SHORT COMMUNICATION

Isolation and *in silico* characterization of cDNA encoding cyclophilin from etiolated *Vigna mungo* seedlings

Kalika Kuhar[1], Varun Kumar Gupta[1], Rekha Kansal[2], Vijay Kumar Gupta[1]*

[1]Department of Biochemistry, Kurukshetra University, Kurukshetra, Haryana, India.

[2]National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute (PUSA), New Delhi, India.

*Corresponding author: vkgupta59@rediffmail.com

Received: 13 February, 2012; Accepted: 23 June, 2012

ABSTRACT

A full-length cDNA clone encoding cyclophilin gene of 848 bp, including a 519 bp open reading frame, has been isolated from the cDNA library constructed from etiolated seedlings of *Vigna mungo* (GenBank FN668732). The cDNA sequence showed 97% identity with *Vigna radiata* cyclophilin mRNA. The sequence was GC rich and lacked introns. The open reading frame encoded 172 amino acid polypeptide with molecular weight 18.3 kDa and theoretical pI 8.61. BlastP analysis indicated that its putative amino acid sequence shared 100% identity with several plant cyclophilins particularly legumes. The conserved seven amino acid residues region in *V. mungo* cyclophilin was RSGKPLH (present in legumes) instead of KSGKPLH, indicating its similarity to the cyclophilins of other legumes. This novel *V. mungo* cyclophilin gene will broaden the pool of plant cyclophilin genes for further studies.

Keywords: gene library, open reading frame.

Cyclophilins (CyPs), first identified as the cellular receptors of the immunosuppressant drug cyclosporin A in mammalian cells (Handschumacher et al., 1984), constitute a family of highly conserved proteins which appear to be ubiquitous in organisms, ranging from bacteria to plants and animals (Galat, 1999). CyPs and FKBPs (receptors for the FK506 and rapamycin proteins) are collectively referred to as immunophilins. These proteins have peptidyl prolyl *cis-trans* isomerase activity that facilitates protein folding *in vivo* by catalyzing the isomerization of peptide bonds preceding proline residues (Brandts et al., 1975). Transcriptional analysis showed they are mostly 18–23 kDa proteins expressed throughout the plant, and different isoforms are present in all parts of the cell including cytosol, nucleus, mitochondria, secretory pathway, and chloroplast (Romano et al., 2004). The abundance and diversity of CyP isoforms suggests that plant CyPs are involved in a wide variety of cellular processes, including protein folding, mRNA processing, protein trafficking and maturation, signal transduction, and abiotic and biotic stress responses (Marivet et al., 1992; Chou and Gasser, 1997; Godoy et al., 2000; Sekhar et al., 2010; Que et al., 2011).

Plant CyPs were first identified in 1990 with the isolation of CyP cDNA sequences from tomato, maize, and *Brassica napus* (Gasser et al., 1990). Subsequently, CyP genes have been isolated from *Vicia faba* (Kinoshita and Shimazaki, 1999), *Glycine max* (Kan et al., 2002), *Arabidopsis thaliana* (Romano et al., 2004), maize (Marivet

et al., 1995), rice (Kumari et al., 2009), *Cajanus cajan* (Sekhar et al., 2010) and sugarcane (Que et al., 2011). While the *A. thaliana* genome contains the largest number (29) of CyP genes (Romano et al., 2004), the number of plant CyPs available in databases is small compared to that of other organisms.

To our knowledge, there is no report on isolation of CyP genes from *Vigna mungo*, which is an important food legume crop of the Indian subcontinent and its tolerance to various stresses is a major concern because of the comparatively less genetic diversity in this crop. In this study, we report the isolation and *in silico* characterization of a cDNA clone encoding CyP gene from the etiolated seedlings of *V. mungo*.

