectively.

Discussion

In the present study, we aimed to enhance the heavy metal tolerance of *A. annua* hairy roots by expression of *Anabaena* PCC7120 phytochelatin synthase (*anaPCS*) gene. In their natural habitat, cyanobacteria including *Anabaena* are constantly exposed to adverse environmental condition such as high temperature, salinity, varied pH range and heavy

metals (Rai et al. 2014) which make them an ideal system for extracting stress-tolerant genes. Cloning of *anaPCS* was achieved in *E. coli* (Chaurasia et al. 2008) and *anaPCS* over-expression in *Anabaena* had shown multiple stress tolerance in the cyanobacteria (Chaurasia et al. 2017). The present study shows first-ever report of expression of cyanobacterial gene in *A. annua*.

We began with three different tissue samples, i.e., normal roots (derived from plantlets), hairy roots and transformed hairy roots (expressing *anaPCS* gene). Presence of *rolB*, presence of *rolB* and *anaPCS* and absence of both *rolB* and *anaPCS* confirmed the type of tissues as hairy roots, transformed hairy roots and normal roots, respectively (Figs. 1 and 4B).

Transformed hairy roots show significantly high amount of PCS expression owing to presence of two copies of PCS gene: one its own and another from *Anabaena* PCC 7120. As shown in Fig. 4, both As- and Cd-stress significantly induced the expression of *anaPCS* and *AaPCS* in transformed hairy roots, and as a result, As- and Cd-accumulation capacity was drastically enhanced in transformed hairy roots which accumulated 143.4 and 191% higher As and Cd than control. The value of bioconcentration factor (BCF) which gives an idea about capacity to accumulate heavy metals was significantly higher for transformed hairy roots. The BCF values of up to 21 and 9.7 for Cd and As, respectively, clearly explain that Cd is accumulated more than As; however, BCF values for both the heavy metals were significantly higher for transformed hairy roots as compared to non-transformed hairy roots or normal roots. In support of the present study, heterologous expression of *NnPCSI* from lotus (with proved phytoremediation potential) conferred Cd tolerance in *A. thaliana* (Liu et al. 2012). Similarly, PCS gene from many other plants such as garlic and Indian mustard has been expressed in *A. thaliana* (Guo et al. 2008; Gasic and Korban 2007) and enhanced Cd and As tolerance significantly. In addition, heterologous expression of PCS gene from an aquatic macrophyte has provided better As and Cd tolerance in Tobacco (Shukla et al. 2012).

Both As and Cd are non-redox active metals which, however, cause significant oxidative damage to cell. Plant cells possess excellent defense system comprising of non-enzymatic and enzymatic antioxidant to mitigate their adverse effects.

Differential responses of four antioxidative enzymes were observed under As- and Cd-stress in all tissue samples. In transformed hairy roots, all antioxidative enzymes except catalase were drastically induced under heavy metal stresses. At different locations in the cell, H₂O₂ is converted to water and oxygen through APX, GPX and catalase. Only a slight increase in catalase activity after As and Cd exposure in all tissue samples (normal roots, hairy roots and transformed hairy roots) may probably be due to its high sensitivity

towards superoxide radicals (Cakmak 2000). Catalase sensitivity under heavy metal stresses have also been reported in several plant species (Cao et al 2004; Gupta et al 2009) including *A. annua* (Rai et al 2011). Lesser catalase activity in heavy metal stressed transformed hairy roots was compensated by APX which possessed 31.9 and 34% greater activity compared to normal roots and hairy roots, respectively. Transformed hairy roots possess better defense against heavy metal stress attributed to its capability to maintain high level of APX, GR and POD in unstressed and stressed conditions. High activity of GR has probably contributed towards better As and Cd tolerance in transformed hairy roots of *A. annua*, because GR plays a significant role in heavy metal stresses by maintaining high GSH:GSSG ratio in the cell via recycling GSSG to GSH. As- and Cd- stress induced antioxidative enzyme activities including that for APX, GR and POD in several plants (Kumar et al 2014; Rai et al 2011; Zhang et al 2007).

