



00551

Instituto de Biotecnología

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Identificación, clonación y análisis ontogenético del ADNc
de la α 2,6-Sialiltransferasa de *Anopheles albimanus*

T E S I S

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS BIOQUÍMICAS

P R E S E N T A:

IBQ GERMÁN | PLASCENCIA VILLA



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

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

NOMBRE: German Pascencia Villo

FECHA: 15 - junio - 05

FIRMA: German Pascencia V.

El presente trabajo fue realizado en el Departamento de Medicina Molecular y Bioprocesos del Instituto de Biotecnología de la Universidad Nacional Autónoma de México, bajo la asesoría del Dr. Sandino Estrada Mondaca, en el grupo de investigación del Dr. Octavio T. Ramírez. Durante la realización de este trabajo se contó con el apoyo económico del Consejo Nacional de Ciencia y Tecnología (CONACyT) a través de la beca 171180, de la DGEP/UNAM por beca complementaria y del proyecto IN220503 DGAPA/UNAM.

Indice

	Índice	i
	<i>Índice de Figuras</i>	iii
	<i>Índice de Tablas</i>	iv
	<i>Nomenclatura</i>	v
	RESUMEN	1
1.	INTRODUCCIÓN	3
2.	ANTECEDENTES	5
2.1.	<i>Glicosilación</i>	5
2.1.1.	<i>N-Glicosilación</i>	5
2.2.	<i>Glicoproteínas</i>	7
2.2.1.	<i>Funciones de los glicanos</i>	7
2.2.2.	<i>Función del ácido siálico</i>	8
2.2.3.	<i>Sistemas de expresión de glicoproteínas</i>	9
2.2.4.	<i>Glicosilación en insectos</i>	10
2.3.	<i>α2,6-Sialiltransferasa</i>	12
2.3.1.	<i>Actividad Enzimática de la α2,6-Sialiltransferasa</i>	12
2.3.2.	<i>Organización de la α2,6-Sialiltransferasa</i>	13
2.3.3.	<i>ADNc que Codifican para la α2,6-Sialiltransferasa</i>	15
	OBJETIVOS	18
3.	MATERIALES Y MÉTODOS	19
3.1.	<i>Búsqueda y Análisis de Secuencias</i>	19
3.2.	<i>Diseño de oligonucleótidos</i>	19
3.3.	<i>Material Biológico</i>	20
3.4.	<i>Extracción de ARN total</i>	20
3.5.	<i>Tratamiento de ARN con DNasa I</i>	21
3.6.	<i>Síntesis de ADNc</i>	22
3.7.	<i>Amplificación por PCR</i>	23
3.8.	<i>Purificación de Amplificaciones</i>	26
3.9.	<i>Clonación de Amplificaciones</i>	27
3.10.	<i>Purificación de Plásmidos</i>	30
3.11.	<i>Secuenciación de ADN</i>	31
3.12.	<i>RACE (Rapid Amplification of cDNA Ends)</i>	32
3.12.1.	<i>Diseño de oligonucleótidos para RACE</i>	32
3.12.2.	<i>5'-RACE (Rapid Amplification of cDNA Ends)</i>	33
3.12.3.	<i>3'-RACE (Rapid Amplification of cDNA Ends)</i>	35

4.	RESULTADOS Y DISCUSIÓN	37
4.1.	Análisis bioinformático para la identificación de la α2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	37
4.1.1.	<i>Búsqueda de α2,6-Sialiltransferasas en el genoma de <i>Anopheles gambiae</i></i>	38
4.1.2.	<i>Análisis de la α2,6-Sialiltransferasa de <i>Anopheles gambiae</i></i>	41
4.2.	Identificación de la α2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	43
4.2.1.	<i>Amplificaciones ADNc dT18</i>	45
4.2.2.	<i>Amplificaciones ADNc SE</i>	46
4.3.	Identificación de los extremos 5' y 3' de la α2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	49
4.3.1.	<i>5'-RACE (Rapid Amplification of cDNA Ends)</i>	49
4.3.2.	<i>3'-RACE (Rapid Amplification of cDNA Ends)</i>	53
4.3.3.	<i>Clonación ADNc α2,6-Sialiltransferasa de <i>Anopheles albimanus</i></i>	56
4.4.	Análisis de la secuencia de la α2,6-Sialiltransferasa de <i>Anopheles albimanus</i> (AaST6)	57
4.4.1.	<i>Identidad de la secuencia de la α2,6-Sialiltransferasa de <i>Anopheles albimanus</i> (AaST6)</i>	57
4.4.2.	<i>Topología de la α2,6-Sialiltransferasa de <i>Anopheles albimanus</i></i>	59
4.4.3.	<i>Modificaciones postraduccionales de AaST6</i>	63
4.5.	Análisis ontogenético de expresión de AaST6	68
5.	CONCLUSIONES	70
6.	PERSPECTIVAS	72
7.	REFERENCIAS	73
	ANEXOS	80

Índice de Figuras

Figura 2.1.	Vía de <i>N</i> -Glicosilación.	6
Figura 2.2.	Estructura del ácido siálico.	8
Figura 2.3.	Vía de síntesis del ácido siálico (CMP-Neu5Ac)	8
Figura 2.4.	Representación esquemática de la α 2,6-Sialiltransferasa	14
Figura 2.5.	Región del <i>Sialylmotif L, S y VS</i> en las α 2,6-Sialiltransferasas	14
Figura 3.1.	Procedimiento <i>seminested</i> -PCR	25
Figura 3.2.	Procedimiento <i>nested</i> -PCR	26
Figura 3.3.	Mapa del vector pCR-Blunt II-TOPO (Invitrogen)	30
Figura 3.4.	Representación de procedimientos RACE y localización de oligonucleótidos específicos sobre fragmento de ADNc conocido.	32
Figura 3.5.	Procedimiento 5'-RACE (<i>Rapid Amplification of cDNA Ends</i>)	33
Figura 3.6.	Procedimiento 3'-RACE (<i>Rapid Amplification of cDNA Ends</i>)	35
Figura 4.1.	Alineamiento de secuencias nucleotídicas de α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	39
Figura 4.2.	Alineamiento de secuencias aminoacídicas de α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	40
Figura 4.3.	Identificación de los <i>sialylmotifs</i> en la α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	42
Figura 4.4.	Amplificación G6PD de <i>Anopheles albimanus</i>	44
Figura 4.5.	Amplificaciones sobre ADNc-dT18	44
Figura 4.6.	Amplificaciones sobre ADNc SE	45
Figura 4.7.	Reamplificaciones <i>seminested</i> -PCR	46
Figura 4.8.	Secuencia de amplificación con oligos AgL-AaR2 sobre ADNc SE	46
Figura 4.9.	Alineamiento de fragmento de las α 2,6-Sialiltransferasas de insectos	47
Figura 4.10.	Amplificación 5'-RACE	48
Figura 4.11.	Amplificación 5'-RACE con NUP-STNGSP1 .	49
Figura 4.12.	Reamplificación 5'-RACE	49
Figura 4.13.	Reamplificaciones 5'-RACE	50
Figura 4.14.	Secuencia de extremo 5' de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	51
Figura 4.15.	Amplificación 3'-RACE	52
Figura 4.16.	Reamplificaciones 3'-RACE	52
Figura 4.17.	Reamplificaciones <i>seminested</i> -PCR 3'-RACE	53
Figura 4.18.	Secuencia de extremo 3' de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	53
Figura 4.19.	Secuencia del ADNc de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i> (1746pb)	55
Figura 4.20.	Amplificación ADNc α 2,6-Sialiltransferasa <i>Anopheles albimanus</i> (1354pb)	56
Figura 4.21.	Alineamiento de las α 2,6-Sialiltransferasas de insectos	59
Figura 4.22.	Alineamiento de las α 2,6-Sialiltransferasas de insectos y vertebrados	60

Figura 4.23.	Análisis de hidrofobicidad de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i> mediante el método Kyte&Doolittle y representación de la región transmembranal	61
Figura 4.24.	Localización de los sitios de N-glicosilación en las α 2,6-Sialiltransferasas	63
Figura 4.25.	Predicción de estructura secundaria de las Sialiltransferasas mediante PSIPRED	65
Figura 4.26.	Árbol filogenético basado en las secuencias proteicas de las α 2,6-Sialiltransferasas (GeneBee)	66
Figura 4.27.	PCR semicuantitativo de actina y <i>AaST6</i> de <i>A.albimanus</i>	67
Figura 4.28.	Expresión relativa a actina de <i>AaST6</i> (<i>A.albimanus</i>)	68

Índice de Tablas

Tabla 2.1.	α 2,6-Sialiltransferasas identificadas, clonadas y caracterizadas	15
Tabla 3.1.	Oligonucleótidos para secuenciación	31
Tabla 3.2.	Secuencia de oligonucleótidos 5'-RACE	34
Tabla 3.3.	Condiciones de reacción recomendadas en 5'-RACE (<i>Touchdown-PCR</i>)	35
Tabla 3.4.	Secuencia de oligonucleótidos APT18 y AP	35
Tabla 4.1.	Bases de datos del genoma de <i>Anopheles gambiae</i> .	38
Tabla 4.2.	Secuencias de las sondas de búsqueda (<i>sialylmotif L</i>).	38
Tabla 4.3.	Números de acceso de las secuencias de <i>Anopheles gambiae</i> identificadas.	38
Tabla 4.4.	Análisis de identidad de la α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	41
Tabla 4.5.	Análisis de identidad de la región catalítica de la α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	41
Tabla 4.6.	Secuencias de oligonucleótidos del consenso de <i>Anopheles-Drosophila</i>	43
Tabla 4.7.	Secuencias de oligonucleótidos con el uso de codones de <i>Anopheles albimanus</i>	43
Tabla 4.8.	Amplificaciones esperadas con los diferentes oligonucleótidos	43
Tabla 4.9.	Identidad del fragmento de 471 pb de <i>A.albimanus</i> respecto a las α 2,6-Sialiltransferasas	47
Tabla 4.10.	Secuencias de oligonucleótidos para RACE de α 2,6-Sialiltransferasa de <i>A.albimanus</i>	48
Tabla 4.11.	Amplificaciones esperadas en reacciones 5'-RACE	50
Tabla 4.12.	Análisis de identidad de la α 2,6-Sialiltransferasa putativa de <i>Anopheles albimanus</i> (<i>AaST6</i>)	57
Tabla 4.13.	Análisis de identidad de la región catalítica de la α 2,6-Sialiltransferasa putativa de <i>Anopheles albimanus</i>	58
Tabla 4.14.	Modificaciones postraduccionales de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	62
Tabla 4.15.	Etapas de desarrollo de <i>A.albimanus</i> analizadas	67

Figura 4.23.	Análisis de hidrofobicidad de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i> mediante el método Kyte&Doolittle y representación de la región transmembranal	61
Figura 4.24.	Localización de los sitios de N-glicosilación en las α 2,6-Sialiltransferasas	63
Figura 4.25.	Predicción de estructura secundaria de las Sialiltransferasas mediante PSIPRED	65
Figura 4.26.	Árbol filogenético basado en las secuencias proteicas de las α 2,6-Sialiltransferasas (GeneBee)	66
Figura 4.27.	PCR semicuantitativo de actina y <i>AaST6</i> de <i>A.albimanus</i>	67
Figura 4.28.	Expresión relativa a actina de <i>AaST6</i> (<i>A.albimanus</i>)	68

Índice de Tablas

Tabla 2.1.	α 2,6-Sialiltransferasas identificadas, clonadas y caracterizadas	15
Tabla 3.1.	Oligonucleótidos para secuenciación	31
Tabla 3.2.	Secuencia de oligonucleótidos 5'-RACE	34
Tabla 3.3.	Condiciones de reacción recomendadas en 5'-RACE (<i>Touchdown-PCR</i>)	35
Tabla 3.4.	Secuencia de oligonucleótidos APT18 y AP	35
Tabla 4.1.	Bases de datos del genoma de <i>Anopheles gambiae</i> .	38
Tabla 4.2.	Secuencias de las sondas de búsqueda (<i>sialylmotif L</i>).	38
Tabla 4.3.	Números de acceso de las secuencias de <i>Anopheles gambiae</i> identificadas.	38
Tabla 4.4.	Análisis de identidad de la α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	41
Tabla 4.5.	Análisis de identidad de la región catalítica de la α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	41
Tabla 4.6.	Secuencias de oligonucleótidos del consenso de <i>Anopheles-Drosophila</i>	43
Tabla 4.7.	Secuencias de oligonucleótidos con el uso de codones de <i>Anopheles albimanus</i>	43
Tabla 4.8.	Amplificaciones esperadas con los diferentes oligonucleótidos	43
Tabla 4.9.	Identidad del fragmento de 471pb de <i>A.albimanus</i> respecto a las α 2,6-Sialiltransferasas	47
Tabla 4.10.	Secuencias de oligonucleótidos para RACE de α 2,6-Sialiltransferasa de <i>A.albimanus</i>	48
Tabla 4.11.	Amplificaciones esperadas en reacciones 5'-RACE	50
Tabla 4.12.	Análisis de identidad de la α 2,6-Sialiltransferasa putativa de <i>Anopheles albimanus</i> (<i>AaST6</i>)	57
Tabla 4.13.	Análisis de identidad de la región catalítica de la α 2,6-Sialiltransferasa putativa de <i>Anopheles albimanus</i>	58
Tabla 4.14.	Modificaciones postraduccionales de la α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	62
Tabla 4.15.	Etapas de desarrollo de <i>A.albimanus</i> analizadas	67

Nomenclatura

ARN	Ácido Ribonucleico
ARNm	Ácido Ribonucleico mensajero
ADN	Ácido Desoxirribonucleico
ADNc	Ácido Desoxirribonucleico complementario
PCR	<i>Polymerase Chain Reaction</i> o Reacción en Cadena de la Polimerasa
T_m	Temperatura de fusión (<i>melting</i>)
T_a	Temperatura de alineamiento (<i>annealing</i>)
Glc	Glucosa
Man	Manosa
Fuc	Fucosa
Gal	Galactosa
GlcNAc	<i>N</i> -Acetilgalactosamina
Neu5Ac	Ácido Acetilneuramínico o Ácido Siálico

RESUMEN

El gran impulso e interés puesto en la secuenciación de los genomas de diversos organismos ha revelado una importante cantidad de información útil en el estudio de diversos aspectos del metabolismo celular. Particularmente, con la disponibilidad de los genomas de los insectos *Drosophila melanogaster* y *Anopheles gambiae* (Adams et al, 2000; Holt et al, 2002), es posible desarrollar estudios sobre la capacidad de N-glicosilación de los insectos.

Este procesamiento es de gran importancia en la producción de glicoproteínas recombinantes por líneas celulares de insecto ya que el patrón de glicosilación que posean impactará directamente en sus propiedades farmacocinéticas.

Con la identificación y caracterización de la α 2,6-Sialiltransferasa de *Drosophila melanogaster* se confirmó el potencial de N-glicosilación de los insectos (Fabini et al, 2001; Fabini et al, 2002; Koles et al, 2004). Haciendo uso de herramientas bioinformáticas y de las secuencias de las ST6Gal de *D.melanogaster* y de humano, se identificó en el genoma del mosquito *Anopheles gambiae* una secuencia ortóloga putativa que posee los dominios funcionales y estructurales de la región catalítica de las α 2,6-Sialiltransferasas. El mosquito *Anopheles albimanus* está estrechamente relacionado con *An.gambiae*, se sabe es capaz de producir sialoglicoproteínas y a diferencia de éste último su genoma no está secuenciado. A partir de las secuencias de *D.melanogaster* y *An.gambiae* se diseñaron oligonucleótidos degenerados en las regiones conservadas llamadas *sialylmotifs*. Usando ADNc de *An.albimanus* como templado se amplificó un fragmento de 471pb. Una vez identificados los extremos 5' y 3' del transcrito, se determinó que éste tiene una longitud total de 1746pb, y posee una identidad de 49 a 64% con las secuencias ortólogas de las α 2,6-Sialiltransferasas conocidas.

La secuencia de este ADNc AaST6 (AY955495) posee un marco abierto de lectura de 440aa, en la región identificada como el dominio catalítico tiene identidades de 38 a 41% con las proteínas ortólogas de vertebrados, 57% con las ortólogas de invertebrados y 87% con la α 2,6-Sialiltransferasa ortóloga putativa de *An.gambiae*. Mediante **BLAST** y **CD-Search** se identificó que posee los dominios funcionales y estructurales que indican que esta secuencia pertenece a la Familia 29 de las Glicosiltransferasas: Sialiltransferasas (EC 2.4.99.). Dichos dominios son: el dominio citoplásmico N-terminal de 32aa, el dominio transmembranal (17aa) y un tallo que termina en la región del dominio catalítico, este último presenta las regiones conservadas importantes catalíticamente conocidas como *sialylmotifs*. La estructura secundaria del dominio catalítico de la sialiltransferasa de *An.albimanus* presenta una organización de α -hélices y hojas- β presente en otras enzimas de esta familia.

El análisis ontogenético de expresión de AaST6 de *An.albimanus* mostró dos máximos, éstos ocurren en la etapa de larvas *instar 2* y en pupas, también se observó que se tiene un nivel basal de expresión en todas las etapas de desarrollo analizadas, esto sugiere la presencia de sialoglicoproteínas a través de toda la ontogenia de *An.albimanus*.

Con la identificación y clonación del transcrito *AaST6* de *Anopheles albimanus*, confirmamos que este organismo posee el potencial genético necesario para la sialilación de glicoproteínas. La información disponible hasta el momento indica que los géneros *Drosophila* y *Anopheles* pueden realizar sialilación, mecanismo que se pensaba ausente en insectos. *AaST6* es la segunda sialiltransferasa de insecto en ser clonada, la primera del género *Anopheles* y la quinta identificada de insectos (géneros *Drosophila* y *Anopheles*).

1. INTRODUCCIÓN

Muchas proteínas y péptidos de mamíferos poseen un elevado valor farmacéutico, terapéutico y de diagnóstico. Sin embargo, estas proteínas suelen estar presentes en cantidades muy pequeñas en sus tejidos de origen y por ello es sumamente costosa su obtención y purificación, varias de estas proteínas son glicoproteínas. La glicosilación de las proteínas puede influir en muchas de sus propiedades, incluyendo movimiento intracelular, función biológica, estabilidad bioquímica, prevención de agregación y degradación proteolítica y solubilidad (Jenkins *et al.*, 1994). La mayoría de las proteínas producidas para terapia humana son glicoproteínas derivadas de células animales, por lo que es esencial caracterizar totalmente y si es posible incluso controlar el patrón de glicosilación de estos productos, ya que la *Food and Drug Administration* (FDA) en los Estados Unidos y el *Committee for Proprietary Medical Productions* (CPMP) en la Comunidad Europea (Jenkins *et al.*, 1994; Jenkins *et al.*, 1996), demandan análisis detallados sobre los carbohidratos de todas las glicoproteínas destinadas para consumo humano.

La necesidad de sistemas de expresión heterólogos que sean capaces de llevar a cabo *N*-glicosilación en forma similar a su forma nativa, ha intensificado el interés en la comprensión de esta vía en los insectos, ya que han sido usados ampliamente como sistema de expresión para la producción de glicoproteínas (Jarvis *et al.*, 1998; Marchal *et al.*, 2001). Los análisis de las glicoproteínas endógenas o recombinantes producidas en insectos realizados por diversos grupos de investigación indicaron que éstas raras veces poseen ácido siálico terminal, por lo que los autores concluyen que la sialilación de glicoproteínas constituye un aspecto altamente especializado que ocurre sólo en algunos tipos de insectos.

La sialilación es llevada a cabo en la última etapa de la vía de *N*-glicosilación, esta consiste en la transferencia del ácido siálico o CMP-ácido acetilneuramínico (CMP-Neu5Ac) hacia un residuo de galactosa, este proceso es catalizado por una sialiltransferasa. Estructuralmente, las sialiltransferasas de eucariontes poseen una topología similar, son proteínas de membrana tipo II residentes en Golgi. Consisten de un dominio citoplásmico *N*-terminal corto, un segmento transmembranal y un tallo de longitud variable seguido del dominio catalítico C-terminal (Breton *et al.*, 1999; Koles *et al.*, 2004). El dominio catalítico contiene tres regiones de péptidos altamente conservados llamados *sialylmotifs*, que están implicados en el reconocimiento de sustrato donador y aceptor, así como en la formación de puentes disulfuro y catálisis de la enzima (Datta *et al.*, 1997; Datta *et al.*, 2001).

Debido a que en general las glicosiltransferasas (incluyendo a las sialiltransferasas) son proteínas poco abundantes, ha hecho difícil su purificación con fines de identificación y clonación de sus respectivos ADNc (Breton *et al.*, 1999; Harquin-Lepers *et al.*, 1995). Algunas de las estrategias que se han utilizado para la identificación y clonación de ADNc y genes que codifican para sialiltransferasas incluyen: análisis de bancos de ADNc, uso de oligonucleótidos degenerados, búsquedas por homología, diversos métodos basados en PCR, el uso de anticuerpos policlonales, entre otros (Tsuji, 1996). El rápido progreso en la secuenciación de los genomas de diversos organismos ha permitido la identificación y clonación de diversas glicosiltransferasas, lo cual difícilmente habría sido posible por otros métodos.

Al completarse y publicarse el genoma de la mosca de la fruta (*Drosophila melanogaster*) en el año 2000 (Adams *et al.*, 2000), fue posible la identificación, clonación y caracterización enzimática de diversas proteínas de la vía de N-glicosilación de este insecto, así como de las vías de síntesis y transporte de diversos azúcares activados. Confirmando así su potencial como organismo modelo para el estudio del proceso de glicosilación en insectos. Se determinó que los insectos y en particular *D.melanogaster*, posee las enzimas necesarias para la síntesis del ácido siálico (CMP-Neu5Ac) y para su transferencia sobre glicoproteínas (Kim *et al.*, 2002; Fabini *et al.*, 2001; Fabini *et al.*, 2002; Koles *et al.*, 2004).

En el año 2002 se publicó el genoma completo del mosquito *Anopheles gambiae* (Holt *et al.*, 2002), hecho que sumado a la disponibilidad de diversas herramientas bioinformáticas abre una posibilidad importante al contar con otro insecto modelo para realizar diversos estudios. Basándose en la información del genoma de *An.gambiae* se han clonado algunos genes de *Anopheles albimanus*, ya que estos organismos comparten gran homología. El genoma de *An.albimanus* no ha sido descrito, y es de resaltar que éste organismo es capaz de producir sialoglicoproteínas (Nok *et al.*, 2002). Fue así como se planteó como objetivo la identificación y clonación de la α 2,6-Sialiltransferasa de *Anopheles albimanus*, basándose en la información de los genomas de *An.gambiae* y de *D.melanogaster*. Demostrando de esta forma que la sialilación de glicoproteínas es llevada a cabo por este organismo.

La identificación y clonación del ADNc de la α 2,6-Sialiltransferasa de *An.albimanus* permite tener un punto de partida para realizar estudios concernientes a la expresión y regulación de este gen. Además de ser una posible herramienta biotecnológica en estrategias de ingeniería de vías metabólicas, pudiendo complementar o sobreexpresar esta actividad en diversas líneas celulares de insecto donde se producen glicoproteínas recombinantes, las cuales carecen o tienen niveles no detectables de actividad sialiltransferasa endógena.

2. ANTECEDENTES

2.1. Glicosilación

La *Glicosilación* es la modificación co- o postraducciona l más importante y compleja de las células eucariotas, consiste en la adición covalente de carbohidratos a proteínas por medio de un grupo de enzimas (Jenkins *et al*, 1994; Spiro, 2002). Se han identificado varias rutas de glicosilación en diferentes organismos, éstas conducen hacia el procesamiento de los glicanos de las glicoproteínas que serán secretadas o formarán parte de las membranas, citoplasma o núcleo. La naturaleza y el tipo de glicanos formados por la maquinaria enzimática tendrán influencia sobre la actividad biológica de la proteína procesada.

Esta modificación está influenciada por diversos factores, los principales son: la conformación global de la proteína, la conformación local de segmentos de la proteína y sobre todo, de las enzimas y los sustratos disponibles para este procesamiento en diversos tipos de células. El tipo y patrón de glicosilación es de tipo proteína dependiente, sitio específica, y particular de algunos tejidos o células (Dwek, 1995).

Se han descrito diversos tipos de enlaces carbohidrato-péptido presentes en las glicoproteínas, estos pueden agruparse en cinco tipos que pueden ocurrir en la misma proteína dependiendo de la maquinaria enzimática disponible y de la secuencia y conformación de la proteína, estos son: *N*-Glicosilación, *O*-Glicosilación, *C*-Manosilación, Fosfoglicosilación y Glipiación. De los anteriores, se le ha dado mayor importancia a *N*-glicosilación al ser la modificación más común, por su implicación en el reconocimiento celular y estar presente en diversas proteínas de interés comercial.

2.1.1. *N*-Glicosilación

La *N*-glicosilación es el tipo de glicosilación más ampliamente distribuida, ha sido observada en varios tipos de proteínas eucariotes incluyendo proteínas plasmáticas, hormonas, enzimas, receptores celulares, inmunoglobulinas y lectinas (Butters, 2001; Spiro, 2002). En todas ellas los glicanos tienen importancia biológica, influyendo directamente sobre su actividad, estabilidad, conformación y otras propiedades farmacocinéticas.

La vía de *N*-glicosilación en mamíferos es un proceso que ocurre en múltiples etapas en varios compartimentos celulares. Comienza con la transferencia del oligosacárido precursor presintetizado ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) unido a dolicol fosfato. Dicha transferencia la realiza la Oligosacariltransferasa (OST), hacia el péptido naciente en un residuo asparagina (Asn) de la proteína que se encuentre en el consenso Asn-X-Ser/Thr, donde X puede ser cualquier aminoácido excepto prolina (**Figura 2.1.-A**) (Jarvis *et al*, 1998; Jenkins *et al*, 1994).

Este precursor es secuencialmente procesado por Glucosidasas (α -Glucosidasa I y α -Glucosidasa II) (**Figura 2.1.-B**) y por α -Manosidasa I (**Figura 2.1.-C**), enzimas localizadas en retículo endoplásmico. Si los glicanos son procesados parcial o totalmente por la α -Manosidasa I y estos ya no sufren otra modificación, serán glicanos altos en manosa (con cinco a nueve manosas terminales). Estas primeras etapas están asociadas con el correcto plegamiento de la proteína, siendo esencial para el transporte al

aparato de Golgi. Estas primeras etapas de la vía de *N*-glicosilación están conservadas en todos los eucariontes (Brooks *et al*, 2004; Tomiya *et al*, 2004).

Al entrar al aparato de Golgi, los glicanos son procesados secuencialmente por la *N*-acetil Glucosaminiltransferasa I (GlcNAcT I) (Figura 2.1.-D), la α -Manosidasa II (Figura 2.1.-E), la Fucosiltransferasa (FucT) (Figura 2.1.-E), la *N*-acetil Glucosaminiltransferasa II (GlcNAcT II) (Figura 2.1.-F), Galactosiltransferasa (GalT) (Figura 2.1.-G) y Sialiltransferasa (SialT) (Figura 2.1.-H). Produciéndose de esta forma glicanos complejos (glicanos con galactosa y ácido siálico terminales) (Brooks *et al*, 2004; Tomiya *et al*, 2004).

Alternativamente, la β -*N*-acetilglucosaminidasa (GlcNasa) puede remover residuos de *N*-acetil glucosamina, desviando la vía hacia glicanos llamados paucimanosídicos (glicanos con tres manosas terminales) (Figura 2.1.-I) (Watanabe *et al*, 2002; Tomiya *et al*, 2004). Este proceso es muy común en insectos.

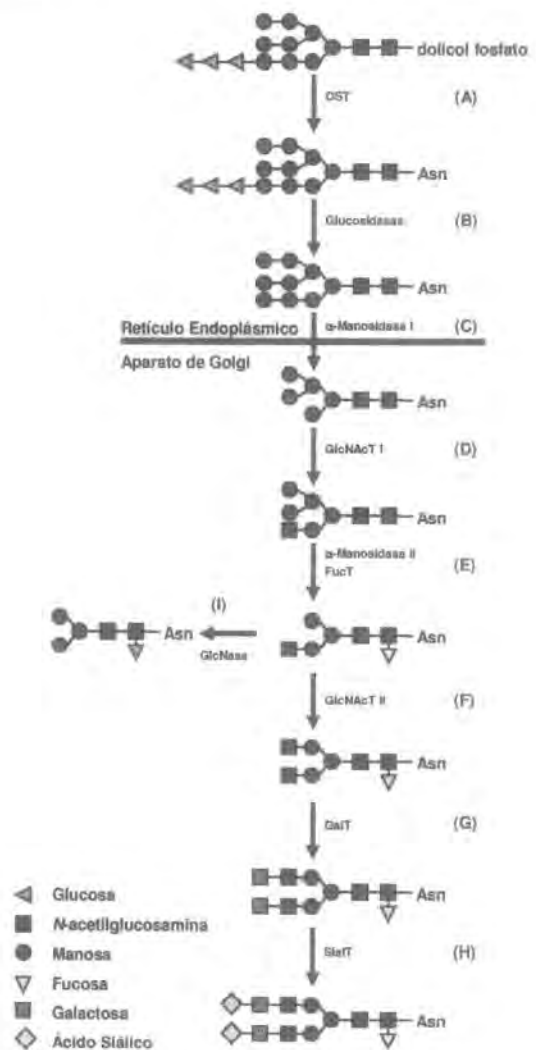


Figura 2.1. Vía de *N*-Glicosilación. Procesamiento de glicanos a través del retículo endoplásmico y aparato de Golgi. **GlcNAcT I**= *N*-acetil Glucosaminiltransferasa I, **FucT**= Fucosiltransferasa, **GlcNasa**= β -*N*-acetil Glucosaminidasa, **GlcNAcT II**= *N*-acetil Glucosaminiltransferasa II, **GalT**= Galactosiltransferasa y **SialT**= Sialiltransferasa.

Dependiendo de los niveles de actividad de las diferentes enzimas de la vía de *N*-glicosilación se producirán glicanos altos en manosa, paucimanosídicos o complejos (Jarvis *et al.*, 1998). El tipo de *N*-glicosilación es específico para cada organismo ya que depende de la presencia o ausencia de los genes que codifican para las glicosiltransferasas, así como de su regulación que es de tipo tejido-específica.

2.2. Glicoproteínas

Las glicoproteínas pueden contener oligosacáridos unidos por medio de diferentes formas (*N*-glicosilación principalmente) y tener más de un glicano por molécula. En conjunto, estos carbohidratos tendrán diversos efectos sobre las propiedades de la proteína sobre la cual están unidos. El mismo tipo de glicano en diferentes proteínas puede tener diferentes propiedades, dependiendo de su orientación y localización en el péptido.

2.2.1. Funciones de los glicanos

Desde el mismo momento de síntesis de la proteína y su paso por el retículo endoplásmico, los glicanos tienen como función promover el correcto plegamiento al orientar las cadenas peptídicas. Sirviendo como un sistema de control de calidad durante el procesamiento a través de retículo endoplásmico y aparato de Golgi (Helentius *et al.*, 2001).

Los glicanos pueden afectar la estabilidad y estructura tridimensional de las glicoproteínas (Dwek, 1995; Butters, 2001). Promueven solubilidad y previenen agregación e inactivación, si los glicanos son alterados o están incompletos ocurre agregación (Jenkins *et al.*, 1994).

Una de las principales funciones de los glicanos a nivel extracelular es servir como sitios de reconocimiento al ser estructuras compactas y versátiles (Dwek, 1995). Los carbohidratos pueden afectar la inmunogenicidad de las proteínas al ser sitios de reconocimiento o enmascarar sitios antigénicos de la proteína, regulando de esta forma la respuesta inmune del organismo (Rademacher *et al.*, 1982; Jenkins *et al.*, 1994).

2.2.2. Ácido Siálico

Los ácidos siálicos comprenden una familia diversa de α -ceto ácidos de nueve carbonos. El ácido *N*-acetil neuramínico (Neu5Ac) es el más común en mamíferos. Generalmente está ocupando las posiciones terminales de los *N*-glicanos de proteínas o gangliosidos (Angata *et al.*, 2002).

Lleva a cabo diversas interacciones con otras moléculas o células. Puede afectar las propiedades biológicas de glicoproteínas, tales como: mantenimiento de glicoproteínas en circulación, vida media de proteínas séricas, modulación de actividad biológica, adhesión celular, comunicación intracelular, estabilidad y conformación, entre otras (Katsutoshi, 1996; Andersen *et al.*, 2002). El ácido siálico terminal protege a las proteínas de ataques proteolíticos previniendo su degradación (Rademacher *et al.*, 1988). Si las glicoproteínas no contienen glicanos con ácido siálico terminal, tendrán propiedades biológicas variables o subóptimas.

En algunas proteínas, la ausencia de glicanos complejos con galactosa y ácido siálico terminal resulta en agregación o reconocimiento por el sistema inmune, con la consecuente inactivación de la proteína. De

esta forma el sistema inmune puede distinguir entre proteínas propias o ajenas al organismo de acuerdo al patrón de ácidos siálicos (Traving *et al*, 1998; Andersen *et al*, 2002).

