CHRISTIANE ELIZA MOTTA DUARTE

BRASSICA OLERACEA LECTIN: ISOLATION, CHARACTERIZATION, AND FUNCTIONAL ASSESSMENT OF THE FIRST LECTIN WITH MATH DOMAINS

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Biologia Celular e Estrutural, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2016

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T

D812b
2016

Duarte, Christiane Eliza Motta, 1986-
Brassica oleracea lectin : isolation, characterization, and functional assessment of the first lectin with MATH domains / Christiane Eliza Motta Duarte. – Viçosa, MG, 2016.
xiii, 70f. : il. (algumas color.) ; 29 cm.

Orientador: Leandro Licursi de Oliveira.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Lectinas. 2. Proteínas. 3. *Brassica oleracea*.
I. Universidade Federal de Viçosa. Departamento de Biologia Geral. Programa de Pós-graduação em Biologia Celular e Estrutural. II. Título.

CDD 22 ed. 571.96

CHRISTIANE ELIZA MOTTA DUARTE

BRASSICA OLERACEA LECTIN: ISOLATION, CHARACTERIZATION, AND FUNCTIONAL ASSESSMENT OF THE FIRST LECTIN WITH MATH DOMAINS

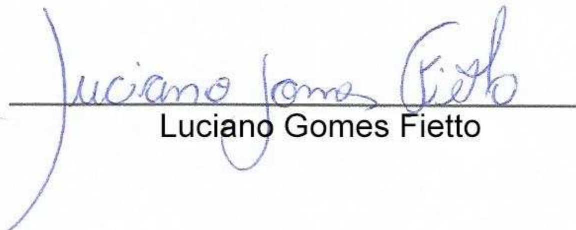
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Biologia Celular e Estrutural, para obtenção do título de *Doctor Scientiae*.

APROVADA: 16 de setembro de 2016

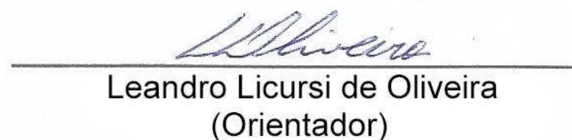
Elyabe Monteiro de Matos

Anesia Aparecida dos Santos

Eduardo de Almeida Marques da Silva



Luciano Gomes Fietto



Leandro Licursi de Oliveira
(Orientador)

ACKNOWLEDGMENTS

I would like to thank the Universidade Federal de Viçosa and graduate program Biologia Celular e Estrutural for the technical and scientific support and the quality teaching without which the execution of this project would not be possible.

I thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, the Fundação de Amparo à Pesquisa do Estado de Minas Gerais and the Conselho Nacional de Desenvolvimento Científico e Tecnológico for the financial support.

I thank the ERASMUS$^+$ ceiA3 program and the European Union for the opportunity to carry out an exchange period. I would thank the Universidad de Córdoba and all the teachers and employees that made possible the accomplishment of my academic mobility.

I am grateful my advisor Leandro Licursi for his guidance and advice, for his ability to understand without limits and mainly for the freedom that has always been given me during these years of doctorate and that enabled me to grow as a person and as a researcher.

I would like to thank the teachers participating in my thesis defense: Eduardo Marques, Anésia dos Santos, Elyabe Matos and Luciano Fietto for the collaboration and all the teachers and employees who directly or indirectly contributed to my training.

Thanks to the teachers Sérgio Oliveira de Paula, from the Laboratório de Imunovirologia and Wagner Campos Otoni, from the Laboratório de Cultura de Tecidos, for receiving me promptly in their laboratories so that I could carry out much of the experiments presented here. Special thanks to teacher Wagner, my co-advisor Andrea Koehler, Elyabe Matos and Diego Batista who helped me a lot in the conduct of stress experiments and all the friends I made in the Laboratório de Cultura de Tecidos.

I am very thankful to Dr. Josefa Muñoz Allamillo, from the Universidad de Córdoba, for excellent guidance during my exchange and for being an example of motivation and love for research. The opportunity to work with you renewed my expectations in the science and certainly your way of "make science" will always be remembered by me during my career.

I thank the friends and colleagues of the Laboratório de Fisiología Molecular y Biotecnología de Plantas: Fran, Oscar, Cristina, Marta, Ana and Gody and the teachers Pedro Piedras and Manuel Pineda for sharing with me the experiences and discoveries of the exchange.

I want to thank my friends at the Laboratório de Imunoquímica e Glicobiologia: Patrick Fernandes, Wendeo Silveira, Alessandra Teixeira, Sabrina Emerick, Thaís Fialho e Patrizia Coelho and the technician Jorge Luiz Santos for the enjoyable company.

Thanks also to my friends and classmates. Especially Iara Calil, for being a friendly shoulder, even thousands of miles away. Thanks to the friends of kickboxing by the affection and the support.

A special thanks to my friends of floor Nieves Rayo, Salvador Polvillo, Jérôme Mouret and Antonio Gomez who welcome me so kindly in Córdoba “¡muchas gracias por tornar mi estancia en la ciudad inolvidable e increíble!”

I thank my parents Liliam da Conceição Motta and Eduardo Gomes Duarte, my brother Marcus Eduardo Motta Duarte and all my relatives for their unconditional love and for always supporting me in the search for the fulfillment of my dreams. To them I owe much of this achievement.

I am very thankful to my husband Carlos José Ferreira Lopes for being by my side in the most difficult hours. During these four years you were always present and it was my safe harbor, your love and encouragement reinvigorated my strength, helping me never give up. Thank you for your understanding and support and for believing in my potential.

I thank God above all, for without him none of this would be possible. The Lord gave me the strength and motivation not to lose my target and give up my aims. To Him I dedicate this work, "because of Him, and through Him, and to Him are the all things".

“Pon en manos del Señor todas tus obras,

y tus proyectos se cumplirán”.

Proverbios16:3

BIOGRAPHY

CHRISTIANE ELIZA MOTTA DUARTE, daughter of Eduardo Gomes Duarte and Liliam da Conceição Motta, was born in Belo Horizonte, Minas Gerais, on 12 January 1986.

In 2010, she graduated from the Federal University of Viçosa with a Bachelors of Science in Biology, when it was honored with the gold medal Presidente Bernardes, awarded to the student with the highest academic standing on graduation. Throughout the graduation, Christiane received a fellowship of the Tutorial Education Program (PET/Biology) and of the PIBIC/CNPq and worked in the Cytogenetics and Cytometry Laboratory. Also was Teacher's Assistant in the Departments of Animal and General Biology.

In 2010, she entered the Ph.D. program at the Federal University of Viçosa with a fellowship from Conselho Nacional de Desenvolvimento e Pesquisa (CNPq). In 2012 received a M.Sc. in Genetics and Embreeding with the dissertation “Fra(X) site identification by cytogenetic techniques using antagonists and different culture mediums”.

In 2012, Christiane then shifted his academic focus towards isolation and characterization of proteins and began her Ph.D. work in Cell and Structural Biology from the same University. She was tutor in Cell Biology and Cell Biology Laboratory disciplines, and received Ph.D. research fellowships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). She was awarded an honorable mention with the work “Cloning and sequencing of a TRAF-like protein isolated from *Brassica oleracea* ssp. *botrytis* with lectin activity” presented at the 7th Academic Production Symposium of the School of Biological and Health Sciences, in 2015. In May 2016, she was selected by the European program ERASMUS+ ceiA3 and held an exchange at the University of Cordoba, Spain, where developed part of the results presented in her thesis.

SUMMARY

LIST OF FIGURES ..... **viii**

LIST OF TABLES ..... **ix**

ABREVIATIONS ..... **x**

RESUMO ..... **xii**

ABSTRACT ..... **xiii**

GENERAL INTRODUCTION **1**

Plant lectins: an overview *1*

Brief historical background *1*

Lectin classification and occurrence *2*

Constitutively Expressed and Inducible Lectins *5*

Interaction between lectins and animal immune system *6*

Physiological role of plant lectins *7*

Aim and scope of the present investigation *9*

References *10*

CHAPTER 1 ..... **17**

Introduction **17**

Materials and methods **18**

Biological material *18*

Soluble protein extraction procedures *18*

Hemagglutination activity assay *18*

Protein purification *18*

SDS-PAGE *19*

Protein concentration *19*

Determination of protein glycosylation *19*

N-terminal Sequencing *19*

Mass spectrometry analysis and protein sequencing by tandem mass spectrometry *20*

Inhibition of hemagglutination *20*

Glycoproteins proteolysis *21*

Effects of temperature, pH and divalent cations on lectin activity *21*

Phagocytic activity of the peritoneal macrophages *21*

NO production by peritoneal macrophage assay *22*

H_2O_2 production by the peritoneal macrophage assay *22*

Statistical Analyses *22*

Results **22**

Protein Purification *22*

Properties of purified lectin ... 24
Carbohydrate specificity of the purified lectin ... 26
Effect of temperature and pH ... 27
Cauliflower lectin activates macrophages and promotes phagocytosis ... 29
Discussion ... 29
Conclusion ... 32
CHAPTER 2 ... 37
Introduction ... 39
Materials and Methods ... 40
Bacterial strains and culture media ... 40
In vitro *seed germination of* Brassica oleracea *ssp.* botrytis ... 40
RNA extraction and cDNA cloning ... 41
Sequence analysis ... 41
Gene expression analysis using semi quantitative and real time quantitative PCR ... 42
Effect of light and darkness in the Bol expression ... 42
Effects of different biotic and abiotic stimuli on the Bol expression ... 43
Construction of expression vector ... 43
Expression and purification of rBOL ... 43
SDS-PAGE ... 44
Biological activity ... 44
Thermal stability of recombinant lectin ... 45
Western Blot Analysis ... 45
Analysis of the predicted three-dimensional structure of the BOL protein ... 45
Circular dichroism ... 45
Statistical analyses ... 46
Results ... 46
Cloning and sequence analysis of Bol cDNA ... 46
The Bol promoter contains regulatory elements in response to light and stress signals ... 48
Comparative tissue specific expression of Bol ... 50
Bol expression is regulated by light ... 50
Bol response to biotic stimuli ... 51
Recombinant expression of pET21b(+)-Bol ... 52
Purification of recombinant lectin ... 53
Properties of recombinant lectin ... 54
Discussion ... 56
GENERAL CONCLUSIONS ... 70

LIST OF FIGURES

GENERAL INTRODUCTION

Figure 1 Ribbon drawing gallery of representative lectins or lectin-like proteins from plant…………………………................................ 4

CHAPTER 1

Figure 1 Purification of cauliflower lectin………………………………………… 23
Figure 2 Properties of the purified BOL……………………………………….. 25
Figure 3 Mass spectrometric analysis of BOL…………………………………… 26
Figure 4 Physicochemical characterization of BOL…………………………….. 28
Figure 5 Macrophage activation…………………………………………………... 29

CHAPTER 2

Figure 1 Nucleotide and deduced amino acid sequences of BOL……………. 47
Figure 2 Alignment of deduced amino acid sequences between BOL and uncharacterized proteins of *Brassicas*……………………………… 48
Figure 3 Comparative expression level of *Bol* in various tissues……………… 50
Figure 4 Effect of light on *Bol* expression…..………………………………… 51
Figure 5 Expression analysis of *Bol* gene under different environmental conditions in *B. oleracea* ssp *botrytis*………………………………….. 52
Figure 6 Solubility comparison of rBOL expressed at different temperatures and purification of recombinant protein………………………………... 53
Figure 7 Properties of rBOL……………………………………………………….. 55
Figure 8 The structure of BOL……………………………………………….…… 56
Figure S1 Predicted *Bol* promoter nucleotide sequence of *Brassica rapa* FPsc v1.3 extracted using Phytozome database....……………………….... 66

LIST OF TABLES

GENERAL INTRODUCTION

Table 1 Classification of lectins families based on evolutionary and structurally related domains…………………………………………… 5

CHAPTER 1

Table I Specific hemagglutination activities and chromatographic fraction yields obtained at different stages of lectin purification……………... 24

Table II Effects of the various carbohydrates and glycoproteins on the hemagglutination induced by the *B. oleracea* lectin…………………. 27

CHAPTER 2

Table 1 Elements present in the predicted promoter region of *Bol* according to the PLANT CARE database………………………………………… 49

Table 2 Hemagglutination activity of BOL and rBOL………………………….. 59

Table S1 List of primers used in this study..…………………………………... 65

ABREVIATIONS

AAA	*Anguilla anguilla* agglutinin
ABL	*Agaricus bisporus* lectin
ACA	*Amaranthus caudatus* agglutinin
ACC	1-aminocyclopropane-1-carboxylic acid
ANOVA	analysis of variance
ArtinM	lectin of *Artocarpus heterophyllus*
BALB/c	inbred strain of mouse
BOL	*Brassica oleracea* ssp. *botrytis* lectin
bp	base pair
BSA	bovine serum albumin
BTB	brie a brae domain
CD	circular dichroism
cDNA	complementary deoxyribonucleic acid
ConA	concanavalin A
CRD	carbohydrate recognition domains
CV-N	Cyanovirin-N
DAB	diaminobenzidine
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DTT	dithiothreitol
EeI	*Euonymus europaeus* lectin
ET	Ethylene
FPLC	fast protein liquid chromatography
g	centrifugal force
HA	hemagglutinating activity
HPLC	high-performance liquid chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
JA	jasmonic acid
LB	Luria Bertani Broth
LysM	lysin motif
LNP	lectin-nucleotide phosphohydrolase

MALDI	matrix-assisted laser desorption ionization
MATH	meprin and TRAF homology
MeJa	methyl jasmonate
MHC	major histocompatibility complex
MM	molecular marker
mRNA	messenger ribonucleic acid
MS	Murashige and Skoog medium
m/z	mass-to-charge ratio
NLR	NOD-like receptors
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PHA	phytohemagglutinin
PNA	peanut agglutinin
rBOL	recombinant *Brassica oleracea* ssp. *botrytis* lectin
RIP	ribosome-inactivating protein
RNA	ribonucleic acid
RT-qPCR	quantitative reverse transcription PCR
SA	salicylic acid
SalT	salt-inducible protein
ScLL	*Synadenium carinatum* lectin
SBA	soybean agglutinin
SDS	sodium dodecyl sulfate
TLR	Toll-like receptor
TOF	time-of-flight
TRAF	TNF-receptor associated factor
TRIM	tripartite motif
UBQ	ubiquitin
v/v	volume per unit volume
WGA	wheat germ agglutinin
w/v	weight per unit volume
w/w	weight per unit weight

RESUMO

DUARTE, Christiane Eliza Motta, D.Sc., Universidade Federal de Viçosa, Setembro de 2016. **Lectina de *Brassica oleracea*: Isolamento, caracterização e avaliação funcional da primeira lectina com domínios MATH.** Orientador: Leandro Licursi de Oliveira. Coorientadores: Sérgio Oliveira de Paula e Andréa Dias Koehler.

Lectinas estão envolvidas em uma ampla variedade de processos biológicos, inclusive atuando como agentes imunomoduladores capazes de ativar a imunidade inata. Nós purificamos e caracterizamos uma nova lectina de couve-flor (*Brassica oleracea* ssp. *botrytis* – BOL) através de três etapas sequenciais de cromatografia, cuja pureza foi confirmada por SDS-PAGE. Além disso, avaliamos o papel dessa lectina na imunidade inata por meio de um ensaio de fagocitose, produção de H_2O_2 e NO. BOL foi caracterizada como uma proteína não glicosilada com uma massa molecular de ~ 34 kDa em SDS-PAGE. Para otimizar o processo de obtenção da lectina e possibilitar um estudo mais aprofundado de sua estrutura e função, realizou-se a clonagem molecular e expressão heteróloga de BOL. Para isso utilizamos RNA total extraído de plântulas de couve-flor e isolamos a sequência de 1053 pb do cDNA codificante. Ferramentas de bioinformática foram utilizadas para determinar a sequência promotora de 1000 pb de *Bol*, a qual revelou vários elementos cis-regulatórios conhecidos por estarem envolvidos em várias condições de estresse na planta. A análise comparativa tecido-específica da expressão de *Bol* demonstrou níveis mais elevados do transcrito nas folhas, em comparação aos tecidos do caule e da raiz. A análise da sequência de aminoácidos e o alinhamento com proteínas homólogas deduzidas nos permitiu determinar que a proteína madura é constituída por 301 aminoácidos. A estrutura tridimensional predita confirmou que esta lectina apresenta uma estrutura em forma de cúpula com dois domínios MATH. Este é o primeiro relato do isolamento, clonagem e expressão bacteriana de uma lectina com domínios MATH e pode ser de interesse significativo para compreender o papel regulador desta proteína como agente imunoestimulante bem como para à fisiologia da própria planta.

ABSTRACT

DUARTE, Christiane Eliza Motta, D.Sc., Universidade Federal de Viçosa, September, 2016. ***Brassica oleracea* lectin: Isolation, characterization, and functional assessment of the first lectin with MATH domains**. Adviser: Leandro Licursi de Oliveira. Co-advisers: Sérgio Oliveira de Paula and Andréa Dias Koehler.

Lectins are involved in a wide range of biological mechanisms, including act as immunomodulatory agent able to activate the innate immunity. We purified and characterized a new lectin from cauliflower (*Brassica oleracea* ssp. *botrytis* - BOL) by three sequential chromatographic steps and confirmed the purity by SDS-PAGE. Additionally, we evaluated the role of the lectin in innate immunity by a phagocytosis assay, production of H_2O_2 and NO. BOL was characterized as a non-glycosylated protein with a molecular mass of ~34 kDa in SDS-PAGE. To optimize the process of the lectin obtaining and allow further study of their structure and function, the molecular cloning and heterologous expression of BOL were carried out. Using total RNA extracted from cauliflower seedlings a *Bol* coding cDNA sequence of 1053 bp was isolated. Bioinformatics tools were used to determine a promoter sequence of 1000 bp of *Bol* which revealed several key cis-regulatory elements known to be involved in various plant stresses. Comparative expression analysis of tissue specific *Bol* demonstrated the highest transcript levels in leaves, as compared to stem and root tissues. Analysis of amino acid sequence and alignment with deduced homologous proteins allowed us to determine that the mature protein comprises 301 amino acids. Predicted three-dimensional structure confirmed that this lectin had an overall dome-like structure with two MATH-domains. This is the first report of isolation, cloning and bacterial expression of a lectin with MATH-domains and may be of significant interest to understand the regulatory role of this protein as immunostimulatory agent as well as to the physiology of the plant itself.

GENERAL INTRODUCTION

Plant lectins: an overview

Abstract

Lectins are involved in a range of biological mechanisms. In this review, we approach historical aspects of the origin and discovery of lectins. Since the Stillmark's work with castor bean extract in 1888 to present. The actual classification based on CDR structure e conservation. Likewise, the physiological role of lectins in defense of plant and animal immunity. All with special focus on plant lectins.
Keywords: agglutinin, hemagglutinin, lectin, lectin-like proteins

Brief historical background

Erythrocyte-agglutinating proteins has been known in nature since the turn of the 19th century. In the 1960s it became apparent that such proteins also agglutinate other types of cells, and that many of them are sugar-specific (Sharon and Lis, 2004). These cell-agglutinating and sugar-specific proteins have been named lectins: proteins with ability to bind specifically and reversibly to carbohydrates, agglutinate cells and/or precipitate glycoconjugates (Boyd and Shapleigh, 1954). This class of proteins was first described in 1888 by Stillmark, who observed that the seed extract of castor bean (*Ricinus communis*) could agglutinate erythrocytes. In 1891, Hellin identified a toxic hemagglutinin present in bean seed extract of *Abrus precatorius*. These lectins were used to immunize mice by Ehrlich (1899), a pioneer in immunological research. This author has established some of the basic principles of immunology and showed that plants agglutinins were antigenic models more useful than the bacterial toxins (Ehrlich, 1899).

In 1899, Camus found the first animal lectin secreted by the albumin gland of snail *Helix pomatia*. The term hemagglutinin (of german *Blutkörperchenagglutinin*) was introduced in 1898 by Elfstrand to describe this group of proteins capable of causing cell agglutination. Landsteiner and Raubitschek (1908) were the first to discuss the specificity of lectins, and observed that legume seed extracts show different hemagglutination properties when combined with different animal erythrocytes. Watkins and Morgan (1952) demonstrated that lectins can bind to monosaccharides and that this binding has a key role to agglutinating activity promoted by these proteins.

The ability of plant agglutinins to discriminate between erythrocytes of different blood types led Boyd and Shapleigh (1954) to use the term lectin (from the Latin '*lectus'*—chosen, selected). Since then several definitions have been proposed to lectins. The three distinct requirements for a protein to qualify as a lectin are: i) contain a carbohydrate recognition domain (CRD) which sole purpose is the effective differentiation between various oligosaccharides; ii) is not the product of the immune system like antibodies are and iii) biochemically does not modify the carbohydrate moiety which it binds, different than carbohydrate-specific enzymes – like glycosyltransferases, glycoside hydrolases and trans glycosylases, that modify their substrates (Goldstein *et al*., 1980; Rüdiger and Gabius, 2001).

Lectin classification and occurrence

Based on the overall domain architecture and properties of agglutination, lectins are classified into: merolectins, hololectins, chimerolectins and superlectins. Merolectins have a single binding domain carbohydrates. They are simple polypeptide proteins, which due to its monovalent structure are unable to precipitate glycoconjugates or agglutinate cells. Hololectins present two or more homologous or identical CRDs. This group comprises all lectins that have multiple binding sites to the same sugar or structurally similar sugars and, because of this, are able to agglutinate cells or precipitate glycoconjugates. The chimerolectins are composed of one or more CRDs tandemly arrayed with an unrelated domain with catalytic or biological activity, which acts independently of the carbohydrate binding domain. Depending on the number of CRDs, these molecules may or not show hemagglutination activity (Peumans and Van Damme, 1995). Superlectins are a special class of hololectins, which have two or more CRDs with specificity for structurally different sugars (Cammue *et al*., 1986).

Lectins represents a diversified group of proteins with respect to size, composition and structure (Sharon *et al*., 1974). Each lectin domain has its own characteristic overall fold with one or more carbohydrate-binding sites. In general, the three-dimensional structure of lectins is composed of a high content of β-sheets with little contribution from α-helixes. The β-sheets are connected by loops forming antiparallel chains (Sharon and Lis, 1990). A ribbon drawing gallery of the crystal and solution structures of representative fold of lectins or lectin-like proteins from plants is shown

in figure 1. Different classification systems have been used in an attempt to categorize this heterogeneous group of proteins. In principle, they were grouped based on the carbohydrate binding specificity (Peumans and Van Damme, 1995), but this system proved to be artificial and uninformative about the evolutionary point of view (De Schutter and Van Damme, 2015). With the progress in the purification and characterization of lectins, evidences were accumulated of that lectins are a very heterogeneous group of proteins, artificially grouped together simply based on the capability of cell agglutination.