Total RNA was isolated from 8-day-old etiolated seedlings of *V. mungo* by LiCl method (Menke et al., 1996) and used for purification of mRNA using Oligotex mRNA isolation Kit (Qiagen). cDNA was synthesized using SMART™ PCR cDNA synthesis kit (Clontech). The cDNA fragments (0.5–2.0 kbp) were eluted from the gel, ligated to pGEMT Easy vector (Promega) and transformed into *Escherichia coli* DH5α competent cells to generate cDNA library. The transformants were spread on LA plates containing ampicillin, X-gal, and IPTG and incubated at 37°C for 18 to 24 hours. The screening of the cDNA library was carried out using pigeon pea protease inhibitor gene as probe according to the method of Sambrook et al. (1989). The positive colonies were got sequenced from Techno Concept, New Delhi, India and after analysis submitted to EMBL Nucleotide Data Base.

The nucleotide sequence was analyzed *in silico* using VecScreen program of National Center for Biotechnology Information (NCBI), Basic Local Alignment Search Tool (BLAST), FGENESH program of Softberry web server and BioEdit. Restriction map analysis of complete open reading frame (ORF) of *V. mungo* CyP gene sequence was generated using software available at www.nebcutter.com. Multiple sequence alignment was done on CLUSTAL W server (Thompson et al., 1994) available at www.genome.ad.jp. The physical and chemical parameters, exact mass and hydrophobic nature of the deduced protein sequence were computed using ExPASy web server. The phylogenetic tree of the deduced amino acid sequence was generated using software available on http://www.123genomics.com.

To isolate a full-length CyP gene from etiolated seedlings of *V. mungo*, a cDNA library was constructed which contained 95% recombinant colonies. On screening of this library by hybridization using the pigeon pea protease inhibitor as probe, four positive clones were identified and got sequenced. The BLASTN search analysis of these sequences revealed one unique full-length cDNA sequence of 848 bp, including an ORF of 519 bp in length with no introns (Figure 1). The sequence showed 97, 89, 85, 83, 82, 82, 80 and 79% identity with CyP mRNA (cDNA) of *Vigna radiata* (ABO20612.1), *G. max* (AF456323.1), *Phaseolus vulgaris* (X74403.1), soybean (BT092190.1), *Arachis digoi* (EU170616.1), *V. faba* (AB012947.1), *Lupinus luteus* (AF178458.1), and *A. thaliana* (AK226392.1) respectively indicating that the isolated gene belonged to a legume CyP gene family. The cDNA sequence encoding *V. mungo* CyP was submitted to the Genbank with the accession number (FN668732). The length of ORF (519 bp) for *V. mungo* CyP cDNA sequence was identical to the ORFs of pigeon pea (Sekhar et al., 2010), sugarcane (Que et al., 2011) and rice (Kumari et al., 2009). The nucleotide composition revealed that the *V. mungo* CyP gene was GC rich as it contained 21.39% A, 30.44% C, 29.87% G and 18.30% T residues. Restriction map analysis of the gene sequence showed the absence of restriction sites for commonly used restriction enzymes such as *Eco*RI, *Eco*RV, *Hind*III, *Hpa*I, *Not*I, *Pst*I, *Sal*I, *Sma*I, *Xba*I, *Xmn*I. The CyP gene was released from the clone using *Eco*RI and *Not*I enzymes as they were present at the flanking sites of MCS region but not internally in the ORF.

The ORF encoded 172 amino acid polypeptide (Figure 2) with molecular weight 18.3 kDa and theoretical pI 8.61. These values are in accordance with the previous reports on CyPs (Ye and Ng, 2001; Kumari et al., 2009). The putative CyP protein encoded by *V. mungo* cDNA had a highly conserved functional site Trp128 (W128) present in all CyPs in eukaryotes (Liu et al., 1991). The conserved seven amino acid residues region KSGKPLH (48–54), which was specific to cytoplasmic CyPs in plants (Lippuner et al., 1994), was also found but with one variation i.e. K (lysine) at position 48 being replaced by R (arginine). Similar to *V. mungo*, the CyPs of other legumes also contained RSGKPLH (48–54) sequence (Figure 2). In contrast, CyPs of non-legumes such as sugarcane contained the sequence KSGKPLH (Que et al., 2011). This observation implied that the seven conserved amino acid residues regions in CYPs of legumes were RSGKPLH.