La sialilación (adición de ácido siálico en glicoproteínas) requiere la síntesis de CMP-Neu5Ac (ácido acetilneuramínico o ácido siálico) (**Figura 2.2.**). La posterior transferencia sobre el oligosacárido es por medio de la Sialiltransferasa..

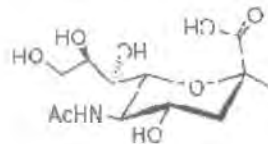


Figura 2.2. Estructura del ácido siálico.

El ácido siálico es sintetizado a partir de UDP-GlcNAc (UDP-*N*-acetil Glucosamina) a través de múltiples reacciones enzimáticas (**Figura 2.3.**). La primer etapa es catalizada por la UDP-GlcNAc 2-epimerasa/ManNAc cinasa, una enzima bifuncional que convierte UDP-GlcNAc en ManNAc-6-P (**Figura 2.3.-1**). Posteriormente es convertida en Neu5Ac-9-P por medio de la Neu5Ac fosfato sintasa (**Figura 2.3.-2**), y luego procesada por la Neu5Ac-9-P fosfatasa obteniendo el ácido acetilneuramínico (Neu5Ac)(**Figura 2.3.-3**). Finalmente ocurre al adición de CMP para obtener el CMP-Neu5Ac, acción mediada por la CMP-Neu5Ac sintasa (**Figura 2.3.-4**).

Se ha observado que el suministro exógeno de ManNAc (*N*-acetil Manosamina) promueve la síntesis de ManNAc-6-P mediada por la enzima ManNAc cinasa (**Figura 2.3.** reacción 5). Ésta es procesada de la misma forma hasta la obtención de CMP-Neu5Ac.

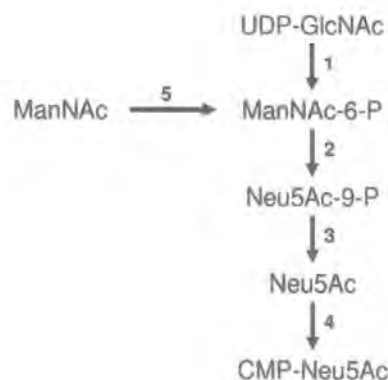


Figura 2.3. Vía de síntesis del ácido siálico (CMP-Neu5Ac). 1, UDP-GlcNAc 2-epimerasa/ManNAc cinasa; 2, Neu5Ac fosfato sintasa; 3, Neu5Ac-9-P fosfatasa; 4, CMP-Neu5Ac sintasa; 5, ManNAc cinasa. (Modificada de Tomiya *et al*, 2003a).

Diversos estudios demostraron que el CMP-Neu5Ac es sintetizado en niveles apenas detectables en los insectos (Tomiya *et al*, 2003a; Tomiya *et al*, 2004). Únicamente ha sido detectado en *Drosophila melanogaster*, *Philaenus spumarius*, *Trichoplusia ni* y *Spodoptera frugiperda* (Kim *et al*, 2002). Esto en parte, debido a que la actividad de algunas de las enzimas de la vía de síntesis del CMP-Neu5Ac en insectos (células Sf9 de

Spodoptera frugiperda), es entre 30 y 50 veces menor respecto a la actividad específica detectada en vertebrados (Efferz *et al.*, 1999).

Se mantenían diversas interrogantes respecto a si los insectos poseen la capacidad de biosintetizar ácido siálico debido a los pocos datos que se habían podido recabar. Se reportó la clonación y caracterización de la Neu5Ac fosfato sintasa de *Drosophila melanogaster* y la identificación de otros genes de la vía de síntesis del ácido siálico en este organismo (Kim *et al.*, 2002). Confirmando que los insectos son capaces de sintetizar este compuesto. Los mecanismos de regulación de la expresión de los genes de la vía de síntesis de ácido siálico en insectos permanecen sin resolver. No se han identificado tampoco genes que codifican para transportadores de ácido siálico. Los insectos poseen parte de la infraestructura para la síntesis de glicanos complejos con ácido siálico terminal.

2.2.3. Sistemas de expresión de glicoproteínas

La expresión de glicoproteínas recombinantes ha sido realizada en diferentes sistemas con diversos resultados. Varios sistemas de expresión alternativos han sido desarrollados y evaluados. Particularmente en el caso de glicoproteínas recombinantes es importante considerar la capacidad de glicosilación del sistema de expresión empleado.

Escherichia coli ha sido el sistema de expresión procarionte más utilizado para la producción de proteínas. Tiene como limitante que en la mayoría de los casos no realiza un plegamiento correcto, dando proteínas insolubles, además de no poseer capacidades para realizar glicosilación. Entre los sistemas con levaduras, son *Saccharomyces cerevisiae*, *Schizosaccharomices pombe* y *Pichia pastoris* de los más utilizados. Los glicanos que producen son típicamente con terminales altas en manosa, por lo que en el caso de proteínas que requieren glicanos complejos no resulta un sistema de expresión adecuado (Andersen, 2002; Butters, 2002; Betenbaugh *et al.*, 2004).

Algunas plantas transgénicas, así como cultivos de células vegetales han sido utilizados para producir diferentes proteínas recombinantes. Se sabe que las plantas tienen capacidades limitadas de glicosilación al producir glicanos altos en manosa y en algunos casos complejos pero sin ácido siálico. Adicionan residuos de fucosa y xilosa que podrían ser antigénicos para humanos (Andersen, 2002; Butters, 2002). Hasta el momento no se ha considerado como un buen sistema de expresión para glicoproteínas recombinantes.

La mayoría de las líneas celulares CHO (*Chinese Hamster Ovary*) y BHK (*Baby Hamster Kidney*) usadas para la producción de proteínas recombinantes son capaces de realizar glicosilación compleja y con mínima heterogeneidad. Tienen como limitante la baja productividad de proteína recombinante en comparación con otros sistemas de producción (Brooks, 2004). Se tienen reportes de que algunas glicoproteínas producidas en este sistema son potencialmente antigénicas al incorporar glicanos con enlaces no presentes en proteínas humanas (Jenkins *et al.*, 1994).

Se esperaría que especies que son filogenéticamente cercanas a los humanos tuvieran una ruta de *N*-glicosilación común. Sorprendentemente hay algunas diferencias en la capacidad de glicosilación de las líneas celulares de roedores y humanos. Estas diferencias producen respuesta inmune y variabilidad en las propiedades farmacocinéticas de las proteínas. Algunas líneas celulares de humano han mostrado alteraciones en su capacidad de glicosilación debido al proceso de establecimiento de líneas estables (Jenkins *et al*, 1996; Andersen *et al*, 2002).

El uso de células de insecto (líneas celulares de lepidópteros principalmente) como sistema de expresión de proteínas heterólogas usando baculovirus como vectores, ha puesto interés sobre las capacidades de glicosilación de las células de insecto. Diversos estudios sugirieron que el procesamiento de los glicanos en insectos es similar o idéntico al que realizan los mamíferos en las primeras etapas. Este sistema ofrece otras ventajas como seguridad, facilidad de uso y alta productividad. La aparente incapacidad de glicosilación compleja parece deberse a la falta de algunas enzimas de la vía de *N*-glicosilación y la presencia de actividades no presentes en mamíferos, obteniendo generalmente glicanos altos en manosa y paucimanosídicos. Limitando su uso en la producción de glicoproteínas que requieren la presencia de glicanos complejos con ácido siálico terminal (Palomares *et al*, 2002; Tomiya *et al*, 2003a).

Recientes estudios muestran que algunas líneas celulares de insecto son capaces de producir proteínas recombinantes con glicosilación compleja e incluso con ácido siálico terminal. Se propone que algunos insectos tienen el potencial de producir glicanos complejos con ácido siálico terminal, aunque los mecanismos y condiciones para que se lleve a cabo esta actividad no han sido deducidos hasta ahora (Jarvis *et al*, 1998; Andersen *et al*, 2002; Butters, 2002; Palomares *et al*, 2002).

2.2.4. Glicosilación en Insectos

En comparación con la información disponible sobre el procesamiento de glicoproteínas en mamíferos, es muy poco lo que sabemos sobre las enzimas y otros factores que intervienen en la *N*-glicosilación en las células de insecto. La mayoría de los glicanos producidos en insectos son del tipo paucimanososa o altos en manosa. Las primeras etapas de la vía de *N*-glicosilación son similares a las de los mamíferos, pero en pocos casos el procesamiento llega hasta la sialilación de los glicanos, proceso que es común en humanos y mamíferos. Se cree que esto es debido a niveles insuficientes de las glicosiltransferasas de las últimas etapas de la vía, incluyendo la Sialiltransferasa.

La vía de *N*-glicosilación en insectos comúnmente sufre una divergencia hacia glicanos de tipo paucimanososa, debido a la actividad de la β -*N*-acetilglucosaminidasa. La remoción de GlcNAc del glicano, por la actividad de esta enzima, evita que siga su procesamiento hasta la adición de galactosa y ácido siálico (Donaldson *et al*, 1999; Licari *et al*, 1993; März *et al*, 1995; Watanabe *et al*, 2002). Esta actividad enzimática ha sido detectada en cultivos de células Sf21 (*Spodoptera frugiperda*), Bm-N (*Bombyx mori*), Mb0503 (*Mamestra brassicae*) y Tn-5B1-4 (*Trichoplusia ni*).

Diversas líneas celulares de insecto, incluyendo Sf9, Sf21, Mb0503, Bm-N, Ea4 (*Estigmene acrea*) y Tn-5B1-4, contienen niveles apenas detectables de actividad *N*-acetil Glucosaminiltransferasa II (GlcNAcT II)

(Figura 2.2.-F). Así como niveles apenas detectables de actividad Galactosiltransferasa. Recientemente, se reportó la clonación y caracterización de la β -N-acetil Galactosaminiltransferasa de *Trichoplusia ni* (Vadaie et al, 2004). Confirmando que algunos insectos poseen las glicosiltransferasas necesarias para formar glicanos complejos.

La actividad sialiltransferasa no ha sido detectada en diversas líneas celulares de insecto, tales como Sf9, Sf21, Tn-5B1-4, Mb0503 y Ea4 (Lopez et al, 1999; Hooker et al, 1999; Tomiya et al, 2004). Algunas de ellas producen ácido siálico, aunque en niveles sumamente bajos y únicamente en ciertas etapas de desarrollo. En *D.melanogaster* el ácido siálico está presente en las primeras etapas de desarrollo, en embriones desde blastodermo hasta larvas de *instar* 3 (Roth et al, 1992). Localizado principalmente en el sistema nervioso.

Las líneas celulares de insecto tienen diferentes capacidades para la producción y procesamiento de glicoproteínas recombinantes. Hasta ahora solo las líneas celulares DpN1 (*Danaus plexippus*), Sf21, Tn4h, Tn4s (*Trichoplusia ni*) y Mb0503 han sido capaces de producir cantidades significativas (5-63%) de proteínas con glicanos complejos y con terminales sialiladas (13-33%) (Palomares et al, 2002; Palomares et al, 2003; Joosten et al, 2003a; Joosten et al, 2003b).

Aunque *D.melanogaster* también contiene glicoproteínas del tipo paucimanosa o altas en manosa principalmente (Seppo et al, 2000). Al identificar los genes Neu5Ac fosfato sintasa y α 2,6-Sialiltransferasa (ST6Gal) y demostrar que son funcionales, se confirmó que este organismo posee la maquinaria necesaria para realizar el procesamiento hasta la obtención de glicanos complejos (Fabini et al, 2001; Fabini et al, 2002; Kim et al, 2002; Koles et al, 2004). Sintetizando y transfiriendo el ácido siálico en glicoproteínas.

La expresión de la ST6Gal de *Drosophila* está regulada en forma espacio-temporal. Ocurre únicamente en algunas etapas del desarrollo embrionario en grupos de células del sistema nervioso de este insecto (Fabini et al, 2002; Beitenbaugh et al, 2004; Koles et al, 2004). Estos datos indican que la sialilación ocurre en algunos insectos y que este proceso es específico de ciertos tipos de células y/o de algunas etapas de desarrollo.

La expresión de los genes de las últimas etapas de la vía de N-glicosilación (incluyendo la Sialiltransferasa) y los que realizan la síntesis del ácido siálico (CMP-Neu5Ac), parecen estar reprimidos o altamente controlados. Es un evento altamente especializado, ocurriendo únicamente en ciertas etapas de desarrollo y tejidos específicos de los insectos (Medvedova et al, 2003).

Recapitulando, la incapacidad de los insectos de producir glicoproteínas similares a las de mamíferos es atribuida a tres factores principales: (1) la presencia de una alta actividad β -N-acetilglucosaminidasa, (2) que no son capaces de producir cantidades significativas de algunos precursores, incluyendo el CMP-Neu5Ac y (3) expresión insuficiente de ciertas glicosiltransferasas implicadas en la generación de glicanos complejos.

Identificadas ya las principales limitantes, en diferentes grupos de investigación se han llevado a cabo trabajos tendientes a la modificación de la vía de *N*-glicosilación en insectos, con el fin de obtener glicoproteínas con terminales de ácido siálico.

Al suprimir la actividad *N*-acetilglucosaminidasa en células Tn-5B1-4 y Sf9 con un inhibidor químico, se produjeron glicanos complejos sialilados (Watanabe *et al.*, 2002). Intuyendo así que estas células poseen actividad sialiltransferasa intrínseca.

Ya que los insectos tienen capacidades limitadas en la producción de CMP-Neu5Ac, sobreexpresaron enzimas de su vía de síntesis en células Sf9. Se observó generación de cantidades importantes de ácido siálico aún en cultivos en medios libre de suero fetal bovino, pero se requirió suplementar con el precursor *N*-acetil Manosamina (Lawrence *et al.*, 2001; Hollister *et al.*, 2003).

Utilizando ingeniería de vías metabólicas, sobreexpresaron GlcNAcT II y GalT de mamífero en células Tn5B1-4 (*Trichoplusia ni*). Esto resultó en la producción de transferrina con glicanos complejos (Tomiya *et al.*, 2003b). El incluir GlcNAcT II, GalT y α 2,6-Sialiltransferasa bovina, permitió la obtención de glicoproteínas sialiladas en células Sf9 (*Spodoptera frugiperda*) (Hollister *et al.*, 1998; Hollister *et al.*, 2001; Hollister *et al.*, 2002). El medio de cultivo contenía suero fetal bovino, que sirvió como suministro de ácido siálico para las células.

Al combinar las glicosiltransferasas (GlcNAcT I, GlcNAcT II, GalT, α 2,3-Sialiltransferasa y α 2,6-Sialiltransferasa) con enzimas de la ruta de síntesis de ácido siálico (Neu5Ac fosfato sintasa y CMP-Neu5Ac sintasa) en la línea celular Sf9 se logró producir glicoproteínas sialiladas (Hollister *et al.*, 2002; Aumiller *et al.*, 2003). Resultados similares se observaron en la línea celular Ea4 (*Estigmene acrea*), confirmando la utilidad de este tipo de metodologías (Chang *et al.*, 2003).

Podemos deducir con base en los antecedentes revisados, que algunos insectos poseen los genes necesarios para el procesamiento de *N*-glicanos hasta su forma compleja, con terminales de galactosa y ácido siálico. Los niveles de expresión no son suficientes para producir o procesar glicoproteínas en niveles significativos. Este es un proceso sumamente regulado que ocurre bajo condiciones específicas, ya sea de cultivo, etapa de desarrollo o de disponibilidad de nutrientes.

2.3. α 2,6-Sialiltransferasa

Las Sialiltransferasas son un subgrupo de la familia de las glicosiltransferasas. Utilizan CMP-Neu5Ac como azúcar activada donadora para catalizar la transferencia de ácido siálico en las cadenas de oligosacáridos de glicoproteínas y glicolípidos (Paulson *et al.*, 1989; Harduin-Lepers *et al.*, 1995). Cada enzima puede caracterizarse enzimáticamente por la especificidad del aceptor y por el tipo de enlace que forma.

2.3.1. Actividad Enzimática de la α 2,6-Sialiltransferasa

Han sido clonadas más de veinte diferentes sialiltransferasas de vertebrados. Se dividen en subfamilias de acuerdo al tipo de enlace con el que unen el Neu5Ac transferido a su aceptor. Estas son: α 2,3-

Sialiltransferasa (ST3Gal), α 2,6-Sialiltransferasa y α 2,8-Sialiltransferasa (ST8Gal). Pueden diferir también en la especificidad por el aceptor. De los tres tipos de enlaces, el α 2,6 es el más común en las glicoproteínas de mamíferos, principalmente sobre un residuo de galactosa (ST6Gal) y ocasionalmente sobre una *N*-acetilgalactosamina (ST6GalNAc) (Harduin-Lepers *et al.*, 1995; Katsutoshi, 1996; Koles *et al.*, 2004).

La α 2,6-Sialiltransferasa (**EC 2.4.99.**) cataliza la siguiente reacción:



La transferencia del ácido siálico hacia galactosa (o *N*-acetil galactosamina) ocurre en el grupo hidroxilo del carbono 6 en la posición terminal de un disacárido Gal β 1-4GlcNAc. Este puede ser un disacárido libre o un oligosacárido de un *N*-glicano. Las Sialiltransferasas son enzimas altamente específicas en el tipo de aceptor y enlace producido durante la transferencia.

Se han clonado y caracterizado las α 2,6-Sialiltransferasas de diferentes vertebrados tales como: bovino (*Bos taurus*), rata (*Rattus norvegicus*), ratón (*Mus musculus*), pollo (*Gallus gallus*) y humano (*Homo sapiens*) (Harduin-Lepers *et al.*, 1995; Datta *et al.*, 1997). En invertebrados únicamente la α 2,6-Sialiltransferasa de la mosca de la fruta (*Drosophila melanogaster*) (Koles *et al.*, 2004; Harduin-Lepers *et al.*, 1995).

2.3.2. Organización de la α 2,6-Sialiltransferasa

Estructuralmente, las sialiltransferasas tienen una conformación similar y comparten la característica de ser glicoproteínas transmembranales de tipo II que residen en el compartimento trans-Golgi. Poseen un dominio citoplasmático corto *N*-terminal (**Figura 2.4-A**), un segmento transmembranal para su retención en Golgi (**Figura 2.4-B**) y un tallo de longitud variable (**Figura 2.4-C**) de entre 20 y 200 aminoácidos de longitud que es seguido del dominio catalítico C-terminal (**Figura 2.4-D**) (Paulson *et al.*, 1989; Breton *et al.*, 1999; Breton *et al.*, 2001). Al igual que otras glicosiltransferasas, la Sialiltransferasa puede también ser encontrada en forma soluble al sufrir un corte proteolítico en la región del tallo (Fenteany *et al.*, 2005). Sin afectar su actividad enzimática.

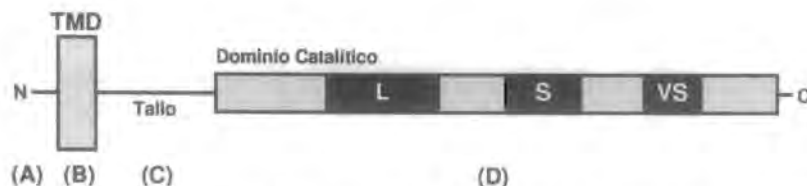


Figura 2.4. Representación esquemática de la α 2,6-Sialiltransferasa. El dominio citoplasmático *N*-terminal (**A**), el dominio transmembranal (**B**), el tallo (**C**) y el dominio catalítico (**D**), donde se localizan los *sialylmotifs* (**L**, **S** y **VS**). Adaptado de Breton *et al.*, 1999.

Las Sialiltransferasas de diferentes organismos comparten el mismo sustrato donador (CMP-Neu5Ac) y reconocen los mismos sustratos aceptores. Es de esperarse que compartan parte de su secuencia proteínica. Las secuencias de aminoácidos de las sialiltransferasas clonadas muestran poca homología con excepción de pequeñas regiones de secuencias en la región catalítica llamadas *Sialylmotifs* (Drickamer,

1893; Livingston *et al.*, 1993; Harduin-Lepers *et al.*, 1995). El **Sialylmotif L** (*Large*) es de aproximadamente 47 aminoácidos de longitud de los cuales 8 son invariables, el **Sialylmotif S** (*Small*) de alrededor de 26 aminoácidos y el **Sialylmotif VS** (*Very Small*) de 6 aminoácidos, estos últimos con 2 residuos invariables cada uno (Figura 2.5.) (Datta *et al.*, 1995; Datta *et al.*, 1997; Geremia *et al.*, 1997; Datta *et al.*, 1998; Jeanneau *et al.*, 2004).

Sialylmotif L		
mST6Gal	189	- CTKCAVVSSAGSLKNSQLCREIDNHDAVLRFNFGAPTDNFQODVGTKT - 236
hST6Gal	192	- WGRCAVVSSAGSLKSSQLCREIDHDAVLRFNFGAPTANFQODVGTKT - 239
rST6Gal	189	- WQRCVVSSAGSLKNSQLCREIDNHDAVLRFNFGAPTDNFQODVGSKT - 236
bST6Gal	189	- WGRCAVVSSAGSLKSSRLCREIDHDAVLRFNFGAPTVKFQODVGTKT - 236
cST6Gal	198	- LGRCAVVSSAGSLKSSHLPEIDSHDAVLRFNFGAPVKGQEDVGOKT - 245

Sialylmotif S		Sialylmotif VS		
mST6Gal	318	- PSSCMLGIIIMMTLQDQVDIYEF - 341	367	- HPLLE - 373
hST6Gal	321	- PSSCMLGIIIMMTLQDQVDIYEF - 344	370	- HPLLYE - 376
rST6Gal	318	- PSSCMLGIIIMMTLQDQVDIYEF - 341	367	- HPLLE - 373
bST6Gal	318	- PSSCMLGIAIMMSLQDQVDIYEF - 341	367	- HPLLE - 373
cST6Gal	328	- PSSCMLGIVIMMTLQDEVVYEF - 351	377	- HPLLE - 383

Figura 2.5. Región del **Sialylmotif L**, **S** y **VS** en las α 2,6-Sialiltransferasas. El alineamiento de secuencias muestra la región más conservada en todas las sialiltransferasas. Los residuos que son estrictamente invariables se muestran sombreados en gris oscuro. Los residuos altamente conservados están sombreados en gris claro. mST6Gal (*Mus musculus*), hST6Gal (*Homo sapiens*), rST6Gal (*Rattus norvegicus*), bST6Gal (*Bos taurus*), cST6Gal (*Gallus gallus*).

Estos *sialylmotifs* parecen ser específicos de enzimas eucariotas, ya que no están presentes en las enzimas que catalizan reacciones similares en procariontas (Jeanneau *et al.*, 2004).

La importancia de estos motivos conservados fue analizada mediante mutación sitio dirigida. Usando como modelo la α 2,6-Sialiltransferasa de rata, reemplazando los residuos conservados por alanina permitió concluir que el *sialylmotif L* está implicado en la unión del sustrato donador (CMP-Neu5Ac). El *sialylmotif S* está implicado en la unión y reconocimiento tanto del sustrato donador como del aceptor. La mutación de ambos *sialylmotifs* produjo enzimas inactivas soportando la hipótesis de que los residuos en estos motivos están implicados en la unión y reconocimiento de sustratos donador y aceptor (Datta *et al.*, 1995; Datta *et al.*, 1997; Datta *et al.*, 1998; Breton *et al.*, 1999). Las cisteínas invariables en los *sialylmotifs L* y *S* participan en la formación de un puente disulfuro que es esencial para la conformación de formas activas de la proteína (Drickamer, 1993; Datta *et al.*, 2001). El *sialylmotif VS* está implicado en el mecanismo de catálisis de la enzima (Geremia *et al.*, 1997). La región comprendida entre los *sialylmotifs S* y *VS* contiene cuatro residuos altamente conservados de importancia funcional, al substituirlos por alanina se produjeron enzimas inactivas (Jeanneau *et al.*, 2004). Además de estar implicados en la correcta conformación de la proteína y en la unión del sustrato aceptor.

Ha sido propuesto que las sialiltransferasas se glicosilan mediante un mecanismo autocatalítico, siendo una forma de control de la actividad enzimática. La α 2,6-Sialiltransferasa requiere al menos glicanos galactosilados para una actividad enzimática óptima *in vivo*, ya que de lo contrario se observa un

decremento de hasta un 60% de la actividad, siendo un componente crítico (Breen, 2002). En la caracterización de la α 2,6-Sialiltransferasa recombinante de *D.melanogaster*, ésta presentó actividad de autosialilación *in vivo* e *in vitro* (Koles *et al.*, 2004). Aunque no se determinó si esta modificación tiene algún efecto en la actividad, estabilidad o localización intracelular de la proteína.

2.3.3. ADNc que Codifican para la α 2,6-Sialiltransferasa

Hasta ahora se han identificado, clonado y caracterizado las α 2,6-Sialiltransferasas (ST6Gal) de cinco vertebrados y de un invertebrado. Los números de acceso se muestran en la **Tabla 2.1.**

TABLA 2.1. α 2,6-Sialiltransferasas identificadas, clonadas y caracterizadas.

Organismo	No. Acceso (GenBank)	
	ADNc	Proteína
Humano (<i>Homo sapiens</i>)	X62822	CAA44634
	X17247	CAA35111
	X54363	CAA38246
Rata (<i>Rattus norvegicus</i>)	M18769	AAA41196
	M54999	AAA41196
	M83143	AAB07233
	M73987	AAB00269
	NM147205	NP671738
Ratón (<i>Mus musculus</i>)	D16106	BAA03680
Pollo (<i>Gallus gallus</i>)	X75558	CAA53235
Bovino (<i>Bos taurus</i>)	Y15111	CAA75385
Mosca (<i>Drosophila melanogaster</i>)	NM166684	NP726474
	NM079129	NP523853

Los análisis entre secuencias de Sialiltransferasas sugieren un origen genético común, a partir de duplicaciones sucesivas de un gen ancestral y evolución divergente (Harduin-Lepers *et al.*, 2005). Estas ST6Gal transfieren ácido siálico (CMP-Neu5Ac) formando un enlace α 2,6 sobre la galactosa terminal de un *N*-glicano Gal β 1-4GalNAc de glicoproteínas.

El análisis de evolución de las Sialiltransferasas realizado a partir de los genomas de diversos organismos reveló que las α 2,6-Sialiltransferasas están presentes únicamente en vertebrados y artrópodos (insectos). Los organismos equinodermos, urocordados y cefalocordados poseen solamente los genes ST6GalNAc, ST3Gal y ST8Gal, respectivamente. Sugiriendo que el gen precursor de las Sialiltransferasas de vertebrados puede ser la ST6Gal de insectos, al no encontrar estos genes en otros invertebrados (Harduin-Lepers *et al.*, 2005).

Los genes y ADNc de α 2,6-Sialiltransferasas han sido clonados mediante diferentes metodologías, ya que varían en abundancia y localización dentro de sus organismos de origen. Debido a que en general las glicosiltransferasas y las sialiltransferasas son proteínas poco abundantes ha sido difícil la purificación de estas enzimas. Por esta razón las estrategias de clonación de genes o ADNc basados en secuencias de proteínas purificadas han sido posibles en pocos casos (Harduin-Lepers *et al.*, 1995; Tsuji, 1995; Tsuji, 1996).

Otras metodologías empleadas han sido: a partir de un banco de ADNc y utilizando como sonda un oligonucleótido correspondiente al N-terminal, lograron identificar y aislar la α 2,3-Sialiltransferasa de cerdo. Utilizando oligonucleótidos degenerados, se aisló el ADNc de la α 2,3-Sialiltransferasa de rata (Weinstein *et al.*, 1982; Gillespie *et al.*, 1992; Wen *et al.*, 1992a). Con el uso de anticuerpos policlonales y una librería bacteriana preparada a partir de ARNm de hígado de rata, una clona fue identificada y purificada conteniendo la α 2,6-Sialiltransferasa de este organismo (Weinstein *et al.*, 1987). Posteriormente, a partir de dos oligonucleótidos diseñados con la secuencia de la sialiltransferasa de rata, se analizó una librería bacteriana y se aisló la α 2,6-Sialiltransferasa de humano (Lance *et al.*, 1989; Grundmann *et al.*, 1990).

La α 2,3-Sialiltransferasa humana fue clonada al expresar en células Burkitt-Namalwa una librería de ADNc e identificando células resistentes a la lectina RCA120 (*Ricinus communis* aglutinina) que detecta Gal β 1-4GlcNAc (Sasaki *et al.*, 1993).

Analizando las secuencias de las sialiltransferasas previamente clonadas se logró identificar regiones conservadas que ahora conocemos como *sialylmotifs*. Estas regiones pueden ser usadas para identificar nuevos genes o ADNc mediante técnicas basadas en PCR. Siendo una alternativa atractiva para la obtención de nuevos miembros de esta familia.

Tres α 2,6-Sialiltransferasas fueron clonadas de una librería de ADNc de embrión de pollo, usando oligonucleótidos degenerados diseñados en los *sialylmotifs* (Kurosawa *et al.*, 1994; Tsuji, 1996). De igual forma, utilizando oligonucleótidos degenerados diseñados en los *sialylmotifs* se clonó la α 2,3-Sialiltransferasa humana, de ratón y rata (Kitagawa *et al.*, 1994a; Kitagawa *et al.*, 1994b; Lee *et al.*, 1994; Tsuji, 1996). La clonación de la α 2,6-Sialiltransferasa de ratón se llevó a cabo mediante el uso de oligonucleótidos degenerados en el *sialylmotif L* (Hamamoto *et al.*, 1993). Las α 2,6-Sialiltransferasas de vertebrados son muy similares, aprovechándose de esto, a partir del ADNc humano se identificó y clonó su ortólogo bovino (Mercier *et al.*, 1999).

La primera Sialiltransferasa de invertebrados en ser identificada, clonada y caracterizada fue la de *Drosophila melanogaster* (Fabini *et al.*, 2001; Fabini *et al.*, 2002; Koles *et al.*, 2004). Fue posible ya que su genoma está secuenciado. Actualmente se dispone de la secuencia completa del ADNc, su organización intrón/exón, su especificidad por sustrato, constantes catalíticas, además se conoce como se regula su expresión así como los tejidos y etapas de desarrollo donde este transcrito está presente.

Identificados y clonados los ADNc de Sialiltransferasas de diversos organismos, fue posible el desarrollo de estrategias para determinar su localización intracelular, especificidades por sustratos y análisis de expresión genética de estas enzimas.

El ADNc de la α 2,6-Sialiltransferasa humana fue aislado de placenta y de células B. Al comparar estos dos transcritos se observó que poseen la misma región codificante, pero difieren en la región no traducida 5'. Posteriormente, se identificó un tercer transcrito que no contiene dos exones de la región no traducida 5' en células de hepatocarcinoma humano (Grundmann *et al*, 1990; Aasheim *et al*, 1993; Wang *et al*, 1993; Harduin-Lepers *et al*, 1995). Cada uno de los transcritos se produce a partir de diferentes promotores dentro del mismo gen, y estos se usan en forma tejido-específica. Con la α 2,3-Sialiltransferasa humana sucede un fenómeno similar ya que se han identificado transcritos de diferentes tamaños que están bajo regulación tejido-específica.

El control de la *N*-glicosilación de glicoproteínas producidas en huéspedes no humanos es actualmente uno de los mayores retos en biotecnología. Recientemente, uno de los objetivos en nuestro laboratorio ha sido el estudio del procesamiento de glicoproteínas producidas en el sistema células de insecto-baculovirus (SCIBV). Tomando en cuenta las ventajas que ofrece y las limitaciones en cuanto a sus capacidades de *N*-glicosilación, al no realizar sialilación constitutivamente al producto recombinante. Se han desarrollado diversas estrategias para obtener las glicoproteínas recombinantes con el patrón de glicosilación lo más cercano a su contraparte nativa.

Por estas razones, para evaluar el potencial real de sialilación que poseen los insectos nos fijamos por objetivo el identificar y clonar la última enzima de la vía de *N*-glicosilación: la Sialiltransferasa. Como material biológico se utilizó el mosquito *Anopheles albimanus*, insecto que es capaz de producir sialoglicoproteínas. Además se dispone de la información del genoma de *An.gambiae* organismo cercano filogenéticamente a *An.albimanus*.

La identificación de dicha enzima y de su ADNc se puede constituir en una herramienta importante complementando la vía de *N*-glicosilación en células de insecto tradicionalmente usadas para la producción de glicoproteínas mediante baculovirus recombinantes.

Objetivos

Objetivo General

Buscar evidencias de la presencia del gen que codifica para la α 2,6-Sialiltransferasa en tejido de mosquito *Anopheles albimanus*.

Objetivos Específicos

- a) Obtener la totalidad de la secuencia codificante del ADNc de la α 2,6-Sialiltransferasa de *Anopheles albimanus*.
- b) Identificar en la secuencia proteica deducida los dominios y estructuras característicos que son comunes a las sialiltransferasas, mediante el uso de herramientas bioinformáticas.
- c) Realizar un análisis ontogenético de la expresión del transcrito de la α 2,6-sialiltransferasa identificado.

3. MATERIALES Y MÉTODOS

3.1. *Búsqueda y Análisis de Secuencias*

Se tienen reportados ADNc de α 2,6-sialiltransferasas de diferentes organismos, éstos se muestran en la **Tabla 2.2**. A partir de estas secuencias de ADNc y de sus respectivas proteínas, se realizaron búsquedas por medio de **BLAST** en diferentes bases de datos de acceso público, que contienen almacenados los resultados del programa de secuenciación del genoma del mosquito *Anopheles gambiae* (**GenBank**, **ENSEMBL** y **TIGR**).

