



ELSEVIER

Contents lists available at ScienceDirect

Comptes Rendus Biologies

www.sciencedirect.com



Molecular biology and genetics/Biologie et génétique moléculaires

Structural identification of putative USPs in *Catharanthus roseus*

Ahmed Bahieldin^{a,b,*}, Ahmed Atef^a, Ahmed M. Shokry^{a,c}, Saleh Al-Karim^a, Sanaa G. Al Attas^a, Nour O. Gadallah^{a,d}, Sherif Edris^{a,b,e}, Magdy A. Al-Kordy^{a,d}, Abdulkader M. Shaikh Omer^a, Jamal S.M. Sabir^a, Ahmed M. Ramadan^{a,c}, Abdulrahman S.M. Al-Hajar^a, Rania M. Makki^a, Sabah M. Hassan^{a,b}, Fotouh M. El-Domyati^{a,b}

^a Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), P.O. Box 80141, Jeddah 21589, Saudi Arabia

^b Department of Genetics, Faculty of Agriculture, Ain Shams University, Cairo, Egypt

^c Agricultural Genetic Engineering Research Institute (AGERI), Agriculture Research Center (ARC), Giza, Egypt

^d Genetics and Cytology Department, Genetic Engineering and Biotechnology Division, National Research Center, Dokki, Egypt

^e Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), Faculty of Medicine, King Abdulaziz University (KAU), Jeddah, Saudi Arabia

ARTICLE INFO

Article history:

Received 13 July 2015

Accepted after revision 18 July 2015

Available online 28 August 2015

Keywords:

Universal stress proteins

Domain

1MJH

1JMV

ABSTRACT

Nucleotide sequences of the *C. roseus* SRA database were assembled and translated in order to detect putative universal stress proteins (USPs). Based on the known conserved USPA domain, 24 Pfam putative USPA proteins in *C. roseus* were detected and arranged in six architectures. The USPA-like domain was detected in all architectures, while the protein kinase-like (or PK-like), ^{tyr}PK-like and/or U-box domains are shown downstream it. Three other domains were also shown to coexist with the USPA domain in *C. roseus* putative USPA sequences. These domains are tetratricopeptide repeat (or TPR), apolipoprotein III (or apoLp-III) and Hsp90 co-chaperone Cdc37. Subsequent analysis divided USPA-like domains based on the ability to bind ATP. The multiple sequence alignment indicated the occurrence of eight *C. roseus* residues of known features of the bacterial 1MJH secondary structure. The data of the phylogenetic tree indicated several distinct groups of USPA-like domains confirming the presence of high level of sequence conservation between the plant and bacterial USPA-like sequences.

© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

* Corresponding author at: Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), P.O. Box 80141, Jeddah 21589, Saudi Arabia.

E-mail addresses: bahieldin55@gmail.com (A. Bahieldin), ahmed_atefaig2@yahoo.com (A. Atef), ahmedshokry1@gmail.com (A.M. Shokry), Skarim@kau.edu.sa (S. Al-Karim), sangtas@mit.edu (S.G. Al Attas), nouromar71@yahoo.com (N.O. Gadallah), sedris@aucegypt.edu (S. Edris), m_alkordy@yahoo.com (M.A. Al-Kordy), aomar@kau.edu.sa (A.M.S. Omer), jsabir2622@gmail.com (Jamal S.M. Sabir), ahmedramadan782@yahoo.com (A.M. Ramadan), ahajarr@gmail.com (Abdulrahman S.M. Al-Hajar), rania_aryah@hotmail.co.uk (R.M. Makki), sabmahmoud@yahoo.com (S.M. Hassan), fm_domyati@hotmail.com (F.M. El-Domyati).

<http://dx.doi.org/10.1016/j.crvi.2015.07.008>

1631-0691/© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Catharanthus roseus is a medicinal plant species that has many pharmaceutical applications. This plant species possesses over 100 terpenoid indole alkaloids (TIAs) including the two important anticancer bisindole alkaloids namely vinblastine and vincristine [1,2]. As being able to grow in the wild habitat, this plant species is considered as an attractive target for characterizing genetic makeup and detecting important genes that might be involved in the tolerance to a wide range of biotic and abiotic stresses [3,4].