BLASTP search analysis of the deduced *V. mungo* CyP amino acid sequence showed 100% homology to the CyPs of other legumes viz. *V. radiata*, *P. vulgaris*, *G. max* and so on. ClustalW analysis calculated the best match for the *V. mungo* cyp and lined them to identify their similarities and differences with the sequences available on NCBI Genbank (Figure 2). The sequence was found to be rich in glycine, threonine, valine, lysine and

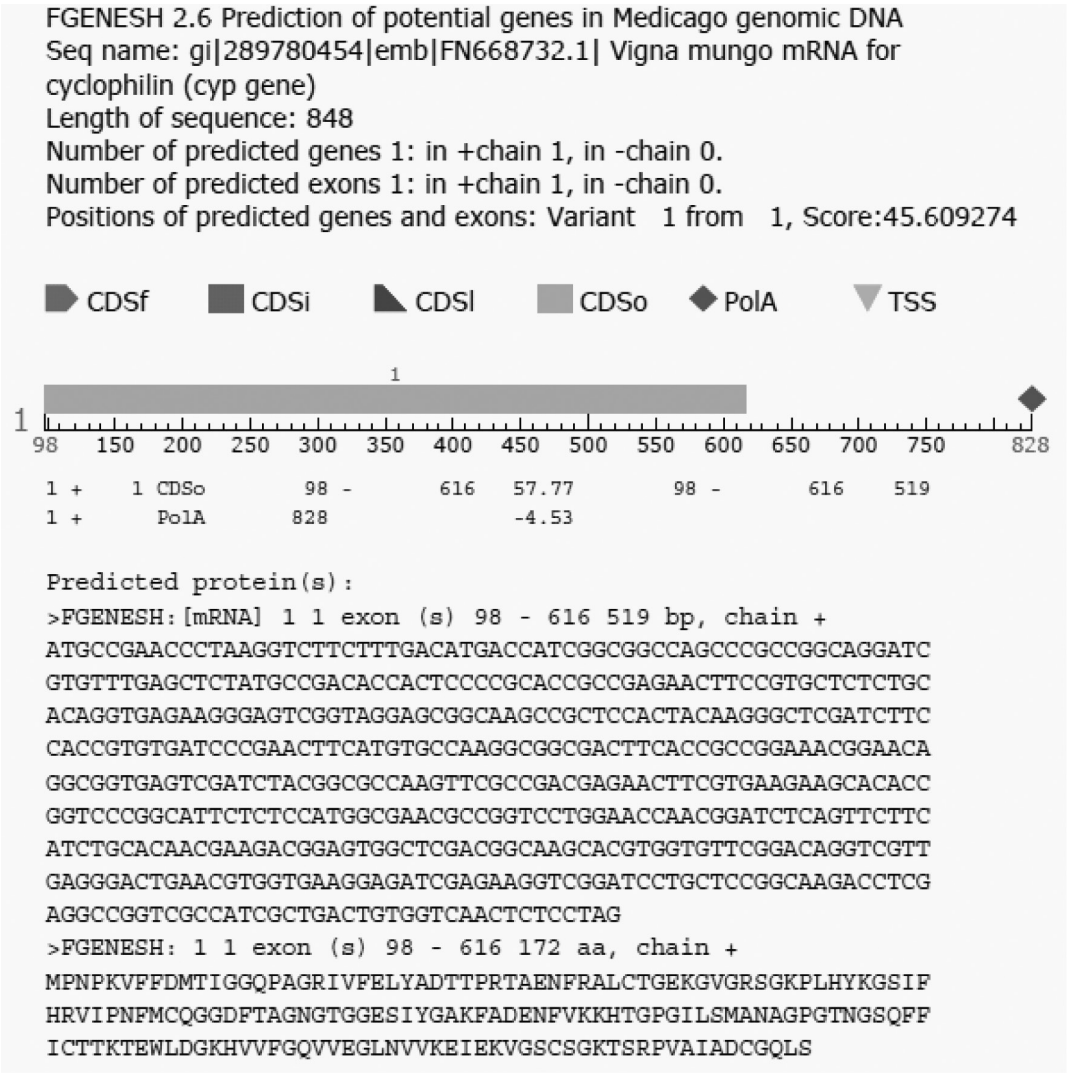

Figure 1. Nucleotide and deduced amino acid sequence of a cDNA encoding *Vigna mungo* cyclophilin protein by the National Center for Biotechnology Information (NCBI), Basic Local Alignment Search Tool (BLAST) program. The vector sequence was recognized by VecScreen program of NCBI. Open reading frame, number of exons and introns in the sequence were predicted using FGENESH program of Softberry web server.

phenylalanine. The instability index depicted a value of 20.09, which suggested the stable nature of this protein. The hydropathy index predicted hydrophilic nature of this protein implying its cytosolic/extracellular location. A dendrogram displaying phylogenetic relationship showed that *V. mungo* CyP sequence was closely related to CyPs from *V. radiata* and *P. vulgaris* but distant from the CyPs of *Triticum*, *Oryza*, and maize (Figure 3). The close relationship among legume CyPs might be due to the presence of similar conserved region of seven amino acids in the region of 48-54. The role of *V. mungo* CyP gene needs investigation for its exploitation in improving plant characteristics.

Overall, a novel full-length cDNA (848 bp) containing 519 bp ORF encoding CyP has been isolated from the cDNA library generated from etiolated seedlings of *V. mungo*. This gene was characterized using bioinformatics tools. It showed substantial homology to CyP genes from other legumes. The ORF encoded 172 amino acids polypeptide having a highly conserved functional site Trp128 and RSGKPLH (48-54) sequence conserved in CyPs of legumes. Phylogenetic analysis also predicted close relationship among CyP genes of *V. mungo*, *V. radiata* and *P. vulgaris*. The isolated novel gene will add to the resource of plant CyP genes for further use after investigation of its role.

```
