

00346

8



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS

CARACTERIZACION DE LA ORGANIZACION ESTRUCTURAL
DE UNA AGLUTININA ESPECIFICA PARA GALACTOSIDOS
DEL MOLUSCO DE AGUA DULCE *Pomacea flagellata*.

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE

MAESTRA EN CIENCIAS (BIOLOGIA CELULAR)

P R E S E N T A

MARIA MARGARITA ZAYIL SALAZAR CAMPOS

DIRECTOR DE TESIS: DR. ROBERTO ARREGUIN ESPINOSA DE LOS MONTEROS

CO-DIRECTOR DE TESIS: DR. ENRIQUE GARCIA HERNANDEZ

MEXICO, D. F.

2002

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

ÍNDICE

presento a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Salazar Campos

Ma. Margarita Zayas

FECHA: 30-OCT-02

FIRMA: [Firma]

RESUMEN

2

1 INTRODUCCIÓN

3

1. LECTINAS

3

1.1 Características generales

3

1.2 Carbohidratos unidos covalentemente a las lectinas

3

1.3 Especificidad por el carbohidrato

4

2. LECTINAS ANIMALES

5

2.1 Lectinas del Tipo-C

6

2.2 Lectinas Tipo-S o Galectinas

8

2.3 Lectinas de invertebrados

10

2.4 Aglutininas de *Pomacea flagellata*

12

3. IDENTIFICACIÓN DE *Pomacea flagellata*

13

3.1 *Pomacea flagellata*

13

3.2 Huella genómica de *Pomacea flagellata*

15

4 OBJETIVO GENERAL

17

OBJETIVOS PARTICULARES

17

HIPÓTESIS

17

4.1 Cromatografía de afinidad

18

4.2 Cromatografía de intercambio iónico

19

4.3 Espectrometría de masas

23

4.4 Secuenciación de aminoácidos

28

4.5 Dicroísmo Circular

31

TESIS CON
FALLA DE ORIGEN

5 METODOLOGÍA	38
5.1 Diagrama de flujo metodológico para determinar la huella genómica de <i>Pomacea flagellata</i>	38
5.2 Diagrama de flujo metodológico para la purificación y caracterización de PFA	39
5.3 Extracción de DNA de <i>P. flagellata</i>	40
5.4 Cuantificación de DNA	40
5.5 Recuperación de DNA de alto peso molecular (APM)	41
5.6 Amplificación de DNA de APM por RAPDS	41
5.7 Extracción de PFA-I	42
5.8 Pruebas de hemaglutinación	42
5.9 Purificación de PFA	43
5.10 Espectrometría de masas	43
5.11 Secuenciación de aminoácidos	43
5.12 Caracterización estructural mediante dicroísmo circular	44
6 RESULTADOS Y DISCUSIÓN	45
6.1 Colecta	45
6.2 Extracción y recuperación de DNA	46
6.3 Amplificación con RAPDS	48
6.4 Pruebas de hemaglutinación	52
6.5 Fraccionamiento mediante columna de afinidad	52
6.6 Separación de PFA-I y PFA-II mediante intercambio iónico	57
6.7 Espectrometría de masas	58
6.8 Secuenciación de aminoácidos	58
6.9 Dicroísmo circular	61
7 CONCLUSIONES	63
8 BIBLIOGRAFÍA	65

TESIS CON
FALLA DE ORIGEN

Índice de tablas

	Pag.
Tabla 1. Lectinas específicas para grupo sanguíneo	5
Tabla 2. Clasificación de las lectinas animales.	6
Tabla 3. Lectinas del Tipo-C. Clasificación de acuerdo a sus funciones y localización	7
Tabla 4. Datos de recuperación de DNA obtenido del pie o músculo de <i>P. flagellata</i> .	46
Tabla 5. Análisis de resultados de amplificación mediante densitómetro.	51
Tabla 6. Comparación de la actividad hemaglutinante de los distintos extractos de <i>P. flagellata</i> con los 4 tipos sanguíneos.	52
Tabla 7. . Actividad aglutinante de las fracciones colectadas a partir de la columna de afinidad	54
Tabla 8. Especificidad por azúcar de PFA	56
Tabla 9. Comparación de las estructuras secundarias de PFA y RA.	62

TESIS CON
FALLA DE ORIGEN

Índice de figuras

	Pag.
Figura 1. Tipos de matrices para intercambio iónico	20
Figura 2. Principio de la cromatografía de intercambio iónico	21
Figura 3. ESI MS de una solución de proteína dispersada como gotas eléctricamente cargadas, después de pasar por una aguja sometida a un campo de alto voltaje.	24
Figura 4. Transformación computarizada del espectro generado de la medición m/z.	25
Figura 5. Tandem MS o MS/MS.	26
Figura 6. Esquema de luz polarizada a la derecha	32
Figura 7. Rotación del plano de luz polarizada	33
Figura 8. Rotación y absorción del plano de luz polarizada	33
Figura 9. Fotografía de <i>P. flagellata</i>	45
Figura 10. Gel de agarosa al 0.8% en TBE 1X. Extracción de DNA de <i>P. flagellata</i> .	47
Figura 11. Gel de acrilamida para DNA, amplificación de DNA de <i>P. flagellata</i> con los distintos primers del kit de RAPD's	48
Figura 12. Gráfica de recuperación a partir de cromatografía de afinidad, D.O. vs fracciones eludías.	53
Figura 13: Estructura química de los determinante A y B basados en las Cadenas de carbohidratos Tipo-1 y Tipo-2	55

Figura 14. Purificación de PFA-I y PFA-II mediante cromatografía de Intercambio iónico.	57
Figura 15. Secuencia parcial de aminoácidos obtenida de PFA.	58
Figura 16. Análisis de alineación de secuencias aminoácidas de PFA y RA.	60
Figura 17. Espectro de DC en U. V. obtenido de PFA.	61

ABREVIATURAS USADAS EN EL TEXTO

PFA: Aglutinina de *Pomacea flagellata* (*Pomacea flagellata* aglutinin)

PFA-I: Aglutinina de *Pomacea flagellata* tipo I

PFA-II: Aglutinina de *Pomacea flagellata* tipo II

RA: Cadena A (tóxica) de la ricina (Ricin A Chain)

RB: Cadena B de la ricina (Ricin B Chain)

DRC: Dominio o sitio de reconocimiento al carbohidato

MBP-A: Proteína tipo A que se une a la manosa (Manose Binding Protein A)

EGF: Factor de crecimiento epidérmico (Epidermis Growth Factor)

CTLD: Dominio de unión de las lectinas tipo-C (C-type lectin-like domains)

LPS: Lipopolisacáridos

BDL: Lectina de *Blaberus discoidalis* (*Blaberus discoidalis* lectin)

RAPDS: Marcadores genéticos que amplifican DNA genómico con primers de secuencia nucleotídica arbitraria que permiten detectar polimorfismos interespecie (Random Amplified Polimorphic DNA)

MALDI-MS: Espectrometría de masas (Matrix Assisted Laser Desorption/Ionization Mass Spectrometry)

ESI-MS: Espectrometría de masas mediante ionización en electroespray (Electrospray Ionization Mass Spectrometry)

DC: Dicroísmo Circular

DRO: Dispersión Rotatoria Óptica

U.V. : Luz Ultravioleta

MW: Peso molecular (Molecular Weight)

TBE: Tris, ácido bórico y EDTA

TAE: Tris, ácido acético y EDTA

DNA: Ácido Desoxirribonucleico

RNA: Ácido Ribonucleico

APM: Alto Peso Molecular

TBS: Tris, Borato de Sodio

HPLC: Cromatografía líquida de Alta Resolución (High Performance Liquid Chromatography)

**CARACTERIZACIÓN DE LA ORGANIZACIÓN ESTRUCTURAL
DE UNA AGLUTININA ESPECÍFICA PARA GALACTOSIDOS DEL
MOLUSCO DE AGUA DULCE *Pomacea flagellata*.**

RESUMEN:

Uno de los grupos más estudiados de proteínas son **las lectinas**. Cuando en 1973 apareció la primera revisión de estas proteínas en el *Annual Review of Biochemistry*, el interés por estas proteínas con especificidad de unión para carbohidratos comenzaba a ganar impulso y, desde entonces, las publicaciones y el interés por las mismas no ha disminuido (Lis y Sharon 1986).

Recientemente, en el laboratorio de Bioquímica del Instituto de Química de la U.N.A.M., fueron aisladas dos aglutininas del caracol de agua dulce *Pomacea flagellata*, PFA-I y PFA-II (Arreguín-Espinosa y Arreguín-Lozano; 1997). Ambas lectinas tienen un peso de alrededor 30kDa y presentan afinidad por la D-galactosa, N-acetil-D-galactosamina, y por disacáridos de galactosa. Sin embargo, presentan diferencias en su punto isoeléctrico y contenido en carbohidratos, lo cual sugiere que son proteínas alelomórficas de la misma lectina, tal como se ha demostrado en las lectinas de otros invertebrados (Arreguín-Espinosa y Arreguín-Lozano; 1997). En este trabajo se presenta la caracterización de la lectina más abundante, PFA-I, la cual mediante espectrometría de masas se determinó que posee un peso molecular de 32,444 Da del cual el 7.4% corresponde a carbohidrato. Así mismo, se determinó la secuencia de los primeros 38 aminoácidos en el extremo amino y 11 del extremo carboxilo de la cadena peptídica de PFA encontrándose que esta proteína presenta homología en su extremo amino con la subunidad catalítica de la Ricina, una citotoxina aislada de las semillas de la “planta castor” (Castor bean). Por otro lado, al obtener el espectro de dicroísmo circular de PFA y compararlo con el de la subunidad catalítica de la ricina (RA) se observó que estos son muy similares y que comparten características generales y estructurales en su estructura secundaria y supersecundaria. Además se determinó que la afinidad de PFA es mayor por N-acetil galactósidos que por galactósidos simples.

INTRODUCCIÓN

1. LECTINAS

1.1 Características generales

Las lectinas son proteínas o dominios proteicos que generalmente se encuentran glicosilados, no son proteínas de transporte, y no tienen función enzimática. Las lectinas son proteínas capaces de reconocer con alta especificidad estructuras glicosídicas presentes en la superficie de las células, característica funcional que las ubica como moléculas clave en eventos de reconocimiento celular (Rini, 1995). Al estudiar las lectinas provenientes de distintos organismos, se ha visto que dicha “función de reconocimiento” varía de acuerdo al organismo y la lectina de la cual se trate. Por ejemplo, algunos patógenos las utilizan como el medio de unión a la superficie de células eucarióticas, tal como las hemaglutininas de los virus y las toxinas de las bacterias. En ciertos animales, las lectinas son parte del sistema de defensa del huésped iniciando los pasos de neutralización de los patógenos a través de la unión selectiva a los virus o a la membrana de las bacterias infectantes (Weis y Drickamer 1996). En otros casos, se les reconoce como mitógenos, ya que estimulan la multiplicación de células cuando éstas se encuentran en medios de cultivo (García, 1997). Es así que las lectinas son un grupo muy amplio de proteínas o dominios proteicos que se ha intentado clasificar de diversas maneras. A pesar de la gran variedad de lectinas que existen, éstas se pueden identificar mediante su purificación a través de cromatografías de afinidad con carbohidratos inmobilizados, y/o por la capacidad de “aglutinar” eritrocitos. Se ha observado que una misma lectina se puede componer de varias subunidades, las cuales por lo general se unen al mismo carbohidrato y presentan sitios de unión equivalentes (Lis y Sharon 1986).

1.2 Carbohidratos unidos covalentemente a las Lectinas

Se ha mencionado que muchas lectinas son glicoproteínas o dominios glicoproteicos, es decir, que poseen carbohidratos unidos de manera covalente. De acuerdo al carbohidrato que poseen, Lis y Sharon (1986) identifican dos clases de lectinas

vegetales: a) las que contienen principalmente manosa y N-acetilglucosamina, y b) las que contienen galactosa y L-arabinosa. Varias lectinas dentro del primer grupo contienen L-fucosa y xilosa, sin embargo, este último monosacárido no se ha encontrado asociado con manosa ni con N-acetilglucosamina en el mismo oligosacárido de glucoproteínas de levaduras o de animales.

1.3 Especificidad por el Carbohidrato

Se ha visto que las constantes de asociación de una lectina al interactuar con una serie de carbohidratos coinciden con la actividad de inhibición de la aglutinación de los mismos carbohidratos (Lis y Sharon, 1986). Existen además algunas lectinas que reconocen al carbohidrato y la secuencia de aminoácidos inmediatos del péptido al cual el carbohidrato se encuentra unido. Ejemplo de este tipo de lectinas es la de *Vicia gramínea*, una lectina específica para el grupo sanguíneo humano N, y que para su unión requiere de un grupo de NAc β 3Gal unido a una serina (o treonina) en cercanía al extremo amino de una leucina (Lis y Sharon 1986). Otra característica, es que muchas lectinas requieren de la presencia de cationes divalentes como Ca²⁺ y Mg²⁺ para que se de la unión al carbohidrato.

El empleo de métodos fisicoquímicos, como los usados en el estudio de sitios enzimáticos y de reconocimiento de anticuerpos, ha revelado que las lectinas tienen la capacidad de **unirse con alta afinidad a los carbohidratos**, especialmente oligosacáridos y glicopéptidos (García-Hernández y Hernández-Arana, 1999; García-Hernández *et al.*, 2000). Es así que las lectinas se han utilizado para separar glicoproteínas y glicopéptidos los cuales difieren muy poco en la composición o en la estructura de sus unidades de oligosacáridos mediante cromatografías de afinidad (Lis y Sharon, 1986b). El ejemplo más representativo de la especificidad de las lectinas y su uso en la bioquímica y la clínica fue su capacidad de aglutinación selectiva para los grupos sanguíneos (Watkins *et al.*, 1981; Lis y Sharon, 1986) tal como se muestra en la tabla 1.

Tabla 1. Lectinas específicas para grupo sanguíneo

Especificidad	Fuente de donde se obtuvo la lectina
Anti-A	<i>Griffonia (Bandeirae) simplicifolia</i> I (A ₄)
	<i>Helix Pomatia</i>
	<i>Phaseolus lunatus</i>
	<i>Vicia Cracca</i>
Anti-A ₁	<i>Dolichos biflorus</i>
Anti-B	<i>Griffonia (Bandeirae) simplicifolia</i> I (B ₄)
	<i>Anguilla anguilla</i>
Anti-O(H)	<i>Lotus tetragonolobus</i>
	<i>Ulex europeus</i>
Anti-A + N	<i>Mollucella laevis</i>
Anti-N	<i>Vicia graminea</i>
Anti-T	<i>Arachis hypogaea</i>
Anti-Tn	<i>Salvia sclarea</i>

Tomado de Lis y Sharon (1986).

Molecularmente, la interacción entre las lectinas y el carbohidrato al que se unen radica en una alta eficiencia energética que depende de su complementariedad estereoquímica. En las zonas de contacto del complejo lectina-carbohidrato, las regiones con una alta hidropatía interactúan con las superficies polares del carbohidrato de tal forma que la cooperatividad permite aumentar la fuerza de los puentes de hidrógeno y del resto de interacciones interatómicas. De tal forma, la complementariedad que existe entre una lectina y su ligando permite la formación de complejos muy estables a pesar del reducido número de contactos en la interfase (García-Hernández *et al.*, 2000).

2. LECTINAS ANIMALES

Aunque inicialmente existía mayor información de lectinas vegetales, el campo de las lectinas de origen animal se ha incrementado de manera considerable debido a que estas proteínas han demostrado que conforman un sistema que tiene la capacidad de reconocer los carbohidratos endógenos propios del mismo organismo, o aquéllos que se encuentran presentes en invasores patógenos (Drickamer y Taylor, 1993).

Una primera descripción de cómo las lectinas animales pueden discriminar de una manera adecuada entre varias estructuras glicosídicas se obtuvo a partir del

Tabla 1. Lectinas específicas para grupo sanguíneo

Especificidad	Fuente de donde se obtuvo la lectina
Anti-A	<i>Griffonia (Bandeirae) simplicifolia</i> I (A ₄)
	<i>Helix Pomatia</i>
	<i>Phaseolus lunatus</i>
	<i>Vicia Cracca</i>
Anti-A ₁	<i>Dolichos biflorus</i>
Anti-B	<i>Griffonia (Bandeirae) simplicifolia</i> I (B ₄)
	<i>Anguilla anguilla</i>
Anti-O(H)	<i>Lotus tetragonolobus</i>
	<i>Ulex europeus</i>
Anti-A + N	<i>Mollucella laevis</i>
Anti-N	<i>Vicia graminea</i>
Anti-T	<i>Arachis hypogaea</i>
Anti-Tn	<i>Salvia sclarea</i>

Tomado de Lis y Sharon (1986).

Molecularmente, la interacción entre las lectinas y el carbohidrato al que se unen radica en una alta eficiencia energética que depende de su complementaridad estereoquímica. En las zonas de contacto del complejo lectina-carbohidrato, las regiones con una alta hidropatía interactúan con las superficies polares del carbohidrato de tal forma que la cooperatividad permite aumentar la fuerza de los puentes de hidrógeno y del resto de interacciones interatómicas. De tal forma, la complementaridad que existe entre una lectina y su ligando permite la formación de complejos muy estables a pesar del reducido número de contactos en la interfase (García-Hernández *et al.*, 2000).

2. LECTINAS ANIMALES

Aunque inicialmente existía mayor información de lectinas vegetales, el campo de las lectinas de origen animal se ha incrementado de manera considerable debido a que estas proteínas han demostrado que conforman un sistema que tiene la capacidad de reconocer los carbohidratos endógenos propios del mismo organismo, o aquéllos que se encuentran presentes en invasores patógenos (Drickamer y Taylor, 1993).

Una primera descripción de cómo las lectinas animales pueden discriminar de una manera adecuada entre varias estructuras glicosídicas se obtuvo a partir del

receptor de la asialoglicoproteína en el hígado de los mamíferos (Drickamer y Taylor, 1993).

Las lectinas animales han sido clasificadas de acuerdo a la naturaleza de sus ligandos carbohidrato, los procesos biológicos en los que participan, su localización subcelular y su dependencia de cationes divalentes. La comparación de las secuencias y estructuras tridimensionales de los sitios o dominios de reconocimiento al carbohidrato (DRCs) muestra que las lectinas se pueden clasificar tal como se muestra en la siguiente tabla.

Tabla 2: Clasificación estructural de las lectinas animales

Grupo de lectinas	Estructura del DRC	Longitud	Ligandos típicos	Ejemplos de función
Tipo-L	β -plegada (sandwich)	230	Varios	Envío de proteínas en el RE
Tipo-P	Únicamente estructura β	130	Manosa 6-P	Envío de proteínas fuera del aparato de Golgi
Tipo-C	Estructura mixta α/β	115	Varios	Unión celular (selectinas) "limpieza" de glicoproteínas
Galectinas o Tipo S	β -sandwich	125	β -galactósidos	Unión de glicanos en la matriz extracelular
Tipo-I	Superfamilia tipo inmunoglobulinas	120	Acido siálico	Unión celular
Tipo-R	β -Trifoliada	125	Varios	Señalamiento de enzimas Encendido de hormonas glicoproteicas

Tomado de Dood y Drickamer (2001)

Sin embargo, las más estudiadas y de las que más información se tiene son las lectinas Tipo-C y lectinas Tipo S. A continuación, se presentan algunas de las características generales y ejemplos para ambos grupos .

2.1 Lectinas del Tipo-C

Las lectinas del Tipo-C son el grupo más diverso dentro de las lectinas animales. Están caracterizadas por un dominio de reconocimiento de aproximadamente 15-kDa, el cual es dependiente de calcio para realizar su actividad de reconocimiento y que generalmente se encuentra unido a otro dominio proteico accesorio. Ejemplos de estas lectinas son: los receptores glucoproteicos endocíticos, las selectinas, el receptor de manosa que se encuentra en los macrófagos y las colectinas solubles (Rini, 1995).

Dada la variedad de estas lectinas, Drickamer y Taylor (1993) hacen una clasificación de las mismas de acuerdo a su solubilidad es decir, si se encuentran libres en suero, en la matriz extracelular o en membranas intracelulares (tabla 3).

