



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Carrera de Licenciatura en Ciencias Biológicas

Estudio estructural y funcional de PIGA, la proteína deficiente en la Hemoglobinuria Paroxística Nocturna

Tesinista: Samantha Diamela D'Amico

Director: Dr. Martín Ezequiel Noguera

Co-directora: Dra. Cecilia D'Alessio

Trabajo realizado en el iB3 (Instituto de Biociencias, Biotecnología y Biología translacional), Pabellón II,
Facultad de Ciencias Exactas y Naturales

Buenos Aires, Julio 2023

Martín Ezequiel Noguera

Dra Cecilia D'Alessio

Samantha D'Amico

INDICE

RESUMEN	5
Estudio estructural y funcional de PIGA, la proteína deficiente en la Hemoglobinuria Paroxística Nocturna	6
AGRADECIMIENTOS	7
ABREVIATURAS	9
INTRODUCCIÓN	11
1. Anclas de glicosilfosfatidilinositol (GPI)	12
2. Ruta de síntesis de proteínas ancladas a GPI.....	14
3. PIGA y el complejo GPI-GnT.....	16
4. Hemoglobinuria Paroxística Nocturna (HPN).....	20
OBJETIVOS E HIPÓTESIS	23
Objetivo general.....	24
Objetivos particulares.....	24
Hipótesis	24
MATERIALES Y MÉTODOS	25
3.1. Métodos de laboratorio.....	26
Cepas y medios de cultivo	26
Preparación de bacterias electrocompetentes	27
Electroporación de bacterias competentes	27
Transformación química de bacterias competentes	27
Preparación de levaduras <i>S. cerevisiae</i> competentes y transformación.....	28
Ensayo de complementación en cepa <i>S. cerevisiae</i> termosensibles.....	28
3.1.1. Procedimientos realizados con ADN	31
3.1.2. Procedimientos realizados con proteínas.....	34
3.2. Métodos bioinformáticos.....	37
3.2.1. Búsqueda de ortólogos de las subunidades del complejo GnT en organismos modelo.....	37

3.2.2. Generación de alineamientos múltiples de secuencias y análisis de distribución taxonómica	38
3.2.3. Construcción del árbol taxonómico	38
3.2.4 Obtención de modelos estructurales y modelado del complejo GPI-GnT	38
RESULTADOS	40
4.1. Primera parte: Clonado de PIGA y estudio de la expresión de su dominio catalítico en bacterias	41
4.1.1. Fundamento del diseño de la plataforma para la expresión de hPIGA en procariontes y eucariotas	41
4.1.2. Clonado de hPIGA en un vector de expresión de bacterias con HisTag y sitio TEV.....	43
4.1.3. Clonado del dominio citoplasmático de hPIGA para su expresión en bacterias	44
4.1.4. Expresión de PIGAc en bacterias	45
4.2. Segunda Parte: Estudio funcional en levaduras <i>S. cerevisiae</i>	54
4.2.1 Ensayos de complementación con hPIGA en <i>S. cerevisiae</i> termosensibles.....	56
4.2.2 Estudio de la expresión de hPIGA en <i>S. cerevisiae</i> termosensibles.....	57
4.3. Tercera parte: estudio de la estructura y función de PIGA y el complejo GPI-GnT por métodos bioinformáticos	61
4.3.1. Características de las secuencias de las distintas subunidades del complejo GPI-GnT.....	61
4.3.2. Conservación del complejo GPI-GnT en eucariotas.....	62
4.3.3. Estructura del complejo GPI-GnT	67
4.3.4. Mutaciones en las subunidades del complejo GPI-GnT y sus implicancias clínicas	74
4.3.5. Modelado del complejo GPI-GnT	79
4.3.6. Modelado de la unión del ligando al sitio activo de PIGA	80
DISCUSIÓN	83
5.1. Discusión.....	84
5.1.2. Expresión recombinante de PIGA: un target difícil para la expresión en <i>E. coli</i>	84
5.1.3. Avances hacia el desarrollo de un sistema para el estudio de variantes de PIGA en levaduras	85

5.1.4. La composición del complejo GPI-GnT está muy conservada en mamíferos, pero es más divergente en otros eucariotas.....	87
5.1.5. Los modelos estructurales de las subunidades del complejo GPI-GnT permiten entender la base estructural de algunas mutaciones clínicas	88
5.1.6. El modelado con glicosiltransferasas resueltas con ligandos unidos y similares a PIGA permitió conocer cómo ésta interactúa con sus sustratos.	88
CONCLUSIONES	90
5.2. Conclusiones.....	91
REFERENCIAS BIBLIOGRÁFICAS	92
Referencias bibliográficas	93
ANEXO	104
Análisis de las estructuras de las subunidades del complejo GPI-GnT: imágenes suplementarias	105
Predicción de la estructura secundaria de PIGA realizada con NetSurfP3.0	112
Enfermedades/síndromes asociados a las subunidades del complejo GPI-GnT	113
Gráficos de entropía en función de la posición de las mutaciones en la secuencia	115

RESUMEN

Estudio estructural y funcional de PIGA, la proteína deficiente en la Hemoglobinuria Paroxística Nocturna

Las moléculas de glicosilfosfatidilinositol (GPI) permiten el anclaje de proteínas sin dominio transmembrana a la membrana celular. Se encuentran en organismos eucariotas y son, en muchos casos, esenciales para la viabilidad. Existen numerosas proteínas en la superficie celular que se unen mediante estas anclas y cumplen diversos roles, tales como la transducción de señales o en la respuesta inmune. El anclaje de proteínas a la superficie celular mediante GPI involucra 12 pasos secuenciales, desde la síntesis del ancla hasta su unión a la proteína. El primero de ellos es la formación de *N*-acetilglucosamina-fosfatidilinositol, en una reacción catalizada por la enzima PIGA, una glicosil transferasa que forma parte de un complejo multiproteico anclado a la membrana del retículo endoplasmático.

Actualmente, se conocen diferentes enfermedades asociadas a la pérdida de función de genes involucrados en el armado de las anclas de GPI. Entre ellas se encuentra la Hemoglobinuria Paroxística Nocturna (HPN), una enfermedad hematológica poco frecuente en la que se produce la lisis de las células sanguíneas. Esto se produce por el ataque del sistema inmune de complemento a dichas células debido a la falta de una inhibición que en células normales se produce por proteínas ancladas a la superficie por GPI y que están ausentes en la HPN.

El objetivo de este trabajo fue obtener información a través de estudios *in silico*, *in vitro* e *in vivo* sobre la estructura de PIGA y su mecanismo de acción, de modo de comprender mejor su función a nivel molecular. Mediante métodos bioinformáticos se realizó un amplio estudio de la estructura y función de PIGA y del complejo del que forma parte. Se encontró que PIGA está presente en casi la totalidad de un grupo diverso de organismos eucariotas especialmente seleccionados por su diversidad taxonómica, mientras que las otras subunidades están menos conservadas. Se identificaron características importantes de PIGA para el reconocimiento de sus ligandos y se usaron modelos estructurales para racionalizar el impacto de algunas mutantes clínicas. Por otro lado, se intentó producir de manera recombinante el dominio catalítico de PIGA en bacterias y estudiar la función de PIGA usando un ensayo de complementación en levaduras *Saccharomyces cerevisiae* mutantes condicionales para el gen homólogo de PIGA. No se obtuvieron resultados concluyentes en este último punto, pero se generó un sistema que a futuro permitirá evaluar la capacidad de complementar funcionalmente a esta cepa con un defecto de crecimiento, realizando el ensayo con distintas mutantes de PIGA halladas en la clínica. Esto representó un avance para empezar a llenar el vacío de conocimiento sobre el funcionamiento de esta enzima clave a nivel estructural, algo necesario para el desarrollo de nuevas terapias basadas en la estructura para enfermedades como la HPN.

AGRADECIMIENTOS

Me gustaría comenzar agradeciendo a mis directores Martín y Cecilia, esta tesis no hubiera sido posible sin ustedes. Gracias por ayudarme a dar mis primeros pasos en la investigación. A Martín quiero agradecerle por su paciencia infinita, por guiarme y por estar siempre al pie del cañón. A Cecilia le agradezco por su dedicación, siempre recibíendome con una sonrisa y dispuesta a enseñarme una nueva técnica o experimento cada vez que lo requería.

A todo mi querido grupo de compañeros del IB3, hermoso grupo humano, siempre me trataron con una gran calidez, incluso los que menos conocía. Gracias por enseñarme y por permitirme aprender. En especial quisiera agradecer a Flor, Hernán y Javi que siempre me escucharon, me aconsejaron y me dieron una mano cuando lo necesitaba.

A mi amiga del alma Juli, que cada vez que le escribía temerosa contándole que tenía que rendir un examen siempre me alentaba y confiaba en mí. Aunque ambas nos dedicamos a dos ciencias exactas totalmente distintas, de alguna manera trataba de ayudarme cuando te lo pedía. Hermosos recuerdos me llevo de haber cursado parte del CBC con ella, de juntarnos para estudiar física o análisis matemático. A Román que siempre escuchó interesado los datos curiosos sobre biología que tiraba en las reuniones. A Fran y Martis que desde la distancia siempre estuvieron para sacarme una sonrisa y para apoyarme en todo. A Miranda o más bien Miri (como yo le suelo decir), uno de los regalos más lindos que me dejó esta carrera fue su amistad, gracias por tantos lindos recuerdos y momentos vividos, tanto afuera como adentro de la facu. Siempre estuvo cuando la necesitaba. Creo que la mitad de mi título se lo tengo que dar a ella, por todas las veces que me ayudó/salvó.

A toda mi familia. A mi hermano, que aunque ahora está lejos siempre creyó en mí. A mis padres... qué decirles...para expresar todo lo agradecida que estoy necesitaría utilizar más hojas que las que utilicé para escribir esta tesis. A pesar de los momentos difíciles que pasamos siempre me sostuvieron y me levantaron cuando me derrumbaba. Sin ellos no podría haber terminado esta carrera, nada de esto hubiera sido posible, gracias por bancarme tanto. Siempre fueron ese aliento que me faltaba para poder seguir.

¡¡¡GRACIAS TOTALES!!!

(Caer está permitido, levantarse es una obligación...)

ABREVIATURAS

RE: Retículo Endoplasmático

HPN: Hemoglobinuria Paroxística Nocturna

PI: Fosfatidil-inositol

GPI: Glicosilfosfatidilinositol

GlcNAc: *N*-Acetilglucosamina

UDP-GlcNAc: Uridina difosfato *N*-acetilglucosamina

GPI-GnT: complejo de la GPI-*N*-acetilglucosaminiltransferasa

EtNP: etanolamina fosfato

PI-PLC: fosfatidilinositol fosfolipasa C

PM: Peso Molecular

SDS-PAGE: electroforesis en gel de poliacrilamida con dodecilsulfato sódico

hPIGA/PIGA: Fosfatidilinositol *N*-acetilglucosaminiltransferasa subunidad A

PIGAc: Región citoplasmática de la proteína PIGA

IPTG: Isopropil- β -D-1-tiogalactopiranosido

TB: medio de cultivo Terrific Broth

URA: Uracilo

SC-URA: medio completo sintético carente de uracilo

ON: durante la noche (“Overnight”)

YPDA: medio de cultivo de levadura conteniendo extracto de levadura, peptona, dextrosa y adenina (“Yeast Peptone dextrose Adenine”)

NTA- NI: Níquel acoplado a ácido Nitrilotriacético

ORF: Marco de lectura abierta

DO: Densidad óptica leída a 600 nm

SASA: Área superficial accesible a solvente

INTRODUCCIÓN

1. Anclas de glicosilfosfatidilinositol (GPI)

Las anclas de glicosilfosfatidilinositol (GPI) son estructuras de glicolípidos que se agregan como modificaciones postraduccionales en el extremo C terminal de algunas proteínas eucariotas (Ferguson, 1999; Ferguson et al., 1985, 1988; Homans et al., 1988; Low, 1999; Nosjean et al., 1997) (**Figura 1**). Gracias a esta modificación, las proteínas pueden ubicarse en la superficie celular permaneciendo unidas a la membrana citoplasmática mediante una estrategia alternativa a la de poseer dominios transmembrana.

Las anclas de GPI están distribuidas en muchos organismos eucariotas como las plantas, insectos, mamíferos, fungi y protozoos (Nosjean et al., 1997). Son esenciales para el desarrollo embrionario en mamíferos y para la viabilidad de levaduras, siendo críticas para el mantenimiento de la morfología de las células (Peter & Menon, 2007). Por otro lado, se han descrito aproximadamente más de 250 proteínas ancladas a GPI (Ferguson, 1999; Tiede et al., 1999). Las proteínas con anclas GPI tienen diversas funciones biológicas, entre las que se encuentran la adhesión y el contacto célula-célula, algunas actividades enzimáticas, o son transportadores, o receptores, etc. Una muestra no exhaustiva y a modo de ejemplo de tales funciones se encuentra en la **Tabla 1**.

El esqueleto de la molécula de GPI está muy conservado entre especies. Su estructura está conformada por una porción lipídica de fosfatidilinositol (PI) que se inserta en la cara externa de la bicapa lipídica de la membrana celular, un núcleo glicosídico formado por *N*-acetil glucosamina (GlcNAc) y tres residuos de manosa (Man) (**Figura 1**). A continuación, se encuentra un residuo de fosfoetanolamina (EtNP), que es el que se une al extremo C de la proteína. Los residuos que se encuentran dentro del núcleo glicosídico pueden ser modificados con otros grupos de EtNP u otros azúcares (Low, 1999; Nosjean et al., 1997). Existe una gran variabilidad en la estructura de las anclas de GPI, tanto en la porción de glicano como en la parte lipídica. Por ejemplo, la primera manosa del núcleo de glicano puede contener una ramificación con otro residuo de manosa, y los residuos internos de manosa pueden ser modificados por EtNP. Las cadenas de ácidos grasos unidas al PI pueden ser insaturadas o saturadas, también pueden variar en el largo y en la saturación, pudiendo presentarse como diacilglicerol, alquilacilglicerol o ceramida. Además, el anillo de inositol también puede estar acetilado en las posiciones 2 o 3. En algunos casos estas modificaciones pueden conferirle alguna propiedad a las anclas, por ejemplo el agregado de un ácido graso como el ácido palmítico en uno de los oxhidrilos de anillo de inositol le confiere resistencia al clivaje producido por la enzima fosfatidil inositol fosfolipasa C (Paulick & Bertozzi, 2008).

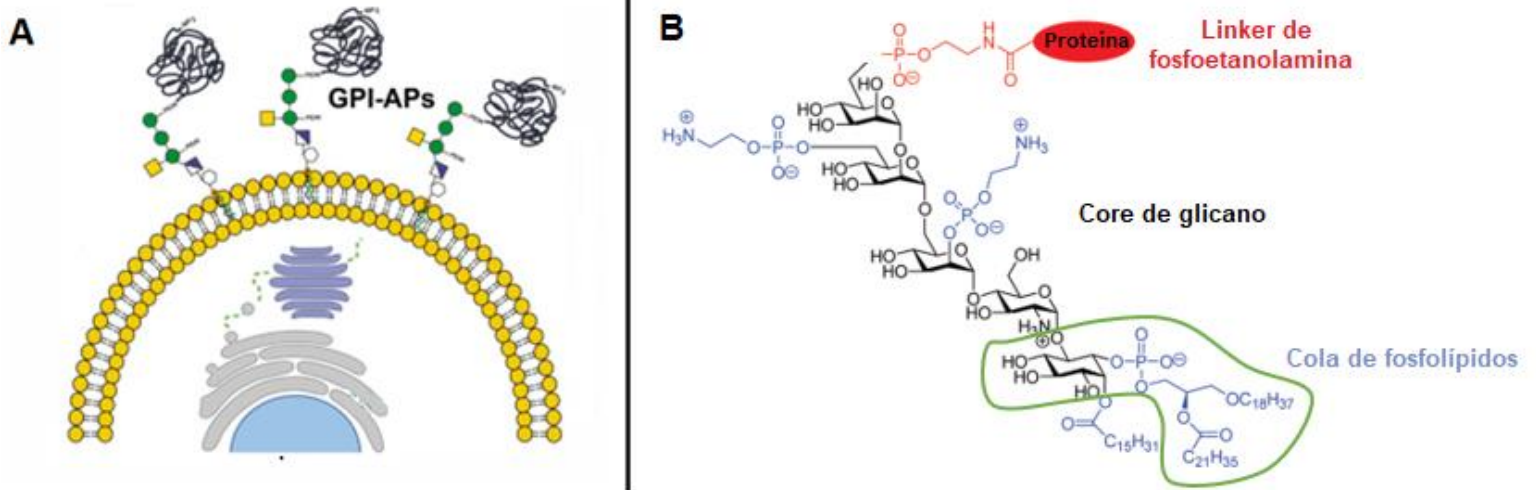


Figura 1. Proteínas ancladas a la membrana plasmática mediante GPI. **A)** Las moléculas de GPI permiten que proteínas sin dominio transmembrana se encuentren ancladas a la cara externa de la membrana. **B) Estructura química de un ancla de GPI.** La estructura se puede dividir en tres partes: un *linker* entre la fosfoetanolamina (roja) y la proteína (roja), el núcleo de glicano (negro) y una cola de fosfolípidos (azul). Los apéndices en azul son variables, así como los lípidos de la cola. En verde se resalta la región correspondiente al fosfatidilinositol. **Panel A:** adaptado de Paula A. Guerrero et al., 2021. **Panel B:** adaptado de Margot G. Paulick & Carolyn R. Bertozzi, 2008.

<i>Rol biológico</i>	<i>Proteína</i>	<i>Procedencia</i>
Enzimas	Fosfatasa alcalina 5'-nucleosidasa Acetilcolinesterasa	Tejidos de mamíferos, <i>Schistosoma</i> . Tejidos de mamíferos Células sanguíneas de mamíferos
Interacción célula-célula	LFA-3 NCAM	Células sanguíneas de humanos Mamíferos y gallinas
Regulación del sistema de complemento	CD55 (DAF), CD59	Células sanguíneas de humanos
Antígenos de mamíferos	Thy-1 Qa-2 CD14 CEA CD52	Cerebro de mamíferos y linfocitos Linfocitos de ratón Monocitos humanos Células tumorales (humanos) Linfocitos humanos

Tabla 1. Funciones de las proteínas ancladas a GPI. Adaptado de Margot G. Paulick & Carolyn R. Bertozzi, 2008.

Se han reportado un número de enfermedades genéticas relacionadas a mutaciones de pérdida de función en genes involucrados en el armado, la adición a proteínas y el remodelado de las anclas de GPI. Estas enfermedades, son llamadas en su conjunto “deficiencias heredadas de GPI”, e ilustran que cantidades normales de GPI y su maduración son muy importantes para la salud humana (Kinoshita, 2016). Además, la deficiencia en células somáticas del gen PIGA, codificante para la enzima que cataliza la primera etapa de biosíntesis de anclas de GPI, es la causa de la Hemoglobinuria Paroxística Nocturna, como se detalla en la Sección 4.

2. Ruta de síntesis de proteínas ancladas a GPI

El proceso de anclaje de las proteínas de superficie a GPI es una de las modificaciones postraduccionales más costosas en el sentido metabólico. Incluye al menos 12 pasos secuenciales que ocurren en el retículo endoplasmático catalizados por enzimas codificadas por genes PIG (Levin & Stamper, 2009).

El primer paso ocurre por la acción del complejo enzimático GPI N-acetilglucosaminiltransferasa (GPI-GnT) en el lado citoplasmático de la membrana del RE. En este paso, se transfiere el azúcar GlcNAc proveniente de su forma activada UDP-GlcNAc al PI, generando el primer intermediario: GlcNAc-PI, (paso 1 de la **Figura 2**). La reacción química que tiene lugar se muestra en la **Figura 3**.

A continuación, el grupo *N*-acetilo del GlcNAc es removido por la deacetilasa PIGL (paso 2 en la **Figura 2**) (Nakamura et al., 1997; Watanabe et al., 1999) para dar GlcNPI, que luego es translocado a la cara luminal del retículo endoplasmático por un mecanismo desconocido (paso 3 de la **Figura 2**). Se cree que esto ocurre gracias a una flipasa que no ha sido identificada (Vishwakarma & Menon, 2005).

En la cara luminal, el GlcNPI es acetilado por la acetiltransferasa PIGW generando GlcN-(acil)PI (paso 4 de la **Figura 2**) (Doerrler et al., 1996; Murakami et al., 2003). Luego el GlcN-(acil)PI se somete a una remodelación (paso 5 en la **Figura 2**) (Kanzawa et al., 2009). La reacción de remodelado y la enzima que cataliza la reacción no están del todo dilucidadas aún.

Los siguientes pasos son la transferencia de dos residuos de manosa al GlcN-(acil) PI, reacciones que, como las siguientes de la vía, transcurren en la cara luminal del RE. La primera manosa, Man1, es transferida al GlcN-(acil)PI, generando Man-GlcN-(acil)PI (paso 6, **Figura 2**) (DeGasperi et al., 1990). El complejo PIG-M:PIGX lleva a cabo la reacción (Ashida et al., 2005; Maeda et al., 2001). La segunda manosa, Man2 se transfiere gracias a PIGV, generando Man-Man-GlcN-(acil)PI (paso 7 de la **Figura 2**) (Kang et al., 2005).

La siguiente modificación es la adición de un residuo de EtNP a la segunda posición de la Man1, generando Man-(EtNP)Man-GlcN-(acil)PI (paso 8 de la **Figura 2**), catalizada por PIGN (Hong, Maeda, et al., 1999). Luego un tercer residuo de manosa es transferido del Dol-P-Man por PIGB, una manosiltransferasa 3, generando Man-Man-(EtNP)Man-GlcN-(acil)PI (paso 9 de la **Figura 2**) (Takahashi et al., 1996). Seguidamente el EtNP, que conectará al ancla de GPI con las proteínas, se une a la Man3 en la posición 6 (paso 10 de la **Figura 2**). Este paso es llevado a cabo por PIGO:PIGF (Hong et al., 2000; Inoue et al., 1993; Menon et al., 1993). El resultado, EtNP-Man-Man-(EtNP)Man-GlcN-(acil)PI, es llamado H7 y ya está en condiciones de unirse a proteínas.

El intermediario H7 puede modificarse generando EtNP-Man-(EtNP)Man-(EtNP)ManGlcN-(acil)PI, conocido como H8 (paso 11 de la **Figura 2**) (Hirose et al., 1992). Es el precursor maduro y es el que usualmente se usa como ancla para las proteínas. Tanto H8 como H7 son unidos post-traduccionalmente a las proteínas por el extremo C (paso 12 de la **Figura 2**) (Maxwell et al., 1995; Sharma et al., 1999). Las pre-proteínas ancladas a GPI tienen una secuencia que funciona como señal para su translocación al retículo endoplasmático en el extremo N, y una secuencia que es identificada para el ataque del GPI en el extremo C terminal (Maxwell et al., 1995).

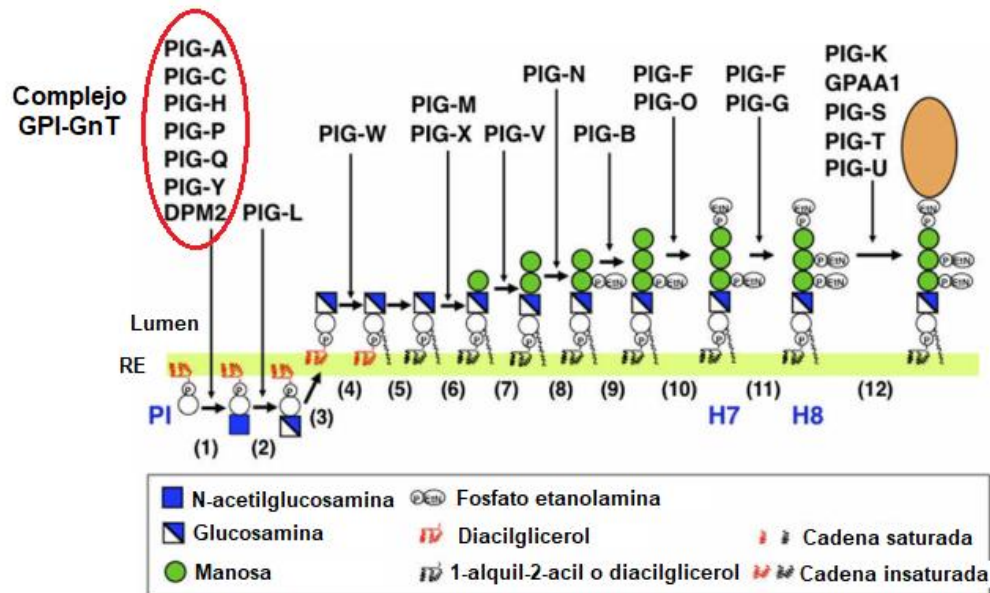


Figura 2. Biosíntesis de GPI y reacción de anclaje a proteínas. El precursor de GPI llamado H8 es sintetizado a partir de PI y GINAc a través de una secuencia de 12 pasos. Se encuentran involucrados en el proceso más de 20 genes. Adaptado de Taroh Kinoshita & Morihisa Fujita, 2016.

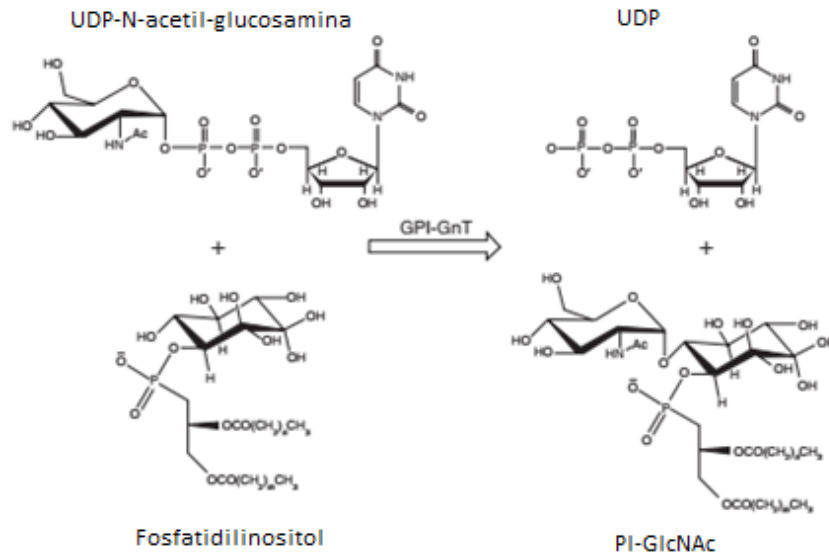


Figura 3. Primer paso de síntesis de GPI, formación del intermediario GlcNAc-PI. La *N*-Acetilglucosamina (GlcNAc) es transferida desde el UDP-*N*-acetil-glucosamina al fosfatidilinositol, esta reacción es llevada a cabo por la GPI-GnT. Adaptado de David E. Levin & Ronald J. Stamper, 2009.

3. PIGA y el complejo GPI-GnT

El complejo enzimático GPI-GnT (enmarcado en rojo en la **Figura 2**) es el responsable del primer paso de la síntesis de GPI. Se encuentra muy bien caracterizado funcionalmente en mamíferos y levaduras, específicamente en levadura *Saccharomyces cerevisiae*. Dicho complejo se encarga de transferir el GlcNAc del UDP-GlcNAc a un PI que se encuentra en la membrana del RE (Levin & Stamper, 2009). En total son siete las proteínas que conforman dicho complejo, tanto en mamíferos como en *S.cerevisiae* (se indica nombre del gen codificante en mamíferos/nombre del gen en levaduras): PIGA/Gpi3, PIGC/Gpi2, PIGH/Gpi15, PIGP/Gpi19, PIGQ (hGpi1)/Gpi1, PIGY/Eri 1 y DPM2/ Yil102c-A. A la fecha, no se ha reportado la estructura tridimensional de ninguna de las subunidades presentes en ninguna especie. Los modelos teóricos de las distintas subunidades se describen en detalle en la sección de Resultados. Una representación topológica de las subunidades y su orientación en la membrana del RE se muestra en la **Figura 4**.

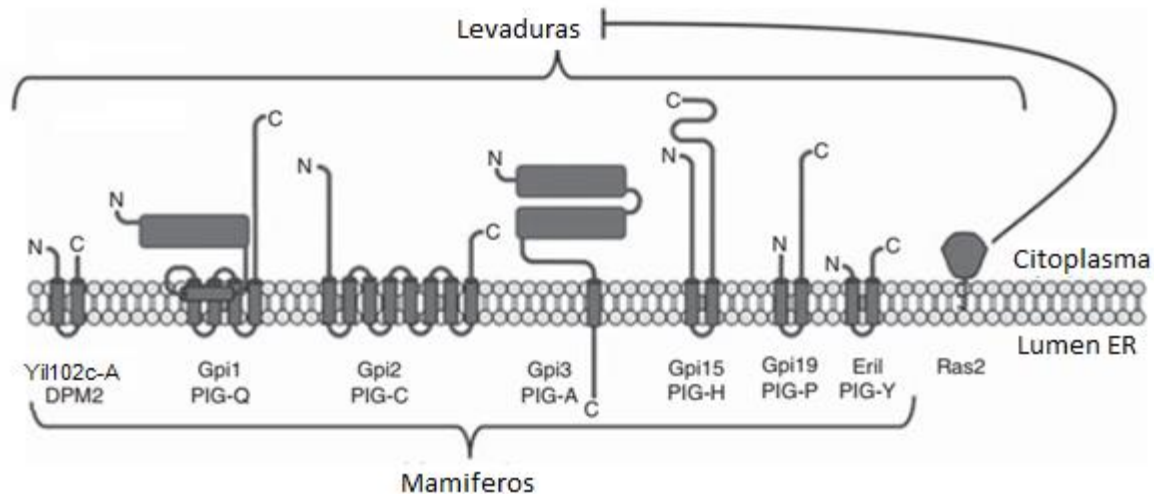


Figura 4. Esquema de la composición y topología del complejo GPI-GnT en mamíferos y levaduras. La mayor parte del complejo está orientado hacia el lado citoplasmático del RE, donde ocurre la síntesis de las anclas. Adaptado de David E. Levin & Ronald J. Stamper, 2009.

PIGA/Gpi3 es una proteína esencial en la biosíntesis de GPI ya que es la subunidad catalítica del complejo GPI-GnT encargado del primer paso de su síntesis (Coutinho et al., 2003; Kostova et al., 2003). Esto se demostró experimentalmente porque un análogo fotoactivable de uno de sus sustratos, el UDP-GlcNAc (Kostova et al., 2000). Es además una glicosiltransferasa (GT), que pertenece a la familia 4 de la clasificación de CAZy, donde estas proteínas son agrupadas según su identidad de secuencia, mecanismo de acción, tipo de plegamiento, azúcar que transfieren y la estereoquímica del producto catalizado (Cruz et al., 2019). La proteína PIGA humana tiene un total de 484 aminoácidos (Miyata et al., 1993). Gran parte de la proteína se encuentra del lado citoplasmático del RE (incluyendo el extremo N), donde tienen lugar los primeros pasos de la biosíntesis de GPI. Luego posee una región transmembrana (residuos 422-442, un solo paso) y le sigue pequeña porción muy polar del lado del lumen del RE (el extremo C), que es importante para la localización de PIGA en el RE (Levin & Stamper, 2009): mutantes de PIGA en las cuales se eliminaron diferentes longitudes del extremo carboxilo (incluyendo la región transmembrana y luminal) modificaban su localización subcelular (Watanabe et al., 1996). El segmento citoplasmático de 420 aminoácidos es globular (Eisenhaber et al., 2003) y sus propiedades estructurales se describen en detalle en la sección Resultados. Se caracterizó a Gpi3 como el ortólogo de PIGA en *S. cerevisiae*, presentando una identidad de secuencia con la variante humana del 47% (Levin & Stamper, 2009). Posee un motivo EX₇E que está muy conservado entre los miembros de varias subfamilias de glicosiltransferasas (Campbell et al., 1997), y que ha mostrado ser importante para la función de Gpi3, como se ha revelado en ensayos de complementación con mutantes sitio

específicas donde se evaluó su capacidad de rescatar el defecto de crecimiento de una cepa mutante condicional sin Gpi3 funcional (Kostova et al., 2003).

La proteína **PIGH**, también parte del complejo inicial GPI-GnT, presenta tanto su extremo amino como carboxilo del lado citosólico del RE (Watanabe et al., 1996). Esta subunidad parece ser fundamental para lograr buenos niveles de biosíntesis de GPI (Eisenhaber et al., 2003). Experimentos de co-precipitación sugieren que PIGH tiene una asociación más directa y estrecha con PIGA que las otras subunidades (Levin & Stamper, 2009). En cuanto a su estructura, se ha postulado que posee un extremo N desestructurado (Eisenhaber et al., 2003). Se demostró experimentalmente que Gpi15 es el ortólogo funcional de PIGH en levaduras (Yan et al., 2001).

En cuanto a **PIGC**, se demostró que se une al complejo PIGA/PIGH a través de PIGQ, es decir, de manera indirecta (Levin & Stamper, 2009), como se esquematiza en la **Figura 5**. Tanto PIGC como su ortólogo en *S. cerevisiae* (Gpi12) tienen una secuencia muy hidrofóbica y se ha sugerido que tienen ocho regiones transmembrana y sus extremos N y C del lado citoplasmático (Inoue et al., 1996). Se demostró experimentalmente que la delección de Gpi12 en levaduras causa letalidad (Levin & Stamper, 2009). Se presume que PIGC funciona como proteína de andamiaje, anclando el complejo GnT al retículo y permitiendo la interacción con otras proteínas del complejo (Eisenhaber et al., 2003).

Respecto a **PIGQ**, se observó que la pérdida de función del gen codificante en *S. pombe* genera letalidad (Colussi & Orlean, 1997) y en *S. cerevisiae* produce un fenotipo termosensible (Leidich & Orlean, 1996), mientras que en células humanas resulta en una actividad enzimática residual (Hong, Ohishi, et al., 1999; Tiede et al., 2001). La falta de esta subunidad está asociada a una menor cantidad de PIGC y PIGH (Eisenhaber et al., 2003). Además, PIGQ/Gpi1 parece facilitar la interacción de PIGC con el complejo PIGA/PIGH (Levin & Stamper, 2009). Por todo lo mencionado anteriormente, se cree que el rol principal de esta subunidad es estabilizar el complejo GPI-GnT.

La subunidad **PIGP** en humanos es una proteína pequeña e hidrofóbica de unos 158 aminoácidos. Fue descubierta como un componente del complejo GPI-GnT purificado. Interactúa con PIGA y en menor medida con PIGQ, pero no se observó interacción con PIGH o PIGC (Eisenhaber et al., 2003). Se descubrió que es esencial para el primer paso de biosíntesis de GPI (Watanabe et al., 2000) al igual que las subunidades PIGA, PIGC y PIGH, ya que las células que contienen mutaciones en estas subunidades son deficientes en la expresión de anclas de GPI (Murakami et al., 2005).

PIGY es una subunidad muy pequeña del complejo GPI-GnT. Se cree que tiene dos segmentos transmembrana y que solo una porción muy pequeña sobresale de la membrana (Levin & Stamper,

2009). Se demostró que tiene sus extremos amino y carboxilo orientados del lado citoplasmático. PIGY parece ser esencial para el funcionamiento del complejo GPI-GnT, ya que células deficientes en esta subunidad no pueden sintetizar las anclas de GPI. Al igual que PIGH, parece asociarse directamente con PIGA, como se observa en la **Figura 5**. A partir de experimentos de coexpresión y coprecipitación que se realizaron con PIGA, PIGH, PIGP, PIGQ y PIGC en ausencia o en presencia de PIGY, se concluyó que el complejo GPI-GnT puede formarse en ausencia de PIGY (Murakami et al., 2005). Se ha postulado que el homólogo para PIGY en *S.cerevisiae* es Eri1p, ya que tienen un perfil de hidrofobicidad semejante y la falta de este gen también afecta la síntesis de anclas de GPI (Murakami et al., 2005), pero no presentan una similitud de secuencia apreciable, por lo que podría tratarse de otra proteína que tomó el rol de PIGY en este organismo.