The universal stress proteins (USPs) were reported to play a role in tolerating prolonged exposure to heat, starvation, different biotic/abiotic stresses or DNA damage [5]. However, the exact function of USP remains to be elucidated. The USP domain is a protein structure that is widespread in prokaryotes. This structure originally refers to USPA (universal stress protein A) of *Escherichia coli* due to its prominence in the stationary phase. However, the structure of USP (namely 1MJH or MJ0577) is also resolved in *Methanococcus jannaschii* [6]. Biochemical analysis showed that MJ0577 binds ATP, hence, it mediates ATP-dependent functions molecular switching. Another type of USP that exists in *Haemophilus influenza* (1JMV) cannot bind ATP. Many studies indicated the lineage of the plant USPA-like domain from a 1MJH-like ancestor [5,7]. The USPA-like domain of different organisms is either found solely as a small protein, or linked to another domain, likely a protein kinase, downstream the N-terminus of the protein [7]. Multiple sequence alignment and topographic cladogram of the USPs in plant indicated the presence of protein kinase domain types 1.3.1, 1.3.2 or 1.4.1. Annotation of these domains indicated similarity in function to the tomato Pto interactor protein or Pti1 [8]. However, very little is known about the function of the *uspA* gene family in plants.

The number of pathway databases (PDBs) constructed from annotated genomes of biological systems is increasing [9–13]. More recently, Van Moerkercke et al. [14] have conducted an RNA-Seq analysis to construct the CathaCyc database of different organs and hairy roots genotypes of the *C. roseus* transcriptomes. This database was utilized in the present work for the identification of putative USPA proteins in *C. roseus*.

In the present study, we have detected and characterized a number of 24 putative USPs in *C. roseus* from the SRA database available in the NCBI in order to detect topology and possible functions.

2. Materials and methods

A number of 19 RNA-seq raw data of *C. roseus* was retrieved from SRA database (<http://www.ncbi.nlm.nih.gov/sra>, experiment SRP005953). RSEM (v1.1.6) in conjunction with Bowtie aligner (Bowtie v0.12.1) were used to estimate the relative abundances and expected read counts for the transcripts and to map the reads against the transcripts. Then, data were merged and *de novo*-assembled using Trinity-RNAseq v (r2013_08_14).

Based on the conserved USPA-like domain of 1MJH from *Methanococcus jannaschii*, putative USPs harboring USPA-like domain in *C. roseus* were characterized. Conserved USPA-like domain sequences were examined by using the CDD of the NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) in conjunction with BLAST searches. USPA-like domains of the recovered putative USPs in *C. roseus* were, then, compared to the two respective bacterial domains as well as 24 randomly selected USPA-like domains from *Arabidopsis*. Sequences of resulted *uspA*-like transcripts and proteins were deposited in the NCBI and USPA domains were identified by Transcoder program (<http://www.wowza.com/>

[transcodingsoftware](#)). For comparison, the *Arabidopsis* sequences were retrieved from the Institute for Genomic Research (TIGR).

Multiple sequence alignment and construction of phylogenetic tree were done for the USPA-like domain of different putative USPs from *C. roseus* after removal of all the other up- and downstream protein sequences. Multiple sequence alignments among USPA-like domains in *C. roseus* and those known in *Methanococcus jannaschii* (protein MJ0577 or 1MJH), *Haemophilus influenza* (1JMV) as well as the randomly selected USPA-like domains in *Arabidopsis* were done by using DNAMAN 5.2.2 (<http://www.lynnon.com>). Then, the results obtained were utilized in constructing phylogenetic tree using the neighbor-joining (NJ) method based on MEGA version 3.1 [15] with bootstrap support at critical nodes (1000 replicates). ClustalW [16] was used to make the alignment, and the profile was generated using the method of Gribskov and Veretnik [17]. Alignments were reshuffled by bootstrap resampling (1000×) and bootstrap support of 40% or higher were used.

3. Results and discussion

3.1. Putative USPA structure in *C. roseus*

Putative USPAs in *C. roseus* were recovered from the SRA database of the NCBI and their structures were compared to those of *Methanococcus jannaschii* MJ0577 (1MJH) and the *Haemophilus influenzae* (1JMV). Accordingly, 24 Pfam putative USP proteins with the USPA-like domains (PF00582) in *C. roseus* were identified (Table 1). Twenty one of these proteins harbor full-length USPA domain sequences, two sequences (AID16052 & AID16053) are truncated in the N terminus and one sequence is truncated in the N and C termini (AID16048). These putative USPAs of *C. roseus* were arranged in six architectures (Table 1 and Fig. 1). Expectedly, the USPA-like domain is present in all the putative USPs analyzed. The PK-like (PF00069), ^{tyr}PK-like (PF07714) and U-box (PF04564) domains are found downstream four, one and two *C. roseus* putative USPs, respectively (Table 1). The latter numbers do not reflect the real presence of these domains in the *C. roseus* putative USPAs as few of them were truncated in the sequences downstream the USPA-like domain. Therefore, our subsequent comparisons was based only on the USPA-like domains of the different putative USPs under study. Besides, we have used USPA-like domains from *Arabidopsis* that are well-known for the presence of different protein kinase (PK) types (designated 1.3.1, 1.3.2, 1.4.1 or “1MJH-like, plant” in Kerk et al. [7]) downstream the USPA domain. PK-like and/or protein tyrosine kinase-like (or ^{tyr}PK) sequences in putative USPAs of *C. roseus* exist downstream of the USPA-like domain [18]. U-Box domain is usually detectable at extreme C terminus of PK. PK/U-Box structures were reported as modified zinc fingers that might play an important role in the process of protein ubiquitination. The U-Box portion of the molecule performs a coiled-coil conformation, which mediates protein-protein interactions [19,20].