Las secuencias obtenidas como resultado del **BLAST** se compararon con secuencias de las α 2,6-sialiltransferasas ya reportadas.

Los resultados de las comparaciones por medio de alineamientos arrojaron porcentajes de identidad y presencia de *sialylmotifs*. Se determinó cual secuencia corresponde a un gen ortólogo de la α 2,6-sialiltransferasa en el genoma de *Anopheles gambiae*.

Las direcciones electrónicas de las bases de datos y de los programas en línea utilizados se muestran en el **Anexo** de esta tesis.

3.2. *Diseño de Oligonucleótidos*

Se diseñaron oligonucleótidos específicos a partir de secuencias de ADN ya caracterizadas. Se diseñaron oligonucleótidos degenerados a partir del consenso de secuencias de ADN o de un consenso de secuencias de aminoácidos.

Para el diseño de oligonucleótidos específicos es necesario tomar en cuenta: la longitud de los mismos, el tamaño de la amplificación, la T_a (temperatura de alineamiento), el %GC (que debe ser entre 40 y 60%), la formación de horquillas y evitar la formación de dímeros. Todo este tipo de cálculos pueden realizarse con la ayuda de programas disponibles en internet, tales como: **Oligonucleotide Properties Calculador**, **Biopolymer calculador**, **OligoAnalyzer 3.0**, y algunos otros.

En el caso de oligonucleótidos degenerados se requiere de un alineamiento de secuencias, ya sea de nucleótidos o de aminoácidos. Este tipo de análisis se realizan con programas especializados tales como: **MultAlin**, **LALIGN** y **THE MEME/MAST SYSTEM**.

En el caso de un consenso de nucleótidos, se localizan las regiones más conservadas y de menor grado de degeneración. Se analizan segmentos de talla especificada en sus propiedades de T_a , %GC, formación de horquillas y dímeros, hasta encontrar aquellas que posean las mejores características. En el caso de un consenso de aminoácidos, se localizan las regiones conservadas y dentro de ellas, aquellos residuos que estén codificados por el menor número de codones de acuerdo al código genético. Una vez definida una zona se identifican los codones correspondientes a cada aminoácido, a partir del código

genético y del uso de codones específico. discriminando aquellos codones que son poco frecuentes y prefiriendo aquellos más frecuentes, definiendo así las secuencias preferenciales de ADN.

Todos los oligonucleótidos diseñados y usados en este trabajo (salvo aquellos explícitamente especificados) fueron sintetizados por la Unidad de Síntesis y Secuenciación de ADN del Instituto de Biotecnología/UNAM.

3.3. Material Biológico

Mosquitos *Anopheles albimanus* en las etapas de huevecillos, larvas, pupas, adultos machos y adultos hembras, fueron amablemente donados por el Bioterio del I.N.S.P. (Cuernavaca, Morelos), estos fueron cultivados a 25-28°C, 80% humedad, bajo ciclos de 12 h luz-oscuridad, y mantenidos en solución 10% sacarosa durante las etapas adultas.

Las muestras que no se procesaron inmediatamente, se almacenaron congeladas a -70°C.

3.4. Extracción de ARN total

Se utilizó el método de purificación de ARN por extracción con fenol ácido - tiocianato de guanidina - cloroformo. Modificación del método Chomczynski-Sacchi (Chomczynski *et al*, 1987).

MATERIALES UTILIZADOS:

1. **Mezcla cloroformo-alcohol isoamílico.** Mezcla 49:1 v/v de cloroformo y alcohol isoamílico.
2. **Etanol.** Grado biología molecular libre de nucleasas.
3. **Isopropanol.** Grado biología molecular libre de nucleasas.
4. **Nitrógeno líquido.**
5. **Fenol ácido.** Equilibrado con buffer de citratos a pH 4.
6. **Acetato de sodio 2M, pH 4.0.**
7. **Solución D.** 4 M tiocianato de guanidina, 25 mM citrato de sodio·2H₂O, 0.5% (w/v) lauril sarcosina de sodio, 0.1 M β-mercaptoetanol.
8. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μm.
9. **Mortero y pistilo.** Tratados con DEPC-H₂O.

PROCEDIMIENTO:

Usar guantes y cubreboca, trabajar en campana de flujo laminar.

1. Preparar la muestra para purificación de ARN.
 - a) Colectar la muestra (fresca o congelada a -70°C) y colocarla en mortero.
 - b) Adicionar nitrógeno líquido y pulverizar la muestra con pistilo.
 - c) Transferir la muestra pulverizada a un tubo limpio estéril libre de nucleasas y adicionar 2 ml de solución D por cada 50 larvas de mosquito o 1x10⁶ células.
 - d) Homogenizar la muestra en vortex por 15-30 segundos.

2. Transferir la muestra a un tubo limpio y secuencialmente adicionar 0.1 ml de acetato de sodio 2 M pH 4.0, 1 ml de fenol ácido, y 0.2 ml de mezcla cloroformo-isoamílico por mililitro de solución D. Después de la adición de cada reactivo, tapar el tubo y mezclar completamente por inversión.
3. Mezclar vigorosamente con vortex por 10 segundos. Incubar el tubo por 15 min en hielo.
4. Centrifugar el tubo a 10000 g por 20 min a 4°C, y después transferir la fase superior acuosa que contiene el ARN extraído a un tubo nuevo.
5. Adicionar 1 volumen de isopropanol a la fase acuosa de ARN. Mezclar por inversión y permitir precipitación del ARN por 1 hora a -20°C.
6. Recolectar el ARN precipitado por centrifugación a 10000 g por 30 min a 4°C.
7. Cuidadosamente decantar el isopropanol y disolver el pellet en 0.3 ml de solución D por cada mililitro de esta solución usada en paso 1.
8. Transferir la solución a un tubo nuevo, mezclar por vortex, y precipitar el ARN con 1 volumen de isopropanol por 1 hora a -20°C.
9. Recolectar el ARN precipitado por centrifugación a máxima velocidad por 10 min a 4°C.
10. Decantar el sobrenadante y lavar el *pellet* con etanol 75%, centrifugar a máxima velocidad por 10 min a 4°C.
11. Secar la muestra hasta que el etanol se haya evaporado, pero no permitir que el pellet se seque completamente.
12. Adicionar 50-100 µl de agua libre de nucleasas, mezclar hasta disolver el pellet de ARN y almacenarlo a -70°C.
13. Determinar la concentración de ARN por medio de espectrofotometría midiendo absorbancia a 260 nm de una alícuota del ARN purificado.

3.5. Tratamiento de ARN con DNasa I

Para remover ADN contaminante de las muestras de ARN y subsecuentemente remover la DNasa I y los cationes divalentes de la muestra, se utilizó el producto **DNA-free™ (DNAase Treatment & Removal Reagents, Ambion)**. Con este procedimiento, es posible remover cantidades pequeñas o moderadas de ADN contaminante (hasta 10 µg/µl) del ARN purificado a un nivel indetectable por RT-PCR.

COMPONENTES UTILIZADOS:

1. **DNasa I, libre de RNasa (2 unidades/µl)**. Una unidad de DNasa I es definida como la cantidad de enzima que degradará 1 µg de ADN en 10 min a 37°C.
2. **10X DNase Buffer**. 100 mM Tris-Cl pH 7.5, 25 mM MgCl₂, 1 mM CaCl₂
3. **DNase Inactivation Reagent**.
4. **Agua libre de nucleasas**. Agua desionizada grado biología molecular filtrada a 0.1 µm.

PROCEDIMIENTO:

Se siguió el protocolo del fabricante, haciendo algunas modificaciones aquí mostradas.

1. Una vez cuantificada la muestra de ARN diluir la alícuota a tratar a 100 µg/ml con agua libre de nucleasas. Adicionar 0.1 volúmenes de 10X DNase Buffer y 1 µl de DNasa I por cada 50 µl de muestra de ARN.
2. Mezclar cuidadosamente e incubar a 37°C por 1 hora.
3. Resuspender el DNase Inactivation Reagent por medio de vortex. Adicionar 0.2 volúmenes a la muestra.
4. Mezclar con cuidado e incubar por 10 min a temperatura ambiente. Mezclar cada minuto durante el tiempo de incubación para evitar precipitación del DNase Inactivation Reagent.
5. Centrifugar a 10000 g por 2 min hasta formación de un precipitado que contiene el DNase Inactivation Reagent, el ADN contaminante, cationes divalentes y la DNasa I.
6. Transferir el ARN tratado a un tubo nuevo.
7. Determinar la concentración de ARN por medio de espectrofotometría midiendo absorbancia a 260 nm de una alícuota del ARN tratado.

3.6. Síntesis de ADNc

La síntesis de la primera cadena del ADNc a partir de un templado de ARN es realizada por una Transcriptasa Reversa. En este caso, una versión modificada genéticamente de *Moloney Murine Leukemia Virus Reverse Transcriptase* (RevertAid™ M-MuLV RT, Fermentas) con baja actividad de RNasa H. La síntesis de la primera cadena del ADNc puede realizarse a partir de dos tipos de oligonucleótidos diferentes:

- ▶ **Oligo(dT)₁₈**: en el extremo 3' de los ARNm poli(A)⁺, sólo los ARNm con tallo poli(A) son templados para la síntesis del ADNc.
- ▶ **Oligo Secuencia Específica**: a partir de un sitio específico de unión de los oligonucleótidos utilizados sobre el ARNm.

El ADNc sintetizado puede ser utilizado como templado en PCR, ya que las condiciones de síntesis de ADNc son compatibles puede adicionarse directamente a la mezcla de reacción PCR. La primera cadena de ADNc sintetizada, puede ser templado para síntesis de la segunda cadena. En este caso, se utilizó el **RevertAid™ First Strand cDNA Síntesis Kit (Fermentas Life Sciences)**.

COMPONENTES UTILIZADOS:

1. **Revert Aid™ M-MuLV Reverse Transcriptase (200u^{**}/µl)**. La enzima en solución en buffer 50 mM Tris-HCl (pH 8.3), 0.1 M NaCl, 1 mM EDTA, 5 mM DTT, 0.1% Triton[®] X-100 y 50% glicerol.
2. **RiboLock™ Ribonuclease Inhibitor (20u^{**}/µl)**. La enzima en solución en buffer 20 mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 8 mM DTT, 0.5 mM ELUGENT[®] Detergent y 50% glicerol.
3. **5X Reaction Buffer**. 250 mM Tris-HCl (pH 8.3), 250 mM KCl, 20 mM MgCl₂ y 50 mM DTT.
4. **10mM dNTP Mix**. Solución acuosa de 10 mM de dGTP, dATP, dTTP y dCTP.
5. **Oligo(dT)₁₈**. Solución acuosa a 0.5 µg/µl.
6. **Oligo Secuencia Específica**. Solución acuosa a 20 pmol/µl de oligonucleótido o mezcla de oligonucleótidos.

7. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μm .

* Una unidad de Revert Aid™ M-MuLV Reverse Transcriptase incorpora 1 nanomol de dTMP en una fracción de polinucleótido en 10min a 37 °C.

** Una unidad de RiboLock™ Ribonuclease Inhibitor inhibe la actividad de 5ng de RNase H en 50%.

PROCEDIMIENTO:

Procedimiento para la síntesis de la primera cadena del ADNc que servirá como templado en PCR.

1. Preparar la siguiente mezcla de reacción en un tubo en hielo:

Templado ARN	ARN total	0.1-5 μg
	ó ARN poli(A) ⁺	10ng-0.5 μg
Oligonucleótido	Oligo(dT) ₁₈ (0.5 $\mu\text{g}/\mu\text{l}$)	1 μl
	u Oligo Secuencia Específica (20pmol/ μl)	1 μl
Agua libre de nucleasas		completar 12 μl

Mezclar con cuidado y coleccionar las gotas con una breve centrifugación.

2. Incubar la mezcla a 70°C por 10 minutos, colocar en hielo y coleccionar las gotas con una breve centrifugación.

3. Colocar el tubo en hielo y adicionar los siguientes componentes en el orden indicado:

5X Reaction Buffer	4 μl
RiboLock™ Ribonuclease Inhibitor (20 μl)	1 μl
10mM dNTP Mix	2 μl

Mezclar cuidadosamente y coleccionar las gotas con una breve centrifugación.

4. Incubar a 37°C por cinco minutos.

5. Adicionar 1 μl de RevertAid™ M-MuLV Reverse Transcriptase (200 μl) para obtener un volumen final de 20 μl .

6. Incubar la mezcla a 43°C por 60 minutos.

7. Detener la reacción calentando a 70°C por 10 minutos. Colocar en hielo.

La primera cadena de ADNc sintetizada puede ser usada directamente para amplificación por PCR.

3.7. Amplificación por PCR

La Reacción en Cadena de la Polimerasa (PCR, *Polymerase Chain Reaction*) permite la producción de millones de copias de la secuencia de un ADN templado a partir de unas pocas copias. La especificidad de la técnica radica en que la muestra no debe estar contaminada con otros ADN o productos previamente amplificados, además del tipo de oligonucleótidos utilizados para la secuencia del templado que se desea amplificar. Se utilizaron la **Taq DNA Polymerase** (NEB), la **Vent_r DNA Polymerase** (NEB) y **ProofStart DNA Polymerase** (QIAGEN). Dependiendo del tipo de templado y de las condiciones de reacción deseadas, se utilizaron estas ADN polimerasas en las diferentes aplicaciones.

COMPONENTES UTILIZADOS:

1. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μm .

2. **10mM dNTP Mix.** Solución acuosa de 10 mM de dGTP, dATP, dTTP y dCTP (Fermentas).
3. **Oligonucleótidos.** Sintetizados por la Unidad de Síntesis y Secuenciación del Instituto de Biotecnología/UNAM. Resuspendidos en agua libre de nucleasas a una concentración de 5 a 25pmol/ μ l.
4. **10X ThermoPol Reaction Buffer.**
5. **100mM MgSO₄.** En solución acuosa.
6. **10X ProofStart DNA Buffer.**
7. **25 mM MgSO₄.** En solución acuosa.
8. **Taq DNA Polymerase.** 5 u/ μ l.
9. **Vent_R DNA Polymerase.** 2 u/ μ l.
10. **ProofStart DNA Polymerase.** 2.5 u/ μ l.

PROCEDIMIENTO:

Para la realización de varias reacciones en forma paralela, se prepara una mezcla que contiene agua libre de nucleasas, 10X Buffer, dNTPs y DNA polimerasa, una vez alicuotada en tubos individuales, se adicionan los oligonucleótidos y templado correspondientes.

1. Descongelar y homogenizar todas las soluciones.
2. Adicionar en un tubo de 0.5ml en hielo lo siguiente:

a) Para reacciones con *Taq* DNA Polymerase:

Reactivo	Concentración Final	Cantidad por 25 μ l de mezcla de reacción
Agua libre de nucleasas	----	variable
10X ThermoPol Buffer	1X	2.5 μ l
10mM dNTP Mix	0.2mM	0.5 μ l
Oligonucleótido Sentido	0.1-2 μ M	variable
Oligonucleótido Antisentido	0.1-2 μ M	variable
<i>Taq</i> DNA Polymerase (5u/ μ l)	1.25u	0.25 μ l
Templado ADN	10pg-1 μ g	variable

b) Para reacciones con Vent_R DNA Polymerase:

Reactivo	Concentración Final	Cantidad por 25 μ l de mezcla de reacción
Agua libre de nucleasas	----	variable
10X ThermoPol Buffer	1X	2.5 μ l
10mM dNTP Mix	0.2mM	0.5 μ l
Oligonucleótido Sentido	0.1-2 μ M	variable
Oligonucleótido Antisentido	0.1-2 μ M	variable
Vent _R [®] DNA Polymerase (2u/ μ l)	1.25u	0.25 μ l
100mM MgSO ₄	2-6mM	variable
Templado ADN	10pg-1 μ g	variable

c) Para reacciones con ProofStart DNA Polymerase:

Reactivo	Concentración Final	Cantidad por 25µl de mezcla de reacción
Agua libre de nucleasas	-----	variable
10X ThermoPol Buffer	1X	2.5µl
10mM dNTP Mix	0.3mM	0.75µl
Oligonucleótido Sentido	0.1-2µM	variable
Oligonucleótido Antisentido	0.1-2µM	variable
ProofStart DNA Polymerase (2.5u/µl)	2.5u	1µl
25mM MgSO ₄	2mM	2µl
Templado ADN	10pg-1µg	variable

3. Mezclar cuidadosamente y centrifugar brevemente para coleccionar las gotas pegadas en las paredes del tubo.

4. Colocar los tubos en termociclador (**PCR Sprint HYBAID**).

5. Cargar, verificar y correr programa de termociclador:

Los tiempos de las diferentes etapas, así como la temperatura de alineamiento dependerán de los oligonucleótidos utilizados, del templado, la longitud de la amplificación y del tipo de DNA polimerasa utilizada en la reacción PCR, que serán específicos en las diferentes aplicaciones.

6. Una vez concluido el programa en el termociclador, verificar productos de amplificación en gel de agarosa 1% teñido con bromuro de etidio y visualizado con transiluminador UV.

REAMPLIFICACIONES *seminested*-PCR

Para generar productos de amplificación específicos o confirmar identidad de amplificaciones, se recurrió a *seminested*-PCR, en las cuales se repite la amplificación, pero con la adición de un nuevo oligonucleótido manteniendo uno de los oligonucleótidos con los cuales se realizó la amplificación original. Esto se ejemplifica en la **Figura 3.1**.

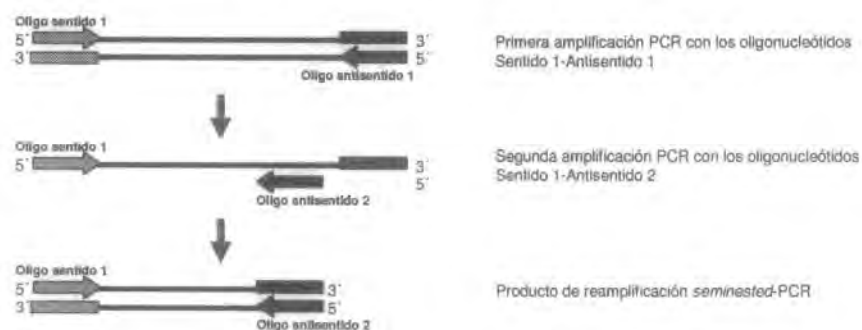


Figura 3.1. Procedimiento *seminested*-PCR

En la **Figura 3.1**, se ejemplifica una *seminested*-PCR donde se mantiene el oligonucleótido sentido que se utiliza en la primera amplificación, y sustituyendo en oligonucleótido antisentido 1 por otro que se encuentra hacia su 3' (oligonucleótido antisentido 2) en una segunda amplificación PCR. Esto mismo

puede realizarse manteniendo el oligonucleótido antisentido 1 y sustituyendo el oligonucleótido sentido, por otro que se encuentre hacia el 3'.

REAMPLIFICACIONES *nested-PCR*

Al igual que la *seminested-PCR*, en este procedimiento se pretende generar productos de amplificación específicos o confirmar identidad de amplificaciones, en este caso ambos oligonucleótidos (sentido y antisentido) son sustituidos en la PCR por otros que se encuentran hacia el 3' de sus respectivas orientaciones sobre el templado, dando un producto de amplificación de talla específica, por lo que este procedimiento nos sirve como método de discriminación de amplificaciones inespecíficas.

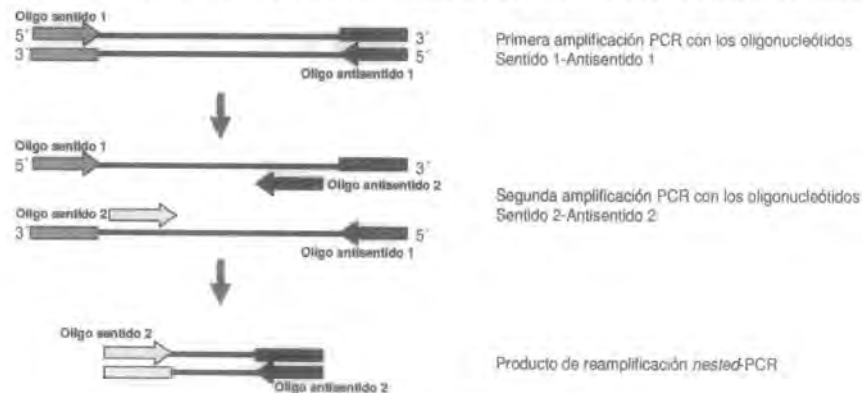


Figura 3.2. Procedimiento *nested-PCR*

En la **Figura 3.2.**, se observa como se realiza la *nested-PCR*, sustituyendo los oligonucleótidos originales (sentido 1 y antisentido 1) por oligonucleótidos internos (sentido 2 y antisentido 2), durante la reamplificación.

3.8. Purificación de Amplificaciones

La purificación de fragmentos de ADN a partir de geles de agarosa se realizó con el **Rapid Gel Extraction System** (Marligen BioScience Inc). Utilizando columnas que contienen una membrana de sílica para capturar y purificar fragmentos de ADN. La muestra purificada de esta forma puede ser utilizada en diferentes aplicaciones incluyendo clonación y amplificación por PCR.

COMPONENTES UTILIZADOS:

1. **Gel Solubilization Buffer (L1).** Contiene perclorato de sodio, acetato de sodio y solubilizador TBE.
2. **Wash Buffer (L2).** Contiene NaCl, EDTA, Tris-HCl y etanol.
3. **Columnas de sílica.**
4. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μm .

PROCEDIMIENTO:

Modificación del protocolo del manual del producto. Las centrifugaciones se realizan a temperatura ambiente.

1. Cortar el fragmento(s) de ADN que se desean purificar directamente del gel de agarosa 1-2% con la ayuda de un bisturí, visualizando el ADN en transiluminador UV.
2. Pesar el fragmento de gel cortado que contiene el ADN a purificar, y adicionar 100 μ l de Gel Solubilization Buffer (L1) por cada 10 mg de gel.
3. Solubilizar el gel en el Buffer L1 incubando a 55°C por 15 min en baño de agua, mezclando cada 3 min para asegurar completa solubilización del gel. Una vez que el gel se disuelva, incubar 10 min adicionales.
4. Colocar una columna de sílica en un tubo eppendorf. Adicionar la mezcla del Paso 3 dentro de la columna y centrifugar a 10000 g por 2 min. Eliminar el líquido colectado en el tubo.
5. Colocar la columna de sílica dentro de otro tubo eppendorf y adicionarle 500 μ l de Buffer L1. Incubar por 2 min a temperatura ambiente y centrifugar a 10000 g por 2 min. Eliminar el líquido colectado en el tubo.
6. Colocar la columna de sílica dentro de otro tubo eppendorf y añadir 700 μ l de Wash Buffer (L2). Incubar por 5 min a temperatura ambiente y centrifugar a 12000 g por 2 min. Eliminar el líquido colectado en el tubo y centrifugar nuevamente a 10000 g por 2 min.
7. Eluir el ADN de la columna de sílica añadiendo 25-50 μ l de agua libre de nucleasas calentada a 70°C. Incubar a temperatura ambiente por 2 min y centrifugar a 12000 g por 2 min.
8. Cuantificar la muestra de ADN por espectrofotometría a 260 nm.

3.9. Clonación de Amplificaciones

La clonación de amplificaciones se realizó por medio del **Zero Blunt® TOPO® PCR Cloning Kit for Sequencing** (Invitrogen Life Technologies). Por medio de ligación directa de productos de PCR con extremos romos (*blunt-end*) en el plásmido vector para secuenciación. Estos son generados por una DNA polimerasa termoestable con *proofreading* (por ejemplo: Vent_R DNA Polymerase) y purificados por medio de **Rapid Gel Extraction System** (Marligen BioScience Inc).

COMPONENTES UTILIZADOS:

1. **TOPO vector pCR®-Blunt II-TOPO®.** 10 ng/ μ l en 50% glicerol, 50 mM Tris-HCl (pH 7.4 @ 25°C), 1 mM EDTA, 2 mM DTT, 0.1% Triton X-100, 100 mg/ml BSA y 30 μ M Bromophenol blue.
2. **Salt Solution.** 1.2M NaCl y 0.06 M MgCl₂.
3. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μ m.
4. **Medio SOC.** 2% Tryptone, 0.5% extracto de levadura, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ y 20 mM glucosa.
5. **Células Competentes TOP-10.** *E. coli* químicamente competentes, preparadas de acuerdo al protocolo mostrado en el **Anexo**.
6. **Placas Medio LB.** Placas de medio LB con 50 μ g de ampicilina/ml o 50 μ g de kanamicina/ml.

PROCEDIMIENTO:

CLONACIÓN

1. Preparar la mezcla de reacción en un tubo de 0.2 ml estéril, adcionando lo siguiente:

Reactivo	Cantidad
Producto PCR a clonar	0.5-4 μ l (20-50 ng de ADN)
Salt Solution	1 μ l
Agua libre de nucleasas	completar 5 μ l
TOPO [®] vector	1 μ l
Volumen Final	6 μl

Mezclar bien los componentes antes de adicionar el TOPO[®] vector, pipetear el TOPO[®] vector con cuidado y adicionarlo a la mezcla, mezclando cuidadosamente con la punta de la pipeta solo con movimientos circulares, no por pipeteo de la mezcla.

2. Incubar la mezcla anterior por 5-30 min a temperatura ambiente (22-25°C).

TRANSFORMACIÓN DE CÉLULAS QUÍMICAMENTE COMPETENTES

3. Colocar la mezcla en hielo y preparar las células competentes a transformar. Esta mezcla puede almacenarse a -20°C para su posterior uso.
4. Descongelar un vial de células competentes TOP-10 al colocarlo en hielo por 5 min aproximadamente.
5. En un tubo, añadir 50 μ l de células competentes TOP-10 y 3 μ l de la mezcla de reacción (Paso 3) y mezclando cuidadosamente con la punta de la pipeta solo con movimientos circulares, no por pipeteo de la mezcla.
6. Incubar en hielo por 5-30 min.
7. Realizar choque térmico incubando la mezcla a 42°C por 30 seg en baño de agua, procurando no agitar la mezcla.
8. Inmediatamente poner el tubo en hielo.
9. Adicionar 250-300 μ l de medio SOC.
10. Agitar colocando el tubo horizontalmente en incubadora a 200 rpm y 37°C por 1-2h.
11. Sembrar en placa de medio LB con antibiótico (50 μ g de kanamicina/ml) 50 μ l del cultivo de células transformadas.
12. Incubar a 37°C por 8-12 h, hasta observar colonias separadas.
13. Proceder a análisis de colonias transformantes.

ANÁLISIS DE TRANSFORMANTES

El vector pCR-Blunt II-TOPO utilizado permite la selección mediante el gen *ccdB* que es letal para *E.coli*. La ligación del inserto no permite la expresión de *ccdB*, permitiendo solo el crecimiento de células que sean transformantes positivas.

Para analizar las transformantes positivas, e identificar las que tienen el inserto en la forma adecuada, se pueden seguir dos alternativas no excluyentes que pueden complementarse y que se pueden aplicar, ya sea para una búsqueda rápida o una más detallada.

1. ANÁLISIS DE TRANSFORMANTES POR PCR

- a) Para cada colonia a analizar, preparar una mezcla de reacción PCR conteniendo lo siguiente:

Reactivo	Concentración Final	Cantidad por 25µl de mezcla de reacción
Agua libre de nucleasas	-----	variable
10X ThermoPol Buffer	1X	2.5 µl
10 mM dNTP Mix	0.2 mM	0.5 µl
Oligonucleótido M13-Reverse	0.4 µM	1 µl @ 10 pmol/µl
Oligonucleótido M13-Forward	0.4 µM	1 µl @ 10 pmol/µl
Taq DNA Polymerase (5 u/µl)	1.25 u	0.25 µl
Templado ADN	10 pg-1µg	variable

Se utilizan los oligonucleótidos M13 (**Tabla 3.1.**) para amplificar una región del vector que contiene el inserto.

- b) Picar las colonias con un palillo estéril y resuspender en la mezcla de reacción PCR.
- c) Colocar los tubos en termociclador (**PCR Sprint HYBAID**).
- d) Cargar, verificar y correr programa de termociclador:
- e) Una vez concluido el programa en el termociclador, verificar productos de amplificación en gel de agarosa 1% teñido con bromuro de etidio y visualizado con transiluminador UV.
- f) Se darán como transformantes positivas aquellas que den amplificación de la talla esperada de la sumatoria del fragmento del vector desde los sitios M13 (Figura) más el tamaño aproximado del inserto clonado.
- g) Se verifican estas amplificaciones por PCR con oligonucleótidos específicos del inserto clonado y/o con patrón de restricción con *EcoRI*.
- h) Las colonias positivas se cultivan por 8-12 h en medio LB o SOC con 50 µg de ampicilina/ml o 50 µg de kanamicina/ml para purificación de plásmido para secuenciación.

2. ANÁLISIS DE TRANSFORMANTES POR PATRÓN DE RESTRICCIÓN

- a) Seleccionar algunas colonias y realizar cultivo en medio LB o SOC con 50 µg de ampicilina/ml o 50 µg de kanamicina/ml por 8-12 h.
- b) Purificar plásmidos.
- c) Realizar análisis de patrón de restricción (con enzima *EcoRI* o alguna dentro del vector o inserto) y comparar tamaños de los fragmentos obtenidos respecto a los esperados con la enzima utilizada.
- d) Se darán como transformantes positivas aquellas que den el patrón de restricción esperado.
- e) Como alternativa, se puede hacer análisis por PCR de los plásmidos purificados.

Una vez identificadas las transformantes positivas, se procede a purificación de plásmido para secuenciación del inserto. Para secuenciar, se dispone de cuatro oligonucleótidos que están dentro de la secuencia del vector, estos son: M13 Forward(-20) y M13 Reverse, además de los oligonucleótidos que generaron el fragmento clonado. **Figura 3.3.**

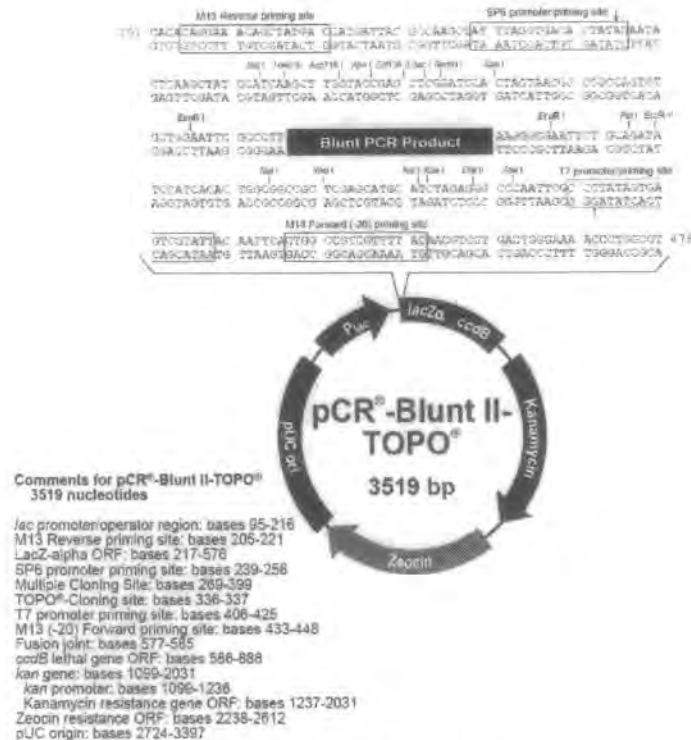


Figura 3.3. Mapa del vector pCR®-Blunt II-TOPO® (Invitrogen).

3.10. Purificación de Plásmidos

La purificación de ADN plasmídico, se realizó con un procedimiento de hidrólisis alcalina/SDS modificado por medio del **Rapid Plasmid Miniprep System** (Marligen BioScience Inc) con columnas de membranas de sílica donde el ADN plasmídico es adsorbido selectivamente. Los plásmidos purificado por este método son aptos para varias aplicaciones, incluyendo secuenciación automatizada de ADN.

COMPONENTES UTILIZADOS

1. **Cell Suspensión Buffer (G1).** 50 mM Tris-HCl (pH 8.0), 10 mM EDTA.
2. **RNasa A.** 20 mg/ml en Cell Suspensión Buffer (G1).
3. **Cell Lysis Solution (G2).** 200 mM NaOH, 1% SDS (w/v).
4. **Neutralization Buffer (G3).** Contiene acetato y guanidina (Formulación del fabricante).
5. **Wash Buffer (G4).** Contiene NaCl, EDTA y Tris-HCl (pH 8.0) (Formulación del fabricante).
6. **Optional Wash Buffer (GX).** Contiene acetato, guanidina hidrocloreto, EDTA y etanol (Formulación del fabricante).
7. **Columnas de sílica.**
8. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μ m.

PROCEDIMIENTO:

Modificación del protocolo del manual del producto. Las centrifugaciones se realizan a temperatura ambiente.