Finalmente, la subunidad **DPM2** participa también en otra ruta de biosíntesis en eucariotas superiores, precisamente la del dolicol, con un rol regulatorio del complejo enzimático DPMS (dolicol-fosfato-manosa sintasa) (Maeda et al., 2000). En el caso de la biosíntesis de GPI, se observó que estimula la actividad del complejo GPI-GnT, pero su presencia no es fundamental para la actividad enzimática. DPM2 interactúa débilmente con PIGA y más débilmente aún con PIGC y PIGQ, y no interacciona con PIGH o PIGP (Watanabe et al., 2000). Si bien durante mucho tiempo se postuló que esta subunidad no se encuentra presente en *S. cerevisiae*, recientemente se ha reportado que existe un gen que codifica para una proteína llamada Yil102c-A que es un homólogo de DPM2. Se comprobó mediante un análisis funcional que Yil102c-A interactúa con el complejo GPI-GnT y también directamente con DPM1 de la ruta del dolicol. Su delección es letal y se ha comprobado que el agregado del gen DPM2 proveniente de *Trichoderma reeseide* suprime este fenotipo letal (Pilsyk et al., 2020).

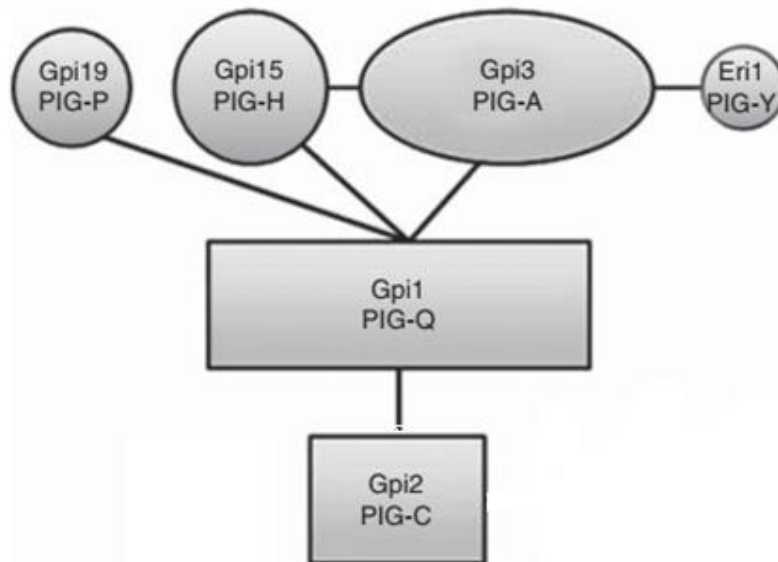


Figura 5. Asociación inferida de las subunidades del complejo GPI-GnT. Este modelo se realizó basándose en experimentos de coprecipitación de las subunidades, tanto en mamíferos como en levaduras. Las líneas que conectan a las subunidades indican una posible interacción directa. Las proteínas con más de dos dominios transmembrana están marcadas con rectángulos, mientras que las que tienen uno o dos dominios están en círculos/elipses. Tomado de David E. Levin & Ronald J. Stamper, 2009.

4. Hemoglobinuria Paroxística Nocturna (HPN)

La HPN es una enfermedad hematológica poco frecuente en la que las células sanguíneas son lisadas por ataque del sistema inmune de complemento (Hill et al., 2017). Como indica su nombre, tiene carácter paroxístico y nocturno, es decir que se manifiesta en forma de crisis generalmente durante la noche. El color oscuro de la orina es signo de la descomposición de los glóbulos rojos, que liberan hemoglobina en el torrente sanguíneo y finalmente en la orina (Carretero Colomer, 2009). Se origina por una mutación somática en el gen que codifica PIGA en una célula precursora sanguínea (CPS). Esta célula mutada (así como toda su descendencia) puede proliferar y diferenciarse, pero carece de las proteínas normalmente ancladas a la superficie celular por GPI, incluyendo a las proteínas CD59 y CD55, que regulan al sistema de complemento impidiendo que las células propias sean destruidas por el mismo. Debido a esto, las células sanguíneas portadoras de mutaciones severas (que llevan a la ausencia o a niveles extremadamente bajos de anclaje por GPI) son lisadas (**Figura 6**). Los eritrocitos son altamente susceptibles a la lisis por el complemento, resultando así en una anemia hemolítica crónica (Hill et al., 2017).

Hay muchos tipos de mutaciones somáticas que se dan en pacientes con HPN, la mayoría son inserciones o deleciones (indels) de una o varias bases, así también como sustituciones, siendo las más

frecuentes aquellas que generan un corrimiento en el marco de lectura (Nafa et al., 1995). Como el gen de PIGA se encuentra en el cromosoma X, una sola mutación del mismo es suficiente para causar deficiencia en la síntesis de GPI, pues solo un alelo es funcional tanto en mujeres como en hombres. Los demás genes involucrados en la síntesis de GPI son genes autosómicos recesivos, por eso la pérdida de función se produce generalmente solo cuando ambos alelos están mutados, evento que ocurre con poca frecuencia (Hill et al., 2017).

Las manifestaciones clínicas más comunes de la HPN son: trombosis (principal causa de muerte), insuficiencia renal (causada por la continua hemólisis intravascular), falla de la médula ósea y síndromes mielodisplásicos (etapa tardía de HPN). La HPN conlleva una importante disminución de la calidad de vida de los pacientes, ya que suele provocarles fatiga severa, disnea y fuertes dolores abdominales, entre otros síntomas (Hill et al., 2017).

La HPN tiene una incidencia estimada es de 1-1.5 por millón de personas alrededor del mundo (Socié et al., 2016; Yu et al., 2016) y es más frecuente en Asia (Japón, Corea, China) que en los países occidentales (Estados Unidos, España, Reino Unido) (Muñoz-Linares et al., 2014; Socié et al., 2016; Yu et al., 2016). Tiende a manifestarse durante la adolescencia y es rara en niños, además hay una leve predominancia en las mujeres en general (54.4%) (Hill et al., 2017). El promedio de supervivencia de pacientes con HPN desde el diagnóstico era de aproximadamente 10 años en los años 1990s (de Latour et al., 2008; Hillmen et al., 1995; Socié et al., 1996), pero se incrementó en los 2000s con la aparición de un tratamiento con un anticuerpo monoclonal llamado Eculizumab, que inhibe la lisis por parte del sistema inmune del complemento (**Figura 6**) (de Latour et al., 2008). Antes de la aparición de Eculizumab, los tratamientos se basaban principalmente en transfusiones de sangre por la hemólisis recurrente, así como suplementación con hierro y anti-trombóticos para la anemia y la prevención de la trombosis, respectivamente. En casos más avanzados con complicaciones en la médula, se optaba por trasplantes alogénicos de médula ósea (Shah & Bhatt, 2022). Se comprobó que el tratamiento con Eculizumab detiene la hemólisis intravascular, reduce la trombosis y estabiliza la función renal, entre otras mejoras del cuadro (Hill et al., 2017).

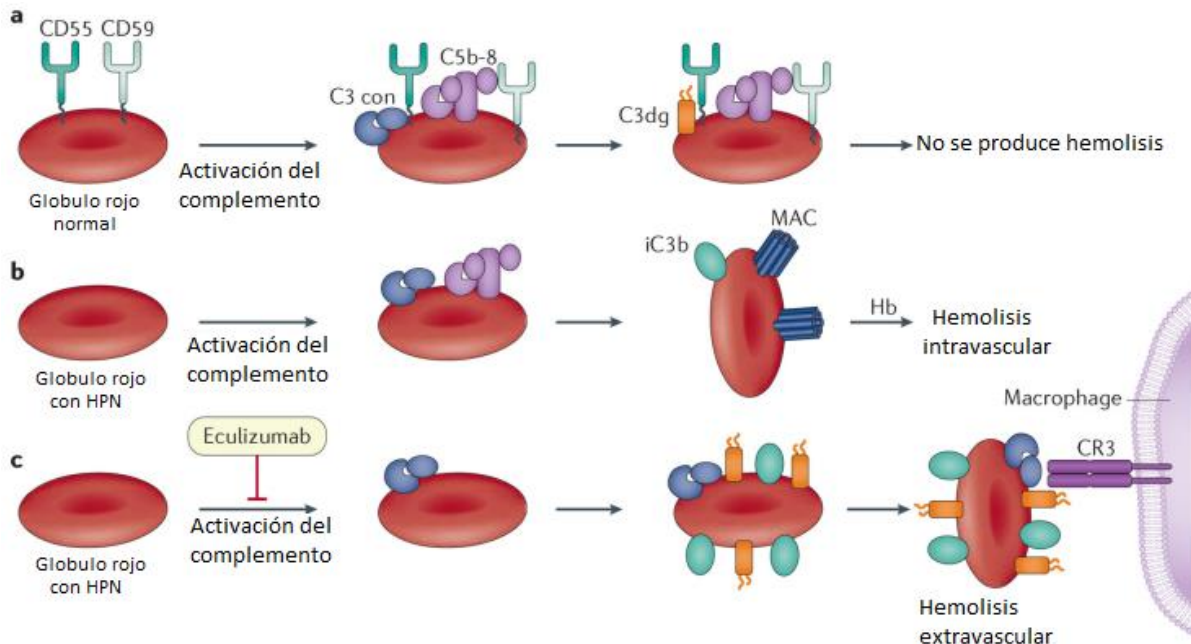


Figura 6. Hemólisis extravascular e intravascular en la HPN. Caso a: eritrocito normal protegido del sistema de complemento. Caso b: eritrocito con HPN, sensible al sistema de complemento causa hemólisis intravascular. Caso c: eritrocito con HPN tratado con Eculizumab, la deficiencia en las anclas de GPI es compensada y se previene por lo menos la hemólisis intravascular. (MAC: complejo de ataque a la membrana). Adaptado de Hill et al., 2017.

En el año 2018, la FDA aprobó el tratamiento con otro anticuerpo monoclonal llamado Ravulizumab. Este anticuerpo tiene un tiempo de vida media 3 o 4 veces mayor que Eculizumab y requiere una aplicación cada 8 semanas. Se demostró que es más conveniente en términos de costo y efectividad comparado con Eculizumab. Produce menos episodios de hemólisis irruptiva y no es inferior en términos de eficacia y seguridad. Sin embargo, como es una nueva droga, todavía no hay mucha información disponible. Cabe destacar que hay muchos otros anticuerpos monoclonales bajo investigación, además de terapias que se están desarrollando que se enfocan en distintos *targets* de la vía de complemento (Shah & Bhatt, 2022).

Si bien los tratamientos con anticuerpos monoclonales parecen ser efectivos, son también extremadamente caros, requieren aplicaciones frecuentes y algunos pacientes no responden al tratamiento. Además, estos tratamientos no se encuentran disponibles en todos los países. Adicionalmente, se comprobó que el bloqueo del sistema de complemento por el uso de Eculizumab aumenta el riesgo de infección con la bacteria *Neisseria*, causante de meningitis (Hill et al., 2017). Por todo esto, es importante seguir explorando opciones terapéuticas alternativas para la HPN, lo cual es una de las motivaciones para estudiar en profundidad al complejo GPI-GnT, y en particular a la subunidad catalítica PIGA como posible blanco terapéutico.

OBJETIVOS E HIPÓTESIS

Objetivo general

El objetivo general de este trabajo es obtener información sobre la estructura de PIGA y de las otras subunidades del complejo GPI-GnT, de modo de comprender mejor su función a nivel molecular.

Objetivos particulares

1. Desarrollar una plataforma que permita obtener tanto a la variante silvestre de PIGA como a mutantes puntuales, en bacterias o levaduras, con un mínimo número de pasos de clonado. Expresar a PIGA de manera recombinante en *E. coli*, de modo de generar un sistema para producir la proteína en cantidad y calidad para estudios *in vitro*.
2. Desarrollar un sistema *in vivo* que permita evaluar el impacto de distintas mutante de PIGA, tanto de relevancia clínica como de diseño, por complementación en levaduras que carecen del gen PIGA funcional.
3. Realizar un estudio bioinformático de PIGA y de las otras subunidades del complejo GPI-GnT, evaluando su conservación en distintos organismos eucariotas. Identificar regiones conservadas en las estructuras predichas que puedan ser importantes para la integridad estructural y actividad biológica, e integrar dicha información con las mutantes patogénicas conocidas

Hipótesis

Dado que las proteínas ancladas a GPI cumplen funciones cruciales en eucariotas, tanto la enzima PIGA como las proteínas accesorias que asisten el primer paso de la síntesis del ancla deberían estar altamente conservadas, en mayor medida cuanto más relevantes sean las subunidades para la función. Además, siendo PIGA una enzima eucariota cuyo dominio catalítico es análogo a otras glicosiltransferasas no eucariotas solubles que son homólogos remotos y no forman complejos multiproteicos, el dominio catalítico de PIGA debiera ser lo suficientemente robusto para ser expresado de manera aislada en *E. coli*. Finalmente, dada la importancia y ancestralidad de PIGA, las características principales que definen su función deben estar muy conservadas, haciendo posible la complementación funcional al expresar la variante humana en el contexto celular de una levadura que carece de PIGA funcional.

MATERIALES Y MÉTODOS

3.1. Métodos de laboratorio

Cepas y medios de cultivo

A lo largo de este trabajo se usaron distintas cepas de *Escherichia coli*. Para los procedimientos de clonado se usó la cepa DH5 α . Para la expresión de proteínas recombinantes se trabajó principalmente con la cepa RosettaGami2 (Novagen). Otras cepas utilizadas se listan en la **Tabla 2**.

En el caso de *S. cerevisiae* se usaron las cepas WT HH3 de genotipo MAT α , *lys2*, *ade2*, *his3*, *trp1*, *ura3*, *leu2* y la cepa mutante condicional GPI3 termosensible con genotipo: *spt14-1-10C::KanR*; *his3 Δ 1*, *leu2 Δ 0* *ura3 Δ 0*, *met15 Δ 0* proveniente de la colección de cepas termosensibles de Boone_TS-Array-v6-1536 (gentilmente donada por el Dr. Alejandro Colman Lerner, IFIBYNE, FBMC, FCEN, UBA).

Cepa bacteriana	Característica principal	Resistencia a antibiótico	Temperatura de inducción
Rosetta Gammi2	Permite la formación de puentes de disulfuro, contiene tRNAs con codones de bajo uso en <i>E.coli</i> .	Cloranfenicol	37°C
Rosetta2	Optimizada para poder traducir bien genes con codones poco frecuentes en <i>E.coli</i> por agregado de sus tRNAs (AGA, AGG, AUA, CUA, GGA, CCC, and CGG).	Cloranfenicol	37°C
RIL (BL21-CodonPlus(DE3)-RIL)	Optimizada para la expresión de proteínas eucariotas con codones de bajo uso en <i>E. coli</i> por el agregado de los tRNAs AGA, AGG, AUA, CUA.	Cloranfenicol	37°C
SHuffle	Reductasas gor y trxB citoplasmáticas mutadas, y DsbC expresada en citosol: se pueden formar puentes disulfuro en el citosol.	No tiene	30°C

Tabla 2. Cepas bacterianas utilizadas para los ensayos de inducción de PIGAc

Para los procedimientos de clonado en bacterias, éstas se cultivaron a 37°C con agitación constante de 250 rpm. El medio estándar utilizado fue LB (NaCl 1%, triptona 1%, extracto de levadura 0.5%) con ampicilina 100 μ g/ml. Para los procedimientos de expresión, se probaron distintos medios de cultivo, temperatura y otras condiciones, que se detallan en la sección Resultados para cada caso. Los medios de cultivo empleados, además del LB, fueron: "Terrific Broth" (TB: extracto de levadura 2.4%, triptona 1.2%, glicerol 0.4%, KH₂PO₄/K₂HPO₄ 90 mM) (Sambrook et al., 1989) y medio ZYP5052 de autoinducción (triptona 1%, extracto de levadura 0,5%, Na₂HPO₄ 50 mM, K₂HPO₄ 50 mM, NH₄SO₄ 25 mM, glicerol 0,5%, glucosa 0,05%, α -lactosa 0,2%, MgSO₄ 2 mM, FeCl₃ 50 μ M, CaCl₂ 20 μ M, MnCl₂ 10

μM , ZnSO_4 10 μM , CoCl_2 10 μM , CuCl_2 2 μM , NiCl_2 2 μM , Na_2MoO_4 2 μM , Na_2SeO_3 2 μM , H_3BO_3 2 μM (Studier, 2005).

Las levaduras fueron crecidas en medios líquidos a 24°C en agitación constante de 200 rpm o en placas. Para los ensayos de termosensibilidad, la temperatura de incubación se elevó a 37°C. Los medios de cultivo utilizados fueron: medio rico YPDA (extracto de levadura bacteriológico 1%, peptona bacteriológica 2%, glucosa 2%, sulfato de adenina 0.004%) o medio selectivo sintético completo carente del componente de la auxotrofia (uracilo): “*synthetic complete medium – ura*” (SC -URA, base nitrogenada de levadura sin aminoácidos 0,67%, glucosa 2%, *drop-out mix* sin uracilo 0.2%, adenina 0.0076%) (Sambrook et al., 1989). Para los medios sólidos se agregó agar bacteriológico 2%.

Preparación de bacterias electrocompetentes

El día anterior a la preparación de las competentes, se inocularon 20 ml de pre-cultivo en medio LB, a partir del stock congelado a -70°C, y se dejó creciendo ON a 37°C. Con el mismo se inocularon 500 ml de medio LB y se dejó creciendo a 37°C hasta una DO_{600} de 0.6. Luego se siguió el protocolo del electroporador “BIORAD *E. coli* pulser”, excepto que se usó agua para los lavados, en lugar de solución de glicerol en agua. Brevemente, se cosecharon las células 10 minutos a 4°C a 5000 x g. Se descartó el sobrenadante y se resuspendió en 500 ml de agua fría estéril. Se repitió la centrifugación, se descartó el sobrenadante y se resuspendió en 250 ml de agua fría. Se repitió la centrifugación y el pellet se resuspendió en 125 ml de agua fría. Nuevamente se centrifugó, pero esta vez se resuspendió el pellet en 1.5 ml de glicerol 10% (en hielo). Se hicieron alícuotas de 80 μl en tubos eppendorf de 1.5 ml fríos y se congelaron a -70°C.

Electroporación de bacterias competentes

Se agregaron 100 ng de ADN (en un volumen de no más de 3 μl) a 80 μl de células competentes, en esterilidad. Se mezcló y se colocó en la cuba de electroporación de 0.2 cm. Se usó un electroporador “BIORAD Gene pulser II” y se electroporó usando los parámetros: resistencia 200 ohms, capacitancia 25 μF y voltaje de 2.5 KV. Inmediatamente se transfirió a un tubo de ensayo con 900 μl de medio SOC o LB y se dejó incubando por una hora a 37°C. Se plaquearon 150 μl de la transformación en LB- agar con el antibiótico necesario para la selección en cada caso.

Transformación química de bacterias competentes

Se transformaron competentes químicas, preparadas previamente en el laboratorio, siguiendo el método descrito en Sambrook (Sambrook et al., 1989): en esterilidad, se agregaron 200 ng de ADN a 100 μl de células y se mezcló. Se incubó 30 minutos en hielo. Luego se incubó 90 segundos a 42°C y

se volvió a incubar 2 minutos en hielo. Se agregaron 900 µl de medio SOC y se incubaron 1 hora a 37°C a 250 rpm. Se plaquearon en medio LB-agar con el antibiótico necesario para la selección.

Preparación de levaduras *S. cerevisiae* competentes y transformación

Dos días antes de la transformación, se largó un pre inóculo partiendo del repique de una colonia de *S. cerevisiae* GPI3, en 2 ml de medio YPDA. Se lo dejó incubando ON a 24°C. Al día siguiente se llevó a un volumen de 15 ml con YPDA y se lo dejó nuevamente toda la noche a 24°C. El día de la transformación se dejó creciendo el inóculo hasta una DO de aproximadamente 0.8. El cultivo se centrifugó a 3000 x *g* durante 5 minutos a 4°C. Se descartó el sobrenadante, preservando un pequeño volumen para resuspender el pellet. Se lo lavó 2 veces agregando 25 ml de agua estéril fría y centrifugando 5 minutos a 3000 x *g* en frío cada vez. Se resuspendió el pellet en 1 ml de agua estéril fría y se lo transfirió a un eppendorf de 1.5 ml. Se centrifugó a 1500 x *g* durante 1 minuto a 4°C. Se descartó el sobrenadante, se resuspendió en 500 µl de acetato de litio 100 mM y se separó en 4 alícuotas de 100 µl, que se incubaron a 30°C por 10 minutos. Luego se centrifugó a 15000 x *g* por 1 minuto a 4°C y se descartó el sobrenadante. Se le agregó a cada uno, en el siguiente orden: 240 µl de PEG 3350 50%, 36 µl de acetato de litio 1 M, 50 µl de ADN de esperma de salmón 2 mg/ml, el ADN a transformar y finalmente suficiente agua para llevar a un volumen añadido de 360 µl. Las transformaciones se hicieron con 300 ng de DNA, para el caso de transformaciones con vectores circulares, o 300 ng de vector linealizado y 1 µg de inserto para el caso de ligaciones *in vivo*. Se vortexearon todos los tubos hasta resuspender el pellet completamente. Se incubaron 30 minutos a 30°C y luego otros 30 minutos a 40°C. Luego se centrifugaron 1 minuto a máxima velocidad a temperatura ambiente. Se descartaron los sobrenadantes y se agregó a cada tubo 1 ml de YPDA, para luego plaquear 250 µl en medio sólido selectivo SC-URA.

Ensayo de complementación en cepa *S. cerevisiae* termosensibles

Se prepararon dos placas SC-URA para incubar una a 35°C y otra a 24°C. A partir de una misma colonia proveniente de la ligación *in vivo* en *S. cerevisiae* se hizo una estría en cada placa del medio selectivo para incubar a las diferentes temperaturas. Se pasó el tip delicadamente primero por la placa que se incubaría a 35°C y luego por la que se dejaría a 24°C. Se repicaron también 3 colonias transformadas solo con el vector, como control negativo. Como control positivo se realizó un repique de una cepa wild type de *S. cerevisiae* HH3 que crece a ambas temperaturas.

Técnicas de extracciones de ADN

Extracciones de ADN a partir de geles de agarosa:

Las purificaciones de las bandas provenientes de la electroforesis se realizaron utilizando el kit “Wizard SV gel and PCR clean up system” de Promega o el kit “QIA quick gel extraction kit” de Qiagen, en ambos casos siguiendo el protocolo del fabricante.

Extracciones de ADN plasmídico de bacterias mediante miraprep:

Se picó con la punta de un tip estéril una colonia, se la sumergió en 10 ml de LB con el antibiótico correspondiente y se incubó ON a 37°C. Al día siguiente se realizó una miniprep con el kit “ADN Puriprep P-Kit” de Inbio Highway, con algunas modificaciones del protocolo para mejorar el rendimiento. Se centrifugó el cultivo, se descartó cuidadosamente el sobrenadante y se resuspendió el pellet en 400 µl de la solución de resuspensión mediante pipeteo. Se agregaron 2 µl de RNAsa A 10 mg/ml libre de DNAsa. Luego se adicionaron 400 µl de solución de lisis alcalina, se agitó suavemente por inversión 4-5 veces y se incubó a temperatura ambiente hasta que la suspensión se aclaró (1 a 5 minutos). Luego se adicionaron 560 µl de solución de neutralización y se mezcló inmediatamente por inversión durante 5 segundos. Se centrifugó a máxima velocidad por 10 minutos a temperatura ambiente. Se transfirió el sobrenadante a un tubo eppendorf con cuidado y se dividió en dos volúmenes idénticos. Se agregó a cada tubo un volumen de etanol absoluto y se mezcló por inversión durante 5 segundos. Se cargaron 700 µl de la muestra en la columna del kit ensamblada en el tubo colector y a partir de ese punto se siguieron las instrucciones del kit para una miniprep convencional. Brevemente, se fue cargando toda la muestra en la columna mediante centrifugaciones breves (30 segundos) en las que se fue descartando el líquido no retenido. Luego se añadieron 500 µl de la solución de lavado de columna N°1, se repitió la centrifugación, luego 700 µl de la solución de lavado N°2, centrifugando de nuevo. Luego de dejar secando la columna 15 min, se agregaron 50 µl de agua libre de nucleasas, colectando el líquido que pasa a través de la columna que contiene el ADN.

Extracción de ADN plasmídico de bacterias a escala intermedia:

Para purificar plásmidos en mayor escala, se usó un protocolo adaptado de maxiprep descrito en Sambrook (Sambrook et al., 1989). El cultivo se hizo en 50 ml de medio LB con el antibiótico correspondiente (crecido ON a 37°C). Se centrifugó a 5000 rpm por 5 minutos y se resuspendió el pellet en 1.9 ml de la solución I (Tris-HCl 25 mM pH 8, EDTA 10 mM, glucosa 50 mM). Se agregó 0.1 ml de lisozima 10 mg/ml y se dejó incubando durante 5 minutos a temperatura ambiente. Se adicionó 4 ml de solución II (NaOH 0.2 M, SDS 1%, preparada al momento de usar), se agitó suavemente y se incubó en hielo 5 minutos. Se agregaron 3 ml de solución III (acetato de potasio 3 M) y se dejó incubando 15 minutos con previa agitación leve. Se centrifugó a 7500 rpm por 7 minutos, luego se tomó el

sobrenadante y se agregaron 0.6 volúmenes de isopropanol, incubando a temperatura ambiente por 10 minutos. A continuación, se centrifugó a 7500 rpm por 5 minutos. Se lavó el pellet con etanol 75% y se resuspendió en 0.4 ml de buffer TE (Tris-HCl 10 mM, EDTA 1 mM, pH 7.5) y se pasó a un tubo eppendorf de 1.5 ml. Luego se agregó LiCl hasta una concentración final de 2.5 M, incubando 20 minutos a 45°C. Se centrifugó a 10000 rpm por 10 minutos y se tomó el sobrenadante, al que se agregaron 2 volúmenes de etanol absoluto y 0.1 volúmenes de acetato de sodio 3 M. Se llevó a -70°C por 20 minutos o a -20°C ON. Se centrifugó a 15000 rpm por 15 minutos y el pellet se resuspendió en 0.2 ml de buffer TE. Luego se agregaron 40 µl de RNAsa A 10 mg/ml y se incubó 20 minutos a 45°C. Se extrajo la fase acuosa con cloroformo/isoamílico (24:1), se agitó bien y se centrifugó a 5000 rpm durante 5 minutos. Este paso se repitió 2 veces. Se agregó acetato de sodio a la muestra hasta una concentración final de 0.3 M. Se precipitó el DNA agregando 2.5 volúmenes de etanol absoluto frío y se dejó 1 hora incubando a -20°C, luego se centrifugó a 10000 rpm por 15 minutos. Se lavó el pellet 2 veces con etanol 75% sin vortexear y se dejó secar el tiempo necesario. Finalmente se resuspendió en 25-50 µl de buffer TE.

Extracción de ADN genómico de *S. cerevisiae*

Se picó una colonia de *S. cerevisiae* y se sumergió en 2 ml de medio SC-URA (para levaduras termosensibles transformadas) o YPDA (cepa salvaje). Se dejó creciendo ON a 24°C o 28°C, según sean levaduras termosensibles o salvajes. Al día siguiente se diluyó el inóculo hasta un volumen de 20 ml y se dejó incubando nuevamente. Finalizada la incubación, se centrifugó el cultivo a 5000 rpm durante 5 minutos a 4°C. Se descartó el sobrenadante y se resuspendió el pellet en 500 µl de agua destilada. Se transfirió a un eppendorf de 1.5 ml y se centrifugó 5 minutos, se descartó el sobrenadante. Se añadió 200 µl de una solución que contenía 2% triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8.0) y 1 mM EDTA. Además se añadió 200 µl de fenol:cloroformo:alcohol isoamílico (25:24:1). Se agregaron bolitas de vidrio estériles hasta el nivel del menisco de la solución. Se vortexeó 5 minutos con intervalos de descanso de 30 segundos en hielo. Se añadieron 200 µl de buffer TE (10 mM Tris-HCl, 1 mM EDTA pH 8) y se transfirió el líquido a un tubo eppendorf nuevo. Se centrifugó a 15000 rpm durante 5 minutos y se transfirió el sobrenadante a un nuevo tubo. Se añadió 1 ml de etanol absoluto. Se mezcló por inversión y se centrifugó otros 5 minutos. Se resuspendió el pellet en 400 µl de TE, procurando mojar las paredes y se agregó 3 µl de RNAsa A [10 mg/ml]. Se incubó 5 minutos a 37°C. Se añadieron 400 µl de cloroformo y se vortexeo cuidadosamente por unos segundos. Se centrifugó por 5 minutos a 15000 rpm en frío y se transfirió la fase superior a un tubo nuevo. Se añadieron 10 µl de acetato de amonio 4 M y 1 ml de etanol. Se mezcló por inversión. Se centrifugó 5 minutos a máxima velocidad a 4°C y se lavó el pellet una vez con etanol 70% frío. Se centrifugó nuevamente. Se dejó secar al aire 2 minutos y se resuspendió en 25 µl de buffer TE.

3.1.1. Procedimientos realizados con ADN

Electroforesis en geles de agarosa

Las muestras de ADN se analizaron en geles de agarosa 1 % en buffer TAE (Tris-acetato 40 mM; EDTA 1 mM) y se utilizó un marcador (*ladder*) de 1000 pb de Inbio Highway. Las bandas fueron visualizadas con bromuro de etidio.

Digestiones con enzimas de restricción y ensayos de restricción enzimática para identificar clones positivos

Para las digestiones realizadas con las enzimas de restricción NdeI (New England Biolabs) y XhoI (Invitrogen) se utilizó el buffer CutSmart (New England Biolabs) y se incubó durante 3 horas a 37°C. Se hizo un control incubando el vector en las mismas condiciones, pero sin enzimas.

Para las digestiones realizadas con las enzimas de restricción Sall y HindIII (ambas de Promega) se utilizaron los buffers D o E (Promega), respectivamente. Cada una se hizo por separado ya que requerían diferentes buffers. Se dejó incubando en buffer D por 1.5 h a 37°C con Sall y luego se diluyó 1/3 con el agregado de agua, buffer E 10X, BSA 10X y HindIII, y se incubó 1.5 h adicionales a 37 ° C.

En todos los casos, las enzimas de restricción se inactivaron por calentamiento durante 15 minutos a 65°C. Las digestiones se analizaron en geles de agarosa y luego se purificaron como se indica en la sección de “Extracciones de ADN a partir de geles de agarosa”.

En cuanto a los ensayos de restricción enzimática para identificar clones positivos antes de enviar a secuenciar, se verificó que los clones seleccionados presenten un patrón de restricción compatible con la construcción deseada. Se realizó en primer lugar una miniprep con el kit “ADN Puriprep P-Kit” de Inbio Highway. Luego se procedió a realizar un ensayo de restricción de la miniprep. Para eso se utilizaron: 10 µl de miniprep, 1.5 µl de buffer H, 1.5 µl de BSA 10 X, 1 µl de XbaI y 1 µl de EcoRI. Se dejó incubando 1.5 horas a 37°C.

Precipitaciones de ácidos nucleicos

Las precipitaciones de los vectores y fragmentos de ADN se realizaron por la adición de 0.1 volúmenes de acetato de sodio 3 M seguido de 2.5 volúmenes de etanol absoluto frío. Se vortexeó y se dejó incubando a -20°C por 1 hora. Se centrifugó a 12000 rpm por 30 min a temperatura ambiente y se descartó el sobrenadante. Luego se añadieron 200 µl de etanol 70% a -20°C y se centrifugó a 12000 rpm durante 10 minutos (se repitió dos veces). Se descartó el sobrenadante cuidadosamente y se dejó secar el pellet al aire. Finalmente, se resuspendió en agua o en buffer TE.

Reacciones de PCR preparativas

Para las reacciones de PCR preparativas se utilizaron como templados ADN plasmídico o genómico. Las reacciones se realizaron en un volumen final de 50 µl conteniendo (se detallan concentraciones finales): MgSO₄ 1 mM, 0.3 µM de cada oligonucleótido (*primer*), dNTPs 0.3 mM de cada uno, buffer Pfx 1x y 1.25 unidades de polimerasa Platinum Pfx (Invitrogen). Los tiempos de ciclado fueron variando según el procedimiento a realizar, como se detalla en los párrafos siguientes. Los *primers* utilizados se detallan en la **Tabla 3**.

Para el caso del clonado de hPIGA en el vector pET22b(+), se usó como templado el plásmido pME-hPIGA (cedido por Taroh Kinoshita, diseñado para la expresión en células eucariota). El ciclado de PCR se compone de una desnaturalización inicial a 94°C por 5 minutos, seguida de 5 ciclos de desnaturalización (94°C, 30 seg), hibridación (50°C, 45 segundos) y extensión (68°C, 90 seg). Luego se realizaron otros 25 ciclos semejantes a los anteriores, pero cambiando la temperatura de hibridación a 53°C. Finalmente, se hizo un calentamiento a 68°C como extensión final.

Para el clonado de la región citoplasmática de PIGA (PIGAc), se empleó el kit “Q5 site-directed Mutagenesis kit” de New England Biolabs, siguiendo las instrucciones del fabricante. Se partió de la construcción conteniendo el gen completo de PIGA clonado en el vector pET22b(+). El ciclado se compuso de una desnaturalización inicial (98°C) de 30 segundos, seguida de 25 repeticiones de desnaturalización (98°C, 10 segundos), hibridación (57°C, 30 segundos) y extensión (72°C, 3 minutos y 20 segundos). Finalmente, se hizo una extensión adicional (72°C) de 2 minutos.

Para la expresión en levaduras, se amplificó el gen de hPIGA con el Hisx6 Tag de la construcción clonada en el vector pET22b(+). El templado fue desnaturalizado a 94°C por 5 minutos. Se realizaron 30 ciclos de desnaturalización (94°C, 30 seg), hibridación (58°C, 30 segundos) y extensión (68°C, 90 seg). Se hizo un calentamiento a 68°C durante 5 minutos como extensión final.

Para la obtención del gen *gpi3* de *S. cerevisiae*, se usó como material de partida una preparación de ADN genómico de la cepa HH3 (WT). En este caso, la mix de PCR se dividió en dos partes de 24 µl cada una. A una parte se le agregó 1 µl del ADN y a la restante, que se usó como control negativo, se le agregó 1 µl de agua. El ciclado de PCR se compuso de una desnaturalización inicial a 94°C por 5 minutos, luego 5 ciclos de desnaturalización (94°C, 30 seg), hibridación (48°C, 30 segundos) y extensión (68°C, 2 min). A continuación, se realizaron otros 25 ciclos semejantes a los anteriores, excepto que la temperatura de hibridación se elevó a 58°C. Finalmente, se hizo un calentamiento a 68°C como extensión final. Para esta reacción de PCR se adicionó *Pfx Enhancer Solution* (Invitrogen) a una concentración relativa de 1X.

Colony PCR

Para los propósitos de identificación de clones positivos, se recurrió a la técnica de “Colony PCR”. Esta metodología se aplicó tanto a colonias de bacterias como de levaduras, con ligeras modificaciones según el caso. Los *primers* utilizados se detallan en la **Tabla 3**.

Para *S. cerevisiae*, se resuspendió una colonia de levadura aislada en 4 µl de NaOH 20 mM. Las colonias se calentaron durante 10 minutos a 95°C. A cada tubo se le agregó 21 µl de mix de PCR compuesta de (se detallan concentraciones finales): *primers* 0.6 µM, MgCl₂ 2.4mM, dNTPs 0.36 mM, DNA polimerasa Taq Pegasus (Embiotec) y su respectivo buffer. Como control positivo se utilizó DNA plasmídico (hPIGA-pET22b (+)) o cromosómico (con el gen *gpi3 wild type*) según corresponda. Como control negativo se utilizaron 4 µl de NaOH 20 mM y mix de PCR. El ciclado de PCR se compone de una desnaturalización inicial a 94°C por 2 minutos, seguida de 35 ciclos de desnaturalización (94°C, 40 seg), hibridación (50°C, 40 segundos) y extensión (72°C, 80 seg). Finalmente, se realizó una extensión final a 72°C por 5 minutos.

Para la colony PCR de bacterias, se resuspendió una colonia en 15 µl de agua MilliQ. Luego se calentaron los tubos a 100 °C durante 5 minutos. Se centrifugó a máxima velocidad por 10 minutos y se tomaron 3 µl de cada sobrenadante, los cuales se usaron como molde para la PCR. Los componentes de la mix fueron los mismos que se utilizaron para la colony PCR en *S. cerevisiae*, cambiando solamente la concentración de los dNTPs a 0.3 mM finales. A todos los tubos se le agregaron 20 µl de la mix. Las temperaturas de desnaturalización y de extensión fueron las mismas que las utilizadas para la colony PCR de levaduras. Los tiempos de ciclado (35 ciclos) consistieron en: 1 ciclo de desnaturalización de 2 minutos, seguido de 35 ciclos de desnaturalización de 30 segundos, hibridación (55.5 °C, 30 segundos) y 2 minutos de extensión, seguido de un paso de extensión final a la misma temperatura durante 5 minutos.

