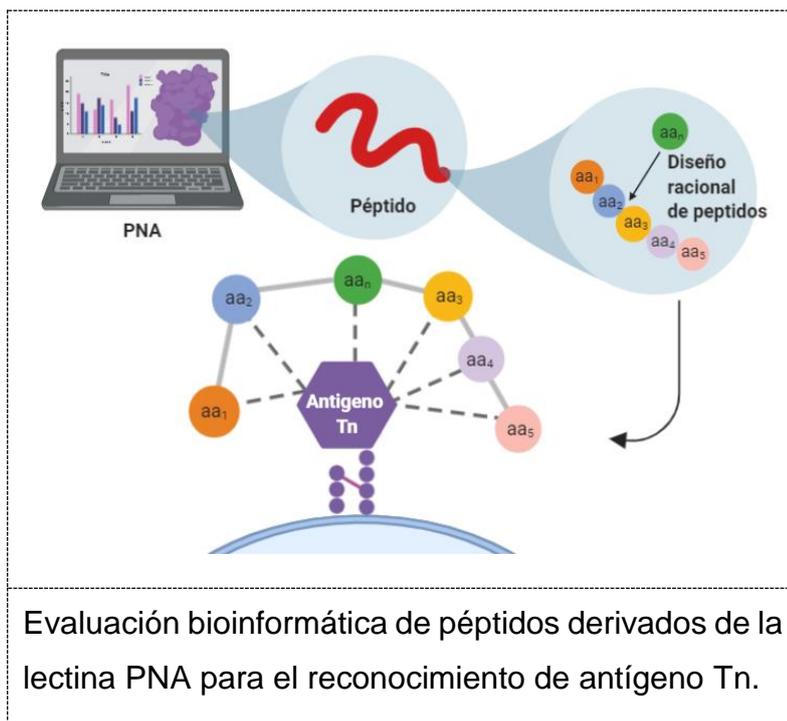


Diseño de secuencias peptídicas derivadas de la lectina de *Arachis Hypogaea* para el reconocimiento del antígeno Tn.

Design of peptide sequences derived from *Arachis Hypogaea* lectin for Tn antigen recognition



Diseño de secuencias peptídicas derivadas de la lectina de *Arachis Hypogaea* para el reconocimiento al antígeno Tn.

Design of peptide sequences derived from *Arachis Hypogaea* lectin for Tn antigen recognition

Edith Vanessa Silva Yate*

** Programa de Bioquímica. Facultad de Ciencias. Universidad Antonio Nariño.*

Dirigido por: MSc Germán Antonio García Contreras

Resumen

Los métodos de diagnóstico del cáncer actualmente se basan en biomarcadores que se encuentran glucosilados como lo es el antígeno Tn y pueden ser detectados por tecnologías basadas en lectinas como son los biosensores. Sin embargo, hay una limitación por la falta de disponibilidad comercial de lectinas que reconocen estas estructuras, además de que puede presentarse baja estabilidad en los biosensores por las capas de proteínas asociadas a este. En el presente se trabajó se analizó la lectina de maní (*Arachis Hypogaea*), en conjunto de otras lectinas de leguminosas pertenecientes la familia *Fabaceae* como *Glycine max* (PDB 4D69), *Wisteria Floribunda* (PDB 5KXB), *Psophocarpus Tetragonolobus* (PDB 2D3S), e *Isolectina B4 Vicia Villosa* (1N47); en donde basándose en sus características se diseñaron y modelaron péptidos los cuales fueron analizados por medio de pruebas de docking molecular. Se obtuvieron doce péptidos candidatos para la librería los cuales, once provienen de la segmentación estructural de nuestra proteína de interés *Arachis Hypogaea* y uno obtenido en las comparaciones y alineamientos de todas

las proteínas fabáceas. Los péptidos analizados mediante docking molecular se consideraron estables, de longitud adecuada, con propiedades energéticas similares a la lectina de maní cuando se encuentra acoplado con el antígeno Tn, considerándose así unos excelentes candidatos para que posteriormente puedan ser sintetizados y analizados para ser posibles métodos de diagnóstico de cáncer.

Abstract

Cancer diagnostic methods are currently based on glycosylated biomarkers such as Tn antigen and can be detected by lectin-based technologies such as biosensors. However, there is a limitation due to the lack of commercial availability of lectins that recognize these structures, in addition to the fact that low stability may occur in biosensors due to the layers of proteins associated with them. In the present work we analyzed peanut lectin (*Arachis Hypogaea*), together with other lectins from legumes belonging to the Fabaceae family such as *Glycine max* (PDB 4D69), *Wisteria Floribunda* (PDB 5KXB), *Psophocarpus Tetragonolobus* (PDB 2D3S), and *Isolectin B4 Vicia Villosa* (1N47); where, based on their characteristics, peptides were designed and modeled and analyzed by means of molecular docking tests. Twelve candidate peptides were obtained for the library, eleven of which come from the structural segmentation of our protein of interest *Arachis Hypogaea* and one obtained from the comparisons and alignments of all the fabaceous proteins. The peptides analyzed by molecular docking were considered stable, of adequate length, with energy properties similar to peanut lectin when coupled with the antigen Tn, being considered excellent candidates to be later synthesized and analyzed for possible cancer diagnostic methods.

Keywords: Antígeno Tn, PNA, mucinas, glicosilación, cáncer, acetilgalactosamina

Introducción

Las lectinas son proteínas que se encuentran en una amplia variedad de organismos como hongos, animales, bacterias, virus y plantas; las lectinas se caracterizan por la capacidad de unirse a carbohidratos de forma reversible ya que poseen en su estructura dominios no catalíticos específicos para el reconocimiento de monosacáridos y oligosacáridos [1], [2].

La clasificación de las proteínas se fundamenta en el organismo donde se encuentre, actividad biológica que posea, afinidad que esta tenga por un determinado sustrato [3], o por su estructura dependiendo de su número de dominios y sus características. Las lectinas mayoritariamente estudiadas son las vegetales y se dividen en cuatro clases de proteínas según su estructura: merolectinas, hololectinas, quimerolectinas y las superlectinas[4], además presentan un peso estimado que varía entre 60 a 400 kDa por la heterogeneidad que se presenta dentro del grupo de glicoproteínas [5].