1. Realizar cultivo en medio LB hasta tener una densidad de 1 a 1.5 unidades A_{600} por ml.
2. Tomar de 3-5 ml de medio y centrifugarlo a máxima velocidad por 5 min, eliminar el sobrenadante.
3. Adicionar 250 μ l de Buffer G1 (que contiene la RNasa A) al pellet de células, resuspender las células en el buffer.
4. Adicionar 250 μ l de Buffer G2, mezclar con cuidado por inversión del tubo (5 veces). Incubar a temperatura ambiente por 5 min.
5. Adicionar 350 μ l de Buffer G3, mezclar inmediatamente por inversión del tubo (5 veces). Centrifugar la mezcla a 12000 g por 10 min.
6. Colocar una columna de silica dentro de un tubo eppendorf y añadir el sobrenadante del Paso 5. Centrifugar a 12000 g por 1 min, eliminar el líquido colectado en el tubo.
7. Adicionar 500 μ l de Buffer GX e incubar a temperatura ambiente por 1 min. Centrifugar a 12000 g por 1 min, eliminar el líquido colectado en el tubo.
8. Adicionar 700 μ l de Buffer G4 y centrifugar a 12000 g por 1 min. Eliminar el líquido colectado en el tubo. Centrifugar otra vez a 12000 g por 1 min para remover todo el buffer de la columna.
9. Colocar la columna en un tubo nuevo, para eluir los plásmidos de la columna adicionar 75-100 μ l de agua libre de nucleasas calentada a 65-70°C. Incubar a temperatura ambiente por 2 min y centrifugar a 12000 g por 2 min.
10. Cuantificar el plásmido eluido por espectrofotometría (A_{260}).

3.11. Secuenciación de ADN

La secuenciación de ADN se realizó en la Unidad de Síntesis y Secuenciación de ADN del Instituto de Biotecnología/UNAM, por el método *Taq FS Dye Terminator Cycle Sequencing Fluorescence-Based Sequencing* (Perkin Elmer/Applied Biosystems Modelo 377-18E).

El ADN plasmídico purificado por medio del **Rapid Plasmid Miniprep System** (Marligen BioScience Inc) es verificado en gel de agarosa 1% y cuantificado por espectrofotometría (A_{260}/A_{280}). Se entregan 750-1000 ng de ADN plasmídico en un volumen de 15 μ l. Los oligonucleótidos son adicionados en la Unidad de Síntesis y Secuenciación de ADN del Instituto de Biotecnología/UNAM de acuerdo a la información proporcionada en la solicitud del servicio de cada muestra a secuenciar. **Tabla 3.1.**

TABLA 3.1. Oligonucleótidos para secuenciación.

Oligonucleótido	Secuencia
M13 Forward (-40)	5'-GTTTTCCCAGTCACGTTGTA-3'
M13 Reverse	5'-TTGTGAGCGGATAACAATTTC-3'
T3	5'-TAATACGACTCACTATAGGG-3'
T7	5'-CGCATTTAACCCTCACTAAAG-3'

Después del proceso de secuenciación automatizada por el método ya mencionado, se obtiene el electroferograma resultado de la secuenciación que contiene la secuencia de ADN.

Las secuencias obtenidas de los electroferogramas se analizan por medio de **BLAST** al analizar el porcentaje de identidad, la expectancia, presencia de *gaps*, presencia de *sialylmotifs* (L, S y VS) y algunos otros residuos importantes característicos, respecto a las secuencias de las α 2,6-sialiltransferasas ya reportadas. También se realiza la traducción a aminoácidos para realizar este mismo análisis, determinando de esta forma si las secuencias obtenidas corresponden o no a una α 2,6-sialiltransferasa.

3.12. RACE (Rapid Amplification of cDNA Ends)

Rapid Amplification of cDNA Ends (RACE) es un procedimiento para la amplificación de secuencias de ácidos nucleicos de un templado de ARN mensajero entre un sitio interno definido y cualquiera de los extremos del ARNm, el 3' o el 5'. Esta metodología de amplificación con especificidad hacia algún extremo del templado ha sido descrita como *one-sided PCR* o *anchored PCR*.

El procedimiento RACE ha sido usado para amplificación y clonación de ARNm poco abundantes o análisis de librerías de ADNc. Los productos generados por el 3' y 5' RACE pueden ser combinados para generar ADNc de longitud completa, así como la caracterización de secuencias codificantes desconocidas. Es estrictamente necesario conocer la secuencia específica de algún fragmento del ARNm del cual se desea amplificar los extremos 3' y 5', ya que en este es donde se realizará el diseño de oligonucleótidos específicos a partir de los cuales se harán las amplificaciones hacia los extremos.

3.12.1. Diseño de oligonucleótidos para RACE

Para la amplificación de alguno de los extremos del ADNc de interés, es necesario diseñar dos oligonucleótidos específicos (GSP) para cada procedimiento: 3'-RACE o 5'-RACE. Estos oligonucleótidos son diseñados sobre un fragmento de ADN conocido, al que se le llama *región de solapamiento*, para la amplificación del extremo 5' se utilizan oligonucleótidos antisentido específicos (GSP1 y NGSP1) y para la amplificación del extremo 3' se utilizan oligonucleótidos sentido específicos (GSP2 y NGSP2) (Figura 3.4.).

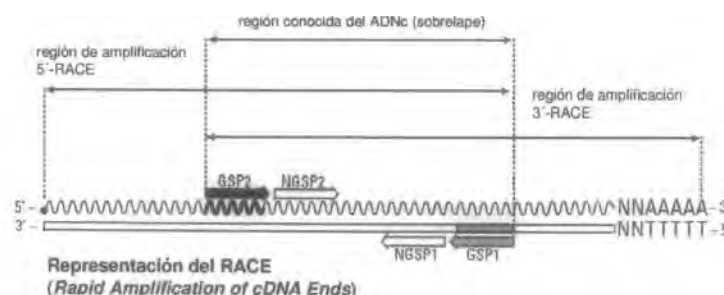


Figura 3.4. Representación de procedimientos RACE y localización de oligonucleótidos específicos sobre fragmento de ADNc conocido (Modificado de BD SMART RACE cDNA Amplification Kit User Manual).

Los oligonucleótidos específicos sentido y antisentido diseñados deben ser de 23 a 28 bases de longitud, con un 50-70% GC y una T_m de al menos 65°C o si es posible, cercana a los 70°C. Evitar además que sean secuencias complementarias entre los mismos oligonucleótidos y con las secuencias de los adaptadores añadidos en los extremos del ADNc.

3.12.2. 5'-RACE (*Rapid Amplification of cDNA Ends*)

Para la síntesis del ADNc para el 5'-RACE se utilizó el **BD SMART RACE cDNA Amplification Kit** (BD Biosciences Clontech) que utiliza la tecnología SMART (Switching Mechanism At 5' end of RNA Transcript) en la síntesis del ADNc, mediante el oligonucleótido BD SMART II™ A y la transcriptasa reversa BD PowerScript™, que exhibe actividad dNTP transferasa en el extremo del ARNm adicionando 3 a 5 residuos (predominantemente dC) en el extremo 3' de la primera cadena del ADNc (**Figura 3.5.**). El oligonucleótido BD SMART II™ A contiene un extremo con residuos G que se alinean con las dC adicionadas en el extremo del ADNc. Sirviendo como extensión para la síntesis del ADNc por parte de la transcriptasa reversa y teniendo de esta forma un adaptador en el extremo 5' del transcrito.

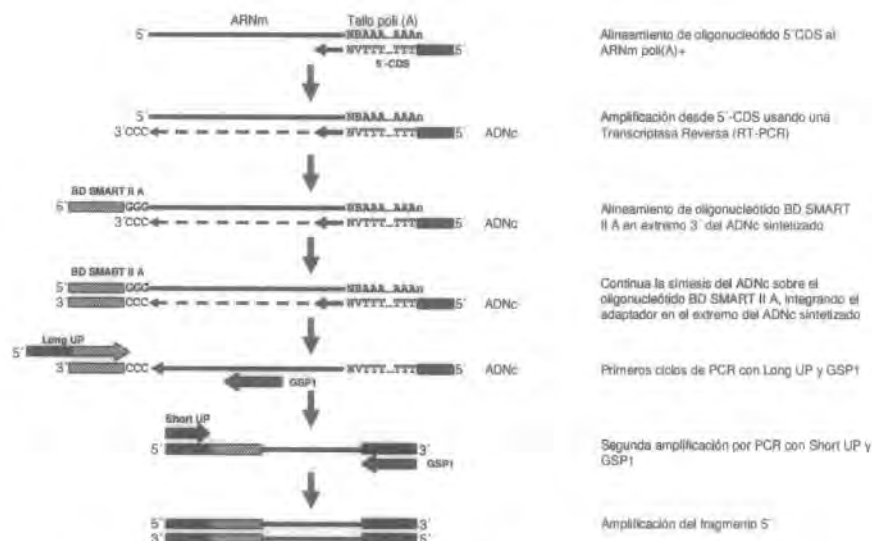


Figura 3.5. Procedimiento 5'-RACE (*Rapid Amplification of cDNA Ends*)

COMPONENTES 5'-RACE

Componentes utilizados para la síntesis de ADNc 5'-RACE:

1. **BD SMART II™ A Oligonucleotide.** Oligonucleótido de ARN en solución acuosa a 12 μ M.
2. **5'-RACE CDS Primer.** Solución acuosa a 12 μ M.
3. **BD PowerScript™ Reverse Transcriptase.**
4. **5X First-Strand Buffer.** 250 mM Tris-HCl (pH 8.3), 375 mM KCl, 30 mM MgCl₂.
5. **Dithiothreitol (DTT).** 20 mM.
6. **Buffer Tricine-EDTA.** 10 mM Tricine-KOH (pH 8.5), 1.0 mM EDTA.
7. **Agua libre de nucleasas.** Agua desionizada grado biología molecular filtrada a 0.1 μ m.
8. **UPM (10X).** Oligonucleótidos en solución acuosa 0.4 μ M de Long UPM y 2 μ M de Short UPM (Tabla 3.3.).

9. **NUP.** Oligonucleótido en solución acuosa 10 μM .

TABLA 3.2. Secuencia de oligonucleótidos 5'-RACE

Oligonucleótido	Secuencia
BD SMART II A	5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'
5'-RACE CDS	5'-TTTTTTTTTTTTTTTTTTTTTTTTTTVN-3'
Long UPM	5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3'
Short UPM	5'-CTAATACGACTCACTATAGGGC-3'
NUP	5'-AAGCAGTGGTATCAACGCAGAGT-3'

PROCEDIMIENTO SÍNTESIS DE ADNc 5'-RACE

Síntesis de la primera cadena del ADNc para 5'-RACE.

1. Preparar la siguiente mezcla de reacción en un tubo en hielo:

Templado ARN	ARN total	hasta 1 μg
	ó ARN poli(A) ⁺	hasta 1 μg
Oligonucleótidos	5'-CDS Primer	1 μl
	BD SMART II [™] A	1 μl
Agua libre de nucleasas		completar 5 μl

2. Mezclar con cuidado y coleccionar las gotas con una breve centrifugación.
3. Incubar la mezcla a 70°C por 2 min, colocar en hielo 2 min y coleccionar las gotas con una breve centrifugación.

4. Colocar el tubo en hielo y adicionar los siguientes componentes en el orden indicado:

5X First-Strand Buffer	2 μl
DTT (20 mM)	1 μl
10 mM dNTP Mix	1 μl
BD PowerScript Reverse Transcriptase	1 μl
Volumen Final	10 μl

5. Mezclar cuidadosamente y coleccionar las gotas con una breve centrifugación.
6. Incubar la mezcla a 42°C por 1.5 h.
7. Diluir el producto de reacción de síntesis de ADNc con Buffer Tricine-EDTA:
Si se partió de <200 ng de ARN total Adicionar 20 μl
Si se partió de >200 ng de ARN total Adicionar 100 μl
Si se partió de ARNm poli(A)⁺ Adicionar 250 μl
8. Incubar la dilución a 72°C por 7 min y coleccionar las gotas con una breve centrifugación.
9. Almacenar muestra a -20°C.

En este punto se tiene ya el ADNc con el adaptador que servirá para la amplificación por extremo 5'.

AMPLIFICACIONES 5'-RACE

La amplificación del extremo 5' del ADNc se realiza siguiendo lo establecido en **Materiales y Métodos 3.6.**, utilizando el oligonucleótido antisentido **GSP1 (Figura 3.2.)** y la mezcla **UPM(10X)** de

oligonucleótidos sentido **UPM Long** y **UPM Short** (Tabla 3.2). Las condiciones de reacción recomendadas por el fabricante del producto son las siguientes:

Tabla 3.3. Condiciones de reacción recomendadas en 5'-RACE (*Touchdown-PCR*)

Etapa	Tiempo	Temperatura	Ciclos
Desnaturalización	0.5min	94 °C	5X
Extensión	3min	72 °C	
Desnaturalización	0.5min	94 °C	5X
Alineamiento	0.5min	70 °C	
Extensión	3min	72 °C	
Desnaturalización	0.5min	94 °C	20-25X
Alineamiento	0.5min	68 °C	
Extensión	3min	72 °C	

Los productos de amplificación se visualizan en gel de agarosa 1% teñido con bromuro de etidio y los tamaños se estiman por comparación con marcador de peso de ADN.

3.12.3. 3'-RACE (*Rapid Amplification of cDNA Ends*)

El procedimiento 3'-RACE consiste en la síntesis de la primera cadena del ADNc iniciando en el tallo poli(A) de los ARNm usando un oligonucleótido poli(T) con un adaptador antisentido (**APT18**) (Figura 3.4.) (Tabla 3.2.) y una Transcriptasa Reversa. Después de esto, la amplificación por PCR se realiza usando dos oligonucleótidos: un oligonucleótido diseñado específicamente para el ADNc a amplificar y con el oligonucleótido adaptador antisentido (**AP**) que esta junto al tallo poli(A) y que fue introducido durante la síntesis del ADNc (Figura 3.6.).

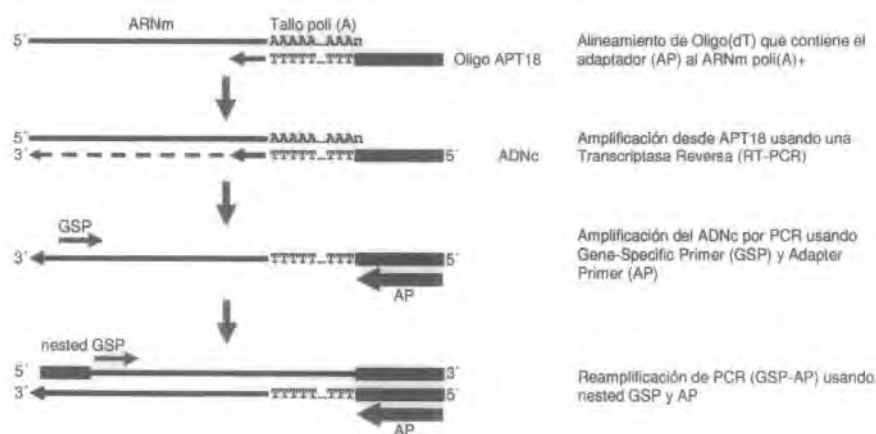


Figura 3.6. Procedimiento 3'-RACE (*Rapid Amplification of cDNA Ends*)

TABLA 3.4. Secuencia de oligonucleótidos APT18 y AP

Clave	Secuencia
APT18	5´-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTTTT-3´
AP	5´-GGCCACGCGTCGACTAGTAC-3´

COMPONENTES 3´-RACE

Los componentes utilizados para la síntesis de ADNc para el 3´-RACE fueron los mismos utilizados en la síntesis de ADNc (**Materiales y Métodos 3.5.**), excepto en el tipo de oligonucleótido utilizado, que en este caso en el **APT18** (**Tabla 3.2.**).

PROCEDIMIENTO SÍNTESIS DE ADNc 3´-RACE

El procedimiento de síntesis de ADNc 3´-RACE fue de acuerdo a lo establecido en **Materiales y Métodos 3.5.**

AMPLIFICACIONES 3´-RACE

La amplificación del extremo 3´ del ADNc se realiza siguiendo lo establecido en **Materiales y Métodos 3.6.**, utilizando como oligonucleótido antisentido el oligonucleótido **AP** (**Tabla 3.2.**) y como sentido el oligonucleótido específico **GSP2** (**Figura 3.2.**).

4. RESULTADOS Y DISCUSIÓN

La capacidad de síntesis y compartimentalización de ácido siálico, así como su incorporación en glicoproteínas, ha sido un aspecto controversial en diversos estudios realizados en insectos (organismos completos o líneas celulares) (Marchal *et al.*, 2001). La mayoría de los resultados reportados indican que en general las glicoproteínas producidas por insectos no contienen ácido siálico. Aunque algunos estudios indican que este tipo de procesamiento puede llevarse a cabo en condiciones muy particulares en algunos tipos de organismos, por lo que se piensa que este tipo de procesamiento en insectos es altamente especializado (Marchal *et al.*, 2001).

Se creía que la capacidad de sialilación era exclusiva de los vertebrados y algunas bacterias patógenas, algunos estudios han revelado la presencia de ácido siálico en ciertos invertebrados (*Drosophila melanogaster*, *Philaenus spumarius*, *Trichoplusia ni*, *Anopheles albimanus*) (Roth *et al.*, 1992; Malykh *et al.*, 1999; Marchal *et al.*, 2001; Nok *et al.*, 2002). La mayoría de los análisis indican que los insectos producen principalmente *N*-glicanos truncos del tipo alto en manosa o paucimanosa.

Al completar la secuenciación de los genomas de *Drosophila melanogaster* y *Anopheles gambiae* se tiene acceso a una vasta cantidad de información, en la cual se puede analizar el potencial genético para sialilación de glicoproteínas en insectos (Adams *et al.*, 2000; Holt *et al.*, 2002). El presente trabajo se realizó a partir de información proveniente de estos dos organismos, extrapolando hacia *Anopheles albimanus*, organismo cercano filogenéticamente a *An.gambiae* para realizar la identificación, clonación y análisis ontogenético de la respectiva α 2,6-Sialiltransferasa.

Esta sección de **Resultados y Discusión** está organizada en una primera parte que contiene la información sobre el análisis bioinformático para la identificación de la α 2,6-Sialiltransferasa de *An.gambiae*. La segunda parte presenta la identificación, clonación y análisis de secuencia de la α 2,6-Sialiltransferasa de *Anopheles albimanus*, así como el análisis ontogenético de expresión.

4.1. Análisis bioinformático para la identificación de la α 2,6-Sialiltransferasa de *Anopheles gambiae*

La secuenciación del genoma del mosquito *Anopheles gambiae* inició en el año de 1999 por **The Internacional Anopheles Genome** por una colaboración entre **Celera Genomics**, **Genoscope**, **University of Notre Dame**, **EBI/Sanger Institute**, **EMBL**, **Institut Pasteur**, **IMBB** y **TIGR**, con el patrocinio de: **Nacional Institute of Allergy and infectious Diseases (NIAID/NIH)**, **French Ministry of Research** y **World Health Organisation (WHO)**. La descripción del genoma del mosquito *Anopheles gambiae* fue completado y publicado en el 2002 (Holt *et al.*, 2002). La importancia de este organismo radica en su capacidad de transporte y transmisión del protozoario causante de la malaria (*Plasmodium spp.*), siendo un importante problema de salud pública.

La secuencia del genoma de *Anopheles gambiae* ha sido depositada en diferentes bases de datos administradas por algunas de las instituciones participantes. **Tabla 4.1.**

Tabla 4.1. Bases de datos del genoma de *Anopheles gambiae*.

Base de Datos Genoma de <i>Anopheles gambiae</i>	Dirección electrónica
Anopheles Database	http://skonops.imbb.forth.gr/AnoBase/
GenBank	www.ncbi.nlm.nih.gov/
Ensembl	www.ensembl.org/
TIGR	www.tigr.org/

Usando las secuencias de ADNc y proteína de las Sialiltransferasas de humano y *Drosophila melanogaster*, se realizaron búsquedas por homología para identificar la secuencia de la α 2,6-Sialiltransferasa ortóloga putativa de *Anopheles gambiae* en su genoma. A continuación se muestran los resultados obtenidos.

4.1.1. Búsqueda de α 2,6-Sialiltransferasas en el genoma de *Anopheles gambiae*

Se realizaron búsquedas por homología mediante **BLAST** en las diferentes bases de datos (Tabla 4.1.) donde está depositado el genoma de *Anopheles gambiae*. Como sonda de búsqueda se utilizó la región conocida como *sialylmotif L* de la α 2,6-Sialiltransferasa de humano (CAA35111) y de *Drosophila melanogaster* (NP726474), siendo ésta región un dominio característico y mediante el cual es posible la identificación de las proteínas de esta familia. Las secuencias se muestran en la Tabla 4.2.

Tabla 4.2. Secuencias de las sondas de búsqueda (*sialylmotif L*).

Organismo	Secuencia <i>Sialylmotif L</i>
Humano	192 - WGRCAVVSSAGSLKSSQLGREIDDHDAVLRFNAPTANFQQDVGTKT - 239
<i>D.melanogaster</i>	215 - IKTCAIVSSAGSLAGSKLGRFIDTHDIVMRFNHAPTQGEVDVGSKT - 262

Como resultado de estas búsquedas por homología en el genoma de *An.gambiae* se identificaron secuencias en las diferentes bases de datos. Los números de acceso de las secuencias identificadas se muestran en la Tabla 4.3.

Tabla 4.3. Número de acceso de las secuencias de *Anopheles gambiae* identificadas.

Base de Datos	Secuencia en aminoácidos	Secuencia en nucleótidos
GenBank	EAAO4038	AAAB01008807
GenBank	CAH25390	AJ821850
ENSEMBL	ENSANGP00000021696	ENSANGT00000021696
TIGR	TC104766	Q683N9

Todas ellas localizadas en el cromosoma 2L (39241405-39242735 pb) como copia única en el genoma. En forma similar a la secuencia ortóloga de *D.melanogaster* que también es copia única localizada en el cromosoma 2R (Adams et al, 2000; Fabini et al, 2001; Fabini et al, 2002; Holt et al, 2002; Koles et al, 2004).

Se observó que la longitud de las secuencias varía, para tener una visión clara de esta variabilidad se alinearon las secuencias obtenidas de las diferentes bases de datos. Las secuencias de aminoácidos alineadas se muestran en la **Figura 4.1.**, mientras que el alineamiento de las secuencias nucleotídicas se muestra en la **Figura 4.2.**

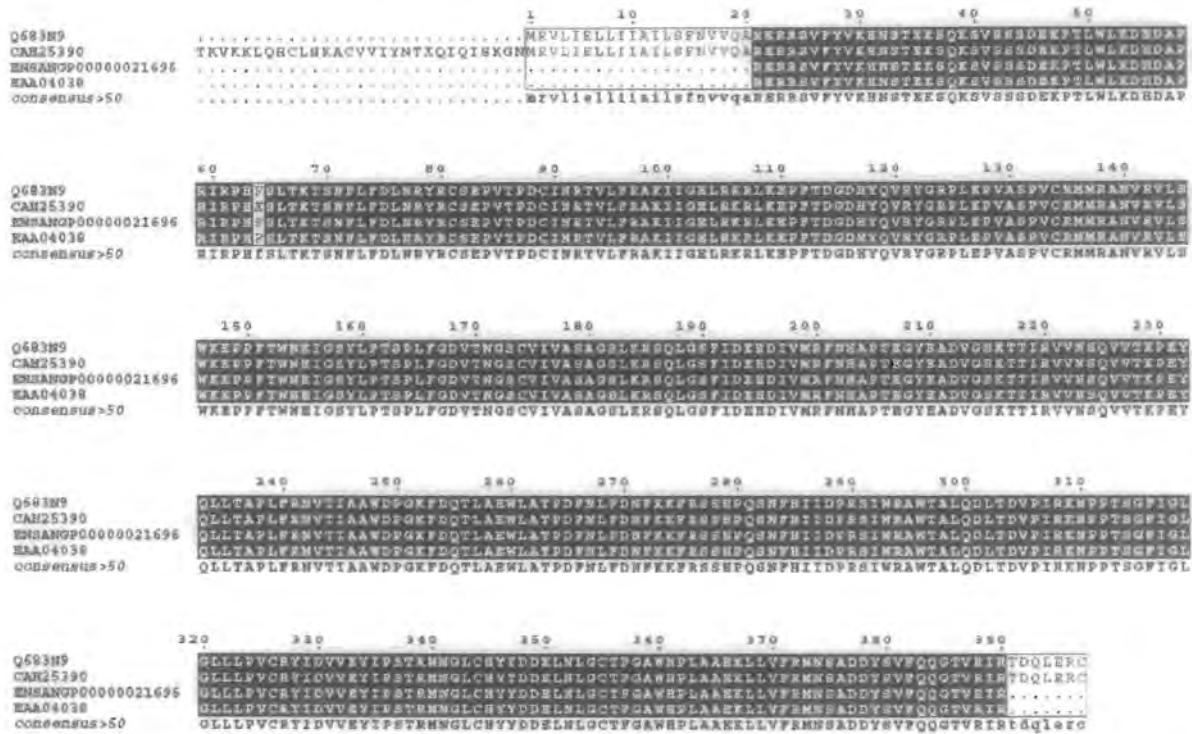


Figura 4.1. Alineamiento de secuencias aminoácidas de α 2,6-Sialiltransferasa de *Anopheles gambiae*. Q693N9 (TIGR), CAH25390 (GenBank), ENSANGP00000021696 (ENSEMBL), EAA04038 (GenBank). La región consenso entre las secuencias se muestra sombreada.

En el alineamiento de la **Figura 4.1.** se identificó un consenso de 370aa entre las secuencias, este se muestra sombreado. La variabilidad ocurre en los extremos de las secuencias, el extremo carboxilo está conservado en dos de las tres secuencias y la longitud corresponde con la de sialiltransferasas de otros organismos. En el extremo amino tenemos variabilidad en las tres secuencias, esto debido a que no se ha identificado donde inicia la traducción de la proteína. Cabe señalar que la región amino terminal es variable y específica para cada organismo, variando en secuencia y longitud (Breton *et al*, 1996). El dominio catalítico está completo, el análisis de su secuencia se muestra en la siguiente sección (**4.1.2.**).

Se identificó un consenso de 1110pb en el alineamiento de la **Figura 4.2.**, correspondiente a los 370aa del consenso de la **Figura 4.1.** En el genoma de *An.gambiae* se pueden localizar tres intrones dentro de esta secuencia, que varían en longitud entre 70 y 77pb (Holt *et al*, 2002).

Las diferencias entre las secuencias obtenidas puede deberse a que las bases de datos aún están en proceso de depuración y las anotaciones respectivas en cada secuencia están también en ese proceso.

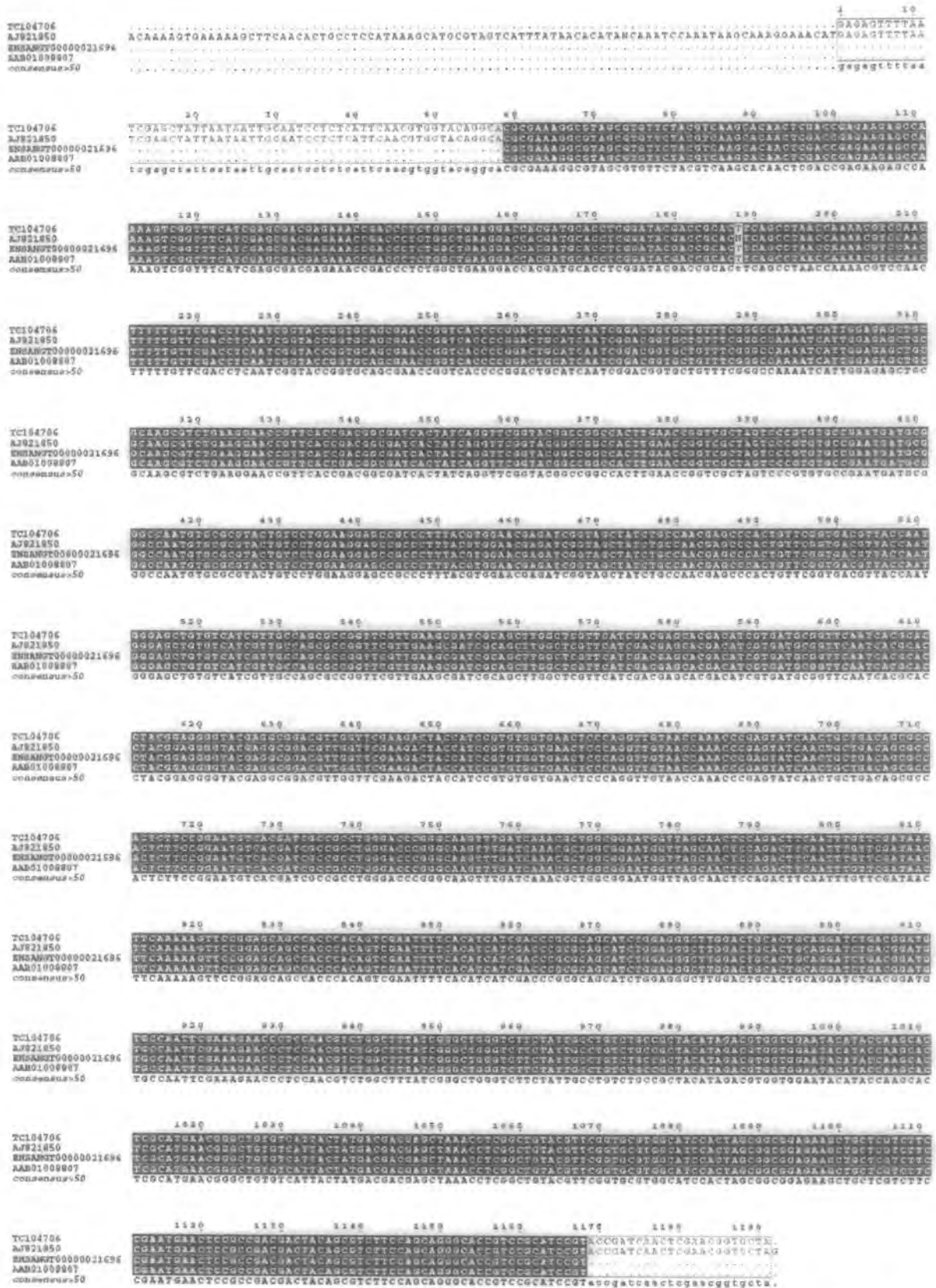


Figura 4.2. Alineamiento de secuencias nucleotídicas de α 2,6-Sialiltransferasa de *Anopheles gambiae*. TC104706 (TIGR), A821850 (GenBank), ENSANGT0000021696 (ENSEMBL), AAB01008807 (GenBank). La región consenso entre las secuencias se muestra sombreada.

Hasta el momento no se han reportado datos experimentales que confirmen el tamaño del ADNc o de la proteína, ni de la actividad enzimática de las proteínas obtenidas a partir de los diferentes ADNc de *An.gambiae*.

Se tomo como referencia las secuencias **TC104766** y **Q683N9** (**Tabla 4.3.**) de *An.gambiae* para realizar las siguientes etapas del presente proyecto.

4.1.2. Análisis de la α 2,6-Sialiltransferasa de *Anopheles gambiae*

Las α 2,6-Sialiltransferasas hasta ahora clonadas y caracterizadas poseen una organización particular que las distingue. Se realizaron análisis de las secuencias **TC104766** y **Q683N9** de *An.gambiae*, comparándolas con las Sialiltransferasas ya conocidas. Determinando así si estas poseen los residuos y la organización característicos de las enzimas de esta familia.

Usando los programas **LALIGN** y **BLAST** se realizó un análisis de identidad entre las secuencias de *An.gambiae* identificadas y las secuencias mostradas en la **Tabla 2.2.** Obteniendo los porcentajes de identidad indicados en la **Tabla 4.4.**

Tabla 4.4. Análisis de identidad de la α 2,6-Sialiltransferasa de *Anopheles gambiae*.

Organismo	Identidad α 2,6-Sialiltransferasa de <i>Anopheles gambiae</i>	
	ADNc	Proteína
Humano (<i>Homo sapiens</i>)	50.2 %	29.2 %
Rata (<i>Rattus norvegicus</i>)	50.6 %	30.1 %
Ratón (<i>Mus musculus</i>)	52.0 %	29.2 %
Pollo (<i>Gallus gallus</i>)	49.4 %	29.8 %
Bovino (<i>Bos taurus</i>)	49.8 %	28.1 %
Mosca (<i>Drosophila melanogaster</i>)	52.4 %	41.2 %

El grado de identidad entre las secuencias analizadas es relativamente bajo, en comparación con las identidades entre las sialiltransferasas de vertebrados, que van de 61 hasta 93% de identidad. La identidad de la única sialiltransferasa de invertebrado caracterizada (*D.melanogaster*) respecto a las de organismos vertebrados es entre 29 y 30% de identidad, por lo que podemos ver que estos datos son similares a los obtenidos con *An.gambiae*. Sabemos que la región del tallo es altamente variable, tanto en secuencia como en longitud en los diferentes organismos, por lo cual estas diferencias se ven reflejadas en el análisis de identidad.

