V. DISCUSSION

The present investigation aims to screen the presence of forisome gene in five Fabaceae plants, to isolate and characterize the full length forisome gene sequence. The spatiotemporal expression study of forisome gene was also carried out to analyse the gene expression in leaves, stem, roots and meristem at various stages by quantitative RT-PCR. The expression study of forisome gene had not been studied before and forisome protein being a prime candidate for developing biomotors and nanorobotics, it became important to clearly understand the forisome gene expression. Hence this study was carried out to isolate full length forisome gene from Phaseolus vulgaris and to clone it in E.coli to aid further studies involving forisome protein.

Forisomes are novel, biological, non-living, smart, ATP-independent plant-protein which is capable of sensing and actuating. Forisomes act as a stopcock in the phloem of higher plants and also function as “wound healers” of the plant world. The in vitro studies of forisome proteins show that, by the shift in pH and calcium concentration, these proteins can be repeatedly stimulated to contract and expand anisotropically by 30% strain along their longitudinal axis while radial axis increases by more than 200% strain. Forisomes exert similar mechanical forces in expansion and contraction with a minimum of 0.1 N. According to Mavroidis and Dubey, (2003) these features make forisomes a prime candidate to outperform current smart materials as the building blocks for sensors, mechanoreceptors, nerve fibres, scaffolds for tissue engineering, biomolecular machines and ultraminiature robotic systems like nano robots. Forisomes are ‘smart’, or responsive materials which can be integrated in monitoring devices and can allow conversion of input
signals into appropriate actions. Calvert, (2009) reported that the forisome-based smart materials can be used for health monitoring of structural integrity in civil infrastructure and aerospace hardware. He also claimed that forisomes can also replace materials such as synthetic hydrogels or shape memory in future.

Knoblauch et al. (2003) identified that sieve elements of Fabaceae contain elongate protein bodies called forisomes which is in accordance with Tuteja et al. (2010). Behnke, (1991) reported that forisomes consist of fibrils and which were previously classified as “non-dispersive P-protein bodies”. He also stated these non-dispersive P-protein bodies were believed to undergo structural transformations, from a crystalloidal state with co-aligned fibrils to a “slime-body” with dispersed fibrils which was earlier reported by Palevitz and Newcomb, (1971). According to Peters et al. (2007, 2008) the transition is a rapid and reversible conformational change in forisomes as it shortens longitudinally and expands radially with a several-fold increase in volume. Knoblauch and Van Bel, (1998), reported that forisomes disperse upon wounding and the same was confirmed again by Knoblauch et al. (2012). Tuteja et al. (2010) mentioned that scanning electron microscopic images illustrated the composition of forisomes to be large with identical subunits named as forisomettes which confirms Noll, (2005) observation. Transmission electron microscopy studies indicate that forisomettes consist of strictly ordered arrays of forisome proteins. Noll et al. (2007) reported existence of three proteins in formation of Medicago truncatula forisomes (forisomettes) and named them as sieve element occlusion 1 (SE01), SE02, and SE03. Rüping et al. (2010) has first reported the presence of SEOs in plant families that do not possess forisomes, e.g., rosaceae, solanaceae, and brassicaceae which was in agreement with Ernst et al. (2011).
In India Tamil nadu has an average temperature of 30°C as cited by Aiyappan (2013). Plants that grow at low temperatures can only be grown under controlled conditions. Most plants of Fabaceae family show good growth in lower temperature of around 20-25°C and this is in accordance with Hall (1994). They grow best at 16 - 18°C which was also reported by Singh and Jauhar (2005). In the present study plants were chosen in Fabaceae family which can grow well at 25-30°C. The five plants chosen that can grow well at 30 °C were Vigna mungo (L.), Cicer arietinum (L.), Vigna unguiculata (L.), Phaseolus vulgaris (L.) and Vigna angularis (L.). Baligar et al. (2007) reported in his publication that Vigna mungo and Cicer arietinum (L.) can sustain temperatures of around 30°C which. Thomas Jefferson Agricultural Institute (2010) has published a report stating that Vigna unguiculata can resist high temperature of 30- 35°C which was also reported in this present investigation. Ahmed et al. (1992) had checked the growth parameters of Phaseolus vulgaris and Vigna angularis (L.) at 30°C and observed considerable growth which is in accordance with our results. The above mentioned growth conditions were in agreement with our growth parameters as we have grown all the Fabaceae plants at around 30°C.