VmCyc           MPNPKVFFDMTIGGQPAGRIVFELYADTTPRTAENFRALCTGEKGVGRSGKPLHYKGSIF 60
Vigna           MPNPKVFFDMTIGGQPAGKIVFELFADTTPRTAENFRALCTGEKGVGRSGKPLHYKGSIF 60
Phaseolus       MPNPKVFFDMTIGGQPAGRIVFELYADVTPRTAENFRALCTGEKGVGRSGKPVHFKGSIF 60
Vicia           MSNPKVFFDMTVGGQNAGRIIFELFADVTPRTAENFRALCTGEKGVGRSGKPLHFKGSSF 60
Glycine         MPNPKVFFDMTIGGQSAGRIVMELYADVTPRTAENFRALCTGEKGVGRSGKPLHYKGSSF 60
Brassica        ---PKVYFDMTVGDKAAGRIVMELYADTVPETAENFRALCTGERGIGKSGKPLHYKGSAF 57
Oryza           MSNTRVFFDMTVGGAPAGRIVMELYAKDVPRTAENFRALCTGEKGVGKSGKPLHYKGSTF 60
Zea             MANPRVFFDMTVGGAPAGRIVMELYANEVPKTAENFRALCTGEKGVGKSGKPLHYKGSTF 60
Triticum        MANPRVFFDMTVGGAPAGRIVMELYKDAVPRTVENFRALCTGEKGVGKSGKPLHYKGSSF 60
                   .:*:****:*.  **:*::**: . .*.*.**********:*:*:****:*:*** *

VmCyc           HRVIPNFMCQGGDFTAGNGTGGESIYGAKFADENFVKKHTGPGILSMANAGPGTNGSQFF 120
Vigna           HRVIPNFMCQGGDFTAGNGTGGESIYGAKFADENFVKKHTGPGILSMANAGPGTNGSQFF 120
Phaseolus       HRVIPNFMCQGGDFTAGNGTGGESIYGAKFADENFVKKHTGPGILSMANAGPGTNGSQFF 120
Vicia           HRVIPNFMCQGGDFTAGNGTGGESIYGSKFADENFIKKHTGPGILSMANAGPGTNGSQFF 120
Glycine         HRVIPSFMCQGGDFTAGNGTGGESIYGAKFADENFVKKHTGPGILSMANAGPGTNGSQFF 120
Brassica        HRVIPKFMCQGGDFTAGNGTGGESIYGMKFKDENFVKKHTGPGILSMRNAGSNTNGSQFF 117
Oryza           HRVIPEFMCQGGDFTRGNGTGGESIYGEKFADEVFKFKHDSPGILSTANAGPNTNGSQFF 120
Zea             HRVIPEFMCQGGDFTRGNGTGGESIYGEKFPDEKFVRKXTAPGVLSMANAGPNTNGSQFF 120
Triticum        HRVIPDFMCQGGDFTKGNGTGGESIYGEKFADEKFVHKHTKPGILSMANAGPNTNGSQFF 120
                *****.********* ********** ** ** *  *   **:**  ***..*******

VmCyc           ICTTKTEWLDGKHVVFGQVVEGLNVVKEIEKVGSCSGKTSRPVAIADCGQLS 172
Vigna           ICTTKTEWLDGKHVVFGQVVEGLNVVKEIEKVGSSSGKTSRPVAIADCGQLS 172
Phaseolus       ICTTKTEWLDGKHVVFGQVVEGLDVVKDIEKVGSGSGKTARPVAIADCGQLS 172
Vicia           ICTAKTDWLDGKHVVFGQVVDGLNVVRDIEKVGSGSGKTSKPVVIANCGQL- 171
Glycine         ICTEKTEWLDGKHVVFGQVIEGLNVVKDIEKVGSGSGRTSKPVVIANCGQPS 172
Brassica        ICTEKTSWLDGKHVVFGQVVEGMDVVRDIEKVGSDSGRTSKKVVTCDCGQL- 168
Oryza           ICTVPCSWLDGKHVVFGRVVEGMDVVKAIEKVGSRGGSTAKPVVIADCGQLS 172
Zea             ICTVATPWLDGKHVVFGQVVEGMDVVKAIEKVGTRNGSTSKVVKVADCGQLS 172
Triticum        ICTVPCNWLDGKHVVFGEVVEGMDVVKNIEKVGSRSGTCSKQVVIADCGQL- 171
                ***    **********.*::*::**: *****: .*  :: *  .:***
```

