



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Posgrado en Ciencias Bioquímicas  
Instituto de Química

**PRODUCCIÓN RECOMBINANTE Y CARACTERIZACIÓN BIOFÍSICA DE LA AGLUTININA DE GERMEN DE TRIGO Y SUS CUATRO DOMINIOS AISLADOS EN BÚSQUEDA DE ELUCIDAR LAS BASES DE SU ESTABILIDAD ESTRUCTURAL Y RECONOCIMIENTO DE LIGANDOS.**

## TESIS

QUE PARA OPTAR POR EL GRADO DE:  
**DOCTOR EN CIENCIAS**

PRESENTA:

**M. en C. Eduardo Leyva Hernández**

TUTOR PRINCIPAL:

**Dr. Enrique García Hernández**  
Instituto de Química

MIEMBROS DEL COMITÉ TUTOR:

**Dr. Miguel Antonio Costas Basín**  
Facultad de Química, UNAM  
**Dra. Nuria Victoria Sánchez Puig**  
Instituto de Química, UNAM

Ciudad Universitaria, Cd. Mx., Agosto, 2019



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

La presente tesis fue realizada en el Departamento de Química de Biomacromoléculas del Instituto de Química de la Universidad Nacional Autónoma de México, bajo la dirección del Dr. Enrique García Hernández. Durante la realización de este trabajo, se contó con el apoyo económico de la Dirección General de Asuntos del Personal Académico (DGAPA) a través del proyecto PAPIIT IN205018 y del Consejo Nacional de Ciencia y Tecnología (CONACyT) proyecto 235831. Se agradece el apoyo de las técnico-académicas Dra. Patricia Cano Sánchez, Q. Ma. del Rocío Patiño Maya, M. C. Lucía del Carmen Márquez Alonso, M. C. Lucero Ríos Ruiz y Q. Eréndira García Ríos del Instituto de Química, UNAM. Eduardo Leyva Hernández recibió la beca no. 176935 otorgada por CONACyT para la realización de estudios de doctorado.

De esta tesis derivó el siguiente artículo de investigación:

Leyva, E., Medrano-Cerano, J. L., Cano-Sánchez, P., López-González, I., Gómez-Velasco, H., del Río-Portilla, F. y García-Hernández, E. (2019) "Bacterial expression, purification and biophysical characterization of wheat germ agglutinin and its four hevein-like domains", *Biopolymers*, 110(1), p. e23242. doi: 10.1002/bip.23242.

## AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México (UNAM) y al Programa de Maestría y Doctorado en Ciencias Bioquímicas.

Al Dr. Enrique García Hernández, por creer en mí en un momento crítico y permitirme integrar a su grupo de investigación; por todo el esfuerzo que implicó la dirección de este proyecto y por darme la libertad de seguir con mi labor docente a la par del doctorado. Por seguir compartiéndome sus conocimientos, así como su confianza y amistad.

A la Dra. Patricia Cano Sánchez, por ser el otro pilar que hizo posible el desarrollo de este trabajo de investigación, así como por su dirección e invaluable apoyo durante todo este tiempo.

A los miembros de mi comité tutor, el Dr. Miguel Costas Basín y a la Dra. Nuria Victoria Sánchez Puig, por sus valiosas aportaciones durante el curso de mi doctorado.

A los miembros del jurado, el Dr. Rogelio Rodríguez Sotres, la Dra. Lilian González Segura, el Dr. Alejandro Sosa Peinado, la Dra. Amelia Farrés González y el Dr. Horacio Reyes Vivas. Por el tiempo y el esfuerzo invertido en mejorar el presente trabajo, el cual representa una gran contribución a mi formación.

Al Dr. José Federico del Río Portilla, por su importante colaboración y por todos sus aportes en el apartado de Resonancia Magnética Nuclear.

A la M. en C. Lucía del Carmen Márquez Alonso, a la M. en C. Lucero Mayra Ríos Ruíz y a la Q. Eréndira García Ríos del Laboratorio de Espectrometría de Masas del Instituto de Química de la UNAM y a la Q.F.B. Ma. del Rocío Patiño Maya del Laboratorio de Espectroscopía y Polarimetría del Instituto de Química de la UNAM, por toda la confianza y el apoyo técnico que me brindaron.

A Jorge L. Medrano Cerano, quien también ha estado involucrado en la investigación de esta proteína y ha sido una pieza clave en el avance del proyecto. Por toda su ayuda en la parte experimental, por su trabajo y, sobre todo, por compartirme su amistad.

A Homero Gómez Velasco, porque además de su amistad, siempre tuvo el tiempo y la disposición para brindarme su valiosa ayuda, tanto en lo laboral como en lo personal.

A Itzel López González y a Ana Anzures Mendoza, es un orgullo haberlas tenido como alumnas de licenciatura y agradezco todo lo que aportaron con su esfuerzo.

A Luis Fernando Cofas Vargas y a Luis Pablo Ávila Barrientos, por ser unos grandes amigos y brindarme su ayuda de forma incesante.

A Manuel García Curiel por darme ánimos y por compartirme su visión y camaradería.

A mis otros camaradas de laboratorio, Beatriz A. Baquero, Alfonso Labra, Paola Mendoza, Alan Juárez, Alinne López e Israel Mares.

A la Dra. Alicia Cea Bonilla, a la M. en C. Deyamira Matuz Mares, a la M. en C. Laura Iliana Álvarez Añorve y al Dr. Gerardo Hernández Puga de la Facultad de Medicina de la UNAM. Quienes me brindaron la oportunidad de incorporarme a la Facultad, por todo su apoyo durante la ejecución de mi labor docente y en particular, por ayudarme a reencontrar el goce de esta tarea, la cual ha sido parte fundamental de mi formación.

A Leticia García Gutierrez, a Adelina González Pérez y a Julio I. Palacios Ordoñez del Posgrado de Ciencias Bioquímicas, por auxiliarme con todos los trámites necesarios durante mi doctorado.

A Valeria Guzmán Luna y Lorena de Anda Aguilar, porque siempre están ahí para ayudarme o darme consejo. C

A Adriana Montiel Leyva, Gustavo Montiel Leyva, David Misael Mercado Peña, Fermín David Cazares Cañal, Etzel Huízar Sánchez, Julio César González Laurrabaquio, Huilver Nolasco Aguilar y Rubén Daniel González Cervantes. Gracias por continuar siendo tan importantes en mi vida.

A Mariana Castro Ramos, por brindarme la palabra y tener la disposición de escucharme. Por compartir mi alegría, mi angustia y darme siempre su opinión.

A mi padre, Emilio Leyva Méndez, por estar ahí para apoyarme cuando lo he necesitado.

A mi madre María Mónica Hernández Medina y a mi abuela Victoria Hernández Medina, porque fueron las dos grandes artífices que, con su esfuerzo, se encargaron de forjar esta oportunidad para mí.

# ÍNDICE

Abreviaturas.....	1
Resumen .....	2
Abstract.....	3
Introducción.....	4
La relevancia del desarrollo de la glicómica .....	4
Lectinas, definición, distribución e importancia para la vida .....	7
Historia de las lectinas .....	9
La diversidad estructural de las lectinas .....	11
Termodinámica de la unión entre proteína y ligando .....	13
Bases moleculares de la formación de los complejos lectina-carbohidrato.....	17
Lectinas, aplicaciones y potencial biotecnológico .....	19
Aglutinina de Germen de Trigo (AGT): Estructura y unión a carbohidratos.....	22
Antecedentes.....	25
Planteamiento del problema y justificación del proyecto.....	30
Hipótesis .....	32
Objetivos .....	33
Objetivo general.....	33
Objetivos específicos .....	33
Materiales y métodos .....	34
Purificación de la AGT silvestre.....	34
Clonación de los genes de la AGT y de sus cuatro monodominios.....	34
Expresión y purificación de las proteínas recombinantes .....	36
Espectroscopia de dicroísmo circular .....	37
Calorimetría diferencial de barrido.....	38
Experimentos de Resonancia Magnética Nuclear .....	38
Calorimetría de titulación isotérmica .....	39
Ensayos de hemaglutinación .....	40
Resultados.....	41
Purificación de la AGT nativa .....	41
Clonación y expresión de la AGT recombinante.....	42
Purificación de la AGT recombinante .....	46

Purificación de los dominios recombinantes de la AGT.....	48
Análisis de espectrometría de masas MALDI-TOF .....	49
Caracterización biofísica de la aglutinina recombinante completa.....	51
Caracterización de los monodominios recombinantes. ....	56
Discusión .....	61
Conclusiones .....	64
Perspectivas .....	66
Referencias.....	67
Anexo .....	73

## ABREVIATURAS

AGT	Aglutinina de germen de trigo	P	Proteína
°C	Grados celsius	pb	Pares de bases
cm	Centímetro	PBS	Solución salina amortiguada por fosfatos
CNBr	Bromuro de cianógeno	PCR	Reacción en cadena de la polimerasa
Con A	Concanavalina A	PDB	Protein Data Bank
$\Delta C_p U$	Cambio en la capacidad calorífica de unión	PHA	Fitohemaglutinina
CRM	Curcumina	pl	Punto isoeléctrico
CSP	Parche de oligosacáridos agrupados	P · L	Complejo proteína-ligando
Da	Daltones	PMSF	Fluoruro de fenilmetilsulfonilo
DC	Dicroísmo circular	R	Constante universal de los gases ideales
dmol	Decimol	RNA	Ácido ribonucleico
DRC	Dominio de reconocimiento a carbohidrato	rAGT	Aglutinina de germen de trigo recombinante
DSC	Calorimetría diferencial de barrido	rAGT.A	Monodominio A recombinante de la AGT
EGFR	Receptor del factor de crecimiento epidérmico	rAGT.B	Monodominio B recombinante de la AGT
<i>E. coli</i>	<i>Escherichia coli</i>	rAGT.C	Monodominio C recombinante de la AGT
EK	Enterocinasa	rAGT.D	Monodominio D recombinante de la AGT
Fuc	Fucosa	RE	Reticulo endoplásmico
$\Delta G$	Cambio en la energía libre de Gibbs	Rha	Ramnosa
$\Delta G^\circ$	Cambio en la energía libre estándar	RIP	Proteínas inactivadoras del ribosoma
Gal	Galactosa	RP-HPLC	HPLC de fase reversa
GalNAc	N-acetilgalactosamina	s	Segundos
GBPs	Proteínas de unión a glicanos	S	Svedberg
Glc	Glucosa	$\Delta S$	Cambio en la entropía
GlcNAc o NAG	N-acetilglucosamina	T	Temperatura en grados Kelvin
gor	Glutación reductasa	TFA	Ácido trifluoroacético
GPCR	Receptor acoplado a proteínas G	tRNA	Ácido ribonucleico de transferencia
h	Horas	UDA	Lectina de <i>Urtica dioica</i>
$\Delta H$	Cambio en la entalpía	UV	Ultra violeta
HPLC	Cromatografía líquida de alto rendimiento	wAGT	Aglutinina de germen de trigo silvestre
IDC	Calorimetría de dilución isotérmica	WBAI	Aglutinina básica del frijol alado
IPTG	Isopropil- $\beta$ -D-tiogalactopiranosido	WBAll	Aglutinina ácida del frijol alado
ITC	Calorimetría de titulación isotérmica		
$K_A$	Constante de asociación		
$K_D$	Constante de disociación		
kJ	kilojulios		
L	Ligando		
LiCl	Cloruro de litio		
ln	Logaritmo natural		
M	Molar		
MALDI-TOF	Espectrometría de masas de desorción/ionización mediante láser asistida por matriz, acoplada a un analizador de tiempo de vuelo		
Man	Manosa		
mg	Miligramo		
MHz	Megahercio		
min	Minutos		
ml	Mililitro		
mM	Milimolar		
$\mu$ M	Micromolar		
MRW	Masa molar media por residuo		
NaCl	Cloruro de sodio		
Neu5Ac o NANA	Ácido N-acetilneuramínico		
NGF	Factor de crecimiento nervioso		
nm	Nanometro		
NMR	Resonancia magnética nuclear		
NOESY	Espectroscopia de efecto nuclear de Overhauser		

## RESUMEN

La aglutinina de germen de trigo (AGT) es una lectina emblemática no solo por ser una de las primeras lectinas en ser purificada y caracterizada, sino también porque se ha convertido en un modelo cuyo estudio ha recuperado mucho interés recientemente. La AGT une principalmente residuos de N-acetilglucosamina y en el trigo (*Triticum vulgare*) desempeña un importante papel protegiéndolo de insectos, levaduras y bacterias. Fue la primera lectina con la que se describió la glicosilación aberrante en células malignas y desde ese momento se le han encontrado numerosas aplicaciones como la detección de agentes patógenos en fluidos humanos, el marcaje con lectinas, la identificación temprana de diferentes tipos de cáncer, la cuantificación de la fibrosis en tejido cardíaco, su implementación en métodos de ingeniería de tejidos y en el rastreo neuroanatómico. De manera destacable, se ha utilizado en la mejora de la administración de fármacos por vía oral, aprovechando su capacidad para atravesar el tracto digestivo e inducir transcitosis en el epitelio intestinal. A pH neutro, la AGT forma homodímeros débiles con un peso molecular de 34 kDa. Cada subunidad está compuesta por cuatro dominios homólogos a la heveína y cada uno, a su vez, está estabilizado por 4 puentes disulfuro. En el dímero, las 2 subunidades están asociadas de manera " cabeza-cola", estableciendo 4 pares de contactos entre dominios y 8 sitios de unión no idénticos, ubicados a lo largo de la intercara de la proteína. Las condiciones ácidas favorecen la monomerización de la AGT, pero aún en su forma monomérica, todos los dominios mantienen su capacidad de unión a carbohidratos. Las firmas termodinámicas correspondientes a cada dominio aún no se han determinado y tampoco se ha probado si los sitios de unión de la intercara del dímero muestran algún tipo de cooperatividad. El paisaje de plegamiento de la AGT también es complejo, puesto que cada monómero presenta transiciones de desplegamiento de múltiples pasos, que a su vez están fuertemente influenciadas por la disociación del dímero. La presencia de muchos intermediarios de desplegamiento estables dificulta la deconvolución directa de los perfiles de desplegamiento e impide evaluar la estabilidad individual de cada dominio, así como la comunicación entre ellos. Para elucidar las bases moleculares de la función de la AGT y facilitar el desarrollo de más aplicaciones basadas en esta lectina, en el presente estudio se estableció la expresión de la AGT y de sus 4 dominios aislados en un sistema bacteriano. Todas las proteínas se produjeron como proteínas de fusión, unidas en la región N-terminal a una tiorredoxina que se removió mediante proteólisis enzimática. La AGT recombinante mostró propiedades estructurales y de unión prácticamente idénticas a la AGT obtenida de su fuente natural. Los dominios recombinantes también se obtuvieron plegados y funcionales, indicando que cada dominio constituye una unidad cooperativa de plegamiento independiente. Además de establecerse un método económico y eficiente para obtener construcciones recombinantes de la AGT, la caracterización biofísica de estas proteínas permitió elucidar propiedades de unión y estabilidad necesarias para lograr entender mejor el comportamiento de la proteína.

## ABSTRACT

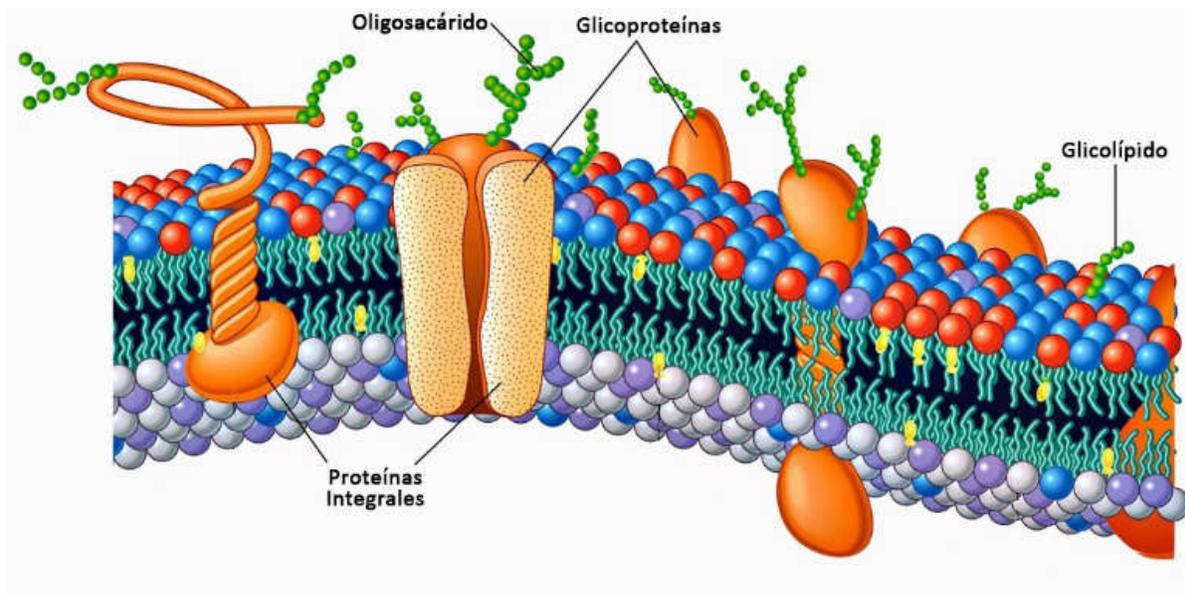
Wheat germ agglutinin (WGA) is an emblematic lectin, not only because it is one of the first lectins to be purified and characterized, but also because it has become a model whose study has recently regained much interest. This lectin primarily binds N-acetylglucosamine residues and within the wheat (*Triticum vulgare*) plays an important role protecting it from invading insects, yeast and bacteria. It was the first lectin used to reveal the aberrant glycosylation of malignant cells. It has been shown to have plenty of promising biomedical applications, like detection of pathogenic agents in human fluids, lectin blotting, early identification of different types of cancer, quantification of fibrosis in cardiac tissue, improvement in methods of tissue engineering and neuroanatomical tracing. Remarkably, it has been used to improve peroral drug delivery systems, taking advantage of its ability to go through the digestive tract and to undergo transcytosis across the intestinal epithelium. At neutral pH, WGA forms weak homodimers with a molecular weight of approximately 34 kDa. Each subunit is composed of four hevein-like domains, each domain in turn is stabilized by 4 highly conserved disulfide bridges. In the dimer, the 2 subunits associate to each other in a "head-to-tail" manner, establishing 4 pairs of domain-domain contacts and 8 non identical interfacial extended binding sites. Acidic conditions favor WGA monomerization, though all the domains in the monomeric subunit keep up their carbohydrate binding ability. The corresponding thermodynamic signatures for each domain have not yet been determined, neither has been proved whether the interfacial binding sites of the dimer show any kind of cooperativity binding effects. WGA also presents a complex folding landscape, with each monomer exhibiting multi-step unfolding transitions that are increasingly influenced by the dissociation of the dimer as the pH changes from neutral to acidic. The presence of many stable unfolding intermediates makes it difficult to perform a direct deconvolution of the unfolding profiles, which precludes a robust determination of the individual domain stabilities, as well as the evaluation of interdomain communication effects. In order to achieve experimental conditions that allow us to clarify the molecular basis of WGA function and facilitate the development of further utilities based on this lectin, we established the heterologous expression of WGA and its 4 isolated domains in a bacterial system. All the proteins were produced as fusion constructs, linked to a thioredoxin domain in the N-terminal region that was lately released by enzymatic proteolysis. Purified recombinant WGA showed indistinguishable conformational and ligand-binding properties as compared to WGA purified from its natural source. The recombinant domains were also folded and capable of binding chitotriose, indicating that each domain constitutes an independent cooperative folding unity. Thus, herein we established a cheap and efficient method for obtaining recombinant derived constructs of WGA. Furthermore, we carried out the biophysical characterization of these recombinant proteins allowed to elucidate binding properties and stability necessary to improve our understanding of this protein behavior.

## INTRODUCCIÓN

### La relevancia del desarrollo de la glicómica

Los carbohidratos son el tipo más diverso de moléculas de los cuatro tipos fundamentales presentes en los seres vivos (Marth, 2008). Pueden estar presentes en su forma más sencilla, como monosacáridos, o constituir moléculas formadas por un mayor número de unidades que van desde disacáridos y trisacáridos hasta oligosacáridos y polisacáridos. Estas moléculas desempeñan papeles biológicos fundamentales, merece especial mención el caso de monómeros como la glucosa, la galactosa y la fructosa que sirven como la principal fuente de energía de muchos organismos. Polisacáridos como la amilosa, la amilopectina y el glucógeno sirven de reserva energética en plantas y animales, respectivamente. Otros polisacáridos de gran importancia son la celulosa y la quitina. La primera es el homopolisacárido (formado exclusivamente de moléculas de glucosa) más abundante en el planeta, ya que se encuentra formando parte estructural de la pared celular de plantas y algas. Mientras que la quitina, un homopolisacárido de N-acetilglucosaminas, se encuentra como componente del exoesqueleto de los artrópodos y ha sido pieza clave del éxito evolutivo de este filo, que es el más numeroso y diverso del reino animal (Kuriyan, Konforti y Wemmer, 2013; Gabius y Roth, 2017; Lovegrove *et al.*, 2017).

También, podemos encontrar glicanos de manera abundante en proteínas secretadas, en la matriz extracelular y de forma cuantiosa, alrededor de muchos tipos de células en la naturaleza. En este último caso, lo usual es encontrar a los carbohidratos en forma de glicoconjugados. Esta definición engloba a todos los carbohidratos que se encuentran unidos a través de enlaces covalentes a proteínas (glicoproteínas y proteoglicanos) o lípidos (glicolípidos). En su conjunto, los glicocojulgados forman una capa densa, asimétrica y compleja que cubre toda la superficie celular y es comúnmente conocida como glicocáliz (Fig. 1), palabra que literalmente significa "cobertura de azúcar". El glicocáliz se ha encontrado presente en todas las células animales, en varias bacterias y en muchas otras células. Inicialmente solo se le consideraba como un elemento cuyas principales funciones eran proteger a la célula y servir como un elemento de filtración selectivo. Pero con el tiempo, se vio que el glicocáliz era mucho más que eso, además de que presenta una composición y una topografía que son únicas para cada tipo celular. Incluso se le ha llegado a comparar con el intrincado dosel arbóreo de un bosque, ya que ambos presentan varias dimensiones o capas. Además de las funciones mencionadas anteriormente, los oligosacáridos del glicocáliz, son importantes portadores de información. De manera que, no solo rivaliza con la información que nos pueden brindar los ácidos nucleicos o las proteínas, sino que la termina superando con creces debido a su enorme variabilidad estructural (Cohen y Varki, 2014; Gabius, 2018).



**Figura 1. Representación gráfica del glicocáliz.** Se muestra de forma general, los distintos tipos de glicoconjugados que conforman esta capa azucarada de gran importancia en los procesos de reconocimiento celular (Harayama y Riezman, 2018).

Esta variabilidad surge debido a que los carbohidratos presentan un inmenso número de confómeros y variaciones estructurales, puesto que pueden estar constituidos por diferentes tipos de monosacáridos, contar con cadenas que pueden ser lineales o ramificadas, formar enlaces glucosídicos con distinta estereoquímica ( $\alpha$  o  $\beta$ ) y pueden sufrir derivatizaciones en cualquier grupo funcional de cada uno de sus monosacáridos por la adición de grupos sustituyentes, como aminas, amidas, sulfatos, entre muchos otros (Varki y Sharon, 2009; Cohen y Varki, 2010).

Estas características estructurales le otorgan las cualidades necesarias para ser considerado como un sistema de codificación de alta densidad, por lo que se le ha dado el nombre de código azucarado. La gran complejidad de este código permite divisar que su función se diversifica aún más de lo que podemos describir con nuestro conocimiento actual. De manera general, se puede decir que cumple la importante función de compilar importante información de señalización y reconocimiento molecular, algo que abarca desde lo que ocurre en el medio intracelular hasta las interacciones intercelulares que pueden darse entre diferentes tejidos o con células de otros organismos. Por mencionar algunos ejemplos, se ha visto la participación del glicocáliz en procesos como la adhesión y la diferenciación celular, la migración de las células durante el desarrollo, la infección por patógenos, la diferenciación de lo propio y de lo ajeno, entre muchos otros (Allen y Snary, 1972; Gabius y Roth, 2017; Gabius, 2018).

Es importante mencionar que los cambios postraduccionales que pueden sufrir las proteínas, como las glicosilaciones, suelen afectar su plegamiento y actividad. Esto, generalmente, amplía el intervalo de su funcionalidad (Gabiús *et al.*, 2011). De la misma manera, los cambios que ocurren en los patrones del glicocáliz en combinación con otros elementos celulares, pueden desencadenar una enorme y variada cantidad de respuestas fisiológicas, llegando incluso a desembocar en el desarrollo de alguna patología indeseable (Ghazarian, Idoni y Oppenheimer, 2011; Ushiyama, Kataoka y Iijima, 2016; Kuo *et al.*, 2018). El ejemplo, que quizás tiene mayor relevancia es el del cáncer. La evidencia experimental que se ha colectado durante casi cinco décadas muestra que casi todos los tipos de células malignas tienen glicosilaciones aberrantes en comparación con sus contrapartes normales. Estas alteraciones en los patrones de glicosilación aparecen en todo tipo de glicoconjugados. No obstante, las alteraciones en los dominios de carbohidratos de glicolípidos se han podido aislar y describir con mayor precisión que los cambios de glicoproteínas, ya que estos últimos suelen ser más heterogéneos. A pesar de que la glicosilación aberrante no es una característica distintiva de todos los tipos de cáncer, se ha visto que cuando aparece, normalmente desempeña un papel crucial en varios de los pasos clave para la progresión y el desarrollo de la enfermedad. Como lo son la disociación del tejido tumoral, la interacción con la matriz celular, angiogénesis, mediación del sistema inmune, invasión y metástasis (Hakomori, 1985; Shida *et al.*, 2009; Ghazarian, Idoni y Oppenheimer, 2011; Munkley y Elliott, 2016; Kuo *et al.*, 2018).

La glicómica es un campo que emerge en una era posterior a la aparición de la genómica y la proteómica, su objetivo se centra en caracterizar el glicoma de los seres vivos. Entiéndase por glicoma al conjunto total de carbohidratos de un organismo. Cabe destacar que esta disciplina no sólo busca realizar una caracterización estructural de los carbohidratos, sino que también pretende elucidar su función y sus implicaciones en los fenómenos biológicos (Cummings y Pierce, 2014). Sin embargo, la tarea de descifrar el código azucarado no es para nada trivial. Por las razones anteriormente mencionadas, el código muestra un nivel de complejidad bastante elevado y aunado a esto, se tiene que tener en cuenta que no podemos mapear la presencia y la estructura de los glicanos directamente del genoma de la misma forma que se utiliza con proteínas. La razón subyacente a esto, es que los patrones de glicosilación dependen de la actividad de múltiples enzimas como glicosiltransferasas y glicosidasas que actúan en el retículo endoplásmico (RE), en el aparato de Golgi y en la superficie celular y que pueden estar o no sometidas a algún tipo de regulación (Feizi y Mulloy, 2003; Ghazarian, Idoni y Oppenheimer, 2011). Esto le da todavía más plasticidad al glicoma y lo hace dependiente de las condiciones del medio en el que se desarrolle el organismo. Tan solo en organismos eucariontes, se ha estimado que la proporción de proteínas glicosiladas supera el 50% (Apweiler, Hermjakob y Sharon, 1999). Si a esto se le suma la gran variedad de datos recabados por distintos grupos de investigación acerca de la relevancia de los carbohidratos para la célula, queda claro que el glicoma es esencial para la vida y debe empezarse a considerar con una importancia similar a la que se les da a conjuntos como el genoma y el proteoma (Cohen y Varki, 2014).

Así, aunque la glicómica sea una disciplina reciente, es indispensable seguir ampliando sus límites a la par de la genética y la biología celular. Algo que en definitiva nos puede brindar información vital para la prevención y el tratamiento de enfermedades (Hirabayashi, 2008; Cummings y Pierce, 2014).

## Lectinas, definición, distribución e importancia para la vida

Buena parte de las funciones fisiológicas que desempeñan los carbohidratos están mediadas por las proteínas que han evolucionado para poder reconocerlos. A este conjunto de proteínas se le conoce como proteínas de unión a glicanos (o GBPs, del inglés glycan-binding proteins). Excluyendo a los anticuerpos específicos de glicanos, es posible clasificar de manera general a las GBPs en dos grandes grupos, las proteínas de unión a glicosaminoglicanos y las lectinas (Varki, 2007). El presente trabajo se centra en este último grupo.

La palabra lectina proviene de *lectus*, el pasado participio de *legere*, que significa seleccionar o escoger. Las lectinas son un grupo de proteínas de origen no inmune que están presentes en todos los tipos de seres vivos, tanto en el reino vegetal como en el animal, así como en todo tipo de microorganismos, lo que incluye bacterias, protozoarios y virus. Se caracterizan por su capacidad de reconocer y unir carbohidratos, de manera reversible (a través de interacciones no covalentes) y muy específica, sin alterar la estructura covalente de éstos durante el proceso. Cada lectina presenta su propio perfil de especificidad por ciertos monosacáridos y por los polímeros derivados de éstos. Antaño, se pensaba que todas las lectinas contenían dos o más sitios de unión a carbohidratos; por lo que su interacción con los azúcares de la superficie celular causaba el entrecruzamiento entre células (comúnmente eritrocitos) y su subsecuente precipitación. A este fenómeno se le ha dado el nombre de aglutinación (hemaglutinación para el caso particular de los glóbulos rojos). Durante mucho tiempo se consideró a esta capacidad de aglutinar como uno de los principales atributos de la actividad de las lectinas y aún se sigue explotando esta propiedad de manera rutinaria como uno de los métodos clásicos para su detección y caracterización. Sin embargo, actualmente se sabe que no todas las lectinas aglutinan y que incluso hay algunas que únicamente cuentan con un solo sitio de unión a carbohidratos. Por lo tanto, la capacidad de aglutinación ya no se considera una característica general de las lectinas (Barondes, 1988; Ambrosi, Cameron y Davis, 2005).

La amplia distribución de las lectinas ya es, en sí mismo, un fuerte argumento de su relevancia fisiológica. Aun así, cada tipo de organismo parece tomar provecho de sus características de distintas formas. Muchos agentes patógenos como virus, bacterias, hongos y protozoarios, expresan una gran cantidad y variedad de lectinas. Con frecuencia, estas lectinas son utilizadas por estos con el fin de reconocer a los azúcares de la célula anfitriona durante el proceso de infección. La hemaglutinina del virus de la influenza fue la primer lectina purificada a partir de un virus (aproximadamente por 1950) y actualmente es considerada una de las mejor descritas (Wilks *et al.*, 2012). Las lectinas de bacterias pueden ser clasificadas como adhesinas si ayudan a la adhesión eficiente al hospedero, o como toxinas si son proteínas secretadas al medio. Un ejemplo de este último tipo es la citotoxina subtilasa (SubAB) producida por la cepa de *Escherichia coli* Shiga-Toxigénica (Kitov *et al.*, 2000).

Las lectinas de animales suelen desempeñar funciones como aglutinar los elementos y agentes extraños a nuestro cuerpo (virus, bacterias, parásitos, etc.) para facilitar la actuación de nuestras propias defensas naturales, regular la vida media de las glicoproteínas en el sistema circulatorio, participar en el reclutamiento de leucocitos a sitios de inflamación, etc. (Ambrosi, Cameron y Davis, 2005; Gabius, 2013).

En las plantas, la mayoría de las lectinas están presentes en los cotiledones y endospermos de las semillas y constituyen del 2 al 10% del total de proteínas. Se ha propuesto que las lectinas pueden proteger a las semillas de ataques de insectos, bacterias, hongos y virus fitopatogénicos a lo largo de sus diferentes etapas de germinación y desarrollo. Uno de los primeros trabajos, en lo que a esto respecta, se hizo sometiendo a unas larvas de escarabajo de la especie *Callosobruchus maculatus* a una dieta que contenía lectina de frijol negro (*Phaseolus vulgaris* L.), lo que resultó en la inevitable muerte de las larvas (Janzen, Juster y Liener, 1976). Pero las lectinas de plantas no son exclusivas de las semillas, de hecho, se han logrado aislar lectinas de otras fuentes vegetales como hojas, bulbos, corteza y raíces. En estas últimas, se les ha visto involucradas en el reconocimiento de las bacterias fijadoras de nitrógeno del género *Rhizobium* (Hirsch, 1999). También, se les ha asociado con otras funciones como aglutinación de proteínas, transporte de carbohidratos, estimulación mitogénica, entre otras. Entre las lectinas vegetales mejor conocidas se encuentran las siguientes: la Concanavalina A (Con A), que es una proteína que se obtiene de la planta *Canavalia ensiformis* que actúa específicamente uniéndose a residuos de glucosa y manosa. La fitohemaglutinina (PHA), que une galactosa y es obtenida de diferentes variedades del frijol común pertenecientes a la especie *Phaseolus vulgaris* y la aglutinina de germen de trigo (AGT). Esta última, es el sistema de estudio de este trabajo y se describirá más adelante (Gabius *et al.*, 2004; Van Damme, Lannoo y Peumans, 2008).

## Historia de las lectinas

Las lectinas son herramientas valiosas en el desarrollo de la glicómica contemporánea, además de que tienen un vínculo muy estrecho con los orígenes de esta disciplina. Es gracias a estas proteínas que se logró distinguir la presencia y la importancia de los carbohidratos del glicocáliz. Su historia se remonta a finales del siglo XIX, aproximadamente por 1860, cuando se hizo la primera observación documentada de la actividad de una lectina, al describirse que el veneno de una serpiente cascabel causaba la aglutinación de los eritrocitos presentes en la sangre de una paloma. Fue hasta 1888, cuando Peter Stillmark en su tesis doctoral realizada en la universidad de Dorpa (Estonia), presentó al mundo la primera proteína con capacidad de hemaglutinar. Se trataba de la ricina, la cual fue obtenida a partir de las semillas del ricino (*Ricinus communis*), planta que coloquialmente es conocida como castor o higuera infernal. Cabe mencionar que la capacidad hemaglutinante de esta proteína fue opacada en gran medida por la toxicidad que presentaba. Actualmente, se sabe que la ricina está compuesta por dos cadenas polipeptídicas, la cadena B que presenta actividad de lectina y la cadena A que presenta actividad de N-glicosidasa. La toxicidad de la ricina se debe principalmente a esta última cadena, la cual se encarga de eliminar una adenina del RNA ribosomal 28 S, lo que causa la inhibición de la síntesis proteica. Esta proteína ha sido utilizada como veneno en algunos de los casos más controversiales de la historia, incluyendo el famoso "asesinato del paraguas" en el que perdió la vida el periodista búlgaro Georgi Markov en 1978, por lo que se le suele recordar más como veneno que como la proteína que marcó el inicio de la incursión de las lectinas en la investigación (Sharon y Lis, 2004; Espino-Solis, 2015).

En el mismo año de 1888, H. Hellin (también de la universidad de Dorpa) logró describir a la abrina, obtenida de la planta conocida por el nombre común de regaliz americano (*Abrus precatorius*). Para finales de 1890, tanto la ricina como la abrina logran volverse comercialmente disponibles. Se cuenta que su uso comenzó a popularizarse cuando Emmanuel Merck, boticario cuyos descendientes son los fundadores de la gigante farmacéutica Merck & Co., logró convencer al inmunólogo Paul Ehrlich de utilizar a la ricina y a la abrina como antígenos en sus ensayos, algo que le haría merecedor del premio Nobel en medicina en 1908, al establecer la base química para la especificidad de la respuesta inmunológica.

A lo largo del siglo XX, se dieron otros avances fundamentales en el campo. En 1919, J. B. Sumner, quien ya había ganado algo de fama por cristalizar a la ureasa, logró obtener la primer lectina pura, la Con A. Ésta fue aislada a partir de *Canavalia ensiformis*, especie conocida comúnmente como frijol espada o judía (Sharon y Lis, 2004). Sumner tardó casi dos décadas más en lograr encontrar su función. Por lo que hasta 1936 fue capaz de describir su capacidad de aglutinar glóbulos rojos, células de levadura y glucógeno en solución. Sumner también encontró que esta aglutinación podía ser inhibida con sacarosa, con lo que se demuestra por primera vez la especificidad de las lectinas por los azúcares (Sumner y Howell, 1936).

Durante los siguientes años, se siguió describiendo la actividad de varias lectinas, encontrándose que cada lectina tenía su propio perfil de hemaglutinación, el cual es distinto para cada tipo sanguíneo. Pero no fue hasta 1950, que Walkins y Morgan demostraron que estas propiedades hemaglutinantes podían ser inhibidas por diversos tipos de azúcares. Estos experimentos fueron la primera evidencia contundente de que en las superficies celulares hay carbohidratos y que estos funcionan como importantes marcadores de identidad celular.

Finalmente, en 1954 nace el término lectina, acuñado por Boyd y Shapleigh (Boyd y Shapleigh, 1954). Aunque no es hasta 1972, cuando Sharon y Lis (Sharon y Lis, 1972) generalizaron su uso a todas las aglutininas sin importar su origen. Durante este período acontecieron otros dos experimentos de gran relevancia. Primero fue C. Nowell (Nowell, 1960), quien, en 1960, encontró que la PHA inducía la mitosis en linfocitos. A este le siguió Joseph C. Aub (Aub, Sanford, & Cote, 1965a; Aub, Sanford, & Wang, 1965b), quien descubrió que la aglutinina de germen de trigo (AGT) aglutina preferencialmente células malignas sobre sus contrapartes normales.

