Int.J.Curr.Microbiol.App.Sci (2018) *7*(6): 959-967

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 7 Number 06 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.706.114

In Silico Analysis and Molecular Docking Studies of *Cajanus cajan* Lectin against Aminopeptidase-N Receptor from *Acyrthosiphon pisum*

Rakesh Kumar Prajapat[1], Puja Singh[1], Poonam Tiwari[1], Pawan Mainkar[1], Sarika Sahoo[2], A.R. Rao[2] and Rekha Kansal[1*]

[1]ICAR-National Research Center for Plant Biotechnology, Pusa, New Delhi-12, India
[2]ICAR- Indian Agricultural Statistics Research Institute, Pusa New Delhi-12, India

**Corresponding author*

Keywords

Cajanus cajan, Lectins, Functional domain, Molecular docking, Molecular modelling

Article Info

Accepted:
06 May 2018
Available Online:
10 June 2018

A B S T R A C T

Cajanus cajan lectins (CCL) is a promising candidate molecules for the protection against sap sucking insect pests like *Lipaphis erysimi* (Kaltenbach). Legume lectin toxicity against hemipteran has been proven by various experiments. This study explains the insecticidal potential of CCL through molecular modeling and docking with receptor alanyl Aminopeptidase N from *Acyrthosiphon pisum* membrane.The functional domain analysis of CCL revealed metal binding and N-linked glycosylation site. The physico-chemical parameter like theoretical isoelectric point (pI) is 5.59, instability index is 24, aliphatic index is 80.80 and GRAVY value is -0.070. The secondary structure attributes of CCL protein explain its structural behavior and biological activity.The CELLO2GO tools revealed subcellular localization of CCL protein was plasma membrane with a reliability score of 2.289. The present study using various bio-computational tools could, therefore, help in our understanding of CCL protein structure and prove to potential candidate gene for generating transgenic for increased aphid resistance.

Introduction

Legume lectins are structurally and evolutionarily related to a well-defined group of lectins that were originally discovered in the seeds of legumes like jack bean, common bean, pea, peanut and soybean. Many of these lectins have been purified and characterized with respect to their structure, sugar-binding specificity and biological activity (Sharon and Lis, 1990; Van Damme *et al.,* 1998; Peumans and Van Damme, 1999). Legume lectins exhibited fairly homogeneous molecular structure with ~30kDa subunits (Van Damme *et al.,* 2007). These lectins shared extensive sequence homology and three dimensional structural similarities, but differed in carbohydrate specificity (Rouge *et al.,* 1991). The monomer structure is characterized by presence of the "jelly roll" motif that is often associated with carbohydrate-binding activity. The "jelly roll" is characterized by the presence of three sets of antiparallel β-sheets. The sheets are connected by several loops of varying lengths (Loris *et al.,* 1998; Vijayan and Chandra 1999). Two metal ions, calcium and a transition metal, found in all the legume lectin structures are essential for the

carbohydrate binding. The amino acid residues which bind Ca^{2+} and Mn^{2+} metal ions are highly conserved, while the residues which constitute the sugar-binding site are less conserved but exhibit similar properties (Lis and Sharon, 1998). ConA, pea, lentil and lathyrus lectins are dimers formed by side-by-side alignment of two monomers such that the two rear ß-sheets form a contiguous 12-stranded ß-sheet.

The four different region A, B, C and D, associated with the concave face of the seven-stranded curved ß-sheet at the top front-side of the subunit, form the binding site (Sharma and Surolia, 1997). The conserved Asp and Gly/Arg residues are present in loops A and B, whereas Asn and the hydrophobic residues (Phe/Tyr/ Trp/ Leu) are located in loop C. Size of the backbone of C loop determines carbohydrate specificity of the lectin. Aligned sequences of legume lectins showed 4 to 7 gaps in the binding loop D, indicating variation in the loop size (Sharma and Surolia, 1997) which contributes to broad specificity of legume lectins.

