(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 July 2013 (04.07.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/098852 A2

(51) **International Patent Classification:**
C07K 14/42 (2006.01)

(21) **International Application Number:**
PCT/IN20 12/000822

(22) **International Filing Date:**
17 December 2012 (17. 12.2012)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
3850/DEL/201 1 28 December 201 1 (28. 12.201 1) IN

(71) **Applicant: COUNCIL OF SCIENTIFIC & INDUSTRI¬AL RESEARCH** [IN/IN]; Anusandhan Bhawan, Rafi Marg, New Delhi 110 001 (IN).

(72) **Inventors: SINGH, Pradhyumna, Kumar;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **RAI, Preeti;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **SINGH, Rahul;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **UPADHYAY, Santosh, Kumar;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **SAURABH, Sharad;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **SINGH, Harpal;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **VERMA, Praveen, Chandra;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **KRISHNAPPA, Chandrashekar;** National Botanical Research Institute, PO Box No 436, Rana Pratap Marg, Lucknow 226 001 (IN). **TULI, Rakesh;** National Agri-food Biotechnology Institute, C-127, Industrial Area, S.A.S. Nagar,Phase 8, Mohali 160 071 (IN).

(74) **Agents: LAKSHMIKUMARAN, Varadachari** et al; B-6/10, Safdarjung Enclave, New Delhi 110029 (IN).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/098852 A2

(54) Title: *ALLIUM FISTULOSUM* LEAF AGGLUTININ RCOMBINANT PROTEIN, ITS ENCODING POLYNUCLEOTIDE, PRIMER AND PROCESS FOR PREPARATION THEREOF

(57) Abstract: Nucleic acid sequence encoding *Allium fistulosum* leaf agglutinin (AFAL) is disclosed. The invention provides *Allium fistulosum* leaf agglutinin (AFAL) recombinant protein, its encoding nucleotides, primers and the process of preparation thereof, said recombinant protein is useful for insect control and haemagglutination activity. AFAL is found more toxic to sap sucking insect pest *Aphis gossypii* (cotton aphid) and *Bemisia tabaci* (whiteflies) as compared to known *Allium sativum* leaf agglutinin. AFAL can be used in the development of transgenic plants for resistance against sap sucking and chewing pests.

ALLIUM FISTULOSUM LEAF AGGLUTININ RCOMBINANT PROTEIN, ITS ENCODING POLYNUCLEOTIDE, PRIMER AND PROCESS FOR PREPARATION THEREOF

FIELD OF THE INVENTION

The invention relates to *Allium flstulosum* leaf agglutinin (AFAL) recombinant protein, its encoding polynucleotide, primers and process of preparation thereof said *Allium flstulosum* leaf agglutinin (AFAL) protein is useful for insecticidal activity and haemagglutination activity. In particular the present invention relates to nucleic acid sequence *(afal)* encoding for *Allium flstulosum* leaf agglutinin (AFAL) applicable for haemagglutination and insect control.

BACKGROUND AND PRIOR ART OF THE INVENTION

Plant lectins, also known as "agglutinins", are heterogeneous group of carbohydrate binding proteins, which are able to bind simple sugars and/or complex carbohydrates reversibly (Van Damme et al. 1998; *CRC Crit Rev Plant Sci* 17:575-692). They show a marked heterogeneity with respect to their molecular structures, sugar binding specificity and temporal and spatial regulation. Mannose binding lectins are widely found in higher plants and play a significant role in defense due to their ability to recognize high-Mannose-type glycans of microbial pathogens and plant predators (Van Damme et al. 1998, 1998, *CRC Crit Rev Plant Sci* 17:575-692; Van Damme et al.2004 *Trends Plant Sci* 9:484-9). Mannose binding lectin from garlic leaf *[Allium sativum* leaf agglutinin (ASAL)] is a 25 kDa homodimeric protein, structurally and evolutionarily related to *Galanthus nivalis* agglutinin (GNA) (Smeets et al. 1997a, *Plant Mol Biol* 35: 531-535; Van Damme et al. 1992, . *Eur J Biochem* 206:413-420). Some of the biological properties of ASAL are - (1) it readily agglutinates rabbit erythrocytes but does not affect human erythrocytes (Bandhopadhyay *et. al.* 2001, *Plant Sci.* 161:1025-1033; Smeets et al. 1997a, *Plant Mol Biol* 35: 531-535 ;), (2) it has inhibitory effect against retrovirus (HIVl and HIV2) induced cytopathicity in MT-4 cells (Smeets *et. al.* 1997b, *Plant Mol Biol* 33:223-234) and (3) it is toxic (growth inhibitory) against a

spectrum of insects of order Homoptera, Lepidoptera and Coleoptera. Some important pests inhibited by ASAL are aphids [mustard aphid, peach potato aphid, tobacco aphid (Bandhopadhyay *et al.* 2001, *Plant Mol Biol* 33:223-234 Hossain *et al.* 2006, *Crop Sci* 46:2022-2032; Smeets *et al.* 1997a, *Plant Mol Biol* 35: 531-535), red cotton bug (Bandhopadhyay *et al.* 2001, *Plant. Mol. Biol.* 33:223-234), brown plant hopper, green leaf hopper (Saha *et al.* 2006, *Plant Mol. Biol.* 62:735-52) and cotton leaf worm (Sadeghi *et al.* 2008, *Transgenic Res.* 17:9-18). Although exact mechanism of insecticidal action of lectins is still not well understood, three different modes of action have been proposed- (1) binding of the lectins to the peritrophic matrix of the midgut, inhibiting nutrient absorption (Harper et al 1998, *Tissue Cell* 30: 166-176), (2) binding to glycoproteins on epithelial cells of the midgut and disrupting tissue integrity (Powell et al. 1998, *J Insect Physiol,* 44: 529-539; Sauvion et al. 2004, *J Insect Physiol.* 50: 1137-1 150) and (3) binding to carbohydrate moieties of the sensory receptors of insect mouth parts, disrupting membrane integrity and interfering in the food detection ability of insects (Murdock et al. 2002, *J Agric Food Chem.* 50: 6605-661 1). All these mechanisms result in decreased ability of insect to ingest food or absorb nutrients, leading to delayed development and premature death.