Hasta este punto, ya se habían descrito las actividades de varias lectinas de extractos celulares de plantas, aunque sólo se habían logrado purificar a la Con A (Sumner y Howell, 1936) y a la AGT (Burger y Goldberg, 1967). Por lo que los experimentos de Nowell y Aub constituyen los primeros testimonios del importante papel fisiológico que desempeñan los glicanos de la superficie celular y, por ende, fueron los causantes de que se intensificara el interés por las lectinas justo antes de que comenzara la década de los setenta. A lo anterior, se le suma la aportación de Agrawal y Goldstein (Agrawal y Goldstein, 1965), quienes en 1967 introducen la cromatografía de afinidad implementada específicamente para la purificación de lectinas, con lo que aumentó considerablemente el número de lectinas purificadas durante los siguientes años. Es importante mencionar que, en 1968, 1969 y 1971, se purificaron las primeras lectinas de animales (de cangrejo herradura, caracol y anguila, respectivamente) y para 1974 se logró purificar al "receptor de asialoglicoproteína", siendo éste la primera lectina de mamíferos en caracterizarse (Gabiús *et al.*, 2004; Sharon y Lis, 2004; Varki y Sharon, 2009).

## La diversidad estructural de las lectinas

Las lectinas presentan dominios especializados para llevar a cabo su función, estas unidades estructurales independientes contienen el sitio de unión a carbohidratos y se conocen comúnmente como dominios de reconocimiento a carbohidratos (DRC). Las diferencias y similitudes estructurales entre estos dominios ha servido para clasificar a las lectinas en diversas familias y superfamilias (Ghazarian, Idoni y Oppenheimer, 2011). El número de familias descritas para cada reino de los seres vivos ha ido aumentando de manera vertiginosa conforme se han encontrado y descrito nuevas lectinas.

Por poner un ejemplo, se puede mencionar el caso de las lectinas animales. Hasta 1988, se creía que las lectinas animales (incluyendo las del humano) solo podían pertenecer a una de las dos grandes familias estructurales que eran las lectinas tipo C y las tipo S. Las lectinas tipo C recibieron su nombre debido a su dependencia por el ion calcio ( $\text{Ca}^{2+}$ ), mientras que las lectinas tipo S fueron clasificadas como tal debido a que necesitaban de condiciones reductoras para mantener su función. En este caso, la letra S hace referencia de los grupos sulfhidrilo involucrados. Actualmente, las lectinas tipo C se siguen considerando como una familia estructural independiente. Sin embargo, se encontró que solo un número reducido de miembros de las lectinas tipo S son dependientes de condiciones reductoras, por lo que se cambió su nombre a galectinas (Kilpatrick, 2002; Yang, Rabinovich y Liu, 2008). Desde ese entonces, el número de patrones de plegamiento en lectinas animales aumentó considerablemente, pasado de ser solo 2 a 14 familias estructurales (Solís *et al.*, 2015).

De manera similar a lo observado en las lectinas animales, los patrones de plegamiento encontrados en lectinas de otros organismos se ha incrementado durante los últimos años. En el caso de lectinas de plantas y hongos se ha logrado reconocer hasta 12 tipos distintos de plegamiento (Manning *et al.*, 2017). Aunque aparentemente el conjunto de patrones de plegamiento que se observa en las lectinas es bastante diverso, si se compara con el número total de plegamientos que se puede encontrar en todo el universo de proteínas, nos podemos dar cuenta de que en realidad se trata de un grupo relativamente pequeño. En especial si se tiene en cuenta que, en menos de una veintena de motivos, las lectinas logran el cometido de reconocer la enorme diversidad de epítomos estructurales presentada por el código azucarado. Acción que también les permite modular su acción fisiológica. Todo esto resulta una prueba contundente del discurrir evolutivo de las lectinas por adaptarse a la enorme diversidad de los glicanos.

Debido a que nuestro modelo de estudio es una lectina de trigo y con el fin de profundizar en la comparación, en la Fig. 2 se describen de manera general los tipos de plegamientos más comúnmente observados en el reino vegetal.

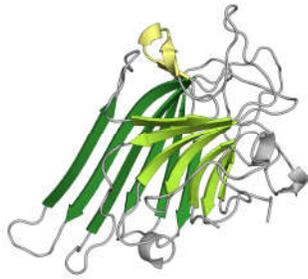
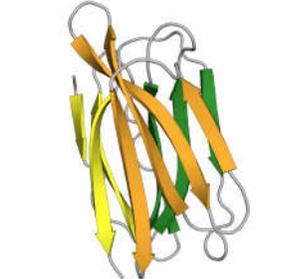
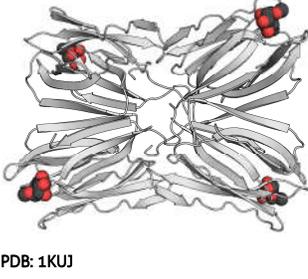
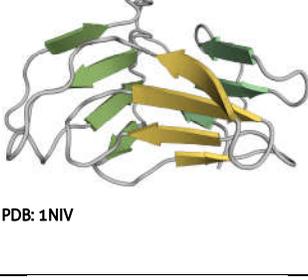
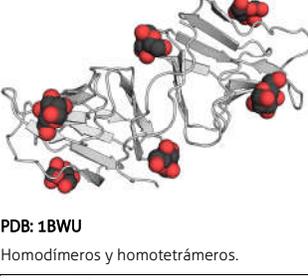
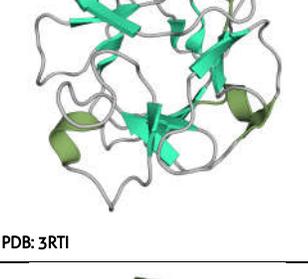
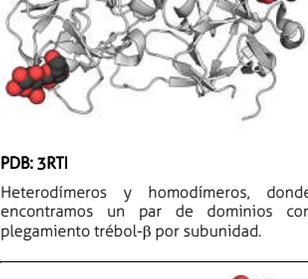
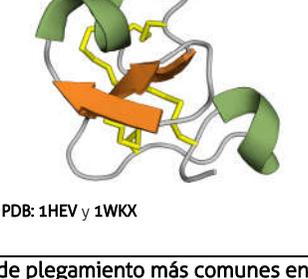
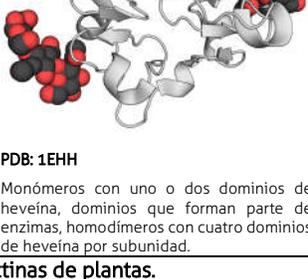
Plegamiento	Estructura cristalográfica	Estados oligoméricos más comunes	Ligandos principales	Ejemplos de lectinas
Sandwich-β I	 PDB: 1JBC	 PDB: 2PEL Homodímeros y homotetrámeros.	Man, Glc, Gal y GalNAc. 1 ligando por dominio.	Concanavalina A ( <i>Canavalia ensiformis</i> ): <b>1JBC</b>  Aglutinina de cacahuate ( <i>Arachis hypogaea</i> ): <b>2PEL</b>  Aglutinina básica (WBAI) y aglutinina ácida (WBAII), ambas del frijol alado ( <i>Psophocarpus tetragonolobus</i> ).
prisma-β I	 PDB: 1KUJ	 PDB: 1KUJ Homodímeros, homooctámeros y homotetrámeros.	Gal o Man. 1 o 2 ligandos por dominio.	Jacalina y Artocarpina, ambas de la yaca ( <i>Artocarpus integrifolia</i> ): <b>1KUJ</b>  Lectina de plátano ( <i>Musa paradisiaca</i> )  Lectina de calsepa ( <i>Calystegia sepium</i> )  Lectina de tupinambo ( <i>Helianthus tuberosus</i> )
prisma-β II	 PDB: 1NIV	 PDB: 1BWU Homodímeros y homotetrámeros.	Man. 3 ligandos por dominio.	Lectina del galanto o campanilla de invierno ( <i>Galanthus nivalis</i> ): <b>1NIV</b>  Lectina del ajo ( <i>Allium sativum</i> ): <b>1BWU</b>
trébol-β	 PDB: 3RTI	 PDB: 3RTI Heterodímeros y homodímeros, donde encontramos un par de dominios con plegamiento trébol-β por subunidad.	Gal y GalNAc. 1 ligando por dominio.	Ricina de la semilla de ricino ( <i>Ricinus communis</i> ): <b>3RTI</b> Aglutinina de la semilla de la calabaza serpiente ( <i>Trichosanthes anguina</i> )  Aglutinina de amaranto ( <i>Amaranthus caudatus</i> )
Dominios homólogos a la heveína	 PDB: 1HEV y 1WKX	 PDB: 1EHH Monómeros con uno o dos dominios de heveína, dominios que forman parte de enzimas, homodímeros con cuatro dominios de heveína por subunidad.	GalNAc y Neu5Ac. 1 ligando por dominio.	Heveína ( <i>Hevea brasiliensis</i> ): <b>1HEV y 1WKX</b>  Lectina de <i>Urtica dioica</i> (UDA): <b>1EHH</b>  Dominios de quitinasas de clase I.  Aglutinina de germen de trigo ( <i>Triticum vulgare</i> )

Figura 2. Patrones de plegamiento más comunes en lectinas de plantas.

El último grupo de la Fig. 2 comprende a los dominios homólogos a la heveína. La heveína es una proteína extraída del árbol del caucho (*Hevea brasiliensis*), cuya estructura se definió por primera vez a través de cristalografía de rayos X a una resolución de 2.8 Å (Rodríguez-Romero, Ravichandran y Soriano-García, 1991). Posteriormente, se resolvió la estructura con una resolución de 1.5 Å (Reyes-López *et al.*, 2004) y se ha analizado a través de NMR (Andersen *et al.*, 1993; Asensio *et al.*, 1995). Esto ha permitido definir de forma consensual la estructura de la heveína e identificar un grupo de proteínas que comparten su plegamiento (Drenth *et al.*, 1980; Berthelot, Peruch y Lecomte, 2016). El patrón de plegamiento se caracteriza por mostrar una hoja-β central que está formada por hebras que son cortas y antiparalelas, la cual se encuentra rodeada por dos pequeñas hélices estabilizadas por cuatro puentes disulfuro. Los dominios tipo heveína son uno de los motivos de unión a quitina más comunes. Se les ha visto en diversas lectinas como la pseudoheveína, la lectina de *Urtica dioica*, en dominios que forman parte de enzimas como las quitinasas de clase I y en cada uno de los cuatro dominios de la AGT, que se describirá más adelante.

## Termodinámica de la unión entre proteína y ligando

Las proteínas no operan como entidades aisladas y estáticas. Por el contrario, requieren involucrarse en numerosas interacciones con otras macromoléculas, como otras proteínas, ácidos nucleicos, membranas, pequeñas moléculas de ligando y de manera constante, con las moléculas de solvente. Estos eventos de reconocimiento molecular dictaminan los procesos biológicos que son críticos para la vida, lo que incluye el reconocimiento entre enzimas y sustratos, la transmisión de señales celulares, el reconocimiento de una célula por otra, los procesos de replicación, transcripción y traducción, etc. (Bronowska, K., 2011). En este tipo de interacciones, dos macromoléculas se asocian de forma no covalente gracias a la participación de puentes de hidrógeno, interacciones hidrofóbicas, electrostáticas y de van der Waals. Estos complejos no covalentes suelen disociarse hasta cierto grado, lo que lleva a una mezcla de moléculas libres y acomplejadas. Esto se reconoce como el equilibrio del proceso de unión y en su forma más simple se puede representar como se muestra en la ecuación 1, donde P representa a la proteína libre, L al ligando libre y P·L al complejo proteína-ligando.



Si se obtienen las proporciones de estas especies al equilibrio, se pueden utilizar para obtener una constante de equilibrio de la misma forma que se hace para cualquier reacción reversible de tipo químico. La constante de equilibrio para la reacción de asociación mostrada en la ecuación 1 es referida como la

constante de asociación ( $K_A$ ) y puede definirse de la siguiente manera, en donde  $[P \cdot L]$ ,  $[P]$  y  $[L]$  son las concentraciones de cada una de las especies al equilibrio:

$$K_A = \frac{[P \cdot L]}{[P][L]} \quad (2)$$

Es más común describir la afinidad de una proteína por su ligando con la constante de disociación ( $K_D$ ). Esta tan solo es el inverso de la  $K_A$ :

$$K_D = \frac{[P][L]}{[P \cdot L]} = \frac{1}{K_A} \quad (3)$$

Las interacciones no covalentes biológicamente relevantes normalmente se caracterizan por valores de  $K_D$  que van del orden de picomolar a milimolar. En este espectro, las interacciones de mayor afinidad se consideran aquellas que presentan  $K_D$  dentro del rango de picomolar a nanomolar ( $10^{-12}$  a  $10^{-9}$  M), mientras que las interacciones más débiles tienen  $K_D$  del orden de milimolar ( $10^{-3}$  M) (Kuriyan, Konforti y Wemmer, 2013). Aunque la  $K_D$  sirve para estimar la afinidad de este tipo de interacciones, hace falta un método que sirva para describirlas de forma cuantitativa.

La asociación no covalente entre proteína y ligando está gobernada por la termodinámica general de manera similar a cualquier otro proceso espontáneo al equilibrio, por lo que se puede determinar su cambio en la energía libre de Gibbs de unión ( $\Delta G_U$ ). Este parámetro nos indica que la asociación ocurre únicamente cuando esta se encuentre descrita por un  $\Delta G_U$  negativo, el cual resulta de la suma de un componente entálpico y un componente entrópico. En la ecuación 4, se puede ver que el  $\Delta G_U$  se puede expresar con el cambio en la entalpía ( $\Delta H_U$ ), el cambio en la entropía ( $\Delta S_U$ ) y la temperatura ( $T$ ). Las contribuciones entálpicas reflejan un  $\Delta H_U$  negativo (proceso exotérmico) si se da la formación de interacciones no covalentes entre átomos que sean energéticamente favorables y un  $\Delta H_U$  positivo (proceso endotérmico) si se pierden interacciones no covalentes energéticamente favorables. Mientras que los cambios en las contribuciones entrópicas son una medida de la pérdida (o ganancia) de grados de libertad que resulta de la suma de la entropía de solvatación, la entropía conformacional y la de rototraslación (Perozzo, Folkers y Scapozza, 2004; Bronowska, K., 2011; Du *et al.*, 2016). Las firmas termodinámicas del proceso de unión pueden ser distintas para diferentes complejos proteína-ligando. En algunos casos la asociación puede estar entálpicamente dirigida y en otros, es el componente entrópico el que la dirige.

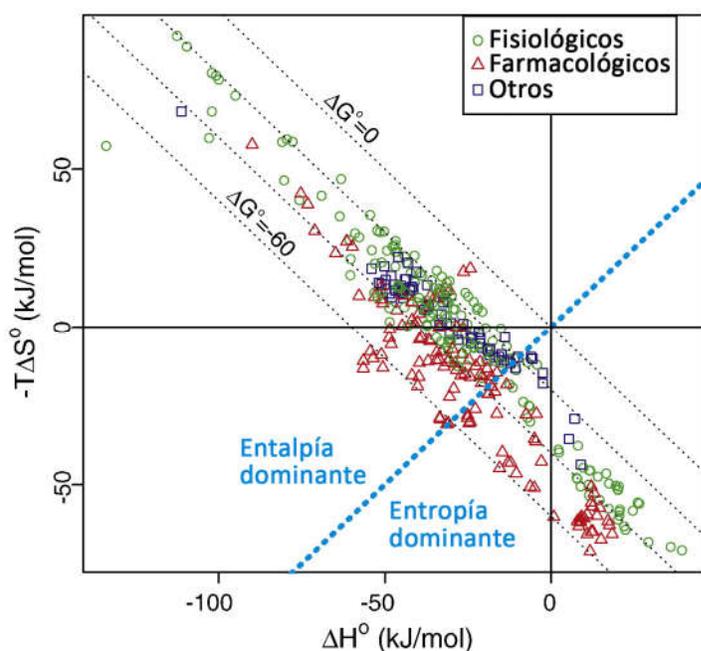
$$\Delta G_U = \Delta H_U - T\Delta S_U \quad (4)$$

A su vez, la constante de equilibrio está relacionada con el cambio en energía libre:

$$\Delta G^\circ_U = -RT\ln(K_D) \quad (5)$$

La  $K_D$  y los otros parámetros termodinámicos de unión se pueden obtener con algunos métodos experimentales, dentro de los que destaca por su precisión, la calorimetría de titulación isotérmica (ITC). Aunque estos experimentos suelen ser complejos debido a algunos factores, como la cantidad de proteína

requerida o el número de repeticiones necesarias para determinar la entropía y la entalpía de unión, esto no ha impedido que se incorpore una buena cantidad de datos a la literatura y nos permita visualizar algunas tendencias. Un ejemplo de esto es la compensación que se observa entre la entalpía y la entropía en los procesos de unión. La Fig. 3 muestra como al comparar el  $\Delta H_U$  con el  $-T\Delta S_U$  aparece una clara correlación, donde la contribución de un componente es, hasta cierto punto, contrarrestada por su contraparte. Esta compensación entalpía-entropía es útil para obtener información del mecanismo que controla la formación del complejo proteína-ligando (Olsson *et al.*, 2008; Reynolds y Holloway, 2011).



**Figura 3.** Gráfica de entalpía ( $\Delta H_U$ ) vs entropía ( $-T\Delta S_U$ ). Se muestra una recopilación de los datos de entalpía y entropía obtenidos para varias interacciones proteína-ligando y almacenadas en la base de datos SCORPIO. Las interacciones están clasificadas en las que involucran al ligando natural de las proteínas, las que ocupan algún ligando sintético producto de una investigación químico-médica y en las que no caen en ninguna de estas dos categorías. Las líneas punteadas en negro indican los valores de  $\Delta G_U$  espaciados a intervalos de  $20 \text{ kJ mol}^{-1}$ . La línea discontinua azul representa el punto donde  $\Delta H_U = T\Delta S_U$ , las interacciones por encima de esta línea están dirigidas por la entalpía y las que están por debajo se dirigen por el componente entrópico (Olsson *et al.*, 2008).

El cambio en la capacidad calorífica a presión constante ( $\Delta C_{p_U}$ ) es otra propiedad fundamental en la caracterización termodinámica de la unión molecular. El  $\Delta C_{p_U}$  es la razón de cambio del  $\Delta H_U$  con respecto a la temperatura:

$$\Delta C_{p_U} = \frac{\partial \Delta H_U}{\partial T} \quad (6)$$

Este parámetro es un importante sensor de los rearrreglos que pueden ocurrir en las interacciones intermoleculares que existen entre los solutos (proteína y ligando) y el solvente. Los cambios en la capacidad calorífica normalmente son causados por las variaciones que se dan en el área accesible al solvente de la proteína tras la unión del ligando (Pulido *et al.*, 2008).

Comprender mejor las fuerzas que conducen la interacción y el reconocimiento, requiere una descripción detallada de la termodinámica de unión, pero también necesita que se correlacione los parámetros termodinámicos con la estructura de los elementos involucrados y con los cambios que éstos puedan tener durante la interacción. Alcanzar un mejor entendimiento del fenómeno de reconocimiento

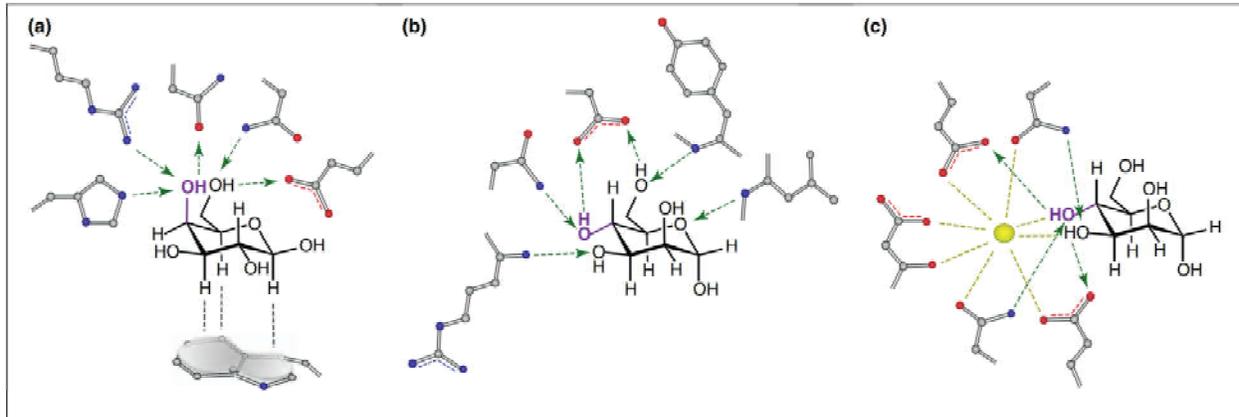
es de gran importancia para la Biología y para la Química medicinal, donde abre la posibilidad de mejorar significativamente el diseño racional de fármacos basado en la estructura (Su y Xu, 2018).

## Bases moleculares de la formación de los complejos lectina-carbohidrato

En la unión y reconocimiento de carbohidratos de las lectinas, una pieza fundamental es el establecimiento de puentes de hidrógeno entre los grupos polares de los aminoácidos de la superficie de reconocimiento y los grupos hidroxilo de los carbohidratos. En comparación con los puentes de hidrógeno intracadena que podemos encontrar dentro de las proteínas, se ha visto que en los complejos lectina-carbohidrato los puentes de hidrógeno que involucran a grupos cargados de la proteína son más abundantes, además de que presentan una mejor geometría de enlace y mayor densidad de superficie (García-Hernández y Hernández-Arana, 1999; Garcia-Hernandez *et al.*, 2000).

Muchas veces los carbohidratos difieren únicamente en la posición relativa de dos o de uno solo de sus grupos hidroxilo, caso en el que se les denomina epímeros. Es fácil pensar que una diferencia así de minúscula pueda pasar desapercibida, pero no es así. La epimerización tiene consecuencias sustanciales en el reconocimiento y las lectinas son capaces de discernir entre epímeros. Después de comparar distintas series de homólogos de lectinas, queda claro que los puentes de hidrógeno son los principales determinantes de esta estereoselectividad (García-Hernández y Hernández-Arana, 1999; Gabius *et al.*, 2011). En la Fig. 4 se puede ver que distintos tipos de lectinas hacen uso de los grupos polares de sus residuos de aminoácidos para establecer puentes de hidrógeno con la geometría necesaria para distinguir entre distintos epímeros de la glucosa.

Otra fuerza intermolecular de gran importancia en el reconocimiento de carbohidratos, es la interacción de la parte apolar de un carbohidrato con los residuos aromáticos de las lectinas (Tyr, Phe y Trp). Esto permite la estabilización de los complejos proteína-carbohidrato a través de interacciones de apilamiento (stacking o interacciones C-H/ $\pi$ ) y contactos de van der Waals. Esta interacción hidrofóbica entre los anillos aromáticos y los azúcares, es algo que se observa en la inmensa mayoría de las lectinas. Se han hecho algunos estudios comparativos para ver la diferencia entre las interacciones de apilamiento de distintos tipos de aminoácidos aromáticos. A pesar de ello, todavía es complicado determinar puntualmente sus aportaciones al proceso de unión, pero al menos en algunos casos ya se han logrado identificar a los aminoácidos aromáticos clave en el reconocimiento, así como sus posiciones relativas (Asensio *et al.*, 1995; Espinosa *et al.*, 2000; Wimmerová *et al.*, 2012).



**Figura 4. Selección de epímeros por tres tipos de lectinas.** Los puentes de hidrógenos están representados como flechas, donde la cabeza representa al grupo aceptor del H y el extremo opuesto al donador del H. (a) Galectina-1 de humano uniendo galactosa a través de varios puentes de hidrogeno, 3 de estos puentes se establecen con el hidroxilo axial de la posición 4 de la galactosa. (b) Concanavalina A uniendo manosa. En contraste con la galectina, la concanavalina puede formar puentes de hidrógeno con el hidroxilo ecuatorial de la posición 4 de la glucosa y la manosa, lo que le permite unir a estos azúcares específicamente. (c) Lectina tipo-C uniendo manosa a través de puentes de hidrógeno y un ion de  $Ca^{2+}$ (Gabiús *et al.*, 2011).

La participación de aminoácidos como Arg, Asp y Glu en la formación de puentes de hidrógeno, en coordinación con los aminoácidos aromáticos capaces de establecer las interacciones de apilamiento, es lo que determina mayoritariamente el que las lectinas tengan distintos perfiles de reconocimiento a monosacáridos y les permita distinguir entre oligosacáridos con diferentes topologías (Gabiús *et al.*, 2011).

De los análisis que se han realizado de distintos DRC, se ha encontrado que por lo general las firmas termodinámicas de la formación de complejos proteína-carbohidrato suelen tener la misma pauta. Aunque las magnitudes de los parámetros de unión pueden ser distintas, normalmente se observan cambios en la entalpía favorables ( $\Delta H_U$  negativo), los cuales son compensados parcialmente por cambios en la entropía desfavorables ( $\Delta S_U$  negativo).

En lo que respecta a la afinidad, se sabe que comúnmente las interacciones de unión de lo DRC individuales a los carbohidratos es de baja afinidad ( $K_D = 10^{-3}$  a  $10^{-7}$  M) en comparación con las interacciones antígeno-anticuerpo ( $K_D = 10^{-6}$  a  $10^{-9}$  M) o de híbridos DNA-RNA (Cooper, 2002; Hirabayashi, 2008). Para compensar esta baja afinidad, se ha visto que durante su interacción con los glicanos, las lectinas recurren frecuentemente a estrategias de multivalencia. Es decir, en lugar de interactuar únicamente a través de un solo sitio de unión con baja afinidad, estas proteínas suelen involucrar a más de un sitio (Gabiús, 2002). Con esto logran incrementar su afinidad y alcanzar una unión biológicamente relevante. Existen tres formas a través de las cuales logran esto:

- 1) Multivalencia simple. Se puede mencionar el caso de la hemaglutinina del virus de la influenza. Aunque el sitio de unión de esta proteína presenta una alta especificidad por carbohidratos sialilados (oligosacáridos que contienen residuos de ácido siálico), su afinidad por éstos es muy baja. Sin embargo, esta proteína se encuentra distribuida en grandes cantidades alrededor de la cápside del virus. Por lo que durante la unión aporta más de un sitio de reconocimiento y así logra aumentar la avidéz de la unión.
- 2) Agregación multivalente. En este caso las subunidades de la misma proteína se asocian o forman agregados entre sí, para lograr aportar un mayor número de sitios de unión. Las subunidades B de la toxina del cólera actúan de este modo, formando un pentámero que se une al gangliósido GM1 con mayor afinidad que la que presentarían las subunidades por separado.
- 3) Unión a un parche de oligosacáridos agrupados (o CSP del inglés clustered saccharide patch). En este caso se forma un epítopo o parche por la agrupación de varios oligosacáridos que pueden estar situados dentro de un mismo glicoconjugado o en distintos glicoconjugados, siempre y cuando estos se encuentren lo suficientemente cercanos entre sí. En esta estrategia, al igual que en las anteriores, se establecen varios sitios de unión entre la proteína y los carbohidratos (Cohen y Varki, 2014).

## Lectinas, aplicaciones y potencial biotecnológico

En la actualidad, la glicómica tiene a su alcance algunas herramientas como la cromatografía líquida, la espectrometría de masas o el uso de anticuerpos específicos. Esto ha permitido purificar y caracterizar una variedad considerable de carbohidratos, tanto aislados como en forma de glicoconjugados. Sin embargo, estas técnicas se enfocan más en la detección específica, por lo que son poco apropiadas para caracterizar el extremadamente diverso espectro de carbohidratos del glicoma. Si se quisiera preparar un conjunto de anticuerpos que fuera lo suficientemente grande y específico para poder detectar todos los elementos del glicoma, nos embarcaríamos en una lista infinita de tareas que sería prácticamente imposible de terminar. Además, también se debe tener en cuenta que los oligosacáridos del glicocáliz no sólo presentan una extensa diversidad en sus patrones de glicosilación, sino que también se pueden encontrar formando distintos parches de carbohidratos o epítopos debido a que se encuentran a distancias relativamente cercanas entre sí (Hirabayashi, 2004, 2008).

No fue hasta hace poco que se empezó a ver el potencial de las lectinas como herramientas auxiliares en la resolución de este problema. Durante mucho tiempo, su capacidad fue subestimada y fueron relegadas por los anticuerpos que, de forma común, presentan mayor afinidad y especificidad por las

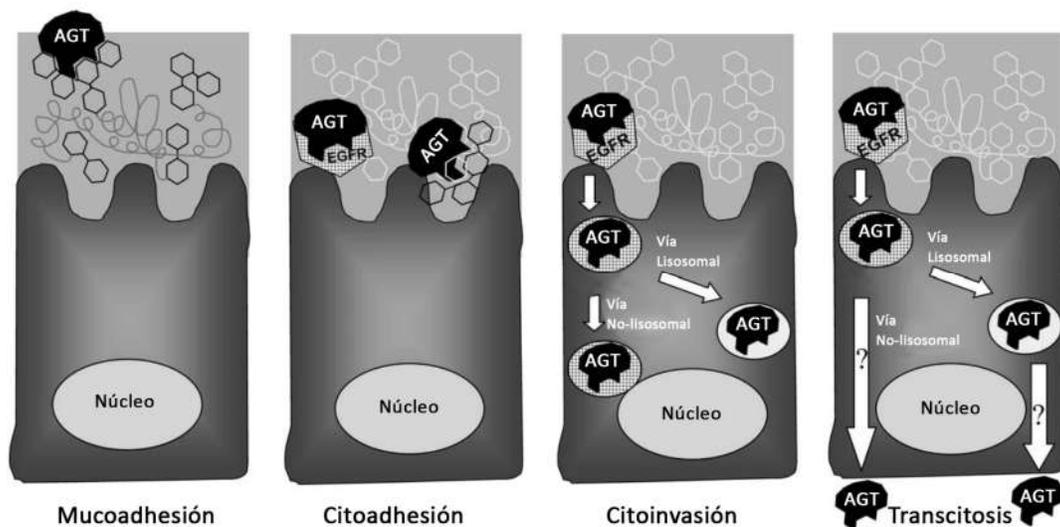
moléculas que unen. Uno podría pensar que su menor afinidad representa una desventaja en todo tipo de circunstancias, pero en realidad no es así. La especificidad de las lectinas ha permitido que se puedan utilizar para leer series de carbohidratos con diferencias estructurales sutiles entre sí y que se muestren altamente competentes en la identificación de los distintos tipos de epítomos presentes en el glicocáliz. Por lo que las lectinas se pueden utilizar de manera fiable no sólo en la identificación individual de carbohidratos, sino también en el reconocimiento y descripción de los perfiles de glicosilación celular (Hirabayashi, 2008; Tang *et al.*, 2015). Buena parte de la información recabada durante los últimos años, deja clara la posibilidad de que las lectinas empiecen a ser utilizadas, de forma más amplia y común, como una herramienta alternativa a los anticuerpos en el campo de reconocimiento molecular. Como ejemplo, podemos encontrar la implementación de microarreglos de lectinas, sin duda una de las aplicaciones más novedosas en la identificación y el análisis de patrones de glicosilación de diversas muestras biológicas como extractos celulares, tejidos y células completas. Esto ha permitido vincular determinados cambios en la composición de carbohidratos a enfermedades muy particulares (Gemeiner *et al.*, 2009; Fu *et al.*, 2011). La aparición de una glicosilación aberrante se ha observado en desordenes hereditarios (Durand y Seta, 2000), cáncer (Munkley y Elliott, 2016), inmunodeficiencia (Ji *et al.*, 2006), neurodegeneración y diabetes (Hart, Housley y Slawson, 2007).

En el campo del desarrollo de agentes antimicrobianos, es común encontrar la aparición de microorganismos multiresistentes y encontrar alternativas para combatir las infecciones que éstos causan se ha vuelto otro reto a superar. Las lectinas son otro medio que puede ayudar con esta tarea, ya que de forma natural son potentes agentes antimicrobianos y su capacidad para unirse a los carbohidratos presentes en las superficies de los microbios es sólo el primer paso en el despliegue de numerosos mecanismos de defensa contra patógenos. Pueden causar la aglutinación de células bacterianas facilitando su asimilación por macrófagos, formar poros en la membrana, desintegrar la pared celular o detener la síntesis *de novo* de ésta. También, se han encontrado capaces de inhibir el crecimiento y la germinación de esporas fúngicas (Bar-Shavit y Goldman, 1976; Gallily *et al.*, 1984; Kumar *et al.*, 2012; Coelho *et al.*, 2018). La actividad antiviral de varias lectinas es otra área que sigue en extenso desarrollo (Hamid *et al.*, 2013).

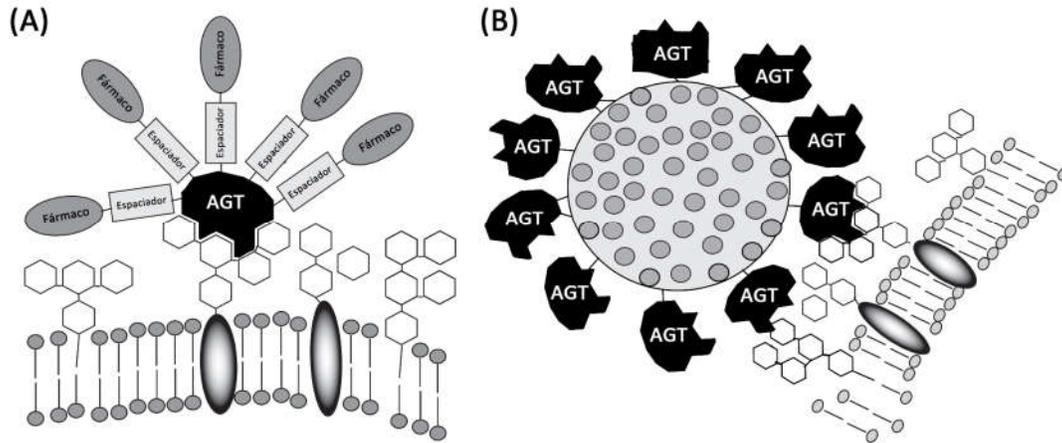
Así mismo, se ha renovado el interés por continuar explorando las capacidades de lectinas que en el pasado han mostrado propiedades potencialmente aplicables. Este es precisamente el caso de la aglutinina de germen de trigo (AGT), lectina objeto de estudio del presente trabajo. La AGT no sólo es emblemática desde la perspectiva histórica, sino que también ha mostrado cualidades prometedoras en el campo de la biomedicina. Desde que se llevó a cabo el estudio en el que se observó que la AGT podía reconocer células malignas, esta proteína se ha seguido utilizando como auxiliar en el diagnóstico de diversos tipos de cáncer como el carcinoma de estómago (Koyama *et al.*, 1999), cáncer de colon (Gabor *et al.*, 2004), cáncer de vejiga (Yang *et al.*, 2011), cáncer de próstata (Yang *et al.*, 2015), entre otros. De

manera similar, se ha usado en la detección de agentes patogénicos (Murakami *et al.*, 2014; De Brito Ximenes *et al.*, 2015) y se ha implementado en métodos de histoquímica y de marcaje con lectinas (lectin-blotting) para la identificación de alteraciones en los glicoconjugados causadas por diversas afecciones como la enfermedad celíaca, artritis reumatoide, septicemia, e incluso por contaminación ambiental (Mislovičová *et al.*, 2009). Entre otras aplicaciones que se le han dado, tenemos su inclusión en la mejora de algunos procedimientos de ingeniería de tejidos (Teuschl *et al.*, 2014), en la cuantificación de fibrosis en el tejido cardiaco después de un infarto de miocardio (Emde *et al.*, 2014), en la inhibición del hongo *Trichoderma viride* (Kumar *et al.*, 2012) e incluso como marcador transináptico (Libbrecht *et al.*, 2017).

La AGT también ha mostrado ser capaz de adherirse al glicocáliz del epitelio intestinal y de realizar la transcitosis a través de este tejido, como se puede ver en la Fig. 5 (Gabor, 2004). Esto ha permitido considerar a la AGT como una herramienta que sirva para mejorar la absorción de fármacos administrados por vía peroral, e incluso para guiar su entrega de manera tejida específica (Gabor *et al.*, 2004; He *et al.*, 2019). En este respecto, se han hecho algunos estudios con liposomas que presentan una cobertura de AGT y con varios tipos de micropartículas conjugadas con AGT (Fig. 6). Por mencionar algunos ejemplos, se ha utilizado a la AGT en la entrega de varios antineoplásicos como la daunorubicina para tratar el cáncer de mama (Liu *et al.*, 2017), la gemcitabina en el tratamiento de cáncer de vejiga (Neutsch *et al.*, 2011, 2012, 2013) y la oridonina para el tratamiento de cáncer de colon (Liu *et al.*, 2016), en la entrega de curcumina (CRM) y del factor de crecimiento nervioso (NGF) para el tratamiento de Alzheimer (Kuo *et al.*, 2017), de antibióticos como la claritromicina (Meng *et al.*, 2016) y en la administración de la insulina y otros péptidos (Zhang *et al.*, 2005; Makhlof *et al.*, 2011). También, se han evaluado las propiedades anticancerígenas intrínsecas de la AGT en células Jurkat y en líneas celulares de cáncer de páncreas y de colon (Schwarz *et al.*, 1999; Gastman *et al.*, 2004; Heinrich *et al.*, 2005).



**Figura 5. Posibles pasos involucrados en la transcitosis de la AGT en los enterocitos.** Se observa de izquierda a derecha, la unión de la AGT a la mucina, al receptor del factor de crecimiento epidérmico (EGFR), la acción citoinvasiva de la AGT y su transcitosis (Gabor *et al.*, 2004).



**Figura 6. Estrategias comúnmente utilizadas en la entrega de fármacos mediada por la AGT.** (A) Conjugados de la AGT unidos a fármacos a través de regiones espaciadoras. (B) Liposomas que portan fármacos en su interior y están cubiertos por varias AGT (Gabor *et al.*, 2004).