Pigeon pea (*Cajanus cajan* (L.) Mill sp.) is an important food legume crop that is predominantly cultivated in tropical and subtropical regions of the world. It is diploid (2n=22) crop with a genome size of 808 Mb. India is the primary pigeonpea growing country, accounting for 4.42 M ha area and 2.86 million tons of production (FAO, 2012). *Cajanus cajan* lectin is a dimer composed of identical subunit with N- and C-terminal residues of threonine and alanine respectively. Its amino acid composition is characterized by presence of hign content of acidic amino acids (Siddiqui *et al.*, 1995). The present study was carried out to study the physico-chemical properties, secondary structure attributes and 3D model of CCL protein. The insect receptor binding potential with CCL protein was evaluated through docking study.

Materials and Methods

Sequence retrieval and analysis

The amino acid sequence of *Cajanus cajan* lectin (CCL) (accession #KU382473.1) protein was retrieved from NCBI database. The functional domains of lectin were determined using the InterPro tool available on the EBI web page (www.ebi.ac.uk/interpro/). The physico-chemical properties like amino acid composition, pI, molecular weight, half-life and instability index were determined using Protparam (http://web.expasy.org/protparam/). Probability of protein disorder was determined by the PrDOS (Protein disorder prediction server) tool (http://prdo s.hgc.jp). The subcellular location and molecular functions of protein were predicted by using CELLO2GO (http://cello.life.nctu.edu.tw/ cello2go/) web server.

Structural analysis and homology-based modelling

The secondary structure and solvent accessibility of CCL was determined by the RaptorX protein structure server (http://raptorx.uchicago.edu/StructurePredictio n/predict/). The 3D structure of the target protein CCLwas generated using SWISS Model tool (https://swissmodel.expasy.org/) using Pea Lectin as a template through homology based modelling. The authenticity of the predicted models was further validated employing RAMPAGE tool (http://mordred. bioc.cam.ac.uk/~rapper/rampage.php).

Molecular docking and active site mapping

In order to accomplish the docking studies, *Acyrthosiphon pisum* membrane alanyl Aminopeptidase N (APN; Accession # DQ440823), was used as receptor for CCL. ClusPro Docking server

(http://cluspro.bu.edu/) was used and results were viewed through Discovery Studio 4.1 visualizer. The amino acid residues forming the cleft area involved in interacting with APN were also identified using ClusPro Docking server (http://cluspro.bu.edu/). PDBSum tool (http://www.ebi.ac.uk/pdbs um) was used to further identify the amino acid residues binding with the ligand present on the lectin binding proteins which undergo glycosylation as a part of post translational modification for functional activation.

Results and Discussion

Sequence analysis and characterization

The functional domain of CCL protein sequence were defined using InterPro tool (Fig. 1). The residue annotation identifies metal binding sites of CCL are GLU^{149}, $ASP^{151}$and HIS^{166}. In case of legume lectins, the presence of metal ions i.e., Mn^{2+} and Ca^{2+} was documented to be very important (Sharon and Lis, 1990). This is signified by the evolutionarily conserved amino acid residues that bind to the metal ions. For example, Con A requires Mn^{2+} and Ca^{2+} for its activity (Hardman and Ainsworth, 1972). The residue annotation using InterPro tools also revealed that, ASN^{135} of CCL involved in N-linked glycosylation. The CCL is characterized as acidic protein based on computed pI value 5.59 (pI<7). The CELLO2GO tools revealed that the CCL protein is localised with plasma membrane with a reliability score of 2.289. This protein play important molecular functions in mannose binding, carbohydrate binding and metal ion like Ca^{+2} and Mn^{+2} binding. Recently Moraes Filho *et al.*, 2017, reported that most of the legume lectins are located in the extracellular medium or associated with the plasma-membrane and play important function in ion binding, kinase activity and enzyme regulator. The instability index of CCL was 24.00, classifying it as a stable protein which is also justified with the result obtained from PrDOS tool. Two disordered regions were predicted in the protein sequence, of which the longest disordered region was found between Gly^{264} to Ala^{275} comprising 12 amino acid residues (Fig. 2). The estimated half-life in mammalian reticulocytes was 30 h, while in yeast and *Escherichia coli* was more than 20 and 10 h, respectively. The aliphatic index of CCL was 80.80. The Aliphatic index (Ai) of proteins determine its thermo stability under changing climatic conditions (Gupta *et al.*, 2012).