La región del dominio catalítico de las Sialiltransferasas es la más conservada entre los diferentes organismos, ésta comprende los tres *sialylmotifs* (*L*, *S* y *VS*) hasta el extremo carboxilo terminal. El análisis de identidad de los dominios catalíticos se muestra en la **Tabla 4.5.**

Tabla 4.5. Análisis de Identidad de la región catalítica de la $\alpha 2,6$ -Sialiltransferasa de *Anopheles gambiae*.

Organismo	Identidad $\alpha 2,6$ -Sialiltransferasa de <i>Anopheles gambiae</i>	
	ADNc	Proteína
Humano (<i>Homo sapiens</i>)	55.8 %	39.6 %
Rata (<i>Rattus norvegicus</i>)	56.1 %	38.3 %
Ratón (<i>Mus musculus</i>)	58.1 %	37.9 %
Pollo (<i>Gallus gallus</i>)	53.0 %	41.0 %
Bovino (<i>Bos taurus</i>)	55.2 %	39.3 %
Mosca (<i>Drosophila melanogaster</i>)	64.8 %	59.9 %

El grado de identidad en la región del sitio catalítico es mayor respecto a la secuencia completa. Particularmente, respecto a la $\alpha 2,6$ -Sialiltransferasa de *D.melanogaster*, la identidad es de 64.8% en nucleótidos y 59.9% en aminoácidos. Estos porcentajes concuerdan con el análisis comparativo entre el total de las enzimas ortólogas de *D.melanogaster* y *An.gambiae*, que son en promedio 60% idénticas (Holt et al, 2002; Zdobnov et al, 2002).

Se realizó un alineamiento múltiple entre las $\alpha 2,6$ -Sialiltransferasas de humano, *D.melanogaster* y la ortóloga putativa de *An.gambiae*, mostrado en la **Figura 4.3**. Se localizaron los aminoácidos característicos de los *sialylmotifs* L, S y VS, así como aquellos relevantes en catálisis y formación del puente disulfuro (Drickamer, 1993; Datta et al, 1995; Datta et al, 1997; Geremia et al, 1997; Datta et al, 1998; Breton et al, 1999; Datta et al, 2001). Por lo que esta secuencia de *An.gambiae* posee la topología característica de la familia de las Sialiltransferasas, identificándola entonces como $\alpha 2,6$ -Sialiltransferasa putativa de este organismo.

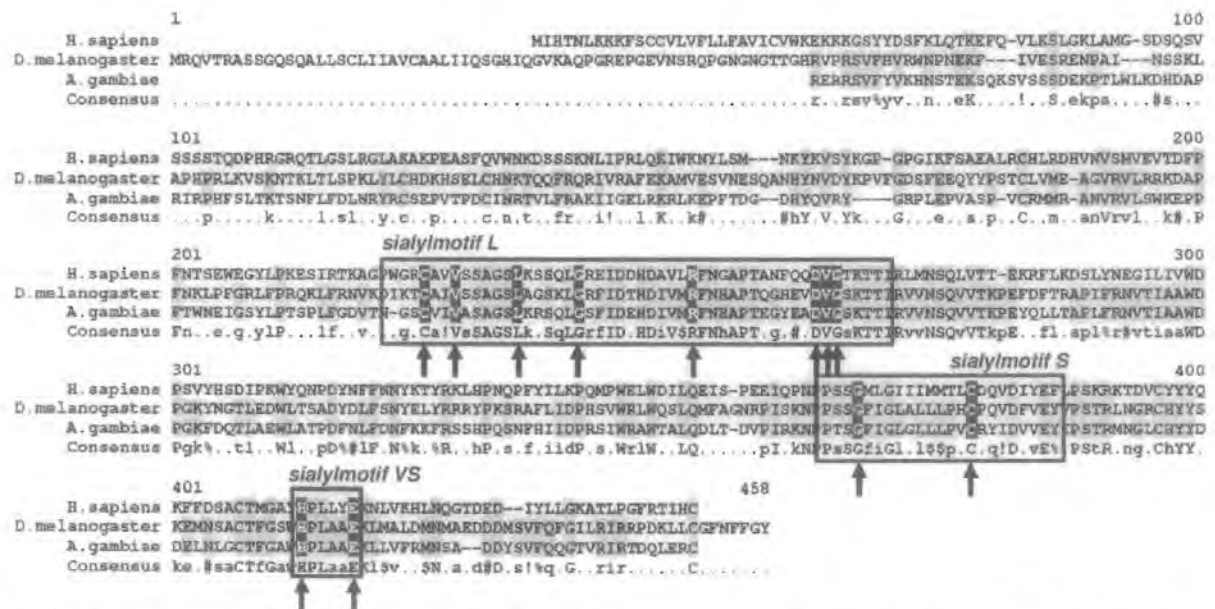


Figura 4.3. Identificación de los *sialylmotifs* en la $\alpha 2,6$ -Sialiltransferasa de *Anopheles gambiae*. El alineamiento de secuencias muestra las regiones conservadas de los *sialylmotifs* L, S y VS enmarcadas por recuadros. Los residuos que son invariantes se señalan con flechas y los residuos altamente conservados están sombreados.

Hasta el momento de escritura de este trabajo, la única Sialiltransferasa de insectos en ser clonada y caracterizada bioquímicamente es la de *D.melanogaster* (Koles *et al.*, 2004). Se confirmó que ésta es una proteína funcional del tipo ST6Gal, tal como había sido predicho por análisis bioinformáticos (Fabini *et al.*, 2001; Fabini *et al.*, 2002). Debido al alto grado de identidad entre el ortólogo putativo identificado en el genoma de *An.gambiae* y la enzima caracterizada de *Drosophila* sugerimos que la secuencia de *A.gambiae* codifica para una proteína del tipo ST6Gal.

En reportes recientes publicados durante el desarrollo de este trabajo, se identifica esta misma secuencia de ADNc de *Anopheles gambiae* como la α 2,6-Sialiltransferasa ortóloga putativa de mosquito (Koles *et al.*, 2004; Harduin-Lepers *et al.*, 2005). Respaldao lo que aquí se sugiere.

4.2. Identificación de la α 2,6-Sialiltransferasa de *Anopheles albimanus*

Ya identificada por análisis bioinformáticos la secuencia de la α 2,6-Sialiltransferasa ortóloga putativa de *An.gambiae*, se procedió al diseño de oligonucleótidos para la búsqueda del gen ortólogo en ADNc de *Anopheles albimanus*.

En contraste con *Anopheles gambiae* cuyo genoma está completamente secuenciado, hasta el momento se han reportado en **GenBank** únicamente 68 secuencias nucleotídicas y 70 aminoácidas de *Anopheles albimanus*. Estas poseen cerca de 90% de identidad respecto a *An.gambiae*.

En el alineamiento de la **Figura 4.3.** se señalan con flechas los residuos invariables y sombreados aquellos altamente conservados. Se localizan principalmente en el dominio catalítico. Sobre estas regiones conservadas se realizó el diseño de dos grupos de oligonucleótidos degenerados mediante **THE MEME/MAST SYSTEM**.

El primer grupo de oligonucleótidos a partir de un consenso de secuencias nucleotídicas de *An.gambiae* y *D.melanogaster* (**Tabla 4.6.**) y el segundo grupo de oligonucleótidos fue diseñado en regiones conservadas de aminoácidos pero con el uso de codones preferencial de *An.albimanus* (**Codon Usage Database**). Estos se muestran en la **Tabla 4.7.**

Tabla 4.6. Secuencias de oligonucleótidos del consenso de *Anopheles-Drosophila*.

Clave	Secuencia
AgL	5'-GCACGACAT GTGATG G TTCAATCA-3'
AgLS	5'-TTCCGGAA GTCACGATCGCCGCTGGG-3'
AgVS	5'-CAGCTTCTCCGCCGCTAGTGGATGCC-3'

Las bases degeneradas se muestran sombreadas.

R = A o G; M = A o C; Y = C o T

AgL y AgLS son oligonucleótidos sentido.

AgVS es oligonucleótido antisentido.

Tabla 4.7. Secuencias de oligonucleótidos con el uso de codones de *Anopheles albimanus*.

Clave	Secuencia
AaF1	5'-ATCGA GAGCA GA ATCGT ATG-3'
AaF2	5'-GT ACGATCGC GC TGGGA CC-3'
AaR1	5'-AGCTCATC TCGTAGTA TG CA-3'
AaR2	5'-TTCTCCGCCG AG GG TGCCA-3'

Las bases degeneradas se muestran sombreadas.

R = A o G; Y = C o T; S = C o G

AaF1 y AaF2 son oligonucleótidos sentido.

AaR1 y AaR2 son oligonucleótidos antisentido.

El uso de oligonucleótidos degenerados para identificar y aislar ADNc de Sialiltransferasas se utilizó con éxito en el caso de las α 2,3-Sialiltransferasas de cerdo, rata, ratón y humano, así como de las α 2,6-Sialiltransferasas de humano, pollo, bovino y ratón (Weinstein *et al.*, 1982; Weinstein *et al.*, 1987; Lance *et al.*, 1989; Grundmann *et al.*, 1990; Gillespie *et al.*, 1992; Wen *et al.*, 1982; Kitagawa *et al.*, 1994a; Kitagawa *et al.*, 1994b; Kurosawa *et al.*, 1994; Lee *et al.*, 1994; Tsuji, 1996; Mercier *et al.*, 1999). Aunque en el caso de varios de estos transcritos se utilizó una librería de ADNc para el aislamiento de las respectivas secuencias completas.

A partir de la localización de los oligonucleótidos sobre la secuencia de la α 2,6-Sialiltransferasa de *Anopheles gambiae* se calculó el tamaño de las amplificaciones esperadas con los diferentes pares de oligonucleótidos. **Tabla 4.8.**

Tabla 4.8. Amplificaciones esperadas con los diferentes oligonucleótidos

Oligo Sentido	Oligo Antisentido	Amplificación
AgL	AaR1	471pb
AgL	AaR2	527pb
AgL	AgVS	523pb
AgLS	AaR1	335pb
AgLS	AaR2	383pb
AgLS	AgVS	387pb
AaF1	AaR1	479pb
AaF1	AaR2	519pb
AaF2	AaR1	326pb
AaF2	AaR2	374pb

El ADNc usado para la búsqueda de la sialiltransferasa de *An.albimanus* fue sintetizado a partir de ARN total extraído de larvas de segundo *instar* (donadas amablemente por el Bioterio del INSP, Cuernavaca, Mor.). Se eligió esta etapa de desarrollo teniendo como antecedente el hecho de que en *D.melanogaster* la incorporación de ácido siálico está regulada desde blastodermo hasta larvas de tercer instar (Roth *et al.*, 1992).

Se sintetizaron dos ADNc, uno a partir de oligo **dt18** y otro de tipo Secuencia Específica (SE). Para verificar la calidad del ARN extraído y del ADNc sintetizado se utilizó un control interno, el gen Glucosa-6-fosfato deshidrogenasa (G6PD) de *Anopheles albimanus* (GenBank AF318224). Se diseñó un par de oligonucleótidos: 5´-GCCGCCAATATATGAAGGTAGA-3´ y 5´-CCCAGATAGTGATCGATACGG-3´, que producen una amplificación de 505pb (Figura 4.4.). Obteniendo amplificaciones específicas de la talla esperada, confirmando calidad de ARN y ADNc.



Figura 4.4. Amplificación G6PD de *Anopheles albimanus*.

1. Marcador de peso molecular
2. Amplificación G6PD (505pb)

4.2.1. Amplificaciones ADNc dt18

Usando el ADNc **dt18** como templado y los oligonucleótidos de la **Tabla 4.6.**, se realizaron amplificaciones PCR con **Vent_R DNA Polymerase** (94°C 1min, 60°C 1min, 72°C 2min; 30 ciclos). El resultado de estas amplificaciones se muestra en la **Figura 4.5.**



Figura 4.5. Amplificaciones sobre ADNc-dt18.

- | | Amplificación |
|--------------|---------------|
| 1. Marcador | |
| 2. AgLS-AgVS | 387pb |
| 3. AgL -AgVS | 523pb |

Las amplificaciones obtenidas en el carril 2 no corresponden al tamaño esperado. En el carril 3 se observó un producto próximo al esperado de 523pb. Esta banda fue purificada y clonada en pCR-Blunt II-TOPO para su secuenciación.

La secuencia resultante de 530pb de longitud fue sometida a una búsqueda por homología en el genoma de *Anopheles gambiae* mediante **BLAST**, sin que se encontraran similitudes significativas. Al comparar esta secuencia y sus marcos abiertos de lectura respecto a la secuencia de la α 2,6-Sialiltransferasa putativa de *An.gambiae*, se encontró que tienen parecido únicamente en pequeñas regiones. No fue posible identificar los residuos característicos de los *sialylmotifs*, descartando así que esta secuencia pudiera corresponder a una sialiltransferasa pese a ser próxima a la talla esperada.

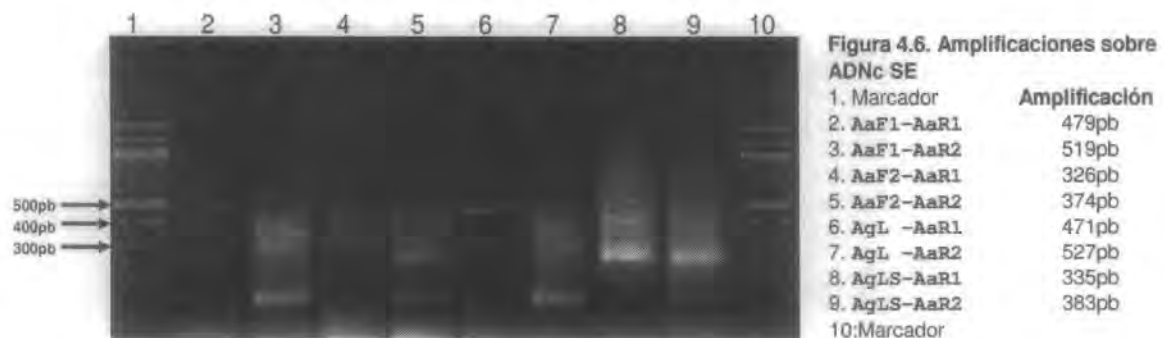
Las glicosiltransferasas tienen en general bajos niveles de expresión. La síntesis de ADNc con el oligonucleótido **dT18** en combinación con el uso de oligonucleótidos degenerados en las reacciones PCR no permitió amplificar el transcrito de interés. Se emprendieron nuevas estrategias para lograr este objetivo, estas se muestran a continuación.

4.2.2. Amplificaciones ADNc SE

Tratando de enriquecer el ADNc con el transcrito de la α 2,6-Sialiltransferasa se utilizó el oligonucleótido antisentido **AaR2** (Tabla 4.7.) para la síntesis de ADNc del tipo Secuencia Específica (ADNc SE), a partir de ARN total de larvas de *An.albimanus instar 2*. Este hibrida en la región catalítica de la sialiltransferasa de *An.gambiae* y *D.melanogaster*. A partir de este ADNc SE se realizaron amplificaciones PCR.

Se ha utilizado esta metodología para la amplificación de genes de baja expresión, tal como genes de resistencia a drogas en tejidos cancerosos y varios receptores olfatorios (Fuqua *et al*, 1990; Yang *et al*, 2002). En todos los casos utilizando oligonucleótidos específicos. En nuestro caso, debido a la escasez de información de *An.albimanus*, el ADNc fue sintetizado a partir de oligonucleótidos degenerados.

Este ADNc SE fue usado como templado en reacciones PCR con **Vent_R DNA Polymerase** (94°C 1min, 60°C 1min, 72°C 2min; 30 ciclos), con los oligonucleótidos degenerados mostrados en la Tabla 4.7.. Se obtuvieron las amplificaciones mostradas en la Figura 4.6.



Los productos de los carriles 2, 4 y 6 son de particular interés, ya que en estas muestras se obtuvieron amplificaciones cercanas a las tallas estimadas de 479pb, 326pb y 471pb, respectivamente. Estas tres bandas fueron purificadas y utilizadas como templado en *seminested*-PCR para confirmar que estas contenían la secuencia buscada. Las reamplificaciones *seminested*-PCR se muestran en la Figura 4.7.

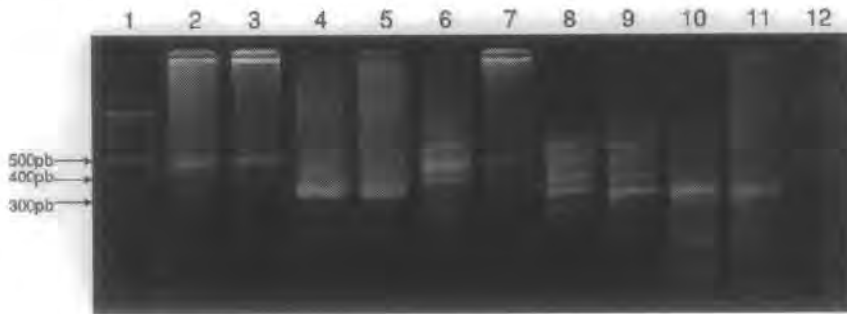


Figura 4.7. Reamplificaciones *seminested*-PCR

	Amplificación
1. Marcador	
2. AgL - AaR1	471pb
3. AaF1 - AaR1	479pb
4. AgLS - AaR1	335pb
5. AaF2 - AaR1	326pb
6. AgL - AaR1	471pb
7. AaF1 - AaR1	479pb
8. AgLS - AaR1	335pb
9. AaF2 - AaR1	326pb
10. AgLS - AaR1	335pb
11. AaF2 - AaR1	326pb
12. Marcador	

Al observar las amplificaciones de la **Figura 4.7.** podemos distinguir que en la mayoría de los casos son bandas discretas y específicas de las tallas esperadas. Los carriles 2 a 5 de esta misma figura corresponden a las reamplificaciones de la banda de 471pb mostrada en el carril 6 de la **Figura 4.6.** Son amplificaciones únicas y de las tallas estimadas. En los carriles 6 a 9 se muestran las reamplificaciones de la banda de 479pb (carril 2 **Figura 4.6.**), pese a obtener las reamplificaciones de las tallas estimadas con los diferentes oligonucleótidos se tiene gran cantidad de amplificaciones inespecíficas. Finalmente, en los carriles 10 y 11 tenemos las reamplificaciones de la banda de 326pb (carril 4 **Figura 4.6.**) que también son discretas y de las tallas estimadas.

Ya que las reamplificaciones (*seminested*-PCR) de algunas de las bandas obtenidas se obtuvieron en la forma esperada, se procedió a confirmar la identidad de la amplificación mediante secuenciación. Para ello se realizó la clonación de la banda de 471pb (carril 2 **Figura 4.6.**) obtenida con los oligonucleótidos **AgL-AaR1** en el vector pCR-Blunt II-TOPO, se seleccionaron clonas con el inserto y se purificaron plásmidos para secuenciación. La secuencia de esta amplificación se muestra en la **Figura 4.8.**

```

GCACGACATTGTGATGCGGTTCAATCATGCACCAACCGAGGGCTTCGAAGCGGACGTC
H D I V M R F N H A P T E G F E A D V

GGAGCGAAGACGACGATCCGCGTACTCAACTCGCAGGTCGTCAGAACCTGAGTACCGG
G A K T T I R V L N S Q V V T K P E Y R

TTGCTGACGGCACCCTGTTCGGAAACGTTACGATAGCGGCCCTGGGATCCGGGCAAGTTC
L L T A P L F R N V T I A A W D P G K F

GATCAACCGCTGTCCGAGTGGTGGCCAATCCGGACTTTAATCTGTTTCGAGAACTTCAAG
D Q P L S E W L A N P D F N L F E N F K

AAGTTTCGCAGCCAGTACCGGGATGCCAACTTCCACATCATCGATCCGCGCAGCATCTGG
K F R S Q Y R D A N F H I I D P R S I W

CGGGCGTGGACGGCCCTTCAGGATCTCACTCCCTTTCCGATCCGACAGAACCCTCCAACA
R A W T A L Q D L T P F P I R Q N P P T

TCCGTTTTCATCGGCTTGGACTGCTGCTGCCCGTGTGCCGGTACATTGACATTGTGGAG
S G F I G L G L L L P V C R Y I D I V E

TACATTCCGAGCACTCGCATGAACGGTCTCTGTCACTACTACGACGATGAGCT
Y I P S T R M N G L C H Y Y D D E

```

Figura 4.8. Secuencia de amplificación con oligos AgL-AaR2 sobre ADNc SE.

Las secuencias de los oligonucleótidos se muestran sombreadas.

La traducción se muestra bajo la secuencia nucleotídica.

La secuencia de ADNc *An.albimanus* identificada, clonada y secuenciada, corresponde a un fragmento de 471pb. Este fragmento posee un alto grado de identidad con las secuencias de las α 2,6-Sialiltransferasas ortólogas de *An.gambiae* y *D.melanogaster*. La secuencia aminoacídica deducida del producto de PCR secuenciado y mostrado en la **Figura 4.8.**, produce un péptido de 156aa. El análisis de identidad de este ADNc y de su traducción en aminoácidos respecto a las Sialiltransferasas realizado mediante **BLAST** y **LALIGN** se muestra en la **Tabla 4.8.**

Tabla 4.9. Identidad del fragmento de 471pb de *A.albimanus* respecto a las α 2,6-Sialiltransferasas

Organismo	Identidad ADNc	Identidad Proteína
<i>Anopheles gambiae</i>	78%	89%
<i>Drosophila melanogaster</i>	72%	61%
<i>Homo sapiens</i>	55%	43%
<i>Gallus gallus</i>	55%	42%
<i>Mus musculus</i>	56%	42%
<i>Rattus norvegicus</i>	56%	42%
<i>Bos taurus</i>	55%	42%

Al analizar este péptido de 156aa mediante **BLAST**, se detectó la presencia de dominios putativos funcionales y estructurales conservados mediante **Conserved Domain Search**, correspondientes a la Familia 29 de las Glicosiltransferasas: Sialiltransferasas (Marchler-Bauer *et al*, 2004; Marchler-Bauer *et al*, 2005). Entre estos tenemos parte del *sialylmotif L*, el *sialylmotif S* y las regiones entre los mismos, que como podemos observar en el alineamiento de la **Figura 4.9.** se encuentran conservados. Además de localizar los residuos invariables implicados en el reconocimiento de sustrato donador y aceptor.

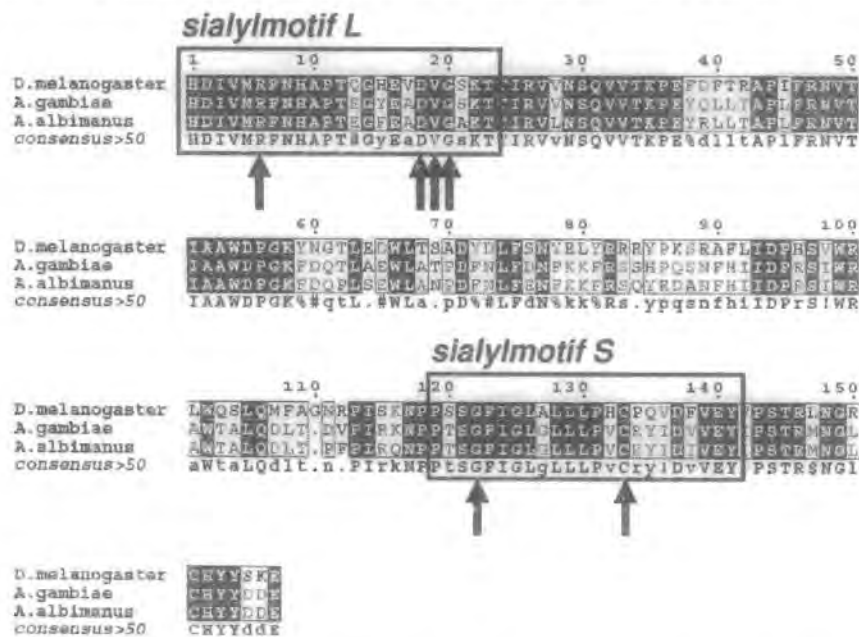


Figura 4.9. Alineamiento de fragmento de las α 2,6-Sialiltransferasas de insectos.

Los residuos conservados se muestran resaltados, los *sialylmotifs* se encuentran enmarcados con recuadros y los residuos implicados en el reconocimiento de sustrato donador y aceptor se señalan con flechas.

En base a estos datos, deducimos que esta amplificación obtenida corresponde a un fragmento de 471pb del ADNc de la α 2,6-Sialiltransferasa de *An. albimanus*. A partir del cual se procedió a la identificación de los extremos 5' y 3' del transcrito.

4.3. Identificación de los extremos 5' y 3' de la α 2,6-Sialiltransferasa de *Anopheles albimanus*

La identificación del fragmento de 471pb de la α 2,6-Sialiltransferasa putativa de *Anopheles albimanus* previamente descrito, permitió el diseño de oligonucleótidos específicos para la amplificación de los extremos 5' y 3' del transcrito mediante *Rapid Amplification of cDNA Ends (RACE)*. Las secuencias de los oligonucleótidos específicos diseñados se muestran en la **Tabla 4.10**.

Tabla 4.10. Secuencias de oligonucleótidos para *RACE* de α 2,6-Sialiltransferasa de *An. albimanus*

Clave	Secuencia
STGSP1	5' -CATGCGAGTGCTCGGAATGTA CTCCA-3'
STNGSP1	5' -AAACCGGATGTTGGAGGGTTCTGTCG-3'
STGSP2	5' -TGATGCGGTTCAATCATGCACCAAC-3'
STNGSP2	5' -TACCGGGATGCCAACTTCCACATCA-3'

4.3.1. 5'-RACE (*Rapid Amplification of cDNA Ends*)

La síntesis del ADNc para el 5'-RACE se realizó a partir de ARNm purificado de adultos hembras de *Anopheles albimanus* (18 horas post-alimentación). Este material fue donado por el Dr. Jesús Martínez (INSP). Existen evidencias de que en este estado del desarrollo del mosquito sí ocurre la transferencia de ácido siálico a glicoproteínas (Nok *et al.*, 2002). Ya que una sialoglicoproteína (gp45) fue purificada a partir del tracto digestivo de mosquitos adultos.

Tal como sugiere el fabricante, se realizó la amplificación PCR utilizando la **Vent_R DNA Polymerase** con los oligonucleótidos **UPM (10X)** (**Tabla 3.3.**) y **STGSP1** (**Tabla 4.10.**). Con las siguientes condiciones de reacción: 95°C 5min 1 ciclo, 94°C 45seg, 68°C 45seg, 72°C 2.5min 35 ciclos, 72°C 10min. El gel de agarosa no mostró ninguna amplificación, por lo cual sobre este producto se realizó una *nested*-PCR con los oligonucleótidos **NUP-STNGSP1**. Las condiciones utilizadas fueron: 94°C 45seg, 68°C 45seg, 72°C 2.5min 35 ciclos con **Vent_R**. El producto de la amplificación se muestra en la **Figura 4.10**.



Figura 4.10. Amplificación 5'-RACE con NUP-STNGSP1
 1. Marcador
 2. NUP-STNGSP1

La obtención de un barrido (carril 2 **Figura 4.10.**) sin amplificaciones específicas pudo deberse a la actividad exonucleasa de autocorrección íntegra 3'→5' intrínseca de la polimerasa **Vent_R**. Misma que pudo degradar los oligonucleótidos en los primeros ciclos de PCR afectando el rendimiento de la amplificación. El uso de esta polimerasa se decidió en función de la alta fidelidad de los productos que se obtienen, sin embargo en esta estrategia su uso no resultó favorable.

Como alternativa al uso de la polimerasa **Vent_R** se optó por el uso de una ADN polimerasa del tipo *Hot-Start* (**ProofStart DNA Polymerase**), que además es de alta fidelidad. Evitando así la degradación de los oligonucleótidos en los primeros ciclos de PCR. A partir del mismo templado de ADNc sintetizado para 5'-RACE se realizaron amplificaciones con esta enzima con los oligonucleótidos **UPM (10X)** (**Tabla 3.3.**) y **STGSP1** (**Tabla 4.10.**). Con las siguientes condiciones de reacción: 94°C 5min 1 ciclo, 94°C 30seg, 65°C 1min, 72°C 2min 35 ciclos, 72°C 10min. De igual forma el gel de agarosa no mostró ninguna amplificación (**Figura 4.11.**), tal como ocurrió en las amplificaciones anteriores.



Figura 4.11. Amplificación 5'-RACE con NUP-STNGSP1
 1. Marcador
 2. NUP-STNGSP1

El producto de la primer PCR fue sometido a reamplificación con los oligonucleótidos **NUP-STGSP1** y **NUP-STNGSP1**, las condiciones utilizadas fueron: 94°C 45seg, 65°C 45seg, 72°C 2min, 35 ciclos con **Vent_R**. Debido a la variabilidad de la región 5' del transcrito de *Anopheles gambiae* registrada en las diferentes bases de datos, es difícil predecir la talla de la amplificación esperada. Adicionalmente, la longitud y secuencia de la región del tallo de las α 2,6-Sialiltransferasas son específicas para cada organismo. Los productos de la amplificación se muestran en la **Figura 4.12.**

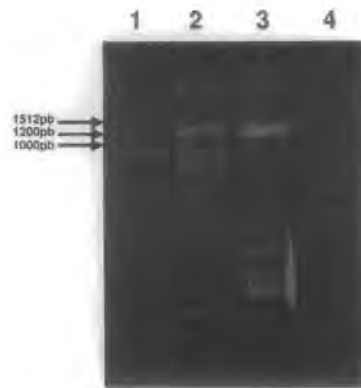


Figura 4.12. Reamplificación 5'-RACE

1. Marcador
2. NUP-STGSP1
3. NUP-STNGSP1
4. Marcador

De acuerdo a la localización de los oligonucleótidos antisentido **STGSP1** y **STNGSP1** sobre el fragmento de 471pb ya conocido, se espera que el tamaño de las amplificaciones producidas difiera en tan solo 72pb. Las amplificaciones que se observan alrededor de 1200pb en los carriles 2 y 3 de la **Figura 4.12**, parecen cumplir esta condición. Ambas bandas fueron purificadas y usadas como templado en reacciones *nested-PCR* (94°C 45seg, 65°C 45seg, 72°C 2min, 35 ciclos con **Vent_R**) para confirmar que estas correspondían al extremo 5' del transcrito. Los oligonucleótidos empleados y los tamaños de amplificación esperados se muestran en la **Tabla 4.11**. La **Figura 4.13**, muestra los productos obtenidos de dichas reacciones.

Tabla 4.11. Amplificaciones esperadas en reacciones 5'-RACE

Oligo Sentido	Oligo Antisentido	Amplificación
STGSP2	STGSP1	427pb
STNGSP2	STGSP1	185pb
STGSP2	STNGSP1	355pb
STNGSP2	STNGSP1	113pb

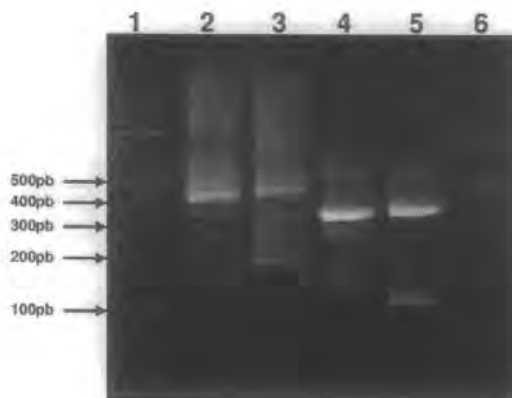


Figura 4.13. Reamplificaciones 5'-RACE

- | | |
|--------------------|----------------------|
| 1. Marcador | Amplificación |
| 2. STGSP2 -STGSP1 | 427pb |
| 3. STNGSP2-STGSP1 | 185pb |
| 4. STGSP2 -STNGSP1 | 355pb |
| 5. STNGSP2-STNGSP1 | 113pb |
| 6. Marcador | |

En los carriles 2 y 3 de la **Figura 4.13**, se muestran los productos obtenidos sobre la banda de 1200pb purificada del carril 2 **Figura 4.12**. En los cuales se esperaban amplificaciones de 427pb y 185pb, respectivamente. Dichos productos están acompañados de amplificaciones inespecíficas. Los carriles 4 y 5 de la **Figura 4.13**, corresponden a las amplificaciones obtenidas al usar como templado la banda de

1200pb del carril 3 **Figura 4.12.**, se observan los productos esperados de 355pb y de 113pb, respectivamente, junto con amplificaciones inespecíficas.