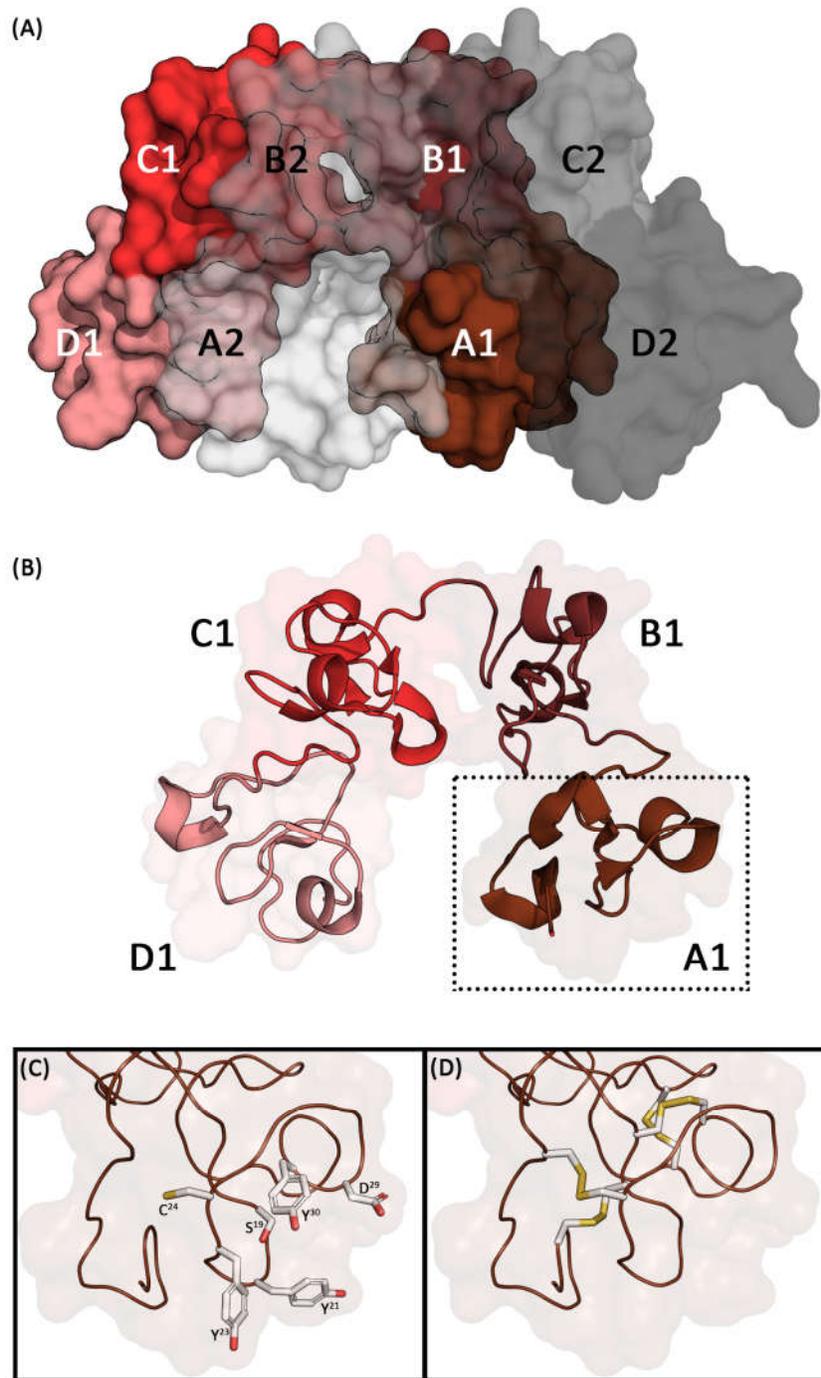
### Aglutinina de Germen de Trigo (AGT): Estructura y unión a carbohidratos

La AGT es una lectina abundante en las semillas de trigo (*Triticum aestivum* o *T. vulgare*), donde se encuentra presente en 3 isoformas (AGT 1, AGT 2 y AGT 3). Se cree que estas isoformas son producto de eventos de duplicación y fusión génica. Las 3 isoformas de la AGT presentan cadenas polipeptídicas de 171 residuos de aminoácidos y tienen una estructura primaria casi idéntica. Con respecto a la isoforma 1, la isoforma 2 es distinta sólo en 5 residuos de su secuencia, mientras que la isoforma 3 difiere en 8 residuos (Wright y Raikhel, 1989). Estas diferencias se pueden apreciar en el alineamiento presentado en la Fig. 7. El presente trabajo se enfoca en la isoforma 1 de la AGT, por lo que se procederá a describir sólo a esta proteína, aunque es de esperarse que muchas de estas características las comparta cercanamente con las otras 2 isoformas.

	10	20	30	40	50	60
AGT1	QRCGEQGSNM	ECPNNLCCSQ	YGYCGMGGDY	CGKGCQNGAC	WTSKRCGSQA	GGATCTNNQC
AGT2	.....	.....	.....	.....	.....	.....P..H.
AGT3	.....G.	.....	.....	.....	.....	..K..P..H.
	70	80	90	100	110	120
AGT1	CSQYGYCGFG	AEYCGAGCQG	GPCRADIKCG	SQAGGKLCPN	NLCCSQWGFC	GLGSEFCGGG
AGT2	.....H....	.....	.....	..S.....	.....	.....
AGT3	.....H....	.....	.....	.....	.....Y.	.....E.
	130	140	150	160	170	
AGT1	CQSGACSTDK	PCGKDAGGRV	CTNNYCCSKW	GSCGIGPGYC	GAGCQSGGCD	G
AGT2	.....	.....	.....	.....	.....	A
AGT3	..N.....	.....	.....	.....	.....	..

Figura 7. Alineamiento de las secuencias de aminoácidos de las tres isoformas de la AGT. El símbolo • denota identidad con respecto a la secuencia de la isoforma 1 de la AGT. La isoforma 2 tiene un 97% de identidad con la isoforma 1, con tan solo 5 residuos de diferencia. La isoforma 3 presenta una identidad del 95% con la isoforma 1, con 8 diferencias en su secuencia.

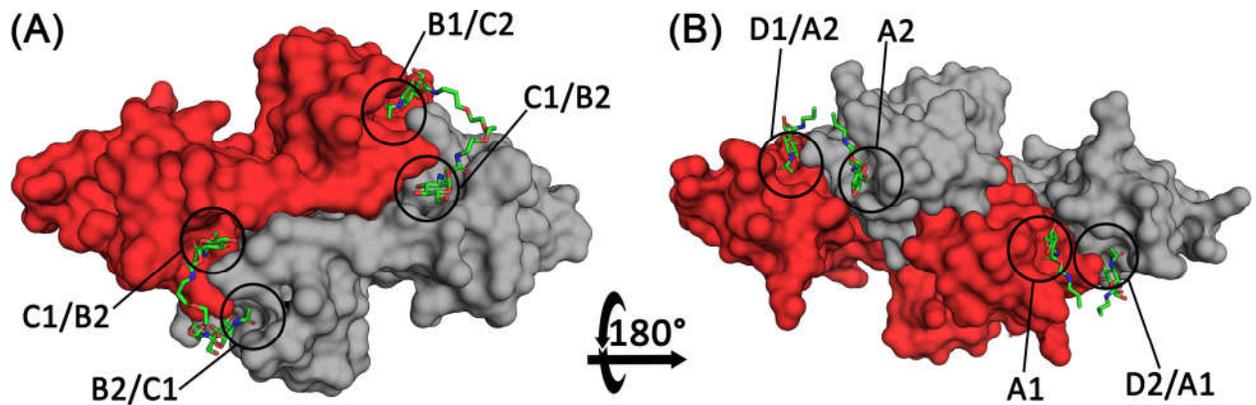
La AGT puede unir, de manera específica y reversible, monosacáridos de N-acetilglucosamina (GlcNAc o NAG) y de ácido N-acetilneuramínico (Neu5Ac o NANA). También puede unir a los oligosacáridos derivados de éstos, pero se sabe que son los oligosacáridos de GlcNAc unidos por enlaces  $\beta$ 1,4 para los que presenta mayor afinidad (Bains *et al.*, 1992). Esta proteína forma un homodímero de 34 kDa, en el que cada subunidad (de 17 kDa cada una) se encuentra compuesta de cuatro dominios homólogos a la heveína (de 42 a 43 aminoácidos cada uno). Para distinguirlos, cada dominio se etiqueta con las letras A, B, C y D, seguidas por un número que indica la subunidad a la cual pertenecen (Wright, 1989, 1992). En la estructura dimérica, los monómeros se asocian de forma cabeza-cola (Fig. 8A), de manera que se forman 4 pares de contactos entre dominios contrapuestos de los monómeros (Wright y Kellogg, 1996). En términos de estructura primaria, el dominio C de la AGT es el más parecido a la heveína, con un 56% de identidad. La estructura secundaria de la AGT se puede apreciar en la Fig. 8B. Resulta evidente la similitud de la estructura secundaria de cada uno de los dominios con la heveína, ya que cada dominio está compuesto por el mismo patrón de pequeñas hélices alfa unidas a través de asas. Cada dominio de heveína presenta 4 puentes disulfuro (Fig. 8D), dando un total de 16 puentes disulfuro por subunidad (Rodríguez-Romero, Arreguín y Hernández-Arana, 1989; Andersen *et al.*, 1993).



**Figura 8. Estructura cristalográfica de la AGT (Código PDB: 2UWG).** (A) Representación de superficie en la que se aprecia el arreglo cabeza-cola entre las dos subunidades de la AGT. Se presenta a la cadena polipeptídica 1 en rojo y a la cadena polipeptídica 2 en gris. Las letras en blanco corresponden a los dominios de la subunidad 1 y las letras en negro a los dominios de la subunidad 2. (B) Representación esquemática de la estructura secundaria de la AGT. Se muestra un solo monómero, donde se pueden distinguir las asas y las pequeñas hélices alfa que forman a cada uno de los dominios homólogos a la heveína. (C) Acercamiento del dominio A en el que se pueden observar los residuos más importantes del sitio de unión de este dominio. Estos residuos son similares en los otros dominios, aunque no idénticos. (D) Utilizando la misma perspectiva, se muestran los 4 puentes disulfuro intracadena del dominio A. Estos puentes se encuentran conservados en cada uno de los dominios (Schwefel *et al.*, 2010).

## ANTECEDENTES

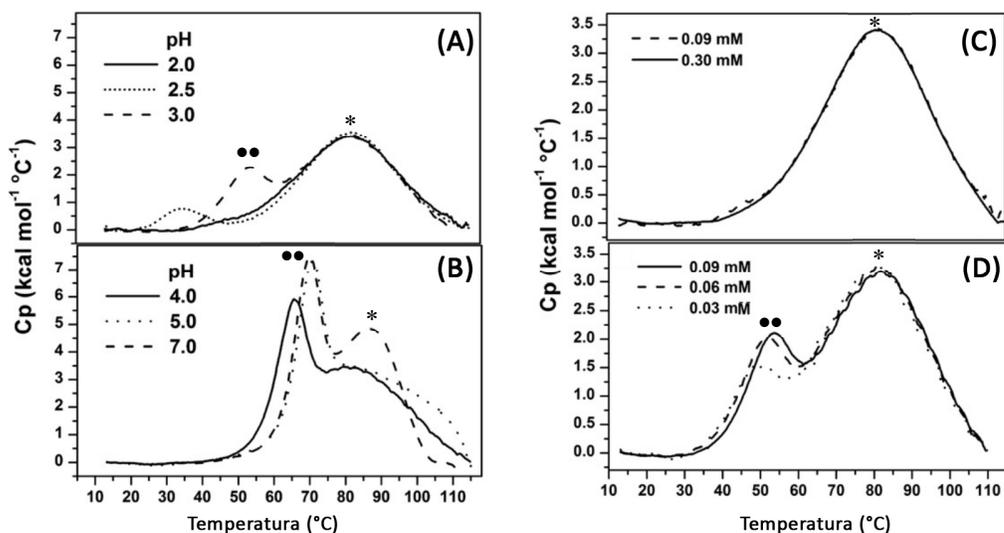
En un principio, se describió a la AGT como una proteína con cuatro sitios de unión a carbohidratos por dímero (Nagata y Burger, 1974). Sin embargo, posteriormente se encontraron evidencias de que el dímero de esta proteína podía presentar hasta ocho sitios de unión (Wright y Kellogg, 1996). Esta información se confirmó recientemente con un estudio de cristalografía en el que se obtuvo una estructura cristalográfica de la isoforma 3 de la AGT que muestra la ocupación simultánea de los 8 sitios (Fig. 9). En el mismo estudio, se obtuvo una estructura cristalográfica de la isoforma 1 de la AGT con condiciones prácticamente similares, lo que demostró que todos los sitios de unión del dímero de la AGT son funcionales. En este estudio, se utilizó una serie de ligandos divalentes capaces de ocupar dos sitios de unión al mismo tiempo (Schwefel *et al.*, 2010). Todos los sitios de unión se encuentran en la intercara dimérica y se pueden agrupar estructuralmente en 4 tipos no equivalentes: (1) B1/C2 y B2/C1. (2) C1/B2 y C2/B1. (3) D1/A2 y D2/A1 y (4) A1 y A2. A pesar de encontrarse en la intercara, los sitios de unión no se encuentran formados de manera equivalente por los dos dominios. La cavidad del sitio de unión está compuesta mayoritariamente por los aminoácidos de uno solo de los dos dominios, razón por la cual a esta parte se le ha designado como sitio de unión principal. Los residuos clave en el sitio de unión principal, son tres residuos aromáticos conservados, de los cuales el segundo se encarga de interactuar con el azúcar por apilamiento. El sitio de unión principal también cuenta con hasta 3 aminoácidos polares, que se encargan de formar importantes puentes de hidrógeno con el carbohidrato (Fig. 8C). El dominio tipo heveína de la subunidad opuesta solo aporta algunos residuos polares más, lo que permite la formación adicional de puentes de hidrógeno con el ligando. A esta parte se le conoce como sitio de unión auxiliar y no es estrictamente necesaria para la unión del carbohidrato a la proteína. De hecho, 2 de los sitios de unión (A1 y A2) están compuestos únicamente por el sitio de unión principal, puesto que carecen de los aminoácidos polares necesarios en el dominio opuesto. En los sitios de unión D1/A2 y D2/A1 encontramos otra diferencia de gran importancia, aquí el sitio de unión principal en los dominios D sólo presenta 2 de los 3 residuos aromáticos que encontramos en los otros dominios. Esto ha hecho que, tradicionalmente, se nombren a los sitios de unión B1/C2, B2/C1, C1/B2 y C2/B1 como sitios de unión de alta afinidad y a D1/A2, D2/A1, A1 y A2 como sitios de unión de baja afinidad, como se puede observar en la Fig. 9 (Nagahora *et al.*, 1995; Schwefel *et al.*, 2010).



**Figura 9. Los 8 sitios de unión de la AGT (Código PDB: 2X52).** La cadena polipeptídica 1 se muestra en rojo, la cadena polipeptídica 2 se muestra en gris y el ligando divalente se muestra como un modelo de bastones principalmente en verde. (A) Se muestra una vista superior, con los sitios B1/C1, C2/B1, C1/B2 y B2/C2. (B) Se observa a la AGT por abajo, con los sitios de unión D1/A2 y D2/A1 y los sitios de unión A1 y A2 que están formados únicamente por un dominio (Schwefel *et al.*, 2010).

En nuestro laboratorio se realizaron algunos experimentos de calorimetría diferencial de barrido (DSC) y de calorimetría de dilución isotérmica (IDC), que arrojaron información de gran utilidad sobre la disociación y el desplegamiento de la AGT. Se encontró que la AGT forma homodímeros débiles a pH neutro y que éstos tienden a disociarse a monómeros conforme se va acidificando el pH (Portillo-Téllez *et al.*, 2011; Afrin y Ikai, 2014).

En un primer análisis por DSC, se estudió a la proteína a una sola concentración (0.09 mM), pero a diferentes valores de pH (2.0, 2.5, 3.0, 4.0, 5.0, 6.0 y 7.0). De acuerdo a los trazos obtenidos, el proceso de desplegamiento se mostró reversible ( $\geq 90\%$ ) entre valores de pH de 2.0 a 4.0, pero a partir de pH 5.0 la reversibilidad disminuyó ( $\sim 60\%$ ), llegando a desaparecer por completo a  $\text{pH} \geq 6.0$ . En todas las endotermas obtenidas a pH 3.0 o mayor, se observó la aparición de 2 picos (Fig. 10A y 10B). Para entender mejor la naturaleza de estos picos, se hizo otra serie de barridos utilizando únicamente 2 valores de pH (2.0 y 3.0), aunque variando la concentración de proteína de 0.03 mM a 0.09 mM. A pH 2.0 se observó un solo pico que mostró un comportamiento independiente de la concentración de proteína (Fig. 10C), por lo que presumiblemente corresponde al desplegamiento del monómero. Conforme se incrementó el pH, empezó a aparecer un segundo pico en el trazo que, a diferencia del otro, sí mostró dependencia de la concentración de proteína. Esta dependencia ratificó que este pico correspondía al proceso de disociación del dímero (Fig. 10D). Todos estos resultados están de acuerdo con algunos estudios previos de la AGT (Nagata y Burger, 1974; Chavelas *et al.*, 2004), lo que confirma la idea de que la lectina se encuentra como monómero a valores de pH ácido. Mientras que, conforme se recorre el pH hacia valores más cercanos al punto isoeléctrico de la lectina ( $\text{pI} = 7.3$ ), el dímero se estabiliza más (Portillo-Téllez *et al.*, 2011).



**Figura 10. Endotermas de DSC para la AGT.** (A) [proteína] = 0.09 mM, pH 2.0, 2.5 y 3.0. (B) [proteína] = 0.09 mM, pH 4.0, 5.0 y 7.0. (C) [proteína] = 0.03 mM y 0.09 mM, pH 2.0. (D) [proteína] = 0.03 mM y 0.09 mM, pH 3.0. (●●) Indica los picos correspondientes a la disociación del dímero. (\*) Señala los picos correspondientes al desplegamiento del monómero. (Portillo-Téllez *et al.*, 2011).

La deconvolución del pico correspondiente al proceso de desnaturalización de las subunidades a pH 2.0, mostró que los datos se podían ajustar a un modelo de tres transiciones independientes de dos estados cada una. No pudiendo ajustarse a cualquier otro modelo que presentara un número diferente de transiciones, incluyendo el modelo de cuatro transiciones, que era el esperado dada la composición de cada subunidad por cuatro dominios isoestructurales. Para explicar este comportamiento, se plantearon tres posibles situaciones. La primera, que uno de los cuatro dominios fuese mucho más resistente a la desnaturalización que los otros tres. La segunda, que dos de los dominios se comportaran como unidades cooperativas de plegamiento. Y la tercera, que uno de los dominios presentara una estructura que mantuviera gran parte de la estructura secundaria, pero con una estructura terciaria perturbada. Se realizaron otros estudios de DSC y de fluorescencia en los que se utilizó cloruro de guanidino; éstos sirvieron para descartar las primeras dos propuestas, dejando abierta la tercera opción.

Los ensayos de IDC permitieron calcular las entalpías de disociación a pH 3.0 en función de la temperatura y a 37 °C en función del pH. En general, los datos calorimétricos obtenidos se describen bien con un modelo de disociación de un dímero e indican que el proceso de dimerización es entálpicamente dirigido (Tabla 1), al menos a temperaturas por arriba de los 30 °C. De manera importante, se encontró que a temperaturas menores a los 35 °C la entalpía de dimerización varía poco con respecto a la temperatura y que lo hace de forma lineal, además de que el cambio en la capacidad calorífica de asociación ( $\Delta Cp_A$ ) no es tan grande. Esto en correlación con los cambios de área accesible al solvente, permite asumir que la asociación de la AGT en este intervalo de temperatura ocurre como una asociación de tipo cuerpo rígido. Sin embargo, por arriba de los 35 °C esta dependencia lineal desaparece y los cambios en el  $\Delta Cp_A$  se vuelven más grandes y negativos (Fig. 11). Probablemente, este efecto se debe a que los valores de entalpía observados corresponden no sólo a la asociación, sino también al plegamiento de uno de los dominios. Lo que apoya la idea de que uno de los cuatro dominios se encuentra parcialmente desplegado y que éste no termina de plegarse hasta que ocurre la asociación. Por lo que, en este caso, ocurre una asociación de tipo ajuste inducido (Portillo-Téllez *et al.*, 2011).

Tabla 1. Energética de dimerización de la AGT medida a través de IDC. <sup>a</sup>

Temperatura (°C)	$K_A$ (M <sup>-1</sup> x 10 <sup>-3</sup> )	$\Delta G_A$ (kcal mol <sup>-1</sup> )	$\Delta H_A$ (kcal mol <sup>-1</sup> )	$T\Delta S_A$ (kcal mol <sup>-1</sup> )
30	12.1 ± 0.2	-5.7	-3.4 ± 0.2	2.3
33	8.3 ± 0.5	-5.5	-5.8 ± 0.2	-0.3
35	7.3 ± 0.4	-5.5	-7.4 ± 0.4	-1.9
37	6.5 ± 0.3	-5.4	-9.6 ± 0.5	-4.2
40	6.2 ± 0.4	-5.4	-14.8 ± 0.2	-9.2
43	3.5 ± 0.7	-5.1	-23.6 ± 0.3	-18.5
46	0.8 ± 0.1	-4.3	-34.2 ± 3.1	-29.9

<sup>a</sup> Adaptado de (Portillo-Téllez *et al.*, 2011).

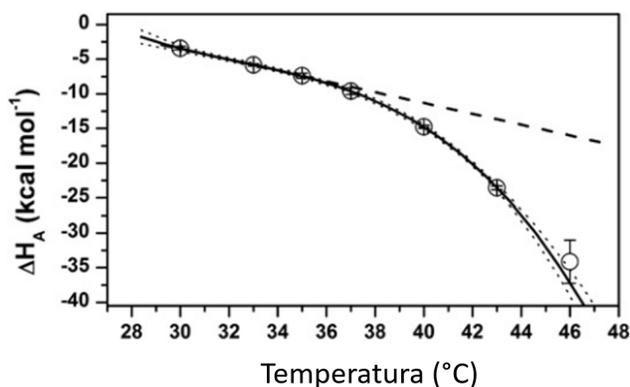


Figura 11. Entalpía de dimerización de la AGT a pH 3.0 como función de la temperatura. Se observa que a temperaturas menores a los 35°C la entalpía varía de manera lineal, linealidad que se pierde a temperaturas mayores (Portillo-Téllez *et al.*, 2011).

Vale la pena mencionar los experimentos realizados por el grupo de Jiménez-Barbero (Espinosa *et al.*, 2000). Ellos lograron producir y aislar al dominio B de la AGT de manera recombinante y a través de resonancia magnética nuclear (NMR) e ITC, lograron determinar las constantes de asociación para este dominio. Con esto se comprueba que los sitios de unión de los dominios aislados son totalmente independientes y capaces de unir carbohidratos sin necesidad de la región auxiliar del sitio de unión.

La gran mayoría de las lectinas que han sido caracterizadas a través de métodos de desnaturalización química, han mostrado ser oligómeros obligados. Es decir, para poder alcanzar la estructura nativa o funcional es completamente indispensable que sus subunidades se asocien y se alcance la estructura cuaternaria. En la Tabla 2, podemos ver los valores de cambio en la energía libre para el plegamiento y asociación ( $\Delta G_A$ ) de varias lectinas de plantas. En general, éstas muestran una elevada estabilidad estructural, exhibiendo valores de  $\Delta G_A$  que se encuentran dentro de un intervalo de -13 a -30 kcal mol<sup>-1</sup> y constantes de dimerización grandes (10<sup>10</sup> – 10<sup>40</sup> M<sup>-1</sup>). En contraste con estas lectinas, el valor que se encontró para la constante de dimerización de la AGT (5 x 10<sup>4</sup> M<sup>-1</sup>) es bajo, indicando que esta forma homodímeros débiles. Además de que la AGT sigue presentando un nivel significativo de actividad aún en su forma monomérica y no se comporta como un homodímero obligado como la gran mayoría de estas lectinas (Portillo-Téllez *et al.*, 2011). Si bien la asociación es crucial para que la AGT alcance su máxima afinidad por el ligando, la habilidad de comportarse como un homodímero débil no obligado hace a esta lectina apta para asistir la administración de fármacos por vía oral, ya que durante su tránsito por el estómago permanecería como monómero, pero recuperaría su forma dimérica al llegar al intestino (Gabor *et al.*, 2003).

**Tabla 2. Energética y mecanismo de formación de homooligómeros de lectinas de plantas.** (Portillo-Téllez *et al.*, 2011).

Lectina	Ensamble cuaternario	$\Delta G_A$ (kcal mol <sup>-1</sup> )	$K_A$ (M <sup>-1</sup> )	Temperatura (°C)	pH
<b>Lectinas de legumbres</b>					
Aglutinina ácida del frijol alado (WBAII)	2D → N <sub>2</sub>	-13	1 x 10 <sup>9</sup>	35	5.0
Concanavalina A	4D → N <sub>4</sub>	-15	2 x 10 <sup>20</sup>	39	7.4
Lectina de <i>Erythrina corallodendron</i>	2D → N <sub>2</sub>	-16	2 x 10 <sup>11</sup>	37	7.4
Lectina de chícharo	2D → N <sub>2</sub>	-18	4 x 10 <sup>12</sup>	37	7.2
Lectina de <i>Erythrina indica</i>	2D → N <sub>2</sub>	-23	1 x 10 <sup>16</sup>	40	7.2
Lectina de cacahuate	4I → N <sub>4</sub>	-15	2 x 10 <sup>20</sup>	37	7.4
Aglutinina de haba de soya	4D → N <sub>4</sub>	-30	2 x 10 <sup>41</sup>	37	7.4
<b>Lectinas emparentadas con la jacalina</b>					
Frutalina	4D → N <sub>4</sub>	-6	3 x 10 <sup>8</sup>	20	7.4
Lectina de plátano	2D → N <sub>2</sub>	-18	4 x 10 <sup>12</sup>	25	7.0
<b>AGT</b>	2D → N <sub>2</sub>	~ -7	5 x 10 <sup>4</sup>	37	4.5

Todos los datos se obtuvieron con ensayos de desnaturalización química. D: Estado desplegado; N: Estado nativo; I: Estado intermedio de plegamiento. Todo se expresa en términos de mol de dímero. <sup>a</sup> Adaptado de (Portillo-Téllez *et al.*, 2011).

## PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACIÓN DEL PROYECTO

La evidencia que se mostró en el apartado anterior, deja claro que la estructura multimodular de la AGT es compleja y dota a esta proteína con intrincados mecanismos de plegamiento y unión del ligando. Además, estos procesos pueden variar significativamente dependiendo de las condiciones del ambiente.

En lo que se refiere a su unión y reconocimiento a carbohidratos, todavía se desconoce buena parte de los mecanismos moleculares que conducen y regulan este proceso. En su forma monomérica, la AGT puede unir un carbohidrato por dominio. Sin embargo, a excepción del dominio B, no se cuenta con las firmas termodinámicas correspondientes a los otros dominios. Como dímero, los sitios de unión de la AGT se encuentran a distancias relativamente cercanas y en las regiones intersubunidad de la proteína, por lo que es concebible la idea de que muestren cooperatividad durante la unión, ya sea por la coordinación simultánea de 2 o más sitios en la unión de la misma molécula de ligando (Schwefel *et al.*, 2010) o mediante comunicaciones internas a través de la proteína. Esta idea no se ha explorado lo suficiente y lo único que de momento es seguro, es que la afinidad del dímero es mayor que la del monómero (Espinosa *et al.*, 2000).

El panorama que se tiene para poder resolver el plegamiento y la auto-asociación de la AGT tampoco se encuentra del todo definido. El desplegamiento de los monómeros ha exhibido múltiples transiciones que dependen en buena medida de la disociación del dímero, la cual se va dando conforme el pH cambia de neutro a ácido. La presencia de varios intermediarios estables durante el desplegamiento, sin duda dificulta la deconvolución de los perfiles de desplegamiento e imposibilita la determinación de las estabilidades individuales para cada dominio y la evaluación de cualquier comunicación entre dominios. En la Fig. 12 se muestra una simplificación de este comportamiento; usando este modelo, se intenta ejemplificar la afectación mutua entre las constantes de plegamiento/auto-asociación y las constantes de unión a carbohidratos.

Con el propósito de establecer condiciones experimentales que permitan seguir avanzando y sirvan en la elucidación de las bases moleculares del funcionamiento de la AGT, en este trabajo se estableció la producción recombinante de la AGT completa (rAGT) y la de sus 4 dominios de manera independiente (rAGT.A, rAGT.B, rAGT.C y rAGT.D). Previo a este trabajo, dos grupos reportaron la expresión de la rAGT. Uno de ellos utilizó los genes que codifican para la prepro-AGT y para la pro-AGT para producir a la aglutinina recombinante en *Saccharomyces cerevisiae*, obteniendo un rendimiento de ~1 mg por litro de medio de cultivo (Nagahora H., Ishikawa K, Niwa Y, Muraki M, 1992; Nagahora *et al.*, 1995). El otro grupo utilizó un sistema de baculovirus para lograr la expresión heteróloga en larvas de *Rachiplusia nu* y en líneas celulares del insecto *Sodoptera frugiperda*. De los dos métodos, el que dio mejor rendimiento fueron las larvas de *R. nu*, con ~0.35 mg por gramo de larva (Urtasun *et al.*, 2015). En lo que respecta a los dominios, solo se tiene reportado la expresión del rAGT.B en sistemas bacterianos de *Escherichia coli*, donde se ha

logrado producir como una proteína de fusión junto con un segmento trunco del antígeno de la hepatitis B (Siebert *et al.*, 1996) o alternativamente con el segmento carboxilo terminal del dominio C-LytA (Espinosa *et al.*, 2000).

La producción de todos los dominios recombinantes es algo que no se había reportado, además de que la aglutinina completa nunca se había tratado de obtener a través de un sistema de expresión bacteriano. Lograr la producción de estas proteínas de manera recombinante en *E. coli*, no sólo cumple con llenar los nichos no explorados por otros grupos de investigación, sino que también permite que nos hagamos de un sistema accesible y económico, con la gran ventaja de permitirnos producir a la rAGT y a sus dominios en cantidades lo suficientemente grandes como para seguir analizándolos a través de métodos espectrofotométricos y calorimétricos. Además, la obtención de las construcciones génicas para este proyecto facilitará la posterior obtención de los genes necesarios para la expresión de bidominios y tridominios de la AGT, facultándonos para que a futuro se puedan ampliar los horizontes de este estudio. Sin duda, esto supone un paso necesario para mejorar nuestro entendimiento de esta lectina, algo esencial para lograr la implementación de cambios estructurales que posibiliten mejorar su aplicación en el campo de la biomedicina.

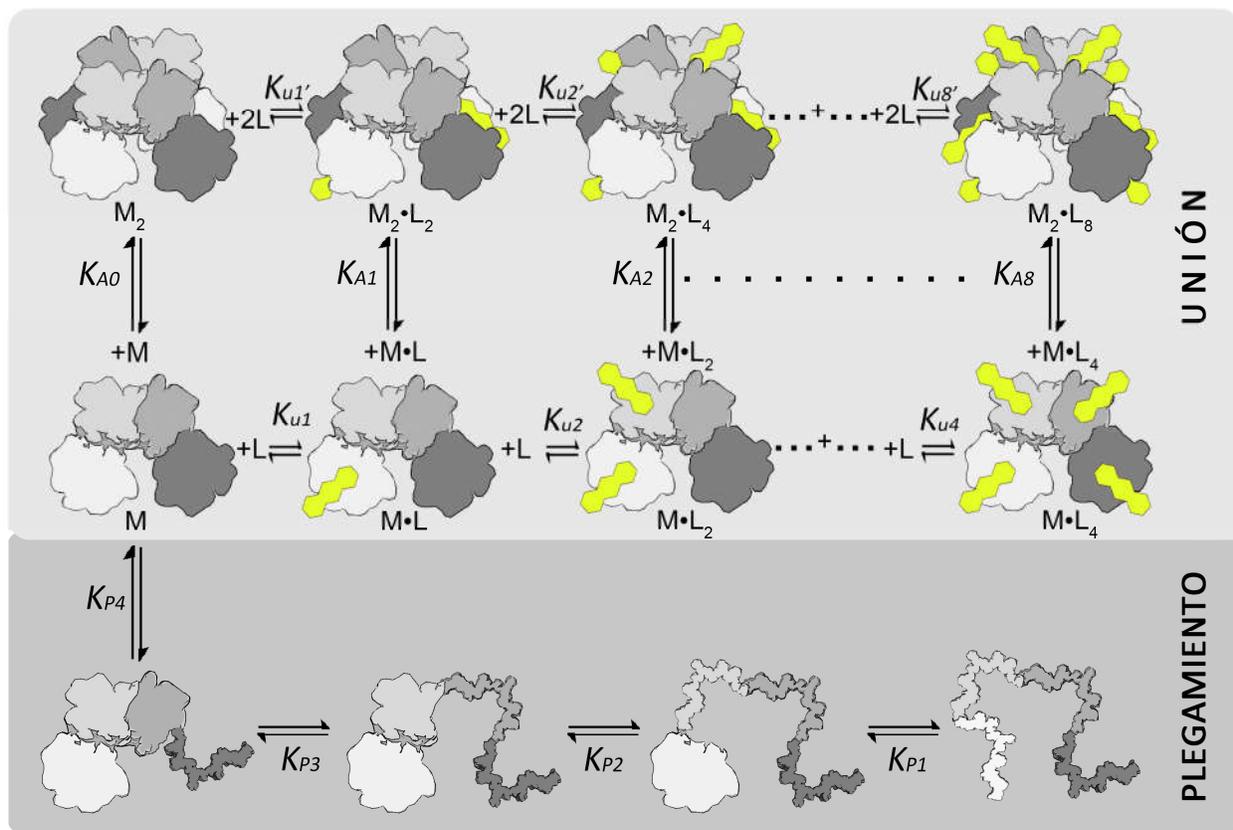


Figura 12. Representación simplificada de los equilibrios involucrados en los eventos de plegamiento, de ensamblaje y de reconocimiento de ligandos de la AGT.  $K_p$  = Constantes de plegamiento,  $K_A$  = Constantes de auto-asociación,  $K_u$  = Constantes de unión.

## HIPÓTESIS

La especificidad del reconocimiento de carbohidratos, su alta estabilidad estructural y su habilidad para realizar transición y atravesar la barrera hematoencefálica, son propiedades que han hecho de la aglutinina de germen de trigo un sistema modelo de gran interés para el desarrollo de diversas aplicaciones biomédicas y biotecnológicas. No obstante, la compleja y dinámica organización estructural de esta lectina ha impedido esclarecer la termodinámica de interacción con ligandos, así como su interdependencia con los procesos de autoasociación y plegamiento. A través del estudio de la aglutinina de germen de trigo completa y de sus dominios aislados producidos de manera recombinante, será posible elucidar propiedades de unión y estabilidad elementales que forman parte del comportamiento global de la proteína.

## OBJETIVOS

### Objetivo general

Elucidar las bases energéticas del proceso de unión a ligandos y estabilidad estructural de la aglutinina de germen de trigo a partir de la caracterización de la lectina completa y de sus cuatro dominios isoestructurales aislados.

### Objetivos específicos

- Producir de manera recombinante a la AGT y a sus 4 dominios aislados utilizando un método de expresión basado en *E. coli*.
- Purificar a homogeneidad cada una de las proteínas recombinantes.
- Realizar una caracterización biofísica a través de métodos espectrofotométricos y calorimétricos de la AGT recombinante y compararla con su correspondiente versión silvestre.
- Mediante una caracterización biofísica, determinar el estado de plegamiento y funcionalidad de los dominios aislados.
- Inferir nuevos aspectos sobre los mecanismos de plegamiento y unión de ligando de la AGT a través del análisis comparativo de la lectina completa y de sus 4 dominios.

## MATERIALES Y MÉTODOS

### Purificación de la AGT silvestre

La aglutinina de germen de trigo silvestre se compró a Sigma Chemical co (St. Louis Missouri). El frasco contiene un liofilizado que comprende una mezcla de las tres isoformas de la lectina. Para separar estas variantes, se llevó a cabo un intercambio catiónico con una columna HiTrap HP SP de 5 ml acoplada a un equipo de cromatografía líquida de alto rendimiento (HPLC), a un flujo de 1 ml min<sup>-1</sup>. La columna se equilibró previamente con un amortiguador 20 mM de ácido fórmico, pH 4. Para cada inyección, se resuspendieron de 10 a 30 mg por cada ml del mismo amortiguador. Cada muestra se centrifugó a 15,000 g por 10 min y se filtró con un corte de paso de 0.22 µm de poro. Para la elución se usó un gradiente lineal de 0 a 750 mM de cloruro de litio (LiCl). La detección se hizo con una lámpara de luz UV operando a una longitud de onda de 280 nm. Se colectaron las fracciones correspondientes a cada una de las de las isoformas, se dializaron en agua pH 4 y se sometieron a un proceso de liofilización.

### Clonación de los genes de la AGT y de sus cuatro monodominios

La secuencia codificante para la isoforma 1 de la AGT se sintetizó comercialmente (Epoch Life Science, Inc., Missouri City, Texas). Se usó como base la secuencia registrada con el número de acceso M25536.1 del GenBank, la cual se optimizó para su expresión en *E. coli*. El gen se insertó en el plásmido de clonación pBSK GS 45510 (Fig. 13A), a partir del cual se amplificaron mediante PCR las secuencias correspondientes a la AGT completa y a sus cuatro monodominios. Para cada caso se utilizó un par específico de oligonucleótidos. Cada oligonucleótido está conformado por una secuencia complementaria al gen de la AGT y por un sitio de restricción para *NcoI* o para *BamHI*. Esto se hizo para insertar estas secuencias de restricción en los extremos 5' y 3' de cada uno de los genes de la AGT. El oligonucleótido Fwd para los genes de la rAGT y la rAGT.A contiene un sitio de corte adicional para enterocinasa (EK) y 4 residuos de alanina para facilitar el corte proteolítico. La secuencia de los oligonucleótidos utilizados se puede revisar en la Tabla 3.



se identificaron como positivos, fueron purificados utilizando el "QIAprep Spin miniprep kit" (Qiagen, USA) y secuenciados (Laragen Inc., USA).