GRAVY indices for CCL was -0.070, indicates the possibility of better interaction with water i.e. hydrophilic nature of the protein which is attributed to charged amino acid residues present in the protein sequence (25 negatively charged and 21 positively charged), suggesting that CCL might be associated as extrinsically in plasma-membrane.

Structural analysis and homology-based modelling

The secondary structure of CCL generated with RaptorX predicted a total of 7% α helices, 42% β pleated sheets and 50% Coil (Fig. 3A). It also revealed solvent accessibility of this protein as 27% residues are buried into structure, 30% residues are exposed and 41% were medium (Fig. 3B). Most of legume lectin proteins contain about 40–50% b-sheet, 35–45 % β-turn and 0–10% α-helix, and thereby fall into a structurally distinct class of proteins. The β-sheet, α-helical and β-turn content predicted here agrees well with X-ray structure determination of Con A by Reeke *et al.*, (1975). The 3D model of CCL was generated by SWISSMODEL using pea lectin (ID - 2bqp.1) as template with identity of 88.41% and coverage was 85% (Fig. 4A).

Fig.1 Functional domains analysis of CCL protein sequence through InterPro

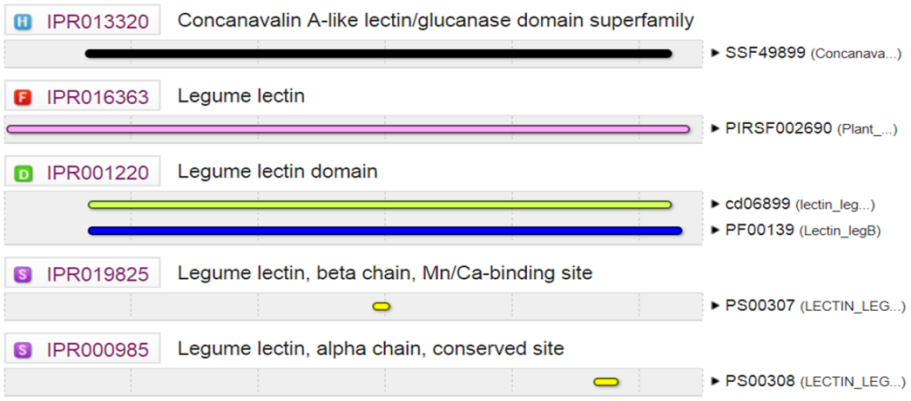

Fig.2 Prediction of the disordered amino acid residues present in CCL protein (shown in red) using PrDOS tool

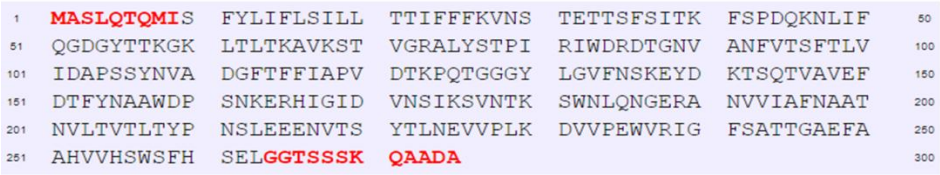

Fig.3A 3-class secondary structure displayed by individual amino acid residue of CCL protein