Recently, biochemical and transcriptome analyses revealed that lectins can be divided into multiple families of structurally and evolutionary related proteins based on a conserved CRD. In plants, lectins were classified into twelve families (Van Damme *et al.*, 2008) that with new findings about the three-dimensional structure were enlarged to fifteen (Fujimoto *et al.*, 2014) – Table 1. However, there are many lectins that can not be included in this classification system because they have a unique structure that can belong to others well-established protein families that are generally not linked to sugar-binding activity (De Schutter and Van Damme, 2015).

Despite the structural diversity of lectins, hitherto lectins with MATH domains not have been described. The meprin and TRAF-C homology (MATH) domain is a fold of seven or eight anti-parallel β-sheets that form β-sandwich structure containing a surface crevice responsible for protein-protein interactions (Zapata *et al.*, 2007), and carbohydrate-binding activity has not been identified in this protein domain. Proteins with MATH domains, known as TRAFs, are commonly associated with other protein domains, apart from a large number of hypothetical proteins identified by conceptual translation of genomes displaying one or multiple MATH-domains arranged in tandem which their function is not yet known (Zapata *et al.*, 2007).

The importance of carbohydrate-binding proteins is shown by their occurrence in all kingdoms of life. Lectins are widely distributed in nature and are found in all classes and families of living organisms, from viruses, bacteria and fungi to plants and animals (Sharon and Lis, 1987). Tissue and cellular distribution of the lectins is varied and depends on the developmental stage and pathological condition (Sharon and Lis, 2004). In plants are predominantly isolated from seeds where they constitute about 10% of the total soluble protein of the seed extracts (Van Damme *et al.*, 1998a). Lectins are also found in vegetative tissues such as leaves, fruits, roots,

tubers, rhizomes, bulbs, bark, stem, phloem sap and even nectar (Peumans and Van Damme, 1995).

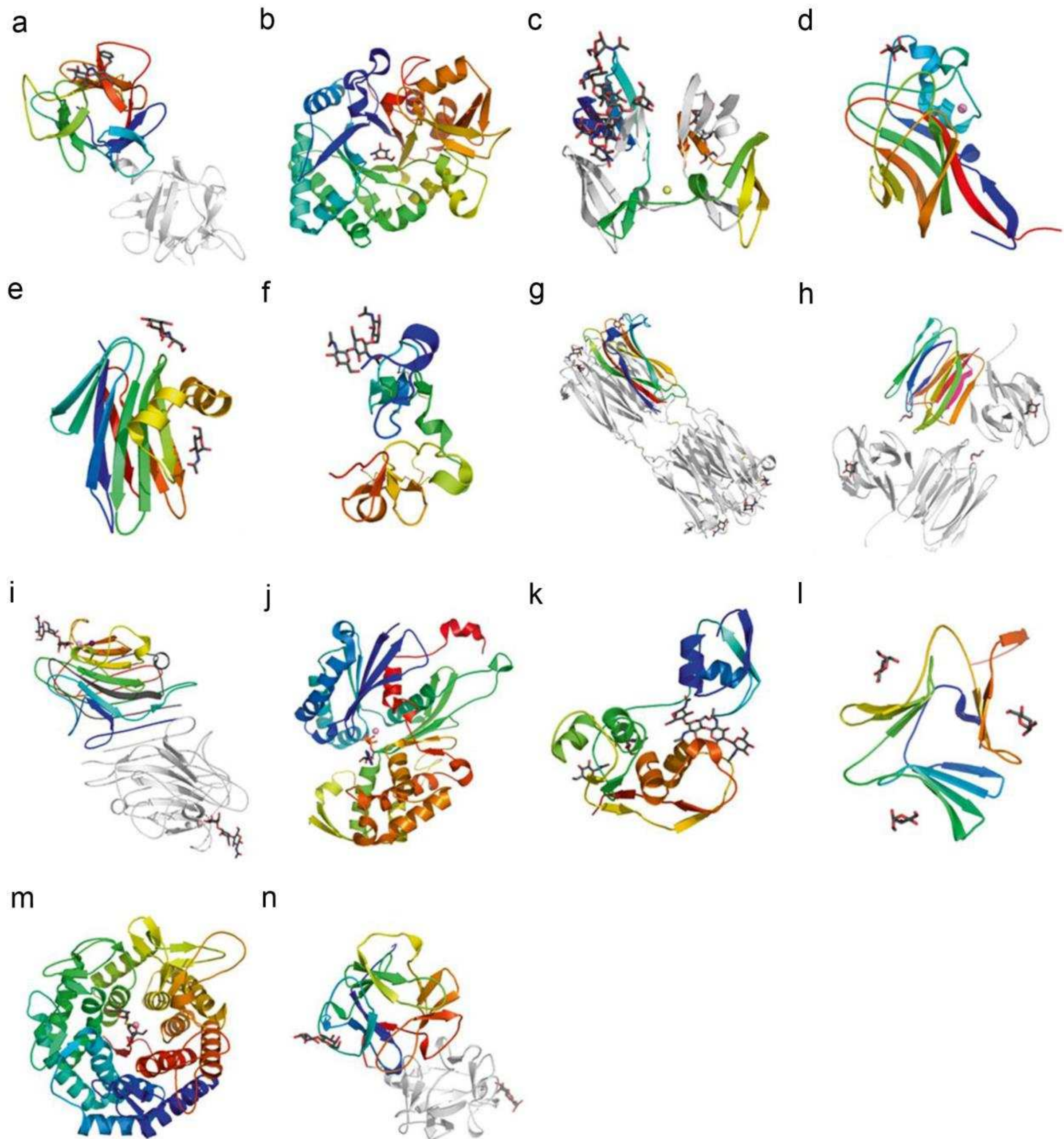


Figure 1 – Ribbon drawing gallery of representative fold of lectins or lectin-like proteins from plants. 3D-structures of lectins were extracted from the Protein Data Bank. Shown as monomer. Some lectins share the same protein motifs, e.g. R-type, the β-Trefoil motif that is present in animal, plant, fungal, bacterial and viral lectin. (a) ACA-like. (b) Chitinase-like. (c) CV-N family. (d) F-type. (e) ABL-like. (f) Hevein family. (g) H-type. (h) Jacalin family. (i) Legume family and L-type-like (j). LNP-type. (k) LysM family. (l) Monocot family. (m) M-type. (n) Ricin-B family. Spheres show ions and sticks show ligands. Adapted from Fujimoto *et al.* (2014).

Table 1. Classification of lectins families based on evolutionary and structurally related domains (Adapted from Van Damme *et al*., 2008; Fujimoto *et al*., 2014)*

Plant lectin family	Fold
Amaranthin family (ACA-like)	β-Trefoil
Chitinase-like	$(\beta/\alpha)_8$-barrel
CV-N family	(3-stranded β-sheet and β-hairpins$)_2$
F-type (AAA-like, Eel-lectin, fucolectins)	β-Sandwich with Ca ion
Fungal fruit-body (Actinoporin-like, ABL-like)	α/β-Sandwich (actinoporin-like)
Hevein family	Hevein-like cystine knot motif Dimer
H-type	Six-stranded antiparallel β-sandwich
Jacalin related family	β-Prism I
Legume family (L-type)	β-Sandwich
LNP-type (N-type)	RNAseH-like α/β-fold
L-type-like	β-Sandwich
LysM family	LysM βααβ-fold (lysin motif)
Monocot family(bulb-type lectin)	β-Prism II
M-type	$(\alpha/\alpha)_7$ –barrel
Ricin-B family (R-type)	β-Trefoil

*Lectin families, such as X-lectin, Nictaba-like lectin, and *Euonymus europaeus* lectin, whose members' structures have not been determined, are not included in this table.

Constitutively Expressed and Inducible Lectins

The lectins of constitutive expression are known as "classical lectins" and are expressed in high amounts in seeds and vegetative storage tissues. Most of the constitutively expressed lectins are synthesized with a signal peptide, and accumulate in vacuoles or related organelles, or are secreted to the extracellular compartment (Van Damme *et al*., 2003). It is generally believed that the storage proteins like lectins have multiple functionalities (Cândido *et al*., 2011). In normal conditions, lectins remain inactive as storage proteins but when the plant is challenged they act as defense proteins because when an invading microbe or insect disrupts the plant's cellular contents, the stored lectin, which may agglutinate or sicken the pathogen or predator, is released (Chrispeels and Raikhel 1991; Peumans and Van Damme 1995; Vandenborre *et al*., 2011).

In addition, lectin expression also can be induced as particular responses toward specific biotic/abiotic stimuli, such as salt stress, drought, light, heat or cold shock, wounding or treatment with abscisic acid, jasmonic acid and gibberellins. In the absence of plant stress, the inducible lectins generally are not expressed at detectable levels. Most strikingly, all inducible lectins identified thus far reside in the nucleus and the cytoplasm of the cells (Lannoo and Van Damme, 2010). Based on these observations, the concept was developed that lectin–mediated protein–carbohydrate interactions in the cytoplasm and in the nucleus play an important or possibly even crucial role in the stress physiology of the plant cell (Van Damme *et al.*, 2004). During the last decade, this new class of plant lectins has been extensively studied (Macedo *et al.*, 2015). The first purified inducible lectin was a mannose-specific jacalin-related lectin (called Orysata) from NaCl-treated rice seedlings (Zhang *et al.*, 2000). Orysata was designated as a lectin only in 2000. However, ten years earlier Claes *et al.* (1990) showed that Orysata is a salt-inducible protein (SalT) and that the salT mRNA accumulates in sheaths and roots of mature plants and seedlings upon salt or drought stress. SalT expression was also induced by abscisic and jasmonic acid treatment (Moons *et al.*, 1997).

Interaction between lectins and animal immune system

It is well documented that lectins show immunomodulatory effects that are initiated by their interaction with glycan's moieties on the surface of immune cells. Such interaction initiate a signal-transduction system, which culminating in the cytokines production and induces efficient immune responses against tumors or pathogens infections (Souza *et al.*, 2013). Some lectins induce lymphocyte proliferation such as concanavalin A (Kilpatrick, 1999). Others are able to stimulate the innate immune response, as soluble lectins that direct the elimination of the foreign agent and help in the phagocytosis by macrophages and dendritic cells. These proteins bind to oligomannosides of infectious microorganisms, causing activation of complement system without participation of antibody, and subsequent lysis of the pathogens, thus acting in innate immunity (Ghazarian *et al.*, 2011). Furthermore, lectins present on the cell surface of dendritic cells and macrophages are involved in the immune surveillance, as endocytic receptors and of cell signaling. Lectins type C can act as mannose, galactose or N-acetyl-galactose receptors, which are involved in the

internalization of antigens, processing and subsequent presentation to T lymphocytes as a peptide-major histocompatibility complex (MHC) (Aarnoudse *et al*., 2006; Denda-Nagai *et al*., 2010). This antigen presentation activates specific immune responses then, lectins participate indirectly in the adaptive immune system (Gorelik *et al*., 2001)

Therefore, these proteins are also directly involved in adaptive immunity. Leukocytes express L-selectins, which aid in the lymphocyte homing and trafficking of leukocyte to sites of inflammation (Tedder *et al*., 1995). They mediate the binding of leukocytes to endothelial cells and thereby initiate a rolling phase, in which the lectins interact transiently with glycan ligands, leading eventually to their extravasation. The T lymphocytes once activated have reduced L-selectin expression or lacking altogether, which allows them to migrate and exit from the site of inflammation via high affinity interaction between integrins and their specific ligands (Gorelik *et al*., 2001).

Another mechanism of action of the lectins as agonists of the immune system was proposed by Coltri *et al*. (2008). It occurs via Toll-like receptor (TLR). The Th1 cytokines production is induced by plant lectins through interaction with glycosylated receptors on macrophages and/or dendritic cells, such as type 2 and 4 Toll-like receptors (TLR2 and TLR4). Several plant lectins may act as TLR agonists. The soybean (SBA), peanut agglutinin (PNA), ConA, and phytohemagglutinins (PHA-L and PHA-P) stimulate extracellular TLRs: TLR-4 for SBA and PNA; TLR-2/6 for ConA; TLR-2/6, -4 and for PHA-L, whereas wheat germ agglutinin WGA is able of activate all tested receptors, except TLR-3 and -4 (Unitt and Hornigold, 2011). Plant lectins can also regulate Th2 immunity. The ScLL, a lectin from *Synadenium carinatum* reduces the leukocyte trafficking and activates Th2 cytokine production in mice (Rogerio *et al*., 2007). This type of response can be induced by the lectin B-chain of type-2 ribosome-inactivating proteins (type-2 RIPs) from *Ricinus communis* through binding to D-galactose containing glycans present on the surface of enterocytes (Carter *et al*., 2010).

Physiological role of plant lectins

In the past lectins were widely used as specific tools in studies of carbohydrates chemistry and histology. However, in recent years they have received recognition as

a class of communication proteins that are involved in interactions between plants and their environment and therefore can play an important role in the physiology of the plant (Van Damme *et al*., 1998b). This question of the possible physiological role of lectins has intrigued researchers since 1970s (Etzler, 1986). In spite of the extensive knowledge about the molecular biology and structure of lectins, their physiological role with respect to the plant still not well understood (De Hoff *et al*., 2009).

The first concepts proposed that lectins could function as antibodies to protect plants against harmful soil bacteria, control seed germination or, be involved in the transport and storage of sugars, but no evidence to confirm these propositions were found (Van Damme *et al*., 1998b). Biological events such as seed maturation or maintenance of seed dormancy (Howard *et al*. 1972; Peumans and Stinissen, 1983), mitogenic stimulation of plant embryonic cells (Howard *et al*., 1972), packaging or mobilization of storage materials (Weber and Neumann, 1980) and N_2 fixation and source for developing embryo (Díaz *et al*., 1995; Peumans and Van Damme, 1995) have been associated with lectins. An advance in the search of physiological role for plant lectin was obtained when these proteins were recognized not only to play a role in the plant itself (e.g. nitrogen storage or as a specific recognition factor) but to be able to interact with foreign organisms recognizing glycoconjugates present on the surface or in the digestive tract of these organisms (Van Damme *et al*., 1998b). The most attractive hypothesis to date for plant lectin function is that these proteins may have a role in plant defense, and protect plants against phytopathogenic microorganisms and insects as well as herbivores (Sharon and Lis, 2004).

The first evidences that reinforce the idea that lectins could act as plant defense proteins was based on the observation that WGA, PNA, and SBA inhibited the sporulation and growth of fungi such as *Trichoderma viride*, *Penicilium notatum*, and *Aspergillus niger* (Barkai-Golan *et al*., 1978). Posteriorly, the potato lectin also showed a similar effect on the phytopathogen fungal *Botrytis cinerea* (Callow, 1977) and the anti-fungal properties of several others lectins has been shown (Guan *et al*., 2008; van Deenen *et al*., 2011; Wang *et al*., 2012).

Lectins also act in the mechanisms of plant defense against insect herbivores. The insecticidal action of lectin found in black bean was reported in 1976. The feeding of bruchid beetles with a diet containing this lectin resulted in the death of

their larvae, so the main role of lectins in legumes could be protect them from attack by insect seed predators (Janzen *et al*., 1976). Several other lectins were shown to be insecticidal, among them WGA, *Galanthus nivalis* lectin and jacalin (Sharon and Lis, 2004). The genes encoding *Allium sativum* leaf and bulb lectin (Sadeghi *et al*., 2003), *Galanthus nivalis* agglutinin (Nagadhara *et al*., 2004), and pea lectin (Melander *et al*., 2003) have been introduced into tobacco, wheat, and rice to reduce predation by insects. In a recent study, Orysata, a mannose-specific lectin from the jacalin-related family, was expressed in transgenic tobacco and its insecticidal activity against three pest insects: beet armyworm (*Spodoptera exigua*), green peach aphid (*Myzus persicae*) and pea aphid *(Acyrthosiphon pisum*) were evaluated. *S. exigua* larvae fed on transformed plants showed higher mortality, reduced weight and extension of larval development. Similar effects also were observed in *M. persicae* and *A. pisum* (Al Atalah *et al*., 2014). However mechanisms, whereby lectins confer resistance to plant bacterial and fungal pathogens, remain poorly defined (De Hoff *et al*., 2009).

Aim and scope of the present investigation

Since its discovery lectins have been used as a research tool in areas such as: biochemistry, cell biology and immunology, as they provide a model system to study the molecular basis of biological events involving recognition between proteins and carbohydrates. Our research group works with lectins and their biological functions and isolated a new lectin from cauliflower florets (*Brassica oleracea* ssp. *botrytis*). In preliminary experiments this protein was shown to be a potential immunomodulatory agent. So to increase the applicability and understanding of the action of this lectin, we investigated its biochemical properties and determined its gene sequence. From these results we were able to express the lectin cauliflower in bacterial system, predict the three-dimensional structure of this lectin and compare it with other proteins of the same group. Through knowledge of the gene sequence and determination of regulatory elements present in the promoter it was also possible to infer a possible physiological role of this protein to the plant itself. The specific aims of this study were:

1. Purification, physicochemical characterization, and molecular cloning of lectin from *Brassica oleracea* ssp. *botrytis;*

2. Evaluation of the effects of the cauliflower lectin on the physiology of macrophages (phagocytosis, production of NO, H_2O_2 and superoxide anion);
3. Development of a system of expression, solubilization and purification to produce recombinant lectin;
4. Deduction of the possible physiological role of this lectin for plant.

References

Aarnoudse, C.A., J.J.G. Vallejo, E. Saeland, and Y. Kooyk (2006). Recognition of tumor glycans by antigen presenting cells. Current Opinion in Immunology, 18:105-11.

Al Atalah, B., G. Smagghe, and E.J.M. van Damme (2014). Orysata, a jacalin-related lectin from rice, could protect plants against biting-chewing and piercing-sucking insects. Plant Science, 221-222:21-28.

Barkai-Golan, R., D. Mirelman, and N. Sharon (1978). Studies on growth inhibition by lectins of Penicillia and Aspergilli. Archives of Microbiology, 116:119–121.

Boyd, W.C., and E. Shapleigh (1954). Specific precipitating activity of plant agglutinins (Lectins). Science, 119:419.

Callow, J.A. (1977). Recognition, resistance and the role of plant lectins in host-parasite interactions. Advances in Botanical Research, 4:1-49.

Cammue, B.P., B. Peeters., and W.J. Peumans (1986). A new lectin from tulip (Tulipa) bulbs. Planta, 169:583-588.

Camus, M.L. (1899). Recherches experimentales sur une agglutinine produite par la glande de l'albumen chez l' *Helix pomatia*. Comptes rendus hebdomadaires des séances de l'Académie des sciences. 129:233-234.

Cândido, E.S., M.F. Pinto, P.B. Pelegrini, T.B. Lima, O.N. Silva, R. Pogue, M.F. Grossi-de-Sá, and O.L. Franco (2011). Plant storage proteins with antimicrobial activity: novel insights into plant defense mechanisms. FASEB Journal, 25:3290-3305.

Carter, J.E.3rd, O. Odumosu, and W.H. Langridge (2010). Expression of a ricin toxin B subunit: insulin fusion protein in edible plant tissues. Molecular Biotechnology, 44:90–100.

Chrispeels, M.J., and N.V. Raikhel (1991). Lectins, lectin genes, and their role in

plant defense. Plant Cell, 3:1-9.
Claes, B., R. Dekeyser, R. Villarroel, M. Van den Bulcke, G. Bauw, M. Van Montagu, and A. Caplan (1990). Characterization of a rice gene showing organ-specific expression in response to salt stress and drought. Plant Cell, 2:19-27.
Coltri, K.C., L.L. Oliveira, C.F. Pinzan, P.E. Vendruscolo, R. Martinez, M. H. Goldman, A. Panunto-Castelo, and M.C. Roque-Barreira (2008). Therapeutic administration of KM+ lectin protects mice against *Paracoccidioides brasiliensis* infection via Interleukin-12 production in a Toll-like receptor 2-dependent mechanism. The American Journal of Pathology, 173:423-432.
De Hoff, P.L., L.M. Brill, and A.M. Hirsch (2009). Plant lectins: the ties that bind in root symbiosis and plant defense. Molecular Genetics and Genomics, 282:1-15.
De Schutter, K., and E.J. Van Damme (2015). Protein-carbohydrate interactions as part of plant defense and animal immunity. Molecules, 20:9029-9053.
Denda-Nagai, K., S. Aida, K. Saba, K. Suzuki, S. Moriyama, S. Oo-Puthinan, M. Tisuiji, A. Morikawa, Y. Kumamoto, D. Sugiura, A. Kudo, Y. Akimoto, H. Kawakami, N.V. Bovin, and T. Irimura (2010). Distribution and function of macrophage galactose-type C-type lectin 2 (MGL2/CD301b): efficient uptake and presentation of glycosylated antigens by dendritic cells. The Journal of Biological Chemistry, 285:19193-19204.
Díaz, C.L., T.J.J. Logman., H.C. Stam, and J.W. Kijne (1995). Sugar-binding activity of pea lectin expressed in white clover root hairs. Plant Physiology, 109:1167-1177.
Ehrlich, P. (1899). Croonian Lecture on immunity with special reference to cell life. Proceedings of the Royal Society (London). 66: 424-448.
Elfstrand, M. (1898). About blood cell agglutinating proteins (Kobert R. Eds.). In: Görberdorfer Publications. Germany. pp 1-159.
Etzler, M. E. (1986). Distribution and function of plant lectins. In: Liener, I.E., N. Sharon, and I.J. Goldstein (Eds.). The lectins: properties, functions and applications in biology and medicine. Academic Press, Orlando, FL, pp. 371-435.
Fujimoto, Z., H. Tateno, and J. Hirabayashi (2014). Lectin structures: classification based on the 3-D structures. Methods in Molecular Biology, 1200:579-606.
Ghazarian, H., B. Idoni, and S.B. Oppenheimer (2011). A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. Acta Histochemica, 113:236-247.
Goldstein, I.J., R.C. Hughes, M. Monsigny, T. Osawa, and N. Sharon (1980). What

should be called a lectin? Nature, 285:66.

Gorelik, E., U. Galili, and A. Raz (2001). On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. Cancer Metastasis Review, 20:245-277.

Guan, Y., S. Ramalingam, D. Nagegowda, P.W.J. Taylor, and M.L. Chye (2008). *Brassica juncea* chitinase BjCHI1 inhibits growth of fungal phytopathogens and agglutinates Gram-negative bacteria. Journal of Experimental Botany, 59: 3475-3484.

Hellin, H. (1891). Study of toxic protein from *Abrus precatorius*. Doctoral thesis, University of Dorpat. Estonia.

Howard, I.K., H.J. Sage, and C.B. Horton. (1972). Studies on the appearance and location of hemagglutinins from a common lentil during the life cycle of the plant. Archives of Biochemistry and Biophysics, 149:323-326.

Janzen, D.H., H.B. Juster, and I.E. Liener (1976). Insecticidal action of the phytohemagglutinin in black beans on a bruchid beetle. Science, 192: 795-796.

Kilpatrick, D.C. (1999). Mechanisms and assessment of lectin-mediated mitogenesis. Molecular Biotechnology, 11:55-65.