## Expresión y purificación de las proteínas recombinantes

Las construcciones que se elaboraron con el vector pET-32a(+), se transformaron en células quimiocompetentes de Rosetta-gami (DE3) pLysS y Origami. Las transformantes de Rosetta-gami se crecieron en medio LB suplementado con ampicilina ( $100 \mu\text{g ml}^{-1}$ ), cloranfenicol ( $34 \mu\text{g ml}^{-1}$ ), kanamicina ( $25 \mu\text{g ml}^{-1}$ ) y tetraciclina ( $15 \mu\text{g ml}^{-1}$ ), hasta alcanzar una D.O.<sub>600</sub> de 0.6. Las células Origami se trataron de la misma forma, aunque omitiendo la adición de cloranfenicol. Con estos cultivos se hicieron varias pruebas de expresión a  $30^\circ\text{C}$  con  $0.5 \text{ mM}$  de Isopropil- $\beta$ -D-tiogalactopiranosido (IPTG), utilizando distintos tiempos de inducción. Para la purificación, los cultivos de Rosetta-gami se escalaron a volúmenes de 2 o 3 l de medio LB suplementado con ampicilina ( $100 \mu\text{g ml}^{-1}$ ), cloranfenicol ( $34 \mu\text{g ml}^{-1}$ ), kanamicina ( $25 \mu\text{g ml}^{-1}$ ) y tetraciclina ( $15 \mu\text{g ml}^{-1}$ ). La inducción de la proteína se realizó cuando los cultivos llegaron una D.O.<sub>600</sub> de 0.6, adicionando  $0.5 \text{ mM}$  de IPTG y manteniendo una temperatura de  $30^\circ\text{C}$  durante 16 h. Al término de la inducción, los cultivos se cosecharon centrifugando a  $6,000 \text{ g}$  durante 10 min y las pastillas se resuspendieron en amortiguador de lisis (pH 8, Tris  $50 \text{ mM}$ ,  $250 \text{ mM NaCl}$ ,  $1 \text{ mM PMSF}$ ). Al amortiguador de lisis que se ocupó para la pastilla de la lectina recombinante completa se le adicionó  $250 \text{ mM}$  de urea y se le ajustó el pH a 9 para evitar problemas de agregación. Las muestras se lisaron con un sonicador (Misonix 3000), utilizando intervalos de 10 s con espacios de 45 s hasta completar 10 min, todo a  $4^\circ\text{C}$ . El lisado se centrifugó a  $15,000 \text{ g}$  durante 45 min. Las fracciones solubles de estos lisados se filtraron con una membrana con tamaño de poro de  $0.22 \mu\text{M}$  y se sometieron a cromatografía de afinidad con metales inmovilizados (IMAC) para aprovechar la afinidad de la etiqueta de histidinas de las proteínas de fusión por el níquel. Para esto se utilizó una columna de níquel HisTrap FF (GE, Healthcare), previamente equilibrada con amortiguador de lisis al que se le añadió  $30 \text{ mM}$  de imidazol, con un flujo de  $1 \text{ ml min}^{-1}$ . Las proteínas recombinantes se eluyeron de la columna aumentando la concentración de imidazol a  $300 \text{ mM}$ . Las fracciones correspondientes a las eluciones se dializaron en amortiguador (pH 8, Tris  $50 \text{ mM}$ ,  $50 \text{ mM NaCl}$ ) y se cortaron con EK (Invitrogen, Carlsbad, California) para remover la tiorredoxina y la etiqueta de histidinas. Las reacciones de corte se realizaron a una concentración final de  $0.5 \text{ mM}$  de  $\text{CaCl}_2$ ,  $0.1 \%$  de tween-20 y una proporción de  $0.001 \text{ U}$  de EK, por cada  $20 \text{ mg}$  de proteína de fusión. Esta mezcla se incubó a  $37^\circ\text{C}$ , por 16 h, para las proteínas de fusión de los dominios, mientras que, para la proteína de fusión de la aglutinina completa, la reacción de corte se llevó a cabo a una temperatura de  $24^\circ\text{C}$ , por 24 h, para evitar problemas de agregación. Los productos de las reacciones de corte se sometieron a una segunda columna de níquel después de filtrarse con una membrana con tamaño de poro de  $0.22 \mu\text{M}$ . De

esta segunda columna de níquel, se colectaron las fracciones no retenidas, se dializaron en agua destilada y liofilizaron. Los dominios rAGT.B, rAGT.C y rAGT.D se resuspendieron en ácido trifluoroacético (TFA) al 50 % para tratarse con bromuro de cianógeno (CNBr) en relación molar 10:1 para remover los residuos extra de alanina y metionina. La reacción se dejó a temperatura ambiente por 12 h. Finalmente, estas fracciones se purificaron mediante cromatografía líquida de fase reversa (RP-HPLC) en un cromatógrafo modelo HP-1100 (HP, Palo Alto, California). Se utilizó una columna C4 para la AGT completa y una columna XB-C18 (Ambas de Phenomenex, Torrance, California) para los monodominios recombinantes. La elución se realizó con un gradiente lineal de agua-acetonitrilo (con 0.05 % de ácido trifluoroacético) del 15 al 100 % a un flujo de 1 ml min<sup>-1</sup>. La detección se hizo con una lámpara de luz UV operando a una longitud de onda de 280 nm. Las muestras de estos últimos pasos de purificación se analizaron con espectrometría de masas de desorción/ionización mediante láser asistida por matriz, acoplada a un analizador de tiempo de vuelo (MALDI-TOF) para comprobar su pureza.

## Espectroscopia de dicroísmo circular

Los espectros de dicroísmo circular (DC) se tomaron con un espectropolarímetro JASCO J-720, equipado con un peltier PTC-348WI para controlar la temperatura. Se ocupó una celda de cuarzo con una longitud de trayectoria óptica de 0.1 cm. Cada medición se realizó a una temperatura de 20 °C en la región del UV lejano (195 a 250 nm), con una velocidad de barrido de 50 nm min<sup>-1</sup> y un tiempo de respuesta de 16 s. Las muestras se analizaron a concentraciones cercanas a los 0.5 mg ml<sup>-1</sup> de proteína, en amortiguadores de glicina-HCl (para pH de 2.0), acetato (para pH de 4.7) o citrato (para pH de 6.0), todos a una concentración de 50 mM y suplementados con 100 mM de NaCl. Para cada espectro, se obtuvo el promedio de tres repeticiones y se le restó la señal del blanco correspondiente. Los espectros se reportaron en grados dmol<sup>-1</sup> cm<sup>-2</sup>, que son las unidades correspondientes a la elipticidad molar media por residuo ( $[\theta]_{mrw}$ ), la cual se calculó utilizando la ecuación 7:

$$[\theta]_{mrw} = \frac{\theta * 100 * MRW}{10 * [P] * L} \quad (7)$$

donde  $\theta$  es la elipticidad en grados, L es el paso de luz de la celda en cm, [P] es la concentración de la proteína en mg ml<sup>-1</sup> y MRW es la masa molar media por residuo. El MRW puede calcularse dividiendo la masa molar (M) entre el número de aminoácidos en la cadena polipeptídica (n) menos 1, como se indica en la ecuación 8:

$$MRW = \frac{M}{n - 1} \quad (8)$$

Para obtener las transiciones por desnaturalización térmica, las muestras se sometieron a barridos de temperatura que abarcaron de los 20 a los 90 °C. Se utilizó una velocidad de barrido de 1 °C min<sup>-1</sup> y se siguió el cambio en la señal de elipticidad a 225 nm, que corresponde al máximo positivo del espectro de dicroísmo de la aglutinina de germen de trigo. Para evaluar el grado de reversibilidad del desplegamiento de las proteínas, se tomaron mediciones (espectros de 195 a 250 nm) al alcanzar los 90 °C y a los 20 °C, una vez que las muestras regresaran a la temperatura inicial.

### **Calorimetría diferencial de barrido**

La isoforma 1 de la AGT silvestre (wAGT) y la recombinante (rAGT) fueron analizadas mediante calorimetría diferencial de barrido (DSC). Estas mediciones se llevaron a cabo con un microcalorímetro capilar VP-DSC (Malvern Panalytical Inc., Westborough, Massachusetts). Las muestras se prepararon a concentraciones de proteína dentro de un intervalo de 0.05 a 0.11 mM. Todos los experimentos se llevaron a cabo a una velocidad de barrido de 1 °C min<sup>-1</sup>, con una temperatura inicial de 20 °C y una temperatura final de 110 °C. Las líneas base se obtuvieron de corridas de amortiguador contra amortiguador, bajo las mismas condiciones experimentales. Estas líneas fueron restadas de las endotermas obtenidas para cada muestra. Para garantizar que el instrumento alcanzara una historia térmica apropiada, se corrieron varias líneas base amortiguador-amortiguador previo a las mediciones de las soluciones con proteína.

### **Experimentos de Resonancia Magnética Nuclear**

El dominio rAGT.D se disolvió a una concentración de 0.5 mM, en H<sub>2</sub>O con D<sub>2</sub>O al 5% (Cambridge Isotopes Laboratories, Tewsbury, Massachusetts). Se obtuvieron datos de Resonancia Magnética Nuclear (NMR) para esta muestra a una temperatura de 25 °C en un espectrómetro Bruker ASCEND 700 (Bruker BioSpin, Billerica, Massachusetts) a una frecuencia de 700 MHz con una criosonda de 3 núcleos (TCI 700 H-C/N-D 5 mm). Se realizó "Espectroscopia de efecto nuclear de Overhauser" (NOESY). El espectro de NOESY se registró a un tiempo de mezclado de 300 ms y se colectó empleando 1024 datos de matrices complejas obtenidos de 32 escaneos. Los datos de NMR fueron procesados usando el programa MestreC (MestreC Lab Research, Santiago de Compostela, España) y una función de ventana seno desplazada y un llenado cero aplicado previo a la transformada de Fourier.

## Calorimetría de titulación isotérmica

Las mediciones de calorimetría de titulación isotérmica (ITC) se realizaron con un equipo MicroCal ITC<sub>200</sub> (GE Healthcare). Se hicieron mediciones para la wAGT, la rAGT y para los cuatro monodominios recombinantes. Todas las muestras se prepararon en amortiguador 0.05 M de acetato suplementado con 100 mM de NaCl, a pH 4.7 y temperatura de 30 °C. Los experimentos con los dominios se llevaron a cabo a una concentración aproximada de 1 mM de proteína, mientras que los experimentos para las versiones completas de la aglutinina se realizaron con concentraciones cercanas a 0.15 mM del monómero. El protocolo consistió de 20 a 25 inyecciones consecutivas de ligando, con intervalos de espera de 5 minutos entre cada inyección, todo realizado con una velocidad de agitación de 750 rpm. El calor de dilución del ligando se obtuvo adicionando éste a la solución amortiguadora en condiciones similares a las utilizadas para las muestras de proteína y ocupando los mismos tiempos de inyección que se ocuparon para éstas. Todas las muestras se degasificaron por 10 min antes de realizar el experimento.

Los parámetros de unión se determinaron ocupando un único e independiente modelo de unión:

$$Q = \frac{nP_t \Delta H_U V_0}{2} \left[ 1 + \frac{L_t}{nP_t} + \frac{1}{nK_U P_t} - \sqrt{\left(1 + \frac{L_t}{nP_t} + \frac{1}{nK_U P_t}\right)^2 - \frac{4L_t}{nP_t}} \right]$$

donde  $Q$  es el calor normalizado por mol de ligando al final de la  $i$ -ésima inyección,  $K_U$  es la constante de unión,  $\Delta H_U$  es el cambio en la entalpía,  $n$  es la estequiometría,  $V_0$  es el volumen de trabajo de la celda y  $L_t$  y  $P_t$  son las concentraciones totales de ligando y proteína, respectivamente. El calor liberado en la  $i$ -ésima inyección es:

$$\Delta Q(i) = Q(i) + \frac{v_i(Q(i) + Q(i-1))}{2V_0} - Q(i-1) + q_{dil}$$

donde  $v_i$  es el volumen de la alícuota añadida en la  $i$ -ésima inyección y  $q_{dil}$  es un término de ajuste introducido para tomar en cuenta calores de dilución experimentales no corregidos (Velazquez-Campoy, Leavitt y Freire, 2004; Pulido *et al.*, 2008). El ajuste de los datos se hizo ocupando Origin 7.0 (OriginLab Co., Northampton, Massachusetts).

## Ensayos de hemaglutinación

Para tratar los eritrocitos de humano (tipo O+), se utilizó una solución salina amortiguada por fosfatos (PBS), que contiene 50 mM de fosfatos y 150 mM de NaCl a pH 7.4. Este amortiguador se ocupó para realizar varios lavados de un volumen aproximado de 20 ml de sangre hasta obtener un sobrenadante transparente. Finalmente, los eritrocitos se resuspendieron a una concentración final de 20 % v/v. Los ensayos se llevaron a cabo en placas de 96 micropozos con fondo cóncavo. Se preparó la rAGT a una concentración de  $0.1 \text{ mg ml}^{-1}$ , se agregaron 100  $\mu\text{l}$  de esta muestra en el primer pozo y a partir del cual se realizaron múltiples diluciones seriadas con factor de dilución 1:2 hasta completar la fila. De manera similar, se prepararon los pozos para el control positivo y para el control negativo. Como control positivo se utilizó a la wAGT (también a una concentración de  $0.1 \text{ mg ml}^{-1}$ ) y como control negativo solo se adicionó amortiguador de PBS. Por último, a cada pozo se le añadió 50  $\mu\text{l}$  de la solución con eritrocitos y se dejó a temperatura ambiente durante 1.5 h.

## RESULTADOS

### Purificación de la AGT nativa

En la Fig. 14 se puede observar uno de los cromatogramas obtenidos para las corridas de HPLC en las que se realizó intercambio catiónico para separar las isoformas de la mezcla comercial de AGT. Los picos del cromatograma aparecen bien definidos y la distancia entre picos es apropiada, por lo que se puede decir que se logró separar de manera eficiente las tres isoformas. El primer pico corresponde a la isoforma 1, la cual eluyó a los 30 min a una concentración aproximada de 30 % de LiCl. El segundo y tercer pico corresponden a las isoformas 2 y 3, respectivamente. La isoforma 2 eluyó a los 54 min a una concentración de 36 % de LiCl, mientras que la isoforma 3 eluyó a los 76 min a una concentración de 42 % de LiCl. Las fracciones correspondientes a cada isoforma se colectaron por separado, se dializaron y se liofilizaron para su uso o almacenamiento. Antes de ocupar alguna de estas muestras en otros análisis, estas se resuspendieron y se analizaron por dicroísmo circular para corroborar que no hubiesen ocurrido cambios en la estructura secundaria durante la liofilización y/o el almacenamiento.

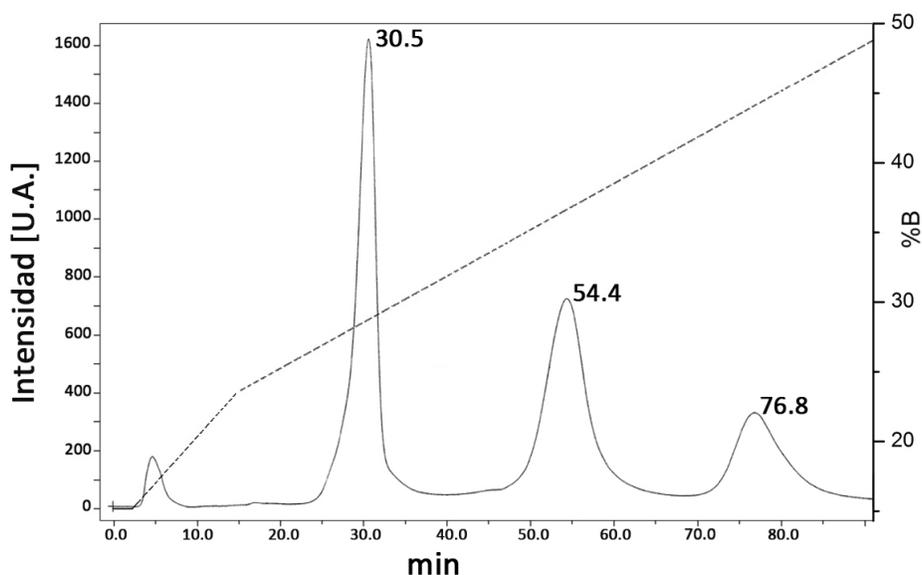


Figura 14. Perfil cromatográfico de las isoformas de la AGT separadas por cromatografía líquida de intercambio catiónico. Purificación realizada en amortiguador de ácido fórmico 20 mM a pH 4.0. Se puede ver en el cromatograma los tres picos correspondientes a cada una de las isoformas de la AGT. La línea punteada indica el gradiente de LiCl.

## Clonación y expresión de la AGT recombinante

Los genes que codifican para la rAGT (de 525 pb) y para sus dominios rAGT.A, rAGT.B, rAGT.C y rAGT.D (de 156, 129, 129 y 129 pb, respectivamente) se amplificaron mediante PCR utilizando el gen sintético de la AGT dentro del vector pBSK y pares específicos de oligonucleótidos que contenían las secuencias de restricción para *NcoI* y *BamHI* en los extremos 5' y 3', respectivamente. Los amplificados se introdujeron en el vector de expresión pET-32a(+), conforme a la estrategia experimental descrita en Materiales y Métodos.

Se decidió utilizar el vector pET-32a(+) debido a que, además de ser un vector apropiado para la expresión por contar con el promotor T7, tiene al gen *trxA* que codifica para la tiorredoxina. Esto se hizo con la finalidad de expresar a la AGT y a sus dominios aislados como proteínas de fusión junto a esta proteína. La tiorredoxina se ha usado previamente en la expresión de otras proteínas ricas en puentes disulfuro, donde se ha visto que su presencia favorece la correcta formación de los puentes disulfuro. El vector también cuenta con una región que codifica para una etiqueta de histidinas, la cual se utilizó para facilitar la purificación con una columna de níquel. El vector cuenta con los sitios de corte para trombina y EK, siendo este último el que se escogió para llevar a cabo la eliminación del segmento que contiene la tiorredoxina y la etiqueta de histidinas en la proteína de fusión. El diseño final de estas construcciones se puede observar en la Fig. 15. Se puede apreciar que para la AGT completa y el dominio A se introdujeron 4 alaninas después del sitio de corte para EK, con el fin de facilitar el corte proteolítico. Mientras que para los dominios B, C y D, sólo se quedaron 2 residuos después del sitio de corte para EK, una alanina y una metionina.

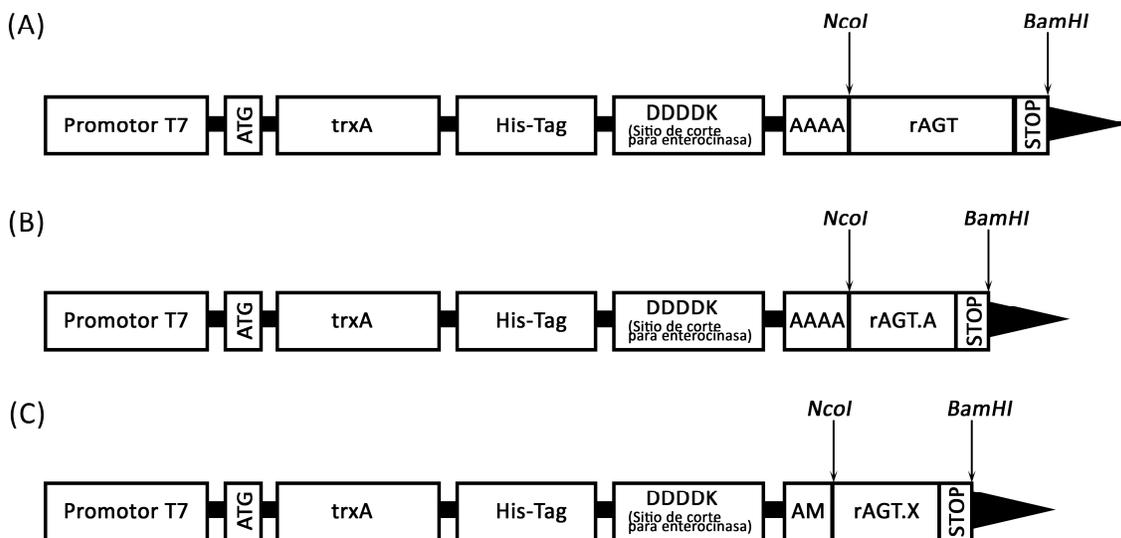
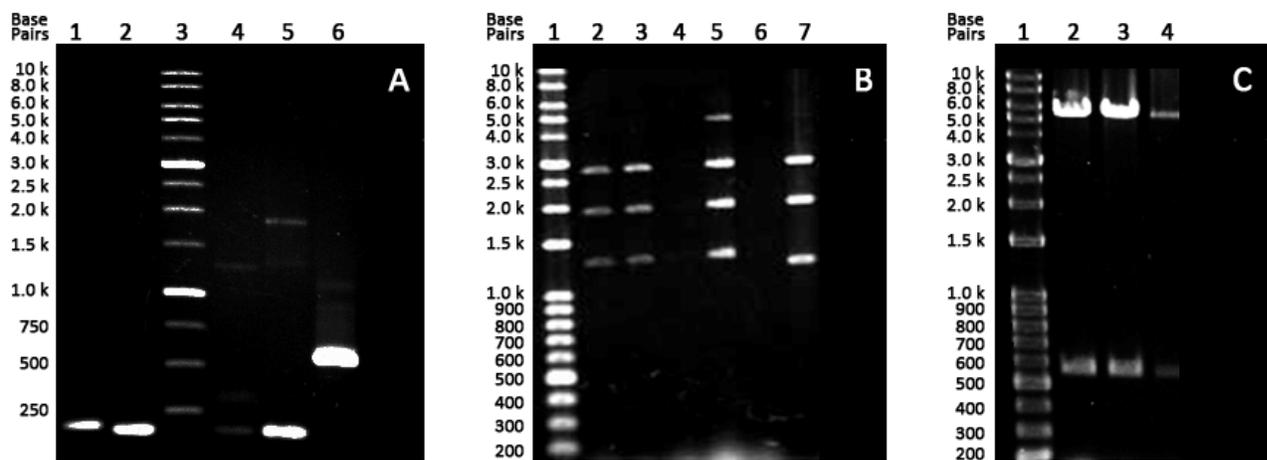


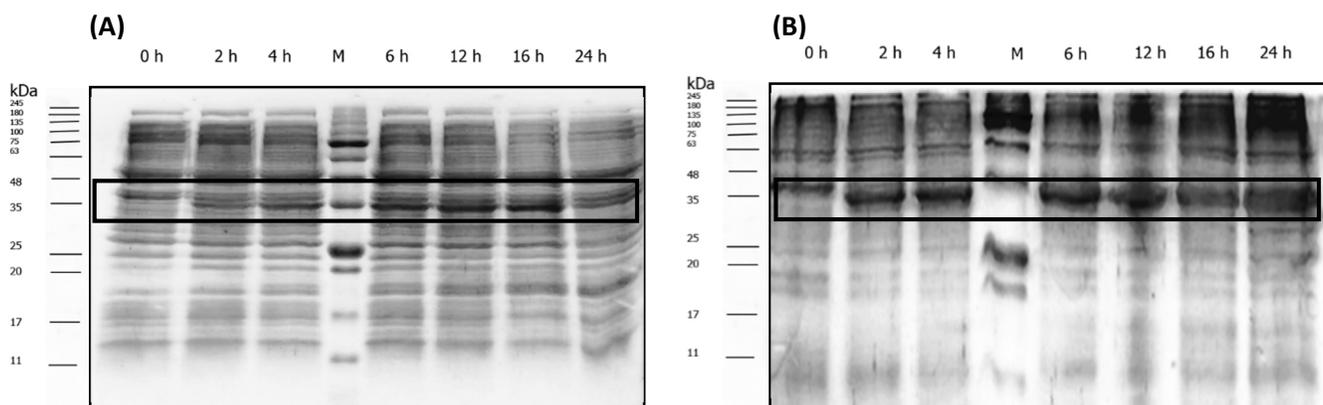
Figura 15. Construcciones para la AGT completa y sus dominios aislados dentro del vector de expresión pET32-a(+). (A) Construcción génica de la proteína de fusión de la AGT completa. (B) Construcción génica de la proteína de fusión del dominio A. (C) Construcción génica de las proteínas de fusión de los otros dominios. Donde X puede ser B, C o D.

En la Fig. 16A se puede ver que el peso obtenido para cada producto de amplificación corresponde al esperado. Una vez obtenidas las construcciones a partir de estos productos de PCR, para cada una de estas se obtuvieron transformantes de *E. coli* DH5 $\alpha$  seleccionadas en cajas suplementadas con ampicilina. Después de la extracción del plásmido, se realizaron las correspondientes reacciones de digestión para confirmar el tamaño de los plásmidos. Las construcciones de los dominios recombinantes se digirieron con *XmnI*. En la Fig. 16B, los carriles 2, 3, 5 y 7, presentan bandas que corresponden en peso molecular a la restricción de los plásmidos de cada dominio recombinante. Para cada uno se obtuvo un patrón de tres o más bandas, debido a la presencia de más de un sitio de corte para *XmnI* dentro de sus secuencias. Los plásmidos con el gen que codifica para la AGT completa fueron analizados con las enzimas de restricción *NcoI* y *BamHI*. En la Fig. 16C, se puede observar que los carriles 2, 3 y 4 dieron resultado positivo, mostrando la presencia de una banda con un peso aproximado de 561 pb, correspondiente al gen completo de la AGT y otra banda con un peso de 5,900 pb correspondiente al vector.



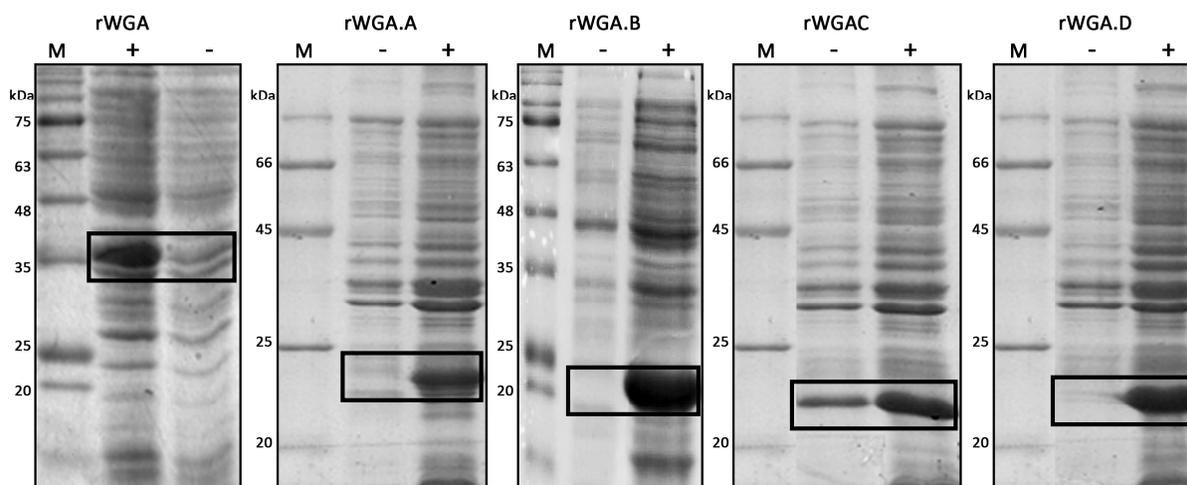
**Figura 16. Proceso de clonación de los genes de AGT y sus dominios aislados.** (A) Amplificación por PCR de los genes de la AGT. Carriles 1: rAGT.A, 2: rAGT.B, 3: marcador de peso molecular, 4: rAGT.C, 5: rAGT.D, 6: rAGT. (B) Reacción de digestión con *XmnI*. Carriles 1: marcador de peso molecular, 2: rAGT.C, 3: rAGT.D, 5: rAGT.A, 7: rAGT.B. (C) Reacción de digestión con *NcoI* / *BamHI*. Carriles 1: marcador de peso molecular, 2, 3 y 4: rAGT.

Las construcciones obtenidas con el vector pET-32a(+) se transformaron en células quimiocompetentes de Rosetta-gami y Origami. Ambas cepas presentan mutaciones en los genes de la tiorredoxina reductasa (*trxB*) y glutatión reductasa (*gor*), lo que genera condiciones oxidantes en el citoplasma que sirven para que los puentes disulfuro se formen de manera posterior al plegamiento y se favorezca la correcta conectividad entre éstos. La única diferencia entre estas cepas es que Rosetta-gami cuenta con el suministro de tRNAs para codones raros, permitiéndole a la célula mejorar el proceso de traducción al no estar limitada por el uso de codones. Esto no debería afectar la expresión en nuestro caso, dado que el gen se encuentra optimizado para su expresión en *E. coli*. A pesar de ello, fue en las células de Rosetta-gami en las que se obtuvo una mayor cantidad de proteína en la fracción soluble. Las mejores condiciones de inducción se alcanzaron utilizando 0.5 mM de IPTG por 16 h a una temperatura de 30 °C. En la Fig. 17 se puede ver la banda correspondiente a la proteína de fusión para la aglutinina completa (rAGT-Trx-His), con un peso aproximado de 34 kDa. De acuerdo a lo observado, se estima que una cuarta parte de la proteína permanece en la fracción soluble (Fig. 17A), mientras que el resto queda atrapado en la fracción insoluble (Fig. 17B).



**Figura 17. Cinética de expresión de la rAGT en Rosetta-gami.** Se resalta en un cuadro la posición de la proteína de fusión. (A) Fracción soluble. (B) Fracción no soluble.

La mejor expresión para las proteínas de fusión de los dominios recombinantes también se logró utilizando células de Rosetta-gami inducidas con 0.5 mM de IPTG por 16 h a 30 °C. En la Fig. 18, se puede observar la inducción para cada uno de estos dominios. Basándose en su estructura primaria, la proteína de fusión de la aglutinina completa tiene un peso teórico aproximado de 35 kDa, mientras que las proteínas de fusión de los dominios pesan de 21 a 22 kDa. En la Tabla 4 se muestran de manera más precisa los pesos teóricos de cada una de las proteínas de fusión.



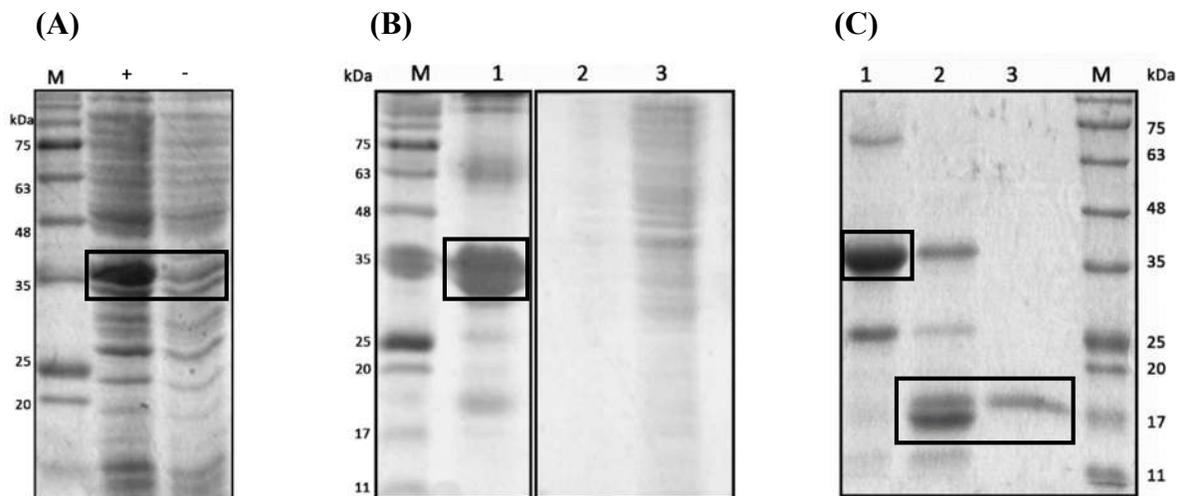
**Figura 18. Inducción de la rAGT y sus dominios recombinantes tras 16 h de inducción con IPTG.** De izquierda a derecha se muestra la recombinante completa, seguida de cada uno de sus 4 dominios. (-) Sin inducir. (+) Con IPTG al 0.5 mM.

**Tabla 4. Pesos teóricos de las proteínas de fusión de la AGT y de sus dominios recombinantes.**

Proteína de fusión	Peso teórico estimado (Da)
rAGT-Trx-His	35,262
rAGT.A-Trx-His	22,610
rAGT.B-Trx-His	21,535
rAGT.C-Trx-His	21,495
rAGT.D-Trx-His	21,453

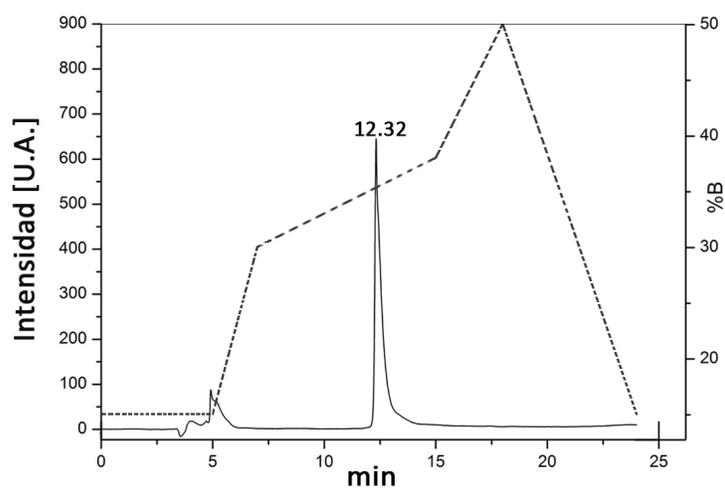
## Purificación de la AGT recombinante

Los cultivos de *E. coli* Rosetta-gami, que contenían el plásmido para producir la AGT completa, se escalaron a volúmenes de 2 y 3 l. Estos cultivos se indujeron con 0.5 mM de IPTG, a 30 ° C por 16 h, que fueron las mejores condiciones de inducción que se encontraron durante los ensayos de expresión. En la Fig. 19<sup>a</sup>, se puede ver la inducción de la proteína de fusión (rAGT-Trx-His) de un cultivo de 2 l. Después de cosechar, lisar la pastilla y obtener la fracción soluble por centrifugación, ésta se pasó a través de una columna níquel. La proteína de fusión retenida en la columna, se eluyó con una solución de imidazol a una concentración de 300 mM. La elución correspondiente a esta primera columna se puede observar en el carril 1 de la Fig. 19B. Esta fracción se mostró enriquecida con una proteína de peso aproximado de 34 kDa, que coincide con el peso teórico de la proteína de fusión de la rAGT. El segundo y tercer carril corresponden al lavado y a la fracción no retenida de la columna, respectivamente.



**Figura 19. Expresión y purificación de la AGT recombinante.** (A) Inducción de la proteína de fusión. (-) Sin inducir. (+) Con IPTG al 0.5 mM. (B) Primer columna de níquel. Carriles 1: elución, 2: lavado, 3: no retenido. (C) Reacción de corte con EK y segunda columna de níquel. Carriles 1: proteína de fusión, 2: producto de la reacción de corte, 3: no retenido de la segunda columna de níquel.

Posteriormente, se dializó la elución de esta primera columna en un amortiguador similar al amortiguador de lisis, pero con una fuerza iónica menor, con el fin de obtener las condiciones necesarias para llevar a cabo la reacción de corte con EK y así separar la rAGT de la tiorredoxina y de la etiqueta de histidinas. Una vez dializada, se dejó la reacción de corte a 24 °C durante 24 h. Este proceso se hizo a una temperatura menor y durante más tiempo que el utilizado para los dominios, para evitar problemas de agregación. Después de realizado el corte, la muestra se centrifugó a 15,000 g por 15 min y se filtró con un corte de paso de 0.22 nm, antes de pasarla por una segunda columna de afinidad de níquel. La Fig. 19C muestra en el primer carril a la proteína de fusión, en el segundo carril la reacción de corte con EK y en el tercer carril el no retenido de la segunda columna de níquel, que corresponde a la AGT recombinante.

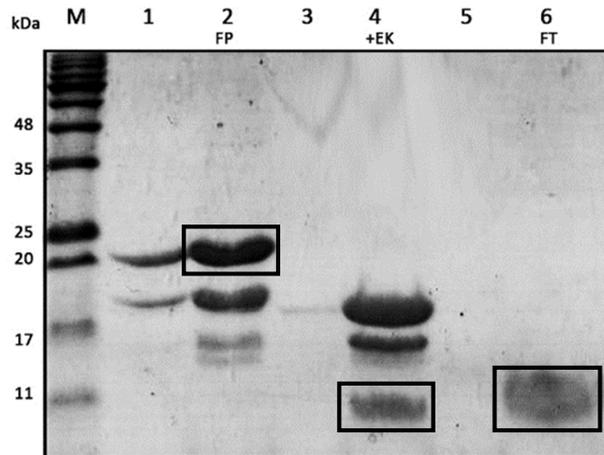


**Figura 20. Cromatograma de la purificación de la rAGT realizada por HPLC de fase reversa en una columna C4.** El gradiente lineal de acetonitrilo se muestra con una línea punteada. La elución de la rAGT se logró a una concentración de aproximadamente 36% de acetonitrilo.