After cultivation the presence of forisome gene in plants was screened. To identify and confirm the presence of forisome gene the total RNA was isolated and cDNA was synthesised from the plants. Fontanellaz, (2006) and Rüping et al. (2010) had also screened for forisome gene in Vicia faba and Glycine max. respectively. Total RNA was isolated from the plants using Acid-Guanidinium-Phenol-Chloroform Method as described by Chomczynski & Sacchi, (2006). RNA extraction performed by Hu et al. (2002), Iandolino et al. (2004), Tattersall et al.
(2005) and Vasanthaiah et al. (2006) by using acid-guanidinium-phenol- chloroform method had also yielded sufficient amount of good quality RNA from different plant tissues. Inclusion of an additional re-extraction step with phenol and chloroform helped to remove contaminants and recover high quantity of RNA. Logemann et al. (1987) observed that inefficient removal of polysaccharides and polyphenolic compounds results in co-precipitation with RNA, which affects the yield and quality of the RNA. The purity of the isolated RNA was analysed by measuring the absorbance of a diluted RNA sample at 260 and 280 nm. Acercbi et al. (2009) stated that good-quality RNA should have an $A_{260}/A_{280}$ ratio in the 1.7–2.0 range. This was in accordance with the reports of this investigation. The $A_{260}/A_{280}$ ratio of all the isolated RNAs used in this study was between 1.7 - 2.0. The integrity of the isolated RNA was verified by electrophoresis on a formaldehyde agarose gel.

Rogozin et al. (2012) reported that eukaryotic genes contains introns (intervening sequences), which are not coding sequences (exons), and makes the genomic DNA sequence bulky. So complementary DNA (cDNA) sequence of RNA has to be constructed which does not include portions of introns. This investigation is in agreement with the above statement so cDNA was synthesised from total RNA. Reverse transcriptase (RT)-PCR was carried out to generate cDNA using gene specific degenerate primers. A PCR primer sequence is called degenerate if one or more nucleotide positions can be occupied by one of several possible nucleotides. Bossolini and Keller, (2006) suggested that degenerate primers are useful for amplifying homologous genes from different and highly related organisms which was later reported by Zuiter et al. (2012). Homologous genes display regions of highly conserved and also regions where they have evolved and are divergent. In
this study the forisome gene has shown highly conserved regions between the related species, hence the fore mentioned statement is in accordance with the present investigation. Primers were designed based on the highly conserved regions. Thus, degenerate primers were used to screen for forisome genes encoding forisome proteins that belonged to Fabaceae family. Previously published forisome gene sequence collected from National Center for Biotechnology Information (NCBI) were Canavalia gladiata (Noll et al., 2007), Vicia faba (Noll et al., 2007), Medicago truncatula (Noll et al., 2007), Pisum sativum (Subramanian and Tuteja, 2008), Glycine max (Ruping et al., 2010). The above mentioned gene sequences were aligned and degenerate primers were designed based on the conserved sequences.

Four different degenerate forward and reverse primers were designed based on the conserved sequences of the existing forisome genes. Amplification of cDNA with these designed primers gave different multiple bands. Fori_F2 primer gave distinct single band of ~1700 bp in all the five plants hence it was chosen for further studies. Noll et al. (2007) has reported the size of forisome gene in V. faba and Canavalia gladiata as 2055 bp and 2007 bp respectively. The PCR amplified gel bands were then eluted from the agarose gel and sequenced to confirm the presence of forisome gene in the plants. Among the five plants screened, the sequences obtained from Phaseolus vulgaris was found to be best based on the gene coverage and gene sequence quality. Simultaneously the total chromosomal DNA was isolated from the plants for molecular identification. Hollingsworth et al. (2009); Hollingsworth (2011) had also reported the importance of DNA barcoding in identification of species of plants.
Barcoding is a method of identifying species using short DNA sequences. As reported by Chen et al. (2010) the definitive goal is to identify a region or a combination of regions capable of discriminating all species. The barcoding method has been extremely useful in species identification, cryptic species identification, biodiversity studies, forensic analysis, and phylogenetics and it is in agreement with Ronning et al. (2005) and Ward et al. (2005). Kress et al. (2005) reported that in plants, the most favorable choices are chloroplast DNA regions, as they have been used as a means to identify species. The sequences used, known as DNA barcodes, are usually short (300-800 bp). Although chloroplast DNA barcoding is mainly used to identify plant species, its application could be extended to the food industry, evolution studies, and forensics as stated by Ferri et al. (2009). Research groups have their preferred plant barcode regions for study, but no consensus has emerged on the use of a standard region as mentioned by Yao et al. (2010). Wojciechowski et al. (2004) have used the matK region to construct the phylogeny of legumes. Asmussen and Liston (1998) have used rpoC1 and psbA-trnH to construct the phylogeny of the genus Lathyrus. Gao et al. (2011) have also used matK to barcode members of the Fabaceae family. Since Phaseolus vulgaris was chosen for the further study DNA barcoding (Nicole et al., 2011) was performed using matK gene (Yu et al., 2011), rpoC1 gene (Shaw et al., 2007) and trnH gene (Maskova et al., 2012) for identifying and confirming the plant species (Phaseolus vulgaris).