United States Patent 5,545,820 (Gatehouse, et al., 1996) discloses the use of lectins having specific mannose-binding ability, derived from family Amaryllidaceae or Alliaceae for the control of insect pests. WO/1992/002139 relates to the use of lectins having specific mannose-binding ability, derived from family Amaryllidaceae or Alliaceae for the control of insect pests and the development of transgenic plants expressing such lectins. US patent 5407454 relates to selected plant lectins having larvicidal activity against a number of common insect pests of agricultural crops. Insect resistance in the transgenic plants is due to insertion of larvicidal lectin gene in all the cells of the plants.

US patent 6,127,532 (Raikhel) refers to transgenic plants containing cDNA encoding Gramineae lectin. Such transgenic plants expressed barley lectin and stored in in the

leaves. The transgenic plants, particularly the leaves exhibit insecticidal and fungicidal properties.

Alliumfistulosum

Allium fistulosum L. (Welsh onion, Japanese bunching onion) is a perennial onion. Other names that may apply to this plant include green onion, spring onion, escallion, and salad onion. These names are ambiguous, as they may also be used to refer to any young green onion stalk, whether grown from Welsh onions, common bulb onions, or other similar members of the genus *Allium.* The species is very similar in taste and odor to the related bulb onion, *Allium cepa,* and hybrids between the two (tree onions). The Welsh onion, however, does not develop bulbs, and possesses hollow leaves ("fistulosum" means "hollow") and scapes. Large varieties of the Welsh onion resemble the leek, such as the Japanese 'negi', whilst smaller varieties resemble chives. Many Welsh onions can be multiplied by forming perennial evergreen clumps. Next to culinary use, it is also grown as an ornamental plant. Historically, the Welsh onion was known as the *cibol.*

The name "Welsh onion" has become a misnomer in modern English, as *Allium fistulosum* is not indigenous to Wales. "Welsh" preserves the original meaning of the Old English word "welisc", or Old German "welsche", meaning "foreign" (compare *wal-* in "walnut", of the same etymological origin). The species originated in Asia, possibly Siberia or China.

Culinary use

In the West, the Welsh onion is primarily used as a scallion or salad onion, but is widely used in other parts of the world, particularly East Asia.

Russia: Welsh onion is used in Russia in the spring for adding green leaves to salads.

Asia: The Welsh onion is an ingredient in Asian cuisine, especially in East and Southeast Asia. It is particularly important in China, Japan, and Korea, hence the other English name for this plant, 'Japanese bunching onion'. Bulb onions were introduced to

East Asia in the 19th century, but *A. fistulosum* remains more popular and widespread. In Japan, it is used in miso soup, negimaki (beef and scallion rolls), among others, and it is widely sliced up and used as a garnish on teriyaki or takoyaki.

Jamaica: Known as escallion, the Welsh onion is an ingredient in Jamaican cuisine, in combination with thyme, scotch bonnet pepper, garlic and allspice (called pimenta). Recipes with escallion sometimes suggest leek as a substitute in salads. Jamaican dried spice mixtures using escallion are available commercially. The Jamaican name is probably a variant of scallion, the term used loosely for the spring onion and various other plants in the genus *Allium.*

Lacking in the prior art

Allium lectin disclosed in present invention shows high insecticidal activity and therefore novel. Homology of the nucleotide sequence with available nucleotide sequences in database shows more than 90% homology.

OBJECTIVES OF THE INVENTION

a) The main objective of the invention is to provide nucleic acid sequence which encodes for *Allium fistulosum* leaf agglutinin

b) Another objective of the present invention is to provide gene specific premiers GSP1 (5'-ATGGACAGTACTCCATCTCCTAAAC-3') and GSP2 (5'-TTAGCCCCTTGGCCTCCTGCA-3 '), useful for amplification of the gene.

c) Another objective of the present invention is to provide agglutinin recombinant protein having high insecticidal activity

d) Another objective of the present invention is the application of *Allium fistulosum* leaf agglutinin protein for insect control.

SUMMARY OF THE INVENTION

Accordingly, the present invention relates to *Allium fistulosum* leaf agglutinin (AFAL) recombinant protein, its encoding nucleotides, primers and the process of preparation thereof, said protein is useful for insect control and haemagglutination activity.

In an embodiment of present invention, the process of preparation of *Allium fistulosum* leaf agglutinin (afal) recombinant protein, useful for insecticidal activity and haemagglutination activity, comprising of following steps-

a) Extracting of total RNA from the *Allium fistulosum* leaves

b) Synthesizing of cDNA from total RNA extracted from leaves of *Allium fistulosum*

c) Designing of primers GSP1 and GSP2 from 5'- and 3'-RACE fragment of c DNA to clone full-length protein encoding DNA, designing of primers GSP3 and GSP4 for cloning of the DNA encoding mature polypeptide of AFAL protein,

d) Expressing the DNA encoding the mature AFAL in *E. coli* SUMO expression vector where SUMO peptide is fused with AFAL at the N-terminus and expressed in *E. coli* under T7 promoter to get desired AFAL recombinant protein.

In another embodiment of the present invention the primers GSP1 and GSP2 comprising

1. GSP1 represented by sequence I.D:5
2. GSP2 represented by sequence I.D:6

Allium fistulosum leaf agglutinin (AFAL) recombinant protein is 8- 10 fold more toxic to cotton aphid *{Aphis gossypii)* and 6-8 fold to whiteflies *(Bemisia tabaci)* as compared to well-known *Allium sativum* leaf agglutinin.