Del grupo de las lectinas vegetales, las provenientes de legumbres han sido ampliamente estudiadas y se ha identificado que a pesar de que estas glicoproteínas pueden ser obtenidas de cualquier parte de la planta, se encuentran mayoritariamente en las semillas y pueden constituir hasta el 15% de las proteínas totales de estas [5],[6]. Las lectinas son utilizadas en las plantas como mecanismo de defensa para virus, bacterias y hongos, además de mediar procesos celulares como el reconocimiento célula-célula y la división celular [7],[8]. Por lo tanto, se han realizado investigaciones que han considerado a las lectinas vegetales como una herramienta clave en el diagnóstico y tratamiento del cáncer; debido a que las lectinas tienen afinidad por glicosilaciones asociadas a células potencialmente cancerosas y son considerados antígenos, los cuales son necesarios en los componentes de la membrana por que pueden actuar como mecanismo regulador de procesos fisiopatológicos y metástasis [9], [10].

Los antígenos más comunes en células cancerígenas solo contienen de uno a tres carbohidratos unidos a proteínas de membrana, sin embargo, el precursor de todos ellos se da mediante la formación del antígeno Tn, el cual consiste en solo la unión de GalNAc a residuos de Serina y/o Treonina[11]. El antígeno Tn es el resultado de la glicosilación incompleta y elongación truncada caracterizándose, así como una de las estructuras específicas asociadas al cáncer ya que a partir de él se pueden formar otros antígenos asociados a células cancerosas como el antígeno sTn, T, sT, que se diferencian estructuralmente por presentar galactosa y ácido N-acetilneurámico, respectivamente [12].

Los péptidos han sido identificados como dianas de interés farmacológico, debido a que se han encontrado péptidos tanto naturales como sintéticos que cumplen funciones de neurotransmisores, factores de crecimiento, hormonas, ligandos, entre otras funciones [13]. Durante los últimos 30 años se han desarrollado numerosas investigaciones relacionadas con el desarrollo farmacéutico de péptidos para el diagnóstico y tratamiento de enfermedades debido a la gran versatilidad que presentan estas moléculas peptídicas en el contexto clínico, considerándose de esta forma seguros, eficaces y tolerables por los seres humanos [13], [14].

El diseño racional de péptidos ha sido una herramienta ampliamente utilizada para realizar estudios a partir de proteínas de interés para obtener moléculas peptídicas derivadas de menor peso molecular y menor complejidad estructural manteniendo y/o potenciando las propiedades específicas de la proteína precursora para el desarrollo de nuevas herramientas terapéuticas, ahorrando tiempo de síntesis y análisis en laboratorio, además de permitir la evaluación de una gran cantidad de moléculas candidatas con una finalidad específica [13], [15]. Por lo cual, se ha sugerido que el diseño de péptidos derivados de

proteínas vegetales pueden desempeñar un papel importante en el diagnóstico y/o detección de glicosilaciones asociadas a células cancerosas [16],[17]. Algunos trabajos han demostrado que péptidos orientados al diagnóstico y tratamiento de cáncer resultan menos inmunogénicos que anticuerpos monoclonales recombinantes usados generalmente para atacar células tumorales con drogas anticancerígenas, siendo menos tóxicos y produciendo menor cantidad de efectos secundarios en comparación con los agentes quimioterapéuticos actuales [18].

Por lo anterior, en el presente trabajo, se estudió la lectina de *Arachis hypogaea* (PNA), extraída del maní (FIGURA 1), que también suele llamarse proteína aglutinina de maní[19]. Esta proteína se caracteriza por presentar cuatro cadenas polipeptídicas de 236 aminoácidos, con estructura homotetramérica, que posee una especificidad para el disacárido Gal β 1-3GalNAc y GalNAc asociados a un factor antigénico de tumores malignos, otra de sus características es su estructura no glicosilada de plegamiento de tipo abierto, que le proporciona mejor estabilidad respecto a las otras lectinas tetraméricas registradas con características estructurales similares [20],[21].



Figura 1. La estructura general de PNA donde se evidencian las cuatro subunidades en colores diferentes.(PDB 1CIW)[20].

La lectina PNA ha sido estudiada en complejo con monosacáridos o carbohidratos encontrando que reconoce glicosilaciones, además de tener una buena estabilidad en comparación con otras proteínas que también reconocen estos carbohidratos [22]. Por otro lado, la lectina de maní se ha utilizado como proteína específica para identificar carcinomas desde hace más de cuatro décadas cuando descubrieron que la PNA reacciona con células de carcinoma embrionario, pero no con sus derivados diferenciados sanos [23], y en estudios, donde se utilizaron en conjunto con otras lectinas para probar su capacidad de reconocimiento hacia células de glaucoma por las características específicas encontradas en la membrana de estas, además de ser probado para reconocer biomarcadores de células cancerosas provenientes de cáncer de seno, siendo eficiente en cada uno de los estudios [24],[25].

La lectina de maní también fue utilizada para diseñar nanoesferas fluorescentes inmovilizadas con PNA como un nuevo agente de imagenología y obtuvieron que las esferas reconocieron tumores de tamaño milimétrico en la mucosa del intestino grueso [28], con el valor agregado asociado a la baja toxicidad de las nanoesferas ofreciendo el desarrollo de un nuevo un agente de imagenología endoscópica [26].

Desde hace algunos años, que se han estudiado diferentes lectinas vegetales para apoyar los métodos de diagnóstico de cáncer y asociarse a biomarcadores glucosilados como lo han hecho distintos investigadores[27-34]. Sin embargo, hasta el momento hay una limitación por la falta de disponibilidad comercial de lectinas y proteínas aglutinantes que reconocen estas estructuras glúcidas o glucoproteínas de membrana, por sus limitaciones al momento de la extracción, baja estabilidad, y alto costo de producción [35].

Actualmente se usan biosensores basados en lectinas como los microarrays para la detección de glicanos aberrantes asociados al cáncer para ayudar al diagnóstico, pronóstico y evaluación del tratamiento de estos pacientes, pero estos presentan algunas dificultades como la obtención y extracción de proteínas con las cuales se trabaja; hallar un agente bloqueador que permite una buena cobertura de la superficie del microarray sin causar un alto ruido de fondo debido a la presencia de glicanos contaminantes, y por último la generación de capas en la superficie del arreglo que genera que sean menos compactas y menos estables, lo que puede afectar a toda la estabilidad del biosensor [36], [37].

Además, las glicosilaciones celulares son variables y complejas, y por la técnica de microarrays de lectinas que es la más común, no se producen bases poliméricas las cuales cuenten con una superficie con diferentes tipos de lectinas inmovilizadas para lograr un análisis más diferenciado y característico de glicanos debido a la complejidad de sus estructuras, y tampoco se cuenta con una amplia variedad de lectinas con alta afinidad a distintas moléculas consideradas antígenos [38], [39].