The forisome protein from Phaseolus vulgaris was isolated using sucrose gradient centrifugation. Sucrose gradient centrifugation is often used to purify enveloped viruses (with densities 1.1-1.2 g/cm³), ribosomes, membranes and exosomes as stated by Neil and Michelle (2008). Forisome has an equilibrium
density of 1.12 g/cm³ in sucrose and was resolved on a 20–80% wt/vol sucrose
density gradient. Forisome separation was carried out in sucrose gradient instead of
Nycodenz gradient as used by the inventor of the forisome protein Knoblauch
(2003). Both Nycodenz and sucrose solution serve the same purpose of forming a
gradient but Nycodenz was expensive. The distribution of the constituents after
ultracentrifugation in sucrose density gradient agrees with what has been found
previously by means of sucrose-density-gradient centrifugation of Pranav et al.
(2013) and Yang et al. (2014). The centrifugal fraction from the density gradient
fraction of 1.2 g/cm³ of sucrose contained the forisome protein was isolated with the
help of a micropipette.

The purified forisomes from Phaseolus vulgaris were separated on SDS -
PAGE, which was performed according to Laemmli U.K. (Laemmli, 1970) using a
4% stacking gel and a 10% separation gel. It is an improved method of gel
electrophoresis used for studying unknown proteins found in bacteriophage T4 and
have been identified with specific gene products. This protocol has been widely used
for separating proteins at good resolution. The forsiome protein bands were expected
to be in the range of 70-80 kDa (Nol, 2005 and Fontanellaz 2006) and isolated
forisomes from Phaseolus vulgaris was observed to have molecular weight of 75
kDa. The present study provides information that tailed and tailless forisomes can be
isolated and studied by similar methods as highlighted by Peters et al. (2007).

The spatial and temporal expression of forisome gene in Phaseolus vulgaris
was studied by isolating the total RNA from different developmental stages of
Phaseolus vulgaris and synthesising the cDNA strands using gene specific primers.
Kenneth and Thomas, (2001) has highly recommended that the internal control gene
should be validated for the experiment. The expression of internal control genes vis., GAPDH, ACTIN, EF1A and EIF4E at various parts of the plants was studied using qRT-PCR. Alexandre et al. (2011) has reported GADPH gene to be the most variably expressed gene. In our investigation GAPDH gene has not been found to be highly expressed gene nor was it highly stable in expression. Hence this study is in accordance with the fore stated report. ACTIN gene was reported to be unsuitable by Selvey et al. (2001) as its expression is affected by external factors like matrigel. This present study found the expression of actin gene to be unsatisfactory. Zhu et al. (2012) has reported that EIF4E gene expression is not stable and the present study is in accordance to this statement. Xiang et al. (2011) reported that the EF1A gene was found to be stable in expression in wheat plants and our study confirmed his report. The expression of forisome gene at various stages of the plants and various parts of the plants was studied using qRT-PCR. This is the first report of forisome gene expression in Phaseolus vulgaris.

In the present study the expression of forisome gene in the phloem cells of Phaseolus vulgaris of leaves, stems and roots were extensively studied which indicated highly tissue-specific regulation. Studies were carried out to determine the developmental regulation of expression pattern studies on young seedlings of the plant at different stages of growth.