Landsteiner, K., and L. Raubitschek (1908). Observations about hemolysis and hemagglutination. Zentralblatt für Bakteriologie und Parasitenkunde, Infektionskrankheiten und Hygiene, 45:660-667.

Lannoo, N., and E.J. Van Damme. (2010). Nucleocytoplasmic plant lectins. Biochimica et Biophysica Acta, 1800:190-201.

Macedo, M.L.R., C.F.R. Oliveira, and C.T. Oliveira (2015). Insecticidal Activity of Plant Lectins and Potential Application in Crop Protection. Molecules, 20:2014-2033.

Melander, M., I. Ahman, I. Kamnert, A.C. Strömdahl (2003). Pea lectin expressed transgenically in oilseed rape reduces growth rate of pollen beetle larvae. Transgenic Research, 12:555-567.

Moons, A., E. Prinsen, G. Bauw, and M. van Montagu (1997). Antagonistic effects of abscisic acid and jasmonates on salt-stress-inducible transcripts in rice shoots. Plant Cell, 9:2243-2259.

Nagadhara, D., S. Ramesh, I.C. Pasalu, Y.K. Rao, N.P. Sarma, V.D. Reddy, and K.V. Rao (2004). Transgenic rice plants expressing the snowdrop lectin gene (gna) exhibit high-level resistance to the whitebacked planthopper (*Sogatella furcifera*). Theoretical and Applied Genetics, 109:1399-1405.

Peumans, W.J., and H.M. Stinissen (1983). Gramineae Lectins: Occurrence, Molecular Biology and Physiological Function. In: Goldstein, I. J. and Etzler, M. E. (Eds.), Chemical Taxonomy, Molecular Biology, and Function of Plant Lectins. New York: Alan R. Liss, pp. 99-116.

Peumans, W.J., and E.J. Van Damme (1995). Lectins as plant defense proteins. Plant Physiology, 109:347-352.

Rogerio, A.P., C.R. Cardoso, C. Fontanari, M.A. Souza, S.R. Afonso-Cardoso, E.V. Silva, N.S. Koyama, F.L. Basei, E.G. Soares, J.B. Calixto, S.R. Stowell, M. Dias-Baruffi, and L.H. Faccioli (2007). Anti-asthmatic potential of a D-galactose-binding lectin from *Synadenium carinatum* latex. Glycobiology, 17:795-804.

Rüdiger, H., and H.J. Gabius (2001). Plant lectins: occurrence, biochemistry, functions and applications. Glycoconjugate Journal, 18:589-613.

Sadeghi, A., G. Smagghe, S. Broeders, J.P. Hernalsteens, H. De Greve, W.J. Peumans, and E. J. Van Damme (2003). Ectopically expressed leaf and bulb lectins from garlic (*Allium sativum* L.) protein transgenic tobacco plants against cotton leafworm (*Spodoptera littoralis*). Transgenic Research, 17:9-18.

Sharon, N., and H. Lis (1987). A century of lectin research (1888-1988). Trends in Biochemical Sciences, 12:488-491.

Sharon, N., and H. Lis (1990). Legume lectins - A large family of homologous proteins. FASEB Journal, 4:3198-3208.

Sharon, N., and H. Lis (2004). History of lectins: from hemagglutinins to biological recognition molecules. Glycobiology, 14:53R-62R.

Sharon, N., H. Lis, and R. Lotan (1974). On the structural diversity of lectins. Colloque International du C.N.R.S, 221:693-709.

Souza, M.A., F.C. Carvalho, L.P. Ruas, R. Ricci-Azevedo, and M.C. Roque-Barreira (2013). The immunomodulatory effect of plant lectins: a review with emphasis on ArtinM properties. Glycoconjugate Journal, 30:641-657.

Stillmark, P.H. (1888). Ricin, a toxic ferment obtained from the seeds of *Ricinus communis* L. and some other Euphorbiaceae species. Doctoral thesis, University of Dorpat. Estonia.

Tedder, T.F., D.A. Steeber, A. Chen, and P. Engel (1995). The selectins: vascular adhesion molecules. FASEB Journal, 9:866-73.

Unitt, J., and D. Hornigold (2011). Plant lectins are novel Toll-like receptor agonists. Biochemical Pharmacology, 81:1324-1328.

Van Damme, E.J., N. Lannoo, and W.J. Peumans. (2008). Plant lectins. Advances in Botanical Research, 48:107-209.

Van Damme, E.J., W.J. Peumans, A. Barre, and P. Rougé (1998a). Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. Critical Reviews in Plant Sciences, 17:575-692.

Van Damme, E.J., W.J. Peumans, A. Pusztai, and S. Bardocz (1998b). Handbook of Plant Lectins: Properties and Biomedical Applications, Chichester, John Wiley and Sons. 445 p.

Van Damme, E.J., A. Barre, P. Rougé, and W.J. Peumans (2004). Cytoplasmic/nuclear plant lectins: a new story. Trends Plant Science, 9:484-489.

Van Damme E.J., N. Lannoo, E. Fouquaert, and W.J. Peumans (2003). The identification of inducible cytoplasmic/nuclear carbohydrate-binding proteins urges to develop novel concepts about the role of plant lectins. Glycoconjugate Journal, 20:449-460.

Van Deenen, N., D. Prüfer, and C. Schulze Gronover (2011). A latex lectin from *Euphorbia trigona* is a potent inhibitor of fungal growth. Biologia Plantarum, 55:335-339.

Vandenborre, G., G. Smagghe, and E.J. Van Damme (2011). Plant lectins as defense proteins against phytophagous insects. Phytochemistry, 72:1538-1550.

Wang, Q., Y. Zhang, Y. Hou, P. Wang, S. Zhou, X. Ma, and N. Zhang (2012). Purification, characterization of a CkChn134 protein from *Cynanchum komarovii* seeds and synergistic effect with CkTLP against *Verticillium dahliae*. Protein Science, 21:865-875.

Watkins, W.M., and W.T.J. Morgan (1952). Neutralization of the anti-H agglutinin in eel serum by simple sugars. Nature, 169:825-826.

Weber, E., and D. Neumann (1980). Protein bodies, storage organelles in plant seeds. Biochemie und Physiologie der Pflanzen, 175:279-306.

Zhang, W., W.J. Peumans, A. Barre, C. Houles Astoul, P. Rovira, P. Rougé, P. Proost, P. Truffa-Bachi, A.A. Jalali, and E.J. Van Damme (2000). Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. Planta, 210:970-978.

Zapata, J.M., V. Martínez-garcía, and S. Lefebvre (2007). Phylogeny of the TRAF/MATH Domain, Advances in Experimental Medicine and Biology, 597:1–24.

CHAPTER 1

Article published in International Journal of Biological Molecules 94(2017) 508-514 http://dx.doi.org/10.1016/j.ijbiomac.2016.10.061.

A new TRAF-like protein from *B. oleracea* ssp. *botrytis* with lectinic activity and its effect on macrophages

Christiane E. M. Duarte[a,¶], Monise V. Abranches[b,¶], Patrick F. Silva[a], Sérgio O. de Paula[a], Silvia A. Cardoso[c], Leandro L. Oliveira[a,*]

[a] Departamento de Biologia Geral, Universidade Federal de Viçosa, 36570-900 Viçosa, MG, Brasil.

[b] Departamento de Nutrição e Saúde, Universidade Federal de Viçosa, 38810-000 Rio Paranaíba, MG, Brasil.

[c] Departamento de Medicina e Enfermagem, Universidade Federal de Viçosa, 36570-900 Viçosa, MG, Brasil.

¶These authors contributed equally to this work

*Corresponding author: Leandro Licursi de Oliveira, Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa - 36570-900, Brazil. Phone: +55 31 3899-3142. E-mail: leandro.licursi@ufv.br

Abstract

Lectins are involved in a wide range of biological mechanisms, like immunomodulatory agent able to activate the innate immunity. In this study, we purified and characterized a new lectin from cauliflower (*Brassica oleracea* ssp. *botrytis* - BOL) by three sequential chromatographic steps and confirmed the purity by SDS-PAGE. Additionally, we evaluated the role of the lectin in innate immunity by a phagocytosis assay, production of H_2O_2 and NO. BOL was characterized like a non-glycosylated protein that showed a molecular mass of ~34 kDa in SDS-PAGE. Its N-terminal sequence (ETRAFREERPSSKIVTIAG) did not reveal any similarity to the other lectins; nevertheless, it showed 100% homology to a putative TRAF-*like* protein from *Brassica rapa* and *Brassica napus*. This is a first report of the TRAF-protein with lectinic activity. The BOL retained its complete hemagglutination activity from 4ºC up to 60ºC, with stability being more apparent between pH 7.0 and 8.0. Moreover, the lectin was able to stimulate phagocytosis and induce the production of H_2O_2 and NO. Therefore, BOL can be explored as an immunomodulatory agent by being able to activate the innate immunity and favor antigen removal.

Keywords: agglutinin, lectin, innate immunity

1. Introduction

Brassicaceae is one of the major groups of the plant kingdom, composed of 348 genera and more than 3,700 species, distributed worldwide [1]. *Brassica oleracea* is a very morphologically diverse species, including the common heading cabbage (*B. oleracea* ssp. *capitata* L.), cauliflower (*B. oleracea* ssp. *botrytis* L.), broccoli (*B. oleracea* ssp. *italica* L.), kale and collards (*B. oleracea* ssp. *acephala*), kohlrabi (*B. oleracea* ssp. *gongylodes* L.), Chinese kale (*B. oleracea* ssp. *alboglabra*), and Brussels sprouts (*B. oleracea* ssp. *gemmifera* DC) [2]. Brassica species form an important human food crop plant with great economic value as vegetables and as sources of edible and industrial oil, animal fodder, and green manure [3].

Plants are a rich source of lectins predominantly isolated from seeds, which comprise 10% of the total protein content in a mature seed [4]. They are also present in the vegetative tissues such as leaves, fruits, roots, tubers, rhizomes, bulbs, bark, stems, phloem sap and even nectar [5]. Lectins are a group of carbohydrate-binding proteins found in viruses, bacteria and eukaryotes, which are involved in various biological processes, such as cell–cell interaction, folding of glycoproteins, host defense, self/non-self-recognition and intracellular routing [6]. Although lectins possess several biological properties in common, they represent a diversified protein group with respect to size, composition and structure.

The use of lectins for biomedical applications has grown because of research studies that indicate their antitumor properties [7] and antimicrobial activities [8] beyond the potential use of these proteins as diagnostic markers [9]. It has been shown that lectins exert an immunostimulating action such as Concanavalin A, functioning as a mitogenic agent, enabling the study of the interaction of lectin with the lymphocyte cells *in vitro* [10]. In innate immunity, soluble lectins are able to direct the antigen elimination and assist in the phagocytic action by the macrophages and dendritic cells. These proteins enhance the immune system by opsonization, culminating in the activation of the adaptive immune response, like the mannose-binding lectin present in humans [11].

In this study, we report the first isolation and characterization of a lectin from cauliflower and evaluate its biological effects on macrophage activation. Additionally, this is the first report of a lectin with only MATH-domains.

2. Materials and methods

2.1. Biological material

The cauliflowers (*Brassica oleracea* ssp. *botrytis*) were purchased from different suppliers in Viçosa, Brazil. The Federal University of Viçosa provided goat, horse and ox erythrocytes and BALB/c mice. The adult male BALB/c mice (20-25 g) were maintained in a photoperiod (12 h light: 12 h dark) controlled ambient environment (25 °C), with free access to water and food. This study was performed in strict accordance with the Ethical Principles in Animal Research adopted by the Brazilian College of Animal Experimentation and Brazilian Society of Animal Science Laboratory. The Ethics Committee on Animal Research of the Federal University of Viçosa approved the protocol (Permit Number: 21/2012). Human blood group A, B and O erythrocytes were collected from healthy donors at the Health Center of Federal University of Viçosa in accordance with the Committee on the Ethics of Humans of the Federal University of Viçosa (Permit Number:108/2012/CEPH/wmt).

2.2. Soluble protein extraction procedures

The cauliflowers were ground and homogenized with a vegetable crusher, in phosphate-buffered saline (PBS) (pH 7.4) in the ratio of 1:1 (w/v), and set aside at 4ºC for 8 hours. The supernatant was filtered through a 0.45 µm membrane (Schleicher & Schull, German) to obtain the crude extract.

2.3. Hemagglutination activity assay

A serial two-fold dilution of the lectin (50 µL) was mixed with 25 µL of a 2% of erythrocyte suspension in microtiter U-plates. The hemagglutination titer is defined as the reciprocal of the highest dilution exhibiting hemagglutination. Specific activity is defined as the number of hemagglutination units per mg protein [12]. Hemagglutination activity was evaluated using goat, horse, ox and human A, B, and O erythrocytes.

2.4. Protein purification

The crude extracts were loaded on a HiTrap Blue HP column (0.7 cm x 2.5 cm, GE Healthcare), which had been equilibrated prior with 50 mM Tris-HCl buffer (pH

7.4) at a flow rate of 1.0 mL/min. The bound proteins were eluted with 1 M NaCl in 50 mM Tris-HCl buffer (pH 7.4). After dialysis, the sample was subjected to ion exchange chromatography on a HiTrap Capto S column (0.7 cm x 2.5 cm, GE Healthcare) equilibrated with 50 mM Tris-HCl buffer (pH 7.4) at a flow rate of 1.0 mL/min. The bound proteins with the hemagglutination activity were eluted with 100 mM NaCl in the 50 mM Tris-HCl buffer (pH 7.4). In the final "polishing" step, we used a Protein-Pak column (7.8 mm x 300 mm, Waters) equilibrated with 0.9% (w/v) NaCl at a flow rate of 0.7 mL/min. The absorbance in all the chromatographic steps was monitored at 280 nm.

2.5. SDS-PAGE

SDS-PAGE was performed in the presence or absence of 2-Mercaptoethanol using a 12% resolving gel and 5% stacking gel [13]. The gel was stained with 2% (w/v) Coomassie Brilliant Blue R-250. The Protein Marker 6.5 -200 kDa (SERVA, Germany) was used as the standard molecular mass marker.

2.6. Protein concentration

The protein concentration was determined using the BCA *Protein Assay* kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions, using bovine serum albumin (BSA) as the standard.

2.7. Determination of protein glycosylation

In order to determine if BOL is a glycoprotein, a bioinformatics analysis was undertaken using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc 4.0 server (http://www.cbs.dtu.dk/services/NetOGlyc/) for the presence of predicted N-glycosylation sites and O-glycosylation sites [14], respectively.

2.8. N-terminal Sequencing

To determine N-terminal amino acid sequence, purified protein was separated by 12% SDS-PAGE and Electroblotted at 100 mA for 1 h on to ProBlott membranes (Applied Biosystems, USA) then stained with 0.1% Coomassie for 30 seconds, destained with 50% methanol, washed with distilled water and dried overnight. The desired fragments were excised and sequenced. Automatic Edman degradation

analyses were performed on the protein sequencer model PPSQ-33A (Shimadzu, Japan).

2.9. Mass spectrometry analysis and protein sequencing by tandem mass spectrometry

Briefly, the BOL band was cleaned of the SDS-PAGE gel, destained with 50% acetonitrile/25 mM ammonium bicarbonate, in-gel reduced with dithiothreitol (DTT) 65 mM for 30 min at 56ºC and alkylated with iodoacetamide 200 mM for 30 min at room temperature. It was then dried with acetonitrile followed by SpeedVac™. Samples were digested using 50 µL of 2.5 µg/ml trypsin (Sigma) in 10% acetonitrile/40 mM ammonium bicarbonate pH 8 at 37ºC overnight. Peptides were extracted with 50% acetonitrile/5% acid formic, dried in SpeedVac™ and redissolved in 8 µL 0.1% formic acid. The samples were desalted using Zip Tip C18 (Sigma). The sample matrix used was Universal MALDI-Matrix (Sigma). Mass spectra were acquired in the reflector ion mode in the *m/z* range of 640–3240 using an Ultraflex III MALDI-TOF/TOF mass spectrometer controlled by flexAnalysis software v. 2.0 (Bruker Daltonics). The instrument was equipped with a smartbeam laser (Bruker Daltonik), and the acquisition laser power was optimized using the PS calibration mixture before collection of the sample data. The peptide masses were sought against the NCBI database employing Mascot (in-house MASCOT-server) for protein identification.

2.10. Inhibition of hemagglutination

The hemagglutination inhibition tests used various 400 mM carbohydrate solutions (D-glucose, D-galactose, D-arabinose, D-xylose, N-acetyl glucosamine, D-fructose, D-mannose, D-ribose, melibiose, maltose, D-lactose, D-cellobiose, D-trehalose, saccharose and D-raffinose) and glycoproteins at a concentration of 0.5 mg/mL (asialofetuin, fetuin and casein) were performed in a manner analogous to the hemagglutination test. A serial two-fold dilution of each sugar sample was prepared in PBS. All the dilutions were mixed with an equal volume (25 µL) of the lectin solution with one hemagglutination unit. The mixture was allowed to stand for 30 minutes at room temperature and then mixed with 25 µL of a 2% goat erythrocyte suspension. The minimum concentration of the sugar which completely inhibited one hemagglutination unit of lectin was calculated [12].

2.11. Glycoproteins proteolysis

Fetuin and asialofetuin (1 mg/ml) were digested with 50 µg/ml proteinase-K (Promega, USA) at a 1:1 (w/w) enzyme is to substrate ratio in 50 mM Tris-HCl (pH 8.0), 10 mM $CaCl_2$ at 45ºC overnight. The complete digestion was confirmed by 12% SDS-PAGE. Then, the inhibition of the lectin-induced hemagglutination was tested using digested and non-digested glycoproteins.

2.12. Effects of temperature, pH and divalent cations on lectin activity

Aliquots of lectin were incubated at different temperatures (4ºC to 100ºC) for 30 minutes and cooled in ice. The hemagglutination activity of the aliquots was tested. The pH stability of the lectin was measured by dialyzing the lectin aliquots against the following buffers for 6 h at 4ºC: 100 mM glycine buffer (pH 2.0 and 3.0), 20 mM acetate buffer (pH 4.0 and 5.0), 100 mM phosphate buffer (pH 6.0 and 7.0), and 100 mM glycine–NaOH buffer (pH 10.0, 11.0 and 12.0). The pH of the lectin solution was adjusted to 7.0 by the addition of 0.1 N HCl or 0.1 N NaOH before the hemagglutination activity was determined. To determine the metal ion dependence, the protein was dialyzed against 100 mM Tris-HCl, 10 mM EDTA at pH 7.4 for 12 hours. Following this period, the lectin was dialyzed once again, but this time against 100 mM Tris-HCl at pH 7.4, followed by the hemagglutination assay. Additionally, the dialyzed protein fractions were dialyzed against 50 mM $CaCl_2$, 50 mM $MgCl_2$, 50 mM $MnCl_2$ or 50 mM $ZnCl_2$, followed by the hemagglutination assay.

2.13. Phagocytic activity of the peritoneal macrophages

Macrophages from the peritoneal cavity of the BALB/c mice were suspended with RPMI culture medium (Gibco, USA), supplemented with 10% fetal bovine serum, 100 units penicillin /mL and 100 mg streptomycin /mL. A 200 µL aliquot of this cell suspension (10^5 cells/100 µL/well) was seeded into a well of a 6-well plate and covered with a coverslip. This was followed by incubation for 2 hours at 37ºC in a humidified atmosphere of 5% CO_2. Different concentrations of the lectin in 200 µL of complete RPMI medium were then added to the wells followed by incubation for 30 minutes. After that, a *Pichia pastoris* (5 x 10^5 cells/well) suspension was added and the plates were incubated for 2 hours. The supernatant was removed and 400 µL of 10% formaldehyde in PBS was added. The coverslips were stained with HEMA 3

Panoptic dye (Renylab, Brazil) and analyzed with a light field optical microscope (Olympus, Japan).

2.14. NO production by peritoneal macrophage assay

Macrophage from the BALB/c mice peritoneal cavity were washed and resuspended in the RPMI culture medium supplemented with 10% fetal bovine serum, 100 units penicillin/mL and 100 mg streptomycin /mL. The cells were seeded in a 96-well culture plate (2 x 10^5 cells/well) and incubated at 37 °C in a humidified atmosphere with 5% CO_2 for 2 hours. The cells were stimulated with different concentrations of lectin or (2.5 mg/mL) Zymosan (positive control), followed by incubation for 48 hours. The supernatant was collected and the amount of nitric oxide in the culture medium was determined by the colorimetric method [15].

2.15. H_2O_2 production by the peritoneal macrophage assay

Macrophages from the BALB/c mice peritoneal cavity were washed with PBS and suspended in phenol red buffer (140 mM NaCl, 10 mM potassium phosphate, 5.5 mM dextrose, 0.56 mM phenol red and 0.01 mg/mL peroxidase type II, pH 7.0). The cell aliquots (100 µL) were seeded in a 96-well culture plate and incubated with different concentrations of lectin or (2.5 mg/mL) Zymosan for 1 hour at 37ºC in a humidified atmosphere with 5% CO_2. The reaction was stopped by the addition of 10 µL/well of 1 M NaOH. The H_2O_2 present in the medium was determined by the colorimetric method [16].

2.16. Statistical Analyses

The statistical significance was analyzed using the analysis of variance (ANOVA), followed by the Dunnett test, using the GraphPad Prism® version 5.0 software. Differences with $p < 0.05$ were considered statistically significant. All experiments were performed in triplicate.

3. Results

3.1. Protein Purification

Purification of the cauliflower lectin involved the initial extraction in PBS (pH 7.4) and three-step chromatography including affinity chromatography on the HiTrap Blue

HP column, ion-exchange chromatography on the Mono S column, and gel filtration on the Protein-Pak column. Fractionation of the crude extract using HiTrap Blue HP revealed the presence of a slightly smaller adsorbed fraction, designated as PI (Fig. 1A). This fraction, with hemagglutination activity, was subsequently applied on the Mono S column, by means of which a fraction designated as PII (Fig. 1B) was obtained. The adsorbed fraction with hemagglutination activity was resolved into a large peak (PIII) by gel filtration on the Protein-Pak column (Fig. 1C). The purified lectin, represented by PIII, appeared as a single band with an apparent molecular mass of 34 kDa on SDS-PAGE (Fig. 1D) and a 36.8 kDa on gel filtration (Figure 1E), these results reinforce the observation that lectin is a monomeric protein. A gradually enriched lectin was purified and then designated as *Brassica oleracea* ssp. *botrytis* lectin (BOL). An almost 139-fold purification and a recovery of 12% were achieved through the purification process (Table I).