Table 1

Description of the *C. roseus* 24 putative USPA proteins and *uspa*-like transcripts recovered from SRA database (<http://www.ncbi.nlm.nih.gov/sra>, experiment SRP005953) and deposited in the NCBI.

No.	Nucleotide accession no.	Protein accession no.	aa	Description	Group ^d
1	KJ634212	AID16030 ^a	175	USPA-like	I
2	KJ634213	AID16031	910	USPA-like, CDC37_N, PK-like, U-Box	VI
3	KJ634214	AID16032	166	USPA-like	I
4	KJ634215	AID16033	200	USPA-like	I
5	KJ634216	AID16034	770	USPA-like, TyrPK-like	III
6	KJ634217	AID16035	215	USPA-like	I
7	KJ634218	AID16036	240	USPA-like	I
8	KJ634219	AID16037 ^a	163	USPA-like	I
9	KJ634220	AID16038	694	USPA-like, PK-like	II
10	KJ634221	AID16039 ^a	163	USPA-like	I
11	KJ634222	AID16040 ^a	177	USPA-like	I
12	KJ634223	AID16041 ^a	169	USPA-like	I
13	KJ634224	AID16042 ^a	164	USPA-like	I
14	KJ634225	AID16043 ^a	167	USPA-like	I
15	KJ634226	AID16044 ^a	163	USPA-like	I
16	KJ634228	AID16045 ^a	180	USPA-like	I
17	KJ634229	AID16046 ^a	277	USPA-like	I
18	KJ634230	AID16047 ^a	277	USPA-like	I
19	KJ634231	AID16048 ^{a,b,c}	103	USPA-like	I
20	KJ634232	AID16049	478	TRP, USPA-like	IV
21	KJ634233	AID16050	799	USPA-like, ApoLp-III, PK-like, U-Box	V
22	KJ634234	AID16051	497	USPA-like, PK-like	II
23	KJ634235	AID16052 ^b	177	USPA-like	I
24	KJ634236	AID16053 ^b	148	USPA-like	I

^a Putative USPA proteins with 1MJH-like USPA domains.

^b USPA-like domain missing amino acids at the C terminus.

^c USPA-like domain missing amino acids at the N terminus.

^d See Fig. 1.

Interestingly, three other domains were also shown to coexist with the USPA-like domain in three *C. roseus* putative USP sequences. These domains are tetratricopeptide repeat (TPR, PF00515), apolipoprotein III (apoLp-III, PF07464) and Hsp90 co-chaperone Cdc37 (PF08565). TPR is a structural motif found in tandem arrays of 316 motifs [21], which fold together to produce a single, linear, protein-protein interaction module, solenoid domain to facilitate interactions with other protein(s). Proteins with such domains can involve protein kinase inhibitor with an evidence of functioning in chaperone, cell cycle, transcription, and protein transport complexes. The TPR motif was also reported to represent an ancient protein-protein interaction module used by other proteins to trigger specific functions. Apolipoprotein III proteins are a family of exchangeable apolipoproteins functionally important in

lipid transport and storage and lipoprotein metabolism in eukaryotes including insects [22]. Cdc37 domain is a Hsp90 co-chaperone protein that was reported to participate in cell cycle and signal transduction. It has been shown to form a complex with Hsp90 and protein kinase to direct Hsp90 to its target kinase [23,24].

3.2. Multiple sequence alignment and phylogenetic classification of USPA-like domains

The alignment of USPA-like domains of the 24 *C. roseus* was done including those of *Methanococcus jannaschii*, *Haemophilus influenzae* and *Arabidopsis* (Fig. 2). The alignment was annotated based on the known features of 1MJH secondary structure with five β strands, alternating with four α helices [6]. The general structure

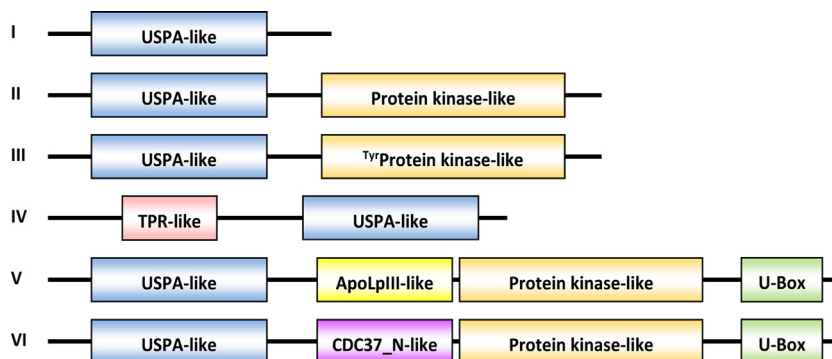


Fig. 1. (Color online.) Architectures of the described putative USPA proteins in *C. roseus* from database retrieved from SRA (<http://www.ncbi.nlm.nih.gov/sra>, experiment SRP005953).

