



UNIVERSIDAD NACIONAL DE COLOMBIA

VICERRECTORÍA DE INVESTIGACIÓN
DIRECCIÓN DE INVESTIGACIÓN SEDE BOGOTÁ

Elucidación parcial de la estructura primaria de la lectina LGL-P2 y purificación y caracterización parcial de la lectina LGL-P4 presentes en semillas de *Galactia lindenii*.

Partial primary structure elucidation of the GLL-P2 lectin and partial purification and characterization of GLL-P4 lectin present in *Galactia lindenii* seeds.

Resumen

Las lectinas son proteínas que se encuentran ampliamente distribuidas en la naturaleza, muchas de ellas han sido aisladas y purificadas de la familia de las Leguminosas, Gramíneas y Lamiaceas entre otras. Estas proteínas reconocen carbohidratos con gran afinidad y especificidad y por lo tanto son unas herramientas poderosas para estudiar la relación entre los glicanos de la superficie celular y las funciones celulares. En la subtribu *Diocleinae*, el grupo de investigación en proteínas (GRIP) ha detectado, aislado y caracterizado dos lectinas diferentes (P2 y P4) de *Diocleae lehmanni* (LDL), *Dioclea sericea* (LDS), *Dioclea grandiflora* (LDG), *Canavalia ensiformis* (LCE) y *Galactia lindenii* (LGL). Las dos lectinas están dentro de los cuerpos proteicos en las semillas y tienen diferentes propiedades fisicoquímicas. La LGL-P2 tiene alta afinidad por el determinante H-2 (Fuc α -(1-2) β -D-Gal(1-4) β -D-GlcNAc) y por residuos de lactosa. Para determinar la estructura primaria de la LGL-P2, la proteína pura fue hidrolizada con tripsina, quimotripsina y Lys-C; los péptidos obtenidos de cada una de las digestiones fueron secuenciados por espectrometría de masas. Por otra parte la LGL-P4 fue aislada de semillas maduras y las extracciones se hicieron con PBS conteniendo tiourea 5 mM. La proteína fue purificada sobre Sephacryl S-200 y la fracción retenida fue eluida con Glucosa 0.2 M. La lectina pura aglutinó eritrocitos humanos y el perfil electroforético fue similar al obtenido para otras lectinas P4 de otros géneros. La proteína es una glicoproteína con un punto isoeléctrico 6.5. La determinación cuantitativa (mg lectina/ 100 g semilla) de estas dos lectinas (P2 y P4) nos permitió ver que se presenta una diferencia en el contenido de proteína en las semillas dentro de los diferentes géneros de la subtribu *Diocleinae*.

Palabras clave: *Diocleinae*, lectina, estructura, péptidos, purificación, propiedades.

Abstract

Lectins are widespread in nature and most of them have been isolated and characterized from Leguminosae, Gramineae and Lamiaceae families among others. These proteins have the ability to recognize carbohydrates with high affinity and specificity, therefore, they are a powerful tool to establish the relationship between glycan structure and cell-specific functions. In the subtribe *Diocleinae*, PRG has detected, isolated and characterized two different lectins (P2 and P4) from *Diocleae lehmanni* (DLL), *Dioclea sericea* (DSL), *Dioclea grandiflora* (DGL), *Canavalia ensiformis* (CEL) and *Galactia lindenii* (GLL). These lectins are present in the same seeds and have different physicochemical properties; this allows us to question about the physiological role of these proteins. GLL-P2 has high affinity towards H-Type 2 antigen (Fuc $\alpha(1-2)\beta$ -D-Gal(1-4) β -D-GlcNAc) and lactose residues. In order to determine partial primary sequence of GLL-P2, the pure protein was hydrolyzed using Trypsin, chymotrypsin and Lys-C; the sequence of peptides was achieved using mass spectrometry (MSn). To isolate GLL-P4 assays were carried out using mature seeds; the crude extract was obtained using PBS 5 mM Thiourea and after that the protein was purified on Sephacryl S-200 by eluting it with Glucosa 0.2 M. GLL-P4 lectin recognized Glc/Man and agglutinated human red blood cells. SDS-PAGE profile was similar to other P4 lectins from other species. It was also determined that it is a glycoprotein with an isoelectric point 6.5. The quantitative determination of these two types of lectins allowed us to establish that there is a different lectin content in *Diocleinae* genders.

Key words: *Diocleinae*, lectin, structure, peptides, purification, properties.

Elucidación parcial de la estructura primaria de la lectina LGL-P2 y purificación y caracterización parcial de la lectina LGL-P4 presentes en semillas de *Galactia lindenii*.

Mónica Quintero Moreno

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de:

Magíster en Ciencias Bioquímica.

Director (a):

Nohora Angélica Vega Castro Dr Sc

Línea de Investigación:

Bioquímica de Proteínas

Grupo de Investigación:

Grupo de Investigación en Proteínas (GRIP)

Universidad Nacional de Colombia

Facultad de Ciencias, Escuela de Química

Bogotá, Colombia

2014

A todas las personas que durante este camino me acompañaron de una u otra forma, ya fuera con su alegría, sus consejos o su presencia y constante apoyo.

“Un sociólogo norteamericano dijo hace más de treinta años que la propaganda era una formidable vendedora de sueños, pero resulta que yo no quiero que me vendan sueños ajenos, sino sencillamente que se cumplan los míos”.

Mario Benedetti.

Agradecimientos

A Dios, por mantener en mí ese sentimiento de alegría, tranquilidad y serenidad y por bendecir cada una de las cosas que realicé para culminar esta parte de mi proyecto de vida.

A Colciencias y su apoyo al proyecto “Estructura de lectinas de Lamiaceae, primera fase”. Código 110148925106.

A mi Madre, por haberme educado y soportar mis errores. Gracias por tus consejos, por el amor que siempre me has brindado, por cultivar e inculcar ese sabio don de la responsabilidad, la dedicación, el trabajo incansable y la confianza en mí misma.

A mi Padre, por haberme dado la comprensión, tranquilidad y confianza que me ayudaron a superar los momentos difíciles en el desarrollo de este trabajo.

A mi hermano, por hacerme reír siempre y ayudarme a olvidar mis preocupaciones, por ser un gran amigo.

A Daniel Orozco, porque sin tu compañía no hubiera podido superar las dificultades externas, gracias por estar ahí y por nunca dejarme desfallecer.

Al Dr Mario Amzel, y los compañeros de la Facultad de Biofísica y Química Biofísica de la Universidad Johns Hopkins por su apoyo.

A mis compañeros, Andrea, Edwin, Nury y Camilo, porque con su felicidad, confianza y apoyo continuo, era más fácil superar las dificultades diarias que se presentaban. Muchas gracias por su compañía.

A mi profesor, Edgar Reyes, por el apoyo, los consejos y su compañía en el desarrollo de mi tesis.

A mi Directora de Tesis, Nohora Vega, porque gracias a sus consejos, enseñanzas, comprensión, apoyo, paciencia y dedicación, logramos un muy buen resultado, del cual me siento orgullosa, y por el que trabajamos arduamente.

Resumen

Las lectinas son proteínas que se encuentran ampliamente distribuidas en la naturaleza, muchas de ellas han sido aisladas y purificadas de la familia de las Leguminosas, Gramíneas y Lamiaceas entre otras. Estas proteínas reconocen carbohidratos con gran afinidad y especificidad y por lo tanto son unas herramientas poderosas para estudiar la relación entre los glicanos de la superficie celular y las funciones celulares. En la subtribu *Diocleinae*, el grupo de investigación en proteínas (GRIP) ha detectado, aislado y caracterizado dos lectinas diferentes (P2 y P4) de *Diocleae lehmanni* (LDL), *Dioclea sericea* (LDS), *Dioclea grandiflora* (LDG), *Canavalia ensiformis* (LCE) y *Galactia lindenii* (LGL). Las dos lectinas están dentro de los cuerpos proteicos en las semillas y tienen diferentes propiedades fisicoquímicas. La LGL-P2 tiene alta afinidad por el determinante H-2 (Fuc α -(1-2) β -D-Gal(1-4) β -D-GlcNAc) y por residuos de lactosa. Para determinar la estructura primaria de la LGL-P2, la proteína pura fue hidrolizada con tripsina, quimotripsina y Lys-C; los péptidos obtenidos de cada una de las digestiones fueron secuenciados por espectrometría de masas. Por otra parte la LGL-P4 fue aislada de semillas maduras y las extracciones se hicieron con PBS conteniendo tiourea 5 mM. La proteína fue purificada sobre Sephacryl S-200 y la fracción retenida fue eluida con Glucosa 0.2 M. La lectina pura aglutinó eritrocitos humanos y el perfil electroforético fue similar al obtenido para otras lectinas P4 de otros géneros. La proteína es una glicoproteína con un punto isoeléctrico 6.5. La determinación cuantitativa (mg lectina/ 100 g semilla) de estas dos lectinas (P2 y P4) nos permitió ver que se presenta una diferencia en el contenido de proteína en las semillas dentro de los diferentes géneros de la subtribu *Diocleinae*.

Palabras clave: *Diocleinae*, lectina, estructura, péptidos, purificación, propiedades.

Abstract

Lectins are widespread in nature and most of them have been isolated and characterized from Leguminosae, Gramineae and Lamiaceae families among others. These proteins have the ability to recognize carbohydrates with high affinity and specificity, therefore, they are a powerful tool to establish the relationship between glycan structure and cell-specific functions. In the subtribe *Diocleinae*, PRG has detected, isolated and characterized two different lectins (P2 and P4) from *Diocleae lehmanni* (DLL), *Dioclea sericea* (DSL), *Dioclea grandiflora* (DGL), *Canavalia ensiformis* (CEL) and *Galactia lindenii* (GLL). These lectins are present in the same seeds and have different physicochemical properties; this allows us to question about the physiological role of these proteins. GLL-P2 has high affinity towards H-Type 2 antigen (Fuc $\alpha(1-2)\beta$ -D-Gal(1-4) β -D-GlcNAc) and lactose residues. In order to determine partial primary sequence of GLL-P2, the pure protein was hydrolyzed using Trypsin, chymotrypsin and Lys-C; the sequence of peptides was achieved using mass spectrometry (MSn). To isolate GLL-P4 assays were carried out using mature seeds; the crude extract was obtained using PBS 5 mM Thiourea and after that the protein was purified on Sephacryl S-200 by eluting it with Glucosa 0.2 M. GLL-P4 lectin recognized Glc/Man and agglutinated human red blood cells. SDS-PAGE profile was similar to other P4 lectins from other species. It was also determined that it is a glycoprotein with an isoelectric point 6.5. The quantitative determination of these two types of lectins allowed us to establish that there is a different lectin content in *Diocleinae* genders.

Key words: *Diocleinae*, lectin, structure, peptides, purification, properties.

Contenido

	Pág.
Resumen	VII
Abstract	VIII
Lista de figuras	XII
Lista de tablas	XV
Lista de abreviaturas	XVII
Introducción.....	XXI
1. Marco teórico.....	1
1.1. Lectinas de origen vegetal.....	2
1.2. Lectinas de leguminosas.....	4
1.3. Género <i>Galactia</i> y lectinas del género <i>Galactia</i>	5
1.4. Determinación de la estructura primaria de lectinas.....	7
2. Objetivos	9
2.1. Objetivo general.....	9
2.2. Objetivos específicos.....	9
3. Materiales y Métodos	10
3.1. Extracción y purificación de la LGL-P2.....	10
3.2. Electroforesis preparativa para obtener el monómero de la LGL-P2.....	13
3.3. Ensayos de digestión enzimática para la LGL-P2.....	14
3.4. Secuenciación de péptidos de la LGL-P2 por espectrometría de masas.....	16
3.5. Determinación de la secuencia parcial de aminoácidos de la lectina P2 de <i>Galactia</i> <i>lindenii</i>	18
3.6. Ensayos exploratorios de cristalización de la LGL-P2.....	20
3.7. Purificación de la lectina de <i>Galactia lindenii</i> P4 (LGL-P4).....	24

3.8. Caracterización parcial de la lectina LGL-P4.....	25
3.9. Purificación de otras lectinas de la subtribu <i>Diocleinae</i>	26
4. Análisis de resultados	27
4.1. Extracción y purificación de la lectina P2 de <i>Galactia lindenii</i>	28
4.2. Electroforesis preparativa para obtener el monómero de la LGL-P2.....	33
4.3. Ensayos de digestión enzimática para la LGL-P2.....	37
4.4. Aproximación a la estructura primaria de LGL-P2.....	42
4.5. Análisis bioinformático de las secuencias de péptidos encontrados contra la secuencia de la LDL-P2.....	51
4.6. Estudios exploratorios de cristalización de la LGL-P2.....	67
4.7. Purificación de la LGL-P4.....	73
4.8. Caracterización parcial de la lectina LGL-P4.....	81
5. Conclusiones.....	88
6. Perspectivas.....	90
7. Aportes al estudio de lectinas.....	90
8. Anexos.....	922
1. Método del Ácido Bicinconínico (BCA) para cuantificación de proteínas.....	92
2. Preparación de geles SDS-PAGE en Tris-Tricina para separación de proteínas de bajo peso molecular.....	93
3. Tinción con plata de geles SDS-PAGE	94
4. Características básicas de los kits utilizados en los ensayos de cristalización..	95
5. Detección cualitativa de glicoproteínas.....	96
6. Isoelectroenfoque de la LGL-P4	97
7. Perfiles de purificación de las lectinas P4 de <i>Dioclea lehmanni</i> , <i>D. grandiflora</i> y <i>Canavalia ensiformis</i>	98
8. Electroforesis del proceso de purificación de lectinas P2 y P4.....	100
9. Composición de aminoácidos LGL-P2	101

10. Resultados identificación de la muestra con base en los péptidos obtenidos en digestión con tripsina contra base de datos PD1.3/Mascot protein/peptide	102
11. Resultados identificación de la muestra con base en los péptidos obtenidos en ...gestión con quimotripsina contra base de datos PD1.3/Mascot protein/peptide.....	103
12. Resultados identificación de la muestra con base en los péptidos obtenidos en digestión con LysC contra base de datos PD1.3/Mascot protein/peptide	104
13. Alineamiento de la LDL-P2 con la LCR-II.....	105
14. PROCHECK statistics: LGL-P2	106
15. PROCHECK statistics: LDL-P2	107
16. Predicción estructura secundaria para LGL-P2.....	108
17. Curva de calibración pI (punto isoeléctrico).....	109
9. Bibliografía.....	111

Lista de figuras

	Pág.
Figura 1: Cromatografía sobre DEAE-Sephacel. Extracto crudo (15 mg de proteína total).....	[30].
Figura 2: Cromatografía sobre Sepharosa 4B-lactosa. Fracción I DEAE –Sephacel (93 mg de proteína total).....	[31].
Figura 3: SDS-PAGE purificación de la lectina de <i>Galactia lindenii</i> P2.....	[31].
Figura 4: Recromatografía para la fracción de LGL-P2.....	[32].
Figura 5: Electroforesis preparativa. Curva de calibración.....	[34].
Figura 6: Electroforesis preparativa (Patrones)	[35].
Figura 7: Electroforesis preparativa de LGL-P2.. ..	[36].
Figura 8: Electroforesis Tris-Tricina para péptidos derivados de Tripsina (20 µg/ml).....	[38].
Figura 9: Electroforesis Tris-Tricina, tripsina, quimotripsinas y pepsina.....	[39].
Figura 10: SDS-PAGE LGL-P2 para análisis en LC/MS/MS.....	[41].
Figura 11: Resultados identificación de la muestra con base en los péptidos obtenidos utilizando BLAST (blastp) con los péptidos obtenidos con: A: Tripsina B: Quimotripsina. C: Lys C.....	[48].
Figura 12: Resultados alineamiento de péptidos obtenidos con tripsina contra la secuencia de Cicer Arietinum utilizando BLAST (blastp).. ..	[44].
Figura 13: Secuencia parcial de la lectina P2 de <i>Dioclea lehmanii</i> (LDL-P2).....	[51].
Figura 14: Dominios encontrados en la comparación de la secuencia parcial de la LGL-P2 utilizando BLAST	[52].

Figura 15: Resultados comparación de la secuencia parcial de la LDL-P2 con los péptidos A: trípticos B: Quimotripticos C: Lys C obtenidos de LGL-P2 utilizando BLAST.....	[53].
Figura 16: Secuencia parcial con variaciones de la LGL-P2	[54].
Figura 17: Superposición de péptidos ensamblados para determinar secuencia LGL-P2.....	[56].
Figura 18: Comparación de las tres secuencias de Lectinas tipo P2 de la subtribu <i>Diocleinae</i>	[57].
Figura 19: Secuencia parcial de la lectina P2 de <i>Dioclea lehmanni</i>	[61].
Figura 20: Secuencia parcial de la LGL-P2.	[62].
Figura 21: LDL-P2: 0.85 ± 0.08 (TM-score): $3.3\pm 2.3\text{Å}$ (RMSD).	[63].
Figura 22: LGL-P2: 0.79 ± 0.09 (TM-score): $4.1\pm 2.7\text{Å}$ (RMSD).	[63].
Figura 23: Diagrama de Ramachandran LDL-P2.	[64].
Figura 24: Diagrama de Ramachandran LGL-P2	[64].
Figura 25: Topología estructural para el modelo estructural de LDL-P2.....	[64].
Figura 26: Monómero típico de una lectina de leguminosa (Lectina de lenteja, PDB: 1LES, Casset et al., 1995).	[65].
Figura 27: Sitio de unión para LDL-P2: Metil- β -galactosa.	[66].
Figura 28: Sitio de unión para LGL-P2: Metil- β -galactosa.	[66].
Figura 29: Ensayos preliminares- Gota de cristalización.....	[67].
Figura 30: Gotas de cristalización positivas.....	[69].
Figura 31: Difracción de rayos X. Cristal en condición 3 del kit Crystal Screen I-II.....	[69].

Figura 32: Optimización de condiciones.....	[70].
Figura 33: Optimización de condiciones	[71].
Figura 34: Imagen de difracción de rayos X Cristal en gota C4:1 µl del ensayo 3 de optimización (pH 8.5, 62% MPD).....	[71].
Figura 35: Seeding de cristales	[72].
Figura 36: Lectinas de la Subtribu <i>Diocleinae</i>	[74].
Figura 37: Purificación de la LGL-P4.....	[75].
Figura 38: Cromatografía de afinidad de LGL-P4 sobre Sephacryl S-200.....	[76].
Figura 39: SDS-PAGE –LGL-P4.....	[78]
Figura 40: SDS-PAGE: LGL-P4 (Detección de bandas)	[79]
Figura 41: Determinación del carácter glicoproteico para LGL-P4.....	[84].
Figura 42: Determinación del Punto isoeléctrico para LGL-P4.....	[85].

Lista de tablas

	Pág.
Tabla 1: Enzimas empleadas en la hidrólisis de la LGL-P2.....	[15]
Tabla 2: Condiciones utilizadas para el ensayo base de cristalización	[21].
Tabla 3: Ensayos con fosfato de amonio como agente precipitante sin adición de buffer ni sales al sistema de reservorio	[22].
Tabla 4: Ensayos de optimización con 0.2 M fosfato de amonio monobásico como sal, Tris-HCl 0.1 M a diferentes pHs, y MPD a diferentes concentraciones.....	[22].
Tabla 5: Ensayos de optimización variando la concentración de MPD con respecto a las anteriores.. ..	[23].
Tabla 6: Extracción de proteína a partir de la harina de <i>Galactia lindenii</i>	[29].
Tabla 7: Purificación de la lectina de <i>Galactia lindenii</i> P2 (LGL-P2).....	[32].
Tabla 8: Secuencias de péptidos obtenidos con tripsina por LC/MS/MS de LGL- P2.....	[42].
Tabla 9: Péptidos obtenidos con quimotripsina por LC/MS/MS de LGL-P2.....	[43].
Tabla 10: Péptidos obtenidos con LysC por LC/MS/MS de LGL-P2.....	[46].
Tabla 11: Propiedades moleculares de algunas lectinas de la Subtribu <i>Diocleinae</i> que reconocen Lactosa y el determinante H2.....	[50]
Tabla 12: Secuencia de los péptidos seleccionados para la determinación de estructura primaria de LGL-P2.....	[55].
Tabla 13: Comparación del análisis de aminoácidos para LGL-P2 Y LDL-P2.....	[59].

Tabla 14: Purificación de la LGL-P4.....	[77]
Tabla 15: Contenido de las lectinas P2 y P4 en las especies estudiadas.....	[80].
Tabla 16: Evaluación de la inhibición de la actividad de la LGL-P4 con azúcares	[82].
Tabla 17: Mínima concentración inhibitoria de carbohidratos para la LGL-P.....	[83].
Tabla 18: Propiedades moleculares de algunas lectinas P4 Subtribu <i>Diocleinae</i> .	[86].
Tabla 19: Secuencia N-terminal Lectinas subtribu <i>Diocleinae</i>	[86].
Tabla 20: Preparación del gel SDS-PAGE en Tris-Tricina.....	[93].
Tabla 21: Preparación buffer de muestra y buffer de corrido geles SDS-PAGE en Tris-Tricina.	[93].
Tabla 22: Preparación del gel para isoelectroenfoque.....	[97].
Tabla 23: Composición aminoácidos LGL-P2.	[101].

Lista de abreviaturas

Abreviatura	Término
ACN	Acetonitrilo
ADN	Ácido Desoxirribonucleico
ADNc	ADN Copia
BCA	Ácido Bicinconico
BLAST	Basic Local Alignment Search Tool
BSA	Albumina de Suero Bovino
Ca ²⁺	Iones de calcio
CaCl ₂	Cloruro de Calcio
CEL	<i>Canavalia ensiformis</i> lectin
CID	Collision - induced dissociation
DEAE	Dietilaminoetilo
DGL	<i>Dioclea grandiflora</i> lectin
DLL	<i>Dioclea lehmanni</i> lectin
DLS	<i>Dioclea sericea</i> lectin
DTT	Ditiotreitolo
ECL	<i>Erythrina cristagalli</i> lectin

Elucidación parcial de la estructura primaria de la lectina LGL-P2 y purificación y caracterización parcial de la lectina LGL-P4 presentes en semillas de *Galactia lindenii*.

<i>ECorL</i>	<i>Erythrina corallodendron</i> lectin
<i>EM</i>	Espectrometría de Masas
<i>ESI-MS</i>	Electrospray Ionization-Mass Spectroscopy
<i>ExPASy</i>	Expert Protein Analysis System
<i>Fuc</i>	Fucosa
<i>Gal</i>	Galactosa
<i>GalNAc</i>	N-acetilgalactosamina
<i>GlcNAc</i>	N-acetilglucosamina
<i>GLL</i>	<i>Galactia lindenii</i> lectin
<i>Glu</i>	Glucosa
<i>GRIP</i>	Grupo de Investigación en Proteínas
<i>GTasL</i>	<i>Galactia tashiroi</i> lectin
<i>GTenL</i>	<i>Galactia tenuiflora</i> lectin
<i>H2</i>	Determinante H tipo 2 (Fuc $\alpha(1-2)\beta$ -D-Gal(1-4) β -D-GlcNAc)
<i>kDa</i>	Kilo Daltons
<i>LCE</i>	Lectina de <i>Canavalia ensiformis</i>
<i>LCR-II</i>	Lectina II de <i>Cymbosema roseum</i>
<i>LDG</i>	Lectina de <i>Dioclea grandiflora</i>
<i>LDL</i>	Lectina de <i>Dioclea lehmanni</i>
<i>LDS</i>	Lectina de <i>Dioclea sericea</i>
<i>LGL</i>	Lectina de <i>Galactia lindenii</i>
<i>mA</i>	miliamperios
<i>Man</i>	Manosa
<i>Mn²⁺</i>	Manganeso (iones)

Elucidación parcial de la estructura primaria de la lectina LGL-P2 y purificación y caracterización parcial de la lectina LGL-P4 presentes en semillas de *Galactia lindenii*.

<i>MnCl₂</i>	Cloruro de Manganeso
<i>MPD</i>	(+/-)-2-metil-2,4-pentanodiol
<i>Mr</i>	Masa relativa
<i>MS</i>	Mass Spectroscopy
<i>MS/MS</i>	Espectrometría de Masas en Tandem
<i>MWCO</i>	Molecular Weight Cut Off
<i>NaCl</i>	Cloruro de Sodio
<i>Na₂HPO₄</i>	Fosfato Disódico
<i>NaH₂PO₄</i>	Fosfato Monosódico
<i>nano-LC-LTQ-Orbitrap-Velos en FTFT</i>	Nano-Cromatografía líquida – Trampa lineal en cuadrupolo-Orbitrap- Velos en Transformada de Fourier
<i>NaOH</i>	Hidroxido de Sodio
<i>pBLAST</i>	Protein BLAST
<i>PBS</i>	Buffer Fosfato Salino
<i>PDB</i>	Protein Data Bank
<i>PEG</i>	Polietilenglicol
<i>PM</i>	Peso Molecular
<i>PRG</i>	Protein Research Group
<i>PVDF</i>	Polivinilidenofluoruro
<i>SDS-PAGE</i>	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
<i>TCA</i>	Ácido tricloroacético
<i>TEMED</i>	Tetrametiletildiamino
<i>TPCK</i>	L-(tosilamido-2-fenil) etil clorometil cetona
<i>V</i>	Voltios
<i>V*H</i>	Voltios / Hora

Elucidación parcial de la estructura primaria de la lectina LGL-P2 y purificación y
caracterización parcial de la lectina LGL-P4 presentes en semillas de *Galactia*
lindenii.

W	Vatios
---	--------

Introducción

Las lectinas son un conjunto de proteínas de origen no inmune que reconocen carbohidratos con gran afinidad y especificidad, sin llegar a modificarlos (Lis y Sharon, 1986). Estas moléculas están ampliamente distribuidas en la naturaleza; se encuentran en virus, bacterias, plantas y animales. Se cree que las lectinas vegetales están implicadas en mecanismos de defensa contra microorganismos, almacenamiento de proteínas, transporte de carbohidratos, procesos de regulación fisiológica y de estimulación mitogénica (Gabiús et al, 2002), mientras que las lectinas de origen animal han sido relacionadas con funciones de chaperonas, direccionamiento intracelular de glicoconjugados, transporte intracelular, mediación de endocitosis, reconocimiento de glicanos foráneos, inducción o supresión de moléculas efectoras, entre otras. El estudio de las lectinas comenzó hace muchos años debido al efecto aglutinante de éstas sobre diversas células como eritrocitos, linfocitos, plaquetas, espermatozoides y células tumorales.

La glicosilación es la modificación covalente postraducciona l más frecuente en las proteínas, ocurre por la unión de una o más cadenas de oligosacáridos a la cadena polipeptídica, confiriéndole a las proteínas diferentes cualidades estructurales y funcionales (Sharon y Lis, 2004); este proceso no es un evento poco común en la célula, por el contrario, más de la mitad de las proteínas se encuentran en forma de glicoproteínas. El patrón de glicosilación de las proteínas es dependiente del tipo de célula en el que se produce y puede modificarse ante diversas situaciones biológicas. En este sentido, las lectinas han sido un instrumento valioso como moléculas de bioreconocimiento. Las propiedades descritas en lectinas, han convertido a estas moléculas en una herramienta fundamental en glicobiología, así mismo, su distribución, interacciones e importancia han sido ampliamente documentadas (Van Damme et

al, 1998). Como es el caso de los procesos de diferenciación celular y transformación maligna de células cancerosas, durante los cuales la biosíntesis de oligosacáridos puede verse afectada; y este tipo de anomalías celulares son detectables mediante el uso de lectinas (Sharon y Lis, 1986). Actualmente muchos estudios se han enfocado en caracterizar las interacciones proteína-carbohidrato por medio de Cromatografía Frontal de Afinidad (FAC) (Hirabayashi, 2003). Las lectinas presentes en las semillas de la familia *Leguminosae* (Leguminosas) han sido ampliamente estudiadas tanto a nivel estructural como nivel funcional (Rini, 1995). El grupo de Investigación en Proteínas (GRIP) se ha dedicado a la detección, purificación y caracterización de proteínas especialmente las lectinas provenientes de Leguminosas y Lamiaceas (Pérez, 1993; Almanza et al, 2004; Melgarejo et al, 2005; Pérez y Vega, 2007). Se han aislado y caracterizado lectinas de los géneros *Galactia* sp., *Dioclea* sp., *Salvia* sp. y *Lepechinia* sp entre otros. La lectina P2 de *Galactia lindenii* (LGL P2) ha sido purificada y caracterizada molecularmente y se ha delineado su especificidad por cromatografía frontal de afinidad observándose que reconoce específicamente el determinante H-2 (Fuc $\alpha(1-2)\beta$ -D-Gal(1-4) β -D-GlcNAc) y con menor afinidad la estructura glicosídica Core 3 (GlcNAc (1-3)- β -D-GalNAc) presente en glicoproteínas de tipo mucina (Vega et al, resultados no publicados); por lo tanto se podrían adelantar estudios enfocados al análisis de glicoproteínas que presenten el Core 3 en su estructura en células normales y tumorales.

Teniendo en cuenta que algunas de estas glicoproteínas son empleadas como marcadores tumorales y sirven para monitorear procesos metastáticos, es importante tener a la mano herramientas que permitan hacer su detección tanto en fluidos biológicos como en tejidos. Estas lectinas tienen un gran potencial como herramienta diagnóstica y pronóstica en clínica, pero aún no se cuenta con información sobre su estructura primaria, lo que dificulta la comprensión de sus mecanismos de acción a nivel molecular y el estudio de las interacciones con los carbohidratos. Las lectinas del género *Galactia* se caracterizan por reconocer específicamente el antígeno H-2, pero no hay suficiente información estructural que permita entender esta interacción; tampoco se sabe cómo las lectinas podrían ser herramientas que pueden interactuar con otras glicoproteínas. El establecimiento de la estructura terciaria podría explicar su especificidad, forma de

interacción con el antígeno H-2, y para ello es importante conocer la estructura primaria de estas proteínas.

Hasta la fecha no se conoce la estructura primaria de ninguna lectina del Género *Galactia*, en este trabajo por primera vez se presenta la secuencia parcial de la LGL-P2, la cual fue obtenida por superposición de péptidos y además se hizo una aproximación de su estructura terciaria empleando herramientas bioinformáticas. Por otra parte, se ha establecido la presencia de dos lectinas con diferente especificidad hacia carbohidratos dentro de la subtribu *Diocleinae* (Melgarejo et al, 2005; Almanza et al., 2004), en *Galactia lindenii* se detectó la LGL-P4, la cual fue purificada y caracterizada parcialmente, por primera vez en este género.

1. Marco teórico

El término lectina se aplica a proteínas o glicoproteínas de origen no inmune que reconocen de manera específica carbohidratos en superficie celular o en suspensión y tienen la capacidad de aglutinar células y precipitar glicoconjugados. Poseen por lo menos un sitio de reconocimiento a carbohidratos y se encuentran en mayor cantidad en las semillas de las plantas, especialmente en leguminosas y gramíneas (Azevedo-Moreira et al. 1991; Pusztai, 1991; Wright, 1997; Sharon y Lis, 2004; Hernández et al. 2005; Jiang et al. 2010; Ghazarian et al. 2011).

La habilidad que poseen las lectinas de plantas para diferenciar los grupos sanguíneos llevó a Boyd y Shapleigh en 1954 a proponer el término lectina, del latín “Legere” que significa seleccionar; posteriormente fue generalizado por Sharon como aglutininas independientemente de la fuente y la especificidad de grupo sanguíneo. El interés en las lectinas se incrementó luego de que se demostrara que son herramientas valiosas para la detección, aislamiento y caracterización de glicoconjugados así como la evaluación de los cambios en el perfil de glicosilación de los carbohidratos de la superficie celular, los cuales se alteran frente a diferentes procesos fisiológicos y patológicos.

Las actividades aglutinantes y precipitantes de las lectinas son similares a las de los anticuerpos; ambas pueden ser inhibidas específicamente por compuestos de bajo peso molecular (en el caso de las lectinas azúcares o haptenos que contienen azúcares), sin embargo, las lectinas difieren de los anticuerpos en varios aspectos importantes. Los anticuerpos son estructuralmente similares, mientras que las lectinas tienen diversidad estructural y son proteínas oligoméricas. Cada subunidad contiene el sitio de unión a

carbohidratos y varían en tamaño, composición de aminoácidos, requerimiento de metales, organización de dominios y especificidad, entre otros (Sharon y Lis, 2007).

Estas proteínas se encuentran ampliamente distribuidas en la naturaleza y han sido identificadas en diversos organismos como virus, bacterias, hongos, plantas, vertebrados superiores e invertebrados (Sharon y Lis, 2004; Hernández et al. 2005, Jiang et al. 2010). En Plantas, se conoce que no sólo están presentes en semillas, sino también en tejido vegetativo como raíces, hojas, tallo, corteza, flores y bulbos (Pusztai, 1991).

1.1 Lectinas de origen vegetal

La mayoría de lectinas de plantas están presentes en las semillas de las cotiledóneas, donde se encuentran ya sea en el citoplasma o en cuerpos proteicos (Azevedo- Moreira et al. 1991). Entre algunas de las funciones se han propuesto las siguientes hipótesis:

- Participan en las interacciones que se presentan entre las bacterias fijadoras de nitrógeno con la raíz de plantas de leguminosas.
- Pueden tener efectos protectores en contra de la acción patogénica de ciertos microorganismos como es el caso de los nemátodos herbívoros (Azevedo- Moreira et al. 1991, Hernández et al. 2005).
- Participan en procesos de reconocimiento que permiten la diferenciación celular, la organogénesis, la migración de linfocitos y como factores determinantes en la metástasis (Sharon y Lis, 2004, Ghazarian et al. 2011).

La principal característica de las lectinas vegetales es la especificidad de los sitios de unión a azúcares; lo que hace evidente que esta propiedad no es casual y que es relevante para la fisiología de las plantas siendo necesaria la existencia de receptores glicosídicos que serían el blanco de reconocimiento de las lectinas (Azevedo- Moreira et al. 1991). El gran impacto de las lectinas vegetales usadas en investigaciones biológicas y biomédicas se debe a dos razones básicas (Peumans y Van Damme, 1998):

- Las semillas de plantas son una buena fuente de lectinas.
- Son adecuadas para el aislamiento y análisis de glicoconjugados animales y humanos.

Las lectinas se clasifican según su sitio de reconocimiento a carbohidrato (Peumans y Van Damme, 1998, Jiang et al. 2010):

- Merolectinas: tienen un solo dominio de unión a carbohidratos por lo tanto son incapaces de precipitar glicoconjugados y/o aglutinar células.
- Hololectinas: Poseen dos dominios de unión a carbohidratos idénticos o muy similares, permitiéndoles precipitar glicoconjugados y/o aglutinar células. La mayoría de las lectinas de plantas aglutinantes son hololectinas.
- Superlectinas: compuestas al menos de dos dominios de unión a carbohidratos. A diferencia de las hololectinas estos dominios no son idénticos o similares pero reconocen azúcares estructuralmente diferentes.
- Quimerolectinas: son proteínas de fusión que poseen un dominio de unión a carbohidratos ubicado al azar y un dominio no relacionado. El último, posee una actividad biológica o catalítica bien definida pero con acción independiente al dominio de unión a carbohidratos.

Además de lo anterior, las lectinas vegetales se clasifican según: la especificidad hacia el monosacárido que inhibe su actividad hemaglutinante (Hernández et al. 2005), según el carbohidrato que reconocen (Manosa, N-acetilglucosamina/glucosa, N-acetilgalactosamina/galactosa, Fucosa, etc) (Azevedo- Moreira et al. 1991, Hernández et al. 2005); y según su estructura (Azevedo- Moreira et al. 1991; Wright, 1997; Peumans y Van Damme, 1998; Hernández et al. 2005; Jiang et al. 2010; Ghazarian et al. 2011). Sin embargo las lectinas de plantas se han clasificado molecularmente en 12 familias; 6 de las cuales ya habían sido descritas previamente: Lectinas de leguminosas, Lectinas tipo amaranto, Lectinas tipo jacalina (o prisma), Lectinas tipo heveína, Lectinas de monocotiledoneas específicas de manosa, Lectinas de quitinasa, Lectina de *Agaricus bisporus* (ABA), Cianovirinas, Lectina de *Euonymus europaeus* (EEA), Lectinas con motivos de lisina (LysM), Nictaba y la Familia de Ricina B. (Wright, 1997; Peumans y Van Damme, 1998; Hernández et al. 2005; Jiang et al. 2010; Ghazarian et al. 2011); hasta ahora pocos datos han sido reportados sobre la

caracterización general del genoma y la evolución molecular de la superfamilia de lectinas en plantas.

1.2 Lectinas de leguminosas

Se han detectado principalmente en semillas y representan hasta el 10 % de proteína soluble; en hojas, tallos y cortezas están en bajas concentraciones. Muchas de las especies de leguminosas poseen dos o más lectinas que se encuentran presentes en diferentes cantidades; en el caso de la especie *Ulex europaeus* hay dos lectinas con diferentes especificidades de unión a carbohidrato y en *Dolichos biflorus* hay 4 lectinas diferentes (Schnell y Etzler, 1981; Van Damme et al. 1998).

Las lectinas presentes en la familia *Leguminosae* han sido las más estudiadas, debido a su posible función de simbiosis entre leguminosas y bacterias fijadoras de nitrógeno; por otra parte presentan una especificidad muy amplia de unión a carbohidratos a pesar de su relación estructural cercana (Van Damme et al. 1998; Gabius, 2002). Los métodos de secuenciación de proteínas y ADN han mostrado que todas las lectinas de leguminosas están formadas por dos o cuatro protómeros, donde cada monómero tiene un peso molecular entre 25-30 kDa. Pueden formar desde dímeros hasta tetrámeros y se asume comúnmente que todas deben presentar una función común, o mejor una familia de funciones comunes, por la similitud de sus secuencias primarias, a pesar de que algunas de ellas presenten permutaciones circulares, como en el caso de la Concanavalina A (Van Damme et al. 1998). También se conoce que todas las lectinas de leguminosas están indudablemente relacionadas a un nivel molecular pero que muestran diferencias marcadas en cuanto a la especificidad de unión a carbohidratos. Todas las lectinas de leguminosas poseen sitios de unión a metales para iones divalentes (Mn^{2+} y Ca^{2+}) que son esenciales para la actividad de unión a carbohidratos (Peumans y Van Damme, 1998, Van Damme et al. 1998). En la interacción de las lectinas con los carbohidratos intervienen fuerzas intermoleculares débiles como la formación de puentes de hidrogeno, interacciones hidrofóbicas y electrostáticas; donde el ácido aspártico (D) y la asparagina (N) por lo general son los aminoácidos que participan en el reconocimiento de los carbohidratos; los iones divalentes como Ca^{+2} y Mn^{+2} presentes en algunas lectinas participan indirectamente en el reconocimiento de carbohidratos, formando enlaces de

coordinación que estabilizan la arquitectura del sitio de unión del carbohidrato (Elgavish y Shaanan, 1997).