Tabla 3. Lectinas del Tipo-C. Clasificación de acuerdo a sus funciones y localización

Función	Localización	Organización	Ligandos	Lectinas
Receptores endocíticos	Membrana plasmática	Tipo II Transmembranal	Endógenos	Receptor de la asiaglicoproteína
			Endógenos	Receptor de las células Kupffer
			Endógenos	Lectina hepática
		Tipo I Transmembranal	Endógenos	Receptor del Linfocito para la IgE Fc
			Exógenos	Receptores de las células asesinas
Moléculas de adhesión	Membrana plasmática	Tipo I Transmembranal	Exógeno	Receptor de manosa
			Endógeno	Selectinas
Defensa Humoral	Extracelular	Soluble, colagenosa	Exógeno	Proteínas de unión a Manosa
			Endógeno	Agreganos, versicanos y Neurocanos
Proteoglicanos	Matriz celular	Extendida		
CRDs libres	secretados	Solubles	desconocidos	Proteínas del hepatoma pancreático

Tomado de Drickamer y Taylor (1993).

En el caso de las colectinas, se trata de un grupo de proteínas tipo colágena en cuyo extremo carboxilo se encuentra el DRC. Éstas se dividen en dos grupos, el primero consiste en péptidos tipo colágeno que forman estructuras semejantes a un ramillete, mientras que en el segundo, los péptidos colágeno forman una estructura cruciforme (Drickamer y Taylor, 1993). La primera estructura tridimensional de una lectina Tipo-C se obtuvo a partir del cristal de la proteína-A unidora de manosa (mannose-binding protein A, MBP-A), una de las colectinas solubles responsables de la defensa inmunológica independiente de anticuerpo. La estructura cristalina reveló que la interacción del DRC con el azúcar se limita a una pequeña porción cercana a los iones calcio, que a su vez interactúan con los grupos hidroxilo de los carbonos C3 y C4 de la manosa. Este complejo es estabilizado por dos puentes de hidrógeno para cada grupo hidroxilo los cuales involucran a la asparagina y al ácido glutámico que también funcionan como unidores de calcio.

Las selectinas son otro ejemplo de lectinas Tipo-C. Son proteínas transmembranales en cuyo extremo amino se encuentra el DRC, seguido de un dominio semejante al del factor de crecimiento epidérmico (EGF-like), un número variable de

repeticiones de unión del complemento y el extremo carboxilo que se encuentra al interior de la membrana. Estas proteínas son las encargadas de regular la fase inicial de unión entre los leucocitos y el endotelio. Por ejemplo, la selectina-L de las células T permite el reconocimiento de dichas células con los nódulos linfáticos periféricos, mientras que las selectinas E y P del endotelio interactúan con los neutrófilos y monocitos (Drickamer y Taylor 1993). Por otra parte, la selectina-P de las plaquetas regula la unión de éstas con el endotelio. Aunque aún no se determina el espectro completo de ligandos endógenos a los cuales se unen la selectinas, es claro que todas presentan afinidad por estructuras que presenten ácido siálico y fucosa.

Sin embargo, en literatura reciente (Drickamer, 1999) se señala que algunos módulos proteicos que contienen al dominio de unión de las lectinas tipo-C (CTLDs C-type lectin-like domains) presentan cambios en el modo de reconocimiento a los sácaridos así como otras funciones además del reconocimiento de los mismos. Por ejemplo, los DRC de la subfamilia de los CTLDs que poseen la misma especificidad de unión a sus ligandos, lo hacen de distinta manera, es decir, aunque en todas las estructuras disponibles de los ligandos de DRC, el carbohidrato se une a la proteína mediante la formación de puentes coordinados con el Ca^{2+} , la unión de azúcares como la manosa y la N-acetilglucosamina con el DRC pueden presentarse en dos posiciones que difieren en una orientación de 180° .

Dado lo anterior, y debido a que las lectinas Tipo-C son el grupo más amplio, se puede esperar que surja una nueva clasificación de las mismas con base no sólo en su localización, sino también a su reconocimiento al ligando, CTLDs, e incluso dependencia de unión al Ca^{2+} .

2.2 Lectinas Tipo-S o Galectinas

El nombre de lectinas Tipo-S se debe a que en un principio se creyó que éstas requerían de tioles o agentes reductores para mantener su actividad. Posteriormente, se observó que esta propiedad no es compartida por todas ellas. Sin embargo, todas poseen la capacidad de unir β -galactósidos, por lo que el término galectinas se consideró como más adecuado y evocativo de su afinidad de unión (Drickamer K.1995). Las galectinas se han caracterizado en diferentes especies animales y usualmente se les localiza

intracelularmente. De acuerdo con su estructura primaria, Drickamer y Taylor (1993) proponen que las galectinas se pueden subdividir en 3 grupos, las L-14 que son lectinas compuestas de subunidades de 14,000 Da; las L-30 que son aquellas que poseen pesos moleculares de aproximadamente 30,000 Da y que existen como monómeros; y las L-36 que consisten de dos CRDs del tipo-S. Barondes *et al.* (1994), proponen una nueva clasificación para galectinas obtenidas a partir de mamíferos (1, 2, 3, y 4) en las cuales se encuentran incluidas las galectinas L-14, L-30 y L-36, sin embargo, esta clasificación sólo se aplica a mamíferos y por lo mismo, la más común es L-14, L-30 y L-36.

Las galectinas L-14, al igual las galectinas-1 y 2 (Barondes, *et al.* 1994), están constituidas por homodímeros que poseen subunidades de 130 aminoácidos los cuales se pliegan a manera de dominios globulares y que se definen como los DRCs de las mismas. Estas galectinas presentan la misma afinidad de unión por la lactosa y los β -galactósidos. Su estructura de rayos X muestra que casi siempre se unen al residuo galactosa de la lactosa, aunque también existe una interacción significativa con la glucosa, puesto que se ha observado que su afinidad por la lactosa es 100 veces mayor que la que existe sólo por la galactosa. Aunque algunas galectinas son claramente secretadas al exterior celular, ninguna de ellas presenta o posee péptidos señaladores de secreción. Así, se pueden encontrar galectinas tanto en el citosol como en el medio extracelular (Barondes *et al.*, 1994).

Se cree que las galectinas L-14 están involucradas en el desarrollo celular, ya que su presencia varía del interior al exterior celular durante los procesos de diferenciación. Al igual que las galectinas L-14, las galectinas L-30 también son secretadas al exterior puesto que se les encuentra como el principal antígeno de superficie en los macrófagos (Barondes *et al.*, 1994). Su peso molecular promedio es de aproximadamente 30 kDa y usualmente existen como monómeros, se les localiza tanto en el interior como en el exterior de la célula. En el caso de L-30 extracelular ésta ha sido identificada como un antígeno de reconocimiento en la superficie de los macrófagos, como lectina de superficie en células tumorales y como proteína de unión a la Ig-E en las células leucémicas basófilas (Frigeri *et al.* 1993). Así mismo, la afinidad de la galectina-3 por la Ig-E y/o su receptor puede disparar la activación de los mastocitos y de los basófilos jugando un

papel importante en la respuesta inflamatoria (Barondes *et al.*, 1994). Respecto de como se encuentra en el interior celular se ha observado que en los fibroblastos la proteína se mueve del citoplasma al núcleo en forma de un complejo proteína-RNA conforme las células pasan de su estado de proliferación a su estado quiescente (Drickamer y Taylor, 1993). Esta observación junto con la homología que las lectinas L-30 presentan en su región amino con otras ribonucleoproteínas las ubica como componentes de los complejos ribonucleoproteicos (Barondes *et al.*, 1994).

Por su parte, las galectinas L-36 o gal-4 son monómeros que poseen 2 DRC (similares a los de las galectinas L-14 y L-30) en una misma cadena peptídica, dichos DRCs se encuentran conectados por una secuencia aminoácida que es homóloga a la secuencia repetida de las Gal-3 (Barondes *et al.*, 1994). Hasta el momento no existe mucha información de estas galectinas las cuales se han caracterizado en su mayoría a partir del hígado de rata.

Se puede observar que en los animales existe una gran variedad y estudios sobre lectinas, sin embargo en los invertebrados ha tomado un auge el estudio de las mismas tal como se describe líneas abajo.

2.3 Lectinas de invertebrados

El interés que existe en las lectinas de animales se debe a la gran diversidad de funciones en las cuales están involucradas, por ejemplo su participación en la defensa humoral y celular, el papel que juegan en la adhesión celular, la fertilización, la organización de tejidos, y la metamorfosis de los organismos. Hasta el momento se ha visto que existen varias clases de lectinas, y que éstas se encuentran distribuidas en todos los animales. Desde el punto de vista inmunológico, las lectinas son parte del mecanismo más antiguo de defensa, la inmunidad innata, la cual consiste en la producción de proteínas de reconocimiento ya sean solubles o de membrana a partir de una línea germinal de codificación, a diferencia de la inmunidad adquirida, que se basa en la producción de anticuerpos y receptores para antígenos a partir de un arreglo genético de células somáticas (Shun-ichiro y Sadaaki 1999).

Ya que la inmunidad innata es más antigua que la adquirida, ésta se encuentra en todos los organismos multicelulares; sin embargo, en los invertebrados se constituye como el único sistema de defensa el cual, a pesar de su antigüedad, es muy complejo y eficiente, tal como lo demuestran los estudios llevados a cabo en el cangrejo de herradura *Tachypleus tridentatus* y en la cucaracha *Blaberus discoidalis* (Shun-ichiro y Sadaaki, 1998; Wilson, *et al.* 1999).

En la hemolinfa del cangrejo herradura *T. tridentatus* se han identificado 5 lectinas denominadas taquilectinas las cuales participan activamente como elementos de defensa. Cuatro de estas lectinas se localizan en los hemocitos y son liberadas cuando éstos tienen contacto con lipopolisacáridos (LPS); por su parte, la quinta taquilectina se encuentra libre en el plasma. Cada una de estas proteínas presenta un reconocimiento específico para diferentes estructuras de carbohidratos y, por lo mismo cada una presenta afinidad de unión por diferentes patógenos. La taquilectina-1, por ejemplo, interactúa con los LPS derivados del tipo silvestre y rugoso de *Salmonella minnesota*, probablemente a través del disacárido 2-ceto-3-deoxioctanato, que es uno de los componentes principales de los LPS de la mayoría de las bacterias Gram-negativas; la taquilectina-2 se une específicamente a D-GlcNAc o D-GalNac y además reconoce ácidos lipoteicoicos de estafilococos y varios LPS. En contraste, la taquilectina 4 se une específicamente a los LPS tipo-S de *E. coli* O111:B4 a través de la mitad de un azúcar del antígeno-O. Por su parte, la taquilectina-3 presenta una mayor especificidad por los antígenos O en los LPS de las bacterias Gram-negativas que la taquilectina-4. Por otro lado, las taquilectinas 5 (a y b) muestran afinidad por todas las sustancias que poseen un grupo acetilo. De esta manera, el sistema inmune del cangrejo herradura es capaz de reconocer los patógenos invasores a través de un sistema de combinaciones que utiliza lectinas con estricta y con amplia especificidad para reconocer las moléculas expuestas en los patógenos.

Otro ejemplo de lectinas que intervienen en la inmunidad innata lo presenta la cucaracha *B. discoidalis* de la cual se han aislado 4 lectinas, BDL1, BDL2, BDL3 y GSL (Wilson, *et al.*, 1999). Se ha demostrado que BDL1 y GSL provocan un aumento en la respuesta fagocítica de los plasmocitos de *B. discoidalis*. Dicho efecto se relaciona con los carbohidratos presentes en la superficie de los microorganismos y la especificidad de

las lectinas por los azúcares. Así mismo, se demostró que las cuatro lectinas eran específicas para diferentes carbohidratos de tal manera que BDL-1 es específica para manosa y provoca un aumento en la fagocitosis de *E. coli* y la levadura de pan; BDL2 es específica para N-acetilglucosamina y N-acetilgalactosamina e induce la fagocitosis de *Bacillus cereus* y *E. coli*. Por su lado, GSL es específica para β -1-,3-glucanos y N-acetil-D-galactosamina; BDL3 también presenta afinidad para N-acetil-D-galactosamina, pero además provoca un aumento en la fagocitosis de *B. cereus* y de la levadura.

Es claro que las lectinas desarrollan un papel fundamental en el sistema inmune de los invertebrados. Empero, pocos ejemplos en la literatura mencionan sobre lectinas provenientes de moluscos, aunque se reconoce que éstas se encuentran ampliamente distribuidas en los invertebrados (Lis y Sharon, 1986).

2.4 Aglutininas de *Pomacea flagellata*

Dado que, como se mencionó líneas arriba, las lectinas participan en múltiples procesos biológicos que incluyen la fertilización, el desarrollo, y el reconocimiento inter e intracelular además de que algunas presentan propiedades tóxicas, se les considera una herramienta muy importante de la investigación biomédica (Arreguín-Espinosa y Arreguín-Lozano; 1997). Por lo mismo, el conocimiento de la estructura de las lectinas y su modo de acción permitirá conocer más acerca de los mecanismos de la fisiología celular.

En nuestro laboratorio se aislaron dos aglutininas del caracol *P. flagellata*, PFA-I y PFA-II (Arreguín-Espinosa y Arreguín-Lozano; 1997). Las características de éstas son un peso de alrededor 30kDa y afinidad por la D-galactosa, N-acetil-D-galactosamina, y por disacáridos de galactosa. Presentan diferencias en su punto isoeléctrico y contenido en carbohidratos, lo que sugiere que son proteínas alelomórficas de la misma lectina, tal como se ha demostrado en las lectinas de otros invertebrados (Arreguín-Espinosa y Arreguín-Lozano; 1997).

La actividad de hemaglutininas de PFA-I y PFA-II, se ve afectada por la mucina del estómago del cerdo (un complejo de O-glicanos), el orosomucoide (α_1 -glicoproteína

ácida), la lactoferrina y la transferina humanas (proteínas N-glicosiladas con glicanos lactosamínicos), y la fetuina (la cual contiene N- y O-glicanos). Así mismo, estas glicoproteínas pierden actividad después de tratarlas con β -galactosidasa. Estos resultados sugieren la importancia de la β -galactosa como determinante de la especificidad de las isolectinas de *P. flagellata*.

Debido a que el conocimiento y caracterización de las lectinas de invertebrados es un objetivo para el estudio de los procesos bioquímicos que involucran interacciones receptor-ligando, reconocimiento celular, respuesta inmunológica, y que hasta el momento son muy pocas las lectinas de moluscos cuyas propiedades conformacionales han sido caracterizadas, su estudio se presenta como un campo abierto para la investigación bioquímica y biológica ya que después de los artrópodos los moluscos son el grupo más abundante de organismos en los animales. En este trabajo, se plantea la caracterización de las propiedades estructurales de la aglutinina I de *P. flagellata*, PFA-I, que es la isolectina más abundante, como una manera de aumentar el conocimiento de las lectinas de invertebrados, sus propiedades y sus relaciones con las lectinas de otros organismos.

3. IDENTIFICACIÓN DE *Pomacea flagellata*

3.1 *Pomacea flagellata* (Say, 1827)

Actualmente diversas disciplinas como la morfo-anatomía, fisiología paleontología, la bioquímica y la biología molecular proporcionan diversos métodos de análisis para la sistemática y clasificación de los organismos. Sin embargo, estos estudios no les restan importancia a la identificación taxonómica tradicional, basada en la caracterización anatómica escrupulosa de los organismos. Por lo anterior, a continuación se presenta una descripción del hábitat y la caracterización taxonómica del caracol *P. flagellata*, organismo a partir del cual se obtuvieron las proteínas objetivo de este estudio.

Los caracoles de *P. flagellata* pertenecen al *Phyllum Mollusca* y a la clase Gasteropoda (Cuvier, 1797). En el género *Pomacea* (Perry, 1810), único de la familia

ácida), la lactoferrina y la transferina humanas (proteínas N-glicosiladas con glicanos lactosamínicos), y la fetuina (la cual contiene N- y O-glicanos). Así mismo, estas glicoproteínas pierden actividad después de tratarlas con β -galactosidasa. Estos resultados sugieren la importancia de la β -galactosa como determinante de la especificidad de las isolectinas de *P. flagellata*.

Debido a que el conocimiento y caracterización de las lectinas de invertebrados es un objetivo para el estudio de los procesos bioquímicos que involucran interacciones receptor-ligando, reconocimiento celular, respuesta inmunológica, y que hasta el momento son muy pocas las lectinas de moluscos cuyas propiedades conformacionales han sido caracterizadas, su estudio se presenta como un campo abierto para la investigación bioquímica y biológica ya que después de los artrópodos los moluscos son el grupo más abundante de organismos en los animales. En este trabajo, se plantea la caracterización de las propiedades estructurales de la aglutinina I de *P. flagellata*, PFA-I, que es la isolectina más abundante, como una manera de aumentar el conocimiento de las lectinas de invertebrados, sus propiedades y sus relaciones con las lectinas de otros organismos.

3. IDENTIFICACIÓN DE *Pomacea flagellata*

3.1 *Pomacea flagellata* (Say, 1827)

Actualmente diversas disciplinas como la morfo-anatomía, fisiología paleontología, la bioquímica y la biología molecular proporcionan diversos métodos de análisis para la sistemática y clasificación de los organismos. Sin embargo, estos estudios no les restan importancia a la identificación taxonómica tradicional, basada en la caracterización anatómica escrupulosa de los organismos. Por lo anterior, a continuación se presenta una descripción del hábitat y la caracterización taxonómica del caracol *P. flagellata*, organismo a partir del cual se obtuvieron las proteínas objetivo de este estudio.

Los caracoles de *P. flagellata* pertenecen al *Phyllum Mollusca* y a la clase Gasteropoda (Cuvier, 1797). En el género *Pomacea* (Perry, 1810), único de la familia

Ampullaridae registrado en América, la última espira de la concha está comprimida y es moderadamente ovoide con una abertura elíptica.

El caracol del género *Pomacea* es un organismo anfíbio que se localiza principalmente en las lagunas de agua dulce de los estados de Veracruz, Tabasco y Quintana Roo. Durante el día se encuentra sumergido en el fondo o sobre la vegetación acuática, y sólo en la noche es cuando se le puede localizar en las orillas de los lagos. Esta especie es preferentemente herbívora, aunque en varias ocasiones se le ha visto alimentarse de peces muertos (Diupotex, 1998). La concha, de acuerdo a la descripción de Rangel (1988), es dextrógira, subglobosa, con un ángulo en el ápice que va de 93.5 a 136.5° y posee una coloración que va de pardo a pardo verdosa. Las vueltas de la espira son cortas y son 3 ó 4, se encuentran ligeramente aplanadas en su parte superior.

Respecto a su complejo cabeza-pie, presenta un par de tentáculos grandes, filiformes, y en la base de cada uno, poseen un omatóforo bien desarrollado en el cual se implantan los ojos. El pie es bastante grande y musculoso, anteriormente es ligeramente espatulado y en su región posterior es puntiagudo.

La boca está localizada en el centro del extremo distal de la trompa, a los lados de ésta se encuentran un par de palpos labiales filiformes muy desarrollados. En el extremo anterior de la masa bucal se encuentra la mandíbula, fuerte, de constitución acelular y detrás de ésta se encuentra el complejo radular sobre la masa bucal. La rádula es de tipo taeniogloso y las cúspides en general tienen forma triangular (Diupotex, 1998). El esófago comunica directamente con el intestino el cual presenta tres porciones, siendo la última en donde desemboca el recto que se abre en el ano, el cual está situado a la derecha del inicio de la branquia cerca del borde de la cavidad del manto. La glándula digestiva ocupa las 3 primeras vueltas de la concha y desemboca en la segunda porción del intestino (Diupotex, 1998). El aparato excretor está representado por un solo riñón, el cual tiene forma alargada y fusiforme y se encuentra localizado en la base de la branquia.

Finalmente, *P. flagellata* es una especie dioica y no presenta dimorfismo sexual marcado ni en forma ni en tamaño, el sexo sólo se puede conocer después de la disección.