La detección temprana de cáncer ha demostrado que contribuye a que el tratamiento acorde a el padecimiento sea eficiente, ya que permite que la probabilidad de vida aumente hasta 90% en los diferentes tipos de cáncer en etapas iniciales, también la morbilidad se reduce y el tratamiento es más económico [40]; sin embargo, la falta de estrategias de diagnóstico en etapas iniciales y tratamiento oportuno son problemas frecuentes en lo que concierne al cáncer (Figura 2) [39]. Desde el 2017, la OMS ha reportado que solo el 26% de los países con índice de desarrollo humano (IDH) bajo contaba con servicios de patología para atender y diagnosticar a la población en general y menos del 30% de estos países ofrecen tratamiento a enfermos oncológicos a comparación de los países con un

IDH alto en donde en más del 90% de los países tienen la capacidad de tratar a pacientes oncológicos, aunque sean tratamientos costosos y netamente paliativos [41].

¿Cómo y cuando se diagnostican los pacientes con cáncer ?

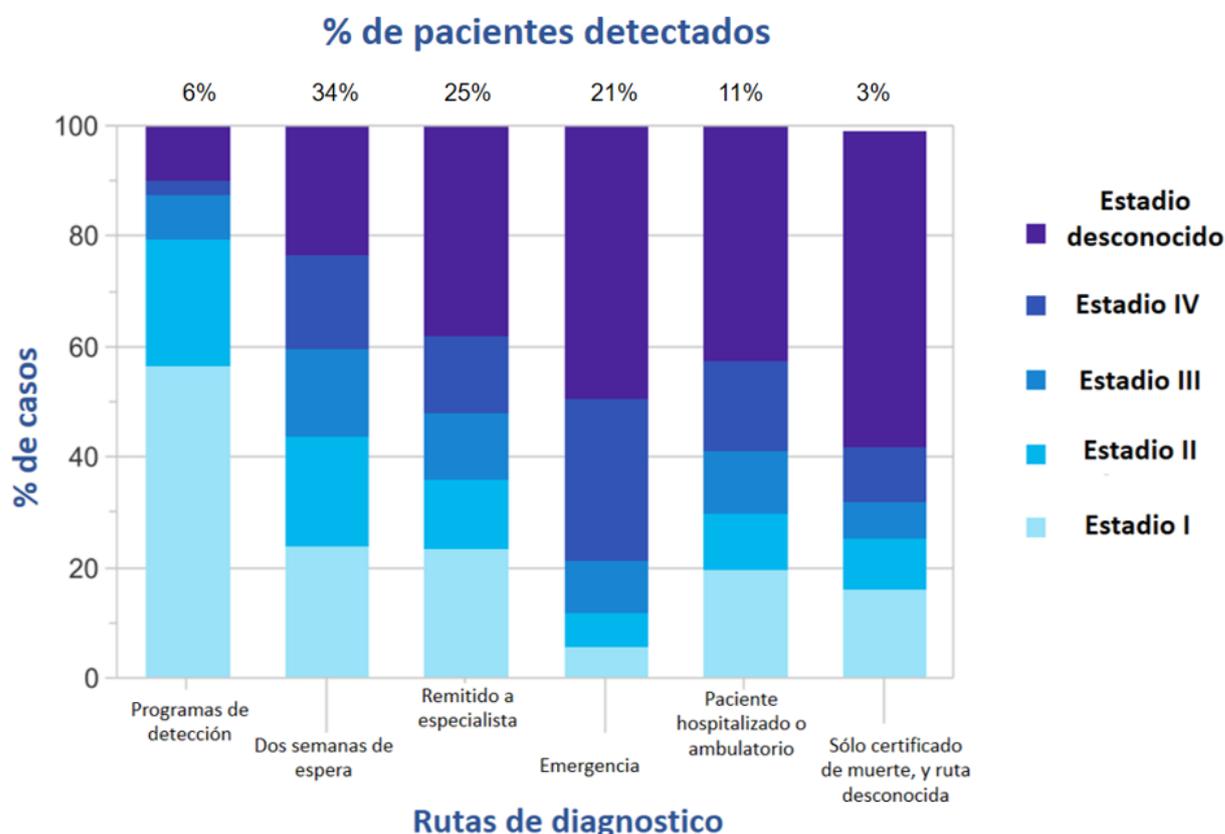


Figura 2. Porcentaje de casos de todos los tipos de cáncer por diagnóstico en adultos de 15 a 99 años, Inglaterra, 2012-2013. Adaptada de Cancer Research UK.[42]

Por lo tanto, es importante investigar proteínas mediante el uso de herramientas bioinformáticas para obtener secuencias peptídicas mediante el diseño racional de péptidos e identificar moléculas bioactivas. que sean altamente estables, de bajo costo, accesibles, y que posean características fisicoquímicas apropiadas [14], [15]. La formulación *in silico* es una tecnológica cada vez más sofisticada para simplificar y acelerar el proceso de obtención de moléculas, además permite ampliar la diversidad química y aumentar el panorama de la selección de nuevas entidades moleculares con actividades específicas y potentes contra objetivos con relevancia terapéutica [43].

Resumiendo lo anteriormente planteado, obtener péptidos mediante las anteriores estrategias, permite que estos puedan apoyar métodos que sirvan para la detección temprana del cáncer [44], y de esta manera ayuden a disminuir la morbilidad y mortalidad por cáncer, el cual es de 198 personas por cada 100 000 hombres y mujeres por año (con base en datos de muertes desde el 2018) [45], [46].

El diseño racional de péptidos sintéticos derivados de la lectina de *Arachis hypogaea* busca generar una librería peptídica con posibles candidatos que reconozcan el antígeno Tn como una estrategia para contribuir con el desarrollo de técnicas para el diagnóstico y/o detección de glicosilaciones asociadas a células cancerosas [36].

1. Materiales y Métodos

La siguiente metodología se baso en la realizada por Romero [47], para analizar lectinas de proteínas vegetales. Se realizaron ajustes para llevar acabo los objetivos del presente trabajo.

1. Búsqueda bibliográfica de estructuras

Se trabajó con la lectina *Arachis hypogaea* y proteínas de la familia *Fabaceae* que tuvieran estructura terciaria reportada en Protein Data Bank (PDB, <https://www.rcsb.org/>) con el antígeno Tn. Se realizaron alineamientos de secuencias multiples con BLAST (The Basic Local Alignment Search Tool, <https://blast.ncbi.nlm.nih.gov/Blast.ci>) y se obtuvieron los perfiles de hidrofobicidad, inestabilidad, punto isoeléctrico, GRAVY , entre otros a partir del uso de ExPASy.