1.3 Género *Galactia* y lectinas del género *Galactia*

Dentro de la familia *Leguminosae*, la subfamilia *Papilionoideae* y la tribu *Phaseoleae*, se encuentra la subtribu *Diocleinae*, dividida en 13 géneros: *Canavalia*, *Cleobulia*, *Pachyrhizus*, *Collea*, *Camptosema*, *Dioclea*, *Cratylia*, *Galactia*, *Cymbosema*, *Herpiza*, *Calopogonium*, *Macropsicanthus*, y *Luzonia*; de los cuales sólo los últimos dos no están presentes en zonas neotropicales (Azevedo- Moreira et al. 1995). El género *Galactia*, posee una amplia distribución en el trópico, cerca de 209 especies para 1993. Agrupando unas 50 especies descritas en las regiones tropicales o templado-cálidas de ambos hemisferios de América, principalmente en el hemisferio occidental. Se conocen 45 especies distribuidas en América tropical, de las cuales el mayor número está agrupado en Cuba con el 46% (Beyra et al, 2005). Dentro del género se conocen 36 especies en Sudamérica, donde la especie *Galactia lindenii* es de crecimiento endémico en Colombia, aparte de su descripción botánica y la caracterización de una de sus lectinas no hay más estudios enfocados en la búsqueda de otras propiedades (Almanza et al. 2004).

Dentro del género de las Diocleas, filogenéticamente relacionado con *Galactia* (Varela et al, 2004), tal como se mencionó en el caso de otras leguminosas, se ha encontrado la presencia de dos lectinas con diferentes especificidades (Pérez G., 1998; Sierra y Pérez., 1999; Melgarejo., et al, 2005; Dam et al., 2011). En algunos casos denominadas P2 y P4, y su mayor diferencia es el reconocimiento a carbohidratos, de galactósidos en el caso de la primera y de manosa/glucosa para la segunda. Muchas de ellas han sido purificadas y caracterizadas con algunas diferencias bioquímicas y similitudes moleculares. Por otro lado, se conocen las secuencias de aminoácidos de muchas de las lectinas P4 y por tanto se ha logrado elucidar la estructura terciaria de varias de ellas (Calvete et al, 1998; Wah et al, 2001; Delatorre et al, 2006; Nagano et al, 2008; Delatorre et al, 2011; da Nóbrega et al, 2012; Bezerra et al, 2013), utilizando $MnCl_2$ y/o $CaCl_2$ como iones requeridos por la proteína y además el ligando para cristalizar. De las lectinas P2 solo se conoce hasta fecha, la secuencia de aminoácidos de la lectina de

Cymbosema roseum (LCR-II); sin embargo aunque se ha cristalizado no se conoce su estructura terciaria (Rocha et al., 2009).

Las lectinas del género *Galactia* han sido poco estudiadas y algunas especies no presentan actividad aglutinante como *Galactia filiformis* y *Galactia striata* (Boyd et al., 1961). Hasta el momento solo se han aislado y caracterizado las lectinas de *Galactia lindenii* (LGL) (Almanza et al, 2004), *Galactia tenuiflora* (GTenL) (LePendou et al. 1986) y *Galactia tashiroi* (GTasL) (Fukuda et al. 1994). La GTenL luego de ser aislada mostró especificidad por el determinante de grupo sanguíneo humano H-tipo 2 (α -L-Fuc (1–2)- β -D-Gal (1–4)- β -D-GlcNAc-O-R), pero las propiedades moleculares excepto su peso molecular no han sido determinadas (LePendou et al., 1986). Respecto a la GTasL se determinó que es una glicoproteína con un contenido de carbohidratos de 3.2% y es inhibida por la N-acetil-D-galactosamina. Es un tetrámero con un peso molecular entre 90.000-95.000 kDa, conformado por cuatro subunidades iguales de 24.000 kDa y es activa en un rango de pH entre 4-10, y con un punto isoeléctrico de 8.0; sin embargo no hay datos concernientes a la interacción con oligosacáridos o neoglicoconjugados, aunque aglutina eritrocitos humanos (Fukuda et al. 1994). La lectina de *Galactia lindenii* (LGL), es la única que ha sido descrita dentro de este género con la caracterización bioquímica más completa. (Almanza et al., 2004). Por ensayo de ELLSA mostró gran especificidad por el determinante antigénico H-2 (α -L-Fuc (1–2)- β -D-Gal (1–4)- β -D-GlcNAc), y por Cromatografía frontal de Afinidad, se determinó su constante de afinidad ($K_a=1.1 \times 10^5$ M); y su especificidad por el Core 3 (GlcNAc (1–3)- β -D-GalNAc) (Vega et al., resultados no publicados). LGL-P2 es un tetrámero de 104 256 Da que se disocia en 4 subunidades iguales de 26 064 Da. Tiene un punto isoeléctrico de 8,3, un contenido de carbohidratos del 7% y es activa en un rango de pH entre 4-10 y entre 20-40 °C. Aglutina eritrocitos tipo O⁺ y en menor extensión A⁺ y B⁺. Su secuencia N-terminal fue establecida y el porcentaje de identidad con *Galactia tenuiflora* es del 52%. Para las lectinas de este género no se tienen datos de estructura primaria, secundaria o terciaria.

1.4 Determinación de la estructura primaria de lectinas

Las lectinas poseen diferentes propiedades biológicas en común, y representan un grupo diverso de proteínas en cuanto a tamaño, composición y estructura. La disponibilidad de la estructura primaria de numerosas lectinas, ha permitido la identificación de homologías entre las secuencias de lectinas de fuentes relacionadas de forma taxonómica, originalmente determinadas en lectinas derivadas de dos familias conservadas filogenéticamente *Leguminosae* y *Gramineae* (Wright. 1997; Sharon et al, 2004). En el caso específico de la subtribu *Diocleinae* (a la que pertenece *Galactia*), se han encontrado similitudes en la estructura primaria de las lectinas P4 de *Dioclea grandiflora*, *Dioclea Lehmanii*, *Canavalia ensiformis*, *Canavalia gladiata* y *Canavalia Maritima* entre otras (Azevedo- Moreira et al, 1995). En los últimos años, los avances en cristalografía de rayos X han llevado a un incremento en la obtención de estructuras cristalinas de lectinas (Wright. 1997; Ghazarian et al. 2011). Para la elucidación de la estructura primaria de cualquier proteína es necesaria la hidrólisis enzimática, secuenciación de sus péptidos y superposición de fragmentos; en algunos casos se debe tener una gran cantidad de proteína pura (para realizar digestiones con diferentes proteasas según la composición de aminoácidos de la proteína), para generar péptidos de diferentes tamaños y secuencias. Los péptidos generados se purifican por medio de cromatografía de exclusión molecular y HPLC en fase reversa. La determinación de la secuencia se puede realizar por la degradación de Edman o por espectrometría de masas en tandem (MS^n) como ha sido reportado en varios trabajos (Paroutaud et al., 1987; Herrmann et al., 1993; Saito et al., 1995; Nakano et al., 1999; De Souza et al. 2003; Rameshwaram et al., 2009; Rocha et al., 2009; Darville et al., 2012). Otros autores han elucidado la estructura primaria con ayuda de secuenciadores de proteínas en fase gaseosa o fase sólida (Konami et al. 1990, Konami et al. 1991, Lee et al. 1991, Fujii et al. 2011). Es importante mencionar que el desarrollo de la biología molecular también ha permitido la elucidación de la estructura primaria de proteínas basado en la obtención del ADNc y posterior traducción en la secuencia de péptidos (Higgins et al., 1983, Nakano et al., 1999). La dificultad de este método radica en el diseño de primers cuando no se tienen secuencias base; tampoco se pueden determinar las modificaciones postraduccionales entre otros.

2. Objetivos

2.1 Objetivo general

Determinar parcialmente la estructura primaria de la lectina de *Galactia lindenii* –P2 (LGL- P2) y purificar y caracterizar parcialmente la lectina de *Galactia lindenii* –P4 (LGL- P4).

2.2 Objetivos específicos

- Purificar la lectina presente en semillas de *Galactia lindenii* –P2 (LGL-P2) específica para el determinante H-2.
- Aislar el monómero de la lectina *Galactia lindenii* –P2 (LGL-P2) por medio de electroforesis preparativa en condiciones denaturantes.
- Obtener péptidos derivados del monómero de la LGL-P2 a partir de diferentes digestiones enzimáticas.
- Obtener la secuencia de algunos de los péptidos de LGL-P2 por espectrometría de masas.
- Determinar la secuencia parcial o total de los aminoácidos de la LGL-P2 superponiendo los peptidos derivados de las diferentes digestiones.
- Purificar la lectina presente en las semillas de *Galactia lindenii* –P4 (LGL-P4) específica para glucosa y manosa.
- Caracterizar parcialmente la LGL-P4.

3. Materiales y Métodos

El desarrollo de la parte experimental de la tesis se realizó en ocho fases:

1. Extracción y purificación de la LGL-P2.
2. Electroforesis preparativa para la obtención del monómero de la LGL-P2.
3. Obtención de péptidos a partir de LGL-P2 por medio de digestiones enzimáticas.
4. Obtención de la secuencia de algunos péptidos por espectrometría de masas.
5. Determinación parcial de la estructura primaria de la LGL-P2 por superposición de péptidos y comparación con otras lectinas tipo P2 de la Subtribu *Diocleinae*.
6. Ensayos exploratorios de cristalización para la LGL-P2.
7. Purificación de la LGL-P4.
8. Caracterización parcial de la LGL-P4.

3.1 Extracción y purificación de la LGL-P2.

3.1.1 Obtención del extracto proteico.

Se siguió la metodología descrita por Almanza et al., 2004. Las semillas maduras de *Galactia lindenii* (COL 15115) colectadas previamente en Fúquene (Cundinamarca), se seleccionaron y molieron hasta obtener un polvo fino. El extracto se preparó mezclando la harina en buffer Acetato Acético (100 mM, 150 mM NaCl, 5 mM tiourea) pH 5.0 en relación 1:10 w/v, y se dejó de 8 - 12 h a 4°C con agitación continua. Posteriormente se centrifugó a 36.000xg por 30 minutos a 4°C, se separó el pellet y con el mismo volumen

de buffer se realizó el procedimiento dos veces más. Finalmente se hizo un pool con los tres extractos, se concentró y se cambió el buffer a PBS 1X pH 7.2-7.4 (Na_2HPO_4 20 mM, NaH_2PO_4 20 mM, NaCl 150 mM) por medio de ultrafiltración (10 MWCO).

3.1.2 Cromatografía de intercambio iónico.

Para eliminar algunos pigmentos del extracto crudo y otras proteínas, se incluyó dentro del esquema de purificación una cromatografía de intercambio iónico sobre DEAE-Sephacel. En una columna de 16.2 cm x 1.7 cm se sembró el extracto crudo; la fracción no retenida se eluyó con el buffer de equilibrio PBS 1X pH 7.2-7.4 (Na_2HPO_4 20 mM, NaH_2PO_4 20 mM, NaCl 150 mM) y la fracción retenida con el mismo buffer con NaCl 500 mM. El seguimiento se hizo midiendo absorbancias a 280 nm y las fracciones colectadas fueron concentradas en Amicon PM 10 kDa.

3.1.3 Cromatografía de afinidad sobre Lactosa-Sepharosa

Preparación del soporte de afinidad: se siguió la metodología descrita en Hermanson et al, (1992); se realizaron dos preparaciones cada una de 10 mL de gel. La lactosa previamente lavada en etanol se acopló a Sepharosa 4B activada con divinilsulfona. Para la preparación del soporte de afinidad se tomaron 10 mL de la Sepharosa 4B, se lavaron con 100 mL de agua desmineralizada y se equilibraron con 100 mL de carbonato de sodio 0.5 M; el gel se resuspendió en 10 mL de la misma solución. Luego se adicionó gota a gota 1 mL de DVS (divinilsulfona) con agitación muy suave y se dejó la mezcla a temperatura ambiente durante dos horas. Posteriormente se hicieron varios lavados con agua desmineralizada y carbonato de sodio 0.5 M. El soporte activado se resuspendió en 10 mL de carbonato de sodio 0.5 M conteniendo 0.3 M de lactosa y se dejó en agitación suave durante 24 h para el acople. Finalmente se realizaron varios lavados consecutivos con agua desmineralizada y el gel se resuspendió en 10 mL de carbonato de sodio 0.5 M con 0.5 mL de β -mercaptoetanol (con el fin de bloquear los grupos que no reaccionaron); la mezcla reaccionó durante dos horas a temperatura ambiente con agitación muy suave. El gel acoplado se lavó con agua desmineralizada seguido de una solución de NaCl 0.1 M; el gel se equilibró en PBS 1X pH 7.2-7.4 (Na_2HPO_4 20 mM, NaH_2PO_4 20 mM, NaCl 150 mM).

Cromatografía de afinidad: se siguió la metodología descrita por Almanza et al., 2004. 93 mg de las fracciones no retenidas obtenidas a partir de cromatografía de intercambio iónico (3.1.2) se sembraron sobre la columna de afinidad (18.5 x 1 cm, 10 mL de gel); la fracción no retenida se eluyó con el buffer de equilibrio, PBS 1X pH 7.2-7.4 y la fracción retenida con el mismo buffer con lactosa de 0.2 M (PBS 1X-lactosa 0.2 M). La velocidad de elución fue 1 mL/12 min y el seguimiento de la cromatografía se hizo midiendo absorbancia a 280 nm. La fracción retenida se dializó contra bicarbonato de amonio 20 mM (realizando tres cambios cada 8 h) y se liofilizó. Cada una de las fracciones retenidas fue evaluada por SDS-PAGE y aglutinación. Algunas de las fracciones retenidas se recromatografiaron usando una nueva columna de afinidad de 1 mL de gel. Se realizaron otras cromatografías usando el mismo soporte pero sembrando 120 mg del extracto crudo en lugar de las fracciones de intercambio iónico.

3.1.4 Ensayos de aglutinación

La actividad de cada una de las fracciones puras se evaluó en microplacas realizando ensayos de aglutinación con eritrocitos O⁺ al 2% en PBS 1X. La suspensión de eritrocitos se preparó según la técnica descrita por Navarro y Pérez (1978), 500 µl de eritrocitos empaquetados (O⁺) se lavaron con 1 mL de PBS 1X y fueron centrifugados a 5.000 rpm durante 5 minutos a 4°C, se hicieron en total tres lavados. La suspensión de eritrocitos se preparó al 2% en PBS. En placas para aglutinación se sembraron 50 µl de los extractos crudos + 50 µl de PBS 1X + 50 µl de la suspensión eritrocitos; en el caso de la proteína pura se emplearon 25 µl de solución + 25 µl de PBS 1X + 25 µl de la suspensión eritrocitos. La interacción se dejó durante 30 minutos a temperatura ambiente y luego se evaluó cualitativamente la actividad empleando una escala de 0 a +4 según el grado de aglutinación. Los blancos se prepararon usando en 100 µl de PBS 1X + 50 µl de la suspensión de células.

3.1.5 Cuantificación de proteína

La concentración de proteínas se evaluó por el método del ácido Bicinconínico (BCA) descrito por Smith et al, (1985). El ensayo se realizó en microplaca según la metodología

descrita Almanza, (1999) empleando una curva de calibración con solución patrón de Albúmina Sérica Bovina (BSA) a una concentración de 1.6 mg/ mL (Anexo 1).

3.1.6 Electroforesis unidimensional en geles de poliacrilamida

Las fracciones obtenidas en cada uno de los pasos de purificación fueron evaluadas por electroforesis. Se siguió el método descrito por Laemmli (1970) empleando geles del 12.5%. Las muestras fueron preparadas con calentamiento durante 10 minutos a 100°C y los geles fueron fijados y teñidos con azul de Coomassie R-250.

3.2 Electroforesis preparativa para obtener el monómero de la LGL-P2

Se prepararon geles en condiciones denaturantes siguiendo el manual del equipo BIO RAD (Model 491 Prep Cell) y la metodología descrita por Sánchez (2003). Se usaron geles con un tamaño total de 10.5 ± 0.5 cm, donde el gel concentrador al 5% midió 0.5 cm y el gel separador al 10% midió 10 ± 0.1 cm. Todas las corridas se hicieron a un voltaje de 330 V, corriente de 32 mA y potencia constante de 12 W. Se ajustó un flujo de colección de 0.5 mL/min y un volumen de fracciones de 3-3.5 mL. Para determinar el tiempo de elución del monómero de la LGL-P2 se hizo una curva de calibración donde se utilizaron como patrones lisozima (14 kDa), mioglobina (17 kDa), ovalbúmina (45 kDa), albúmina de suero bovino (BSA) (66 kDa) e IgG (150 kDa), cada uno en una concentración de 2 mg/mL en solución de NH_4HCO_3 20 mM, sembrando en total 1 mg de proteína. De LGL-P2 se sembró 1 mg en SDS al 1% y previo calentamiento a 100 °C durante 10 minutos. El control de elución se realizó midiendo absorbancia a 280 nm y las fracciones colectadas fueron dializadas contra bicarbonato de amonio 20 mM y posteriormente liofilizadas. Las fracciones obtenidas fueron cuantificadas y evaluadas por electroforesis unidimensional (Smith et al, 1985; Laemmli, 1970).

3.3 Ensayos de digestión Enzimática para la LGL-P2

3.3.1 Digestiones enzimáticas en solución

Las digestiones enzimáticas para la proteína se llevaron a cabo según el procedimiento descrito por Darville et al., (2012). La lectina liofilizada se resuspendió en bicarbonato de amonio pH 8.5-9.0, posteriormente se le adicionó SDS a una concentración final de 1% y se incubó a 95°C durante 30 min, al final del proceso se diluyó la muestra hasta obtener una concentración final de SDS al 0.1%. La enzima se adicionó en diferentes proporciones y la mezcla de reacción se incubó a 37°C por 12, 24 y 48 h en diferentes ensayos; la reacción se detuvo calentando la mezcla a 95 °C durante 10 minutos (Tabla 1). En todos los casos se utilizaron los siguientes controles:

- **Control stock de enzimas:** Solución de la enzima sometida a calor sin SDS y/o DTT y sin adición de sustrato.
- **Control actividad enzimática:** enzima en solución y condiciones de digestión (calentamiento, y adición SDS y/o DTT) sin sustrato para corroborar autólisis.
- **Control solución LGL-P2:** Lectina con SDS sometida a condiciones de digestión sin la adición de enzima.
- **Control solución lisozima:** se incluyó DTT utilizando una concentración inicial de 2 mM y diluyendo antes de adicionar las enzimas de digestión hasta 0.4 mM. Se adicionó SDS 1% (final) y se calentó a 80°C por 4 horas, posteriormente se disminuyó la concentración del SDS a 0.1%. El control se utilizó sin enzima en las condiciones de digestión.

Tabla 1. Enzimas empleadas en la hidrólisis de la LGL-P2

Enzima*	Marca	Características	Concentración inicial LGL-P2 (mg/mL)	Concentración final LGL-P2 (mg/mL)	Relación E:L**	t digestión
Tripsina	SIGMA	Tipo III. Páncreas bovino	4.636	0.431 0.417 0.357	1:50 10:50 50:50	24 h
Tripsina	Amresco	Grado proteómico. Páncreas bovino.	0.100	0.047	10:10	48h
Tripsina	Difco		0.383	0.125	100:50	12h
Tripsina	SIGMA	TPCK tratada. Páncreas bovino.	0.243	0.042	100:25	12h
α - Quimotripsina 1	SIGMA	TPCK tratada. Tipo VII. Páncreas bovino.	0.383	0.125	100:50	12h
α - Quimotripsina 2	Merck	Páncreas bovino.	0.383	0.125	100:50	12h

** Se refiere a la proporción μg enzima: μg lectina; concentración inicial de la solución de cada enzima 2 mg/mL.

3.3.2 Electroforesis en geles de poliacrilamida

Para evaluar las hidrólisis realizadas en 3.3.1 se hizo electroforesis en geles Tris-Tricina siguiendo la metodología descrita por Schägger y Von Jagow (1987), empleando geles al 16.5%T3%C (Anexo 2). Los geles fueron fijados con ácido tricloroacético al 10% y 1% y fueron teñidos con azul de Coomassie G-250 o con tinción de plata (Anexo 3).

3.3.3 Digestiones enzimáticas en geles de poliacrilamida

3.3.3.1 Se cortó la banda de proteína del gel de digestión preparado en las condiciones mostradas en 3.3.2 y teñido con azul de coomassie G-250. Para eliminar la tinción se hicieron dos lavados de 30 min con 40% acetonitrilo/200 mM bicarbonato de amonio; dos lavados más con 25 mM bicarbonato de amonio durante 5 min y finalmente con 25 mM bicarbonato de amonio/50% acetonitrilo por 15 min. El gel se deshidrató con 100% acetonitrilo por 5 min y se hidrató empleando tripsina de alta pureza (Marca SIGMA proveniente de páncreas bovino, tratada con TPCK, 13.500 u/ mg sólido preparada a una concentración de 2 mg/ mL) durante 10 min a temperatura ambiente. Se hizo la digestión durante 12 horas a pH 8.0 según la metodología descrita (Shevchenko et al, 1996 y Shevchenko et al, 2006).

3.3.3.2. La proteína pura se resuspendió en buffer MES 20 mM pH 7.4/100 mM NaCl /glicerol 4% y se calentó a 100°C durante 10 minutos. Se corrió un gel en condiciones denaturantes de la casa comercial Invitrogen con un rango de concentración de acrilamida de 8-16%. Al final de la corrida el gel se tiñó con azul de coomassie R-250. Posteriormente se cortó la banda de interés y se fraccionó en dos piezas (a la mitad) que fueron enviadas al Instituto de Espectrometría de Masas y Proteómica del Johns Hopkins University School of Medicine (se tomaron 3 bandas de interés, una para cada enzima utilizada). Las bandas de proteína fueron reducidas con DTT, alquiladas con iodoacetamida y sometidas a varios lavados para eliminar la tinción con coomassie. Las digestiones se realizaron con tripsina (Promega, proveniente de porcino, tratada con TPCK, modificada resistente a cortes proteolíticos, grado secuenciación), quimotripsina (Promega, derivada de páncreas bovino, grado secuenciación) y LysC (Promega, aislada de *Lysobacter enzymogenes* grado secuenciación) en bicarbonato de amonio 25 mM a 37°C durante 16 h (Shevchenko et al, 1996). La concentración final de las enzimas fue: tripsina 12.5 ng/μl, LysC 0.01-0.05 μg/μl y quimotripsina 0.5-1 μg/μl.

3.4 Secuenciación de péptidos de la LGL-P2 por espectrometría de masas.

Los péptidos obtenidos fueron analizados en un nano-LC-LTQ-Orbitrap-Velos en FTFT con una resolución de precursor de 30.000 y una resolución de fragmento de 15.000. Los análisis se realizaron haciendo una búsqueda en Mascot-Scaffold: Mascot 2.2 via PD1.3 usando RefSeq2012 como base de datos, teniendo en cuenta como parámetros de búsquedas todas las especies disponibles; tripsina, quimotripsina y LysC como enzimas, un valor del error en clivaje de 2, teniendo en cuenta oxidación de metioninas y deamidación como modificaciones variables. Una tolerancia al precursor de 12 ppm, y al fragmento de 0.03 Da. Estos ensayos se realizaron en el instituto de espectrometría de masas en la Universidad Johns Hopkins (Mass Spectrometry and Proteomics Facility de Johns Hopkins University School of Medicine).

3.4.1 Preparación de la muestra para el análisis de MS/MS

Los digeridos fueron extraídos del gel con 50% acetonitrilo/0.1% TFA y se secaron. Para el digerido con tripsina, el 15% de la muestra se limpió utilizando ziptip C18, mientras que en el caso de los digeridos de LysC y Quimotripsina el 50% de la muestra fue utilizado. La mezcla de péptidos se reconstituyó en 10 μ l 2%ACN/0.1%FA y se inyectaron 5 μ l o 25% del reconstituido (en el caso del digerido triptico y con los digeridos de las otras dos enzimas respectivamente).

3.4.2 Obtención de los péptidos de la LGL-P2 por Cromatografía Líquida-Espectrometría de masas en tándem (LCMS/MS)

Los péptidos se sembraron en una columna C18 (YMC*GEL ODS-A 12nm S-10 μ m) con un flujo de 600 nL/min de 0.1% ácido fórmico (solvente A) y separándolos a 300 nL/min en una columna en fase reversa C18 de 75 μ m x 100 mm Magic AQ (5 μ m, 120Å, Microm Bioresources) usando un gradiente de 3-10% del solvente B (90% acetonitrilo en 0.1% ácido fórmico) durante los primeros 10 min, luego incrementando a 30% solvente B por 55min, 45% solvente B por 65 min y 100% solvente B por 70min. Los péptidos eluidos fueron enviados en forma de spray hacia el espectrómetro de masas LTQ OrbitrapVelos a través de una aguja inyectora (New Objective) de 1 μ m a 2.0 kV. Los resultados fueron adquiridos en un rango de 350-1700 m/z con máximo 10 masas peptídicas (iones precursores) aisladas individualmente a IW1.9Da, y fragmentadas utilizando energía de activación por colisión HCD 35. El ión precursor y los fragmentos de iones fueron analizados a una resolución de 30.000 y 15.000, respectivamente. Con parámetros de inclusión dinámica de 30 s, MIPS (monoisotopic ion precursor selection) encendido, la opción m/z apagada y la opción bloqueo de peso molecular encendida con corte en el silocsano (371 kDa).

3.4.3 Análisis de los datos

Cuando se compilaron los resultados de la búsqueda en PD1.3/Mascot protein/peptide identification se creó un reporte de Scaffold (Anexos 10, 11 y 12); ya que provee una visión más global de todos los datos obtenidos, donde las proteínas listadas fueron

identificadas en al menos una de las muestras con un nivel de confianza $\geq 95\%$ basado en la identificación de al menos dos péptidos con un nivel de confianza $\geq 95\%$. Los códigos de color representan el nivel de confianza de la identificación de la proteína. Los números en los cuadros de color son los números de péptidos únicos identificados con un nivel de confianza $\geq 95\%$. Los datos de los espectros obtenidos de la espectrometría de masas en tándem fueron procesados con ayuda del Proteome Discoverer (v1.3 Thermo Fisher Scientific), en tres formas diferentes utilizando 3Nodes: “common”, “Xtract” y “MS2 Processor”.

Los resultados obtenidos de 3Nodes fueron analizados con Mascot v.2.2.2 Matrix Science utilizando:

- “RefSeq2012” como base de datos, con base de datos concatenados especificando todas las especies, las enzimas utilizadas como “enzymes” y un error de corte de valor 2.
- “Custom sns4” como base de datos, sin especificar la enzima.

Otros parámetros utilizados en ambas búsquedas fueron: tolerancia a la masa del precursor de 12 ppm, tolerancia a la masa de los fragmentos de 0.03 kDa, y la oxidación (M: Metioninas) y carbamidometil (C: cisteínas) como modificaciones variables.

3.5 Determinación de la secuencia parcial de aminoácidos de la lectina P2 de *Galactia lindenii*.

3.5.1 Análisis Bioinformático de los datos obtenidos

Con las secuencias de los péptidos se hizo un análisis bioinformático, para disminuir la redundancia se utilizó la herramienta Decrease redundancy de EXPASY (Artimo et al., 2012). Con la herramienta blastp de BLAST (Altschul et al., 1990) en NCBI, utilizando algoritmos diseñados para secuencias cortas (NCBI-BLAST help, <http://www.ncbi.nlm.nih.gov/blast/Why.shtml>) y proteínas de leguminosas se buscaron secuencias relacionadas para realizar el ensamblaje de los péptidos.

3.5.2 Análisis de las secuencias de péptidos de la LGL-P2 contra secuencias de péptidos derivados de la DLL- P2 (P2 de *Dioclea lehmannii*)

Con ayuda de blastp de BLAST en NCBI, y utilizando algoritmos diseñados para secuencias cortas (NCBI-BLAST help, <http://www.ncbi.nlm.nih.gov/blast/Why.shtml>) se compararon las secuencias de los péptidos contra secuencias de lectinas reportadas y no reportadas, en el caso de la P2 *D. lehmannii* (Vega et al., Datos no publicados).

3.5.3 Superposición de péptidos para la obtención de la secuencia

Con cada uno de los grupos de péptidos obtenidos con las tres enzimas se hizo una comparación en blastp para secuencias cortas (NCBI-BLAST help, <http://www.ncbi.nlm.nih.gov/blast/Why.shtml>) contra la secuencia de LDL-P2. Con los péptidos que alinearon con regiones de la LDL-P2 se hizo un alineamiento total, tomando los péptidos que alinearan con un menor valor E y con un mayor porcentaje de identidad y cobertura de secuencia. Luego de tener alineamientos por bloques de péptidos con cada una de las enzimas, se hizo superposición para las tres enzimas teniendo en cuenta las modificaciones puntuales de aminoácidos que posteriormente afectarían el ensamblaje de la secuencia. Con las regiones superpuestas se definieron los tipos de cambios en los aminoácidos para obtener la secuencia parcial de la LGL-P2. Para los gaps en el alineamiento total de la secuencia, se utilizaron las secuencias de las lectinas de dioclea con un alto porcentaje de identidad con LDL-P2.

3.5.4 Acercamiento a la estructura terciaria por modelamiento de las lectinas P2 de la Lectina de *Galactia lindenii* y *Diocleae lehmannii*

Teniendo en cuenta los estudios preliminares de estructura primaria parcial de las dos lectinas por análisis bioinformático, se realizó un acercamiento a la estructura terciaria, para ver si los dominios son característicos de las lectinas de leguminosas.

3.5.4.1. Modelamiento por Homología

Para el modelamiento por homología se escogió el programa I-Tasser Server <http://zhanglab.ccmb.med.umich.edu/> el cual permite el modelado comparativo utilizando

una combinación de múltiples estructuras de referencia y una optimización iterativa de alineaciones alternativas. Se realizó la comparación y selección de modelos de acuerdo con el mejor resultado (score).

3.5.4.2. Evaluación del modelo.

Para la evaluación del modelo se utilizaron los parámetros de I-TASSER, el cual arroja un algoritmo de puntuación (C-score) que va de -5 a 2; para la estimación de la calidad de los modelos de predicción, una mayor puntuación indica una mejor predicción en la estructura. Igualmente el TM-score y el (RMSD) son parámetros estándar que se utilizan para indicar la cercanía del modelo que se predice con la estructura nativa del molde. TM-score > 0,15 significa un modelo que presenta una topología correcta (Banerjee et al, 1996; Zhang 2008; Roy et al., 2010; Roy et al., 2012). También se realizó el análisis de estructura secundaria por diagramas de Ramachandran mediante la herramienta PROCHECK ANALYSIS <http://www.ebi.ac.uk/>.

3.6 Ensayos exploratorios de cristalización de la LGL-P2

Para cristalizar la proteína se utilizó el método de la gota suspendida (McPherson 1990) en presencia de $MnCl_2$ y/o $CaCl_2$ como iones y de lactosa según la metodología descrita por varios autores para lectinas de *Diocleae* tipo P4 (Wah et al., 2001., Delatorre et al, 2006., de Oliveira et al., 2008; Nagano, et al, 2008; Delatorre et al, 2011., Da Nóbrega et al, 2012 ., Bezerra et al., 2013). La lectina pura se solubilizó en buffer MES pH 7.4 (20 mM, 100 mM, NaCl, glicerol 4%) hasta una concentración de 14.44 mg/mL.

3.6.1 Parámetros del primer grupo de ensayos de cristalización

Se llevaron a cabo en el buffer MES pH 7.4 y buffer acetato pH 4.6, a 16°C, utilizando como precipitante PEG 400 en un amplio rango de concentraciones: 20%, 25%, 28%, 30%, 32%, 36%, 38%, 40%, 42%, 45%, 48%, 50%. Se trabajó en presencia y/o ausencia de sales (5 mM $MnCl_2$ y 5 mM $CaCl_2$), y 100 mM de lactosa. El agente

precipitante, las sales y el ligando fueron adicionados al buffer del reservorio. Se utilizó una relación de 1:1 (μL de proteína: μL buffer de reservorio) y se dejó el ensayo durante 60 días.

Tabla 2: Condiciones utilizadas para el ensayo base de cristalización

pH	% PEG	Iones 5 mM	Lactosa 100 mM
4.6 /7.4	20,25,28,30,32,36,38 40,42,45,48, 50	Ca ⁺⁺	+
4.6 /7.4	20,25,28,30,32,36,38 40,42,45,48, 50	Mn ⁺⁺	+
4.6 /7.4	20,25,28,30,32,36,38 40,42,45,48, 50	Ca ⁺⁺	-
4.6 /7.4	20,25,28,30,32,36,38 40,42,45,48, 50	Mn ⁺⁺	-

Se realizaron ensayos individuales de pH, de agente precipitante y de adición de iones
 Por ejemplo : Ensayo 1: pH 4.6, 20% PEG, Ca⁺⁺ y 100 mM lactosa

3.6.2 Parámetros del segundo grupo de ensayos de cristalización

Se utilizaron los kits Crystal Screen I – II (CSI-II) de Hampton Research y Morpheus Screen MD1-46 de Molecular Dimensions (Anexo 4). En el caso de CSI-II se ensayaron dos condiciones: sólo la proteína en buffer (MES pH 7.4, 20 mM, 100 mM, NaCl, glicerol 4%) y la proteína en buffer con 5 mM MnCl₂, 5 mM CaCl₂ y lactosa 100 mM; en ambos casos se utilizó una relación 0.3:0.3 (μL de proteína: μL de buffer reservorio).

En el caso del kit Morpheus Screen MD1-46, sólo se realizó un ensayo con la proteína en el buffer (MES pH 7.4, 20 mM, 100 mM, NaCl, glicerol 4%) conteniendo 5 mM MnCl₂, 5 mM CaCl₂ y 100 mM de lactosa en la misma relación, los ensayos se dejaron 27 días a 16 °C.

3.6.3 Optimización de las condiciones de cristalización

Para obtener cristales de mayor tamaño se optimizaron las condiciones de acuerdo con los resultados obtenidos en 3.6.2. En la tabla 3, se muestran las condiciones utilizadas usando fosfato de amonio como agente precipitante (condición en la que el kit Cristal Screen produjo crecimiento de cristales) sin adición de buffer ni sales al sistema de reservorio, donde el pH y las sales eran las presentes en la muestra de proteína (buffer MES pH 7.4, 20 mM, 100 mM, NaCl, glicerol 4% conteniendo 5 mM MnCl₂, 5 mM CaCl₂ y 100 mM de lactosa). La otra condición ensayada fue la del kit Morpheus, donde se variaron las concentraciones del precipitante MPD (+/- 2-metil-2, 4-pentanodiol) y el pH del buffer Tris para analizar cuál funcionaba mejor, sin variar la concentración de la sal en el reservorio (fosfato de amonio). En las tablas 4 y 5 se repitieron las condiciones montadas en la columna 5 y 6 de la primera y la 1 y 2 de la segunda, debido a que los cristales formados fueron los que más rápido y mejor se desarrollaron. Los ensayos se realizaron sembrando dos gotas por pozo, una gota en una relación de 0.6:0.6 (μL de proteína: μL buffer de reservorio), y la otra en una relación 1:1 (μL de proteína: μL buffer de reservorio). Las tablas hacen la representación de las cajas multipozos en las cuales se realizó el experimento.

Tabla 3. Ensayos con fosfato de amonio como agente precipitante sin adición de buffer ni sales al sistema de reservorio.

\	Columna 1	Columna 2	Columna 3	Columna 4	Columna 5	Columna 6
Fila A	NH ₄ PO ₃ 0.8 M	NH ₄ PO ₃ 0.7 M	NH ₄ PO ₃ 0.6 M	NH ₄ PO ₃ 0.5 M	NH ₄ PO ₃ 0.4 M	NH ₄ PO ₃ 0.3 M
Fila B	NH ₄ PO ₃ 0.2 M	NH ₄ PO ₃ 0.1 M	NH ₄ PO ₃ 0.05 M	----	-----	-----

Tabla 4. Ensayos de optimización con 0.2 M fosfato de amonio monobásico como sal, Tris-HCl 0.1 M a diferentes pHs, y MPD a diferentes concentraciones.

\	Columna 1	Columna 2	Columna 3	Columna 4	Columna 5	Columna 6
Fila A	pH 7.5 /46% MPD	pH 7.5 /48% MPD	pH 7.5 /50% MPD	pH 7.5 /52% MPD	pH 7.5 /54% MPD	pH 7.5 /56% MPD
Fila B	pH 8.0/46% MPD	pH 8.0/48% MPD	pH 8.0/50% MPD	pH 8.0/52% MPD	pH 8.0/54% MPD	pH 8.0/56% MPD
Fila C	pH 8.5/46% MPD	pH 8.5/48% MPD	pH 8.5/50% MPD	pH 8.5/52% MPD	pH 8.5/54% MPD	pH 8.5/56% MPD
Fila D	pH 9.0/46% MPD	pH 9.0/48% MPD	pH 9.0/50% MPD	pH 9.0/52% MPD	pH 9.0/54% MPD	pH 9.0/56% MPD

Tabla 5. Ensayos de optimización variando la concentración de MPD con respecto a las anteriores

\	Columna 1	Columna 2	Columna 3	Columna 4	Columna 5	Columna 6
Fila A	pH 7.5 /54% MPD	pH 7.5 /56% MPD	pH 7.5 /58% MPD	pH 7.5 /60% MPD	pH 7.5 /62% MPD	pH 7.5 /64% MPD
Fila B	pH 8.0/54% MPD	pH 8.0/56% MPD	pH 8.0/58% MPD	pH 8.0/60% MPD	pH 8.0/62% MPD	pH 8.0/64% MPD
Fila C	pH 8.5/54% MPD	pH 8.5/56% MPD	pH 8.5/58% MPD	pH 8.5/60% MPD	pH 8.5/62% MPD	pH 8.5/64% MPD
Fila D	pH 9.0/54% MPD	pH 9.0/56% MPD	pH 9.0/58% MPD	pH 9.0/60% MPD	pH 9.0/62% MPD	pH 9.0/64% MPD

3.6.4 Siembra de cristales

Se sembraron los cristales obtenidos en 3.6.3 (tabla 5) y un cristal de los formados se colocó en la mezcla buffer: proteína, bajo las mismas condiciones del cristal original, con el fin de obtener cristales de mayor tamaño. Se colocaron los cristales (tabla 5, relacionando primero la fila y luego la columna) A3, A4, B3, B6, C4, C6, D3 en la gota en una relación 1:1 (μL de proteína: μL de buffer reservorio). Los cristales B2 y B5 en la gota en una relación 0.6:0.6 (μL de proteína: μL de buffer reservorio). Las mezclas se incubaron durante 8 días para permitir el crecimiento de los cristales.

3.6.5 Difracción de rayos X

Con algunos de los cristales obtenidos se realizó difracción de rayos X, en el equipo Rigaku Saturn 944+ (Instituto de rayos X de la facultad de Biofísica y Química Biofísica del Johns Hopkins University School of Medicine) con las siguientes condiciones: distancia al goniómetro de 80 mm, 1° de oscilación, con un ángulo de inicio de -90° y ángulo final de 1° , tomando dos marcos cada 90° sobre el eje de rotación Omega, y con tiempo de exposición de 30 o 60 segundos. Algunos de los cristales obtenidos se sometieron a difracción.

Nota: Con estos ensayos se obtuvo un cristal de la celda C6 que se congeló para enviar a analizar en ciclotrón y obtener un espectro de difracción con la resolución necesaria para determinar la estructura terciaria de la proteína.

3.7 Purificación de la lectina de *Galactia lindenii* –P4 (LGL-P4).

Ensayos dentro del grupo de investigación en proteínas (GRIP) mostraron que en las semillas de *Galactia lindenii* estaba presente una segunda lectina como era característico en algunas de las especies que están en la subtribu *Diocleinae* (Pérez G, 1998; Sierra y Pérez, G., 1999,; Melgarejo et al. 2005; Dam et al., 2011; Quintero et al., 2013). En esta fase se purificó y caracterizó parcialmente esta nueva lectina.