Nombre	Secuencia
PIGA_Ndel-F	5'- CAGTCATATGGCATGTAGAGGTGGGGCTGGGAAT-3'
PIGA_XhoI-R	5'-AACGCTCGAGGGATCCTTGAAATACAGTTTTTCAGAACgggttcagatatCT -3'
PIGAc-F	5'- TCTGAAAACCTGTATTTCC-3'
PIGAc-R	5'-GTGAGAAATAAGTCTGTCC-3'
416PIGA-F	5'-AGAACTAGTGGATCCCCGGGCTGCAGGAATTCGATATCAATGGCATGTAGAGGTGGGGC-3'
416PIGA-R	5'- ATGTAAGCGTGACATAACTAATTACATGACTCGAGGTCGATCAGTGGTGGTGGTGGT -3'
416gpi3-F	5'-AGAACTAGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATGGGCTTCAATATAGCGATGCTGTGCGACTTTTTTATCCAC-3'
416gpi3-R	5'-ATGTAAGCGTGACATAACTAATTACATGACTCGAGGTCGACTTAAGTTTCTCTTCTTCTTCTTCTCTC-3'
Sk primer	5'- TCTAGAACTAGTGGATC-3'
Seq PIGA _r	5'- CCACTAAGCAAATCGATC-3'
T7 promoter	5'- TAATACGACTCACTATAGGG-3'

Tabla 3. Primers utilizados en las reacciones de PCR preparativas y analíticas.

Ligación *in vitro*

Para la ligación *in vitro* se utilizaron el vector y el inserto digeridos con las mismas enzimas de restricción, como se indica en la sección “Digestiones con enzimas de restricción”. Las reacciones se llevaron a cabo con una relación inserto: vector de 6:1, empleando la enzima Anza T4 ligase master mix (Invitrogen). Se dejó incubando a 25°C por 30 minutos y luego ON a 4°C. Al día siguiente se realizó la transformación por electroporación.

3.1.2. Procedimientos realizados con proteínas

Electroforesis en geles de poliacrilamida

Los extractos totales de bacterias y levaduras correspondientes a los ensayos de inducción o Western Blot se analizaron en geles de poliacrilamida preparados según el sistema Laemmli (Laemmli, 1970). Los geles contenían un porcentaje de acrilamida del 10% (para separar proteínas de entre 30 a 150 kDa) y con un ancho de 0.75 mm. A las muestras de proteínas se les agregó el volumen necesario de buffer de siembra concentrado conteniendo β -mercaptoetanol (concentración final 4%) (Sambrook et al., 1989) o DTT (concentración final 2 mM). Los marcadores de peso molecular utilizados fueron: marcador preteñido PAGE- Ruler de ThermoFischer y marcador SDS7 de Sigma. Para las condiciones de corrida el voltaje se fijó a 90 V mientras las bandas migran en el gel de stacking, luego durante la migración en el gel de separación se aumentó a 140 V.

Obtención de extractos de proteínas totales de *S. cerevisiae*

Se picó con un tip estéril una colonia y se sumergió en 2 ml de medio líquido selectivo SC-URA. Se lo dejó incubando ON a 24°C. Al día siguiente se tomó 1 ml del inóculo y se lo diluyó nuevamente en 15 ml de medio selectivo SC-URA, se incubó ON a 24°C. Se dejó creciendo hasta una DO de entre 1 y 1.5. Se centrifugó a 3000 g durante 5 minutos a 4°C. Se descartó el sobrenadante y el pellet de células se lavó con 15 ml de agua y se centrifugó nuevamente a 3000 g por 5 minutos a 4°C. Se descartó el sobrenadante y el pellet se resuspendió en 300 μ l de buffer HB (Tris 25 mM pH 7.5, 15 mM Cl_2Mg , 1% Tritón), con el agregado de inhibidores de proteasas (fluoruro de fenilmetilsulfonilo (PMSF) 100 μ M, L-1-tosilamida-2-feniletíl-clorometicetona (TPCK) 10 μ M, N α -p-tosil-L-lisina clorometil cetona (TLCK) 10 μ M, leupeptina 10 μ M, pepstatina 10 μ M y E64 10 μ M). Se transfirió todo a un tubo eppendorf de 1.5 ml. Se centrifugó a 15000 rpm durante 1 minuto a 4°C. Se descartó el sobrenadante y el pellet se resuspendió en 20 μ l de buffer HB con inhibidores de proteasas. Se agregó un volumen de 400 μ l de bolitas de vidrio lavadas y 400 μ l de buffer HB. Se vortexeó durante 2 minutos con intervalos de 30 segundos incubando en hielo por unos segundos entre cada intervalo. Se transfirió el sobrenadante a

un tubo eppendorf de 1.5 ml y se centrifugó durante 15 minutos en frío a 12000 rpm. Se guardó el sobrenadante para realizar un ensayo de Western Blot.

Pull down de hPIGA con agarosa-NTA-Ni a partir de extractos celulares

Se trabajó con extractos de *S. cerevisiae* preparados en un buffer sin EDTA ni DTT. Luego de medir la concentración de proteínas totales por Bradford, se tomó de cada muestra el volumen necesario para que tengan 250 µg de proteínas totales. Se llevó al mismo volumen final por el agregado de buffer de preparación de extractos (sin EDTA ni DTT) o buffer de equilibrado de la matriz de NTA-Ni (20 mM Tris-HCl, pH 7.4, 250 mM NaCl). Se separó un volumen equivalente a 25 µg de proteína total de cada muestra como control. Al volumen remanente se le agregó resina de agarosa-NTA-Ni equilibrada (120 µl de matriz equilibrada en 20 mM Tris-HCl, pH 7.4, 250 mM NaCl) y se incubó 1 hora a 4°C con agitación. Se repitió la centrifugación y se guardó el sobrenadante (muestra no retenida). Se añadió nuevamente buffer de equilibrado, se repitió la centrifugación y se guardó el sobrenadante (lavado). Finalmente se añadieron 80 µl de buffer de elución (buffer de equilibrado con 500 mM imidazol) y se incubó 10 minutos a 4° C con agitación. Se repitió la centrifugación y se guardó el sobrenadante (elución). Como lavado final se usó imidazol 1 M en lugar de buffer de elución.

Western Blots

Se tomaron 30 µl de extractos proteicos y se le agregaron 7.5 µl de buffer de muestra 5X con DTT 2 mM final. Se calentó por 5 minutos a 95°C y luego se centrifugó por 2 minutos a 15000 rpm a temperatura ambiente.

Las muestras se sembraron en un SDS-PAGE 10% que se corrió aproximadamente 1.5 horas, aumentando el voltaje progresivamente desde 90 volts hasta 180 volts. Al terminar la corrida se transfirieron las proteínas a una membrana de PVDF con un tamaño de poro de 0,45 µm (Thermo Scientific). La membrana se activó previamente en metanol aproximadamente 30 segundos. Se transfirió una hora a 100 volts en baño de agua-hielo en buffer de transferencia (Tris 20 mM, pH=8.8, 200 mM glicina, metanol 20%). Luego se desarmó el cassette y se embebió la membrana en TTBS (500 ml de TBS 1x, 250 µl de Tween 0.05%). Se descartó el TTBS y se tiñó la membrana con rojo Ponceau para comprobar la correcta transferencia. Se lavó 2 veces con TTBS esperando 5 minutos entre lavado y lavado con agitación. Se agregó a la membrana 20 ml de buffer de bloqueo (leche descremada 5% en TTBS) y se incubó 1 hora con agitación. Luego se hicieron lavados con TTBS con agitación, esperando 5 minutos entre lavado y lavado. Se agregó una dilución 1/1000 del anticuerpo primario anti-His mouse (Thermo Fisher) diluido en buffer de bloqueo 0.5X. Se dejó incubando con el anticuerpo primario ON a 4°C con agitación. Se realizaron 3 lavados de 10 minutos con TTBS. Se dejó incubando 1 hora a

temperatura ambiente con una dilución 1/3000 del anticuerpo secundario anti Fc mouse (Thermo Fisher) diluido en buffer de bloqueo 0.5X. Se agregó a la membrana una solución de reactivos del kit “Clarity Western ECL” de Bio Rad. Se reveló la membrana utilizando el equipo quimioluminiscente “Amersham Imager 600” de General Electric.

Ensayos de expresión de hPIGAc en *E. coli*

Se realizaron ensayos de expresión de hPIGA probando distintos medios y temperaturas, así como distintas concentraciones de IPTG.

Se realizó un pre cultivo ON (37°C) de 10 ml de LB con cloranfenicol 35 µg/ml y ampicilina 100 µg/ml, a partir del repique de la placa con colonias de Rosetta Gami 2 transformadas con hPIGAc o de las mismas células transformadas pero de un stock almacenado a -70°C. El proceso general para los distintos ensayos de inducción fue el siguiente: se inocularon varios tubos falcon conteniendo 10 ml de medio (LB, TB o ZYP-5052) con 200 µl de precultivo. Se dejaron crecer a 37°C hasta llegar a una DO de 0.8 aproximadamente. Se guardó 1 ml de cultivo no inducido de cada condición. Se agregó IPTG a cada tubo en la concentración deseada y se dejó induciendo por 4 horas a la temperatura seleccionada. En el caso del medio ZYP-5052, por ser de autoinducción, no se agregó IPTG. Luego del tiempo de inducción, se centrifugaron a máxima velocidad a 4°C durante 15 minutos y se resuspendieron los pellets en 0.1 volúmenes de cultivo de buffer de lisis (Tris-HCl 20 mM pH 7.4, 250 mM NaCl, imidazol 0.5 M). Luego se sonicaron en las siguientes condiciones: 4 pulsos de 20 segundos y 50 segundos de descanso, con una amplitud de 30%. De las muestras de lisado inducido se separó la fracción soluble centrifugando 5 minutos a máxima velocidad, en frío. Se tomaron 30 µl de muestra del lisado total y de la fracción soluble. El material insoluble se resuspendió en 0.01 volumen de cultivo de sample buffer con betamercaptoetanol, y se diluyó 1/10 previo a la siembra. Las muestras se sembraron en geles de poliacrilamida 10% de 0.75 mm.

Pruebas de purificación de PIGAc con columna de NTA-Ni

Se resuspendieron las células correspondientes a 1 litro de cultivo en 20 ml de buffer de equilibrado (Tris 20 mM pH 7.4, 0.25 mM NaCl, Imidazol 10 mM). Al momento de sonicar se agregaron 20 µl de cóctel de inhibidores de proteasa (Set III, Roche) y 200 µl de de PMSF 100 mM. Se sonicó en 5 ciclos de 30 segundos y 60 segundos de descanso con una amplitud de 30%. Se centrifugó lo sonicado durante 20 minutos a 12000 rpm en frío. Se tomó una muestra de la fracción soluble para SDS-PAGE (“siembra”). Se sembró lo soluble en aproximadamente 5 µl de matriz NTA-Ni y se juntó lo no retenido. Se lavó la columna con 40 ml de buffer de equilibrado y se recolectó el lavado. Luego se realizó la elución

con buffer de equilibrado pero suplementado con imidazol en concentraciones crecientes (80 mM, 300 mM, 500 mM y 1 M, 10 ml de cada buffer).

3.2. Métodos bioinformáticos

3.2.1. Búsqueda de ortólogos de las subunidades del complejo GnT en organismos modelo

Se hicieron búsquedas mediante HMM (Hidden Markov Models) usando el programa Jackhmmer en la web del EBI (<https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer>), usando las secuencias de las variantes humanas de las 7 subunidades recuperadas de Uniprot. La base de datos empleada fue Uniprot Reference Proteomes. La búsqueda se hizo usando los parámetros default y el número de iteraciones necesarias para llegar a convergencia (típicamente 3-5). Luego, mediante un script de Python, se retuvieron aquellas secuencias correspondientes a un conjunto de 142 organismos modelo eucariotas listados en la base de datos OrthoInspector (<https://lbgi.fr/orthoinspectorv3/database/Eukaryota/2019>). En el caso de PIGA, dado que el número de hits obtenidos en esta búsqueda fue muy grande, las secuencias no pudieron descargarse de la web (la descarga está restringida a menos de 10000 hits), por lo que la búsqueda y recuperación de secuencias se llevó a cabo usando un script en basado en las funciones NCBIWWW.qblast y Entrez.efetch de BioPython, usando como parámetros de BLAST: matrix = blosum62, gap open penalty = -11, gap extensión penalty = -1, expect=0.05, word_size=6, database = "nr". En aquellos casos donde se obtuvo más de una secuencia por organismo, se retuvo la de mayor similitud con la variante humana.

En los casos donde no se recuperaron secuencias de los organismos modelo, se realizaron búsquedas manuales usando la versión web de BLAST, con la estrategia tBlastN y los parámetros default, restringiendo la búsqueda al organismo en cuestión y probando las distintas bases de datos. En pocos casos, se recuperaron secuencias de esta forma que no pudieron identificarse por BlastP. En el caso de PIGA, también se hicieron búsquedas en bases de datos específicas, según se indica en el texto, de modo de confirmar la ausencia del gen.

Finalmente, la ortología de las secuencias se infirió mediante Reciprocal Best Hits: se usó cada secuencia recuperada para realizar una búsqueda en Blast usando los genomas de referencia de 6 especies, considerándose un auténtico ortólogo cuando el mejor hit de esta búsqueda es la proteína que se usó como *query* para recuperar la secuencia en cuestión. Los genomas de referencia corresponden a *Homo sapiens*, *Dictyostellium discoideum*, *Glycine max*, *Plasmodium falciparum*, *Saccharomyces cerevisiae*, *Schyzosaccharomyces pombe*, *Drosophila melanogaster* y *Trypanosoma cruzi*.

3.2.2. Generación de alineamientos múltiples de secuencias y análisis de distribución taxonómica

Para generar los alineamientos, las secuencias recuperadas como se describe en la sección anterior se subieron al servidor MAFFT y se generó un alineamiento inicial. El mismo se refinó eliminando aquellas secuencias muy cortas o con caracteres ambiguos o muy divergentes. Luego se repitió el alineamiento en MAFFT y de ser necesario se hicieron correcciones manualmente con Aliview. Estos alineamientos bien curados, generados con secuencias de organismos modelo, se usaron como guía para la construcción de alineamientos de todo el dominio eucariota (opción `-add` en MAFFT), luego de realizar la búsqueda con BlastP como se describe en la sección anterior, pero sin restringir a organismos modelo sino a todo el dominio eucariota.

3.2.3. Construcción del árbol taxonómico

El árbol taxonómico se construyó usando el portal Open Tree of Life (OTL) mediante el paquete `rotl` implementado en R. Este programa permite realizar la síntesis de un árbol a partir de distintos estudios filogenéticos. En primer lugar, se recuperaron los identificadores taxonómicos que usa OTL, llamados `ott_id`, usando la página Taxonomic Name Resolution Service (TNRS), para los organismos modelo de OrthoInspector. Las asociaciones se “curaron” manualmente, de modo de encontrar aquellas que no se recuperaron inicialmente, considerando nombres sinónimos. Con la lista de identificadores, se procedió a la síntesis del árbol en `rotl`. Algunos organismos no fueron incluidos en el árbol debido a que no hay un consenso en la ubicación de los mismos relativo a los otros organismos investigados (*incertae sedis*). Debido a que esto ocurrió con muy pocos organismos, no se modificó el árbol para incluirlos a posteriori. Las tareas de edición del árbol (conversiones de formato y edición de etiquetas) se realizaron en TreeGraph2 (<http://treegraph.bioinfweb.info>) y la visualización final se armó en Interactive Tree of Life (iTOL, <https://itol.embl.de>).

3.2.4 Obtención de modelos estructurales y modelado del complejo GPI-GnT

Los modelos estructurales de cada subunidad se obtuvieron a partir del servidor de Uniprot (<https://www.uniprot.org/>). Para cada una, se seleccionó en los resultados de búsqueda la variante humana (por ejemplo, PIGA_HUMAN).

En la sección “Structure” se descargó la estructura predicha por AlphaFold2 en formato `.pdb`. Para comparar estructuras obtenidas de diversas fuentes se calcularon algunas mediante el programa ESMFold (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/ESMFold.ipynb#scrollTo=JM5ciSmeTZKO>). Se siguieron las instrucciones de la página.

Para modelar el complejo GPI-Gnt en un principio se usó la opción “Multimer” de AlphaFold2 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). Dado que

el programa impone un límite en cuanto al número de residuos a modelar, para explorar la factibilidad de este enfoque se modelaron todos los posibles heterodímeros usando las combinaciones de dos subunidades. También se probó generar el modelado del complejo con RaptorX (<http://raptorx.uchicago.edu/ComplexContact/>). Este servidor predice los contactos interfaciales a partir de dos secuencias de proteínas potencialmente interactuantes (solo heterodímeros) mediante técnicas de coevolución y *Deep learning*.

RESULTADOS

4.1. Primera parte: Clonado de PIGA y estudio de la expresión de su dominio catalítico en bacterias

4.1.1. Fundamento del diseño de la plataforma para la expresión de hPIGA en procariontes y eucariotas

El primer objetivo consistió en generar una construcción genética general de PIGA, a partir de la cual se pudiera obtener tanto a la variante silvestre como las mutantes puntuales, y tanto para sistemas de expresión en bacterias como de levaduras, tratando de reducir al mínimo el número de pasos de clonado (**Figura 7**). A continuación, se mencionarán los puntos clave de cada paso, mientras que los detalles de la estrategia y los análisis que se hicieron en cada paso para confirmar identidad de la construcción buscada se detallarán en las secciones siguientes.

En primer lugar, se realizó el clonado del gen hPIGA en el vector pET22b(+) para fusionarle la secuencia correspondiente a un Hisx6 Tag (secuencia de 6 histidinas en tándem) que ya se encuentra en dicho vector (construcción A, **Figura 7**). A partir de esta construcción se obtuvo la región citoplasmática (excluye el dominio transmembrana) fusionada al Hisx6 Tag para expresar la proteína en *E. coli* (construcción B, **Figura 7**).

Partiendo de la construcción A de la **Figura 7** se obtuvo mediante PCR el *ORF* de hPIGA flanqueado por extremos complementarios a un vector de expresión de levaduras (p416-GPD), permitiendo de este modo realizar una ligación *in vivo* del gen entero fusionado a Hisx6 en *Saccharomyces cerevisiae* por *gap repair* (construcción C, **Figura 7**).

Mediante PCR inversa es posible introducir mutaciones puntuales en el gen clonado en la construcción A (construcción D, **Figura 7**). Luego aplicando las mismas estrategias descritas en los dos párrafos precedentes, pueden derivarse las construcciones para la expresión tanto en bacterias como levaduras de las mutantes correspondientes. En este trabajo de tesis no se estudiaron mutantes puntuales, pero podrían obtenerse las construcciones necesarias fácilmente, haciendo uso de la plataforma generada en este trabajo.

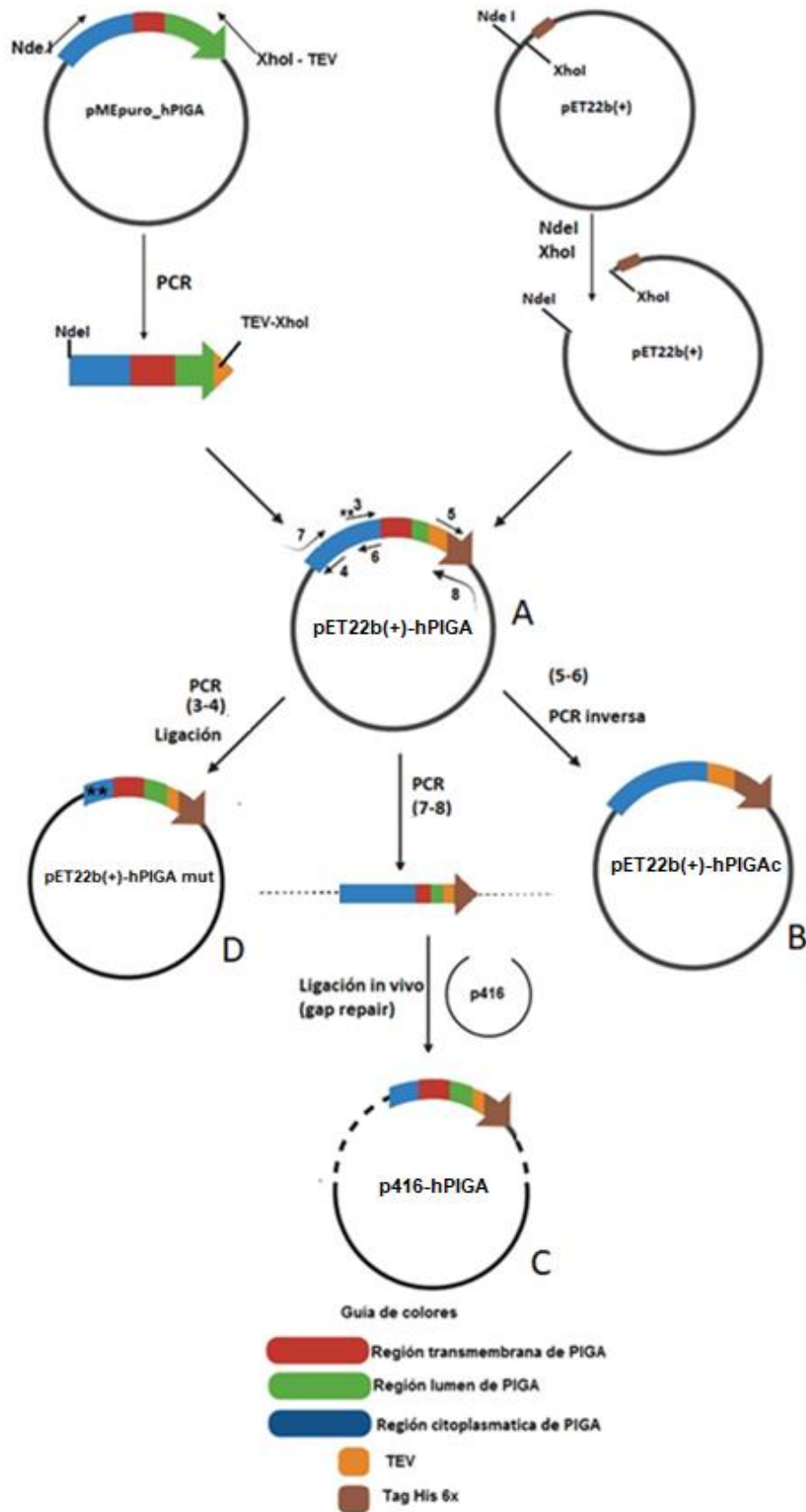


Figura 7. Esquema simplificado de la estrategia de clonado de PIGA. En la guía de colores se muestran las distintas regiones que componen el inserto. Las diferentes flechas numeradas sobre el vector pET22b(+)-PIGA representan los primers utilizados para cada caso. Los asteriscos sobre la construcción D representan los aminoácidos mutados.

4.1.2. Clonado de hPIGA en un vector de expresión de bacterias con HisTag y sitio TEV

El marco de lectura codificante para hPIGA se obtuvo del vector de expresión en células de mamífero pMEpuro_hPIGA, gentilmente cedido por el Dr. Taroh Kinoshita de la Universidad de Osaka, Japón. La región codificante se obtuvo por PCR utilizando los *primers* PIGA_NdeI-F y PIGA_XhoI-R, que contenían en sus extremos sitios para las enzimas de restricción NdeI y XhoI, respectivamente. Al primer PIGA_XhoI-R también se le añadió la secuencia de reconocimiento para la proteasa TEV, para poder eventualmente eliminar el Hisx6 Tag una vez expresada y purificada la proteína. Como se observa en la **Figura 8**, el inserto correspondiente al gen de PIGA se amplificó correctamente, obteniéndose la banda indicada con una flecha en ambas calles del tamaño esperado de aproximadamente 1500 pb. La banda se aisló y se digirió con las enzimas NdeI y XhoI, dando un producto que fue separado en un gel de agarosa y recuperado del mismo.

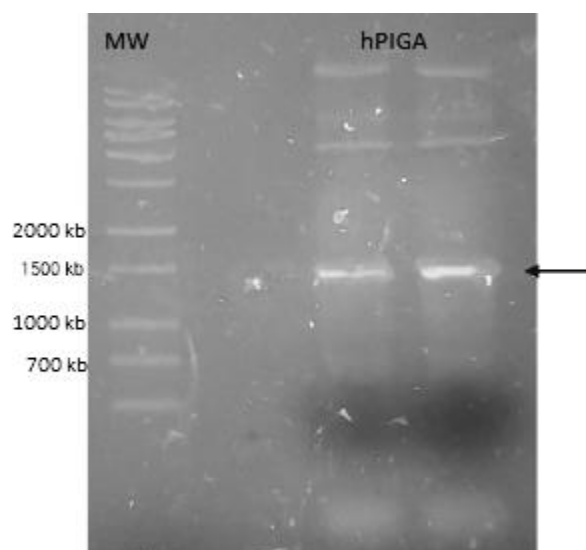


Figura 8. Obtención del gen hPIGA por PCR a partir de la construcción pMEpuro_hPIGA. Electroforesis en gel de agarosa 1% en buffer TAE 1x. Los nombres en las calles corresponden a: MW, marcador de peso molecular (Marker Ladder 1 Kb de InbioHighway), hPIGA: ambas calles corresponden a la amplificación por PCR del gen hPIGA.

Paralelamente, se digirió el vector pET22b(+) de expresión en bacterias con las enzimas de restricción NdeI y XhoI para realizar la ligación con hPIGA. Se utilizaron dos enzimas de restricción diferentes para evitar, en lo posible, la religación de pET22b(+) y para lograr que el inserto se ligue en la dirección correcta y fusionado al Hisx6 tag. De esta forma se obtuvo la construcción A de la **Figura 7** que fue verificada por secuenciación.

4.1.3. Clonado del dominio citoplasmático de hPIGA para su expresión en bacterias

Dado que uno de los objetivos era estudiar a la región citoplasmática de hPIGA, las regiones transmembrana y luminal debieron ser eliminadas de la construcción para obtener la región de interés (hPIGAc). Esto se realizó mediante una PCR inversa a partir de pET22b(+)_hPIGA (construcción A de la **Figura 7**), preservando la región codificante para el segmento citoplasmático del aminoácido 1 al 419, incluyendo a la metionina inicial. Se utilizó un kit de mutagénesis (basado en PCR inversa) que no requiere de digestiones enzimáticas para generar extremos compatibles ni de purificación de bandas luego de la electroforesis. Así se obtuvo la construcción B de la **Figura 7**.

Antes de enviar a secuenciar, se realizó una prueba de digestión enzimática para verificar que los clones obtenidos luego de usar el kit de mutagénesis correspondían a la construcción buscada. Para eso se aprovechó el hecho de que con la eliminación de los segmentos transmembrana y luminal, también se remueve un sitio de corte para la enzima EcoRV (**Figura 9**). El gen completo contiene 2 sitios de corte espaciados por 2714 pb, mientras que el segmento citoplasmático sólo tiene uno. Luego de hacer la reacción, se observó el patrón de 2 bandas esperado para el vector conteniendo el inserto hPIGA (2714 + 4132 pb) y una única banda del tamaño esperado para los clones con hPIGAc (6651 pb) (**Figura 10**). Finalmente, el resultado fue confirmado por secuenciación.

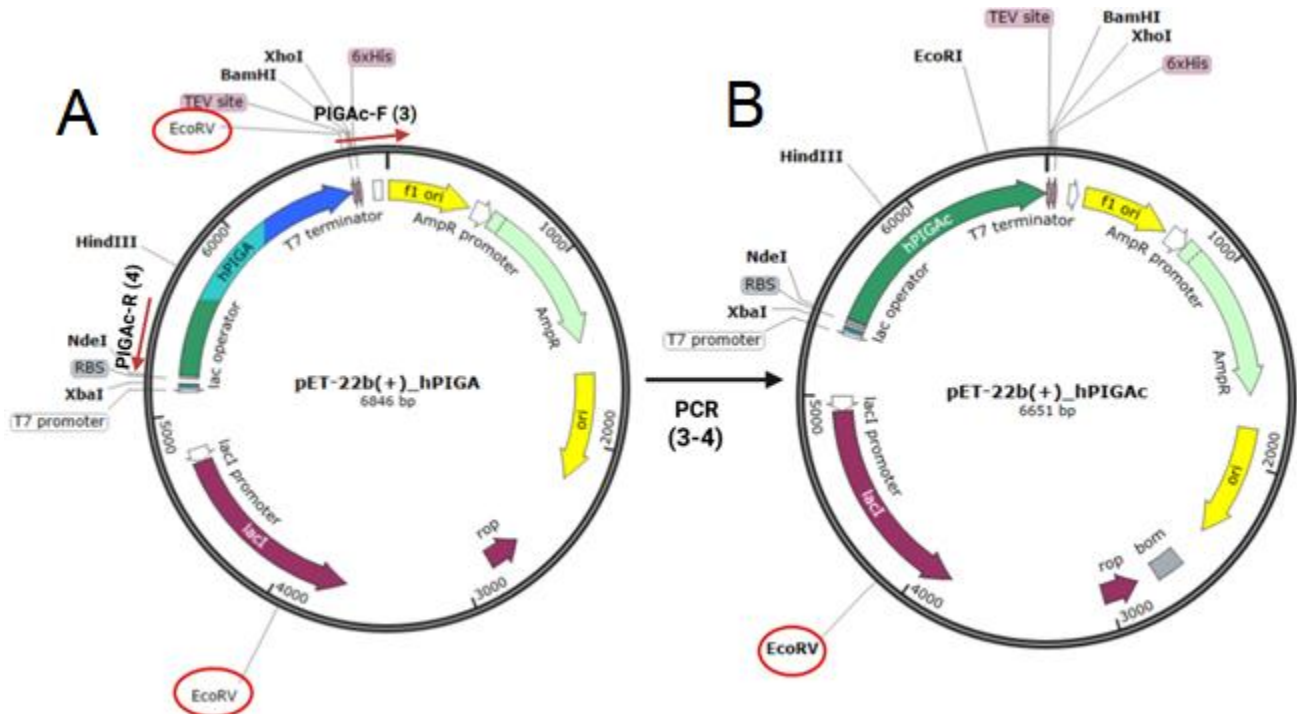


Figura 9. Estrategia de clonado de PIGAc en bacterias. A partir de la construcción A con los primers PIGAc-R y PIGAc-F (flechas en rojo) y utilizando el kit "Q5 site-directed Mutagenesis kit" se realizó la mutagénesis para obtener la región que codifica para la parte citoplasmática de la proteína (B).

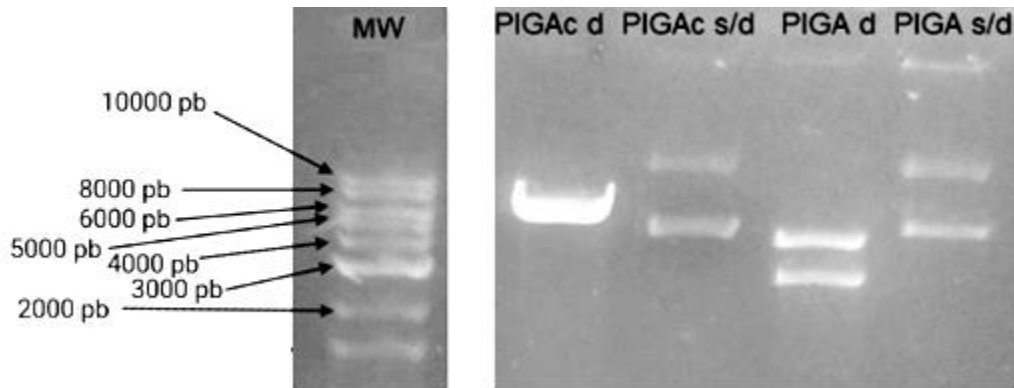


Figura 10. Resultado de la restricción enzimática de PIGA completa (PIGA) y de la parte citoplasmática (PIGAc) clonadas en pET22b(+). Electroforesis en gel de agarosa. Las abreviaturas corresponden a: MW, marcador de peso molecular (Marker Ladder 1 Kb de InbioHighway), PIGAc d, región citoplasmática de PIGA digerida con EcoRV PIGAc s/d, región citoplasmática de PIGA sin digerir, PIGA d proteína entera digerida con EcoRV, PIGA s/d proteína entera sin digerir. Los plásmidos sin digerir mostraron una migración diferencial (vector superenrollado respecto al lineal).

4.1.4. Expresión de PIGAc en bacterias

A la hora de elegir un sistema de expresión en *E. coli*, se tuvo en cuenta que la proteína posee 12 residuos de cisteína que podrían potencialmente formar enlaces disulfuro y, además, que el gen es de origen humano, con una frecuencia de uso de codones diferente a la de bacterias. Se empleó como primera opción a la cepa RosettaGami 2, la cual posee mutaciones en los genes *trxB* y *gor* que codifican para la tiorredoxina reductasa y la glutatión reductasa citoplasmáticas, respectivamente. Esta cepa mutante posee un citoplasma menos reductor que favorece la formación de enlaces disulfuro, así como también presenta una expresión aumentada de codones de bajo uso en *E. coli* (pero de mayor uso en humanos) debido a la presencia de tRNAs poco frecuentes.

La cepa RosettaGami2 fue transformada con el vector pET22b(+)_hPIGAc y luego se realizó una prueba de inducción estándar en medio líquido LB, con el agregado de IPTG 1 mM como inductor, durante 4 horas a 37 °C con agitación. Se probaron varios clones y se observó una potencial banda de inducción de baja intensidad (**Figura 11**). El PM esperado para el fragmento citosólico hPIGAc es de 49 kDa. En la **Figura 11** se observa una banda de mayor intensidad que el control correspondiente sin inducir, con el peso aproximado al de la proteína (señalado con un círculo).

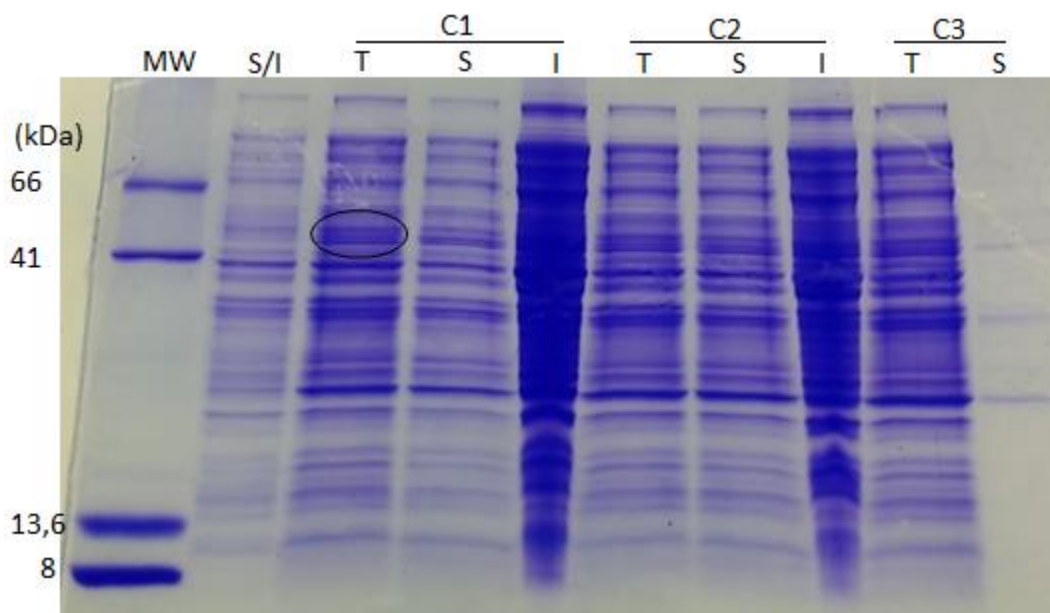


Figura 11. Prueba de inducción estándar de clones transformados con hPIGAc. Los nombres en las calles corresponden a: MW, marcador de peso molecular, S/I, muestra sin inducir. C1, C2 y C3 son tres clones transformados e inducidos, de los cuales se sembró la fracción total (T), soluble (S) e insoluble (I). Gel SDS-PAGE 10%. El círculo señala la banda de mayor intensidad que el control sin inducir, cercana al peso molecular de la proteína.