In Phaseolus vulgaris, strong expression of forisome gene in developing phloem tissue and, especially, in the meristematic and elongation zones at root tips of germinating seedlings was detected. During root elongation, this expression pattern was restricted to the meristematic and elongation zones at root tips. In the tips of fully developed roots, forisome expression was no longer detectable due to
the cession of developing phloem synthesis. The most apical 2mm of the root tip, which showed strong β-glucuronidase activity, is a region of the root that includes the root meristem, the elongation zone, and the differentiation zone where lateral root formation is initiated (Schiefelbein and Benfey, 1994; Dolan and Davies, 2004). The primary root meristem has many cells undergoing rapid cell division (Beeckman et al. 2001; Birnbaum et al. 2003). It has also been observed (Van der Weele et al. 2003) that the root tip contains two distinct zones with constant elongation rates, lower in the meristematic zone (0 to 0.5 mm from the tip) and much higher in the elongation zone (0.5 to 1.2 mm from the tip). In the differentiation zone, the growth rate decreases to zero. Some roots are genetically determined to have the capacity to undergo additional radial growth, but other roots do not. Only dicotyledons have this capacity and it is more likely to occur in thicker, lower-order lateral roots. This radial growth is called secondary growth to distinguish it from primary growth at the root apex. Secondary growth occurs when new meristematic tissue forms in a ring around the vascular cylinder of roots and produces new xylem inwards and new phloem outwards (Dean, et al., 2004). This would explain the forisome expression observed in growing lateral root. During the early stages of lateral root formation, forisome expression seemed to be restricted to the developing phloem tissue. These results provide convincing evidence that indicate not only a spatial but also a temporal regulation of the expression pattern in roots of Phaseolus vulgaris plants.

A considerable expression of forisome gene was detected in the shoot apical meristem, roots, stem of young phloem tissue of Phaseolus vulgaris plants. In addition, a particular expression pattern of forisome gene was observed based on the
veins of *Phaseolus vulgaris* leaves. The leaf vein classification of dicotyledonous plants is based largely on anatomical observations as described by Avery, (1933). In anatomical terms, minor veins do not have rib tissues that protrude beneath the surface of the lamina (Esau, 1965). However, from a physiological and developmental perspective, minor veins are those that are immature in sink leaves and do not participate in phloem unloading but mature during the sink-source transition. The anatomical, developmental, and functional roles of leaf venation have been well studied in tobacco. The veins of tobacco leaves have been subdivided into classes, based on cell numbers rather than branching pattern, which can be misleading as stated by Baker et al. (2013).

Forisome gene expression in leaves was found to be comparatively less. According to Wright et al. (2003), it seems likely that structural vein maturation is developmentally programmed. Taken together, the overall temporal and tissue-specific regulation of the forisome gene is indicative of the expression of forisomes mainly in the first stage of phloem development. The biological function of these proteins in the sieve elements of *V. faba* is related to as a part of a defence mechanism against the loss of assimilates upon wounding (Knoblauch and Peters, 2004).

In phloem tissue, it has been suggested by Sophie et al. (2001) that the earliest differentiation takes place in the external phloem at the tip of the midrib and that maturation progresses basipetally. However, other reports claim that maturation of the midrib and major veins occurs acropetally. Although the present work does not provide any information regarding the structural maturation of the main stem, it does demonstrate a temporal forisome gene expression pattern. Taken together, this
can be correlated with the presence of forisome gene expression in seedlings of Phaseolus vulgaris with cotyledon and the cessation of forisome gene expression in mature stems suggests that activation of an accurate marker which turns off the expression in structurally mature phloems as stated by Brandy et al. (2009). Emst et al. (2012) confirmed in his report that expression of SEO genes are restricted to immature sieve elements. However, further experiments aiming to identify the factors controlling forisome gene expression will be required to clarify the control points in the expression pattern in developing stems.

Full length cDNA of the forisome gene from Phaseolus vulgaris was obtained using Random Amplification of cDNA Ends (RACE). The full length cDNA obtained was sequenced using Sanger Coulson di-deoxynucleotide sequencing method. The obtained full length sequence was submitted in NCBI and accession number was obtained (KF924554).

The alignment of the forisome gene against the completely sequenced genome of Arabidopsis thaliana and Oryza sativa could not retrieve any information. The identification of significantly homologous genes and/or proteins could not be achieved from A. thaliana. Blast search was performed in sequence databases which showed closest homology 89 % to Glycine max.(SEO-F4, 2001 bp) which is approximately of similar length and 72 % with Cicer arrietinum uncharacterised gene which shares 18 % similar length.