According to the invention a DNA fragment of 651 bp was cloned from total cDNA of *A. fistulosum* leaves, consisted of 585 bp long open reading frame encoding AFAL precursor protein of 194 amino acid residues with 28 amino acid long N-terminal signal peptide and 56 amino acid long C-terminal peptide. The amino acid sequences of AFAL

are different from the other reported *Allium* lectin sequences. The cloned genomic DNA sequence of *afal* showed absence of intron.

In another embodiment of the present invention the nucleic acid sequence having sequence ID. 2, encodes a polypeptide as represented in Seq ID No 4, which is a 110 amino acid residue long mature peptide of *Allium fistulosum* leaf agglutinin(AFAL), having a molecular weight of ~ 12 kDa.

In yet another embodiment of the present invention the polypeptide has the minimum haemagglutination value of in the range of 6-10 ng/ml.

In yet another embodiment of the present invention the polypeptide exhibits insecticidal activity selected from the group comprising haemagglutination, insecticidal, antifungal and anti-prolific activity.

In still yet another embodiment of the present invention the polypeptide exhibits haemagglutination activity selected from the group comprising haemagglutination, insecticidal, antifungal and anti-prolific activity.

Brief description of figures:

Fig. 1A. Lane 1, Molecular weight markers; Lane 2, Total soluble protein from leaves of *A. fistulosum;* Lane 3, Unbound total soluble protein; Lane 4, Column wash (before elution); Lanes 5-8, eluted fractions. **Fig.lB.** Lane 1, Molecular weight markers; Lane 2, Purified AFAL concentrated on 10 kDa cut-off filtration device.

Fig. 2. Peptide mass fingerprinting of AFAL

Fig 3. V-bottom plate showing haemagglutination of rabbit erythrocytes by AFAL

Fig. 4 Clustalw analysis of AFAL with other *Allium lectins.* Amino acid identity of AFAL with other mannose binding lectins vary from 67% to 75%. Dark area represents mannose binding domain.

Fig. 5. Lane 1, Molecular weight markers; lane 2. uninduced bacterial lysate; lane 3-5, sample after 1 h, 2 h and 3 h induction; lane 6, supernatant of 3 hr induced culture; lane 7, protein in inclusion.

Fig 6. Lane 1, molecular weight markers; lane 2, total *E. coli* protein containing SUMO-AFAL; lanes 3-4 represent purified fusion protein.

DETAILED DESCRIPTION OF THE INVENTION

The leaves of *Allium fistulosum* were collected from the National Bureau of Plant Genetic Resources, Bhowali, Nainital, Uttarakhand, India, and used for the purification of *Allium fistulosum* leaf agglutinin (AFAL). It was purified on mannose-agarose affinity column, followed by cut-off filtration device (Example 1, Fig. 1 & 2). The purified protein was used for brief characterization.

AFAL shows several times better insecticidal activity against sap sucking pest *Aphis gossypii* (cotton aphids), *Bemisia tabaci* (whitefly) as compared to ASAL (Example 3). Large amount of purified protein is required for further characterization. The purification of AFAL from plant leaves in bulk amount is difficult due to unavailability of plant material and very low accumulation of the lectin in leaves. Expression of protein in a re combinant system is an alternative approach to produce the desired protein in large amount, which required the cloning of AFAL encoding gene. The gene *(afal)* was cloned from cDNA prepared from total RNA of leaves of *A. fistulosum.* The cloning was done by RACE (Rapid Amplification of cDNA Ends) using degenerate primers designed from the conserved mannose binding domain.

The cloned DNA fragment gene was of 651 bp, consisted of 585 bp open reading frame, 66 bp 5' untranslated leader sequence. The full-length gene encoding AFAL precursor protein of 194 amino acid residues had 28 amino acid long N-terminal signal and 56 amino acid long C-terminal peptide. It contains three mannose binding domains as reported in the case of other mannose binding lectins (Example 5). The cloned genomic DNA sequence of *afal* had no intron. The amino acid sequences of AFAL are different from the other reported *Allium* lectin sequences (Example 5, Fig. 4).

The gene encoding insecticidal protein was cloned in *E. coli* expression vector in fusion with SUMO peptide and the recombinant insecticidal protein was expressed. The protein was purified on Ni-NTA column. The recombinant protein showed the insecticidal activity against cotton aphids and whiteflies. The gene encoding the mature

peptide was cloned in plant expression vector pBI121 under CaMVE35S promoter (Example 6-8).

In the embodiment of the invention, the cloned full-length gene sequence of *Allium fistulosum* leaf agglutinin *(afal)* similar to Seq. ID No. 1. The Seq. ID no. 1 contains DNA sequence which encodes N-terminal signal peptide, mature peptide and C-terminal peptide.

In another embodiment of the invention, nucleotides encoding N-terminal signal peptide and C-terminal peptide are removed from the Seq. ID No. 1, to obtain the mature protein encoding gene sequence, similar to the Seq. ID No. 2.

In another embodiment of the invention, the gene sequence of Seq. ID No. 1 was translated to obtain the full-length amino acid sequences of AFAL similar to Seq. ID No. 3.

In yet another embodiment of the invention, the gene sequence of Seq. ID No. 2 was translated to obtain mature AFAL with amino acid sequence similar to Seq. ID No. 4.

In yet another embodiment of the invention, the gene encoding the mature AFAL is cloned in *E. coli* SUMO expression vector where SUMO peptide is fused with AFAL at the N-terminus and expressed in *E. coli* under T7 promoter.

In yet another embodiment of the invention, the AFAL agglutinated rabbit erythrocytes.