1.1 Análisis de estructuras de proteínas a trabajar

Para el análisis de la estructura primaria, se realizaron alineamientos múltiples de la lectina *Arachis Hypogaea* y otras lectinas pertenecientes a la misma familia (*Fabaceae*) con los algoritmos de ClustalOmega, ClustalW, MAFFT y T-COFFEE para encontrar zonas conservadas en las lectinas a partir diferentes softwares para el análisis de las secuencias proteicas. Además, se realizó una segmentación de la proteína *Aglutinina de maní*, de acuerdo a los alineamientos previamente realizados. Los resultados fueron visualizados en Jalview, para sus respectivos análisis.

2. Diseño de péptidos.

La información obtenida anteriormente se usó para para identificar los aminoácidos involucrados en el reconocimiento del antígeno Tn y los aminoácidos conservados en las proteínas precursoras, esta información fue útil para el diseño de los péptidos sintéticos, los cuales deberán incluir los aminoácidos asociados a la interacción con el antígeno Tn, los aminoácidos conservados y deberán conservar una aproximación estructural similar al sitio de reconocimiento en la proteína precursora.

2.1 Creación de la librería de péptidos.

La librería de péptidos se diseñó a partir de las conservaciones identificadas en los alineamientos obtenidos de las proteínas y la segmentación de la PNA. Se realizaron modificaciones a las secuencias más promisorios para aumentar el score. Se analizaron los péptidos promisorios obtenidos según criterios específicos expuestos en el trabajo por AutoDock Tools evaluando y comparando la afinidad con la proteína precursora.

3. Modelamiento de péptidos

Los péptidos de interés fueron modelados en PEP-FOLD SERVER (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/>), donde se obtuvo la respectiva estructura para el acoplamiento péptido-antígeno, en el programa AutoDock v4.2. La visualización de los resultados se realizó mediante AutoDockTools v1.5.4 (ADT), a fin de observar el grado de similitud existente entre las conformaciones generadas y las conformaciones nativas.

4. Preparación de las moléculas para Docking molecular.

La proteína PNA, se simplificó mediante UCFS Chimera a una sola cadena de las cuatro que conforman la proteína, debido a que no se tuvieron en cuenta las cadenas B, C y D para realizar los análisis. Posteriormente a las estructuras proteicas se añadieron hidrógenos polares y cargas pertenecientes a la estructura mediante Autodock Tools, además de eliminar los hidrógenos sin carga.

El antígeno Tn, se descargó de la base de datos PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) en estructura 2D formato (Structure 84265) y fue preparada en Avogadro mediante una minimización de energía y asignación de los enlaces (ángulos y distancias) mediante la base de datos del software. Posteriormente fueron añadidos hidrógenos y cargas de la molécula mediante Autodock Tools.

El diseño del Grid Box para la proteína se realizó teniendo en cuenta la región asociada a para la interacción antígeno-proteína y los aminoácidos relacionados en esta [47]; para los péptidos propuestos se realizó un docking global y se tuvo en cuenta toda la estructura de la molécula.

4.1. Análisis de las conformaciones de los péptidos seleccionados.

Se analizarán todas las conformaciones de los péptidos seleccionados teniendo en cuenta los criterios mencionados en el punto anterior, el análisis se realiza con los programas AutoDock Tools y USFC Chimera.

2. Resultados y Discusión

2.1. Búsqueda bibliográfica de estructuras

Se realizó una búsqueda bibliográfica de lectinas pertenecientes a la misma familia de la lectina de maní (PNA) la cual corresponde a las fabaceas o *Fabaceae*, que son capaces de reconocer el antígeno Tn.

Las proteínas seleccionadas para realizar el estudio a parte de la de *Arachis Hypogaea* (PDB 1CIW) correspondieron a 4 lectinas las cuales fueron, *Glycine max* (PDB 4D69), *Wisteria Floribunda* (PDB 5KXB), *Psophocarpus Tetragonolobus* (PDB 2D3S) e *Isolectina B4 Vicia Villosa* (1N47), debido a que poseen estructuras homotetraméricas similares a la PNA, además todas se encuentran reportadas en PDB asociadas al reconocimiento del antígeno.

Tabla 1. Características fisicoquímicas de las proteínas de estudio

Proteína	Peso molecular (Da)	Punto isoeléctrico	Índice de inestabilidad	Índice hidropaticidad (GRAVY)	Índice alifático
SBA	27571,93	5,15	34,75	-0,107	89,05
WFA	26670,54	4,64	38,08	-0,216	85,39
WBL	26614,93	5,70	29,61	-0,129	82,11
VVA B4	24695,51	4,53	31,68	0,145	85,97
PNA *	25189,07	4,99	24,78	-0,038	84,15

SBA: Aglutinina de soya; WFA: Aglutinina de *Wisteria Floribunda*; WBL: Lectina de *Psophocarpus tetragonolobus*; VVA B4: Isolectina B4 de *Vicia Villosa*; PNA: Aglutinina de maní*proteína de estudio

Las características generales de las proteínas muestran que el peso molecular se encuentra entre 25.000 a 27.000 Da y poseen carga neta de cero en pH 4 y 6 considerándose así una homogeneidad entre ellas. Sus puntuaciones GRAVY por debajo de 0 indican que son posiblemente proteínas globulares y/o hidrofílicas, además, de ser termoestables debido a su alto índice alifático; se confirman su tendencia estable, debido a sus índices de inestabilidad calculados en Protparam tool (<https://web.expasy.org/protparam/>) ya que sus índices son menores a 40.

Los alineamientos por pares correspondientes a las cinco proteínas de estudio indican que la proteína de interés PNA, en la matriz de sustitución guarda mayor coincidencia y conservación con la lectina VVA B4, como se muestra en la Tabla 2.

Tabla 2. Alineamiento de proteínas en Clustal Omega

<i>Proteína</i>	<i>SBA</i>	<i>WFA</i>	<i>WBL</i>	<i>VVA B4</i>	<i>PNA</i>
<i>SBA</i>	100,00	46,72	42,36	40,69	40,360
<i>WFA</i>	46,72	100,00	48,73	48,10	49,57
<i>WBL</i>	42,36	48,73	100,00	59,58	53,91
<i>VVA B4</i>	40,690	48,10	59,58	100,00	59,230
<i>PNA</i>	40,36	49,57	53,91	59,230	100,00

*SBA: Aglutinina de soya; WFA: Aglutinina de Wisteria Floribunda ;WBL:Lectina de Psophocarpus tetragonolobus;;VVA B4: Isolectina B4 de Vicia Villosa;PNA: Aglutinina de maní*proteína de estudio*

2.2. Análisis de estructuras primarias de las proteínas seleccionadas

El análisis de las estructuras demostró que las cinco proteínas estudiadas tienen en común siete aminoácidos consecutivos conservados presentando una única diferencia en el aminoácido 129, en donde 4D69, 5KXB y 2D3S presentan una fenilalanina y 1N47 y 1CIW presentan una tirosina; siendo relevante la presencia de los aminoácidos Gln127 y Asn130 en la secuencia conservada los cuales se encuentran estrechamente relacionados con la interacción del antígeno con la lectina [47].