Using the forisome gene sequence, the amino acid sequence was deduced using ExPASy tool (http://web.expasy.org/translate/).
Conventional Ca\(^{2+}\) - binding motifs could not be predicted for forisome protein, which renders the Ca\(^{2+}\) sensing to be unique. This finding is in accordance with Tuteja, 2009. There are a number of speculations as to the apparent absence of Ca\(^{2+}\) binding sites: (i) essential forisome proteins including one containing a Ca\(^{2+}\)-binding site still await detection; (ii) the Ca\(^{2+}\)- binding site of forisome proteins is not yet recognized; (iii) forisome proteins provide the matrix for a Donnan phase in which Ca\(^{2+}\) is accumulated; or (iv) Ca\(^{2+}\) binding sites are formed by residues of more than one protein and thus remain invisible in searches for known Ca\(^{2+}\)binding motifs (Pelissier, 2008).

Another Ca\(^{2+}\) - sensitive protein aggregate, the spasmoneme from Vorticella sp. is known to behave in a similar manner (Moriyama et al. 1998; Moriyama et al. 1999). This structure contracts strongly on exposure to Ca\(^{2+}\) and swells when the Ca\(^{2+}\) ions are sequestered by a chelating agent such as EDTA. However, there appears to be no obvious homology between the forisome P protein (Noll, 2006) and the known spasmin proteins which form spasmoneme (Maciejewski, 1999). There also appears to be no homology between forisome proteins other EF- hand Ca\(^{2+}\) binding proteins such as calmodulin and troponin C.

Computer analysis (http://www.ch.embnet.org/software/COILS_form.html; Lupas, 1996) revealed another structural motif, an amphipathic \(\alpha\)-helix within the coiled-coil domain of forisome amino acid sequence. The coiled-coil domain is responsible for oligomerization of protein subunits and results in folding of the proteins. It has been described for the first time in the intermediate filament keratin protein by Burkhard et al. (2001). Later, a dimerization domain of a family of transcription factors, the leucine zipper was also described comprising this structural
motif. Some amphipatic helices are arranged as interwined helices that are termed as coiled coils or super-helices. Although structurally simple, coiled coils have the ability to form a variety of different assemblies ranging from dimers to pentamers. Such coils were mentioned by Walshaw and Woolfson, (2001); Lupas and Gruber, (2005) in their reports. Furthermore, coiled coils can form homomers or heteromers with their chains arranged in a parallel or antiparallel fashion (Lupas, 1992). Sequences of parallel left-handed coiled coils are characterized by a seven-residue periodicity (heptad repeat), with the occurrence of apolar residues preferentially in the first (a) and fourth (d) positions of the ‘heptad’ as described by Lupas, (1996). The interaction between these apolar amino acids denote the formation of the hydrophobic core, which induces coiled coils. Charged amino acid residues, frequently at positions outside the hydrophobic core, can affect upon changes of the pH in the surrounding medium, after phosphorylation of the AA side chains and after interaction with ions, the hydrophobic interaction of the coiled-coil domains. It is also generally acknowledged that the detailed packing geometry of hydrophobic core residues correlates with the oligomerization state.

The 28 AA coiled-coil domain of the forisome protein is relatively short, however, this has been already described for some other globular proteins with less than 20 AA by Rieker and Hu, 2000; Eckert and Kim, 2001; Ryadnov et al., 2003 and Cho et al., 2004. Moreover and in contrast to the characterized apolar amino acid residues found in the first (a) and fourth (d) positions of the heptad pattern, in the forisome coiled-coil domain a polar amino acid residue is present within one of the heptads. In this context, it has been shown that distinct coiled-coil ‘trigger sites’ within heptad-repeat-containing amino acid sequences might be necessary to
mediate coiled coil formation (Frank et al., 2000). As a hallmark, the coiled-coil trigger site of the actin bundling protein cortexillin I contains a distinct inter- and intrahelical salt-bridge pattern, which includes charged residues even at core positions was reported by Burkhard. (2000a). Other recent results suggest that ionic interactions are not only important for proper alignment, orientation and selectivity of coiled coils, as shown previously but can also contribute considerably to stability of coiled coils as mentioned by Burkhard et al. (2000b) thereby modulating assembly of coiled coils in a pH-dependent manner.