In yet another embodiment of the invention, the AFAL was tested against insect pests like cotton aphid *(Aphis gossypii)* and whiteflies *(Bemisia tabaci).*

In yet another embodiment of the invention, the gene encoding the mature AFAL was cloned in plant expression vector under constitutive promoter CaMV35S and phloem specific promoters CoYMoV, RSS1, RolC etc. and expressed in transgenic plants for insect control.

(5) Examples

Example 1. Isolation and purification of the protein

Allium fistulosum leaves were collected from National Bureau of Plant Genetic Resources, Regional Centre, Bhawali, Uttaranchal, India. The protein was purified

according to the protocol described (Smeets et al., 1997b). The purified protein was further purified on 50 kDa cut-off filtration device. The purified protein was concentrated on 10 kDa cut-off filter and stabilized in PBS for experiments.

Example 2 Peptide Mass finger printing

The purified protein was resolved on SDS-PAGE. The protein band was excised and digested with trypsin and used for peptide mass finger printing. The data was analyzed on MASCOT search. The peptides were found matching to the mannose binding lectins (fig.2).

Example 3. Haemagglutination assay

Haemagglutination assays with rabbit RBCs were carried out in V-bottomed microtitre plates. Total volume of assay was 100 µï, 50 µï aliquot of two-fold serially diluted lectin in PBS was mixed with 50 µï of 2% trypsinized rabbit erythrocytes suspension. Microtitre plate was incubated for 1 hour at room temperature. Agglutination was assessed visually. Reciprocal of the highest dilution of lectin showing detectable agglutination was taken as titer of the haemagglutination

Table 1. Haemagglutination assay of AFAL and its comparison with ASAL

S.N	Agglutinin	Haemagglutination
1	*Allium sativum* agglutinin (standard)	200 ng/ml
2	*Allium fistulosum* leaf agglutinin	8 ng/ml

Example 4. Insect bioassay

Insect bioassay was carried out against sap sucking pest, cotton aphid *{Aphis gossypii)* and whiteflies *(Bemisia tabaci).* The known amount of purified protein was mixed in synthetic diet and insect mortality data was recorded at different time interval. The data was used for the calculation of LC_{50} using probit analysis.

Table 2.1nsecticidal activity of AFAL purified from leaves of *A. fistulosum.*

S.N	**Agglutinin**	Aphids *{Aphis gossypii),* **LC50**	Whiteflies *(Bemisia tabaci),* $\mathbf{LC_50}$
1	*Allium sativum* agglutinin (Standard)	68 μg/ml	76 μg/ml
2	*Galanthus nivalis* agglutinin	51.39 μg/ml	53.39 μg/ml
3	*Allium fistulosum* agglutinin	7.1 μg/ml	8.5 μg/ml

Example 5 Gene cloning and charecterization

cDNA was synthesized following standard protocol. The 3' RACE was performed with degenerate primer {5'-ATGCA(A/G)(C/G)A(G/T)GACTGCAACC-3'} (primer sequence was derived from the mannose-binding site, QXDXNXVXY, conserved among most of the monocot mannose-binding lectins) and universal primer. For 5' RACE, RNA was reversely transcribed with the 5' -RACE CDS Primer. Based on the 3' and 5' RACE results, primers were designed for amplification of full length gene, GSP1 (5'-ATGGACAGTACTCCATCTCCTAAAC-3 ') GSP2 (5'-GCCCCTTGGCCTCCTGCA-3'). The full-length gene was amplified and cloned. The mature AFAL encoding DNA was amplified with primers GSP3 (5'-AGAAACGTATTGGTG AACAACG-3') and GSP4 (5'-TTATCTTCTGTA GGTACCAGTAGAC-3').

The amino acid sequence of AFAL was deduced with Expasy translate tool. The analysis and comparison of the deduced amino acid sequences and nucleotide sequences obtained in RACE was performed with blast p (Standard Protein-Protein BLAST), blastn (Standard Nucleotide-Nucleotide BLAST) on NCBI (www.ncbi.nlm.nih.gov) and clustal W.

Seq. ID no. 1 Nucleic acid sequences encoding full-length AFAL

Seq. ID no. 2 Nucleic acid sequences encoding mature AFAL.

Seq. ID no. 3 Amino acid sequence of the *Allium fistulosum* leaf agglutinin

Seq. **ID** no. 4 Amino acid sequence of the mature *Allium fistulosum* leaf agglutinin

Seq. ID no. 1 Nucleic acid sequences encoding full length AFAL

```
ATGGACAGTA CTCCATCTCC TAAACTAATG AGCATGACCA CTGTAGCCAC CATCCTAACC        60
ATTTTGGCAT CTACATGCAT GGCCAGAAAC GTATTGGTGA ACAACGAAGG ACTGTACGCA       120
GGCCAATCCC TAGTCGTAGA ACAGTACACT TTTACAATGC AGGATGACTG CAACCTTGTA       180
CTCTACGAAT ACTGCGCCCC AATCTGGGCC TCAAACACGG GCGTCACCGG CAAAAATGGG       240
TGCAGGGCCG TGATGCAGGC TGATGGCAAC TTTGTGGTCT ACGATGTTAA CGGGCGTGCC       300
GTCTGGGCCA GTAACAGCAG AAGAGGGAAC GGAAACTATA TCCTGGTGCT TCAGGAGGAC       360
AGGAACGTTG TTATTTACGG ATCTGATATT TGGTCTACTG GTACGTACAG AAGAGGGCCC       420
GGTCCTGGTC CTGGTGCCGC CTGCAAGTGC GATGACGATG GTCCTGACAT TCGCAGTGCT       480
ACTTTGACAG GCACTGTCGA TTTGGGAAGC TGCAACGAGG GATGGGAGAA GTGCGCATCT       540
TTCTACACCA TCCTCGCGGA TTGCTGCAGG AGGCCAAGGG GCTAA                       585
```