In addition to antioxidative enzymes, low-molecular-weight thiols also play critical role in maintaining the overall oxidative balance in the cell. Their maximum induction in transformed hairy roots (with *anaPCS* gene) under As- and Cd- stresses (Fig. 5a) is indicative of better stress tolerance. GSH is one of the low-molecular-weight thiols with at least two crucial functions: synthesis of PCs and maintaining the ascorbate–glutathione cycle, both as part of antioxidative defense (Cobbett 2000). In the present study, transformed hairy roots expressing *anaPCS* showed approximately four-fold and twofold higher PCs level as compared to normal roots and non-transformed roots. As PCs are known to form metal-complexes which are sequestered into the vacuoles, boosted PCs level in transformed hairy roots suggests their potential to efficiently detoxify As and Cd. This can also be correlated with higher accumulation of As and Cd in transformed hairy roots than other samples and high GR activity which generates enough GSH to be used for PCs synthesis. Overexpression of *AaPCS* and *anaPCS* in transformed hairy roots supports the higher PCs level. However, in the present study, GSH levels were dropped under As- and Cd-stress. As product of PCS gene is being synthesized in high amount, lower GSH level in transformed hairy roots was probably because of the overconsumption for PCs synthesis and not due to its inhibition. The result finds support with Samane et al. (2007), Ben Ammar et al. (2008) and Meng et al. (2019) who have found lower GSH but higher PCs level in heavy metal stressed *Arabidopsis thaliana*, tomato, and lettuce, respectively.

In-silico analysis gave several confirmatory pieces of evidences for the proper functioning of a cyanobacterial *anaPCS* in a higher plant *A. annua*. The presence of conserved catalytic domain (Fig. 9) as well as functional domain structures (Supplementary Fig. 7) among *anaPCS* and *AaPCS* may have led cyanobacterial *anaPCS* to function

properly in *A. annua* hairy roots. Protein–protein interaction analysis confirmed *anaPCS* involvement in heavy metal transport as their functional interacting partners belong to transmembrane protein families having metal transport activity. This further get confirmation through gene ontology analysis where proportions of *anaPCS* proteins have been shown to be involved in biological processes like phytochelatin biosynthesis, arsenic transport and detoxification of cadmium ions.

In addition to better tolerance towards As and Cd, transformed hairy roots also showed the significantly greater synthesis of its popular antimalarial compound artemisinin (Fig. 7). Artemisinin biosynthesis have been shown to be improved under both arsenic (Kumari et al. 2018) and cadmium stress (Li et al. 2012). Owing to its immense therapeutic potential as antimalarial, anticarcinogenic, antimicrobial and many others, its global demand is always high and needs to be supplied enough to meet high demand (Pandey and Pandey-Rai 2016). Recently artemisinin has been suggested to be repurposed for the treatment of COVID-19 and other related respiratory disorders which will further increase its demand (Cheong et al. 2020). Higher artemisinin biosynthesis with enhanced heavy metal tolerance in hairy roots of *A. annua* expressing *anaPCS* gene can serve a dual purpose.

Conclusion

Heterologous expression of a cyanobacterial phytochelatin synthase (*anaPCS*) in hairy roots of *A. annua* was successfully achieved which led to enhanced heavy metal (arsenic and cadmium) tolerance. Under conditions of heavy metal stresses, greater accumulation of arsenic and cadmium, higher biosynthesis of phytochelatin, and activation of enzymatic and non-enzymatic antioxidants have made transformed hairy roots (expressing *anaPCS*) more tolerant to arsenic and cadmium stresses. Due to conserved catalytic domain and functional domain among *anaPCS* and *AaPCS*, the former could have performed in a different system (higher plant). In-silico analysis revealed involvement of *anaPCS* protein in heavy metal transport, phytochelatin biosynthesis and detoxification of cadmium and arsenic. Enhanced artemisinin biosynthesis in transformed hairy roots further gave scope of their probable use in phytoremediation and extraction of artemisinin from used hairy roots.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11816-021-00682-5>.

Acknowledgements The authors acknowledge Department of Science and Technology (DST- SERB), Govt. of India, for financial support (EMR/2017/001343). They also acknowledge the facility provided by Department of Biotechnology funded Sub-Distributed Information Centre, Centre for Bioinformatics, School of Biotechnology, Banaras

Hindu University. DST-PURSE, FIST and Central Instrument facility, Department of Botany, Banaras Hindu University is also acknowledged for providing financial and instrument support during the experiments. Authors also sincerely acknowledge Prof. L C Rai, Department of Botany, Institute of Science, Banaras Hindu University, Varanasi, for his support.

Author contributions NP, KKR, SKR and SPR designed the experiments. NP and KKR conducted the experiments. NP, KKR, SKR and SPR analyzed the data and wrote the manuscript.