Identificación taxonómica:

Phyllum: Mollusca, (Cuvier, 1797)

Clase: Gastropoda, (Cuvier, 1797)

Sub-clase: Prosobranchia, (Edwards, 1848)

Orden: Mesogastropoda, (Pennak, 1974)

Superfamilia: Viviparacea, (Taylor y Schi, 1962)

Familia: Ampullariidae, (Guilding, 1925)

Familia: Pilidae, (Bequaer y Clench, 1937)

Género: Pomacea, (Perry, 1810)

Especie: Flagellata, (Say, 1827)



3.2 HUELLA GENÓMICA DE *Pomacea flagellata*

En la actualidad diversas disciplinas se utilizan como herramientas fundamentales para llevar a cabo la sistemática moderna, y por lo tanto existen estudios que complementan la información acerca de la identificación taxonómica del organismo de interés. En el caso de *P. flagellata*, aunque ha sido identificado taxonómicamente de acuerdo a sus características anatómicas y se han realizado estudios cromosómicos y electroforéticos (Diupotex, 1998), no existe hasta el momento ningún estudio de tipo molecular que permita identificar a la especie mediante pruebas de DNA genómico.

Una alternativa para detectar el polimorfismo genómico de la especie de caracol *P. flagellata* con respecto a otras especies es el uso de RAPDS (por sus siglas en inglés de Random Amplified Polimorphic DNA). Los RAPDS son marcadores genéticos que se basan en la amplificación de DNA genómico con primers (o iniciadores) de secuencia nucleotídica arbitraria, que permiten detectar polimorfismos interespecie en ausencia de información nucleotídica, de tal manera que el polimorfismo funciona como marcador genético y puede usarse para construir mapas genéticos (Williams *et al.*, 1990). Una de las principales ventajas de los RAPDS como marcadores es que casi todos son dominantes, es decir, los segmentos de DNA con una determinada longitud que amplifican para el individuo de una especie, no amplifican para otra especie. Además, con un marcador RAPD no es posible distinguir si el segmento de DNA amplificado de un locus corresponde a un heterócigo o a un homócigo, aunque en algunas pocas ocasiones se han encontrado segmentos de diferente tamaño de DNA amplificado que pertenecen al mismo locus se trata de un evento muy poco frecuente (Williams, *et al.*, 1990).

El mapeo genético usando RAPDs tiene varias ventajas sobre otros métodos; por ejemplo, un grupo universal de primers que pueden ser usados en diferentes especies. No se requiere de trabajo preliminar, tal como el aislamiento y clonación de sondas, así como la preparación de filtros para la hibridación, tampoco es necesario hacer la secuencia nucleotídica; y finalmente, mediante los RAPDS los mapas genéticos se pueden automatizar. Por las ventajas que esta técnica ofrece para la identificación genética de una especie, se propone su uso en *P. flagellata*, así el organismo a partir del cual se extraerán las lectinas de estudio quedará plenamente identificado.

4 OBJETIVO GENERAL

El objetivo de este trabajo es la caracterización estructural de PFA-I, una aglutinina específica para β -galactosidos, mediante técnicas bioquímicas y espectroscópicas.

OBJETIVOS PARTICULARES

1. Determinar la huella digital de *P. flagellata* mediante la extracción de DNA somático y su posterior amplificación utilizando distintos primers para RAPDS
2. Determinar el peso molecular de PFA por medio de espectroscopía de masas de alta resolución, y conocer el porcentaje de carbohidratos que contribuye al peso total de la proteína.
3. Determinar la secuencia a partir de los extremos carboxilo y amino terminal de la cadena polipeptídica, mediante degradación secuencial y espectrometría de masas, así como de la secuenciación automatizada de Edman, respectivamente.
4. Determinar la estructura secundaria de PFA y su clase estructural terciaria mediante dicroísmo circular.
5. Comparar la secuencias aminoácidas obtenidas con el banco de datos de SWISS-PROT para encontrar si existe similitud con alguna otra proteína.

Para cumplir con los objetivos planteados fue necesario recurrir a técnicas bioquímicas de extracción, separación y caracterización, las cuales serán explicadas en las siguientes páginas.

HIPÓTESIS

Perteneciendo PFA al grupo de las lectinas, un conocimiento más detallado de su estructura permitirá acercarse a una mejor comprensión de las bases moleculares del funcionamiento de este tipo de proteínas, y así mismo determinar si existe relación filogenética con alguna otra lectina antes reportada.

4.1 Cromatografía de afinidad:

Como ya se mencionó antes, las lectinas son proteínas que se pueden purificar mediante cromatografías de afinidad con carbohidratos inmovilizados. Este método de separación proteica no sólo se aplica a las lectinas, sino a cualquier otra proteína que presente unión no covalente y reversible con otra molécula llamada ligando. Cuando se desea separar una proteína (que se ha demostrado que posee ligando), de otras proteínas mediante la cromatografía de afinidad, su ligando específico se une de manera covalente a un grupo funcional situado sobre la superficie de las partículas de una columna de material poroso. Así, la molécula proteica de interés se une específicamente al ligando inmovilizado, se adhiere a las partículas de la columna que contienen dicho ligando, mientras que todas las demás proteínas pasan sin ser retenidas a través de la columna.

Este método depende en gran medida de la afinidad biológica de la proteína por su ligando característico y, por lo tanto, las condiciones de la solución amortiguadora deben optimizarse para asegurar que las moléculas blanco interactúen con el ligando y sean retenidas mientras que el resto son lavadas a través de la columna. Posteriormente, la proteína unida específicamente a las partículas de la columna puede ser eluida mediante un cambio de solución que debilite la interacción con las moléculas ligando y, que al mismo tiempo mantenga la estructura y actividad biológica de la proteína que interesa eluir. La solución eluyente puede basarse en cambios de pH, cambios en la fuerza iónica, y/o soluciones con eluyentes competitivos (ligando en forma libre, o inhibidores competitivos de la unión con el ligando) (Lowe, 1974 y *Amersham Pharmacia Biotech* 2001).

Algunas de las interacciones biológicas típicas que se aprovechan en las cromatografías de afinidad son:

Enzima	↔	sustrato, inhibidor, cofactor
Anticuerpo	↔	antígeno, virus, célula
Lectina	↔	Polisacárido, glicoproteína, receptor de superficie celular, célula
Acido nucleico	↔	secuencia de bases complementarias, histonas, DNA polimerasa o proteínas de unión a ácidos nucleicos
Hormonas, vitaminas	↔	receptor, o proteína acarreadora
proteínas de fusión Poli (His), proteínas nativas con histidina	↔	Iones metálicos

Las interacciones biológicas que pueden existir entre las proteínas y sus ligandos pueden ser del tipo interacciones hidrofóbicas o electrostáticas, fuerzas de van der Waals y/o puentes de hidrógeno.

Mediante este método, algunas proteínas se pueden aislar en una sola etapa a partir de mezclas muy complejas y alcanzar un grado de purificación muy elevado. La cromatografía de afinidad es una técnica que se emplea para aislar proteínas tan diversas como enzimas, moléculas receptoras de las membranas celulares que captan hormonas específicas o bien, proteínas con afinidad por los carbohidratos como son las lectinas. Sin embargo, en algunos casos es necesario llevar a cabo otro tipo de separación que permita tener un mayor grado de pureza y homogeneidad de la proteína, del tal forma que basándose en la naturaleza química de las proteínas que se desean obtener se utilizan otras estrategias de separación como son las columnas de intercambio iónico.

4.2 Cromatografía de intercambio iónico

Esta separación cromatográfica se basa en la adsorción reversible de moléculas con carga eléctrica que se fijan a grupos iónicos de carga opuesta inmovilizados.

Un intercambiador iónico consiste de una matriz insoluble a la cual se le ha unido de manera covalente grupos cargados. Estos grupos cargados se asocian con los iones-

contables. Estos iones-contables se pueden intercambiar con otros iones de la misma carga sin alterar la matriz.

Es posible tener intercambiadores cargados tanto positivamente como negativamente (figura 1). Los intercambiadores cargados positivamente tienen iones-contables cargados negativamente (aniones) disponibles para ser intercambiados y por lo tanto se llaman intercambiadores aniónicos. Los intercambiadores cargados negativamente tienen iones cargados positivamente (cationes) y se denominan intercambiadores catiónicos.

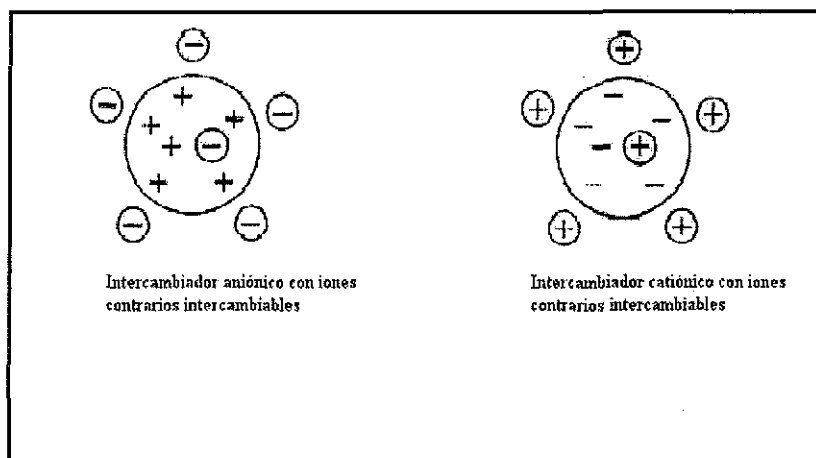


Figura 1: Tipos de matrices para intercambio iónico. Tomado de *Amersham Pharmacia Biotech 2001*

La matriz de las columnas se basa en compuestos inorgánicos, resinas sintéticas o polisacáridos. Las características de la matriz determinan sus propiedades cromatográficas tales como la eficiencia, capacidad y recuperación así como su estabilidad química, fuerza mecánica y propiedades de flujo. La naturaleza de la matriz afecta su comportamiento hacia las sustancias biológicas y su actividad biológica (*Amersham Pharmacia Biotech 2001*).

La mayoría de los experimentos en una cromatografía de intercambio iónico se suceden en cinco etapas, las cuales son ilustradas a continuación.

TESIS CON
FALLA DE ORIGEN

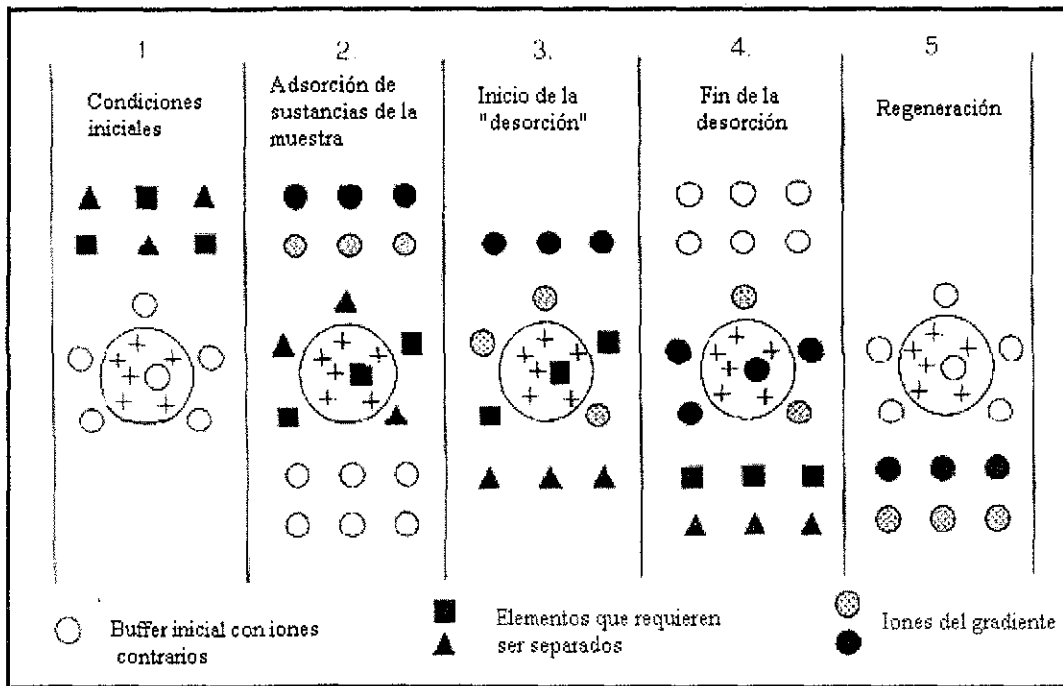


Figura 2: Principio de la cromatografía de intercambio iónico. Tomado de *Amersham Pharmacia Biotech 2001*

La primera etapa consiste en equilibrar el intercambiador iónico llevándolo a las condiciones iniciales en términos de pH y fuerza iónica que permiten la unión de las moléculas de soluto deseadas (figura 2.1). Los grupos intercambiadores se asocian con los iones-contrarios que son intercambiables (usualmente aniones o cationes simples, tales como el cloro o el sodio).

La segunda etapa es la aplicación de la muestra y la adsorción del soluto cuando sus moléculas poseen la carga adecuada para desplazar los iones-contrarios y se unen de manera reversible al gel (figura 2.2). Las sustancias que no se unen pueden lavarse y excluirse de la cama intercambiadora usando el buffer inicial.

En la tercera etapa, los solutos de interés se remueven de la columna al cambiar las condiciones para que sean desfavorables a la unión iónica. Esto normalmente involucra el aumento de la fuerza iónica de la solución amortiguadora de elución o cambio de su pH. En la figura 2.3, la desorción se alcanza a través de un gradiente ascendente en la concentración de sales, de tal manera que las moléculas del soluto son liberadas de la columna en el orden de sus fuerzas de unión, así las moléculas con unión más débil son las primeras en eluir.

TESIS CON
FALLA DE ORIGEN

La cuarta y quinta etapas corresponden a la remoción de sustancias que no eluyeron bajo las condiciones experimentales y el reequilibrio a las condiciones iniciales para la siguiente purificación (figuras 2.4 y 2.5).

La separación que se obtiene en estas condiciones se debe a los diferentes grados de interacción con el intercambiador iónico debido a las diferencias en sus cargas, densidades de carga y distribución de carga en su superficie. Estas interacciones se pueden controlar a través de la variación de las condiciones tales como la fuerza iónica y el pH. Las diferencias de carga son propiedades de los compuestos biológicos que a menudo deben considerarse ya a que la cromatografía de intercambio iónico es capaz de separar especies con diferencias muy pequeñas en sus propiedades, por ejemplo, dos proteínas que difieren en la carga de un solo aminoácido.

En la cromatografía de intercambio iónico se puede elegir si se unen las sustancias de interés y se permite que las mezclas complejas pasen a través o bien que los contaminantes se unan y que las sustancias de interés pasen a través. Generalmente, el primer método es más utilizado ya que permite un mayor grado de fraccionamiento o separación y concentra las sustancias de interés.

4.3 Espectrometría de masas

La espectrometría de masas es una herramienta indispensable en química. Las moléculas que se analizan por este método se conocen como analitos y son ionizadas en condiciones de vacío. Cuando estas moléculas, recién cargadas, se introducen en un campo eléctrico y/o magnético, la ruta que siguen a través del campo esta en función de su proporción masa/carga, m/z . Esta propiedad mensurable de las especies ionizadas puede ser usada para deducir la masa (M) del analito con alta precisión.

Aunque la espectrometría de masas se ha usado durante muchos años, ésta no se había aplicado a macromoléculas como las proteínas o los ácidos nucleicos debido a que en la técnica las mediciones de m/z se hacen en la fase gaseosa de las moléculas y, el calentamiento y otros tratamientos que se llevaban a cabo para llevar a la molécula al estado gaseoso causaban una rápida descomposición de estas macromoléculas. En 1988, se desarrollaron 2 técnicas que superaron este problema. En una de ellas, las proteínas se colocan en una matriz que absorbe la luz, de tal manera que con un pulso de luz láser las proteínas se ionizan y posteriormente se separan de la matriz en un sistema de vacío. Este proceso se conoce como MALDI-MS (de las siglas en inglés matrix assisted laser desorption/ionization mass spectrometry) y ha sido utilizado para medir la masa de una gran variedad de macromoléculas. Además, de acuerdo a Jensen *et al.* (1998), esta técnica es más simple y sensible. La segunda técnica consiste en forzar el paso de las moléculas en solución líquida a la fase gaseosa. Es decir, una solución de analitos se pasa a través de una aguja cargada que se encuentra sometida en un alto impacto electrónico, de tal forma que la solución que sale de la aguja es una fina “nube de microgotas cargadas eléctricamente”. El solvente que rodea a las macromoléculas se evapora y el resultado son macromoléculas ionizadas en fase gaseosa. Esta técnica se conoce como ESI-MS (de la siglas en inglés electrospray ionization mass spectrometry). Los protones añadidos durante el paso a través de la aguja proporcionan una carga adicional a la macromolécula y ésta se puede analizar en la cámara de vacío (figura 3).

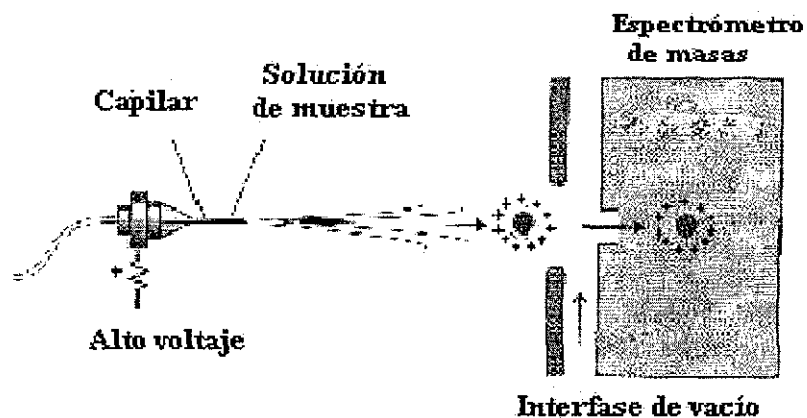


Figura 3. ESI MS de una solución de proteína dispersada como gotas eléctricamente cargadas, después de pasar por una aguja sometida a un campo de alto voltaje. Las gotas se evaporan y los iones (con protones añadidos) entran al espectrómetro para la medición m/z . Tomado de Nelson y Cox, 2000.

La espectrometría de masas en la investigación proteómica es una herramienta muy útil ya que al proporcionar la masa exacta de la proteína que se estudia, característica esencial en la identificación de la misma, permite determinar cambios que puedan ocurrir por la unión de cofactores, iones metálicos o modificaciones covalentes que alteran la masa de ésta. Además, para realizar este tipo de estudio se requiere de mínimas cantidades de proteína (fentomoles).

El proceso por el cual a una proteína se le determina su masa mediante ESI MS consiste en que durante la inyección la proteína adquiere un número variable de protones y por lo tanto de cargas positivas del solvente creando un espectro de especies con diferentes proporciones de masa-carga. Cada pico que se da sucesivamente corresponde una especie que difiere del pico vecino por una diferencia de carga de 1 y una diferencia de masa de 1 (1 protón). La masa de la proteína se determina a partir de cualquier par de picos vecinos. Así, si el valor de m/z de un pico es:

$$(m/z)_2 = \frac{M + n_2 X}{n_2}$$

donde, M es la masa de la proteína, n_2 es el número de cargas y X es la masa de los grupos añadidos (protones en este caso). De la misma manera para el pico vecino

$$(m/z)_1 = \frac{M + (n_2 + 1)X}{n_2 + 1}$$

Luego entonces, se tienen dos incógnitas (M y n_2), y dos ecuaciones. Si se despeja primero para n_2 y luego para M se tiene:

$$n_2 = \frac{(m/z)_2 - X}{(m/z)_2 - (m/z)_1}$$

$$M = n_2 [(m/z)_2 - X]$$

Este cálculo usando los valores de m/z para cualquier par de picos de un espectro cualquiera como el que se muestra en la figura de abajo, proporciona la masa de la proteína con un error de sólo el $\pm 0.01\%$. Si se tienen varios grupos de picos y se repite el cálculo, un promedio de los resultados proporciona un valor más exacto de M . Los algoritmos de la computadora son capaces de transformar el espectro de m/z en un pico único que proporciona un grado de medición más agudo.