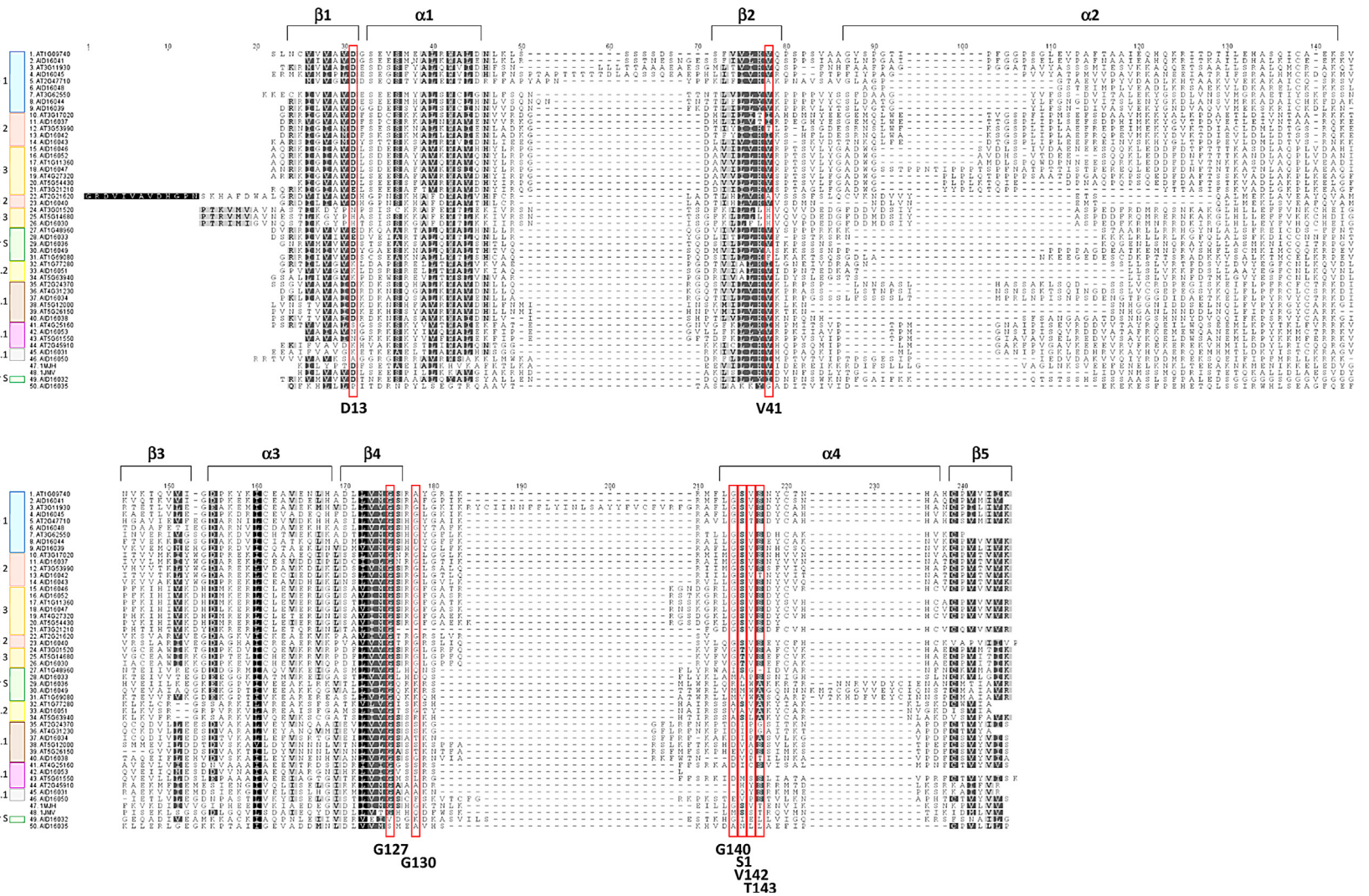


Fig. 2. (Color online.) Multiple sequence alignment (MSA) of the USP A-like domains of the putative USP A proteins shown in Tables 1 and S1. Bacterial 1MJH from *Methanococcus jannaschii* and 1JMV of *Haemophilus influenzae* were also used in the comparison. Shading was done using Clustal X [30] reflecting conserved features of 1MJH secondary structure (five β strands and four α helices [6]). Red boxes indicate residues for ATP binding sequences (D13, V41, G127, G130, G140, S141, V142 and T143). Types of the USP A-like domain are indicated as described in Table 1.