Por esa razón, se procedió a realizar una exploración sistemática de las condiciones de expresión, ensayando distintos medios de cultivo, concentraciones de inductor (IPTG) y diferentes temperaturas (**Figura 12**). Como puede observarse en un gel de SDS-PAGE representativo (**Figura 13**), ninguna de las condiciones resultó en un nivel elevado de sobreexpresión. Sin embargo, en medio TB a 37°C con 0.25 mM de IPTG la banda de interés presentaba mayor intensidad relativa que en las otras condiciones ensayadas (señalada con asterisco en la **Figura 13**) y se usó como condición para las siguientes purificaciones.

También se ensayó la expresión en el medio ZYP5052, ya que según la experiencia con otras proteínas con las que se ha trabajado en el laboratorio, el uso de autoinducción resultó en un muy marcado incremento del nivel de sobreexpresión. Este medio de cultivo contiene una cantidad de glucosa que reprime la expresión de la proteína bajo control del promotor T7, permitiendo el crecimiento de las células hasta una determinada DO en la cual la glucosa se agota y entonces se usa la lactosa para continuar con el crecimiento. La lactosa también induce la expresión de la proteína de interés incluso a una muy alta DO a la vez que reprime la expresión indeseada (*leaky*) en la fase inicial de cultivo. En este caso el uso del medio ZYP5052 no resultó en una mejora significativa (no mostrado).

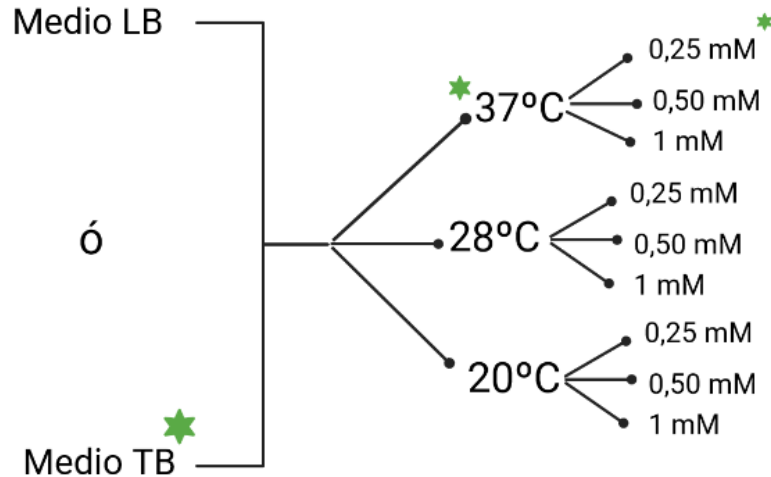


Figura 12. Variables ensayadas para la optimización de la expresión de PIGAc. Las variables utilizadas fueron: medio de inducción (LB o TB), temperatura (37, 28, 20 °C) y concentración de IPTG (0.25, 0.5, 1 Mm). Los asteriscos verdes en cada variable indican las condiciones óptimas de inducción (medio TB, 37°C y 0,25 mM).

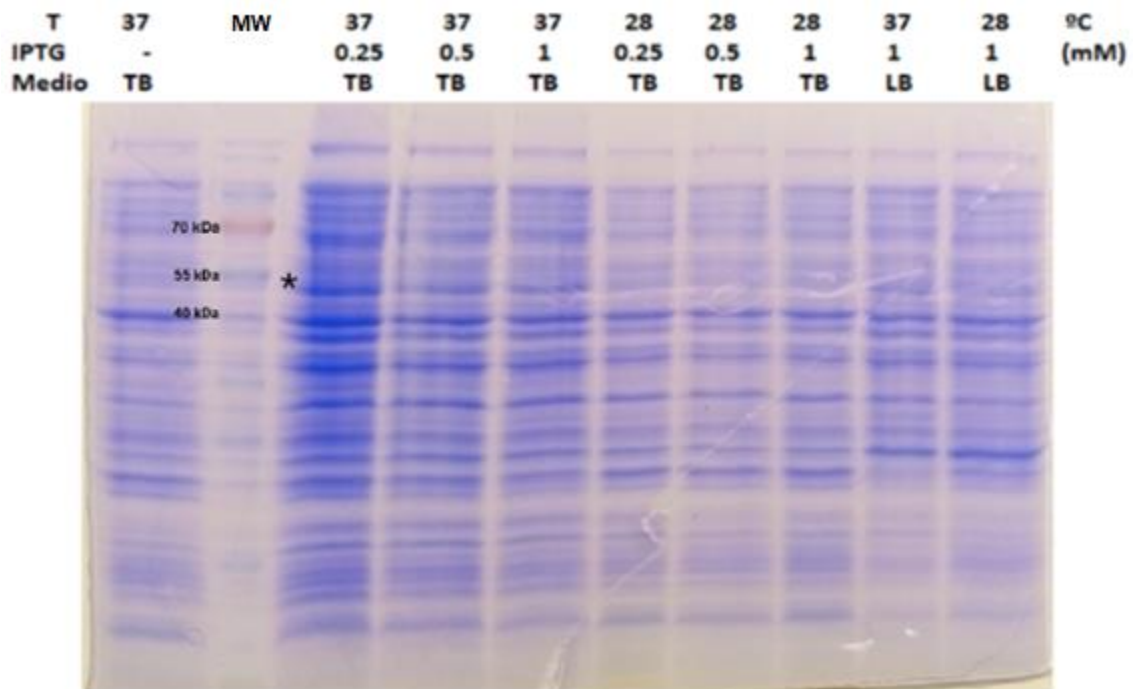


Figura 13. Ensayo de optimización de la expresión de PIGAc en Rosetta Gami2. Se muestran las fracciones solubles de lisis de los cultivos crecidos en las condiciones indicadas. MW: marcador de peso molecular. Gel SDS-PAGE 10%. El asterisco indica la banda de interés.

Dado que la purificación por NTA-Ni es sencilla y fácilmente escalable, se intentó purificar a PIGAc usando la mejor condición de cultivo encontrada a pesar de que las pruebas de expresión no mostraron niveles importantes de expresión. Para ello, las células se lisaron por sonicación agregando previamente un cóctel de inhibidores de proteasas (Set III, Roche y PMSF). En este ensayo, se realizó una purificación por afinidad en columna de NTA-Ni aprovechando que la proteína cuenta con un Hisx6 Tag. El uso de este tag permite que se una a la resina formando un complejo. La elución de la fase estacionaria se realiza utilizando concentraciones crecientes de imidazol, que desplaza a las histidinas unidas al níquel. De esta manera se recogen diferentes fracciones del eluido, entre las cuales están las que contienen la proteína. En este caso la primera elución se hizo con el buffer conteniendo imidazol 500 mM y luego se hizo un lavado con imidazol 1 M. Luego de la elución, se observó en un análisis por SDS-PAGE una banda más enriquecida alrededor del peso molecular de PIGAc (cerca a la del marker de 45 kDa) en la calle correspondiente a la elución 1 (**Figura 14**).

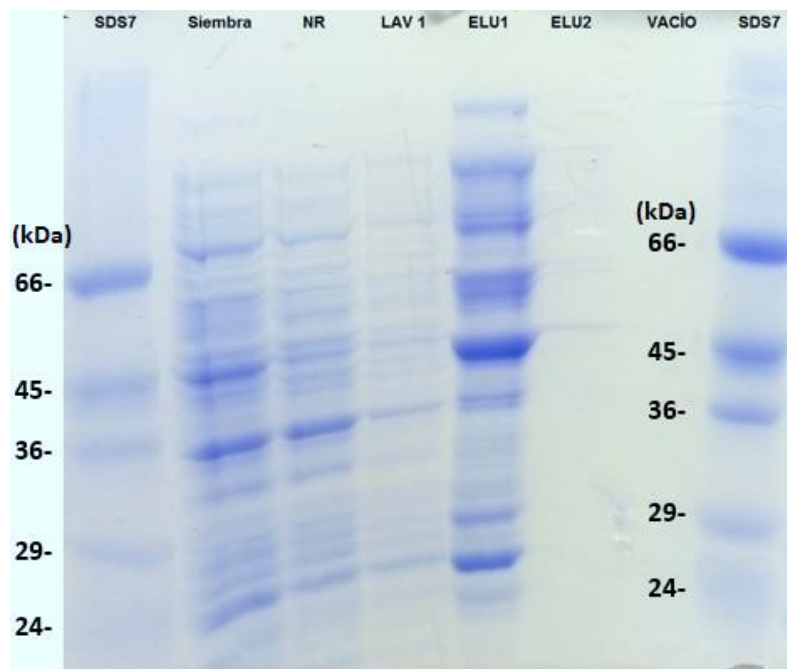


Figura 14. Purificación de muestras de inducción de Rosetta Gami2 transformadas con PIGAc en columna de NTA-Ni. Gel correspondiente a las muestras solubles. Los nombres de las calles corresponden a: SDS7: marcador de peso molecular (kDa), Siembra: antes de pasar por la columna, NR: no retenido en la columna, LAV1: lavado sin imidazol, ELU1: elución imidazol 500 mM, ELU2: elución imidazol 1 M, Vacío: calle sin sembrar. Gel SDS-PAGE 10%.

Sin embargo, en intentos posteriores de mejorar la purificación, no se logró el mismo enriquecimiento de la banda de interés. Por ejemplo, como se muestra en **Figura 15**, luego de la elución no se obtuvo casi proteína. Tampoco fueron exitosos otros intentos de mejorar la expresión, como

cambiando la DO al momento de inducir o iniciando el cultivo a partir de una placa de transformación nueva en lugar de un stock congelado a -70 °C (**Figura 15**).

En paralelo, también se ensayó la expresión con otras cepas de *E. coli*, en particular Rosetta2, RIL y SHuffle. Las dos primeras tienen la expresión aumentada de codones de bajo uso en *E. coli* que favorecen la traducción de proteínas humanas, al igual que la cepa Rosetta Gami. La última, contiene mutaciones en los genes que codifican para las reductasas *gor* y *trxB*. Además expresa DsbC en el citosol, esta enzima presente en el retículo endoplasmático cataliza la formación y ruptura de puentes disulfuro entre cisteínas dentro de una proteína mientras esta se está plegando. Esto permite que la proteína encuentre rápidamente su conformación tridimensional correcta, por lo tanto la enzima funciona catalizando el plegamiento proteico. Igualmente, no se obtuvo algún resultado que haya sido significativo en comparación a los descritos en los párrafos precedentes.

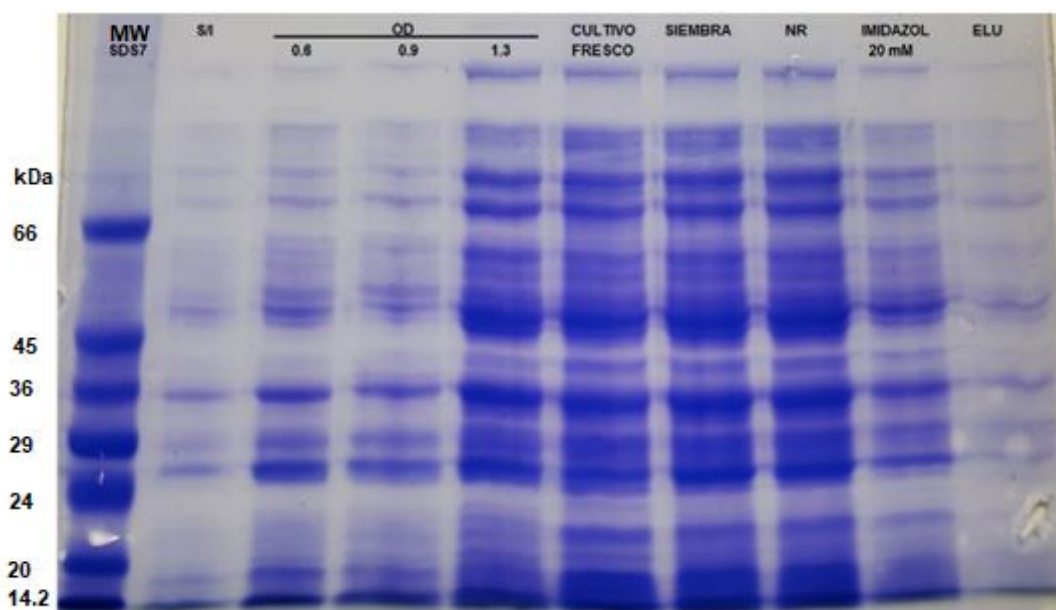


Figura 15. Purificación de Rosetta Gami2 transformadas con PIGAc (inducidas a diferentes O.D y a partir de nueva transformación) en columna de NTA-Ni. Gel correspondiente a las muestras solubles. Los nombres de las calles corresponden a: SDS7: marcador de peso molecular, S/I: muestra sin inducir, O.D: muestras inducidas a las respectivas O.D (0.6, 0.9, 1.3), Cultivo fresco; muestra correspondiente a la placa de transformación nueva, Siembra: antes de pasar por la columna, NR: no retenido en la columna, Imidazol 20mM: concentración de imidazol que se usó para una primera elución, ELU: elución imidazol 500 mM. Gel SDS-PAGE 10%.

Debido a que los nuevos intentos de purificar a PIGAc de la fracción soluble no fueron exitosos, se hipotetizó que en la prueba inicial (**Figura 14**) la proteína pudo expresarse de manera soluble debido a un factor no controlado que no se replicó en los intentos posteriores, pero que la proteína sí se estaría expresando, aunque de manera insoluble. Para probar esto, las células fueron lisadas en un buffer que contenía urea en una concentración de 8M, que solubiliza la mayor parte de las proteínas. Luego la elución se hizo por una reducción del pH de los buffers conteniendo urea (**Figura 16**). En este caso tampoco se detectó una banda del tamaño esperado, aunque sí hay bandas enriquecidas de menor tamaño.

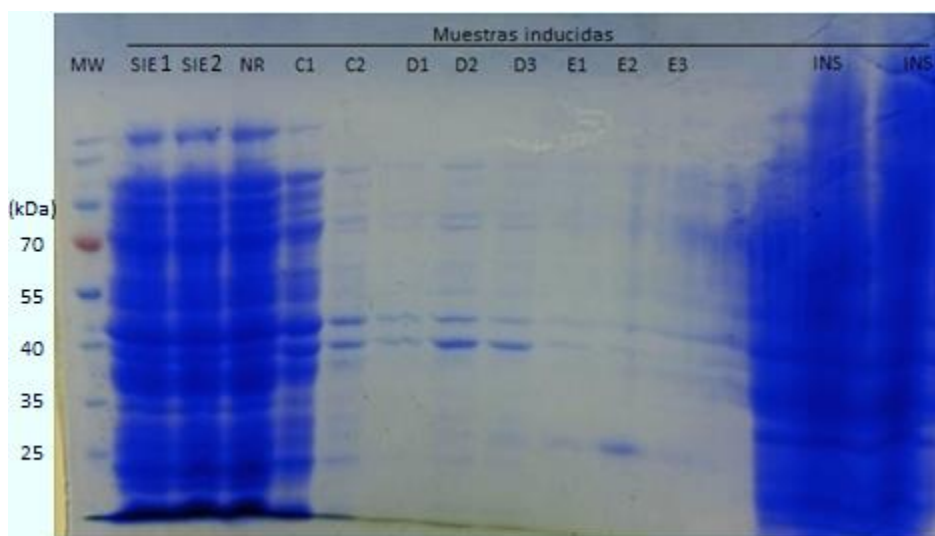


Figura 16. Purificación en condiciones desnaturalizantes de PIGAc. En el gel superior están las muestras inducidas. Los nombres de las calles corresponden a: SIE1 y SIE 2: 5 o 12.5 ul, respectivamente, de la muestra antes de sembrarla en la matriz, NR: no retenido en la matriz, C1 y C2: lavados con buffer C (pH= 6.19), D1, D2, D3: eluciones con buffer D (pH= 5.73), E1, E2 y E3: eluciones con buffer E (pH= 4.34), MW: marcador de peso molecular, INS: corresponde a las fracciones insolubles. Gel SDS-PAGE 10%.

Los resultados de todas estas pruebas de expresión y purificación empezaron a ser consistentes con la idea de que la banda de expresión que se estaba siguiendo podía tratarse de una proteína endógena de *E. coli*. Para verificar esto se realizó un ensayo en donde tanto células Rosetta Gami2 sin transformar como transformadas con PIGAc, se indujeron con IPTG. En paralelo se realizaron controles donde las células sin inducir se cultivaron el mismo tiempo. Luego se midió la DO 600 de cada cultivo, se tomó el volumen necesario para dar la misma masa de células y se sonicó en presencia de inhibidores de proteasas. Finalmente se purificó de la fracción soluble en resina de NTA-Ni como se realizó anteriormente.

En la **Figura 17** se presenta el resultado obtenido. Se puede observar que alrededor de los 55 kDa, (nótese el cambio de markers) hay unas bandas que eluyeron tanto con 80 mM como con 300 mM de imidazol, en las calles que pertenecen a las bacterias transformadas pero no inducidas. Estas bandas también se notaron muy levemente en las calles correspondientes a las Rosetta Gami2 sin transformar. En base a esto, se concluyó que la proteína cuya banda de expresión se estuvo siguiendo era una proteína de *E. coli*. Se desconoce la causa por la que su nivel de expresión es tan variable (notar las diferencias de intensidad en el gel entre las distintas condiciones), lo que provocó que se asociara erróneamente con PIGAc en los primeros ensayos.

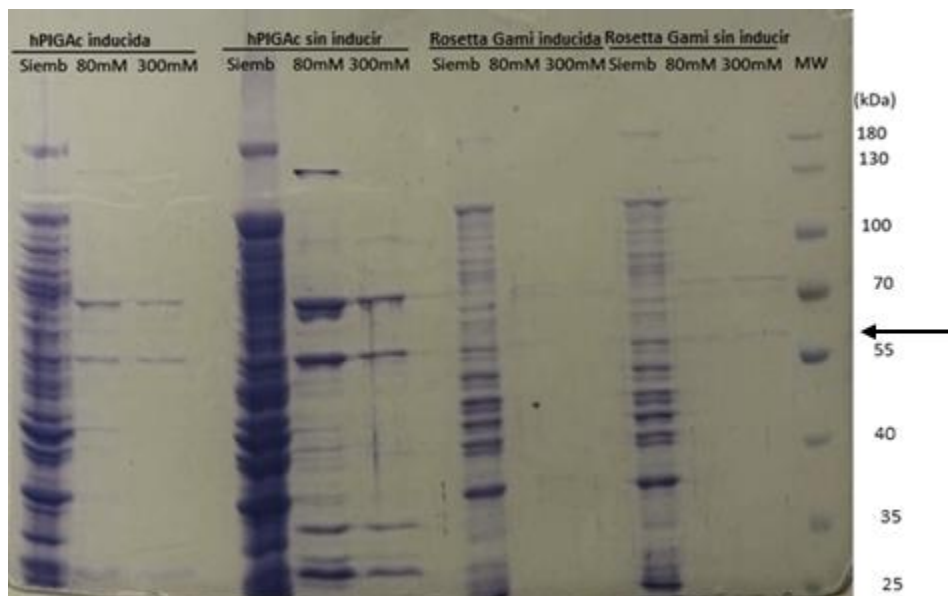


Figura 17. Purificación con NTA-Ni a partir de Rosetta Gami2 transformadas o sin transformar con PIGAc. Gel correspondiente a las muestras solubles. De izquierda a derecha las calles se agrupan de a tres de acuerdo a la condición: bacteria transformada con hPIGAc e inducida (hPIGAc inducida), transformada y no inducida (hPIGAc sin inducir), sin transformar e inducida (Rosetta Gami2 inducida) y sin transformar y no inducida (Rosetta Gami2 sin inducir). Siemb: corresponde a lo sembrado en la matriz de NTA-Ni, 80 mM y 300 Mm corresponden a las concentraciones de imidazol que se usaron para eluir. MW: marcador de peso molecular.

Análisis computacional de la expresión de PIGAc en bacterias

Una causa común de falla en la expresión heteróloga es que el organismo de origen del gen tenga una frecuencia de uso de codones diferente a la del sistema de expresión, y que el ORF a expresar tenga codones de bajo uso. En este trabajo se usó un gen humano no optimizado para la expresión en *E. coli*. Es por eso que se analizó la secuencia para buscar codones de bajo uso. En la **Figura 18** se graficó la frecuencia en el uso de codones en función de la secuencia de PIGAc. Como se puede ver, en una cepa como *E. coli* K12 habría muchos residuos que no serían traducidos con eficiencia

dificultando obtener la proteína funcional. En cambio, si se analizan los gráficos correspondientes a las cepas que se utilizaron para la expresión de PIGAc: Rosetta Gammi2, Rosetta y RIL, se observa que se corrige bastante la limitación en la traducción debido a que éstas poseen un pool de codones de bajo uso aumentado ya que los tRNA correspondientes son sobreexpresados desde un plásmido, aunque quedan unos pocos codones de muy bajo uso (CGA y CGG, ambos codifican para la arginina) que igualmente podrían afectar la expresión de PIGAc.

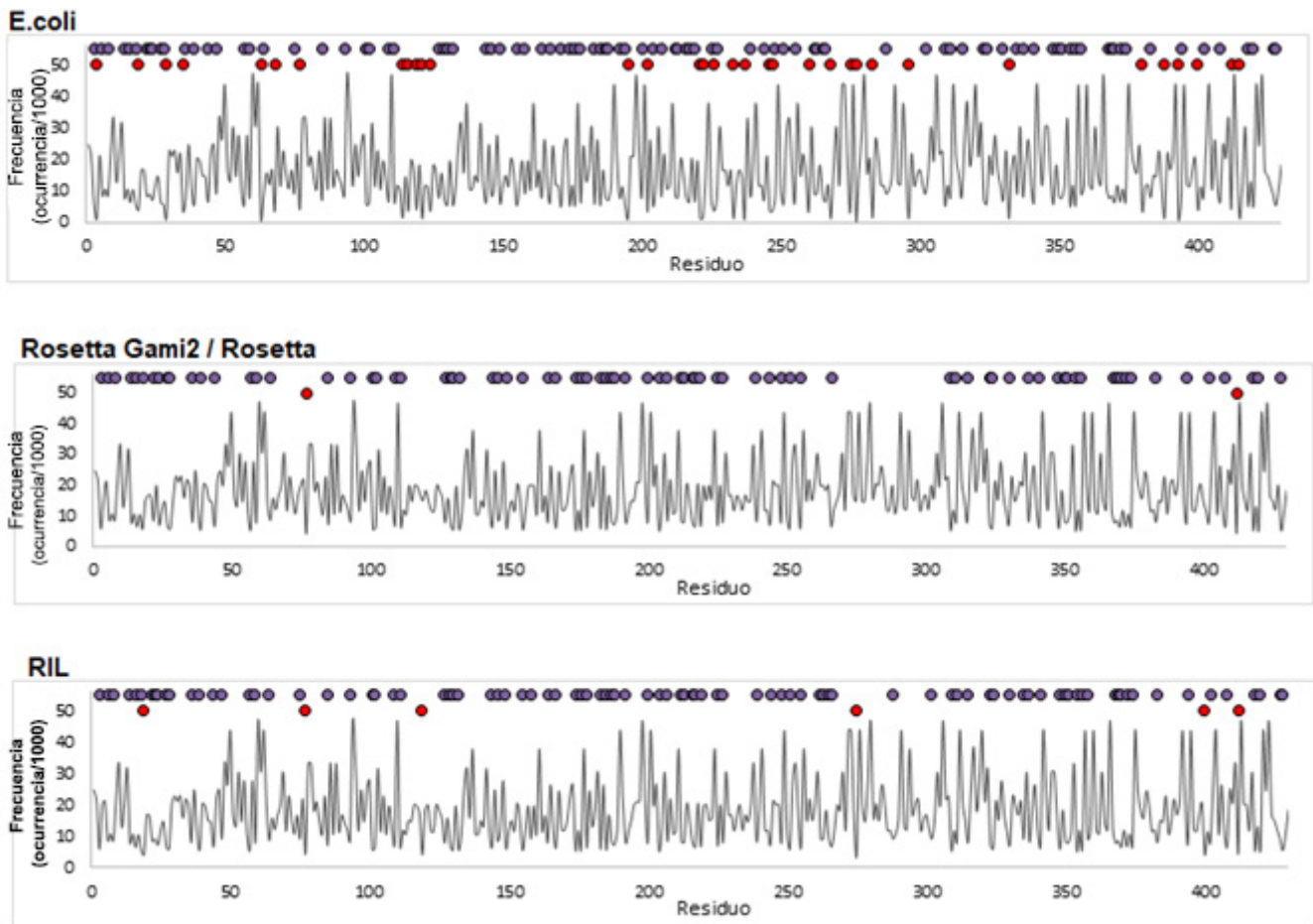


Figura 18. Gráficos de frecuencia de uso de codones en función de la secuencia de PIGAc. Los círculos azules representan los codones que pueden ser traducidos por la bacteria con alta frecuencia, mientras que los círculos rojos son los codones que tienen menos frecuencia en la cepa dada.

Otra causa común de problemas de expresión puede estar en la secuencia de ARNm. Esto puede ser su inestabilidad que hace que sea más rápidamente degradado, pero también la formación de estructuras que hacen más ineficiente el reclutamiento de la maquinaria de traducción. Esto fue analizado computacionalmente con el algoritmo TISIGNER. Este servidor permite ajustar la expresión de proteínas optimizando la accesibilidad del ARNm a los sitios de iniciación de la traducción (Bhandari

et al., 2021), mediante el cálculo de una energía de apertura de una estructura del ARNm que ha sido correlacionada con la tasa de éxito de expresión de un gran número de proteínas. El servidor reporta un éxito de expresión como un puntaje entre 0 (falla de expresión) y 100 (éxito). Se observó que los primeros codones tienen un drástico efecto en la tasa de éxito predicha de expresión, ya que con la secuencia completa de PIGAc el índice de expresión es de 9.27, pero al eliminar los primeros 19 codones mejora a un valor de 88.53 y si se acepta la optimización de los primeros codones, el índice de expresión llega a 96.13. Así, la expresión de una variante de PIGAc donde el extremo N esté recortado es una posible estrategia para mejorar la expresión en el futuro.

4.2. Segunda Parte: Estudio funcional en levaduras *S. cerevisiae*

El objetivo principal en este caso fue evaluar si la adición del gen hPIGA mediante un plásmido revierte el fenotipo termosensible de levaduras *S. cerevisiae* que son mutantes condicionales para GPI3, el gen homólogo a PIGA. Este fenotipo termosensible permite que las levaduras con la mutación en dicho gen crezcan a 24°C pero no a 35°C, como se observa en la **Figura 19**.

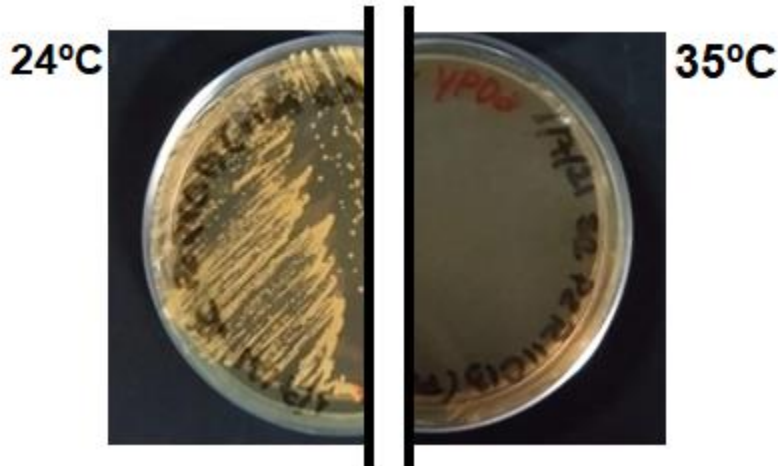


Figura 19. Cepas *S. cerevisiae* termosensibles donde el gen *gpi3*, homólogo de hPIGA, fue mutado haciéndolo no funcional. Las levaduras crecen a 24°C (izquierda) pero no a 35°C (derecha).

A partir de la construcción pET22b(+)_hPIGA (Construcción A, **Figura 7**), se subclonó el gen hPIGA en el vector *shuttle* (tiene origen de replicación y selección tanto en bacterias como en levaduras) p416 por ligación *in vivo* (*gap repair*) en *S. cerevisiae* termosensibles obteniéndose la construcción C de la **Figura 7**. Debe notarse que en este caso se usó el gen codificante para proteína completa, con su segmento transmembrana y región luminal y con el Hisx6 Tag obtenido de la construcción A. Para lograr el clonado por *gap repair* es necesario que el inserto presente secuencias idénticas al vector en sus extremos. Las mismas fueron introducidas por PCR usando los primers 416PIGA-F y 416PIGA-R, como se lista en la **Tabla 3**. El vector p416, de expresión constitutiva en levaduras, se digirió previamente con las enzimas HindIII y Sall. Se procedió luego a realizar la transformación de levaduras competentes agregando el inserto y el vector lineales realizando además algunos controles: transformación sólo con vector superenrollado, sólo con vector digerido, y con dos relaciones vector:inserto correspondientes a 1 µg de vector digerido y 300 ng inserto o 1 µg de vector digerido y 600 ng de inserto.

Las placas de transformación se dejaron creciendo a 24°C en medio selectivo carente de uracilo, auxotrofia que permite la selección de las células que incorporaron el vector p416. Se obtuvo una mayor cantidad de colonias en las placas correspondientes a la transformación con relación vector:inserto 1 µg:300 ng (**Figura 20**), seguidas de las placas que contenían la relación 1 µg:600 ng. Se obtuvieron muy pocas colonias en las placas correspondientes a las transformaciones con vector vacío superenrollado y ninguna en la placa correspondiente a la transformación con sólo el vector linealizado.

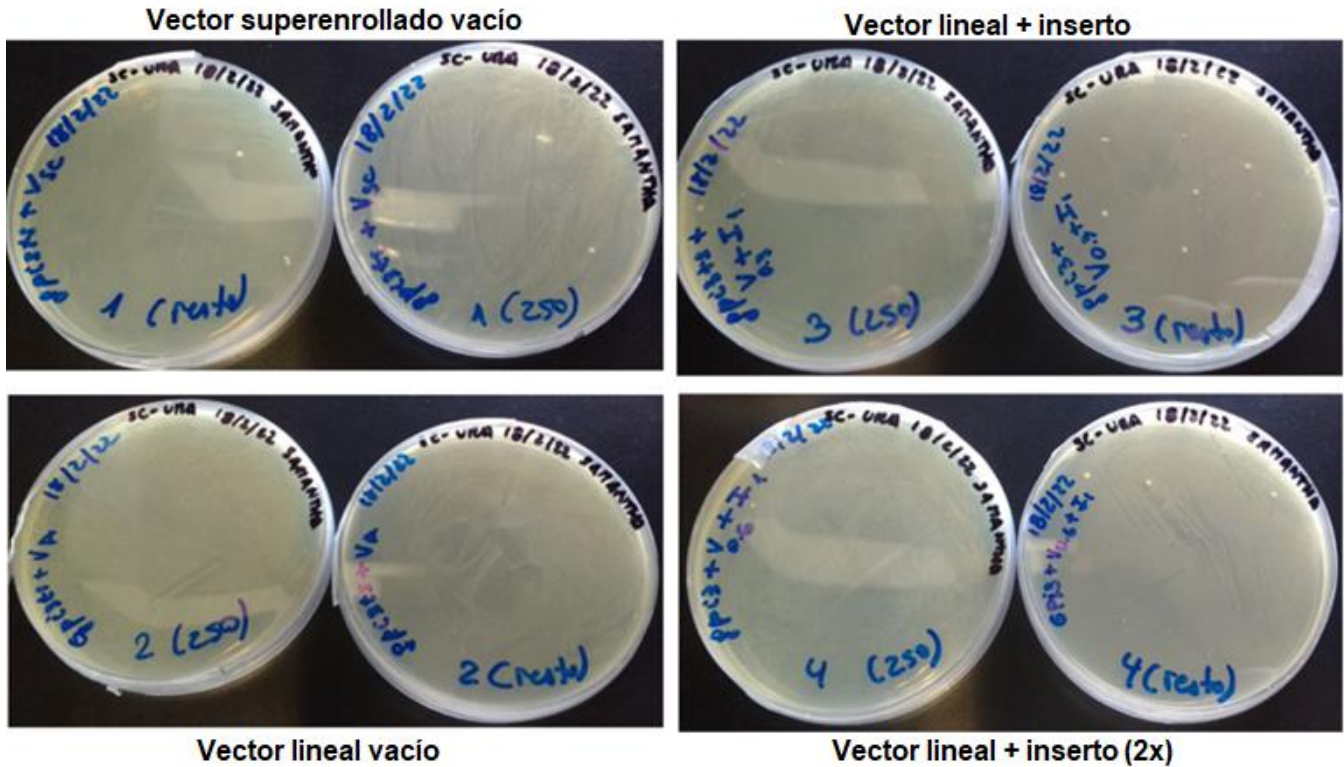


Figura 20. Transformación de *S. cerevisiae* termosensibles con p416 + hPIGA. Se agregaron las cantidades de vector e inserto mencionadas en el texto y se crecieron en el medio selectivo a 24°C.

Para verificar la correcta ligación, se realizó una *colony* PCR de todas las colonias que crecieron en las placas. Se usó un primer que hibrida en el vector y otro que lo hace en el inserto, por lo que la amplificación sólo es posible si se produjo la ligación *in vivo* (**Figura 21 A**). Como control positivo de amplificación se utilizó DNA plasmídico pET22b(+)_hPIGA con 2 primers que se unen a los extremos del ORF de hPIGA. También se incluyó un control negativo sin ADN. El tamaño de banda esperado para los clones positivos (con el vector e inserto ya ligados) es de 1500 pb aproximadamente, y el control positivo amplifica un tamaño semejante. En la **Figura 21 (panel B)**, correspondiente a *colony* PCR sobre levaduras transformantes, se pueden observar varios clones positivos para las ligaciones *in vivo*

(indicadas con las flechas). A partir de este resultado, se repicaron y aislaron algunas de estas colonias para su posterior uso, como se describe a continuación.

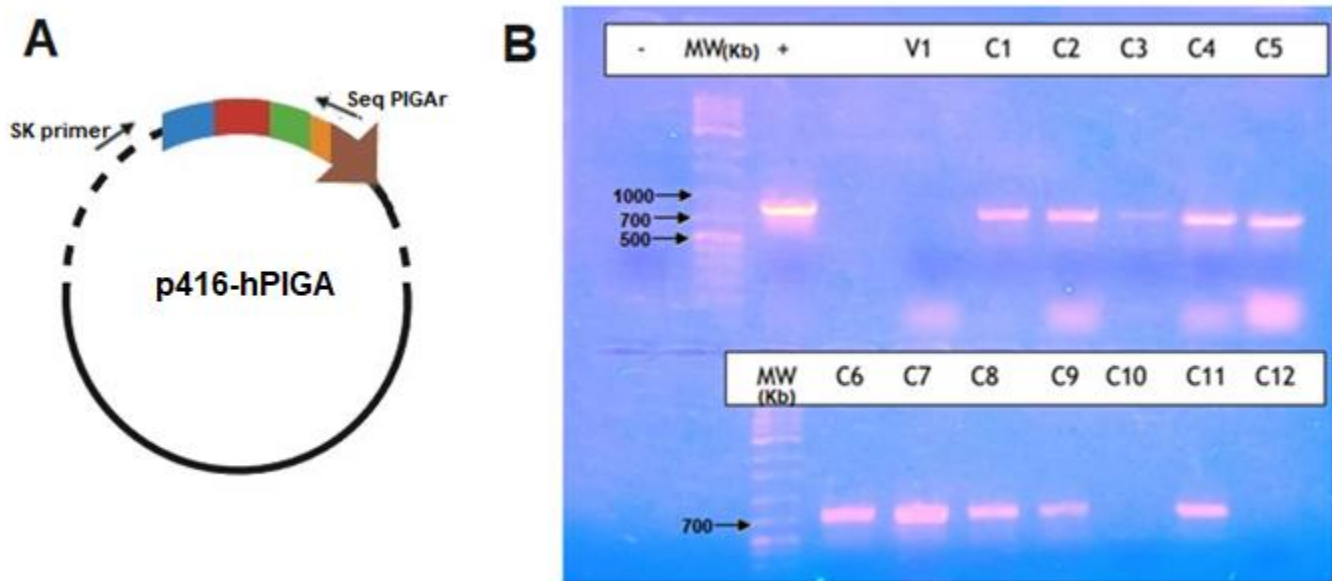


Figura 21. Colony PCR de la transformación de *S. cerevisiae* termosensibles con hPIGA. Panel A: esquema del vector p416-hPIGA con los primers utilizados para la *colony* PCR. Panel B: se muestran los resultados de la *colony* PCR. Los nombres de las calles corresponden a: (-): control negativo, MW: marcador de peso molecular (Ladder 1 Kb de InbioHighway), (+): control positivo, V1: colonia transformada con el vector vacío superenrollado, C1, C2, C3, C4, C5, C6, C7, C12: colonias transformadas con 300ng de inserto y 1 µg de vector, C8, C9, C10, C11: colonias transformadas con 600 ng de inserto y 1 µg de vector. Electroforesis en gel de agarosa 1x en buffer TAE 1x.