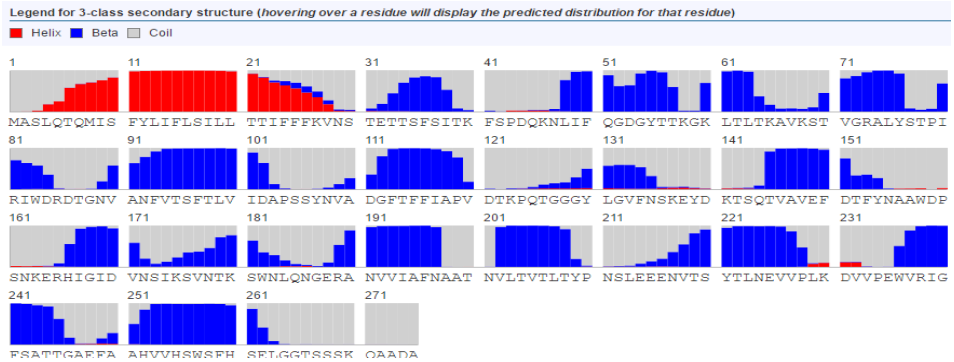

Fig.3B Solvent accessibility of individual amino acid residue of CCL protein

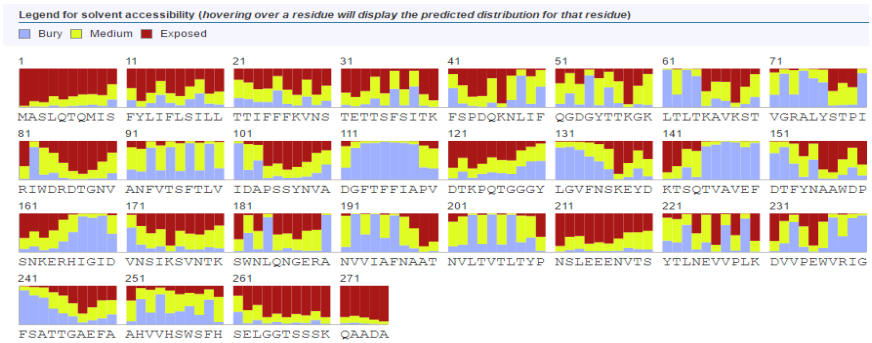

Fig.4 3D model of (A) CCL (B) APN generated via homology-based modeling using SWISS MODEL depicting various secondary structures—α helices, β pleated sheets and random coils

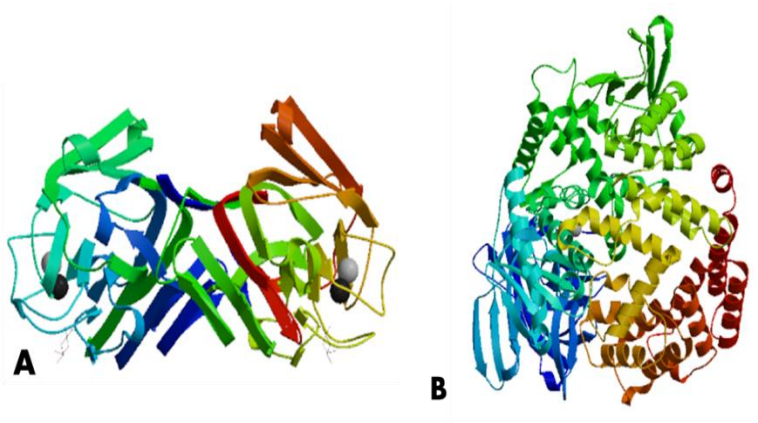

Fig.5 Validation of 3D predicted structure using RAMPAGE (A) CCL (B) APN

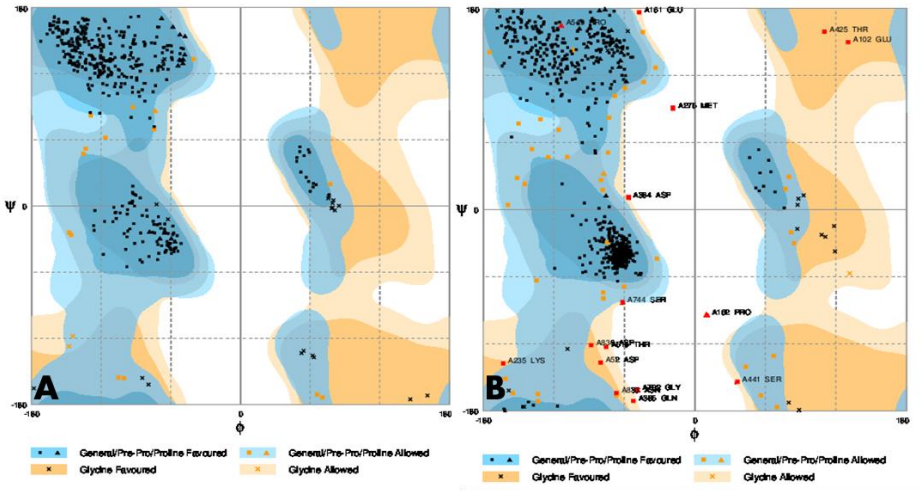

Fig.6 Molecular docking of CCL with APN, showing hydrogen bond (green dotted line) at interaction sites