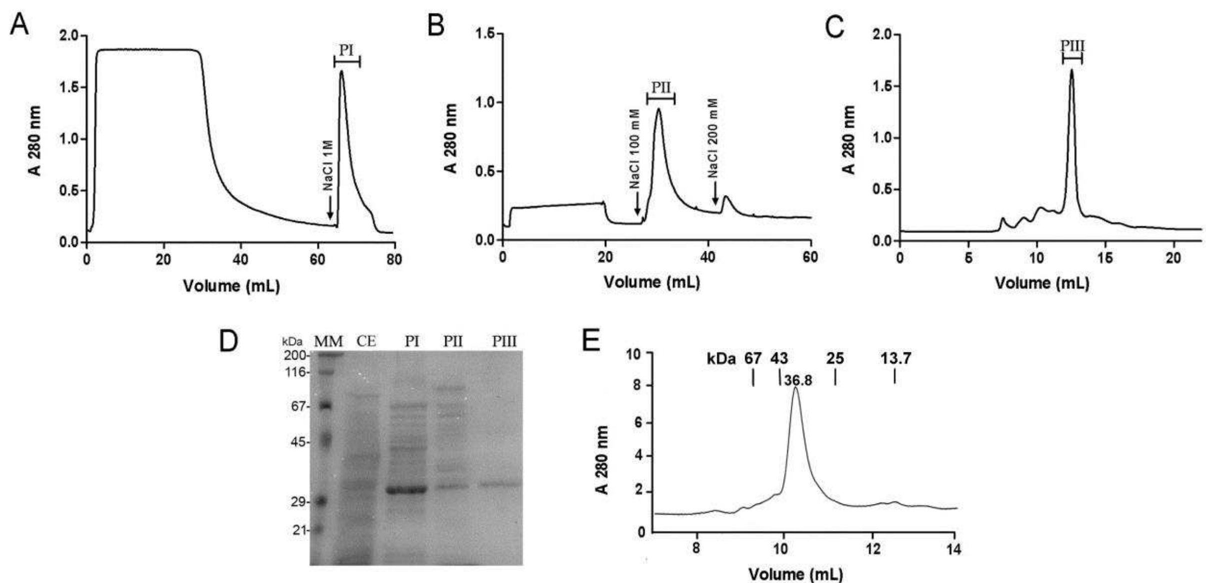


FIGURE 1. Purification of cauliflower lectin. (A) Affinity chromatography of crude extract of cauliflower on a HiTrap Blue HP column. The peak labeled PI exhibited HA. (B) Ion exchange chromatography of fraction PI on a HiTrap Capto S column. The peak labeled PII exhibited HA. (C) Molecular size exclusion chromatography of fraction PII on a Protein-Pak column. The peak labeled PIII exhibited HA. All the elutions were monitored at 280 nm. (D) The SDS-PAGE of the fractions obtained in the chromatography steps. MM: molecular weight marker, CE: crude extract, PI-fraction obtained by affinity chromatography, PII-fraction obtained by ion exchange chromatography, PIII-fraction obtained by gel filtration. (E) Estimation of molecular weight by gel filtration, using BSA, ovalbumin, chymotrypsinogen A and ribonuclease A as calibration standard.

Table I. Specific hemagglutination activities and chromatographic fraction yields obtained at different stages of lectin purification

Purification steps	Total Protein (mg)	Total Activity (HA)	Specific Activity (HA/mg)	Purification fold	Recovery (%)
Crude extract	10642.50	180000	17	1.0	100
HiTrap Blue HP	395.28	43200	109	6.4	24
HiTrap Capto S	80.28	28800	359	21.1	16
Protein-Pak	8.91	21120	2370	139.4	12

HA – Hemagglutination Activity Unit, corresponds to the minimum quantity of protein capable of inducing agglutination; HA/mg corresponds to the amount of hemagglutination units per milligram of protein.

3.2. Properties of purified lectin

The physical and biochemical properties of the lectin were investigated. BOL migrated as a single band on the SDS–PAGE under reducing and non reducing conditions (Fig. 2A). Taken together with gel filtration results we can conclude that BOL is a monomeric protein. The *in silico* prediction of possible sites of glycosylation show that BOL has no N-glycosylation sites and a minimal probability of being O-glycosylated, so BOL is a non-glycosylated protein (Fig. 2B). The N-terminal amino acid sequence of BOL was obtained by the automated Edman degradation. The first 19 amino acid residues were determined (ETRAFREERPSSKIVTIAG) which showed significant homology by the BLAST to predict, and uncharacterized proteins from *Brassica rapa* (XP_009111696.1) and *Brassica napus* (CDY19775.1 and CDX87054.1) with 100% identity with a putative TRAF-*like* protein.

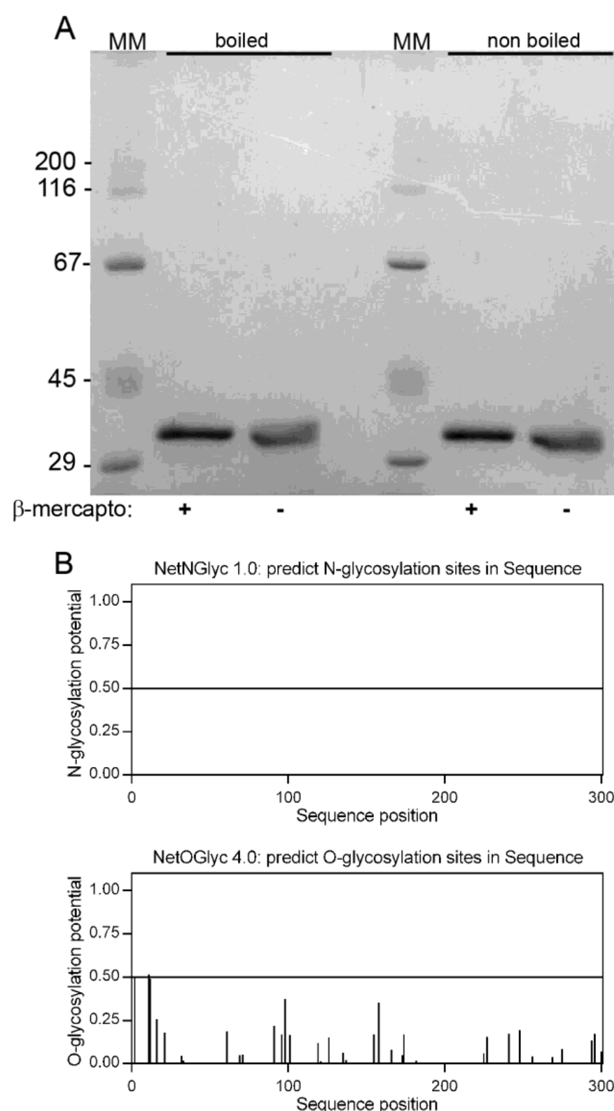


FIGURE 2. Properties of the purified BOL. BOL is a monomeric lectin (A) SDS–PAGE under the reducing and non-reducing conditions. Lanes 1 and 4, molecular weight markers, Lanes 2 and 3: samples boiled. Lane 2, reducing condition and Lane 3 non-reducing condition. Lanes 5 and 6: not boiled samples. Lane 5 reducing condition, Lane 6 non-reducing condition. BOL is a non-glycosylated protein (B) *In silico* prediction of possible sites of glycosylation using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc 4.0 server (http://www.cbs.dtu.dk/services/NetOGlyc/).

The purified protein was digested by trypsin and the resultant peptides were analyzed by mass spectrometry (MALDI-TOF/TOF). Fig. 3 shows the monoisotopic masses of the five peptides identified, which were used to identify the homologous proteins in the NCBI database through the MASCOT server. Matching the same set of peptides aligned with the homologous sequences of the 39.5 kDa putative TRAF-*like* protein protein of *Brassica napus* (CDX87054.1) and *Brassica rapa* (XP_009111696.1) was achieved with the 309 MASCOT score.

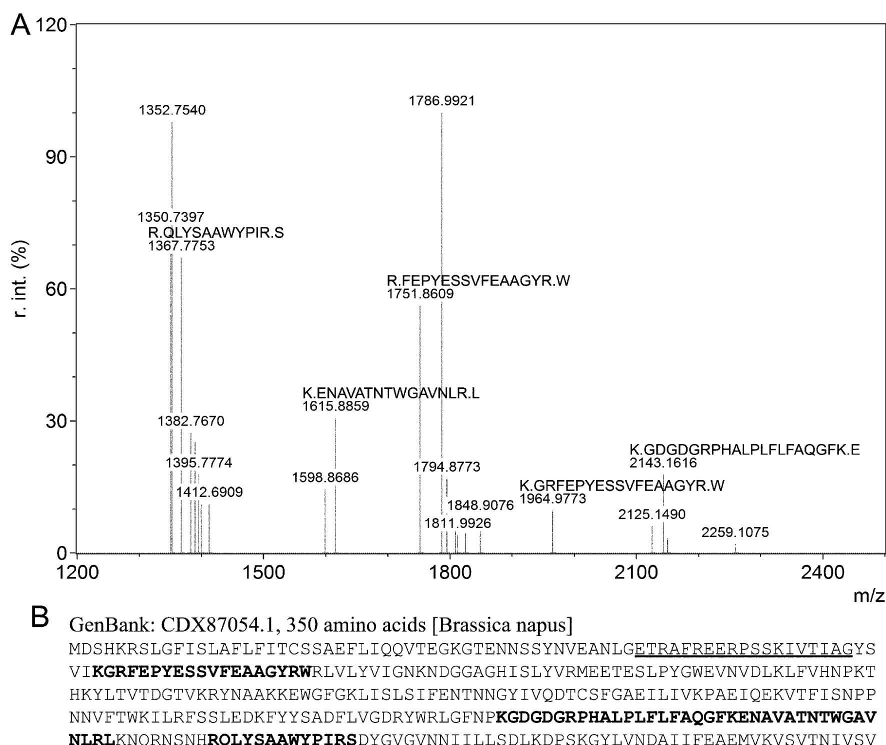


FIGURE 3. Mass spectrometric analysis of BOL. A) Mass fingerprint obtained from tryptic digested BOL was analyzed by MALDI-TOF/TOF mass spectrometer scanning from 640 to 3240 amu in the positive ion mode for detection of protonated peptides. Each tryptic peptide was subjected to LIFT dissociation to produce a fragment ion pattern and the amino acid sequence was deduced. The tryptic peptides are listed above the ion pattern. B) Amino acid sequence of putative protein of Brassica napus (CDX87054.1). N-terminal (residues 1–19) determined by Edman degradation (underline) and tryptic peptides by mass spectrometry (bold).

3.3. Carbohydrate specificity of the purified lectin

The blood specificity of BOL was determined by use of erythrocytes from different species (goat, horse, ox) and humans from the ABO system. The lectin showed more selective for horse and goat erythrocytes than others (Fig. 4A). The hemagglutination activity of the purified cauliflower lectin was not observed to be inhibited by any of the simple sugars tested at 400 mM; however, it was inhibited by the glycoproteins, asialofetuin > fetuin > ferritin > casein, but not for ovalbumin (Table II). To determine if the interaction of the BOL-glycoproteins was mediated by a carbohydrate-protein or protein-protein interaction, fetuin and asialofetuin complete proteolysis were performed (Data not show) and the hemagglutination activity of the BOL was noted to continue to be inhibited by oligosaccharides (Table II).

Table II. Effects of the various carbohydrates and glycoproteins on the hemagglutination induced by the *B. oleracea* lectin

Inhibitor	mM
Simple sugars[a]	ND
Asialofetuin	0.12
Asialofetuin + proteinase K[b]	0.24
Fetuin	0.23
Fetuin + proteinase K[b]	0.46
Casein	1.30
Ferritin	0.57
Ovalbumin	ND

[a] Lactose, galactose, arabinose, melibiose, xylose, cellobiose, N-acetyl glucosamine, fructose, glucose, maltose, mannose, saccharose, ribose, trehalose, raffinose were non-inhibitory at 400mM concentration.
[b] Glycoprotein (1mg/mL) was digested with 50 µg/ml proteinase-K (overnight, 45ºC)
ND- inhibition non-detected

3.4. Effect of temperature and pH

The thermal stability of BOL was determined in the temperature range between 4 and 100ºC. The results indicated that BOL was stable between 4 and 60ºC. The lectin was totally inactivated when incubated at 70ºC for 30 min (Fig. 4B). The pH sensitivity profile of the lectin is shown in Fig. 4C, in which the stability was more apparent between pH 7.0 and 8.0. The hemagglutination activity of the native lectin was not affected either by the sequential dialysis (with EDTA followed by Tris-HCl) or by the addition of Ca^{2+} and Zn^{2+} to the dialyzed lectin. The lectinic activity was slightly inhibited in the presence of Mn^{2+} and increased after the addition of Mg^{2+} (Fig. 4D).

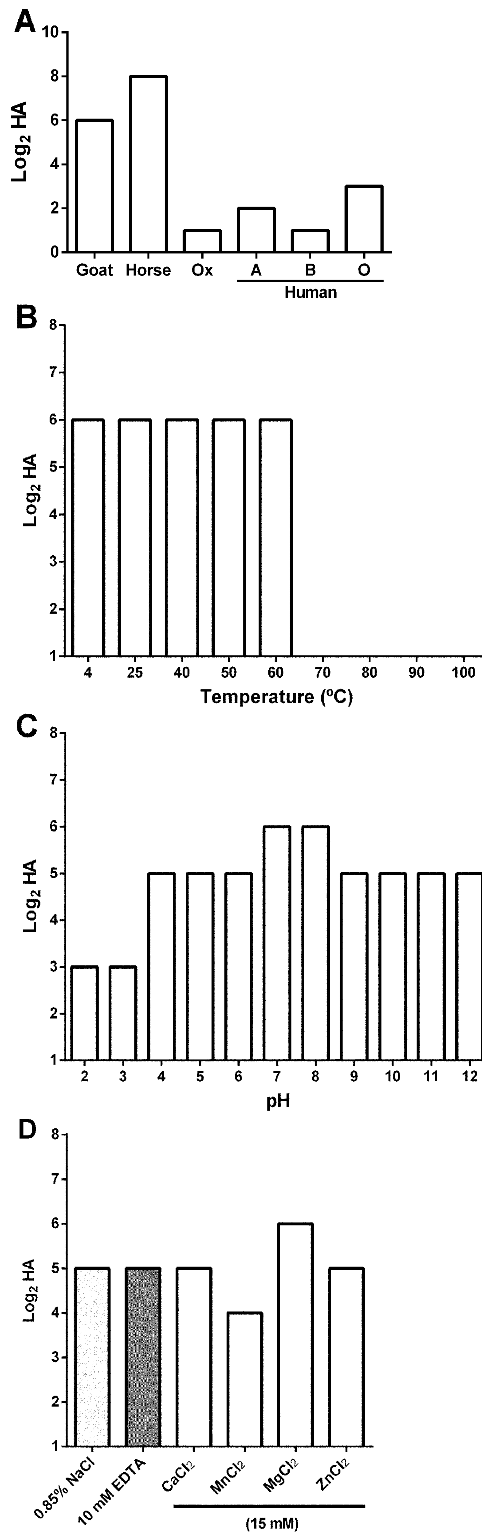


FIGURE 4. Physicochemical characterization of BOL. (A) Specificity of agglutination activity, red blood cells of goat, horse, ox and human, A, B and O groups were tested. (B) Thermal stability of BOL. The lectin was incubated at an elevated temperature (4–100ºC). (C) pH stability of BOL. The lectin was incubated with buffers ranging from pH 2.0 to 12.0. (D) Influence of the divalent cations on BOL. After treatment with a chelating agent, the lectin was incubated with the indicated various divalent cations. The bars represented the HA of BOL.

3.5. Cauliflower lectin activates macrophages and promotes phagocytosis

In order to verify whether BOL was capable of acting as an immunostimulator, we evaluated the induction of macrophage activation promoted by the BOL. The effect of cauliflower lectin on the phagocytic activity of the peritoneal macrophage in engulfing yeast cells is shown in Figs. 5A and 5B. The phagocytosis of the yeast cells is increased by two-fold ($p < 0.05$) when compared with the control. The BOL induced a significant increase in the production of inflammatory mediators compared with the untreated cells. The results of the nitrite and H_2O_2 production are shown in Figs. 5C and 5D. Taken together we can see that the macrophages were activated by lectin, phagocytizes more, better and with greater capacity microbicide.

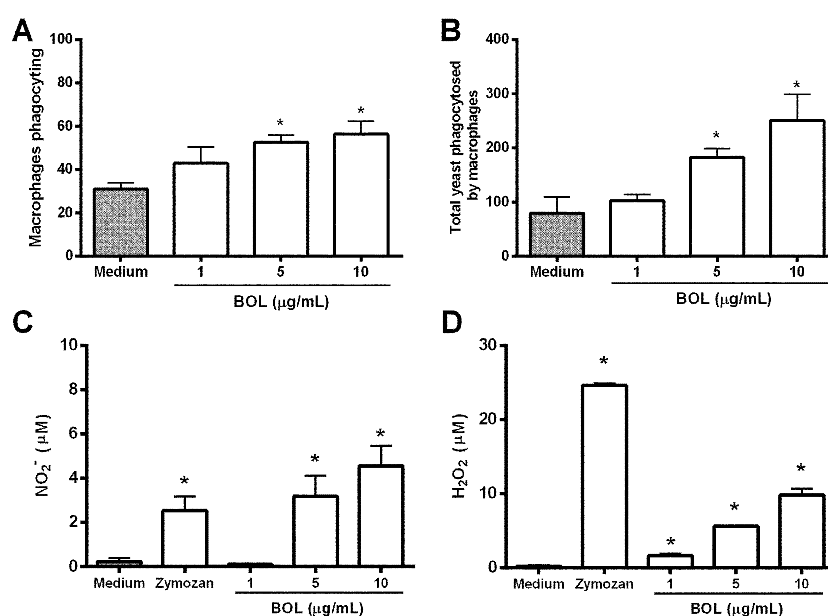


FIGURE 5. Macrophage Activation. Peritoneal macrophages of BALB/c mice were treated with the lectin obtained from cauliflower. (A) The number of macrophages that exhibited phagocytosis in each of the 200 cells analyzed. (B) Phagocytic index in each of the 200 macrophages analyzed. (C) Nitric oxide production by the macrophages. (D) Hydrogen peroxide production by the macrophages. Assays were performed in triplicate; the results represent the average ± SD of three independent experiment; $^*p < 0.05$ compared with the control group.

4. Discussion

The isolation, purification, characterization and biological applications of plant lectins have been the focal point of several studies over the last few years [17–19]. The protocol used in the purification can be distinguished into three sequential chromatography steps, enabling the recovery of 12% of the total hemagglutination

activity present in the crude extract, with 139-fold purification. BOL was found to be a non-glycosylated monomer, with a molecular mass estimated at 34 kDa protein by SDS-PAGE and 36.8 kDa by gel filtration.

BOL does not exhibit sequence similarity with the other earlier reported lectins, including the lectin isolated from broccolini [20]. According to the search results from N-terminal homology and MALDI-TOF/TOF the lectin showed 100% identity with a putative TRAF-*like* protein of *Brassica rapa* and *Brassica napus.* The TRAF family, a tumor necrosis factor of receptor-associated factors, was first identified as a group of mammalian adaptor proteins. TRAF proteins physically and functionally connect the cell surface receptors to the signaling pathways involved in the regulation of diverse cellular responses, which include activation, differentiation and survival [21]. This type of protein also was seen in plants (*Arabidopsis*, *Medicago*, *Oryza*, and *Sorghum*), lower eukaryota (*Trypanosoma*, *Dictyostelium*, *Theileria* and *Plasmodium*), and lower metazoa (*Caenorhabditis elegans*) [22]. However, until now, TRAF-like proteins with lectin activity have not been described. BOL was related as a hypothetical protein that possess only two MATH-domains without any other protein domain, the function of these kind of proteins not yet know. BOL can be the first TRAF-like protein capable of recognizing carbohydrates, and we hypothesize that the MATH domain can be a new carbohydrate-recognition domain (CRD).

The *B. oleracea* lectin was not inhibited by the mono-, di- or tri-saccharides, but complex carbohydrate structures inhibited its activity. Asialofetuin and fetuin were found to be strong inhibitors of the lectin from *B. oleracea* suggesting that the BOL binds the complex N-linked oligosaccharides. To demonstrate that the inhibition of hemagglutination activity of BOL was caused by oligosaccharide moieties, the asialofetuin and fetuin were completely hydrolyzed by proteases and the lectin-inhibition was maintained. Wright *et al*. [23], also observed the inhibition of the hemagglutination activity by the fetuin and asialofetuin, which in turn inhibits the action of the lectin obtained from *Scilla campanulata*. Although plant lectins have specificity toward monosaccharides, they show high specificity to the more complex glycans that are found in animals and humans but absent from plants [5]. Recent high performance analytical techniques (like glycan microarray analysis) demonstrated that plant lectins have a preferential binding to oligosaccharides and glycans rather than to monosaccharides. Even lectins classic as GNA, which was originally considered a mannose-specific lectin, interact only weakly with mannose

but exhibit a strong afinity to high-mannose N-glycans [24]. This property may be related to the fact that lectins are capable of recognizing the glycoconjugates present on the microorganism surface or in digestive tracts of insects and herbivorous animals and are possibly part of plant defense pathways [5, 25].

The lectin retained its whole hemagglutination activity from 4ºC up to 60ºC. The BOL activity was also maintained in wide pH variation, with stability being more evident between pH 7.0 and 8.0, while 50% activity remained at pH 4-6 and 9-12. A similar case is observed with the lectins from other plant species, for example *Glycine max* [26] and *Phaseolus coccineus* [27]. Lectins are mostly the defense proteins [6] which are known for their stability under various physicochemical conditions [28]. On the other hand, some lectins have been reported whose activities decrease above pH 9.0 [29] or below pH 5.0 [30]. The stability shown by BOL increases the applications of this protein.

Purified lectin does not need bivalent cations to reveal the hemagglutination action. Lectin activity remained unaltered even after metal ion chelation with EDTA or in the presence of Ca^{+2} and Zn^{+2} ions; however, it was affected by the Mg^{2+} and Mn^{2+} ions. The hemagglutination activity of *Inocybe umbrinella* lectin was also depressed by Mn^{2+} [31] while the activity of Con A was potentiated by the Ca^{2+} and Mn^{2+} ions [32]. Divalent cations although it does not required for the formation of heterodimers, may increase the stability of the complex formed by decreasing the dissociation rate [33]. In addition, antimicrobial activity of peptides can also be increased in the presence of divalent ions which excess may induce conformational changes in the peptide [34]. Therefore, the influence of the divalent cations in the binding of BOL to the carbohydrates could be explained in the light of the appropriate conformational recognition.