4.2.1 Ensayos de complementación con hPIGA en *S. cerevisiae* termosensibles

Para determinar si la variante humana de PIGA complementa el fenotipo de la cepa termosensible GPI3(ts) (permisiva a 24°C y restrictiva a 35°C), se repicaron las colonias que fueron utilizadas para el ensayo de *colony* PCR, haciendo estrías en dos placas de medio SC-URA como se indica en la sección de Materiales y Métodos. Una de las placas se dejó creciendo a 24°C (temperatura en la que crece la cepa termosensible) y la otra se dejó creciendo a 35°C (temperatura en la que crece la cepa *wild type* y *levaduras complementadas pero no la termosensible*). Se estrilaron tres colonias pertenecientes a las transformaciones con vector vacío como control negativo, y como control positivo se estrió una cepa *wild type* de *S. cerevisiae*.

Se puede observar que las colonias incubadas a 24°C tuvieron un crecimiento similar al control positivo (izquierda, **Figura 22**). Por otro lado, a 35°C no se observó un crecimiento diferencial de las colonias positivas para la ligación con inserto respecto a las transformadas con el vector solo (derecha,

Figura 22). Debido a esto se cree que la transformación con hPIGA no complementó el defecto de las levaduras. Esto puede deberse a que la proteína humana no puede sustituir a GPI3 en el complejo GPI-GnT de levaduras o a que los niveles de expresión no son los adecuados.

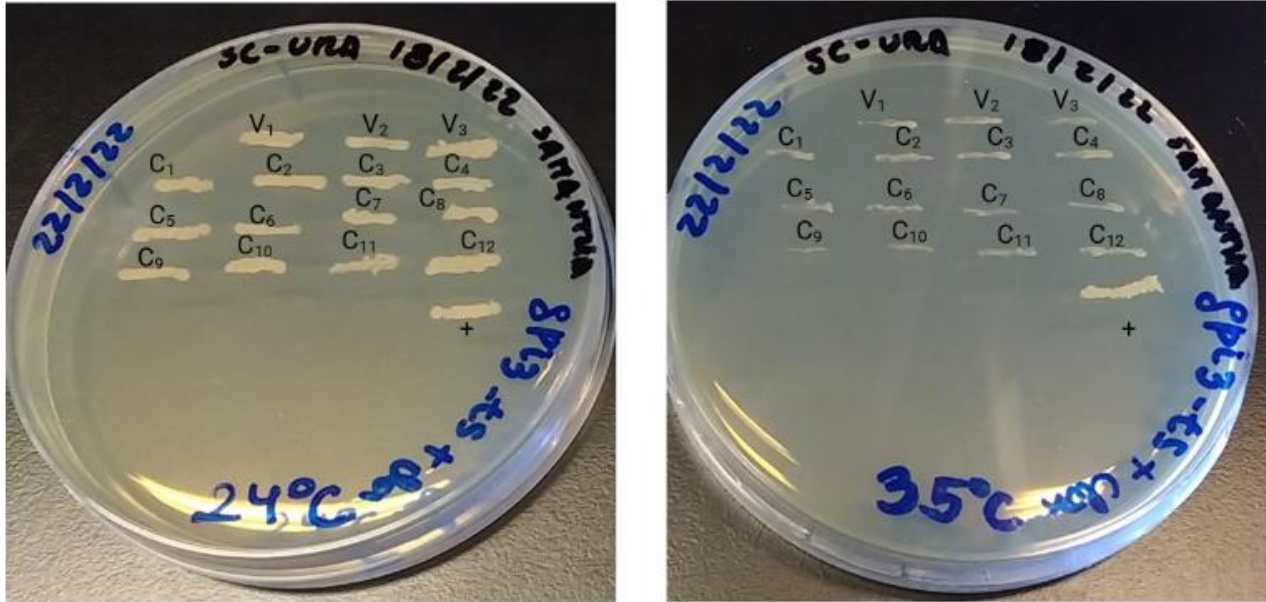


Figura 22. Resultados del ensayo de complementación con hPIGA en *S. cerevisiae* termosensibles. A la izquierda se observa la placa incubada a 24°C y a la derecha la placa correspondiente a las mismas colonias pero incubada a 35°C. Las letras sobre las estrías significan: V1, V2 y V3 son colonias transformadas con vector vacío superenrollado. Mientras que C1 a C12 son las colonias transformadas con vector e inserto, (+) es el control positivo, una cepa *wild type* de *S. cerevisiae*.

4.2.2 Estudio de la expresión de hPIGA en *S. cerevisiae* termosensibles

Se evaluó si las colonias de *S. cerevisiae* termosensibles transformadas con hPIGA expresaban la proteína mediante Western Blot. Para eso se prepararon lisados de algunas colonias aisladas positivas para la ligación según *colony* PCR (las llamadas C2, C3, C4, C5, C6 en la **Figura 21**). Además, se agregó como control negativo un lisado de una colonia transformada solamente con el vector vacío (V).

La concentración proteica total de las fracciones se determinó por el método de Bradford, utilizando una curva de calibración con albúmina de suero bovino (BSA) como estándar, de modo de sembrar en cada calle del gel la misma masa de proteína. Para revelar, se usó un anticuerpo monoclonal anti Hisx6 tag generado en ratón, y un secundario anti ratón acoplado a la enzima HRP, que permite revelar las bandas por quimioluminiscencia. Puede observarse en todos los lisados, inclusive de la transformación con vector solo, una banda de aproximadamente 55 kDa (**Figura 23**, flecha), cercana a los 56.2 kDa esperados para PIGA. Esto es indicativo de una reacción cruzada con proteínas de *S.*

cerevisiae, implicando a una proteína endógena de aproximadamente el mismo tamaño, lo que impide evaluar la expresión de hPIGA por este método.

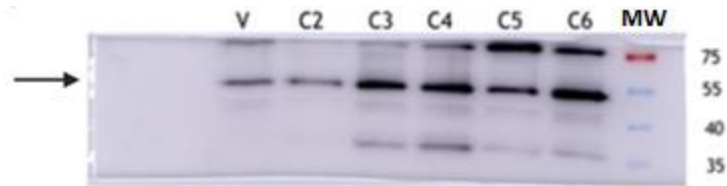


Figura 23. Western Blot de lisados de *S. cerevisiae* termosensibles GPI3 transformados con p416- hPIGA.

Los nombres en las calles indican: V: lisado de colonia transformada con vector superenrollado vacío, C2 – C6 lisados de colonias transformadas con vector e inserto (300 ng), MW: marcador de peso molecular. Nota: la calle C2 tiene sembrada un tercio de la masa de las otras muestras. El gel es un SDS-PAGE 10%.

En un intento de eliminar, o al menos atenuar, esa reacción cruzada, se realizó otro Western Blot pero esta vez, se realizó una captura de las proteínas presentes en la fracción soluble de los lisados mediante una columna de agarosa-NTA-Ni. En este caso, para cargar aproximadamente la misma masa proteica en cada calle del gel, se realizó la lisis y purificación de cada cultivo partiendo de la misma masa celular estimada por DO a 600 nm. También en esta prueba se agregó una cepa *wild type* que sirvió para comparar con las muestras transformadas, además del control negativo correspondiente a la transformación con el vector superenrollado vacío. Se sembraron las fracciones de lisado total, soluble luego de centrifugar el lisado y el eluído de la columna (**Figura 24**).

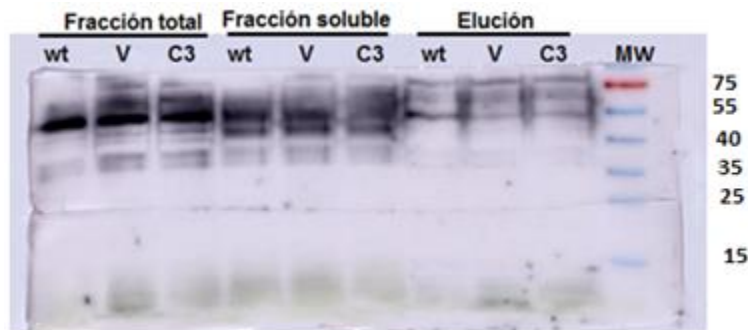


Figura 24. Western Blot de lisados de *S. cerevisiae* termosensibles transformadas con hPIGA antes y después de pasar por resinas de NTA-Ni.

wt: corresponde al lisado de la cepa *wild type*, V: lisado de la cepa termosensible transformada con vector superenrollado vacío, C3: lisado de la cepa termosensible transformada con hPIGA y el vector, MW: marcador de peso molecular. Gel SDS-PAGE 10%.

Analizando el resultado de la **Figura 24**, se observa nuevamente que la banda de aproximadamente 55 kDa aparece en todas las calles, también en la cepa *wild type*. Además, la intensidad de la banda es mucho menor en las fracciones solubles, con lo cual no se enriquece al pasar esta fracción por la resina de NTA-Ni. Persiste de todas formas una señal tenue en las eluciones de la

resina, que es comparable en las tres cepas. En definitiva, este protocolo no permitió enriquecer a PIGA en el extracto, en caso de que se expresara, por encima de las proteínas de *S. cerevisiae* lo que no permite definir si la proteína se estaba expresando efectivamente o si no se expresa y por eso no se observa una complementación de la cepa termosensible.

El resultado de colony PCR indica una ligación exitosa, sin embargo el Western Blot parece indicar la ausencia de expresión de PIGA. Existen varias posibilidades que expliquen la falta de expresión, una de las primeras cosas que se buscó verificar fue que la ligación ocurriera de la forma esperada, para eso se intentó recuperar el plásmido de las cepas termosensibles transformadas para luego secuenciarlo. Para este fin, se realizó la extracción de ADN total a partir de algunos clones de una colonia aislada que resultó positiva en la *colony* PCR (colonia C3) y se transformaron bacterias DH5 α por transformación química. A partir de las colonias obtenidas se realizaron miniprep, pero en lugar del patrón de bandas habitual de un plásmido, se obtuvieron bandas difusas y tenues (**Figura 25**). A modo de comparación, se muestra en el mismo gel la siembra de una muestra de plásmido p416 vacío obtenido por miniprep de *E. coli* transformadas con este vector. El procedimiento se repitió pero aumentando la escala (maxiprep, **panel B** de la **Figura 25**), obteniéndose nuevamente un patrón de bandas no esperable para el plásmido, por lo que las muestras no se pudieron secuenciar.

En definitiva, no se observó la complementación de la cepa termosensible por hPIGA agregada mediante el vector, pero no se llegó a explicar si esto fue resultado de que la proteína no puede expresarse adecuadamente en el sistema ensayado o que el plásmido no se generó correctamente en la ligación in vivo. Algunas posibles estrategias para avanzar en el estudio de la función de PIGA por complementación en levaduras, se desarrollan en la Discusión.

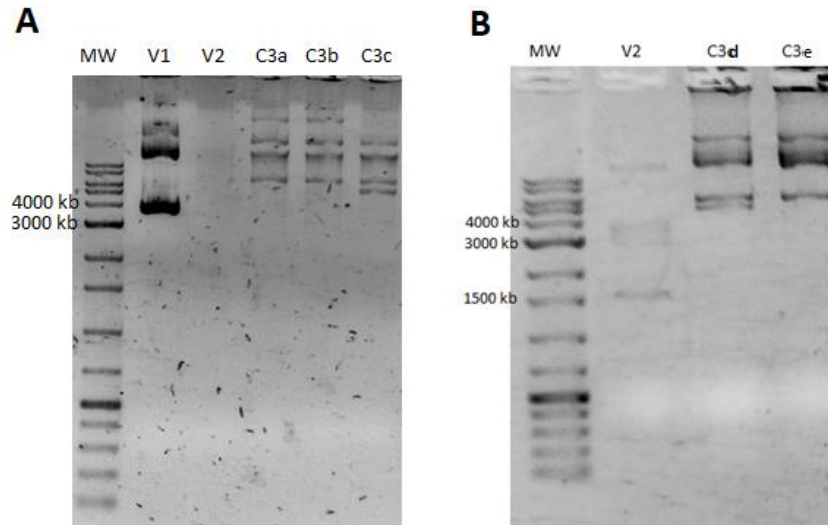


Figura 25. Recuperación del plásmido p416_hPIGA de *S. cerevisiae* termosensibles transformadas. Se realizaron miniprep (A) o maxiprep (B) de distintas colonias de *E. coli* DH5alfa transformadas con ADN total de *S. cerevisiae*. Los nombres en las calles indican: MW: marcador de peso molecular ("Ladder" 1 Kb de InbioHighway), V1: vector p416 vacío, V2: *E. coli* transformada con ADN de *S. cerevisiae* a su vez transformada con vector p416 vacío, C3: (a,b,c,d,e), son diferentes clones de *E. coli* transformados con ADN de *S. cerevisiae* a su vez transformada con vector e inserto (colonia C3 en figuras 19-23).

4.3. Tercera parte: estudio de la estructura y función de PIGA y el complejo GPI-GnT por métodos bioinformáticos

Como se mencionó en la Introducción, muy poco se sabe desde un punto de vista estructural sobre la primera etapa de biosíntesis de anclas de GPI, realizada por el complejo GPI-GnT del cual forma parte PIGA. Para conocer un poco más sobre la relación estructura-función en este complejo, y para tratar de dilucidar el impacto que tienen las mutaciones halladas en la clínica, se realizaron una serie de análisis bioinformáticos de cada una de las subunidades en diferentes organismos eucariotas. Se estudiaron los patrones de presencia/ausencia de cada uno de los genes correspondientes, como una forma de interrogar acerca de si todas las subunidades son esenciales para que se lleve a cabo la reacción química catalizada por PIGA. También, con las secuencias recolectadas se construyeron alineamientos que permitieron identificar los residuos que se encuentran conservados en cada subunidad y analizar este patrón en el contexto de los modelos estructurales predichos por AlphaFold2 y de las mutaciones halladas en la clínica.

4.3.1. Características de las secuencias de las distintas subunidades del complejo GPI-GnT

Uno de los primeros enfoques para obtener información sobre las distintas subunidades del complejo GPI-GnT es analizar las secuencias de cada una, en este caso, del complejo humano. En la **Tabla 4** se resume la información relevante sobre sus características. Como se puede observar, el rango de tamaños proteicos es muy amplio, yendo desde unas decenas de aminoácidos (PIGY y DPM2, con 71 y 84 residuos, respectivamente) a varios cientos (PIGA y PIGQ con 484 y 760 residuos, respectivamente). Una de las características más notables de estas secuencias es la alta proporción de residuos hidrofóbicos, que llegan a valores >60% en los casos de PIGY y DPM2, algo que también queda capturado en su índice de hidropaticidad (GRAVY), calculado como el valor promedio de hidrofobicidad según la escala de Kyte & Doolittle (Kyte & Doolittle, 1982). En este caso, los valores positivos de todas las proteínas indican que son fundamentalmente hidrofóbicas, siendo DPM2 la de mayor índice.

También se observa que algunas de las subunidades presentan varios residuos de cisteína. Esto es particularmente notorio en PIGH, cuyo contenido de cisteínas (4.8%) es más de dos veces superior al del promedio de las proteínas humanas (2.26%) (Miseta & Csutora, 2000).

Como se mencionó anteriormente, el complejo GPI-GnT está anclado a la membrana del RE, lo que podría lograrse mediante regiones transmembrana de sólo una o algunas subunidades. Sin embargo, en este caso se observa que todas las subunidades tienen como mínimo un paso transmembrana, según lo calculado con el servidor Protter (Omasits et al., 2014).

	PIGA	PIGC	PIGH	PIGQ	PIGP	PIGY	DPM2
N° aminoácidos	484	297	188	760	158	71	84
Residuos hidrofóbicos	242 (50%)	171 (57.58%)	92 (48.94%)	447 (58.82%)	90 (56.96%)	44 (61.97%)	56 (66.67%)
Índice de hidropaticidad (GRAVY)	0.060	0.476	0.296	0.162	0.305	0.918	1.042
N° de cisteínas	12 (2.5%)	6 (2%)	9 (4.8%)	24 (3.2%)	0 (0%)	2 (2.8%)	0 (0%)
N° pasos transmembrana	1	8	2	5	2	2	2

Tabla 4. Características fisicoquímicas de las proteínas que forman el complejo GPI-GnT (Gasteiger et al., 2005; Omasits et al., 2014; *Protein Parameters*, s. f.)

4.3.2. Conservación del complejo GPI-GnT en eucariotas

Como se mencionó en la Introducción, todas las subunidades del complejo son importantes para la función biológica, lo que se refleja en que la deficiencia de cualquiera de ellas ocasiona alguna patología descrita. De todas formas, puede que sólo algunas de las subunidades cumplan un rol esencial, mientras que otras tengan un rol accesorio, ajustando los niveles de biosíntesis de anclas GPI a las necesidades celulares. Con el objetivo de investigar si todas las subunidades que integran el complejo GPI-GnT son esenciales, se estudió el perfil de presencia/ausencia de los genes correspondientes en los genomas de distintos organismos eucariotas representativos. Partiendo de la variante humana de cada subunidad, se realizaron búsquedas con HMMER o BlastP en la base de datos de secuencias del NCBI. En los casos donde la subunidad estaba ausente, se repitieron las búsquedas con tBlastN, que hace una búsqueda en los genomas traduciéndolos en los 6 marcos de lectura posibles. Esto sirvió para verificar que la ausencia no sea debido a que no estén anotados los ORF. Los resultados se integraron en un árbol filogenético y en una tabla (**Figura 26** y **Tabla 5**, respectivamente).

Examinando la **Tabla 5**, se observa que PIGA está presente en casi la totalidad de los organismos analizados (94% de Eukarya). Se debe tener en cuenta que el conjunto de organismos incluye especies cercanas a la raíz de Eukarya, como Trypanosoma y Leishmania pertenecientes a Discoba. Esto es consistente con una aparición temprana de esta vía de biosíntesis en eucariotas. Por el contrario, PIGY es la subunidad menos representada, presente en sólo un 31% del conjunto de organismos. Debería tenerse en cuenta que para el caso de PIGY, el 60% de sus 71 aminoácidos corresponden a hélices transmembrana, que por tratarse de regiones de baja complejidad pueden resultar en una menor sensibilidad en la detección de ortólogos. Sin embargo, DPM2 también es una proteína pequeña (84 aminoácidos) con el 50% de su secuencia predicha como transmembrana, y pudo detectarse en un porcentaje mayor de organismos (50%, en comparación con el 31% de PIGY), con lo cual es posible que efectivamente PIGY sea la subunidad más prescindible.

Notablemente hay 8 organismos donde PIGA no fue encontrada (**Figura 26**). Estos organismos son *Perkinsella sp.*, *Trichomonas vaginalis*, *Galdieria sulphuraria*, *Emiliania huxleyi*, *Aereococcus anophagefferens*, *Vavraia culicis*, *Thelohanellus kitauei*, *Cyanidioschyzon merolae*. Para asegurar que este resultado no correspondiera a una deficiencia en la base de datos del NCBI, se realizaron búsquedas *ad hoc* en los genomas publicados en Ensembl o PhycoCosm para estos organismos, empleando tanto BlastP como tBlastN, no encontrándose hits significativos (algunas búsquedas arrojaron hits donde unos pocos aminoácidos exhibían similitud, distribuidos en segmentos de aproximadamente un 10% de la longitud de PIGA, en genes hipotéticos anotados como otras glicosiltransferasas). Cuando PIGA está ausente, el resto de los componentes del complejo también lo está, reforzando la idea de que es una verdadera ausencia del complejo y no un problema de búsqueda. Estos organismos no están relacionados filogenéticamente, algunos son parásitos intracelulares obligados de genomas muy reducidos, y en general son todos unicelulares. En resumen, PIGA es una proteína conservada en todo el dominio eucariota y unos pocos organismos han perdido este gen de manera independiente. Para estos últimos se encontró que algunos no presentan anclas de GPI, perdieron la capacidad de producir las durante su evolución o poseen la vía de síntesis reducida (Desjardins et al., 2015; Noël et al., 2010; Van Etten et al., 2023).

En cuanto a la distribución taxonómica, analizando la **Figura 26**, se puede observar que hay organismos que presentan todas las subunidades a la vez, como es el caso de la mayoría de los animales (particularmente mamíferos). Se observa que PIGY es la subunidad menos presente dentro de los Ascomycota y Basidiomycota (al igual que lo que ocurre con el Reino Vegetal). Según puede verse en la **Tabla 5**, en Microsporidia están ausentes la mayoría de las subunidades (uno de los organismos listados carece de todo el complejo y los dos restantes poseen PIGA acompañada de PIGC), en línea con que estos organismos han experimentado eventos masivos de reducción y compactación de su genoma.

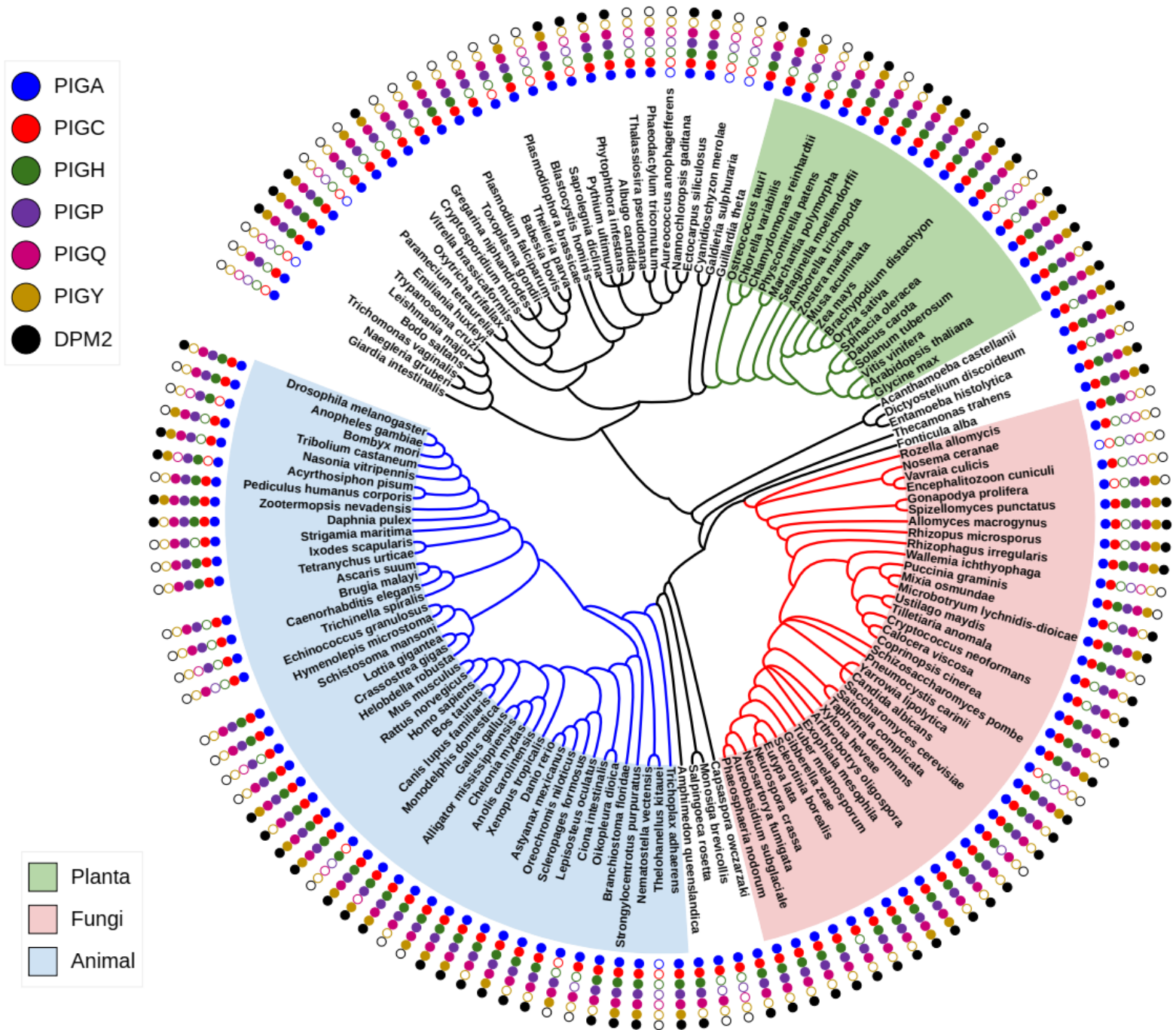


Figura 26. Representación de todas las subunidades en árbol filogenético. Se resaltan algunos grupos de organismos en color. Los círculos alrededor de árbol representan las distintas subunidades según el código de colores que se puede ver en el recuadro a la izquierda del árbol. Filogenia recuperada de Open Tree of Life. Figura generada con iTOL.

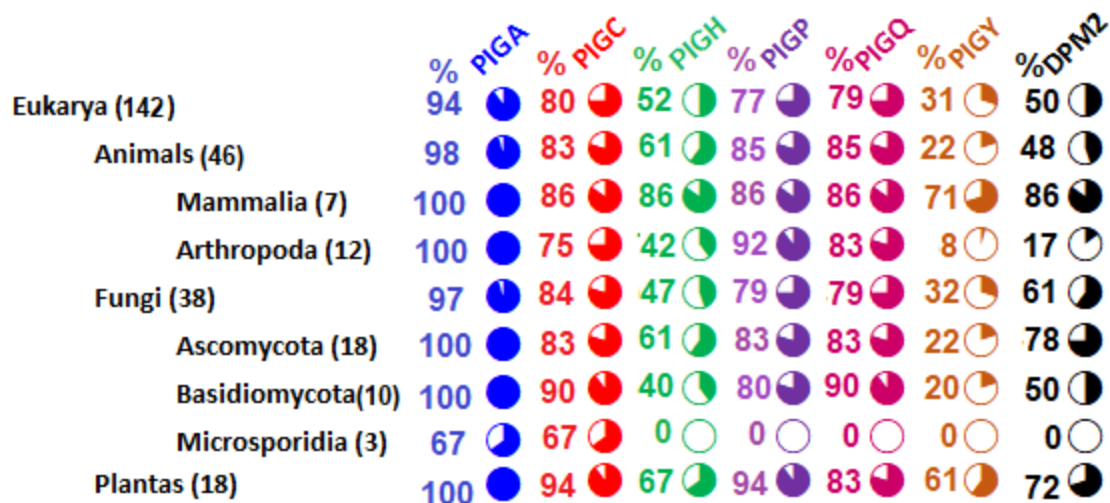


Tabla 5. Presencia de las subunidades del complejo GPI-GnT en los distintos organismos. Los colores de cada subunidad se corresponden con los de la Figura 24. Se muestra como pie chart el porcentaje de presencia de cada subunidad en cada grupo de organismos. El número entre paréntesis indica cantidad de organismos de esa clase.

Utilizando la información sobre la distribución de las subunidades en los distintos organismos, se buscó inferir cuál podría ser el núcleo funcional evolutivamente conservado del complejo. Para eso, se contó el número de veces en que dos subunidades están ambas presentes o cuando sólo está una de ellas (**Tabla 6**). Se observa que cuando PIGA está presente también se encuentran, para la mayoría de los casos, las subunidades PIGC, PIGP y PIGQ (con coexistencias aproximadas de 80 %). Si se analiza además la coexistencia entre estas tres últimas también se observan porcentajes altos (aproximadamente 70%), lo cual sugiere que PIGQ, PIGP y PIGC serían las subunidades auxiliares más importantes.

Salvo unas pocas excepciones, PIGH no es detectada en ausencia de PIGC, PIGP y PIGQ. También se obtuvo una muy baja presencia de PIGY en ausencia de estas 3 subunidades (**Tabla 6**), algo semejante ocurre en el caso de DPM2. Este último resultado es notable, ya que como se mencionó en la Introducción, DPM2 está compartida entre dos rutas de biosíntesis, pues también participa en la formación de dolicolfosfato-manosa (DPM). En esta vía, el DPM es sintetizado por un complejo que incluye 3 subunidades, en las cuales DPM1 es la catalítica, DPM3 es responsable de la localización de DPM1 en el RE, y DPM2 aparentemente estabiliza el complejo (Maeda et al., 2000). Esto plantea el interrogante de si evolutivamente DPM2 fue seleccionada inicialmente como una subunidad regulatoria para la ruta de biosíntesis de GPI, para luego formar parte de la vía de síntesis de DPM o viceversa. Para responder a esta pregunta son necesarios más análisis, como por ejemplo una comparación entre los perfiles filogenéticos de los complejos GPI-GnT y el de síntesis de DPM.

Los resultados obtenidos hasta ahora permiten postular una jerarquía en la composición de subunidades del complejo GPI-GnT, en la cual la subunidad catalítica PIGA es asistida primariamente por las subunidades PIGC, PIGP y PIGQ (coexistencias con PIGA de aproximadamente 80%). PIGH se presenta como la siguiente subunidad auxiliar (coexistencias con PIGA de 52%) y luego DPM2 (coexistencias con PIGA de 50%). Finalmente, la biosíntesis de GPI parece alcanzarse en niveles adecuados para las necesidades metabólicas sin necesidad de PIGY en gran parte de los organismos.

Sub 1	Sub 2	Ambas (%)	Solo Sub 1 (%)	Solo Sub 2 (%)	Ninguna (%)
PIGA	PIGC	79,6	14,8	0,0	5,6
	PIGH	52,1	42,3	0,0	5,6
	PIGP	77,5	16,9	0,0	5,6
	PIGQ	78,9	15,5	0,0	5,6
	PIGY	31,0	63,4	0,0	5,6
	DPM2	50,0	44,4	0,0	5,6
PIGC	PIGH	50,0	29,6	2,1	18,3
	PIGP	71,8	7,7	5,6	14,8
	PIGQ	73,2	6,3	5,6	14,8
	PIGY	30,3	49,3	0,7	19,7
	DPM2	47,9	31,7	2,1	18,3
PIGH	PIGP	51,4	0,7	26,1	21,8
	PIGQ	50,0	2,1	28,9	19,0
	PIGY	23,9	28,2	7,0	40,8
	DPM2	38,7	13,4	11,3	36,6
PIGP	PIGQ	72,5	4,9	6,3	16,2
	PIGY	31,0	46,5	0,0	22,5
	DPM2	50,0	27,5	0,0	22,5
PIGQ	PIGY	29,6	49,3	1,4	19,7
	DPM2	48,6	30,3	1,4	19,7
PIGY	DPM2	23,9	7,0	26,1	43,0

Tabla 6. Perfilado filogenético de las subunidades del complejo GPI-GnT

4.3.3. Estructura del complejo GPI-GnT

Para entender un poco más acerca del mecanismo de funcionamiento del complejo GPI-GnT, realizamos un análisis de las estructuras tridimensionales de las subunidades. Debido a que ninguna de ellas presenta una estructura determinada experimentalmente, se utilizaron los modelos predichos por el servidor AlphaFold2. Este programa se basa en la técnica de *Deep Learning* y utiliza un alineamiento múltiple de secuencias de proteínas homólogas en la predicción de contactos intramoleculares, que luego son usados para construir un modelo tridimensional. Además, para cada residuo calcula un valor de confianza (pLDDT): valores bajos de pLDDT indican regiones presentes en pocos homólogos, lo que genera una menor fiabilidad de los modelos, mientras que valores altos de pLDDT sugieren lo contrario (Jumper et al., 2021).

En este análisis, se buscó responder las siguientes preguntas:

1) ¿Las distintas subunidades presentan estructura estable? En ese caso: ¿Existen dominios globulares?

2) Como resultado de la formación de estructura: ¿Se forman regiones con propiedades fisicoquímicas particulares, como superficies hidrofóbicas o cargadas, que podrían ser importantes para la interacción proteína-proteína?

3) Usando la información anterior: ¿Es posible construir un modelo del complejo entero? ¿Puede modelarse la unión de los ligandos?

Propiedades estructurales de las subunidades

Las secuencias de cada subunidad fueron analizadas en estudios previos proponiéndose una descripción topológica para cada una, incluyendo las regiones transmembrana y la orientación de los extremos en el RE (Eisenhaber et al., 2003; Levin & Stamper, 2009). Aquí se profundizó el análisis estructural de estas proteínas, al analizar sus modelos de alta resolución, para lo cual se recurrió a distintas herramientas que revelaron muchas características nuevas que se describen a continuación.

En el caso de **PIGA (Figura 27)**, se observa en su estructura predicha por AlphaFold2 que la mayor parte de la secuencia adquiere un plegado globular correspondiente al dominio citoplasmático, el cual se encuentra separado del dominio luminal por una única hélice (transmembrana). El dominio globular está formado por dos plegados tipo Rossmann consecutivos y enfrentados, consistente con que esta enzima ha sido clasificada como una glicosiltransferasa-B (Cruz et al., 2019). En medio de estos dos subdominios Rossmann hay una cavidad que, por comparación con otras enzimas análogas, se correspondería con el sitio activo (ver sección “Modelado de unión del ligando al sitio activo de PIGA”).

La orientación relativa de la hélice transmembrana resulta incorrecta, ya que se encuentra posicionada sobre el dominio globular de la enzima, en lugar de sobresalir para insertarse en la membrana. Por otro lado, el extremo N se observa desorganizado, lo que es consistente con la predicción de desorden realizada con el programa NetSurfP3 (Høie et al., 2022) (**Suplementaria 8**, ANEXO). Este segmento está formado mayormente por residuos polares (aprox. 50%) así como glicinas y prolinas (aprox. 20%), por lo que es esperable su baja propensión a formar estructura estable. Además, allí se encuentran 2 residuos de cisteína, los cuales podrían encontrarse oxidados formando un enlace disulfuro, aunque no hay evidencia de ello en el modelo predicho por AlphaFold2. La porción luminal es predicha como una hélice alfa, aunque con baja confiabilidad ($pLDDT < 50$).

El grado de conservación de secuencias de PIGA (al igual que el de las demás subunidades) se evaluó en el contexto del modelo estructural, mediante el cálculo de la tasa de evolución con el servidor ConSurf, usando alineamientos múltiples de secuencia cuidadosamente construidos y curados a mano. La mayor parte de la proteína está muy conservada (**Figura 27**), las regiones más variables se encuentran en los extremos N y C (**Suplementaria 1**, ANEXO). En el caso del extremo N, el cálculo de conservación no se realizó, debido a que su alta variabilidad impedía realizar un alineamiento confiable. Como podría anticiparse, la región más conservada se ubica en la interfaz entre los dos subdominios, correspondiente al sitio activo.

En cuanto a **PIGC** la mayor parte de la secuencia corresponde a las 8 hélices transmembrana, donde hay gran abundancia de residuos hidrofóbicos, como se espera (**Figura 27**). Cuando la estructura es analizada con el servidor PPM 3.0 (Lomize et al., 2022) para orientarla en una membrana simulada con parámetros específicos para el RE, se observa que esencialmente todo ese dominio de hélices se encuentra embebido en la membrana, sobresaliendo hacia el citosol una pequeña porción entre las hélices 2 y 3 (predicha con estructura helicoidal) y los extremos N y C, y hacia el lumen sólo un residuo cargado (**Figura 28**). Los extremos amino y carboxilo se encuentran desorganizados. Estos extremos están compuestos principalmente por residuos polares, muchos de ellos cargados, que podrían ser importantes para la interacción entre subunidades. La proteína posee 6 residuos de cisteína, de los cuales 5 se encuentran en la región transmembrana y uno en la región luminal. Según este modelo estructural las cisteínas ubicadas en hélices diferentes no se encuentran próximas entre sí, por lo que no tendrían un rol en la estabilización del arreglo de hélices en la membrana. Por otro lado, si se analiza la conservación de la estructura (**Figura 27**), en el dominio transmembrana compuesto por las hélices se distinguen regiones más conservadas y algunas otras más variables (**Suplementaria 2**, ANEXO). El extremo amino está muy poco conservado.