in *C. roseus* USPA-like domains involved the eight common residues required for ATP binding (D13, V41, G127, G130, G140, S141, V142 and T143) in 1MJH. This preliminary data indicates the possibility that most of the plant USP sequences might be arisen from a 1MJH-like ancestor. Referring to the classification made by Kerk et al. [7], USPA-like domain of each of the three types 1MJH-like, 1.3.2 or 1.4.1 aligned as a separate group of USPAs, except for those of types 1.3.1 and S (sequences not present in the PlantsP plant phosphorylation database [25]) that integrated in one group. USPA-like domains of *C. roseus* aligned close to those of USPA-like domain types indicated above, except for AID16032 and AID16035 that were not aligned close to any of these types. We speculate that the latter two accessions might not be USPs. Generally speaking, 1MJH was closely related to USPA-like domains of *C. roseus* than 1JMV (Fig. 2).

The eight residues in the USPA structure of 1MJH for ATP binding were also studied in USPA-like domains of *C. roseus* (Fig. 2). Accession no. AID16035 has none of the eight ATP binding residues. Accession AID16032 has only two residues, the first at position D13 and the second at position G140. The first position (e.g., D13) coordinates a Mn^{2+} ion, which binds the phosphate groups. This residue is conserved in different groups of USPA-like domains in *C. roseus* except for AID16030 of 1MJH-like 3 type, AID16051 of 1.3.2 type, and AID16053, AID16031 and AID16050 of 1.4.1 type. Position V41 of 1MJH, which binds adenine, is not conserved in AID16037 and AID16042 of 1MJH-like 2 type, AID16030 of 1MJH-like 3 type, AID16049 of 1.3.1 or S type and AID16050 of 1.4.1 type. The G127 position, which binds Rib, is conserved in all the sequences, except for AID16023 and AID16035. Position G130 of 1MJH, which binds beta phosphate, is conserved in the accessions of the three MJH1-like groups, except for AID16041 and AID16044 of MJH1-like 1. It is also conserved in AID16051 of 1.3.2 type, and AID16038 of 1.4.1 type. The G140 position is conserved only in accessions of the three 1MJH-like groups as well as accession AID16032 of 1.3.1 or S type. This residue does not bind to ATP, but it is conserved in all the MJH-like bacteria and plant. Position S141 in 1MJH, which binds gamma phosphate, is conserved only in the all MJH1-like sequences, except for AID16030 of 1MJH-like 3. The V142 position in 1MJH is conserved in all *C. roseus* accessions of 1MJH-like type. It is also conserved in accession AID16051 of 1.3.2 type. The T143 position, which binds alpha phosphate, is conserved only in accessions AID16042 of 1MJH-like 2 type and AID16038 of 1.4.1 type. Otherwise, this residue is replaced by S in accessions of the 1MJH-like type or A/S in accessions of the other types. In summary, most accessions of 1MJH-like type contain all ATP binding residues.

Based on the known nine features (five β strands and four α helices) of 1MJH secondary structure, numbers of 2 ((I/V)---D), 3 ((A/S)--A(L/V)), 6 (VILLH(V/I)), 0, 1 (I/V/L), 2 ((I/L)----- (V/I/L)), 6 ((V/I/L)(I/L)(I/V)(M/L/V)GS), 2 (S-(T/S/A)) and 3 (V-(V/A)(V/I)) were common in USPA-like domains of MJH1 and plant for structures β 1, α 1, β 2, α 2, β 3, α 3, β 4, α 4 and β 5, respectively. There are few residues that are common in USPA-like domain of 1JMV

and plant in structures β 1 (V-(V/I)), α 1 (A(V/L/I)) and α 3 ((D/E)----- (I/V)). In addition, there are few residues that are plant-specific, e.g., K/R in β 1 structure, W--- (N/H) in α 1 structure, I/V in α 3 structure, L/V/I between α 3 and β 4 structures, S/A in structure α 4 and C----- (R/K) in structure β 5 (Fig. 2).