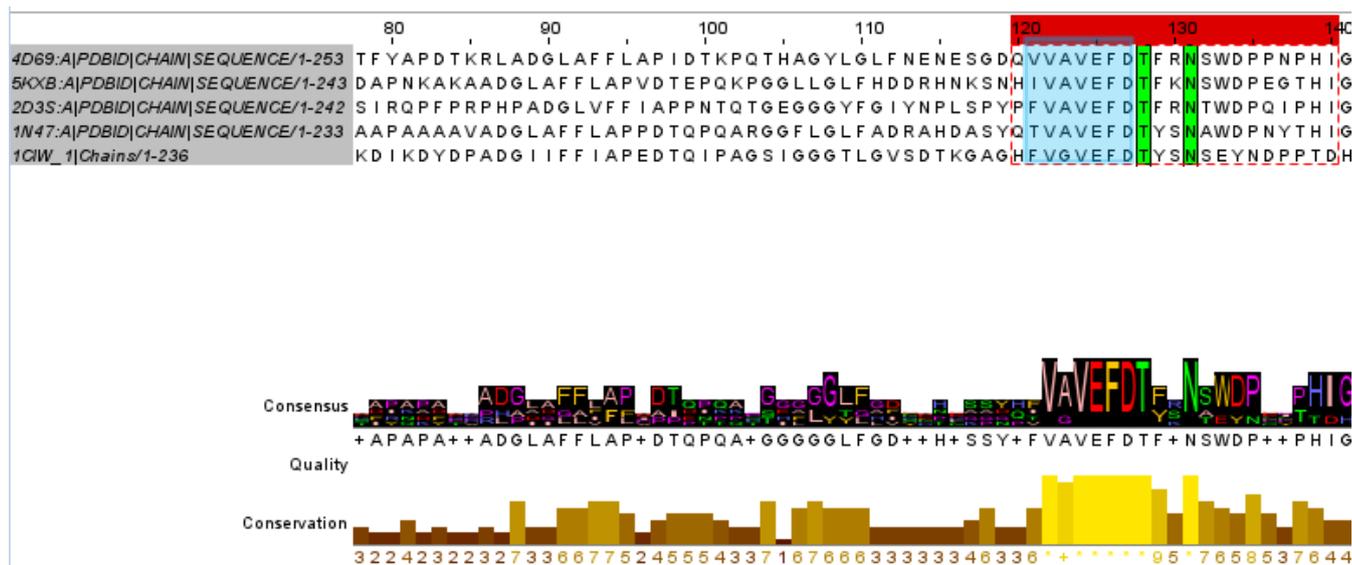


Figura 3. Alineamiento múltiple de las secuencias de las proteínas de estudio: En azul se encuentra la secuencia mayoritariamente conservada en las cinco secuencias; en verde se encuentran los residuos que interactúan con el antígeno Tn.

De acuerdo con la figura 1, las proteínas estudiadas presentan características estructurales similares asociadas a la presencia del aminoácido 123 el cual es alifático y se caracteriza por presentar baja hidrofobicidad y en la posición 129 se encuentra un aminoácido aromático. Las variaciones respectivas se presentan en la lectina PNA, sugiriendo así que estos podrían indicar la mediana afinidad por antígeno Tn, y adicional variar la estabilidad de este fragmento de la proteína.

La segmentación de PNA permitió obtener diecisiete secuencias peptídicas constituidas por trece residuos y una secuencia adicional de quince, de las cuales se descartaron cuatro debido a sus altos índices de inestabilidad que predicen la baja estabilidad de los péptidos [48].

Tabla 3. Características fisicoquímicas de los peptidos obtenidos de la segmentación de PNA

Código	Secuencia peptídica	Aminoácidos	Peso molecular (Da)	Punto isoeléctrico	Índice de inestabilidad	Índice Alifático	Índice hidropaticidad (GRAVY)
EVSY1	AETVSFNFSFSE	13	1478,64	3,80	12,48 (E)	30,00	-0,208
EVSY2	GNPAINFQGDVTV	13	1331,45	3,80	-2,78 (E)	82,31	0,031
EVSY3	LSNGNIQLTNLNK	13	1428,61	8,75	31,59 (E)	120,00	-0,569
EVSY4*	VNSVGRVLYAMPV	13	1404,69	8,72	69,53 (I)	126,92	0,938
EVSY5	RIWSSATGNVASF	13	1395,54	9,75	10,98 (E)	67,69	0,208
EVSY6	LTSFSFEMKDIKD	13	1560,78	4,56	2,70 (E)	60,00	-0,369
EVSY7*	YDPADGIIFFIAP	13	1438,64	3,56	44,02 (I)	105,38	0,831
EVSY8	EDTQIPAGSIGGG	13	1201,26	3,67	38,27 (E)	67,69	-0,338
EVSY9	TLGVSDTKGAGHF	13	1289,41	6,41	-23,37 (E)	60,00	-0,108
EVSY10	VGVEFDTYSNSEY	13	1509,55	3,57	15,45 (E)	44,62	-0,623
EVSY11	NDPPTDHVGIDVN	13	1392,44	3,93	10,98 (E)	74,62	-0,931
EVSY12	SVDSVKTPWNSV	13	1417,58	5,55	32,36 (E)	89,23	0,023
EVSY13	SGAVVKVTVIYDS	13	1337,54	5,55	20,45 (E)	126,92	0,900
EVSY14	STKTLVAVTNDN	13	1349,46	5,55	-8,86 (E)	82,31	-0,315
EVSY15	GDITIAQVVDLK	13	1372,58	4,21	-14,66 (E)	142,31	0,523
EVSY16	AKLPERVKFGFSA	13	1449,72	9,99	28,67 (E)	67,69	-0,108
EVSY17*	SGSLGGRQIHLIR	13	1393,61	12,00	43,03 (I)	120,00	-0,146
EVSY18*	SWSFTSTLITTRRS	15	1743,94	12,00	76,81(I)	52	-0,367

Se encuentran resaltados los peptidos con las mejores estabilidades y (*) peptidos descartados debido a sus deficientes características fisicoquímicas. Estable;(I):Inestable

Los alineamientos de pares de peptidos derivados de la proteína PNA, muestran que las secuencias peptídicas EVSY2 y EVSY8 presentan la mayor puntuación de los alineamientos realizados a todas las secuencias peptídicas propuestas.