3.7.1 Cromatografía de intercambio iónico

S corrió en el soporte Sephacryl S-200, en una columna de dimensiones 16.2 cm x 1.7 cm, donde se sembró el extracto crudo obtenido en 1.1.1; la fracción no retenida se eluyó en el buffer de equilibrio, PBS 1X pH 7.2-7.4 (Na_2HPO_4 0.2 M, NaH_2PO_4 0.2 M, NaCl 150 mM), y la fracción retenida en el mismo buffer conteniendo NaCl 500 mM. El seguimiento se hizo midiendo absorbancias a 280 nm y las fracciones colectadas fueron concentradas en Amicon PM 10 kDa.

3.7.2 Cromatografía de afinidad

Las fracciones retenidas obtenidas en 1.7.1, fueron sometidas a cromatografía de afinidad sobre un soporte de Sephacryl S-200 (6.9 x 2.6 cm; 2 cm altura del soporte). La fracción no retenida se eluyó con buffer PBS 1X pH 7.2-7.4 y la fracción retenida en el mismo buffer conteniendo glucosa 0.2 M. La velocidad de elución fue constante (1 mL/12 min) y el seguimiento se hizo midiendo absorbancias a 280 nm. Se colectó la fracción retenida, se dializó contra bicarbonato de amonio 20 mM (realizando tres cambios cada 8 h) y se liofilizó. La concentración de proteínas se evaluó por el método del ácido Bicinconínico (BCA) y las fracciones obtenidas en cada uno de los pasos de purificación fueron evaluadas por electroforesis (Laemmli, 1970) empleando geles del 12.5%.

3.7.3 Actividad eritroaglutinante

La actividad de cada una de las fracciones purificadas se evaluó en microplaca realizando ensayos de aglutinación con eritrocitos al 2% en PBS 1X. La suspensión de eritrocitos se preparó según la técnica descrita por Navarro y Pérez (1978); 500 μ l de células empaquetadas fueron resuspendidas en 1 ml de PBS y posteriormente fueron centrifugadas a 5.000 rpm, 5 minutos a 4°C. Las células se lavaron tres veces más y se preparó la suspensión celular al 2% en PBS. Los blancos se prepararon con 50 μ l de PBS 1X + 50 μ l de suspensión eritrocitaria. Las muestras fueron evaluadas empleando 25 μ l de solución, 25 μ l de PBS 1X y 50 μ l de suspensión eritrocitaria. La actividad se evaluó empleando una escala cualitativa de 0 a +4 según el grado de aglutinación. Se realizaron ensayos con eritrocitos humanos A+, B+ y O+. Para determinar el título de la lectina se realizaron diluciones seriadas de los extractos crudos y de la lectina pura. Las aglutinaciones fueron corroboradas con un microscopio de luz.

3.8 Caracterización parcial de la lectina LGL-P4.

3.8.1 Ensayos de inhibición con carbohidratos

Se realizaron según el método reportado por Vega (2004); 50 μ l de la lectina pura (0.0125 μ g/ μ l) se incubaron por 30 min con 50 μ l de cada una de las soluciones de carbohidratos: Glucosa, Manosa, Lactosa, Galactosa, Fucosa, GalNAc, 1-O-metil- β -galactopiranosido, 1-O-metil- β -glucopiranosido, β -metil D-glucosa, N-acetil- α -D-glucosamina, α -metil O- manosido, ácido D-galacturónico todos con una concentración inicial 0.3 M; p-Nitrofenil- β -D-glucopiranosido, p-Nitrofenil- α -D-glucopiranosido, p-Nitrofenil- β -D-glucosamina, p-Nitrofenil-N-acetil- α -D-glucosamina, p-Nitrofenil- β -D-manopiranosido, p-Nitrofenil- β -D-galactopiranosido, p-Nitrofenil-N-acetil- β -D-galactosamina, p-Nitrofenil-N-acetil- β -D-galactopiranosido todos en una concentración inicial 37 mM. Posteriormente se adicionaron 50 μ l de eritrocitos A⁺ y se incubaron por 2 h a temperatura ambiente. Al cabo de este tiempo evaluó la aglutinación empleando una escala cualitativa de 0 a +4. Con los carbohidratos que inhibieron la actividad de la

lectina se determinó la concentración mínima inhibitoria realizando diluciones seriadas con 50 μl de cada una de las soluciones de carbohidratos y 50 μl de la lectina pura (0.0125 $\mu\text{g}/\mu\text{l}$), las mezclas se incubaron por 30 minutos y al final se adicionaron 50 μl de eritrocitos A⁺. Los ensayos se dejaron una hora a temperatura ambiente y se evaluó la aglutinación.

3.8.2 Detección cualitativa de glicoproteínas

Se siguió la metodología descrita por (Vega, 1997) realizando PAGE-SDS de la proteína pura y transferencia sobre una membrana de Nitrocelulosa. Como control del ensayo se utilizó ovalbúmina a una concentración de 5 $\mu\text{g}/\mu\text{L}$ (Anexo 5).

3.8.3 Determinación del punto isoeléctrico

Se siguió la metodología descrita por Bollag y Edelstein (1994), el ensayo se realizó en condiciones nativas empleando anfolitos en un rango de pH 3-10. Para realizar la curva de calibración se trabajó con patrones de un rango de pH 3-10 (Farmacia). Como control de corrido se empleó mioglobina 5 $\mu\text{g}/\mu\text{l}$ y como soluciones en el ánodo ácido fosfórico 0.1 M y en el cátodo NaOH 0.1 M (Anexo 6).

3.9 Purificación de otras lectinas de la subtribu *Diocleinae*

Se extrajeron y purificaron las lectinas de *Dioclea lehmanni* y *Dioclea grandiflora* P2 y P4 según la metodología descrita por Pérez, (1998) y (1990); y *Canavalia ensiformis* P2 y P4 según la metodología de Melgarejo et al., 2005. La harina de cada una de la especies se encontró disponible en el laboratorio de proteínas (GRIP). La concentración de proteínas se determinó para los extractos crudos y la proteína pura por el método de microKjeldahl. Con las fracciones de proteína pura se realizaron electroforesis según el método descrito por Laemmli (1970) en geles del 12.5%.

4. Análisis de resultados

Para entender la función de una proteína es necesario conocer su estructura tridimensional y para llegar a determinarla se debe conocer su secuencia de aminoácidos; en el caso de la LGL-P2 hasta la fecha no se ha establecido la estructura primaria ni la estructura terciaria de ninguna de las lectinas de este género. Estas lectinas son interesantes porque reconocen específicamente el determinante H tipo 2 y presentan diferencias muy sutiles en reconocimiento con respecto a otras lectinas de la subtribu *Diocleinae* (Almanza et al., 2004; Melgarejo et al., 2005; Dam et al, 2011). Por otra parte y teniendo en cuenta los prerequisites que se necesitan para establecer la estructura primaria de una proteína, la LGL-P2 es una buena candidata para realizar estudios estructurales. De esta lectina previamente se ha establecido:

- Los ensayos de purificación permiten obtener un material homogéneo que garantiza que solo hay una entidad molecular.
- Se conoce su secuencia N-terminal
- Luego del proceso de purificación se obtienen buenas cantidades de proteína, la cual es soluble en altas concentraciones.
- Se conoce el análisis de aminoácidos.
- Como no se conoce nada acerca de la secuencia se puede hacer la identificación de secuencias idénticas de péptidos en más de un digerido.
- Se conocen datos completos de su caracterización bioquímica y de secuencias parciales o completas de las lectinas tipo P2 presentes en el

género *Dioclea*, *Canavalia* y *Cymbosema* que nos permiten hacer una aproximación a la secuencia de la lectina de *Galactia lindenii* P2.

4.1 Extracción y purificación de la lectina P2 de *Galactia lindenii*.

4.1.1 Obtención del extracto proteico.

Los extractos crudos presentaron una coloración tenue y una buena actividad de aglutinación con eritrocitos O⁺. Experimentos con *S. bogotensis* (Jiménez, 2000) y *L. bullata* (Sanabria, 1999) mostraron que la adición de tiourea 5 mM (Van Driessche et al, 1983) al buffer de extracción (PBS) reduce sustancialmente la oxidación enzimática, por lo tanto es importante incluirlo dentro del proceso de extracción. La pigmentación aumento un poco luego del proceso de concentración y del cambio de buffer; por lo tanto antes del proceso de purificación por cromatografía de afinidad se incluyó una cromatografía de intercambio iónico para eliminar la coloración; en el esquema reportado por Almanza et al. 2004, incluyeron una cromatografía de intercambio en batch empleando DEAE celuosa. Al final del proceso de extracción, los extractos crudos fueron completamente solubles y no presentaron una viscosidad alta como sucede en el caso de la extracción de las lectinas de labiadas (Vega, 2004). En la tabla 6 se presentan los datos de cuantificación y actividad eritoaglutinante de los extractos crudos derivados de las semillas de *Galactia lindenni*; el proceso se extracción se hizo seis (6) veces con cantidades diferentes de harina, 5, 10 y 20 gramos.

Tabla 6. Extracción de proteína a partir de la harina de *Galactia lindenii*

Extracción*	Harina (g)	Volumen (mL)	Concentración (mg/mL)	Cantidad total de proteína (mg)	Rendimiento (mg proteína/g harina)	Aglutinación (O)**
E2	5	22,50	12,1	279,0	55,80	+4
E3	10	60,00	4,49	269, 4	26,94	+4
E6	20	100,0	1,74	174, 2	8,71	+4

* Se presentan los datos del pool de extractos concentrados por ultrafiltración y con el cambio de buffer a PBS

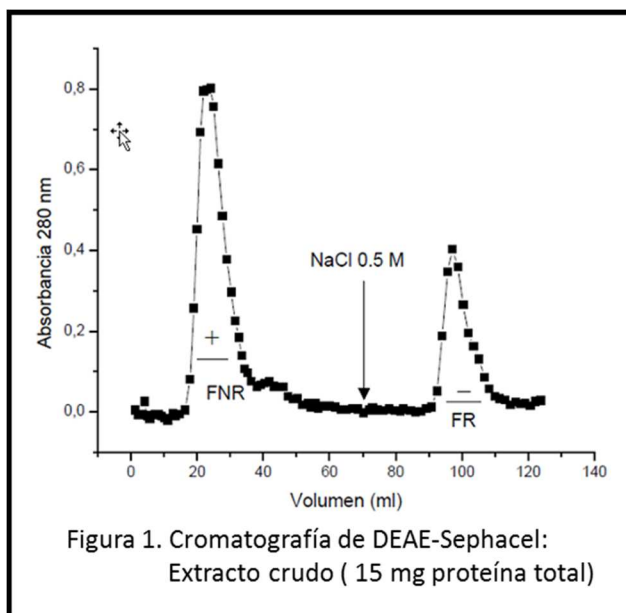
** La actividad eritroaglutinante se determina cualitativamente empleando una escala donde +4 es la máxima aglutinación, significa que la proteína aglutina todos los eritrocitos. En todos los casos se usaron 100 µl de extracto crudo.

A partir de 5 g de harina se obtuvieron 279,0 mg de proteínas totales lo que indica que el proceso de extracción fue eficiente con esta cantidad de harina; Almanza et al, 2004 reportaron un total de 161 mg de proteínas totales. Empleando una mayor cantidad de harina, al final del proceso de concentración se pierde una gran cantidad de proteínas por insolubilización.

Aquí es importante aclarar que la determinación de la concentración se hizo por el método del ácido bicinoninico y que se puede presentar una sobrestimación de la cantidad de proteína por interferentes del extracto como azúcares reductores que no se eliminan completamente luego del proceso de concentración y diálisis. En el caso de Almanza et al., 2004 la determinación se hizo por microKjeldajhl por lo tanto no es conveniente comparar los resultados obtenidos.

4.1.2 Purificación de la LGL-P2.

Para obtener extractos más traslúcidos se realizó una cromatografía sobre DEAE-Sephacel para el extracto concentrado. La actividad de la lectina se obtuvo en la fracción no retenida (Fig.1), la cual se caracterizó por ser completamente soluble y sin pigmentos visibles; algunas proteínas contaminantes, polisacáridos y los pigmentos que presentaba el extracto fueron retenidos por la columna.



En el transcurso de los ensayos se observó que la segunda lectina de *Galactia lindenii* (LGL-P4) se encontraba en la fracción retenida, así que este paso de purificación a la vez que nos permite obtener una fracción apropiada para realizar cromatografía de afinidad y purificar la LGL-P2, también nos permite separar las dos lectinas, las cuales deben presentar puntos isoeléctricos diferentes. La fracción I obtenida de la cromatografía sobre DEAE-Sephacel se sembró sobre la columna de Lactosa-Sepharosa 4B (Figura 2), la elución de la lectina se hizo con lactosa 0.2 M como había sido reportado por Almanza et al., 2004. Se observa una fracción retenida con una absorbancia apreciable a 280 nm, lo que nos indica que la cantidad de proteína retenida es alta y tiene alto contenido de aminoácidos aromáticos, lo anterior está de acuerdo con el coeficiente de extinción determinado para la proteína pura $A^{1\%1\text{cm}} 12.00$; este último también nos permite determinar la concentración de las soluciones de la lectina pura. La electroforesis, los ensayos de actividad y la cuantificación para la fracción retenida se realizaron después de dializarla y liofilizarla; los resultados se presentan en la Tabla 7. Al reconstituir el liofilizado se observa muy poca precipitación o pigmentación y se obtiene una concentración del orden de 2.6 mg/ml y una actividad de eritroaglutinación alta; para la fracción no retenida (I) no se detectó actividad de lectina. Por SDS-PAGE con

calentamiento se observa (Figura 3, carril 5) para la fracción II una banda mayoritaria alrededor de 24 kDa (70%) que corresponde al monómero de la LGL-P2. Algunas veces se observa una banda alrededor de 50 kDa que corresponde al dímero de la lectina, el cual aparece por disociación incompleta del tetrámero. Por otra parte también se observan algunas bandas minoritarias (8%) de bajo peso molecular correspondientes a contaminantes.

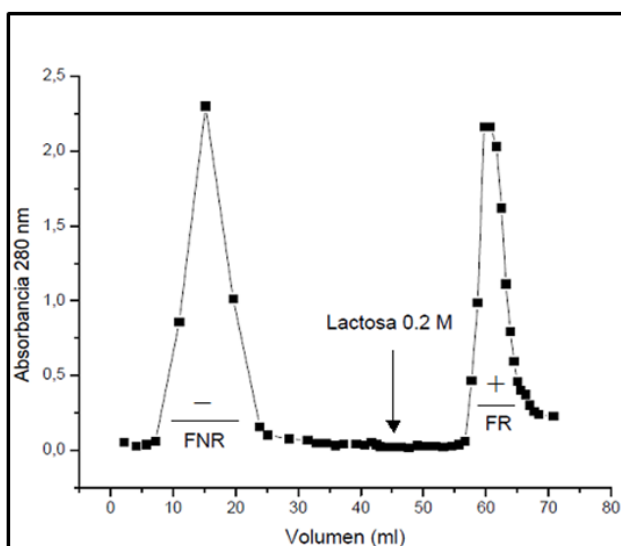


Figura 2. Cromatografía de afinidad sobre Sepharosa 4B-Lactosa. Fracción I-DEAE sephacel (93 mg de proteína total).

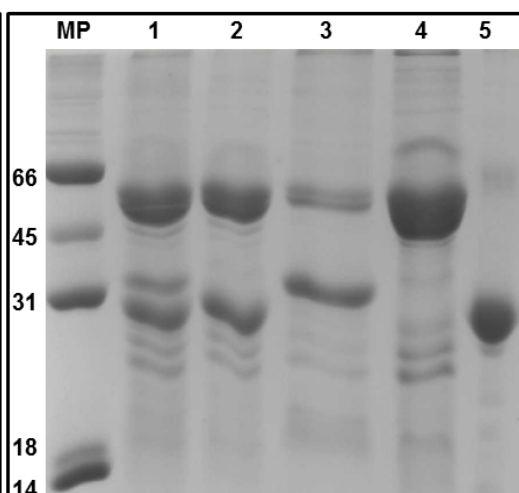


Figura 3. SDS-PAGE - Purificación de la lectina LGL-P2 (20 µg de proteína / carril). Carril MP: Marcador de peso molecular (kDa). Carriles: 1. Extracto crudo. 2. Fracción no retenida DEAE-Sephael. 3. Fracción retenida DEAE-Sephael. 4. Fracción no retenida de cromatografía de afinidad Sepharosa 4B-Lactosa. 5. Fracción retenida de cromatografía de afinidad Sepharosa 4B-Lactosa (LGL-P2).

Como uno de los objetivos es realizar la hidrólisis de la lectina para obtener péptidos no es conveniente tener estas impurezas, por lo tanto la proteína se repurificó. En la figura 4 se observa que con cualquiera de los dos métodos ensayados, la fracción de proteínas contaminantes es muy baja y por SDS PAGE no se observan las bandas contaminantes.

El extracto crudo purificado por afinidad sobre Lactosa-Sepharosa 4B mostró esencialmente el mismo perfil de elución (Figura no mostrada) sin cambio en el volumen de elución de la fracción obtenida con Lactosa 0.2 M. La fracción retenida después de dializarla y liofilizarla, mostró actividad y una concentración apreciable de proteína (18 mg/ml) (Tabla 7); sin embargo estas fracciones se precipitaron y presentaron una pigmentación residual; en el soporte de afinidad quedan retenidos algunos de estos pigmentos que no se pueden eliminar y bloquean la Sepharosa irreversiblemente.

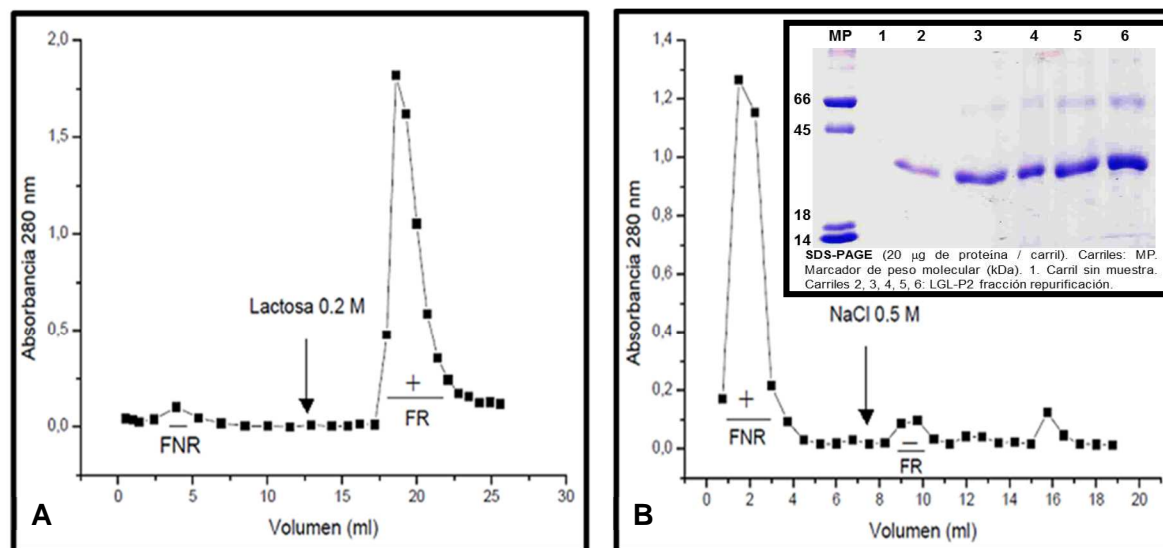


Figura 4. Recromatografía para la Fracción de LGL-P2 : A. Afinidad sobre Sepharosa 4B-Lactosa. B. Intercambio iónico sobre DEAE – Sephacel

Tabla 7. Purificación de la lectina de *Galactia lindenii* P2 (LGL-P2).

Procedimiento	LGL-P2 (mg/mL)	Volumen (mL)	(mg) Totales	Aglutinación O ⁺
Extracto crudo sobre Sepharosa 4B-lactosa	18,0	0,50	9,00	+4
FNR DEAE Sephacel sobre Sepharosa 4B-Lactosa	2,60	0,50	1,30	+4
LGL-P2- pura sobre Sepharosa 4B-Lactosa	4,10	1,20	4,90	+4
LGL-P2 pura sobre DEAE Sephacel	0,18	2,00	0,36	+4

Tres extracciones consecutivas fueron suficientes para obtener soluciones donde fuese observable la actividad de la lectina. Con el esquema de purificación empleado se obtuvieron aproximadamente 100 mg de proteína pura /100 gramos harina, lo cual es menor que lo obtenido por Almanza et al., 2004. Durante el proceso de purificación se presentaron pérdidas de la proteína, en algunos casos se precipitó. Con el esquema de purificación las soluciones de LGL-P2 alcanzaron una concentración en un rango de 8-10 mg/ml, por lo tanto las determinaciones por UV y los perfiles cromatográficos se ajustaron al contenido real de proteína. Para las lectinas de *Galactia Tashiroi* y *Galactia tenuiflora* no existen reportes acerca de la cantidad de lectina pura / 100 gramos de Harina o semilla aunque las dos fueron purificadas por cromatografía de afinidad

(LePendou et al, 1986 Fukuda et al, 1994). En las lectinas del genero *Dioclea* se han reportado entre 1.210 – 2.000 mg de lectina pura/100 g harina seca en el caso de las lectinas tipo P4 que reconocen Glc/Man y en el caso de las lectinas tipo P2 que reconocen lactosa entre 100-870 mg /100 g de harina seca (Pérez G., 1998; Calvete et al, 1999; Sierra y Pérez, 1999; Melgarejo et al, 2005). El contenido de lectina en especies de la familia *Lamiacea* es muy bajo con respecto a lo que ha sido reportado para lectinas de leguminosas, entre 5 -20 mg lectina / 100 g de semilla (Pérez y Vega, 2007).

4.2 Electroforesis preparativa para la obtención del monómero de la LGL-P2.

Uno de los objetivos del proyecto se enfoca en la obtención de la secuencia parcial o completa de la LGL-P2, lo anterior significa que la proteína debe ser pura y se debe asegurar una buena accesibilidad de los agentes proteolíticos para realizar la hidrolisis y obtener un número de péptidos que sean adecuados para estudios de secuencia. Los ensayos de electroforesis preparativa se hicieron con el fin de obtener el monómero de la LGL-P2 y que la digestión enzimática fuera más eficiente. Inicialmente se trabajó en la obtención de la curva de calibración empleando proteínas de pesos moleculares intermedios con el fin de identificar el tiempo de elución del monómero de la LGL-P2. Inicialmente se hizo una corrida con cinco patrones de proteína de peso molecular intermedio, pero no se obtuvo una buena resolución, en la figura 5A, se observa que en la fracción I, se encuentran la lisozima y el monómero de la mioglobina que son de bajo peso molecular y que por sus valores de peso molecular tan cercanos (14.3 kDa y 18.2 kDa respectivamente) no hubo la suficiente resolución del método empleando un gel del 10%; sin embargo este comportamiento fue generalizado para todas las proteínas analizadas (BSA, Ovoalbúmina, tripsinogeno, β -lactoglobulina y lisozima). Por SDS PAGE (Figura 5B) se observó que proteínas de bajo peso molecular estaban presentes en volúmenes de elución cercanos a 150 ml.

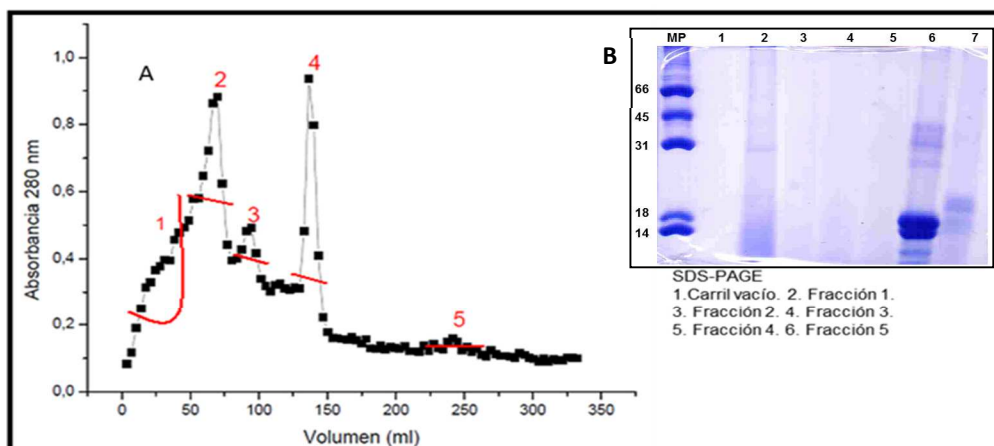


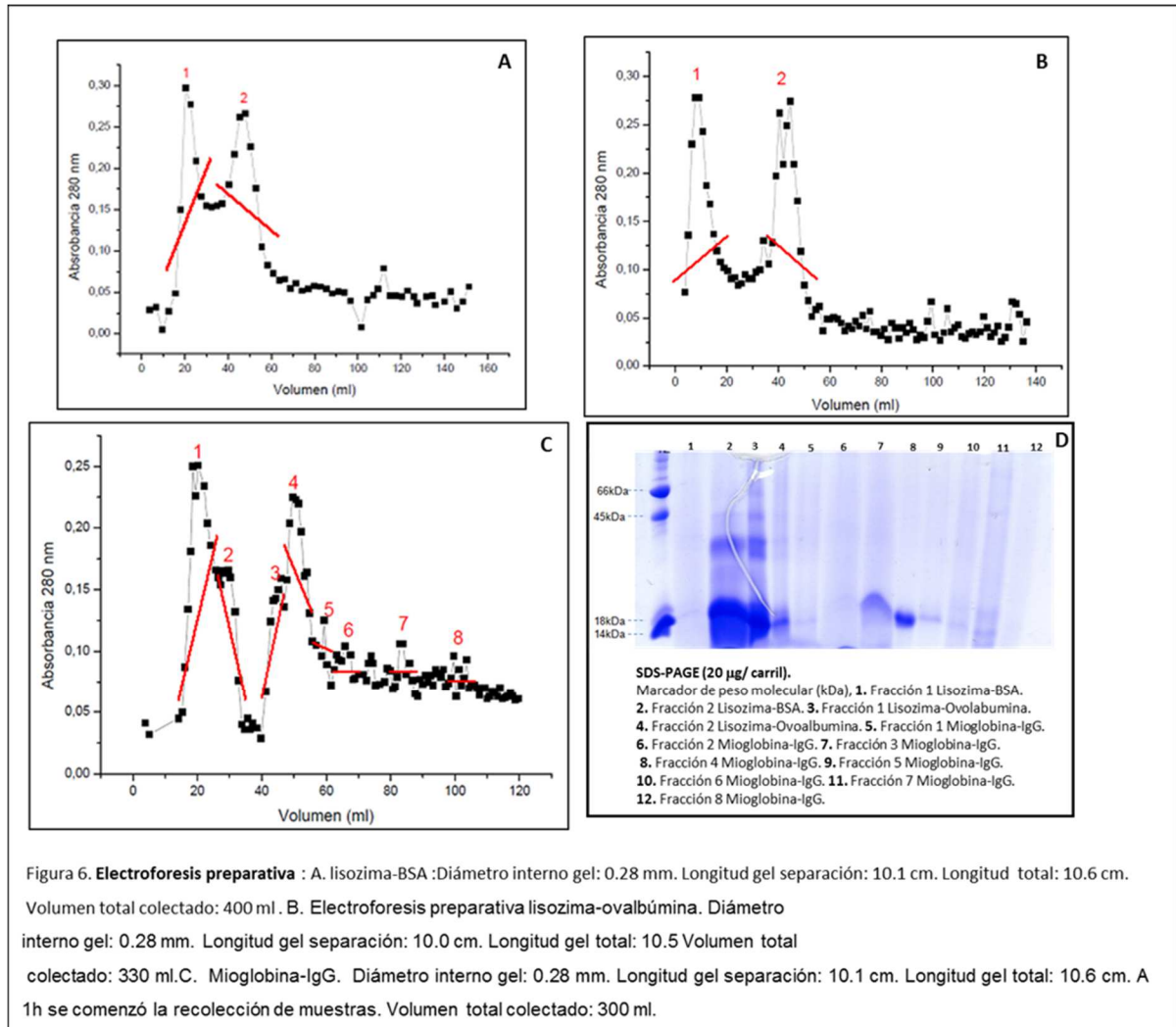
Figura 5. Electroforesis Preparativa –Curva de Calibración

Patrones (kDa): Lisozima (14.3), β -lactoglobulina (18.8), Tripsinogéno (24), Ovoalbumina (45), BSA (66).

Condiciones: Diámetro interno gel: 0.28 mm. Longitud gel separación: 10.3 cm.

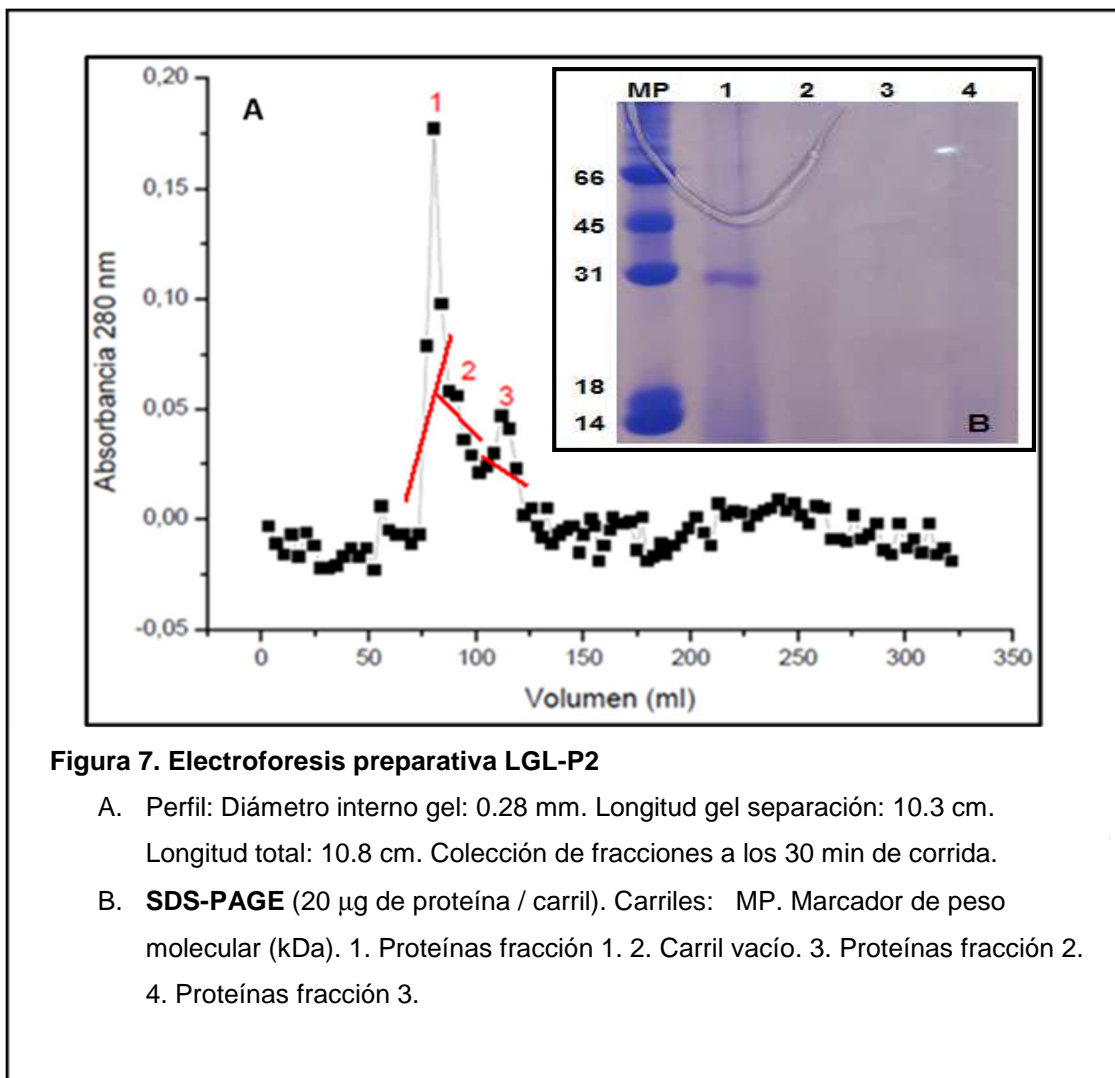
Longitud total del gel : 10.75 cm. A las 2h 11 min se comenzó la recolección de muestras .

Con estos resultados no se pudo asociar ninguna de las fracciones y por supuesto los volúmenes de elución con ninguna de las proteínas presentes en la mezcla. Posteriormente y teniendo en cuenta los resultados se hicieron diferentes corridas ensayando con mezclas pares de proteínas. En todos los casos (Figura 6 A, B, C) se observaron dos fracciones no resueltas y por SDS-PAGE (Figura 6D) se corroboraron los resultados. En consecuencia no fue posible elaborar una curva de calibración bajo las condiciones propuestas; el método de electroforesis preparativa teóricamente puede separar proteínas con diferencias en peso molecular del orden 1kDa pero se deben estandarizar las condiciones adecuadas como longitud del gel, cantidad de muestra, tamaño de poro y diámetro interno de columna entre otros. Sin embargo la LGL-P2 pura se corrió en electroforesis preparativa bajo las mismas condiciones descritas para los patrones de peso molecular.



La lectina se detectó en la Fracción I (Figura 7B, Carril 1) y se recuperaron 100 μg de monómero (350 μl ; 0.3 mg/ml), el 10 % de la proteína pura inicial. Este porcentaje de recuperación tan bajo se puede atribuir a pérdidas durante el proceso de diálisis y concentración de la proteína. Se realizaron otras electroforesis con el objetivo de recuperar más LGL-P2 pura (perfiles no presentados), se colectaron tres fracciones en las cuales no se observó la presencia de proteínas por la baja cantidad de recuperación en cada pico (0.1 mg/ml, fracción I). Al final del proceso de electroforesis preparativa se obtuvieron 150 μg del monómero de la LGL-P2. Estos resultados nos muestran que por este método los porcentajes de recuperación (10.5%) son muy bajos y menores a lo

reportado para otros ensayos de electroforesis preparativa corrida en condiciones muy similares (Fountoulakis y Juranville., 2003) donde el porcentaje de recuperación fue del 25%, sembrando 50 mg de proteína. En los perfiles también se observó que bajo condiciones denaturantes y un calentamiento previo de la muestra, se presenta una muy baja proporción del dímero y no se evidencia el tetrámero (Figura 6A).

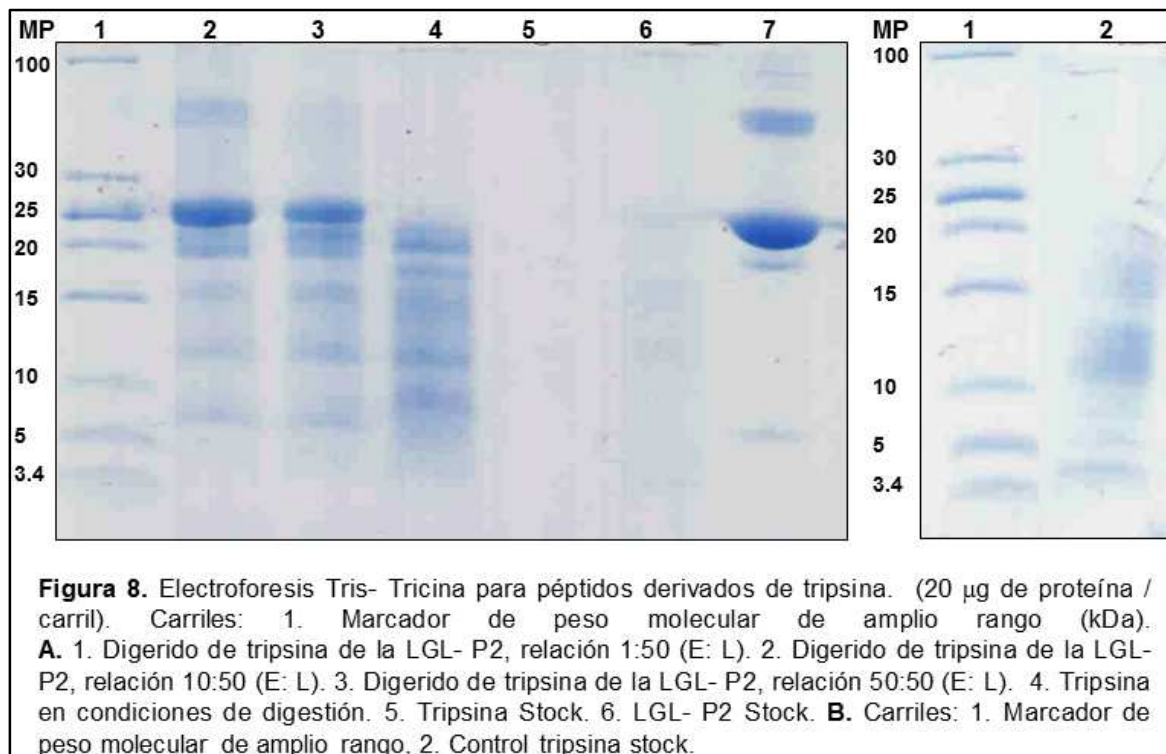


Para realizar los ensayos de hidrolisis enzimática también se trabajó con LGL-P2 repurificada (Figura 4) en condiciones denaturantes y con calentamiento antes de las digestiones para asegurar una buena accesibilidad de los agentes proteolíticos.

4.3 Ensayos de digestión enzimática para la obtención de péptidos de la LGL-P2

4.3.1 Digestiones enzimáticas con la solución de LGL-P2

Se siguió el procedimiento de Darville et al., 2012 con cada una de las enzimas propuestas. Según la composición de aminoácidos para LGL-P2 (Almanza et al., 2004, anexo 9) al realizar digestión con tripsina se esperaban 19 fragmentos. Se evaluaron diferentes relaciones enzima: lectina (E: L) con diferentes preparaciones de tripsina y la fragmentación fue evaluada por electroforesis empleando el método de tris-tricina. En el caso de la relación 50:50, se observó que la proteína fue hidrolizada en su totalidad ya que se observa una ausencia de la banda de 25 kDa con respecto a las otras dos mezclas (1:50; 10:50) (Figura 8A , carril 1,2,3 y 6) sin embargo el control electroforético de la tripsina Sigma tipo III de páncreas bovino (Figura 8B, carril1) presentó bandas con pesos moleculares entre 20-25 kDa ,lo que indica que la enzima sufrió autólisis y que por lo tanto no es adecuada para estos ensayos ya que sus péptidos pueden interferir en los procesos de purificación y secuenciación de los péptidos derivados de la LGL-P2.