La estructura de **PIGH** (**Figura 27**) presenta dos hélices correspondientes a la región transmembrana, cuyos ejes son aproximadamente paralelos, y los extremos N y C de la proteína se integran en un único dominio globular formado por una lámina beta antiparalela de 6 hebras, sobre la que apoya una hélice alfa del extremo C y una lámina beta menor, de tres hebras. Este tipo de plegado resultaría novedoso, ya que la búsqueda de análogos estructurales mediante el programa Dali (Holm & Sander, 1995) no reveló ninguna estructura experimentalmente determinada con el mismo plegado. Esta proteína presenta 6 residuos de cisteína, pero en el modelo se encuentran todos separados a distancias mayores a los 10 Å, por lo que no es posible la formación de enlaces disulfuro sin realizar grandes alteraciones al modelo. En cuanto a su grado de conservación, la lámina beta central está altamente conservada, mientras que la lámina beta menor y la hélice lo están en menor medida. Pueden observarse diversos “parches” hidrofóbicos en su superficie (**Suplementaria 3 B**, ANEXO), los cuales podrían ser importantes para la interacción proteína-proteína en el complejo.

El modelo estructural de **PIGP** presenta 2 hélices alfa centrales, que por su composición aminoacídica dominada por residuos hidrofóbicos, son los dos pasos transmembrana. Los dos extremos presentan segmentos helicoidales y regiones desestructuradas. Las hélices transmembrana corresponden a la región más confiable del modelo ($pDDT > 90$) y se observa que las mismas no están interactuando, ya que el ángulo entre las hélices es de aproximadamente 45° (**Figura 27**). Para corroborar esto, se analizó otro modelo realizado con el programa ESMFold (Lin et al., 2023), que también es un método basado en Deep Learning aunque no necesita de alineamientos de secuencias homólogas. Todos los modelos coinciden en la posición relativa de las hélices transmembrana (ángulo entre las hélices), aunque el posicionamiento de los extremos N y C difiere. La proteína contiene una gran cantidad de leucinas distribuidas a lo largo de la secuencia, 3 de ellas se ubican en el extremo N con un motivo LxxLxxL, lo que las posiciona alineadas en la misma cara de la hélice, y sugiere que dichos residuos podrían formar interacciones específicas con otra proteína interactuante. Analizando la conservación de secuencia en la estructura, se observa que las hélices están compuestas por residuos variables (**Suplementaria 4**, ANEXO). En general hay pocos residuos conservados, sobre todo en el extremo C, donde el cálculo de conservación no se realizó debido a su alta variabilidad, que impedía realizar un alineamiento confiable.

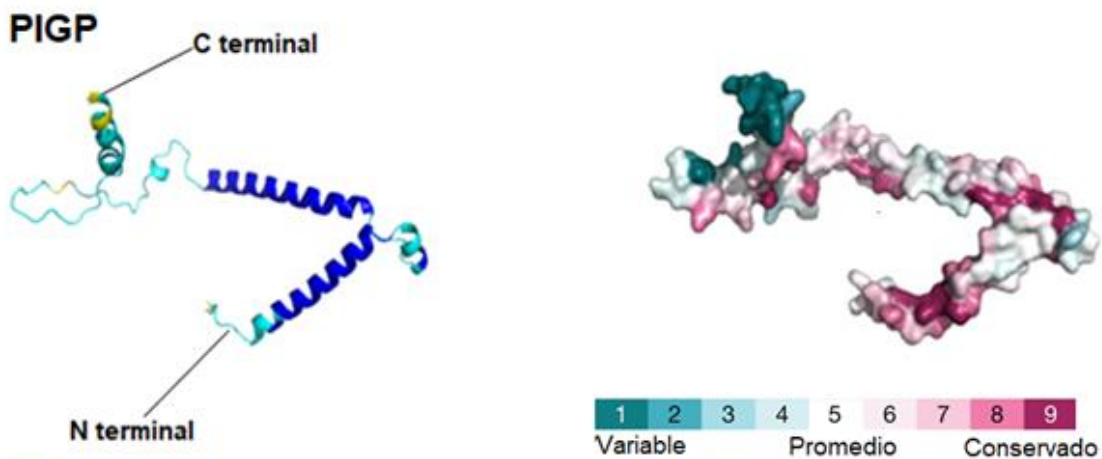
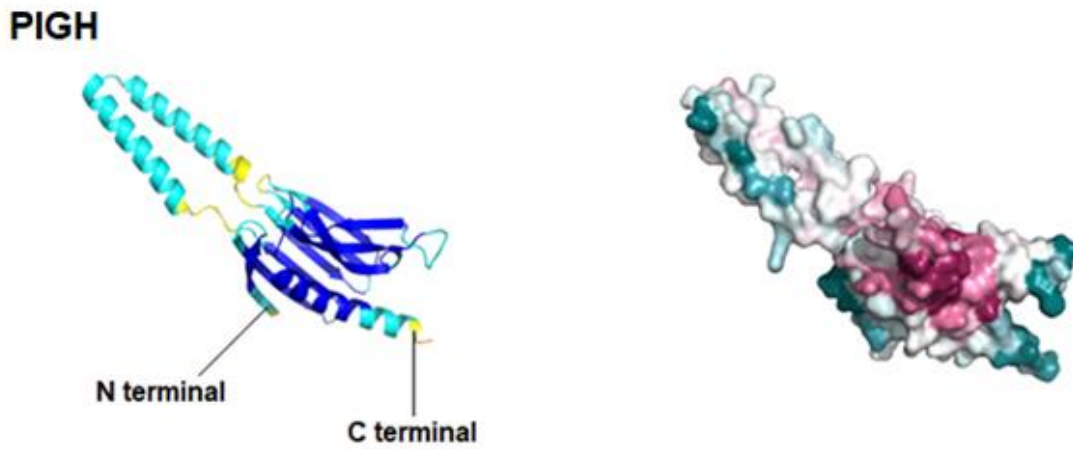
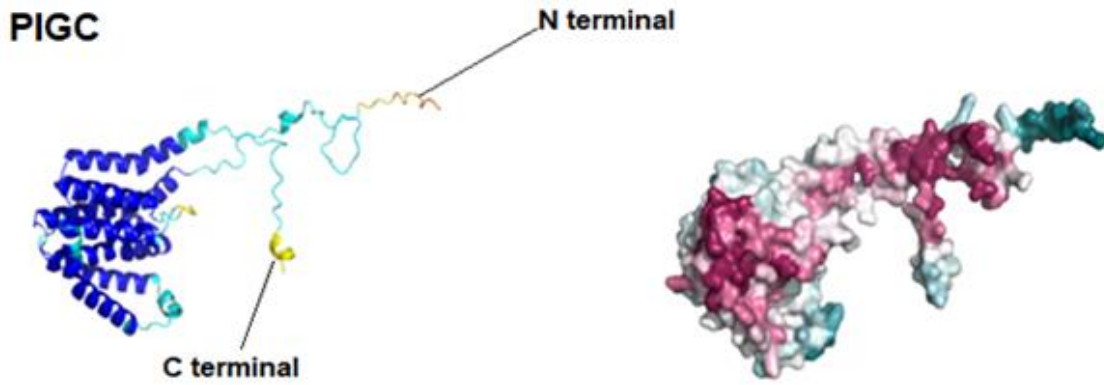
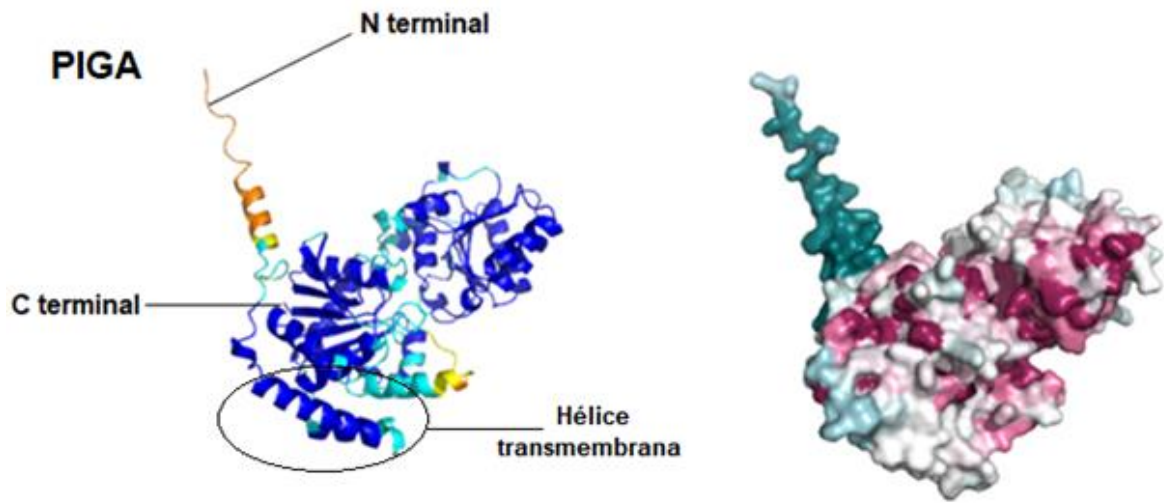
PIGQ es la subunidad más grande del complejo y su estructura predicha presenta cierta complejidad (**Figura 27**). Las características más notables son: un arreglo de hélices, que incluye a la región transmembrana, y un extremo N dominado por una lámina beta altamente curvada, casi formando un barril. Se ha postulado que PIGQ posee 4 hélices transmembrana y otra hélice completamente embebida en la membrana, más hacia el extremo N relativo a las otras (Levin & Stamper, 2009). Al

intentar calcular el modo de inserción de PIGQ, usando el servidor PPM y el modelo de AlfaFold2, del cual se preservó únicamente el arreglo de hélices (residuos 191-506) para remover las regiones mal modeladas, se obtuvo como resultado un modelo donde la mayor parte de las hélices se encuentran completamente embebidas entre ambas capas de la membrana disponiéndose de manera paralela (**Figura 28**). Esta disposición estaría en conflicto con lo postulado en bibliografía y no es una forma común en la que se ubica un paso transmembrana. Como una forma de validar este resultado, se realizó una búsqueda de análogos estructurales del arreglo de hélices con el programa DaliLite. La idea era identificar algún análogo estructural en la base de datos PDB, y en caso de tratarse de una proteína de membrana, examinar su interacción con la membrana ya sea a partir de la propia estructura experimental (estructura determinada en un sistema de membranas) o por bibliografía. Lamentablemente, la búsqueda no arrojó ningún resultado significativo. Por otro lado, la estructura es principalmente hidrofóbica (~59%), conteniendo una alta cantidad de leucinas. Para el dominio del extremo N sí se encontró similitud estructural parcial en la disposición de la lámina beta central, con proteínas involucradas en la ubiquitinación (ejemplos: PDBID 6FIJU y 1OIO) y que por lo tanto no es claro si tienen relación con PIGQ. Hacia el extremo carboxilo la estructura se vuelve muy desordenada, aunque la predicción es de baja confiabilidad ($pLDDT < 50$). Las hélices transmembrana, así como el dominio del extremo N, se encuentran muy conservados (**Figura 27** y **Suplementaria 5**, ANEXO), pero el extremo C es demasiado variable y no se pudo alinear con confianza, por lo que se omitió del cálculo de conservación.

PIGY (Figura 27) está formada por dos alfa-hélices aproximadamente paralelas que forman una estructura parecida a un *hairpin* al igual que PIGP. Los extremos amino y carboxilo están orientados hacia el citoplasma, dejando solo una pequeña región entre las dos hélices orientada hacia el lumen. La hélice correspondiente al extremo carboxilo posee un *kink*, y en esa región se encuentran dos prolinas, las cuales son disruptoras de hélices y provocan que la misma se tuerza. Para analizar este detalle se examinó otro modelo predicho por ESMFold (no mostrado), en este caso también se encuentra el *kink* (se produce en los mismos residuos de prolinas) pero la hélice se tuerce en la dirección contraria respecto a la de AlphaFold2. Por otro lado, en la estructura predicha por AlphaFold2 se observa una pequeña hélice en la región que da hacia el lumen que no se observa en el modelo de ESMFold. La estructura está compuesta principalmente por aminoácidos hidrofóbicos, lo cual tiene sentido ya que es una proteína pequeña que presenta la mayor parte de su estructura insertada en la membrana. Tiene dos cisteínas, las cuales se encuentran muy próximas entre sí, formando un puente disulfuro. La estructura se encuentra bastante conservada (**Figura 27**), en particular el segmento que da hacia el

lumen del RE y el extremo carboxilo, y además las prolinas que conforman el *kink* están muy conservadas (**Suplementaria 6**, ANEXO).

Por ultimo **DPM2** tiene una estructura topológicamente semejante a la de PIGY, formada por dos hélices alfa cuyos ejes no son paralelos, sino que se encuentran formando un ángulo de 45 grados (**Figura 27**). Los dos extremos están orientados hacia el citoplasma con un pequeño *loop* que apunta hacia el lumen. Mediante el servidor PPM 3.0 se calculó el modo de inserción de la estructura en la membrana, encontrándose la mayor parte de la proteína embebida en la misma, con los dos extremos sobresaliendo hacia el citoplasma y la región entre las hélices que sobresale hacia el lumen del RE (**Figura 28**). En general, la estructura está muy poco conservada (**Suplementaria 7 A**, ANEXO).



■ Muy alta (pLDDT>90)
■ Confiable (90>pLDDT>70)
■ Baja (70>pLDDT>50)
■ Muy baja (pLDDT<50)

1	2	3	4	5	6	7	8	9
Variable			Promedio			Conservado		

Información insuficiente

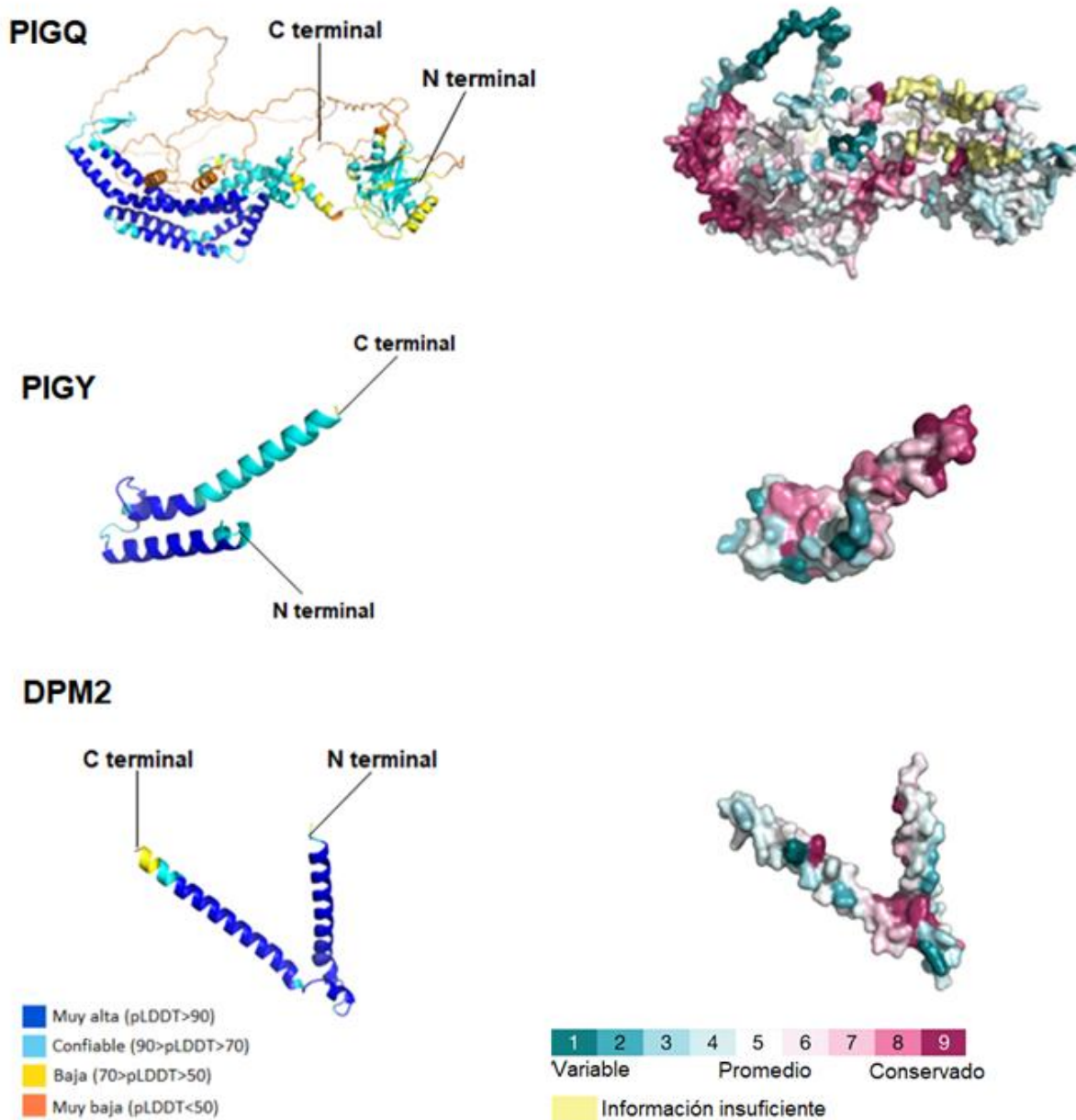


Figura 27. Estructuras predichas de las subunidades del complejo GPI-GnT y mapeo de la conservación de secuencia (comienza en la página anterior). Se muestran a la izquierda las estructuras predichas para cada subunidad por AlphaFold2 con la escala de colores que indica el grado de confiabilidad. A la derecha se muestra el grado de conservación de las secuencias correspondientes según ConSurf (con la escala de conservación de los residuos en colores). En el caso de FIGA y FIGP se eliminaron de la estructura los extremos N y C respectivamente, ya que son regiones muy variables y el cálculo de conservación no se pudo realizar en estos casos.

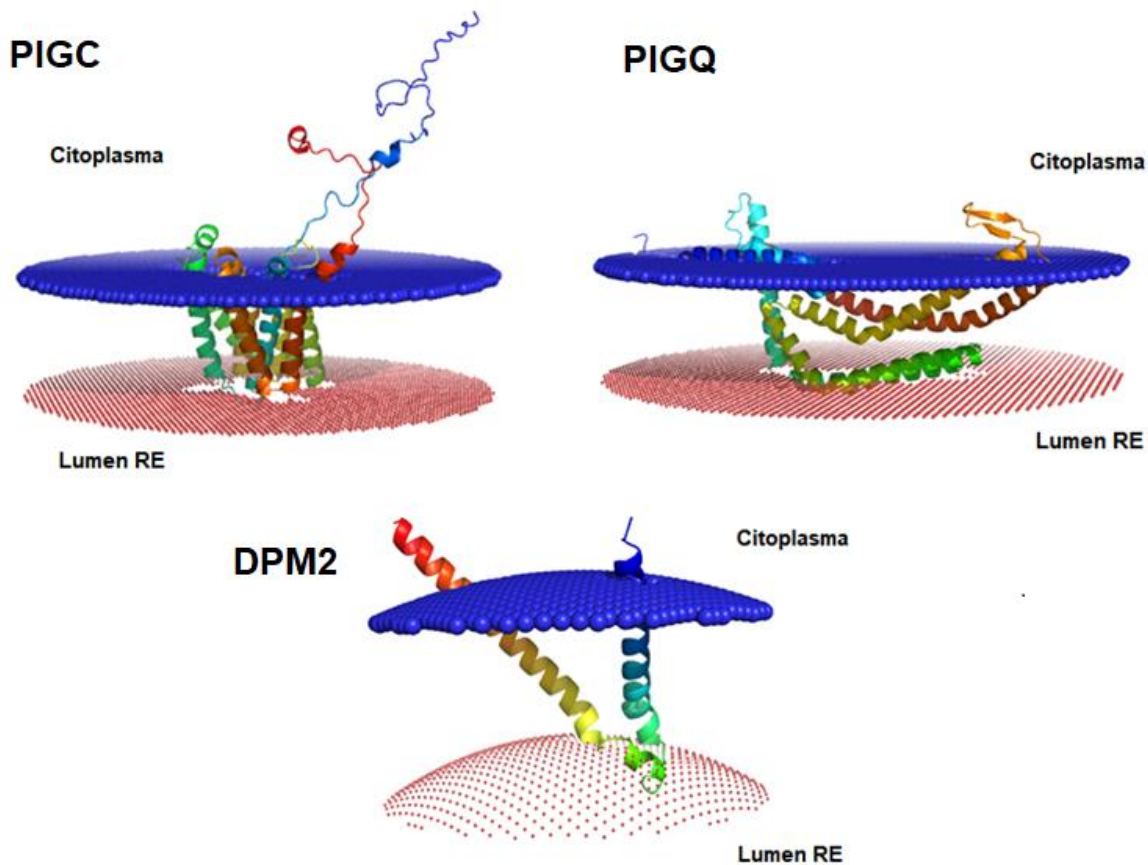


Figura 28. Disposición de algunas subunidades en la membrana del RE. A la izquierda: PIGC, a la derecha: Arreglo espacial de PIGQ (hélices, residuos 191-506), abajo: Arreglo espacial de DPM2 en la membrana del RE. Esquemas realizados con el servidor PPM 3.0.

4.3.4. Mutaciones en las subunidades del complejo GPI-GnT y sus implicancias clínicas

Las anclas de GPI son esenciales para la embriogénesis, el desarrollo, la neurogénesis, la fecundación y para el sistema inmune de los mamíferos (Kinoshita & Fujita, 2016). Como se ha mencionado en la Introducción, se reportaron enfermedades genéticas relacionadas a mutaciones de pérdida de función en la mayoría de los genes involucrados en la biosíntesis de las anclas de GPI. Estas enfermedades/síndromes, son llamadas en su conjunto “deficiencia heredada de GPI” o IGD por sus siglas en inglés (Carmody et al., 2020; Kinoshita, 2016). Aquellas que se producen debido a mutaciones en los genes que codifican para las subunidades del complejo GPI-GnT se muestran en la **Tabla 7** y se describen en más detalle en el “Anexo”. Los principales síntomas incluyen: discapacidad intelectual, epilepsia, rasgos faciales toscos y anomalías en múltiples órganos. Estos síntomas varían en gravedad dependiendo del grado del defecto y/o de la posición en la vía del gen afectado (Murakami & Kinoshita, 2015).

Para esta tesis, se analizaron las mutaciones clínicas en el complejo GPI-GnT desde una perspectiva estructural. Se recurrió a los datos de mutaciones halladas en pacientes que se reportaron en la base de datos Clinvar, donde se encuentran las variantes genéticas humanas e interpretaciones de su importancia clínica (Landrum et al., 2020). Las variantes se clasifican de acuerdo a su importancia clínica según: patogénicas, probablemente patogénicas, interpretaciones conflictivas de patogenicidad, significancia incierta, benignas o probablemente benignas. Estos datos se complementaron con variantes clínicas obtenidas de otras bases de datos, en particular de Human Gene Mutation Database (Stenson et al., 2020), Uniprot (The UniProt Consortium, 2023) y VarSite (Laskowski et al., 2020).

Gen	Enfermedad/síndrome asociado	Herencia
PIGA	Síndrome de anomalías congénitas múltiples - hipotonía - convulsiones tipo 2 (MCAHS 2) Hemoglobinuria Paroxística Nocturna (HPN) Trastorno del Neurodesarrollo con Epilepsia y Hemocromatosis (NEDEPH)	Recesivo ligado al X
PIGC	Defecto en la biosíntesis de GPI 16 (retraso global del desarrollo y trastorno convulsivo)	Autosómico recesivo
PIGH	Defecto en la biosíntesis de GPI 17 (retraso en el desarrollo, epilepsia y microcefalia)	Autosómico recesivo
PIGP	Encefalopatía Epiléptica Infantil Temprana 55 (EIEE 55)	Autosómico recesivo
PIGQ	Encefalopatía Epiléptica Infantil Temprana 77 (EIEE 77)	Autosómico recesivo
PIGY	Síndrome de Hiperfosfatasa con Retraso Mental 6 (HPMRS 6)	Autosómico recesivo
DPM2	Trastorno Complejo de la Glicosilación Tipo Ie	Autosómico recesivo

Tabla 7. Enfermedades/síndromes asociados a los genes que codifican para las subunidades del complejo GPI-GnT(EIEE77 - ClinVar - NCBI, s. f.; *Multiple Congenital Anomalies-Hypotonia-Seizures Syndrome 4 disease: Malacards - Research Articles, Drugs, Genes, Clinical Trials*, s. f.; Swoboda et al., 2014). Adaptado de Carmody et al., 2020.

Se estudiaron las mutaciones que producen sustituciones (*missense*) y que son clasificadas como patogénicas o probablemente patogénicas (**Figura 29**). Los residuos intervinientes fueron mapeados en los modelos estructurales para analizar si las mutaciones se ubican en regiones variables o conservadas, si se afectan interacciones intramoleculares que podrían alterar la estabilidad o en la superficie proteica, lo que podría afectar a las interacciones intermoleculares (**Figura 30**).

En primer lugar, se estudió a PIGA, para la cual existe el mayor número de variantes clínicas reportadas. La mayor parte de las mutaciones se ubican en el dominio que da al citoplasma, mientras que sólo una mutación se localiza en la región del lumen. Puede notarse (**Figura 30**) que del dominio citoplasmático la mayoría de las mutaciones (70%) se encuentran en el subdominio que está más cercano a la membrana, aunque este subdominio es sólo un 20% más largo que el otro. Algunas posiciones mutadas se localizan en la interfaz entre los subdominios: V49, G48, S155 e I206, que se encuentran en las inmediaciones del sitio activo, por lo que es esperable que su efecto fenotípico esté

relacionado con deficiencias en la unión de sustratos o eficiencia catalítica. Para investigar si las posiciones mutadas están conservadas entre las secuencias de distintos organismos eucariotas, se construyó un alineamiento múltiple de secuencias y se calculó la conservación en forma de entropía, la cual es mayor cuanto más variable es la secuencia. Como se observa en la **Figura 31**, la mayor parte (74%) de las posiciones mutadas corresponden a regiones conservadas. Además, de las 38 posiciones mutadas, 10 se encuentran ocultas al solvente (SASA relativa < 75%), por lo que es esperable que con la mutación se altere el empaquetamiento de la proteína resultando en su desestabilización. Asimismo, el hecho de que la mayor parte de dichas mutaciones se encuentren expuestas, podría implicar que su impacto en el fenotipo resulta de la alteración de interacciones intermoleculares con otras subunidades del complejo.

En cuanto a PIGC, sólo hay reportadas 3 variantes patogénicas que se localizan en las hélices transmembrana, aunque una de ellas (L220) se encuentra justo en el límite de la membrana en la cara luminal (**Figura 30**). En la variante silvestre, los residuos correspondientes a estas tres posiciones se orientan hacia el interior del arreglo de hélices, por lo que su impacto en el fenotipo podría ser debido a una desestabilización de la estructura.

Por otro lado, para PIGH se ha informado sólo una variante patogénica, que involucra la sustitución de un residuo de serina medianamente conservado y oculto al solvente, localizado sobre la hoja beta de la región citoplasmática (**Figura 30**). Dicha serina se encuentra mutada a prolina, que es incompatible con una geometría beta ideal (Das & Basu, 2012), por lo cual el fenotipo patológico de esta mutación podría ser resultado de una desestabilización de la estructura.

Para los casos de PIGP, PIGQ y PIGY, también hay reportada una sola variante patológica para cada subunidad. Para PIGP, la mutación involucra un residuo conservado que se localiza en el lumen, en una región bisagra entre 2 hélices transmembrana. En PIGQ se trata una posición en la región transmembrana de un residuo muy oculto (3% SASA relativa) correspondiente a una glicina conservada que cambia a arginina. La introducción de una carga en un contexto hidrofóbico se espera que sea desestabilizante. Respecto a PIGY, la mutación ocurre en las inmediaciones del *kink* de la hélice alfa transmembrana, siendo una sustitución por prolina, lo que llevaría a 2 residuos de prolina consecutivos que distorsionarían la hélice.

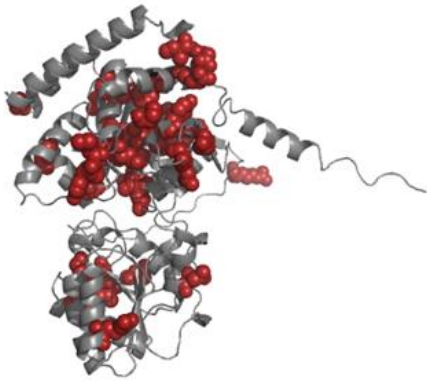
Finalmente, en el caso de DPM2, se han reportado dos mutaciones, correspondientes a dos residuos de las hélices alfa completamente embebidos en la membrana. Una de las mutaciones es el remplazo de la glicina 58 por glutamato, resultando en la introducción de un residuo cargado en un contexto hidrofóbico. La otra mutación es el remplazo Y23C, que puede ser desestabilizante ya que la

Y23 junto con W27 y Y24 (**Suplementaria 7 B**, ANEXO) forman un grupo de residuos aromáticos muy cercanos en el espacio que interactúan y están muy conservados.

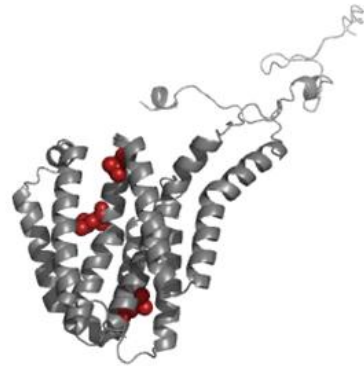
Mutaciones patogénicas y probablemente patogénicas						
PIGA	R19W	R19Q	H33R	M37V	M37T	V38A
	D40H	G48A	G48D	G48V	V49M	H65Y
	R77L	R77Q	R77G	R81H	Y90H	P93L
	M97K	Y98C	F117S	R119W	R119Q	S127L
	H128R	S132C	A135V	A142T	K143E	S155F
	N168H	N179Y	K189E	I206F	G239R	R275W
	N297D	V329A	S330N	L344P	L344-	L355S
	E395K	G421R	I451T	D452V	D452N	D452Y
PIGC	L189W	L212P	L220P			
PIGH	S103P					
PIGP	M25T					
PIGQ	G449R					
PIGY	L46P					
DPM2	Y23C	G58D				

Figura 29. Mutaciones halladas para cada subunidad del complejo GPI-GnT. Las siguientes mutaciones son de tipo *missense* (patogénicas y probablemente patogénicas) halladas en las distintas bases de datos.

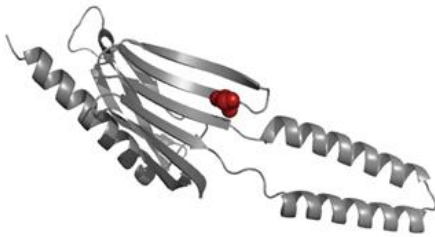
PIGA



PIGC



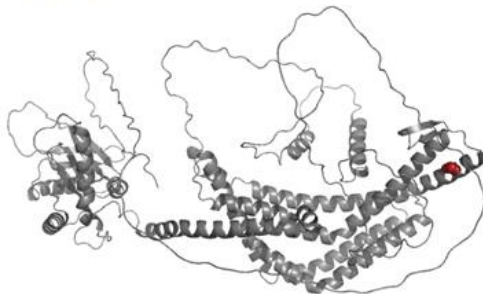
PIGH



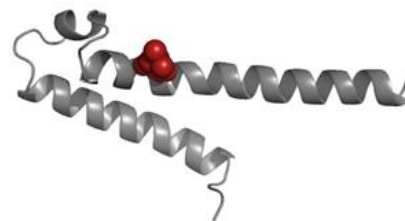
PIGP



PIGQ



PIGY



DPM2

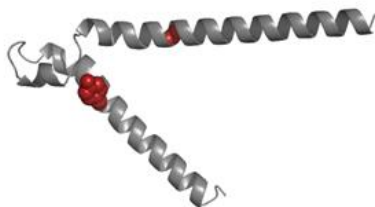


Figura 30. Mapeo de las mutaciones en las estructuras de las subunidades. A partir de las estructuras teóricas predichas por AlphaFold2, se mapearon las mutaciones listadas en la Fig.29. Sólo se detallan las mutaciones patogénicas y probablemente patogénicas.

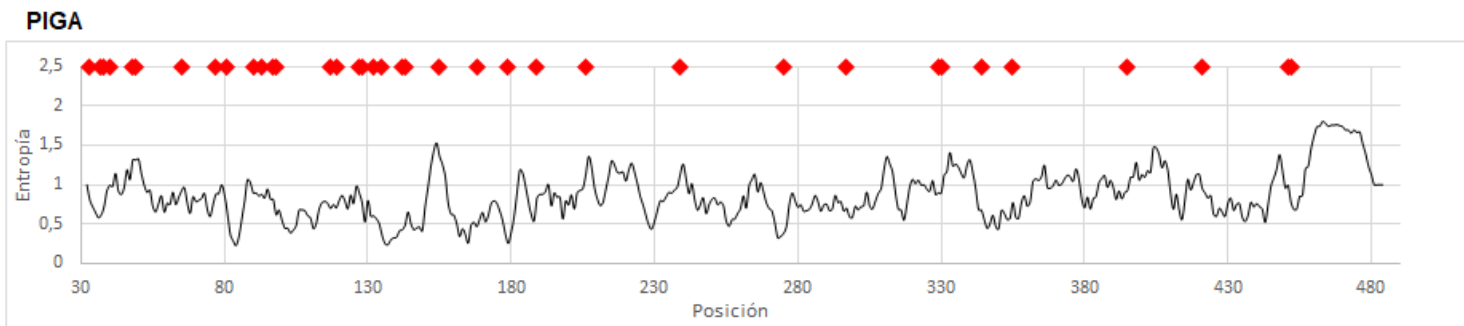


Figura 31. Conservación de secuencia asociada a las posiciones mutadas en PIGA. Se graficó la entropía (relacionada con el grado de conservación de la secuencia) en función de la posición y las mutaciones halladas en las bases de datos. Los puntos (en rojo) en la parte superior del gráfico representan la posición de las mutaciones. Se graficaron solo las mutaciones patológicas y probablemente patológicas. Gráficos semejantes para las demás subunidades pueden encontrarse en la sección Anexo en la Figura Suplementaria 9.

4.3.5. Modelado del complejo GPI-GnT

Para tratar de entender cómo interactúan las subunidades entre sí y como están dispuestas tridimensionalmente en la membrana del RE, se intentó construir un modelo del complejo GPI-GnT partiendo de las estructuras predichas y utilizando varias estrategias. En un principio se intentó obtener un modelo del complejo entero utilizando una extensión de AlphaFold2 (opción “Multimer”), pero por limitaciones de cómputo no se obtuvieron resultados. También se intentó analizar las subunidades de a pares con el mismo programa, pero los resultados que se obtuvieron no tenían sentido físico como ser que partes hidrofóbicas (dominios de membrana) de algunas estructuras estaban en contacto con regiones solubles de otra.

Por otro lado, se intentó ensamblar el complejo mediante medidas de coevolución. Este método necesita hacer alineamientos de a pares, donde las dos subunidades estén presentes en el mismo organismo. De cómo varían las apariciones de los distintos aminoácidos en las distintas posiciones en el par se determina que hay regiones que están acopladas y podrían estar en contacto, y esos acoplamientos se utilizan como restricciones en algoritmos de docking que van a tratar de ensamblar las dos subunidades tridimensionalmente. Para realizar esto es necesario tener miles de secuencias, donde las dos subunidades estén presentes en cada organismo, y en este caso para cada subunidad sólo se obtuvieron cientos de pares de secuencias. Esto complica el análisis ya que cuantas menos secuencias, menor es la confianza de que las variaciones observadas se correspondan con contactos, como ocurrió en este análisis, utilizando el servidor RaptorX (Wang et al., 2016), donde las probabilidades de los contactos eran muy bajas y no resultaban de utilidad.