Tabla 4. Alineamientos de péptidos obtenidos de la segmentación de PNA

Péptido	EVSY1	EVSY2	EVSY3	EVSY5	EVSY6	EVSY8	EVSY9	EVSY10	EVSY11	EVSY12	EVSY13	EVSY14	EVSY15	EVSY16
EVSY1	100,00	50,00	0,00	0,000	0,000	30,00	22,22	0,000	14,29	0,000	11,11	0,000	0,000	22,22
EVSY2	50,00	100,00	25,00	0,00	16,67	60,00	22,22	0,00	16,67	0,000	0,000	0,000	11,11	25,00
EVSY3	0,000	25,00	100,00	20,00	37,50	-nan	25,00	0,00	11,11	23,08	20,00	18,18	16,67	-nan
EVSY5	0,000	0,00	20,00	100,00	27,27	0,000	0,000	33,33	9,090	16,67	25,00	8,330	9,090	0,000
EVSY6	0,000	16,67	37,50	27,27	100,00	20,00	11,11	12,50	8,330	20,00	10,00	10,00	33,33	0,000
EVSY8	30,00	60,00	-nan	0,000	20,00	100,00	33,33	10,00	0,000	0,000	0,000	0,000	0,000	16,67
EVSY9	22,22	22,22	25,00	0,000	11,11	33,33	100,00	33,33	0,000	0,000	0,000	0,000	0,000	12,5
EVSY10	0,000	0,000	0,000	33,33	12,50	10,00	33,33	100,00	0,000	20,00	0,000	0,000	0,000	0,000
EVSY11	14,29	16,67	11,11	9,090	8,330	0,000	0,000	0,000	100,00	36,36	11,11	27,27	10,00	0,000
EVSY12	0,000	0,000	23,08	16,67	20,00	0,000	0,000	20,00	36,36	100,00	14,29	30,77	15,38	0,000
EVSY13	11,11	0,000	20,00	25,00	10,00	0,000	0,000	0,000	11,11	14,29	100,00	57,14	33,33	28,57
EVSY14	0,000	0,000	18,18	8,330	10,00	0,000	0,000	0,000	27,27	30,77	57,14	100,00	25,00	0,000
EVSY15	0,000	11,11	16,67	9,090	33,33	0,000	0,000	0,000	10,00	15,38	33,33	25,00	100,00	-nan
EVSY16	22,22	25,00	-nan	0,000	0,000	16,67	12,50	0,000	0,000	0,000	28,57	0,000	-nan	100,00

Los valores mas altos de puntuacion en la matriz de sustitucion se encuentran resaltados en la tabla

2. Diseño de los péptidos y creación de librería

La tabla 3 sugiere que los péptidos que presentan índices de inestabilidad por debajo de 40 son viables para analizar la interacción entre el péptido y antígeno Tn mediante docking molecular, por esta razón se decidió seleccionar un péptido que cumpliera este parámetro y se descartaron los péptidos con Índices de inestabilidad cercanos a 40 (ESVY3, EVSY8, EVSY17 Y EVSY18), adicionalmente se descartó el péptido EVSY9.

La librería de péptidos obtenidos (Tabla 5), corresponden a diecisiete péptidos, sin embargo, el análisis se centró en estudiar péptidos con las características fisicoquímicas más prometedoras correspondientes a EVSY, EVSY14 Y EVSY 15, además de los resultantes de alineamientos y sustituciones que corresponden a EVSY20, EVSY10 y EVSY10A.

Tabla 5. Librería de péptidos propuesta del estudio de PNA y otras lectinas Fabaceas

Péptido	Secuencia	Aminoácidos
<i>EVSY1</i>	AETVSFNFNSFSE	13
<i>EVSY2</i>	GNPAINFQGDVTV	13
<i>EVSY5</i>	RIWSSATGNVASF	13
<i>EVSY6</i>	LTSFSFEMKDIKD	13
<i>EVSY10</i>	VGVEFDTYSNSEY	13
<i>EVSY11</i>	NDPPTDHVGIDVN	13
<i>EVSY12</i>	SVDSVKTPWNSV	13
<i>EVSY13</i>	SGAVVKVTVIYDS	13
<i>EVSY14</i>	STKTLSSAVTNDN	13
<i>EVSY15</i>	GDITTIAQVVDLK	13
<i>EVSY20</i>	VAVEFDTY	8
<i>EVSY10A</i>	VAVEFDTYSNSEY	13

3. Modelamiento de péptidos

Para evaluar la afinidad de los péptidos derivados hacia el antígeno, se realizó docking molecular global entre el péptido y el antígeno. La metodología utilizada fue la misma que se usó para realizar el modelamiento proteína-antígeno.

Los resultados de los modelamientos moleculares entre los péptidos-antígeno sugieren que el péptido *EVSY10*, presenta afinidad por el antígeno Tn, se realizó una modificación estructural, sustituyendo la Gly2 por Ala, para obtener la secuencia *EVSY10A*. El modelamiento molecular realizado a esta secuencia peptídica sugiere que la secuencia *EVSY10*, puede tener otras modificaciones estructurales que pueden mejorar la afinidad del péptido por el antígeno Tn.

Tabla 6. Propiedades energéticas de los péptidos acoplados al antígeno Tn

<i>Peptidos</i>			
<i>Propiedades</i>	<i>EVSY15</i>	<i>EVSY10</i>	<i>EVSY10A</i>
<i>Energia de enlace</i>	-2,31	-2,95	-3,22
<i>Eficiencia de ligando</i>	0,15	0,20	0,21

El modelamiento molecular de la proteína y el péptido, muestra que existen diez conformaciones posibles para el acoplamiento del péptido y la proteína, siendo la conformación 1 (*Conf. 1*) la mas viable.

Tabla 7. Conformaciones posibles Proteína PNA-antígeno Tn (AutoDock Tools, ADT)

<i>Conformaciones</i>										
<i>Propiedades</i>	<i>Conf.1</i>	<i>Conf.2</i>	<i>Conf.3</i>	<i>Conf.4</i>	<i>Conf.5</i>	<i>Conf.6</i>	<i>Conf.7</i>	<i>Conf.8</i>	<i>Conf.9</i>	<i>Conf.10</i>
<i>Energia de enlace</i>	-3,71	-3,50	-3,37	-3,34	-3,06	-2,96	-2,79	-2,65	-2,62	-2,56
<i>Eficiencia de ligando</i>	0,25	0,23	0,22	0,22	0,20	0,20	0,19	0,18	0,17	0,17

3. Conclusiones

Los resultados del docking molecular muestran que la secuencia EVSY10 presenta una afinidad hacia el antígeno Tn similar a la proteína precursora lo cual lo hace una secuencia promisoría para el diseño de estrategias para la detección de antígeno Tn.