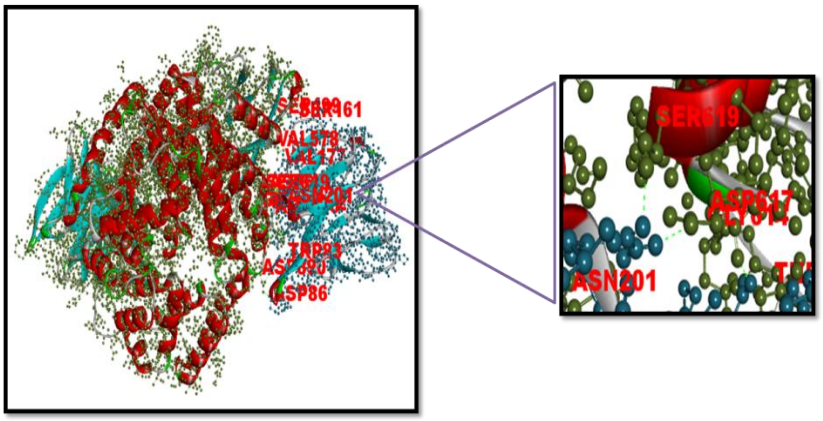

Fig.7 LigPlot of CCL protein and its interaction with ligand XMM303 (A)

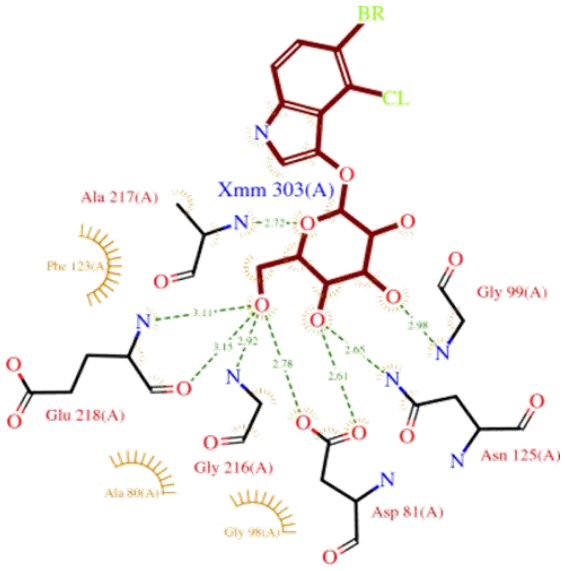

Table.1 Amino acid residues with their respective positions involved in interaction through H-bonding in molecular docking

PPL	APN
ASN201	SER619
ASN201	ASP617
ASP810	ASP86
ASP890	TRP83
GLY614	ASN172
VAL578	VAL177
THR510	ASP140
SER499	SER160

Table.2 Amino acid residue and their respective position involved in interaction with ligand XMM with H-bond length in LigPlot

Target protein (CCL)		Ligand (XMM)	
Res. name	Res.position	Res. position	Distance (A^0)
ASP	81	303	2.61
ASP	81	303	2.78
GLY	99	303	2.98
ASN	125	303	2.65
GLY	216	303	2.92
ALA	217	303	2.72
GLU	218	303	3.11
GLU	218	303	3.15

The 3D model of APN with sequence identity 30.59% and coverage 89% generated by same server using APN1 from *Anopheles gambiae* as a template (ID-4wz9.1.A) (Fig. 4B). The generated 3D models were further validated with RAMPAGE programme. The torsion angles ψ and ϕ were examined to access the reliability of the protein model. The result obtained in the validation of CCL protein, 96.3% of the amino acid residues were found in the most favoured region, while 3.7% of amino acid residues were found in allowed region (Fig. 5A). The validation result for APN revealed that 93.5% of amino acid residues were found in most favoured region, while 4.7% and 1.8% of amino acid residues were present in allowed and outlier region (Fig. 5B).