Los resultados obtenidos nos llevaron a evaluar otras tripsinas con el objetivo de eliminar contaminantes o que estuvieran en una cantidad muy baja, para ello se evaluó una tripsina grado proteómica con una mejor actividad (Amresco) en una relación 10:10 (E: L). Esta enzima tenía la ventaja de no sufrir autólisis ya que sus lisinas están metiladas; sin embargo no se observó fragmentación alguna de la LGL-P2 (Figura 9 A). Lo anterior nos indicaba que se debía aumentar la concentración de enzima o mejorar las condiciones de accesibilidad de la enzima a los sitios de corte. Por lo tanto la LGL-P2 se incubó con SDS al 10% durante 4 horas a 70-80°C y se emplearon otras enzimas para verificar fragmentación de la LGL-P2; también se sometió a digestión enzimática la lisozima para evaluar las condiciones experimentales. En la figura 9B se observan los resultados; se encontró que las muestras tratadas con quimotripsina no arrojaron buenos resultados. Se observó fragmentación en los controles (Carril 1-5) y no hubo ninguna digestión de LGL-P2 (carril 7, 8) ya que las bandas tanto del monómero como del dímero estaban presentes. Para el caso de la lisozima (carril 4) se observaron bandas de peso molecular similar a las observadas en el control de digestión de la enzima, por lo que no se pudo confirmar la digestión de la proteína.

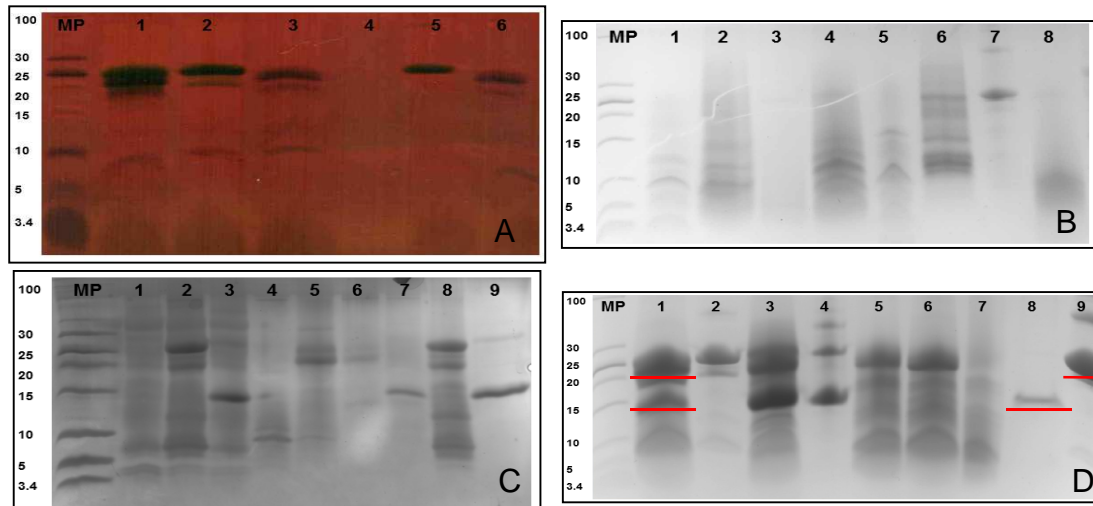


Figura 9. Electroforesis Tris-Tricina. Tripsina, quimotripsina, pepsina. (20 µg de proteína / carril). Carriles: MP. Marcador de peso molecular de amplio rango (kDa). **A.** 1. Digerido de tripsina de la LGL- P2, relación 10:10 (E: L). 2. Tripsina Stock. 3. LGL-P2 en condiciones de digestión. 4. Carril vacío. 5. LGL- P2 Stock. 6. Tripsina Stock. **B.** (50 µg/ carril). 1. Quimotripsina SIGMA Stock. 2. Quimotripsina SIGMA Control digestión. 3. Digerido de quimotripsina SIGMA de la LGL-P2, relación 50:100 (L: E). 4. Digerido de quimotripsina SIGMA de la lisozima, relación 50:100 (P: E). 5. Quimotripsina MERCK Stock. 6. Quimotripsina MERCK Control digestión. 7. Digerido de quimotripsina MERCK de la LGL-P2, relación 50:100 (L: E). 8. Digerido de quimotripsina MERCK de la lisozima, relación 50:100 (P: E). **C.** Digestión enzimática con tripsina Difco, quimotripsina MERCK, quimotripsina SIGMA y pepsina MERCK. (50 µg/ carril). Carriles: MP. Marcador de peso molecular de amplio rango (kDa). 1. Carril sin interés. 2. Carril sin interés. 3. Carril sin interés. 4. Tripsina Difco Stock. 5. Tripsina Difco Control digestión. 6. Digerido de Tripsina Difco de la LGL-P2, relación 50:100 (L: E). 7. Digerido de Tripsina Difco de la lisozima, relación 50:100 (P: E). 8. LGL-P2 control digestión. 9. Lisozima control digestión. **D.** (50 µg/ carril). 1. Digerido de tripsina de la LGL-P2, relación 50:100 (L: E). 2. LGL-P2 control digestión. 3. Digerido de Tripsina de la lisozima, relación 50:100 (P: E). 4. Lisozima control digestión. 5. Tripsina Control digestión con SDS únicamente. 6. Tripsina Control digestión con SDS v DTT. 7. Tripsina Stock. 8. Lisozima Stock.

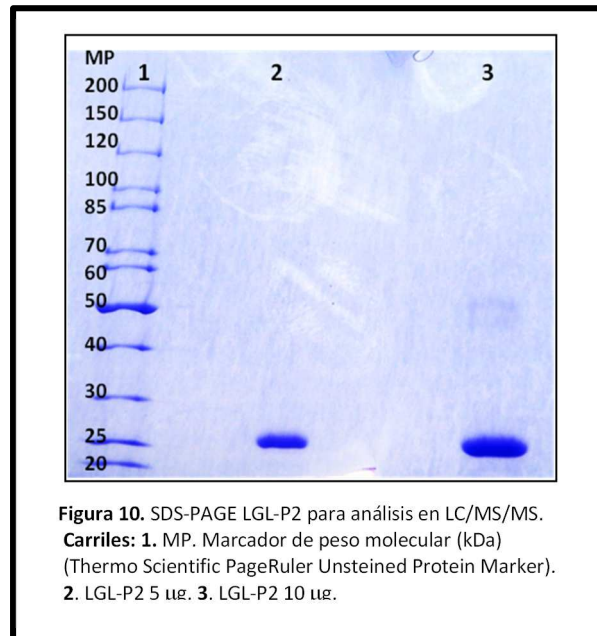
En la figura 9C se observó que la tripsina marca Difco había sufrido una autólisis marcada (ya que se observaron dos bandas en el carril 4 de 10 y 8 kDa aproximadamente) y con la LGL-P2 (carril 6), se observaron dos bandas, una de 20 kDa y una de 13 kDa; estas últimas bandas también se evidencian en los controles de digestión de la tripsina (carril, 5). Con estos resultados no es posible asegurar que las bandas encontradas en el digerido de la LGL-P2 correspondan a péptidos de la lectina. Otros ensayos empleando una tripsina (tratada con TPCK) arrojaron una hidrólisis muy parcial, donde se observó únicamente un fragmento de 15 kDa (Figura 9D, carril 1 y 2). En el control de la enzima (Figura 9D, carril 7) se observa que hay degradación, razón por la cual la actividad de proteólisis fue muy baja y adicionalmente hay fragmentos de bajo peso molecular que contaminan los péptidos obtenidos a partir de la LGL-P2. En el caso de la lisozima no hubo hidrólisis.

Los ensayos de hidrolisis realizados para la LGL-P2 en solución mostraron una gran dificultad para fragmentar la molécula debida principalmente a la actividad de las enzimas empleadas, por lo tanto se procedió a ensayar otros métodos de fragmentación y enzimas con otras características.

4.3.2 Digestiones enzimáticas en gel de la LGL-P2

Para obtener un mayor número de fragmentos, se cortó la banda de 15 kDa previamente obtenida y se sometió de nuevo a digestión con tripsina (SIGMA, tratada con TPCK) (Figura 9D, bandas señaladas en color rojo bajo las mismas condiciones descritas previamente). Los resultados no mostraron la presencia de proteínas ni fragmentos derivados de hidrolisis. A pesar de seguir los métodos reportados para la hidrolisis de lectinas de *Diocleae* (Richardson et al, 1984; Almeida., et al, 2011), no se obtuvieron resultados favorables; como se mencionó previamente las enzimas empleadas no fueron activas.

Se realizaron otros ensayos corriendo la LGL-P2 en geles de poliacrilamida (Figura 10); las bandas de interés se fraccionaron y fueron sometidas a digestiones con tripsina, quimotripsina y LysC durante 16 h. Para evaluar la eficiencia de la hidrolisis los digeridos fueron sometidos de una vez a cromatografía líquida. Como veremos más adelante en este caso las enzimas empleadas hidrolizaron la lectina, lo que nos indica una vez más la baja actividad de las enzimas antes empleadas en los ensayos anteriores.



4.3.3 Análisis de los péptidos de la LGL-P2 por Cromatografía Líquida-Espectrometría de masas en tándem (LC MS/MS).

Se encontraron 123 péptidos DeNovo en el caso de la Tripsina, 573 con Quimotripsina y 168 con LysC, generados por PEAKS 6 DeNOVO, dentro de los cuales había redundancias que fueron eliminadas posteriormente.

4.3.4 Péptidos derivados de digestión con tripsina

La comparación de la secuencia de los péptidos con la base de datos para identificar la proteína, se realizó contra la base de datos “PD1.3/Mascot protein/peptide identification” y se crearon los reportes en Scaffold. Los resultados estuvieron relacionados con queratina humana, tripsina y BSA para los tres digeridos. Estos resultados se obtuvieron debido a la falta de una secuencia del genoma de la planta o una secuencia similar a la de la lectina, que nos diera un punto de comparación y así facilitar el ensamblaje de los péptidos. En los anexos 10, 11,12 se presentan los resultados obtenidos contra la base de datos PD1.3/Mascot protein/peptide, a partir de los cuales con un reporte en Scaffold, las proteínas listadas fueron identificadas en al menos una de las muestras con un nivel de confianza $\geq 95\%$ basado en la identificación de al menos dos péptidos con un nivel de confianza de $\geq 95\%$. Los números en los cuadros de color son los números de péptidos

únicos identificados con un nivel de confianza $\geq 95\%$. Se observó que se encontraron péptidos pertenecientes a secuencias de proteínas humanas, las cuales se presentan por contaminación con queratina y otras proteínas humanas que son contaminantes durante el proceso de preparación y análisis de la muestra. Sin embargo la LGL-P2 hidrolizada estaba en un gran porcentaje (gel, Figura 10) por lo tanto es poco probable que estén ausentes los péptidos de la LGL-P2.

4.4 Aproximación a la estructura primaria de LGL-P2.

4.4.1 Análisis Bioinformático de los datos obtenidos

Empleando el programa Decrease redundancy de Expasy (Artimo P, et al 2012) se eliminaron las secuencias de los péptidos “iguales” con un nivel de identidad mayor o igual al 90%. Los resultados se presentan en las Tablas 8,9 y 10.

Tabla 8. Secuencias de péptidos obtenidos con tripsina por LC/MS/MS de LGL-P2		
TLLDLNTR	LANFVSSLTFR	LAGNGMLSLTSK
VNVQQLEGR	GNGMLSLTSK	VGLSASTSSGGVQR
NYLGAVNPK	SLASSLPGQTK	VGLSAATSGNGVQK
NVAVPLYNR	LSLSTSGGSFR	AALQGDGMLSLSK
STMQELNSR	LTGLLGAPFSK	VVGLSASTSSNVQR
YLDGLTAER	VASFVSSLTLR	SNGFSSGSAGLLNYQR
RMPATLNSR	TGGYLGAVNPK	VAGGGGGFGAAGGFNR
MGLSLTSK	NVKVNPQLQNVK	PSGSAGGGALGTPNPK
QFSSSYLSR	LNNLEALQQAK	PSGSTNYLGALGAVNPK
GMLSLTSK	AGGGALGVVNP	VVGLSASTSSGGVQR
KNPSYQLGR	LSASTSSGGVQR	LPSGSTGGYLGAVNPK
SSPATLNSR	TSLKNPSYQLGR	SLVNLGGSKLSLSVAR
LASYLDKVR	SAGGQLGVVNP	LQPSGSANGALGVVNP
FASFLDKVR	TSYKNPSSNLGR	PSGSTGGYLGALGAVNPK

SFVSSLTLR	SSGGFSGGSFSR	SLGGGGFSSGGFSGGSFSR
SGFSSVSVSR	SKAQASLYQSK	PSGSAGGQLGVVSGVNP
TQAQSWYQTK	LSAATSGNGVQK	GPYQSGSGHSSGLGHQQR
LKNPSYQLGR	LQGNLLALTSYK	AQAESLYQSKYEELQLTAGR
AEAESLYQSK	GSTGGYLGAVNPK	PSYPHLGLNVNSLLSLQTA
LLAKVFGYQD	LQGNMGLSLTSK	LGEHNLNVLEGNEQFLNAK
YKNPSSNLGR	NKLNLDLQDALQAK	SKAEASLYQSKYEELQLTAGR
NHEQQMKNLR	SLNNQFASFLNKVR	LSSKGLGGGFSSGGFSGGSFSR
VAGFVSSLTFR	TLNNKFASFLNKVR	ALHSSPVTLCFAHGNVASFVSSLTLR

Tabla 9. Péptidos obtenidos con quimotripsina por LC/MS/MS de LGL-P2

LPSGSTGGY	QVSLQDKTGF	SPGSAGGQLGVVNPDAFNNF
LSNTNCKKYW	LGAVNPKDAGW	HNLLEGGQEDFESSGAGKGL
VTLKDVGDY	VGVGPKDAFNNF	SAVENDGQLLVAQMVDKL
TSLKNPSYF	NGVQKHDVY	VSAVENDGELLTVAQMVDKL
GPSGDGVQKHDVY	SLKNPSYQGL	MHTTSDVVVAGEFDQGSSEKLQKL
LQLDNARLAADD	EAHPVTLWDEATGNVASF	EGDGGEVASVLLTVAQMVDKL
VVTQVSLQDKTGF	NPKDAFNNF	VQALEEANNDLENKLQDWKDY
SAATSGDGVAGKHDVY	APDVAVRLEDVSGLM	VGVEFDDYKNPSSNLGRVLY
STSSGGVQRHDVY	LKNPSSNLGVRLY	VGVEFDDYSNSWDPSYAKF
QVSLQDKTGFHF	KNPSEPAGRVLY	GAPDAGTLALPSGSTGGY
SPGSTGGYL	NPSSNLGRVLY	NPKDAFNEPLTVAQMVDKL
ELEGKLEW	LGAVNPKDAFDNF	APGADTGALPSGSTGGY
VTVKAALLY	SDSWDPSYPHLGLDVNSL	VGVEFDDYSNSWDPSYSF
DPSYPHLGLDVNSL	VGVEFDDYF	LWDEATGNLANF
SAGPSGDGVQKHDVY	HSSPVTWDEATGNAW	LMLVNPDAFNNF
SLTSLKNPSY	NLGGFGGSRF	HSSVPTWDEATGNVASFY
QNLNDRASY	AADPSSNLGRVLY	VNGSAEDSVVNPVSSYAGF
DGVQKHDVY	LGAVNPKDAF	EMEQGGLRQGVADNL

SPYPHLGLDVNSL	LPEAAFLVKAALLY	VPDPSVTLWDEATGNVASF
WDEASGNVAGF	HSSPVTWDEATGNVASF	HSSPVTWDEATGNVADF
LNGTAEANF	LGAVNPKDAFNNFF	QPDAGTLFALPSGSTGGY
KTVLPSKVRRLF	VNPKDAFNNF	HSSVPTWDEATGGGVASF
LGAVNPTAAAFF	LLGAVNPRSAFNNF	VPTLSSMEATGNVACASF
WSDPSYPHLGLDVNSL	VRALEESNY	SSPAPLWDACATGNLANF
ALKQSLEASLAETEGRY	SPGSTLALGAPYY	SHSVPTWDEATGNVASF
YSSPVGGLW	RTAPSVSGYVPADGVL	SPGSAGGQLLALGAPYY
SGDGVQKSM	VPADGLVYF	SAATSGNRQWHDVY
VGVEFDDYSNSW	LKNPSSNLRGVLY	VGVEFDDYSNSWDPGYAY
DQGSSEKLQKL	TSDVVVAGEF	VDLKTVLPSKVRVGL
NLNDRLASY	KAVLPSKVVGGLF	SASTGAVNPKDAFNNF
NPKDAFDGW	SAATSGNGVSPLHDVY	AEPVNLGEDEATGNLANF
ELEGKPAAGF	SAATSGNGVPPSSGDVY	TKNPSSNLGRVLY
VDLKTVLPSKVRRLF	VSDVVAGEF	SASTSGAVNPKDAFNNF
YKTLDDLKNQLL	AATPAFDGW	VPTLDADEATGNVASSSF
GDGVQKHDVY	QPALQDKTGF	LGNPSSNLGRVLY
VDLGAVNPKDAFNNF	TTLALQPYY	VDLKAVLPSKVVGGL
KDPSSNLRGVLY	SPSNLGRVLF	SPGSAGGKMLFALQPYY
SGGVQRHDVY	SPGLFALGAPYY	SLTKNPSSNLGRVL
SSNLGRVLY	SLNPSSNLGRVLY	SPGSALELFMALQPYY
VSLQDKTGF	KAVLPSKVVGGLF	HSSPVTLGEDEATGGGVASF
SAATSGDGVQKHDVY	TSGDGVQKHDVY	SASTSSGGVQRGNDVY
SPGSAGGQLF	VGVNPKDANF	SSPVNLWDEATGNLAGW
WDEATGNVASF	VPADGLVFL	STSDGGVQRHDVY
VDLKTVLPSKVVRGL	AFHLAATVVM	YNTLDDLSRQLVDLTVGKPM
KDPSSNLRRLY	SASTSLRNTGSL	HSSPVTWDEATW
VNGSAEDAGF	APDVARLEDVSGPF	SAATSGNGKGGKHDVY
VPADGLVFF	AATPAFNPM	LNVPSYWDEATGNLANF
SPVTWDEATGNVASF	VVGGGPSAVW	VPSSYNLWDEATGNLANF
LDKVRALEESNY	VSAVENDGQLL	SPYPHLGLDVNSAFYCY
SPGCLPSGSTGGY	LLALGAPYY	VGVGPKDADGW
DDYSNSWDPSYSF	SPLFALGAPYY	VSAVENDGELLAQAM

LGAVNPKDAFPM	SLQDKTGFHF	VPTLDADEATGGGVSSASF
AALLYDCHAKTL	SLGGGGFSSGGF	NSSRHLGLDVNSLDCNSL
EPALQDKTGF	APDGLELFF	SPGRAPQDTGALPSGSTGGY
SGNGVQKHDVY	SPKVRVGLF	HSSPVTLCVDEATGNVASF
WAASSGGVQRHDVY	EPLAAFLVKAALLY	HSSPVTLMSDEATGNVASF
TVLPSKVRVGLF	AGDVKKLTVYM	LGAVNPKDAFNPM
SPSNLGRVLY	VGVGPKDAFDGW	GTSTSSGGVLWHDVY
LFALGAPYY	YTLKHALTVYM	LEGVFDDYSNSWDPSYSF
VDLKTVLPSSL	VGVNPKDAF	EGALAEQNRKDAAEW
KSDQSRLDSEL	SFYSKLDTSRSF	VGVNPKDAFNFF
VDLKNPSSNLGRVLY	VGDGQKVSM	SSPVNLWDEATGNLANF
SLTGAVNPKDAFNFF	KAVLPSKVVGNL	VVGGVGVKAVLY
ATSGDGVQKSM	LDKVRFLEQQNQL	VPNLSSMEACATGNLANF
HSSPVTLWDEATGNGL	SAATSGNGVQSLHDVY	QPDAGTLFALPGSSTGGY
LGAVNPKDAGGF	ANGALGVVNPKDAFNFF	QPDAGTLPSTGGY
LGVVNPKDAF	LGDAVYVF	SSPVNLWDEATGNKAPM
DEATGNVASF	VSAVENDGELLTVAQM	VPNLMSDEACATGNLANF
SPGSAFLALGAPYY	LNGTADEANF	NSSRHLGLDVDCDGLCL
RLEDVSGYVPADGLVF	QLGAVNPSRAFNNF	HSSPVTLGEDEATGNVASF
AEEGNVLVSGPWPW	SSGGVQRHDVY	QPDAGTLAALLPSGSTGGY
TSASTSSGGVQRHDVY	APDVARLQDVSGLM	LAGTDQPALPSGSTGGY
ATSGDGVQAHTAVY	GGVQRHDVY	ATDGNGVQAHTAVY
SKLDTSSRF	LGAVNPSRSFNFF	VSAVENDGQLLTVQAM
GVVNPKDAF	WSTSSGGVQRHDVY	HSSVPTLWDEATGNVASF
SDSWDPSYSF	QVGGPSRAFNNFF	SPSTSSGGVQRHDVY
LTVPSSTGF	VNPKDAFDGW	SAATSGNGVQEASVY
VPADGLVYF	SLDLDSLLADKL	SAATSGDGVQKGHVY
VGVGPKDAGW	VGAFPLSPGPEL	TRAPSVSGYVPADGLVF
SAATSGDRQKHDVY	TLWDEATGNVASF	APGDATGALPSGSAGGQLGVVNPKDAF
SPGSTFLALQPY	KNPSSNLRRLY	SPGEAAGLFLPSGSAGGQLGVVNPKDAF
SGSLVKAALLY	-----	-----

Tabla 10. Péptidos obtenidos con LysC por LC/MS/MS de LGL-P2

LLNDEDPVVVTK	LENLFEALNNK	PSYPHLGLDVNSLLSLQTAK
VLEHLMEDLDTNADK	EEANNDLENK	TLLDLNTRMTLDDFRLK
ATSGDGVQK	YYKTLDDLK	ELETYHNLLEGGQEDFESSGAGK
PSGSTGGYLGAVNPK	VRALEEANADLEVK	AFSAVDTDGNGLTNAQELGAALK
VRLEESNYELEGK	YNSPYYNTLDDLK	VTPGQNEIDLVTGYQVVK
LTVAQMVDLK	ALEEANNDLENK	LTSVAVENDWLLTVAQMVDLK
AATSGDGVQK	SAATSGNGVQK	WLTAVENDGQLLTVAQMVDLK
VGLSAATSGNGVQK	VDLTVGNK	LPSGSAGGQLGVVNP
VDALMDELNFMK	VQALEEANNDLENK	VTMQNLNDRLASYLK
LNDLEEALQQA	MGLSLTSLK	YEELQVTVGRHGDSLK
VDPELQNVK	PYYNTLDDLK	EHLVLMEDLDTNADK
NKLEGLDALQK	STMQELNSRLASYLK	RVVGLSAPGSGNGVQK
RADLDSLLAEVK	LTTLEVEPSDTLENVK	QNLNDRLASYLK
AALLYDCDAK	ATSGNGVQK	LDVNSLLSLQTAK
LAEESNYELEGK	DAEAWFNEK	RVVGLSAATDGDGVQK
SLDLSLLAEVK	GYLGAVNPK	LQGNGLLALTSYK
PFHSLSEK	LVDLTVGNK	NQFASFLDK
NGMLSLTSLK	ALNSLLDVYHK	RAPLADQGSSHHTVK
TGGYLGAVNPK	SDQSRLDSELK	WTLLQEQQTK
ELLTVAQMVDLK	HDVYSMSFYK	RVVGLSAGPSGNGVQK
DDGGQLGVVNP	TVAQMVDLK	RVVGLSAATSGNGVQK
LEGLDALQK	RVALEEANADLEVK	VRALEESNYELEGK
AEAESLYQSK	VDLLNQELEFLK	AALQGDGMLSLTSLK
LAELEEALQK	NNQFASFLDK	LETTELDNNLEQLSSYK
HDVYSMSFYK	QGLLTVAQMVDLK	VRVGLSAATSGNGVQK
DAEAMAFNEK	VNSLLSLQTAK	AALQGDGLLALTSYK
SLNNQFASFLDK	AALLYDCHAK	GFFDPNTEENLTYLQLK
HDVYSWSFYK	TVGRHGDSLK	LQPSGSANQLGVVNP
RTTAENEFVMLK	-----	-----

Con esta metodología no es probable que se diferencien pares o trios de aminoácidos o aminoácidos que tengan un peso molecular similar. El método seguido (secuenciación De Novo) compara y determina diferencia en pesos moleculares, por lo tanto el análisis en algunos casos puede dejar dos péptidos similares siendo bastante subjetivo; sin embargo la secuenciación es más rápida y menos costosa que la degradación de Edman.

Los péptidos fueron seleccionados por medio de BLASTp y posteriormente se realizó la búsqueda de secuencias de lectinas de leguminosas donde los péptidos encajaran y nos permitiera tener un molde inicial para determinar la secuencia de la LGL-P2 (Altschul et al., 1990). Los parámetros de la búsqueda en BLAST se modificaron para analizar secuencias pequeñas, se utilizó un tamaño de palabra de 2, la matriz PAM30 en lugar de la BLOSUM62 utilizada por defecto en BLAST, un costo de gaps de 9 para existencia y de 1 para extensión, se eliminaron los ajustes por composición, se eliminó el filtro para regiones de baja complejidad y se corrió contra secuencias referenciadas y pertenecientes a organismos de leguminosas (*Leguminosae*). El dato más representativo lo entregó la búsqueda con el péptido más largo y un conjunto de los más largos, donde se encontraron secuencias de aglutininas (o lectinas) dentro de la familia *Leguminosae* (figuras 11 A, B, C) para cada una de las enzimas utilizadas. Se pudo determinar que según el porcentaje de identidad, unido al porcentaje de cobertura y los valores más bajos, que de las secuencias encontradas era una buena decisión utilizar las secuencias relacionadas a lectinas de *Cicer arietinum* (XP_004516638.1, XP_004514830.1) y *Medicago truncatula* (XP_003599063.1, XP_003599808.1) para analizar todos los posibles alineamientos de los péptidos de LGL-P2 contra ambas lectinas. Las dos plantas pertenecen a la familia *Leguminosae*, poseen lectinas de 227 aminoácidos (*Cicer arietinum*) y de 277 aminoácidos (*Medicago truncatula*), y la última presenta especificidad por Glucosa/Manosa [NCBI XP_003599063.1 y XP_004516638.1]. Se alinearon ambas secuencias con los péptidos de LGL-P2 y se encontró que la mayoría de ellos alineaban con las dos lectinas escogidas por lo que si había una gran cobertura de la secuencia, de esta forma podrían llegar a superponerse los tres conjuntos de péptidos y por lo tanto podríamos tener la secuencia parcial de la LGL-P2.

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0 A

Alignments [Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> PREDICTED: putative bark agglutinin LECRPA3-like [Cicer arietinum]	27.8	27.8	92%	1.4	47%	XP_004516638.1
<input type="checkbox"/> Vegetative lectin [Medicago truncatula]	26.5	26.5	33%	3.8	89%	XP_003599063.1
<input type="checkbox"/> Lectin [Medicago truncatula]	26.5	26.5	33%	3.8	89%	XP_003599008.1
<input type="checkbox"/> PREDICTED: vacuolar-processing enzyme-like [Cicer arietinum]	26.5	26.5	55%	3.8	65%	XP_004498922.1
<input type="checkbox"/> PREDICTED: uncharacterized protein LOC101489114 [Cicer arietinum]	26.1	26.1	81%	5.2	43%	XP_004506092.1

B

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> PREDICTED: L-type lectin-domain containing receptor kinase IX.1-like [Cicer arietinum]	28.6	28.6	70%	0.30	59%	XP_004509794.1
<input type="checkbox"/> PREDICTED: L-type lectin-domain containing receptor kinase IX.1-like [Glycine max]	28.6	28.6	70%	0.30	59%	XP_003551295.1
<input type="checkbox"/> Lectin-like protein [Medicago truncatula]	28.2	42.8	70%	0.41	58%	XP_003628862.1
<input type="checkbox"/> Lectin receptor-like kinase Tg-20 [Medicago truncatula]	26.5	26.5	70%	1.6	53%	XP_003628872.1
<input type="checkbox"/> PREDICTED: L-type lectin-domain containing receptor kinase IX.1-like [Glycine max]	26.5	40.3	50%	1.6	80%	XP_003532592.1
<input type="checkbox"/> PREDICTED: putative zinc finger protein At1g68190-like [Cicer arietinum]	25.7	25.7	80%	3.2	63%	XP_004497022.1
<input type="checkbox"/> PREDICTED: uncharacterized protein LOC100776828 [Glycine max]	25.2	25.2	50%	4.4	80%	XP_003517880.1
<input type="checkbox"/> PREDICTED: L-type lectin-domain containing receptor kinase IX.1-like [Glycine max]	25.2	39.0	55%	4.6	73%	XP_003532590.1
<input type="checkbox"/> PREDICTED: E3 ubiquitin-protein ligase At3g02290-like [Glycine max]	24.8	24.8	50%	6.4	60%	XP_003531881.1
<input type="checkbox"/> PREDICTED: CBL-interacting serine/threonine-protein kinase 11-like [Glycine max]	24.8	24.8	80%	6.4	56%	XP_003551605.1
<input type="checkbox"/> PREDICTED: agglutinin-2-like [Glycine max]	24.4	24.4	70%	9.1	53%	XP_003551297.1
<input type="checkbox"/> PREDICTED: protein TWIN LOV 1-like isoform X1 [Cicer arietinum]	24.4	39.5	70%	9.1	73%	XP_004489990.1
<input type="checkbox"/> PREDICTED: probable L-type lectin-domain containing receptor kinase VII.2-like [Glycine max]	24.4	38.2	70%	9.2	75%	XP_003541907.1
<input type="checkbox"/> PREDICTED: probable L-type lectin-domain containing receptor kinase VII.2-like [Cicer arietinum]	24.4	24.4	40%	9.2	75%	XP_004487272.1
<input type="checkbox"/> PREDICTED: putative bark agglutinin LECRPA3-like [Cicer arietinum]	24.0	43.3	65%	13	75%	XP_004514830.1
<input type="checkbox"/> putative bark agglutinin LECRPA3-like precursor [Cicer arietinum]	24.0	39.5	65%	13	75%	NP_001269244.1

C

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Lectin alpha chain [Medicago truncatula]	33.7	33.7	95%	0.006	63%	XP_003628870.1
<input type="checkbox"/> PREDICTED: putative bark agglutinin LECRPA3-like [Cicer arietinum]	32.9	32.9	75%	0.012	67%	XP_004514830.1
<input type="checkbox"/> Lectin alpha chain [Medicago truncatula]	31.2	31.2	95%	0.043	63%	XP_003628871.1
<input type="checkbox"/> Vegetative lectin [Medicago truncatula]	31.2	31.2	80%	0.043	50%	XP_003599063.1
<input type="checkbox"/> Lectin [Medicago truncatula]	29.9	29.9	90%	0.11	56%	XP_003628866.1
<input type="checkbox"/> Peptide transporter [Medicago truncatula]	29.5	29.5	45%	0.16	100%	XP_003613570.1
<input type="checkbox"/> putative bark agglutinin LECRPA3-like precursor [Cicer arietinum]	29.1	29.1	100%	0.21	55%	NP_001269244.1
<input type="checkbox"/> PREDICTED: agglutinin-2-like [Cicer arietinum]	29.1	29.1	60%	0.21	75%	XP_004509711.1
<input type="checkbox"/> PREDICTED: agglutinin-2-like [Glycine max]	28.6	28.6	75%	0.30	67%	XP_003530577.1
<input type="checkbox"/> Lectin [Medicago truncatula]	27.4	27.4	75%	0.82	60%	XP_003599008.1

Figura 11. Resultados identificación de la muestra con base en los péptidos obtenidos utilizando BLAST (blastp) con los péptidos obtenidos con: A: Tripsina B: Quimotripsina. C: Lys C

En la figura 12, se muestran los resultados de las comparaciones con la secuencia de *Cicer arietinum* para los péptidos derivados de la digestión con tripsina; las secuencias de color verde, rojo y azul dentro la secuencia principal (secuencia completa sin espacios, color negro) son los péptidos que se alinearon parcial o totalmente. El mismo procedimiento se llevó a cabo con la secuencia de la lectina de *Medicago truncatula* (datos no presentados).

Se realizó el proceso con los péptidos derivados de quimotripsina y Lys C y luego de ubicar cada péptido sobre cada una de las secuencias, se observó que los péptidos abarcaban una gran parte de las secuencias encontradas, pero que era necesario tener un mejor molde para la superposición de los péptidos obtenidos. Se encontró que el porcentaje de cobertura de la secuencia es bajo.



Figura 12. Resultados alineamiento de péptidos obtenidos con tripsina contra la secuencia de Cicer Arietinum utilizando BLAST (blastp).

Los anteriores resultados nos muestran que a pesar de no contar con una secuencia base de *G. lindenii* se logró relacionar el conjunto de péptidos, obtenidos del tratamiento con cada una de las enzimas, con secuencias de lectinas de leguminosas caracterizadas parcialmente. Aunque los péptidos cubrieron un amplio espectro de las secuencias utilizadas como molde, los bajos porcentajes de identidad no representaron un alto grado de confianza para formar la secuencia de la LGL-P2, por lo tanto empleamos otras alternativas para alcanzar nuestro objetivo.

Dentro del grupo de investigación (GRIP) se ha trabajado en la purificación, caracterización y elucidación de la estructura primaria y terciaria de las lectinas presentes

en el género *Dioclea* (Melgarejo y Pérez, 1997; Rojas, 1998; Pérez, 1998; Garzón, 1999; Pérez y Sierra, 1999; Megarejo et al., 2005). Trabajos previos con *C. ensiformis*, *D. grandiflora*, *D. sericea*, *D. lehmanni* y más recientemente *C. roseum* mostraron que en estas especies hay una segunda lectina análoga a LGL-P2 (Sierra y Pérez, 1999; Almanza et al, 2004; Melgarejo et al., 2005; Dam et al., 2011). Estas lectinas se caracterizan porque reconocen el determinante H-2 y tienen características moleculares particulares como alta similitud en la secuencia N-terminal, en el patrón de inhibición de carbohidratos y tienen estructura dimérica y/o tetramérica entre otras (Tabla 11).

Tabla 11. Propiedades moleculares de algunas lectinas de la Subtribu *Diocleinae* que reconocen lactosa y el determinante H2

Lectina	Proteína Nativa (kDa)	Subunidades (kDa)	PI	Secuencia N-terminal	A ^{1%} _{1cm}	Especificidad	Ka (M ⁻¹)
LGL-P2 ¹	104.256 26.064	25	8.3	AKVTSIKYTSIS-SSXGKPLQGNA	12	H-Tipo 2 ^{7,8}	1x10 ⁵ , Fuc α1-2(Galβ1-4Glc) ⁶
LDL-P2 ²	58.4	29-30	6.5	AGAVYFYFTKFTTSSSSLTLQGSAAEVSNNDHK--TNLKNPTNKVGRALY	24	H-Tipo 2 ^{7,8}	2.4x10 ⁵ , Fuc α1-2(Galβ1-4Glc) ⁶
LDG-P2 ³	58.9	29-30	5.4 5.3 5.1	AGAVYFXFTKFTTSSSSLTLQGSAAEVSNNGILALTNLKNPTNKVGRALY	22	H-Tipo 2 ⁷	
LCE-P2 ⁴		25		+GAVYF+FTKFTTS+S+LTL+Q+AEV+N		H-Tipo 2 ⁷	
LCR-II ⁵	25.652 ¹⁰			SGAVHFSFTKFTTSSSDLTLQGSAAVSSKGLKKNPSKKGKPVVDHSGRALY		H-Tipo 2 ⁹	1.2x10 ⁵ , Fuc α1-2(Galβ1-4GlcNac)

1. Almanza et al, 2004; 2. Pérez, 1998; 3.4. Melgarejo et al, 2005; 5. Rocha et al, 2010; 6. Vega et al, resultados no publicados; 7. Determinación por ensayos de ELLSA; 8. Determinación por cromatografía frontal de afinidad; 9. Determinación por microcalorimetría (ITC); 10. Rocha et al, 2010.

Estas lectinas son en general inhibidas por galactosyl derivados, siendo el mejor inhibidor GalNAc y me(α,β)-galactopiranosido; también se pueden purificar por cromatografía de afinidad sobre Sepharosa 4B-Lactosa, eluyendo con lactosa 200 mM. La secuencia SSSS está presente en todas las lectina tipo P2 y ha sido propuesta como un rasgo característico de estas lectinas; por lo tanto se pueden hacer alineamientos (Melgarejo et al., 2005). En la región N-terminal entre las lectinas LGL-P2 y LDL-P2 hay un 52 % de identidad y un 66 % de similitud. Las lectinas LGL-P2 y LCR-II tienen un 30 % de identidad y entre las lectinas LDL-P2 y LCR-II un 70% de identidad. En PDB hay dos estructuras de lectinas del género *Erythrina* que unen lactosa, ECL (PDB 1GZC) y EcorL (PDB 1AX1); otras como la lectina de *Arachis hypogaea* (PNA) (PDB 1CR7) y la de *Cymbosema roseum* (LCR-II) (UniProtKB/Swiss-Prot, P86795), esta última es la única reportada a la fecha de las lectinas tipo P2, que están en la subtribu *Diocleinae*.

Para la LCR-II inicialmente se determinaron 189 aminoácidos y un cubrimiento del 80% de la secuencia con una masa molecular de 20.528 Da, y una masa no secuenciada de

5.000 Da (Rocha et al., 2009). Sin embargo en un trabajo posterior se secuenciaron en total 239 aminoácidos con una masa molecular de 25,701 Da (Rocha et al., 2010, última modificación enero 2011). En el grupo de investigación se ha estado trabajando en la determinación de la estructura primaria de lectina de *Dioclea lehmanni* P2 (LDL-P2); para ello se prepararon hidrolizados de la proteína con tripsina, quimotripsina y proteasa V8, los péptidos fueron purificados y su secuencia se determinó por degradación de Edman (Datos no publicados). Posteriormente por superposición manual se determinaron en total 227 aminoácidos con un cubrimiento del 81% de la secuencia (Figura 13, Datos no publicados), el análisis de aminoácidos arrojó un total de 279 residuos de aminoácidos (Pérez, 1998). Como las dos lectinas, LGL-P2 y LDL-P2, tienen características bioquímicas similares y un mayor % identidad en el N-terminal fue seleccionada como punto de partida para comparar su secuencia con la secuencia de los péptidos obtenidos de la LGL-P2.

10	20	30	40	50	60
AGAVYFKFTK	FTTSSSSLTL	QGSAEVSNDL	KLTLNKNPTN	KVGRALYSTP	ITIWDETTGK
70	80	90	100	110	120
VASFVTSITL	RMEDVSSFVP	ANGDVNFFIA	PDASLPKDSI	-GGQLGVVNP	NYFNFTNFVG
130	140		160	170	180
VEFDDYSNSW	DPSYS-HIGI	DVN	FSKGGL	VQAAALYNSV	LSVAVENDGR
190	200	210	220		
TITVAQVVDL	KDVLPSKVRV	GLSAATSGNG	VQKHDLYSWF	FMSNFDAT	

Figura 13. Secuencia parcial de la lectina P2 de *Dioclea lehmanni* (LDL-P2). Vega et al., Datos no publicados.