Funding Department of Science and Technology (DST-SERB), Govt. of India, (EMR/2017/001343).

Availability of data and materials All raw data have been provided as supplementary material.

Declarations

Conflict of interest Authors declare no conflict of interest.

References

- Agostini E, Talano MA, González PS, Oller ALW, Medina MI (2013) Application of hairy roots for phytoremediation: what makes them an interesting tool for this purpose? *Appl Microbiol Biotechnol* 97(3):1017–1030
- Anderson ME (1985) Determination of glutathione and glutathione disulfide in biological samples. *Methods Enzymol* 113:548–555
- Bailey TL, Williams N, Misleh C, LWW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34:369–373
- Ben Ammar W, Mediouni C, Tray B, Ghorbel MH, Jemal F (2008) Glutathione and phytochelatin contents in tomato plants exposed to cadmium. *Biol Plant* 52:314–320. <https://doi.org/10.1007/s10535-008-0065-9>
- Benkert P, Kunzli M, Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* 37:W510–W514
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H et al (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
- Bleeker PM, Hakvoort HW, Blik M, Souer E, Schat H (2006) Enhanced arsenate reduction by a CDC25-like tyrosine phosphatase explains increased phytochelatin accumulation in arsenate-tolerant *Holcus lanatus*. *The Plant J* 45(6):917–929
- Cakmak I (2000) Tansley review No. 111 possible roles of zinc in protecting plant cells from damage by reactive oxygen species. *New Phytol* 146(2):185–205
- Cao X, Ma LQ, Tu C (2004) Antioxidant responses to arsenic in the arsenic hyperaccumulator Chinese brake fern (*Pteris vittata* L.). *Environ Pollut* 128:317–325. <https://doi.org/10.1016/j.envpol.2003.09.018>
- Castrignano T, De Meo PD, Cozzetto D, Talamo IG, Tramontano A (2006) The PDB protein model database. *Nucleic Acids Res* 34:306–309
- Chaurasia N, Mishra Y, Rai LC (2008) Cloning expression and analysis of phytochelatin synthase (pcs) gene from *Anabaena* sp. PCC 7120 offering multiple stress tolerance in *Escherichia coli*. *Biochem Biophys Res Commun* 376(1):225–230
- Chaurasia N, Mishra Y, Chatterjee A, Rai R, Rai YS, LC, (2017) Over-expression of phytochelatin synthase (pcs) enhances abiotic stress tolerance by altering the proteome of transformed *Anabaena* sp. PCC 7120. *Protoplasma* 254(4):1715–1724
- Cheong DH, Tan WD, Tran WWF, T, (2020) Anti-malarial drug, artemisinin and its derivatives for the treatment of respiratory diseases. *Pharmacol Res* 158:104901
- Cobbett CS (2000) Phytochelatin and their roles in heavy metal detoxification. *Plant Physiol* 123(3):825–832
- Colovos C, Yeates TO (1993) ERRAT: an empirical atom-based method for validating protein structures. *Protein Sci* 2:1511–1519
- Das S, Sillitoe I, Lee D, Lees JG, Dawson NL, Ward J, Orengo CA (2015) CATH FunFMMer web server: protein functional annotations using functional family assignments. *Nucleic Acids Res* 43(W1):W148–W153. <https://doi.org/10.1093/nar/gkv488>
- de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E et al (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 1:362–365
- Gasic K, Korban SS (2007) Transgenic Indian mustard (*Brassica juncea*) plants expressing an *Arabidopsis* phytochelatin synthase (AtPCS1) exhibit enhanced As and Cd tolerance. *Plant Mol Biol* 64(4):361–369
- Guo J, Dai X, Xu W, Ma M (2008) Overexpressing GSH1 and AsPCS1 simultaneously increases the tolerance and accumulation of cadmium and arsenic in *Arabidopsis thaliana*. *Chemosphere* 72(7):1020–1026
- Gupta M, Sharma P, Sarin NB, Sinha AK (2009) Differential response of arsenic stress in two varieties of *Brassica juncea* L. *Chemosphere* 74:1201–1208. <https://doi.org/10.1016/j.chemosphere.2008.11.023>
- Gutierrez-Valdes N, Häkkinen ST, Lemasson C, Guillet M, Oksman-Caldentey KM, Ritala A, Cardon F (2020) Hairy root cultures—a versatile tool with multiple applications. *Front Plant Sci* 11:33. <https://doi.org/10.3389/fpls.2020.00033>
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucleic Acids Symposium Series* [London]: Information Retrieval Ltd., c1979–c2000. 41(41): 95–98
- Hooykaas PJJ, Klapwijk PM, Nuti MP, Schilperoort RA, Rörsch A (1977) Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent agrobacteria and to Rhizobium ex planta. *Microbiology* 98(2):477–484
- Huang B (2009) MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS* 13(4):325–330
- Ismael MA, Elyamine AM, Moussa MG, Cai M, Zhao X, Hu C (2019) Cadmium in plants: uptake, toxicity, and its interactions with selenium fertilizers. *Metallomics* 11(2):255–277
- Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C et al (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240
- Khalid S, Shahid M, Niazi NK, Rafiq M, Bakhat HF, Imran M, Abbas T, Bibi I, Dumat C (2017) Arsenic behaviour in soil-plant system: biogeochemical reactions and chemical speciation influences. Enhancing cleanup of environmental pollutants. Springer, Cham, pp 97–140
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D et al (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645
- Kumar A, Singh RP, Singh PK, Awasthi S, Chakrabarty D, Trivedi PK, Tripathi RD (2014) Selenium ameliorates arsenic induced oxidative stress through modulation of antioxidant enzymes and thiols in rice (*Oryza sativa* L.). *Ecotoxicology* 23(7):1153–1163. <https://doi.org/10.1007/s10646-014-1257-z>
- Kumari A, Pandey N, Pandey-Rai S (2017) Protection of *Artemisia annua* roots and leaves against oxidative stress induced by arsenic. *Biol Plant* 61(2):367–377
- Kumari A, Pandey-Rai PN, S, (2018) Exogenous salicylic acid-mediated modulation of arsenic stress tolerance with enhanced accumulation