La proteína obtenida de la segunda columna de níquel se purificó con un paso más, que consistió en realizar HPLC de fase reversa utilizando una columna C4. Se utilizó un gradiente lineal de agua-acetonitrilo que inicialmente aumenta del 15 al 30% y luego del 30 al 38%. En esta última parte eluyó la proteína, aproximadamente a los 12 min de corrida. En un último paso se llevó la concentración de acetonitrilo, hasta el 50% para liberar cualquier remanente que hubiese quedado pegado a la columna. En la Fig. 20, se puede ver el cromatograma correspondiente a la cromatografía de fase reversa, donde se observa la aparición de un pico mayoritario a los 12 min.

## Purificación de los dominios recombinantes de la AGT

Para los dominios recombinantes se siguió una estrategia de purificación similar al que se utilizó para la lectina recombinante completa. Las fracciones solubles del lisado celular se pasaron a través de una columna de níquel. Tras eluir, las proteínas de fusión se cortaron con EK y se pasaron por una segunda columna de níquel. En la Fig. 21 se puede ver el resultado de esta estrategia para el dominio A.



**Figura 21. Purificación del dominio A recombinante de la AGT.** Carriles 2: proteína de fusión obtenida durante la elución de la primera columna de níquel, 4: Reacción de corte con EK, 6: No retenido de la segunda columna de níquel. Los otros carriles muestran fracciones no retenidas o de lavados de la columna.

Los dominios obtenidos después de la segunda columna de níquel se dializaron en agua destilada y liofilizaron. Los dominios rAGT.B, rAGT.C y rAGT.D fueron tratados con CNBr para remover los residuos de alanina y metionina que se encontraban al inicio de sus cadenas polipeptídicas. Con esto, las secuencias de aminoácidos de estos dominios quedaron sin residuos adicionales e idénticas a la secuencia de la AGT silvestre. Por último, se purificaron mediante HPLC de fase reversa utilizando una columna C18. Se utilizó un gradiente lineal de agua-acetonitrilo del 12 al 30%. El pico correspondiente a los dominios apareció de los 25 a los 30 min de corrida a concentraciones del 20 al 26 % de acetonitrilo. El cromatograma correspondiente al dominio A se puede observar en la Fig. 22.

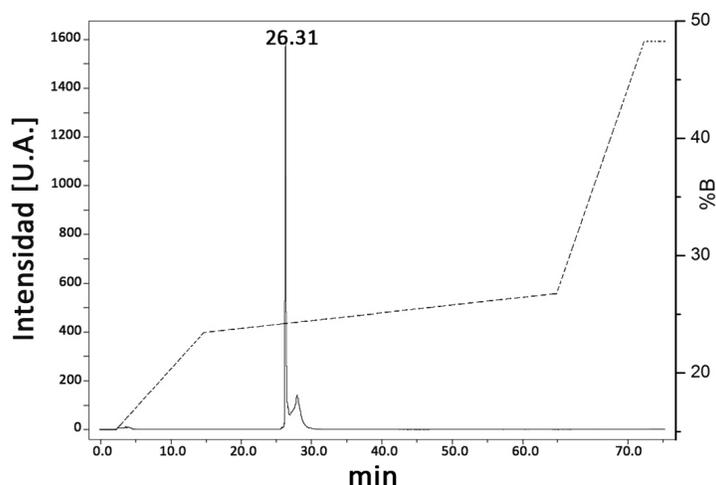


Figura 22. Cromatograma de la purificación del dominio A realizada por HPLC de fase reversa en una columna C18. El gradiente lineal de acetonitrilo se muestra con una línea punteada. La elución de la rAGT.A se logró a una concentración de aproximadamente 24% de acetonitrilo.

## Análisis de espectrometría de masas MALDI-TOF

El análisis de MALDI-TOF MS mostró que las masas moleculares obtenidas para cada una de las proteínas recombinantes purificadas coinciden con el peso teórico esperado. Mismo que se calculó utilizando su secuencia primaria (Fig. 23) y asumiendo que todas las cisteínas están oxidadas. En la Fig. 24, se muestran los perfiles de masas obtenidos para la proteína completa y para uno de los dominios. Todos los resultados teóricos y experimentales (MALDI-TOF MS) que se obtuvieron para cada una de las proteínas, se muestran de forma resumida en la Tabla 5.

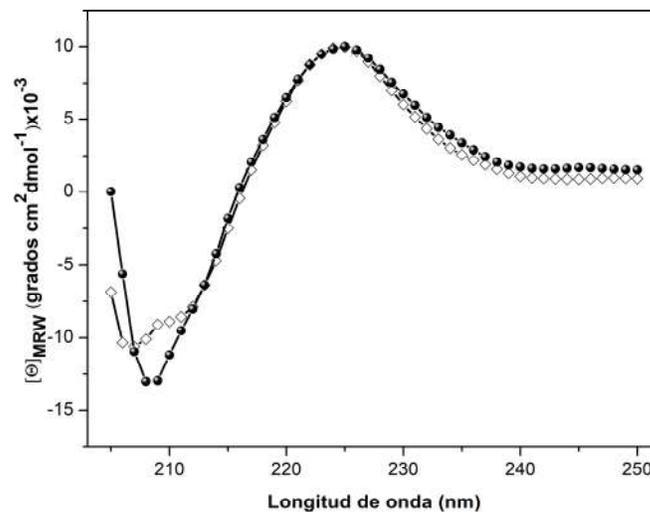
Tabla 5. Resultados de MALDI-TOF MS.

Proteína	MALDI Masa de la Proteína de fusión (Da)	MALDI Masa después del corte con EK/CNBr (Da)	Masa esperada (Da)	Rendimiento (mg por litro de medio de cultivo)
rAGT	35,262	17,382	17,382	4.8
rAGT.A	22,610	4,754	4,754	3.6
rAGT.B	21,535	4,268	4,269	3.0
rAGT.C	21,495	4,228	4,228	3.8
rAGT.D	21,453	4,186	4,186	2.9



## Caracterización biofísica de la aglutinina recombinante completa.

Para evaluar la integridad conformacional de la aglutinina recombinante, se llevó a cabo la caracterización biofísica de ésta y la comparación con la aglutinina silvestre. En la Fig. 25, se muestra el espectro de dicroísmo circular que se obtuvo en el UV-lejano para la rAGT, el cual resultó prácticamente superponible al de la wAGT. Esto indica que las dos proteínas tienen una composición de estructura secundaria muy similar. El pico positivo que aparece a los ~225 nm, es característico de las proteínas tipo heveína. Acorde a lo esperado, esta banda desapareció al incubar a la rAGT en presencia de un agente reductor o a elevadas temperaturas durante el ciclo de calentamiento-enfriamiento al que se sometió (Fig. 27). Gran parte de la señal de dicroísmo a esta longitud de onda se recuperó tras dejar que la muestra se enfriara y regresara a temperatura ambiente. Por lo que, al menos de acuerdo a esta técnica, el desplegamiento de la lectina silvestre y el de la lectina recombinante, son procesos reversibles ya que estos presentan una reversibilidad  $\geq 90\%$ . De acuerdo al ciclo de calentamiento al que fueron sometidas, ambas lectinas mostraron una termorresistencia similar. Esto se puede observar en la Fig. 26, donde se aprecia que la transición de desplegamiento empieza aproximadamente a los 65 °C y que esta no llega a concluir al alcanzar los 90 °C. En este caso, los puentes disulfuro son los principales responsables de la elevada resistencia a la desnaturalización térmica que muestra la aglutinina de germen de trigo.



**Figura 25. Espectros de dicroísmo en el UV-lejano de la rAGT y la wAGT.** Se compara el espectro en el UV-lejano de la AGT comercial con el de su versión recombinante, ambas a 20 °C en amortiguador de acetatos 50 mM con 0.1 M de NaCl a pH 4.7 y a concentraciones de proteína a  $\sim 0.5 \text{ mg ml}^{-1}$ . (●) rAGT, (◇) wAGT.

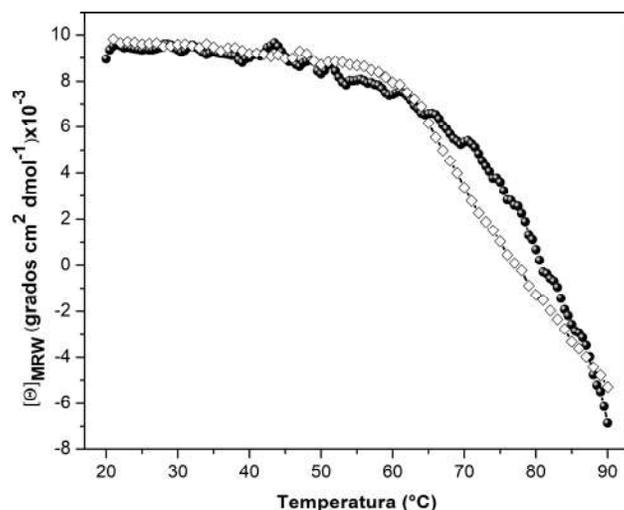


Figura 26. Curvas de perturbación térmica para la rAGT y la wAGT. Las muestras se calentaron de 20 °C a 90 °C, todo el proceso se siguió por monitoreo de la elipticidad a 225 nm, en un amortiguador de acetatos 50 mM con 0.1 M de NaCl a pH 4.7, a concentraciones de proteína a ~0.5 mg ml<sup>-1</sup> y con una velocidad de barrido de 1 °C min<sup>-1</sup>, (●) rAGT, (◇) wAGT.

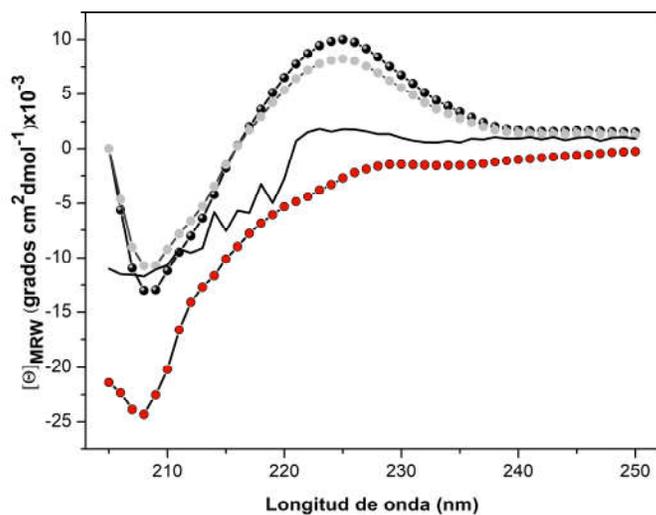


Figura 27. Espectros de dicroísmo en el UV-lejano para la rAGT sometida a un ciclo de calentamiento-enfriamiento. Para evaluar la reversibilidad de la desnaturalización térmica, se obtuvieron los espectros de la rAGT en un amortiguador de acetatos 50 mM con 0.1 M de NaCl a pH 4.7 y a concentraciones de proteína a ~0.5 mg ml<sup>-1</sup>. (●) rAGT a 20 °C, (●) rAGT a 90 °C, (●) rAGT a 20 °C después de enfriar la muestra tras calentarla. (—) espectro correspondiente a la rAGT en presencia de 10 mM de DTT.

En estudios previos, ya se había mostrado que la wAGT muestra curvas calorimétricas complejas al perturbarla térmicamente. Estas curvas presentan dos picos, un primer pico que aparece a temperaturas más bajas y que corresponde a la disociación del dímero de la lectina y un segundo pico que aparece a temperaturas mayores y es de mayor amplitud. Este último consiste en el complejo desplegamiento de los monómeros libres, el cual se da en varios pasos. En la Fig. 28, se puede ver la comparación de las endotermas obtenidas para la rAGT y para la wAGT mediante DSC. Es claro, que la rAGT también exhibe dos picos similares a los de la wAGT. Al igual que para la wAGT, el pico de la derecha de la rAGT no mostró dependencia con la concentración de proteína, por lo que este pico debe corresponder al desplegamiento del monómero. Además, se superpone bastante bien al pico de la wAGT, lo que indica que los monómeros de ambas proteínas exhiben la misma estabilidad térmica y se despliegan a la misma temperatura. El pico izquierdo de la rAGT mostró dependencia con la concentración de proteína, de la misma manera que lo hace el pico correspondiente de la wAGT. Lo que muestra que la enzima recombinante también es capaz de formar homodímeros. Sin embargo, el pico de la rAGT se encuentra desplazado a temperaturas menores con respecto al de la wAGT (aun a la misma concentración de proteína). Por lo que, aparentemente, la rAGT forma dímeros en los que la energía de asociación es ligeramente menor a la de los dímeros de wAGT.

La capacidad de unión a quitotriosa de la rAGT fue caracterizada a través de calorimetría de titulación isotérmica (ITC). Se utilizó un amortiguador con un pH de 4.7 debido a que este es el valor de pH más alto en el que todavía se encuentra una buena proporción de la aglutinina de germen de trigo en estado soluble, lo que permite obtener isotermas de unión con una forma bien definida. La Fig. 29 compara las curvas calorimétricas obtenidas para la wAGT y para la rAGT. Los parámetros termodinámicos se obtuvieron a través de la regresión no lineal del mejor ajuste que se hizo de la curva de titulación a un modelo de un sitio de unión. Se muestra un compendio de estos en la Tabla 6.

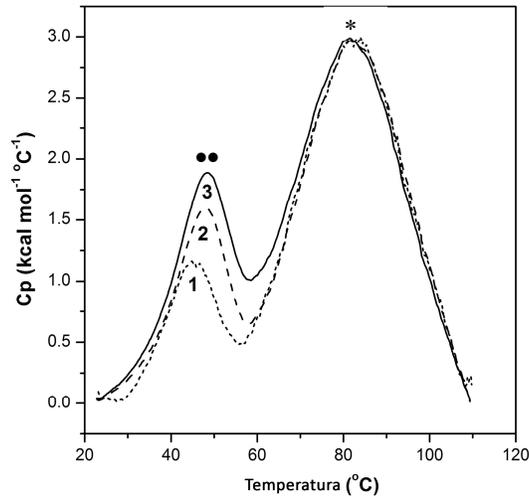


Figura 28. Endotermas de la rAGT y la wAGT obtenidas mediante DSC. Trazos obtenidos a una velocidad de barrido de  $1\text{ }^{\circ}\text{C min}^{-1}$  en amortiguador de citrato 30 mM, pH 3.5. (1) rAGT a una concentración de 0.05 mM, (2) rAGT a una concentración de 0.11 mM, (3) wAGT a una concentración de 0.09 mM. (••) Indica los picos correspondientes a la disociación del dímero. (\*) Señala los picos correspondientes al desplegamiento del monómero.

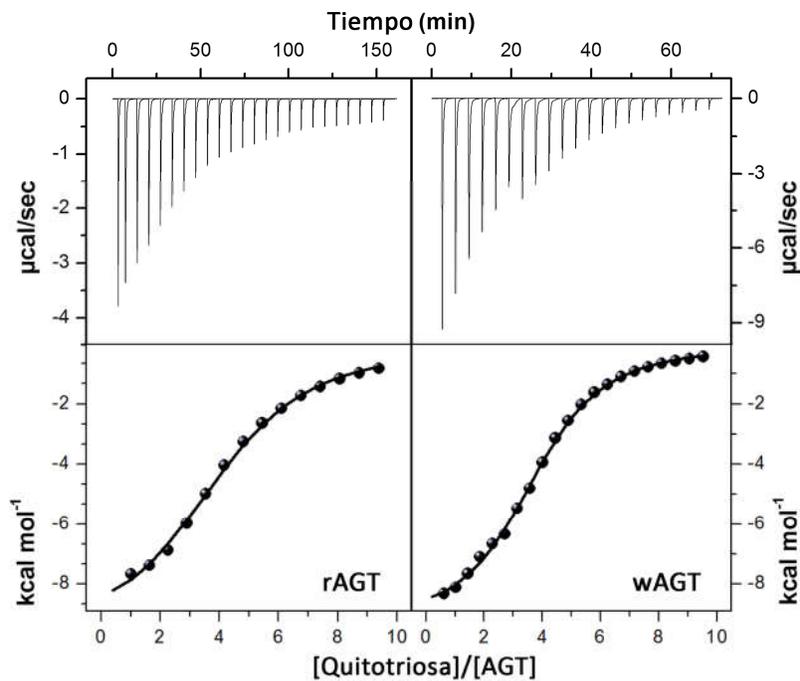


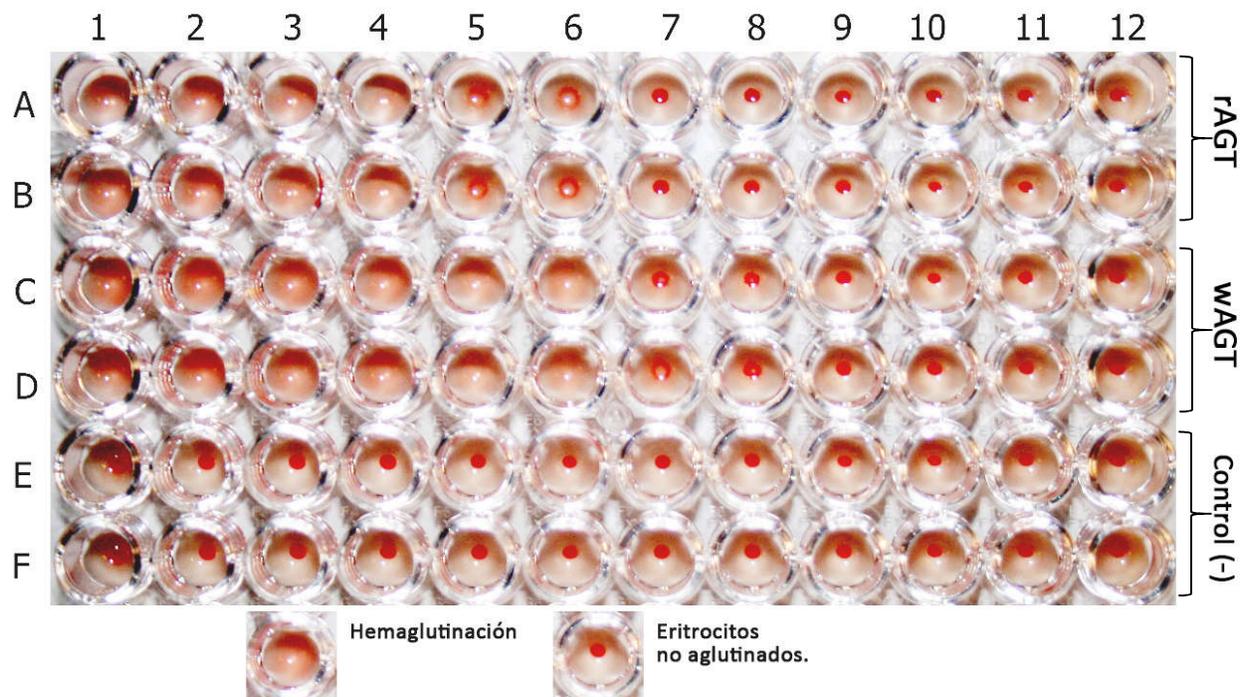
Figura 29. Isotermas de unión a quitotriosa obtenidas para la rAGT y la wAGT. Se obtuvieron a  $30\text{ }^{\circ}\text{C}$ , en un amortiguador de acetatos 50 mM con 0.1 M de NaCl a pH 4.7. De lado izquierdo se observa la isoterma de la rAGT a una concentración de 0.18 mM, del lado derecho se observa la isoterma de la wAGT a 0.10 mM.

Tabla 6. Parámetros termodinámicos de unión de la quitotriosa a la wAGT, a la rAGT y a los monodominios recombinantes.

Proteína	$K_U$ ( $\text{mM}^{-1}$ )	$\Delta G_U$ ( $\text{kcal Mol}^{-1}$ )	$\Delta H_U$ ( $\text{kcal Mol}^{-1}$ )	$T\Delta S_U$ ( $\text{kcal Mol}^{-1}$ )
wAGT	$13.2 \pm 0.6$	-5.7	$-9.4 \pm 0.1$	-3.7
rAGT	$12.5 \pm 0.6$	-5.7	$-9.9 \pm 0.2$	-4.2
rAGT.A	$1.1 \pm 0.1$	-4.1	$-9.1 \pm 0.3$	-5.0
rAGT.B	$1.1 \pm 0.1$	-4.1	$-10.1 \pm 0.7$	-6.0
rAGT.C	$0.5 \pm 0.1$	-3.7	$-7.6 \pm 0.8$	-3.9
rAGT.D	$0.3 \pm 0.1$	-3.5	$-9.9 \pm 1.3$	-6.4

De acuerdo a estos valores, es claro que las dos versiones de la AGT presentan el mismo tipo de firma termodinámica, mostrando que el proceso de unión a quitotriosa es un proceso de baja afinidad que está guiado entálpicamente. Algo común en las interacciones proteínas-carbohidratos. El ajuste es consistente con la presencia de cuatro sitios de unión por monómero para ambas aglutininas y con la ocupación de los 8 sitios de unión que se ha observado en algunas de las estructuras cristalográficas obtenidas recientemente para el dímero de la AGT completa (Schwefel *et al.*, 2010). Los valores de unión a quitotriosa también concuerdan en mayor medida con los obtenidos en otro estudio de microcalorimetría (Bains *et al.*, 1992). El valor reportado para la  $K_U$  en este último estudio ( $K_U=12.2 \text{ mM}^{-1}$ , pH 4.5, 26 °C) es muy similar al que obtuvimos, aunque difiere un poco en el valor de la entalpía de unión ( $\Delta H_U = -160 \text{ kcal mol}^{-1}$ ), siendo este último más negativo.

Los resultados de los ensayos de hemaglutinación se pueden revisar en la Fig. 30. Estos ensayos se llevaron a cabo en PBS (50 mM de fosfatos y 150 mM de NaCl) a pH 7.4. El PBS es una solución comúnmente empleada para procedimientos bioquímicos, ya que su osmolaridad y concentración de iones es muy semejante a la del líquido extracelular de los mamíferos. Por lo que resulta apropiada para el manejo de eritrocitos. La AGT une carbohidratos con residuos de N-Acetilglucosamina, los cuales se encuentran presentes en los oligosacáridos de las membranas plasmáticas de los eritrocitos de los tres tipos sanguíneos (A, B y O). Por lo mismo, aunque para estos ensayos se ocuparon eritrocitos tipo O+, se hubiera podido ocupar cualquiera de los otros tipos sanguíneos. Se partió de una concentración inicial de  $0.1 \text{ mg ml}^{-1}$  de proteína, tanto para la rAGT como para la wAGT, sirviendo esta última como control positivo. Como control negativo se utilizó únicamente amortiguador de PBS. La primera columna corresponde a los pozos a la concentración inicial, a partir de esta columna y hacia la derecha, se realizaron varias diluciones seriadas con factor de dilución 1:2. Los pozos en los que se aprecia un punto rojo al centro, son aquéllos en los que no hubo hemaglutinación, por lo que los eritrocitos precipitaron en el fondo del pozo. Los pozos en los que no se observa este punto y se ve una turbidez distribuida homogéneamente, son aquellos en los que ocurrió la hemaglutinación. Con este experimento se comprueba la capacidad aglutinante de la rAGT, aunque se observa una capacidad para aglutinar ligeramente menor con respecto al mostrado por la wAGT.



**Figura 30. Prueba de hemaglutinación con eritrocitos humanos.** La prueba se realizó con eritrocitos resuspendidos en PBS a pH 7.4. Como se puede ver en el recuadro inferior izquierdo, la prueba es positiva si se forma una capa difusa, además de que no se logra reconocer eritrocitos que hayan precipitado y se puedan distinguir en el centro del pozo. El recuadro inferior derecho, muestra como debe ser el resultado en el caso de que la hemaglutinación no ocurra. En este caso los eritrocitos precipitan, formando una pastilla que aparece en el centro del pozo. (Filas A y B) Ensayos con la rAGT. (Filas C y D) Ensayos con la wAGT, utilizada como control positivo. (Filas E y F) Amortiguador PBS sin proteína como control negativo.

### Caracterización de los monodominios recombinantes.

La conformación de cada dominio recombinante se evaluó por dicroísmo circular en el UV-lejano en función del pH. La Fig. 31A muestra los espectros superpuestos del dominio A, a diferentes valores de pH (2,0, 4,7 y 6,0). Aquí se puede observar que el dominio, independientemente de la acidez de la solución, mantiene una conformación plegada. Esto se basa en la forma del espectro, la cual prácticamente no cambia a ninguno de estos valores de pH y mantiene la forma característica observada para las proteínas tipo heveína en estado plegado. El mismo comportamiento se observó para los dominios B y C, mientras que el dominio D mostró diferencias. Este dominio también mostró una conformación plegada a pH 4,7 y 6,0, en tanto que a pH 2,0 mostró una conformación desplegada (Fig. 31C). Esto concuerda con los datos de desplegamiento anteriormente obtenidos mediante DSC para la wAGT (Portillo-Téllez *et al.*, 2011). A pH 2,0 la lectina se encuentra completamente monomérica. La deconvolución de las trazas de desplegamiento reveló un mecanismo de varios pasos que consiste de tres transiciones independientes de dos estados. Esta información junto con los datos estructurales y espectrofotométricos, llevó a la

conclusión de que, a este valor de pH, solo 3 de los 4 dominios se encuentran en estado plegado y que muestran transiciones cooperativas a altas temperaturas. Mientras que el otro dominio estaba desplegado aun a bajas temperaturas. Para obtener más información sobre el proceso de desplegamiento de la AGT, los dominios A, B y C se sometieron a un ciclo de calentamiento/enfriamiento a pH 2.0 (Fig. 31D). El inicio de la transición de desplegamiento de rAGT.A ocurrió a ~65 °C, mientras que para rAGT.B y rAGT.C la transición comenzó a ~75 °C. El final de la transición no se observó en ninguno de los perfiles de desplegamiento, aunque los 3 dominios se mostraron mayoritariamente desplegados a los 90 °C, a juzgar por el espectro de dicroísmo. Después de enfriar la muestra y volver a temperatura ambiente, las señales de dicroísmo se recuperaron por completo, un comportamiento similar al observado para la lectina completa (Chavelas *et al.*, 2004; Portillo-Téllez *et al.*, 2011).

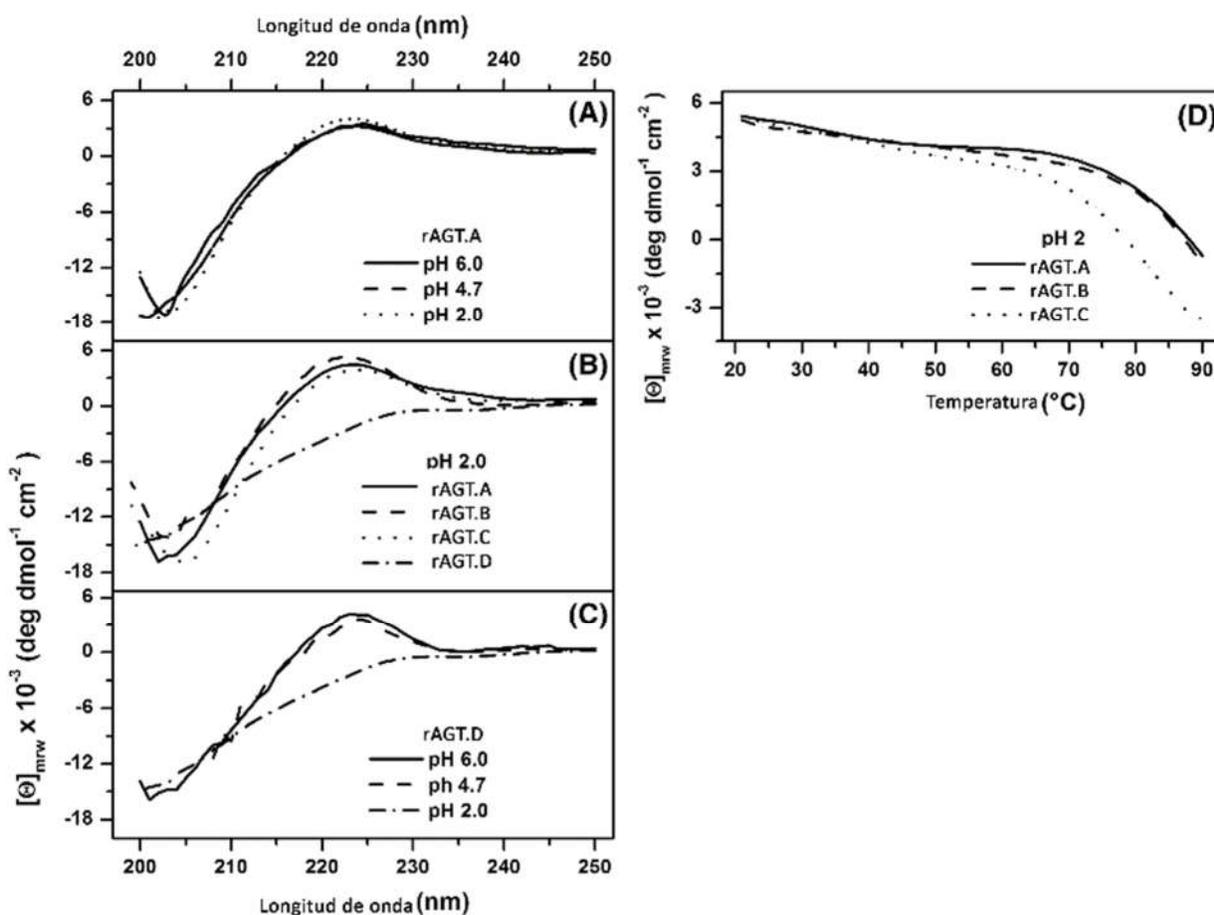
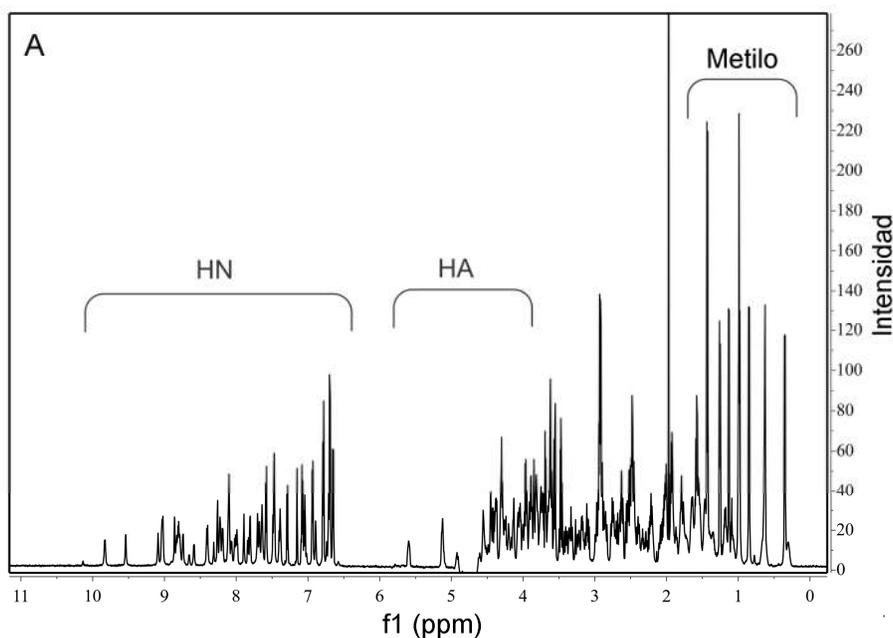
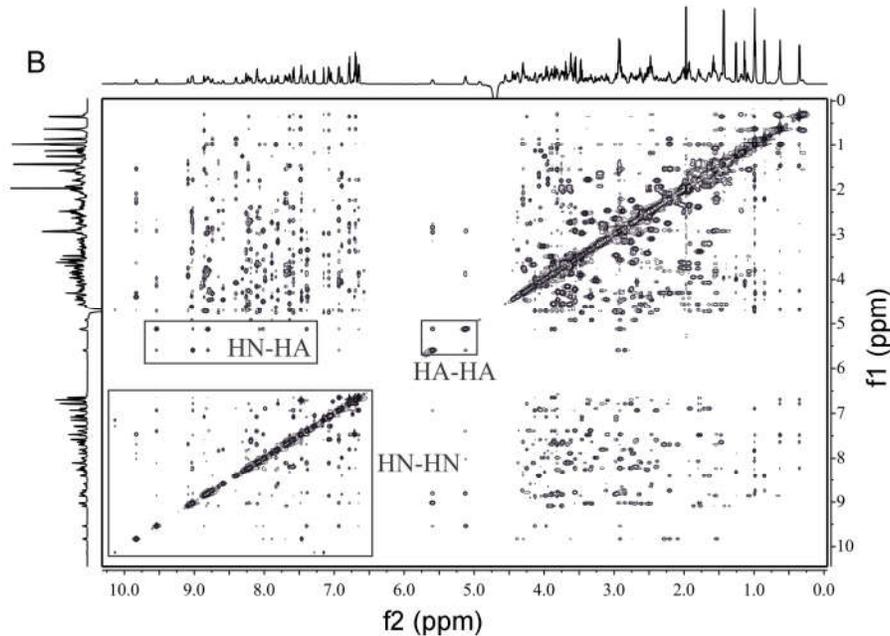


Figura 31. Espectros de dicroísmo circular en el UV lejano de los dominios recombinantes de la AGT. (A) Espectros del dominio rAGT.A a pH 2.0, 4.7 y 6.0. (B) Espectros de los cuatro monodominios recombinantes a pH 2.0. (C) Espectros del dominio rAGT.D a pH 2.0, 4.7 y 6.0. (D) perfiles de desplegamiento seguidos a 225 nm para la rAGT.A, rAGT.B y rAGT.C.

Se obtuvo el espectro de protones del dominio rAGT.D por NMR para elucidar la conformación de este dominio (Fig. 32A). El espectro en 1D mostró señales muy bien dispersadas en todas las regiones HN, HA, grupos metilo y de la cadena principal (Fig. 32). Este experimento confirmó que el dominio está plegado. Además, no se encontró evidencia de una dinámica conformacional significativa. El que las señales HN y HA se observen en un campo por debajo de lo esperado y el que los grupos metilo aparezcan desde los 0.35 ppm, corroboró el estado plegado del dominio. La región HN se separó bastante bien, permitiendo la identificación de la señal NH del triptófano a 10.13 ppm y de otras tres señales de HN a 9.84, 9.83 y 9.53 ppm. Dos señales HA se observaron a 5.60 y 5.12 ppm. En concordancia con la estructura primaria de rAGT.D, se encontraron 8 señales de grupos metilo entre las 1.40 y las 0.35 ppm; varias de ellas mostraron interacción protectora. El espectro NOESY en 2 dimensiones (Fig. 33) mostró un NOE bien definido entre 2 HA-HA (HA de 5.60 ppm y HA de 5.12 ppm) y por lo menos 3 sistemas de spin en la región NH con correlación  $\text{HN}_i\text{-HA}_i$  intraresiduo, correlación  $\text{HN}_i\text{-HA}_{i-1}$  secuencial y  $\text{HN}_i\text{-HA}_j$  dentro de una hoja  $\beta$  (estas regiones están delimitadas por cuadrados en la Fig. 33). Además, varios NOE en varias tiras de HN evidenciaron la presencia de por lo menos una hélice. En general, los datos obtenidos por NMR concuerdan con la conformación del dominio observada en las estructuras cristalográficas de la lectina completa.



**Figura 32.** Espectro de NMR de protón del dominio rAGT.D. Espectro obtenido a 700 MHz en un equipo Bruker 700. El espectro muestra una buena dispersión de las señales, evidenciando que la proteína se encuentra plegada en una sola conformación.



**Figura 33. Espectro NOESY del dominio rAGT.D.** Espectro de NMR en 2 dimensiones, obtenido a 300 ms de tiempo de mezclado. El espectro muestra una correlación HA-HA entre dos protones de una hoja  $\beta$ . Dentro de la misma hoja  $\beta$ , también es posible observar 3 sistemas en la región HN-HA con correlación intra-residuos. La región HN-HN mostró varias correlaciones entre los protones HN, sugiriendo la formación de una hélice.

Se evaluó la capacidad de unir quitotriosa de los 4 dominios a pH 4.7 utilizando ITC. La Fig. 34 muestra las isotermas que se obtuvieron para la titulación de cada uno de los dominios con quitotriosa. Los parámetros de unión para cada uno de los dominios se pueden revisar en la Tabla 6. En general, todos los dominios mostraron una unión entálpicamente dirigida con una afinidad muy baja hacia el trisacárido. Los valores obtenidos para el dominio rAGT.B son ligeramente distintos a los reportados previamente para el mismo dominio ( $\Delta H_{ij} = -8.1$  kcal/mol,  $K_{ij} = 908$  M $^{-1}$ , 25°C) por el grupo de Espinosa, quienes también hicieron uso de ITC para obtener estos parámetros (Espinosa *et al.*, 2000). Todos los dominios mostraron afinidades al menos 10 veces más pequeñas que la observada para la aglutinina completa. De manera contrastante, no se observaron diferencias significativas de las entalpías de unión de los dominios con respecto a la versión completa de la proteína. Estos resultados indican que el incremento en la afinidad mostrado por la lectina en estado dimérico con respecto a los dominios aislados, es de origen entrópico.