4.5 Análisis Bioinformático de las secuencias de péptidos encontrados contra la secuencia de la LDL-P2.

Inicialmente se realizó una comparación de la secuencia parcial de LDL-P2 contra todas las proteínas de leguminosas y se encontró que estaba relacionada con otras lectinas de leguminosas. En la secuencia se encontraron dominios conservados de unión a metales

(Ca⁺⁺ y Mn⁺⁺), de interacción con carbohidratos y sitios de N-glicosilación entre otros (Figura 14). Por BLAST se encontró un porcentaje de identidad mayor al 51% con respecto a *Cicer arietinum* y un E value muy bajo ($3e^{-21}$); esta lectina es del género *Cicer*, tribu *Cicerae*, subfamilia *Papilionoideae* de la familia de las leguminosas mientras que con la LCR-II el porcentaje de identidad fue del 40 % y un E value ($1 e^{-34}$) (Anexo 13), recordemos que esta última especie pertenece a la misma subtribu y por lo tanto son más cercanas. De acuerdo con los resultados anteriores y teniendo en cuenta la cercanía filogenética, tomamos como base la secuencia parcial de la LDL-P2 para el ensamblaje de los péptidos de LGL-P2.

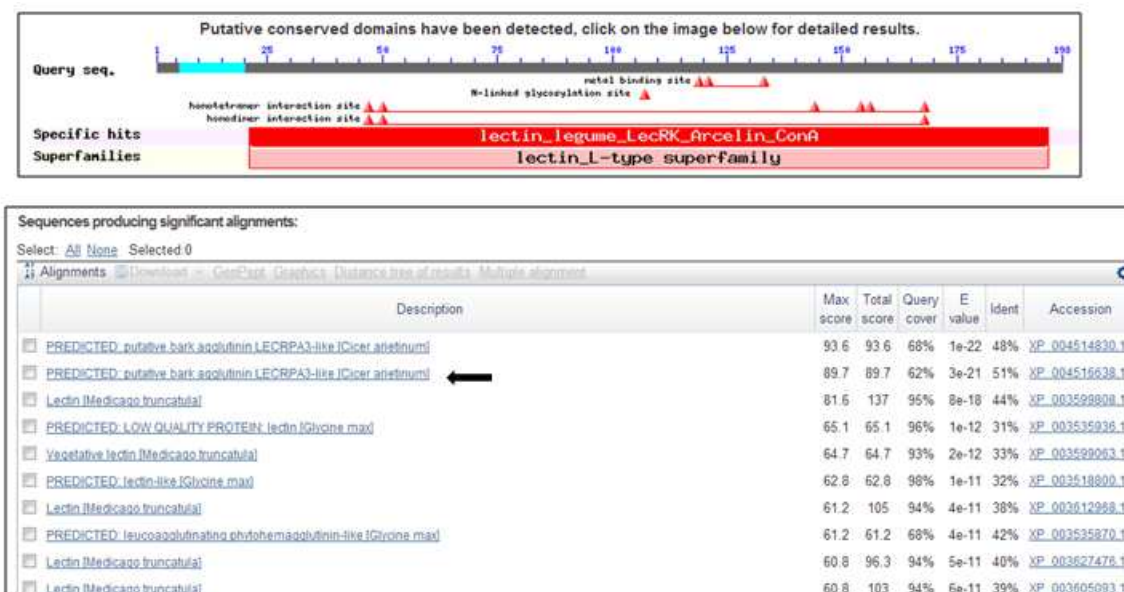


Figura 14. Dominios encontrados en la comparación de la secuencia parcial de la LGL-P2 utilizando BLAST.

4.5.1 Superposición de péptidos para la obtención de la secuencia LGL-P2

Primero se realizaron los alineamientos de los péptidos con la secuencia de la LDL-P2 en BLAST con las modificaciones para secuencias pequeñas. Los tres set de péptidos de LGL-P2 fueron alineados sobre la mayor parte de la extensión de la secuencia LDL-P2 (Figura 15); sin embargo la mayor cobertura de la secuencia y los mejores puntajes de alineamiento se lograron con los péptidos obtenidos con quimotripsina (Figura 15B).

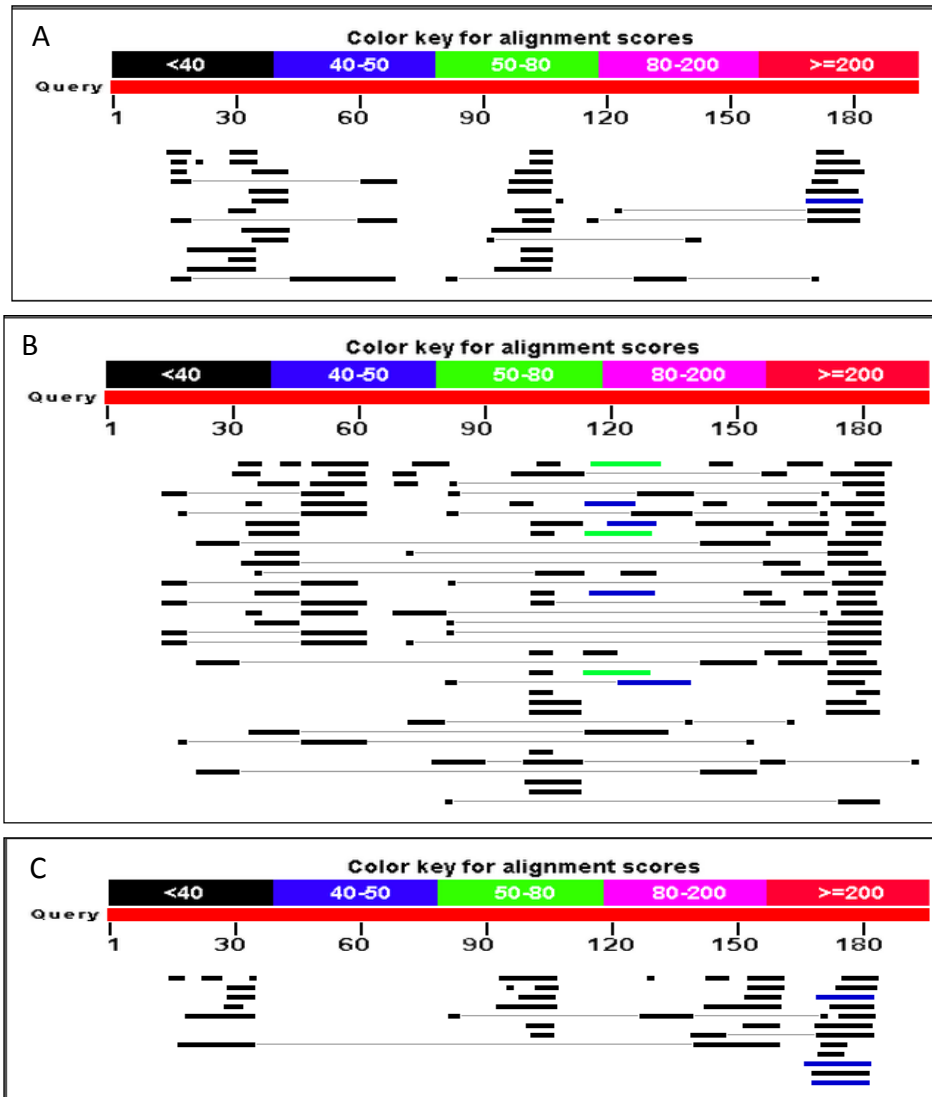


Figura 15. Resultados comparación de la secuencia parcial de la LDL-P2 con los péptidos A: trípticos B: Qimotripticos C: Lys C obtenidos de LGL-P2 utilizando BLAST.

En cada una de las figuras, los colores representan los puntajes de cada alineamiento (hay 5 colores diferentes donde del color negro al rojo aumentan los puntajes) y la línea roja (Query) muestra la longitud de la secuencia que se está analizando. En este caso se tomó toda la secuencia que se tenía de la LDL-P2 (algunas modificaciones se realizaron al momento de analizar cada alineamiento, teniendo en cuenta los cortes de datos en la secuencia base). Los “hits” encontrados por Blast aparecen como líneas horizontales a lo largo de la secuencia Query (en este caso al tratarse de péptidos, se esperaba que se alinearan pequeñas regiones del Query, pero que en conjunto cubrieran la mayor

cantidad como sucedió en el caso de la quimotripsina. Se observó que a pesar de eliminar la redundancia en los péptidos hay regiones que son cubiertas con un alto número de péptidos; los cuales se caracterizaban por tener una o dos modificaciones dentro de la secuencia o tenían uno o dos aminoácidos más con respecto a los otros, es importante resaltar que los mejores alineamientos se produjeron con quimotripsina.

En segundo lugar se hizo la superposición de todos los péptidos derivados LGL-P2 y con ayuda del programa Excel® se hizo el empalme de los péptidos que se complementaban realizando las búsquedas de las secuencias iniciales y finales de cada péptido contra los demás. Dentro de los péptidos evaluados no se encontró la secuencia N-terminal, probablemente este péptido se pudo perder porque estaba glicosilado y la metodología establecida lleva a que durante el análisis de los datos se descarte este péptido. En la figura 16 se presenta la secuencia parcial de la LGL-P2 donde las regiones señaladas en rojo y separadas por “/”, se refieren a los posibles cambios en secuencia que se encontraron en varios de los péptidos obtenidos; “nnn” representan los lugares donde no se pudo determinar la secuencia con respecto a la de LDL-P2.

En la tabla 12 se presenta la secuencia de los péptidos derivados de LGL-P2 que están en la secuencia propuesta (Figuras 16 y 17), indicando de qué enzima provienen, las modificaciones de aminoácidos según cada alineamiento, los valores E y porcentajes de identidad encontrados con BLAST. En la figura 17 se señala en verde la secuencia del N-terminal de la lectina determinada por Almanza et al, 2004.

```
SASTSL/SSPVRNTGSVA/SAEVENDW/GELTVAQMVDL/LTDLKNPSSN/YQLGRVLY/HSSPVTLWDEAT/CCF
AHGNVASF VSSLTLnnnn
RL/FEDVSGYVPAD/NGALGVVNAPD/PKDAGTLLALPSGST/AGGSAGGQLG/QGLV/A/TVNPKDAFNNFVGVE
FDDYSN/DSWDPS/GYSFL/IGLDVNSLLSLQTAKnnnnnn
SVAVENDGQ/ELLTVAQMVDKLT/AVLPSKVRRLF/VGLSA/SAT/SS/GGD/NGVQK/LHDVY
```

Figura 16. Secuencia parcial con variaciones de la LGL-P2.

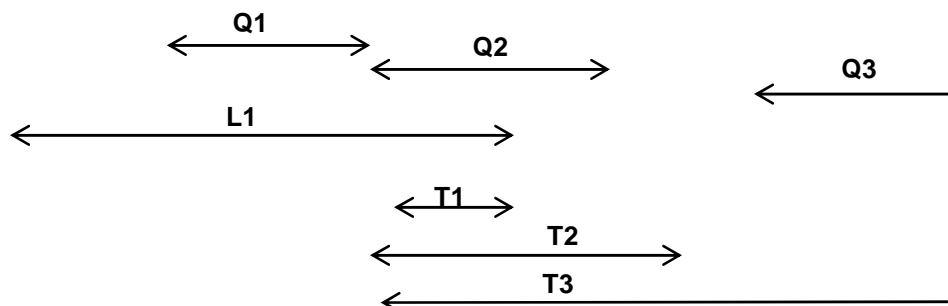
Tabla 12. Secuencia de los péptidos seleccionados para la determinación de estructura primaria de LGL-P2.

Secuencia Péptido	Enzima	Valor E (alineamiento)	% Identidad
VSAVENDGELLTVAQM	Q1	0,53	62%
VDLKNPSSNLGRVLY	Q2	2,00E-06	57%
HSSPVTLWDEATGNVASF	Q3	4,00E-07	65%
RLEDVSGYVPADGLVF	Q4	8,00E-09	77%
ANGALGVVNPKDAFNNF	Q5	1,10	47%
GAPDAGTLLALPSGSTGGY	Q6	0,002	53%
SPGSAGGQLGVVNPKDAFNNF	Q7	1,00E-07	70%
VGVEFDDYSNSWDPSYSF	Q8	4,00E-17	100%
SVAVENDGQLLTVAQMVDKL	Q9	4,00E-09	74%
VDLKTVLPSKVRRLF	Q10	1,00E-07	83%
SAATSGDGVQKHDVY	Q11	3,00E-11	87%
LTSVAVENDWLLTVAQMVDLK	L1	2,00E-09	63%
PSYPHLGLDVNSLLSLQTAK	L2	2,00E-06	69%
VRVGLSAATSGNGVQK	L3	2,00E-13	100%
VASFVSSLTLR	T1	2,00E-06	80%
LQGNGLSLTSLK	T2	0,005	50%
TSLKNPSYQLGR	T3	2,00E-04	58%
ALHSSPVTLCFAHGNVASFVSSLTLR	T4	7,00E-06	38%
PSYPHLGLNVNSLLSLQTAK	T5	2,200	67%
VGLSAATSGNGVQK	T6	2,00E-11	100%
HSSPVTLCVDEATGNASF	Q12	5,00E-04	53%
SNSWDPSYKTLGMGLCL	Q13	3,00E-06	88%
AALLYDCHAKTL	Q14	0,093	67%

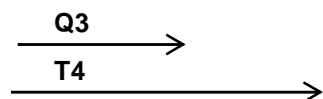
Q: hidrólisis obtenidos luego del tratamiento con quimotripsina T: hidrólisis obtenidos luego del tratamiento con tripsina L: hidrólisis obtenidos luego del tratamiento con Lysil C

Figura 17. Superposición de péptidos ensamblados para determinar secuencia LGL-P2.

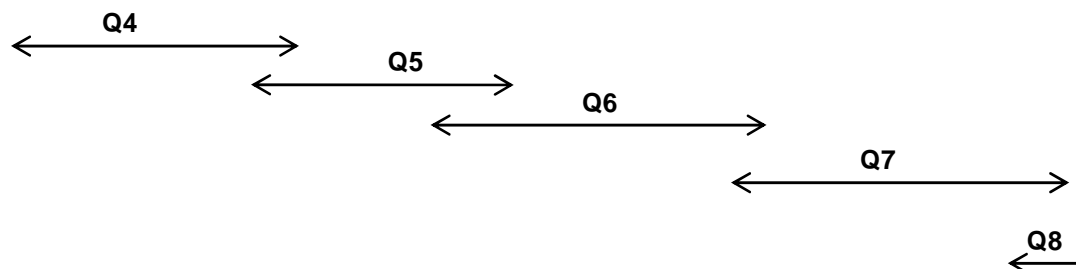
AKVTSIKYTSISSSXGKPLQGNAGELTVAQMVDL/LTDLKNPSSN/YQLGRV/ALY/HSSPVTLWDEA



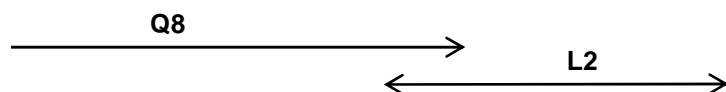
T/CCFAHGNVASFVSSLTLnnnnnn



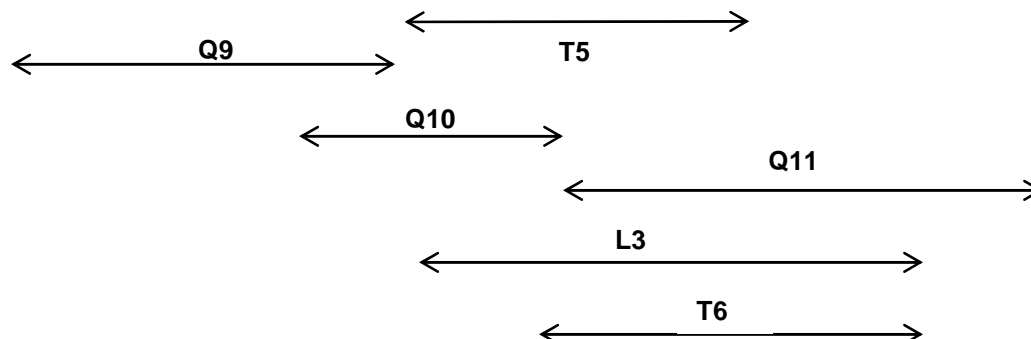
RL/FEDVSGYVPAD/NGALGVVNAPD/PKDAGTLLALPSGST/AGGSAGGQLG/QGLV/A/TVNPK



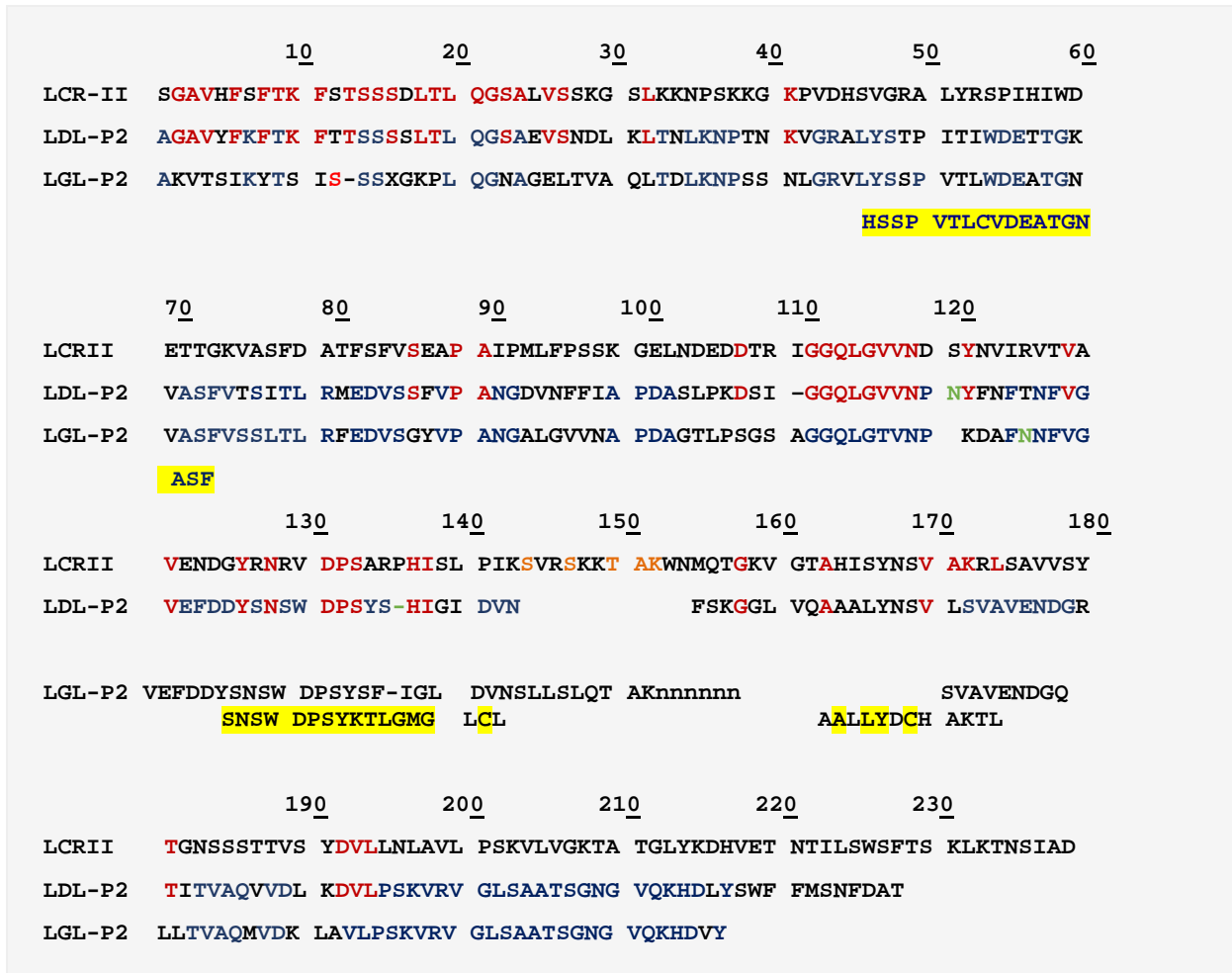
DAFNNFVGVEFDDYSN/DSWDPS/GYSFL/IGLDVNSLLSLQTAKnnnnnn



SVAVENDGQ/ELLTVAQMVDKLT/AVLPSKVRRLF/VGLSA/SAT/SS/GGD/NGVQK/LHDVYnnn



Si alineamos las secuencias de las tres lectinas tipo P2 podemos proponer la siguiente secuencia para la LGL-P2 (Figura 18):



Color Rojo: Residuos de aminoácidos idénticos entre LDL-P2 y LCR-II ; Color Azul: Residuos de aminoácidos idénticos entre LDL-P2 y LGL-P2; Color anaranjado: Algunos residuos idénticos entre LGL-P2 y LCR-II ; Color amarillo : Otras secuencias que se pueden superponer que contienen residuos de cisteína.

Figura 18. Comparación de las tres secuencias de lectinas P2 de la subtribu *Diocleinae*.

A partir de los análisis realizados, se calculó un porcentaje de identidad de 62% entre las lectinas LDL-P2 y LGL-P2 (sin tener en cuenta los péptidos marcados en amarillo,

posibles ubicaciones de cisteínas), encontrando que con una cobertura del 90%, se tenía un valor E de $6,00E-78$ y un porcentaje de similitud de 73%. Valores mucho mayores a los encontrados entre LDL-P2 y LCR-II.

Por otra parte, se consideró la posibilidad de tener dentro de la secuencia los péptidos que permiten tener las dos cisteínas encontradas en los ensayos de Almanza (2004). Por lo que se ajustó la secuencia adicionando los péptidos encontrados con menor valor E (Q12 y Q13) y el porcentaje de identidad entre las dos secuencias disminuye al 59%, con un mismo porcentaje de cobertura de la secuencia y un valor E mayor ($3,00E-70$), disminuyendo también el porcentaje de similitud hasta 71%.

En la tabla 13, se presenta la comparación del número de aminoácidos reportado en el análisis de aminoácidos y los obtenidos a partir de la secuencia. En el caso de LGL-P2, se observa un número menor de residuos reportados en el análisis con respecto a los determinados por la secuencia parcial; si tenemos en cuenta que la determinación más exacta del peso molecular del monómero fue la realizada por espectrometría de masas (MS-ESI) podríamos pensar que durante el análisis de aminoácidos algunos residuos no se determinaron. Además en la secuencia faltarían por determinar aproximadamente 25 aminoácidos, para tener un total de 220 aminoácidos y un 7% de carbohidratos que nos darían un peso molecular aproximado de 26000 Da. Por otra parte dentro de la secuencia y el análisis no se reportaron las dos cisteínas, pero durante el análisis se obtuvieron péptidos con cisteína (Figura 18, resaltados en color amarillo) que se podrían superponer en la secuencia propuesta. Con estos resultados para la LGL-P2 se determinaron 195 aminoácidos y un cubrimiento del 89 % de la secuencia con una masa molecular de 20.213 Da, y una masa no secuenciada de 2750 Da. Por otra parte en el caso de la LDL-P2 faltarían por secuenciar aproximadamente 50 aminoácidos; en este caso el peso molecular determinado por SDS-PAGE es muy inexacto y como la lectina es una glicoproteína se presenta una menor movilidad electroforética. Si se tiene en cuenta el análisis de aminoácidos, el peso molecular del monómero estaría entre 29.000- 30.000 Da.

En general el número de residuos de cada aminoácido no concuerda con el de la secuencia, en el caso de la LGL-P2, la determinación de la secuencia de los péptidos se hizo por masas (diferencia de peso molecular), por lo que es menos exacta y más

subjetiva que la obtención de secuencia por el método Edman, donde se identifica aminoácido por aminoácido.

Tabla 13. Comparación del análisis de aminoácidos para LGL-P2 Y LDL-P2.

Aminoácido analizado	LGL-P2	LGL-P2^a	LDL-P2	LDL-P2^b
Asx (N-D)	19	25	38	31
Glx (Q-E)	6	12	21	10
Ser (S)	21	2	32	25
Gly (G)	14	19	22	16
His (H)	5	2	1	2
Arg (R)	7	3	8	4
Thr (T)	16	12	21	17
Ala (A)	8	17	24	16
Pro (P)	5	9	12	8
Tyr (Y)	2	5	9	7
Val (V)	13	23	25	24
Met (M)	10	1	3	2
Cys (C)	2	0	0	0
Ile (I)	20	3	16	8
Leu (L)	12	21	21	16
Phe (F)	8	6	9	15
Lys (K)	11	9	13	11
Trp (W)	7	2	4	3
Número Total de aminoácidos	186	195	279	227
Peso molecular (Da)	21 116¹	20 213,72⁵	29 970¹	23 153,86⁵
Peso molecular (Da)	24 240²	22 037,48⁶	31 170²	24 383,00⁶
Peso Molecular (Da)	26 064³		29 000-30 000⁴	
Aminoácidos por determinar	-----	~25		~50

a, b datos obtenidos a partir de la secuencia.

1. Calculado del análisis de aminoácidos no presenta carbohidratos (Pérez G, 1998; Almanza et al., 2005).

2. Calculado a partir del análisis de aminoácidos teniendo en cuenta el porcentaje de Carbohidratos, para LGL-P2 7 %. (Almanza et al., 2005). Para LDL-P2: 4.1%
3. Determinado por Espectrometría de Masas (MS-ESI) y con este peso se hizo el cálculo de los residuos en el monómero (Almanza et al., 2005)
4. Determinado por electroforesis (Pérez G, 1998)
5. Los cálculos se hicieron con el servidor: http://web.expasy.org/compute_pi/, a partir de la secuencia.
6. Peso molecular calculado a partir de la secuencia teniendo en cuenta el contenido de carbohidratos para las dos lectinas. Se tomó como valor real el Peso molecular para LGL-P2: 26064 Da y para LDL-P2: 29000 Da.

Aunque se identificaron secuencias idénticas de péptidos en más de un digerido y se encontraron datos de secuencias de proteínas dentro de la misma familia, en este caso *Leguminosae* y la misma subtribu (*Diocleinae*), solo se pudo hacer una aproximación de la secuencia, por lo tanto se requieren más experimentos con otras enzimas y/o una caracterización más minuciosa de los péptidos y si es posible secuenciación por Edman.

Otras fuentes de error se evidencian en la asignación de Glutámico y Aspártico en vez de glutamina y Asparagina por deaminación en medios ácidos durante los procesos de purificación y secuenciación. Para corroborar la secuencia se podrían aislar y secuenciar el gen que codifica para la lectina pero aunque es rutina trabajar con ácidos nucleicos, la secuenciación directa de la proteína es importante por muchas razones. Por ejemplo los puentes disulfuro solamente pueden localizarse por secuenciado de la proteína. Muchas proteínas son modificadas después de su biosíntesis por eliminación o modificación de residuos de aminoácidos, que son esenciales para su función biológica. Estas modificaciones solo pueden identificarse a partir de la secuenciación directa de la proteína.

Por otra parte resulta difícil, con frecuencia aislar e identificar un ácido nucleico que codifique la proteína que interesa. Se hace el diseño de muchos pares de primers basados en secuencias de otras proteínas de la misma familia, de otras proteínas que tengan funciones similares o teniendo en cuenta otros parámetros y por lo general es difícil “pescar” la proteína de interés. Un error corriente en la secuenciación del ADN es la inserción o eliminación de un nucleótido, lo que cambia la pauta de lectura del gen y cambia también las predicciones de todos los residuos de aminoácidos que se hallan

después del punto donde se cometió el error. Esto se puede comprobar frente a la secuencia directa de los péptidos derivados de las hidrólisis de la proteína. Finalmente en este caso el trabajo directo con la proteína es más robusto, confiable y exacto.

4.5.2 Acercamiento a la estructura terciaria por modelamiento de las lectinas P2 de la Lectina de *Galactia lindenii* y *Dioclea lehmanni*

Con las secuencias parciales obtenidas para las lectinas de *Galactia lindenii* (LGL-P2) y *Dioclea lehmanni* (LDL-P2) por medio de análisis bioinformático se hizo una aproximación a la estructura terciaria del monómero por modelado. Este método está basado en la generación de estructuras a partir de otras estructuras de referencia que han sido determinadas experimentalmente.

Las secuencias evaluadas fueron las anteriormente propuestas (Figura 19, 20)

10	20	30	40	50	60
AGAVYFKFTK	FTTSSSSLTL	QGSAEVSNL	KLTNLKNPTN	KVGRALYSTP	ITIWDETTGK
70	80	90	100	110	120
VASFVTSITL	RMEDVSSFVP	ANGDVNFFIA	PDASLPKDSI	-GGQLGVVNP	NYFNFTNFVG
130	140		160	170	180
VEFDDYSNSW	DPSYS-HIGI	DVN	FSKGGI	VQAAALNSV	LSVAVENDGR
190	200	210	220		
TITVAQVVDL	KDVLPSKVRV	GLSAATSGNG	VQKHDLYSWF	FMSNFDAT	

Figura 19. Secuencia parcial de la lectina P2 de *Dioclea lehmanni* (LDL-P2).

<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>	<u>60</u>
AKVTSIKYTS	IS-SSXGKPL	QGNAGELTVA	QLTDLKNPSS	NLGRVLYSSP	VTLWDEATGN
<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>	<u>110</u>	<u>120</u>
VASFVSSLTL	RFEDVSGYVP	ANGALGVVNA	PDAGTLPSGS	AGGQLGTVNP	KDAFNNFVG
<u>130</u>	<u>140</u>	<u>150</u>	<u>160</u>	<u>170</u>	<u>180</u>
VEFDDYSNSW	DPSYSF-IGL	DVNSLLSLQT	AK		SVAVENDGQ
<u>190</u>	<u>200</u>	<u>210</u>			
LLTVAQMVDK	LAVLPSKVRV	GLSAATSGNG	VQKHDVY		

Figura 20. Secuencia parcial de la LGL-P2.

Dado que el programa I-TASSER server, recibe como entrada la secuencia en formato FASTA, y hace una búsqueda en bases de datos que contengan información estructural de proteínas, efectivamente encontró que la estructura reportada cuya secuencia tiene mayor similitud con LGL-P2 y la LDL-P2 fue la lectina de mani (Penaut lectin) con código PDB 1CIW.

Para la LDL-P2, el modelo de estructura terciaria obtenido por I-TASSER con un C-score=1.01 se muestra en la Figura 21. Ese valor de C-score indica que es un modelo adecuado ya que está en el rango [-5,2]. Un C-score de mayor valor significa que es un modelo con un alto grado de confianza; para la LGL-P2 se obtuvo C-score=0.60 (Figura 22).

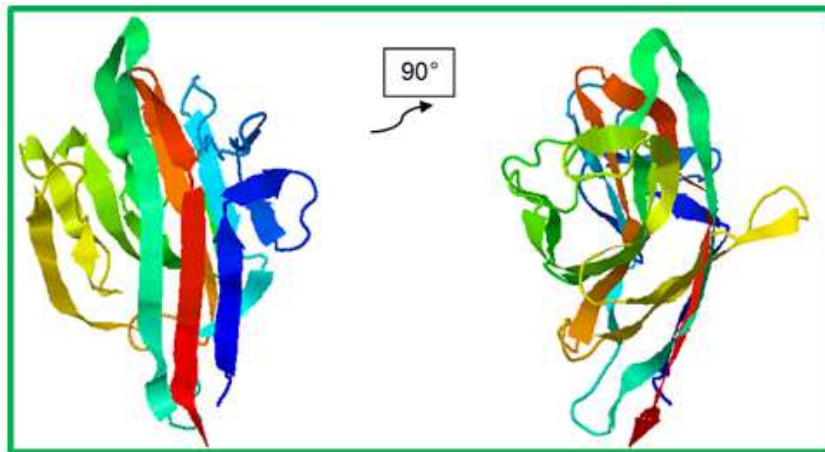


Figura 21. LDL-P2: 0.85 ± 0.08 (TM-score): $3.3 \pm 2.3 \text{ \AA}$ (RMSD)

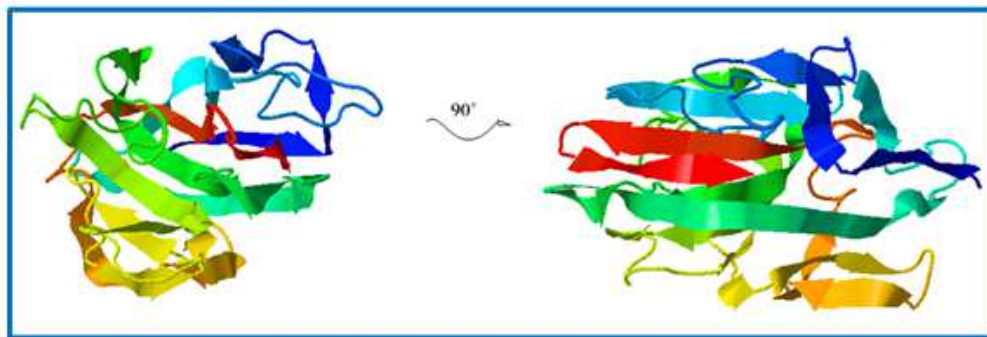


Figura 22. LGL-P2: 0.79 ± 0.09 (TM-score): $4.1 \pm 2.7 \text{ \AA}$ (RMSD)

TM-score y RMSD son parámetros que evalúan la precisión de la estructura de modelado cuando se conoce la estructura nativa del molde; un TM score >0.5 indica un modelo con una topología correcta, en este caso arrojaron TM 0.85 para LDL-P2 y 0.79 para LGL-P2.

Para evaluar la calidad del modelo, se generó el diagrama de Ramachandran donde se analizaron las mejores configuraciones o topologías estructurales (ϕ y ψ). Para LDL-P2, se observa que la mayoría de aminoácidos están en zonas permitidas y que son mayoritarias en las hojas beta (color rojo, Figuras 23, 24; anexos 14,15). La mayoría de regiones más favorables según el diagrama es del 80.5%, teniendo presente que el programa PROCHECK establece que basado en un análisis de 118 estructuras con resolución al menos de 2.0 \AA y un Factor R no más grande que 20, un modelo con buena calidad debe tener 90% de regiones favorables. En el caso de LGL-P2 se observa algo similar pero la mayoría de las regiones más favorables son un 78 %.

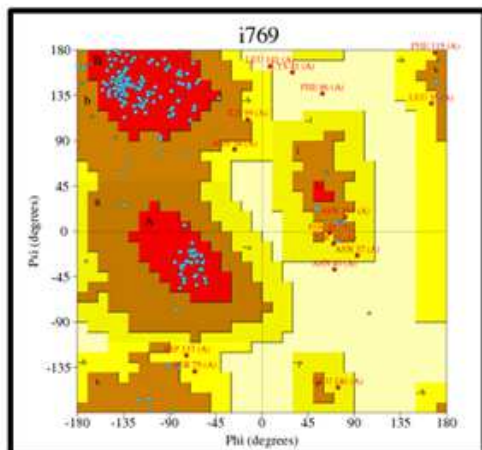


Figura 23. Diagrama de Ramachandran: LDL-P2

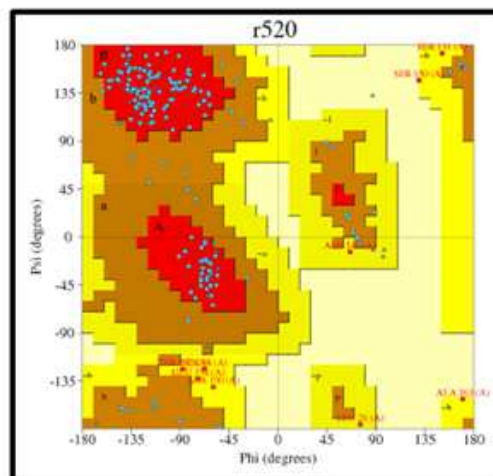


Figura 24. Diagrama de Ramachandran: LGL-P2

Para las dos lectinas la predicción de estructura secundaria muestra un alto contenido de zonas irregulares y hojas plegadas beta; muy poca proporción de hélices alfa (Figura 25, Anexo 16).

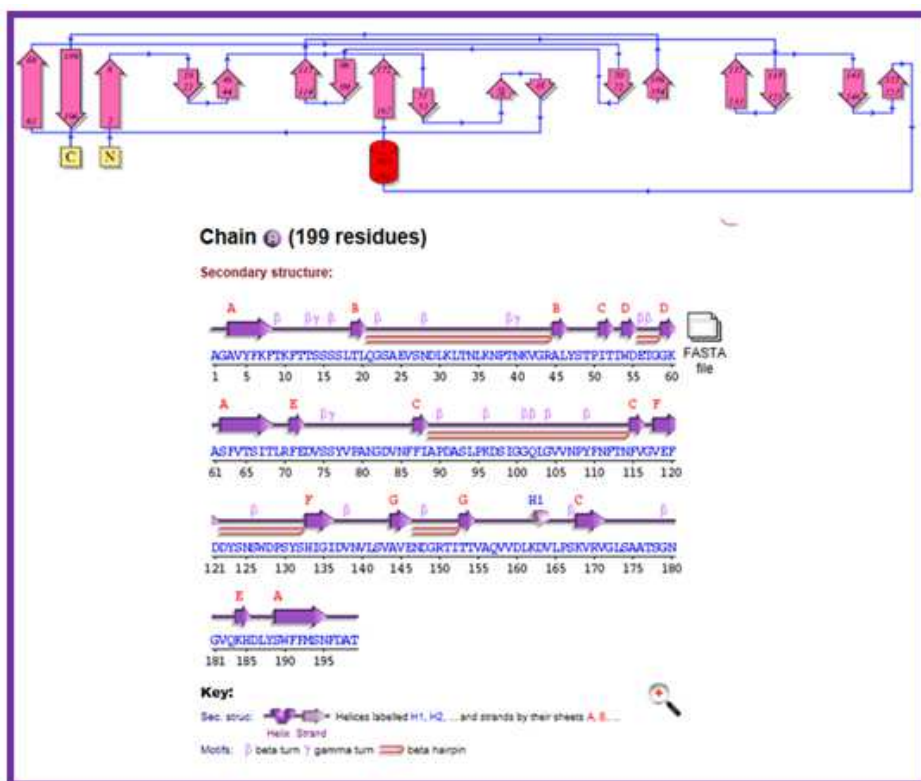


Figura 25. Topología estructural para el modelo estructural de LDL-P2.

Algo característico de las lectinas de leguminosas es la presencia de dominios de hojas beta antiparalelas como se aprecia en este caso; sin embargo es importante aclarar que los modelos fueron diseñados con una secuencia parcial. Estos dominios se llaman “tipo leguminosa” (Sharon y Lis, 2001). Otra característica de estas lectinas es que se asocian en forma dimérica o tetramérica y son hetero u homo oligómeros; su estructura tridimensional es extremadamente similar y su arquitectura es llamada β -sandwich. En dicha estructura hay tres hojas β , la primera formada por seis hebras y está dispuesta como un piso, la segunda hoja con siete hebras está en frente de la primera y la tercera es una hoja de 5 hebras que mantiene unida la primera y segunda (figura 26).