Molecular docking and active site mapping

Molecular docking through ClusPro revealed positions of interaction between *Cajanus cajan* lectin and APN receptor (Fig. 6). The amino acid residues name and positions involved in interaction are summarized in Table 1. The amino acid residue of both the protein involved in inter chain H-bond, without selecting any residue in lectin protein and receptor, were evaluated using ClusPro server. Active site mapping for determining the residues involved in ligand binding i.e. XMM [(5-Bromo-4-Chloro-3-Indolyl)-A-D-Mannose] for *Cajanus cajan* lectin protein is done by LigPlot using PDBSum tool (Fig. 7). The LIGPLOT displays all the interactions in term of hydrogen bond between the ligand and the residues of protein molecules (Table 2). A wide range of lectins, viz., GNA, Con A, PSA and ASA, exhibiting mannose or mannose/glucose sugar binding affinity, revealed profound anti-metabolic effects towards members of the homopteran insects both under in vitro (Habibi *et al.,* 1993; Powell *et al.,* 1993; Rahbe *et al.,* 1995) and in planta conditions (Powell *et al.,* 1995; Gatehouse *et al.,* 1996; Rao *et al.,* 1998). Most common sugar specificities expressed by lectins are towards Mannose, Mannose/Glucose, Mannose/Maltose, Fucose, Galactose/N-acetylgalactosamine, N-acetylglucosamine, sialic acid and complex glycan groups (Peumans and Van Damme, 1998). The nature of lectins and carbohydrates interaction has been detailed in a review by Del Carmen Fernandez-Alonso *et al.,* (2012).

Cajanus cajan lectin protein sequence was retrieved from NCBI database and further subjected for *in silico* determination of insecticidal potential using various computational approaches. The structural attributes like secondary structure, physicochemical properties and the amino acid residues participating in the catalytic activity were determined using bio-computational tools. 3D model of the protein was generated via homology-based modeling which can help in understanding the structural and functional characteristics of the protein. The binding potential of CCL with its probable ligand showed that this protein could serve as best candidate gene for generating transgenic against hemipterans. This model can serve as a template for characterizing various lectins from plant species since no templates are available for the same in the protein model database.

Acknowledgement

RKP and PM are thankful to DBT (Department of Biotechnology), Govt. of India, and PS, PT and RK are thankful to ICAR-NPTC for providing the financial assistance.

References

Del Carmen Fernández-Alonso, M., Díaz, D., Berbis, M., A., Marcelo, F., Cañada J.

and Jiménez-Barbero J., 2012. Protein-Carbohydrate Interactions Studied by NMR: From Molecular Recognition to Drug Design. *Current Protein and Peptide Science*, 13:816-830.

FAOSTAT, Food and Agriculture data (Statistics Division, Food and Agriculture Organization of United Nations, Rome, Italy) 2012 [http://faostate.fso.org/].

Gatehouse, A.M.R., Down, R.E., Powell, K.S., Sauvion, N., Rahbe, Y., Newell, C.A., Merryweather, A., Hamilton, W.D.O. and Gatehouse, J.A., 1996. Transgenic potato plants with enhanced resistance to peach-potato aphid *Myzuspersicae*. *Entomol. Exp. Appl*. 79: 295-307.

Gupta, S.K., Rai, A.K., Kanwar, S.S., Sharma, T.R., 2012. Comparative analysis of zinc finger proteins 553 involved in plant disease resistance. *PLoS One* 7(8): e42578.

Habibi, J., Backus, E.A., and Czapla, T.H., 1993. Plant lectins affect survival of the potato leafhopper (Homoptera: Cicadellidae). *J. Econ. Entomol*. 86: 945-951.

Hardman, K. D. and Ainsworth, C. F. (1972). Structure of concanavalin A at 2.4 Å resolution. *Biochemistry;* 1149: 10-19.

Lis, H. and Sharon, N., 1998. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev*. 98: 637-674.