Seq. ID no. 2 Nucleic acid sequences encoding mature AFAL.

```
AGAAACGTAT TGGTGAACAA CGAAGGACTG TACGCAGGCC AATCCCTAGT CGTAGAACAG        60
TACACTTTTA CAATGCAGGA TGACTGCAAC CTTGTACTCT ACGAATACTG CGCCCCAATC       120
TGGGCCTCAA ACACGGGCGT CACCGGCAAA AATGGGTGCA GGGCCGTGAT GCAGGCTGAT       180
GGCAACTTTG TGGTCTACGA TGTTAACGGG CGTGCCGTCT GGGCCAGTAA CAGCAGAAGA       240
GGGAACGGAA ACTATATCCT GGTGCTTCAG GAGGACAGGA ACGTTGTTAT TTACGGATCT       300
GATATTTGGT CTACTGGTAC GTACAGAAGA                                        330
```

Seq. ID no. 3 Amino acid sequence of AFAL

```
MET ASP SER THR PRO SER PRO LYS LEU MET SER MET THR THR VAL ALA THR ILE LEU THR  20
ILE LEU ALA SER THR CYS MET ALA ARG ASN VAL LEU VAL ASN ASN GLU GLY LEU TYR ALA  40
GLY GLN SER LEU VAL VAL GLU GLN TYR THR PHE THR MET GLN ASP ASP CYS ASN LEU VAL  60
LEU TYR GLU TYR CYS ALA PRO ILE TRP ALA SER ASN THR GLY VAL THR GLY LYS ASN GLY  80
CYS ARG ALA VAL MET GLN ALA ASP GLY ASN PHE VAL VAL TYR ASP VAL ASN GLY ARG ALA 100
VAL TRP ALA SER ASN SER ARG ARG GLY ASN GLY ASN TYR ILE LEU VAL LEU GLN GLU ASP 120
ARG ASN VAL VAL ILE TYR GLY SER ASP ILE TRP SER THR GLY THR TYR ARG ARG GLY PRO 140
GLY PRO GLY PRO GLY ALA ALA CYS LYS CYS ASP ASP ASP GLY PRO ASP ILE ARG SER ALA 160
THR LEU THR GLY THR VAL ASP LEU GLY SER CYS ASN GLU GLY TRP GLU LYS CYS ALA SER 180
PHE TYR THR ILE LEU ALA ASP CYS CYS ARG ARG PRO ARG GLY                         194
```

Seq. ID no. 4 Amino acid sequence of the mature AFAL

```
ARG ASN VAL LEU VAL ASN ASN GLU GLY LEU TYR ALA GLY GLN SER LEU VAL VAL GLU GLN  20
TYR THR PHE THR MET GLN ASP ASP CYS ASN LEU VAL LEU TYR GLU TYR CYS ALA PRO ILE  40
```

```
TRP ALA SER ASN THR GLY VAL THR GLY LYS ASN GLY CYS ARG ALA VAL MET GLN ALA ASP     60
GLY ASN PHE VAL VAL TYR ASP VAL ASN GLY ARG ALA VAL TRP ALA SER ASN SER ARG ARG     80
GLY ASN GLY ASN TYR ILE LEU VAL LEU GLN GLU ASP ARG ASN VAL VAL ILE TYR GLY SER    100
ASP ILE TRP SER THR GLY THR TYR ARG ARG                                            110
```

Seq. ID no. 5.
Primer sequence GSP 1: (5'-ATGGACAGTACTCCATCTCCTAAAC-3 ')

Seq. ID no. 6 .

Primer sequence GSP2: (5'-GCCCCTTGGCCTCCTGCA-3')

Seq.ID no.7.

Primer sequence GSP3: (5'-AGAAACGTATTGGTGAAC AACG-3 ')

Seq. ID no. 8.

Primer sequence GSP4: (5'-TTATCTTCTGTA GGTACCAGTAGAC-3 ').

Example 6. Expression of AFAL with N-terminal fusion of SUMO in *E. coli* and purification of SUMO-AFAL

The gene encoding mature AFAL was cloned in *E. coli* expression vector in fusion with SUMO peptide under T7 promoter. SUMO had (His_6) tag attached which helped in the purification of recombinantly expressed protein on Ni-NTA resin. SUMO-AFAL was expressed after induction with IPTG. The expression of the recombinant protein was observed every hour for 3 hours. After 3 hours of induction, cells were harvested by centrifugation; suspended in 20 mM TrisCl (pH 8). Bacterial cells were lysed by lysozymen and disrupted by sonication. The lysed bacterial cells were spun and supernatant and pellet were collected and electrophorased on denaturing PAGE (Fig. 5). Approximately half of the recombinant protein was in the soluble form and rest as inclusion in the pellet.

PURIFICATION OF SUMO-AFAL from *E.coli*

Recombinantly expressed SUMO-AFAL was purified on metal-affinity column. Total bacterial protein was loaded on Ni-column, pre-equilibrated with the buffer (20 mM Tris pH 8, 300 mM NaCl and 10 mM Imidazole). NaCl and Imidazole were used to prevent the binding of non-specfic proteins to the column. The column was washed with same buffer having 20 mM imidazole to remove low affinity bound proteins. Finally, the protein. was eluted with 200 mM Imidazole (Fig.6).

Example 7. **Insect bioassay with recombinant fusion protein**

Insect bioassay was carried out against sap sucking pest, cotton aphid *(Aphis gossypii)* and whiteflies *(Bemisia tabaci).* The known amount of SUMO-AFAL was mixed in synthetic diet and insect mortality data was recorded at different time interval. The data was used for the calculation of LC_{50} using probit analysis. SUMO-ASAL served as positive control. The results of insect bioassay is shown in the table 3

Table 3

S.N	**Agglutinin**	**LC_{50}**	
		Aphids (*Aphis gossypii*)	Whiteflies(*Bemisia tabaci*)
1	Recombinant ASAL (SUMO-ASAL)	55.05 µg/ml	51.2 µg/ml
2	Recombinant AFAL *(SUMO-AFAL)*	4.97 µg/ml	8.80 µg/ml

ADVANTAGES OF THE INVENTION

The lectin protein being disclosed in the present invention (AFAL) is 6-10 folds more toxic to insects like aphids and whiteflies as compared to the standard *Allium sativum* leaf lectin (ASAL). AFAL also showed 25 folds higher haemagglutination activity as compared to ASAL. AFAL binds to lesser number of carbohydrate residues on glycans

array as compared to ASAL. This assures fewer non-specific binding of AFAL to carbohydrates and therefore expected to be safe as compared to ASAL.