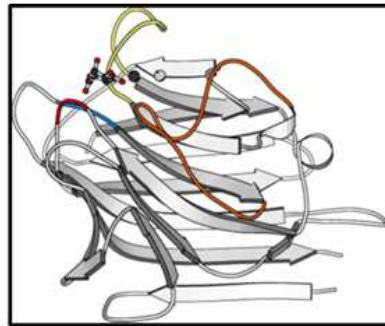
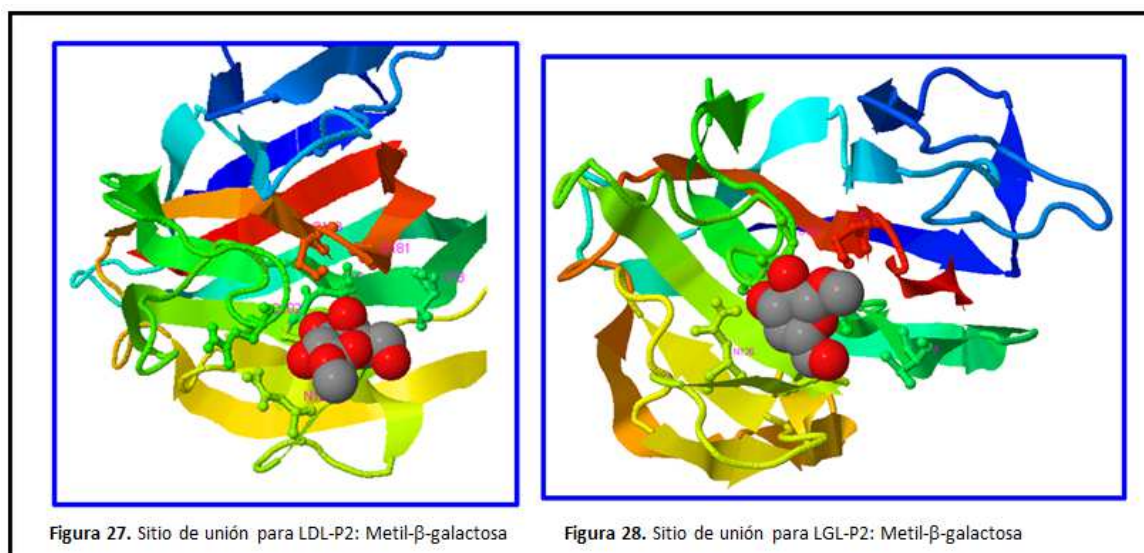


Figura 26. Monómero típico de una lectina de leguminosa (lectina de lenteja, PDB 1LES, Casset *et al.*, 1995). El sitio de unión del carbohidrato está en la parte de arriba en una cavidad. Los cuatro puntos de interacción son : Ala-Asp cis peptide en azul, loop de unión a metal en amarillo, naranja, loop en rojo especificidad del carbohidrato (Tomado, Structural studies of four members of the legume lectin family:Phytohemagglutinin-L, Arcelin-5, *Dolichos biflorus* seed lectin and *Dolichos biflorus* stem and leaf lectin. Thomas W. Hamelryck, 1999, Laboratorium voor Ultrastructuur, Vrije Universiteit Brussel (VUB).

En estas estructuras los iones metálicos forman parte esencial de la estructura nativa y son importantes en la actividad biológica de las proteínas. Los aminoácidos que intervienen en la unión de los iones metálicos son conservados en las lectinas de leguminosas que han sido comparadas (Pusztai, 1999) y se encuentran en la región N terminal de la lectina.

I-TASSER Server proporcionó un modelo de LDL-P2 y LGL-P2 con un sitio de unión a metil-beta-galactosa (Figuras 27 y 28). Para la LDL-P2 los aminoácidos que interactúan con el carbohidrato son la S178, V78, N125, N81, Y123, por medio de interacciones hidrofóbicas y de puentes de hidrógeno, resultados similares fueron obtenidos para la

LGL-P2, donde los aminoácidos están casi en las mismas posiciones (V78, N81, N126, Y124, S185, G101, G102). En general para las lectinas de leguminosas, no importa su especificidad, unen monosacáridos a través de tres cadenas laterales de una constelación de tres aminoácidos invariantes en el sitio de combinación: ácido aspártico (D), asparagina (N), y un aminoácido aromático; este último puede ser reemplazado por leucina, y en otros casos como en la lectina de *Maackia amaurensis* (MAL) la asparagina es reemplazada por ácido aspártico. El cuarto aminoácido invariante es la glicina que participa en la interacción vía puentes de hidrógeno con la amida del enlace peptídico y puede ser reemplazado por arginina o lisina. El anillo del carbohidrato interactúa con el aminoácido aromático; este es un rasgo común de los complejos con carbohidratos y dicha interacción es esencial en el reconocimiento.



Para Concanavalina A, se determinó que varios aminoácidos están implicados en la unión del azúcar y aunque ellos se encuentran en posiciones muy diferentes en la estructura primaria de la subunidad, espacialmente están estrechamente localizados en la conformación nativa de la molécula de lectina. Los residuos de aminoácidos involucrados son: Y-12, N-14, D-16, L-99, Y-100, S-168, D-208 y R-228 (Pusztai, A.1991). Se ha demostrado que las interacciones hidrofóbicas juegan un papel muy importante en el mantenimiento de la estructura nativa de la mayoría de la lectinas estudiadas. En Con A hay una cavidad hidrofóbica donde las cadenas laterales de leucina (L 81), valina (V 89) y tres residuos de fenilalanina (F 11, 191 y 212) constituyen la parte principal de dicha cavidad. Cuando los aminoácidos que contribuyen a la cavidad

hidrofóbica en Con A, no son conservados en otras lectinas, ellos siempre son reemplazados por otros aminoácidos hidrofóbicos. Para las LGL-P2 y LDL-P2 se obtuvieron modelos características de las lectinas de leguminosas (con bueno puntajes de calificación para cada modelo) con las secuencias parciales propuestas.

4.6 Estudios exploratorios de cristalización de la LGL-P2

4.6.1 Preparación de la muestra para ensayos de cristalización

La LGL-P2 se solubilizó por completo en glicerol al 4% alcanzando una concentración final de 14.44 mg/mL; otras lectinas similares también presentaron valores altos de concentración, LCR-II 12 mg/ml (Rocha et al., 2009) lo que es una ventaja para cristalizar proteínas. Para evitar precipitación del stock de LGL-P2 por congelamiento y descongelamiento se almacenó en fracciones de 50 μ L.

Luego de dos meses en las gotas se observó que no hubo crecimiento de cristales pero sí la presencia de precipitados y agregados de proteína (Figura 29 A,B,C).

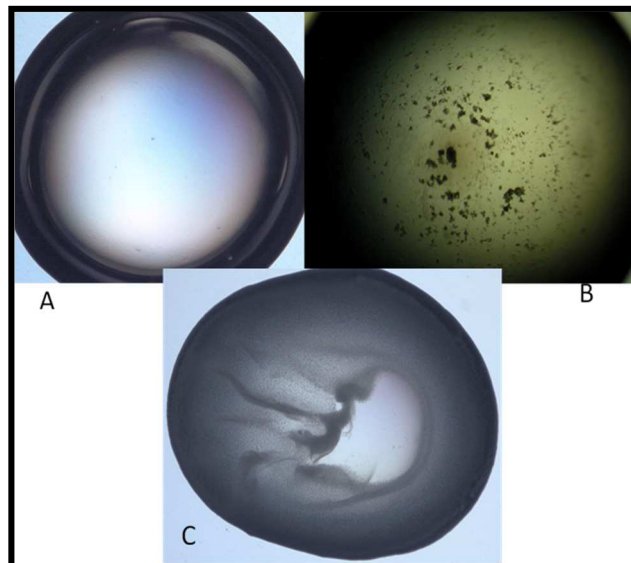


Figura 29. Ensayos preliminares- Gota de cristalización
A. Negativa. B. Aglutinados de proteína. C. Precipitación de proteína.

Inicialmente se trabajaron condiciones descritas para la cristalización de lectinas de *Diocleae* tipo P4, sin embargo cada proteína tiene sus propias condiciones de cristalización y hay muchas condiciones que pueden ser analizadas. En este caso los ensayos iniciales solo se realizaron empleando PEG 400 como agente precipitante y dos condiciones de pH. Muchos autores reportan ensayos con este mismo agente precipitante de peso molecular 6000 y 8000; por ejemplo las lectinas LCR-II y LER fueron cristalizadas empleando PEG 8000 (Rojas y Pérez, 1990; Calvete et al, 1999; Rocha et al., 2009; De Latorre et al, 2011) y la de *Cratylia mollis* con PEG 6000 (Tavares et al, 1996). Condiciones similares fueron empleadas para lectinas de leguminosas específicas por galactosa y/o lactosa (Banerjee et al, 1996). Por otra parte Rojas A (1998) cristalizó la LDL-P4 y determinó su estructura terciaria por difracción de rayos X, luego de varios ensayos obtuvo cristales que crecieron en tres semanas con una resolución de 2.8 Å. En este trabajo se estandarizaron condiciones como el peso molecular y la concentración del agente precipitante, el pH y la molaridad del buffer, y la concentración de proteína; sin embargo se hace énfasis en que el pH y la molaridad del buffer determinan la cristalización de la proteína. Por otra parte lo que se busca es que crezcan cristales de mayor tamaño en un menor número en una sola gota, también De Latorre et al (2011) sugieren que los procesos de cristalización deben ser optimizados en presencia del ligando debido a la estabilidad que se genera en el sitio de unión al carbohidrato.

En el caso de los ensayos con el sistema Crystal Screen I – II con $MnCl_2$, $CaCl_2$ y lactosa, se observó la formación de cristales luego de 12 h en dos gotas, las condiciones 3 y 43 del ensayo (0.4 M fosfato de amonio monobásico y 50% v/v MPD respectivamente) (Figura 30), estos resultados nos permitieron definir cuál era el mejor agente precipitante en el que la proteína cristaliza.

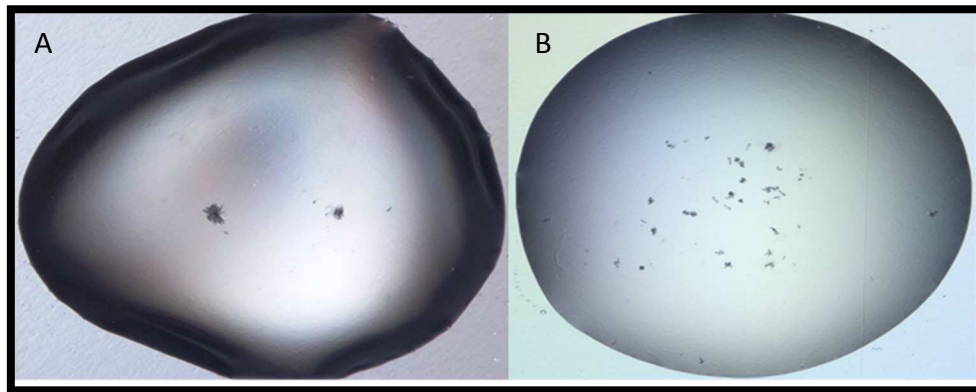
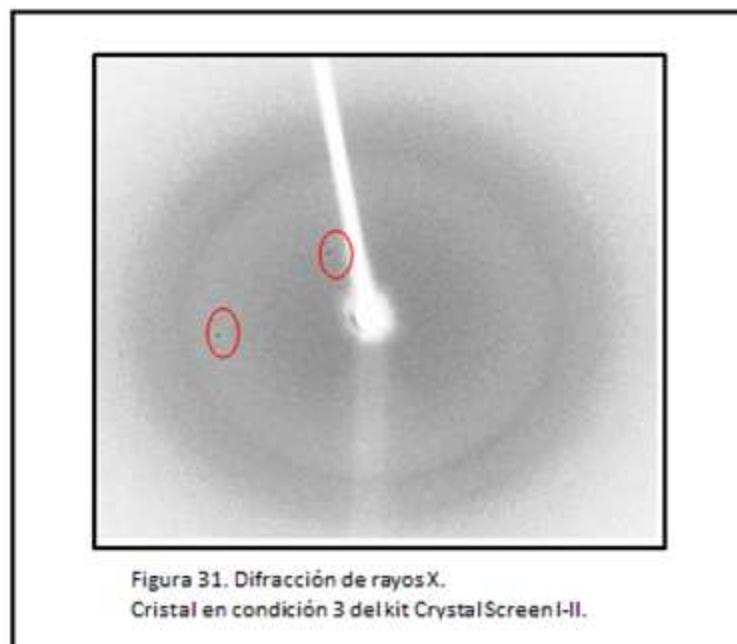


Figura 30. Gotas de cristalización positivas.

A. Condición 3 del Crystal Screen I, **B.** Condición 43 del Crystal Screen.

Para corroborar que se trataba de cristales de proteína y no de cristales de sal se dejaron crecer los cristales por ocho días más para realizar difracción de rayos X. Se observaron puntos con una resolución de 16 Å en el punto más cercano al haz de luz (Figura 31), y no se observaron señales características de cristales de sales; sin embargo al no observarse muchas señales alrededor del haz de luz se trabajó en la optimización de las condiciones de cristalización.



4.6.2 Optimización de condiciones de cristalización de la LGL-P2

Con las dos condiciones en las que crecieron los cristales se diseñó un experimento para obtener cristales de mayor tamaño. Con fosfato de amonio como agente precipitante no se obtuvieron cristales de proteína y no se observó nucleación, solamente precipitación y/o aglomeración de proteína (Figura 32, C). Con MPD como agente precipitante y realizando cambios de pH, se obtuvieron cristales en todas las gotas (Figura 32 A,B), y se observó que mientras más alta la concentración de MPD más separados y grandes eran los cristales (Figura 33 A,B). Los cristales se analizaron por medio de difracción de rayos X y se obtuvo una resolución de 20 Å (Figura 34) más baja que la que se obtuvo inicialmente; tampoco fue adecuada para realizar análisis de cristales en ciclotrón.

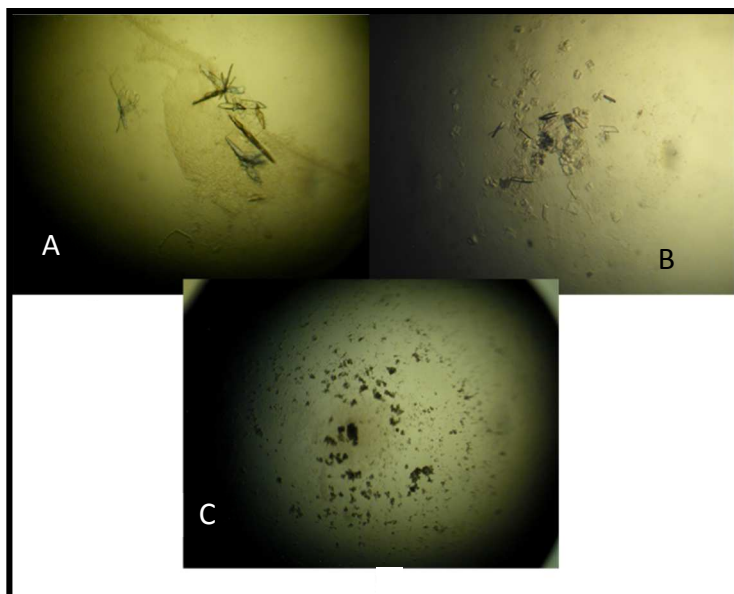


Figura 32. Optimización de condiciones.

A. Ensayo MPD, gota B2: 0.6 µl, **B.** Ensayo MPD, gota C6: 1 µl, **C.** Ensayo fosfato de amonio (gota A4).

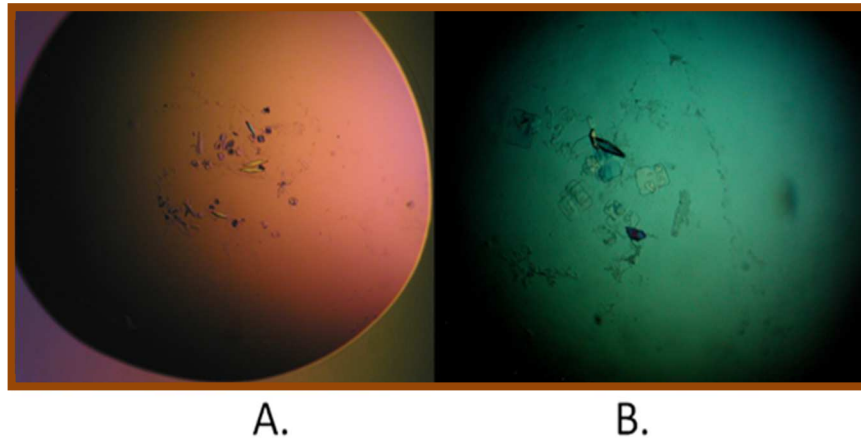


Figura 33. Optimización de condiciones.
A. Ensayo 3, gota A5: 0.6 μl , B. Ensayo 2, gota C4: 1 μl .

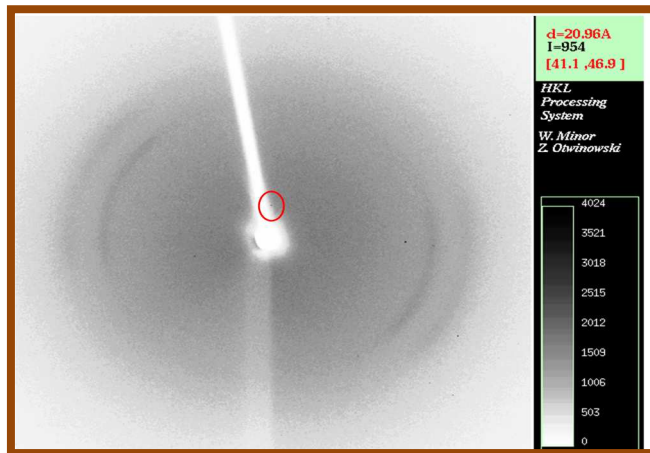


Figura 34. Imagen de difracción de rayos X. Cristal en gota C4:1 μl del ensayo 3 de optimización (pH 8.5, 62% MPD).

Posteriormente y con el fin de obtener cristales de mejor calidad se hizo una resiembra usando los mejores cristales con más LGL-P2, todo en las mismas condiciones que favorecieron la formación de los cristales (Figura 35).

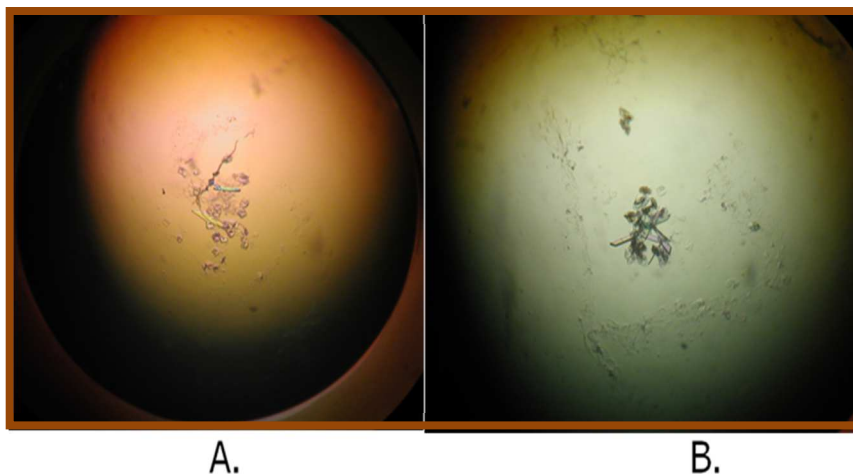


Figura 35. Seeding de cristales.

A. Seeding, gota C6: 1 μ l, **B.** Seeding, gota C4: 1 μ l.

Se obtuvieron cristales clasificados según su forma 2D como platos (gota C4) y 3D como “roca” (gota C6), pero fueron muy pequeños para determinar sus dimensiones. Algunos de ellos se congelaron y almacenaron para realizar posteriormente un análisis de mayor resolución al ciclotrón. Los resultados mostraron que el pH 7.4 de la solución de lectina facilitó la formación de los cristales pero el pH 8.5 al ser más cercano al punto isoeléctrico de LGL-P2 (8.3) mejoró la formación de los cristales sin la formación de precipitados, aunque se recomienda trabajar por lo menos una unidad de pH lejos del punto isoeléctrico (Rojas, 1998). En el caso de lectinas de *Erythrina* (género dentro de la subtribu *Erythrinae* se utilizó también el kit Cristal screen I-II (Svensson et al, 2002) y se obtuvieron cristales usando sulfato de amonio 2M y buffer tris-HCl 0.1 M pH 8.0. Los ensayos de cristalización son importantes ya que por medio de difracción de rayos X, se puede corroborar la estructura primaria de la LGL-P2 con las densidades electrónicas de los aminoácidos y por otra parte con la estructura terciaria se puede entender como es la interacción de la lectina con el carbohidrato y su posible función biológica.

4.7 Purificación de la LGL-P4

4.7.1 Purificación de la LGL-P4.

Dentro de la familia de las leguminosas la tribu *Diocleae* está dividida en subtribus, de las cuales la *Diocleinae* ha sido ampliamente estudiada. Lo más interesante es que en diferentes especies se producen lectinas que están cercanamente relacionadas y tiene un sitio de unión conservado que reconoce oligosacáridos de manosa o son específicas de Glc/Man. Algunas de las mismas especies co-sintetizan una segunda lectina con un sitio de unión conservado que reconoce el trisacárido H2 (Figura 36, tomado Dam et al., 2011, modificado). A pesar de su proximidad filogenética, su secuencia de aminoácidos conservada y similar especificidad por carbohidratos se presentan diferencias muy sutiles en la interacción que se evidencian en las constantes de afinidad. Por otra parte en la especie *Diocleae lehmanni* se estableció la presencia simultánea de las dos lectinas en el mismo organelo o cuerpo proteico (Melgarejo y Pérez, 1997). La comparación de las secuencias de aminoácidos N-terminales de las lectinas tipo P2 mostraron un alto porcentaje de similitud, lo cual nos llevó a buscar en *Galactia lindenii* una lectina similar a P4, proteína que no ha sido descrita hasta ahora. LGL-P4 se encontró en una cantidad muy pequeña y tiene propiedades similares a las lectinas tipo P4 de *Canavalia ensiformis* (Concanavalina A) y de las especies de *Dioclea*. Esta segunda lectina presente en las semillas de *Galactia lindenii* (LGL-P4) inicialmente aglutinó eritrocitos de conejo.

En este trabajo se realizó la purificación y la caracterización parcial de esta segunda lectina partiendo del mismo extracto que se empleó para purificar la LGL-P2. Una ventaja de la cromatografía de intercambio iónico es que las dos lectinas se separan en este paso de purificación; en la fracción no retenida se obtiene la LGL-P2 y en la retenida la LGL-P4 (Figura 37 A). Por electroforesis en geles de poliacrilamida se observa una banda alrededor de 29 kDa para la fracción retenida que también se evidencia en el extracto crudo. En el perfil electroforético (Figura 37B, Carriles 2, 4) es

importante resaltar la proporción de las lectinas, donde la LGL-P2 se encuentra en porcentaje mucho mayor que la LGL-P4 (Carril 3,4). Con las fracciones retenidas se hizo una cromatografía de afinidad sobre Sephacryl S-200 y la fracción retenida se eluyó con glucosa 0.2 M; en este último caso se siguió la metodología descrita para otras lectinas similares tipo P4 afines por Glc/Man (Pérez, et al., 1990; Sierra y Pérez, 1999; Melgarejo et al., 2005).

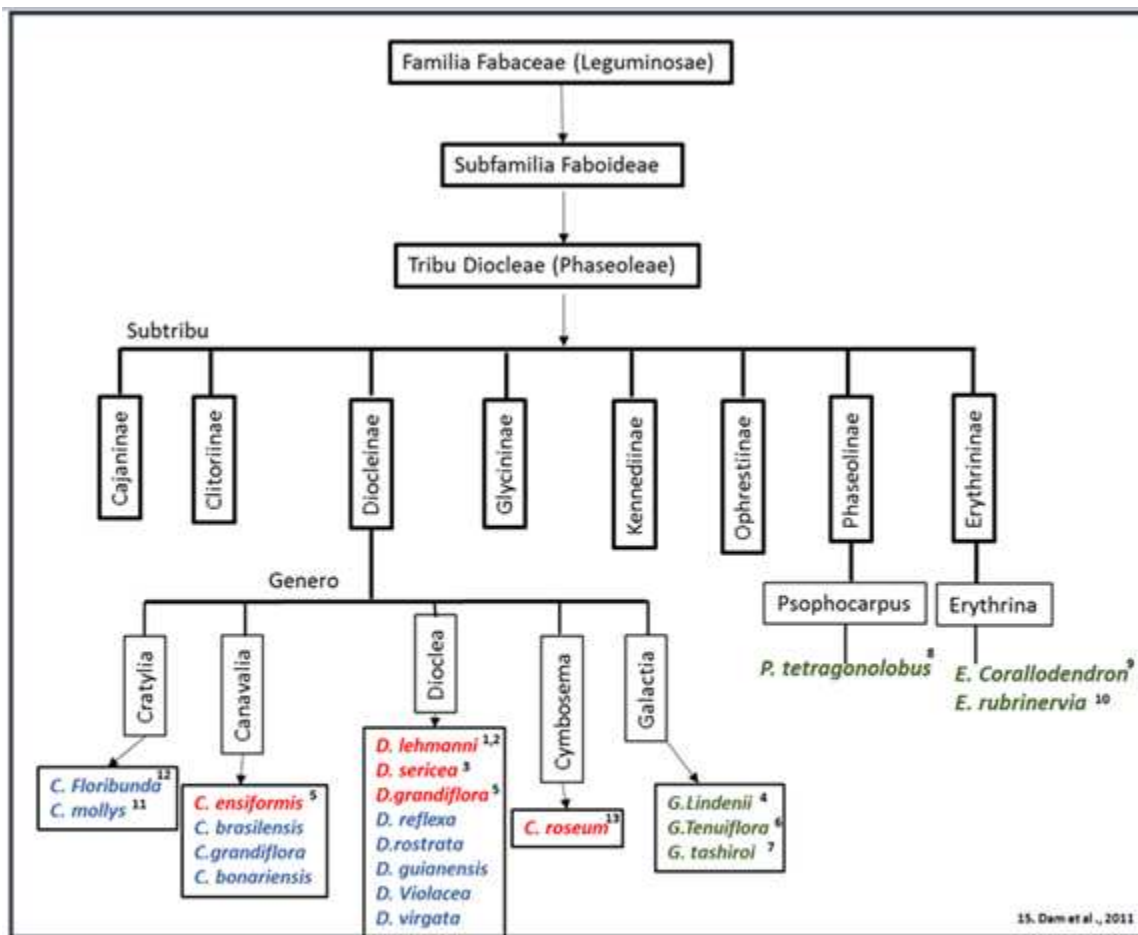


Figura 36. Lectinas de la subtribu Diocleinae

Color Azul : Especies que poseen Lectina Glc/Man. Color Rojo: Especies que poseen dos clases de lectinas : H-tipo 2 y Glc/Man.

Color Verde : Especies que poseen lectina específica para H-tipo 2

1. Pérez et al., 1990; 2. Pérez et al., 1998; 3. Sierra y Pérez, 1999 4. Almanza et al., 2004 5. Melgarejo et al., 2005

6. Le pendu et al., 1986; 7. Fukuda et al., 1994; 8. Srinivas et al., 1999; 9. Moreno et al., 1997

10. Peña et al., 1988; 11. Correia y Coelho, 1995 12. Grangeiro et al., 1997 13. Rocha et al., 2009; 14. Kuku et al., 2007; 15. Dam et al., 2011

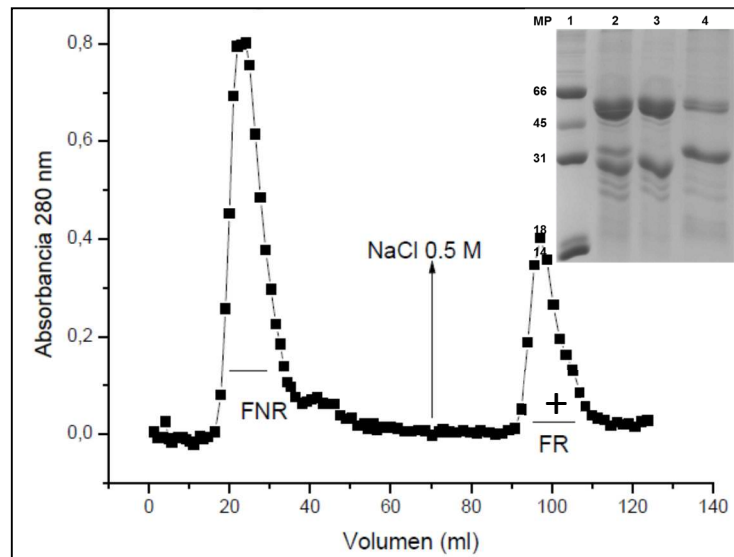


Figura 37. Purificación de la LGL-P4

A. Cromatografía de Intercambio iónico sobre DEAE-Sephacel

B. PAGE –SDS: 20 µg proteína /carril

Carriles: 1. Marcador de peso molecular (kDa), 2. Extracto crudo.

3. Fracción no retenida 4. Fracción retenida

La fracción no retenida de la cromatografía de afinidad en algunos casos fue desprovista de actividad aglutinante ya que previamente se había eliminado del extracto la LGL-P2 (Figura 38 A, B). Los perfiles de afinidad de las dos lectinas, claramente nos muestran que la cantidad de LGL-P4 es mucho menor que la LGL-P2 (Figura 38 C, D); otros ensayos se realizaron purificando directamente el extracto crudo (Figura 38 C).

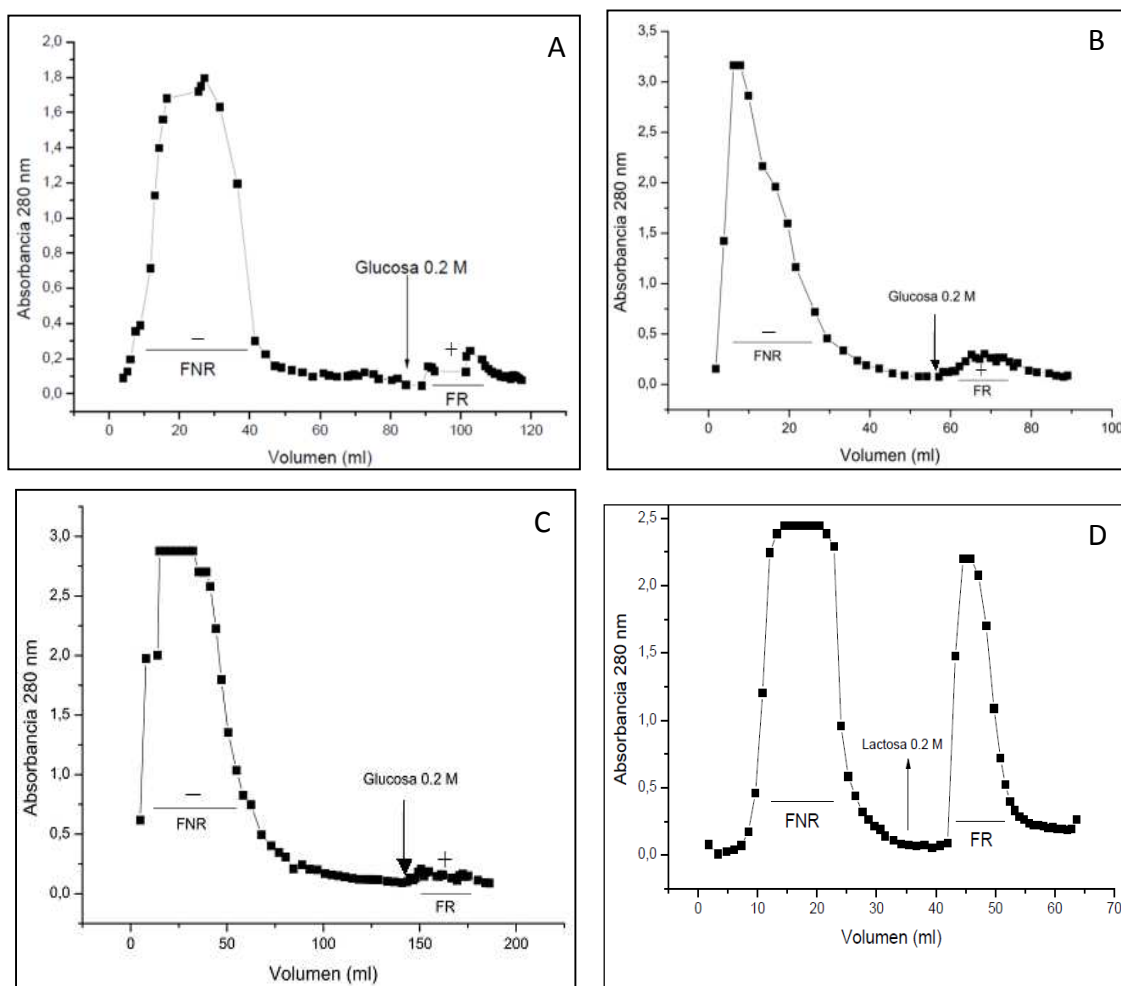


Figura 38. Cromatografía de afinidad de LGL-P4 sobre Sephacryl S-200

A: se sembraron 8.92 mg de la fracción retenida de DEAE –Sephacel

B: Recromatografía: Se sembraron 47.88 mg de las fracciones no retenidas del Sephacryl S-200

C: Se sembraron 162.73 mg del extracto crudo

D: Se sembraron 120mg del extracto crudo sobre Sepharosa 4B- Lactosa

En la tabla 14 se presentan las concentraciones de algunas fracciones de la LGL-P4 obtenidas durante el proceso de purificación. Para evaluar la actividad se hicieron ensayos de aglutinación con eritrocitos humanos A+, B+ y O+ los cuales fueron aglutinados por la proteína; sin embargo La LGL-P4 aglutinó mejor los eritrocitos A+ y B+ (mínima cantidad de lectina para ver aglutinación $0.625 \mu\text{g}$) que los eritrocitos tipo O+ ($1.25 \mu\text{g}$). Las lectinas P4 de *Dioclea lehmanni* y *Dioclea sericea*, también aglutinaron eritrocitos humanos; en este último caso aglutinó mejor los eritrocitos A+ y B+ (Pérez G., 1998; Sierra y Pérez, 1999). La aglutinación con eritrocitos de conejo es menos confiable

ya que en muchos casos se autoaglutinan y por lo tanto es difícil evaluar algunas veces la actividad, pero es importante mencionar que la mayoría de las lectinas que reconocen Glc/Man tienen mayor especificidad o solo aglutinan eritrocitos de conejo como la Concanavalina A (ConA) (Goldstein et al., 1986).

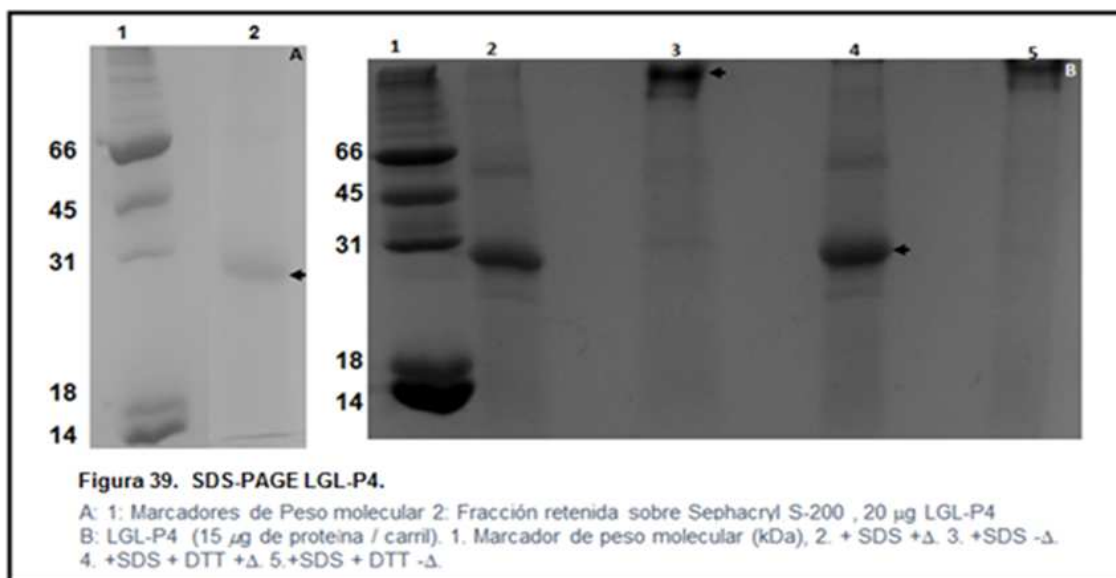
Tabla 14. Purificación de la LGL-P4.

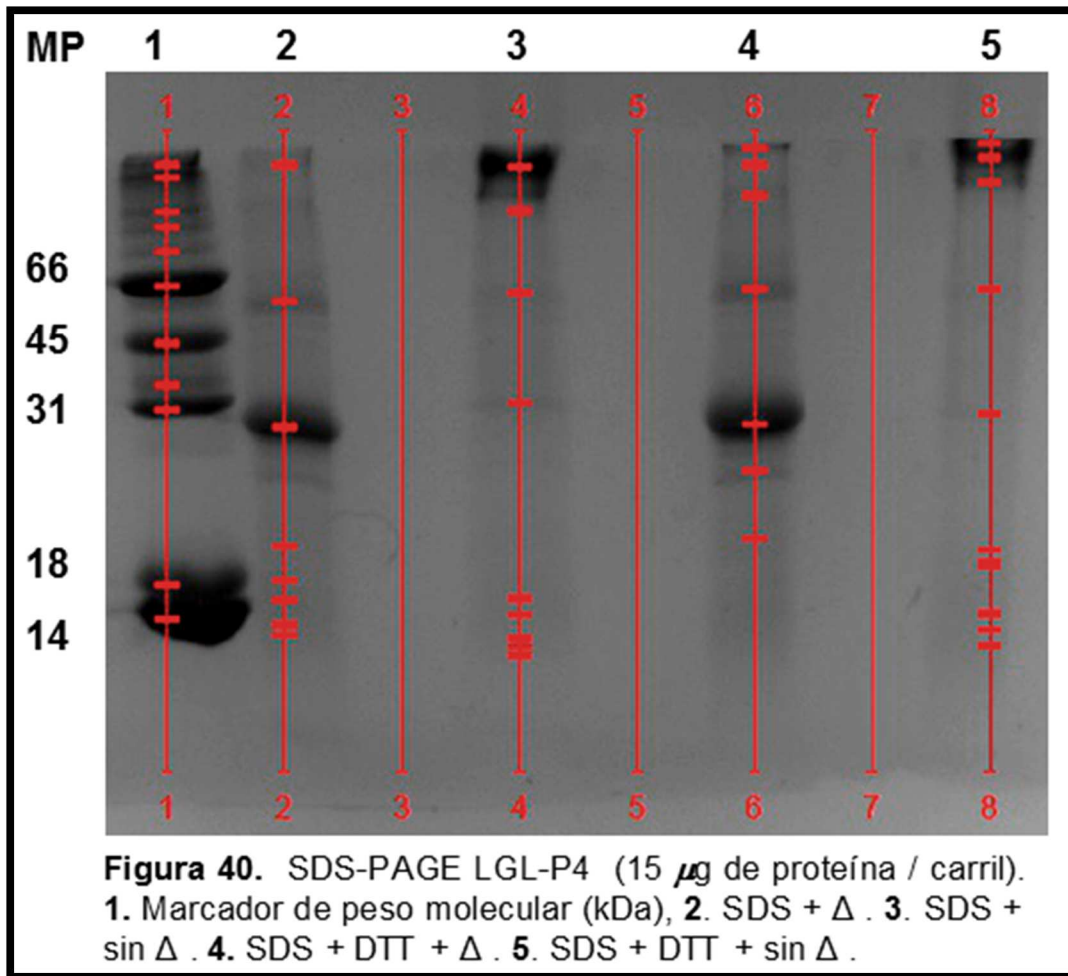
Método	(mg/mL)	Volumen (μL)	μg totales	Aglutinación (A+)
Extracto crudo	0,86	1100	960	+4
Intercambio Ionico	0,446	250	111,5	+4
Recromatografía fracciones no retenidas	0,58	1400	812	+4

4.7.2 Electroforesis unidimensional en geles de poliacrilamida

Por SDS-PAGE en condiciones reductoras se observaron dos bandas predominantes alrededor de 29 y 60 kDa, las cuales corresponden al monómero y el dímero respectivamente (Figura 39 A, carril 1; B). Con el fin de evaluar la interacción entre las subunidades de la proteína se prepararon las muestras en presencia y ausencia de DTT y con y sin calentamiento; estas lectinas se caracterizan por la presencia en su estructura de puentes disulfuro intracatenarios. Se observó claramente la necesidad del calor para la disociación completa del tetrámero, esto se evidenció por la banda de alto peso molecular y una total ausencia del monómero y el dímero (Figura 39B carril 3 y 5). Otros estudios han mostrado que las lectinas P4 presentan tres bandas características que corresponden a la cadena alfa (α) (alrededor de 30 kDa), la cadena beta (β) (alrededor de 16.5 kDa) y la cadena gamma (γ) (alrededor de 13.4 kDa); en estas proteínas se presentan modificaciones postraduccionales (Herman, et al., 1985; Bowles y Pappin, 1988) donde la subunidad α sufre un hidrolisis en condiciones reductoras a dos subunidades β y γ (Moreira et al, 1987; Melgarejo y Pérez, 1997; Pérez G., 1998). Sorpresivamente en el caso de la LGL-P4 no se observaron las bandas β y γ , sin embargo se hizo un análisis de los geles en el analizador (BIORAD VersaDoc 3000) (Figura 40). Se detectaron algunas bandas de pesos moleculares entre 10-20 kDa en una

muy baja proporción, probablemente las subunidades β y γ o en la LGL-P4 la subunidad alfa no se disocia fácilmente y se requieren condiciones más drásticas de reducción. En la electroforesis para la lectina de *Cratylia mollis* tampoco fueron evidentes las subunidades β y γ (Correia y Coelho, 1995) mientras que con la de *Cratylia floribunda* se evidencia para la cadena β dos bandas (16 y 18 kDa) y una para la γ (12-13 kDa) (Grangeiro et al., 1997). Para las lectinas P4 de los géneros *Dioclea*, *Canavalia* y *Cymbosema* las subunidades β y γ se evidencian fácilmente ya que están en una gran proporción con respecto a la cadena α (Moreira et al., 1983; Cavada et al., 2006; Martínez P, 2009).