La obtención de los productos esperados confirmó la identidad de las bandas de 1200pb amplificadas y purificadas mostradas en la **Figura 4.12.** Dichas bandas corresponden al extremo 5' del ADNc de la $\alpha 2,6$ -Sialiltransferasa de *Anopheles albimanus*. Se procedió a la clonación de las mismas en el plásmido pCR-Blunt II-TOPO. Se seleccionaron clones con el inserto y se purificaron plásmidos para someter a secuenciación de ADN. La secuencia obtenida se muestra en la **Figura 4.14.**

```

cgc ggg ggt gtg gct gtg gca gcg aac aaa gtt gct gtt gct gct gct gca
acc gct cca gcg █████ CGT GTA CAC GGG ATA GCA TAC CTT CAG GGG CAT TTC
M R V H G I A Y L Q G H F

CCT AGA AGG CAG AAA CCG ACG GCC ACA AAC CGT ACC GTC GTG CCA TCG CGC
P R R Q K P T A T N R T V V P S A

TCG TGC CTC TGG CTG TTA TTG TTC GCG GCG ATC TTT GTT GCT GGA GGA GTA
S C L W L L L F A A I F V A G G V

GCA GCT GAT GCG CTT GGC AAG CCG AAC GTA TTC TAC GTC AAG CTC AAC ACC
A A D A L G K R N V F Y V K L N T

ACG GAG AAA ACG TAC AAG GCG AGC GAT GAG AAA CCG ACG CTG TGG TCG AAG
T E K T Y K A S D E K P T L W S K

GAA CAG GAC GCA CCC CGA ATA CCG CCG CAC TTC AGT CTG GCC AAA ACG TCG
E Q D A P R I R P H F S L A K T S

CAC TTT CTG TTC GAC CTG AAC CGA TAC CCG TGC GAG GTG CCG AGC GCC GGC
H F L F D L N R Y R C E V A S A G

AAT GGA ACA GCG GCC GTC CCG CTG GTT AGT GCC GAC TGC CGC AAC CGG ACC
N G T A A V P L V S A D C R N R T

GCA CTG TTC CGG GCG AAA ATC ATC GGC GAA CTG AGC CGA CGG TTA AAG GAA
A L F R A K I I G E L S R R L K E

CCG GTC CCG GCG AGC GGA CAG TAT CAG GTG CCG TAT AGC CGG CCG TTC CAG
P V P A S G Q Y Q V R Y S R P F Q

CAG CGA CCC CTC GGT ACG GGT CAG TGC CCG TTG CGG CAG GCG AAT GTG CGC
Q R P L G T G Q C R L R Q A N V R

GTC CTG TCC TGG AAG GAG CCA CCG TTT GCG TGG AAC GAG CTC GGC ACT TAC
V L S W K E P P F A W N E L G T Y

CTA CCG GTG ACG CCA TTG TTC GAT CCG GAT ACG GCT ACC AAC AGG AGC TGC
L P V T P L F D P D T A T N R S C

GTG ATA GTG GCC AGC GCC GGT TCG CTG AAA CGC TCC CGG CTT GGG CCG TTC
V I V A S A G S L K R S R L G P F

ATC GAT GGT CAC GAT ATT GTG ATG CCG TTC AAC CAT GCA CCA ACC GAG GGC
I D G H D I V M R F N H A P T E G

TTC GAA CCG GAC GTC GGA GCG AAG ACG ACG ATC CGC GTA CTC AAC TCG CAG
F E A D V G A K T T I R V L N S Q

GTC GTC ACG AAG CCT GAG TAC CCG TTG CTG ACG GCA CCA CTG TTC CGG AAC
V V T K P E Y R L L T A P L F R N

GTT ACG ATA GCG GCC TGG GAT CCG GGC AAG TTC GAT CAA CCG CTG TCC GAG
V T I A A W D P G K F D Q P L S E

TGG TTG GCC AAT CCG GAC TTT AAT CTG TTC GAG AAC TTC AAG AAG TTT CGC
W L A N P D F N L F E N F K K F R

AGC CAG TAC CGG GAT GCC AAC TTC CAC ATC ATC GAT CCG CGC AGC ATC TGG
S Q Y R D A N F H I I D P R S I W

CGG GCG TGG ACG GCC CTT CAG GAT CTC ACT CCC TTT CCG ATC CGA CAG AAC
R A W T A L Q D L T P F P I R Q N

CCT CCA ACA TCC GGT TTC ATC GGC CTT GGA CTG CTG CTG CCC GTG TGC CGG
P P T S G F I G L G L L L P V C R

TAC ATT GAC ATC GTG GAG TAC ATT CCG AGC ACT CGC AT
Y I D I V E Y I P S T R

```

Figura 4.14. Secuencia de extremo 5' de la $\alpha 2,6$ -Sialiltransferasa de *Anopheles albimanus*. Se muestra sombreado el codón de inicio (ATG) y la región compartida con el fragmento de 471pb previamente descrito. La traducción se muestra bajo la secuencia nucleotídica.

La secuencia obtenida tiene una longitud de 1211pb, comparte 438pb con el extremo 5' del fragmento de 471pb previamente identificado y secuenciado. Esta región se muestra sombreada en la **Figura 4.14**. La traducción a aminoácidos de esta secuencia nos indica que posee un marco abierto de lectura, el codón de inicio (**ATG**) también se muestra sombreado y no posee codones de término. La secuencia de aminoácidos coincide con la deducida del fragmento de 471pb. Se detectó la presencia de dominios estructurales y funcionales característicos de las sialiltransferasas, entre ellos los *sialylmotifs* L y S, así como los residuos conservados dentro de estos. Los detalles de este análisis se presentan en la sección 4.4.

4.3.2. 3'-RACE (Rapid Amplification of cDNA Ends)

Para amplificar el extremo 3' del ADNc de la α 2,6-Sialiltransferasa de *Anopheles albimanus* se usó como templado ADNc de larvas *instar* 2 sintetizado con el oligonucleótido **APT18** (**Tabla 3.2.**). Este agrega un adaptador al tallo poli(A) de los ARNm y a partir del mismo es posible obtener el extremo 3' del transcrito mediante PCR.

Se realizaron PCR con ADNc **APT18** y los oligonucleótidos **STGSP2-AP** y **STNGSP2-AP** (**Tabla 3.2.** y **Tabla 4.10.**) con **ProofStart DNA Polymerase**. Las condiciones de reacción fueron: 94°C 5min 1 ciclo, 94°C 45seg, 65°C 45seg, 72°C 2min 40 ciclos, 72°C 10min. Se obtuvieron múltiples bandas distribuidas entre 100pb y 1200pb (**Figura 4.16.**), por lo que fue necesario hacer reamplificaciones.

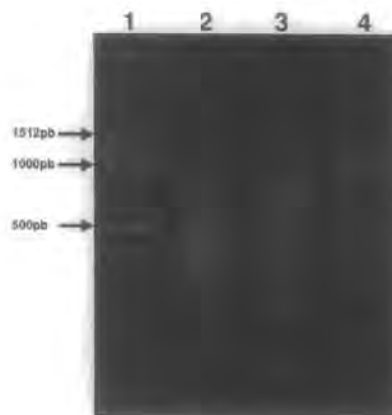


Figura 4.15. Amplificación 3'-RACE

1. Marcador
2. STGSP1 -AP
3. STNGSP1-AP
4. Marcador

Con el fin de aumentar la especificidad, se realizó una reacción PCR usando como templado 1 μ l de las reacciones mostradas en la **Figura 4.15**. Se usaron los oligonucleótidos **STNGSP2-AP** bajo las siguientes condiciones de reacción: 94°C 45seg, 70°C 45seg, 72°C 2min por 40 ciclos con **Vent_R**, los productos de la amplificación se muestran en la **Figura 4.16**.

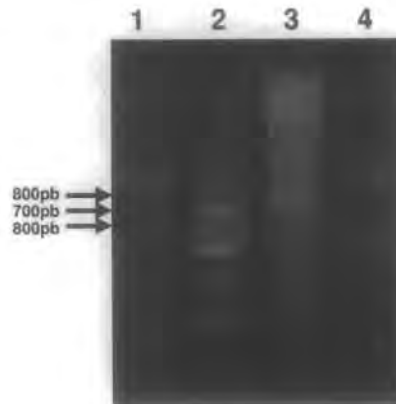


Figura 4.16. Reamplificaciones 3'-RACE

1. Marcador
2. STNGSP2-AP
3. STNGSP2-AP
4. Marcador

Entre los productos de amplificación mostrados en los carriles 2 y 3 (**Figura 4.16.**) se observa una banda de alrededor de 700pb. Estas dos bandas se purificaron y fueron usadas como templado en *seminested*-PCR para confirmar que éstas correspondían al extremo 3' del transcrito buscado. Las reamplificaciones se realizaron con los oligonucleótidos mostrados en la **Tabla 4.10.** y las tallas de las amplificaciones esperadas se muestran en la **Tabla 4.11.** Las condiciones de reacción fueron: 94 °C 45seg, 70 °C 45seg, 72 °C 2min por 40 ciclos con **Vent_R**. Los productos de las reamplificaciones se muestran en la **Figura 4.17.**

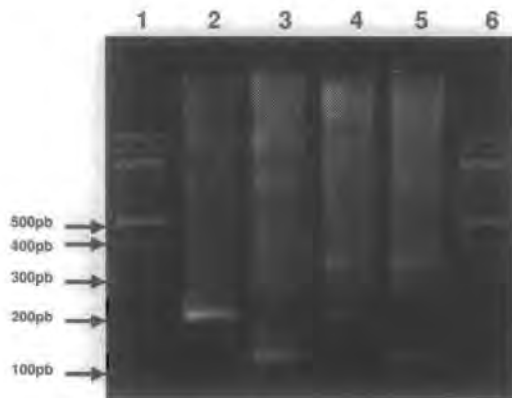


Figura 4.17. Reamplificaciones *seminested*-PCR 3'-RACE

- | | |
|-------------------|----------------------|
| 1. Marcador | Amplificación |
| 2. STNGSP2-STGSP1 | 185pb |
| 3. STNGSP2-STGSP1 | 113pb |
| 4. STNGSP2-STGSP1 | 185pb |
| 5. STNGSP2-STGSP1 | 113pb |
| 6. Marcador | |

Al realizar las reacciones PCR con los oligonucleótidos específicos se obtuvieron los productos previstos. Por lo que las bandas de 700pb mostradas en la **Figura 4.16.** sí corresponden al extremo 3' del transcrito de la α 2,6-Sialiltransferasa putativa de *An.albimanus*. Al clonar y secuenciar esta banda se confirmó que el fragmento contiene el extremo 3' del ADNc. La secuencia nucleotídica de dicho fragmento se muestra en la **Figura 4.18.**, nótese el tallo poli(A) en el extremo 3' de la secuencia.

```

TAC CGG GAT GCC AAC TTC CAC ATC ATC GAT CCG CGC AGC ATC TGG CGG GCG
Y R D A N F H I I D P R S I W R A
TGG ACG GCC CTT CAG GAT CTC ACT CCC TTT CCG ATC CGA CAG AAC CCT CCA
W T A L Q D L T P F P I R Q N P P
ACA TCC GGT TTC ATC GGC CTT GGA CTG CTG CTG CCC GTG TGC CGG TAC ATT
T S G F I G L G L L L P V C R Y I
GAC ATC GTG GAG TAC ATT CCG AGC ACT CGC ATG AAC GGT CTC TGT CAC TAT
D I V E Y I P S T R M N G L C H Y
TAC GAT GAA GAG CTC AAT CTT GGC TGT ACG TTC GGC GCG TGG CAC CCG CTG
Y D E E L N L G C T F G A W H P L
GCG GCC GAG AAG CTG CTC GTC TTC CAG ATG AAT GCC GCC AGC GAT TAC ACC
A A E K L L V F Q M N A A S D Y T
GTC TTT CAG CAG GGC ACC GTC CGC ATC TAC AGT GAT ACG CTG GAT CAA TGC
V F Q Q G T V R I Y S D T L D Q C
■ ttc gtc cgg ccg tcc gtc cgt cca ttc cgt aaa cca ctt ccg gca cct
*
gtc aca cct gtg acc ggg gca tcg gcg cgt gag cta tgg cag gtg gtg gtg
caa aac aat tag caa aca gaa cga att aat taa gca atg cga atc gac gac
agc taa ggt cgt gct gga tcg atg ccc gtc agc gta tta gac att cca gcc
gag gct ctg tgc gcg att tcg atg aca ata tcg acg acg cgg atc aga ccg
aga cca aac agc aac cag cca gga gtg cag tgg gta cgg ttt tgc att tga
agc tag tga cga cga tac ccg ggc acg gag tag caa atc aaa aaa aaa
aaa aaa

```

Figura 4.18. Secuencia de extremo 3' de la α 2,6-Sialiltransferasa de *Anopheles albimanus*. Se muestra resaltado el codón de término (TAG), la señal de poliadenilación (AATAAA) y la región compartida con el fragmento de 218pb previamente descrito. La traducción se muestra bajo la secuencia nucleotídica.

La secuencia obtenida tiene una longitud de 720pb (sin tomar en cuenta la secuencia del adaptador AP). Comparte 218pb con el extremo 3' del fragmento de 471pb previamente identificado y secuenciado y también con la obtenida del 5'-RACE. Esta región se muestra resaltada en la **Figura 4.18**. La traducción a aminoácidos de esta secuencia nos indica que posee un marco abierto de lectura que coincide con la traducción deducida del fragmento de 471pb y con la obtenida del 5'-RACE. Presenta un único codón de término (TAG) y 310pb hacia el 3' del mismo se encuentra la señal de poliadenilación (AATAAA).

Se detectó también la presencia de las regiones características de las sialiltransferasas. Entre ellos los *sialylmotifs* S y VS y los residuos conservados dentro de estos. Los detalles de este análisis se presentan en la sección 4.4.

4.3.3. Clonación ADNc $\alpha 2,6$ -Sialiltransferasa

Habiendo ya obtenido y caracterizado los productos de los extremos 5' y 3' del transcrito de la $\alpha 2,6$ -Sialiltransferasa de *An.albimanus*, se procedió a la amplificación de la región codificante a partir del ADNc utilizado para el 5'-RACE.

El ADNc de la $\alpha 2,6$ -Sialiltransferasa de *Anopheles albimanus* tiene una longitud total de 1746pb. Un marco abierto de lectura identificado por **ORFinder** correspondiente a 440aa de longitud. La secuencia nucleotídica junto con la secuencia de aminoácidos se muestran en la **Figura 4.19**.

```

1      cgcgggggtgtggctgtggcagcgaacaaagtgtgctggtgctgctgctgcaaccgctccagcgATGCGGTACACGGGATAGCATACTT
1      M R V H G I A Y L
91     CAGGGGCATTTCCCTAGAAGGCAGAAACCGACGGCCACAAACCGTACCGTGTGCCATCGGCGTGTGCCTCTGGCTGTTATTGTTCCGG
10     Q G H F P R R Q K P T A T N R T V V P S A S C L W L L L F A
181    GCGATCTTTGTTGCTGGAGGAGTAGCAGCTGATGCGCTTGGCAAGCGGAACGTATTCTACGTCAAGCTCAACACCAGGAGAAAAACGTAC
40     A I F V A G G V A A D A L G K R N V F Y V K L N T T E K T Y
271    AAGGCGAGCGATGAGAAAACCGACGCTGTGGTTCGAGGAACAGGACGCACCCCGAATACGGCCGCACTTCAGTCTGGCCAAAACGTGCGAC
70     K A S D E K P T L W S K E Q D A P R I R P H F S L A K T S H
361    TTTCTGTTCCGACCTGAACCGATAACCGGTGCGAGGTGGCGAGCGCCGGCAATGGAACAGCGCGCTCCCGCTGGTTAGTCCCGACTGCCGC
100    F L F D L N R Y R C E V A S A G N G T A A V P L V S A D C R
451    AACCGGACCGCACTGTTCCGGGGGAAAATCATCGGCCAACTGAGCCGACGTTAAAGGAACCGGTCGCCGGCAGCGGAAGTATCAGGTG
130    N R T A L F R A K I I G E E L S R R L K E P V P A S G Q Y Q V
541    CGGTATAGCCGGCCGTTCCAGCAGCGAACCCCTCGGTACGGGTGAGTGCCTGCGGCGAGCGAATGTGCGCGCTCCTGTCTGGAAGGAG
160    R Y S R P F Q Q R P L G T G Q C R L R Q A N V R V L S W K E
631    CCACCGTTTTCGTTGGAACGAGCTCGGCACTTACCTACCGGTGACGCCATTGTTTCGATCCGGATACGGCTACCAACAGGAGCTGCCGTATA
190    P P F A W N E L G T Y L P V T P L F D P D T A T N R S C V I
721    GTGGCCAGCGCCGGTTCGTTGAAACGCTCCCGGCTTGGGCGGTTTCATCGATGGTCAAGTATTGTGATCGGGTTCAACCATGCACCAACC
220    V A S A G S L K R S R L G P F I D G H D I V M R F N H A P T
811    GAGGGCTTCGAAGCGGACGTCGGAGCGAAGACGACGATCCGCGTACTCAACTCGCAGGTCGTCAGGAAGCCTGAGTACCGGTTGCTGACG
250    E G F E A D V G A K T T I R V L N S Q V V T K P E Y R L L T
901    GCACCTGTTCCGGAAACGTTACGATAGCGGCTCGGGTCCGGCAAGTTCGATCAACCGCTGTCGAGTGGTGGCCAAATCCGGACTTT
280    A P F L F R N V T I A A W D P G K F F D Q P L S E W L A N P D F
991    AATCTGTTTCGAGAACTTCAAGAAGTTTCGAGCCATACCGGGATGCCAACTTCCACATCATCGATCCGCGCAGCATCTGGCGGGCGTGG
310    N L F E N F K K F R S Q Y R D A N F H I I D P R S I W R A W
1081   ACGGCCCTTCAGGATCTCACTCCCTTTCGATCCGACAGAACCCCTCCAAACATCCGGTTTCATCGGCCCTTGGACTGCTGCTGCCCGGTGTC
340    T A L Q D L T P F P I R Q N P P T S G F I G L G L L L P V C
1171   CGGTACATTGACATCGTGGAGTACATCCGAGCACTCGCATGAACGGTCTCTGTCACTATTACGATGAAGAGCTCAATCTTGGCTGTACG
370    R Y I D I V E Y I P S T R M N G L C H Y Y D E E L N L G C T
1261   TTGGCGCGTGGCACCCGCTGGGGGCCGAGAAGTCTGCTCTTCCAGATGAATGCCGCCAGCGATTACACCGTCTTTCAGCAGGGCACC
400    F G A W H P L A A E K L L V F Q M N A A S D Y T V F Q Q G T
1351   GTCCGATCTACAGTACGCTGGATCAATGCTAGTctgctcggcgctcgcgtccgtccattccogtaaaccaactccggcactgtccaca
430    V R I Y S D T L D Q C *
1441   cctgtgaccggggcatcggcggcgtgagctatggcaggtggtggtgcaaaacaattagcaaacgagaacgaattaattaagcaatgogaate
1531   gacgacagctaaggtcgtgctggatcgatgccgtcagcgtattagacattccagccgaggctctgtgocgatttcgatgacaatatcg
1621   scgacgaggatcagaccgagaccgaacagcaaccagccaggagtgcagtggtacggttttgcattgaagctagtAataaagacgatac
1711   ccgggacgggagtgcgaatcaaaaaaaaaaaaaaa

```

Figura 4.19. Secuencia del ADNc de la $\alpha 2,6$ -Sialiltransferasa de *Anopheles albimanus* (1746pb)

Se muestra el marco de lectura abierto de 440aa, así como resaltados el codón de inicio (ATG) y término (TAG), la señal de poliadenilación (AATAAA) y el tallo poll(A).

La amplificación de la región codificante del ADNc conteniendo el ORF de 440aa se realizó con los oligonucleótidos específicos **STAA1** 5'-GCTGCAACCGCTCCAGCGATG-3' y **STAA2** 5'-ACGGCCGGACGAACTAGCATTG-3' (las bases subrayadas en los oligonucleótidos corresponden a los codones de inicio y término). El templado para la reacción PCR fue el ADNc de mosquitos adultos hembras 18 horas postalimentación del que se obtuvo el extremo 5' del transcrito. Las condiciones de reacción fueron las siguientes: 95°C 5min 1 ciclo, 94°C 45seg, 68°C 45seg, 72°C 2min, 35 ciclos, 72°C

10min, con **ProofStart DNA Polymerase**. La reacción produjo la amplificación de 1354pb esperada. La identidad de la misma se confirmó a través de *nested*-PCR con los oligonucleótidos **STGSP1** y **STGSP2**, mismos que producen una amplificación de 427pb, como se observa en la **Figura 4.20**.

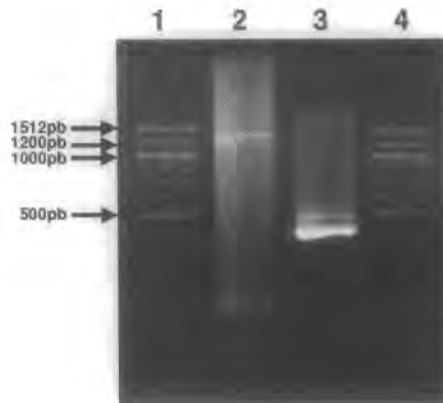


Figura 4.20. Amplificación ADNc α 2,6-Sialiltransferasa *Anopheles albimanus* (1354pb)

	Amplificación
1. Marcador	
2. STAa1-STAa2	1354pb
3. STGSP1-STGSP2	427pb
4. Marcador	

La banda de 1354pb, correspondiente al ADNc de la α 2,6-Sialiltransferasa de *Anopheles albimanus* fue clonada en el plásmido pCR-Blunt II-TOPO. Se seleccionaron clonas con el inserto y se purificaron plásmidos para secuenciación.

El transcrito corresponde a la α 2,6-Sialiltransferasa putativa de *Anopheles albimanus*, su secuencia fue sometida a la base de datos **GenBank** atribuyéndosele el número de acceso **AY955495**. Desde este punto nos referiremos a esta secuencia como **AaST6**.

4.4. Análisis de la secuencia de la α 2,6-Sialiltransferasa de *Anopheles albimanus* (**AaST6**)

Las α 2,6-Sialiltransferasas poseen características topológicas que permiten agruparlas en la familia 29 de las glicosiltransferasas. Dicha familia agrupa glicosiltransferasas que actúan como invertasas y que tienen como azúcar donador al CMP- β -N-acetilneuraminato (CMP-Neu5Ac o ácido siálico) (Harduin-Lepers *et al.*, 1995; Marchler-Bauer *et al.*, 2004).

A continuación se presentan las evidencias que permiten incluir a la secuencia **AaST6** aquí identificada como parte de la Familia 29 de Glicosiltransferasas: Sialiltransferasas.

4.4.1. Identidad de la α 2,6-Sialiltransferasa putativa de *Anopheles albimanus* (**AaST6**)

El análisis de identidad entre la secuencia aquí descrita y las secuencias de las α 2,6-Sialiltransferasas ya reportadas (**Tabla 2.2.**) se realizó con los programas **LALIGN** y **BLAST**. Los resultados se muestran en la **Tabla 4.12**.

Tabla 4.12. Análisis de identidad de la α 2,6-Sialiltransferasa putativa de *Anopheles albimanus* (AaST6)

Organismo	Identidad α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	
	ADNc	Proteína
Humano (<i>Homo sapiens</i>)	49.9 %	30.5 %
Rata (<i>Rattus norvegicus</i>)	50.1 %	27.4 %
Ratón (<i>Mus musculus</i>)	49.0 %	27.4 %
Pollo (<i>Gallus gallus</i>)	49.6 %	28.5 %
Bovino (<i>Bos taurus</i>)	49.0 %	26.3 %
Mosca (<i>Drosophila melanogaster</i> A)	54.1 %	39.1 %
Mosca (<i>Drosophila melanogaster</i> B)	51.9 %	38.2%
Mosca (<i>Drosophila pseudoobscura</i>)	50.4 %	38.4 %
Mosca (<i>Drosophila yakuba</i>)	51.7 %	38.3 %
Mosquito (<i>Anopheles gambiae</i>)	64.0 %	66.1 %

El grado de identidad se obtuvo al comparar las secuencias de ADNc en sus regiones codificantes, ya que las regiones no traducidas de las secuencias de los diferentes organismos pueden ser altamente variables en longitud debido al uso diferencial de promotores que trae consigo la obtención de transcritos de diferentes longitudes (O'Hanlon *et al.*, 1989; Paulson, 1989; Wen *et al.*, 1992b; Harduin-Lepers *et al.*, 1995; Tsuji, 1996).

La traducción del transcrito de *An.albimanus* indica la presencia de un marco de lectura abierto de 440aa en el cual mediante **BLAST**, **Conserved Domain-Search (CD-Search)** y **Conserved Domain Database**, se detectó la presencia de dominios putativos funcionales y estructurales conservados indicando que pertenece a la Familia 29 de las Glicosiltransferasas: Sialiltransferasas (Marchler-Bauer *et al.*, 2004; Marchler-Bauer *et al.*, 2005).

Derivado de los programas de secuenciación, recientemente se reportó la presencia de α 2,6-Sialiltransferasas putativas en los genomas de *Drosophila yakuba* y *D.pseudoobscura* (Harduin-Lepers *et al.*, 2005), cuyas secuencias (AJ821848 y AJ821849) están disponibles en **GenBank**. *Drosophila melanogaster* presenta dos isoformas (A y B) de α 2,6-Sialiltransferasa. Estas difieren en 69pb (23aa) debido a un *splicing* alternativo en la región del tallo de esta proteína, no afectando el dominio catalítico de la misma. Entre los insectos, únicamente se ha reportado la presencia de α 2,6-Sialiltransferasas en miembros de los géneros *Drosophila* y *Anopheles*.

Los porcentajes de identidad de las secuencias de proteínas respecto a *An.albimanus* de la **Tabla 4.12**, son muy parecidos a los mostrados en la **Tabla 4.5**, del mismo análisis respecto a *An.gambiae*. Comparando las secuencias de ambas especies de *Anopheles* se obtiene que comparten 66.1% de identidad.

La longitud de las secuencias proteicas en los diferentes organismos es variable. Entre los vertebrados las sialiltransferasas tienen una longitud de entre 403 y 406aa. En los invertebrados se encuentran proteínas más grandes, por ejemplo 451aa en la isoforma A de *D.melanogaster*, 474aa en la isoforma B de *D.melanogaster*, 473aa en *D.pseudoobscura*, 479aa en *D.yakuba* y 440aa en *An.albimanus*.

En la **Tabla 4.13.** se muestra el análisis de identidad de las regiones catalíticas exclusivamente. Desde el inicio del *sialylmotif L* hasta el extremo 3' de los ADNc o extremo carboxilo de las proteínas.

Tabla 4.13. Análisis de identidad de la región catalítica de la α 2,6-Sialiltransferasa putativa de *Anopheles albimanus*

Organismo	Identidad α 2,6-Sialiltransferasa de <i>Anopheles albimanus</i>	
	ADNc	Proteína
Humano (<i>Homo sapiens</i>)	54.5 %	40.8 %
Rata (<i>Rattus norvegicus</i>)	54.2 %	38.3 %
Ratón (<i>Mus musculus</i>)	54.3 %	38.8 %
Pollo (<i>Gallus gallus</i>)	54.5 %	40.4 %
Bovino (<i>Bos taurus</i>)	53.5 %	39.6 %
Mosca (<i>Drosophila melanogaster A</i>)	61.5 %	57.0 %
Mosca (<i>Drosophila melanogaster B</i>)	61.5 %	57.0 %
Mosca (<i>Drosophila pseudoobscura</i>)	60.1 %	57.9 %
Mosca (<i>Drosophila yakuba</i>)	61.6 %	57.0 %
Mosquito (<i>Anopheles gambiae</i>)	75.9 %	86.9 %

La α 2,6-Sialiltransferasa putativa de *An.gambiae* comparte un 59.9% de identidad respecto a la de *D.melanogaster*, mientras que la secuencia que aquí se reporta de *An.albimanus* es 57% y 86.9% idéntica respecto a *D.melanogaster* y *An.gambiae*, respectivamente. Las regiones catalíticas de las sialiltransferasas de los miembros del género *Anopheles*, comparten cerca del 90% de identidad.

4.4.2. Topología de la α 2,6-Sialiltransferasa de *Anopheles albimanus*

Las α 2,6-Sialiltransferasas se caracterizan por compartir organización y conformación. Dentro de la región catalítica encontramos pequeñas regiones de secuencias que poseen algunos residuos invariables a los que se les ha llamado *Sialylmotifs* (Drickamer, 1993; Livingston et al, 1993; Harduin-Lepers et al, 1995). En vista de que tanto las proteínas de esta familia hasta ahora caracterizadas como la secuencia ortóloga putativa identificada en el genoma de *Anopheles gambiae* poseen estos *sialylmotifs*, se procedió a identificarlos en la traducción de la secuencia AaST6 de *Anopheles albimanus*. Para ello se realizó un alineamiento múltiple mostrado en la **Figura 4.21.** incluyendo las secuencias de las α 2,6-Sialiltransferasas de invertebrados.

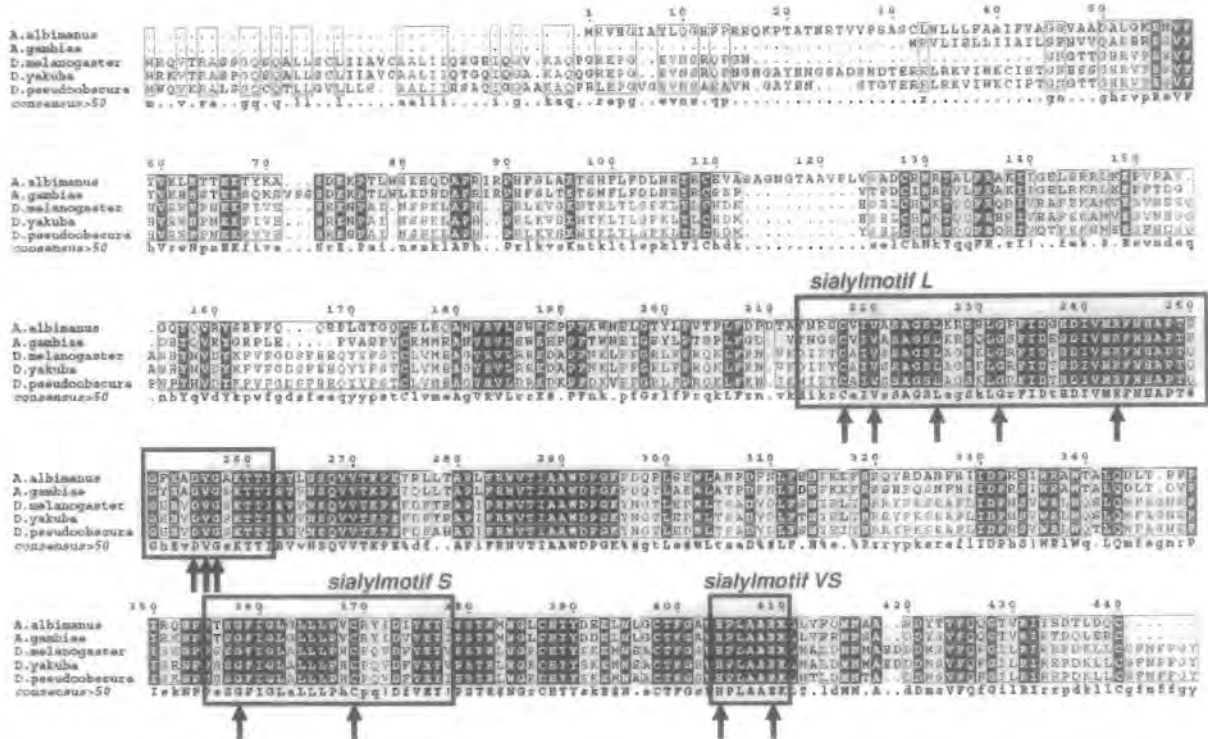


Figura 4.21. Alineamiento de las $\alpha 2,6$ -Sialiltransferasas de insectos.

Los residuos conservados se muestran resaltados, los *sialylmotifs* L, S y VS se encuentran enmarcados con recuadros y los residuos implicados en el reconocimiento de sustrato donador y aceptor se señalan con flechas.

Se sabe que el *sialylmotif* L posee 8 residuos invariables y que los *sialylmotifs* S y VS tienen también dos residuos invariables cada uno (Datta *et al.*, 1995; Geremia *et al.*, 1997; Datta *et al.*, 1998; Jeanneau *et al.*, 2004). En el alineamiento de la Figura 4.21, podemos ver que estos 12 residuos están presentes e indicados con flechas. Lo que indica que la organización dentro de los *sialylmotifs* así como su longitud están conservados entre las sialiltransferasas de insectos aquí mostradas. Las cisteínas conservadas presentes en L y S (Figura 4.21.) están implicadas en la formación de un puente disulfuro, por lo que su presencia es esencial para la correcta conformación enzimática y obtención de proteínas activas (Geremia *et al.*, 1997; Datta *et al.*, 1998; Datta *et al.*, 2001).

Debido a la cercanía filogenética entre los organismos mostrados en el alineamiento de la Figura 4.21, las regiones dentro y entre los *sialylmotifs* muestran gran conservación, esto se ve además reflejado en el porcentaje de identidad entre estas secuencias (Tabla 4.13.). Por lo que es de esperarse que estas proteínas adopten conformaciones muy similares. Estas regiones que están altamente conservadas entre las sialiltransferasas de insectos fueron sobre las cuales se diseñaron los oligonucleótidos con los cuales se obtuvo la primera amplificación del fragmento de 471pb. Podrían ser punto de partida en el diseño de oligonucleótidos para la búsqueda de este tipo de transcritos en otros insectos.

El dominio catalítico de las Sialiltransferasas está altamente conservando, no así el tallo que presenta diferencias en longitud y muy pocos residuos conservados entre los organismos analizados (Figura 4.21.). La región entre los residuos 112 y 124 de AaST6 de *An.albimanus* es exclusiva de este organismo, cabe la posibilidad de que estas 33pb constituyan un intrón retenido o ser producto de un *splicing* alternativo y que su amplificación fue debido a un artefacto de PCR. Su localización respecto al

dominio catalítico nos indica que ésta no tendría efecto sobre la actividad de la enzima, pero podría afectar la conformación de la proteína.