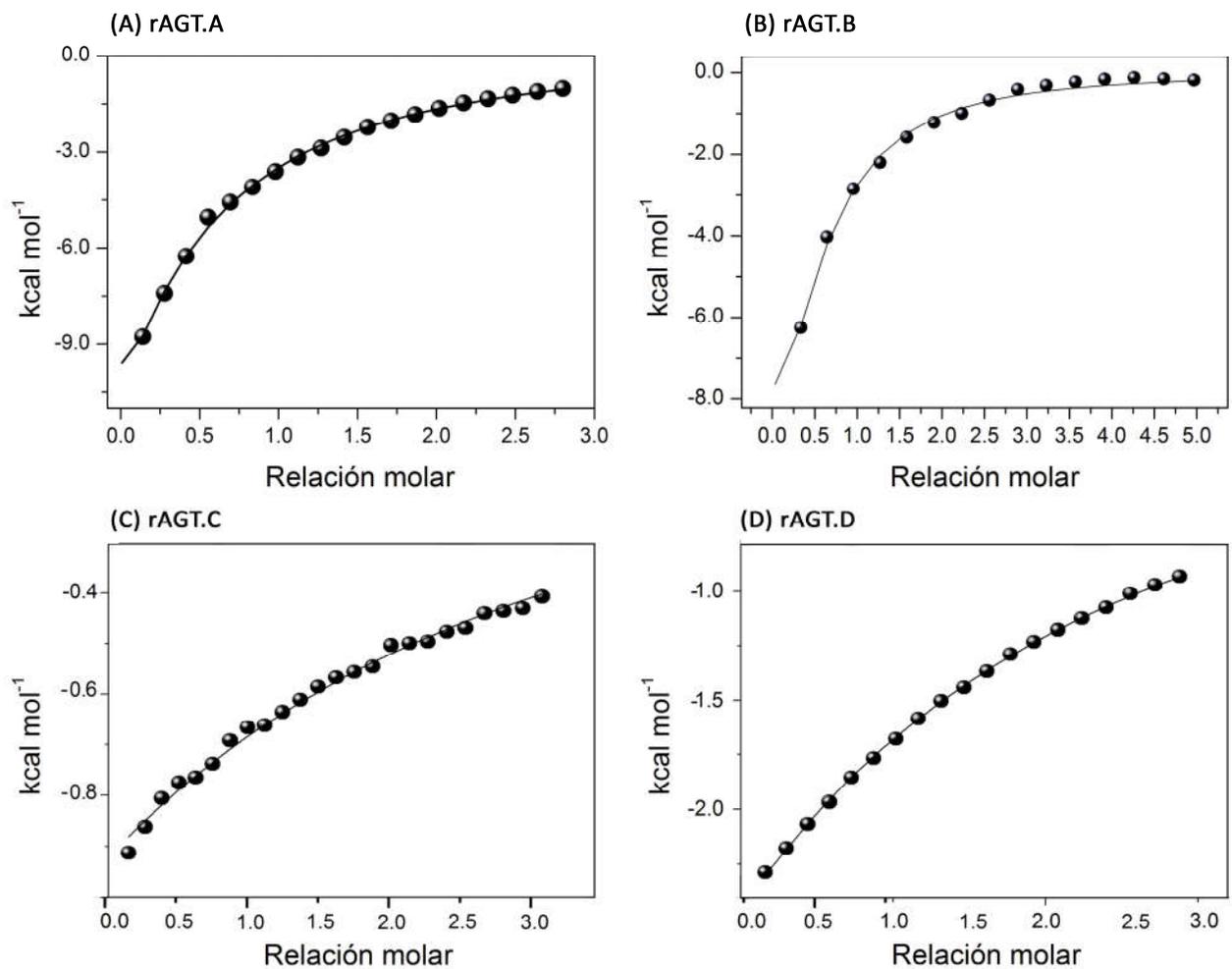


Figura 34. Isothermas obtenidas mediante ITC para los dominios recombinantes de la AGT. Titulaciones hechas con quitotriosa a una temperatura de 30 °C en amortiguador 0.05 M de acetato, 0.1 M de NaCl, pH 4.7.

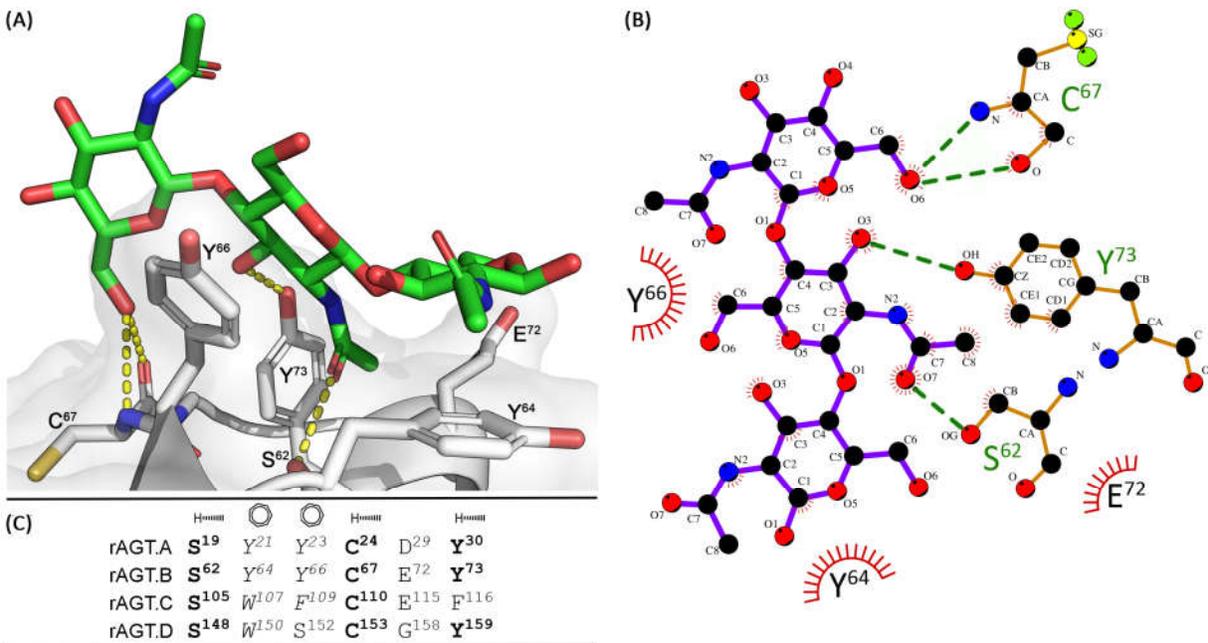
## DISCUSIÓN

Durante décadas, la aglutinina de germen de trigo ha atraído un gran interés tanto en aspectos fundamentales como de glicobiología aplicada. Esta lectina es una herramienta invaluable tanto para el análisis convencional como de vanguardia de glicoconjugados. Numerosos grupos de investigación se encuentran actualmente en la búsqueda de lograr implementar su uso en sistemas de entrega de fármacos. A pesar de esto, hasta el momento no existían reportes de la producción de esta lectina en sistemas de expresión bacterianos. La disponibilidad de un método barato y eficiente para obtener a la AGT recombinante y a sus dominios aislados, facilitará la realización de estudios enfocados en mejorar las propiedades de esta lectina con fines aplicados a la biología y a la biomédica. Además, esto permitirá profundizar nuestro conocimiento acerca del complejo comportamiento que la proteína exhibe en solución.

En este estudio, con el uso de una proteína de fusión con tiorredoxina y una etiqueta de histidinas, logramos obtener un rendimiento aproximado de 4 mg de rAGT por litro de medio de cultivo de *E. coli*. Propiedades, como la estructura secundaria, la resistencia térmica, la capacidad de unión al ligando y la actividad hemaglutinante de la lectina recombinante, resultaron muy similares a las exhibidas por la proteína silvestre. La estrategia que ocupamos fue muy similar a la que se usó en la producción de proteínas con dominios tipo heveína funcionales y plegadas correctamente (Martínez-Caballero *et al.*, 2014) y a la que se ha utilizado en la producción de algunas toxinas pequeñas de arácnidos ricas en puentes disulfuro (Saucedo *et al.*, 2012; Flores-Solis *et al.*, 2016). El uso de constructos con tiorredoxina y una etiqueta de histidinas permitió la producción bacteriana de los 4 dominios independientes de la AGT en su conformación nativa y totalmente competentes, con rendimientos de expresión similares a los obtenidos para la aglutinina completa. Este resultado implica que cada dominio se comporta como una unidad de plegamiento independiente.

De acuerdo a las mediciones de ITC, los 4 dominios aislados difieren entre ellos en sus propiedades de unión a ligandos más allá de la incertidumbre experimental. Por lo que, a pesar de la similitud en la arquitectura de sus sitios de unión, no son equivalentes en términos energéticos. Previamente, se reportaron 2 modos de unión para el complejo rATG.B-quitotriosa resueltos por NMR (Espinosa *et al.*, 2000). En uno de los modos de unión, los tres residuos de GlcNAc están en contacto con el dominio (Fig. 35A); en el otro complejo, uno de los residuos no tiene contacto con el dominio y está completamente expuesto al solvente. La Fig. 35B muestra una representación esquemática de la estructura del primer complejo, resaltando los residuos que se encuentran a 5 Å del ligando. 4 puentes de hidrógeno (en los que participan los residuos Ser<sup>62</sup>, Cys<sup>67</sup> y Tyr<sup>73</sup>) y dos interacciones de apilamiento (formadas por la Tyr<sup>64</sup> y la Tyr<sup>66</sup>) que estabilizan la interacción de rAGT.B con la quitotriosa. Un análisis de la secuencia de los dominios muestra que rAGT.A preserva los mismos residuos que en rAGT.B se encuentran interactuando con el azúcar (Fig. 35C). En concordancia con este parecido, los dos dominios mostraron firmas energéticas

de interacción con la quitotriosa muy similares. En cambio, la posición equivalente a la Tyr<sup>73</sup> de rAGT.B, se encuentra ocupada por la Phe<sup>116</sup> en rAGT.C. Además, este dominio tiene W<sup>107</sup> y F<sup>109</sup> en lugar de los aminoácidos Tyr<sup>64</sup> y Tyr<sup>66</sup> presentes en rAGT.B. Presumiblemente, la pérdida de un puente de hidrógeno y la modificación de las interacciones de apilamiento causan que la fuerza de interacción con el ligando de rAGT.C. sea más débil que la observada para los dominios rAGT.B y rAGT.A. En cuanto a rAGT.D, es el único dominio que carece de 1 de los 2 residuos aromáticos que se encuentran participando en las interacciones de apilamiento y, en consecuencia, fue el que mostró la menor afinidad por el ligando. En la Fig. 36 se pueden observar las diferencias entre los residuos principales de unión de cada dominio dentro de la estructura cristalográfica. Es interesante resaltar que la identidad promedio de la secuencia entre los 4 dominios es de alrededor del 40 %, un valor que es muy similar al observado entre lectinas que unen quitina y que pertenecen a la familia de la heveína. Recientemente, se mostró que la AGT no sólo es capaz de reconocer quito-oligosacáridos, sino que también presentan especificidad por diversos oligosacáridos (Itakura *et al.*, 2017).



**Figura 35 Aminoácidos relevantes en el reconocimiento a carbohidratos de la AGT.** (A) Estructura tridimensional del sitio de unión de rAGT.B en complejo con la quitotriosa (en verde) utilizando las coordenadas de NMR que fueron cordialmente proporcionadas por el Dr. Jesús Jiménez-Barbero del Centro de Investigación Cooperativa en Biociencias (CIC-bioGUNE, España)(Espinosa *et al.*, 2000). Los puentes de hidrógeno se resaltan en líneas punteadas. (B) Representación esquemática del sitio de unión de rAGT.B en complejo con la quitotriosa La figura se generó con el software LigPlus (Laskowski y Swindells, 2011). Los puentes de hidrógeno también se indican con líneas punteadas. (C) Alineamiento de los principales residuos de unión de los 4 dominios de la AGT. Arriba de cada columna se indica si esos residuos están implicados en la formación de puentes de hidrógeno, en la formación de interacciones de apilamiento o de interacciones electrostáticas.

En este aspecto, aunque la AGT une con una considerable especificidad N-glicanos, sobre todo a los que contienen GlcNAc, es tentador especular que las heterogeneidades estructurales y energéticas de unión entre los 4 dominios de la AGT contribuyen a la habilidad de esta lectina de reconocer sacáridos con arquitecturas complejas con alta especificidad.

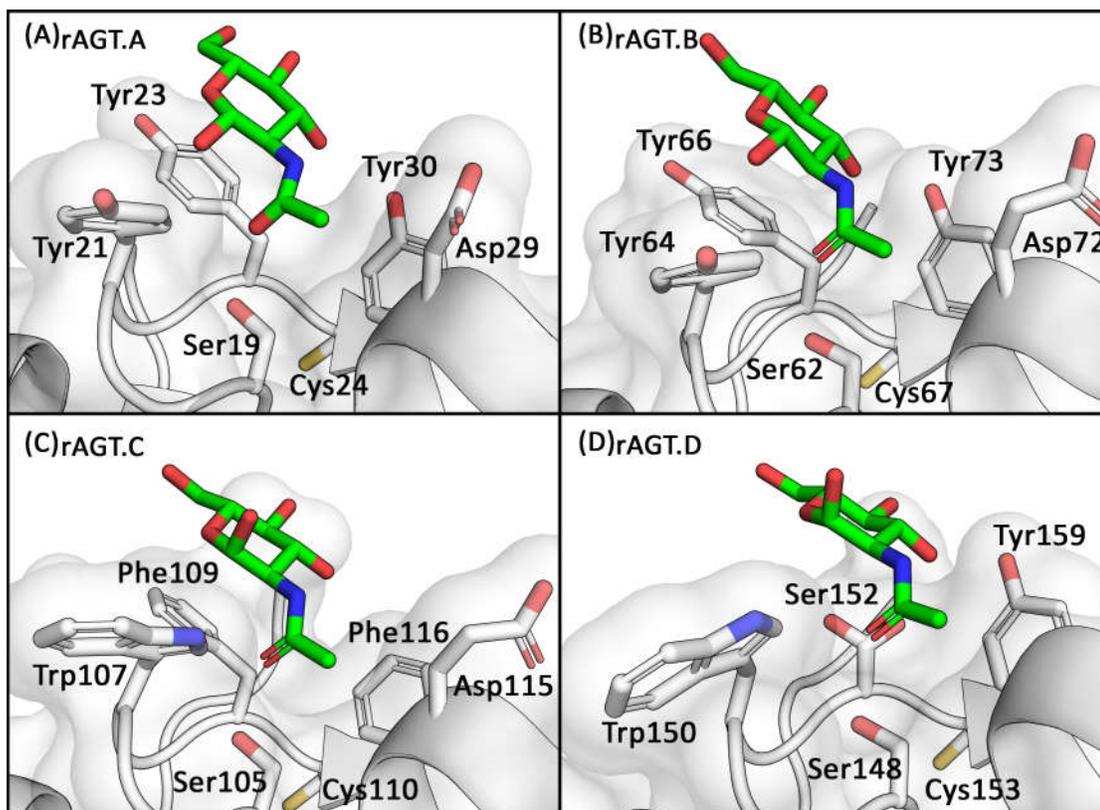


Figura 36. Comparación de los sitios de unión de los dominios de la AGT. (Código PDB: 2UV0). Se resaltan los principales residuos del sitio de unión de cada dominio. En verde se señala el ligando (GlcNAc).

La meta de utilizar a la AGT como un sistema de entrega de fármacos de administración peroral, involucra el que la aglutinina se exponga a las distintas condiciones de pH que encontramos a lo largo del tracto digestivo. Por ello, resulta de gran importancia elucidar la estabilidad estructural de esta lectina en función del pH. Previamente, se mostró que en condiciones ácidas similares a las del estómago humano, la AGT está principalmente monomerizada (Nagata y Burger, 1974; Rice y Etzler, 1974; Chavelas *et al.*, 2004; Portillo-Télliz *et al.*, 2011). Bajo estas condiciones, la aglutinina exhibe perfiles de desnaturalización térmica que consisten de 3 transiciones independientes de 2 estados, las cuales ocurren con diferentes temperaturas medias. Para explicar el hecho de que sólo se encuentran 3 y no 4 transiciones en las trazas calorimétricas, se propusieron los siguientes escenarios de desplegamiento: (1) La existencia de un dominio muy estable, cuyo desplegamiento ocurre a temperaturas mayores a las que son asequibles experimentalmente, (2) 2 dominios se despliegan de manera concertada, o (3) 1 dominio

se encuentra desplegado a temperatura ambiente. Basados en los datos espectrofotométricos y calorimétricos de la aglutinina completa, se consideró al último como el modelo más probable (Portillo-Téllez *et al.*, 2011). Los datos del presente estudio, muestran que rAGT.A, rAGT.B y rAGT.C preservan su conformación tipo heveína aún a pH 2.0 y que exhiben una transición de desplegamiento a temperaturas mayores. En cambio, el dominio rAGT.D se mostró con propiedades propias de proteínas desestructuradas a valores muy bajos de pH. Por lo tanto, el estudio de los dominios aislados refuerza la idea de que bajo condiciones de pH similares a las del estómago, 1 de los 4 dominios, específicamente el dominio D, se encuentra desplegado, mientras que los otros 3 permanecen completamente plegados. Adicionalmente, los perfiles térmicos revelaron que el dominio A es menos termoestable en relación con los dominios B y C. Para seguir desentrañando el complejo mecanismo de desplegamiento de la AGT, se requiere de una caracterización más detallada de los dominios aislados. Adicionalmente, consideramos que el estudio de construcciones compuestas de 2 (AB, BC y CD) y 3 (ABC y BCD) dominios sería de gran valor para elucidar posibles efectos de comunicación inter-dominio tanto para la estabilidad como para la capacidad de reconocimiento de ligandos.

## CONCLUSIONES

- Se logró establecer un método de expresión adecuado para la producción de la AGT completa y para sus 4 dominios aislados utilizando un sistema recombinante de *E. coli*.
- Se alcanzó un rendimiento que va de los 3 a los 5 mg de proteína por cada litro de medio de cultivo, el cual es similar e incluso mejor que en otros sistemas no bacterianos, en especial si se toma en cuenta que este sistema constituye un método más accesible y económico.
- La AGT recombinante mostró propiedades estructurales y de unión prácticamente idénticas a la AGT obtenida a partir de su fuente natural, lo que permite validar su implementación en estudios posteriores, incluyendo aquellos en los que se busque darle un uso que tenga una mayor aplicación biotecnológica.
- En este estudio, se vio que, al menos los dominios rAGT.A, rAGT.B y rAGT.C, preservan una conformación estructurada y parecida a la de la heveína, incluso cuando estos se encuentran en condiciones de pH ácido (a pH 2.0).
- También se encontró que estos tres dominios exhiben perfiles de desplegamiento similares a la AGT completa, con transiciones que empiezan por arriba de los 60 °C.
- De forma contrastante, se encontró que a pH de 2.0, el dominio rAGT.D muestra propiedades que son más parecidas a la de una proteína desestructurada.

- Este comportamiento de los dominios, refuerza la hipótesis de que, durante el tránsito de la AGT por el estómago, uno de sus cuatro dominios (específicamente el dominio D) se encuentra desplegado.
- Los ensayos de ITC sirvieron para corroborar la funcionalidad de todos los dominios, lo que implica que, efectivamente, estos se comportan como unidades que presentan una función y un plegamiento independiente.
- A pesar de mostrar la misma tendencia, los parámetros termodinámicos de unión a la quitotriosa muestran pequeñas diferencias entre cada uno de los dominios. Esto ha servido para distinguir y apreciar mejor a los dos aminoácidos aromáticos y a los cuatro aminoácidos polares que se encuentran en el sitio de unión a ligando y que son los principales encargados de formar interacciones energéticamente favorables con el carbohidrato.
- Por lo mismo, y a pesar de la similitud en la arquitectura entre los sitios de unión, la comparación entre dominios ha servido para estimar la repercusión causada en la fuerza con la que se une al ligando por la pérdida de estas interacciones.
- Gracias al estudio de la proteína completa y de sus dominios aislados, se logró encontrar propiedades fundamentales de la unión de esta proteína a los carbohidratos y de su estabilidad termodinámica. Estos hallazgos han servido para esclarecer un poco el comportamiento global de la AGT y nos acerca un poco más a la meta de mejorar la implementación de esta proteína como un sistema de entrega de fármacos.
- Sin embargo, la interacción de la AGT con los ligandos es un evento complejo en el que es probable que ocurran fenómenos de cooperatividad entre dominios. Aunado a esto, hay que recordar su interdependencia con los procesos de autoasociación y plegamiento. Por lo que, para poder avanzar en su estudio, se requiere una caracterización más detallada de los dominios aislados, así como de constructos de los bidominios (AB, BC y CD) y de los tridominios de la AGT (ABC y BCD).

## PERSPECTIVAS

- Se planea realizar estudios más detallados de los dominios aislados mediante DC y DSC, en los que se analice el perfil de desplegamiento de estos a diferentes valores de pH.
- Utilizando como hebra molde al gen de la proteína completa que se encuentra en el vector pBSK GS 45510, se obtendrán los amplificadores de los constructos correspondientes a los bidominios (AB, BC y CD) y a los tridominios (ABC y BCD) de la AGT. Los cuáles serán transferidos a vectores de expresión para su posterior producción.
- Una vez se hayan obtenido y transformado estas construcciones, se realizarán los correspondientes ensayos de expresión para corroborar si las condiciones ocupadas previamente para la producción de la AGT completa y para sus dominios aislados pueden ser ocupadas en la producción de bidominios y tridominios de la AGT.
- Una vez que se haya logrado la expresión y purificación de estas construcciones, se procederá a realizar la caracterización espectroscópica de estos mediante DC a diferentes condiciones de pH.
- Utilizando la misma técnica de DC, se planea realizar ensayos de perturbación térmica de los bidominios y los tridominios.
- Se procederá a obtener la estabilidad termodinámica de estas construcciones mediante ensayos de DSC.
- Realizar la caracterización estructural de los bidominios y los tridominios a través de RMN.
- Obtener los parámetros termodinámicos de unión a quitotriosa mediante ensayos de ITC y realizar los ajustes, el análisis y la comparación correspondiente para cada una de las construcciones.

## REFERENCIAS

- Afrin, R. y Ikai, A. (2014) "Subunit unbinding mechanics of dimeric wheat germ agglutinin (WGA) studied by atomic force microscopy", *FEBS Letters*, 588(23), pp. 4472–4477. doi: 10.1016/j.febslet.2014.10.018.
- Agrawal, B. B. y Goldstein, I. J. (1965) "Specific binding of concanavalin A to cross-linked dextran gels.", *Biochem J*, 96(3), pp. 23C-25C. doi: 10.1042/bj0960023C.
- Allen, A. y Snary, D. (1972) "The structure and function of gastric mucus.", *Gut*, 13(8), pp. 666–672. doi: 10.1136/gut.13.8.666.
- Ambrosi, M., Cameron, N. R. y Davis, B. G. (2005) "Lectins: Tools for the molecular understanding of the glycode", *Org Biomol Chem*, 3(9), pp. 1593–1608. doi: 10.1039/b414350g.
- Andersen, N. H., Cao, B., Rodríguez-Romero, A. y Arreguin, B. (1993) "Hevein: NMR Assignment and Assessment of Solution-State Folding for the Agglutinin-Toxin Motif", *Biochemistry*, 32(6), pp. 1407–1422. doi: 10.1021/bi00057a004.
- Apweiler, R., Hermjakob, H. y Sharon, N. (1999) "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database", *Biochim Biophys Acta*, 1473(1), pp. 4–8. doi: 10.1016/S0304-4165(99)00165-8.
- Asensio, J. L., Canada, F. J., Bruix, M., Rodríguez-Romero, A. y Jimenez-Barbero, J. (1995) "The Interaction of Hevein with N-acetylglucosamine-containing Oligosaccharides. Solution Structure of Hevein Complexed to Chitobiose", *Eur J Biochem*, 230(2), pp. 621–633. doi: 10.1111/j.1432-1033.1995.tb20604.x.
- Aub, J. C., Sanford, B. H. y Cote, M. N. (1965) "Studies on reactivity of tumor and normal cells to a wheat germ agglutinin.", *Proc Natl Acad Sci U S A*, 54(2), pp. 396–399. doi: 10.1073/pnas.54.2.396.
- Aub, J. C., Sanford, B. H. y Wang, L. H. (1965) "Reactions of normal and leukemic cell surfaces to a wheat germ agglutinin.", *Proc Natl Acad Sci U S A*, 54(2), pp. 400–402. doi: 10.1073/pnas.54.2.400.
- Bains, G., Lee, R. T., Lee, Y. C. y Freire, E. (1992) "Microcalorimetric Study of Wheat Germ Agglutinin Binding to N-Acetylglucosamine and Its Oligomers", *Biochemistry*, 31(50), pp. 12624–12628. doi: 10.1021/bi00165a012.
- Bar-Shavit, Z. y Goldman, R. (1976) "Concanavalin A-mediated attachment and ingestion of yeast cells by macrophages", *Exp Cell Res*, 99(2), pp. 221–236. doi: 10.1016/0014-4827(76)90578-4.
- Barondes, S. H. (1988) "Bifunctional properties of lectins: lectins redefined", *Trends Biochem Sci*, 13(12), pp. 480–482. doi: 10.1016/0968-0004(88)90235-6.
- Berthelot, K., Peruch, F. y Lecomte, S. (2016) "Highlights on Hevea brasiliensis (pro)hevein proteins", *Biochimie*, 127, pp. 258–270. doi: 10.1016/j.biochi.2016.06.006.
- Boyd, W. C. y Shapleigh, E. (1954) "Specific Precipitating Activity of Plant Agglutinins (Lectins)", *Science*, 119(3091), pp. 419–419. doi: 10.1126/science.119.3091.419.
- De Brito Ximenes, P., Beltrão, E. I. C., Macêdo, D. P. C., Buonafina, M. D. S., De Lima-Neto, R. G. y Neves, R. P. (2015) "Targeting the *Cryptococcus neoformans* var. *grubii* cell wall using lectins: Study of the carbohydrate-binding domain", *Molecules*, 20(3), pp. 3776–3782. doi: 10.3390/molecules20033776.
- Bronowska, K., A. (2011) "Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design", en *Thermodynamics - Interaction Studies - Solids, Liquids and Gases*, pp. 1–48. doi: 10.5772/19447.
- Burger, M. M. y Goldberg, A. R. (1967) "Identification of a tumor-specific determinant on neoplastic cell surfaces.", *Proc Natl Acad Sci U S A*, 57(2), pp. 359–366. doi: 10.1073/pnas.57.2.359.
- Chavelas, E. A., Beltrán, A. P., Pérez-hernández, G. y García-hernández, E. (2004) "Spectroscopic Characterization of the Thermal Unfolding of Wheat Germ Agglutinin", *J. Mex. Chem. Soc.*, 48, pp. 279–282.
- Coelho, L., Silva, P., Oliveira, W., Moura, M., Pontual, E., Soares, F., Guedes, P., Napoleão, T. y Correia, M. (2018) "Lectins as antimicrobial agents", *J Appl Microbiol.*, 125(5), pp. 1238–1252. doi: 10.1111/jam.14055.
- Cohen, M. y Varki, A. (2010) "The Sialome—Far More Than the Sum of Its Parts", *OMICS: A Journal of Integrative Biology*, 14(4), pp. 455–464. doi: 10.1089/omi.2009.0148.
- Cohen, M. y Varki, A. (2014) "Modulation of Glycan Recognition by Clustered Saccharide Patches", en *International review of cell and molecular biology*, pp. 75–125. doi: 10.1016/B978-0-12-800097-7.00003-8.
- Cooper, D. N. W. (2002) "Galectinomics: Finding themes in complexity", *Biochimica et Biophysica Acta*, 1572(2–3), pp. 209–231. doi: 10.1016/S0304-4165(02)00310-0.

- Cummings, R. D. y Pierce, J. M. (2014) "The Challenge and Promise of Glycomics", *Chemistry & Biology*, 21(1), pp. 1–15. doi: 10.1016/J.CHEMBIOL.2013.12.010.
- Van Damme, E. J. M., Lannoo, N. y Peumans, W. J. (2008) "Plant Lectins", en *Advances in Botanical Research*, pp. 107–209. doi: 10.1016/S0065-2296(08)00403-5.
- Drenth, J., Low, B. W., Richardson, J. S. y Wright, C. S. (1980) "The toxin-agglutinin fold. A new group of small protein structures organized around a four-disulfide core.", *J Biol Chem.*, 255(7), pp. 2652–2655.
- Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., Ji, X.-L. y Liu, S.-Q. (2016) "Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods.", *Int J Mol Sci*, 17(2). doi: 10.3390/ijms17020144.
- Durand, G. y Seta, N. (2000) "Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring.", *Clin Chem*, 46(6 Pt 1), pp. 795–805.
- Emde, B., Heinen, A., Gödecke, A. y Bottermann, K. (2014) "Wheat germ agglutinin staining as a suitable method for detection and quantification of fibrosis in cardiac tissue after myocardial infarction.", *Eur J Histochem*, 58(4), p. 2448. doi: 10.4081/ejh.2014.2448.
- Espino-Solis, G. P. (2015) "Brief review - Lectins", *Revista Vitae*, 22(1), pp. 9–11. doi: 10.17533/udea.vitae.v22n1a01.
- Espinosa, J. F., Asensio, J. L., García, J. L., Laynez, J., Bruix, M., Wright, C., Siebert, H. C., Gabius, H. J., Cañada, F. J. y Jiménez-Barbero, J. (2000) "NMR investigations of protein-carbohydrate interactions. Binding studies and refined three-dimensional solution structure of the complex between the B domain of wheat germ agglutinin and N, N', N"-triacetylchitotriose", *Eur J Biochem*, 267(13), pp. 3965–3978. doi: 10.1046/j.1432-1327.2000.01415.x.
- Feizi, T. y Mulloy, B. (2003) "Carbohydrates and glycoconjugates. Glycomics: The new era of carbohydrate biology", *Curr Opin Struct Biol*, 13(5), pp. 602–604. doi: 10.1016/j.sbi.2003.09.001.
- Flores-Solis, D., Toledano, Y., Rodríguez-Lima, O., Cano-Sánchez, P., Ramírez-Cordero, B. E., Landa, A., Rodríguez de la Vega, R. C. y del Rio-Portilla, F. (2016) "Solution structure and antiparasitic activity of scorpine-like peptides from *Hoffmannihadrurus gertschi*", *FEBS Letters*, 590(14), pp. 2286–2296. doi: 10.1002/1873-3468.12255.
- Fu, L., Zhou, C., Yao, S., Yu, J., Liu, B. y Bao, J. (2011) "Plant lectins: Targeting programmed cell death pathways as antitumor agents", *Int J Biochem Cell Biol*, 43(10), pp. 1442–1449. doi: 10.1016/j.biocel.2011.07.004.
- Gabius, H.-J., Siebert, H.-C., André, S., Jiménez-Barbero, J. y Rüdiger, H. (2004) "Chemical Biology of the Sugar Code", *ChemBioChem*, 5(6), pp. 740–764. doi: 10.1002/cbic.200300753.
- Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. y Solís, D. (2011) "From lectin structure to functional glycomics: principles of the sugar code.", *Trends Biochem Sci*, 36(6), pp. 298–313. doi: 10.1016/j.tibs.2011.01.005.
- Gabius, H.-J. (2013) "Animal and Human Lectins", en *Antitumor Potential and other Emerging Medicinal Properties of Natural Compounds*, pp. 21–32. doi: 10.1007/978-94-007-6214-5\_3.
- Gabius, H.-J. (2018) "The sugar code: Why glycans are so important", *Biosystems*, 164, pp. 102–111. doi: 10.1016/j.biosystems.2017.07.003.
- Gabius, H.-J. y Roth, J. (2017) "An introduction to the sugar code", *Histochem Cell Biol*, 147(2), pp. 111–117. doi: 10.1007/s00418-016-1521-9.
- Gabius, H. (2002) "The sugar code: functional lectinomics", *Biochim Biophys Acta*, 1572(2–3), pp. 165–177. doi: 10.1016/S0304-4165(02)00306-9.
- Gabor, F., Bogner, E., Weissenboeck, A. y Wirth, M. (2004) "The lectin–cell interaction and its implications to intestinal lectin-mediated drug delivery", *Adv Drug Deliv Rev*, 56(4), pp. 459–480. doi: 10.1016/j.addr.2003.10.015.
- Gallily, R., Vray, B., Stain, I. y Sharon, N. (1984) "Wheat germ agglutinin potentiates uptake of bacteria by murine peritoneal macrophages.", *Immunology*, 52(4), pp. 679–686.
- García-Hernández, E., Zubillaga, R. A., Rodríguez-Romero, A. y Hernández-Arana, A. (2000) "Stereochemical metrics of lectin-carbohydrate interactions: comparison with protein-protein interfaces", *Glycobiology*, 10(10), pp. 993–1000. doi: 10.1093/glycob/10.10.993.
- García-Hernández, E. y Hernández-Arana, A. (1999) "Structural bases of lectin-carbohydrate affinities: Comparison with protein-folding energetics", *Protein Sci*, 8(5), pp. 1075–1086. doi: 10.1110/ps.8.5.1075.
- Gastman, B., Wang, K., Han, J., Zhu, Z., Huang, X., Wang, G.-Q., Rabinowich, H. y Gorelik, E. (2004) "A novel apoptotic pathway as defined by lectin cellular initiation", *Biochem Biophys Res Commun*, 316(1), pp. 263–271. doi: 10.1016/J.BBRC.2004.02.043.
- Gemeiner, P., Mislovičová, D., Tkáč, J., Švitel, J., Pätöprstý, V., Hrabárová, E., Kogan, G. y Kožár, T. (2009) "Lectinomics. II. A highway to biomedical/clinical diagnostics", *Biotechnology Advances*, 27(1), pp. 1–15. doi: 10.1016/j.biotechadv.2008.07.003.