4.3.6. Modelado de la unión del ligando al sitio activo de PIGA

Resulta de gran interés disponer de un modelo de la enzima PIGA interaccionando con sus sustratos. Esto es importante tanto para entender su mecanismo de catálisis como para el diseño de drogas tales como las chaperonas farmacológicas que se unen reversiblemente a la enzima estabilizando su conformación activa y que pueden potencialmente “rescatar” a variantes inestables o mal plegadas (Liguori et al., 2020). El programa AlphaFold2 da modelos de resolución atómica muy confiables, pero no modela a las moléculas pequeñas que pueden estar unidas a la proteína, como cofactores y sustratos. Entonces se usó el modelo de AlphaFold2 para buscar estructuras de otras glicosiltransferasas en la base de datos PDB, que pudieran estar resueltas con sus sustratos unidos. Para eso se usó nuevamente el programa Dali, y como se esperaba, se obtuvieron como resultado a proteínas que también pertenecen a la familia de las glicosiltransferasas tipo B. Muchas de ellas presentan ligandos unidos, mayormente nucleótidos de uridina (como UDP) y en algunos casos un monosacárido (o derivado) o disacárido. En la **Figura 32** se muestra la superposición de la estructura de PIGA con la glicosiltransferasa de *Thermosynechococcus elongatus*, en complejo con sus ligandos UDP y sacarosa-6-fosfato (uno de los sustratos de PIGA es UDP-N-acetilglucosamina). Como puede observarse en la figura, ambas estructuras son muy similares, a pesar de que *T. elongatus* es una cianobacteria termófila por lo que su glicosiltransferasa está muy distantemente relacionada con PIGA (presentan un 18% de identidad de secuencia). A pesar de ser homólogos remotos, las cadenas laterales involucradas en la unión de ligandos están muy conservadas (**Figura 32**). Este resultado es el punto de partida para un posterior modelado y refinamiento del complejo de PIGA con sus ligandos, realizado posteriormente en el laboratorio.

El otro sustrato de PIGA es el fosfatidilinositol (PI), que está formado por el azúcar inositol unido a una porción lipídica. Aunque no ha sido estudiado en detalle, aparentemente las distintas modificaciones involucradas en la síntesis de GPI ocurren sin que el fosfolípido abandone la membrana. En la búsqueda de estructuras semejantes con Dali no aparecieron complejos con PI, ni con moléculas solubles derivados de inositol o lípidos. Por lo tanto, no pudo inferirse el modo de interacción de este sustrato de la forma en que se hizo para UDP-N-acetilglucosamina. Es razonable que este sustrato interaccione con PIGA a través de su subdominio más hidrofóbico pues éste a su vez se espera que esté más cercano a la membrana. Una propuesta es que lo haga a través de la cavidad señalada en la **Figura 33**, lo que se podría explorar mediante docking.

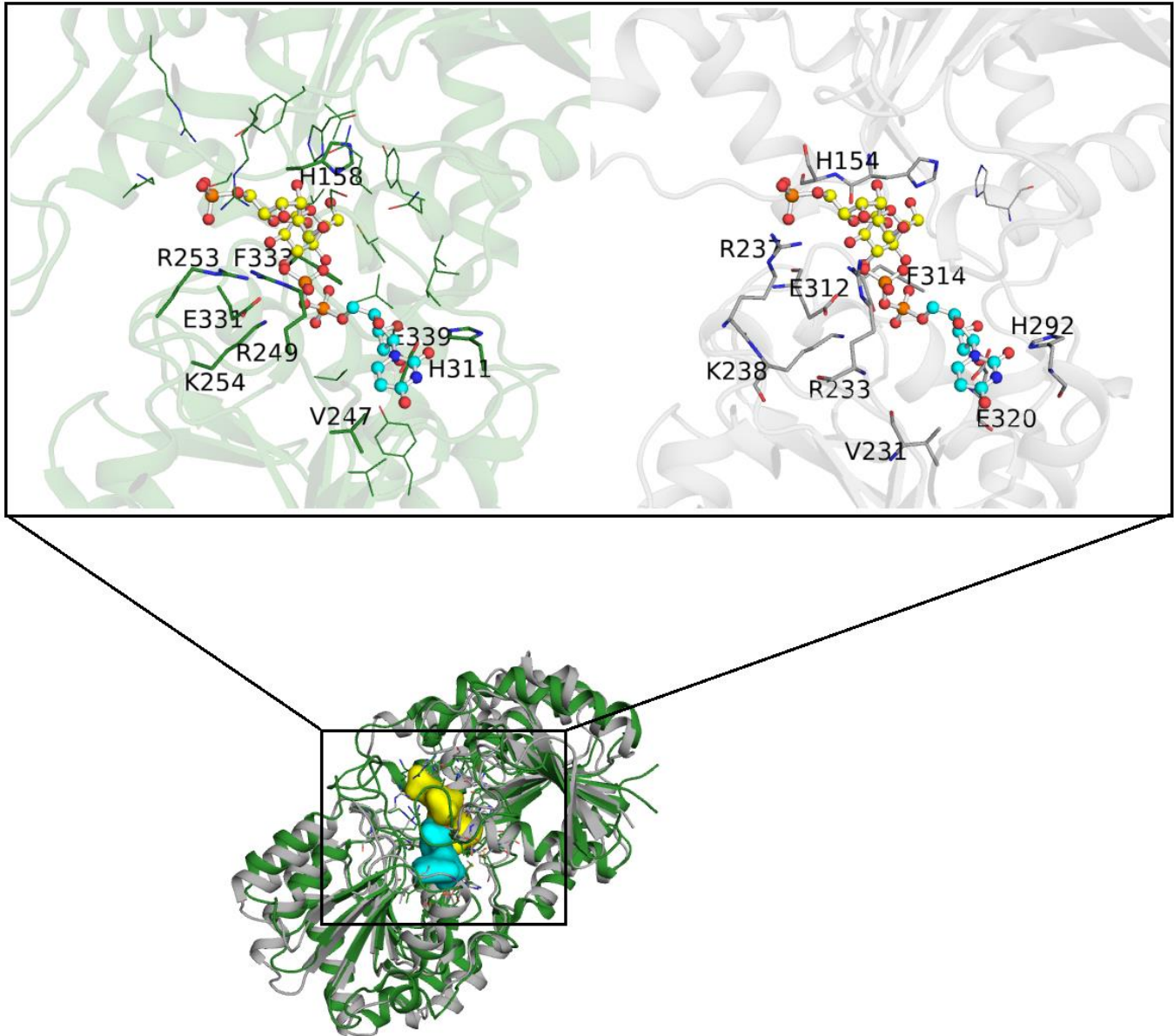


Figura 32. Detalle de las cadenas laterales involucradas en la unión de ligandos de PIGA y de la glicosiltransferasa de *T. elongatus*. Se observan en verde (izquierda) las cadenas laterales correspondientes a la glicosiltransferasa de *T. elongatus* que interactúan con los ligandos (sacarosa-6-P en amarillo y UDP en cyan) y en gris (derecha) las equivalentes en PIGA. Los residuos conservados en ambos casos se encuentran rotulados. Las cadenas laterales que interactúan con ligandos en *T. elongatus* que no se están conservadas en PIGA se representan con líneas. Los sustratos que se muestran en la estructura de PIGA son los correspondientes a la glicosiltransferasa de *T. elongatus*.

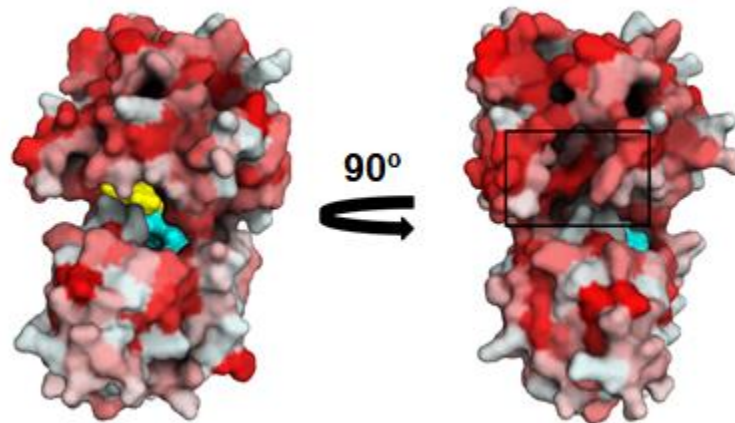


Figura 33. Modelado de unión de ligandos a PIGA. Los ligandos UDP (cyan) y sacarosa-6-P (amarillo) presentes en la glicosiltransferasa de *T. elongatus* se ubicaron en posiciones equivalentes de PIGA (al superponer ambas estructuras). Se muestra la superficie de PIGA coloreada por hidrofobicidad (en rojo las zonas más hidrofóbicas). A la derecha se señala la región donde podría unirse el PI.

DISCUSIÓN

5.1. Discusión

En este trabajo se buscó hacer un aporte a la comprensión de un proceso fisiológico importante, como lo es la síntesis de GPI, y a partir de allí racionalizar sobre las consecuencias de la deficiencia de este proceso cuando hay alteraciones por mutaciones en las proteínas que intervienen en la etapa inicial. El primer paso de la síntesis de GPI ocurre en el RE mediante la acción de la enzima PIGA en el complejo de siete unidades (en mamíferos) GPI-GnT. A la fecha de realización de este trabajo no hay estudios que describan la estructura de PIGA en detalle, por lo que se decidió indagar un poco más en esta cuestión. Para eso se abordó el problema tanto desde un enfoque experimental como computacional. Con la primera estrategia, se buscó producir PIGA de manera recombinante para ser purificada y estudiada *in vitro*, y también realizar estudios funcionales en levaduras. En el aspecto computacional, mediante nuevas herramientas bioinformáticas, que permiten la predicción de la estructura de proteínas con alta confiabilidad, se pudieron analizar diversos aspectos estructurales de cada subunidad, inferir el modo de unión de ligandos y especular sobre la base estructural del efecto fenotípico de las mutaciones.

5.1.2. Expresión recombinante de PIGA: un target difícil para la expresión en *E. coli*

Para intentar producir PIGA recombinante en *E. coli*, una de las precauciones que se tuvieron en cuenta fue eliminar la región transmembrana y el pequeño segmento luminal, ya que la expresión de proteínas de membrana suele ser más dificultosa. Dado que el objetivo a largo plazo en el proyecto era estudiar a las variantes salvajes y mutantes de PIGA, para expresarlas tanto en *E. coli* como en levaduras, se creó una plataforma para generar las variantes con el menor número de pasos posibles. Todas las construcciones necesarias para este fin se consiguieron con éxito.

A lo largo de la experiencia en los intentos de expresar PIGA_c, quedó claro que no es sencillo obtener esta proteína por expresión en *E. coli*. Inicialmente se observó en SDS-PAGE una banda compatible con el tamaño esperado, cuya intensidad era mayor que en la muestra sin inducir. Un intento de purificación mediante NTA-Ni mostró un importante enriquecimiento de una banda de tamaño compatible. Sin embargo, intentos posteriores de optimizar la expresión fueron infructuosos. Se cambiaron los medios de cultivo, temperaturas, las concentraciones de inductor, la densidad óptica de inducción y los tiempos de cultivo. Además, se ensayaron 4 cepas de *E. coli* con distintas propiedades, pero más allá de pequeñas diferencias, no se logró una sobreexpresión significativa.

Finalmente, en ensayos en los que se usó como control a las células no transformadas inducidas, en lugar del control habitual de transformadas sin inducir, se observó que tanto los cultivos no transformados como los transformados presentaban una misma banda, compatible con el tamaño esperado. Debe remarcar que para visualizar esta banda en cultivos no transformados fue necesario

enriquecer la muestra mediante purificación con NTA-Ni, de otra forma dicha banda no alcanzaba a observarse. Como en pruebas posteriores no pudo lograrse el mismo enriquecimiento que en la purificación original, lo que se concluyó de mínima es que la estrategia ensayada no permite producir la proteína recombinante de manera reproducible.

En este punto se cuestionó cual podría ser el fallo y qué se podría hacer para mejorar la expresión. ¿Podría ser que PIGA no pueda expresarse en ausencia de las otras subunidades? Mientras que es posible que PIGA sea inestable en forma aislada, en ese caso sería esperable que se acumule en cuerpos de inclusión, que no fue lo observado.

Tal como se mostró en la sección de resultados la expresión de PIGAc, en las cepas utilizadas (Rosetta Gammi2, Rosetta y RIL) se logra corregir casi completamente la limitación en la traducción con el agregado de codones de bajo uso. Aunque quedan unos pocos codones de muy bajo uso (CGA y CGG, ambos codifican para la arginina) que igualmente podrían afectar la expresión de PIGAc. Mientras que es posible que el único par de codones de bajo uso que persiste al expresar en las cepas ensayadas sea el origen del problema de expresión, en principio se considera difícil que sea la única causa.

A partir de los resultados obtenidos con el servidor TESINGER se observó que los primeros codones de la secuencia del mRNA tienen un drástico efecto en la tasa de éxito de expresión predicha. En ese sentido, eliminar el extremo N tendría como ventaja: 1) quitar un segmento predicho como desplegado por AlfaFold2 que podría ser problemático, 2) eliminar el codón de bajo uso que no se corrige con Rosetta Gami2, y 3) mejorar el reconocimiento del mRNA por la maquinaria de traducción.

De todas formas, también se ha discutido que algunas glicosiltransferasas son targets difíciles para la sobreexpresión en *E. coli* (Welner et al., 2017). No puede descartarse entonces la necesidad de recurrir a sistemas de expresión eucariotas, como levaduras o baculovirus.

5.1.3. Avances hacia el desarrollo de un sistema para el estudio de variantes de PIGA en levaduras

El análisis funcional por complementación en levaduras resultaba muy atractivo por su sencillez: un gen PIGA no funcional en *S. cerevisiae* causa que la síntesis de GPI sea deficiente, y como el GPI es importante para el direccionamiento de proteínas que intervienen en la formación de la pared celular, su ausencia es letal bajo ciertas condiciones. Disponiendo de una mutante termosensible de PIGA que no crece a 35°C pero sí a 24°C, se diseñó una construcción genética para realizar la complementación con PIGA humana. En bibliografía está descrito un experimento relacionado, donde mediante un plásmido se introdujo al homólogo de levadura de PIGA, con ciertas mutaciones de diseño, en una cepa termosensible por deficiencia de PIGA (Kostova et al., 2003).

Como primer paso de ese proceso, debió generarse el plásmido de expresión de PIGA humana en levaduras. Esto se realizó mediante una la ligación *in vivo*. Sin embargo, no observamos complementación: la tasa de crecimiento es baja a 35°, pero el crecimiento es el mismo cuando se transforma con la construcción o el vector solo.

En este punto se pensó que el problema podría deberse a la falla de expresión en *S. cerevisiae*. Para verificar la expresión se realizaron Western Blots de distintos extractos de la cepa termosensible, tanto de colonias transformadas con la construcción como con el vector vacío. Se reveló una banda compatible con el tamaño esperado pero que estaba presente en el control transformado con el vector vacío. Aunque no está descrito en bibliografía, los resultados indican que hay un problema de reconocimiento cruzado del anticuerpo anti Hisx6Tag de ratón con proteínas endógenas de *S. cerevisiae*.

Considerando los resultados obtenidos, se analizaron algunas posibles causas de las dificultades encontradas. Una posibilidad es que la proteína introducida con el plásmido se exprese, aunque no pueda detectarse porque una proteína endógena tiene la misma migración en el gel y también tiene reconocimiento cruzado por el anticuerpo anti Hisx6Tag. En ese caso, además, la proteína no funcionaría como se espera, ya que no detectamos complementación. Eso podría deberse a que la misma no se localice en el RE, donde debe asociarse a las otras subunidades para funcionar. Esto podría ser resultado de la ausencia de señal para el direccionamiento a ese compartimiento, en el caso de que dicha señal sea diferente en el ortólogo humano o esté ausente. No está reportada la presencia de péptido señal en ningún ortólogo de PIGA y los intentos de detección de péptido señal por métodos bioinformáticos no dieron resultados positivos (no mostrado). Se podría suponer que, como es el caso de otras proteínas transmembrana, no existe una secuencia definida para el direccionamiento y tal vez la información está en la presencia del segmento transmembrana. Por otro lado, es relevante para este punto comentar que en un estudio previo se encontró que un segmento de 23 residuos del extremo luminal (Watanabe et al., 1996), bien cercano a la membrana, es importante para la persistencia de PIGA en el RE ya que con su eliminación la proteína pierde esa localización en células humanas JY5. Realizando la comparación de las secuencias de la variante humana y de levadura en esa región, se observa que no está conservada. Eso permite proponer como una posible estrategia, el diseño de una variante quimérica donde el segmento luminal sea el de levadura para lograr su correcto direccionamiento al RE.

Por otro lado, no puede descartarse que la ausencia de complementación funcional sea debida a las diferencias entre la variante humana y levadura, que hagan imposible la consolidación del complejo

GPI-GnT en la membrana del RE. Si bien las estructuras modeladas de las subunidades son semejantes (no mostrado), podría ser que pequeñas diferencias sean cruciales para formar el complejo.

Finalmente, no se descarta la necesidad de cambiar el sistema de expresión en levaduras a otro vector con un promotor más fuerte.

5.1.4. La composición del complejo GPI-GnT está muy conservada en mamíferos, pero es más divergente en otros eucariotas

Se realizó un estudio bioinformático de todas las subunidades del complejo GPI-GnT con especial énfasis en PIGA. Primeramente, se buscó responder si existen subunidades del complejo que, junto con PIGA, sean indispensables para que la síntesis de anclas pueda tener lugar en esta primera etapa de la vía. Este sería el caso, por ejemplo, si alguna de ellas colaborara en el adecuado posicionamiento de alguno de los ligandos en el sitio activo, de modo tal que en su ausencia, PIGA no pudiera llevar a cabo la reacción. Para ello, se compró el perfil de presencia/ausencia de los genes correspondientes en los genomas de distintos organismos eucariotas representativos. Para eso, a partir de un número representativo de organismos eucariotas modelo, se realizó una búsqueda de ortólogos para todas las subunidades del complejo GPI-GnT utilizando HMMER/BlastP y tBlastN. Se observó que PIGA se encontraba presente en el 94% de los organismos y fue la única subunidad detectada en un 8% de los organismos. Además, existe una proporción importante de organismos que carecen de una o varias subunidades. Esto indica que, en principio, de todo el complejo sólo es esencial PIGA y que esta puede funcionar en ausencia de las otras subunidades, aún cuando su actividad estando sola pueda ser insuficiente para las necesidades fisiológicas de muchos organismos.

Un caso llamativo de ausencia se da con la subunidad PIGY en *S. cerevisiae*. A pesar de que en la bibliografía se presenta a Eri1 como un posible homólogo de PIGY (Murakami et al., 2005), realmente no comparten una identidad de secuencia apreciable y aunque poseen un perfil de hidrofobicidad parecido y la ausencia de Eri1 afecta a la síntesis de GPI, no parece correcto asegurar que hay ortología entre ambos si no hay similitud de secuencia detectable. Es por eso que PIGY no aparece en el árbol filogenético para *S. cerevisiae* y especies relacionadas (**Figura 26**).

En el presente trabajo de tesis, se estudió la co-ocurrencia de las subunidades del complejo, lo que fue útil para poder conocer cuál es el núcleo funcional evolutivamente conservado. Se encontró que tanto PIGC, como PIGP y PIGQ, suelen estar presentes cuando PIGA también lo está. Este resultado podría indicar que estas tres subunidades podrían ser los principales actores para ajustar la actividad de PIGA a las necesidades fisiológicas. Además, se observó que tanto PIGH como PIGY y DPM2 coexisten con PIGA en una menor proporción de organismos. Esto contrasta con algunos estudios de

coprecipitación realizados, usados para inferir la estequiometría del complejo humano, donde se encontró que PIGA interactúa más fuertemente con PIGH y también se uniría de manera directa a PIGY y PIGQ, y esta última uniría a PIGC y PIGP (revisado en (Levin & Stamper, 2009)). Una jerarquía basada en la fuerza de interacción de las subunidades del complejo de un organismo en particular no se traduce directamente en cuáles son las subunidades más importantes para el funcionamiento del mismo. Por lo que un modelo posible sería que los organismos empleen primariamente a PIGQ para ajustar el funcionamiento de PIGA por interacción directa, y que PIGQ traiga consigo a PIGC y PIGP, integrándolas al complejo. La siguiente subunidad más importante sería PIGH, uniéndose de manera directa a PIGA. Luego PIGY y DPM2 serían subunidades accesorias empleadas por un número menor de organismos.

5.1.5. Los modelos estructurales de las subunidades del complejo GPI-GnT permiten entender la base estructural de algunas mutaciones clínicas

A partir de las estructuras predichas por AlphaFold2 para cada subunidad del complejo y las mutantes patogénicas conocidas, se estudió su localización en la estructura. Se observó que para PIGA hay algunas mutaciones puntuales de importancia clínica que se ubican cerca del sitio activo, aunque otras no. Entre las últimas, algunas podrían afectar la interacción con las demás subunidades del complejo, otras comprometer la integridad estructural de PIGA e incluso algunas tener un efecto alostérico en el sitio activo. En relación a esto último, hace unos años se publicó un trabajo (Agrahari et al., 2019) donde se hizo un modelado por homología de la estructura de PIGA (usando como templatado otra glicosiltransferasa procariota de baja identidad, 18%) y se analizó el potencial impacto de 4 mutaciones clínicas en la estructura mediante predictores de estabilidad y simulaciones de dinámica molecular. Es de destacar que dichas simulaciones se hicieron con el equivalente al dominio globular citosólico de PIGA, es decir, en ausencia de las otras subunidades y de la membrana plasmática. En las simulaciones algunas alteraciones en la movilidad se interpretaron como una desestabilización de la estructura. En este sentido, si bien esos análisis pueden dar algo de información, es fundamental analizar lo que le ocurre a PIGA en presencia de las otras subunidades y con las restricciones que impone su orientación en la membrana para tener un mejor panorama del efecto de las mutaciones. Además, el uso de un modelo más exacto como el obtenido con AlfaFold2, puede mejorar la calidad de tales análisis.

5.1.6. El modelado con glicosiltransferasas resueltas con ligandos unidos y similares a PIGA permitió conocer cómo ésta interactúa con sus sustratos.

El modelado de unión del ligando al sitio activo de PIGA permitió confirmar que el mismo se localiza en la interfaz entre sus dos subdominios, como se esperaba por analogía con otras glicosil transferasas. Se observó que los residuos involucrados en la interacción con uno de los ligandos, UDP-

GlcNAc, están conservados en comparación con otras proteínas de la familia. Sería necesario hacer un modelado de la estructura unida a PI (su otro ligando) para poder así construir el modelo completo con los sustratos unidos, que a su vez podría servir de material de partida para un estudio teórico del mecanismo catalítico. Aunque los resultados del análisis de conservación de las subunidades indican que PIGA tiene todo lo necesario para ser catalíticamente activa, puede que alguna de las subunidades complete de alguna forma el sitio activo por ejemplo facilitando la interacción con el PI. Sería óptimo contar con un modelo de PIGA junto con las subunidades con las que interacciona directamente.

CONCLUSIONES

5.2. Conclusiones

- PIGA parece ser la única subunidad catalíticamente esencial del complejo, ya que muchos organismos carecen de varias de las otras subunidades. Estas subunidades accesorias podrían tener distintos roles, como moduladores, en la canalización del sustrato, etc.
- Algunas mutaciones puntuales de importancia clínica se ubican cerca del sitio activo afectando de manera directa la función, pero aquellas que se ubican en otras regiones podrían afectar la interacción con las subunidades del complejo, comprometer la integridad estructural de PIGA o tener un efecto alostérico en el sitio activo.
- El modelado de la estructura tridimensional de PIGA y la búsqueda de estructuras semejantes en la base de datos de proteínas de estructura conocida, permitió conocer donde se localiza su sitio activo. Además, se comprobó que varios residuos involucrados en la unión de sustratos están conservados. Todo esto puede ser la base para diseñar nuevos tratamientos o estrategias terapéuticas (como chaperonas farmacológicas) donde la enfermedad sea el resultado de una mutación *missense*.
- No es fácil lograr la sobreexpresión de PIGA en *E. coli*, es en principio necesario optimizar la construcción genética.
- Los resultados obtenidos del ensayo de complementación de PIGA en *S. cerevisiae* no son suficientes para concluir si la introducción de la variante humana a través de un plásmido complementa la ausencia de la proteína funcional de levadura.

REFERENCIAS BIBLIOGRÁFICAS

Referencias bibliográficas

- Agrahari, A. K., Pieroni, E., Gatto, G., & Kumar, A. (2019). The impact of missense mutation in PIGA associated to paroxysmal nocturnal hemoglobinuria and multiple congenital anomalies-hypotonia-seizures syndrome 2: A computational study. *Heliyon*, 5(10), e02709. <https://doi.org/10.1016/j.heliyon.2019.e02709>
- Ashida, H., Hong, Y., Murakami, Y., Shishioh, N., Sugimoto, N., Kim, Y. U., Maeda, Y., & Kinoshita, T. (2005). Mammalian PIG-X and Yeast Pbn1p Are the Essential Components of Glycosylphosphatidylinositol-Mannosyltransferase I. *Molecular Biology of the Cell*, 16(3), 1439-1448. <https://doi.org/10.1091/mbc.E04-09-0802>
- Bhandari, B. K., Lim, C. S., & Gardner, P. P. (2021). TISIGNER.com: Web services for improving recombinant protein production. *Nucleic Acids Research*, 49(W1), W654-W661. <https://doi.org/10.1093/nar/gkab175>
- Campbell, J. A., Davies, G. J., Bulone, V., & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *The Biochemical Journal*, 326 (Pt 3), 929-939. <https://doi.org/10.1042/bj3260929u>
- Carmody, L. C., Blau, H., Danis, D., Zhang, X. A., Gourdine, J.-P., Vasilevsky, N., Krawitz, P., Thompson, M. D., & Robinson, P. N. (2020). Significantly different clinical phenotypes associated with mutations in synthesis and transamidase+remodeling glycosylphosphatidylinositol (GPI)-anchor biosynthesis genes. *Orphanet Journal of Rare Diseases*, 15(1), 40. <https://doi.org/10.1186/s13023-020-1313-0>
- Carretero Colomer, M. (2009). Hemoglobinuria paroxística nocturna. Abordaje de una enfermedad rara. *Offarm*, 28(3), 60-61.
- Colussi, P. A., & Orlean, P. (1997). The essential *Schizosaccharomyces pombe* gpil+ gene complements a bakers' yeast GPI anchoring mutant and is required for efficient cell separation. *Yeast (Chichester, England)*, 13(2), 139-150. [https://doi.org/10.1002/\(SICI\)1097-0061\(199702\)13:2<139::AID-YEA69>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1097-0061(199702)13:2<139::AID-YEA69>3.0.CO;2-S)
- Coutinho, P. M., Deleury, E., Davies, G. J., & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *Journal of Molecular Biology*, 328(2), 307-317. [https://doi.org/10.1016/s0022-2836\(03\)00307-3](https://doi.org/10.1016/s0022-2836(03)00307-3)

Cruz, L. M. G., Juárez, J. H., Rojas, B. F., Antonio, P., & Cruz, H. (2019). *EL PAPEL DE LAS GALNAc-TRANSFERASAS EN EL DESARROLLO DEL CÁNCER DE MAMA*. 9.

Das, M., & Basu, G. (2012). Glycine Rescue of β -Sheets from cis-Proline. *Journal of the American Chemical Society*, 134(40), 16536-16539. <https://doi.org/10.1021/ja308110t>

De Latour, R. P., Mary, J. Y., Salanoubat, C., Terriou, L., Etienne, G., Mohty, M., Roth, S., de Guibert, S., Maury, S., Cahn, J. Y., Socié, G., French Society of Hematology, & French Association of Young Hematologists. (2008). Paroxysmal nocturnal hemoglobinuria: Natural history of disease subcategories. *Blood*, 112(8), 3099-3106. <https://doi.org/10.1182/blood-2008-01-133918>

DeGasperi, R., Thomas, L. J., Sugiyama, E., Chang, H. M., Beck, P. J., Orlean, P., Albright, C., Waneck, G., Sambrook, J. F., & Warren, C. D. (1990). Correction of a defect in mammalian GPI anchor biosynthesis by a transfected yeast gene. *Science (New York, N.Y.)*, 250(4983), 988-991. <https://doi.org/10.1126/science.1978413>

Desjardins, C. A., Sanscrainte, N. D., Goldberg, J. M., Heiman, D., Young, S., Zeng, Q., Madhani, H. D., Becnel, J. J., & Cuomo, C. A. (2015). Contrasting host–pathogen interactions and genome evolution in two generalist and specialist microsporidian pathogens of mosquitoes. *Nature Communications*, 6, 7121. <https://doi.org/10.1038/ncomms8121>

Doerrler, W. T., Ye, J., Falck, J. R., & Lehrman, M. A. (1996). Acylation of Glucosaminyl Phosphatidylinositol Revisited: PALMITOYL-CoA DEPENDENT PALMITOYLATION OF THE INOSITOL RESIDUE OF A SYNTHETIC DIOCTANOYL GLUCOSAMINYL PHOSPHATIDYLINOSITOL BY HAMSTER MEMBRANES PERMITS EFFICIENT MANNOSYLATION OF THE GLUCOSAMINE RESIDUE*. *Journal of Biological Chemistry*, 271(43), 27031-27038. <https://doi.org/10.1074/jbc.271.43.27031>

EIEE77—ClinVar—NCBI. (s. f.). Recuperado 9 de febrero de 2023, de <https://www.ncbi.nlm.nih.gov/clinvar/?gr=1&term=EIEE77>

Eisenhaber, B., Maurer-Stroh, S., Novatchkova, M., Schneider, G., & Eisenhaber, F. (2003). Enzymes and auxiliary factors for GPI lipid anchor biosynthesis and post-translational transfer to proteins. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 25(4), 367-385. <https://doi.org/10.1002/bies.10254>

Ferguson, M. A. (1999). The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *Journal of Cell Science*, 112 (Pt 17), 2799-2809. <https://doi.org/10.1242/jcs.112.17.2799>

Ferguson, M. A., Homans, S. W., Dwek, R. A., & Rademacher, T. W. (1988). Glycosylphosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science (New York, N.Y.)*, 239(4841 Pt 1), 753-759. <https://doi.org/10.1126/science.3340856>

Ferguson, M. A., Low, M. G., & Cross, G. A. (1985). Glycosyl-sn-1,2-dimyristylphosphatidylinositol is covalently linked to *Trypanosoma brucei* variant surface glycoprotein. *The Journal of Biological Chemistry*, 260(27), 14547-14555.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. En J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). Humana Press. <https://doi.org/10.1385/1-59259-890-0:571>

Hill, A., DeZern, A. E., Kinoshita, T., & Brodsky, R. A. (2017). Paroxysmal nocturnal haemoglobinuria. *Nature Reviews. Disease Primers*, 3, 17028. <https://doi.org/10.1038/nrdp.2017.28>

Hillmen, P., Lewis, S. M., Bessler, M., Luzzatto, L., & Dacie, J. V. (1995). Natural history of paroxysmal nocturnal hemoglobinuria. *The New England Journal of Medicine*, 333(19), 1253-1258. <https://doi.org/10.1056/NEJM199511093331904>

Hirose, S., Prince, G. M., Sevelev, D., Ravi, L., Rosenberry, T. L., Ueda, E., & Medof, M. E. (1992). Characterization of putative glycoinositol phospholipid anchor precursors in mammalian cells. Localization of phosphoethanolamine. *Journal of Biological Chemistry*, 267(24), 16968-16974. [https://doi.org/10.1016/S0021-9258\(18\)41879-0](https://doi.org/10.1016/S0021-9258(18)41879-0)

Høie, M. H., Kiehl, E. N., Petersen, B., Nielsen, M., Winther, O., Nielsen, H., Hallgren, J., & Marcatili, P. (2022). NetSurfP-3.0: Accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Research*, 50(W1), W510-W515. <https://doi.org/10.1093/nar/gkac439>

Holm, L., & Sander, C. (1995). Dali: A network tool for protein structure comparison. *Trends in Biochemical Sciences*, 20(11), 478-480. [https://doi.org/10.1016/s0968-0004\(00\)89105-7](https://doi.org/10.1016/s0968-0004(00)89105-7)

- Homans, S. W., Ferguson, M. A., Dwek, R. A., Rademacher, T. W., Anand, R., & Williams, A. F. (1988). Complete structure of the glycosyl phosphatidylinositol membrane anchor of rat brain Thy-1 glycoprotein. *Nature*, 333(6170), 269-272. <https://doi.org/10.1038/333269a0>
- Hong, Y., Maeda, Y., Watanabe, R., Inoue, N., Ohishi, K., & Kinoshita, T. (2000). Requirement of PIG-F and PIG-O for transferring phosphoethanolamine to the third mannose in glycosylphosphatidylinositol. *The Journal of Biological Chemistry*, 275(27), 20911-20919. <https://doi.org/10.1074/jbc.M001913200>
- Hong, Y., Maeda, Y., Watanabe, R., Ohishi, K., Mishkind, M., Riezman, H., & Kinoshita, T. (1999). Pign, a mammalian homologue of yeast Mcd4p, is involved in transferring phosphoethanolamine to the first mannose of the glycosylphosphatidylinositol. *The Journal of Biological Chemistry*, 274(49), 35099-35106. <https://doi.org/10.1074/jbc.274.49.35099>
- Hong, Y., Ohishi, K., Watanabe, R., Endo, Y., Maeda, Y., & Kinoshita, T. (1999). GPI1 stabilizes an enzyme essential in the first step of glycosylphosphatidylinositol biosynthesis. *The Journal of Biological Chemistry*, 274(26), 18582-18588. <https://doi.org/10.1074/jbc.274.26.18582>
- Inoue, N., Kinoshita, T., Orii, T., & Takeda, J. (1993). Cloning of a human gene, PIG-F, a component of glycosylphosphatidylinositol anchor biosynthesis, by a novel expression cloning strategy. *The Journal of Biological Chemistry*, 268(10), 6882-6885.
- Inoue, N., Watanabe, R., Takeda, J., & Kinoshita, T. (1996). PIG-C, one of the three human genes involved in the first step of glycosylphosphatidylinositol biosynthesis is a homologue of *Saccharomyces cerevisiae* GPI2. *Biochemical and Biophysical Research Communications*, 226(1), 193-199. <https://doi.org/10.1006/bbrc.1996.1332>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kang, J. Y., Hong, Y., Ashida, H., Shishioh, N., Murakami, Y., Morita, Y. S., Maeda, Y., & Kinoshita, T. (2005). PIG-V involved in transferring the second mannose in glycosylphosphatidylinositol. *The Journal of Biological Chemistry*, 280(10), 9489-9497. <https://doi.org/10.1074/jbc.M413867200>
- Kanzawa, N., Maeda, Y., Ogiso, H., Murakami, Y., Taguchi, R., & Kinoshita, T. (2009). Peroxisome dependency of alkyl-containing GPI-anchor biosynthesis in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences*, 106(42), 17711-17716. <https://doi.org/10.1073/pnas.0904762106>

Kinoshita, T. (2016). Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. *Journal of Lipid Research*, 57(1), 4-5. <https://doi.org/10.1194/jlr.E065417>

Kinoshita, T., & Fujita, M. (2016). Biosynthesis of GPI-anchored proteins: Special emphasis on GPI lipid remodeling. *Journal of Lipid Research*, 57(1), 6-24. <https://doi.org/10.1194/jlr.R063313>

Kostova, Z., Rancour, D. M., Menon, A. K., & Orlean, P. (2000). Photoaffinity labelling with P3-(4-azidoanilido)uridine 5'-triphosphate identifies gpi3p as the UDP-GlcNAc-binding subunit of the enzyme that catalyses formation of GlcNAc-phosphatidylinositol, the first glycolipid intermediate in glycosylphosphatidylinositol synthesis. *The Biochemical Journal*, 350 Pt 3, 815-822.