Loris, R., Hamelryck, T., Bouckaert, J., Wyns, L., 1998. Legume lectin structure. *Biochem. Biophys. Acta*. 1383: 9-36.

Moraes Filho, R.M., Rossiter, J.G., Cavalcanti, Junior, E.A., Martins, L.S.S., 2017.*In Silico* Comparative Analysis of Legume Lectins. *J Genet Genom* 1:103.

Peumans, W.J. and VanDamme, E.J.M., 1998. Plant lectins; Specific tools for the identification, isolation and characterization of O-linked glycans. *Critic. Rev. Biochem Mol. Biol*. 33: 209-258.

Powell, K., S., Gatehouse, A.M.R., Hilder, V.A., and Gatehouse, I.A., 1993. Antimetabolic effects of plant lectins and plant and fungal enzymes on the nymphal stages of 2 important rice pests, *Nilaparvata lugens* and *Nephotettix virescens*. *Entomol. Exp Appl*. 66: 119-126

Powell, K.S., Gatehouse, A.M.R., Hilder, V.A., and Gatehouse, J.A., 1995. Antifeedant effects of plant lectins and an enzyme on the adult stage of rice brown planthopper, *Nilaparvata lugens*. *Entomol. Exp. Appl*. 75: 51-59.

Rahbe, Y., Sauvion, N., Febvay, G., Peumans, W.J.and Gatehouse, A.M.R., 1995. Toxicity of lectins and processing of ingested proteins in the pea aphid *Acyrthosiphon pisum*. *Entomol. Exp. Appl*. 76: 143-155.

Rao, K.V., Rathore, K.S., Hodges, T.K., Fu, X., Stoger, E., Sudhakar, D., Williams, S., Christou, P., Bharathi, M., Bown, D.P., Powell, K.S., Spence, J., Gatehouse, A.M.R., and Gatehouse, J.A., 1998. Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. *Plant J*. 15: 469-477.

Reeke, G. N., Jr., Becker, J. W. and Edelman, B.M., 1975. The covalent and three-dimensional structure of concanavalin A. IV. Atomic coordinates, hydrogen bonding, and quaternary structure. *J. Biol. Chem*., 250, 1525-1547.

Rouge, P., Cambillu, C. and Bourne, Y., 1991. The three-dimensional structure of legume lectins. In: Lectin Reviews. Kilpatrick DC, Van Driessche E, Bog-Hansen TC (eds). Vol.1 Sigma Chemical co. St.Louis USA pp.143-159

Sharma, V. and Surolia, A., 1997. Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary activity. *J. Mol. Biol*. 267: 433-445.

Sharon, N. and Lis, H., 1990. Legume lectins –a large family of homologous proteins. *The FASEB. J*. 4:3198-3208.

Siddiqui S., Hasan S., and Salahuddin A., 1995. Isolation and characterization of cajanuscajan lectin. *Archives of biochemistry and biophysics*., 319;426-431.

Van Damme, E.J.M., Allen, A.K. and Peumans, W.J., 1988. Related mannose-specific lectins from different species of the family Amaryllidaceae. *Physiol. Plant.*73: 52-57.

Van Damme, E.J.M., Culerrier, R., Barre, A., Alvarez, R., Rouge, P. and Peumans, W.J., 2007. A novel family of lectins evolutionarily related to class V chitinases: an example of neofunctionalization in legumes. *Plant Physiol*. 144: 662-672.

Vijayan, M. and Chandra, N., 1999. Lectins. *Curr. Opin. Struct. Biol*. 9: 707-714.

How to cite this article:

Rakesh Kumar Prajapat, Puja Singh, Poonam Tiwari, Pawan Mainkar, Sarika Sahoo, A.R. Rao and Rekha Kansal. 2018. *In Silico* Analysis and Molecular Docking Studies of *Cajanus cajan* Lectin against Aminopeptidase-N Receptor from *Acyrthosiphon pisum*. *Int.J.Curr.Microbiol.App.Sci.* 7(06): 959-967. doi: https://doi.org/10.20546/ijcmas.2018.706.114