- Ghazarian, H., Idoni, B. y Oppenheimer, S. B. (2011) "A glycochemistry review: Carbohydrates, lectins and implications in cancer therapeutics", *Acta Histochemica*, 113(3), pp. 236–247. doi: 10.1016/j.acthis.2010.02.004.
- Hakomori, S. (1985) "Aberrant Glycosylation in Cancer Cell Membranes as Focused on Glycolipids: Overview and Perspectives", *Cancer Research*, 45(6), pp. 2405–2414.
- Hamid, R., Masood, A., Wani, I. H. y Rafiq, S. (2013) "Lectins: Proteins with diverse applications", *Journal of Applied Pharmaceutical Science*, 3, pp. S93–S103. doi: 10.7324/JAPS.2013.34.S18.
- Harayama, T. y Riezman, H. (2018) "Understanding the diversity of membrane lipid composition", *Nat Rev Mol Cell Biol.*, 19(5), pp. 281–296. doi: 10.1038/nrm.2017.138.
- Hart, G. W., Housley, M. P. y Slawson, C. (2007) "Cycling of O-linked  $\beta$ -N-acetylglucosamine on nucleocytoplasmic proteins", *Nature*, 446(7139), pp. 1017–1022. doi: 10.1038/nature05815.
- He, H., Lu, Y., Qi, J., Zhu, Q., Chen, Z. y Wu, W. (2019) "Adapting liposomes for oral drug delivery", *Acta Pharmaceutica Sinica B*, 9(1), pp. 36–48. doi: 10.1016/j.apsb.2018.06.005.
- Heinrich, E. L., Welty, L. A. Y., Banner, L. R. y Oppenheimer, S. B. (2005) "Direct targeting of cancer cells: A multiparameter approach", *Acta Histochemica*, 107(5), pp. 335–344. doi: 10.1016/j.acthis.2005.06.013.
- Hirabayashi, J. (2004) "Lectin-based structural glycomics: Glycoproteomics and glycan profiling", *Glycoconj J.*, 21(1/2), pp. 35–40. doi: 10.1023/B:GLYC.0000043745.18988.a1.
- Hirabayashi, J. (2008) "Concept, Strategy and Realization of Lectin-based Glycan Profiling", *J Biochem.*, 144(2), pp. 139–147. doi: 10.1093/jb/mvn043.
- Hirsch, A. M. (1999) "Role of lectins (and rhizobial exopolysaccharides) in legume nodulation", *Curr Opin Plant Biol.*, 2(4), pp. 320–326. doi: 10.1016/S1369-5266(99)80056-9.
- Itakura, Y., Nakamura-Tsuruta, S., Kominami, J., Tateno, H. y Hirabayashi, J. (2017) "Sugar-Binding Profiles of Chitin-Binding Lectins from the Hevein Family: A Comprehensive Study", *Int J Mol Sci.*, 18(6), p. 1160. doi: 10.3390/ijms18061160.
- Janzen, D., Juster, H. y Liener, I. (1976) "Insecticidal action of the phytohemagglutinin in black beans on a bruchid beetle", *Science*, 192(4241), pp. 795–796. doi: 10.1126/science.1265481.
- Ji, X., Chen, Y., Faro, J., Gewurz, H., Bremer, J. y Spear, G. T. (2006) "Interaction of human immunodeficiency virus (HIV) glycans with lectins of the human immune system.", *Current protein & peptide science*, 7(4), pp. 317–324. doi: 10.2174/138920306778017990.
- Kilpatrick, D. C. (2002) "Animal lectins: a historical introduction and overview.", *Biochimica et biophysica acta*, 1572(2–3), pp. 187–197.
- Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J. y Bundle, D. R. (2000) "Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands", *Nature*, 403(6770), pp. 669–672. doi: 10.1038/35001095.
- Koyama, S., Terashima, S., Takano, Y., Ohori, T., Kanno, T., Hoshino, Y. y Inoue, H. (1999) "Wheat germ agglutinin binding is a useful prognostic indicator in stomach cancer", *International Journal of Clinical Oncology*, 4(2), pp. 96–100. doi: 10.1007/s101470050034.
- Kumar, K., Reddy, G., Reddy, B., Shekar, P., Sumanthi, J. y Chandra, K. (2012) "Biological role of lectins: A review", *Journal of Orofacial Sciences*, 4(1), pp. 20–25. doi: 10.4103/0975-8844.99883.
- Kuo, J. C.-H., Gandhi, J. G., Zia, R. N. y Paszek, M. J. (2018) "Physical biology of the cancer cell glycocalyx", *Nature Physics*, 14(7), pp. 658–669. doi: 10.1038/s41567-018-0186-9.
- Kuo, Y.-C., Lin, C.-Y., Li, J.-S. y Lou, Y.-I. (2017) "Wheat germ agglutinin-conjugated liposomes incorporated with cardiolipin to improve neuronal survival in Alzheimer's disease treatment", *Int J Nanomedicine*, 12, pp. 1757–1774. doi: 10.2147/IJN.S128396.
- Kuriyan, J., Konforti, B. y Wemmer, D. (2013) *The molecules of life: physical and chemical principles*.
- Laskowski, R. A. y Swindells, M. B. (2011) "LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery", *J Chem Inf Model.*, 51(10), pp. 2778–2786. doi: 10.1021/ci200227u.
- Libbrecht, S., Van den Haute, C., Malinouskaya, L., Gijssbers, R. y Baekelandt, V. (2017) "Evaluation of WGA–Cre-dependent topological transgene expression in the rodent brain", *Brain Struct Funct.*, 222(2), pp. 717–733. doi: 10.1007/s00429-016-1241-x.
- Liu, S., Song, X.-L., Wang, Y.-H., Wang, X.-M., Xiao, Y., Wang, X., Cheng, L. y Li, X.-T. (2017) "The efficacy of WGA modified daunorubicin anti-resistant liposomes in treatment of drug-resistant MCF-7 breast cancer", *J Drug Target.*, 25(6), pp. 541–553. doi: 10.1080/1061186X.2017.1298602.
- Liu, Y., Zhao, Y., Liu, J., Zhang, M., Yu, M., Feng, N., Zhang, Q., Jiang, X., Kim, J., Her, S., Oh, Y. K., Kwon, I. C., Kim, K., Jeong, S. Y., Lee, L. J., Marcucci, G. y Lee, R. J. (2016) "Wheat germ agglutinin modification of lipid–polymer hybrid nanoparticles: enhanced cellular uptake and

- bioadhesion", *RSC Adv.*, 6(42), pp. 36125–36135. doi: 10.1039/C6RA04023C.
- Lovegrove, A., Edwards, C. H., De Noni, I., Patel, H., El, S. N., Grassby, T., Zielke, C., Ulmius, M., Nilsson, L., Butterworth, P. J., Ellis, P. R. y Shewry, P. R. (2017) "Role of polysaccharides in food, digestion, and health.", *Crit Rev Food Sci Nutr.*, 57(2), pp. 237–253. doi: 10.1080/10408398.2014.939263.
- Makhlof, A., Fujimoto, S., Tozuka, Y. y Takeuchi, H. (2011) "In vitro and in vivo evaluation of WGA–carbopol modified liposomes as carriers for oral peptide delivery", *Eur J Pharm Biopharm.*, 77(2), pp. 216–224. doi: 10.1016/j.ejpb.2010.12.008.
- Manning, J. C., Romero, A., Habermann, F. A., García Caballero, G., Kaltner, H. y Gabius, H.-J. (2017) "Lectins: a primer for histochemists and cell biologists", *Histochem Cell Biol.*, 147(2), pp. 199–222. doi: 10.1007/s00418-016-1524-6.
- Marth, J. D. (2008) "A unified vision of the building blocks of life", *Nat Cell Biol.*, 10(9), pp. 1015–1015. doi: 10.1038/ncb0908-1015.
- Martínez-Caballero, S., Cano-Sánchez, P., Mares-Mejía, I., Díaz-Sánchez, A. G., Macías-Rubalcava, M. L., Hermoso, J. A. y Rodríguez-Romero, A. (2014) "Comparative study of two GH19 chitinase-like proteins from *Hevea brasiliensis*, one exhibiting a novel carbohydrate-binding domain", *FEBS Journal*, 281(19), pp. 4535–4554. doi: 10.1111/febs.12962.
- Meng, Y., Hou, X., Lei, J., Chen, M., Cong, S., Zhang, Y., Ding, W., Li, G. y Li, X. (2016) "Multi-functional Liposomes Enhancing Target and Antibacterial Immunity for Antimicrobial and Anti-Biofilm Against Methicillin-Resistant *Staphylococcus aureus*", *Pharm Res.*, 33(3), pp. 763–775. doi: 10.1007/s11095-015-1825-9.
- Mislovičová, D., Gemeiner, P., Kozarova, A. y Kožár, T. (2009) "Lectinomics I. Relevance of exogenous plant lectins in biomedical diagnostics", *Biologia*, 64(1), pp. 1–19. doi: 10.2478/s11756-009-0029-3.
- Munkley, J. y Elliott, D. J. (2016) "Hallmarks of glycosylation in cancer.", *Oncotarget*, 7(23), pp. 35478–89. doi: 10.18632/oncotarget.8155.
- Murakami, Y., Hasegawa, Y., Nagano, K. y Yoshimura, F. (2014) "Characterization of wheat germ agglutinin lectin-reactive glycosylated OmpA-like proteins derived from *Porphyromonas gingivalis*.", *Infect Immun.*, 82(11), pp. 4563–4571. doi: 10.1128/IAI.02069-14.
- Nagahora H., Ishikawa K, Niwa Y, Muraki M, J. Y. (1992) "Expression and secretion of wheat germ agglutinin by *Saccharomyces cerevisiae*", *Eur J Biochem*, 210(3), pp. 989–997. doi: 10.1111/j.1432-1033.1992.tb17504.x.
- Nagahora, H., Harata, K., Muraki, M. y Jigami, Y. (1995) "Site-Directed Mutagenesis and Sugar-Binding Properties of the Wheat Germ Agglutinin Mutants Tyr73Phe and Phe116Tyr", *Eur J Biochem*, 233(1), pp. 27–34. doi: 10.1111/j.1432-1033.1995.027\_1.x.
- Nagata, Y. y Burger, M. M. (1974) "Wheat Germ Agglutinin. Molecular characteristics and specificity for sugar binding", *J. Biol. Chem.*, 249(10), pp. 3116–3122.
- Neutsch, L., Plattner, V. E., Polster-Wildhofen, S., Zidar, A., Chott, A., Borchard, G., Zechner, O., Gabor, F. y Wirth, M. (2011) "Lectin Mediated Biorecognition as a Novel Strategy for Targeted Delivery to Bladder Cancer", *J Urol.*, 186(4), pp. 1481–1488. doi: 10.1016/j.juro.2011.05.040.
- Neutsch, L., Eggenreich, B., Herwig, E., Marchetti-Deschmann, M., Allmaier, G., Gabor, F. y Wirth, M. (2012) "Lectin bioconjugates trigger urothelial cytoinvasion – A glycotargeted approach for improved intravesical drug delivery", *Eur J Pharm Biopharm.*, 82(2), pp. 367–375. doi: 10.1016/j.ejpb.2012.07.016.
- Neutsch, L., Wirth, E.-M., Spijker, S., Pichl, C., Kählig, H., Gabor, F. y Wirth, M. (2013) "Synergistic targeting/prodrug strategies for intravesical drug delivery — Lectin-modified PLGA microparticles enhance cytotoxicity of stearoyl gemcitabine by contact-dependent transfer", *J Control Release*, 169(1–2), pp. 62–72. doi: 10.1016/J.JCONREL.2013.04.004.
- Nowell, P. C. (1960) "Phytohemagglutinin: an initiator of mitosis in cultures of normal human leukocytes.", *Cancer research*, 20(4), pp. 462–6.
- Olsson, T. S. G., Williams, M. A., Pitt, W. R. y Ladbury, J. E. (2008) "The Thermodynamics of Protein–Ligand Interaction and Solvation: Insights for Ligand Design", *J Mol Biol.*, 384(4), pp. 1002–1017. doi: 10.1016/J.JMB.2008.09.073.
- Perozzo, R., Folkers, G. y Scapozza, L. (2004) "Thermodynamics of Protein–Ligand Interactions: History, Presence, and Future Aspects", *J Recept Signal Transduct Res.*, 24(1–2), pp. 1–52. doi: 10.1081/RRS-120037896.
- Portillo-Téllez, M. D. C., Bello, M., Salcedo, G., Gutiérrez, G., Gómez-Vidales, V. y García-Hernández, E. (2011) "Folding and homodimerization of wheat germ agglutinin.", *Biophys J.*, 101(6), pp. 1423–1431. doi: 10.1016/j.bpj.2011.07.037.
- Pulido, N. O., Chavelas, E. A., Turner, F. y García-Hernández, E. (2008) "Current applications of isothermal titration calorimetry to the study of protein complexes", en García-Hernández, E. y Fernández-Velasco, D. A. (eds.) *Advances in Protein Physical Chemistry*, pp. 115–138.
- Reyes-López, C. A., Hernández-Santoyo, A., Pedraza-Escalona, M., Mendoza, G., Hernández-Arana, A. y Rodríguez-Romero, A. (2004) "Insights into a conformational epitope of Hev b 6.02 (hevein).", *Biochem Biophys Res Commun.*, 314(1), pp. 123–130. doi: 10.1016/j.bbrc.2003.12.068.

- Reynolds, C. H. y Holloway, M. K. (2011) "Thermodynamics of Ligand Binding and Efficiency", *ACS Medicinal Chemistry Letters*, 2(6), pp. 433–437. doi: 10.1021/ml200010k.
- Rice, R. H. y Etzler, M. E. (1974) "Subunit structure of wheat germ agglutinin", *Biochem Biophys Res Commun.*, 59(1), pp. 414–419. doi: 10.1016/S0006-291X(74)80222-6.
- Rodríguez-Romero, A., Arreguín, B. y Hernández-Arana, A. (1989) "Unusual far-ultraviolet circular dichroism of wheat germ agglutinin and hevein originated from cystine residues", *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 998(1), pp. 21–24. doi: 10.1016/0167-4838(89)90113-1.
- Rodríguez-Romero, A., Ravichandran, K. G. y Soriano-García, M. (1991) "Crystal structure of hevein at 2.8 Å resolution", *FEBS Letters*, 291(2), pp. 307–309. doi: 10.1016/0014-5793(91)81308-U.
- Saucedo, A. L., del Rio-Portilla, F., Picco, C., Estrada, G., Prestipino, G., Possani, L. D., Delepierre, M. y Corzo, G. (2012) "Solution structure of native and recombinant expressed toxin Cssl from the venom of the scorpion *Centruroides suffusus suffusus*, and their effects on Nav1.5 Sodium channels", *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1824(3), pp. 478–487. doi: 10.1016/j.bbapap.2012.01.003.
- Schwarz, R. E., Wojciechowicz, D. C., Picon, A. I., Schwarz, M. A. y Paty, P. B. (1999) "Wheatgerm agglutinin-mediated toxicity in pancreatic cancer cells", *Br J Cancer.*, 80(11), pp. 1754–1762. doi: 10.1038/sj.bjc.6690593.
- Schwefel, D., Maierhofer, C., Beck, J. G., Seeberger, S., Diederichs, K., Möller, H. M., Welte, W. y Wittmann, V. (2010) "Structural basis of multivalent binding to wheat germ agglutinin.", *J Am Chem Soc.*, 132(25), pp. 8704–8719. doi: 10.1021/ja101646k.
- Sharon, N. y Lis, H. (1972) "Lectins: Cell-Agglutinating and Sugar-Specific Proteins", *Science*, 177(4053), pp. 949–959. doi: 10.1126/science.177.4053.949.
- Sharon, N. y Lis, H. (2004) "History of lectins: from hemagglutinins to biological recognition molecules", *Glycobiology*, 14(11), pp. 53R–62R. doi: 10.1093/glycob/cwh122.
- Shida, K., Misonou, Y., Korekane, H., Seki, Y., Noura, S., Ohue, M., Honke, K. y Miyamoto, Y. (2009) "Unusual accumulation of sulfated glycosphingolipids in colon cancer cells", *Glycobiology*, 19(9), pp. 1018–1033. doi: 10.1093/glycob/cwp083.
- Siebert, H.-C., von der Lieth, C.-W., Kaptein, R., Soedjanaatmadja, U. M. S., Vliegenthart, J. F. G., Wright, C. S. y Gabius, H.-J. (1996) "Role of Aromatic Amino Acids in Carbohydrate Binding. Laser Photo CIDNP(Chemically Induced Dynamic Nuclear Polarisation) and Molecular Modeling Study of Hevein-domain Containing Lectins", *Journal of Molecular Modeling*, 2(9), pp. 351–353. doi: 10.1007/s0089460020351.
- Solís, D., Bovin, N. V., Davis, A. P., Jiménez-Barbero, J., Romero, A., Roy, R., Smetana, K. y Gabius, H.-J. (2015) "A guide into glycosciences: How chemistry, biochemistry and biology cooperate to crack the sugar code", *Biochim Biophys Acta.*, 1850(1), pp. 186–235. doi: 10.1016/j.bbagen.2014.03.016.
- Su, H. y Xu, Y. (2018) "Application of ITC-Based Characterization of Thermodynamic and Kinetic Association of Ligands With Proteins in Drug Design", *Front Pharmacol.*, 9, p. 1133. doi: 10.3389/fphar.2018.01133.
- Sumner, J. B. y Howell, S. F. (1936) "Identification of Hemagglutinin of Jack Bean with Concanavalin A.", *J Bacteriol.*, 32(2), pp. 227–237.
- Tang, H., Hsueh, P., Kletter, D., Bern, M. y Haab, B. (2015) "The Detection and Discovery of Glycan Motifs in Biological Samples Using Lectins and Antibodies: New Methods and Opportunities", en *Adv Cancer Res.*, pp. 167–202. doi: 10.1016/bs.acr.2014.11.003.
- Teuschl, A. H., Neutsch, L., Monforte, X., Rünzler, D., van Griensven, M., Gabor, F. y Redl, H. (2014) "Enhanced cell adhesion on silk fibroin via lectin surface modification.", *Acta Biomater.*, 10(6), pp. 2506–17. doi: 10.1016/j.actbio.2014.02.012.
- Urtasun, N., Baieli, M. F., Cascone, O., Wolman, F. J. y Miranda, M. V. (2015) "High-level expression and purification of recombinant wheat germ agglutinin in *Rachiplusia nu* larvae", *Process Biochemistry*, 50(1), pp. 40–47. doi: 10.1016/j.procbio.2014.11.004.
- Ushiyama, A., Kataoka, H. y Iijima, T. (2016) "Glycocalyx and its involvement in clinical pathophysiology.", *J Intensive Care.*, 4(1), p. 59. doi: 10.1186/s40560-016-0182-z.
- Varki, A. (2007) "Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins", *Nature*, 446(7139), pp. 1023–1029. doi: 10.1038/nature05816.
- Varki, A. y Sharon, N. (2009) "Historical Background and Overview", en *Essentials of Glycobiology*.
- Velazquez-Campoy, A., Leavitt, S. A. y Freire, E. (2004) "Characterization of Protein–Protein Interactions by Isothermal Titration Calorimetry", en *Protein-Protein Interactions*, pp. 35–54. doi: 10.1385/1-59259-762-9:035.
- Wilks, S., de Graaf, M., Smith, D. J. y Burke, D. F. (2012) "A review of influenza haemagglutinin receptor binding as it relates to pandemic properties.", *Vaccine*, 30(29), pp. 4369–4376. doi: 10.1016/j.vaccine.2012.02.076.

- Wimmerová, M., Kozmon, S., Nečasová, I., Mishra, S. K., Komárek, J. y Koča, J. (2012) "Stacking interactions between carbohydrate and protein quantified by combination of theoretical and experimental methods.", *PLoS One*, 7(10), pp. 1–9. doi: 10.1371/journal.pone.0046032.
- Wright, C. S. (1989) "Comparison of the refined crystal structures of two wheat germ isolectins", *J Mol Biol.*, 209(3), pp. 475–487. doi: 10.1016/0022-2836(89)90011-9.
- Wright, C. S. (1992) "Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. Structural basis for cooperative lectin-cell binding", *J Biol Chem.*, 267(20), pp. 14345–14352.
- Wright, C. S. y Kellogg, G. E. (1996) "Differences in hydrophobic properties of ligand binding at four independent sites in wheat germ agglutinin-oligosaccharide crystal complexes.", *Protein Sci.*, 5(8), pp. 1466–1476. doi: 10.1002/pro.5560050803.
- Wright, C. S. y Raikhel, N. (1989) "Sequence variability in three wheat germ agglutinin isolectins: Products of multiple genes in polyploid wheat", *J Mol Evol.*, 28(4), pp. 327–336. doi: 10.1007/BF02103429.
- Yang, H., Li, Z., Shan, M., Li, C., Qi, H., Gao, Q., Wang, J. y Zhang, C. (2015) "Electrogenerated chemiluminescence biosensing for the detection of prostate PC-3 cancer cells incorporating antibody as capture probe and ruthenium complex-labelled wheat germ agglutinin as signal probe.", *Anal Chim Acta.*, 863, pp. 1–8. doi: 10.1016/j.aca.2014.09.001.
- Yang, N., Feng, S., Shedden, K., Xie, X., Liu, Y., Rosser, C. J., Lubman, D. M. y Goodison, S. (2011) "Urinary Glycoprotein Biomarker Discovery for Bladder Cancer Detection using LC-MS/MS and Label-free Quantification", *Clin Cancer Res.*, 17(10), pp. 3349–3359. doi: 10.1158/1078-0432.CCR-10-3121.
- Yang, R.-Y., Rabinovich, G. A. y Liu, F.-T. (2008) "Galectins: structure, function and therapeutic potential", *Expert Rev Mol Med.*, 10, pp. 1–24. doi: 10.1017/S1462399408000719.
- Zhang, N., Ping, Q. N., Huang, G. H. y Xu, W. F. (2005) "Investigation of lectin-modified insulin liposomes as carriers for oral administration", *Int J Pharm.*, 294(1–2), pp. 247–259. doi: 10.1016/j.ijpharm.2005.01.018.

## ANEXO

Leyva, E., Medrano-Cerano, J. L., Cano-Sánchez, P., López-González, I., Gómez-Velasco, H., del Río-Portilla, F. y García-Hernández, E. (2019) "Bacterial expression, purification and biophysical characterization of wheat germ agglutinin and its four hevein-like domains", *Biopolymers*, 110(1), p. e23242. doi: 10.1002/bip.23242.



## ORIGINAL ARTICLE

# Bacterial expression, purification and biophysical characterization of wheat germ agglutinin and its four hevein-like domains

Eduardo Leyva | Jorge L. Medrano-Cerano | Patricia Cano-Sánchez | Itzel López-González | Homero Gómez-Velasco | Federico del Río-Portilla | Enrique García-Hernández

Instituto de Química, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, México, Mexico

**Correspondence**

Enrique García-Hernández, Instituto de Química, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, México 04510, D.F., Mexico.

Email: egarciah@unam.mx

**Funding information**

Consejo Nacional de Ciencia y Tecnología, Grant/Award Number: 0224747235831; Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, Grant/Award Number: PAPIIT, IN205018

**Abstract**

Wheat germ agglutinin (WGA), a chitin binding lectin, has attracted increasing interest because of its unique characteristics such as conformational stability, binding specificity and transcytosis capacity. To pave the way for the study of the molecular basis of WGA's structural stability and binding capacity, as well as to facilitate its use in biomedical and biotechnological developments, we produced recombinant WGA and its 4 isolated hevein-like domains in a bacterial system. All the proteins were expressed as fusion constructs linked to a thioredoxin domain, which was enzymatically or chemically released. The structural and ligand-binding properties of recombinant WGA were similar to the wild lectin. The 4 isolated domains folded and were ligand-binding competent, indicating that each domain constitutes an independent folding unity. The biophysical characterization of the recombinant domains sheds new light on the intricate folding and binding behavior of this emblematic lectin.

**KEYWORDS**

cereal lectin, circular dichroism, hevein-like domain, isothermal titration calorimetry, recombinant lectin

## 1 | INTRODUCTION

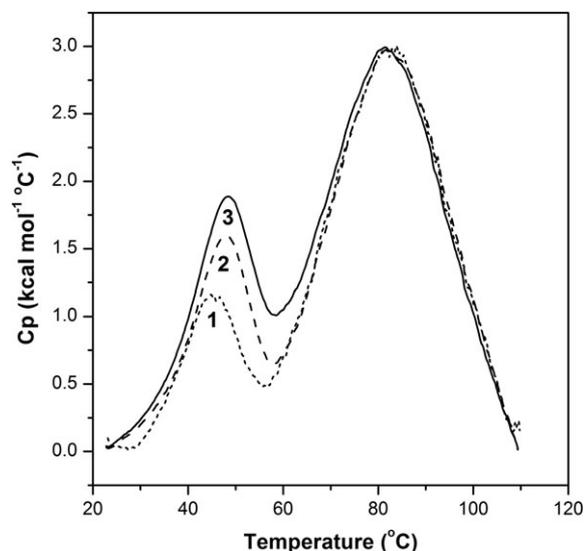
Recent times have witnessed the development of biomedical and biotechnological strategies that increasingly exploit the ability of lectins to recognize glycoconjugates.<sup>[1-5]</sup> The discovery and characterization of new lectins have expanded our ability to decipher the so-called sugar code.<sup>[6-10]</sup> In addition, there is a renewed interest in previously known lectins, with which innovative approaches have been developed.<sup>[7,8,11]</sup> This is the case with wheat germ agglutinin (WGA), which occupies a historic place as the lectin with which the aberrant glycosylation of malignant cells was first revealed.<sup>[12]</sup> Detection of pathogenic agents in human fluids,<sup>[13,14]</sup> lectin blotting,<sup>[15]</sup> early identification of different types of cancer,<sup>[16-18]</sup> quantification of fibrosis in cardiac tissue,<sup>[19]</sup> improvement in methods of tissue engineering,<sup>[20]</sup> and neuroanatomical tracing<sup>[21-23]</sup> are just some of the many recent approaches in which

WGA has demonstrated valuable utility. Given its ability to cross different epithelia and the blood-brain barrier, WGA has been tested to improve drug delivery.<sup>[24-28]</sup> In addition, it has been proposed as an adjuvant to improve the specificity and potency of pharmacological agents against malignant cells and pathogens.<sup>[29,30]</sup>

WGA has a multi-modular architecture.<sup>[31,32]</sup> Its chain of 171 residues is composed of 4 hevein-like domains (labeled A to D).<sup>[33,34]</sup> Each domain is stabilized by 4 highly conserved disulfide bridges.<sup>[35,36]</sup> WGA forms weak homodimers at neutral pH.<sup>[37,38]</sup> With the 2 subunits associated in a "head-to-tail" manner, 4 pairs of domain-domain contacts are established, forming 8 interfacial extended binding sites, structurally grouped into 4 non-equivalent site types.<sup>[39]</sup> The complex architecture of WGA endows this lectin with intricate folding and ligand-binding behaviors. Each domain in the monomeric subunit can bind to a carbohydrate moiety,<sup>[40]</sup> although the corresponding thermodynamic signatures have not yet been determined. In the dimer, the 8 extended binding sites

Eduardo Leyva and Jorge L. Medrano-Cerano contributed equally to this study.



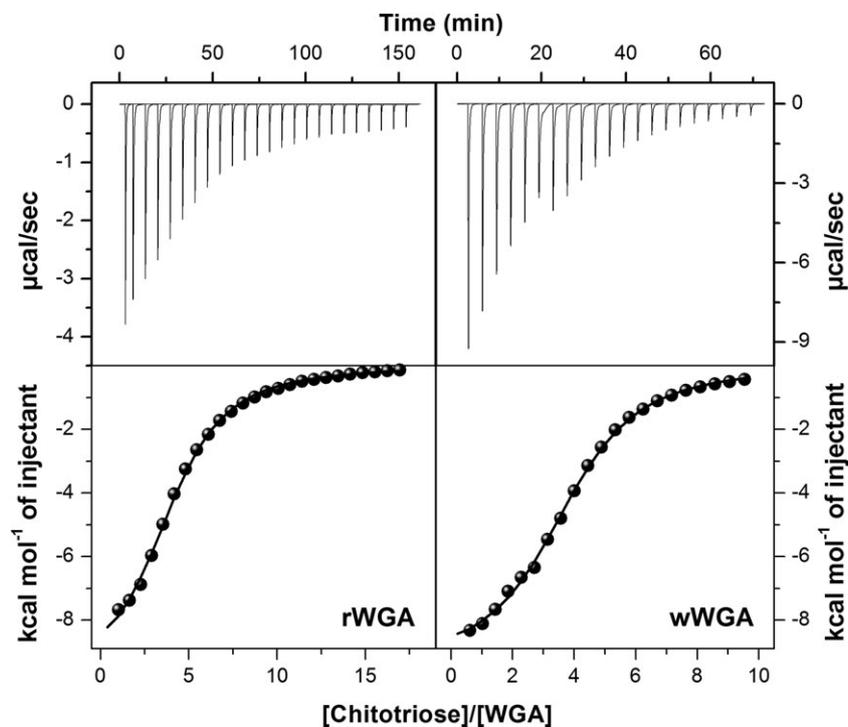


**FIGURE 3** DSC endotherms of rWGA and wWGA at a scan rate of  $1^{\circ}\text{C}/\text{min}$ , 30 mM sodium citrate, pH 3.5. Numerals 1 and 2 in the plot stand for rWGA at a monomer equivalent concentration of 0.05 and 0.11 mM, respectively. Numeral 3 stands for wWGA at a monomer equivalent concentration of 0.09 mM. Abbreviations: DSC, differential scanning calorimetry; WGA, wheat germ agglutinin

site fusion proteins (Supporting Information Figures S1 and S2 and Table S1). The Fwd primer for the rWGA and rWGA.A genes contained an additional enterokinase site and 4 alanine residues necessary for effective proteolytic digestion. Since rWGA.B, rWGA.C, and rWGA.D sequences do not have any methionine, they were treated with CNBr after enterokinase cleavage to get the corresponding wild sequences.

Recombinant protein expression was carried out in *E. coli* Rosetta gami (DE3) pLysS strain (Novagen). The overnight cultures

grown at  $37^{\circ}\text{C}$  in lysogeny broth medium were diluted 1:100 in fresh medium containing ampicillin (100  $\mu\text{g}/\text{ml}$ ), kanamycin (25  $\mu\text{g}/\text{ml}$ ), tetracycline (15  $\mu\text{g}/\text{ml}$ ), and chloramphenicol (34  $\mu\text{g}/\text{ml}$ ), and incubated at  $37^{\circ}\text{C}$  until they reached an  $\text{OD}_{600}$  of 0.5–0.6. Then, 0.5 mM isopropyl 1-thio- $\beta$ -D-galactopyranoside (Goldbio, Olivette, Missouri) was added and incubated at  $30^{\circ}\text{C}$  for 16 hours. Cells from 1.5 L were harvested by centrifugation (10 minutes at 5500g,  $4^{\circ}\text{C}$ ), suspended in 40 ml of lysis buffer (Tris 50 mM, 250 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, pH 8.0) and lysed by sonication with a Misonix 3000 sonicator using a 5.0 output with a sonication pulse rate of 20 seconds on, 40 seconds off, for 8 minutes at  $4^{\circ}\text{C}$ . The soluble fractions were clarified by centrifugation (45 minutes at 30 000g), filtered and subjected to affinity chromatography (His-Trap-FF, GE Healthcare, Northampton, Massachusetts). Before sample injection, the affinity column was equilibrated with 5 column volumes of binding buffer (lysis buffer + 30 mM imidazole). The unbound proteins were washed with 5 column volumes of binding buffer, and the fusion proteins were eluted with 5 column volumes of binding buffer containing 300 mM imidazole. After overnight dialysis against Tris 50 mM, pH 8.0, the fusion proteins were treated with 1 unit of enterokinase per 20 mg of protein (Invitrogen, Carlsbad, California) to remove the thioredoxin and the poly-His tag. Except for rWGA, all digestions were performed at  $37^{\circ}\text{C}$  for 16 hours. For rWGA, it was necessary to proteolyze at  $24^{\circ}\text{C}$  for 24 hours due to aggregation issues. After digestion, a second His-Trap affinity column was used. The unbound fractions were dialyzed against distilled water and then lyophilized. rWGA.B, rWGA.C, and rWGA.D were suspended in 50% trifluoroacetic acid for further treatment with CNBr (200 M excess) at room temperature overnight to remove the extra Ala and Met residues. The final



**FIGURE 4** Calorimetric binding isotherms of chitotriose to rWGA (0.18 mM) and wWGA (0.10 mM). The measurements were performed at  $30^{\circ}\text{C}$ , 0.05 M acetate, 0.1 M NaCl, pH 4.7. Abbreviation: WGA, wheat germ agglutinin

**TABLE 1** Calorimetric parameters for the binding of chitotriose to WGA and its 4 isostructural domains

	$K_b$ ( $\text{mM}^{-1}$ )	$\Delta G_b$ (kcal/Mol)	$\Delta H_b$ (kcal/Mol)	$T\Delta S_b$ (kcal/Mol)
wWGA	$13.2 \pm 0.6$	-5.7	$-9.4 \pm 0.1$	-3.7
rWGA	$12.5 \pm 0.6$	-5.7	$-9.9 \pm 0.2$	-4.2
rWGA.A	$1.1 \pm 0.1$	-4.1	$-9.1 \pm 0.3$	-5.0
rWGA.B	$1.1 \pm 0.1$	-4.1	$-10.1 \pm 0.7$	-6.0
rWGA.C	$0.5 \pm 0.1$	-3.7	$-7.6 \pm 0.8$	-3.9
rWGA.D	$0.3 \pm 0.1$	-3.5	$-9.9 \pm 1.3$	-6.4

Abbreviation: WGA, wheat germ agglutinin.

purification was performed by reverse phase high-performance liquid chromatography (HPLC) with an HP-1100 instrument (HP, Palo Alto, California) on an Aeris Peptide 3.6  $\mu\text{M}$  XB-C18 column (Phenomenex, Torrance, California), using a linear gradient of 12%-30% of water-acetonitrile (containing 0.05% trifluoroacetic acid) for 30 minutes at a flow rate of 1 ml/min.

The presence of recombinant protein was monitored in each purification step by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and the concentration was determined by using the bicinchoninic acid method (Thermo Scientific, Waltham, Massachusetts) and UV absorption using theoretical extinction coefficients calculated with the ExPASy ProtParam tool ( $A_{0.1\%}^{280\text{nm}} = 2.20, 1.16, 1.42, 2.14, \text{ and } 1.75$  for rWGA.A, rWGA.B, rWGA.C, rWGA.D, and rWGA, respectively). The molecular masses of the pure recombinant proteins were verified by matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry (MS) in a Bruker Daltonics Microflex LT equipment using a synapinic acid matrix.

### 2.3 | Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded at 20°C in the far-UV region with a JASCO J-720 spectropolarimeter (Jasco Inc., Easton, Maryland) equipped with a PTC-348WI Peltier-type cell holder for temperature control as described elsewhere.<sup>[47]</sup> Protein solutions of ~0.5 mg/ml in 0.05 M glycine-HCl (pH 2.0), acetate (pH 4.7), or citrate

(pH 6.0) buffer supplemented with 0.1 M NaCl were loaded into a quartz cell of 0.1-cm length path. Each spectrum corresponds to the average of 3 repetitive scans after the subtraction of the buffer signal. Ellipticities are reported as mean residue ellipticity,  $[\theta]_{\text{mrw}}$ . Thermal denaturation transitions were followed by monitoring the ellipticity signal at 225 nm, using a constant scan rate of 1°C/min.

### 2.4 | Differential scanning calorimetry

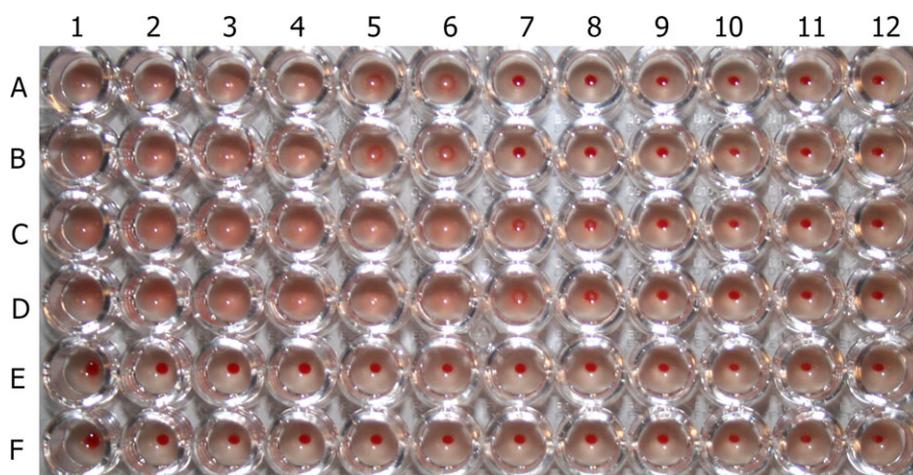
Differential scanning calorimetry (DSC) measurements were carried out with a MicroCal VP capillary DSC equipment (Malvern Panalytical Inc., Westborough, Massachusetts). All experiments were conducted at a scan rate of 60°C/h. Buffer-buffer baselines were obtained under the same experimental conditions and subtracted from sample traces. To achieve a proper thermal history of the instrument, several buffer-buffer baselines were run before recording the protein solution signal.

### 2.5 | NMR experiments

Lyophilized rWGA.D was dissolved at 0.5 mM in H<sub>2</sub>O 5% (v/v) of D<sub>2</sub>O (Cambridge Isotopes Laboratories, Tewksbury, Massachusetts). Nuclear magnetic resonance (NMR) data were acquired at 298 K in a Bruker ASCEND 700 (Bruker BioSpin, Billerica, Massachusetts) spectrometer operating at 700 MHz proton frequency with a 3 nucleus cryoprobe (TCI 700 H-C/N-D 5 mm). The nuclear overhauser spectroscopy (NOESY) spectra were recorded with a mixing time of 300 ms.<sup>[48]</sup> NOESY spectra were collected with 1024 data complex point matrices using 32 scans. NMR data were processed using MestreC program (MestreC Lab Research, Santiago de Compostela, Spain) and a shifted sine window function and zero filling was applied prior to Fourier transformation.

### 2.6 | Isothermal titration calorimetry

Calorimetric measurements for the titration of WGA and the 4 isolated domains with chitotriose were carried out using a MicroCal iTC<sub>200</sub> instrument (GE Healthcare). All experiments were performed in a 0.05 M acetate buffer solution supplemented with 0.1 M NaCl,

**FIGURE 5** Hemagglutination assays of rWGA (rows A and B) and wWGA (rows C and D) using human erythrocytes type O (2% v/v) in PBS pH 7.4. The lectins (25  $\mu\text{g/ml}$ ) were serially 2-fold diluted (lanes 1-10). Negative controls without the lectins were run (rows E and F).

Abbreviation: WGA, wheat germ agglutinin

pH 4.7, at 30°C. Binding experiments with isolated domains contained ~1 mM protein concentration, while the concentration for the whole WGA was ~0.15 mM of monomer equivalent. The titration schedule consisted of 20–25 consecutive injections of the ligand with a 5-minute interval between injections, using a stirring speed of 750 rpm. The dilution heat of the ligand was obtained by adding ligand to a buffer solution under identical conditions and the same injection schedule used with the protein sample. All samples were degassed for 10 minutes before the experiment.

Binding parameters were determined by using an identical and independent binding site model:

$$Q = \frac{nM_t \Delta H_b V_o}{2} \left[ 1 + \frac{L_t}{nM_t} + \frac{1}{nK_b M_t} - \sqrt{\left( 1 + \frac{L_t}{nM_t} + \frac{1}{nK_b M_t} \right)^2 - \frac{4L_t}{nM_t}} \right]$$

where  $Q$  is the normalized heat evolved per mol of ligand at the end of the  $i$ th injection,  $K_b$  is the binding constant,  $\Delta H_b$  is the enthalpy change,  $n$  is the stoichiometry,  $V_o$  is the working volume of the cell, and  $L_t$  and  $M_t$  are the total ligand and macromolecule concentrations, respectively. The heat released in the  $i$ th injection is:

$$\Delta Q(i) = Q(i) + \frac{v_i(Q(i) + Q(i-1))}{2V_o} - Q(i-1) + q_{dil}$$

where  $v_i$  is the aliquot volume added at injection  $i$ , and  $q_{dil}$  is a fitting term introduced to account for experimentally uncorrected dilution heat effects.<sup>[49,50]</sup> Non-linear regressions were carried with Origin 7.0 (OriginLab Co., Northampton, Massachusetts).

## 2.7 | Hemagglutination assay

The hemagglutination activity using human erythrocytes type O was assayed as described elsewhere.<sup>[51]</sup>

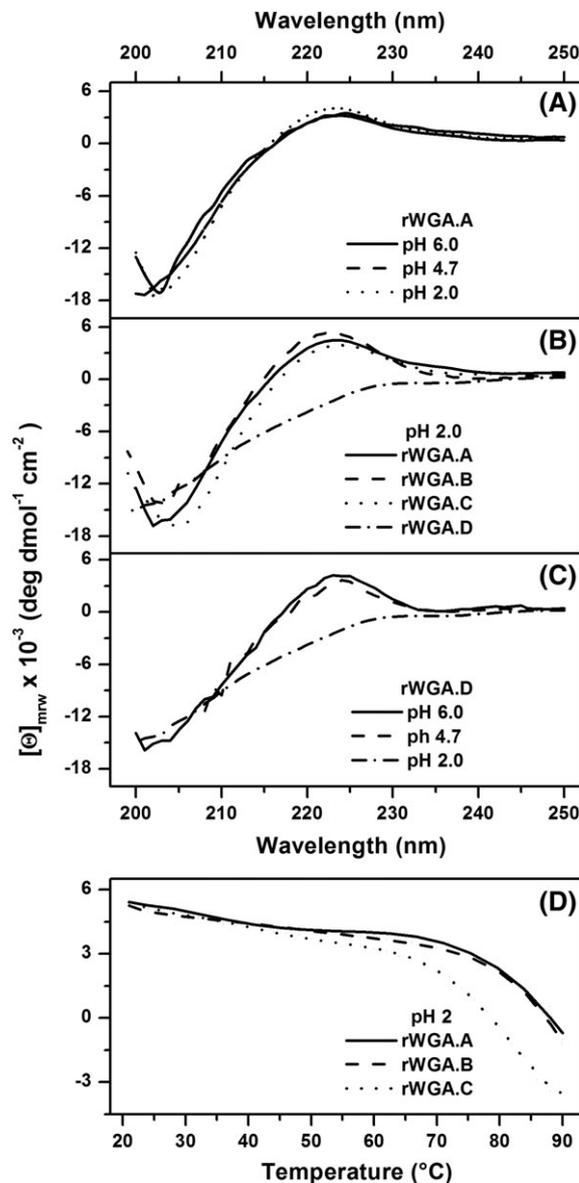
## 3 | RESULTS

### 3.1 | Purification of recombinant proteins

WGA and its 4 isolated domains were expressed in *E. coli* and purified by affinity and reverse phase chromatography (Supporting Information Figure S3). Typically, 1 L of culture medium yielded ~3–5 mg of pure protein (Supporting Information Table S2). Due to the cloning and purification strategies used, rWGA and rWGA.A were produced with 4 additional alanine residues at the N-terminus (Figure 1). According to the MS results, the enterokinase cleavage yielded a minority rWGA product (~10%) containing an additional lysine residue (Supporting Information Figure S2B), that showed the same properties observed for the main recombinant product (data not shown). As rWGA.B, rWGA.C, and rWGA.D have no Met residues, they were produced with 1 Met at position –1. Therefore, treatment with CNBr allowed obtaining wild-type sequences for these 3 domains (Figure 1).