Further analysis of the forisome amino acid sequence was performed, where no well known motifs appeared to contribute to the secondary structure of the protein. A low-complexity region was identified for which no structure can be predicted with the algorithms available (Schultz et al., 1998). One general idea about the functional relevance of these low-complexity regions is that they only acquire a rigid structure (e.g., an α-helix) upon interaction with another protein partner or oligomerization into a protein complex (Wright and Dyson, 1999). This has been reported for soluble Nethylmaleimide- sensitive factor attachment protein receptor, which play a role in vesicle docking and fusion, and which only form an α-helix when they bind to their interaction partner (Fasshauer et al., 1997; Jahn and Sudhof, 1999).

In this study several cysteine residues could also be identified at the carboxy-terminal domain. The reversible conformational change of forisomes from the condensed into the dispersed state after addition of Ca$^{2+}$ leads not to a complete dissolution of the protein complexes. Forisomes still remains as distinctive bodies (Knoblauch et al., 2003). A function of the cysteine residues with the stabilization of the protein complex by the formation of disulfide bridges of the side chains would
be conceivable. Possibly these covalent bonds serve as the linkage of the individual subunits holding them together and facilitating, after Ca$^{2+}$ withdrawal, the back folding of the forisome into the arranged condensed conformation. From the characterization of the different motifs identified within the forisome protein sequence, it is becoming clear that this protein comprise several structural features related to helices and coiled-coil motive formation. These structural characteristics might indicate a possible Ca$^{2+}$ dependent reaction of the forisomes. However, future work focusing on the identification of additional forisome genes and their molecular characterization are needed for the elucidation of mechanisms that influence their conformational change.

The amino acid sequences of forisomes proteins from Cannavalia gladiata, Medicago truncatula, Vicia faba and Pisum sativum were downloaded from NCBI web server (www.ncbi.nlm.nih.gov). Conserved portion of amino acids were aligned with ClustalX software developed by Thompson, (1997). The studies on TryX - NRX domain of the tryparedoxin (TryX) and nucleoredoxin (NRX) subfamily domains were performed using NCBI server (http://www.ncbi.nlm.nih.gov). The distribution of TryX-like domains across the major taxonomic kingdoms was studied using the Superfamily 1.73 (supfam.mrc-lmb.cam.ac.uk/SUPERFAMILY). The content of secondary structure of Phaseolus vulgaris forisome protein was studied for the modelled stretch (amino acids 329 to 447) using the PHYRE web server (www.sbg.bio.ic.ac.uk/phyre) as reported by Kelley, (2009).

The complex 3D structure of Phaseolus vulgaris forisome protein confers a multitude of electro- and hydrostatic interactions within the protein and its hydration shell. The Ca$^{2+}$ induced changes in surface charge and electrical density may allow
massive and sudden binding of water molecules and induce conformation shifts. The structural coherence within forisomes is mysterious. The template on which Phaseolus vulgaris forisome protein was modeled is TryX protein, a member of the TRX-fold family from Trypanosoma brucei based on report of Alpey, (2003). The functional significance of this TryX-NRX subfamily of disulphide oxidoreductases is unclear, particularly in plants. The conserved domain of thioredoxin fold is present in SEO2 (amino acids 307 to 450) and SEO3 (amino acids 323 to 459) of Medicago truncatula, and For1 (353 to 440) of Cannavalia gladiata and Ps. SEO1 (amino acids 320 to 460) of Pisum sativum. TryX and NRX are thioredoxin-like protein disulfide oxidoreductases alter the redox state of target proteins which is present in multitude in forisome.