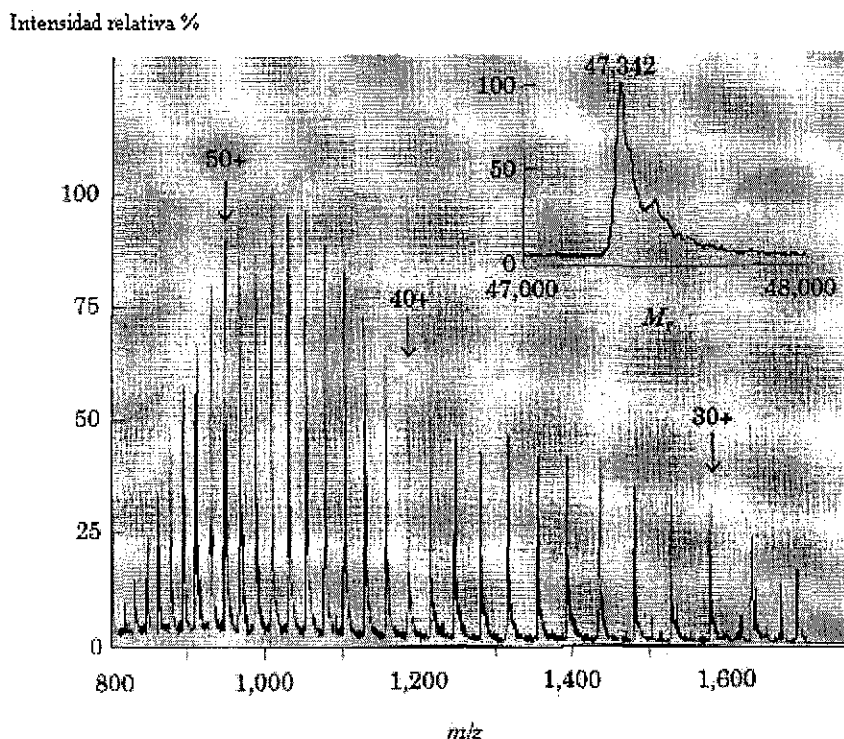


Figura 4. El espectro generado de la medición m/z aparece como una familia de picos, en la que cada pico sucesivo (de derecha a izquierda) corresponde a especies cuya carga se incrementa por 1 tanto en masa como en carga. La transformación computarizada de este espectro se muestra en la figura interior. Tomado de Nelson y Cox, 2000.

La espectrometría de masas de alta resolución también se utiliza para secuenciar pequeños polipéptidos, una aplicación que ha surgido como una herramienta invaluable para la identificación rápida de proteínas desconocidas. La información de la secuencia se obtiene usando una técnica llamada “MS en tandem”, o MS/MS. La proteína que se va analizar mediante este método se somete a la digestión con una proteasa o a un corte hidrolítico con un reactivo químico con el objeto de tener un péptido más corto. La mezcla se inyecta entonces en un aparato que consiste esencialmente en dos espectrómetros en línea (figura 5).

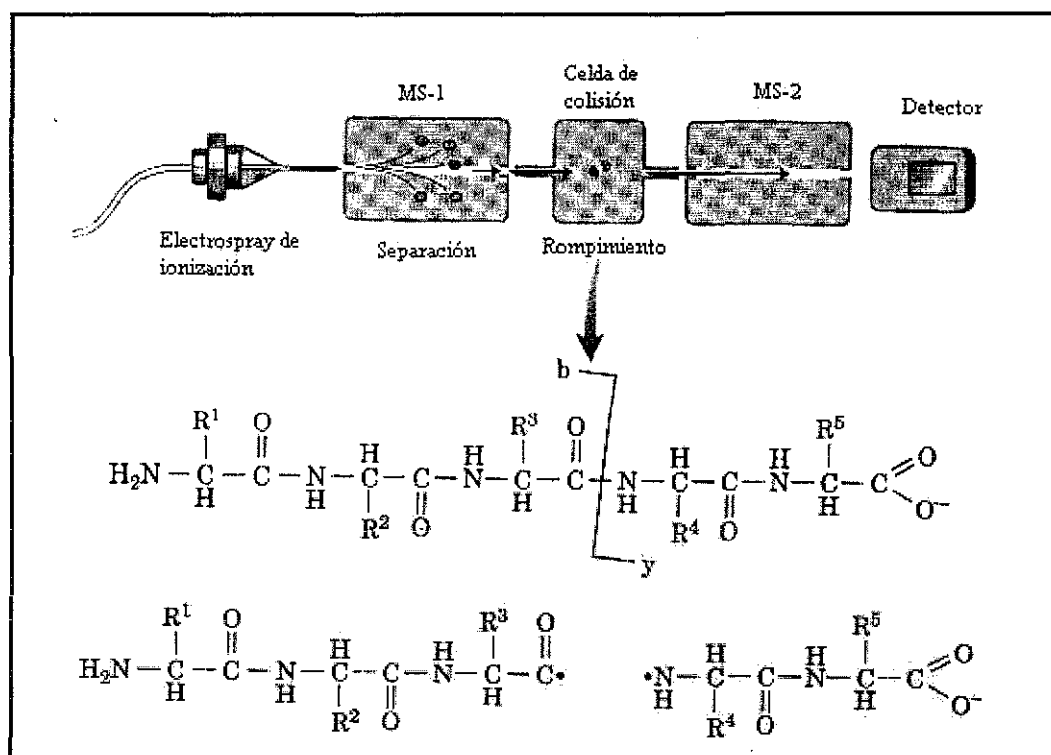


Figura 5. Tandem MS o MS/MS Después de la hidrólisis proteolítica, la solución de la proteína se inyecta en un espectrómetro de masas (MS-1). Los diferentes péptidos son clasificados y solo los de un tipo se seleccionan para ser fragmentados en la cámara entre los dos espectrómetros, la relación m/z de cada fragmento se mide en el segundo espectrómetro de masas (MS-2). Muchos de los iones generados durante la segunda fragmentación son el resultado de la ruptura del enlace peptídico. Estos iones se denominan tipo-b o tipo-y dependiendo de si la carga se retuvo en el extremo amino o el carboxilo. Tomado de Nelson y Cox, 2000

En la primera parte, la mezcla de fragmentos ionizados se manipulan de tal manera que solo uno de los diversos péptidos producidos por el corte salen al otro extremo de la celda. La muestra del péptido seleccionado viaja a través de la cámara de vacío entre los espectrómetros de masas. Posteriormente, el péptido se fragmenta al impactarse con un gas

de colisión, es decir una pequeña cantidad de algún gas noble como el helio o el argón que se ha inyectado en la cámara de vacío. Este procedimiento está diseñado para fragmentar muchas de las moléculas peptídicas en la muestra en la que cada péptido se ha separado o roto en el mismo enlace peptídico. Dado que esta fragmentación no involucra la adición de agua la carga del péptido original se mantiene.

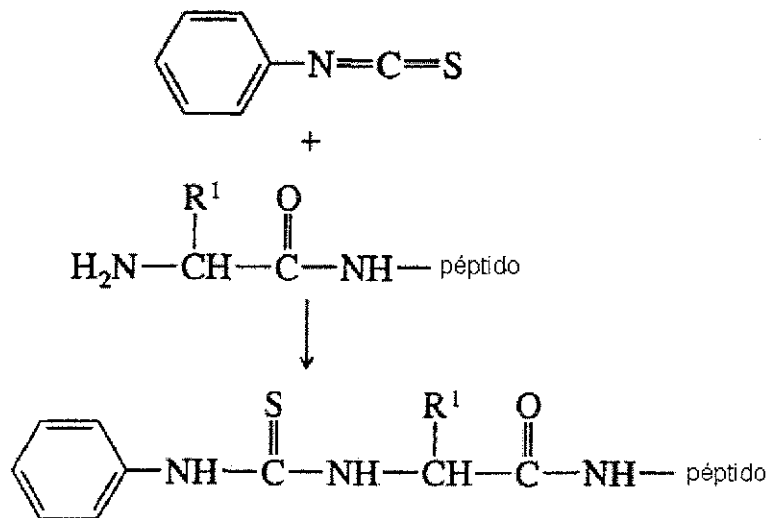
En el segundo espectrómetro se realizan las mediciones de las proporciones m/z de todos los fragmentos cargados, lo cual genera uno o más grupo de picos. Un determinado grupo de picos, consiste de todos los fragmentos cargados que fueron generados al cortar en el mismo tipo de enlace (pero en diferentes lugares del péptido), y derivados del mismo lado en el corte peptídico ya sea por el lado del extremo amino o el extremo carboxi. Cada pico sucesivo en un determinado grupo tiene un aminoácido menos que el pico anterior. La diferencia en masa de pico a pico identifica el aminoácido que se perdió en cada caso, así se obtiene la secuencia del péptido. La única ambigüedad que puede surgir es cuando se trata de leucinas o isoleucinas, las cuales poseen la misma masa *pero pueden identificarse por otras técnicas*.

La carga del péptido puede retenerse ya sea en el fragmento del extremo amino o del extremo carboxi, y otros enlaces peptídicos también se pueden romper durante el proceso de fragmentación resultando así un múltiple grupo de picos. Los grupos más prominentes generalmente consisten de fragmentos cargados derivados del rompimiento de los enlaces peptídicos. El grupo que consiste de fragmentos carboxi terminal puede ser claramente distinguido del grupo de péptidos fragmentados por el extremo amino debido a que los cortes generados entre los dos espectrómetros no rinden grupos completos carboxi y amino en los sitios de ruptura, los únicos grupos α -amino y α -carboxilo intactos en los grupos de los fragmentos peptídicos son aquellos que se encuentran en el mismo extremo (fig. 5). Los dos grupos de fragmentos pueden por lo tanto asignarse por el resultado de ligeras diferencias en la masa. La secuencia aminoácida derivada de un grupo a otro, se puede confirmar por la otra.

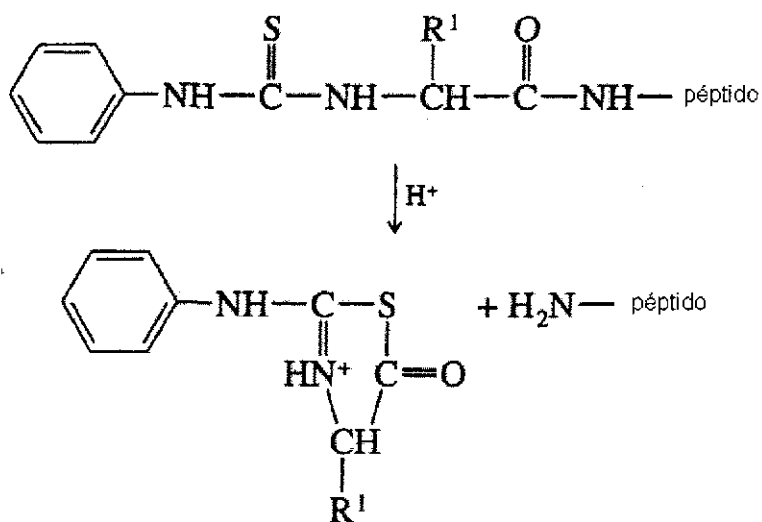
4.4 Secuenciación de aminoácidos mediante la degradación de Edman

Uno de los métodos más exitosos para determinar la secuencia de aminoácidos a partir del extremo amino-terminal es el método de degradación de Edman. Este consiste en remover e identificar aminoácido por aminoácido a partir del extremo amino repitiendo el número de veces necesario para determinar la secuencia de aminoácidos.

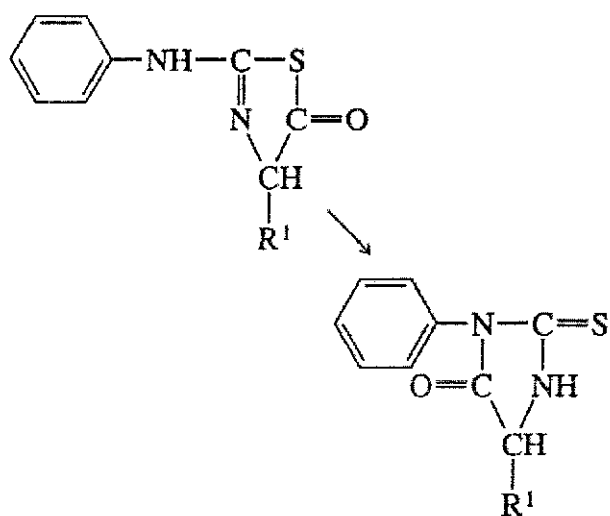
Este procedimiento requiere que el grupo que está en el extremo α -amino terminal se encuentre libre de tal manera que pueda reaccionar con el fenilisotiocianato en un medio alcalino (pH =9) para dar el derivado feniltiocarbamil (PTC):

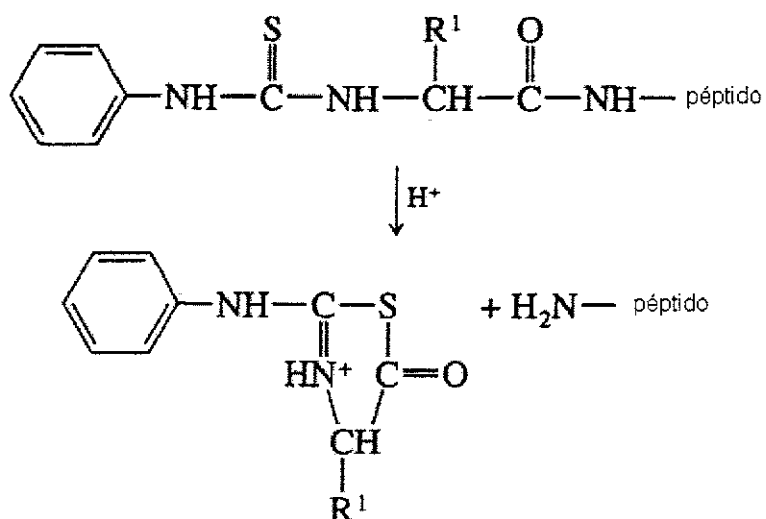


El péptido-PTC se somete entonces a condiciones extremadamente ácidas, como por ejemplo ácido trifluoroacético. La primera unión peptídica, que es menos estable debido a el grupo PTC, se hidroliza bajo condiciones en las cuales las otras uniones peptídicas son estables.



El derivado cíclico del residuo amino terminal liberado, se reacomoda en solución acuosa para formar un derivado de feniltiohidantoina (PTH):





El residuo original del extremo amino se identifica con el aminoácido derivado de PTH mediante HPLC (High Performance Liquid Chromatography). Dentro de las virtudes del método se encuentra la posibilidad de regenerar la cadena peptídica, la cual ha sido alterada únicamente por la pérdida del residuo amino terminal. Este péptido acortado, se encuentra listo para un segundo ciclo con el mismo procedimiento, para identificar el segundo aminoácido que sigue en la secuencia del péptido. No obstante, se pueden presentar algunos errores debido a la presencia de reacciones laterales que impidan la ruptura de las moléculas, o bien que algunos péptidos se disuelvan y pierdan en los solventes orgánicos que se utilizan durante la extracción. La secuenciación automatizada evita estos problemas mediante la fijación de la cadena peptídica de manera covalente a un soporte sólido. No existe un método ideal para hacer la unión pero como las proteínas a menudo se fijan a membranas mediante una electroforesis, en este estado se pueden someter al análisis de secuencia. En los secuenciadores de fase gaseosa la muestra se adhiere en la superficie de una filmina. Algunos de los reactivos se liberan como vapores lo cual minimiza la pérdida del polipéptido y la contaminación de la muestra. En este método se requiere una cantidad mínima de proteína, ya que 5-10 picomoles son suficientes, y la eficiencia del mismo es del 98% por ciclo y permite secuenciar hasta 70 residuos. (Creighton, 1993).

4.5 Dicroísmo Circular

Otra manera de acercarse al conocimiento y caracterización de las proteínas es a través de los estudios espectrofotométricos, los cuales se basan en el cambio de comportamiento de la luz cuando interactúa con las moléculas proteicas. Sin embargo, para entender el principio de cualquier técnica espectrofotométrica es necesario conocer la naturaleza de la luz.

Puede definirse a la luz como una onda que consiste de un campo eléctrico y uno magnético, los cuales son perpendiculares entre sí. De tal forma, la radiación electromagnética presenta una dirección de propagación y propiedades oscilatorias. La energía de la radiación electromagnética se puede definir como:

$$E = hc/\lambda = h\nu,$$

donde E = energía, h =constante de Planck, c = velocidad de la luz,
 λ = longitud de onda y, ν = frecuencia.

Cuando una onda de luz se encuentra con una partícula o molécula, ésta puede ser desviada o absorbida. La absorción de la luz se debe a que la energía del fotón corresponde a las diferencias entre los niveles energéticos de la molécula donde está incidiendo la luz. Así, la energía del fotón lleva a la molécula a un estado de excitación. Por otro lado, las moléculas o sustancias que absorben luz se conocen como **cromóforos**, los cuales presentan espectros de absorción únicos que se definen por las longitudes de onda donde presentan máximos de absorción (λ máx).

Cuando la luz es polarizada, es decir, que todas las ondas presentan un plano común de propagación, puede proporcionar información fina acerca de la estructura de las moléculas. Dos técnicas espectrofotométricas que utilizan la luz polarizada son la Dispersión Rotatoria Óptica (DRO) y el Dicroísmo Circular (DC). En el caso de la DRO lo que se mide es la capacidad de la sustancia de rotar o desviar el plano de propagación, mientras que en el DC se mide la capacidad de absorber diferencialmente luz *circularmente polarizada* a la izquierda y a la derecha. Por lo tanto, para entender el fenómeno del DC es necesario comprender el concepto de luz circularmente polarizada.

Si suponemos que dos haces de luz polarizada se encuentran orientados perpendicularmente uno con respecto al otro, tal como cuando pasan a través de un cristal biaxial y las dimensiones del cristal son tales que provocan un retraso o “diferencia de fase” entre los dos haces de $\frac{1}{4}$ de longitud de onda, entonces la radiación resultante será “circularmente polarizada” (figura 6).

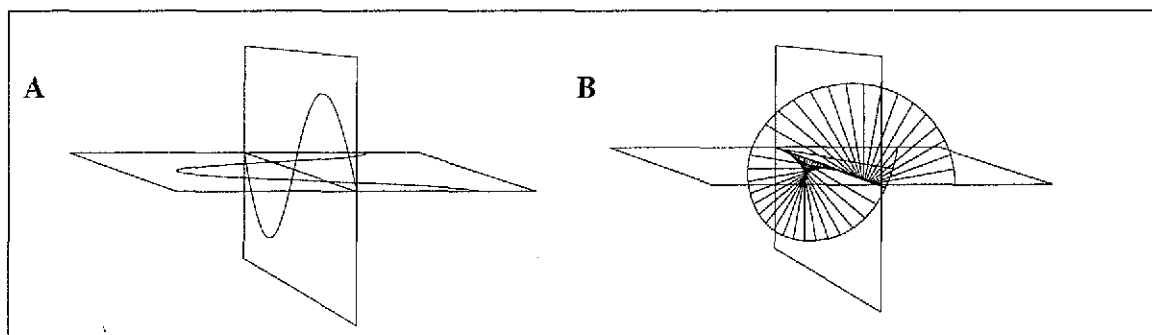


Figura 6: Ejemplo de luz polarizada de manera circular hacia la derecha. La combinación de dos haces perpendicularmente orientados en planos polarizados y con un desfase de $\lambda/4$ (A) genera un haz circularmente polarizado hacia la derecha (B). Tomado de Bayley, 1980.

Dependiendo del signo de retraso será la dirección de la luz emergente, es decir levógira o dextrógira (L o R). Ahora bien, dado que el sentido helicoidal y la quiralidad de una molécula son propiedades intrínsecas a las mismas, independientemente de su orientación molecular o de que se encuentren en movimiento en una solución, entonces es posible que una solución que contenga moléculas asimétricas desvíe y/o absorba de manera diferencial luz polarizada ya sea en sentido levógiro o dextrógiro.

Como se muestra en la figura 7, la suma de 2 haces, uno levógiro y el otro dextrógiro en fase producen una luz plano polarizada ($P_1 + P'_1 = P_0$). La desviación de la luz circularmente polarizada hacia la izquierda y la derecha se define por los coeficientes n_L y n_R respectivamente. Cuando en una solución ópticamente activa hay una desviación del plano polarizado, es decir que $n_L \neq n_R$, entonces la luz polarizada izquierda (LH) y la luz polarizada derecha (LR) se transmiten con diferente velocidad, lo que da por resultado una onda con un plano de polarización rotado con respecto a lo que fue el rayo incidente inicial en la molécula (figura 7) (Bayley, 1980).