El péptido EVSY10A debido a sus resultados de modelamiento molecular sugieren que hay modificaciones que pueden mejorar la afinidad de la secuencia peptídica EVSY10, por lo

cual se considera continuar el estudio de esta secuencia peptídica, a través de una librería basada en de este péptido para obtener moléculas peptídicas probablemente más afines con el antígeno.

Las herramientas bioinformáticas demuestran que ofrecen la posibilidad de analizar interacciones de interés bioquímico evitando los costos innecesarios de reactivos, tiempo y recursos consumibles, ya que se evidencio que los péptidos que se consideraban con mejores características fisicoquímicas no tuvieron resultados de interacción iguales o superiores con el antígeno Tn comparado con los resultados de interacción de la proteína.

Por lo tanto, este trabajo contribuye a sustentar que péptidos derivados de proteínas que reconocen el antígeno Tn pueden ser utilizados para este mismo fin; los péptidos al ser moléculas más pequeñas, son más fáciles de obtener, purificar y/o sintetizar que las proteínas.

Como recomendación se sugiere continuar con el modelamiento molecular de los péptidos con mejores características fisicoquímicas presentados en este trabajo.

Agradecimientos

Quiero agradecer en especial mis padres, mis hermanos y pareja que con su esfuerzo y dedicación me ayudaron a culminar mi carrera universitaria y me dieron el apoyo para no decaer. Asimismo, agradezco infinitamente a mis profesores y compañeros, que me animaron a llenarme de sabiduría y nuevas experiencias; y en especial al profesor German Antonio García Contreras por los conocimientos transmitidos y su paciencia que permitieron llevar a cabo este proyecto.

Referencias

- [1] R. S. Singh, A. K. Tiwary, and J. F. Kennedy, "Lectins: Sources, activities, and applications," *Crit. Rev. Biotechnol.*, vol. 19, no. 2, pp. 145–178, 1999.
- [2] R. Hamid, A. Masood, I. H. Wani, and S. Rafiq, "Lectins: Proteins with diverse applications," *J. Appl. Pharm. Sci.*, vol. 3, no. 4SUPPL.1, 2013.
- [3] A. S. W. Ang, R. C. F. Cheung, X. Dan, Y. S. Chan, W. Pan, and T. B. Ng, "Purification and characterization of a glucosamine-binding antifungal lectin from phaseolus vulgaris cv. Chinese Pinto Beans with antiproliferative activity towards nasopharyngeal carcinoma cells," *Appl. Biochem. Biotechnol.*, vol. 172, no. 2, pp. 672–686, 2014.
- [4] I. Lagarda-Diaz, A. Guzman-Partida, and L. Vazquez-Moreno, "Legume Lectins: Proteins with Diverse Applications," *Int. J. Mol. Sci.*, vol. 18, no. 6, p. 1242, Jun. 2017.
- [5] L. Da Silva and M. Correia, "Plant lectins and Toll-like receptors: implications for therapy of microbial infections," *Front. Microbiol.*, vol. 5, no. 1, 2014.
- [6] E. Hamed El S, M. M. Ibrahim El A, and M. Mounir S, "Antimicrobial Activities of Lectins Extracted from Some Cultivars of Phaseolus vulgaris Seeds," *J. Microb. Biochem. Technol.*, vol. 09, no. 03, pp. 109–116, 2017.
- [7] N. Lannoo and E. J. M. Van Damme, "Lectin domains at the frontiers of plant defense," *Front. Plant Sci.*, vol. 5, no. August, pp. 1–16, 2014.
- [8] X. Dan, W. Liu, and T. B. Ng, "Development and Applications of Lectins as Biological Tools in Biomedical Research," *Med. Res. Rev.*, vol. 36, no. 2, pp. 221–247, Mar. 2016.
- [9] R. Graham, R. Beatson, V. Tajadura-Ortega, J. Taylor-Papadimitriou, and J. M. Burchell, "O-linked mucin-type glycosylation in breast cancer," *Biochem. Soc. Trans.*, vol. 46, no. 4, pp. 779–788, 2018.
- [10] S. S. Pinho and C. A. Reis, "Glycosylation in cancer: mechanisms and clinical implications : Nature Reviews Cancer : Nature Publishing Group," *Nat. Rev. Cancer*, vol. 15, pp. 540–555, 2015.
- [11] A. Medeiros, N. Berois, M. Incerti, S. Bay, L. Franco Fraguas, and E. Osinaga, "A Tn antigen binding lectin from Myrsine coriacea displays toxicity in human cancer cell lines," *J. Nat. Med.*, vol. 67, no. 2, pp. 247–254, 2013.
- [12] G. Poiroux, A. Barre, E. van Damme, H. Benoist, and P. Rougé, "Plant Lectins Targeting O-Glycans at the Cell Surface as Tools for Cancer Diagnosis, Prognosis and Therapy," *Int. J. Mol. Sci.*, vol. 18, no. 6, p. 1232, Jun. 2017.
- [13] K. Fosgerau and T. Hoffmann, "Peptide therapeutics: current status and future directions.," *Drug Discov. Today*, vol. 20, no. 1, pp. 122–8, 2015.
- [14] M. Erak, K. Bellmann-Sickert, S. Els-Heindl, and A. G. Beck-Sickinger, "Peptide chemistry toolbox – Transforming natural peptides into peptide therapeutics," *Bioorganic Med. Chem.*, vol. 26, no. 10, pp. 2759–2765, 2018.
- [15] L. Ning, B. He, P. Zhou, R. Derda, and J. Huang, "Molecular Design of Peptide-Fc Fusion Drugs," *Curr. Drug Metab.*, vol. 20, no. 3, pp. 203–208, Aug. 2018.
- [16] A. Tyagi, P. Kapoor, R. Kumar, K. Chaudhary, A. Gautam, and G. P. S. Raghava, "In Silico Models for Designing and Discovering Novel Anticancer Peptides," *Sci. Rep.*, vol. 3, p. 2984, Oct. 2013.
- [17] K. Valko, G. Ivanova-Berndt, P. Beswick, M. Kindey, and D. Ko, "Application of biomimetic HPLC to estimate lipophilicity, protein and phospholipid binding of potential peptide therapeutics," *ADMET DMPK*, vol. 6, no. 2, pp. 162–175, 2018.
- [18] G. Guidotti, L. Brambilla, and D. Rossi, "Exploring Novel Molecular Targets for the Treatment of High-Grade Astrocytomas Using Peptide Therapeutics: An Overview.," *Cells*, vol. 9, no. 2, 2020.
- [19] J. Sun, Q. L. Yang, J. Bi, C. S. Zhang, L. N. Yu, and F. Zhu, "Purification and Characterization of a Natural Lectin from the Seed of Peanut *Arachis hypogaea*," *Adv. Mater. Res.*, vol. 152–153, pp. 1499–1504, 2010.
- [20] R. Banerjee, K. Das, R. Ravishankar, K. Suguna, A. Surolia, and M. Vijayan, "Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex," *J. Mol. Biol.*, vol. 259, no. 2, pp. 281–296, 1996.
- [21] P. T. Campana, L. R. S. Barbosa, and R. Itri, "Conformational stability of peanut agglutinin using small angle X-ray scattering," *Int. J. Biol. Macromol.*, vol. 48, no. 3, pp. 398–402, 2011.
- [22] A. Movafagh *et al.*, "The structure Biology and Application of Phytohemagglutinin (PHA) in Phytomedicine: With special up-to-date references to lectins," *J. Paramed. Sci. Vol*, vol. 4, no. Winter, pp. 2008–4978, 2013.
- [23] Y. Reisner, G. Gachelin, P. Dubois, J. F. Nicolas, N. Sharon, and F. Jacob, "Interaction of peanut agglutinin, a lectin specific for nonreducing terminal d-galactosyl residues, with embryonal carcinoma cells," *Dev. Biol.*, vol. 61, no. 1, pp. 20–27, 1977.
- [24] H. Tateno *et al.*, "Glycome diagnosis of human induced pluripotent stem cells using lectin microarray," *J. Biol. Chem.*, vol. 286, no. 23, pp. 20345–20353, 2011.
- [25] P. M. Drake *et al.*, "Lectin chromatography/mass spectrometry discovery workflow identifies putative biomarkers of aggressive breast cancers," *J. Proteome Res.*, vol. 11, no. 4, pp. 2508–2520, 2012.
- [26] S. Sakuma *et al.*, "Detection of early colorectal cancer imaged with peanut agglutinin-immobilized fluorescent nanospheres having surface poly(N-vinylacetamide) chains," *Eur. J. Pharm. Biopharm.*, vol. 74, no. 3, pp. 451–460, 2010.