The immunomodulatory effect triggered by the cauliflower lectin was evidenced by its capacity to stimulate the phagocytosis and production of the inflammatory mediators by the peritoneal macrophages. Wong and Ng [15] reported that the banana lectin increased the NO production by the macrophages, in a dose-dependent manner. Similar results were observed with onion lectin, which induced a significant increase in the production of NO, the pro-inflammatory cytokines and phagocytic activity of the yeast cells by the activated macrophages [19]. In our study, the cauliflower lectin activated the macrophages by inducing the NO and H_2O_2 production. Unitt and Hornigold (2011)[35] reported that some plant lectins exhibit

specific patterns of stimulation of human Toll-like receptors, suggesting that the innate immune system can detect and respond to certain lectins. Working in this direction, Mariano *et al.* [36] presented a plausible mechanism of macrophage stimulation: ArtinM, a D-mannose-binding lectin, interacted with Toll-like receptor 2 and its heterodimers in a carbohydrate recognition-dependent manner, which culminated in a larger secretion of cytokines, due to the action of the NFκB nuclear transcriptional factor. Furthermore, several plant lectins exhibit immunomodulatory activities that are initiated by their interaction with the glycan moieties present on the surfaces of the immune cells. Such interactions may trigger signal transduction to produce certain cytokines and induce efficient immune responses against tumors or microbial infections [37].

5. Conclusion

This is the first report of the isolation of a lectin from *Brassica oleracea* ssp. *botrytis* (BOL). In this study we purified, characterised and evaluated the stimulatory effects of BOL, which demonstrated be able activate macrophage improve their clearance capacity. It is also described for the first time a TRAF-like protein with lectin activity, supporting the concept that the lectins are indeed multifunctional and diverse group. The new lectin isolated from cauliflower can favoring the removal of foreign agents, which is a potentially exploitable activity.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Acknowledgments

The authors thank the Núcleo de Análise de Biomoléculas, Universidade Federal de Viçosa-UFV, for the available facilities and technical assistance. This work was supported by Brazilian funding agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [474715/2013-2, 552459/2011-9] and Fundação de Amparo à Pesquisa do estado de Minas Gerais.

Author contributions: LLO and SAC designed and coordinated the study. LLO, CEMD and MVA wrote the paper. MVA, CEMD and PFS performed and analyzed the experiments. SOP and SAC provided technical assistance and contributed to the preparation of the figures and tables. All authors reviewed the results and approved the final version of the manuscript.

References

[1] S.I. Warwick, K. Mummenhoff, C. a. Sauder, M. a. Koch, I. a. Al-Shehbaz, Closing the gaps: Phylogenetic relationships in the Brassicaceae based on DNA sequence data of nuclear ribosomal ITS region, Plant Syst. Evol. 285 (2010) 1–24. doi:10.1007/s00606-010-0271-8.

[2] J. Yu, M. Zhao, X. Wang, C. Tong, S. Huang, S. Tehrim, et al., Bolbase: a comprehensive genomics database for Brassica oleracea., BMC Genomics. 14 (2013) 664. doi:10.1186/1471-2164-14-664.

[3] G. Rakow, Species Origin and Economic Importance of Brassica, in: E.C. Pua, C.J. Douglas (Eds.), Biotechnol. Agric. For., Springer-Verlag, Berlin, 2004: pp. 3–7. doi:10.1007/978-3-662-06164-0.

[4] E.J.M. Van Damme, W.J. Peumans, A. Barre, P. Rougé, Plant Lectins : A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles Plant Lectins : A Composite of Several Distinct Families of Structurally and Evolut, CRC. Crit. Rev. Plant Sci. 17 (1998) 575–692.

[5] W.J. Peumans, E.J. Van Damme, Lectins as plant defense proteins., Plant Physiol. 109 (1995) 347–352. doi:10.1104/pp.109.2.347.

[6] K. De Schutter, E. Van Damme, Protein-Carbohydrate Interactions as Part of Plant Defense and Animal Immunity, Molecules. 20 (2015) 9029–9053. doi:10.3390/molecules20059029.

[7] M. Pervin, Y. Koyama, M. Isemura, Plant Lectins in Therapeutic and Diagnostic Cancer Research, Int. J. Plant Biol. Res. 3 (2015) 1030–1035.

[8] R. Dias, L. Machado, L. Migliolo, O. Franco, Insights into Animal and Plant Lectins with Antimicrobial Activities, Molecules. 20 (2015) 519–541. doi:10.3390/molecules20010519.

[9] L. Klukova, T. Bertok, M. Petrikova, A. Sediva, D. Mislovicova, J. Katrlik, et al., Glycoprofiling as a novel tool in serological assays of systemic sclerosis: A

comparative study with three bioanalytical methods, Anal. Chim. Acta. 853 (2015) 555–562. doi:10.1016/j.aca.2014.10.029.

[10] D.C. Kilpatrick, Mechanisms and assessment of lectin-mediated mitogenesis., Mol. Biotechnol. 11 (1999) 55–65. doi:10.1007/BF02789176.

[11] H. Ghazarian, B. Idoni, S.B. Oppenheimer, A glycobiology review: carbohydrates, lectins, and implications in cancer therapeutics, Acta Histochem. 113 (2012) 236–247. doi:10.1016/j.acthis.2010.02.004.A.

[12] H. Wang, J. Gao, T.B. Ng, A new lectin with highly potent antihepatoma and antisarcoma activities from the oyster mushroom Pleurotus ostreatus., Biochem. Biophys. Res. Commun. 275 (2000) 810–816. doi:10.1006/bbrc.2000.3373.

[13] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4., Nature. 227 (1970) 680–685. doi:10.1038/227680a0.

[14] C. Steentoft, S.Y. Vakhrushev, H.J. Joshi, Y. Kong, M.B. Vester-Christensen, K.T. Schjoldager, et al.. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32(10):1478-1488. doi: 10.1038/emboj.2013.79.

[15] J.H. Wong, T.B. Ng, Isolation and characterization of a glucose/mannose-specific lectin with stimulatory effect on nitric oxide production by macrophages from the emperor banana., Int. J. Biochem. Cell Biol. 38 (2006) 234–243. doi:10.1016/j.biocel.2005.09.004.

[16] E. Pick, Y. Keisari, A simple colorimetric method for the measurement of hydrogen peroxide produced by cells in culture., J. Immunol. Methods. 38 (1980) 161–170. doi:10.1016/0022-1759(80)90340-3.

[17] G.A. Bezerra, R. Viertlmayr, T.R. Moura, P. Delatorre, B.A.M. Rocha, K.S. Do Nascimento, et al., Structural studies of an anti-inflammatory lectin from Canavalia boliviana seeds in complex with dimannosides, PLoS One. 9 (2014) 1–12. doi:10.1371/journal.pone.0097015.

[18] S.C. Gordts, M. Renders, G. Ferir, D. Huskens, E.J.M. Van Damme, W. Peumans, et al., NICTABA and UDA, two GlcNAc-binding lectins with unique antiviral activity profiles, J. Antimicrob. Chemother. 70 (2015) 1674–1685. doi:10.1093/jac/dkv034.

[19] V.K. Prasanna, Y.P. Venkatesh, Characterization of onion lectin (Allium cepa agglutinin) as an immunomodulatory protein inducing Th1-type immune response in vitro, Int. Immunopharmacol. 26 (2015) 304–313. doi:10.1016/j.intimp.2015.04.009.

[20] P. Xu, T. Zhang, X. Guo, C. Ma, X. Zhang, Purification, characterization, and biological activities of broccolini lectin, Biotechnol. Prog. 31 (2015) 736–743. doi:10.1002/btpr.2070.

[21] J.R. Bradley, J.S. Pober, Tumor necrosis factor receptor-associated factors (TRAFs)., Oncogene. 20 (2001) 6482–6491. doi:10.1038/sj.onc.1204788.

[22] J.M. Zapata, V. Martínez-García, S. Lefebvre, Phylogeny of the TRAF/MATH domain, in: H. Wu (Ed.), Adv. Exp. Med. Biol., Springer, Landes, 2007: pp. 1–24. doi:10.1007/978-0-387-70630-6-1.

[23] L.M. Wright, C.D. Reynolds, P.J. Rizkallah, A.K. Allen, E.J.M. Van Damme, M.J. Donovan, et al., Structural characterisation of the native fetuin-binding protein Scilla campanulata agglutinin: A novel two-domain lectin, FEBS Lett. 468 (2000) 19–22. doi:10.1016/S0014-5793(00)01109-1.

[24] E.J.M. Van Damme, N. Lannoo, W.J. Peumans, Chapter 3 Plant Lectins, Adv. Bot. Res. 48 (2008) 107–209. doi:10.1016/S0065-2296(08)00403-5.

[25] N. Lannoo, E.J.M. Van Damme, Lectin domains at the frontiers of plant defense, Front. Plant Sci. 5 (2014) 1–16. doi:10.3389/fpls.2014.00397.

[26] P. Lin, X. Ye, T. Ng, Purification of melibiose-binding lectins from two cultivars of Chinese black soybeans, Acta Biochim. Biophys. Sin. (Shanghai). 40 (2008) 1029–1038. doi:10.1111/j.1745-7270.2008.00488.x.

[27] J. Chen, B. Liu, N. Ji, J. Zhou, H.J. Bian, C.Y. Li, et al., A novel sialic acid-specific lectin from Phaseolus coccineus seeds with potent antineoplastic and antifungal activities, Phytomedicine. 16 (2009) 352–360. doi:10.1016/j.phymed.2008.07.003.

[28] J.M. Khan, A. Qadeer, E. Ahmad, R. Ashraf, B. Bhushan, S.K. Chaturvedi, et al., Monomeric Banana Lectin at Acidic pH Overrules Conformational Stability of Its Native Dimeric Form, PLoS One. 8 (2013) e62428. doi:10.1371/journal.pone.0062428.

[29] A.F.M. Vaz, R.M.P.B. Costa, A.M.M. a Melo, M.L. V Oliva, L. a. Santana, R. a. Silva-Lucca, et al., Biocontrol of Fusarium species by a novel lectin with low ecotoxicity isolated from Sebastiania jacobinensis, Food Chem. 119 (2010) 1507–1513. doi:10.1016/j.foodchem.2009.09.035.

[30] Q. Yan, L. Zhu, N. Kumar, Z. Jiang, L. Huang, Characterisation of a novel monomeric lectin (AML) from Astragalus membranaceus with anti-proliferative activity, Food Chem. 122 (2010) 589–595. doi:10.1016/j.foodchem.2010.03.015.

[31] J.K. Zhao, H.X. Wang, T.B. Ng, Purification and characterization of a novel lectin from the toxic wild mushroom Inocybe umbrinella, Toxicon. 53 (2009) 360–366. doi:10.1016/j.toxicon.2008.12.009.

[32] C.F. Brewer, R.D. Brown, S.H. Koenig, Metal ion binding and conformational transitions in concanavalin A: a structure-function study., J. Biomol. Struct. Dyn. 1 (1983) 961–997. doi:10.1080/07391102.1983.10507497.

[33] L. Vallar, C. Melchior, S. Plançon, H. Drobecq, G. Lippens, V. Regnault, et al., Divalent cations differentially regulate integrin alphaIIb cytoplasmic tail binding to beta3 and to calcium- and integrin-binding protein., J. Biol. Chem. 274 (1999) 17257–17266.

[34] S.G. Dashper, N.M.O. Brien-simpson, K.J. Cross, R. a Paolini, B. Hoffmann, D. V Catmull, et al., Divalent Metal Cations Increase the Activity of the Antimicrobial Peptide Kappacin, Antimicrob. Agents Chemother. 49 (2005) 2322–2328. doi:10.1128/AAC.49.6.2322.

[35] J. Unitt, D. Hornigold, Plant lectins are novel Toll-like receptor agonists, Biochem. Pharmacol. 81 (2011) 1324–1328. doi:10.1016/j.bcp.2011.03.010.

[36] V.S. Mariano, A.L. Zorzetto-Fernandes, T.A. Da Silva, L.P. Ruas, L.L. Nohara, I.C. De Almeida, et al., Recognition of TLR2 N-glycans: Critical role in ArtinM immunomodulatory activity, PLoS One. 9 (2014) 1–9. doi:10.1371/journal.pone.0098512.

[37] M. a. Souza, F.C. Carvalho, L.P. Ruas, R. Ricci-Azevedo, M.C. Roque-Barreira, The immunomodulatory effect of plant lectins: A review with emphasis on ArtinM properties, Glycoconj. J. 30 (2013) 641–657. doi:10.1007/s10719-012-9464-4.

CHAPTER 2

(Article 2)

Molecular cloning, expression and characterization of a new TRAF-like protein with lectin activity from *Brassica oleracea*

Christiane E. M. Duarte[a], Josefa M. Alamillo[b,*], Andréa D. Koehler[c], Manuel Pineda[b], Wagner C. Otoni[c], Leandro L. Oliveira[a]

[a] Departamento de Biologia Geral, Universidade Federal de Viçosa, 36570-900 Viçosa, MG, Brasil.

[b] Departamento de Fisiología Vegetal, Universidad de Córdoba, 14071, Córdoba, Espanha.

[c] Departamento de Biologia Vegetal, Universidade Federal de Viçosa, 36570-900 Viçosa, MG, Brasil.

*Corresponding author: Josefa Muñoz Alamillo, Departamento de Fisiología Vegetal, Universidad de Córdoba, 14071, Córdoba, Spain. Phone: +34 957 218 693. E-mail: bv1munaj@uco.es

Abstract

BOL is a non-glycosylated monomeric lectin extracted from *Brassica oleracea* ssp. *botrytis* that shows no sequence similarity to any other protein group, nevertheless, it showed 100% homology to a putative TRAF-like protein from *Brassica rapa* and *Brassica napus*. In order to ascertain whether the hemagglutination activity was promoted by a possible MATHd-only lectin or other contaminant, we determined the complete sequence of this protein and expressed in a heterologous system. Degenerated primers were designed from the amino acid sequence of tryptic fragments and N-terminal sequencing from the native protein. Using total RNA extracted from cauliflower seedlings, a *Bol* coding cDNA sequence of 1053 bp was

isolated. Bioinformatics tools were used to determine a promoter sequence of 1000 bp of *Bol* which revealed several key cis-regulatory elements known to be involved in various plant stresses. Furthermore, the expression of this gene was checked under biotic and abiotic stress simulating conditions. Our results demonstrate that *Bol* transcript increases in response to methyl jasmonate but decrease in response to salicylic acid. Comparative expression analysis of specific tissue revealed the highest transcript levels in leaves, as compared to stem and root tissues. Analysis of amino acid sequence and alignment with deduced homologous proteins allowed us to determine that the mature protein comprises 301 amino acids. Predicted three-dimensional structure showed that this lectin had an overall dome-like structure with two MATH-domains. The full coding sequence confirmed identity with TRAF proteins and the resulting recombinant protein retained the ability to recognize carbohydrates, thereby BOL is the first TRAF protein with lectin activity and it can show a new interaction mechanism yet unidentified in other TRAF proteins. The production of active rBOL in good yields opens up the possibility of obtaining its tridimensional structure as well as making subsequent investigations about its structure and function feasible. This is the first report of isolation, cloning and bacterial expression of a lectin with MATH domains and may be of significant interest to understand the regulatory role of this protein, as immunostimulatory agent, as well as the physiology of the plant itself.

Keywords: cauliflower, heterologous expression, lectin, MATH-domain, stress response.

Highlights

- BOL is a protein with peculiar features: a TRAF-like protein with lectin activity.
- BOL shares no significant homology with other known lectins.
- *Bol* putative promoter contains regulatory elements in response to light and stress signals.
- BOL can exhibit a new mechanism of TRAF-interaction by carbohydrate moieties.
- Recombinant BOL has structural and functional properties similar to native lectin.

Introduction

Lectins are proteins widely distributed in nature that contain at least one non-catalytic domain that recognizes and binds in a reversible way to specific glycans, either free or as part of glycoproteins or glycolipids [1]. Accordingly, lectins have the ability to agglutinate erythrocytes and other cells, to precipitate polysaccharides and other glycoproteins [2]. Plants express a wide range of lectins, with different molecular structures and carbohydrate-binding specificities [3]. In the past, the most interesting use of lectins was as tools and bioactive proteins in biological and biomedical research. However, in recent years they have received recognition as their own right as class communication proteins that are involved in interactions between plants and their environment and therefore can play an important role in the plant physiology [3]. Biological events, such as seed maturation or maintenance of seed dormancy [4], mitogenic stimulators of plant embryonic cells [4], packaging or mobilization of storage materials [5] and N_2 fixation and source for developing embryos [6,7], have been associated with lectins. Lectins are also capable of interacting with foreign organisms through glycoconjugates present on their surface or in the digestive tract and, therefore protect plants against phytopathogenic microorganisms and insects as well as herbivores [8].

Despite the structural diversity reported for lectin families [3], lectins with MATH domains hitherto have not been described. This protein domain form a β-sandwich structure containing a surface crevice responsible for protein-protein interactions, typical of Tumor necrosis factor receptor (TNFR)-associated factors (TRAFs) [9]. TRAF proteins are required for the activation of various cell surface receptors in mammals and participate in immune response signaling through interaction with receptors as Toll-like receptors (TLRs), NOD-like receptors (NLRs), and T cell receptors [10]. In plants, the family of gene codified TRAF proteins is highly expanded, with more than 50 genes identified in the genome of *Arabidopsis*. The structural similarity among mammalian and plant TRAF protein suggest that the plant counterpart also may be involved in transferring signals to intracellular signaling pathways [11]. Nevertheless, the function and action mechanism of these proteins remain poorly understood in plants.

Our research group identified and characterized a new lectin isolated from *Brassica oleracea* ssp. *botrytis*, designated BOL. This lectin is a non-glycosylated

protein with apparent molecular mass of 34 kDa, thermostable up to 60ºC and optimum activity at pH 7-8. Different from previously identified lectins, partial sequencing of BOL revealed two MATH-domains in tandem arrayed in its structure [12]. As lectin isolated from natural sources may be a heterogeneous mixture of several lectin isoforms with distinct biological activities [13] and, even presenting contaminants, we wanted to ensure that the lectin activity was indeed exerted by a TRAF protein. For this purpose, we determine the full coding sequence of *Bol* to confirm the sequence identity with TRAF-like proteins and developed a successful method to produce higher amounts of pure, biological-active rBOL. In this work, we detail the results of the gene sequencing and *in silico* prediction of the promoter region of the *Bol* with identification of key cis-regulatory elements and gene expression evaluation under different stress conditions. The study also includes a comparative tissue-specific expression profile and establishes methods for the recombinant expression and purification of this lectin that support initial studies of structure and function *in vitro*.

Materials and Methods

Bacterial strains and culture media

We used *E. coli* competent cells, strain DH5α (Invitrogen, USA), for plasmid cloning. BL21 (DE3) and C41 (DE3) competent cells were used for protein expression (Sigma-Aldrich, USA). Luria–Bertani (LB) [Bacto-tryptone 1% (w/v), yeast extract 0.5% (w/v), NaCl 1% (w/v)] was used as culture media.

In vitro *seed germination of* Brassica oleracea *ssp.* botrytis

For surface disinfection, seeds were immersed in 70% (v/v) ethanol for 1 min, and then kept in 1% sodium hypochlorite and 0.1% Tween 20 solution, for 10 min. The seeds were rinsed three times in sterile distilled water and transferred to 250-mL glass flasks containing 40 mL of semisolid medium composed of half-strength MS salts and vitamins [14], 2% sucrose, 100 mg L^{-1} of myo-inositol and solidified with 6.5 g L^{-1} of agar (Merck, Germany). The pH of the media was adjusted to 5.7 ± 0.01, and the media was autoclaved at 120ºC and 108 kPa for 20 min. Flasks were covered with rigid polypropylene lids containing two 11 mm-orifices covered with 0.45-µm-pore size membranes (PTFE; MilliSeal® AVS-045 Air Vent, Tokyo, Japan).

The cultures were kept in growth room at 27 ± 1ºC, 16/8h light/dark photoperiod under 40 μmol m^{-2} s^{-1} irradiance, for 30 days

RNA extraction and cDNA cloning

Total RNA was extracted from *Brassica oleracea* ssp. *botrytis* seedlings using Tris reagent® (Sigma-Aldrich, USA) according to manufacturer instructions. The first-strand cDNA was synthesized from 3 ug of RNA using SuperScript™ First-Strand cDNA Synthesis SuperMix kit (Invitrogen, USA). PCR amplification and DNA cloning were performed in two steps: the first using degenerate oligonucleotide primers (Table S1) designed based on the N-terminal amino acid sequence and tryptic peptides determined previously by mass spectrometry of the native lectin [12]. Amplification reactions were carried out using a C1000 Touch™ Thermal Cycler (Bio-Rad Laboratories Inc., USA) and Platinum Taq DNA Polymerase High Fidelity (Invitrogen, USA) in a total volume of 25 μL, under the following conditions: 1 cycle at 94ºC for 5 min followed by 35 cycles at 94ºC for 1 min, 54ºC for 1 min, and 72ºC for 1 min, and 1 cycle at 72ºC for 5 min. A 856 bp DNA fragment was recovered, cloned into the pGEM®-T Easy cloning vector (Promega Co. USA), and subjected to nucleotide sequencing by Automated DNA Sequencing Service at Macrogen (Seoul, Korea). The second PCR amplification was performed with specific primers (5'-ATG GAT AGT CAC AAA AGG AGT TTG-3' and 5'- CTA AAC GGA GAC GAT GTT GG) designed based in the *Brassica oleracea* var. *oleracea* sequence, GenBank accession no. XM_013750511, covering the coding sequence of the gene. The positive recombinants were confirmed by restriction enzyme analysis, using *Eco*RI enzyme, and DNA sequencing was carried out by Servicio Central de Apoyo a la Investigación, Universidad de Córdoba (Córdoba, Spain) in the sense and antisense directions with universal primers T7 and SP6, respectively.

Sequence analysis

DNA sequencing chromatograms were processed and the consensus sequences assembled using CodonCode™ Aligner (Li-COR Inc., USA). Homology searches were carried out using the BLAST program at the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/). Alignments of multiple protein sequences were performed using Clustal Omega version 1.2.1 [15] (http://www.ebi.ac.uk/Tools/msa/clustalo/) and potential cleavage sites were

identified using PeptideCutter [16] (http://web.expasy.org/peptide_cutter/). One kb upstream (from the start codon) of the gene was extracted as promoter sequences using Phytozome database [17] (http://www.phytozome.net/). To investigate potential regulatory cis-acting elements, the putative *Bol* promoter was analyzed with PlantCARE [18] (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Gene expression analysis using semi quantitative and real time quantitative PCR (RT-qPCR)

Relative expression analysis of *Bol* was carried out from leaf, stem and root tissues using iCycler iQ system (Bio-Rad Laboratories Inc., USA). Total RNA was isolated from different tissues using the TRI REAGENT® RNA/DNA/protein isolation reagent (Sigma-Aldrich, USA). Prior to RT-qPCR, total RNA was treated with RNase-free DNAseI (New England Biolabs, USA) at 37ºC for 30 min to eliminate any traces of genomic DNA. To ascertain successful removal of DNA, a PCR reaction was done with primers UBQ2-F and UBQ2-R to confirm the absence of amplification bands of the ubiquitin-2 (gene ID. KF218594) using the DNase-treated RNAs samples. Equal amounts of RNA extracted (2.5 µg) from all samples were used to synthesize first strand cDNA anchored with Oligo dT in a 20 µl reaction using SuperScript™ First-Strand cDNA Synthesis SuperMix kit (Invitrogen, USA). Semi-quantitative PCR was performed with primers BOLsRT-F and BOLsRT-R; UBQ2-F and UBQ2-R (Table S1), under the following conditions: 1 cycle at 95ºC for 3 min followed by 25 cycles of 95ºC for 45 s, 57ºC for 30 s and 72ºC for 30 s followed by a final extension of 72ºC for 7 min. Real-time qPCR was carried out with the iCycler iQ system using the iQ SYBR Green supermix (Bio-Rad Laboratories Inc., USA) and the *Bol* gene-specific primers BOLqRT-F and BOLqRT-R (Table S1). The PCR amplification parameters were 95ºC for 5 min, followed by 40 cycles of 95ºC for 30 s, 60ºC for 30 s and 72ºC for 30 s. The melting curve (single peak for each primer) was used for primer specificity with a temperature range of 60–100ºC by increments of 0.5ºC. Analysis of relative gene expression was done according to Livak and Schmittgen (2001) [19] using Ubiquitin-9 (gene ID.At4G27960) as reference gene.