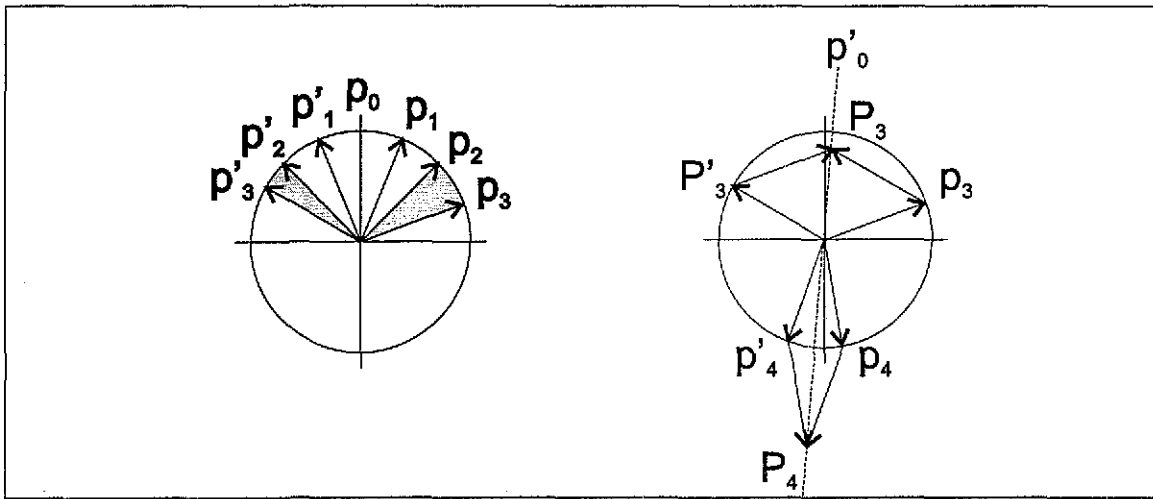


Figura 7. Rotación óptica: en este caso el plano de luz polarizada (p_0) está compuesto de elementos circulares (p_1 y p'_1), los cuales cuando se encuentran dentro de un medio con actividad óptica (entre las posiciones p_2 y p_3) se transmiten con diferentes velocidades. Debido a que $n_L \neq n_R$, entonces la recombinación de p_3 y p'_3 y posteriormente p_4 y p'_4 dan P_3 y P_4 orientadas en la dirección p'_0 y por lo tanto, el plano de polarización ha sido rotado. Tomado de Bayley, 1980.

Si a la desviación de rotación de la luz se le añade la absorción, es decir que en este caso, los coeficientes de absorción molar ϵ_L y ϵ_R son diferentes es decir, $\epsilon_L \neq \epsilon_R$, entonces, la luz emergente ha sido elípticamente polarizada (ver figura 8). Por lo tanto, el dicroísmo circular molar se define como la diferencia de absorción que existe entre ambos coeficientes, es decir $\Delta\epsilon = (\epsilon_L - \epsilon_R)$.

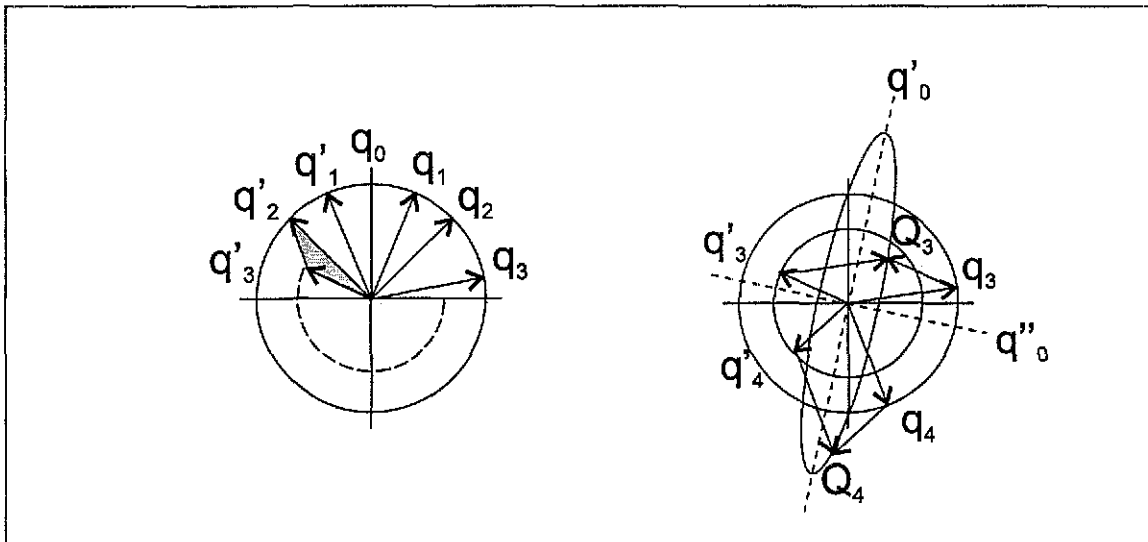


Figura 8. Para el caso del Dicroísmo Circular, los componentes q_1 y q'_1 cuando se encuentran entre las posiciones q_2 y q_3 se absorben en diferente proporción, dado que $\epsilon_L \neq \epsilon_R$; la recombinación de q_3 y q'_3 y posteriormente q_4 y q'_4 dan Q_3 y Q_4 , los cuales se trazan fuera de la elipse con sus ejes mayor y menor en dirección a q'_0 y q''_0 , por lo cual la luz ha sido elípticamente polarizada. Tomado de Bayley, 1980.

Cuando las proteínas se estudian mediante DC, estas muestran señal principalmente en dos intervalos de la región del ultravioleta; en el U.V. lejano, (170-260 nm), los cromóforos son los enlaces peptídicos, ya que aunque como grupos químicos aislados no presentan actividad óptica dado que carecen de centros quirales, la asimetría del medio generada por las conformaciones de la estructura secundaria que adquiere la proteína provoca que un conjunto de enlaces peptídicos se comporte como un centro quiral (Hennessey y Johnson, 1982). En el U.V. cercano (260-320 nm), la señal se debe a la absorción por parte de los grupos aromáticos, que aunque tampoco presentan quiralidad *per se*, su entorno asimétrico en la proteína les proporciona esa actividad (Hooker y Schellman, 1970; mencionado en García-Hernández, E. 1994).

En términos de rotación óptica, la rotación específica $[\alpha]$ de una solución de concentración c_w g ml⁻¹ y una longitud de celda L en decímetros con una rotación observada α_{obs} (en grados) está dada por:

$$[\alpha] = \alpha_{obs}/c_w L.$$

Y de la misma manera, la rotación molar $[\phi]$, está dada por:

$$[\phi] = [\alpha] MW/100$$

donde, MW es el peso molecular y ϕ está expresada en deg cm²dmol⁻¹. En el caso de los biopolímeros el valor de MW se toma como el promedio del peso de los residuos moleculares (MRW) y es igual al peso molecular del polímero dividido por el número de residuos. Entonces, la rotación promedio por residuo $[m]$ se da por:

$$[m] = [\alpha] MRW/100$$

Donde $[m]$ se expresa en grados cm²dmol⁻¹.

La elipticidad, se define como aquél ángulo cuya tangente es el cociente del eje mayor y eje menor de la elipse (ver figura 8). Las unidades de elipticidad son por lo tanto las mismas que las de la rotación óptica de tal manera que la elipticidad específica, la elipticidad molar y la elipticidad promedio por residuo se definen mediante las siguientes ecuaciones:

elipticidad específica

$$[\psi] = \psi_{obs}/c_w L$$

elipticidad molar

$$[\theta] = [\psi] MW/100 \quad y,$$

elipticidad promedio por residuo $[\theta]_{MRW} = [\psi] MRW/100$

La relación entre la elipticidad molar y el dicroísmo circular es:

$$[\theta] = 3300 \Delta\epsilon$$

En la literatura, el dicroísmo circular se reporta generalmente en términos de $[\Theta]_{MRW}$ (y ocasionalmente en términos de ΔA o como miligrados de elipticidad).

Debido a que gran parte de la señal de DC en el U.V. lejano proviene de los enlaces amido, el espectro que se obtiene depende sensiblemente del contenido de estructura secundaria de la proteína (Hennessey y Johnson, 1982). Por lo tanto, existen métodos para determinar el contenido de estructura secundaria dominante a partir del análisis de espectros de DC.

Para analizar la conformación de la proteína en solución mediante DC existen muchos métodos en la literatura, los cuales asumen que el espectro de una proteína se puede representar por una combinación lineal de los espectros de sus elementos estructurales secundarios más efectos no considerados, entre ellos el “ruido” experimental, de tal suerte que lo anterior se puede resumir mediante la siguiente ecuación:

$$\Theta_{\lambda} = \sum F_i S_{\lambda i} + \text{“ruido”}$$

Donde, Θ_{λ} es la señal de DC de la proteína como función de la longitud de onda, F_i es la fracción de cada tipo de elemento de estructura secundaria i , y $S_{\lambda i}$ es la elipticidad a cada longitud de onda de cada i ésimo elemento estructural secundario. En términos estrictos, la suma del peso de todas las fracciones estructurales, $\sum F_i$, debería ser igual a 1.

Para obtener la información estructural de un espectro de DC existen distintos programas, pero en este caso solo se mencionarán los que se utilizaron en este trabajo.

El método de autoconsistencia o **SELCON** (Sreerama y Woody, 1993), consiste en incluir el espectro de la proteína cuya estructura se desconoce en el grupo de 33 espectros de proteínas conocidas, pero restando 3 de los espectros conocidos. Se hace una primera suposición para la estructura desconocida como una primera aproximación. La ecuación matricial se resuelve utilizando el algoritmo de valor único de descomposición

y, la suposición inicial se sustituye por la solución. El proceso se repite restando de 3 en 3 espectros de proteínas conocidas hasta que se llega a una consistencia entre el resultado y la suposición inicial, es decir una autoconsistencia

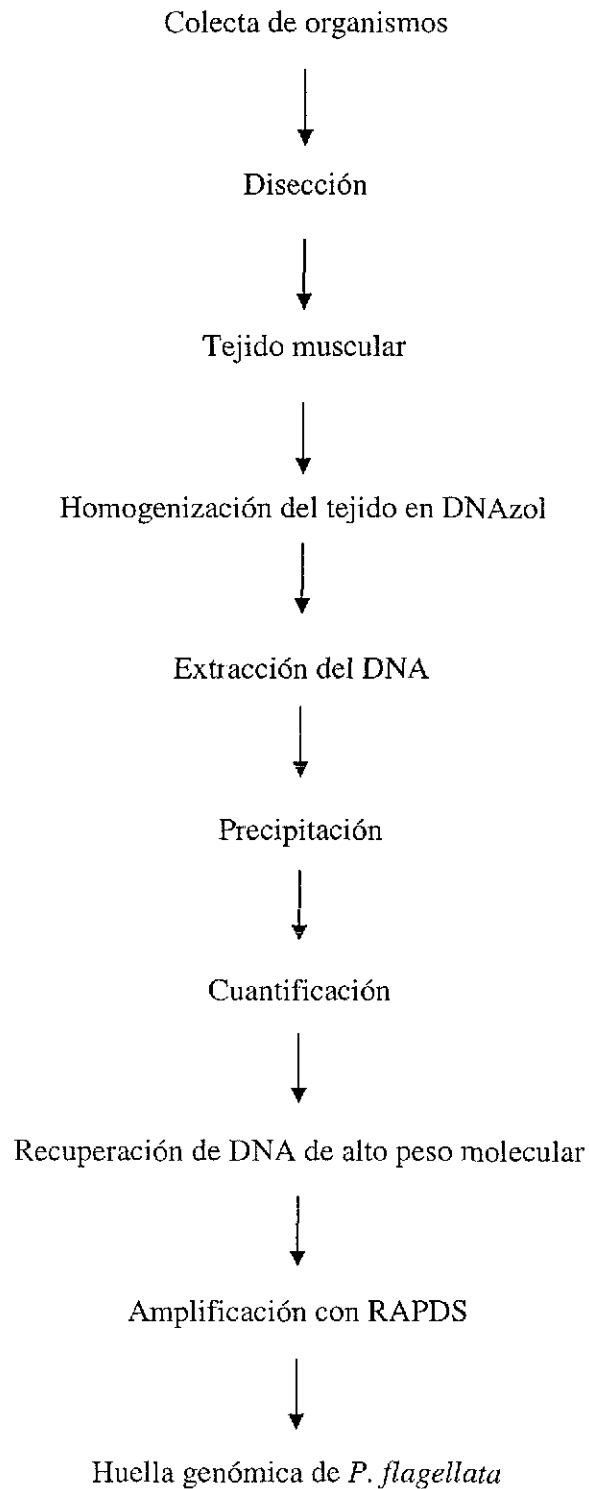
Dada la naturaleza de los datos que se obtienen con el DC, ésta los convierte en una herramienta fundamental para el estudio de las proteínas no sólo como elementos estructurales pasivos, sino como entidades biológicas que pueden sufrir cambios conformacionales de acuerdo a factores físicos, químicos y biológicos como son, variación del pH en el medio, cambios en la temperatura y unión a ligandos o cofactores.

Por otro lado, puede definirse que la descripción del arreglo de los elementos de estructura secundaria de una proteína es su estructura supersecundaria. De tal forma el DC puede acercar al conocimiento de este nivel de organización estructural de las proteínas, tal como lo demostraron Venyaminov y Vassilenko (1994) mediante el programa DEFCLASS. El principio de este programa consiste en que de diferentes proteínas cuya estructura supersecundaria era conocida mediante rayos X, se tomaron sus espectros de 190 a 236nm con intervalos de 2nm los cuales fueron descritos como puntos en un hiperespacio de 24 dimensiones, y los valores de las coordenadas fueron tomados como valores de elipticidad en longitudes de onda fijas. Esto permitió utilizar a los espectros como patrones y posteriormente analizarlos usando algoritmos “patrón” de reconocimiento. Cuando ya se tuvieron suficientes patrones de espectros y sus algoritmos, éstos se llevaron a análisis de clusters o agrupamientos, resultando que la aproximación mediante los clusters fue bastante acertada respecto a la estructura supersecundaria de las proteínas. Dado lo anterior existen métodos que permiten determinar el tipo de estructura supersecundaria a la cual una proteína pertenece a través de tratamientos matemáticos y estadísticos de sus datos de DC. Los tipos de estructura supersecundaria de las proteínas han sido clasificadas en 5 de acuerdo a su tipo de estructura secundaria: aquéllas en las que predomina la estructura α -hélice, denominadas todo- α ; en las que predomina la estructura β son las todo- β , las $\alpha + \beta$ se refieren a aquellas proteínas que tienen estructuras α y β pero, en dominios separados mientras que, las α/β poseen segmentos intercalados de ambas estructuras a lo largo de la cadena, finalmente las proteínas que no presentan un arreglo periódico pero que sin embargo, presentan una estructura “ordenada”(Venyaminov y Yang, 1996). Para cada tipo de

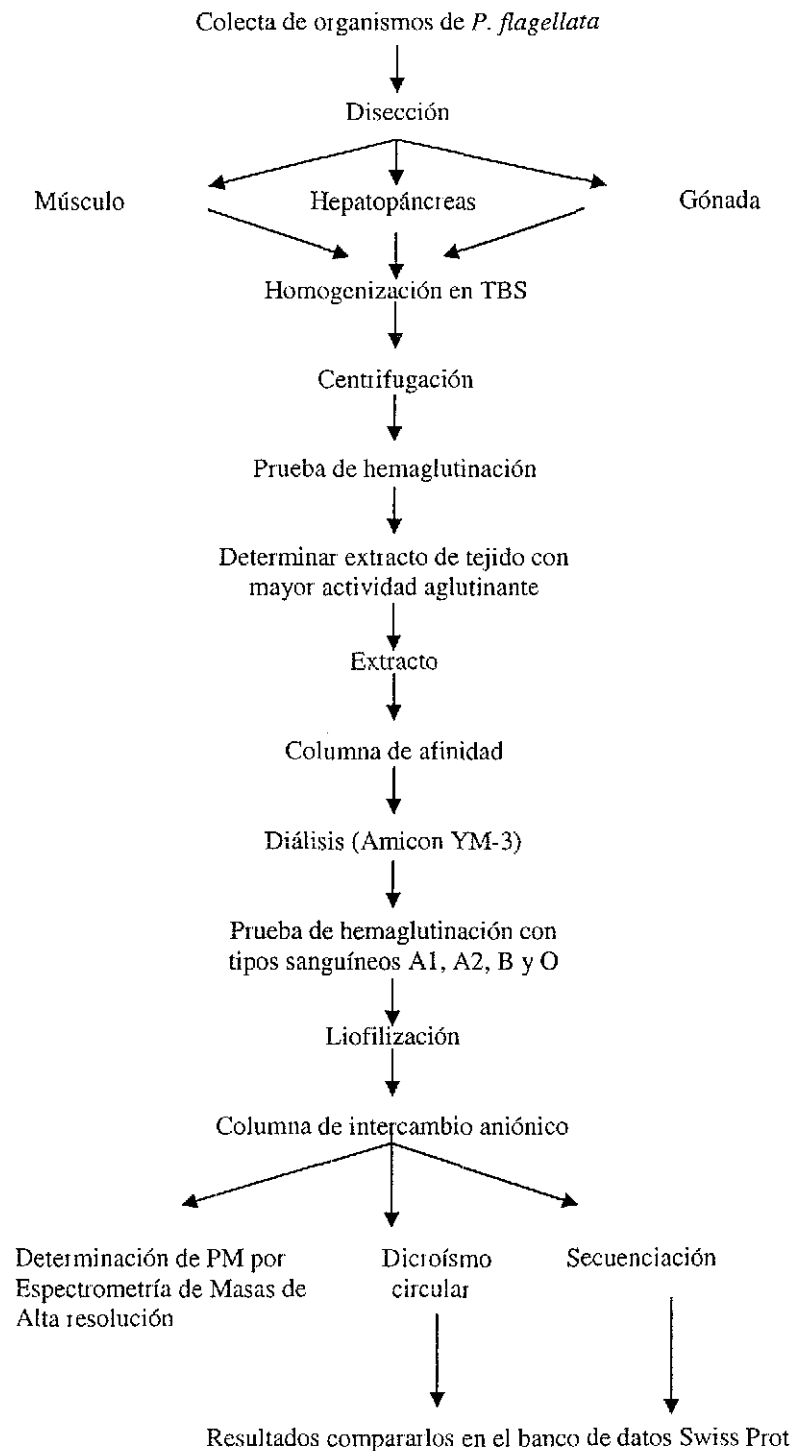
proteína existen espectros característicos que funcionan como patrones que se comparan con los espectros obtenidos de otras proteínas.

5 METODOLOGÍA

5.1 Diagrama de flujo metodológico para determinar la huella genómica de PFA



5.2 Diagrama de flujo metodológico para la purificación y caracterización de PFA



5.3 Extracción de DNA de *P. flagellata*

La extracción de DNA se llevó a cabo utilizando el kit de extracción para DNA DNAzol de LIFE TECHNOLOGIES, y se realizó con algunas modificaciones. El tejido de músculo de *P. flagellata* (0.3g) se homogenizó en un mortero con 0.86ml de solución de DNAzol y se sonicó durante 10 minutos. Posteriormente, el homogenado se centrifugó a $10,000 \times g$ durante 10 minutos a 4°C y se recuperó la fase acuosa. Se precipitó con etanol al 100% en una proporción de 0.5ml de etanol por ml de fase acuosa. Posteriormente, el precipitado se centrifugó a $4,000 \times g$ durante 90 segundos y la pastilla recuperada se lavó dos veces con etanol al 75%. El DNA se solubilizó en una solución de NaOH 8mM y el pH se ajustó a 7.8 con una solución de HEPES.