References Cited

US **patents**

5545820	**August 13, 1996**	**Gatehouse et al**
6,127,532	**October 3, 2000**	**Raikhel**
WO/1992/002139	**February 20, 1992**	**Gatehouse et al**
5407454	**April 18, 1995**	**Anthony et al**

Other references

Bandyopadhyay S, Roy A, Das S (2001) Binding of garlic *(Allium sativum)* leaf lectin to the gut receptors of homopteran pests is correlated to its insecticidal activity. Plant Sci 161:1025-1033

Harper SM, Hopkins TL, Czapala TH (1998) Effect of wheat germ agglutinin on formation and structure of the peritrophic membrane in European corn borer *(Ostrina nubilalis)* larvae. Tissue Cell 30: 166-176

Hossain MA, Maiti MK, Basu A, Sen S, Ghosh AK, Sen SK (2006) Transgenic Expression of Onion Leaf Lectin Gene in Indian Mustard Offers Protection against Aphid Colonization. Crop Sci 46:2022-2032

Murdock LL, Shade RE (2002) Lectins and protease inhibitors as plant defenses against insects. J Agric Food Chem 50: 6605-661 1

Powell KS, Spence J, Bharathi M, Gatehouse JA, Gatehouse AMR (1998) Immunohistochemical and developmental studies to elucidate the mechanism of action of the snowdrop lectin on the rice brown planthopper *Nilaparvata lugens* (Stal). J Insect Physiol 44: 529-539

Sadeghi A, Smagghe G, Broeders S, Hemalsteens JP, De Greve H, Peumans WJ, Van Damme EJ (2008) Ectopically expressed leaf and bulb lectins from garlic *(Allium sativum* L.) protect transgenic tobacco plants against cotton leaf worm *(Spodoptera littoralis).* Transgenic Res 17:9-18

Saha P, Dasgupta I, Das S (2006) A novel approach for developing resistance in rice against phloem limited viruses by antagonizing the phloem feeding hemipteran vectors. Plant Mol Biol 62:735-52

Sauvion N, Nardon C, Febvay G, Gatehouse AMR, Rahbe Y (2004) Binding of the insecticidal lectin concanavalin A in pea aphid, *Acyrthosiphon pisum* (Harris) and induced effects on the structure of midgut epithelial cells. J Insect Physiol 50: 1137—1150

Smeets K, Van Damme EJM, Van Leuven F, Peumans WJ (1997 a) Isolation, characterization and molecular cloning of a leaf-specific lectin from ramsons *(Allium ursinum* L.). Plant Mol Biol 35: 531-535

Smeets K, Van Damme EJM, Verhaert P, Barre A, Rouge, P, Van Leuven F, Peumans WJ (1997 b) Isolation, characterization and molecular cloning of the mannose-binding lectins from leaves and roots of garlic *(Allium sativum* L.). Plant Mol Biol 33:223-234

Van Damme EJ, Barre A, Rouge P, Peumans WJ (2004) Cytoplasmic/nuclear plant lectins: a new story. Trends Plant Sci 9:484-9

Van Damme EJM, Peumans WJ, Barre A, Rouge' P (1998) Plant lectins: a composite of several distinct families of structurally and evolutionaryrelated proteins with diverse biological roles. CRC Crit Rev Plant Sci 17:575-692

Van Damme EJM, Smeets K, Tonekens S, Van Leuven F, Goldstein IJ, Peumans WJ (1992) The closely related homomeric and heterodimeric mannose-binding lectins from garlic are encoded by one-domain and two-domain lectin genes, respectively. Eur J Biochem 206:413-420.

We claim:

1. Set of gene specific primers GSP 1 and GSP 2 useful for amplification of *Allium fistulosum* leaf agglutinin (afal) gene of seq ID No 1 wherein the said primers have sequences of sequence ID No. 5 and sequence ID No.6 respectively as shown below:

   GSP 1 (5'-ATGGACAGTACTCC ATCTCCTAAAC-3 ')

   GSP2 (5'-TTAGCCCCTTGGCCTCCTGC A-3 ')

2. Set of gene specific primers GSP 3 and GSP 4 useful for the amplification of mature afal gene represented by seq ID No 2, wherein the said primers have sequence ID no. 7 and ID NO. 8 respectively as shown below:

   GSP 3(5'-AGAAACGTATTGGTGAACAACG-3 ')

   GSP 4 (5'-TTATCTTCTGTAGGTACCAGTAGAC-3')

3. A process for preparation of *Allium fistulosum* leaf agglutinin (AFAL) recombiriant protein by amplifying the AFAL gene employing the primers as claimed in claim 1 and *2* , wherein the process comprises of following steps-

   a) Extracting of total RNA from the *Allium fistulosum* leaves

   b) Synthesizing of cDNA from total RNA extracted from leaves of *Allium fistulosum*

   c) Designing of primers GSP1 and GSP2 from 5'- and 3'-RACE fragment of c DNA to clone full-length protein encoding DNA, designing of primers GSP3 and GSP4 for cloning of the DNA encoding mature polypeptide of AFAL protein,

   d) Expressing the DNA encoding the mature AFAL in *E. coli* SUMO expression vector where SUMO peptide is fused with AFAL at the N-terminus and expressed in *E. coli* under T7 promoter to get desired AFAL recombinant protein.