El comportamiento electroforético de LGL-P2 y LGL-P4 indica que son dos lectinas que constituyen dos grupos bien diferenciados. Con este trabajo es la primera vez que se reporta la presencia de la segunda lectina Glc/Man dentro del género *Galactia*. Las características bioquímicas son diferentes con respecto a la LGL-P2, más adelante en la tabla 5 se resumen las propiedades moleculares de algunas lectinas P4 estudiadas hasta la fecha; para realizar los ensayos de caracterización parcial se obtuvieron 2.69 mg de la LGL-P4.

A partir de 100 g de harina se obtienen 428 µg de LGL-P4, realizando la purificación a partir del extracto crudo por cromatografía de afinidad sobre Sephacryl S-200; la cuantificación en cada uno de los pasos de purificación se hizo por el método de microKjeldahl. Las lectinas LDL-P4 (Lectina de *Dioclea lehmanni*-P4), LDS- P4 (Lectina de *Dioclea seríceea*-P4) y LDG-P4 (Lectina de *Dioclea grandiflora*-P4) se encuentran en una cantidad entre 1200 -1300 mg de lectina pura /100 g de harina seca. Con estos resultados es evidente que en el género *Dioclea*, la lectina P4 está en mayor proporción que la P2, caso contrario al género *Galactia*, donde la lectina P2 está en mayor proporción que la P4, lo que nos muestra que hay una diferencia en expresión de cada una de las lectinas dentro de la subtribu *Diocleinae*. (Hipótesis presentada por Almanza et al., 2004), por lo tanto sería interesante analizar cuáles son las razones fisiológicas por las cuales se da esta diferencia de expresión y por qué se presenta una correlación inversa (Tabla 15). Se propone que la evaluación cuantitativa de P2 y P4 en cada una de las especies mostró que el contenido decrece y crece en el siguiente orden:

Contenido de P2: LGL > LDG > LDL > LCE

Contenido de P4: LGL < LDL < LCE < LDG

Tabla 15. Contenido de las lectinas P2 y P4 en las especies estudiadas

Especie	Lectina P4 mg/100 g harina	Lectina P2 mg/100 g harina
<i>Canavalia ensiformis</i>	2 400	30,2
<i>Dioclea grandiflora</i>	1 210	21,6
<i>Dioclea lehmanni</i>	1 300	184
<i>Galactia lindenii</i>	4,28	588

Quintero et al, 2013.

Estos resultados nos llevan a plantear varias preguntas relacionadas con la función de estas proteínas que quedan por resolver:

1. Cuáles factores regulan esa modulación?
2. Porque razón se da dicha modulación?
3. Qué ventajas poseen las especies que tienen un mayor contenido de lectina respecto a las otras?
4. Qué interacciones hay entre estas proteínas y otras proteínas presentes en el citoplasma de la célula?
5. Existen complejos multiproteicos con funciones definidas?

Es importante mencionar que hay una gran variedad de ejemplos donde se presenta más de una lectina en las semillas de plantas con especificidades diferentes (Van Damme et al., 1997), lo que indica que no es un rasgo único de la Familia de las leguminosas ni de la subtribu *Diocleinae*.

4.8 Caracterización parcial de la lectina LGL-P4.

4.8.1 Ensayos de inhibición con carbohidratos.

La actividad de la lectina fue inhibida con manosa, galactosa, glucosa y fucosa lo que indica que los OH que están en posición 2 y 3 no son importantes para el reconocimiento del carbohidrato. En general se presentó inhibición con los azúcares que en su estructura tenían manosa o glucosa sin importar la configuración (α , β) de los grupos hidrofobicos metilo y paranitrofenilo. La actividad no fue inhibida por los galactósidos y la lactosa; sin embargo se presentó una inhibición muy leve con la galactosa y el 1-O-metil- β -galactopiranosido (Tabla 16). Los ensayos de concentración mínima inhibitoria mostraron que el mejor inhibidor de la actividad fue el p-Nitrofenil- β -D-manopiranosido con una concentración de 9.5 mM (Tabla 17), lo que nos muestra que el grupo p-nitrofenil es importante en la interacción con la lectina y que esta puede discriminar entre la

configuración α y β . En las lectinas de leguminosas por lo general cerca al sitio de interacción con el carbohidrato hay un bolsillo que promueve la interacción con grupos hidrofobicos; de esta forma se potencia el reconocimiento del carbohidrato por la lectina. La LDS -P4 también fue inhibida por el p-Nitrofenil- α -D-manopiranosido a la misma concentración (Sierra y Pérez, 1999). La lectina mostró actividad inhibitoria frente a p-Nitrofenil-N-acetil- β -D-galactosamina y 1-O-metil- β -galactopiranosido, probablemente son más importantes las interacciones hidrofóbicas que los puentes de hidrogeno, la lectina no fue inhibida por galactosa.

Tabla 16. Evaluación de la inhibición de la actividad de la LGL-P4 con azúcares.

Carbohidrato	Concentración inicial	Concentración final en solución	Inhibición
D-Glucosa	0,3 M	0,150 M	+
D-Manosa	0,3 M	0,150 M	+
D-Lactosa	0,3 M	0,150 M	-
D-Galactosa	0,3 M	0,150 M	+
L-Fucosa	0,3 M	0,150 M	+
GalNAc	0,3 M	0,150 M	-
1-O-metil- β -galactopiranosido	0,3 M	0,150 M	+
1-O-metil- β -glucopiranosido	0,3 M	0,150 M	-
β -metil D-glucosa	0,3 M	0,150 M	+
N-acetil- α -D-glucosamina	0,3 M	0,150 M	±
α -metil O- manosido	0,3 M	0,150 M	±
ácido D-galacturónico	0,3 M	0,150 M	-
p-Nitrofenil- β -D-glucopiranosido**	37 mM	18,5 mM	±
p-Nitrofenil- β -D-glucosamina**	37 mM	18,5 mM	±
p-Nitrofenil-N-acetil- α -D-glucosamina**	37 mM	18,5 mM	+
p-Nitrofenil- β -D-manopiranosido**	37 mM	18,5 mM	+
p-Nitrofenil- α -D-manopiranosido**	37 mM	18,5 mM	+
p-Nitrofenil- β -D-galactopiranosido**	37 mM	18,5 mM	-
p-Nitrofenil-N-acetil- β -D-galactosamina**	37 mM	18,5 mM	+
p-Nitrofenil-N-acetil- β -D-galactopiranosido**	37 mM	18,5 mM	-

+: Fue inhibida la aglutinación

** soluble hasta una concentración de 37 mM

Dam et al., 2011, estudiaron la interacción de monosacáridos y oligosacáridos que incluían manosas para la LCR-II por medio de ensayos de inhibición y titulación calorimétrica isotérmica (ITC); ellos encontraron que la afinidad de la LCR-II se incrementa con los oligosacáridos altos en manosas (core de trimanosas; $K_a = 24.7 \times 10^4 \text{ M}^{-1}$) en comparación con α -metilmanopiranosido; una característica que ha sido observada para otras lectinas P4 como Concanavalina A (ConA) y *Dioclea grandiflora* (LDG) (Dam et al., 1998). La constantes de afinidad que se determinaron los llevaron a sugerir que estas lectinas poseen un sitio de unión extendido que acomoda el core de trimanosa y que es un sitio de unión conservado para estas lectinas. Por glicoarreglos se evaluó la interacción de la LDL-P4 con un panel de oligosacáridos y se observó básicamente el mismo patrón de reconocimiento exhibido por la ConA, además de la unión a glicoproteínas con N-glicanos altos en manosa y ramificados; también reconoció α -Manosa (Vega et al., resultados no publicados). En el caso de la LGL-P4 se deben realizar otros estudios para determinar su patrón de reconocimiento aunque se espera que sea similar al determinado para otras lectinas P4.

Tabla 17. Mínima concentración inhibitoria de carbohidratos para la LGL-P4.

Carbohidrato	Concentración inicial (mM)	Concentración mínima de inhibición (MCI) (mM)	Actividad Inhibitoria Relativa **
D-Manosa	300	150	1
L-Fucosa	300	300	0,5
1-O-metil- β -galactopiranosido	300	37	4
p-Nitrofenil-β-D- manopiranosido	37	9,5	16
p-Nitrofenil- α -D-manopiranosido	37	37	4
p-Nitrofenil-N-acetil- β -D-galactosamina	37	18,5	8

**Actividad Inhibitoria relativa: Es calculada como la proporción MCI de manosa /MCI del monosacárido evaluado

4.8.2 Determinación de la presencia de carbohidratos y del punto isoeléctrico

La LGL-P4 se caracterizó por ser una glicoproteína (Figura 41) al igual que las demás lectinas tipo P4 estudiadas hasta el momento, sin embargo la concanavalina A (ConA) y la *Dioclea grandiflora* se caracterizan por ser no glicosilados (Melgarejo et al., 2005). En este trabajo no se determinó el porcentaje de carbohidratos pero para algunas las lectinas tipo P4 está entre 1.7 -5%, un porcentaje bajo si se compara con lectinas de labiadas estudiadas en nuestro grupo las cuales presentaron porcentajes entre el 13-28% (Pérez et al., 1990; Sierra y Pérez, 1999; Filgueira, 2002; Pérez y Vega, 2007).

La determinación del punto isoeléctrico (pI) arrojó una banda entre 5.85 – 6.55, determinado con la curva de calibración (Anexo 17), cercano a 6.15 (Figura 42). Este pI fue similar al reportado para la LDS-P4, el cual está entre 6.64-6.85, en la tabla 18 y 19 se presentan las propiedades moleculares de algunas lectinas P4 y su secuencia N-terminal; esta secuencia es muy conservada en estas lectinas. Aunque en este trabajo no fue posible determinarla se esperaba que tuviera una alta identidad y similaridad con respecto a las secuencias que ya han sido reportadas.

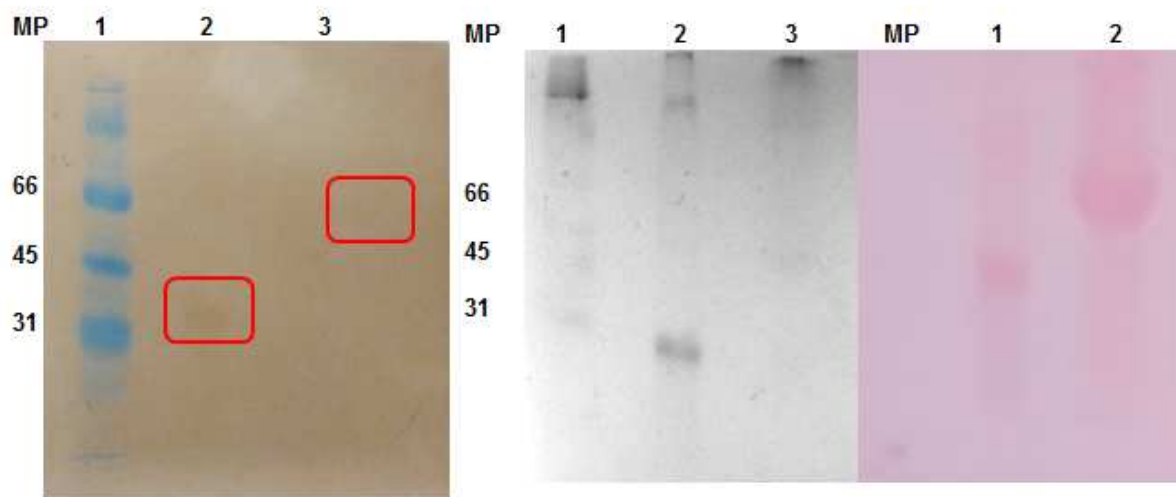


Figura 41. Determinación de glicoproteína para LGL-P4.

A. Membrana de nitrocelulosa tratada para detección de glicoproteínas.

1. Marcador de peso molecular (kDa). 2. Fracción proteína pura LGL P4 (15 μ g). 3. Control de Ovoalbúmina (150 μ g).

B. SDS-PAGE. 1. Marcador de peso molecular (kDa). 2. Fracción proteína pura LGL P4 (15 μ g). 3. Control de Ovoalbúmina (150 μ g).

C. Membrana de nitrocelulosa transferida y teñida con rojo ponceau. 1. Fracción proteína pura LGL P4. 2. Control de Ovoalbúmina.

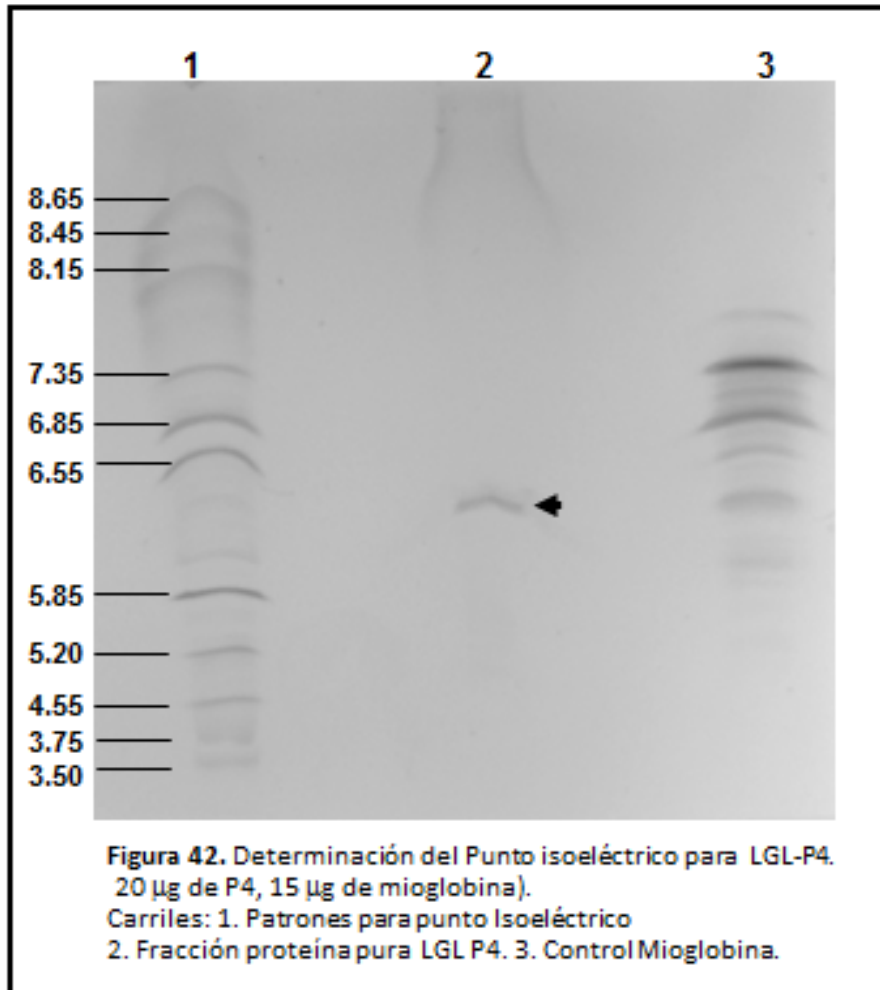


Tabla 18. Propiedades moleculares de algunas lectinas P4 Subtribu *Diocleinae*

	LGL-P4 ¹	LDL-P4 ²	LCR-I ³	ConA ⁴
M _r subunidades (kDa)	---		-----	26 500
M _r proteína (kDa)	100 ⁽⁸⁾	25 300 14 000	-----	106 000
Bandas en SDS-PAGE (kDa) ⁶	29; 60	25; 14	30 ; 18 ; 12	26; 14; 12,5
Azúcares neutros (%)	---	1.7-1.9	-	
pI	6,15	8,0 ; 8,13 8,3 ; 8,42	-	7.1
Inhibición por Manosa (mM)	150	50	19.5	
p-Nitrofenil-β-D-manopiranosido (mM)	16	-----	-----	

¹ Este trabajo.² Pérez et al., 1990³ Cavada et al., 2006⁴ Pérez, 1998⁵ Alperin *et al.*, (1992).⁶ Proteína reducida.⁸ Sin calentamiento.

Tabla 19. Secuencia N-terminal Lectinas subtribu <i>Diocleinae</i>	
<i>Cymbosema roseum</i>	ADTIVAVELDSYPNTDIGDPSYPH
<i>Dioclea lehmanni</i>	ADTIVAVELDSYPNTDIGDPSYPH
<i>Dioclea grandiflora</i>	ADTIVAVELNSYPNTDIGDPNYPH
<i>Canavalia ensiformis</i>	ADTIVAVELDTYPNTDIGDPSYPH
<i>Canavalia brasiliensis</i>	ADTIVAVELDSYPNTDIGDPSYPH
<i>Cratylia floribunda</i>	ADTIVAVELDSYPNTDIGDPNYQH

Tomado de Cavada et al., 2006

Sin embargo aunque las lectinas P4 tienen una alta homología se han observado diferencias en su función biológica, por ejemplo estimulan la producción de citoquinas (IFN- γ , IL-10, TNF- α , IL4 entre otras) en células mononucleares de sangre periférica, se emplean como mitógenos para la diferenciación de linfocitos (estudios citogenéticos), presentan actividad tóxica frente a diferentes insectos plaga y se usan en estudios citoquímicos e histoquímicos entre otros (Cavada et al., 2001). Para evaluar las propiedades biológicas y la función de estas proteínas es importante realizar estudios de caracterización bioquímica y estructural; en el caso de las lectinas de *Galactia lindenii* son proteínas que hasta ahora no se han estudiado desde el punto de vista biológico y estructural.

5. Conclusiones.

- Se determinó la secuencia de 20 péptidos que ensamblaron para la lectina de *Galactia lindenii* P2.
- Por superposición de los péptidos secuenciados se determinó la estructura primaria para 195 aminoácidos con una cobertura de la secuencia del 89% y una masa no secuenciada de 2750 Da. Es la primera estructura primaria reportada a la fecha para lectinas P2 del género *Galactia* (“nnn” representa la parte de la secuencia que hace falta determinar).

AKVTSIKYTSIS-SSXGKPLQGNAGELTVAQLTDLKNPSSNLGRVLYSSPVTLWDEATGNVASFVSSLTL
 RFEDVSGYVPANGALGVVNAPDAGTLPSPGSAGGQLGTVNPKDAFNNFVGVEFDDYSNSWDPSYSFIGLD
 NSLLSLQTaknnnnSVAVENDGQLLTVAQMVDKLAVLPSKVRVGLSAATSGNGVQKHdVYnnn

- La predicción de la estructura terciaria usando las secuencias parciales de LGL-P2 y LDL-P2 mostró la presencia de hojas beta antiparalelas y regiones desordenadas, las cuales son características de lectinas de leguminosas.
- Para la lectina LGL-P2 se obtuvieron cristales clasificados según su forma 2D como platos, empleando fosfato de amonio monobásico 0.4 M y 50% v/v MPD a pH=8.5.
- La LGL-P4 es una glicoproteína con un peso molecular de 30 kDa y un punto isoeléctrico igual a 6,15. Aglutinó mejor los eritrocitos A+ y B+. El mejor inhibidor de su actividad fue p-Nitrofenil- β -D- manopiranosido.

- **Por primera vez se reportó una segunda lectina dentro de las semillas de *Galactia lindenii* la cual reconoció glucosa/manosa y aglutinó mejor los eritrocitos A+.**

6. Perspectivas.

- Realizar el análisis de la expresión y evaluar el efecto fisiológico de ambas lectinas para la planta y qué ventajas poseen las especies que tienen un mayor contenido de lectina respecto a las otras.
- Obtener el cADN de la lectina para comparar la secuencia obtenida por biología molecular con la secuencia parcial determinada por medio de secuenciación de péptidos obtenidos por espectrometría de masas.
- Evaluar la calidad de los cristales obtenidos para la LGL- P2.
- Continuar con la caracterización molecular completa de la lectina LGL-P4.

7. Aportes al estudio de lectinas.

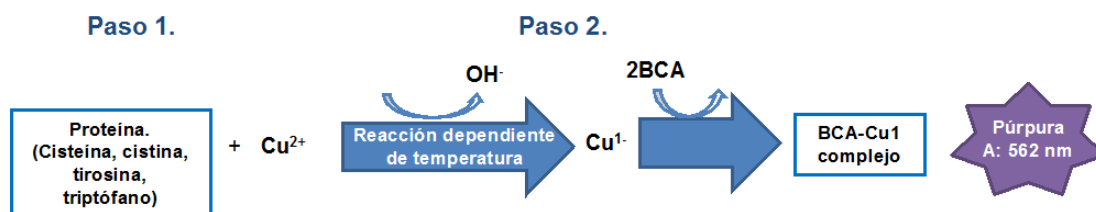
- Desarrollo del método simultáneo de separación de las dos lectinas presentes en las semillas de *Galactia lindenii*; y estandarización del método de purificación de la LGL-P4. Por primera vez se reporta la presencia de una segunda lectina en las semillas de *Galactia lindenii*.
- Desarrollo de la estrategia de análisis informático de secuencias para determinar estructura primaria a partir de péptidos obtenidos por digestión enzimática en caso de no tener el genoma de la planta fuente.
- Estandarización de las condiciones de cristalización para la LGL-P2, lo cual es una base de partida para continuar los estudios de cristalización y estructura terciaria de otras lectinas.
- Participación en Congreso 17th European Carbohydrate Symposium (Eurocarb 17). Tel-Avid (Israel). Abstracts book pag 67 2- 7, Julio 2013. “EXPRESSION OF TWO TYPES *Diocleinae* SUBTRIBE LECTINS”. Quintero Moreno Mónica, Reyes Montaña Edgar Antonio, Pérez Gómez Gerardo, Vega Castro Nohora Angélica.
- Colaboración con la Universidad Johns Hopkins. Departamento de Biofísica y Química Biofísica.
- Artículo en Proceso: Estructura primaria parcial de la lectina de LGL-P2 obtenido por espectrometría de masas MS/MS.

8. Anexos.

Anexo 1: Método del Ácido Bicinconínico (BCA) para cuantificación de proteínas (Smith et al, 1985).

El método se basa en la reducción de Cu^{2+} a Cu^{1+} . Esta reacción está influenciada por residuos de cisteína, tirosina y triptófano; También de los enlaces peptídicos.

La reacción es la siguiente:



Donde, la absorbancia es directamente proporcional a la cantidad de proteína presente en la solución y puede ser estimada por comparación con un estándar de proteína conocido (generalmente Albumina Bovina sérica). El método del BCA es altamente tolerante a detergentes iónicos y no iónicos, pero por ejemplo los agentes reductores pueden interferir con el ensayo. Su límite de detección está entre 2-2000 $\mu\text{g/mL}$.

Anexo 2: Preparación de geles SDS-PAGE en Tris-Tricina para separación de proteínas de bajo peso molecular (Schägger H y Von Jagow G. 1987).

Los geles de Tris-Tricina se prepararon para llevar a cabo las separaciones de los péptidos obtenidos en las digestiones enzimáticas con las diferentes enzimas utilizadas. La preparación de un gel a una concentración 16.5%T, 3%C, se muestra en la tabla 20.

Tabla 20: Preparación del gel SDS-PAGE en Tris-Tricina.

Reactivo	Gel Separador (mL)	Gel Concentrador (mL)
Acrilamida (Acrilamida 48%-Bisacrilamida 1.5%)	1.67	0.25
Buffer Tris-HCl 0.3 M, 0.3% SDS pH 8.45	1.67	0.775
Agua desionizada	1.12	2.072
Glicerol	0.530	-----
Persulfato de amonio 10%	0.01667	0.025
TEMED	0.00167	0.0025

Luego de correr el gel se realiza la fijación de las proteínas con ácido acético 10% durante 30 min y posteriormente se tiñe el gel con coomassie blue G-250.

La preparación de los demás reactivos utilizados se muestra a continuación (Tabla 21).

Tabla 21: Preparación buffer de muestra y buffer de corrido geles SDS-PAGE en Tris-Tricina.

Buffer de muestra	Buffer de corrido
Buffer Tris-HCl 0.45 M pH 8.45	Tris 100 mM
SDS 4%	Tricina 100 mM
Glicerol 12%	SDS 0.1%
Coomassie Blue G-250 0.0025%	----
Rojo de fenol 0.025%	----

Anexo 3: Tinción con plata de geles SDS-PAGE (Vega N, comunicación personal).

Las tinciones de plata se llevaron a cabo según el procedimiento descrito a continuación:

1. Incubar el gel en ácido acético 10% durante 30 min.
2. Eliminar el ácido acético y adicionar al gel una solución de Metanol 10% - Ácido acético 7% - H₂O, durante 20 min.
3. Eliminar la solución anterior y adicionar Metanol 50% durante 10 min.
4. Realizar dos lavados con H₂O, cada uno de 10 min.
5. Adicionar Tiosulfato de sodio 0.016 g/50 mL, e incubar durante 5 min.
6. Realizar dos lavado rápidos con H₂O.
7. Adicionar el Nitrato de plata fresco a una concentración de 0.5 mg/mL. Dejar el gel cubierto de la luz durante 20 min.
8. Eliminar el Nitrato de plata y adicionar una solución tibia y fresca de Carbonato de sodio 1g/50mL – Formaldehído 200 µL.
9. Al instante en que aparezcan bandas, eliminar la solución de carbonato de sodio – formaldehído y adicionar la solución neutralizadora, Ácido acético 1%.
10. Lavar el gel con H₂O y documentar.

Anexo 4: Características básicas de los kits utilizados para los ensayos de cristalización.

Los kits utilizados fueron Cristal screen I y II de Hampton Research, y el Morpheus screen MD1-46 de Molecular Dimensions.

Cristal screen I-II.

El kit presenta un método altamente efectivo y rápido para evaluar las condiciones de cristalización de macromoléculas. Se hacen 96 mezclas únicas de pH, sales, polímeros y compuestos orgánicos para ver las condiciones iniciales del crecimiento de cristales.

Características:

- Para proteínas, péptidos solubles y otras moléculas pequeñas solubles en agua.
- Amplio rango de condiciones.
- Formulación eficiente de mezclas de sales, polímeros, compuestos orgánicos y pH.
- Efectividad con más de 1000 macromoléculas biológicas.

Morpheus screen MD1-46.

El kit facilita un chequeo rápido de 96 condiciones de cristalización de proteínas basado en minería de datos de los resultados reportados en el PDB de proteínas cristalizadas y sus condiciones óptimas. Incorporando componentes de bajo molecular como ligandos, que también han sido reportados en PDB y que van desde el HEPES (PM: 238.30 g/mol) hasta iones de litio (6.94 g/mol).

Características:

- Cubre un amplio rango de condiciones de pH, PEGs y sales (aditivos).
- Incorporación de 49 ligandos de bajo peso molecular que promueven tanto la formación de cristales como la estabilidad de los mismos.
- Todas las condiciones presentan crioprotección.

Anexo 5: Detección cualitativa de glicoproteínas (Vega, 1997).

El tratamiento de la membrana para detectar glicoproteínas fue:

1. Lavado: con PBS 1X filtrado, durante 10 min.
2. Oxidación: con el metaperiodato de sodio 30 mM (0.081 g en 7 mL) en buffer acetato-acético 100 mM pH 5-5.5. Durante 20 min.
3. Lavado: con PBS 1X filtrado por 10 min, repitiendo 3 veces el lavado.
4. Incubación con Biotina-hidrazina a una concentración 1.3 mM, disuelta en 0.1 mL de DMSO y 1 mL buffer acetato-acético 100 mM pH 5-5.5. Incubar por 60 min a temperatura ambiente.
5. Lavado: con PBS 1X filtrado por 10 min, repitiendo 3 veces el lavado.
6. Bloquear la membrana con PBS 1X-BSA 1.3% a 37°C por 30 min. Luego a temperatura ambiente durante 60 min.
7. Lavado: con PBS 1X filtrado por 10 min, repitiendo 3 veces el lavado.
8. Incubar con una solución de estreptavidina- peroxidasa 1: 500 en PBS-BSA 1.3%
9. Lavar con PBS 1X filtrado por 10 min. Repitiendo 3 veces el lavado. Hacer un último lavado rápido con PBS 1X filtrado.
10. Revelar con DAB (50 mg/100 mL PBS 1X filtrado) precalentado a 30 °C. Adicionar 20 µL de peroxidasa (peróxido 30%).
11. Lavado: 2 o 3 veces con PBS 1X filtrado rápidamente.

Las condiciones de la transferencia semihumeda fueron 25 V, 120 mA y 2 W durante 40 minutos para alcanzar un $V \cdot H = 14$. La tinción con rojo ponceau se dejó durante 10 minutos, se eliminó el colorante en exceso y se documentó.

Anexo 6: Isoelectroenfoque de la LGL-P4 (Bollag y Edelstein, 1994).

El proceso se llevó a cabo en condiciones nativas, empleando anfolitos en un rango de pH 3-10 y para la realización de la curva de calibración se trabajó con patrones de punto isoeléctrico en un rango de pH 3-10 (Pharmacia).

Se utilizó como control de punto isoeléctrico mioglobina 15 μg (5 $\mu\text{g}/\mu\text{l}$) y se sembraron 20 μg de LGL-P4 (Tabla 22).

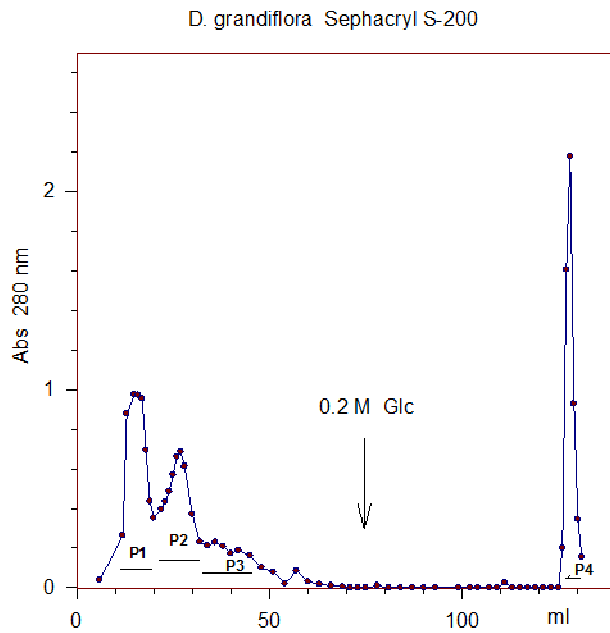
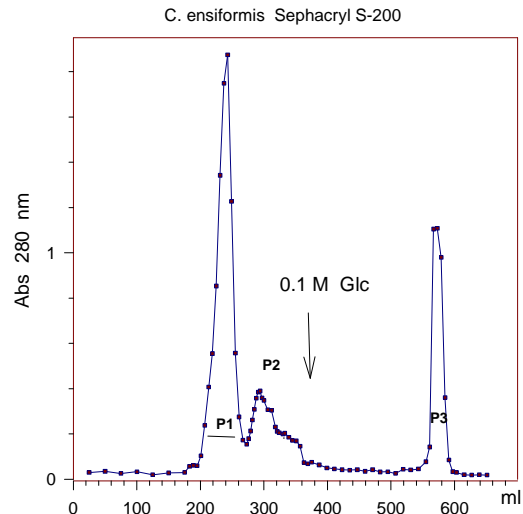
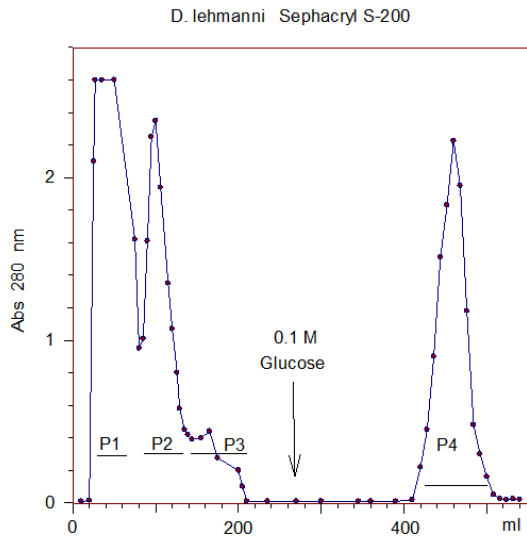
Tabla 22: Preparación del gel para isoelectroenfoque.

Reactivo	Cantidad (mL)
Agua desionizada	8.74
Acilamida (Laemmli) 50%	2.0
Glicerol	1
Anfolitos 3-10	0.250
Persulfato de amonio 10%	0.025
TEMED	0.020

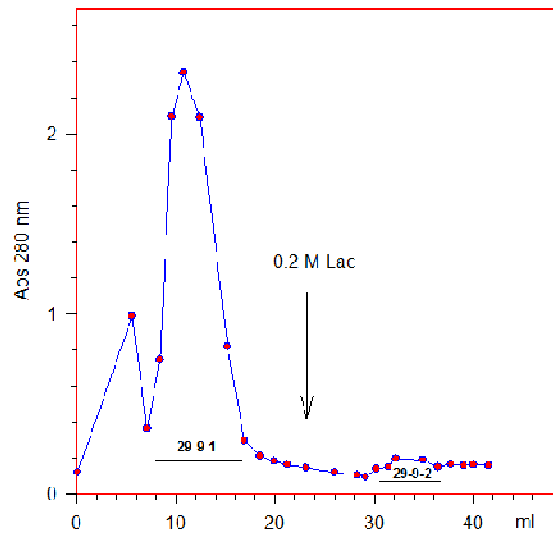
La corrida se lleva a cabo utilizando una solución ácido fosfórico 0.1 M (ánodo) y una NaOH 0.1 M (cátodo). Se hizo una precorrida de 30 min a 500 V para enfocar los anfolitos.

Posteriormente se sembró la muestra en glicerol. Se hizo una corrida a 200 V por 1.5 horas, 220 V por 30 min y 600 V por 2 horas, con un V*H total de 1736. Al se lavó el gel con TCA 10% por 30 minutos y luego se dejó en TCA 1% toda la noche. Finalmente se tiñó con coomassie G-250.

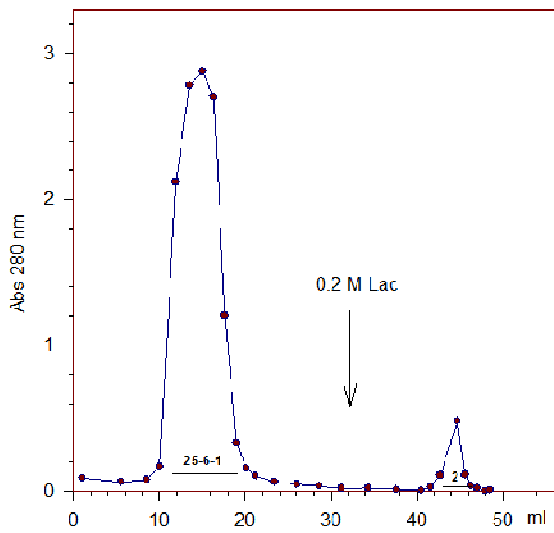
Anexo 7: Perfiles de purificación de las lectinas P4 de *Dioclea lehmanni*, *D. grandiflora* y *Canavalia ensiformis*.



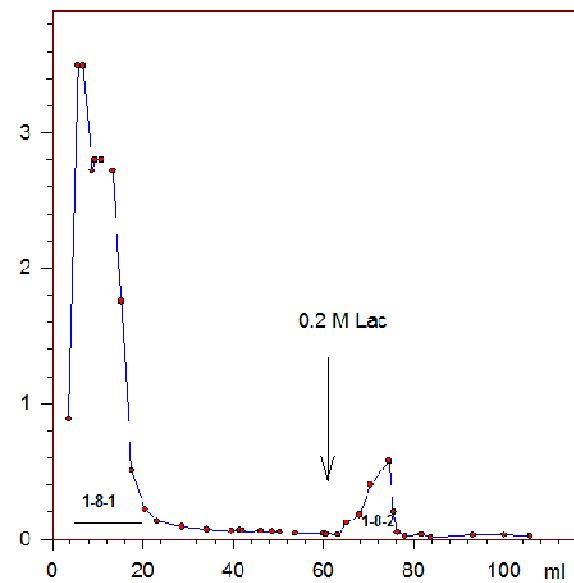
C. ensiformis Lac- Sepharose



D. lehmanni Lac-Sepharosa

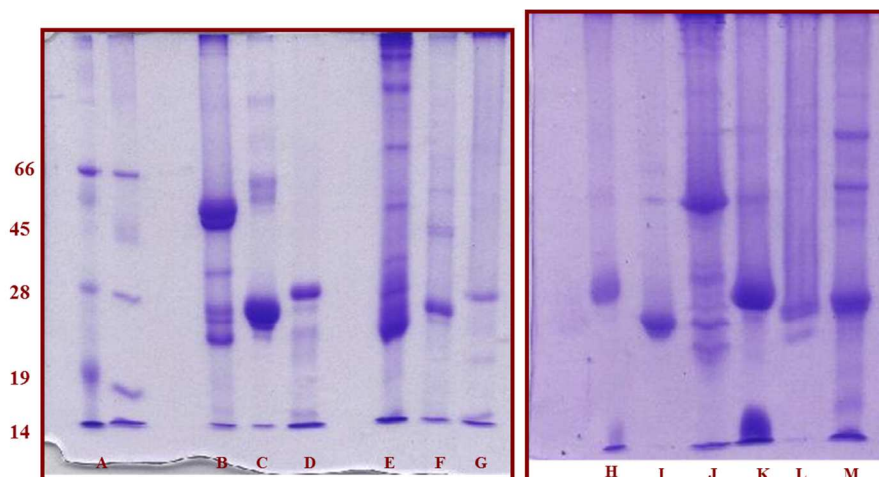


D. grandiflora Lac-Sepharosa



Anexo 8: Electroforesis del proceso de purificación de lectinas P2 y P4.