Si extrapolamos este tipo de análisis e incluimos las secuencias de vertebrados obtenemos el alineamiento mostrado en la **Figura 4.22**.

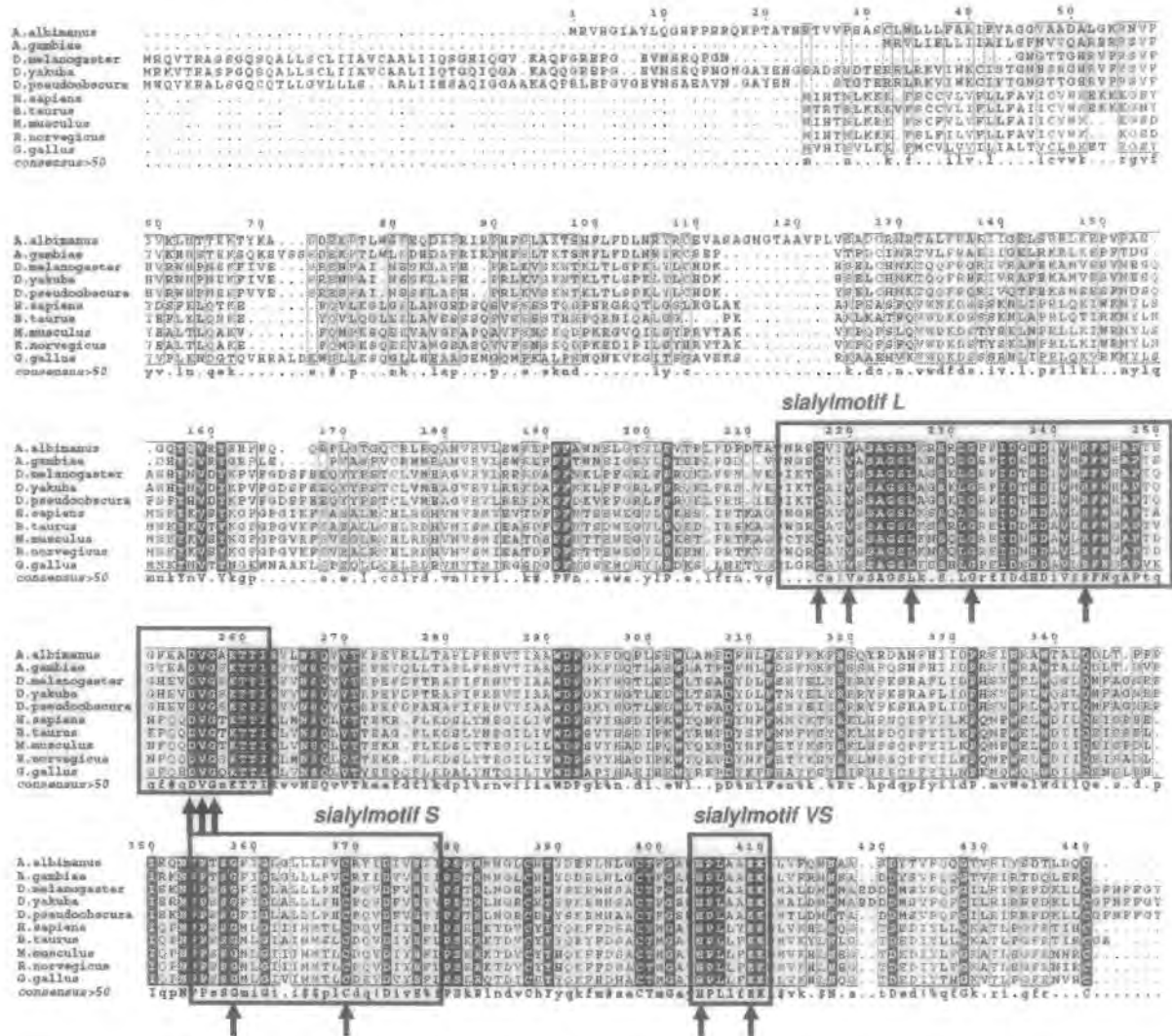


Figura 4.22. Alineamiento de las $\alpha 2,6$ -Sialiltransferasas de insectos y vertebrados. Los residuos conservados se muestran resaltados, los *sialylmotifs* L, S y VS se encuentran enmarcados con recuadros y los residuos implicados en el reconocimiento de sustrato donador y aceptor se señalan con flechas.

En este alineamiento los recuadros indican la localización de los *sialylmotifs* y las flechas los residuos que deben estar conservados para mantener la actividad catalítica y que permiten el reconocimiento de sustratos donador y aceptor. La topología de esta región catalítica hasta el extremo carboxilo es bastante conservada entre los diferentes organismos. En la región entre los *sialylmotifs* tenemos varios residuos 100% conservados, de los cuales los encontrados entre el S y VS parecen estar implicados en la catálisis enzimática (Jeanneau et al, 2004).

El extremo amino presenta mayor variación y es específico para cada organismo. Esta región determina en gran medida la longitud final de la enzima. En ella ocurren diversas modificaciones tal como cortes

proteolíticos que producen enzimas solubles o algunas modificaciones postraduccionales que pueden impactar en la conformación de la proteína (Breton *et al.*, 1999).

Con este análisis se predice que la α 2,6-Sialiltransferasa de *Anopheles albimanus* posee un dominio citoplasmático N-terminal de 32aa de longitud y un segmento transmembranal es de 17aa y participa en la retención en el aparato de Golgi (Paulson *et al.*, 1989; Breton *et al.*, 1999) (Figura 4.23.).

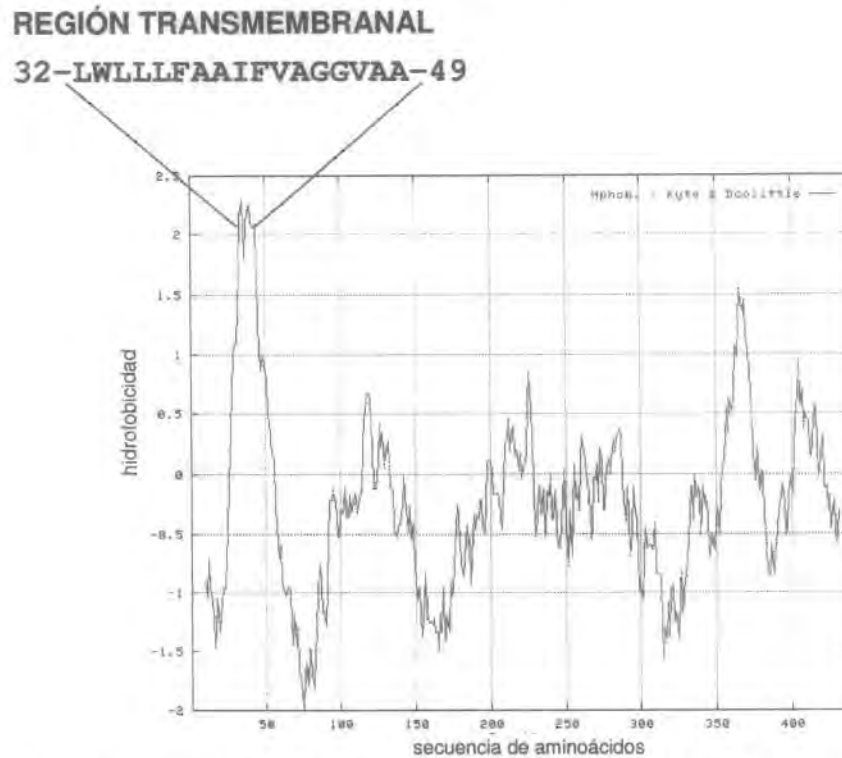


Figura 4.23. Análisis de hidrofobicidad de la α 2,6-Sialiltransferasa de *Anopheles albimanus* mediante el método Kyte&Doolittle y representación de la región transmembranal.

La región transmembranal de la proteína se predijo a partir de un análisis de hidrofobicidad realizado mediante **TMPred** y **PSIPRED** utilizando el algoritmo de Kyte y Doolittle (Kyte *et al.*, 1982) (Figura 4.23.). Las Sialiltransferasas de organismos vertebrados e invertebrados presentan este mismo patrón de hidrofobicidad pese a que esta región es variable en secuencia, tal como podemos verlo en la Figura 4.22.

El patrón de hidrofobicidad indica la presencia de una región transmembranal altamente hidrofóbica entre los primeros 50aa de la proteína que ayuda a la retención de la proteína en la membrana del compartimento *trans*-Golgi (Paulson *et al.*, 1989; Breton *et al.*, 1999). De acuerdo a la escala de hidrofobicidad de Kyte&Doolittle las regiones transmembranales poseen >1.6 unidades en el centro. La secuencia primaria nos indica la presencia de una sola región transmembranal de 17aa de longitud y con 2.2 unidades de hidrofobicidad en su centro (Figura 4.23.).

4.4.3. Modificaciones postraduccionales de AaST6

Al analizar la secuencia primaria de una proteína es posible identificar *sitios* o *consensos* específicos donde pueden ocurrir diferentes modificaciones postraduccionales. Tales como glicosilación, fosforilación, amidación, etc. que pueden tener un impacto importante a nivel funcional y estructural.

Se sometió la secuencia primaria de la α 2,6-Sialiltransferasa de *Anopheles albimanus* (440aa de longitud) a una búsqueda de posibles sitios que indiquen modificaciones postraduccionales mediante el algoritmo **PROSITE**. Se localizaron diferentes sitios potenciales para diferentes modificaciones, mismos que se muestran en la **Tabla 4.14**.

Tabla 4.14. Modificaciones postraduccionales de la α 2,6-Sialiltransferasa de *Anopheles albimanus*

Modificación postraducciona	Localización (aa)
N-Glicosilación	23
N-X-S/T-X	63
X = cualquier aminoácido excepto Pro	116
	130
	214
	285
Fosforilación Proteína Cinasa C	22
S/T-X-R/K	65
	68
	144
	186
	213
	225
	261
	380
	429
Fosforilación Caseína Cinasa II	186
S/T-X-X-D/E X = cualquier residuo	271
Fosforilación Tirosina Cinasa	377
R/K-X-D/E-X-Y X = 2 o 3 residuos cualquiera	
Miristoilación	45
G-[excepto EDRKHPFYM]-X-X-[STAGCN]-	115
[excepto P]	171
	257
	397
Amidación	53
X-G-R/K-R/K	

Entre estas modificaciones postraduccionales, la N-glicosilación es una característica conservada entre las sialiltransferasas de diferentes especies ya que en todas es posible identificar estos sitios potenciales

(Vallejo-Ruiz *et al.* 2001). En la **Figura 4.24**, se señalan los posibles sitios de *N*-glicosilación en las α 2,6-Sialiltransferasas.

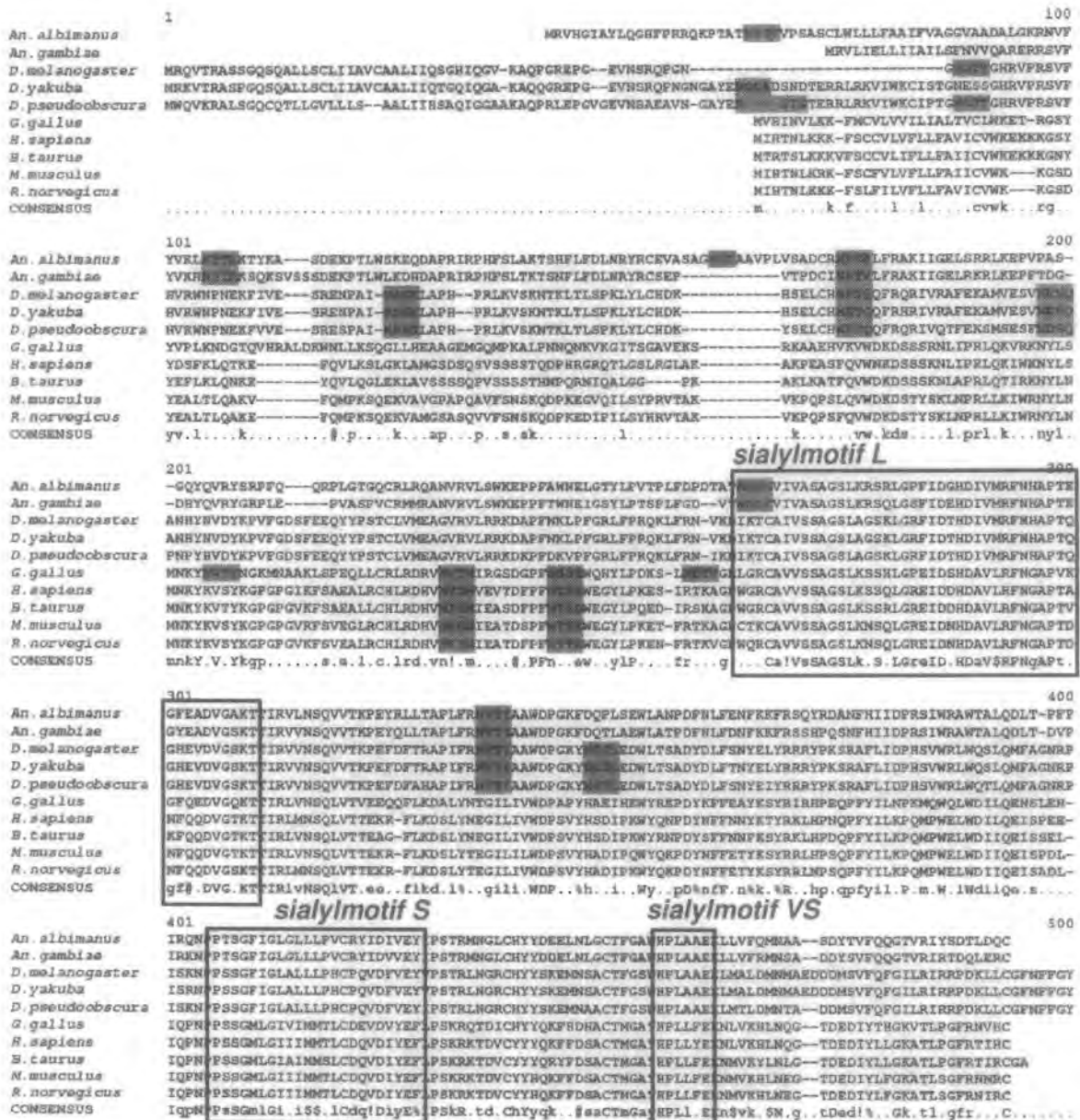


Figura 4.24. Localización de los sitios de *N*-glicosilación en las α 2,6-Sialiltransferasas
Los sitios potenciales de *N*-glicosilación están sombreados, los *sialylmotifs* están enmarcados por recuadros.

Se puede observar que los sitios de *N*-glicosilación se encuentran en su mayoría en la región del tallo, que es una región variable, flexible y específica para cada organismo. Los glicanos presentes en esta región regulan la conformación local de la proteína y promueven su correcta localización intracelular (Chen *et al.* 2000; Breton *et al.* 2001). En algunos casos se ha visto que son necesarios para la actividad enzimática *in vivo* ya que la falta de glicanos sialilados o galactosilados resulta en un decremento significativo de la actividad enzimática. En el caso de la α 2,3-Sialiltransferasa humana (hST3Gal), la presencia de sitios de *N*-glicosilación en el dominio catalítico cerca o entre los *sialylmotifs* ocupados con glicanos complejos sugirió que estos participan en el mecanismo de acción de la enzima y en la estabilidad estructural de la proteína (Vallejo-Ruiz *et al.* 2001; Breen, 2002).

En la **Figura 4.24.** podemos ver que entre las ST6Gal también existen sitios de *N*-glicosilación, sugiriendo un papel similar al descrito para la hST3Gal. Se sabe además que las proteínas de rata, humano y *D.melanogaster* incluidas en el alineamiento están glicosiladas en forma compleja (Koles *et al.*, 2004). Debido a las diversas funciones que los glicanos ejercen y el hecho de que los sitios potenciales de este procesamiento están altamente conservados en las sialiltransferasas, es probable que *AaST6* de *An.albimanus* esté glicosilada al igual que las proteínas ortólogas.

La actividad sialiltransferasa está controlada también por el estado de fosforilación de la enzima. En la línea celular de neuroblastoma humano 6B100 se observó un decremento de hasta 50% en la actividad ST6Gal I al inhibir las cinasas, la actividad sialiltransferasa se restableció al activar la proteína cinasa C (PKC) (Breen *et al.*, 2003). Esto sin alterar los niveles de expresión genética de esta proteína. La ST6Gal humana está fosforilada en el tallo y dominio catalítico por medio de una caseína cinasa, afectando ña actividad enzimática de la proteína (Ma *et al.*, 1999).

En la **Tabla 4.14.** se enlistan los 10 sitios potenciales de fosforilación por Proteína Cinasa C (PKC), 2 sitios de fosforilación por Caseína Cinasa II y un sitio de fosforilación por Tirosina Cinasa. Al localizarlos sobre la secuencia primaria de la proteína vemos que estos sitios están distribuidos en los diferentes dominios de la enzima. Son particularmente aquellos localizados en el tallo y dominio catalítico los que llaman la atención, corresponden a sitios de fosforilación por PKC y Caseína Cinasa II. En la ST6Gal humana actúan estas enzimas afectando su actividad enzimática (Breen *et al.*, 2003). Esto sugiere que la actividad de la sialiltransferasa de *An.albimanus* pueda también estar controlada por medio de fosforilación. Al analizar las α 2,6-Sialiltransferasas de vertebrados e invertebrados se observó la presencia de estos sitios de fosforilación en todos los organismos. La ST6Gal de *Drosophila melanogaster* posee sitios para las cinasas PKA y PKC, la actividad de estas cinasas está controlada a nivel hormonal (Medvedova *et al.*, 2003).

Se identificaron posibles sitios para miristoilación y amidación (**Tabla 4.14.**) tanto en la sialiltransferasa de *An.albimanus* como en el resto de las secuencias de esta familia. Hasta el momento no se tienen datos experimentales sobre si ocurren este tipo de modificaciones postraduccionales en las proteínas de esta familia. La miristoilación es una modificación que ocurre en el extremo amino terminal de las proteínas, exclusivamente en residuos de glicina que se encuentren en regiones hidrofóbicas. Se cree que la transferencia del ácido mirístico (ácido graso de 12 carbonos) interviene en la asociación y retención de las proteínas membranales (Kraft *et al.*, 2001; Martín, 2001; Reithmeier, 2001). De los cinco posibles sitios de miristoilación identificados en *AaST6* (**Tabla 4.14.**), el localizado en la glicina 45 cumple con los requisitos planteados, ya que se encuentra en la región más hidrofóbica de la proteína (la región transmembranal) que está localizada en el extremo *N*-terminal de la sialiltransferasa. Por lo que es posible que esta modificación postraducciona ocurra en esta enzima.

En la **Figura 4.25.** se muestran las predicciones de estructura secundaria de las α 2,6-Sialiltransferasas, obtenidas por **PSIPRED v2.4.**



Figura 4.25. Predicción de estructura secundaria de las Sialiltransferasas mediante PSIPRED. Los cilindros representan α -hélices, las flechas representan hojas- β . Se incluyó la región del dominio catalítico desde el inicio del *sialylmotif L* hasta el carboxilo terminal. Los *sialylmotifs L, S y VS* se muestran enmarcados en recuadros.

En el alineamiento de las estructuras secundarias predichas para las $\alpha 2,6$ -Sialiltransferasas observamos que la región catalítica de las proteínas de invertebrados y de vertebrados mantiene la misma organización estructural, que consiste en α -hélices y hojas- β presentadas en arreglos $\alpha/\beta/\alpha$ (Jeanneau *et al.*, 2004). La organización del dominio catalítico de las $\alpha 2,6$ -Sialiltransferasas presenta regiones altamente conservadas. Particularmente en los *sialylmotifs*, pese a que estas proteínas provienen de especies distantes evolutivamente, tal como podemos observar en el árbol filogenético de la **Figura 4.26**, donde agrupamos estas proteínas (**GeneBee**).

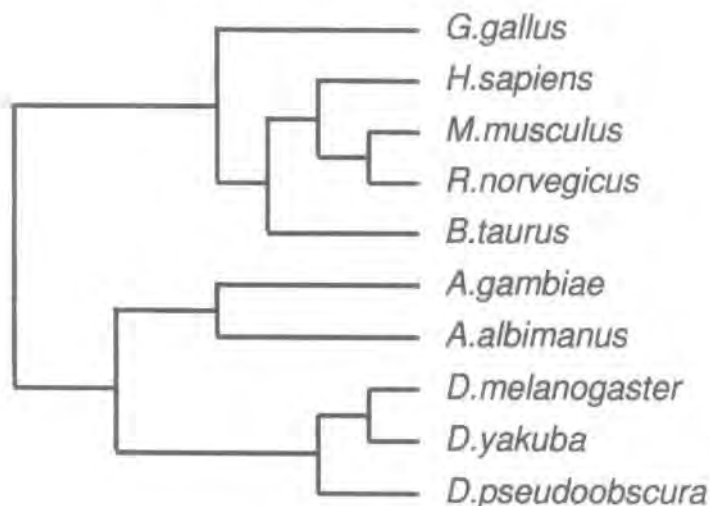


Figura 4.26. Árbol filogenético basado en las secuencias proteicas de las $\alpha 2,6$ -Sialiltransferasas (GeneBee).

Estas regiones que poseen similitud estructural muchas veces no poseen un alto grado de identidad. Se han utilizado otros métodos para predicción de estructura secundaria, tal como *Hydrophobic Cluster Analysis (HCA)* (Goremla *et al.*, 1997). Con el fin de identificar estas regiones de similitud estructural se sometieron las secuencias de la **Figura 4.25**, a análisis *HCA (DrawHCA)* (**Anexo 4**). Se observó la misma organización estructural de α -hélices y hojas- β obtenida con **PSIPRED**, confirmando que la estructura secundaria de las $\alpha 2,6$ -Sialiltransferasas está conservada.

4.5. Análisis ontogenético de expresión de *AaST6*

La expresión de las Sialiltransferasas está controlada en forma espacio-temporal. Con el fin de determinar el patrón de expresión del transcrito *AaST6* a través de las diferentes etapas de desarrollo del mosquito, se purificó ARN total. Las etapas de desarrollo analizadas se muestran en la **Tabla 4.15**. (estas muestra fueron donadas amablemente por el Bioterio del INSP, Cuernavaca, Mor.).

Tabla 4.15. Etapas de desarrollo de *An.albimanus* analizadas

Etapas de desarrollo de <i>Anopheles albimanus</i>	
▶ Huevecillos 24h oviposición	(H24)
▶ Huevecillos 48h oviposición	(H48)
▶ Larvas instar 1	(L1)
▶ Larvas instar 2	(L2)
▶ Larvas instar 3	(L3)
▶ Larvas instar 4	(L4)
▶ Pupas	(P)
▶ Adultos machos	(M)
▶ Adultos hembras-sin alimentación	(F)
▶ Adultos hembras-alimentadas	(Fa)

Se amplificó por PCR un fragmento de 300pb de *AaST6* con los oligonucleótidos *AaST2* 5'-TTCCGGAACGTTACGATAGC-3' y *AaST3* 5'-GCGAGTGCTCGGAATGTACT-3', con las siguientes condiciones de reacción: 94°C 30seg, 55°C 30seg, 72°C 30seg, por 35 ciclos con **Taq DNA Polymerase** tomando 1µl de ADNc como templado. Como control se realizó una amplificación por PCR de un fragmento de 215pb de la actina de *An.albimanus* con los oligonucleótidos *Dact-F* 5'-CCGTAAGGAYCTGTACGCCAA-3' y *Dact-R* 5'-CTCCTGCTTSGAGATCCACAT-3', bajo las condiciones: 94°C 30seg, 55°C 30seg, 72°C 30seg, por 35 ciclos con los mismos reactivos y templado. Las amplificaciones se muestran en la **Figura 4.27**.

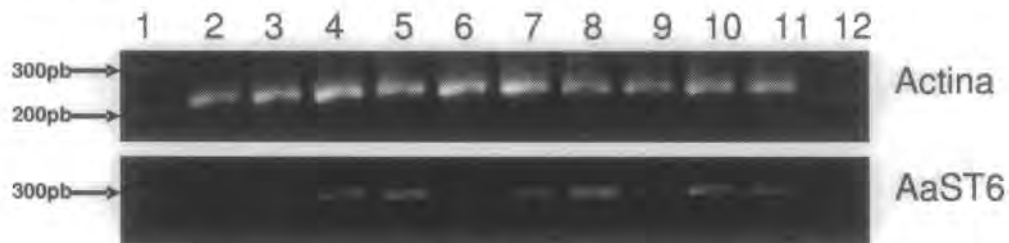


Figura 4.27. PCR semicuantitativo de actina y *AaST6* de *A.albimanus*

Carril 1. Marcador de peso	Carril 7. Larvas instar 4
Carril 2. Huevecillos 24h	Carril 8. Pupas
Carril 3. Huevecillos 48h	Carril 9. Adultos machos
Carril 4. Larvas instar 1	Carril 10. Adultos hembras
Carril 5. Larvas instar 2	Carril 11. Adultos hembras alimentadas
Carril 6. Larvas instar 3	Carril 12. Marcador de peso

La intensidad de las bandas en los geles se analizó por densitometría mediante **Image J 1.33u** (Wayne Rasband, National Institutes of Health, USA). Los datos obtenidos se expresaron como amplificación relativa a actina en cada uno de los estadios analizados. En la **Figura 4.30**, se muestra el perfil de expresión de de *An.albimanus* a través de su ontogenia.

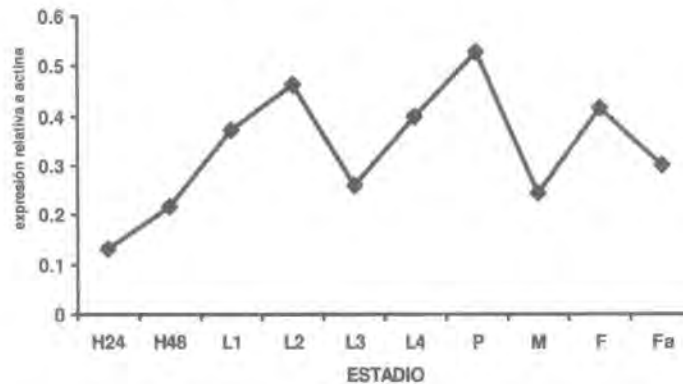


Figura 4.28. Expresión relativa a actina de AaST6 (*An.albimanus*)

Se observan dos etapas de mayor expresión de AaST6, esto es en larvas *instar* 2 y en pupas. Un análisis similar realizado en *D.melanogaster* indicó que la ST6Gal presenta picos de expresión en embriones de 24h y en pupas (Fabini *et al*, 2002). Mediante ensayos de hibridación *in situ* se determinó que la expresión durante las etapas embrionarias ocurre en las glándulas salivales y en el sistema nervioso (Fabini *et al*, 2002; Koles *et al*, 2004). Influyendo en el desarrollo de estos órganos y en la modificación postraduccional de proteínas residentes en estas regiones, tales como N-CAM (*neural cell adhesion molecule*) y los gangliósidos (Roth *et al*, 1992). La mitad de los genes de *Drosophila melanogaster* siguen un patrón de expresión con niveles máximos en las primeras etapas de desarrollo y después de la metamorfosis, demostrado mediante microarreglos (Arbeltman *et al*, 2002). Esto concuerda con los datos de expresión de AaST6 de *An.albimanus*.

El análisis aquí mostrado indica también la presencia de cantidades elevadas del ARNm AaST6 en adultos hembras. Por el momento no contamos con los análisis bioquímicos para confirmar que los niveles de expresión detectados a través de la ontogenia corresponden en la misma proporción de actividad Sialiltransferasa.

ESTA TESIS NO SALE
DE LA BIBLIOTECA

5. CONCLUSIONES

El trabajo de identificación, clonación y análisis ontogenético de expresión de la α 2,6-Sialiltransferasa de *Anopheles albimanus* llevado a cabo y reportado en esta tesis, deriva en las siguientes conclusiones:

- ▶ El mosquito *Anopheles albimanus* posee un gen ortólogo identificado como codificante para una α 2,6-Sialiltransferasa putativa, la traducción posee la topología característica de la familia de las sialiltransferasas incluyendo los *sialylmotifs* L, S y VS así como los residuos conservados implicados en su función catalítica.
- ▶ La identificación del transcrito correspondiente a la α 2,6-Sialiltransferasa de *Anopheles albimanus* (*AaST6*) fue facilitada al usar ADNc de tipo secuencia específico (SE), sintetizado a partir de oligonucleótidos degenerados.
- ▶ Al amplificar los extremos 5' y 3' se determinó que el transcrito *AaST6* tiene una longitud de 1746pb, codifica para una proteína de 440aa y posee identidad con las secuencias de ADNc de α 2,6-Sialiltransferasas en un rango de 49 a 64%.
- ▶ La región del dominio catalítico de la α 2,6-Sialiltransferasa de *Anopheles albimanus* posee identidad de 39% respecto a las proteínas ortólogas de vertebrados, 57% con las ortólogas de invertebrados y 87% respecto a *A.gambiae*.
- ▶ La secuencia de la proteína deducida de *AaST6* presenta dominios putativos funcionales y estructurales conservados que indican que pertenece a la Familia 29 de la Familia de las Glicosiltransferasas: Sialiltransferasas, y la organización característica de las proteínas de esta familia: un dominio citoplasmático N-terminal, la región transmembranal seguida del tallo hasta el dominio catalítico que contiene los *sialylmotifs* L, S y VS.
- ▶ La α 2,6-Sialiltransferasa de *Anopheles albimanus* posee sitios potenciales para modificaciones postraduccionales tales como N-glicosilación, fosforilación, miristoilación y amidación, sugiriendo la existencia de mecanismos reguladores de su acción enzimática.
- ▶ La predicción de estructura secundaria en la región del dominio catalítico de la α 2,6-Sialiltransferasa de *Anopheles albimanus* muestra una estructura de α -hélices y hojas- β intercaladas. Esta estructura se encuentra conservada en la región del dominio catalítico de las Sialiltransferasas conocidas.
- ▶ El análisis ontogenético de expresión del ARNm de la α 2,6-Sialiltransferasa de *A.albimanus* mostró dos picos máximos de expresión en las etapas de larvas *instar* 2 y pupas, se tiene un nivel basal

de expresión en todas las etapas de desarrollo analizadas, sugiriendo que la presencia de sialoglicoproteínas en *A.albimanus* no es un evento aislado.

► La α 2,6-Sialiltransferasa de *A.albimanus* es la segunda en su tipo en ser clonada de insectos, la primera del género Anopheles y la quinta identificada de insectos (incluyendo los géneros *Drosophila* y *Anopheles*).

Además de lo expuesto en los puntos anteriores, es de notar que a través de este trabajo tenemos ya evidencias del potencial para la obtención de glicanos sialilados en insectos, un hecho aún en duda hasta hace algún tiempo. Se ha utilizado a la mosca de la fruta (*Drosophila*) como modelo de estudio la N-glicosilación en insectos, es posible ahora plantear la posibilidad de ampliar este campo de estudio hacia otros insectos tal como

6. PERSPECTIVAS

La identificación, clonación y análisis ontogénico de la α 2,6-Sialiltransferasa de *Anopheles albimanus* permite abordar nuevas rutas y contestar algunas interrogantes que resultan evidentes.

- ▶ El análisis de expresión de *AaST6* se realizó semicuantitativamente, es necesario hacer cuantificación exacta de expresión mediante PCR cuantitativa. Se debe determinar si la expresión sufre algún cambio debido a la ingesta de sangre o por confrontación contra el *Plasmodium vivax*, recordemos que este mosquito es vector de la malaria y las sialoglicoproteínas en el tracto digestivo del hospedero podrían ser receptores. Si este fuera el caso, regular la expresión de *AaST6* podría constituirse en una alternativa para controlar la colonización del mosquito por el protozoario.
- ▶ Además de una regulación temporal se sabe que existe también una regulación espacial, por lo que el análisis de la expresión diferencial en órganos o tejidos del mosquito sería de interés. Especialmente en sistema nervioso, tracto digestivo y glándulas salivales, donde ha sido detectado ácido siálico y algunas sialoglicoproteínas.
- ▶ La caracterización enzimática y preferencia por sustratos deben ser realizados, demostrando así que es una proteína funcional. La producción recombinante mediante baculovirus en células de insecto puede ser la herramienta ideal para dicho análisis.
- ▶ El análisis del ADN genómico permitirá identificar la organización del gen (intrones/exones).
- ▶ Queda también por determinar si existe la presencia de diferentes isoformas de la proteína producidas por *splicing* alternativo del ARNm, tal como ocurre en *Drosophila* y algunos vertebrados.
- ▶ La utilización de un baculovirus recombinante expresando la α 2,6-Sialiltransferasa de *An. albimanus* es una posibilidad de aplicación en ingeniería de vías metabólicas al complementar o sobreexpresar esta enzima en las líneas celulares de insecto deficientes o carentes de la misma, promoviendo de esta forma la obtención de proteínas recombinantes con glicanos sialilados.
- ▶ Se ha planteado la posibilidad de extender la búsqueda de este tipo de genes a otros organismos que han probado tener el potencial de sialilación, tal como *Trichoplusia ni*, *Danaus plexippus* y *Manduca sexta*, aplicando parte de la metodología utilizada en este trabajo.
- ▶ La patente *TAMUS 2081: A Novel Drosophila Sialyltransferase with Unique Enzymatic Activity*. (Technology Licensing Office Industry's Connection to the A&M System), desarrollada por Panin y Koles en el año 2004, demostró su utilidad en la obtención de sialoglicoproteínas recombinantes. Una vez caracterizada la sialiltransferasa de *A. albimanus*, se podrá analizar la utilidad que pueda dar para la obtención de sialoglicoproteínas recombinantes producidas mediante el sistema células de insecto-baculovirus, de forma similar a lo reportado con la sialiltransferasa de *Drosophila*.