The phylogenetic tree is constructed from the multiple sequence alignment (Fig. 3). The data indicated several distinct groups of USPA-like domains confirming the presence of high level of sequence conservation between the plant and bacterial USPA-like sequences. The results indicated the complete distinction among different types of USPA-like domain, except for 1.3.1 and S, where domains of both types were located in the same cluster. Kerk et al. [7] indicated that the sequences of types 1.4.1, 1.3.1, and 1.3.2 are diverged considerably from those of the small (S) *Arabidopsis* proteins. In the present study, we did not find this divergence as the USPA-like sequences of 1.3.1 were in the same cluster with those of S in either *C. roseus* or *Arabidopsis*. The 1MJH-like type of USPA-like domain was subdivided into three groups, e.g., 1MJH-like 1, 2 and 3. There are two *C. roseus* accessions, namely AID16032 and AID16035, that did not fit within any of the known types of USPA-like domains. There are 13 *C. roseus* accessions that were more closely related to the bacterial 1MJH. The rest of them were closely related to USPA-like types 1.3.1, 1.3.2, 1.4.1 and/or S. This indicates that the latter USPA-like types might not be originated from the bacterial 1MJH. Instead, the domain types 1.3.1, 1.3.2 and S were more related to the bacterial 1JMV. The radial phylogenetic tree confirmed the close relationship between plant USPA-like domains and 1JMV, where accession AID16035, followed by the USPA-like type 1.3.1 or S, were the closest to the bacterial 1JMV. However, the overall results indicate the chance that accessions AID16032 and AID16035 might not harbor a USPA domain.

We speculate that functional analysis of the *C. roseus* putative USPAs should involve the accessions related to 1MJH as they proved to harbor basic elements for ATP binding, which is the main feature of USP in conferring the tolerance against biotic and abiotic stresses. Biochemical function(s) of USP is not resolved up to date. There is incidence for the possible role of the small *Arabidopsis* putative USP in stress responses [7]. Hohnjec et al. [26] identified a possible derivative of 1MJH namely, VfENOD18 protein, in the broad bean (*Vicia faba*) that possesses an ATP-dependent function. Kerk et al. [7] indicated the possible functions of AT3G53990 and AT3G17020 in *Arabidopsis*. They indicated their sequence homology with ATE1 and ATE3, respectively. The first was reported to be involved in the post-translational conjugation of arginine to the N-terminal aspartic or glutamic acid of a protein and required for degradation of the protein via the ubiquitin pathway (<http://www.uniprot.org/uniprot/Q9ZT48>). Sawa et al. [27] indicated that the ATE genes are generally responsible for repression of trans-differentiation into xylem cells in *Arabidopsis* (<http://www.arabidopsis.org/servlets/TairObject?name=ATE3&type=locus>). In the present study, AT3G53990 was closely related to AID16042 and AT3G17020 was closely related to AID16037 (Fig. 3). Therefore, we expect that these two *C. roseus* putative