Effect of light and darkness on the Bol expression

To test the effect of light and darkness on *Bol* gene expression, leaves of eight week-old plants were excised and kept in a humidified chamber protected from light

for 4 days. Control leaves were kept in the same conditions, but with 16 hours of light and 8 hours of darkness. Samples were processed and transcript levels were measured by semi-quantitative PCR with BOLsRT-F and BOLsRT-R and Ubiquitin-2 as reference gene. The densitometric value of optical density (OD) units of each band, obtained using the program Quantify One (Bio-Rad Laboratories Inc., USA) was then related to the OD units of the corresponding band of the ubiquitin.

Effects of different biotic and abiotic stimuli on the Bol expression

To investigate if *Bol* may be responsive to biotic and abiotic stimuli, seeds were grown in Petri dishes using MS medium as described above. The seedlings were then transferred, after 5 days, to culture bottles with different MS medium combinations. For simulated biotic stress we used 100 μM methyl jasmonate (MeJa), 200 μM salicylic acid (SA) or 10 μM 1-aminocyclopropane-1-carboxylic acid (ACC). Whereas, for abiotic stress 90 mM NaCl, 100 mM mannitol or MS modified medium lacking potassium phosphate were used. Tissues were collected after 72 h of treatment, and stored until further analysis. Control plants were similarly transferred to only MS medium.

Construction of expression vector

Amino acid sequence (301 residues) of BOL was back-translated *in silico* and DNA encoding of the protein of interest was synthesized by Genscript (Genscript Co., USA) with the *E. coli* preferred coding usage. The 3′-end of the *Bol* gene was extended by addition of six His residues corresponding to the His-Tag. *Nde*I restriction site was incorporated into the 5’-end, and *Sal*I was incorporated into the 3’-end into the expression vector pET21b(+).

Expression and purification of rBOL

The recombinant protein was expressed after transformation of BL21 (DE3) and C41 (DE3) cells with the construct pET21b-*Bol*. Transformed cells were cultured in 20 mL of LB agar-plates supplemented with 100 μg/mL ampicillin. Plates were incubated overnight at 37ºC. Single colonies were inoculated in 5 mL LB medium/100 μg/mL ampicillin, overnight at 37ºC with shaking at 220 rpm. This pre-culture was used to inoculate (1:100) 500 mL LB medium containing 50 μg/mL

ampicillin, which was cultivated at 37ºC or at 20ºC. For expression at 37ºC, 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added when the OD_{600nm} reached between 0.6 – 0.8, then cultures were incubated for another 4 h at 37ºC. For expression at 20ºC, the cultures were induced by adding 1 mM IPTG and grown at 20ºC overnight with shaking at 220 rpm. The cells were placed on ice for 10 min and harvested by centrifugation at 5.000 *g* for 10 min at 4ºC. Pellets were resuspended in lysis buffer (50 mM Tris, 300 mM NaCl, 10 mM imidazole pH 8.0). The homogenates were sonicated 6 times for 10 s on ice, with 10 s interval between each sonication and then, centrifuged at 7000 *g* for 15 min. The supernatant was loaded onto a His Trap column (0.7 cm x 2.5 cm, GE Healthcare). The column was washed with washing buffer (50 mM Tris, 500 mM NaCl, 10 mM imidazole, pH 8.3) to exclude nonspecifically bound contaminants and eluted with elution buffer (50 mM Tris, 500 mM NaCl, 300 mM imidazole, pH 8.3). Finally, recombinant protein was concentrated 10-fold and the buffer was exchanged to PBS buffer (NaH_2PO_4 10 mM, Na_2HPO_4 40 mM, NaCl 150 mM, KCl 2.7 mM) at pH 7.4 using an Amicon® Ultra filter (Merck Milipore, Germany). The concentration of this protein was determined according to the BCA assay (Pierce Biotechnology, USA) with bovine serum albumin as a standard, according to manufacturer instructions.

SDS-PAGE

SDS–polyacrylamide gel electrophoresis (SDS–PAGE) was conducted using 12% (w/v) acrylamide in gels and the protein bands were visualized by 2% (w/v) Coomassie brilliant blue R-250. Protein Ladder Marker (Pierce Biotechnology, USA) was used as the standard molecular mass marker.

Biological activity

Activity of the recombinant protein was assessed by a hemagglutination assay. A serial two-fold dilution of the lectin (50 μL) was gently mixed with 25 μL of a 2% goat erythrocyte suspension in microtiter U-plates, incubated at 37ºC for 1 h, and the agglutination was visualized. The titer is defined as the reciprocal of the highest dilution exhibiting hemagglutination. Specific activity is defined as the number of hemagglutination units per mg protein.

Thermal stability of recombinant lectin

To study the effect of temperature on hemagglutinating activity, purified rBOL was incubated for 30 min at different temperatures (4, 25, 50, 60, 70, 80, 90 and 100ºC). At the end of the incubation period, the samples were gradually cooled to room temperature (25ºC) and measured for their remaining functional activity.

Western Blot Analysis

Protein samples were separated on 12% PAGE/SDS and electrophoretically transferred onto nitrocellulose membrane for 2 h using a Mini Trans-Blot® Cell system (Bio-Rad Laboratories Inc., USA). Non-specific protein-protein interactions were blocked using 3% gelatin in TBS-T buffer (20 mM Tris-HCl, 150 mM NaCl, 0.1% v/v Tween-20) for 1 h at room temperature. The membrane was incubated with a polyclonal BOL antibody produced in our laboratory against the native lectin in mouse (not shown) diluted 1:100 in blocking buffer (1% gelatin) for 2 h at room temperature after that was washed three times for 15 min each time in TBS-T buffer and then incubated with a 1:2000 dilution of secondary HRP-conjugated antibody for 30 min (rabbit anti-mouse IgG, Sigma-Aldrich, USA.). Immunoreactive bands were visualized using 3,3'-diaminobenzidine (DAB) as chromogen (Sigma-Aldrich, USA). We used prestained protein standard broad range as ladder (Bio-Rad Laboratories Inc., USA).

Analysis of the predicted three-dimensional structure of the BOL protein

Prediction of the secondary structure of BOL was done with SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) and a tertiary structure analysis was accomplished using RaptorX [20] (http://raptorx.uchicago.edu/). The predicted model was visualized using the PyMOL Molecular Graphics System (DeLano Scientific, USA).

Circular dichroism

Circular dichroism (CD) measurements of the recombinant lectin were performed in PBS buffer, pH 7.4 using a Jasco J 810 spectropolarimeter (Jasco Inc., USA). Each sample was scanned four times from 190 to 250 nm in a quartz cuvette with a 0.1 cm path length. The baselines (buffer alone) were subtracted from the protein sample spectra.

Statistical analyses

All results are means of three independent experiments with two technical replicates per experiment. Values are mean ± SE. Data were subjected to one-way ANOVA for means comparison using the statistical software package (GraphPad Prism®, La Jolla, USA), and the means were compared using confidence interval by Dunnet's test ($p < 0.05$). In the tissue expression analysis, the letters represent significant differences ($p < 0.05$) based on ANOVA using the Holm–Sidak method. To test the light/darkness effect, a t-test was carried out.

Results

Cloning and sequence analysis of Bol cDNA

A cDNA fragment (approximately 900 bp) of the lectin was amplified using degenerate primers. The sequenced 856 bp *Bol* cDNA cloned in this study was 100% identical to the *Brassica oleracea* var. *oleracea* gene sequence (XM_013750511.1), which encodes a putative TRAF-like protein of 350 amino acids. Specific primers were designed from this sequence. The entire coding DNA sequence (1053-bp) of *Bol* was analyzed and deduced (Fig. 1). The deduced amino acid sequence confirmed an identical protein result with five matched polypeptide sequences from mass spectrometry and the 19 residues of native lectin determined by Edman degradation [12]. BLAST searches against protein databases indicated that the deduced amino acid sequence of the BOL shares no significant homology with other known lectins. On the other hand, BOL showed 100% sequence identity with uncharacterized, 350 amino acids TRAF-like proteins, from *Brassica oleracea* var. *oleracea* (XP_013605965.1), 99% identity to *Brassica rapa* (XP_009111696.1), and 98% identity to *Brassica napus* (CDY19775.1). Alignment of BOL and these TRAF-like proteins is shown in Fig. 2. Analysis of the prediction signal sequences of the proteins of *B. oleracea* var. *oleracea, B. rapa* and *B. napus* showed they contain a signal peptide that is most probably cleaved between residues 49 and 50, resulting in a polypeptide of 301 amino acids. This position corresponded to the beginning of the determined N-terminal region of the lectin of *B. oleracea* ssp. *botrytis*. Therefore, we considered that the mature BOL protein is consisting of 301 amino acids, containing two domains with homology to the TRAF family.

```
atggatagtcacaaaaggagtttgggtttcatatctctcgctttctcttcatcacttgc 60
 M  D  S  H  K  R  S  L  G  F  I  S  L  A  F  L  F  I  T  C
tcttctgctgagttcctcattcaacaggtcacagaaggcaaaggaacagagaacaacagt 120
 S  S  A  E  F  L  I  Q  Q  V  T  E  G  K  G  T  E  N  N  S
tattacaatgtcgaggcgaatcttggagagacaagagcgtttagagaggagcgaccatca 180
 S  Y  N  V  E  A  N  L  G  E  T  R  A  F  R  E  E  R  P  S
agtaagatagtgacaatagcaggctactcggtgattaaagggagattcgaaccctacgaa 240
 S  K  I  V  T  I  A  G  Y  S  V  I  K  G  R  F  E  P  Y  E
tcctctgttttcgaagcagctggttacagatggagattggttttgtacgtgattggtaat 300
 S  S  V  F  E  A  A  G  Y  R  W  R  L  V  L  Y  V  I  G  N
aaaaacgacggtggagctggccatatatcactttacgtgaggatggaagagaccgagtct 360
 K  N  D  G  G  A  G  H  I  S  L  Y  V  R  M  E  E  T  E  S
cttccctatggttgggaagtcaatgttgatctcaaactctttgtccacaatccaaagacc 420
 L  P  Y  G  W  E  V  N  V  D  L  K  L  F  V  H  N  P  K  T
cataagtacttgactgtcacagatggaaccgtgaagcgatacaacgctgcaaaaaaggag 480
 H  K  Y  L  T  V  T  D  G  T  V  K  R  Y  N  A  A  K  K  E
tggggattcggaaaattgatctctctttcaatattcgagaacacgaacaatggctacatt 540
 W  G  F  G  K  L  I  S  L  S  I  F  E  N  T  N  N  G  Y  I
gtgcaagacacttgttcttttggtgctgagatcctcattgttaaaccagctgagatacaa 600
 V  Q  D  T  C  S  F  G  A  E  I  L  I  V  K  P  A  E  I  Q
gagaaagtcacattcatatcaaaccctccacacaatgttttcacttggaagatacttcgt 660
 E  K  V  T  F  I  S  N  P  P  H  N  V  F  T  W  K  I  L  R
ttctccagcttggaagataaattttactattctgccgatttctcgttggagaccgatac 720
 F  S  S  L  E  D  K  F  Y  Y  S  A  D  F  L  V  G  D  R  Y
tggagattaggctttaacccgaaaggggatggagatggacgaccacatgcacttccactc 780
 W  R  L  G  F  N  P  K  G  D  G  D  G  R  P  H  A  L  P  L
ttcttatttgctcaaggctttaaggaaaacgcagttgctacgaacacttggggagcagtt 840
 F  L  F  A  Q  G  F  K  E  N  A  V  A  T  N  T  W  G  A  V
aacttgcggttaaagaatcagcgaaactccaaccacagacaattatattctgcagcttgg 900
 N  L  R  L  K  N  Q  R  N  S  N  H  R  Q  L  Y  S  A  A  W
taccctattcgaagcgattatggtgtaggagtgaacaacatcatattgctatcagattta 960
 Y  P  I  R  S  D  Y  G  V  G  V  N  N  I  I  L  L  S  D  L
aaggatccgtcgaaagggtatttggtgaatgatgccattatctttgaagctgaaatggtt 1020
 K  D  P  S  K  G  Y  L  V  N  D  A  I  I  F  E  A  E  M  V
aaggtctctgtgaccaacatcgtctccgtttag                            1053
 K  V  S  V  T  N  I  V  S  V
```

Figure 1. Nucleotide and deduced amino acid BOL sequences. The N-terminal amino acid sequence determined from the native protein is indicated by bold letters and the tryptic fragments obtained from mass spectrometry are underlined. Degenerated primers were designed based on these amino acid sequences and are indicated by arrows (continuous line). Arrows with dotted line indicate specific primers. Boxes in ATG and TAG mark the positions of the translational start and stop codons, respectively.

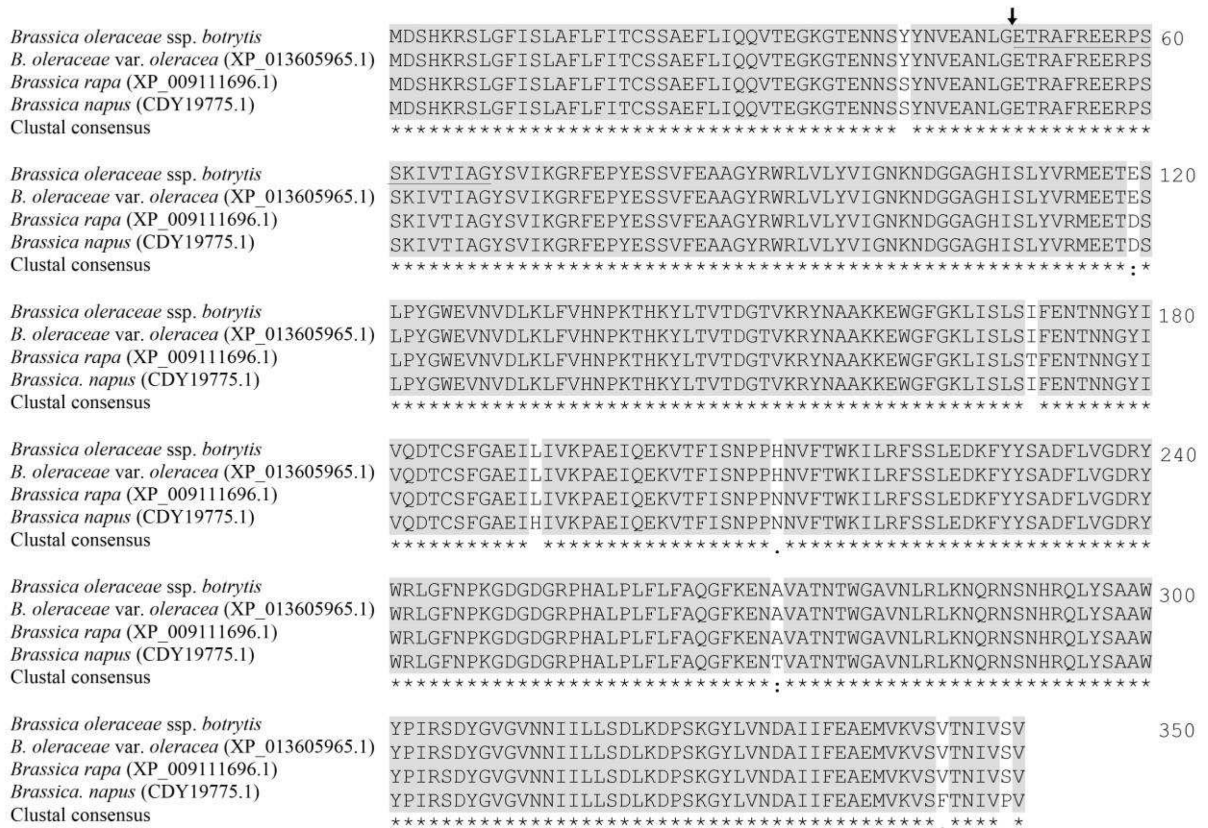


Figure 2. Alignment of deduced amino acid sequences between BOL and uncharacterized proteins of *Brassica oleracea* var. *oleracea* (XP_013605965.1) *B. rapa* (XP_009111696.1) and *B. napus* (CDY19775.1). Sequences were aligned using Clustal Omega version 1.2.1. Identical amino acids are highlighted in gray. The consensus residues are indicated with asterisks below the sequence, highly conserved residues with (**:**) and weakly conserved (**.**).The N-terminal amino acid sequence reported for the *Brassica oleracea* ssp. *botrytis* lectin is underlined. Arrow represent the predicted cleavage site in the TRAF-like proteins.

The Bol promoter contains regulatory elements in response to light and stress signals

To investigate potential regulatory cis-acting elements, the *Bol* putative promoter was analyzed *in silico* with PlantCARE and we considered elements whith matrix score > 5 (Fig. S1). TATA box (TATA) sequences were found at numerous positions 55 (-), 388 (-), 428 (+), 460 (-), 461 (-), 462 (-), 853 (-), 900 (+), 937 (-), 938 (-), 939 (-), 983 (+), 984 (+), 985 (+) and CAAT-box sequences responsible for the tissue specific promoter activity were identified at 596 (+), 679 (-) positions. Several stress-regulatory elements, such as TC-rich repeat element, ABRE, MBS, HSE, LTR, which are homologous to cis-acting elements identified in plants, were found in the *Bol* promoter. There are also eight elements related to light responsiveness. (Table 1).

Table 1. Elements present in the predicted promoter region of *Bol* according to the PLANT CARE database. Selected Matrix score for all elements >5

cis element	Function
ABRE	cis-acting element involved in the abscisic acid responsiveness
ACE	cis-acting element involved in light responsiveness
ARE	cis-acting regulatory element essential for the anaerobic induction
AT-rich sequence	element for maximal elicitor-mediated activation (2 copies)
Box 4	part of a conserved DNA module involved in light responsiveness (2 copies)
Box III	protein binding site
CAT-box	cis-acting regulatory element related to meristem expression
G-Box	cis-acting regulatory element involved in light responsiveness (2 copies)
HD-Zip 1	element involved in differentiation of the palisade mesophyll cells
HD-Zip 2	element involved in the control of leaf morphology development
HSE	cis-acting element involved in heat stress responsiveness (2 copies)
LTR	cis-acting element involved in low-temperature responsiveness
MBS	MYB binding site involved in drought-inducibility
MBSI	MYB binding site involved in flavonoid biosynthetic genes regulation
MNF1	light responsive element
MRE	MYB binding site involved in light responsiveness
P-box	gibberellin-responsive element
Skn-1_motif	cis-acting regulatory element required for endosperm expression
Sp1	light responsive element
TC-rich repeats	cis-acting element involved in defense and stress responsiveness
TCT motif	part of a light responsive element
as-2-box	involved in shoot-specific expression and light responsiveness
circadian	cis-acting regulatory element involved in circadian control

Comparative tissue specific expression of Bol

Tissue-specific expression pattern of *Bol* was detected in root, stem and leaf by using an equal amount of RNA sampled from different tissues of *B. oleracea* ssp. *botrytis* by using semi quantitative and quantitative PCR. Results revealed that *Bol* transcripts were relatively abundant in leaf tissue, followed by stem and lastly in root tissue (Fig. 3).

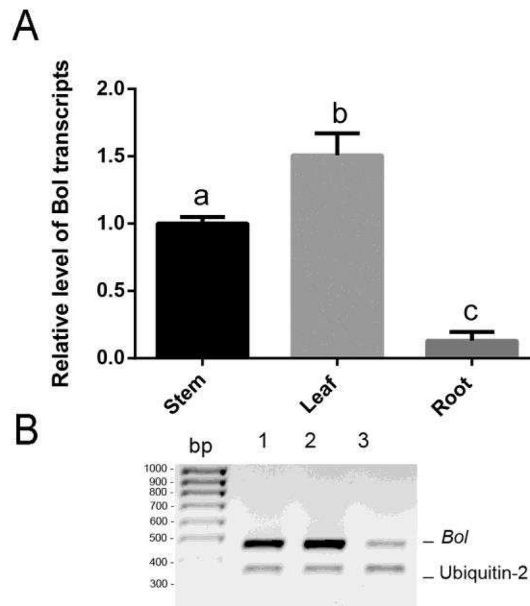


Figure 3. Comparative expression level of *Bol* in various tissues. *Bol* mRNA expression was measured by quantitative RT-PCR (A) and semi-quantitative RT-PCR (B). Lane 1 - stem. Lane 2 - leaf. Lane 3 - root. Error bars represent means ± SD (n = 3). Different letters indicate significant differences ($p<0.01$) based on ANOVA using the Holm–Sidak method.

Bol expression is regulated by light

The effect of light/dark on the expression of *Bol* was evaluated after excided leaves were kept under continuous darkness for 4 days. Results revealed that darkness decreased *Bol* expression in leaves (Fig. 4).

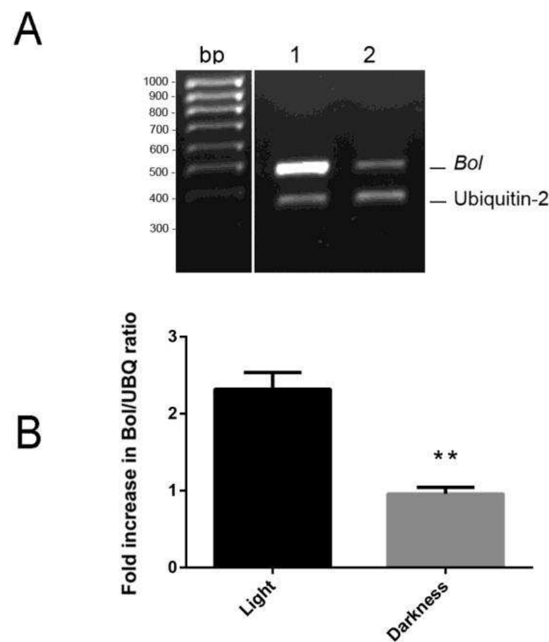


Figure 4. Effect of light on *Bol* expression. (A) Semi- quantitative PCR results are shown from excised leaves kept in light – Line 1 and dark – Line 2 for 4 days; Ubiquitin 2 was used as control. (B) Transcript level of *Bol* in light and dark conditions measured by densitometry. *Bol* transcript accumulation was quantified by densitometry using Bio-Rad Quantity One software; each value was corrected according to the *Ubq* 2 amplification signal. Error bars represent means ± SD (n = 3). An t-test was carried out and a significant P value ≤ 0.01 is indicated by asterisks.