- of secondary metabolites and improved size of glandular trichomes in *Artemisia annua* L. *Protoplasma* 255(1):139–152
- Laskowski RA, Chistyakov VV, Thornton JM (2005) PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res* 1:D266–D268
- LeBlanc MS, McKinney EC, Meagher RB, Smith AP (2013) Hijacking membrane transporters for arsenic phytoextraction. *J Biotechnol* 163(1):1–9
- Li X, Zhao M, Guo L, Huang L (2012) Effect of cadmium on photosynthetic pigments, lipid peroxidation, antioxidants, and artemisinin in hydroponically grown *Artemisia annua*. *J Environ Sci* 24(8):1511–1518
- Liu Z, Gu C, Chen F, Yang D, Wu K, Chen S, Jiang J, Zhang Z (2012) Heterologous expression of a *Nelumbo nucifera* phytochelatin synthase gene enhances cadmium tolerance in *Arabidopsis thaliana*. *Appl Biochem Biotech* 166(3):722–734
- Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, Prisant MG et al (2003) Structure validation by Ca geometry: phi-psi and C-beta deviation. *Proteins* 50:437–450
- Macindoe G, Mavridis L, Venkatraman V, Devignes M, Ritchie D (2010) Hex Server: an FFT-based protein docking server powered by graphics processors. *Nucleic Acids Res* 38:W445–W449
- Maiti R, Van Domeselaar GH, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res* 32:W590–W594
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C et al (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 39:D225–D229
- Meng Y, Zhang L, Wang L, Zhou C, Shangguang Y, Yang Y (2019) Antioxidative enzymes activity and thiol metabolism in three leafy vegetables under Cd stress. *Ecotoxicol Environ Saf* 173:214–224
- Mukhopadhyay R, Bhattacharjee H, Rosen BP (2014) Aquaglyceroporins: generalized metalloid channels. *Biochim Biophys Acta* 1840(5):1583–1591
- Nagalakshmi N, Prasad MNV (2001) Responses of glutathione cycle enzymes and glutathione metabolism to copper stress in *Scenedesmus bijugatus*. *Plant Sci* 160(2):291–299
- Pandey N, Pandey-Rai S (2014) Short term UV-B radiation-mediated transcriptional responses and altered secondary metabolism of in vitro propagated plantlets of *Artemisia annua* L. *Plant Cell Tissue Organ Cult* 116(3):371–385
- Pandey N, Pandey-Rai S (2016) Updates on artemisinin: an insight to mode of actions and strategies for enhanced global production. *Protoplasma* 253(1):15–30
- Patra N, Srivastava AK (2014) Enhanced production of artemisinin by hairy root cultivation of *Artemisia annua* in a modified stirred tank reactor. *Appl Biochem Biotech* 174(6):2209–2222
- Pu GB, Ma DM, Chen JL, Ma LQ, Wang H, Li GF, Ye HC, Liu BY (2009) Salicylic acid activates artemisinin biosynthesis in *Artemisia annua* L. *Plant Cell Rep* 28(7):1127–1135
- Rai R, Pandey S, Rai SP (2011) Arsenic-induced changes in morphological, physiological, and biochemical attributes and artemisinin biosynthesis in *Artemisia annua*, an antimalarial plant. *Ecotoxicology* 20(8):1900–1913
- Rai R, Pandey S, Shrivastava AK, Pandey-Rai S (2014) Enhanced photosynthesis and carbon metabolism favor arsenic tolerance in *Artemisia annua*, a medicinal plant as revealed by homology-based proteomics. *Int J Proteom* 2014:163962
- Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* 111(1):1–61