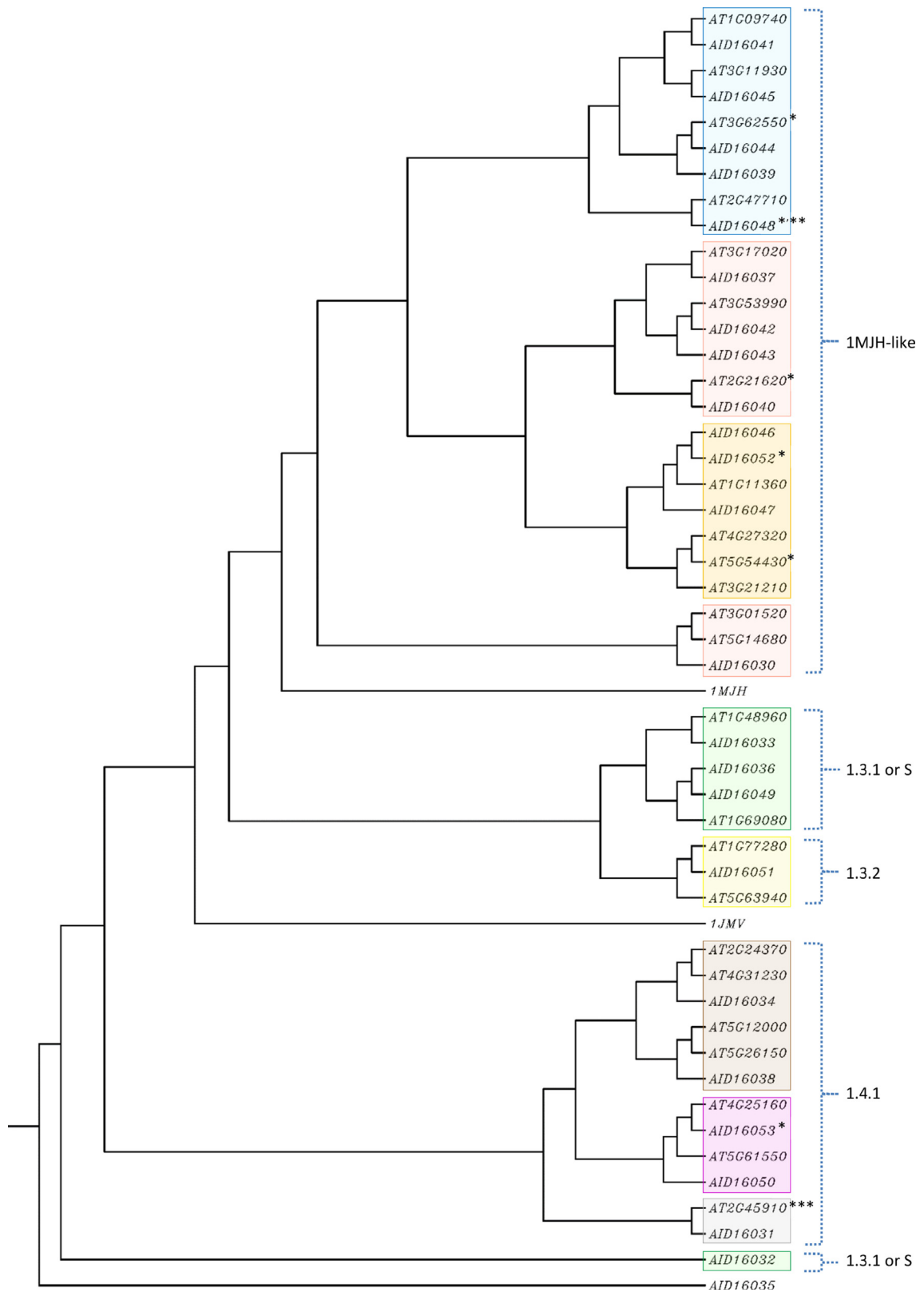


Fig. 3. (Color online.) Phylogenetic tree of the USA-like domains of the putative USA proteins shown in Tables 1 and S1. Bacterial 1MJH and 1JMV were also used in the comparison. Types of the USA-like domain are indicated as described in Table 1. *USA-like domain is missing amino acids at the C terminus. **USA-like domain is missing amino acids at the N terminus. ***No USA-like domain type or description is available.

USPAs might have the same functions. Zegzouti et al. [28], in their study of ethylene-induced gene expression in tomato (*Lycopersicon esculentum*), identified a transcript, namely ER6, that showed up-regulation under ethylene treatment. Kerk et al. [7] indicated that accession AT1G09740 is a homolog of ER6 protein. The present results indicated that the *Arabidopsis* accession is closely related to *C. roseus* accession AID16041 (Fig. 3). Therefore, we speculate that the latter might possess the same function.

There are other *Arabidopsis* protein accessions with suggested functions (Table S1). For example, AT3G62550 (closely related to *C. roseus* AID16044) is an ER6-like, AT3G11930 (closely related to *C. roseus* AID16045) is ethylene-responsive and AT2G21620 (closely related to *C. roseus* AID16040) might respond to severe desiccation and auxin treatment [7,29]. In conclusion, structural and phylogenetic analyses indicated that many USPA-like domains in *C. roseus* originated from a 1MJH-like ancestor. Functions of many of them have been speculated based on the known functions of their analogs in *Arabidopsis*.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgments

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant No. (15-3-1432/HiCi). The authors, therefore, acknowledge with thanks DSR technical and financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.crv.2015.07.008>.