5.4 Cuantificación y determinación de la pureza del DNA

Como todos los compuestos orgánicos, el DNA presenta un espectro de absorción característico cuya máxima absorción se presenta a los 260nm. De tal forma el coeficiente de extinción a esta longitud de onda es de $E_{260} = 20$, lo que indica que el DNA a una concentración de 1mg/ml tiene una absorción de 20 ($A_{260} = 20$). A una absorción $A_{260} = 2$ la relación entre el DNA y la absorción es lineal, por lo que es posible determinar la concentración del DNA en una solución. Por ejemplo $A_{260} = 0.5$ indica una concentración de $25\mu\text{g/ml}$ y $A_{260} = 0.1$ corresponde a $5\mu\text{g/ml}$.

El uso del factor de conversión en el que $50\mu\text{g/ml} = 1$ es una unidad de absorbancia a 260 nm ($A_{260} = 1$) por lo tanto, es posible determinar fácilmente la concentración de la mayoría de las soluciones de DNA.

Por otra parte, como la presencia de RNA, proteínas, detergentes y solventes orgánicos pueden contribuir en la absorbancia de una solución de DNA, para determinar la pureza se utiliza la relación A_{260} / A_{280} , sabiendo que el máximo de absorción del DNA es a 260nm y para proteínas es a 280nm, se considera que un DNA está puro cuando la relación da aproximadamente 1.6 y, excelente cuando da 1.9.

Finalmente para visualizar la calidad del DNA recuperado, éste se sometió a una electroforesis en un gel de agarosa al 0.8% en TBE 1X (Tris 89mM, ácido bórico 89 mM y EDTA 2mM, pH 8.4), se tiñó con bromuro de etidio y se visualizó bajo luz U.V. en un transiluminador.

5.5 Recuperación de DNA de alto peso molecular (APM)

El DNA de APM se recuperó a partir de un gel de agarosa al 0.8% en TAE1X (Tris 40mM, ácido acético 20mM y EDTA 1mM, pH 8.3) utilizando el kit de extracción de GENE CLEAN II BIO 101. Así, la banda del gel donde se visualizó el DNA de APM se cortó y se introdujo en un tubo Ependorff donde se cortó en pedazos más pequeños. Posteriormente, por cada 0.1g de peso de gel se agregaron 300µl de solución de NaI incluida en el kit, y se incubó a 50°C durante 5 minutos agitando intermitentemente. Posteriormente se incorporó la solución de Glass Milk también incluida en el kit, agitando vigorosamente y se dejó incubar a temperatura ambiente durante 2 minutos, agitando algunas veces. Una vez que la solución de Glass Milk aparecía homogénea en el tubo, se centrifugó 5 segundos y la pastilla se lavó tres veces con la solución de New Wash, incluida en el kit, y se verificó que en el último lavado no quedaran restos del sobrenadante. La pastilla que quedó en el tubo se resuspendió en una solución de TE (Tris-EDTA), y se centrifugó para posteriormente recuperar el DNA en el sobrenadante. Los sobrenadantes recuperados se cuantificaron en el espectrofotómetro para determinar la cantidad de DNA disponible para amplificación.

5.6 Amplificación de ADN de alto peso molecular por RAPDS

La amplificación se llevó a cabo con el kit de **Ready to Go RAPD Análisis Beads** (Pharmacia Biotech), el cual posee 6 primers o iniciadores para la amplificación de distintos *loci*.

Las secuencias de los primers o iniciadores son las siguientes:

RAPD Análisis Primer 1-(5'-d[GGTGCGGGAA]-3')

RAPD Análisis Primer 2-(5'-d[GTTTCGCTCC]-3')

RAPD Análisis Primer 3-(5'-d[GTAGACCCGT]-3')

RAPD Análisis Primer 4-(5'-d[AAGAGCCCGT]-3')

RAPD Análisis Primer 5-(5'-d[AACGCGCAAC]-3')

RAPD Análisis Primer 6-(5'-d[CCCGTCAGCA]-3')

La amplificación se llevó a cabo siguiendo las instrucciones del fabricante. En la mezcla de reacción para PCR se tuvo especial cuidado de mantener los 10ng de DNA de templado que recomienda el fabricante así como **los 25pmol de primers o iniciadores**

para evitar otro tipo de patrón de bandeo. Los ciclos del PCR fueron como sigue: 1 ciclo a 95°C durante 5 minutos, seguido por 45 ciclos de 1 minuto a 95°C, 1 minuto de 36°C y 2 minutos a 72°C. La amplificación con todos los primers se llevó a cabo con el DNA obtenido de 3 organismos.

Las amplificaciones se visualizaron en un gel de acrilamida al 5% (teñido con Bromuro de etidio), el cual sirve para tener una resolución óptima de bandas cuyo tamaño va de 500 a 80pb.

5.7 Extracción de PFA-I

La extracción se realizó de acuerdo a lo reportado en Arreguín-Espinosa y Arreguín-Lozano (1992), sólo que en esta ocasión, después de coleccionar los organismos de *P. flagellata*, sus tejidos se separaron de acuerdo a las partes anatómicas más evidentes. Cada parte se desintegró y homogenizó en una solución amortiguadora de TBS (Tris-HCl pH 7.5; NaCl 50mM). Posteriormente, se centrifugaron a 100,000 \times g durante 30 minutos a 4°C.

Con cada extracto se realizaron pruebas de hemaglutinación con el objeto de determinar si PFA se produce en todo el organismo o es particular de alguna zona anatómica en específico.

5.8 Pruebas de hemaglutinación

La sangre de los donadores de los tipos sanguíneos A₁, A₂, B y O, (obtenida del banco de sangre del IMSS), se coleccionó bajo condiciones estériles usando una solución amortiguadora de citrato de sodio al 3% en TBS. Las células se lavaron 2 veces con TBS mediante un pulso de centrifugación de 3,500 r.p.m. y posteriormente los eritrocitos se diluyeron en una solución amortiguadora de fosfatos a una concentración final de 3%. La actividad de hemaglutinación se llevó a cabo en cajas elisa en las cuales, para cada pozo se agregaron 25µl de solución amortiguadora, y 25µl de solución de extractos crudos (extractos de músculo, hepatopáncreas y gónada) con diluciones sucesivas a partir del primer pozo. Después a todos los pozos se les agregaron 25µl de solución de eritrocitos al 3%. Las cajas de elisa se dejaron reposar durante 4 hrs a temperatura ambiente. En cada prueba de hemaglutinación se utilizó un control positivo (hemaglutininas de esponja) y

un control negativo para tener parámetros de comparación de actividad hemaglutinante. Las pruebas de hemaglutinación se repitieron cinco veces.

5.9 Purificación de PFA

El extracto con resultado positivo de hemaglutinación fue sometido a una purificación mediante cromatografía de afinidad en Sefarosa 4-B activada con ácido, y se eluyó con una solución de D-galactosa 200mM en TBS. Las fracciones con actividad hemaglutinante se concentraron en una membrana de ultrafiltración UK-10 y la fracción retenida posteriormente se dializó contra una solución de Tris-HCl 25mM. La separación de las dos isolectinas se logró en este punto mediante un equipo de HPLC usando una columna de DEAE-Toyopearl 650S previamente equilibrada con Tris-HCl 25mM. La proteína se eluyó en un gradiente lineal de NaCl (0 a 300mM). Posteriormente se concentró por ultrafiltración y se dializó contra agua destilada.

5.10 Espectrometría de masas

Los experimentos de espectrometría de masas se llevaron a cabo en los laboratorios a cargo del Dr. Srukroski en la compañía KRATOS en Austin Texas. Los espectros de masas se adquirieron en un espectrómetro de masas de trampa de iones Esquire, tal como se describe en Jensen, O. N. *et al.* 1998. La PFA-I se disolvió en 500µl de 50:50 metanol/agua, con 0.1% de ácido fórmico. Se determinó el contenido de carbohidratos, tratando la aglutinina con F-endoglucosidasa-N y Glucosidasa-O y volviendo a determinar la masa mediante espectrometría de masas.

5.11 Secuenciación de aminoácidos

La secuencia amino-terminal se determinó a través de la degradación en un secuenciador automático Edman (Mod 477A, Applied Biosystems, USA), equipado con un sistema de HPLC (Mod. 120A on-line phenylthiohydantoin analyzer, Applied Biosystems, USA). El análisis del extremo carboxi se llevó a cabo mediante una digestión progresiva de la PFA con una mezcla de carboxipeptidasas P y Y (Boehringer, Germany) en combinación con un análisis de espectrometría de masas (Bonetto, V., *et al.*

1997). Estos experimentos también se llevaron a cabo en la compañía KRATOS (Austin Texas).

5.12 Caracterización estructural mediante dicroísmo circular.

Los espectros de DC se registraron en un espectropolarímetro Jasco J-720, el cual fue previamente calibrado con ácido (+)-camforsulfónico. Los espectros en el UV lejano (180-260nm) se obtuvieron de soluciones de proteína de concentración 0.2mg/ml, usando una celda de 2mm de recorrido óptico. Para el UV cercano (260-320 nm), la concentración de proteína fué de 1mg/ml, usando una celda de 1 cm de recorrido óptico. La elipticidad media residual, Φ_{mrw} , fué calculada usando un peso molecular promedio por residuo de 110g/mol.

Los espectros de DC en el UV lejano se analizaron mediante el programa de deconvolución SELCON (Sreerama & Woody 1993) para estimar el contenido de estructura secundaria de PFA. Adicionalmente, los espectros se analizaron mediante el programa DEFCLASS para definir la clase estructural terciaria de la proteína.

6 RESULTADOS Y DISCUSIÓN

6.1 Colecta

Durante la colecta de los organismos de *P. flagellata* en la laguna de Coba, en Campeche, se observó que los organismos se localizan en la orilla de la laguna, en los lugares ocultos al sol y que en su mayoría están adheridos a la superficie sumergida y oculta de rocas así como en la raíz o tallos sumergidos de los pastos que se encuentran en esta laguna. La colecta se realizó en el mes de septiembre pero al parecer la mejor época de colecta es en los meses de junio, julio y agosto ya que es temporada de lluvias y la población de organismos es mayor.



Figura 9. Imágen de *P. flagellata* después de la colecta y antes de la disección

TESIS CON
FALLA DE ORIGEN

6.2 Extracción y recuperación de DNA

Después de extraer el DNA de 12 organismos de *P. flagellata*, se procedió a determinar su concentración, recuperación y pureza. A continuación en la siguiente tabla se presentan los datos de recuperación, concentración y pureza de DNA para cada uno de los organismos que se trabajaron.

Tabla 4: Datos de recuperación de DNA obtenido del pie o músculo de *P. flagellata*.

Organismo	Recuperación total de DNA (mg)	Recuperación de DNA por gramo de tejido (mg/g)	Concentración DNA (mg/ml)	Pureza (260/280)
1	0.36	0.9	1.2	1.96
2	0.315	1.05	1.05	1.9
3	0.218	0.054	1.09	1.88
4	0.292	0.97	1.46	1.9
5	0.088	0.22	0.44	1.7
6	0.05	0.1	0.25	1.8
7	0.297	0.99	0.99	1.7
8	0.052	0.26	0.26	1.6
9	0.046	0.46	0.23	1.5
10	0.105	0.35	0.35	1.75
11	0.108	0.27	0.36	1.8
12	0.105	0.35	0.35	1.75

Como puede verse en la tabla anterior, la recuperación de DNA fue satisfactoria, y a pesar de las modificaciones que se realizaron al método original, la pureza y cantidad recuperada de DNA fue aceptable. Posteriormente, para verificar que el DNA se encontraba en condiciones para llevar a cabo la amplificación, éste se sometió a una electroforesis en un gel de agarosa al 0.8% y se tiñó con bromuro de etidio, con el objetivo de visualizarlo. En la figura 10, se puede observar que aunque la cantidad de DNA en el gel coincide con las cuantificaciones realizadas en el espectrofotómetro, éste se encuentra en todos los casos degradado, es decir, no se visualiza como una sola banda de alto peso molecular (APM), sino que además de esta banda, muy notoria en algunos carriles, se observa un barrido hacia abajo en todos ellos, lo cual indica que parte del DNA sufrió degradación.

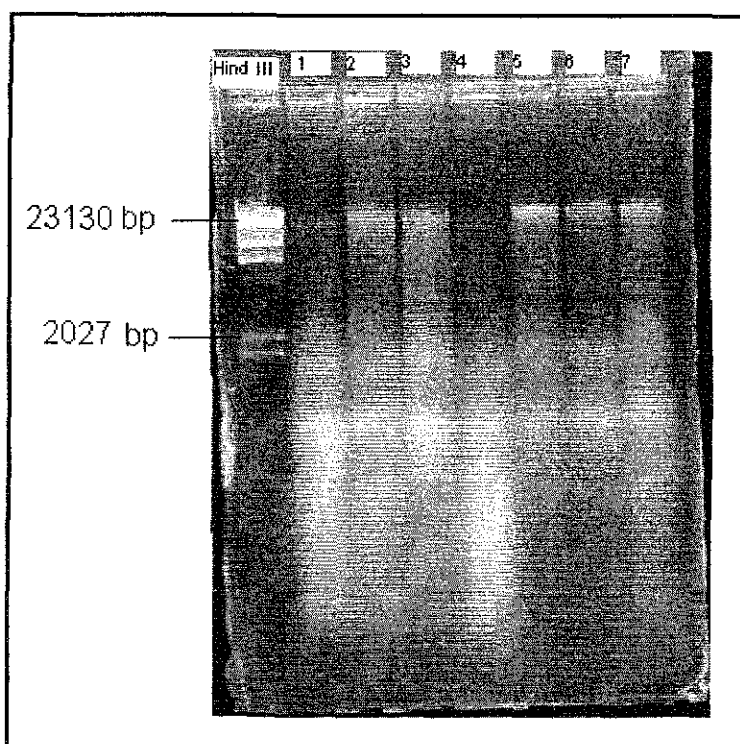


Figura 10. Gel de agarosa al 0.8% en TBE 1X. Carril 1, marcador de PM λ Hind III; carriles 2 al 8 DNA de los organismos 1, 2, 3, 4, 5, 6, y 7 respectivamente. Se observa una banda de APM de aproximadamente 20 Kb en todos los carriles, dicha banda corresponde a DNA no degradado. Sin embargo en todos los carriles también se observa DNA degradado.

Es posible que durante la extracción del DNA el manejo mecánico de los tejidos o la sonicación haya influido en la degradación del mismo. Por otra parte, aunque todo el material que se utilizó para la extracción fue esterilizado, la acción de DNAasas presentes en el mismo tejido (se sugiere que este tipo de moluscos poseen gran cantidad de DNAasas en sus tejidos, comunicación personal), pudo contribuir a la degradación del DNA recuperado.

Debido a que para realizar amplificaciones con RAPDS se requiere que el DNA genómico se encuentre íntegro con el fin de evitar falsas amplificaciones o resultados ambiguos, fue necesario recuperar o aislar el DNA de APM mediante el kit de extracción de GENE CLEAN II BIO 101. Sin embargo, la cantidad de DNA recuperado para cada muestra mediante esta técnica disminuyó en un factor de 10. Aunque esto no fue un obstáculo para llevar a cabo la amplificación, sería conveniente optimizar el método de

TESIS CON
FALLA DE ORIGEN

extracción para disminuir la degradación de DNA. En este caso se propone el uso de EDTA para eliminar la acción de DNAasas, aumentar la concentración de inhibidores de estas proteínas, o bien llevar a cabo otro método de recuperación de DNA de APM.

6.3 Amplificación con RAPDS

Una vez que terminaron los ciclos de amplificación, las muestras se sometieron a una electroforesis en un gel de acrilamida que posteriormente se tiñó con bromuro de etidio, se visualizó bajo luz UV y se procedió a tomar la fotografía (Figura 11).

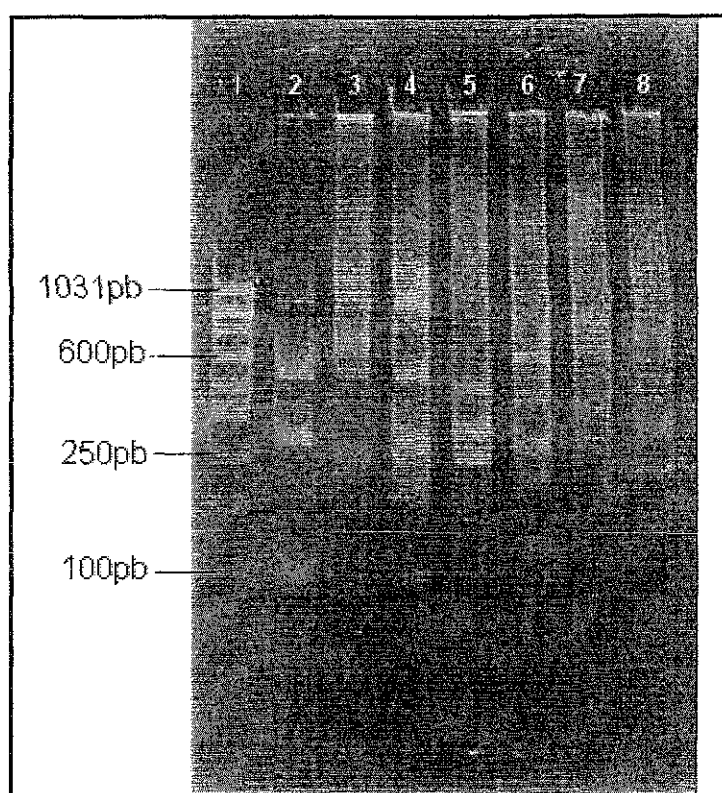


Figura 11. Gel de acrilamida para DNA al 5%. Carril 1, marcador de peso molecular de 50pb; carril 2 control positivo de amplificación con DNA de *E. coli* BL21; carril 3, amplificación de DNA de *P. flagellata* con el primer 1; carril 4, amplificación de DNA de *P. flagellata* con el primer 2; carril 5, amplificación de DNA de *P. flagellata* con el primer 3; carril 6, amplificación de DNA de *P. flagellata* con el primer 4; carril 7, amplificación de DNA de *P. flagellata* con el primer 5, carril 8, amplificación de DNA de *P. flagellata* con el primer 6; carril 9, control negativo de amplificación.

Para determinar el tamaño de las bandas características de *P. flagellata* con estos marcadores genéticos, la imagen de los RAPD's se analizó contra un marcador de 50 pb (Fermentas) en un densitómetro Bio-rad, y los resultados fueron los siguientes. En el primer carril se aplicó el marcador de peso molecular mientras que en el segundo carril se encuentra el control positivo con DNA de la cepa de *E. coli* BL21 amplificada con el primer 2 para comparar lo reportado con la bibliografía y así determinar que las condiciones de amplificación fueron las adecuadas. A partir del carril 3 se trata de amplificación del DNA de APM *P. flagellata* con el primer o iniciador 1, y en el se obtienen 6 bandas visibles; para el carril 4 es la amplificación del DNA de *P. flagellata* con el primer 2 y se obtuvieron 10 bandas, en el carril 5 es la amplificación con el primer 3 obteniéndose sólo 3 bandas, en el carril 6 es la amplificación con el primer 4 y da por resultado 5 bandas, en el carril 7 con el primer 5 se obtienen 5 bandas y, en el carril con el primer 6 se obtienen 6 bandas, finalmente en el carril 8 se aplicó el control negativo con el primer 2, y dado que no aparecen bandas en dicho carril entonces se confirma que no existió la posibilidad de contaminación durante el proceso de amplificación. De acuerdo al análisis realizado con el densitómetro (ver tabla 5), se puede ver que para cada primer o iniciador, la amplificación del DNA de *P. flagellata* dio no sólo distinto número de "bandas marcadoras" sino que cada una posee un peso molecular distintivo, el cual la identifica respecto a cualquier banda obtenida con los otros primers. Además, hay que resaltar que las bandas de este experimento se repitieron en las amplificaciones con los mismos primers, llevadas a cabo con el DNA de otros dos organismos de *P. flagellata*. Lo anterior confirma la utilidad que los RAPD's tienen como marcadores genéticos, ya que mediante amplificaciones con iniciadores arbitrarios dan amplificaciones características de organismos pertenecientes a la misma especie. Hay que aclarar, sin embargo, que las amplificaciones se llevaron a cabo **con las mismas condiciones** de temperatura tiempo, número de ciclos y cantidad de primers utilizados para cada reacción de PCR (25 pmol), ya que si se varía algo en estos parámetros, entonces las bandas resultantes son distintas. Las condiciones óptimas utilizadas en este trabajo son las que reportan y recomiendan utilizar los laboratorios de Pharmacia Biotech.