Figure 2. Alignment of deduced amino acid sequence of *Vigna mungo* cyclophilin gene with other plant cyclophilins: BAB82452.1 from *Vigna radiata*; CAA52414.1 from *Phaseolus vulgaris*; BAA25755.1 from *Vicia faba*; ACU16431.1 from *Glycine max*; AAA62706.1 from *Brassica napus*; AAA57046.1 from *Oryza sativa*; ACF87685.1 from *Zea mays*; ABS85544.1 from *Triticum aestivum*.

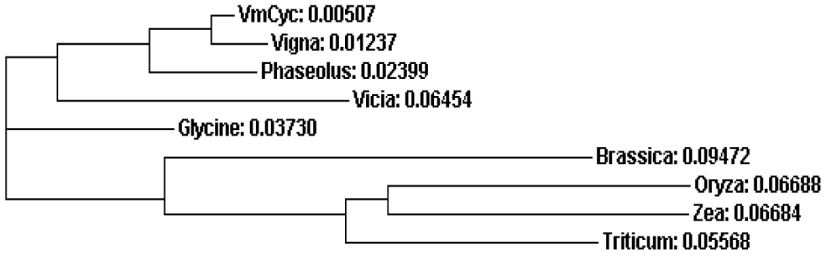

Figure 3. Phylogenetic tree analysis of *Vigna mungo* cyclophilin (VmCyc) with other plant cyclophilins: BAB82452.1 from *Vigna radiata*; CAA52414.1 from *Phaseolus vulgaris*; BAA25755.1 from *Vicia faba*; ACU16431.1 from *Glycine max*; AAA62706.1 from *Brassica napus*; AAA57046.1 from *Oryza sativa*; ACF87685.1 from *Zea mays*; ABS85544.1 from *Triticum aestivum*.