4. Nucleic acid sequence represented by Seq ID No 1, obtained by the process claimed in claim 1 and 3, wherein Seq ID No lcomprised of 1 to 585 nucleotides or the sequence complementary thereto, which encodes full-length *Allium fistulosum* leaf agglutinin polypeptide, wherein nucleotide sequence 583 to 585 is a stop codon.

5. Nucleic acid sequence represented by Seq ID No 2, obtained by PCR amplification with the primers GSP3 and GSP 4 using full-length AFAL encoding gene as template as claimed in claims 1-4 , wherein Seq ID 2 is part of Seq ID 1 having nucleotide sequence from 85 to 414.

6. Polypeptide of Seq ID No 3, encoded by the nucleic acid Seq ID No 1 as claimed in claim 1-5, wherein the said polypeptide is a 194- amino acid residues long precursor of *Allium fistulosum* leaf agglutinin (AFAL) protein.

7. Polypeptide of Seq ID No 4, encoded by the nucleic acid Seq ID No 2 as claimed in claim 5, wherein the said polypeptide is a 110 amino acid residue long mature peptide of *Allium fistulosum* leaf agglutinin(AFAL) recombinant protein , having a molecular weight of ~ 12 kDa.

8. The recombinant protein as claimed in claim 7, wherein the minimum haemagglutinin value of the said polypeptide is in the range of 8 µg/ml.

9. The recombinant protein as claimed in claim 7, wherein the said protein exhibits insecticidal, antifungal, antibacterial and antiprolific activity.

10. The recombinant protein as claimed in claim 7, wherein the said protein exhibits insecticidal activity against *Aphis gossypii, Bemisia tabaci, Helicoverpa armigera* and *Spodoptera litura* or any other crop insects for example insect of order lepidoptera, homoptera, coleopteran, diptera and others.

11. The recombinant protein as claimed in claim 1, wherein the said protein is expressed in *E. coli* and can be expressed in any expression system such as *Pseudomonas,* or in yeast like *Pichia pestoris* or other yeast expression systems.

12. *Allium fistulosum* leaf agglutinin (AFAL) recombinant protein, its encoding polynucleotide, primers and the process of preparation thereof substantially as herein described with reference to the examples and drawings accompanying this specification.

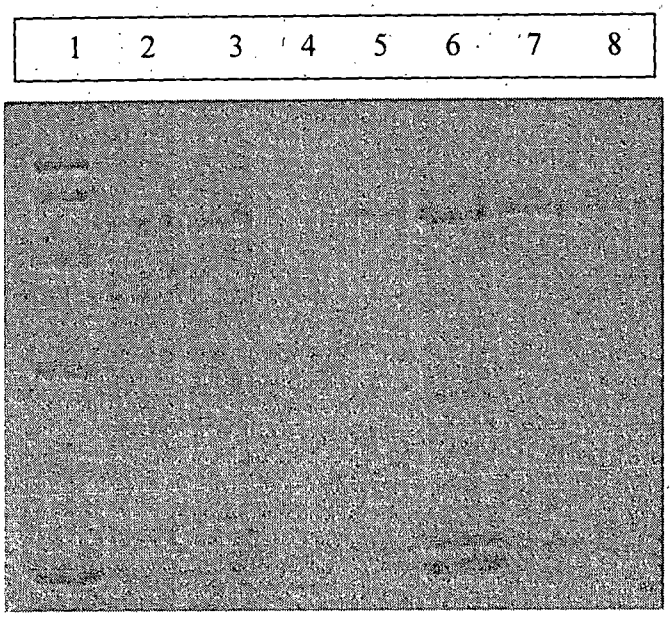

Fig. 1A

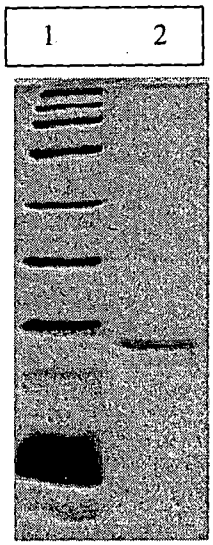

Fig. 1B

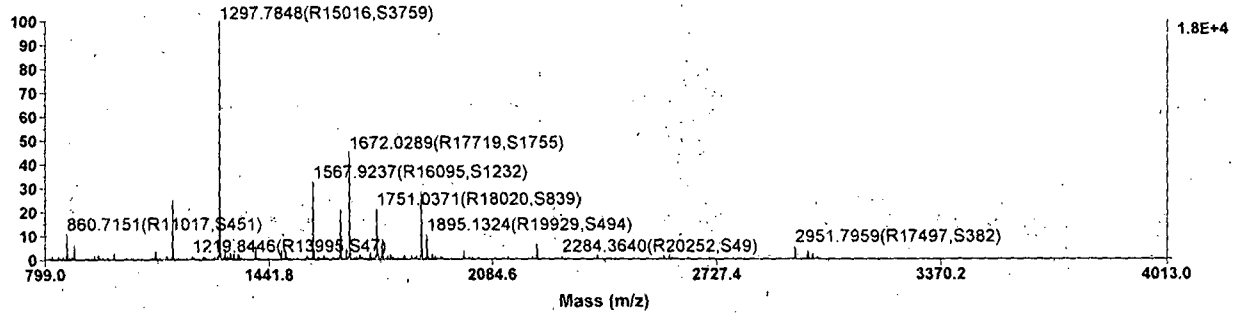

Fig. 2.