El uso de los RAPD's confirmó que los diferentes individuos de caracol utilizados para este estudio pertenecen a la misma especie y por lo tanto, proporciona las bases moleculares para identificar organismos de nuevas colectas, lo cual es importante para identificar al organismo con el que se trabajó puesto que no se basa únicamente en disección y características anatómicas sino que se trata de pruebas de tipo genético moleculares.

Otra ventaja de los RAPD's para este caso es que si se desea volver a extraer PFA y no se está seguro de la identidad taxonómica de los organismos bastará con sólo utilizar uno de los primers y comparar el patrón de bandas obtenido con el patrón de bandas de referencia de *P. flagellata*.

Tabla 5. Análisis de resultados de amplificación mediante densitómetro. Dentro de los cuadros de determina el peso molecular de las bandas en pb en orden descendente

Carril 1	Carril 2	Carril 3	Carril 4	Carril 5	Carril 6	Carril 7	Carril 8	Carril 9
1031	-	1173	3405.5	-	-	-	1256.8	-
850	835.8	1001.6	2055.65	-	-	-	844.35	-
750	-	-	1006.75	-	-	-	-	-
600	-	686.8	812.7	-	-	617	586.4	-
500	406.2	516.2	750.9	-	564.4	498	-	-
350	-	475.8	493	343.3	365	465	365	-
250	-	-	-	277.1	-	336.4	281.4	-
200	212	219.2	205.1	173	188	210.4	-	-
150	-	-	169	-	146.1	-	146.5	-
100	-	-	130	-	75	-	-	-
50	56.5	-	56.46	-	-	-	-	-

6.4 Pruebas de hemaglutinación

Las pruebas de hemaglutinación se realizaron con extractos de tejidos de músculo, hepatopáncreas y gónada, para determinar la distribución y localización de PFA en el organismo. En la tabla 6 es posible observar que la actividad de hemaglutinación es mayor para los 4 tipos sanguíneos cuando el extracto es de músculo, en hepatopáncreas la actividad es mínima, mientras que en gónada se presenta actividad hemolítica. Lo anterior indica una distribución específica de PFA en el caracol *P. flagellata*. Las unidades UHA (Unidades de Hemaglutinación) representan la mínima dilución del extracto hasta donde es posible observar actividad hemaglutinante. Por lo tanto, es posible determinar que la hemaglutinación es mayor en los tipos sanguíneos A₁ y A₂, lo cual indica una afinidad con mayor selectividad hacia los grupos N-acetilgalactosamina.

Tabla 6. Comparación de la actividad hemaglutinante de los distintos extractos de *P. flagellata* con los 4 tipos sanguíneos.

Tipo sanguíneo	UHA de extracto de tejido muscular	UHA de extracto de tejido de hepatopáncreas	UHA de extracto de tejido de gónada
A ₁	341	64	Nulo (hemólisis 1,280 HU)
A ₂	363	80	Nulo (hemólisis 1,280 HU)
B	176	96	Nulo (hemólisis 1,280 HU)
O	128	112	Nulo (hemólisis 1,280 HU)

6.5 Fraccionamiento mediante columna de afinidad

Una vez que se determinó que la mayor actividad hemaglutinante se encuentra en músculo, el extracto se sometió a una columna de afinidad de Sefarosa 4-B tratada con ácido y se eluyó con soluciones cuya concentración de D-Galactosa aumentó de 70mM, a 140mM y terminó con 200mM. Se recolectaron fracciones de 1ml y posteriormente se hicieron lecturas en el espectrofotómetro a 280nm para determinar las fracciones donde existía mayor concentración de proteína. En la gráfica de la figura 12 se muestran con puntos las fracciones recolectadas y sus correspondientes absorbancias

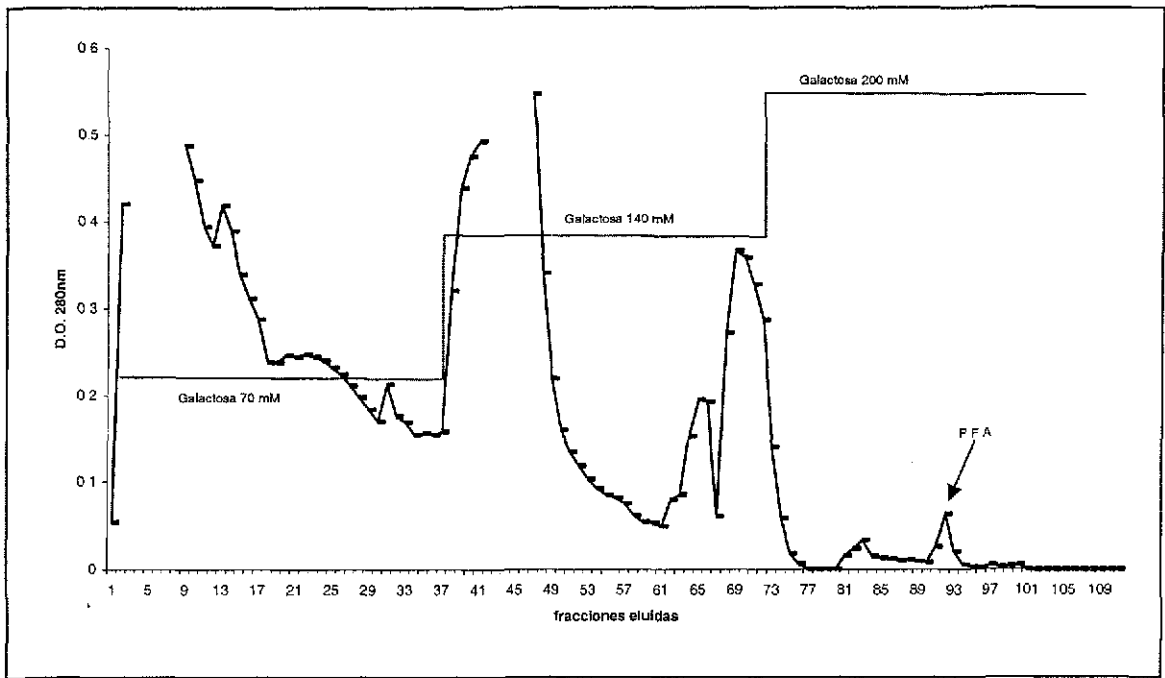


Figura 12. En esta gráfica en el eje de las abscisas se encuentran las fracciones de 1ml que fueron colectadas de la columna de afinidad, en el eje de las ordenadas la absorbancia registrada a 280nm. La recuperación de PFA se obtuvo en las fracciones marcadas como 90 a 94 las cuales se unieron en una misma fracción. El pico de absorbancia correspondiente a PFA es muy pequeño con respecto al resto de proteína total, quizá debido a que en el músculo hay gran cantidad de proteínas, no solo las estructurales sino también las funcionales como las que están involucradas en el movimiento. Así la recuperación y parcial purificación de PFA se puede considerar como exitosa ya que proviene de un extracto crudo con una gran cantidad de proteínas

Para cada pico correspondiente a fracciones de elución se dializaron las muestras y se procedió a determinar en cuál de ellas existía actividad aglutinante. En la siguiente tabla se puede observar que fue en las dos últimas fracciones de recolección donde se observó actividad aglutinante. Por lo anterior, es posible decir que estas aglutininas muestran el más alto grado de afinidad, dado que eluyen en las concentraciones más altas de D-galactosa. Hay que notar que los dos últimos picos de absorbancia son muy pequeños con respecto a los otros y que el último es aquél en donde se encontró la mayor concentración de PFA correspondiendo a un 0.23% del total de la proteína, es decir, $\sim 70\mu\text{g}$ de proteína. Por lo cual el último pico se recirculó para tener la proteína "más pura". Con respecto a los otros picos de elución se puede decir que éstos también presentan afinidad por la D-galactosa (siempre menor a la de PFA). Sin embargo, no poseen actividad hemaglutinante probablemente porque presentan un solo sitio de reconocimiento a la D-galactosa o bien no son proteínas oligoméricas.

Tabla 7. Actividad aglutinante de las fracciones colectadas a partir de la columna de afinidad

Tipo Sanguíneo	UHA frac 92	UHA frac 82	UHA frac 71	UHA frac 65	UHA frac 44	UHA frac 40	UHA frac 33	UHA frac 25	UHA frac 7
A ₁	320	-	-	-	-	-	-	-	-
A ₂	160	160	-	-	-	-	-	-	-
B	160	-	-	-	-	-	-	-	-
O	-	-	-	-	-	-	-	-	-

Por otro lado, al realizar las pruebas de hemaglutinación con las distintas fracciones de la columna se observó que en la fracción 92 es donde se presentó mayor capacidad aglutinante. En esta fracción la afinidad es mayor por los grupos sanguíneos A₁ y A₂ que por el grupo B, mientras que por el grupo sanguíneo O no existe actividad aglutinante. De acuerdo con la figura 13 y los resultados presentados en la tabla 7 se puede decir que la afinidad superior de la proteína por los grupos sanguíneos A con respecto a los grupos B se debe a que en los carbohidratos de membrana de estos eritrocitos se encuentra un grupo acetil unido a el carbono 2 de la galactosa externa mientras que, en los grupos sanguíneos B el carbono 2 de la galactosa externa posee un grupo hidroxilo. Así mismo, se puede observar que la afinidad en los grupos A es diferente, de tal forma que en A₁ es mayor que en A₂, esto se debe quizá a que el sitio de reconocimiento de la lectina es un sitio extendido y por lo mismo, es posible que la estructura del trisacárido influya en la afinidad de reconocimiento. De acuerdo a lo anterior, es posible sugerir que la diferencia de afinidad se debe a la estructura del trisacárido, donde en A₁ la segunda galactosa se une al carbono 3 de de la glucosa, mientras que en A₂ la galactosa se une al carbono 4 de la glucosa, influyendo este cambio de configuración en el reconocimiento a la proteína.

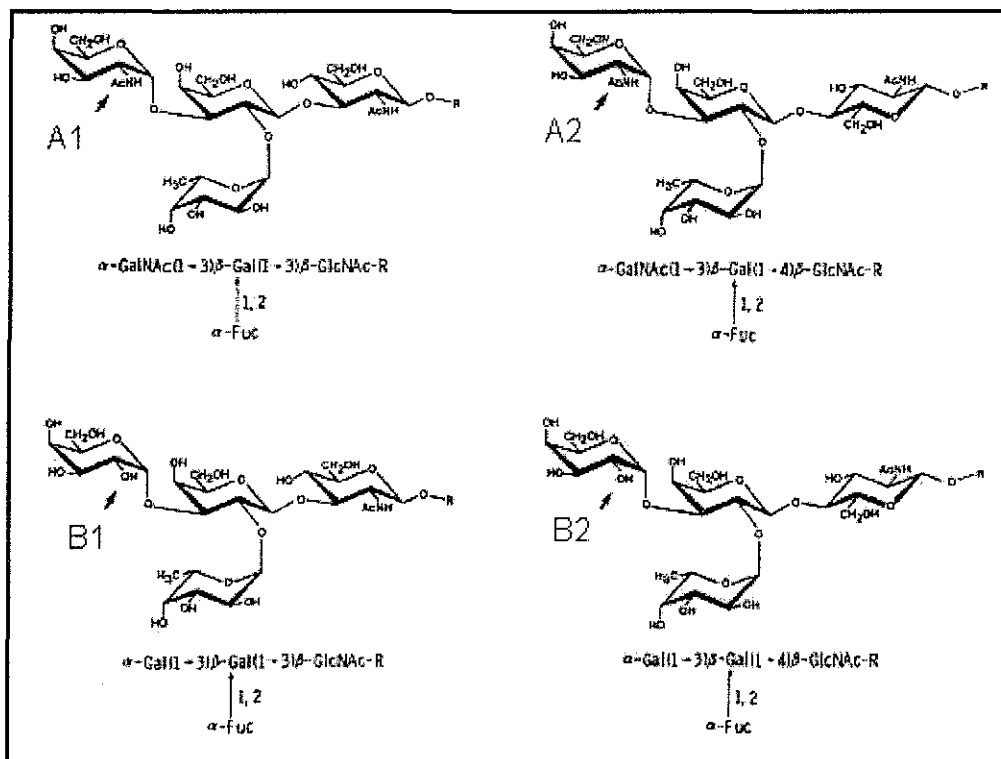


Figura 13. Estructura química de los determinantes A y B basados en las cadenas de carbohidratos Tipo-1 y Tipo-2. Tomado de Watkins *et. al.* (1981)

De acuerdo a la tabla 8 de especificidad por azúcar (reportada anteriormente por Arreguín-Espinosa y Arreguín-Lozano (1997), se observa que la galactosa en sí es una estructura molecular propicia para la unión de PFA, pero que dicha unión se potencia cuando se trata de N-Acetil-galactosamina. Así, la galactosa presenta un factor de inhibición de 1 a una concentración de 200mM mientras que la N-acetil-galactosamina presenta un factor de inhibición de 4 a una concentración de 50 mM. Lo anterior es consistente con los resultados en los grupos sanguíneos A y B los cuales presentan aglutinación en presencia de PFA pero con diferente intensidad. También es posible observar que cuando se trata de polisacáridos aumenta la afinidad de la proteína por éstos.

Aunque la recuperación de PFA mediante la columna de afinidad fue aceptable, quizá sería recomendable realizar ensayos de elución con otros azúcares tales como los β -galactósidos y la N-acetil-galactosamina, de tal manera que permitieran una recuperación más rápida y específica de la PFA.

TESIS CON
FALLA DE ORIGEN

Tabla 8. Especificidad por azúcar de PFA*

Inhibidor	Concentración (mM)	Factor inhibitorio ¹	Inhibidor	Concentración (mM)	Factor inhibitorio ¹
Galactosa	200	1	Lactulosa	200	1
Lactosa	100	2	NeuAc	NI	-
Melibiosa	100	2	Ovoalbumina	NI	-
Lactulosa	100	2	Transferrina	0.02	10,000
Sialolactosa (α 2,3)	100	2	Orosomucoide	0.001	200,000
Sialolactosa (α 2,6)	100	2	Asialo-orosomucoide	0.0001	2,000,000
GalNAc	50	4	Asialo-agalacto-orosomucoide	NI	-
Galactosamina	100	2	Fetuina	0.0005	400,000
Fenil- α -D-galactosa	50	4	Asialo fetuina	0.00002	10,000,000
Fenil- β -D-galactosa	5	40	Asialo-agalacto-fetuina	0.002	200,000
Metil- α -D-galactosa	50	4	mucina	0.00001	200,000,000
Metil- β -D-galactosa	5	40			

(Tomado de Arreguín-Espinoza y Arreguín Lozano, 1997)

*Concentración mínima necesaria para inhibir 4 dosis de hemaglutinación en eritrocitos de conejo.

¹A concentración de 10 mg/ml. NI=No inhibió a una concentración de 200mM para azúcares simple y 10 μ M en presencia de glicoproteínas. Los azúcares simples sin efecto inhibitorio a 200mM son: D-glucosa, D y L-fucosa, GlcNAc, D-Manosa, D-arabinosa, D-fructosa, glucósidos (α y β). Tampoco el ácido colomínico a 10mg/ml fue inhibidor.

6.6 Separación de PFA-I y PFA-II mediante intercambio iónico. Después de concentrar y recircular la proteína por la columna de afinidad, esta se sometió a una separación mediante un HPLC y una columna de intercambio iónico. Puede observarse que ambas proteínas se encuentran muy cerca, siendo el pico más alto el correspondiente a PFA-I y el menor a PFA-II. Se observa que PFA-I comienza a eluir al 75 mM de NaCl mientras que PFA-II comienza a fluir al 140mM, cuando PFA-I ya terminó de eluir.

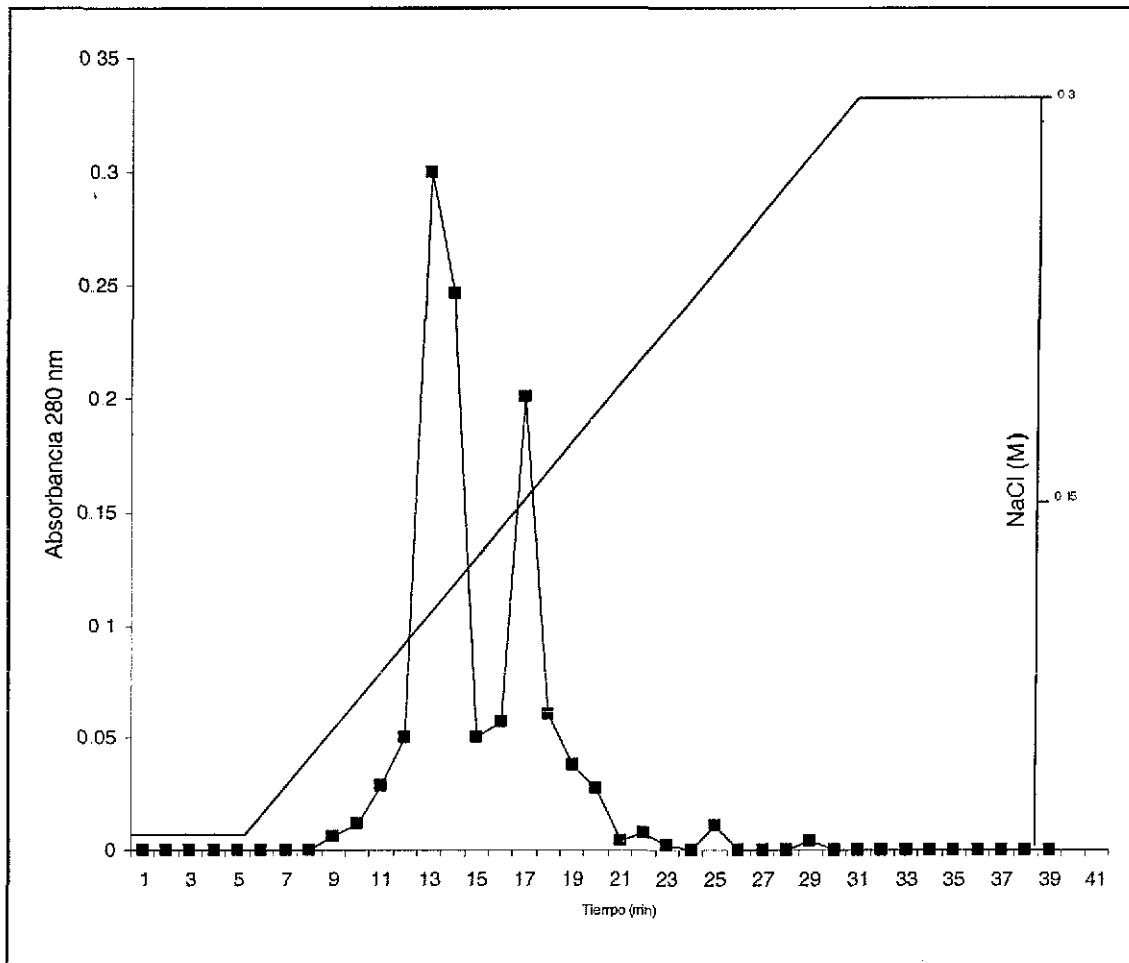


Figura 14. Purificación de PFA-I y PFA-II mediante cromatografía de intercambio iónico. La fracción de proteínas con actividad hemaglutinante se efluyeron en una columna aniónica con un gradiente lineal de NaCl (0-300mM). El primer pico corresponde a PFA-I mientras que el según pico corresponde a PFA-II.