### 3.2 | Characterization of the full agglutinin

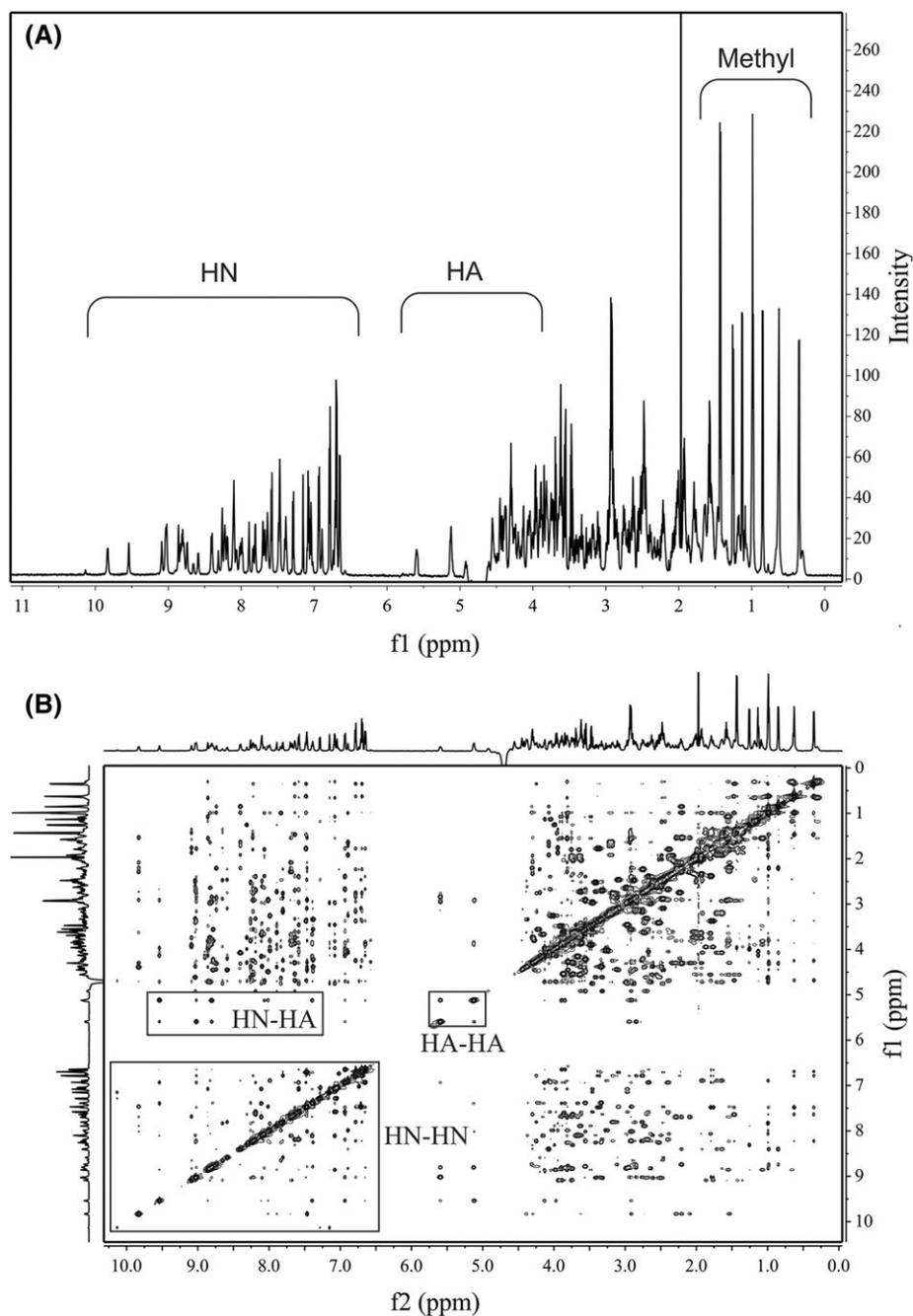
rWGA was characterized biophysically to evaluate its conformational integrity. As shown in Figure 2, the far-UV CD spectra of the recombinant and the wild WGA (wWGA) are nearly superimposable, indicating



**FIGURE 6** A–C, Far-UV CD spectra of the isolated WGA domains at pH 2.0 (0.05 M glycine-HCl), pH 4.7 (0.05 M acetate) and pH 6.0 (0.05 M citrate) supplemented with 0.1 M NaCl, at 20°C. D, Thermal unfolding profiles at pH 2.0 (0.05 M glycine-HCl) monitored by the ellipticity change at 225 nm at a scan rate of 1°C/min. Abbreviation: WGA, wheat germ agglutinin

that the 2 proteins share a very similar content of secondary structure. The positive band centered at ~225 nm is a feature exhibited by hevein-like proteins<sup>[35,46]</sup>. This band was lost when rWGA was incubated in the presence of a reducing agent or at high temperature (Figure 2). Judging by the large recovery of the CD signal after cooling the protein sample to room temperature, the unfolding of the recombinant lectin is a reversible process. The thermal denaturation assays showed that rWGA and wWGA are similarly thermoresistant (Figure 2, inset).

It was previously shown that wWGA displays complex thermal perturbation calorimetric curves, consisting of a low temperature peak corresponding to the dissociation of the dimer, and a wide peak at high temperatures caused by the multi-step unfolding of the free monomers.<sup>[37]</sup> Figure 3 compares the DSC endotherms of rWGA and

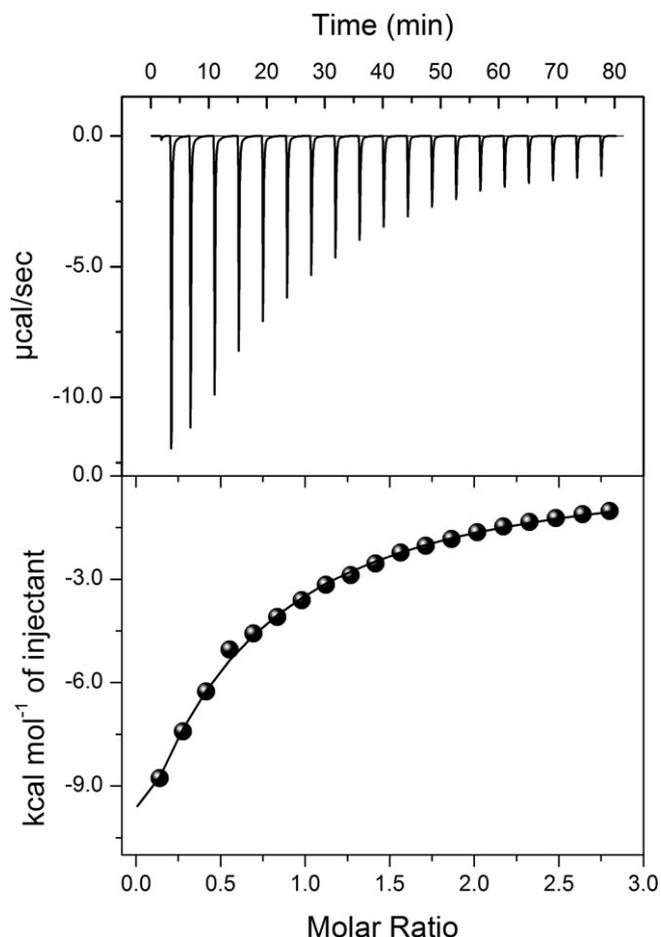


**FIGURE 7** A, 700 MHz NMR  $^1\text{H}$  spectrum of rWGA.D. The spectrum showed great dispersion, evidencing the presence of virtually 1 folded conformation. B, NOESY 2D spectrum at 700 MHz. NOESY experiment obtained at 300 ms mixing time shows HA-HA correlation between 2 protons in a  $\beta$ -sheet. Also, it is possible to observe 3 systems in the HN-HA region with  $\text{HN}_i\text{-HA}_i$  intra residue correlation,  $\text{HN}_i\text{-HA}_{i-1}$  sequential correlation and  $\text{HN}_i\text{-HA}_j$  within the same  $\beta$ -sheet. HN-HN region displayed several correlations between the HN protons, strongly suggesting the formation of at least 1 helix. Abbreviations: NMR, nuclear magnetic resonance; NOESY, nuclear overhauser spectroscopy; WGA, wheat germ agglutinin

wWGA. As can be seen, the recombinant lectin also exhibited 2 clearly defined peaks. The high temperature peak did not show dependence on protein concentration, and overlapped with the corresponding wWGA peak, indicating that the monomers of both proteins exhibit the same thermal stability. The low temperature peak of rWGA was dependent on protein concentration, demonstrating the existence of the dimeric form in the recombinant lectin. However, the dissociation peak of rWGA was smaller and shifted to a lower temperature than the corresponding peak of wWGA at a similar protein concentration. Overall, the DSC results revealed that rWGA forms somewhat weaker

dimers than wWGA, while the monomers of both proteins exhibit very similar thermal stabilities.

To test the binding capacity of rWGA, the recognition of the natural ligand chitotriose ( $\beta 1 \rightarrow 4$  trisaccharide of GlcNAc) was characterized by isothermal titration calorimetry (ITC). We selected pH 4.7 to carry out the calorimetric measurements since this is the highest pH value at which we found agglutinin solubility high enough to obtain well-shaped binding isotherms.<sup>[52]</sup> Figure 4 compares the isotherms obtained for rWGA and wWGA. The parameters resulting from the best non-linear regression fitting of a single binding-site model to the



**FIGURE 8** Calorimetric binding isotherms of rWGA.A titrated with chitotriose. The measurements were performed at 30°C, 0.05 M acetate, 0.1 M NaCl, pH 4.7. Abbreviation: WGA, wheat germ agglutinin

corresponding titration curve are shown in Table 1. Clearly, the 2 WGA versions share the same binding signatures, showing an enthalpically-driven low affinity, which is a recurrent feature of protein-carbohydrate interactions.<sup>[53–55]</sup> The fitting is consistent with the presence of 4 active binding sites per monomer for both agglutinins ( $n = 4.1 \pm 0.1$  and  $3.9 \pm 0.0$  for rWGA and wWGA, respectively). This binding stoichiometry is in agreement with ligand occupancies seen in recently solved crystallographic structures<sup>[41]</sup> as well as in solution binding studies with designed multivalent neoglycopeptides exhibiting nanomolar affinity for WGA.<sup>[56,57]</sup> In an early microcalorimetric study, Bains et al.<sup>[58]</sup> characterized the interaction of wWGA with several chitin-derived oligosaccharides. The binding strength reported in that study for chitotriose ( $K_b = 12.2 \text{ mM}^{-1}$ , pH 4.5, 26°C) is similar to that measured by us, although the binding enthalpy ( $\Delta H_b = -16.0 \text{ kcal/mol}$ ) was more negative. Finally, the hemagglutination activity of rWGA and wWGA were analyzed by serial dilution experiments using human erythrocytes. Figure 5 shows a typical agglutination result. On the average, rWGA showed a 2-fold weaker hemagglutination potency than that exhibited by wWGA.

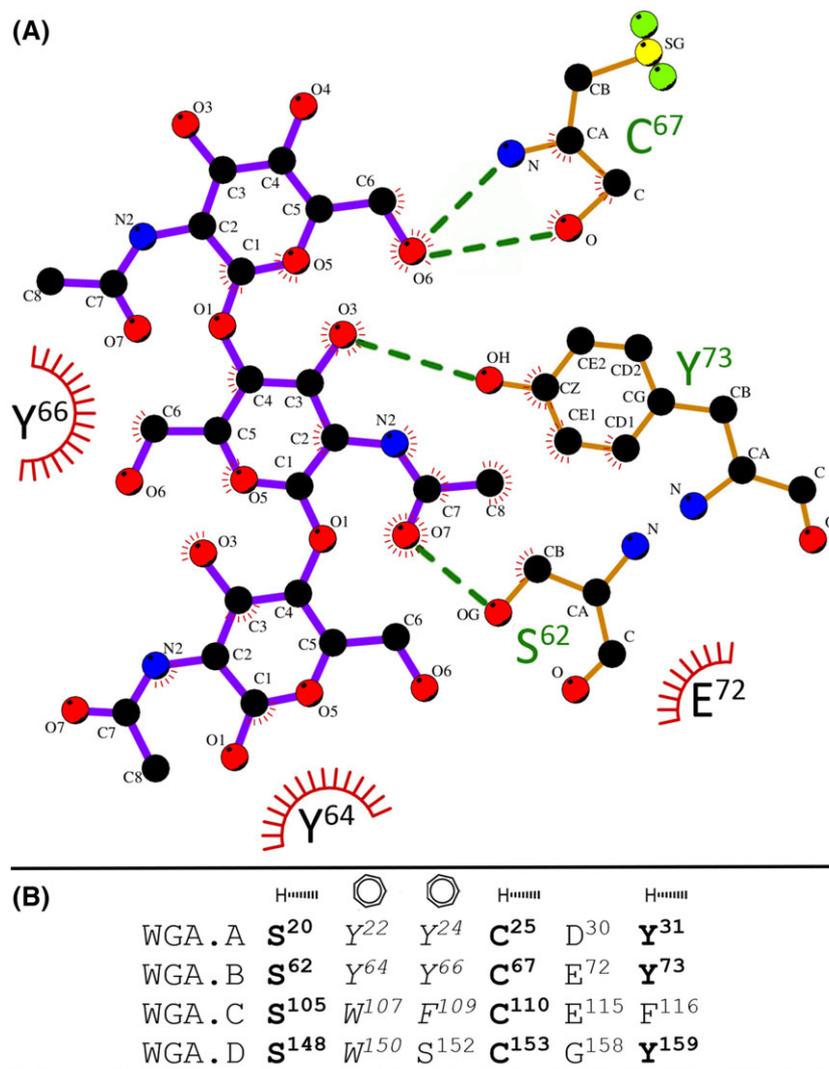
### 3.3 | Characterization of isolated WGA domains

The conformation of the isolated domains as a function of pH was evaluated by far-UV CD spectroscopy. Figure 6A shows the CD

spectra for rWGA.A at 3 different pH values (2.0, 4.7, and 6.0). Regardless of the acidity of the solution, the domain showed the characteristic spectral shape of a hevein-like folded protein. The same behavior was observed for rWGA.B and rWGA.C (Figure 6B). rWGA.D also adopted a native-like conformation at pH 4.7 and 6.0; however, unlike the other 3 domains, it showed an unfolded-like conformation at pH 2 (Figure 6C). The thermal unfolding of wWGA was previously studied by DSC.<sup>[37]</sup> At pH 2, the lectin is completely monomeric. The deconvolution of the unfolding traces revealed a multi-step mechanism, consisting of 3 independent 2-state transitions. Spectroscopic and structural data led to the conclusion that at pH 2, 3 out of the 4 domains of the agglutinin were folded, showing cooperative transitions at high temperatures, while the other domain was unfolded even at low temperature. To further clarify the unfolding mechanism of WGA, the isolated domains were submitted to a heating/cooling process at pH 2 (Figure 6D). The unfolding onset of rWGA.A occurred at  $\sim 65^\circ\text{C}$ , while rWGA.B and rWGA.C started to unfold at  $\sim 75^\circ\text{C}$ . The transition end was not observed in any of the unfolding profiles. Nevertheless, the 3 domains were largely unfolded at  $90^\circ\text{C}$ , judging by the shape of their CD spectra (data not shown). Upon cooling to room temperature, the CD signals were completely recovered, a behavior similar to that observed for the whole WGA.<sup>[37,46]</sup>

<sup>1</sup>H NMR spectrum of rWGA.D was recorded to further assess the conformation of this domain (Figure 7). The 1D spectrum showed well dispersed signals in all HN, HA, backbone and methyl regions (Figure 7A). This experiment confirmed that the domain is folded. Furthermore, no evidence of significant conformational dynamics was found. HN and HA signals observed at a lower field than expected and methyl groups as low as 0.35 ppm corroborated the folded state of the domain. The HN region was well separated, allowing the identification of an NH tryptophan signal at 10.13 ppm and 3 well resolved HN signals at 9.84, 9.83 and 9.53 ppm. Two HA signals were observed at 5.60 and 5.12 ppm. In agreement with the primary sequence of rWGA.D, signals for 8 methyl groups were found between 1.40 and 0.35 ppm; several of them showed protective interaction. NOESY 2D spectrum (Figure 7B) displayed a well-defined NOE between 2 HA-HA (5.60 and 5.12 ppm HA's) and in the NH region at least 3 spin systems with  $\text{HN}_i\text{-HA}_i$  intra residue correlation,  $\text{HN}_i\text{-HA}_{i-1}$  sequential correlation and  $\text{HN}_i\text{-HA}_j$  within a  $\beta$ -sheet (regions delimited by the squares in Figure 7B). These results revealed the existence of a  $\beta$ -sheet in the molecule. Furthermore, several NOE's in many HN strips evidenced the presence of at least 1 helix. Overall, the NMR results agree with the conformation of the domain observed in the crystallographic structures of the entire lectin.

The binding capacity of the 4 isolated domains was evaluated by ITC at pH 4.7. As a representative result, Figure 8 shows the calorimetric isotherm obtained for the titration of rWGA.A with chitotriose. As summarized in Table 1, all domains showed an enthalpically-driven low-affinity towards the saccharide. The values obtained for rWGA.B are somewhat different from those reported previously for the same domain ( $\Delta H_b = -8.1 \text{ kcal/mol}$ ,  $K_b = 908 \text{ M}^{-1}$ , 25°C) by Espinosa et al.<sup>[40]</sup> also using ITC. All the domains showed affinities 10-fold smaller than that observed with the complete agglutinin. By contrast, no significant differences were observed in the binding enthalpies of the whole WGA and its fragmentary constructions. Therefore, these



**FIGURE 9** A, Schematic representation of the binding site of rWGA.B in complex with chitotriose using the NMR coordinates kindly provided by Dr. Jesús Jiménez-Barbero, Center for Cooperative Research in biosciences, CIC-bioGUNE, (Spain).<sup>[40]</sup> Figure was generated with the LigPlus software.<sup>[66]</sup> B, Alignment of the binding residues of the 4 WGA domains. Abbreviations: NMR, nuclear magnetic resonance; WGA, wheat germ agglutinin

results indicate that the increase in affinity shown by the dimeric lectin in relation to its isolated domains is of entropic origin.

## 4 | DISCUSSION

For decades, WGA has attracted keen interest in both fundamental and applied glycobiology. This lectin is a valuable tool for conventional and state-of-the-art analysis of glycoconjugates, and numerous groups are currently pursuing its use in targeted drug delivery applications. Despite this, no report had been presented on the production of WGA in a bacterial expression system. The availability of a cheap and efficient method to obtain recombinant WGA and its isolated domains will facilitate the realization of studies aimed at improving the properties of the lectin for application purposes. Furthermore, it will pave the way to deepen our knowledge about the very complex behavior that the agglutinin exhibits in solution.

In the present study, using a Trx-(poly)His tag fused protein, we obtained ~4 mg of rWGA per liter of *E. coli* culture medium. The

secondary structure content, the thermal resistance, the ligand-binding capacity and the agglutination activity of rWGA were very similar to those exhibited by wWGA. The strategy we used was similar to that used previously by some of us to produce correctly folded and active hevein-like domain containing proteins<sup>[59]</sup> and small arachnid toxins also presenting numerous disulfide bridges.<sup>[60-62]</sup> The use of the Trx-(poly)His tag-fused construct allowed also the bacterial production of the 4 isolated WGA domains in a binding-competent native-like conformation, with expression yields similar to those of the whole agglutinin. This result implies that each domain behaves as an independent folding unit.

According to the ITC measurements, the 4 isolated WGA domains differ somewhat between themselves in ligand binding properties. Thus, despite the similarity in the architecture of their binding sites, they are not energetically equivalent. Espinosa et al.<sup>[40]</sup> reported 2 NMR-solved binding modes for the rWGA.B-chitotriose complex. In 1 binding mode, all 3 GlcNAc units contact the domain; in the other complex, 1 saccharide unit is fully exposed to the solvent. Figure 9 shows a schematic representation of the structure of the first

complex, highlighting residues that are within 5 Å from the ligand. Four hydrogen bonds (formed by residues Ser,<sup>62</sup> Cys,<sup>67</sup> and Tyr<sup>73</sup>) and 2 stacking interactions (formed by Tyr<sup>64</sup> and Tyr<sup>66</sup>) stabilize the interaction of rWGA.B with chitotriose. A sequence analysis shows that rWGA.A preserves the same interacting residues as seen in rWGA.B (Figure 9). Consistently, these 2 domains showed a very similar interaction energetics with chitotriose. By contrast, the position equivalent to Tyr<sup>73</sup> in rWGA.B is occupied by Phe<sup>116</sup> in rWGA.C. In addition, rWGA.C presents W<sup>107</sup> and F<sup>109</sup> instead of Tyr<sup>64</sup> and Tyr<sup>66</sup> in rWGA.B. Overall, the loss of 1 hydrogen bond and the modification of the stacking interactions yield a ligand interaction strength of rWGA.C weaker than those of rWGA.A or rWGA.B. rWGA.D, the only domain lacking 1 of the 2 stacking aromatic residues, showed the weakest ligand affinity. It is interesting to note that the average sequence identity between the 4 WGA domains is ~ 40%, a value that is very similar to that observed among chitin-binding lectins belonging to the hevein family. Recently, it has been shown that hevein-type lectins show diverse oligosaccharide specificity, beyond their ability to recognize chito-oligosaccharides.<sup>[63]</sup> WGA binds with considerable specificity to high-mannose-type N-glycans containing bisecting GlcNAc. It is tempting to speculate that the heterogeneity between the 4 domains of WGA contributes to the ability of this lectin to recognize complex saccharide architectures with high specificity.

The goal of using WGA as a peroral drug delivery device involves the exposure of the agglutinin to the distinctly different pH conditions found throughout the digestive tract. In this sense, the elucidation of the structural stability of the lectin as a function of pH is relevant. Previously, it was shown that in acidic conditions similar to those in the human stomach, WGA is mostly monomerized.<sup>[37,46,64,65]</sup> Under these conditions, the agglutinin exhibits thermally induced profiles consistent with 3 independent 2-state transitions. Different unfolding scenarios were proposed to rationalize the fact that only 3—and not 4—transitions were observed in the calorimetric traces, namely: (1) the existence of a very stable domain, whose unfolding takes place above the temperature experimentally accessible, (2) 2 domains unfold concertedly, or (3) 1 domain is unfolded at room temperature. Based on spectroscopic and calorimetric data of the complete agglutinin, the latter was proposed as the most probable model.<sup>[37]</sup> rWGA.A, rWGA.B, and rWGA.C preserve the typical hevein-like conformation even at pH 2, exhibiting an unfolding transition at high temperatures. By contrast, rWGA.D shows the properties of an unstructured protein at very low pH. Therefore, the study of the isolated domains reinforces the picture that under pH conditions similar to the stomach, 1 of the 4 domains, specifically the domain D, is unfolded, while the other 3 domains are fully folded. In addition, the thermal profiles revealed that domain A is less thermostable in relation to domains B and C. To further clarify the complex unfolding mechanism of WGA, a more detailed characterization of the isolated domains, as well as of constructs composed of 2 (AB, BC, and CD) and 3 (ABC and BCD) domains, is clearly warranted.

## ACKNOWLEDGMENTS

E.L. and J.L.M.C. are doctoral students from Programa de Doctorado en Ciencias Bioquímicas, H.G.V. is a doctoral student from Programa

de Doctorado en Ciencias Biomédicas, both Programs from Universidad Nacional Autónoma de México (UNAM), and received Ph. D. fellowships from CONACyT. We thank the LANEM-IQ-UNAM for the use of the CD Spectropolarimeter and Q.F.B. Ma. del Rocío Patiño Maya for the technical support. We are grateful to M. C. Lucía del Carmen Márquez Alonso, M. C. Lucero Rios Ruiz and Q. Eréndira García Ríos for their assistance in carrying out HPLC analysis and MALDI-TOF spectra. Finally, we also thank Dr. Beatriz Quiroz-García for the support conducting NMR experiments. This work was financed by DGAPA, UNAM [PAPIIT, IN205018] and CONACyT [Grants 235831]. NMR experiments were performed in the UNAM's NMR lab: LURMN at IQ-UNAM which is funded by CONACyT (Project 0224747).

## ORCID

Enrique García-Hernández  <https://orcid.org/0000-0003-4561-4973>

## REFERENCES

- [1] D. Mislovičová, P. Gemeiner, A. Kozarova, T. Kožár, *Biologia (Bratisl)* **2009**, *64*, 1.
- [2] P. Gemeiner, D. Mislovičová, J. Tkáč, J. Svitel, V. Pätöprstý, E. Hrabárová, G. Kogan, T. Kozár, *Biotechnol Adv* **2009**, *27*, 1.
- [3] L. Fu, C. Zhou, S. Yao, J. Yu, B. Liu, J. Bao, *Int J Biochem Cell Biol* **2011**, *43*, 1442.
- [4] B. Islam, A. U. Khan, *Protein Purification*, InTech, London, UK **2012**, p. 167.
- [5] B. Shakhsher, M. Anderson, K. Khatib, L. Tadoori, L. Joshi, F. Lisacek, L. Hirschman, E. Mullen, *J Mol Recognit* **2013**, *26*, 426.
- [6] S. K. Lam, T. B. Ng, *Appl Microbiol Biotechnol* **2011**, *89*, 45.
- [7] R. Hamid, A. Masood, I. H. Wani, S. Rafiq, *J Appl Pharm Sci* **2013**, *3*, S93.
- [8] J. Hirabayashi, M. Yamada, A. Kuno, H. Tateno, *Chem Soc Rev* **2013**, *42*, 4443.
- [9] A. Varrot, S. M. Basheer, A. Imberty, *Curr Opin Struct Biol* **2013**, *23*, 678.
- [10] D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana, H.-J. Gabius, *Biochim Biophys Acta* **2015**, *1850*, 186.
- [11] T. Hayashi, Y. Sun, T. Tamura, K. Kuwata, Z. Song, Y. Takaoka, I. Hamachi, *J Am Chem Soc* **2013**, *135*, 12252.
- [12] J. C. Aub, B. H. Sanford, M. N. Cote, *Proc Natl Acad Sci USA* **1965**, *54*, 396.
- [13] Y. Murakami, Y. Hasegawa, K. Nagano, F. Yoshimura, *Infect Immun* **2014**, *82*, 4563.
- [14] P. de Brito Ximenes, E. I. C. Beltrão, D. P. C. Macêdo, M. D. S. Buonafina, R. G. de Lima-Neto, R. P. Neves, *Molecules* **2015**, *20*, 3776.
- [15] T. Manalo, A. May, J. Quinn, D. S. Lafontant, O. Shifatu, W. He, J. M. Gonzalez-Rosa, G. C. Burns, C. E. Burns, A. R. Burns, P. J. Lafontant, *J Histochem Cytochem* **2016**, *64*, 687.
- [16] J. Baeten, A. Suresh, A. Johnson, K. Patel, M. Kuriakose, A. Flynn, D. Kademani, *Transl Oncol* **2014**, *7*, 213.
- [17] L. Neutsch, B. Eggenreich, E. Herwig, M. Marchetti-Deschmann, G. Allmaier, F. Gabor, M. Wirth, *Pharm Res* **2014**, *31*, 819.
- [18] H. Yang, Z. Li, M. Shan, C. Li, H. Qi, Q. Gao, J. Wang, C. Zhang, *Anal Chim Acta* **2015**, *863*, 1.
- [19] B. Emde, A. Heinen, A. Gödecke, K. Bottermann, *Eur J Histochem* **2014**, *58*, 2448.
- [20] A. H. Teuschl, L. Neutsch, X. Monforte, D. Rünzler, M. van Griensven, F. Gabor, H. Redl, *Acta Biomater* **2014**, *10*, 2506.
- [21] B. Gómez-González, H. M. Larios, A. Escobar, *Int J Dev Neurosci* **2011**, *29*, 839.
- [22] S. L. Levy, J. J. White, R. V. Sillitoe, *Neuromethods*, Vol. 92, Springer, New York, NY **2015**, p. 51.

- [23] S. Libbrecht, C. Van den Haute, L. Malinouskaya, R. Gijsbers, V. Baekelandt, *Brain Struct Funct* **2017**, *222*, 717.
- [24] F. Gabor, E. Bogner, A. Weissenboeck, M. Wirth, *Adv Drug Deliv Rev* **2004**, *56*, 459.
- [25] Q. Liu, X. Shao, J. Chen, Y. Shen, C. Feng, X. Gao, Y. Zhao, J. Li, Q. Zhang, X. Jiang, *Toxicol Appl Pharmacol* **2011**, *251*, 79.
- [26] X.-T. Li, R.-J. Ju, X.-Y. Li, F. Zeng, J.-F. Shi, L. Liu, C.-X. Zhang, M.-G. Sun, J.-N. Lou, W.-L. Lu, *Oncotarget* **2014**, *5*, 6497.
- [27] Y. Liu, Y. Zhao, J. Liu, M. Zhang, M. Yu, N. Feng, Q. Zhang, X. Jiang, J. Kim, S. Her, Y. K. Oh, I. C. Kwon, K. Kim, S. Y. Jeong, L. J. Lee, G. Marcucci, R. J. Lee, *RSC Adv* **2016**, *6*, 36125.
- [28] C. M. Pichl, B. Dunkl, B. Brauner, F. Gabor, M. Wirth, L. Neutsch, *Pathogens (Basel, Switzerland)* **2016**, *5*, 16.
- [29] C. Apfelthaler, M. Anzengruber, F. Gabor, M. Wirth, *Eur J Pharm Biopharm* **2017**, *115*, 131.
- [30] Y. Meng, X. Hou, J. Lei, M. Chen, S. Cong, Y. Zhang, W. Ding, G. Li, X. Li, *Pharm Res* **2016**, *33*, 763.
- [31] C. S. Wright, *J Mol Biol* **1989**, *209*, 475.
- [32] C. S. Wright, *J Biol Chem* **1992**, *267*, 14345.
- [33] A. Rodríguez-Romero, K. G. Ravichandran, M. Soriano-García, *FEBS Lett* **1991**, *291*, 307.
- [34] J. Jiménez-Barbero, F. Javier Cañada, J. L. Asensio, N. Aboitiz, P. Vidal, A. Canales, P. Groves, H.-J. Gabius, H.-C. Siebert, *Adv Carbohydr Chem Biochem* **2006**, *60*, 303.
- [35] A. Rodríguez-Romero, B. Arreguín, A. Hernández-Arana, *Biochim Biophys Acta - Protein Struct Mol Enzymol* **1989**, *998*, 21.
- [36] N. H. Andersen, B. Cao, A. Rodríguez-Romero, B. Arreguín, *Biochemistry* **1993**, *32*, 1407.
- [37] M. D. C. Portillo-Téllez, M. Bello, G. Salcedo, G. Gutiérrez, V. Gómez-Vidales, E. García-Hernández, *Biophys J* **2011**, *101*, 1423.
- [38] R. Afrin, A. Ikai, *FEBS Lett* **2014**, *588*, 4472.
- [39] C. S. Wright, G. E. Kellogg, *Protein Sci* **1996**, *5*, 1466.
- [40] J. F. Espinosa, J. L. Asensio, J. L. García, J. Laynez, M. Bruix, C. Wright, H. C. Siebert, H. J. Gabius, F. J. Cañada, J. Jiménez-Barbero, *Eur J Biochem* **2000**, *267*, 3965.
- [41] D. Schwefel, C. Maierhofer, J. G. Beck, S. Seeberger, K. Diederichs, H. M. Möller, W. Welte, V. Wittmann, *J Am Chem Soc* **2010**, *132*, 8704.
- [42] H. Nagahora, K. Ishikawa, Y. Niwa, M. Muraki, Y. Jigami, *Eur J Biochem* **1992**, *210*, 989.
- [43] H. Nagahora, K. Harata, M. Muraki, Y. Jigami, *Eur J Biochem* **1995**, *233*, 27.
- [44] N. Urtasun, M. F. Baieli, O. Cascone, F. J. Wolman, M. V. Miranda, *Process Biochem* **2015**, *50*, 40.
- [45] H. C. Siebert, C. W. von der Lieth, R. Kaptein, J. J. Beintema, K. Dijkstra, N. van Nuland, U. M. Soedjanaatmadja, A. Rice, J. F. Vliegthart, C. S. Wright, H. J. Gabius, *Proteins* **1997**, *28*, 268.
- [46] E. A. Chavelas, A. P. Beltrán, G. Pérez-hernández, E. García-hernández, *J Mex Chem Soc* **2004**, *48*, 279.
- [47] R. Arreguín-Espinosa, B. Fenton, E. Vázquez-Contreras, B. Arreguín, E. García-Hernández, *Arch Biochem Biophys* **2001**, *394*, 151.
- [48] A. Kumar, R. R. Ernst, K. Wüthrich, *Biochem Biophys Res Commun* **1980**, *95*, 1.
- [49] A. Velazquez-Campoy, S. A. Leavitt, E. Freire, Characterization of protein-protein interactions by isothermal titration calorimetry. In *Protein-Protein Interactions*, Vol. 261, Humana Press, Totowa, NJ **2004**, p. 035.
- [50] N. O. Pulido, E. A. Chavelas, F. Turner, E. García-Hernández, Current applications of isothermal titration calorimetry to the study of protein complexes. In *Advances in Protein Physical Chemistry* (Eds. E. García-Hernández, D. A. Fernández-Velasco), Transworld Research Network, India, 2008, 115.
- [51] B. Fenton-Navarro, R. Arreguín-Espinosa, E. Vázquez-Contreras, B. Arreguín-Lozano, N. Sánchez-Sánchez, E. García-Hernández, E. Zenteno-Galindod, *Indian J Biochem Biophys* **2013**, *50*, 562.
- [52] T. Wiseman, S. Williston, J. F. Brandts, L. N. Lin, *Anal Biochem* **1989**, *179*, 131.
- [53] E. García-Hernández, A. Hernández-Arana, *Protein Sci* **1999**, *8*, 1075.
- [54] T. K. Dam, C. F. Brewer, *Chem Rev* **2002**, *102*, 387.
- [55] E. García-Hernández, R. A. Zubillaga, E. A. Chavelas-Adame, E. Vázquez-Contreras, A. Rojo-Domínguez, M. Costas, *Protein Sci* **2003**, *12*, 135.
- [56] M. Fiore, N. Berthet, A. Marra, E. Gillon, P. Dumy, A. Dondoni, A. Imberty, O. Renaudet, *Org Biomol Chem* **2013**, *11*, 7113.
- [57] P. Rohse, V. Wittmann, *Chem - A Eur J* **2016**, *22*, 9724.
- [58] G. Bains, R. T. Lee, Y. C. Lee, E. Freire, *Biochemistry* **1992**, *31*, 12624.
- [59] S. Martínez-Caballero, P. Cano-Sánchez, I. Mares-Mejía, A. G. Díaz-Sánchez, M. L. Macías-Rubalcava, J. A. Hermoso, A. Rodríguez-Romero, *FEBS J* **2014**, *281*, 4535.
- [60] A. L. Saucedo, F. del Rio-Portilla, C. Picco, G. Estrada, G. Prestipino, L. D. Possani, M. Delepierre, G. Corzo, *Biochim Biophys Acta Proteins Proteom* **2012**, *1824*, 478.
- [61] B. Ramírez-Cordero, Y. Toledano, P. Cano-Sánchez, R. Hernández-López, D. Flores-Solis, A. L. Saucedo-Yáñez, I. Chávez-Urbe, L. G. Briebe, *F Chem Res Toxicol* **2014**, *27*, 960.
- [62] D. Flores-Solis, Y. Toledano, O. Rodríguez-Lima, P. Cano-Sánchez, B. E. Ramírez-Cordero, A. Landa, R. C. Rodríguez de la Vega, *FEBS Lett* **2016**, *590*, 2286.
- [63] Y. Itakura, S. Nakamura-Tsuruta, J. Kominami, H. Tateno, J. Hirabayashi, *Int J Mol Sci* **2017**, *18*, 1160.
- [64] Y. Nagata, M. M. Burger, *J Biol Chem* **1974**, *249*, 3116.
- [65] R. H. Rice, M. E. Etzler, *Biochem Biophys Res Commun* **1974**, *59*, 414.
- [66] R. A. Laskowski, M. B. Swindells, *J Chem Inf Model* **2011**, *51*, 2778.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Leyva E, Medrano-Cerano JL, Cano-Sánchez P, et al. Bacterial expression, purification and biophysical characterization of wheat germ agglutinin and its four hevein-like domains. *Biopolymers*. 2018;e23242. <https://doi.org/10.1002/bip.23242>