Bol response to biotic stimuli

Most of the predicted motifs present in the *Bol* promoter sequence are involved in response to environmental stress. Since lectins have been long associated with a role in defense responses, the expression was checked under biotic and abiotic stress simulating conditions in *B. oleracea* ssp *botrytis*. Our results demonstrate that the expression of *Bol* increases in response to MeJa but decrease in response to SA (Fig. 5A and C). Treatment simulating abiotic stress did not cause significate changes in the *Bol* transcript levels (Fig. 5B).

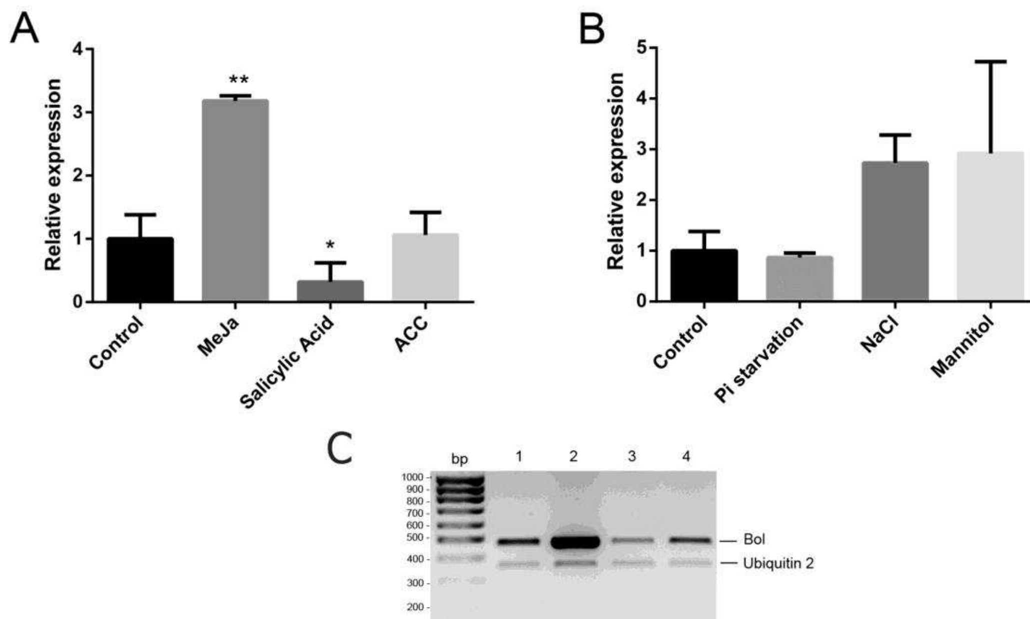


Figure 5. Expression analysis of *Bol* gene under different environmental conditions in *B. oleracea* ssp *botrytis*. Evaluation of *Bol* transcript under biotic stress simulation using quantitative RT-PCR (A) and semi-quantitative RT-PCR analysis (C). Lane 1: control, Lane 2: 100 µM MeJa, Lane 3: 200 µM SA and Lane 4: 10 µM ACC. Evaluation of *Bol* transcript under abiotic stress simulation using quantitative RT-PCR analysis. Error bars represent means ± SD (n = 3). Asterisks indicate stress treatment groups that showed a significant difference in transcript abundance compared with the control group (*p < 0.05; ** p < 0.01).

Recombinant expression of pET21b(+)-Bol

In order to observe the expression of *Bol* in *E. coli*, the mature protein coding cDNA of *Bol* was cloned into an expression vector. The resulting construction was confirmed by restriction analysis and sequencing (not shown) and named pET21b-*Bol*. pET21b-*Bol* was expressed in *E. coli* BL21 (DE3) and C41 (DE3). Initially, we tested different concentrations of IPTG (0.3, 0.4, 0.5, 1, 1.5 and 2 mM), the optimized expression condition was obtained with 1 mM of IPTG. Using this concentration, two expression conditions were compared: 4 h at 37ºC and overnight at 20ºC. A molecular weight of about 35 kDa was expected for the BOL recombinant protein. As is shown in Fig. 6, a protein of the expected molecular mass was expressed upon IPTG induction, whereas no similar size protein was observed in the non-induced *E.coli* transformed with same vector pET21b-*Bol*. The protein was both present in soluble and insoluble (inclusion bodies) fractions, in the two bacterial strains, however, BL21 (DE3) had slower growth and lower recombinant protein

production either at 37ºC or 20ºC. Protein solubility is usually correlated with its correct folding and activity, therefore our aim was to establish the conditions to obtain high production of soluble protein and correctly folded protein. We chose to express in large-scale pET21b-*Bol* at 20ºC overnight in the C41 (DE3) strain, which gave rise to a higher amount of soluble protein.

Purification of recombinant lectin

The recombinant lectin was purified from the bacterial lysate by affinity chromatography on a nickel column, and it was eluted with 300 mM imidazole in 50 mM Tris/500 mM NaCl, resulting in the profile shown in Fig. 6C.

The molecular mass of rBOL was estimated to be 34 kDa using SDS-PAGE. This electrophoretic profile is similar to that provided by native protein (Fig. 6D). The ultimate concentration of recombinant protein analyzed by the BCA protein estimation assay was 600 µg/mL, with an average yield of 2 mg of purified lectin per liter of bacterial culture from C41 (DE3) strain.

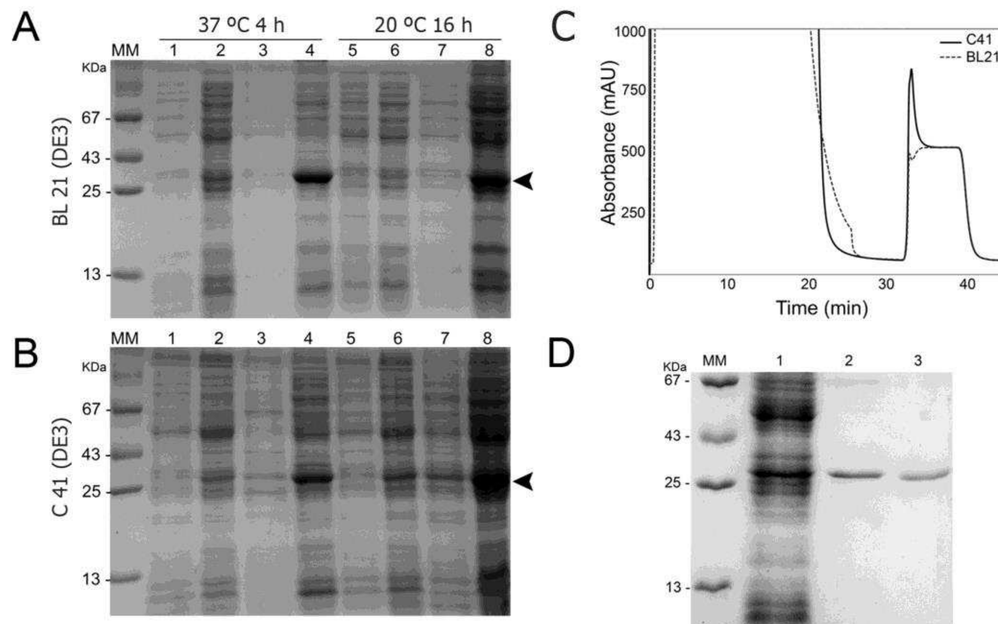


Figure 6. Solubility comparison of rBOL expressed at different temperatures and purification of recombinant protein. (A) BL21 (DE3) strain (B) C41 (DE3) strain. The arrowhead points out the position of target protein. Lanes 1 and 5: control without IPTG; Lanes 2 and 6: induced whole-cell lysate; Lanes 3 and 7: supernatant of the cell lysate; Lanes 4 and 8, pellet of the cell lysate. (C) Elution profile of soluble cytoplasmic fraction of rBOL on Ni-Sepharose affinity column monitored by measuring absorption at 280 nm. Solid line: C41 (DE3). Dotted line: BL21 (DE3). (D) SDS-PAGE analysis of purified rBOL. Lane 1 - whole bacterial cell lysate. Lane 2 - rBOL purified by affinity column. Lane 3 - purified native BOL loaded as control. MM: molecular weight markers.

Properties of recombinant lectin

We showed that recombinant lectin has structural and functional properties similar to those of the native protein. The hemagglutination assay is the most widely used tool to evaluate the carbohydrate-binding ability of lectins. rBOL induced cell agglutination at concentrations 1.875 µg/mL, with specific activities of 533 HA/mg, while the native protein showed agglutination at 2 µg/mL and specific activities of 500 HA/mg (Table 2).

Table 2. Hemagglutination activity of BOL and rBOL. Hemagglutinating Activity Unit (HA), corresponds to the minimum quantity of protein capable of inducing agglutination; HA/mg corresponds to the amount of hemagglutinating units per milligram of protein.

Protein	Total Protein (µg)	Total Activity (HA)	Specific Activity (HA/mg)
rBOL	600	320	533
BOL	160	80	500

Thermal stability of the rBOL was determined in the temperature range from 4ºC to 100ºC (Fig. 7A). The lectin was stable up to 60ºC for 30 min and it was totally inactivated when incubated at 70ºC or above for 30 min. Polyclonal antibodies generated against BOL by immunizing a mouse with native form of the lectin (not shown) were able to recognize native and recombinant lectin (Fig. 7B).

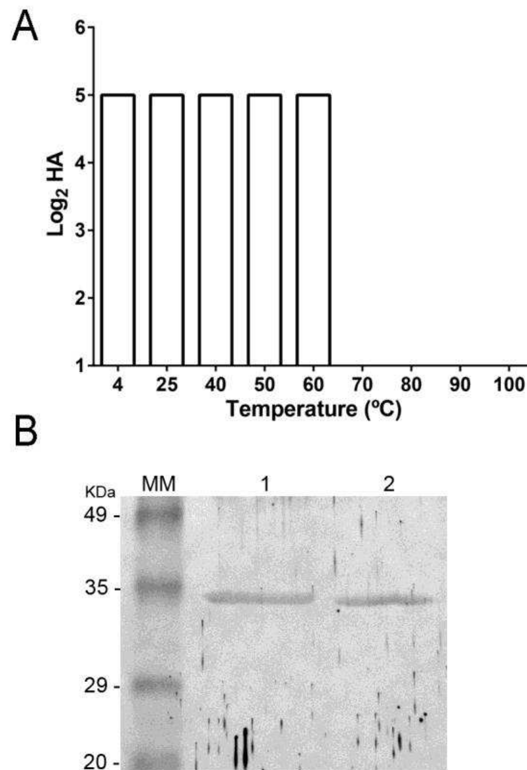


Figure 7. Properties of rBOL. (A) Effects of temperature on the hemagglutination activity of rBOL. The activity was determined after incubation of lectin at different temperatures (4–100ºC) for 30 min. (B) Western blot of the recombinant BOL (Lane 1) and native BOL (Lane 2) developed with polyclonal antiserum raised against the purified natural protein.

Analysis *in silico* of the secondary structure of BOL using SOPMA and Raptor-X programs showed the presence of β-sheets connected with turn and coil. The BOL protein contained 2% helix, 45% β-sheet and 51% loop structures. The solvent accessibility was divided into three states by 2 cut-off values: 10% and 42%. A value of less than 10% was identified as buried, higher than 42% was identified as exposed and between 10% and 42% was identified as medium. Proportions of buried, medium and exposed regions in our protein were 27%, 35% and 37%, respectively. Raptor-X structure prediction showed a three-dimensional model with 100% of the amino acid residues modeled in two in-tandem arrayed MATH-domains: domain 1 consists of residues 1-157 and residues 158-301 compose domain 2, with overall uGDT(GDT) value 214(71) and p-value 1.17e-07. A schematic representation of the structure is shown in Fig. 8A. Both domains exhibited folding mode and spatial configuration similar to PDB 2F1X, a MATH domain of HAUSP/USP7 [21]. Although it did not displayed any sequence homology to any other lectin described, the BOL protein had an overall dome-like structure typical of L-type lectins with a β-sandwich

fold. The secondary structure of recombinant lectin was investigated by circular dichroism and showed a typical spectrum of proteins constituted by β-sheet with a broad negative band near 218 nm (Fig. 8B).

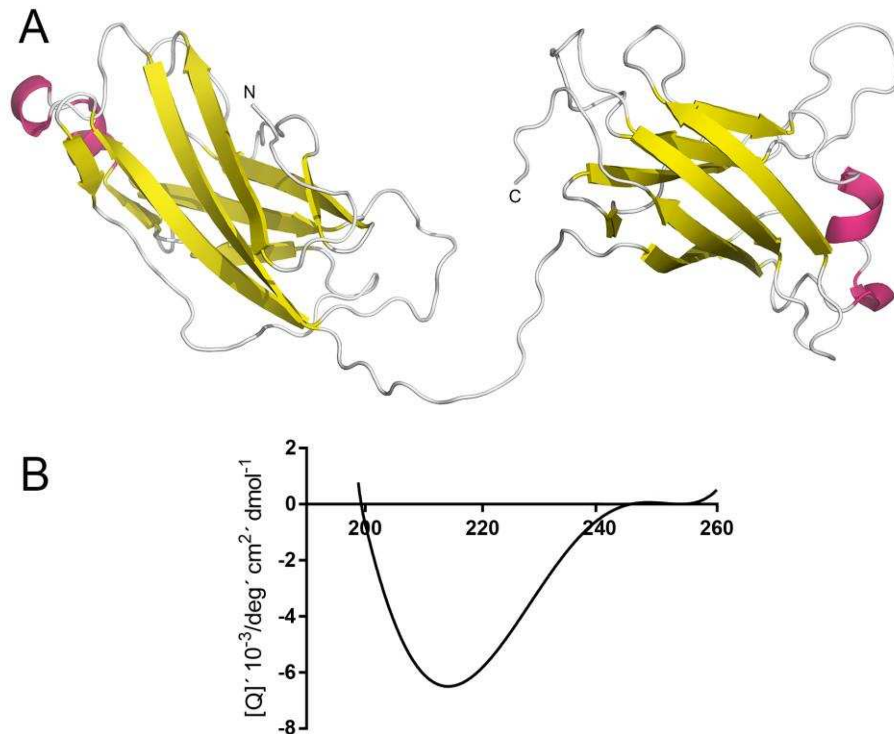


Figure 8. The structure of BOL. (A) Ribbon diagram of the predicted three-dimensional BOL structure. The model was generated with RaptorX using PDB template 2F1X. The lectin contains two MATH-domains and adopts a dome-like fold typical of L-type lectins. Helices are colored red, β-sheets are yellow, and loops are gray. N and C mark the N- and C-terminal ends of the protein. (B) CD spectra of BOL in PBS buffer.

Discussion

The present study describes the cloning, sequencing, promoter *in silico* analysis and heterologous expression of a novel lectin isolated from *Brassica oleracea* ssp. *botrytis* (BOL). BLAST search of deduced amino acid sequence demonstrated that BOL is 99% identical to uncharacterized Tumor Necrosis Factor Receptor-associated Factors (TRAFs)-like protein from *Brassica rapa* (XP_009111696.1) and *Brassica napus* (CDY19775.1), and two conserved MATH-domain were detected in lectin sequence. The TRAF domain, also known as meprin and TRAF-C homology (MATH) domain is a fold of seven or eight anti-parallel β-sheets that form β-sandwich structure containing a surface crevice responsible for protein-protein interactions

[20,21]. We have previously shown that BOL is a lectin, capable of binding to glycoproteins as fetuin and asialofetuin, and that this occurs via oligosaccharide moieties instead of protein-protein interactions [12], which could be a new interaction mechanism, as yet unidentified in TRAF proteins. Lectins able to precipitate glycoconjugates show at least two carbohydrate-recognition domains (CRD) [3]. Based on this and even in the absence of structural data, we hypothesized that CRD domains can be located in the MATH domains or compose the domain itself.

BOL three-dimensional structure exhibited only two MATH-domains. Proteins that feature MATH-domains are commonly associated with other protein domains, including peptidases, filamin and RluA domains, broad-complex, tramtrack and brie-a-brae (BTB) domain, tripartite motif (TRIM), astacin domain and RING and Zinc finger domains. Nevertheless, there are a large number of hypothetical proteins identified by conceptual translation of genomes displaying one or multiple MATH-domains arranged in tandem (MATHd-only). These proteins are common in plants, lower metazoan and lower eukaryotes, but their function is not yet known [9]. Recent studies have been developed to try to clarify the physiological role of this class of proteins. The TRAF-protein of *Arabidopsis* roots is one of the first components to respond to the presence of the fungus *Piriformospora indica*, which interacts with *Arabidopsis* and mimics an arbuscular mycorrhiza [11]. Zhao and colleagues (2013) analyzed the expression pattern of genes coding for possible MATHd-only proteins in *B. rapa* and suggested that these genes may be involved in seed development [23]. Huang *et al*. (2016) reported the identification of two TRAF proteins Arabidopsis as redundant immune regulators [24]. MATHd-only sequences represent the majority of the MATHd-encompassing proteins identified in *Arabidopsis thaliana* and *Caenorhabditis* sp. and have been considered as truncated proteins [9]. Hitherto there have been no studies, at the protein level, that have identified functional MATHd-only proteins, and accordingly BOL is the first functional protein isolated having only MATH-domains.

Bol promoter analysis revealed the presence of numerous TATA box, two CAAT-box and 23 regulatory elements in response to light signals and stress. The presence and the number of TATA boxes can quantitatively increase the gene expression and CAAT-box sequences have a proven positive impact on frequency of transcriptional initiation [25].The promoter analysis also identified TC-rich repeats (ATTCTCTAAC) which are associated with a defense and stress-responsive element of *Nicotiana*

tabacum anionic peroxidase gene [26]. The MBS element is a binding consensus site for a MYB homologue [27], LTR is a low-temperature responsive element of a *Hordeum vulgare* lipid transfer protein gene [28] and HSE is an element involved in heat stress responsiveness in *Brassica oleracea*. In plants, most ABA-responsive genes have the conserved ABREs in their promoters, which are significant cis-elements for genes responsive to abiotic stress in *Arabidopsis* [29].

Additional predicted cis-regulatory elements including, an endosperm expression element, P-box, a Skn-1 motif, a circadian element, and an as-2-box, were also identified in the promoter region. Most of the predicted motifs are involved in response to environmental stress. The presence of these elements suggests that *Bol* promoter may play a core role in defense response regulation. The plant hormones jasmonic acid (JA), salicylic acid (SA), and ethylene (ET) are primarily involved in plant defense [30]. Increased expression in response to JA and ET [31], SA [32] are related to pathogen protection responses. In this study, *Bol* was upregulated by MeJa and downregulated by SA, in agreement with the primary mode of mutually antagonistic interaction between these two phytohormones pathways [33]. Classical lectins have constitutive expression and are expressed in high amounts in seeds and vegetative storage tissues. These plant lectins have not generally been thought to be inducible proteins [34]. Nevertheless, there is increasing evidence that some lectins are induced by biotic or abiotic stimuli, such as salt stress, drought, light, heat or cold [35]. During the last decade, this new class of plant lectins has been extensively studied [36] and the identification of inducible lectin genes provides experimental evidence that reinforces the idea of a potential intersection between lectins and plant stress responses [37]. The results of *Bol* transcript induction presented herein provide new insights into the possible role of BOL in stress responses from *Brassicas*.

Another interesting feature of the *Bol* promoter was the frequent presence of response elements to light and the light/darkness experiments demonstrated the importance of these elements in the gene expression, since there was a 59% reduction in the BOL expression in leaves kept in the dark. Light is one of the most important environmental stimuli for plants, not only as an energy source, but also as a critical signal for controlling physiological and developmental processes [38], including the regulation of light-responsive genes [29,30]. Jung *et al*. (2008) found significant 8-fold or greater induction in the light, relative to dark conditions, in

different genes from rice [41]. Higher expression of *Bol* under light conditions is consistent with the expression pattern observed in the cauliflower tissues. A high level of *Bol* mRNA was found in organs like leaves and stems grown under well-illuminated conditions. These results suggest that *Bol* regulation may be mediated by light.

At the protein level, a better understanding of the BOL structure and functions requires larger quantities of this lectin. Nevertheless, purification of BOL from natural sources includes three chromatography steps and is a time consuming process demanding a large amount of initial biological material. Through these laborious purification steps, we recovered only 12% of total activity [12]. Identification of the whole coding sequence of the lectin plus heterologous expression could easily solve the problem of low protein yield by conventional purification procedures. In order to optimize the process to obtain soluble protein we tested different induction conditions for BL21 (DE3) and C41 (DE3). The amount of soluble protein obtained for the strain BL21 (DE3) was less than 0.1 mg/L in contrast to 2 mg/L for C41 (DE3), most likely due to toxic effects that BOL exert on this strain. Thus, for large scale expression, we choose to cultivate the C41 (DE3) strain at a lower culture temperature. Purified rBOL showed a molecular mass of ~ 34 kDa on SDS-PAGE, which indicated that the monomeric structure is identical to the native state, in agreement with previous results with the lectin extracted from the plant [12]. The recombinant lectin promotes hemagglutination of goat erythrocytes, therefore soluble rBOL retains the ability to bind carbohydrate motifs. Furthermore, it showed a thermal stability pattern similar to that of native BOL. rBOL also was recognized by anti-BOL antibodies raised against the native lectin, showing that the purified rBOL corresponds to the native protein. Taken together, these results indicated that recombinant lectin has structural and functional properties similar to those of the native protein and confirm that the hemagglutination activity is indeed performed by a TRAF protein.

Using bioinformatics tools, we predicted that the BOL three-dimensional structure resembles that of the L-type lectins, despite its primary sequence not being closely linked to any clusters of the other known lectins. The more recent classification of plant lectins was based on studies of genome/transcriptome that enabled the grouping of these proteins into twelve distinct families of evolutionary and structurally related carbohydrate-recognition domains [3,34]. Commonly L-lectins are composed of antiparallel β-sheets connected by short loops and β-bends, and form a dome-like

structure [43]. BOL showed a dome-like structure typical of the L-type lectins. The spectrum of circular dichroism from the recombinant lectin corroborates the data obtained for the prediction of the secondary structure of the BOL. Nevertheless, most of the L-lectins consist of two or four subunits, each with a single, small carbohydrate combining site with the same specificity, and a metal binding site (for calcium and manganese) usually required for their carbohydrate-binding activity. Each subunit has a size of 25–30 kDa and is composed of a single polypeptide chain of 200–300 residues [44]. In the present study, the full-length cDNA sequence of BOL encoded only one subunit with two domains. Moreover, the metal binding site is not crucial for biological activity of this lectin as shown previously [12]. These findings suggest that the BOL might be a new L-type lectin with peculiar features or even be part of a new family of lectin not yet described.