REFERENCES

Brandts JF, Halvorson HR, Brennan M (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry 14:4953-4963.

Chou IT, Gasser CS (1997) Characterization of the cyclophilin gene family of *Arabidopsis thaliana* and phylogenetic analysis of known cyclophilin proteins. Plant. Mol. Biol. 35:873-892.

Galat A (1999) Variations of sequences and amino acid compositions of proteins that sustain their biological functions: an analysis of the cyclophilin family of proteins. Arch. Biochem. Biophys. 371:149-162.

Gasser CS, Gunning DA, Budelier KA, Brown SM (1990) Structure and expression of cytosolic cyclophilin peptidyl-prolyl cis-trans isomerase of higher-plants and production of active tomato cylophilin in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 87:9519-9523.

Godoy AV, Lazzaro AS, Casalongue CA, San Segundo B (2000) Expression of a *Solanum tuberosum* cyclophilin gene is regulated by fungal infection and abiotic stress conditions. Plant Sci. 152:123-134.

Handschumacher RE, Harding MW, Rice J, Drugge RJ, Speicher DW (1984) Cyclophilin: a specific cytosolic binding protein for cyclosporin A. Science 226:544-546.

Kan YC, Liu SW, Guo ZJ, Li DB (2002) Characterization of a cyclophilin cDNA from soybean cells. Acta Bot. Sin. 44:173-176.

Kinoshita T, Shimazaki K (1999) Characterization of cytosolic cyclophilin from guard cells of *Vicia faba* L. Plant Cell Physiol. 40:53-59.

Kumari S, Singh P, Singla-Pareek SL, Pareek A (2009) Heterologous expression of a salinity and developmentally regulated rice cyclophilin gene (OsCyp2) in *E. coli and S. cerevisiae* confers tolerance towards multiple abiotic stresses. Mol. Biotechnol. 42:195-204.

Lippuner V, Chou IT, Scott SV, Ettinger WF, Theg SM, Gasser CS (1994) Cloning and characterisation of chloroplast and cytosolic forms of cyclophilin from *Arabidopsis thaliana*. J. Biol. Chem. 269:7863-7868.

Liu J, Chen CM, Walsh CT (1991) Human and Escherichia coli cyclophilins: Sensitivity to inhibition by the immunosuppressant cyclosporin A correlates with a specific tryptophan residue. Biochemistry 30:2306-2310.

Marivet J, Frendo P, Burkard G (1995) DNA sequence analysis of a cyclophilin gene from maize: developmental expression and regulation by salicylic acid. Mol. Gen. Genet. 2:222-228.

Marivet J, Frendo P, Burkard G (1992) Effects of abiotic stresses on cyclophilin gene expression in maize and bean and sequence analysis of bean cyclophilin cDNA. Plant Sci. 84:171-178.

Menke FLH, Kijne JW, Memelink J (1996) Digging for gene expression levels in *Catharanthus roseus*: nonradioactive detection of plant mRNA levels. In: Leous M, Matter K, Schröder C, Ziebolz B (eds), Biochemica 2, Boehringer-Mannheim, Mannheim, Germany, pp. 16-18.

Que Y, Liu J, Xu L, Guo J, Xu J, Chen R (2011) Molecular cloning and characterization of a cytoplasmic cyclophilin gene in sugarcane. Afr. J. Biotechnol. 10:8213-8222.

Romano PGN, Horton P, Gray JE (2004) The Arabidopsis cyclophilin gene family. Plant Physiol. 134:1268-1282.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, New York.

Sekhar K, Priyanka B, Reddy VD, Rao KV (2010) Isolation and characterization of a pigeonpea cyclophilin (CcCYP) gene, and its over-expression in Arabidopsis confers multiple abiotic stress tolerance. Plant Cell Environ. 33:1324-1338.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Ye XY, Ng TB (2001) Isolation of unguilin, a cyclophilin-like protein with anti-mitogenic, antiviral, and antifungal activities, from black-eyed pea. J. Protein Chem. 20:353-359.