En el caso de las lectinas tipo P4 para algunas especies se observan bandas de bajo peso molecular que corresponden a la reducción del monómero. Por lo tanto se pueden observar en algunos casos tres bandas. En el caso de las P2 se puede observar el dímero y el tetrámero por falta de disociación de la proteína nativa.



SDS-PAGE: Carril A: Patrones de peso molecular; B: extracto crudo de *Galactia lindenii*; C: LGL-P2 ; D: LGL-P4; E: extracto crudo de *Canavalia ensiformis*; F: LCE-P2; G: LCE-P4 ; J: extracto crudo de *Dioclea lehmanni*; I: LDL-P2; H: LDL-P5; M: extracto crudo de *Dioclea grandiflora*; L: LDG-P2; K: LDG-P4.

Anexo 9: Composición de aminoácidos LGL-P2 (Almanza et al., 2004).

La composición de aminoácidos mostró un alto contenido de Asx, Ile y aminoácidos hidroxilados. Además un alto contenido de Met (que es inusual en lectinas de leguminosas). Los cálculos fueron basados en un Mr=26.064 con un 7% de carbohidratos (Tabla 23).

Tabla 23: Composición aminoácidos LGL-P2.

Aminoácido analizado	LGL-P2
Asx (N-D)	19
Glx (Q-E)	6
Ser (S)	21
Gly (G)	14
His (H)	5
Arg (R)	7
Thr (T)	16
Ala (A)	8
Pro (P)	5
Tyr (Y)	2
Val (V)	13
Met (M) ^a	10
Cys (C) ^a	2
Ile (I)	20
Leu (L)	12
Phe (F)	8
Lys (K)	11
Trp (W) ^b	7

^a. Determinados como MetSO₂ y CySO₃, respectivamente.

^b. Determinado espectrofotométricamente.

Anexo 10: Resultados identificación de la muestra con base en los péptidos obtenidos en digestión con tripsina contra base de datos PD1.3/Mascot protein/peptide.

#	Visible?	Starred?	BioView: Identified Proteins (18) Including 1 Decoy	Accession Number	Molecular Weight	Protein Grouping Ambiguity	Probability Legend	
							Blank_prenun_allsp	LGLP2_15pec_allsp
							over 95%	
							80% to 94%	
							50% to 79%	
							20% to 49%	
							0% to 19%	
1	<input checked="" type="checkbox"/>		keratin, type II cytoskeletal 1 [Homo sapiens]	gi 1193957...	66 kDa			39
2	<input checked="" type="checkbox"/>		keratin, type II cytoskeletal 2 epidermal [Homo sapiens]	gi 47132620	65 kDa	-		40
3	<input checked="" type="checkbox"/>		keratin, type I cytoskeletal 10 [Homo sapiens]	gi 1959728...	59 kDa	-		25
4	<input checked="" type="checkbox"/>		keratin, type I cytoskeletal 9 [Homo sapiens]	gi 55956899	62 kDa			26
5	<input checked="" type="checkbox"/>		trypsinogen precursor [Sus scrofa]	gi 2422538...	26 kDa			4
6	<input checked="" type="checkbox"/>		homerin [Homo sapiens]	gi 57864582	282 kDa			11
7	<input checked="" type="checkbox"/>		PREDICTED: keratin, type II cytoskeletal 5-like [Pan troglodytes]	gi 3328393...	55 kDa	-		10
8	<input checked="" type="checkbox"/>		PREDICTED: keratin, type I cytoskeletal 14-like [Pan troglodytes]	gi 3328475...	52 kDa	-		6
9	<input checked="" type="checkbox"/>		dermcidin preproprotein [Homo sapiens]	gi 1675192...	11 kDa			4
10	<input checked="" type="checkbox"/>		beta-casein precursor [Ovis aries]	gi 5716437...	25 kDa			2
11	<input checked="" type="checkbox"/>		alpha-S1-casein precursor [Bos taurus]	gi 30794348	25 kDa			3
12	<input checked="" type="checkbox"/>		PREDICTED: keratin, type II cytoskeletal 1-like isoform 6 [Maca...	gi 2972624...	65 kDa	-		2
13	<input checked="" type="checkbox"/>		keratin, type II cytoskeletal 6C [Homo sapiens]	gi 1559696...	60 kDa	-		3
14	<input checked="" type="checkbox"/>		salivary acidic proline-rich phosphoprotein 1/2 precursor [Hom...	gi 4826944...	17 kDa	-		2
15	<input checked="" type="checkbox"/>		keratin 16 [Pan troglodytes]	gi 2213076...	51 kDa	-		2
16	<input checked="" type="checkbox"/>		gi 386588560-R	gi 3865885...	?			2
17	<input checked="" type="checkbox"/>		PREDICTED: keratin, type II cytoskeletal 6A [Pan troglodytes]	gi 1146445...	60 kDa			3
18	<input checked="" type="checkbox"/>		PREDICTED: keratin, type II cytoskeletal 6A-like isoform 2 [Mac...	gi 2972624...	60 kDa	★		1

Anexo 12: Resultados identificación de la muestra con base en los péptidos obtenidos en digestión con LysC contra base de datos PD1.3/Mascot protein/peptide.

#	Visible?	Starred?	BioView: Identified Proteins (21) Including 0 Decoys	Accession Number	Molecular Weight	Probability Legend		Protein Grouping Ambiguity	
						Blank_prerun	LGLP2_LysC		
1	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: keratin, type I cytoskeletal 9 [Pan troglodytes]	gi 332847598 (+1)	62 kDa	over 95%		13	
2	<input checked="" type="checkbox"/>	<input type="checkbox"/>	keratin, type II cytoskeletal 1 [Homo sapiens]	gi 119395750	66 kDa	80% to 94%		14	
3	<input checked="" type="checkbox"/>	<input type="checkbox"/>	keratin, type II cytoskeletal 2 epidermal [Homo sapiens]	gi 47132620	65 kDa	50% to 79%		14	
4	<input checked="" type="checkbox"/>	<input type="checkbox"/>	keratin, type I cytoskeletal 10 [Homo sapiens]	gi 195972866 (+2)	59 kDa	20% to 49%		8	
5	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: protein S100-A9 isoform 2 [Pan troglodytes]	gi 55593263 (+5)	13 kDa	0% to 19%		5	
6	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: keratin, type II cytoskeletal 5-like [Pan troglodytes]	gi 332839316 (+5)	55 kDa			8	
7	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: keratin, type I cytoskeletal 14-like [Pan troglodytes]	gi 332847588 (+8)	52 kDa			6	
8	<input checked="" type="checkbox"/>	<input type="checkbox"/>	desmoplakin isoform II [Homo sapiens]	gi 58530842 (+8)	260 kDa			9	
9	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: protein S100-A8 [Pan troglodytes]	gi 332810638 (+11)	11 kDa			5	
10	<input checked="" type="checkbox"/>	<input type="checkbox"/>	keratin, type II cytoskeletal 6A [Homo sapiens]	gi 5031839 (+15)	60 kDa			4	
11	<input checked="" type="checkbox"/>	<input type="checkbox"/>	keratin, type II cytoskeletal 78 [Homo sapiens]	gi 89357932	57 kDa			3	
12	<input checked="" type="checkbox"/>	<input type="checkbox"/>	protein S100-A7 [Homo sapiens]	gi 115298657	11 kDa			3	
13	<input checked="" type="checkbox"/>	<input type="checkbox"/>	annexin A2 [Mus musculus]	gi 6996913 (+27)	39 kDa			2	
14	<input checked="" type="checkbox"/>	<input type="checkbox"/>	cystatin-A [Homo sapiens]	gi 4885165 (+1)	11 kDa			2	
15	<input checked="" type="checkbox"/>	<input type="checkbox"/>	ubiquitin-60S ribosomal protein L40 [Mus musculus]	gi 9845265 (+324)	15 kDa			3	
16	<input checked="" type="checkbox"/>	<input type="checkbox"/>	caspase-14 precursor [Homo sapiens]	gi 6912286	28 kDa			2	
17	<input checked="" type="checkbox"/>	<input type="checkbox"/>	alpha-S2-casein precursor [Bos taurus]	gi 27806963	26 kDa			2	
18	<input checked="" type="checkbox"/>	<input type="checkbox"/>	desmocollin-1 isoform Dsc1b preproprotein [Homo sapiens]	gi 4826702 (+5)	94 kDa			2	
19	<input checked="" type="checkbox"/>	<input type="checkbox"/>	glyceraldehyde-3-phosphate dehydrogenase isoform 1 [Homo s...gi 7669492 (+5)	gi 7669492 (+5)	36 kDa			2	
20	<input checked="" type="checkbox"/>	<input type="checkbox"/>	thioredoxin isoform 1 [Homo sapiens]	gi 50592994 (+9)	12 kDa			2	
21	<input checked="" type="checkbox"/>	<input type="checkbox"/>	PREDICTED: keratin, type I cytoskeletal 15-like [Callithrix jacch... gi 296202888	gi 296202888	36 kDa		★	2	

Anexo 14: PROCHECK statistics: LGL-P2

1. Ramachandran Plot statistics

	No. of residues	%-tage
	-----	-----
Most favoured regions [A,B,L]	128	78.0%**
Additional allowed regions [a,b,l,p]	27	16.5%
Generously allowed regions [~a,~b,~l,~p]	8	4.9%
Disallowed regions [XX]	1	0.6%*
	----	-----
Non-glycine and non-proline residues	164	100.0%
End-residues (excl. Gly and Pro)	2	
Glycine residues	19	
Proline residues	9	

Total number of residues	194	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and *R*-factor no greater than 20.0 a good quality model would be expected to have over 90% in the most favoured regions [A,B,L].

2. G-Factors

Parameter	Score	Average Score
-----	-----	-----
Dihedral angles:-		
Phi-psi distribution	-0.89*	
Chi1-chi2 distribution	-0.45	
Chi1 only	0.07	
Chi3 & chi4	0.51	
Omega	-0.58*	
		-0.52*
		=====
Main-chain covalent forces:-		
Main-chain bond lengths	-0.14	
Main-chain bond angles	-1.26**	
		-0.79*
		=====
OVERALL AVERAGE		-0.59*
		=====

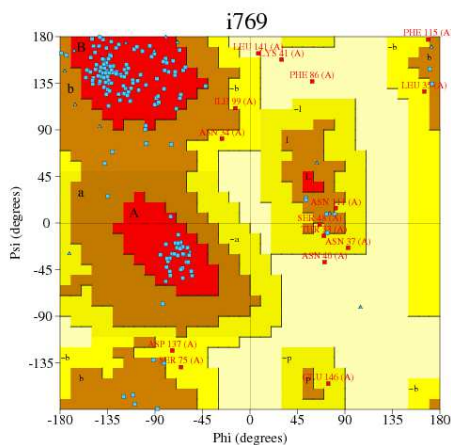
G-factors provide a measure of how unusual, or out-of-the-ordinary, a property is.

Values below -0.5* - unusual

Values below -1.0** - highly unusual

Important note: The main-chain bond-lengths and bond angles are compared with the Engh & Huber (1991) ideal values derived from small-molecule data. Therefore, structures refined using different restraints may show apparently large deviations from normality.

Anexo 15: PROCHECK statistics: LDL-P2.



PROCHECK statistics

1. Ramachandran Plot statistics

	No. of residues	%-tage
Most favoured regions [A,B,L]	140	80.5%*
Additional allowed regions [a,b,l,p]	19	10.9%
Generously allowed regions [-s,-b,-l,-p]	11	6.3%
Disallowed regions [XX]	4	2.3%*
Non-glycine and non-proline residues	174	100.0%
End-residues (excl. Gly and Pro)	2	
Glycine residues	15	
Proline residues	8	
Total number of residues	199	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and *R*-factor no greater than 20.0 a good quality model would be expected to have over 90% in the most favoured regions [A,B,L].

2. G-Factors

Parameter	Score	Average Score
Dihedral angles:-		
Phi-psi distribution	-0.85*	
Chi1-chi2 distribution	-0.34	
Chi1 only	-0.01	
Chi3 & chi4	0.58	
Omega	-0.50	
		-0.47
Main-chain covalent forces:-		
Main-chain bond lengths	-0.22	
Main-chain bond angles	-1.26**	
		-0.92*
OVERALL AVERAGE		-0.58*

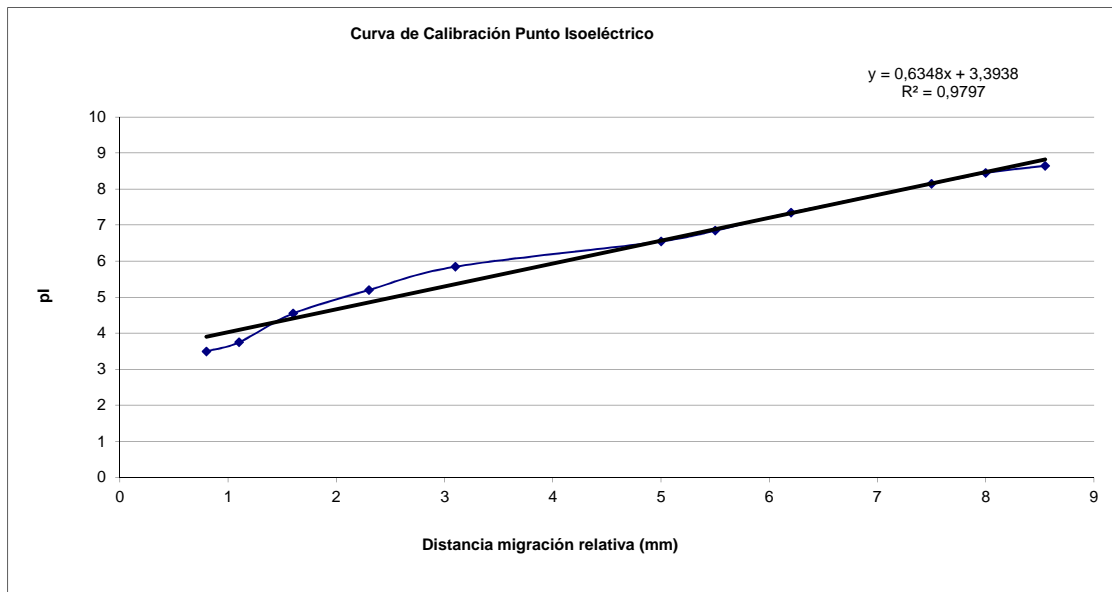
G-factors provide a measure of how unusual, or out-of-the-ordinary, a property is.

Values below -0.5* - unusual
Values below -1.0** - highly unusual

Important note: The main-chain bond-lengths and bond angles are compared with the Engh & Huber (1991) ideal values derived from small-molecule data. Therefore, structures refined using different restraints may show apparently large deviations from normality.

Anexo 17: Curva de calibración pi (Punto isoeléctrico).

La curva de punto isoeléctrico se realizó teniendo en cuenta que se tenía la mioglobina como patrón y que la referencia son las dos bandas predominantes de están a pH 6.85 y 7.35



9. Bibliografía

1. Ainouz I., Moreira R., Campos F., Richardson M., Begbie R., Stewart J., Watt W., Pusztai A. 1987. "The isolation and aminoacid sequence of the β and γ subunits of the lectin from the seeds of *Dioclea grandiflora*". **Phytochemistry** 26:1435-1440.
2. Almanza M. 1999." Purificación y Caracterización parcial de la lectina presente en las semillas de *Galactia lindenii*". Tesis Departamento de Farmacia. Universidad Nacional de Colombia.
3. Almanza M., Vega N., Pérez G. 2004."Isolating and characterizing a lectin from *Galactia lindenii* seeds that recognises blood group H determinants". **Archives of biochemistry and biophysics**. 429: 180-190.
4. Almeida J.L, Fernandes do Nascimento A.S., Cajazeiras J.B., Silva A.C., Pereira R.I, Lopes de Sousa B., Coelho da Silva A.L., Garcia W., Holanda Teixeira E., do Nascimento K.S., da Rocha B.A.M., Nagano C.S., Sampaio A.H., Cavada B.S. 2011. "Molecular Characterization and Tandem Mass Spectrometry of the Lectin Extracted from the Seeds of *Dioclea sclerocarpa* Ducke". **Molecules**. 16: 9077-9089.
5. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. 1990. "Basic local alignment search tool". **J. Mol. Biol.** 215: 403–410.
6. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, and Stockinger H. 2012. "ExpPASy: SIB bioinformatics resource portal". **Nucleic Acids Res.** 40(W1):W597-W603.

7. Azevedo-Moreira R., Franca-Cordeiro E., Sousa-Cavada B., Paula-Nunes E., Gomes A., Abreu J.T. 1995. "Lectins and the chemotaxonomy of the sub-tribe Diocleinae (*Leguminosae-Phaseoleae*)". **R. Bras. Fisiol. Veg.** 7:7-14.
8. Azevedo Moreira R., Lima I., Abreu J.T., Souza B. 1991. "Plant lectins, chemical and biological aspects". **Mem. Inst: Oswaldo Cruz.** 86: 211-218.
9. Banerjee R., Das K., Ravishankar R., Suguna K., Surolia A., Vijayan M. 1996. "Conformation, Protein-Carbohydrate Interactions and a Novel Subunit Association in the Refined Structure of Peanut Lectin-Lactose Complex". **J. Mol. Biol.** 259: 281–296.
10. Beyra A., Herrera P., Reyes G., Hernández L. 2005. "Revisión taxonómica del género *Galactia* P. BR (*Leguminosae-Papilionoideae*) en Cuba". **Rev.Acad. Colomb. Cienc.** 113: 467-494.
11. Bezerra B.M.J., Velloso Fontenelle N., de Freitas A., Bezerra G.A., Bezerra C., de Lima Alencara K.L., Gomes Soares P.M., do Nascimento K.S., Nagano C.S., Martins J.L., Gruber K., Holanda Sampaio A., Delatorre P., Rocha B.A.M., Sampaio Assreuy A.M., Cavada B.S. 2013. "Crystal structure of *Dioclea violacea* lectin and a comparative study of vasorelaxant properties with *Dioclea rostrata* lectin". **The International Journal of Biochemistry & Cell Biology.** 45: 807– 815.
12. Bollag D., Edelstein S. 1994. Protein methods (4a. ed.). New York. **Wiley-Liss Inc.**, pp. 162-171.
13. Bowles D. J. y Pappin D.J, 1988. "Posttranslational processing of concanavalin A precursors in jackbean cotyledons". **Journal of Cell Biology.** 102: 1284-1297, 1986.
14. Calvete J.J., Trole H.H., Raida M., Urbanke C., Romero A., Grangeiro T.B., Ramos M.V., Almeida da Rocha M.I., Guimara F.N., Cavada B.S. 1999. "Molecular characterization and crystallization of *Diocleinae* lectins". **Biochimica et Biophysica Acta.** 1430: 367-375.

15. Cavada B., Barbosa T., Arruda S., Grangeiro T., Barral-Netto. 2001. "Revisiting proteus: Do minor changes in lectin structure matter in Biological Activity? Lessons from and potential Biotechnological Uses of the *Diocleinae* Subtribu lectins. **Current Protein and Peptide Science**. 2: 123-135.
16. Cavada B., Marinho E., Souza E., Benevides R., Delatorre P., Souza L., Nascimento K., Sampaio A., Moreno F., Rustigel J., Canduri F., Azevedo W., Debray H. 2006. "Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from *Cymbosema roseum* sedes". **Acta Cryst.** F62, 235–237.
17. Correia M., Coelho L. 1995. "Purification of a Glucose/Mannose specific lectin, isoform 1, from seeds of *Cratylia mollis* Mart (Camaratu Bean)". **Applied Biochemistry and Biotechnology**. 55: 262-273.
18. Dam T.K., Cavada B.S., Nagano C.S., Rocha B.A.M, Benevides R.G., Nascimento K.S., de Sousa L.A.G., Oscarson S., Brewer C.F. 2011. "Fine specificities of two lectins from *Cymbosema roseum* seeds: a lectin specific for high-mannose oligosaccharides and a lectin specific for blood group H type II trisaccharide". **Glycobiology**. Vol. 21 (7): 925 – 933.
19. Dam TK, Cavada BS, Grangeiro TB, Santos CF, de Sousa FAM, Oscarson S, Brewer CF. 1998. "*Diocleinae* lectins are a group of proteins with conserved binding sites for the core trimannoside of asparagine-linked oligosaccharides and differential specificities for complex carbohydrates. **J Biol Chem**. 273:12082–12088.
20. da Nóbrega R.B., Rocha B.A.M., Gadelha C.A.A., Santi-Gadelha T., Pires A.F., Assreuy A.M.S., Nascimento K.S., Nagano C.S., Sampaio A.H., Cavada B.S., Delatorre P. 2012. "Structure of *Dioclea virgata* lectin: Relations between carbohydrate binding site and nitric oxide production". **Biochimie**. 94: 900-906.
21. Darville L.N.F., Merchant M.E., Maccha V., Siddavarapu V.R., Hasan A., Murray K.K. 2012. "Isolation and determination of the primary structure of a lectin protein from the serum of the American alligator (*Alligator*

- mississippiensis*)". **Comparative Biochemistry and Physiology**, Part B. 161: 161-169.
22. Delatorre P., Nascimento K.S., Melo L.M., de Souza E.P., da Rocha B.A.M., Benavides R.G., de Oliveira T.M., Bezerra G.A., Bezerra M.J.B., da Cunha R.M.S., da Cunha F.A.B., Freired V.N., Cavada B.S. 2006. "Crystallization and preliminary X-ray diffraction analysis of the lectin from *Dioclea rostrata* Benth Seeds". **Acta Cryst.** F62: 166–168.
23. Delatorre P., Rocha B.A.M., Simões R.C., Pereira-Júnior F.N. Silva H.C., Bezerra E.H.S., Bezerra M.J.B., Marinho E.S., Gadelha C.A.A., Santi-Gadelha T., Farias D.L., Assrey A.M.S., Marques-Domingos G.F.O., Nagano C.S., Cavada B.S. 2011. "Mass Spectrometry and X-ray Diffraction Analysis of Two Crystal Types of *Dioclea virgata* Lectin: An Antinociceptive Protein Candidate to Structure/Function Analysis". **Appl Biochem Biotechnol.** 164:741–754.
24. de Oliveira T.M., Delatorre P., da Rocha B.A.M., de Souza E.P., Nascimento K.S., Bezerra G.A., Moura Tales R., Benevides R.G., Bezerra E.H.S., Moreno F.B.M.B., Freire V.N., de Azevedo Jr W.F., Cavada B.S. 2008. "Crystal structure of *Dioclea rostrata* lectin: Insights into understanding the pH-dependent dimer-tetramer equilibrium and the structural basis for carbohydrate recognition in *Diocleinae* lectins." **Journal of Structural Biology.** 164: 177–182.
25. Elgavish S., Shaanan B. 1997. "Lectin-Carbohydrate interactions: Different folds, common recognition principles". **TIBS.** 22: 462-467.
26. Equipo Electroforesis Preparativa. Model 491 Prep Cell. Instruction manual. BIORAD.
27. Filgueira J.J. 2002. "Purificación, caracterización y análisis estructural preliminar de las lectinas presentes en *Salvia palifolia* e *Hyptis mutabilis*". Tesis de Doctorado Departamento de Química. Universidad Nacional de Colombia.
28. Fountoulakis M., Juranville J.F. 2003. "Enrichment of low-abundance brain proteins by preparative electrophoresis". **Analytical Biochemistry.** 313: 267–282.

29. Fujii Y., Kawsar S.M.A., Matsumoto R., Yasumitsu H., Ishizaki N., Dogasaki C., Hosono M., Nitta K., Hamako J., Taei M., Ozeki Y. 2011. "A D-galactose-binding lectin purified from coronate moon turban, *Turbo (Lunella) coreensis*, with a unique amino acid sequence and the ability to recognize lacto-series glycosphingolipids". **Comparative Biochemistry and Physiology, Part B**. 158: 30–37.
30. Fukuda N., Yoshimaru A., Hidaka T., Ohta. Yamamoto, K., Yomo H. 1994. "Isolation and characterization of N-acetylgalactosamine-specific lectin from *Galactia Tashiroi* seeds". **Biosci, Biotech, Biochem**. 58 (2): 423-424.
31. Gabius H-J, Andre S, Kaltner H, Siebert H-C. 2002. "The sugar code: functional lectinomics" **Biochim.Biophys. Acta**. 1572: 165-177.
32. Garzon R. 1999. "Cristalización de la lectina P2 de la *Dioclea lehmanni*". Tesis de pregrado. Departamento de Farmacia. Universidad Nacional de Colombia.
33. Ghazarian H., Idoni B., Oppenheimer S. 2011. "A glycobiology review: Carbohydrates, lectins and implications in cancer therapeutics". **Acta histochemica**.113: 236–247.
34. Goldstein J., Poretz R. D. 1986. "Isolation, Physicochemical Characterization, and Carbohydrate-Binding Specificity of Lectins". En: "The Lectins. Properties, Functions, and Applications in Biology and Medicine": Liener, I. E., Sharon, N., Goldstein, I. J. (eds). **Academic Press Inc.**,Orlando. Pp. 51-67.
35. Grangeiro T., Gatehouse J., Pereira M., Cavada B. 1997. "Investigation on the origin of the naturally occurring fragments of *Cratylia floribunda* lectin". **R. Bras.Fisiol.Veg**.9: 9-13.
36. Hamelryck T. 1999. "Structural studies of four members of the legume lectin family: Phytohemagglutinin-L, Arcelin-5, *Dolichos biflorus* seed lectin and *Dolichos biflorus* stem and leaf lectin Laboratorium voor Ultrastructuur". Vrije Universiteit Brussel (VUB).
37. Hernández P., Pérez E., Martínez L., Ortiz B., Martínez G. 2005. "Las lectinas vegetales como modelo de estudio de las interacciones proteína-carbohidrato". **REB**. 24. 1: 21-27.

38. Herman E. M., Shannon L.M., Chrispeels M.J. 1985. "Concanavalin A is synthesized as a glycoprotein precursor". **Planta** **165**: 23-29.
39. Herrmann, J., Turckll, C.W., Atchison, R.E., Huflejt, M.E., Poulter, L., Gitt, M.A., Burlingame, A., Barondes, S.H., Leffler, H. 1993. "Primary Structure of the Soluble Lactose Binding Lectin L-29 from Rat and Dog and Interaction of Its Non-collagenous Proline-, Glycine-, Tyrosine-rich Sequence with Bacterial and Tissue Collagenase". **The journal of biological chemistry**. 268 (35): 26704-26711.
40. Hermanson G.; Mallia A. y Smith P. 1992. "Immobilized affinity ligand techniques". **Academic Press, INC.**, San Diego, págs. 53-54, 74-75, 88, 253-254.
41. Higgins T., Chandler P., Zurawski G., Butto S., Spencer D. 1983. "The biosynthesis and primary structure of pea seed lectin". **The journal of biological chemistry**. 258 (15): 9544-9549.
42. Hirabayashi J. 2003. "Oligosaccharide microarrays for glycomics". **Trends in Biotechnology**. 21: 141-143.
43. Jiang S-Y., Ma Z., Ramachandran S. 2010. "Evolutionary history and stress regulation of the lectin superfamily in higher plants". **BMC Evolutionary Biology**. 10:79.
44. Jimenez S. 2000. " Obtención y caracterización parcial de las subunidades de la lectina Tn específica de *Salvia bogotensis*". Tesis Departamento de Química. Universidad Nacional de Colombia.
45. Konami Y., Yamamoto K., Osawa T. 1990. "The primary structure of the *Lotus tetragonolobus* seed lectin". **Federation of European Biochemical Societies letters**. 268 (1): 281-286.
46. Konami Y., Yamamoto K., Toyoshima S., Osawa, T. 1991. "The primary structure of the *Laburnum alpinum* seed lectin". **Federation of European Biochemical societies letters**. 286 (12): 33-38.
47. Laemmli U.K. 1970. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". **Nature**. 227: 680-685.

48. Lee Y.M., Leiby K.R., Allar J., Paris K., Lerch B., Okarma T. B. 1991. "Primary Structure of Bovine Conglutinin, a Member of the C-type Animal Lectin Family." *The journal of biological chemistry*. 266 (5): 2715-2723.
49. LePendou J., Gérard G., Lambert F., Mollicone R., Oriol R. 1986. "A new anti-H lectin from the seeds of *Galactia tenuiflora*". **Glycoconjugate Journal**. 3 (2): 203-216.
50. Lis H. y Sharon N. 1986. "Lectins as molecules and as tools". **Annual Review of Biochemistry**. 55: 35-67.
51. Martínez P. 2009. "Estudio de las interacciones entre la vicilina y las lectinas ConA y CEL-II de la semilla de *Canavalia ensiformis*". Tesis de Doctorado. Facultad de Ciencias .Universidad Nacional de Colombia.
52. Matsudaira P. 1987. "Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes". **JBid Chem**. 262: 10035-10038.
53. McPherson A. 1990. "Current approaches to macromolecular crystallization". **Eur J Biochem**. 189(1):1-23.
54. Melgarejo L.M.; Pérez G. 1997. "Immunolocalization of the lectins P2 and P4 from *Dioclea lehmanii* seeds". **Plant Cell Physiol**. 38(4):480-483.
55. Melgarejo L.M.; Vega N. y Pérez G. 2005. "Isolation and characterization of novel lectins from *Canavalia ensiformis* DC and *Dioclea grandiflora* Mart. Ex Benth. Seeds". **Brazilian Journal of Plant Physiology**. 17: 315-324.
56. Moreira R., Barros A., Stewart J., Pustai A. 1983. "Isolation and characterization of a lectin from seeds of *Dioclea grandiflora* (Mart)". **Planta** 158: 63-69.
57. Nagano C.S., Calvete J.J., Baretino D., Pérez A., Cavada B.S., Sanz L. 2008. "Insights into the structural basis of the pH-dependent dimer-tetramer equilibrium through crystallographic analysis of recombinant *Diocleinae* lectins". **Biochem. J**. 409: 417-428.
58. Nakano M., Tabata S., Sugihara K., Kouzuma Y., Kimura M., Yamasaki, N. 1999. "Primary structure of hemolytic lectin CEL-III from marine invertebrate

- Cucumaria echinata* and its cDNA: structural similarity to the B-chain from plant lectin, ricin". **Biochimica et Biophysica Acta.** 1435: 167-176.
59. Navarro Y., Pérez G. 1978. "Detección y caracterización preliminar de lectinas presentes en semillas de leguminosas". **Rev. Col. Quím.**8:25-43.
60. Paroutaud P., Levi G., Teichberg V.I., Strosberg A.D. 1987. "Extensive amino acid sequence homologies between animal lectins". **Proc. Natl. Acad. Sci.** 84: 6345-6348.
61. Pérez G., Hernández M., Mora E. 1990. "A lectin from the seeds of *Dioclea lehmannii*". **Phytochemistry.** 29: 1745-1749.
62. Pérez G. 1993. "*Erythrina* lectins: Structural and physico-chemical properties: A review". **Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.** 18: 545-553.
63. Pérez G. 1998. "Isolation and characterization of a novel lectin from *Dioclea lehmannii* (Fabaceae) seeds". **The International Journal of Biochemistry and Cell Biology.** 30:843-853.
64. Pérez G. y Vega N. 2007. "*Lamiaceae* lectins". **Functional Plant Science and Biotechnology.** 1(2): 288-299.
65. Peumans W., Van Damme E.J.M. 1998. "Plant lectins: versatile proteins with important perspectives in biotechnology". **Biotechnology and genetic engineering reviews.** 15:199-228.
66. Pusztai A. 1991. "Plant lectins". **Cambridge University Press,** Cambridge, pp.4-8, 35-40, 74-75.
67. Quintero M., Reyes Montaña E., Pérez G., Vega N. 2013. "Expression of two types *Diocleinae* Subtribe Lectins". **17th European Carbohydrate Symposium** (Eurocarb 17) Tel-Avid (Israel). Abstracts book pag 67.
68. Rameshwaram N., Kumar N., Scharf C., Völker U., Nadimpalli S.K. 2009. "Complete primary structure of a newly characterized galactose-specific lectin from the seeds of *Dolichos lablab*". **Glycoconjugate Journal.** 26:161-172.
69. Richardson M, Campos F.D.A.P., Moreira R.A., Ainouz I.L., Begbie R., Watt W.B., Pusztai A. 1984. "The complete amino acid sequence of the major a

- subunit of the lectin from the seeds of *Dioclea grandiflora* (Mart)". **Eur. J. Biochem.** 144: 101-111.
70. Rini, J. 1995. "Lectin Structure". **Annual Review of Biochemistry.** 24: 551-577.
71. Rocha B., Moreno F., Delatorre P., Souza E., Marinho E., Benevides R., Rodrigues J., Souza L., Nagano C., Debray h., Sampaio A., De azevedo W., Cavada B. 2009. "Purification, Characterization, and Preliminary X-Ray Diffraction Analysis of a Lactose-Specific Lectin from *Cymbosema roseum* Seeds". **Appl Biochem Biotechnol** 152:383–393.
72. Rocha B.A.M., Nagano C.S., Delatorre P., Calvete J.J., Cavada B.S. 2010, "Primary sequencing of CRLII by Edman's degradation and tandem mass spectrometry". UniProtKB. .
73. Rojas C., Pérez G. 1990. "Preliminary Crystallographic study of the *Erythrina rubrinervia* lectin". **J.Mol. Biol.** 214: 21-22.
74. Rojas A. 1998. "Determinación por difracción de rayos X de la estructura tridimensional de la lectina P4 de la *Dioclea lehmanni*". Tesis de pregrado. Departamento de Física. Universidad Nacional de Colombia.
75. Roy A., Kucukural A, Zhang Y. 2010. "I-TASSER: a unified platform for automated protein structure and function prediction". **Nat. Protoc.** 5:725-738.
76. Roy A., Zhang Y. 2012. "Recognizing Protein-Ligand Binding Sites by Global Structural Alignment and Local Geometry Refinement". **Structure.** 20 (6): 987-997.
77. Saito T., Kawabata S., Hirata M., Iwanaga S. 1995. "A novel type of limulus lectin L-6: Purification, primary structure and antibacterial activity". **The journal of biological chemistry.** 270 (24): 14493-14499.
78. Sánchez M. 2003. "Estandarización de las condiciones de separación de las subunidades que constituyen la lectina presente en las semillas de *Salvia bogotensis*". Tesis Pregrado en Química. Universidad Nacional de Colombia. Bogotá.

79. Sanabria E. 1999. "Aislamiento, purificación y caracterización parcial de la lectina presente en las semillas de *Lepechinia bullata*". Tesis Departamento de Química. Universidad Nacional de Colombia.
80. Schägger H. Von Jagow G. 1987. "Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa". **Analytical biochemistry**. 166: 368-379.
81. Schnell D.J., Etzler M.E. 1981. "Primary structure of the *Dolichos biflorus* seed lectin". **The journal of biological chemistry**. 262 (15): 7220-7225.
82. Sharon N. y Lis H. 2001. "The structural basis for carbohydrate recognition by lectins. In molecular Immunology of Complex Carbohydrates". **Wu A.M. (ed). Academic Press**, págs. 1-16.
83. Sharon, N., Lis, H. 2004. "History of lectins: from hemagglutinins to biological recognition molecules". **Glycobiology**. 14: 53-62.
84. Sharon N., Lis H. 2007. "Lectins". **Springer, Second edition**. pp.1-26.
85. Shevchenko A., Wilm M., Vorm O., Mann M. 1996. "Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels". **Anal. Chem.** 68: 850-858.
86. Shevchenko A., Tomas H., Havlis J., Olsen J.V., Mann M. 2006. "In-gel digestion for mass spectrometric characterization of proteins and proteomes". **Nature Protocols**. 1 (6): 2856-2860.
87. Sierra A.J., Pérez G. 1999. "Extracción, purificación y caracterización de dos lectinas en semillas de *Dioclea serícea*". **REV.ACAD.COLOMB.CIENC. VOL XXIII**. 88: 446-454.
88. Smith P., Krohn R., Hermanson G., Mallia A., Gartner F., Provenzano M., Fujimoto E., Goeke N., Olson B., Klenk D. 1985. "Measurement of protein using Bicinchoninic Acid". **Anal Biochem**. 150: 76-85.
89. Svensson C., Teneberg S., Nilsson C.L., Kjellberg A., Schwarz F.P., Sharon N., Krenzel U. 2002. "High-resolution Crystal Structures of *Erythrina cristagalli* Lectin in Complex with Lactose and 2'- α -L-Fucosyllactose and Correlation with Thermodynamic Binding Data". **J. Mol. Biol.** 321: 69-83.

90. Tavares G., Caracelli I., Burger M., Correia M., Coelho L., Oliva G. 1996. "Crystallization and preliminary X-ray studies on the lectin from seeds of *Cratylia mollis*". **Acta Cryst.** DS52 1046-1047.
91. Van Damme, E.J.M., Peumans, W., Pusztai, A. 1998. "Handbook of plant lectins: properties and biomedical applications". **Wiley VCH.** pp. 1-13.
92. Van Driessche E., Beeckmans S., Dejagere R., Kanarek L. (1983). "Thiourea: The antioxidant of choice for the purification of proteins from phenol-rich plant tissues". **Anal. Biochem.** 141: 184-188.
93. Varela E.S., Lima J.P.M.S., Galdino A.S., Pinto L., Becerra W.M., Nunes E.P., Alves M.A.O., Grangeiro T.B. 2004. "Relationships in subtribe *Diocleinae* (*Leguminosae*; *Papilionoideae*) inferred from internal transcribed spacer sequences from nuclear ribosomal DNA". **Phytochemistry.** 65: 59-69.
94. Vega N. 1997. "Detección, purificación y caracterización parcial de la lectina presente en las semillas del género *Salvia*" Tesis Departamento de Química. Universidad Nacional de Colombia.
95. Vega N. 2004 "Caracterización bioquímica, funcional y biológica de la lectina de *Salvia bogotensis* y evaluación de su aplicación para la detección del antígeno Tn". Tesis de Doctorado. Facultad de Ciencias. Universidad Nacional de Colombia.
96. Wah D.A., Romero A., Gallego del Sol F., Cavada B.S., Ramos M.V., Grangeiro T.B., Sampaio A.H., Calvete J.J. 2001. "Crystal Structure of Native and Cd/Cd-substituted *Dioclea guianensis* Seed Lectin. A Novel Manganese-binding Site and Structural Basis of Dimertetramer Association". **J. Mol. Biol.** 310: 885-894.
97. Web: NCBI-BLAST help, <http://www.ncbi.nlm.nih.gov/blast/Why.shtml>.
98. Wright C.S. 1997. "New folds of plant lectins". **Current Opinion in Structural Biology.** 7:631-636.