- [27] R. Singh, L. Nawale, D. Sarkar, and C. G. Suresh, "Two chitotriose-specific lectins show anti-angiogenesis, induces caspase-9-mediated apoptosis and early arrest of pancreatic tumor cell cycle," *PLoS One*, vol. 11, no. 1, pp. 1–18, 2016.
- [28] T. Ju *et al.*, "Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers," *Proteomics Clin. Appl.*, vol. 7, pp. 618–631, 2013.
- [29] K. L. Bicker *et al.*, "Synthetic lectin arrays for the detection and discrimination of cancer associated glycans and cell lines," *Chem. Sci.*, vol. 3, no. 4, pp. 1147–1156, 2012.
- [30] H. A. Badr, A. I. Elsayed, H. Ahmed, M. V. Dwek, C. Z. Li, and L. B. Djansugurova, "Preferential lectin binding of cancer cells upon sialic acid treatment under nutrient deprivation," *Appl. Biochem. Biotechnol.*, vol. 171, no. 4, pp. 963–974, 2013.
- [31] D. Zupančič, M. E. Kreft, and R. Romih, "Selective binding of lectins to normal and neoplastic urothelium in rat and mouse bladder carcinogenesis models," *Protoplasma*, vol. 251, no. 1, pp. 49–59, 2014.
- [32] K. Nakajima *et al.*, "Establishment of new predictive markers for distant recurrence of colorectal cancer using lectin microarray analysis," *Cancer Med.*, vol. 4, no. 2, pp. 293–302, 2015.
- [33] K. Ikemoto, K. Shimizu, K. Ohashi, Y. Takeuchi, M. Shimizu, and N. Oku, "Bauhinia purpurea agglutinin-modified liposomes for human prostate cancer treatment," *Cancer Sci.*, vol. 107, no. 1, pp. 53–59, 2016.
- [34] N. Vega and G. Pérez, "Isolation and characterisation of a *Salvia bogotensis* seed lectin specific for the Tn antigen," *Phytochemistry*, vol. 67, no. 4, pp. 347–355, 2006.
- [35] M. Kailemia, D. Park, and C. Lebrilla, "Glycans and Glycoproteins as specific biomarkers for Cncer," *Anal. Bioanal. Chem.*, pp. 395–410, 2017.
- [36] M. L. S. Silva, *Lectin biosensors in cancer glycan biomarker detection*, 1st ed., vol. 93. Elsevier Inc., 2019.
- [37] H. Yu, J. Shu, and Z. Li, "Lectin microarrays for glycoproteomics: an overview of their use and potential," *Expert Rev. Proteomics*, vol. 17, no. 1, pp. 27–39, 2020.
- [38] D. Pihíková, P. Kasák, and J. Tkac, "Glycoprofiling of cancer biomarkers: Label-free electrochemical lectin-based biosensors," *Open Chem.*, vol. 13, no. 1, pp. 636–655, 2015.
- [39] H. A. Badr *et al.*, "Lectin approaches for glycoproteomics in FDA-approved cancer biomarkers," *Expert Rev. Proteomics*, vol. 11, no. 2, pp. 227–236, 2014.
- [40] R. Etzioni *et al.*, "The case for early detection," *Nat. Rev. Cancer*, vol. 3, no. 4, pp. 243–252, 2003.
- [41] "Cáncer: Organizacion Mundial de la Salud." [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/cancer>. [Accessed: 26-Mar-2020].
- [42] National Cancer Intelligence Network., "Routes to diagnosis," 2016. [Online]. Available: http://www.ncin.org.uk/publications/routes_to_diagnosis. [Accessed: 26-Apr-2020].
- [43] A. Caporale *et al.*, "Synthetic Peptide Libraries. From Random Mixtures to In Vivo Testing," *Curr. Med. Chem.*, vol. 25, Sep. 2018.
- [44] D. A. Araripe *et al.*, "Partial characterization and immobilization in CNBr-activated Sepharose of a native lectin from *Platypodium elegans* seeds (PELa) and comparative study of edematogenic effect with the recombinant form," *Int. J. Biol. Macromol.*, vol. 102, pp. 323–330, 2017.
- [45] "El cáncer - Instituto Nacional del Cáncer." [Online]. Available: <https://www.cancer.gov/espanol/cancer>. [Accessed: 26-Mar-2020].
- [46] "Global Cancer Observatory." [Online]. Available: <https://gco.iarc.fr/>. [Accessed: 26-Mar-2020].
- [47] D. A. Romero Frenchy, "Diseño , modelamiento y evaluación in silico de péptidos que reconocen antígeno Tn," 2019.
- [48] E. Gasteiger *et al.*, "Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server," 2005.