TESIS CON
FALLA DE ORIGEN

6.7 Espectrometría de masas de PFA-I

A partir de este momento los resultados que se mencionan se refieren únicamente a PFA-I. De acuerdo al trabajo de Arreguín-Espinosa y Arreguín-Lozano (1997), se había determinado mediante electroforesis en geles de poliacrilamida de SDS que el peso molecular de la proteína era de 30, 000 Da. Sin embargo, de acuerdo a las mediciones mediante espectrometría de masas el peso de PFA-I corresponde a 32 444 Da es decir, 2 kDa más grande que lo que se había determinado antes. Cuando la proteína se sometió a digestión con las glicosidasas F,N-glicosidasa y O-glicosidasa y se volvió a determinar su peso molecular mediante masas, éste resultó ser de 30, 040 Da. De acuerdo a lo anterior, se deriva que 7.41% del peso total de la proteína intacta corresponde a carbohidratos. Con este nuevo dato, y utilizando los porcentajes de composición de aminoácidos reportados previamente (Arreguín-Espinosa y Arreguín-Lozano 1997), se recalculó el número total de aminoácidos como 273, es decir 13 aminoácidos más que los que se habían calculado mediante PAGE-SDS.

6.8 Secuenciación de aminoácidos de PFA-I

En el trabajo de Arreguín-Espinosa y Arreguín-Lozano (1997) se determinó la secuencia de 12 aminoácidos en el extremo amino. En este trabajo se presenta la secuencia de 43 aminoácidos del extremo amino mediante la degradación de Edman y 11 del extremo carboxilo mediante digestión progresiva con carboxipeptidasas P y Y. Los experimentos se realizaron dos veces dando los mismos resultados y coincidiendo con el resultado anterior. Por lo tanto, la secuencia de PFA obtenida hasta el momento es la siguiente:

PFA-I L-AKT-N-EP LNFP IHVRDA GVQWSP LGRL TTGADV RHEI PVLSNR 43
263 ARVGHRLARWA 273

Figura 15. Secuencia parcial de aminoácidos obtenida de PFA. En azul se muestran los residuos pertenecientes al extremo amino (1-43) y en negro los pertenecientes al extremo carboxilo (263-273)

TESIS CON
FALLA DE ORIGEN

Durante el proceso de secuenciación no se encontró que alguno de los extremos de la cadena peptídica estuviera químicamente bloqueado. Posteriormente, se buscó si existía similitud de secuencia con alguna otra proteína reportada, y ambos segmentos de PFA fueron comparados contra aproximadamente 575,000 secuencias peptídicas del banco de información SWISS-PROT utilizando el programa BLAST (Altschul, *et al.*, 1997). De acuerdo a este análisis, se encontró que PFA y la ricina, la cual es una citotoxina que proviene de las semillas de la planta del castor, son homólogas en su amino terminal. Cabe aclarar que ninguna otra proteína presentó una alineación o similitud de secuencia con la PFA bajo los estándares de BLAST.

De acuerdo a Peumans y Van Damme (1995), la ricina es un claro ejemplo de quimerolectinas, es decir, proteínas constituidas por un dominio de reconocimiento al carbohidrato más otro dominio que puede tener actividad catalítica (u otra actividad) y que funciona de manera independiente al dominio que reconoce al carbohidrato. Así, la ricina es una N-glicosidasa heterodimérica que está compuesta de dos fracciones proteicas, una cadena A denominada RA que consiste de 267 residuos aminoácidos que se encuentra unida por un puente disulfuro a la cadena B, denominada RB, la cual es una cadena de 262 residuos. RB es una lectina específica para galactosa que cuando se une a la membrana de células eucarióticas permite la interiorización de RA. Una vez que RA se encuentra dentro de la célula cataliza la hidrólisis de la unión N-glicosídica de una adenosina específica que se encuentra en un asa altamente conservada en el RNA ribosomal 28S y, de esta forma, la síntesis de proteínas se inhibe completamente. El mecanismo de toxicidad de RA, es a su vez compartido por otras citotoxinas que se han denominado RIP's (de las siglas en inglés Ribosome Inactivating Proteins), las cuales se dividen en dos grandes grupos, las RIP's-1, que están constituidas por una sola cadena sencilla, y las RIP's-2 las cuales, como la ricina, consisten de dos cadenas polipeptídicas, la cadena A con función catalítica en los ribosomas y la cadena B con propiedad de reconocimiento a receptores galactosilados (Barbieri et al. 1993).

Respecto a la homología de PFA con la ricina, se encontró que la similitud es con respecto a la ricina justamente en el extremo amino de la unidad catalítica de ésta, es decir en la cadena A. De los 43 aminoácidos secuenciados de PFA, 24 se corresponden con la secuencia de RA, y 16 de estos forman un segmento continuo mientras otros 5 presentan mutaciones conservadas (figura 16).

	* *	** **	***** **		
PFA-1	L-AKT-N- <u>FP</u>	-LNF [*] EIHVRDA	GVQWSPL	GRLPTGADVRRHEIPVLSNR 43	263
RA	14	AGA- <u>IVQSYT</u>	NEIRAVR--	- - - GRLPTGADVRIIK-PVLPNR 48	

Figura 16. Análisis de alineación de PFA y RA. Los residuos idénticos para ambas proteínas se marcan mediante asteriscos mientras que los residuos faltantes se marcan por guiones. Los residuos invariables en las RIP's están subrayados y en azul para RA mientras que sus homólogos en PFA sólo se encuentran en fondo azul

Como se mencionaba arriba, RA pertenece a la familia de las RIP's, y se ha identificado que éstas proteínas poseen 13 residuos altamente conservados (Funatsu *et al.* 1991). Tres de estos residuos se encuentran en el extremo amino con un patrón de secuencia YxxFxxxxR (ver figura 16) y aunque ninguno de ellos participa en la catálisis, su ausencia (especialmente F y R) provocan la inactivación de RA, lo cual sugiere que están involucrados en la unión con el sustrato o bien en el plegamiento del centro catalítico (Munishkin y Wool, 1995). Como puede observarse en la figura 16, el patrón en PFA (FxxFxxxxR), es muy semejante al de los tres aminoácidos conservados en las RIP's (YxxFxxxxR), sólo que en este caso se presenta una mutación de Y por F (tirosina por fenilalanina), además de una inserción de un aminoácido con relación a la RA. Por otra parte, se observa que PFA y RA poseen una cadena peptídica muy similar en longitud, 267 aminoácidos para RA y 273 aminoácidos para PFA, aunque ésta última carece de los 13 primeros aminoácidos de RA.

TESIS CON
FALLA DE ORIGEN

6.9 Dicroísmo Circular

Dado que los resultados de secuenciación indicaron similitud entre PFA y RA entonces, se procedió a hacer una comparación de los espectros de DC en el UV lejano de ambas proteínas, encontrándose que ambos prácticamente se superponen, y por lo mismo, puede sugerirse que ambas glicoproteínas comparten características tanto en su nivel de arreglo estructural secundario como en el supersecundario (ver figura 17).

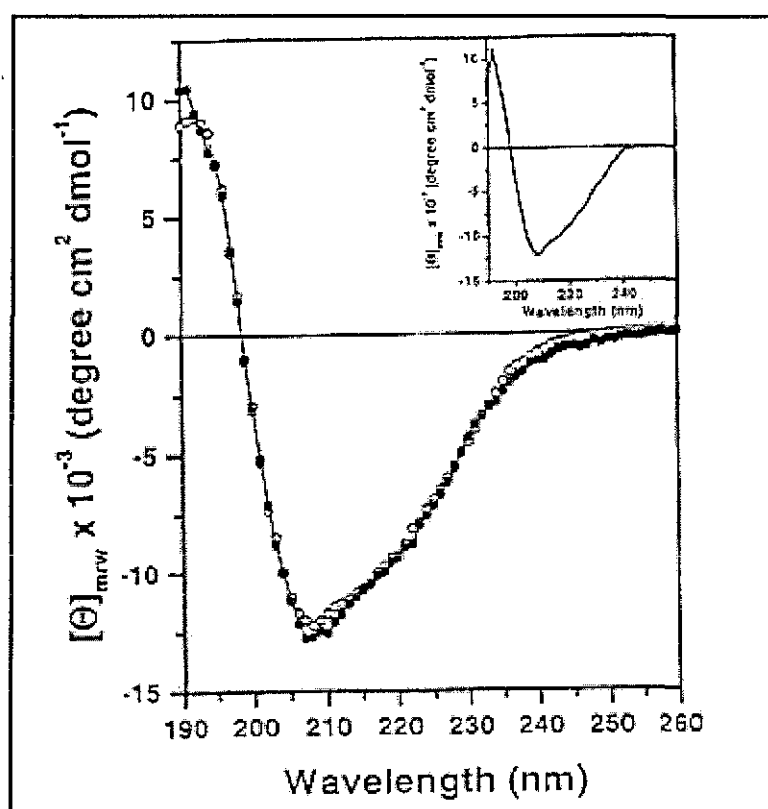


Figura 17. Espectro de DC en U.V. lejano. Los círculos cerrados representan lo datos experimentales y los círculos abiertos los datos calculados con el programa de SELCON. La figura interior corresponde al espectro de RA tomado de Argent, *et al.* (2000)

TESIS CON
FALLA DE ORIGEN

Después analizar el espectro mediante el programa SELCON, se determinó que PFA posee en su cadena peptídica segmentos significativos de estructuras hélice- α y hebras- β , los cuales, de acuerdo al programa DEFCLASS, se encuentran separados como se observa para las proteínas tipo $\alpha + \beta$. Como puede ser observado en la tabla 9, la composición de estructura secundaria y el arreglo supersecundario son muy similares a aquellos observados en la estructura tridimensional de RA (código PDB xxx).

Tabla 9. Comparación de las estructuras secundarias de PFA y RA*

Proteína	Contenido total de hélice- α	Contenido total de hebras- β	Contenido de otras estructuras	Clase de proteína
PFA	30	18	52	$\alpha + \beta$
RA	35	20	45	$\alpha + \beta$

* El contenido de estructura secundaria y el tipo de estructura supersecundaria de PFA fueron estimados del análisis de su espectro de DC mediante los programas SELCON y DEFCLASS, respectivamente. Los datos para RA son los observados en su estructura tridimensional resuelta mediante rayos X (código PDB 1IFS, Katzin, *et al.* 1991)

7 CONCLUSIONES.

De acuerdo a los experimentos de aglutinación se determinó que en el caracol *P. flagellata*, PFA se encuentra en mayor proporción en el músculo o pie del organismo por lo que en siguientes extracciones se puede obtener un extracto crudo sin las proteínas de otros tejidos del organismo.

En este trabajo la PFA demostró tener una mayor afinidad por la N-acetil galactosamina cuando se realizaron las pruebas de hemaglutinación con distintos grupos sanguíneos, aunque se sabe que presenta afinidad por galactosa y β -galactósidos.

De los experimentos de espectrometría de masas se determinó que PFA tiene un peso molecular de 32,444 Da y sin su porción glicosídica de 30,040 Da, por lo que un 7.4% de su peso total corresponde a carbohidratos.

Al comparar la secuencia de aminoácidos obtenida de PFA en el banco de datos SWISS-Prot se encontró que esta proteína presenta homología en su extremo amino con la subunidad activa de la toxina ricina (RA) y que de 43 aminoácidos secuenciados de PFA 25 se corresponden con la secuencia de RA, y 19 de estos forman un segmento continuo mientras otros 4 presentan mutaciones conservadas. Además, presenta la secuencia FxxxFxxxxR la cual es muy similar a la secuencia YxxFxxxxR, dicha secuencia corresponde a aminoácidos cruciales (aunque no involucrados) en la actividad tóxica no sólo de la RA sino del grupo de proteínas conocidas como inactivadoras de ribosomas (RIP's). Así mismo, la longitud de la cadena de PFA (267 residuos) es muy similar a la de RA (273 residuos).

De acuerdo al análisis del espectro de DC, la composición en estructura secundaria de PFA es de 30% para la hélice- α , 18% de hebra- β y 52% de estructura no periódica y, al comparar con RA, las composiciones en estructura secundaria son muy similares. De la misma manera, ambas pertenecen al grupo de proteínas catalogadas como $\alpha + \beta$.

De acuerdo a lo anterior, se determina que existe una gran similitud entre PFA y RA y, por ende, entre PFA y las RIPs. PFA representa la primera lectina animal con un motivo estructural RIP. Dado que el motivo estructural de las RIPs se encuentra en organismos filogenéticamente distantes, este puede ser un caso de evolución divergente de una proteína ancestro común, en la que se mantuvo el motivo estructural. Sin embargo, como sólo se secuenció el 20% del total de aminoácidos en la cadena no fue

posible encontrar proteínas semejantes en otros invertebrados ni por lo mismo inferir árboles filogenéticos.

Sería importante en estudios posteriores llevar a cabo pruebas para conocer si PFA presenta afinidad por patógenos como las bacterias, y si por lo mismo se trata de una proteína de defensa. Por otra parte, sería conveniente determinar si PFA posee actividad catalítica inactivadora de la síntesis proteica. Así mismo, no debe olvidarse que la biología molecular ofrece muchas posibilidades para avanzar en el conocimiento de las proteínas cuando se logra aislar el gen o genes que las codifican.

8 BIBLIOGRAFÍA

Altschul, S. F., Madden, T. L., Schäffer, A. A. Zhang, J. Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search program. *Nucleic Acids Research*. 25: 3389-3402.

Amersham Pharmacia Biotech. 2001. Affinity Chromatography Principles and Methods. Edition AA. Amersham Pharmacia Biotech AB 156 pp.

Amersham Pharmacia Biotech. 2001. Ion Exchange Chromatography Principles and Methods. Edition AA. Amersham Pharmacia Biotech AB 157 pp.

Argent, R. H. , Parrott, A. M., Day, P. J., Roberts, L. M., Stockley, P. G., Lord, J. M., and Radford, S.Z. 2000. Ribosome mediated folding of Partially Unfolded Ricin A-chain. *Journal of Biological Chemistry* 275, 13: 9263-9269.

Arreguín-Espinosa, R. and Arreguín Lozano, B. 1997. Biochemical properties of Hemagglutinins in the molusk *Pomacea flagellata*. *Biochemistry and Molecular Biology International* 43, 6:1241-1251.

Barbieri, L., Battelli, M. G., Stirpe, F. 1993. Ribosome-inactivating proteins from plants. *Biochimica et Biophysica Acta* 1154:237-282.

Barondes, S. H., Cooper, D. N. W., Gitt, M. A. and Leffler, H. 1994. Galectinas. Structure and Function of a large family of Animal Lectins. *The Journal of Biological Chemistry* 269, 33:20807-20810.

Bayley Peter. 1980. Circular Dichroism and Optical Rotation Chapter 5 In: An Introduction to Spectroscopy for Biochemists Edited by S.B. Brown. Academic Press. London pp 1-234.

Bonetto V, Bergman AC, Jornvall H, Sillard R. 1997. C-terminal sequence analysis of peptides and proteins using carboxypeptidases and mass spectrometry after derivatization of Lys and Cys residues. *Analytical Chemistry* 69(7):1315-9.

Creighton, T. E. 1993. PROTEINS, Structures and Molecular Properties. Second edition. Freeman N. Y. 1993 pp 31-35.

Diupotex Chong María Esther. 1998. Estudio Cromosómico y electroforético de especies del género *Pomacea* (Gastropoda-Pilidae) Procedentes del sureste de México. Tesis Doctoral. UNAM.

Dodd, R. B. and Drickamer, K. 2001. Lectin-like proteins in model organismos: implications for evolution of carbohydrate-binding activity. *Glycobiology* 11, 5:71R-79R.

- Drickamer, K.** and Taylor, M. E. 1993 Biology of Animal Lectins. *Annual Review of Cell Biology* 9:237-264.
- Drickamer, K.** 1995. Increasing Diversity of animal lectin structures. *Current Opinion in Structural Biology* 5:612-616.
- Drickamer, Kurt.** 1999. C-Type lectin-like domains. *Current Opinion in Structural Biology*, 9:585-90.
- Frigeri, L.G.,** Zuberi, R.I., and Liu, F. 1993. ϵ BP, a β -Galactoside-Binding Animal lectin, recognizes IgE Receptor (Fc ϵ RI) and activates mas cells. *Biochemistry* 32: 7644-49.
- Funatsu, G.,** Islam, M.R., Minami, Y., Sung-Sil, K., and Kimura, M. 1991. Conserved amino acid residues in ribosome-inactivating proteins from plants. *Biochimie* 73: 1157-1161.
- García-Hernández, E.** 1994. Estabilidad estructural de la β -lactoglobulina. Tesis de Mestría en Biología Experimental U. A. M.
- García-Hernández, E.** and Hernández-Arana, A. 1999. Structural bases of lectin-carbohydrate affinities: Comparison with protein-folding energetics. *Proteins Science* 8:1075-1086.
- García-Hernández, E.,** Zubillaga, R. A., Rodríguez-Romero, A. And Hernández-Arana, A. 2000. Stereochemical metric of lectin-carbohydrate interactions: comparison with protein-protein interfaces. *Glycobiology*.
- García Tamayo, Fernando.** Los antígenos. En: Fundamentos de inmunología 1997.
- Hennessey, J. P.** and Johnson WC. 1981. Information Content in the Circular Dichroism or proteins. *Biochemistry* 20:1085-1094.
- Hennessey, J.P.** and Johnson WC. 1982. Experimental errors and their effect on analyzing circular dichroism spectra of proteins. *Analytical Biochemistry* 125(1):177-88
- Katzin, B. J.,** Collins, E. J., and Robertus, J. D. 1991. *Proteins: Structure Function and Genetics*. 10: 251-259.
- Jensen, O.N.,** Larsen, M. R. and Roepstorff, P. 1998. *Proteins: Structure Function. And Genetics Suppl.* 2: 74-89.
- Lis, H.** and Sharon, N. 1986. Lectins as Molecules and as Tools. *Annual Review of Biochemistry* 55:35-67.

- Lis, H.** and Sharon, N. 1986 b. Applications of lectins Chapter 5 in: The Lectins Edited by Irvin E. Liener, Nathan Sharon and Irwin J. Goldstein. Academic Press, INC.
- Lowe, Christopher.** 1974. In: Affinity Chromatography, ed. Dean Peter D. G. 272pp. New York: Wiley.
- Munishkin, A.** and Wool, I. G. 1995. Systematic Deletion Analysis of Ricin A-Chain Function. SINGLE AMINO ACID DELETIONS. *The Journal of Biological Chemistry*. 270,51:30581-30587.
- Nelson David L. and Cox Michael M. Lehninger Principles of Biochemistry. Third Edition Worth Publishers. New York 2000.*
- Peumans, W. J.** and Van Damme, E. J. M. 1995. Lectins as Plant Defense Proteins. *Plant Physiol*. 109:347-352.
- Rangel, R. L. J.** 1988. Estudio Morfológico de *Pomacea flagellata* Say 1827 (Gastropoda: Ampullaridae) y algunas consideraciones sobre su taxonomía y distribución geográfica en México. Anales del Instituto de Biología 58 (1987) ser. Zool. (1):21-34.
- Rini, J. M.** 1995. Lectin Structure. *Annual Review of Biophysics and Biomolecular Structure*. 24:551-77.
- Shun-ichiro Kawabata** and Sadaaki, Iwanaga. 1999. Role of lectins in the innate immunity of horseshoe crab. *Developmental and Comparative Immunology*. 23:391-400.
- Sreerama, N.,** and Woody, R. W. 1993. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem*. 209: 32-44.
- Venyaminov, S. Y.** and Vassilenko K. S. 1994. Determination of Protein Tertiary Structure Class from Circular Dichroism Spectra. *Analytical Biochemistry* 222: 176-184.
- Venyaminov Sergei Yu.** and Yang Jen Tsi. Determination of Protein Secondary Structure Cap 3 in Circular Dichroism and the Conformational Analysis of Biomolecules. Edited by Gerald D. Fasman . Plenum Press, New York 1996.
- Watkins, W. M.,** Yates, A. D. and Greenwell, P. 1981. Blood group antigens and the enzymes involved in their synthesis: past and present. *Biochemical Society Transactions* vol 9: 186-191.
- Weis, W. I.** and Drickamer, K. 1996. Structural Basis of lectin-Carbohydrate Recognition. *Annual Review of Biochemistry*, 65: 441-73.

Williams, G.K.J., Kubelik, R.A., Livak, J.K., Rafalski, J. A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*. vol.8, 22: 6531-35.

Wilson, R. Chen, Ch. and Ratcliffe, A. N. 1999. Innate immunity in Insects: The role of Multiple, Endogenous Serum lectins in the Recognition of Foreign Invaders in the Cockroach, *Blaberus discoidalis*. *The Journal of Immunology*. 162:1590-1596.