Both vertebrate and plant NRXs show thiol oxidoreductase activity in vitro. Their localization in the nucleus suggests a role in the redox regulation of nuclear proteins such as transcription factors. The TRX-like domain is widely distributed among major taxonomic kingdoms. Secondary Structural Content Prediction (SSCP) tool (http://coot.embl.de/SSCP/) was used to predict secondary structural content. Amino acid composition of forisome protein indicates that α, β and coiled-coil configurations are present at approximately equal levels.

alpha-contents : 35.1 %
beta-contents : 35.3 %
coil-contents : 29.6 %

Noll (2005) hypothesized 15 coiled-coil structures presumably present in Mt.SEO2 might play a role in forisome function whereas Fontanellaz (2006) discussed the potential of Vf.SEO1 to form coiled-coil structures.
SEO genes encode subunits of forisomes as stated by Peters et al. (2010). SEO proteins contain three conserved domains: the SEO N-terminal domain (SEO-NTD), a potential thioredoxin fold, and the SEO C-terminal domain (SEO-CTD). The presence of all the three domains has been found in the Forisome gene of Phaseolus vulgaris which is in accordance with Ernst et al. 2012. Sieve element occlusion, N-terminal domain stretches from amino acid 3 to 274, Thioredoxin fold domain stretches from amino acids 329 to 447 and Sieve element occlusion, C-terminal domain stretches from amino acid 442 to 659.

Rüping et al. (2010) described the SEO genes and stated that these SEO genes also encode for the conventional P-proteins like forisome proteins in M. truncatula (SEO-F genes). The sequence similarity indicates that forisomes and conventional P-proteins have a common evolutionary origin. This notion is supported by several shared structural characteristics, such as their fibrillar composition, their ultrastructure in the condensed and dispersed states, their spatiotemporal expression profiles, and their cellular functions. Forisomes are unique in sealing of sieve plates which is reversible as stated by Knoblauch (2001).

Ruping et al. (2010) has identified a number of conserved motifs of unknown function which is in right conformity with this study. A conserved domain has been found in Phaseolus vulgaris Forisome from amino acid 322 - 343 whose identity is not yet known.

A protein is made up of four different structures, it's primary, secondary, tertiary and quaternary structures. Depending on the sequence, it will form either β sheets, α helices or other complex structures. The sequence determines its shape because of the amino acids forms bonds that hold the protein together. The shape is
important because proteins are fundamental in a living organism. Thus, Structure prediction (both secondary and tertiary) has become an indispensable tool for investigating function of proteins and mechanisms of biological processes. Structure of the protein is partially dictated by the properties of the peptide bond. The peptide bond is resonance hybrid of two inter convertible structures. The resonance causes the peptide bonds to be either less reactive like esters or be quite rigid and nearly planar or exhibits large dipole moment. Ramachandran Plot was also drawn. The torsion (phi-psi) angles from a Ramachandran plot are useful not only for determining amino acids role in secondary structure but can also be used to verify the solution to a crystal structure. Furthermore, it also assists in defining energy functions. This also helps in predicting 3 - D structure of the forisome protein.

The isoelectric point (pI) of the forisome protein was calculated using the amino acid sequence (http://web.expasy.org/compute_pi/). The pI of the forisome was found to be 5.9.

Lawton (1970) had observed the ultrastructure of tailed P-protein crystals (forisome) as mentioned in his paper where he had discussed about the difference in diameter of P-protein tubules in pea and bean in relation to the periodicities of striation in their respective crystals. Applying the criteria that the molecular nature of structural P-proteins has remained poorly understood we carried out a comprehensive characterization of the forisome genes in Phaseolus vulgaris.

Many unsuccessful attempts were made before actually Forisome encoding genes could be identified as it depends on the quality of RNA purification and differentially expressed mRNAs. Previously published SEO sequences as BLAST queries were used to search for homologous sequences of forisome gene.
Degenerate primers were designed from highly conserved regions to amplify the cDNA sequence with template of Phaseolus vulgaris total RNA by RT-PCR. The whole cDNA was cloned, including 5' and 3' UTR fraction of the forisome gene with 2188 bp which was identified for the first time in Phaseolus vulgaris.

The recombinant vector was introduced in E.coli DH5α by transformation. The recombinants were screened by plating it on media with (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) (Xgal), isopropyl-beta-D-thiogalactopyranoside (IPTG) and ampicillin based on Lac Z gene and plasmid-borne ampicillin-resistance gene.

This is the first study in Phaseolus vulgaris to report on the full length forisome gene isolation and cloning in E. coli to aid further studies involving forisome protein. The expression study of forisome gene is also the first report where the spatiotemporal expression study of forisome gene has been carried out to analyse the gene expression in leaves, stem, roots and meristem at various stages by quantitative RT-PCR. This research paves a way for future research work using forisome proteins in developing microfluidic lab on chip systems, artificial muscle, microgrippers, scaffolds in tissue engineering and vectors in drug delivery system.