References

- [1] S.K. Rijhwani, J.V. Shanks, Effect of elicitor dosage and exposure time on biosynthesis of indole alkaloids by *Catharanthus roseus* hairy root cultures, *Biotechnol. Prog.* 14 (1998) 442–449.
- [2] H. Rischer, M. Orešič, T. Seppänen-Laakso, M. Katajamaa, F. Lammertyn, et al., Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells, *Proc. Natl. Acad. Sci. USA* 103 (2006) 5614–5619.
- [3] T. Nystrom, F.C. Neidhardt, Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest, *Mol. Microbiol.* 11 (1994) 537–544.
- [4] R. Van der Heijden, D.I. Jacobs, W. Snoeijer, D. Hallard, R. Verpoorte, The *Catharanthus* alkaloids: pharmacognosy and biotechnology, *Curr. Med. Chem.* 11 (2004) 1241–1253.
- [5] W.T. Li, Y.M. Wei, J.R. Wang, C.J. Liu, X.J. Lan, et al., Identification, localization, and characterization of putative USP genes in barley, *Theor. Appl. Genet.* 121 (2010) 907–917.
- [6] T.I. Zarebinski, L.-W. Hung, H.-J. Mueller-Dieckmann, K.-K. Kim, H. Yokota, et al., Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics, *Proc. Natl. Acad. Sci. USA* 95 (1998) 15189–15193.
- [7] D. Kerk, J. Bulgrien, D.W. Smith, M. Gribskov, *Arabidopsis* proteins containing similarity to the universal stress protein domain of bacteria, *Plant Physiol.* 131 (2003) 1209–1219.
- [8] G. Sessa, M. D'Ascenzo, G.B. Martin, Thr38 and Ser198 are Pto autophosphorylation sites required for the AvrPto–Pto-mediated hypersensitive response, *EMBO J.* 19 (2000) 2257–2269.
- [9] P. Romero, J. Wagg, M.L. Green, D. Kaiser, M. Krummenacker, et al., Computational prediction of human metabolic pathways from the complete human genome, *Genome Biol.* 6 (2004) R2.
- [10] P. Zhang, H. Foerster, C.P. Tissier, L. Mueller, S. Paley, et al., MetaCyc and AraCyc. Metabolic pathway databases for plant research, *Plant Physiol.* 138 (2005) 27–37.
- [11] P. Jaiswal, J. Ni, I. Yap, D. Ware, W. Spooner, et al., Gramene: a bird's eye view of cereal genomes, *Nucleic Acids Res.* 34 (2006) D717–D723.
- [12] P. May, J.-O. Christian, S. Kempa, D. Walther, ChlamyCyc: an integrative systems biology database and web-portal for *Chlamydomonas reinhardtii*, *BMC Genom.* 10 (2009) 209.
- [13] P. Zhang, K. Dreher, A. Karthikeyan, A. Chi, A. Pujar, et al., Creation of a genome-wide metabolic pathway database for *Populus trichocarpa* using a new approach for reconstruction and curation of metabolic pathways for plants, *Plant Physiol.* 153 (2010) 1479–1491.
- [14] A. Van Moerkercke, M. Fabris, J. Pollier, G.J.E. Baart, S. Rombauts, et al., CathaCyc, a metabolic pathway database built from *Catharanthus roseus* RNA-Seq data, *Plant Cell Physiol.* 54 (2013) 673–685.
- [15] S. Kumar, K. Tamura, M. Nei, MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment, *Brief. Bioinform.* 5 (2004) 150–163.
- [16] D.G. Higgins, J.D. Thompson, T.J. Gibson, Using CLUSTAL for multiple sequence alignments, *Methods Enzymol.* 266 (1996) 383–401.
- [17] M. Gribskov, S. Veretnik, Identification of sequence patterns with profile analysis, *Methods Enzymol.* 266 (1996) 198–211.
- [18] R.D. Isokpehi, S.S. Simmons, H.H.P. Cohly, S.I.N. Ekunwe, G.B. Begonia, et al., Identification of drought-responsive universal stress proteins in Viridiplantae, *Bioinform.* 5 (2011) 41–58.
- [19] L. Aravind, E.V. Koonin, The U box is a modified RING finger: a common domain in ubiquitination, *Curr. Biol.* 10 (2000) R132–R134.
- [20] L. Aravind, V. Anantharaman, E.V. Koonin, Monophyly of class I aminoacyl tRNA synthetase, USPA, ETPF, photolyase, and PP-ATPase nucleotide-binding domains: implications for protein evolution in the RNA, *Proteins* 48 (2002) 1–14.
- [21] G.L. Blatch, M. Lässle, The tetratricopeptide repeat: a structural motif mediating protein–protein interactions, *Bioessays* 21 (1999) 932–939.
- [22] J. Wang, B.D. Sykes, R.O. Ryan, Structural basis for the conformational adaptability of apolipoprotein III, a helix–bundle exchangeable apolipoprotein, *Proc. Natl. Acad. Sci. USA* 99 (2002) 1188–1193.
- [23] K. Dai, R. Kobayashi, D. Beach, Physical interaction of mammalian CDC37 with CDK4, *J. Biol. Chem.* 271 (1996) 22030–22034.
- [24] L. Stepanova, X. Leng, S.B. Parker, J.W. Harper, Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4, *Genes Dev.* 10 (1996) 1491–1502.
- [25] M. Gribskov, F. Fana, J. Harper, D.A. Hope, A.C. Harmon, et al., PlantsP: a functional genomics database for plant phosphorylation, *Nucleic Acids Res.* 29 (2001) 111–113.
- [26] N. Hohnjec, H. Kuster, U. Albus, S.C. Frosch, J.D. Becker, et al., The broad bean nodulin VFNOD18 is a member of a novel family of plant proteins with homologies to the bacterial MJ0577 superfamily, *Mol. Gen. Genet.* 264 (2000) 241–250.
- [27] S. Sawa, T. Demura, G. Horiguchi, M. Kubo, H. Fukuda, The ATE genes are responsible for repression of transdifferentiation into xylem cells in *Arabidopsis*, *Plant Physiol.* 137 (2005) 141–148.
- [28] H. Zegzouti, B. Jones, P. Frasse, C. Marty, M. Maitre, et al., Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display, *Plant J.* 18 (1999) 589–600.
- [29] K. Yamaguchi-Shinozaki, M. Koizumi, S. Urao, K. Shinozaki, Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in *Arabidopsis thaliana*: sequence analysis of one cDNA clone that encodes a putative transmembrane channel protein, *Plant Cell Physiol.* 33 (1992) 217–224.
- [30] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25 (1997) 4876–4882.