- Sambrook J, Russell DW (2001) *Molecular Cloning. A laboratory manual*, vol 2, 3rd edn. Cold Spring Harbor Laboratory Press, New York
- Semane B, Cuypers A, Smeets K, Van Belleghem F, Horemans N, Schat H, Vangronsveld J (2007) Cadmium responses in *Arabidopsis thaliana*: glutathione metabolism and antioxidative defence system. *Physiol Plant* 129:519–528. <https://doi.org/10.1111/j.1399-3054.2006.00822.x>
- Shahi SK, Singh VK, Kumar A, Gupta SK, Singh SK (2013) Interaction of dihydrofolate reductase and aminoglycoside adenyl transferase enzyme from *Klebsiella pneumoniae* multidrug resistant strain DF12SA with clindamycin: a molecular modeling and docking study. *J Mol Model* 19:973–983
- Shukla D, Kesari R, Mishra S, Dwivedi S, Tripathi RD, Nath P, Trivedi PK (2012) Expression of phytochelatin synthase from aquatic macrophyte *Ceratophyllum demersum* L. enhances cadmium and arsenic accumulation in tobacco. *Plant Cell Rep* 31(9):1687–1699
- Sillitoe I, Lewis TE, Cuff A, Das S, Ashford P, Dawson NL et al (2015) CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res* 43:D376–D381
- Srivastava AK, Ara A, Bhargava P, Mishra Y, Rai SP, Rai LC (2007) A rapid and cost-effective method of genomic DNA isolation from cyanobacterial culture, mat and soil suitable for genomic fingerprinting and community analysis. *J Appl Phycol* 19(4):373–382
- Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* 6(7):e21800
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J et al (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43:D447–D452
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Usman ARA, Alkredaa RS, Al-Wabel MI (2013) Heavy metal contamination in sediments and mangroves from the coast of Red Sea: *Avicennia* sp. Marina as potential metal bioaccumulator. *Ecotoxicol Environ Saf* 97:263–270. <https://doi.org/10.1016/j.ecoenv.2013.08.009>
- Vatamaniuk OK, Mari S, Lu YP, Rea PA (1999) AtPCS1, a phytochelatin synthase from *Arabidopsis*: isolation and in vitro reconstitution. *Proc Natl Acad Sci U S A* 96(12):7110–7115
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
- Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* 31(13):3316–3319
- Xu D, Zhang Y (2011) Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 101(10):2525–2534
- Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, Lu CH (2014) CELLO2GO: a web server for protein subCELLular LOcalization prediction with functional geneontology annotation. *PLoS One* 9(6):e99368
- Zhang FQ, Wang YS, Lou ZP, Dong JD (2007) Effect of heavy metal stress on antioxidative enzymes and lipid peroxidation in leaves and roots of two mangrove plant seedlings (*Kandelia candel* and *Bruguiera gymnorhiza*). *Chemosphere* 67(1):44–50. <https://doi.org/10.1016/j.chemosphere.2006.10.007>
- Zhang W, Lin K, Zhou J, Zhang W, Liu L, Han X (2013) Spatial distribution and toxicity of cadmium in the joint presence of sulfur in rice seedling. *Environ Toxicol Pharmacol* 36(3):1235–1241
- Zhao SS, Zeng MY (1985) Spectrometric high-pressure liquid chromatographic (hplc) studies on the analysis of qinghaosu. *Planta Med* 51:323–323

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.