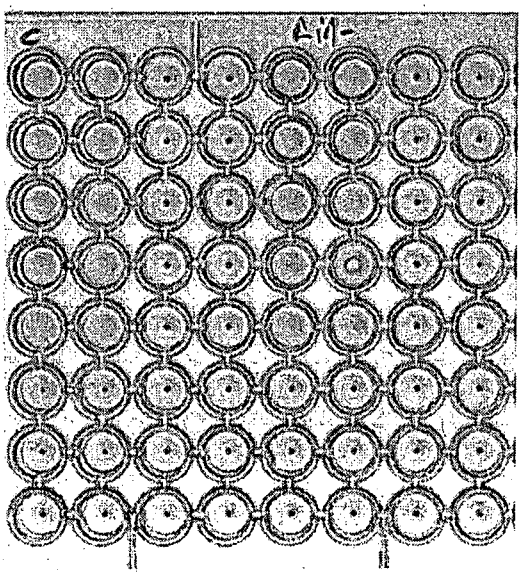

Fig 3.

```
ATA          --MAYSVTCKLIMVCTVGAILSVLTATCMGRNILLNGEGLYAGQSLEEGPYRLAQDDCN 58
AUA          --MAISVNCKIIMVCAVGTILSILTPTSMGRNILLNGEGLYAGQSLEEGSYKLIQDDCN 58
ASA          ------------------------------RNILTNDEGLYAGQSLDVNPYHLIQEDCN 30
ASAL         MGRTTSSPKAMMRIATVAAILTILASTCMARNVLTNGEGLYAGQSLDVEQYKFIQDDCN 60
AMPL         MGRTTPSPKLIMSITTVAAILTILASTCMARNLLTNGEGLYAGQSLDVEQYKFIQDDCN 60
ACA          ---------------TVATILTILASTCMARNVLVNNEGLYAGQSLVVEQYTFIQDDCN 45
AFAL         --MDSTPSPKLMSMTTVATILTILASTCMARNVLVNNEGLYAGQSLVVEQYTFIQDDCN 58
                                        **:* *.*********    * : **:***

ATA          LVLYDEYSRPVWASNTGVTGRNGCRAVMQADGNFVVYDSNSRAVWASNSRKGNGNYILVL 118
AUA          LVLF-EYSTQVWASNTGVSGRNGCRAVMQADGNFVVYDSNSRAVWASQSRRGNGNYILAL 117
ASA          LVLY-DHSTAVWSSNTDIPGKKGCKAVIQSDGNFVVYDAEGASLWASHSVRGNGNYVLVL 89
ASAL         LVLY-EYSTPIWASNTGVTGKNGCRAVMQRDGNFVVYDVNGRPVWASNSVRGNGNYILVL 119
AMPL         LVLY-EYSTPIWASNTGVTGKNGCRAVMQKDGNFVVYDVNGRPVWATNSVRGNGNYILVL 119
ACA          LVLY-EYSTPIWASNTGVTGKNGCRAVMQADGNFVVYDVKGRAVWASNSRRGNGNYILVL 104
AFAL         LVLY-EYCAPIWASNTGVTGKNGCRAVMQADGNFVVYDVNGRAVWASNSRRGNGNYILVL 117
             ***: ::.  :*:***.:.*::**:**:* ******** :. .:**::* :*****:*.*

ATA          QKDRNAVIYGSDIWSTGTYRRGVG--------------GSVVTAMNGTVDAGFAVKNVTT 164
AUA          QEDRNVVIYGTDIWSTGTYRRGVG--------------GTVVTVINGTVDAGSGMENVTA 163
ASA          QEDGNVVIYGSDIWSTNTYK-------------------------------------- 109
ASAL         QKDRNVVIYGSDIWSTGTYRRSVG--------------GAVVMAMNGTVDGGSVIGPVVV 165
AMPL         QQDRNVVIYGSDIWSTGTYRRSAG--------------GPVVMAMNGTVNGGSVVGPVIV 165
ACA          QKDRNVVIYGSDIWSTGTYRKKVG--------------GTVVMAMNGTVDGGSVVGPVTV 150
AFAL         QEDRNVVIYGSDIWSTGTYRRGPGPGPGAACKCDDDGPDIRSATLTGTVDLGSCNEGWEK 177
             *:* *.****:*****.**:

ATA          AAVGDVAIA--------- 173
AUA          TAA--------------- 166
ASA          ------------------
ASAL         NQNVTAAIRKVGTGAA-- 181
AMPL         NQNVT-AIRKVGTSAA-- 180
ACA          NQNVT-AVRKVAAAA--- 164
AFAL         CASFYTILADCCRRPRG- 194
```

Fig. 4.

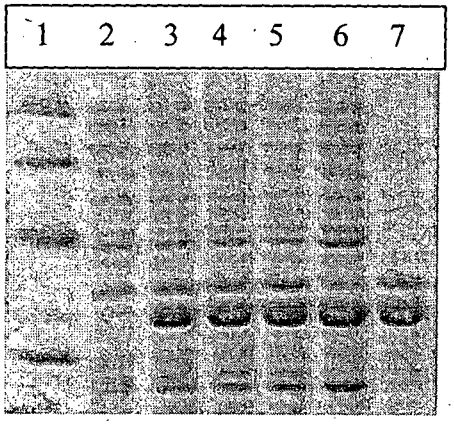

Fig. 5.

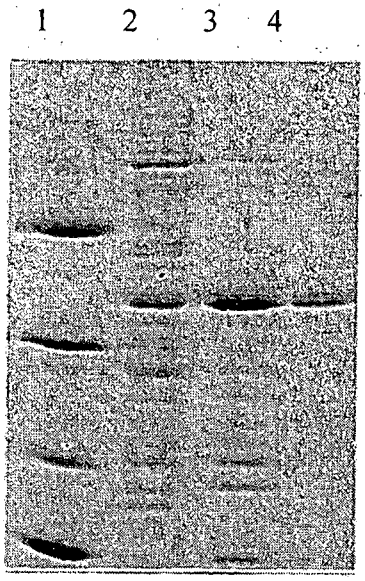

Fig. 6.