In conclusion, the *Bol* cDNA encoding a lectin from *B. oleracea* ssp. *botrytis* was successfully cloned and expressed using the His-tag pET21b(+) vector system. The full coding sequence confirmed identity with TRAF proteins and the resulting recombinant protein retained the ability to recognize carbohydrates, thereby BOL is the first TRAF protein with lectin activity. Furthermore, *Bol* promoter analysis identified motifs involved in response to environmental stress and *Bol* transcript was changed under biotic stress simulation, suggesting that BOL may play a role in defense response regulation of plants. The heterologous expression of the recombinant protein in *E. coli* as showed here is a promising tool for large-scale production of BOL, enabling the advancement of our understanding of this intriguing TRAF-like protein with lectin activity.

Acknowledgments

The financial support through grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico [474715/2013-2] and Fundação de Amparo à Pesquisa do estado de Minas Gerais are gratefully acknowledged. CEM Duarte received PhD research fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

References

[1] N. Lannoo, E.J.M. Van Damme, Lectin domains at the frontiers of plant defense., Front. Plant Sci. 5 (2014) 397. doi:10.3389/fpls.2014.00397.

[2] W.C. Boyd, E. Shapleigh, Specific Precipitating Activity of Plant Agglutinins (Lectins), Science (80-.). 119 (1954) 419. doi:10.1126/science.119.3091.419.

[3] E.J.M. Van Damme, N. Lannoo, W.J. Peumans, Plant Lectins, in: Adv. Bot. Res., 2008: pp. 107–209. doi:10.1016/S0065-2296(08)00403-5.

[4] I.K. Howard, H.J. Sage, C.B. Horton, Studies on the appearance and location of hemagglutinins from a common lentil during the life cycle of the plant, Arch. Biochem. Biophys. 149 (1972) 323–326. doi:10.1016/0003-9861(72)90328-1.

[5] E. Weber, D. Neumann, Protein bodies, storage organelles in plant seeds, Biochem. Und Physiol. Der Pflanz. 175 (1980) 279–306. doi:10.1016/S0015-3796(80)80070-9.

[6] C.L. Diaz, T.J.J. Logman, H.C. Stam, J.W. Kijne, Sugar-Binding Activity of Pea Lectin Expressed in White Clover Hairy Roots, Plant Physiol. 109 (1995) 1167–1177. doi:10.1104/pp.109.4.1167.

[7] W.J. Peumans, E.J. Van Damme, Lectins as plant defense proteins, Plant Physiol. 109 (1995) 347–352. doi:10.1104/pp.109.2.347.

[8] N. Sharon, H. Lis, History of lectins: from hemagglutinins to biological recognition molecules, Glycobiology. 14 (2004) 53R–62R. doi:10.1093/glycob/cwh122.

[9] J.M. Zapata, V. Martínez-garcía, S. Lefebvre, Phylogeny of the TRAF/MATH Domain, Adv. Bot. Res. Exp. Med. Biol. (2007) 1–24. doi:10.1007/978-0-387-70630-6_1.

[10] P. Xie, TRAF molecules in cell signaling and in human diseases, J. Mol. Signal. 8 (2013) 7. doi:10.1186/1750-2187-8-7.

[11] R. Oelmüller, T. Peškan-Berghöfer, B. Shahollari, A. Trebicka, I. Sherameti, A. Varma, MATH domain proteins represent a novel protein family in Arabidopsis thaliana, and at least one member is modified in roots during the course of a plant-microbe interaction, Physiol. Plant. 124 (2005) 152–166. doi:10.1111/j.1399-3054.2005.00505.x.

[12] C.E.M. Duarte, M.V. Abranches, P.F. Silva, S.O. De Paula, S.A. Cardoso, L.L. Oliveira, A new TRAF-like protein from Brassica oleracea ssp. botrytis with lectinic activity and its effect on macrophages, Int. J. Biol. Macromol. 1 (2017)

508–514. doi:10.1017/CBO9781107415324.004.

[13] J.W. Han, M.G. Jung, E.Y. Shim, J.B. Shim, Y.M. Kim, G.H. Kim, Functional recombinants designed from a fetuin/asialofetuin-specific marine algal lectin, Rhodobindin, Mar. Drugs. 13 (2015) 2183–2195. doi:10.3390/md13042183.

[14] T. Murashige, F. Skoog, A revised medium for rapid growth and bioassays with tobacco tissue cultures, Physiol. Plant. 15 (1962) 473–497. doi:10.1021/jf9040386.

[15] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y.M. Park, N. Buso, A.P. Cowley, R. Lopez, Analysis Tool Web Services from the EMBL-EBI, Nucleic Acids Res. 41 (2013). doi:10.1093/nar/gkt376.

[16] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. Wilkins, R. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, Proteomics Protoc. Handb. (2005) 571–607. doi:10.1385/1-59259-890-0:571.

[17] D.M. Goodstein, S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, D.S. Rokhsar, Phytozome: a comparative platform for green plant genomics, Nucleic Acids Res. 40 (2012) D1178–D1186. doi:10.1093/nar/gkr944.

[18] M. Lescot, P. Déhais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouzé, S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Res. 30 (2002) 325–327. doi:10.1093/nar/30.1.325.

[19] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods. 25 (2001) 402–408. doi:10.1006/meth.2001.1262.

[20] M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, Template-based protein structure modeling using the RaptorX web server, Nat. Protoc. 7 (2012) 1511–1522. doi:10.1038/nprot.2012.085.

[21] M. Hu, L. Gu, M. Li, P.D. Jeffrey, W. Gu, Y. Shi, Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: Implications for the regulation of the p53-MDM2 pathway, PLoS Biol. 4 (2006) 228–239. doi:10.1371/journal.pbio.0040027.

[22] J.M. Zapata, K. Pawlowski, E. Haas, C.F. Ware, A. Godzik, J.C. Reed, A diverse family of proteins containing tumor necrosis factor receptor-associated factor domains, J Biol Chem. 276 (2001) 24242–24252.

doi:10.1074/jbc.M100354200.
[23] L. Zhao, Y. Huang, Y. Hu, X. He, W. Shen, C. Liu, Y. Ruan, Phylogenetic Analysis of Brassica rapa MATH-Domain Proteins, Curr Genomics. 14 (2013) 214–223. doi:10.2174/1389202911314030007.
[24] S. Huang, X. Chen, X. Zhong, M. Li, K. Ao, J. Huang, X. Li, Plant TRAF Proteins Regulate NLR Immune Receptor Turnover, Cell Host Microbe. 19 (2016) 204–215. doi:10.1016/j.chom.2016.01.005.
[25] V. Kusnetsov, M. Landsberger, J. Meurer, R. Oelmüller, The Assembly of the CAAT-box Binding Complex at a Photosynthesis Gene Promoter Is Regulated by Light, Cytokinin, and the Stage of the Plastids, J. Biol. Chem. 274 (1999) 36009–36014. doi:10.1074/jbc.274.50.36009.
[26] F. Diaz-De-Leon, K.L. Klotz, L.M. Lagrimini, Nucleotide sequence of the tobacco (Nicotiana tabacum) anionic peroxidase gene, Plant Physiol. 101 (1993) 1117–1118. doi:10.1104/pp.101.3.1117.
[27] T. Urao, K. Yamaguchi-Shinozaki, S. Urao, K. Shinozaki, An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence, Plant Cell. 5 (1993) 1529–1539. doi:10.1105/tpc.5.11.1529.
[28] A.J. White, M. Alison Dunn, K. Brown, M.A. Hughes, Comparative analysis of genomic sequence and expression of a lipid transfer protein gene family in winter barley, J. Exp. Bot. 45 (1994) 1885–1892. doi:10.1093/jxb/45.12.1885.
[29] J. Mundy, N.H. Chua, Abscisic acid and water-stress induce the expression of a novel rice gene, EMBO J. 7 (1988) 2279–2286.
[30] K. Kazan, R. Lyons, Intervention of Phytohormone Pathways by Pathogen Effectors, Plant Cell. 26 (2014) 2285–2309. doi:10.1105/tpc.114.125419.
[31] Y. Xu, P. Chang, D. Liu, M.L. Narasimhan, K.G. Raghothama, P.M. Hasegawa, R.A. Bressan, Plant defense genes are synergistically induced by ethylene and methyl jasmonate, Plant Cell. 6 (1994) 1077–1085. doi:10.1105/tpc.6.8.1077.
[32] T.P. Delaney, S. Uknes, B. Vernooij, L. Friedrich, K. Weymann, D. Negrotto, T. Gaffney, M. Gut-Rella, H. Kessmann, E. Ward, J. Ryals, A central role of salicylic Acid in plant disease resistance, Science (80-.). 266 (1994) 1247–1250. doi:10.1126/science.266.5188.1247.
[33] A. Koornneef, C.M.J. Pieterse, Cross Talk in Defense Signaling, Plant Physiol. 146 (2008) 839–844. doi:10.1104/pp.107.112029.

[34] E.J.M. Van Damme, A. Barre, P. Rougé, W.J. Peumans, Cytoplasmic/nuclear plant lectins: a new story, Trends Plant Sci. 9 (2004) 484–489. doi:10.1016/j.tplants.2004.08.003.

[35] N. Lannoo, E.J.M. Van Damme, Nucleocytoplasmic plant lectins, Biochim. Biophys. Acta - Gen. Subj. 1800 (2010) 190–201. doi:10.1016/j.bbagen.2009.07.021.

[36] M.L.R. Macedo, C.F.R. Oliveira, C.T. Oliveira, Insecticidal activity of plant lectins and potential application in crop protection, Molecules. 20 (2015) 2014–33. doi:10.3390/molecules20022014.

[37] B. Saeed, V.K. Baranwal, P. Khurana, Identification and Expression Profiling of the Lectin Gene Superfamily in Mulberry, Plant Genome. (2016) 1–13. doi:10.3835/plantgenome2015.10.0107.

[38] R.E. Kendrick, L.H.J. Kerckhoffs, A. Van Tuinen, M. Koornneef, Photomorphogenic mutants of tomato, Plant Cell Environ. 20 (1997) 746–751. doi:10.1046/j.1365-3040.1997.d01-109.x.

[39] W.B. Terzaghi, A.R. Cashmore, Light-Regulated Transcription, Annu. Rev. Plant Physiol. Plant Mol. Biol. 46 (1995) 445–474. doi:10.1146/annurev.pp.46.060195.002305.

[40] S. Tanaka, N. Mochizuki, A. Nagatani, Expression of the AtGH3a Gene, an Arabidopsis Homologue of the Soybean GH3 Gene, is Regulated by Phytochrome B, Plant Cell Physiol. 43 (2002) 281–289. doi:10.1093/pcp/pcf033.

[41] K.-H. Jung, J. Lee, C. Dardick, Y.-S. Seo, P. Cao, P. Canlas, J. Phetsom, X. Xu, S. Ouyang, K. An, Y.-J. Cho, G.-C. Lee, Y. Lee, G. An, P.C. Ronald, Identification and Functional Analysis of Light-Responsive Unique Genes and Gene Family Members in Rice, PLoS Genet. 4 (2008) e1000164. doi:10.1371/journal.pgen.1000164.

[42] K. De Schutter, E. Van Damme, Protein-Carbohydrate Interactions as Part of Plant Defense and Animal Immunity, Molecules. 20 (2015) 9029–9053. doi:10.3390/molecules20059029.

[43] M.E. Etzler, A. Surolia, R.D. Cummings, L-type Lectins, in: Essentials Glycobiol., 2009: pp. 415–424.

[44] N. Sharon, H. Lis, How proteins bind carbohydrates: Lessons from legume lectins, J. Agric. Food Chem. 50 (2002) 6586–6591. doi:10.1021/jf020190s.

Table S1. List of primers used in this study.

Name	Forward 5'–3'	Reverse 5'–3'	amplicon size	Primer use
BolDeg1	GARACNMGNGCNTTYMGNGAR	TGTTGGTCACAGAGACCTTAAC	~850 bp	cloning
BolDeg2	GGAGCGACCATCAAGTAAGATAG	CTTAACCATTTCAGCTTCAAAGAT	~850 bp	cloning
BOLSpf	ATGGATAGTCACAAAAGGAGTTTG	CTAAACGGAGACGATGTTGG	1053 bp	cloning
BOLsRT	ACAGATGGAACCGTGAAGCGA	CGCTTCGAATAGGGTACCAAGC	478 bp	SQ-PCR
UBIQ2	ATATTCGTGAAGACGCTG	CTCAACTGGTTGCTGTG	359 bp	SQ-PCR
BOLqRT	CGTGAGGATGGAAGAGACCGA	AGCGTTGTATCGCTTCACGG	133 bp	RT-qPCR
UBQ9	GAAGACATGTTCCATTGGCA	ACACCTTAGTCCTAAAAGCCACCT	160 bp	RT-qPCR

SQ-PCR: semi quantitative PCR; RT-qPCR: quantitative PCR; bp base pair

```
+ TAAATTCATA ATTTGAATTT TTTTTTTTTT TTACATTTTT CGGCTAAGAG ACGTTTCTTA TATCTCTTAT
- ATTTAAGTAT TAAACTTAAA AAAAAAAAAA AATGTAAAAA GCCGATTCTC TGCAAAGAAT ATAGAGAATA

+ TTCAGAAACA GGTTTTCTAT TTTTTAGATA AAAATCTACG GAAAGCTAAT AACCGTCTTT TAGTCGAAAC
- AAGTCTTTGT CCAAAAGATA AAAAATCTAT TTTTAGATGC CTTTCGATTA TTGGCAGAAA ATCAGCTTTG

+ TAAAAACTCC AGTTAAGAGA TTGAGATTAA TTATGGTCTA AGAATGGTTT CCTCCGAGTT GTAGGATTGA
- ATTTTTGAGG TCAATTCTCT AACTCTAATT AATACCAGAT TCTTACCAAA GGAGGCTCAA CATCCTAACT

+ TCAAAACTTG GGTTTGCATC AGTAGTAACC ACAATATTTT TTACAAGAAA ACAAGTTATC ATTCTTAATC
- AGTTTTGAAC CCAAACGTAG TCATCATTGG TGTTATAAAA AATGTTCTTT TGTTCAATAG TAAGAATTAG

+ AAGTCTTTGA ACCCAACAAG GAATCCACTT ACTGGATGTA AGTTTTTTAC TTTAAAATAA CAAAATTGGA
- TTCAGAAACT TGGGTTGTTC CTTAGGTGAA TGACCTACAT TCAAAAAATG AAATTTTATT GTTTTAACCT

+ GACCGTCTTC TCGATATTTC CCAAGTAGAC CCAAAGGTTT TTAAATTGAA CCTGTGAATT GTTACTTTGT
- CTGGCAGAAG AGCTATAAAG GGTTCATCTG GGTTTCCAAA AATTTAACTT GGACACTTAA CAATGAAACA

+ CTGTCTATAC AAAAGTTGAT TGAAAATCAA TACATGAATA TTTTATAATT CTTTGATGAT AATGATATTA
- GACAGATATG TTTTCAACTA ACTTTTAGTT ATGTACTTAT AAAATATTAA GAAACTACTA TTACTATAAT

+ CCTTTGACAT TGGTTCGAAG TCGAACACGC AATTATTTAA GAATCTTTTT TTTTACCTCA ATTATTTAAG
- GGAAACTGTA ACCAAGCTTC AGCTTGTGCG TTAATAAATT CTTAGAAAAA AAAATGGAGT TAATAAATTC

+ TATCTTACTT TAGATGTTCT TCATCTAAGA TTTTCGGCAA TTATTGATTC GAACCTGCCA CTCACCACGT
- ATAGAATGAA ATCTACAAGA AGTAGATTCT AAAAGCCGTT AATAACTAAG CTTGGACGGT GAGTGGTGCA

+ ACTTCATATA GATCATCCAG AAATCCATGC CCATATTGTT TGTTAATAAA ATTGGGGGTG GGGGTTAGAA
- TGAAGTATAT CTAGTAGGTC TTTAGGTACG GGTATAACAA ACAATTATTT TAACCCCCAC CCCCAATCTT

+ TTAAAAAAGC ATAAAACAAC AATAACCAAA ATAATCAATT GTTTTTTTGA TGATTAATTG TCTTATGAAT
- AATTTTTTCG TATTTTGTTG TTATTGGTTT TATTAGTTAA CAAAAAAACT ACTAATTAAC AGAATACTTA

+ ATTTCAACTA AAACCCTGGA GTTTTTAATG CCTTCTTCAA TTCTCTACCT GTCATCATAC ATAGGACTTT
- TAAAGTTGAT TTTGGGACCT CAAAAATTAC GGAAGAAGTT AAGAGATGGA CAGTAGTATG TATCCTGAAA

+ ACGAGAATCT CCTTTTGTAA AGCCTCAATC ATTTTCTCTT TACTAAATTT CAACTTCAAC CTAAAAAAAA
- TGCTCTTAGA GGAAAACATT TCGGAGTTAG TAAAAGAGAA ATGATTTAAA GTTGAAGTTG GATTTTTTTT

+ CGTGTAGAAA TGTTTTATTT AGGAGTATTT TATATTAGTA ACTAACTCAA TAGGTGCCCT GATGTGGCTA
- GCACATCTTT ACAAAATAAA TCCTCATAAA ATATAATCAT TGATTGAGTT ATCCACGGGA CTACACCGAT

+ TCTATATAAA TGTAAAGC
- AGATATATTT ACATTTCG
```

Figure S1. Predicted *Bol* promoter nucleotide sequence of *Brassica rapa* FPsc v1.3 extracted using Phytozome database. Motifs with significant similarity to the previously identified cis-acting elements are shaded in different colors and the correspondent names are given bellow.

Motifs Found

AAGAA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
AAGAA-motif	Avena sativa	538	-	9	gGTAAAGAAA	

ABRE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ABRE	Oryza sativa	624	-	9	AGTACGTGGC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	626	-	6	TACGTG	cis-acting element involved in the abscisic acid responsiveness

ACE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ACE	Petroselinum crispum	907	+	9	AAAACGTTTA	cis-acting element involved in light responsiveness

ARE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ARE	Zea mays	185	+	6	TGGTTT	cis-acting regulatory element essential for the anaerobic induction

AT-rich sequence

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
AT-rich sequence	Pisum sativum	934	-	9	TAAAATACT	element for maximal elicitor-mediated activation (2copies)

Box 4

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Box 4	Petroselinum crispum	166	+	6	ATTAAT	part of a conserved DNA module involved in light responsiveness
Box 4	Petroselinum crispum	753	-	6	ATTAAT	part of a conserved DNA module involved in light responsiveness

Box III

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Box III	Pisum sativum	868	+	11	atCATTTTCACt	protein binding site

CAAT-box

CAT-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
CAT-box	Arabidopsis thaliana	617	+	6	GCCACT	cis-acting regulatory element related to meristem expression

+ G-Box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-Box	Antirrhinum majus	626	+	6	CACGTA	cis-acting regulatory element involved in light responsiveness
G-Box	Pisum sativum	909	-	6	CACGTT	cis-acting regulatory element involved in light responsiveness

+ G-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-box	Daucus carota	626	-	6	TACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Zea mays	909	-	6	CACGTT	cis-acting regulatory element involved in light responsiveness

+ HD-Zip 1

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
HD-Zip 1	Arabidopsis thaliana	598	-	8.5	CAAT(A/T)ATTG	element involved in differentiation of the palisade mesophyll cells

+ HD-Zip 2

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
HD-Zip 2	Arabidopsis thaliana	598	-	8	CAAT(G/C)ATTG	element involved in the control of leaf morphology development

+ HSE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
HSE	Brassica oleracea	15	-	9	AAAAAATTTC	cis-acting element involved in heat stress responsiveness
HSE	Brassica oleracea	917	-	9	AAAAAATTTC	cis-acting element involved in heat stress responsiveness

+ LTR

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
LTR	Hordeum vulgare	38	-	6	CCGAAA	cis-acting element involved in low-temperature responsiveness
LTR	Hordeum vulgare	592	-	6	CCGAAA	cis-acting element involved in low-temperature responsiveness

+ MBS

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
MBS	Arabidopsis thaliana	150	-	6	TAACTG	MYB binding site involved in drought-inducibility

+ MBSI

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
MBSI	Petunia hybrida	120	-	10.5	aaaAaaC(G/C)GTTA	MYB binding site involved in flavonoid biosynthetic genes regulation

+ MNF1

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
MNF1	Zea mays	964	+	6.5	GTGCCC(A/T)(A/T)	light responsive element

+ MRE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
MRE	Petroselinum crispum	898	+	7	AACCTAA	MYB binding site involved in light responsiveness

+ P-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
P-box	Oryza sativa	851	+	7	CCTTTTG	gibberellin-responsive element

+ Skn-1_motif

+ Sp1

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Sp1	Zea mays	686	-	5.5	CC(G/A)CCC	light responsive element

+ TATA-box

+ TC-rich repeats

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
TC-rich repeats	Nicotiana tabacum	574	+	9	ATTTTCTTCA	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	Nicotiana tabacum	810	+	9	ATTCTCTAAC	cis-acting element involved in defense and stress responsiveness

+ TCT-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
TCT-motif	Arabidopsis thaliana	563	+	6	TCTTAC	part of a light responsive element

+ Unnamed__1

+ Unnamed__3

+ Unnamed__4

+ as-2-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
as-2-box	Nicotiana tabacum	478	+	9	GATAatGATG	involved in shoot-specific expression and light responsiveness

+ box E

+ circadian

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
circadian	Lycopersicon esculentum	727	+	6	CAANNNNATC	cis-acting regulatory element involved in circadian control

GENERAL CONCLUSIONS

This is the first report of the isolation of a lectin from *Brassica oleracea* ssp. *botrytis* (BOL). BOL is a monomeric lectin, non-glycosylated with a molecular mass of ~34 kDa. It is active at the temperatures from 4ºC up to 60ºC, pH 7.0 and 8.0. In this study we purified, characterised and evaluated the stimulatory effects of BOL, which demonstrated be able activate macrophage improve their clearance capacity. It is also described for the first time a TRAF-like protein with lectin activity, supporting the concept that the lectins are indeed multifunctional and diverse group. The new lectin isolated from cauliflower can favoring the removal of foreign agents, which is a potentially exploitable activity.

The *Bol* cDNA was successfully cloned and expressed using the His-tag pET21b(+) vector system. The full coding sequence confirmed identity with TRAF proteins and the resulting recombinant protein retained the ability to recognize carbohydrates, thereby BOL is the first TRAF protein with lectin activity. Furthermore, *Bol* promoter analysis identified motifs involved in response to environmental stress and *Bol* transcript was changed under biotic stress simulation, suggesting that BOL may play a role in defense response regulation of plants. The heterologous expression of the recombinant protein in *E. coli* as showed here is a promising tool for large-scale production of BOL, enabling in advancing our understanding of this intriguing TRAF-like protein with lectin activity.