7. REFERENCIAS

1. Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. **Cell-specific expression of the human β -galactoside α 2,6-sialyltransferase transcripts differing in the 5' untranslated region.** *Eur J Biochem* 1993, Vol. 213, 467-475.
2. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YHC, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Miklos GLG, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadiou E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RDC, Scheeler F, Shen H, Shue BC, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM and Venter JC. **The Genome Sequence of *Drosophila melanogaster*.** *Science* 2000, Vol. 287, 2185-2195.
3. Andersen DC and Krummen L. **Recombinant protein expression for therapeutic applications.** *Curr Opin Biotechnol* 2002, Vol. 13, 117-123.
4. Angata T, Varki A. **Chemical diversity in the sialic acids and related α -keto acids: an evolutionary perspective.** *Chem Rev* 2002, Vol. 102, 439-469.
5. Arbeitman MN, Furlong EEM, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP. **Gene expression during the life cycle of *Drosophila melanogaster*.** *Science* 2002, Vol. 297, 2270-2275.
6. Aumiller JJ, Hollister JR and Jarvis DL. **A transgenic insect cell line engineered to produce CMP-sialic acid and sialylated glycoproteins.** *Glycobiology* 2003, Vol. 13, No. 6, 497-507.
7. Betenbaugh MJ, Tomiya N, Narang S, Hsu JTA and Lee YC. **Biosynthesis of human-type N-glycans in heterologous systems.** *Curr Opin Struct Biol* 2004, Vol. 14, 601-606.

8. Breen KC. **The role of protein glycosylation in the control of cellular N-sialyltransferase activity.** *FEBS Lett* 2002, Vol. 517, 215-218.
9. Breen KC, Georgopoulou N. **The role of protein phosphorylation in α 2,6(N)-sialyltransferase activity.** *Biochem Biophys Res Commun* 2003, Vol. 309, 32-35.
10. Breton C and Imberty A. **Structure/function studies of glycosyltransferases.** *Curr Opin Struct Biol* 1999, Vol. 9, 563-571.
11. Breton C, Mucha J and Jeanneau C. **Structural and functional features of glycosyltransferases.** *Biochimie* 2001, Vol. 83, 713-718.
12. Brooks SA. **Appropriate glycosylation of recombinant proteins for human use.** *Mol Biotechnol* 2004, Vol. 28, 241-255.
13. Butters TD. **Glycoproteins.** In: *Nature Encyclopedia of Life Sciences* 2001. London: Nature Publishing Group. <http://www.els.net/> [doi:10.1038/npg.els.0000705].
14. Chang GD, Chen CJ, Lin CY, Chen HC and Chen H. **Improvement of glycosylation in insect cells with mammalian glycosyltransferases.** *J Biotech* 2003, Vol. 102, 61-71.
15. Chen C, Colley KJ. **Minimal structural and glycosylation requirements for ST6Gal I activity and trafficking.** *Glycobiology* 2000, Vol. 10, No. 5, 531-538.
16. Chomczynski P and Sacchi N. **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** *Anal Biochem* 1987, Vol. 162, 156-159.
17. Datta AK, Chammas R and Paulson JC. **Conserved cysteines in the sialyltransferase form an essential disulfide bond.** *J Biol Chem* 2001, Vol. 276, No. 18, 15200-15207.
18. Datta AK and Paulson JC. **The sialyltransferase "sialylmotif" participates in binding the donor substrate CMP-Neu5Ac.** *J Biol Chem* 1995, Vol. 270, No. 4, 1497-1500.
19. Datta AK and Paulson JC. **Sialylmotifs of sialyltransferases.** *Indian J Biochem Biophys* 1997, Vol. 34, 157-165.
20. Datta AK, Sinha A and Paulson JC. **Mutation of the sialyltransferase S-sialylmotif alters the kinetics of the donor and acceptor substrates.** *J Biol Chem* 1998, Vol. 273, No. 16, 9608-9614.
21. Donaldson M, Wood HA, Kulakosky PC and Shuler ML. **Use of mannosamine for inducing the addition of outer arm N-acetylglucosamine onto N-linked oligosaccharides of recombinant proteins in insect cell.** *Biotechnol Prog* 1999, Vol. 15, 168-173.
22. Drickamer K. **A conserved disulphide bond in sialyltransferases.** *Glycobiology* 1993, Vol. 3, 2-3.
23. Dwek AD. **Glycobiology: More Functions for Oligosaccharides.** *Science* 1995, Vol. 269, 1234-1235.
24. Effertz K, Hinderlinch S, Reutter W. **Selective loss of either the epimerase or kinase activity of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase due to site-directed mutagenesis based on sequence alignments.** *J Biol Chem* 1999, Vol. 274, 28771-28778.
25. Fabini G and Wilson IBH. **Glycosyltransferases in *Drosophila melanogaster*.** *Dros Inf Serv* 2001, Vol. 84, 122-129.
26. Fabini G and Wilson IBH. **Stage and tissue specific expression of sialyltransferase-like gene (CG4871) in *Drosophila melanogaster*.** *Dros Inf Serv* 2002, Vol. 85, 45-49.

27. Fenteany FH and Colley KJ. **Multiple signals are required for α 2,6-Sialyltransferase (ST6Gal I) oligomerization and Golgi retention.** *J Biol Chem* 2005, Vol. 280, No. 7, 5423-5429.
28. Fuqua SAW, Fitzgerald SD, McGuire WL. **A simple polymerase chain reaction method for detection and cloning of low-abundance transcripts.** *BioTechniques* 1990, Vol. 9, No. 2, 206-211.
29. Geremia RA, Harduin-Lepers A and Delannoy P. **Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: implications for their mechanism of action.** *Glycobiology* 1997, Vol. 7, No. 2, v-xi.
30. Gillepsie W, Kelm S and Paulson JC. **Cloning and expression of the Gal β 1,3GalNAc: α 2,3-sialyltransferase.** *J Biol Chem* 1992, Vol. 267, 21004-21010.
31. Grundmann U, Nerlich C, Rein T, Zettlmesl G. **Complete cDNA sequence encoding human β -galactoside α 2,6-sialyltransferase.** *Nucleic Acids Res* 1990, Vol. 18, 667.
32. Hamamoto T, Kawasaki M, Kurosawa N, Nakaoka T, Lee YC, Tsuji S. **Two steps single primer mediated polymerase chain reaction; application to cloning of putative mouse β -galactoside α 2,6-sialyltransferase.** *Bioorg Med Chem Lett* 1993, Vol. 1, 141-145.
33. Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. **The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach.** *Glycobiology* 2005, doi:10.1093/glycob/cwi063.
34. Harduin-Lepers A, Recchi MA and Delannoy P. **1994, the year of sialyltransferases.** *Glycobiology* 1995, Vol. 5, No. 8, 741-758.
35. Helenius A and Aebi M. **Intracellular Functions of N-Linked Glycans.** *Science* 2001, Vol. 291, 2364-2369.
36. Hollister JR, Grabenhorst E, Nimtz M, Conradt HO and Jarvis DL. **Engineering the protein N-glycosylation pathway in insect cells for production of biantennary, complex N-glycans.** *Biochemistry* 2002, Vol. 41, 15093-15104.
37. Hollister JR and Jarvis DL. **Engineering lepidopteran insect cells for sialoglycoprotein production b genetic transformation with mammalian β 1,4-galactosyltransferase and α 2,6-sialyltransferase genes.** *Glycobiology* 2001, Vol. 11, No. 1, 1-9.
38. Hollister JR, Shaper JH and Jarvis DL. **Stable expression of mammalian β 1,4-galactosyltransferase extends the N-glycosylation pathway in insect cells.** *Glycobiology* 1998, Vol. 8, 473-480.
39. Hollister JR and Jarvis DL. **Evidence for a sialic acid salvaging pathway in lepidopteran insect cells.** *Glycobiology* 2003, Vol. 13, No. 6, 487-495.
40. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nussekern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, Cai S, Center A, Chaturverdi K, Christophides GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanigan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic

- I, Levitsky A, Liang Y, Lin JJ, Lobo NF, Lopez JR, Malek JA, McIntosh TC, Meister S, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, Zhao S, Zhu SC, Zhimulev I, Coluzzi M, della Torre A, Roth CW, Louis C, Kalush F, Mural RJ, Myers EW, Adams MD, Smith HO, Broder S, Gardner MJ, Fraser CM, Birney E, Bork P, Brey PT, Venter JC, Weissenbach J, Kafatos FC, Collins FH and Hoffman SL. **The genome sequence of the malaria mosquito *Anopheles gambiae*.** *Science* 2002, Vol. 298, pp. 129-149.
41. Hooker AD, Green NH, Baines AJ, Bull AT, Jenkins N, Strange PG and James DC. **Constraints on the transport and glycosylation of recombinant IFN- γ in Chinese Hamster Ovary and Insect Cells.** *Biotechnol Bioeng* 1999, Vol. 63, 473-480.
42. Jarvis DL, Kawar ZS and Hollister JR. **Engineering N-glycosylation pathways in the baculovirus-insect cell system.** *Curr Opin Biotech* 1998, Vol. 9, 528-533.
43. Jeanneau C, Chazalet V, Augé C, Soumpasis DM, Harduin-Lepers A, Delannoy P, Imberty A and Breton C. **Structure-Function analysis of the human sialyltransferase ST3Gal I. Role of N-glycosylation and a novel conserved sialylmotif.** *J Biol Chem* 2004, Vol. 279, No. 14, 13461-13468.
44. Jenkins N, Curling EMA. **Glycosylation of recombinant proteins: Problems and prospects.** *Enzyme Microb Technol* 1994, Vol. 16, 354-364.
45. Jenkins N, Parekh RB and James DC. **Getting the glycosylation right: Implications for the biotechnology industry.** *Nat Biotechnol* 1996, Vol. 14, 975-981.
46. Joosten CE and Shuler ML. **Production of a sialylated N-linked glycoprotein in insect cells: Role of glycosidases and effect of harvest time on glycosylation.** *Biotechnol Prog* 2003a, Vol. 19, 193-201.
47. Joosten CE and Shuler ML. **Effect of culture conditions on the degree of sialylation of a recombinant glycoprotein expressed in insect cells.** *Biotechnol Prog* 2003b, Vol. 19, 739-749.
48. Katsutoshi S. **Molecular cloning and characterization of Sialyltransferases.** *Trends Glycosci Glycotechnol* 1996, Vol. 8, 195-215.
49. Kim K, Lawrence SM, Park J, Pitts L, Vann WE, Betenbaugh MJ, Palter KB. **Expression of a functional *Drosophila melanogaster* N-acetylneuraminic acid (Neu5Ac) phosphate synthase gene: evidence for endogenous sialic acid biosynthetic ability in insects.** *Glycobiology* 2002, Vol. 12, No. 2, 73-83.
50. Kitagawa H, Paulson JC. **Cloning of a novel α 2,3-sialyltransferase that sialylates glycoproteins and glycolipids carbohydrate group.** *J Biol Chem* 1994a, Vol. 269, 1394-1401.
51. Kitagawa H, Paulson JC. **Differential expression of five sialyltransferases genes in human tissues.** *J Biol Chem* 1994b, Vol. 269, 17872-17879.
52. Koles K, Irvine KD and Panin VM. **Functional characterization of *Drosophila* Sialyltransferase.** *J Biol Chem* 2004, Vol. 279, No. 6, 4346-4357.

53. Kraft R, Wittmann-Liebold B. **Proteins: postsynthetic modifications.** In: *Nature Encyclopedia of Life Sciences 2001.* John Wiley & Sons, Ltd: Chichester. <http://www.els.net/> [doi:10.1038/npg.els.0000551].
54. Kurosawa N, Kawasaki M, Hamamoto T, Nakaoka T, Lee YC, Arita M, Tsuji S. **Molecular cloning and expression of chick embryo Gal β 1,4GlcNAc α 2,6-sialyltransferase.** *Eur J Biochem* 1994, Vol. 219, 375-381.
55. Kyte J, Doolittle R. **A simple method for displaying the hydrophatic character of a protein.** *J Mol Biol* 1982, Vol. 157, 105-132.
56. Lance P, Lau KM, Lau JTY. **Isolation and characterization of a partial cDNA for a human sialyltransferase.** *Biochem Biophys Res Commun* 1989, Vol. 164, 225-232.
57. Lawrence SM, Huddleston KA, Tomiya N, Nguyen N, Lee YC, Vann WF, Coleman TA and Betenbaugh MJ. **Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells.** *Glycoconjugate J* 2001, Vol. 18, 205-213.
58. Lee YC, Kojima N, Wada E, Kurosawa T, Nakaoka T, Hamamoto T, Tsuji S. **Cloning and expression of cDNA for a new type of Gal β 1-3GalNAc α 2,3-sialyltransferase.** *J Biol Chem* 1994, Vol. 269, 10028-10033.
59. Licari PJ, Jarvis DL and Bailey JE. **Insect cell hosts for baculovirus expression vectors contain endogenous exoglycosidase activity.** *Biotechnol Prog* 1993, Vol. 9, 146-152.
60. Livingston BD and Paulson JC. **Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family.** *J Biol Chem* 1993, Vol. 268, 11504-11507.
61. Lopez M, Tataert D, Juliant S, Gazon M, Cerutti M, Verbert A and Delannoy P. **O-Glycosylation potential of lepidopteran insect cell lines.** *Biochim Biophys Acta* 1999, Vol. 1427, 49-61.
62. Ma J, Simonovic M, Qian R, Colley KJ. **Sialyltransferase isoforms are phosphorylated in the cis-medial Golgi on serine and threonine residues in their luminal sequences.** *J Biol Chem* 1999, Vol. 274, 8046-8052.
63. Marchal I, Jarvis DL, Cacan R and Verbert A. **Glycoproteins from Insect Cells: Sialylated or Not?.** *Biol Chem* 2001, Vol. 382, 151-159.
64. Marchler-Bauer A, Bryant SH. **CD-Search: protein domain annotations on the fly.** *Nucleic Acids Res* 2004, Vol. 32, 327-331.
65. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH. **CDD: a Conserved Domain Database for protein classification.** *Nucleic Acids Res* 2005, Vol. 33, D192-196.
66. Martin BL. **Regulation by covalent modifications.** In: *Nature Encyclopedia of Life Sciences 2001.* John Wiley & Sons, Ltd: Chichester. <http://www.els.net/> [doi:10.1038/npg.els.0000866].
67. März L, Altmann F, Staudacher E and Kubelka V. **Protein glycosylation in insects.** In *Glycoproteins (Montreuil J, Schachter H, Vliegenthart H, Eds.)* 1995, Vol. 3, Chapter 10, 543-563, Elsevier, B.V.: Amsterdam, The Netherlands.

68. Medvedova L, Knopp J and Farkas R. **Steroids regulation of terminal protein glycosyltransferase genes: Molecular and functional homologies within Sialyltransferase and Fucosyltransferase Families.** *Endocrine Regulations* 2003, Vol. 37, 203-210.
69. Mercier D, Wierinchx A, Oulmouden A, Gallet PF, Palcic MM, Harduin-Lepers A, Delannoy P, Petit JM, Levéziel H, Julien R. **Molecular cloning, expression and exon/intron organization of the bovine β -galactoside α 2,6-sialyltransferase gene.** *Glycobiology* 1999, Vol. 9, No. 9, 851-863.
70. Nok AJ, Njoku GC, Balogun E. **A 45-kDa midgut glycoprotein from *Anopheles albimanus* mosquito mediates the killing of trypanosomes.** *Cell Biochem Funct* 2002; 20:257-262.
71. O'Hanlon TP, Lau KM, Wang X and Lau JTY. **Tissue-specific expresión of β -galactoside α 2,6-sialyltransferase. Transcript heterogeneity predicts a divergent polypeptide.** *J Biol Chem* 1989, Vol. 264, 17389-17394.
72. Palomares LA, Joosten CE, Hugues PR, Granados RR and Shuler ML. **Novel insect cell line capable of complex N-glycosylation and sialylation of recombinant proteins.** *Biotechnol Prog* 2003, Vol. 19, 185-192.
73. Palomares LA and Ramírez OT. **Complex N-glycosylation of recombinant proteins by insect cells.** *BioProcessing* 2002, Fall, 70-73.
74. Paulson JC and Colley KJ. **Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation.** *J Biol Chem* 1989, Vol. 264, No. 30, 17615-17618.
75. Rademacher TW. **Therapeutic challenges - does glycobiology have a role?.** *Trends Biotechnol* 1992, Vol. 10, 227-230.
76. Rademacher TW, Parekh RB and Dwek RA. **Glycobiology.** *Ann Rev Biochem* 1988, Vol. 57, 785-838.
77. Reithmeier RAF. **Membrane proteins.** In: *Nature Encyclopedia of Life Sciences* 2001. John Wiley & Sons, Ltd: Chichester. <http://www.els.net/> [doi:10.1038/npg.els.0000624].
78. Roth J, Kempf A, Reuter G, Schauer R and Gehring WJ. **Occurrence of sialic acids in *Drosophila melanogaster*.** *Science* 1992, Vol. 256, 673-675.
79. Sasaki K, Watanabe E, Kawashima K, Sekine S, Dohi T, Oshima H, Hanai N, Nishi T, Hasegawa M. **Expression cloning of a novel Gal β (1-3/1-4)GlcNAc α 2,3-sialyltransferase using lectin resistance selection.** *J Biol Chem* 1993, Vol. 268, 22782-22787.
80. Seppo A, Tiemeyer M. **Function and structure of *Drosophila* glycans.** *Glycobiology* 2000, Vol. 10, 751-760.
81. Spiro RG. **Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.** *Glycobiology* 2002, Vol. 12, No. 4, 43R-56R.
82. Tomiya N, Betenbaugh MJ and Lee YC. **Humanization of lepidopteran insect-cell-produced glycoproteins.** *Acc Chem Res* 2003a, Vol. 36, 613-620.
83. Tomiya N, Howe D, Aumiller JJ, Pathak M, Park J, Palter KB, Jarvis DL, Betenbaugh MJ and Lee YC. **Complex-type biantennary N-glycans of recombinant human transferring from *Trichoplusia ni* insect cells expressing mammalian β 1,4-galactosyltransferase and β 1,2-acetylglucosaminyltransferase II.** *Glycobiology* 2003b, Vol. 13, No. 1, 23-34.

84. Tomiya N, Narang S, Lee YC, Betenbaugh MJ. **Comparing N-glycan processing in mammalian cell lines to native and engineered lepidopteran insect cell lines.** *Glycoconjugate J* 2004, Vol. 21, 343-360.
85. Traving C and Schauer R. **Structure, function and metabolism of sialic acids.** *CMLS Cell Mol Life Sci* 1998, Vol. 54, 1330-1349.
86. Tsuji S. **Structural and functional analysis of sialyltransferases.** *RIKEN Review* 1995, No. 8, Focused on Glycobiology and Glycotechnology, 5-6.
87. Tsuji S. **Molecular cloning and functional analysis of sialyltransferases.** *J Biochem* 1996, Vol. 120, No. 1, 1-13.
88. Vadaie N and Jarvis DL. **Molecular cloning and functional characterization of a lepidopteran insect β 4-N-Acetylgalactosaminyltransferase with broad substrate specificity, a functional role in glycoprotein biosynthesis, and a potential role in glycolipid biosynthesis.** *J Biol Chem* 2004, Vol. 279, No. 32, 33501-33518.
89. Vallejo-Ruiz V, Haque R, Mir AM, Schwientek T, Mandel U, Cacan R, Delannoy P, Harduin-Lepers A. **Delineation of the minimal catalytic domain of human Gal β 1-3GalNAc α 2,3-sialyltransferase (hST3Gal I).** *Biochim Biophys Acta* 2001, Vol. 1549, 161-173.
90. Wang XC, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JTY. **Chromosome mapping and organization of the human β -galactoside α 2,6-sialyltransferase gene.** *J Biol Chem* 1993, Vol. 268, 4355-4361.
91. Watanabe S, Kokuho T, Takahashi H, Takahashi M, Kubota T and Inumaru S. **Sialylation of N-Glycans on the recombinants proteins expressed by a baculovirus-insect cell system under β -N-Acetylglucosaminidase inhibition.** *J Biol Chem* 2002, Vol. 277, No. 7, 5090-5093.
92. Weinstein J, de Souza SU, Paulson JC. **Purification of a Gal β 1-4GlcNAc α 2,6-sialyltransferase and a Gal β 1-3(4)GlcNAc α 2,3-sialyltransferase to homogeneity from rat liver.** *J Biol Chem* 1982, Vol. 257, 13835-13844.
93. Weinstein J, Lee EU, McEntee K, Lai P and Paulson JC. **Primary structure of β -galactoside α 2,6-sialyltransferase.** *J Biol Chem* 1987, Vol. 262, 17735-17743.
94. Wen DX, Livingston BD, Medzihradzky KK, Kelm S, Burlingame AL and Paulson JC. **Primary structure of Gal β 1,3(4)GlcNAc: α 2,3-sialyltransferase determined by mass spectrometry sequence analysis and molecular cloning.** *J Biol Chem* 1992a, Vol. 267, 21011-21019.
95. Yang X, Marchand JE. **A sensitive technique to clone low abundance receptor transcripts from single microdissected tissue punches.** *Brain Res Brain Res Protoc* 2002, Vol. 9, No. 2, 135-46.
96. Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, Copley RR, Christophides GK, Thomasova D, Holt RA, Subramanian GM, Mueller HM, Dimopoulos G, Law JH, Wells MA, Birney E, Charlab R, Halpern AL, Kokoza E, Kraft CL, Lai Z, Lewis S, Louis C, Barillas-Mury C, Nusskern D, Rubin GM, Salzberg SL, Sutton GG, Topalis P, Wides R, Wincker P, Yandell M, Collins FH, Ribeiro J, Gelbart WM, Kafatos FC, Bork P. **Comparative Genome and Proteome Analysis of *Anopheles gambiae* and *Drosophila melanogaster*.** *Science* 2002 298: 149-159.

Anexos

1. Bases de datos y herramientas bioinformáticas utilizadas

Descripción	Dirección electrónica
Anopheles Database	http://skonops.imbb.forth.gr/AnoBase/
Biopolymer calculador versión 4.1.1.	http://paris.chem.yale.edu/extinct.frames.html
BLAST (Basic Local Alignment Search Tool)	http://www.ncbi.nlm.nih.gov/BLAST/
Codon Usage Database	http://www.kazusa.or.jp/codon/
NCBI Conserved Domain Search A Conserved Domain Database and Search Service, v2.03	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
Conserved Domain Database	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cdd
DrawHCA	http://bioserv.rpbs.jussieu.fr/RPBS/cgi-bin/Ressource.cgi?chzn_lg=fr&chzn_rsrc=HCA
Ensembl (Project Ensembl)	http://www.ensembl.org/
GenBank (National Center for Biotechnology Information)	http://www.ncbi.nlm.nih.gov/
GeneBee Molecular Biology Server	http://www.genebee.msu.su/
LALIGN (LALIGN-find multiple matching subsegments in two sequences)	http://www.ch.embnet.org/software/LALIGN_form.html
MultAlin (Multiple sequence alignment with hierarchical clustering)	http://prodes.toulouse.inra.fr/multalin/multalin.html
OligoAnalyzer 3.0 IDT SciTools	http://207.32.43.70/biotools/oligocalc/oligocalc.asp
Oligonucleotide Properties Calculator ver 3.07	http://www.basic.nwu.edu/biotools/oligocalc.html
ORFinder (Open Reading Frame Finder)	http://www.ncbi.nlm.nih.gov/gorf/gorf.html
PROSITE Database of protein families and domains	http://us.expasy.org/prosite/
PSIPRED (The PSIPRED Protein Structure Prediction Server)	http://bioinf.cs.ucl.ac.uk/psipred/psiform.html
THE MEME/MAST SYSTEM (Motif Discovery and Search, Version 3.0)	http://meme.sdsc.edu/meme/website/intro.html

TIGR (The Institute for Genomic Research)	http://www.tigr.org/
TMPred Prediction of transmembrane regions and orientation	http://www.ch.embnnet.org/software/TMPRED_form.html

2. Preparación de células químicamente competentes

COMPONENTES UTILIZADOS

1. **Placas de medio LB sin antibiótico.**
2. **Medio LB líquido.**
3. **DMSO. Dimetilsulfoxido.**
4. **Crioviales de 2 ml.**
5. **Nitrógeno líquido.**
6. **Buffer TB.** 10mM PIPES, 55 mM MnCl₂, 15 mM CaCl₂·2H₂O, 250mM KCl. Disolver en agua, ajustar el pH a 6.7 con 5 N KOH antes de adicionar el MnCl₂, ajustar al volumen deseado. Esterilizar con membrana de 0.22µM y almacenar a 4°C.

Las células *E.coli* TOP-10 químicamente competentes, se prepararon de acuerdo al siguiente protocolo:

1. Sembrar células TOP-10 en una caja de medio LB sin antibiótico. Incubar a 37°C por doce horas o hasta obtener colonias pequeñas separadas.
2. Colocar 25-50 ml de medio LB líquido en un matraz de 250 ml. Inocular con una colonia pequeña y poner en cultivo agitado a 300 rpm a 22°C (procurando tener buena oxigenación del cultivo).
3. Monitorear el cultivo hasta obtener una OD₆₀₀ de 0.4-0.6.
4. Repartir el cultivo colocando 2 ml por tubo.
5. Incubar en hielo por 10 min.
6. Centrifugar a 3000 rpm por 10 min a 4°C (si se centrifuga el cultivo completo en un tubo de 50 ml debe hacerse a 5000 rpm por 10 min a 4°C).
7. Retirar el sobrenadante y resuspender la pastilla, adicionar 600 µl de buffer TB frío por cada 2ml de cultivo.
8. Incubar por 20 min en hielo.
9. Centrifugar por 10 min a 3000 rpm si son tubos eppendorf de 1.5 ml (o 3000rpm por 10-15 min si es un tubo de 50 ml) a 4°C.
10. Retirar el sobrenadante y resuspender la pastilla, adicionando 150 µl de buffer TB frío por cada 2 ml de cultivo original.
11. Adicionar 10 µl de DMSO por cada 2 ml de cultivo, y colocar en hielo (si se trabajo el cultivo completo, alicuotar en tubos eppendorf de 1.5 ml colocando 150 µl de la mezcla por tubo).
12. Incubar por 10 min en hielo.
13. Congelar rápidamente al pasar los tubos a nitrógeno líquido.
14. Almacenar a -70°C.

3. Localización de oligonucleótidos utilizados

1 cgcgggggtgtggctgtggcagcgcaaaagtgtgctgtgctgctgcaaccgctccagcgatgCGGTGACACGGGATAGCATACCTT
 1 M R V H G I A Y L
 STAa1 GCTGCAACCGCTCCAGCGATG

91 CAGGGGCATTTCCTAGAGGAGGAGAAACCGACGGCCACAACCTACCCTCGTGCCATCGGCGTCTGCCTCTGGCTGTATTGTTCCGG
 10 Q G H F P R R Q K P T A T N R T V V P S A S C L W L L L F A

181 GCGATCTTTGTTGCTGGAGGAGTAGCAGCTGATGCGCTTGGCAAGCGGAACGTATTCTACGTCAAGCTCAACACCACGGAGAAAACGTAC
 40 A I F V A G G V A A D A L G K R N V F Y V K L N T T E K T Y

271 AAGCGAGCGATGAGAAACCGACGCTGTGGTCGAGGAAACAGGACGCCCGAATACGGCCGCACTTCAGTCTGGCCAAAACGTCGCAC
 70 K A S D E K P T L W S K E Q D A P R I R P H F S L A K T S H

361 TTTCTGTTGACCTGAACCGATACCGGTGCGAGGTGGCGAGCGCGCAATGGACAGCGCGCTCCCGCTGGTTAGTGCAGCTGCGCC
 100 F L F D L N R Y R C E V A S A G N G T A A V P L V S A D C R

451 AACCAGCCGCACTGTTCCGGGCGAAATCATCGCGCAACTGAGCCGACGGTTAAAGGAACCGGTCCCGGCGAGCGGACAGTATCAGTGT
 130 N R T A L F R A K I I G E L S R R L K E P V P A S G Q Y Q V

541 CGGTATAGCCGGCCGTTCCAGCAGCGACCCCTCGGTACGGGTGAGTCCGGTTGCGGCGAGCGAATGTGCGGTCTGTCTGGAAGGAG
 160 R Y S R P F Q Q R P L G T G Q C R L R Q A N V R V L S W K E

631 CCACCGTTTGGTGAACGAGCTCGGCACCTTACCTACCGGTGACGCCATTGTTTCGATCCGGATACGGCTACCAACAGGAGCTGCGTGATA
 190 P F F A W N E L G T Y L P V T P L F D P D T A T N R S C V I

721 GTGCCAGCGCCGGTTGCTGAAACGCTCCCGGCTTGGGCGGTTTCATCGATGGTCACGATATTGTGATGCGGTTCAACCATGCACCAAC
 220 V A S A G S L K R S R L G P F I D G H D I V M R F N H A P T
 GCACGATGTGATGCTCAATCA Ag1
 ATCGAGGACAGATCGTATG AaF1
 STGSP2 TGATCGGTTCAATCATGCACCAAC

811 GAGGGCTTCGAGCGGACGCTCGGAGCGAAGACGACGATCCGCGTACTCAACTCGCAGGTCGTCACGAAGCCTGAGTACCGGTTGCTGAGC
 250 E G F E A D V G A K T T I R V L N S Q V V T K P E Y R L L T

901 GCACCACTGTTCCGGAACTTACGATAGCGGCTGGGATCCGGGCAAGTTCGATCAACCGCTGTCGAGTGGTTGGCCAACTCCGACTTT
 280 A P L F R N V T I A A W D P G K F D Q P L S E W L A N P D F
 TTCCGGAAGTACGATCCCGCTGGG AgLS
 GTSACGATCCGCGCTGGGACC AaF2
 TTCCGGAACGTTACGATAGC AaST2

991 AATCTGTTGAGAACTTCAAGAAGTTTCGACGCCAGTACCGGGATGCCAACTTCCACATCATGATCCGCGCAGCATTCGGCGGGCGTGG
 310 N L F E N F X K F R S Q Y R D A N F H I I D P R S I W R A W
 TACCGGATGCCAACTTCCACATCA STNGSP2

1081 ACGCCCTTCAGGATCTCACTCCCTTTCGATCCGACAGAACCCTCCAACAACCGGTTTCAATCGGCTTGGACTGCTGCTGCCGCTGTGC
 340 T A L Q D L T P F P I R Q N P P T S G F I G L G L L L P V C
 AAACCGGATGTTGGAGGTTCTGTG STNGSP1

1171 CGGTACATTGACATCGTGGAGTACATTCGAGCACTCGCATGAACGGTCTCTGTCACTATTACGATGAAGAGCTCAATCTGGCTGTACG
 370 R Y I D I V E Y I P S T R M N G L C H Y Y D E E L N L G C T
 CATGCGAGTCTCGAATGTACTCCA STGSP1 AGCTCATCTCGTAGTATGCA AaR1
 GCGAGTCTCGAATGTACT AaST3

1261 TTCGGCGGTGGCACCCGCTGGCGGCGGAGAAGCTGCTGCTTCCAGATGAATGCCGCCAGCGATTACACCGCTTTTCAGCAGGGCACC
 400 F G A W H P L A A E K L L V F Q M N A A S D Y T V F Q Q G T
 CAGCTTCTCCGCGCTAGTGGATGCC AgVS
 TTCTCCGCGGAGGGTGC AaR2

1351 GTCCGCATCTACAGTATACGCTGGATCAATGCTAGTtctgctcggccgctcggctccattccgtaaacacttcgggcacctgtcaaa
 430 V R I Y S D T L D Q C *
 STAa2 ACGCCGGACGAATAGCATTG

1441 cctgtgacgggggcatcggcggctgagctatggcaggtggtggtgcaaaacaattagcaaacagaaattaattaagcaatgogaatc
 1531 gacgacagctaaggctgctgctgatcgatgcccgtcagcgtattagacattccagccgggctotgtgccgctttogatgacaatcag
 1621 acgacgggatcagaccgagaccaaacggcaaccagccagcaggtgcagtggtggtgctttgctttgaagctagtcaaaagcagatc
 1711 ccgggacggagtgcacaatcaaaazaaazaaazaaaz

4. HCA (Hydrophobic Cluster Analysis)

Región del dominio catalítico de *Anopheles albimanus*