Kostova, Z., Yan, B. C., Vainauskas, S., Schwartz, R., Menon, A. K., & Orlean, P. (2003). Comparative importance in vivo of conserved glutamate residues in the EX7E motif retaining glycosyltransferase Gpi3p, the UDP-GlcNAc-binding subunit of the first enzyme in glycosylphosphatidylinositol assembly. *European Journal of Biochemistry*, 270(22), 4507-4514. <https://doi.org/10.1046/j.1432-1033.2003.03844.x>

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), 105-132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259), Article 5259. <https://doi.org/10.1038/227680a0>

Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., Lyoshin, V., Maddipatla, Z., Maiti, R., Mitchell, J., O'Leary, N., Riley, G. R., Shi, W., Zhou, G., Schneider, V., ... Kattman, B. L. (2020). ClinVar: Improvements to accessing data. *Nucleic Acids Research*, 48(D1), D835-D844. <https://doi.org/10.1093/nar/gkz972>

Laskowski, R. A., Stephenson, J. D., Sillitoe, I., Orengo, C. A., & Thornton, J. M. (2020). VarSite: Disease variants and protein structure. *Protein Science: A Publication of the Protein Society*, 29(1), 111-119. <https://doi.org/10.1002/pro.3746>

Leidich, S. D., & Orlean, P. (1996). Gpi1, a *Saccharomyces cerevisiae* protein that participates in the first step in glycosylphosphatidylinositol anchor synthesis. *The Journal of Biological Chemistry*, 271(44), 27829-27837. <https://doi.org/10.1074/jbc.271.44.27829>

Levin, D. E., & Stamper, R. J. (2009). Chapter 2 The N-Acetylglucosamine-PI Transfer Reaction, the GlcNAc-PI Transferase Complex, and Its Regulation. En *The Enzymes* (Vol. 26, pp. 31-47). Academic Press. [https://doi.org/10.1016/S1874-6047\(09\)26002-1](https://doi.org/10.1016/S1874-6047(09)26002-1)

Liguori, L., Monticelli, M., Allocca, M., Hay Mele, B., Lukas, J., Cubellis, M. V., & Andreotti, G. (2020). Pharmacological Chaperones: A Therapeutic Approach for Diseases Caused by Destabilizing Missense Mutations. *International Journal of Molecular Sciences*, *21*(2), 489. <https://doi.org/10.3390/ijms21020489>

Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., dos Santos Costa, A., Fazel-Zarandi, M., Sercu, T., Candido, S., & Rives, A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, *379*(6637), 1123-1130. <https://doi.org/10.1126/science.ade2574>

Lomize, A. L., Todd, S. C., & Pogozheva, I. D. (2022). Spatial arrangement of proteins in planar and curved membranes by PPM 3.0. *Protein Science: A Publication of the Protein Society*, *31*(1), 209-220. <https://doi.org/10.1002/pro.4219>

Low, M. G. (1999). GPI-anchored biomolecules—An overview. *GPI-anchored membrane proteins and carbohydrates*, 1-14.

Maeda, Y., Tanaka, S., Hino, J., Kangawa, K., & Kinoshita, T. (2000). Human dolichol-phosphate-mannose synthase consists of three subunits, DPM1, DPM2 and DPM3. *The EMBO Journal*, *19*(11), 2475-2482. <https://doi.org/10.1093/emboj/19.11.2475>

Maeda, Y., Watanabe, R., Harris, C. L., Hong, Y., Ohishi, K., Kinoshita, K., & Kinoshita, T. (2001). PIG-M transfers the first mannose to glycosylphosphatidylinositol on the luminal side of the ER. *The EMBO Journal*, *20*(1-2), 250-261. <https://doi.org/10.1093/emboj/20.1.250>

Maxwell, S. E., Ramalingam, S., Gerber, L. D., Brink, L., & Udenfriend, S. (1995). An active carbonyl formed during glycosylphosphatidylinositol addition to a protein is evidence of catalysis by a transamidase. *The Journal of Biological Chemistry*, *270*(33), 19576-19582. <https://doi.org/10.1074/jbc.270.33.19576>

Menon, A. K., Eppinger, M., Mayor, S., & Schwarz, R. T. (1993). Phosphatidylethanolamine is the donor of the terminal phosphoethanolamine group in trypanosome glycosylphosphatidylinositols. *The EMBO Journal*, *12*(5), 1907-1914.

Miseta, A., & Csutora, P. (2000). Relationship Between the Occurrence of Cysteine in Proteins and the Complexity of Organisms. *Molecular Biology and Evolution*, 17(8), 1232-1239. <https://doi.org/10.1093/oxfordjournals.molbev.a026406>

Miyata, T., Takeda, J., Iida, Y., Yamada, N., Inoue, N., Takahashi, M., Maeda, K., Kitani, T., & Kinoshita, T. (1993). The cloning of PIG-A, a component in the early step of GPI-anchor biosynthesis. *Science (New York, N.Y.)*, 259(5099), 1318-1320. <https://doi.org/10.1126/science.7680492>

Multiple Congenital Anomalies-Hypotonia-Seizures Syndrome 4 disease: Malacards—Research Articles, Drugs, Genes, Clinical Trials. (s. f.). Recuperado 29 de junio de 2023, de https://www.malacards.org/card/multiple_congenital_anomalies_hypotonia_seizures_syndrome_4_2

Muñoz-Linares, C., Ojeda, E., Forés, R., Pastrana, M., Cabero, M., Morillo, D., Bautista, G., Baños, I., Monteserín, C., Bravo, P., Jaro, E., Cedena, T., Steegmann, J. L., Villegas, A., & Cabrera, J. R. (2014). Paroxysmal nocturnal hemoglobinuria: A single Spanish center's experience over the last 40 yr. *European Journal of Haematology*, 93(4), 309-319. <https://doi.org/10.1111/ejh.12346>

Murakami, Y., & Kinoshita, T. (2015). [Inherited GPI deficiencies: a new disease with intellectual disability and epilepsy]. *No to Hattatsu = Brain and Development*, 47(1), 5-13.

Murakami, Y., Siripanyaphinyo, U., Hong, Y., Tashima, Y., Maeda, Y., & Kinoshita, T. (2005). The initial enzyme for glycosylphosphatidylinositol biosynthesis requires PIG-Y, a seventh component. *Molecular Biology of the Cell*, 16(11), 5236-5246. <https://doi.org/10.1091/mbc.e05-08-0743>

Murakami, Y., Siripanyapinyo, U., Hong, Y., Kang, J. Y., Ishihara, S., Nakakuma, H., Maeda, Y., & Kinoshita, T. (2003). PIG-W is critical for inositol acylation but not for flipping of glycosylphosphatidylinositol-anchor. *Molecular Biology of the Cell*, 14(10), 4285-4295. <https://doi.org/10.1091/mbc.e03-03-0193>

Nafa, K., Mason, P. J., Hillmen, P., Luzzatto, L., & Bessler, M. (1995). Mutations in the PIG-A gene causing paroxysmal nocturnal hemoglobinuria are mainly of the frameshift type. *Blood*, 86(12), 4650-4655.

Nakamura, N., Inoue, N., Watanabe, R., Takahashi, M., Takeda, J., Stevens, V. L., & Kinoshita, T. (1997). Expression cloning of PIG-L, a candidate N-acetylglucosaminyl-phosphatidylinositol deacetylase. *The Journal of Biological Chemistry*, 272(25), 15834-15840. <https://doi.org/10.1074/jbc.272.25.15834>

Noël, C. J., Diaz, N., Sicheritz-Ponten, T., Safarikova, L., Tachezy, J., Tang, P., Fiori, P.-L., & Hirt, R. P. (2010). *Trichomonas vaginalis* vast BspA-like gene family: Evidence for functional diversity from structural organisation and transcriptomics. *BMC Genomics*, *11*(1), 99. <https://doi.org/10.1186/1471-2164-11-99>

Nosjean, O., Briolay, A., & Roux, B. (1997). Mammalian GPI proteins: Sorting, membrane residence and functions. *Biochimica Et Biophysica Acta*, *1331*(2), 153-186. [https://doi.org/10.1016/s0304-4157\(97\)00005-1](https://doi.org/10.1016/s0304-4157(97)00005-1)

Omasits, U., Ahrens, C. H., Müller, S., & Wollscheid, B. (2014). Protter: Interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*, *30*(6), 884-886. <https://doi.org/10.1093/bioinformatics/btt607>

Paulick, M. G., & Bertozzi, C. R. (2008). The Glycosylphosphatidylinositol Anchor: A Complex Membrane-Anchoring Structure for Proteins. *Biochemistry*, *47*(27), 6991-7000. <https://doi.org/10.1021/bi8006324>

Peter, O., & Menon, A. K. (2007). Thematic review series: Lipid Posttranslational Modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycopospholipids. *Journal of Lipid Research*, *48*(5), 993-1011. <https://doi.org/10.1194/jlr.R700002-JLR200>

Piłyk, S., Perlinska-Lenart, U., Janik, A., Gryz, E., Ajchler-Adamska, M., & Kruszewska, J. S. (2020). Yil102c-A is a Functional Homologue of the DPMII Subunit of Dolichyl Phosphate Mannose Synthase in *Saccharomyces cerevisiae*. *International Journal of Molecular Sciences*, *21*(23), 8938. <https://doi.org/10.3390/ijms21238938>

Protein Parameters. (s. f.). Recuperado 17 de enero de 2023, de <http://www.protparam.net/index.html>

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory.

Shah, N., & Bhatt, H. (2022). Paroxysmal Nocturnal Hemoglobinuria. En *StatPearls [Internet]*. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK562292/>

Sharma, D. K., Vidugiriene, J., Bangs, J. D., & Menon, A. K. (1999). A cell-free assay for glycosylphosphatidylinositol anchoring in African trypanosomes. Demonstration of a transamidation reaction mechanism. *The Journal of Biological Chemistry*, *274*(23), 16479-16486. <https://doi.org/10.1074/jbc.274.23.16479>

Socié, G., Mary, J. Y., de Gramont, A., Rio, B., Leporrier, M., Rose, C., Heudier, P., Rochant, H., Cahn, J. Y., & Gluckman, E. (1996). Paroxysmal nocturnal haemoglobinuria: Long-term follow-up and prognostic factors. French Society of Haematology. *Lancet (London, England)*, *348*(9027), 573-577. [https://doi.org/10.1016/s0140-6736\(95\)12360-1](https://doi.org/10.1016/s0140-6736(95)12360-1)

Socié, G., Schrezenmeier, H., Muus, P., Lisukov, I., Röth, A., Kulasekararaj, A., Lee, J. W., Araten, D., Hill, A., Brodsky, R., Urbano-Ispizua, A., Szer, J., Wilson, A., Hillmen, P., & PNH Registry. (2016). Changing prognosis in paroxysmal nocturnal haemoglobinuria disease subcategories: An analysis of the International PNH Registry. *Internal Medicine Journal*, *46*(9), 1044-1053. <https://doi.org/10.1111/imj.13160>

Stenson, P. D., Mort, M., Ball, E. V., Chapman, M., Evans, K., Azevedo, L., Hayden, M., Heywood, S., Millar, D. S., Phillips, A. D., & Cooper, D. N. (2020). The Human Gene Mutation Database (HGMD®): Optimizing its use in a clinical diagnostic or research setting. *Human Genetics*, *139*(10), 1197-1207. <https://doi.org/10.1007/s00439-020-02199-3>

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification*, *41*(1), 207-234. <https://doi.org/10.1016/j.pep.2005.01.016>

Swoboda, K. J., Margraf, R. L., Carey, J. C., Zhou, H., Newcomb, T. M., Coonrod, E., Durtschi, J., Mallempati, K., Kumanovics, A., Katz, B. E., Voelkerding, K. V., & Opitz, J. M. (2014). A novel germline PIGA mutation in Ferro-Cerebro-Cutaneous syndrome: A neurodegenerative X-linked epileptic encephalopathy with systemic iron-overload. *American Journal of Medical Genetics. Part A*, *164A*(1), 17-28. <https://doi.org/10.1002/ajmg.a.36189>

Takahashi, M., Inoue, N., Ohishi, K., Maeda, Y., Nakamura, N., Endo, Y., Fujita, T., Takeda, J., & Kinoshita, T. (1996). PIG-B, a membrane protein of the endoplasmic reticulum with a large luminal domain, is involved in transferring the third mannose of the GPI anchor. *The EMBO Journal*, *15*(16), 4254-4261.

The UniProt Consortium. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), D523-D531. <https://doi.org/10.1093/nar/gkac1052>

Tiede, A., Bastisch, I., Schubert, J., Orlean, P., & Schmidt, R. E. (1999). Biosynthesis of glycosylphosphatidylinositols in mammals and unicellular microbes. *Biological Chemistry*, *380*(5), 503-523. <https://doi.org/10.1515/BC.1999.066>

Tiede, A., Daniels, R. J., Higgs, D. R., Mehrein, Y., Schmidt, R. E., & Schubert, J. (2001). The human GPI1 gene is required for efficient glycosylphosphatidylinositol biosynthesis. *Gene*, 271(2), 247-254. [https://doi.org/10.1016/s0378-1119\(01\)00510-8](https://doi.org/10.1016/s0378-1119(01)00510-8)

UniProt. (s. f.). Recuperado 14 de julio de 2023, de <https://www.uniprot.org/>

Van Etten, J., Cho, C. H., Yoon, H. S., & Bhattacharya, D. (2023). Extremophilic red algae as models for understanding adaptation to hostile environments and the evolution of eukaryotic life on the early earth. *Seminars in Cell & Developmental Biology*, 134, 4-13. <https://doi.org/10.1016/j.semcd.2022.03.007>

Vishwakarma, R. A., & Menon, A. K. (2005). Flip-flop of glycosylphosphatidylinositols (GPI's) across the ER. *Chemical Communications*, 4, 453-455. <https://doi.org/10.1039/B413196G>

Wang, S., Li, W., Liu, S., & Xu, J. (2016). RaptorX-Property: A web server for protein structure property prediction. *Nucleic Acids Research*, 44(Web Server issue), W430-W435. <https://doi.org/10.1093/nar/gkw306>

Watanabe, R., Kinoshita, T., Masaki, R., Yamamoto, A., Takeda, J., & Inoue, N. (1996). PIG-A and PIG-H, which participate in glycosylphosphatidylinositol anchor biosynthesis, form a protein complex in the endoplasmic reticulum. *The Journal of Biological Chemistry*, 271(43), 26868-26875. <https://doi.org/10.1074/jbc.271.43.26868>

Watanabe, R., Murakami, Y., Marmor, M. D., Inoue, N., Maeda, Y., Hino, J., Kangawa, K., Julius, M., & Kinoshita, T. (2000). Initial enzyme for glycosylphosphatidylinositol biosynthesis requires PIG-P and is regulated by DPM2. *The EMBO Journal*, 19(16), 4402-4411. <https://doi.org/10.1093/emboj/19.16.4402>

Watanabe, R., Ohishi, K., Maeda, Y., Nakamura, N., & Kinoshita, T. (1999). Mammalian PIG-L and its yeast homologue Gpi12p are N-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. *The Biochemical Journal*, 339 (Pt 1), 185-192.

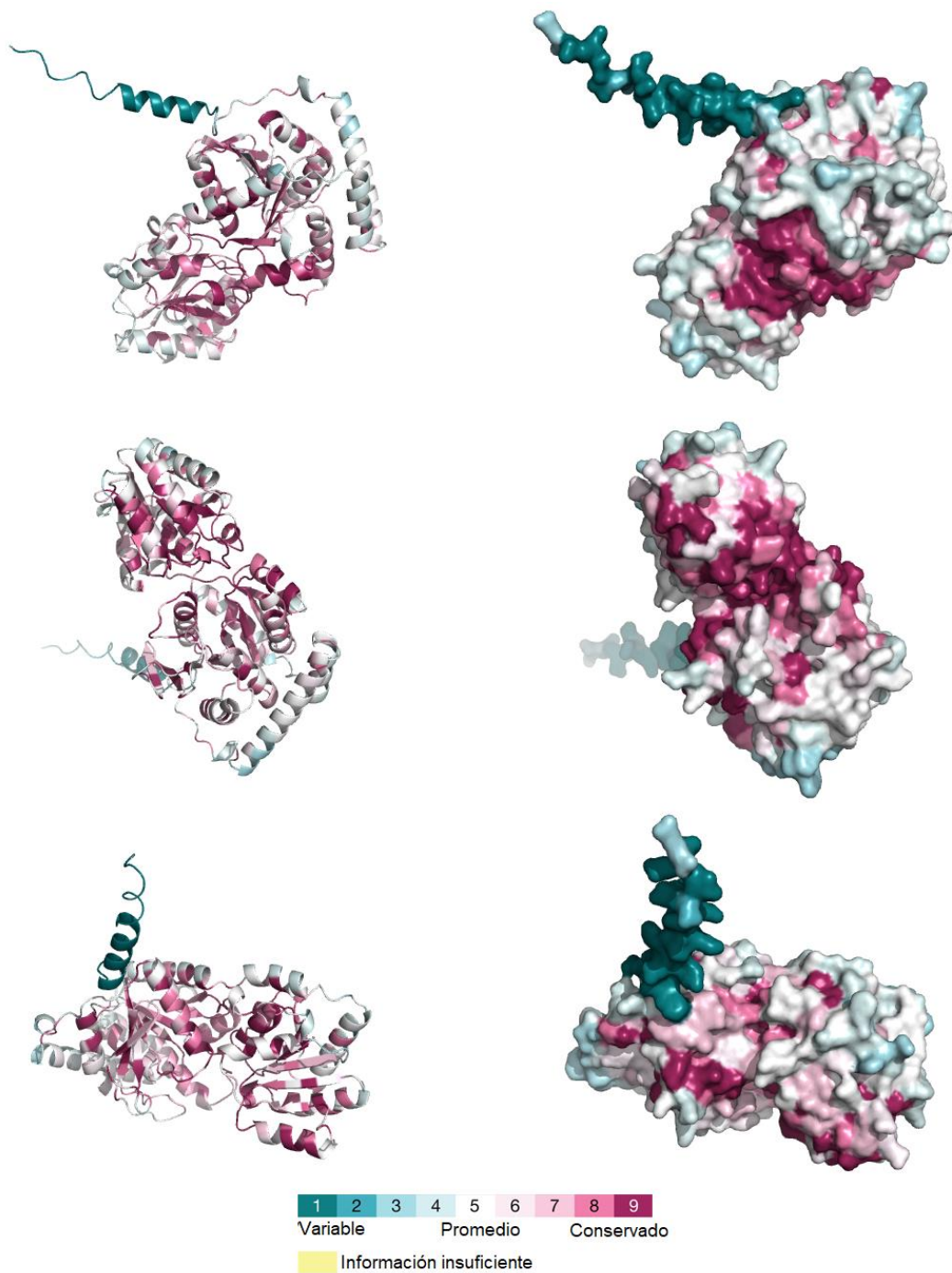
Welner, D. H., Shin, D., Tomaleri, G. P., DeGiovanni, A. M., Tsai, A. Y.-L., Tran, H. M., Hansen, S. F., Green, D. T., Scheller, H. V., & Adams, P. D. (2017). Plant cell wall glycosyltransferases: High-throughput recombinant expression screening and general requirements for these challenging enzymes. *PLOS ONE*, 12(6), e0177591. <https://doi.org/10.1371/journal.pone.0177591>

Yan, B. C., Westfall, B. A., & Orlean, P. (2001). Ynl038wp (Gpi15p) is the *Saccharomyces cerevisiae* homologue of human Pig-Hp and participates in the first step in glycosylphosphatidylinositol assembly. *Yeast*, 18(15), 1383-1389. <https://doi.org/10.1002/yea.783>

Yu, F., Du, Y., & Han, B. (2016). A comparative analysis of clinical characteristics of patients with paroxysmal nocturnal hemoglobinuria between Asia and Europe/America. *International Journal of Hematology*, 103(6), 649-654. <https://doi.org/10.1007/s12185-016-1995-1>

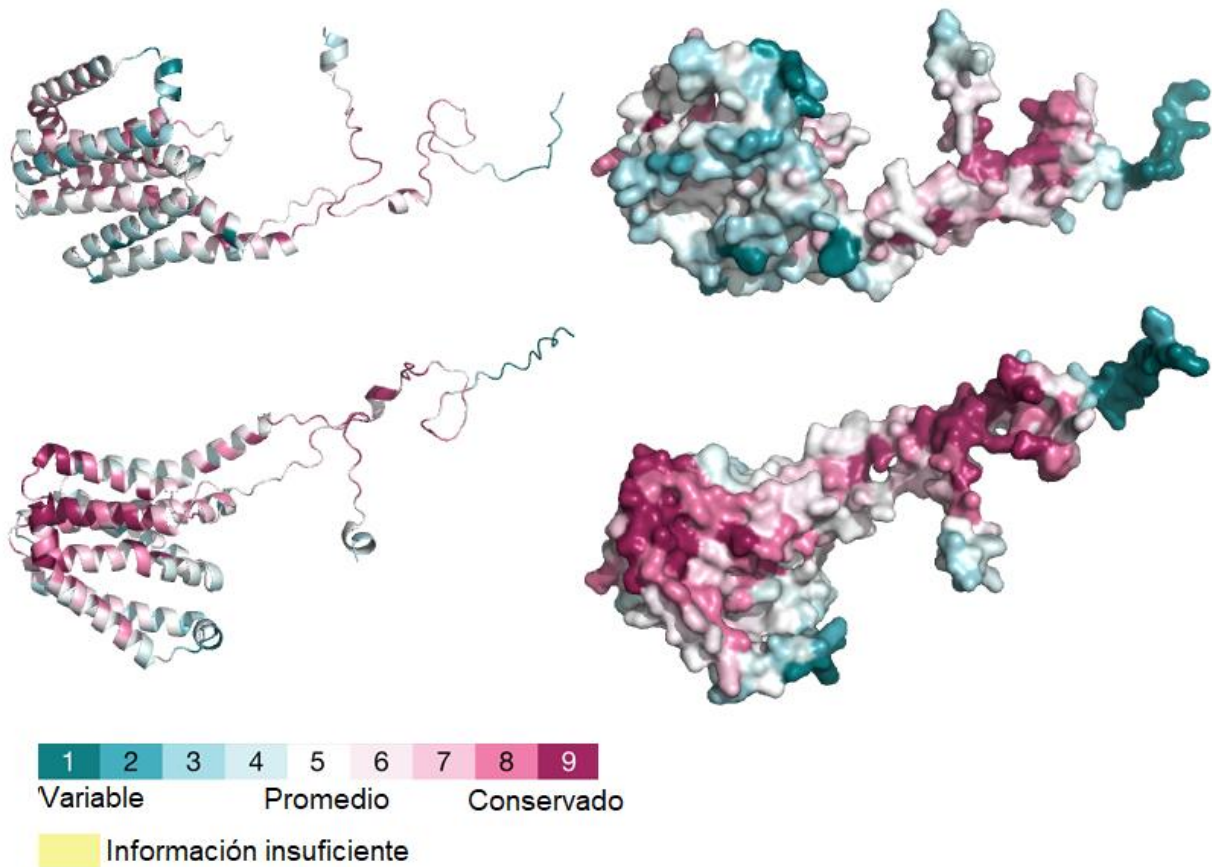
ANEXO

**Análisis de las estructuras de las subunidades del complejo GPI-GnT: imágenes suplementarias
Suplementaria 1**



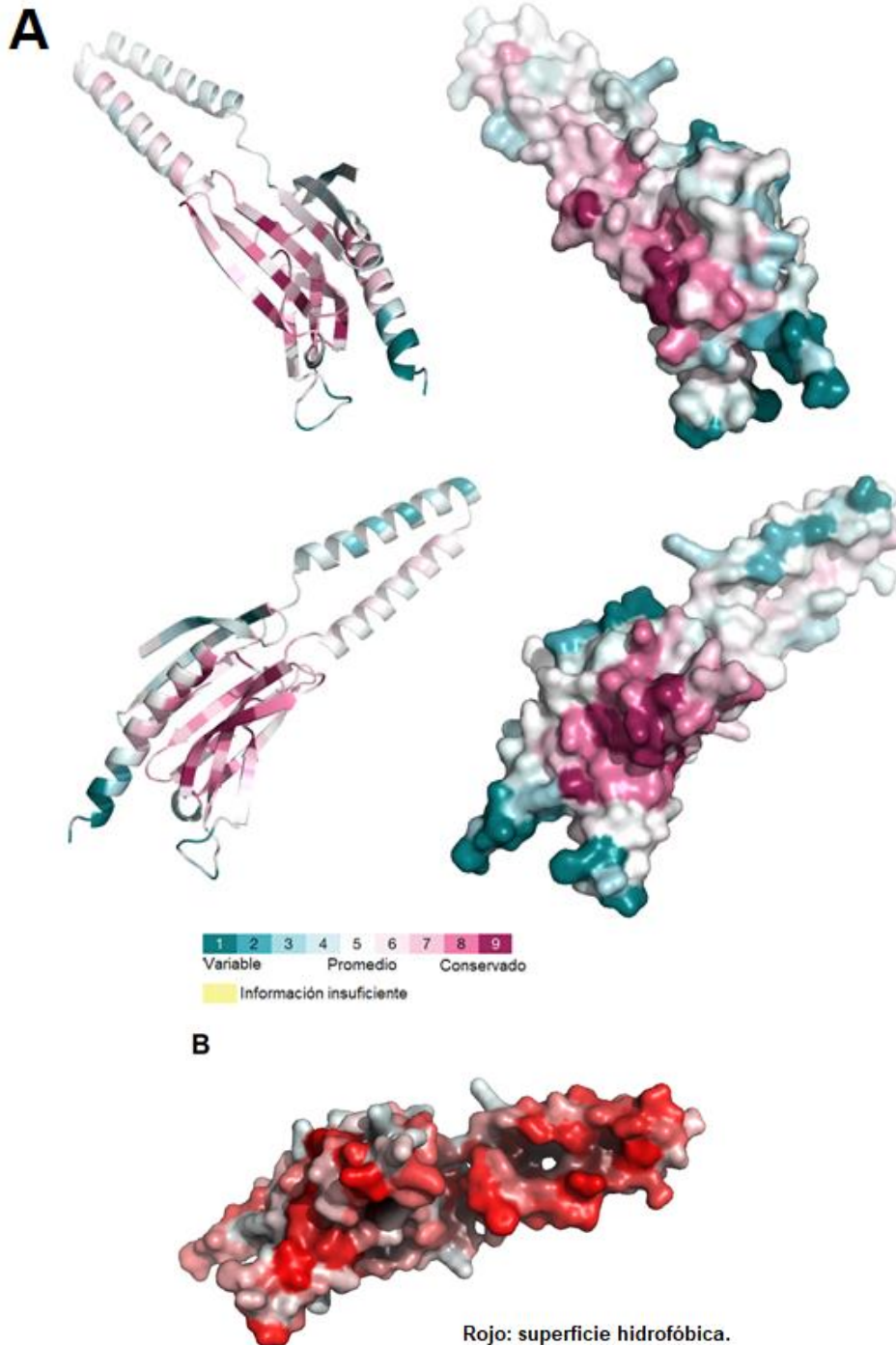
Suplementaria 1. Mapeo de la conservación de la secuencia en la estructura de PIGA. Se muestran distintas orientaciones de la subunidad PIGA. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia).

Suplementarias 2



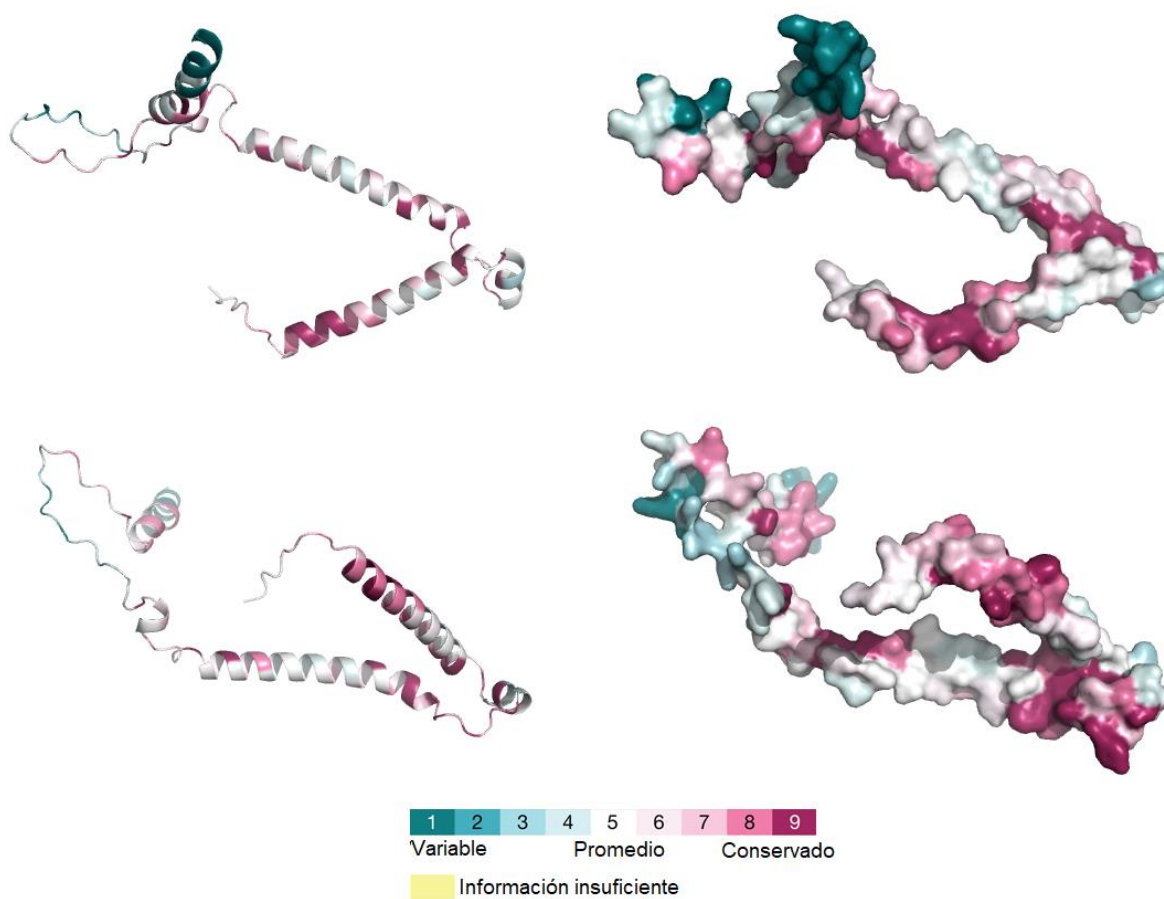
Suplementaria 2. Mapeo de la conservación de la secuencia en la estructura de PIGC. Se muestran distintas orientaciones de la subunidad PIGC. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia).

Suplementarias 3



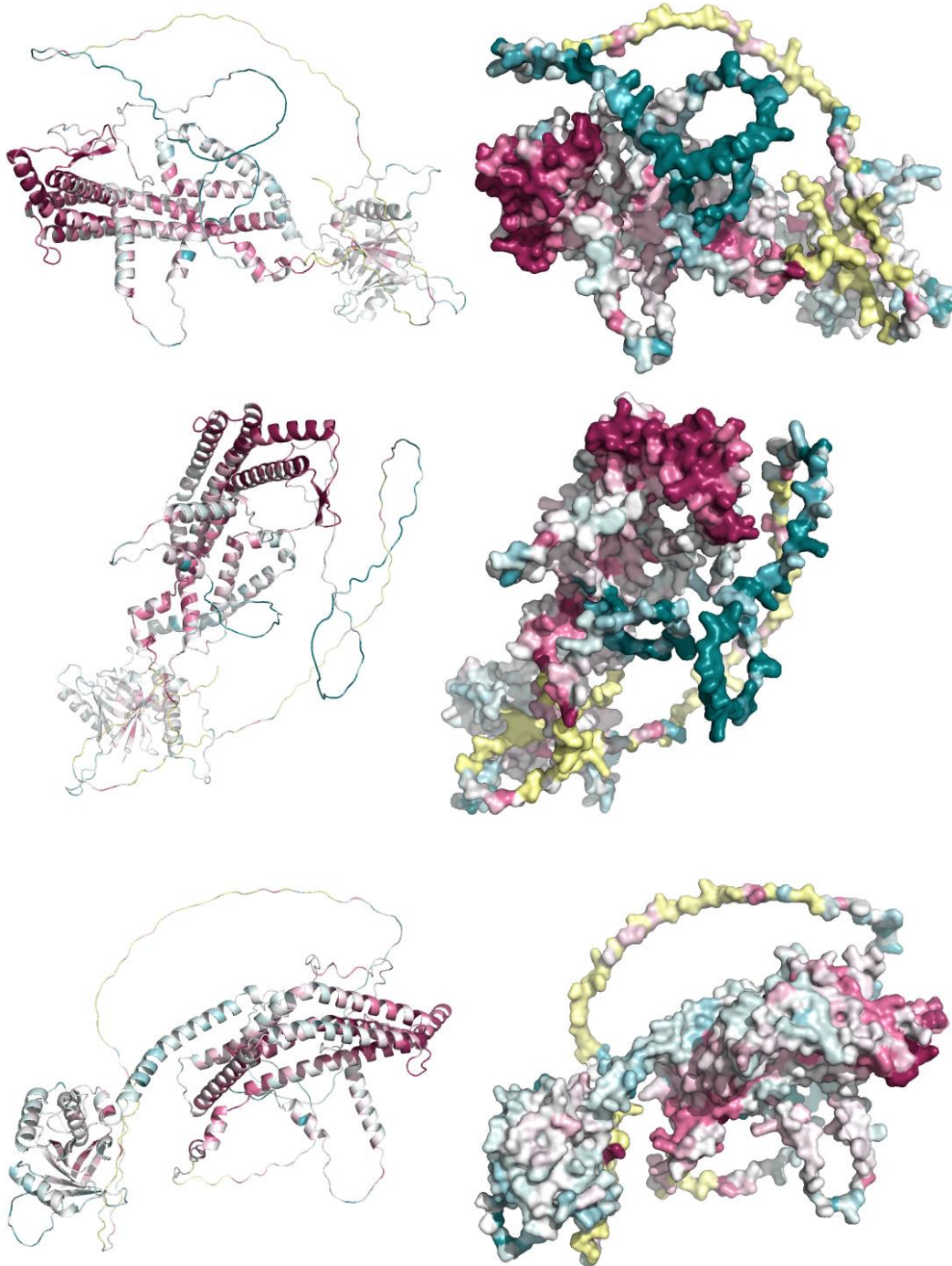
Suplementaria 3. Panel A: Mapeo de la conservación de la secuencia en la estructura de PIGH. Se muestran distintas orientaciones de la subunidad PIGH. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia). **Panel B: Hidrofobicidad de la estructura de PIGH.** La hidrofobicidad varía según la intensidad del color rojo (más intenso: más hidrofóbico, menos intenso: menos hidrofóbico).

Suplementarias 4



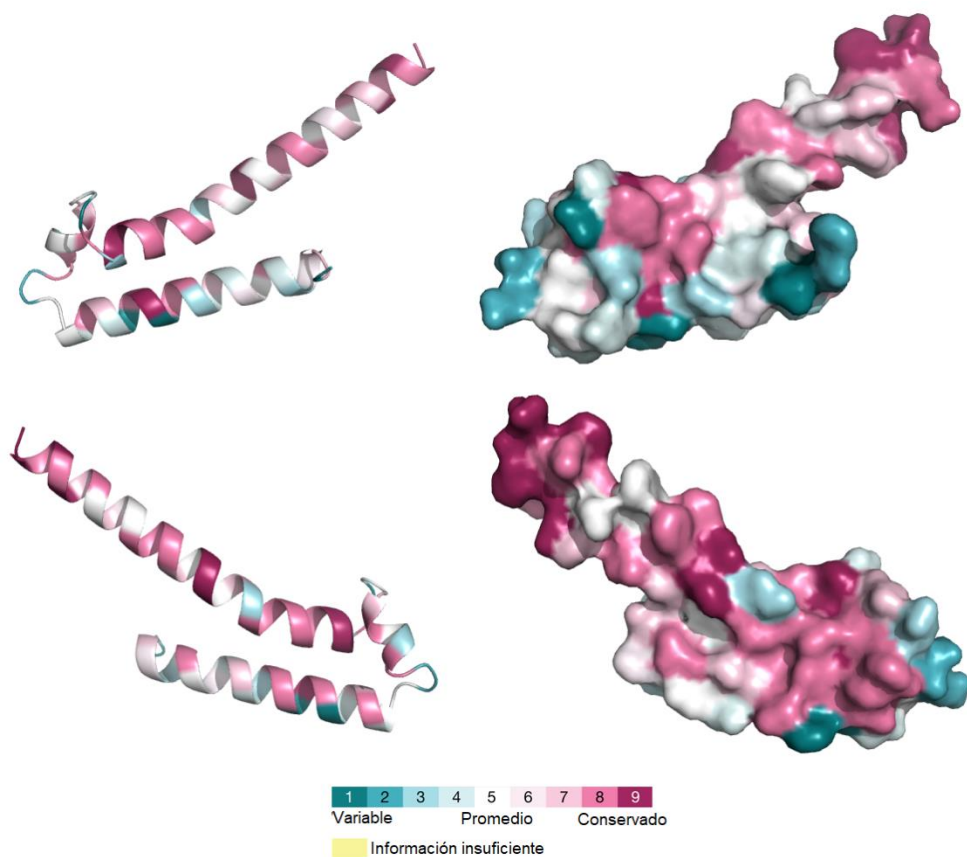
Suplementaria 4. Mapeo de la conservación de la secuencia en la estructura de PIGP. Se muestran distintas orientaciones de la subunidad PIGP. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia).

Suplementaria 5



Suplementaria 5. Mapeo de la conservación de la secuencia en la estructura de PIGQ. Se muestran distintas orientaciones de la subunidad PIGQ. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia).

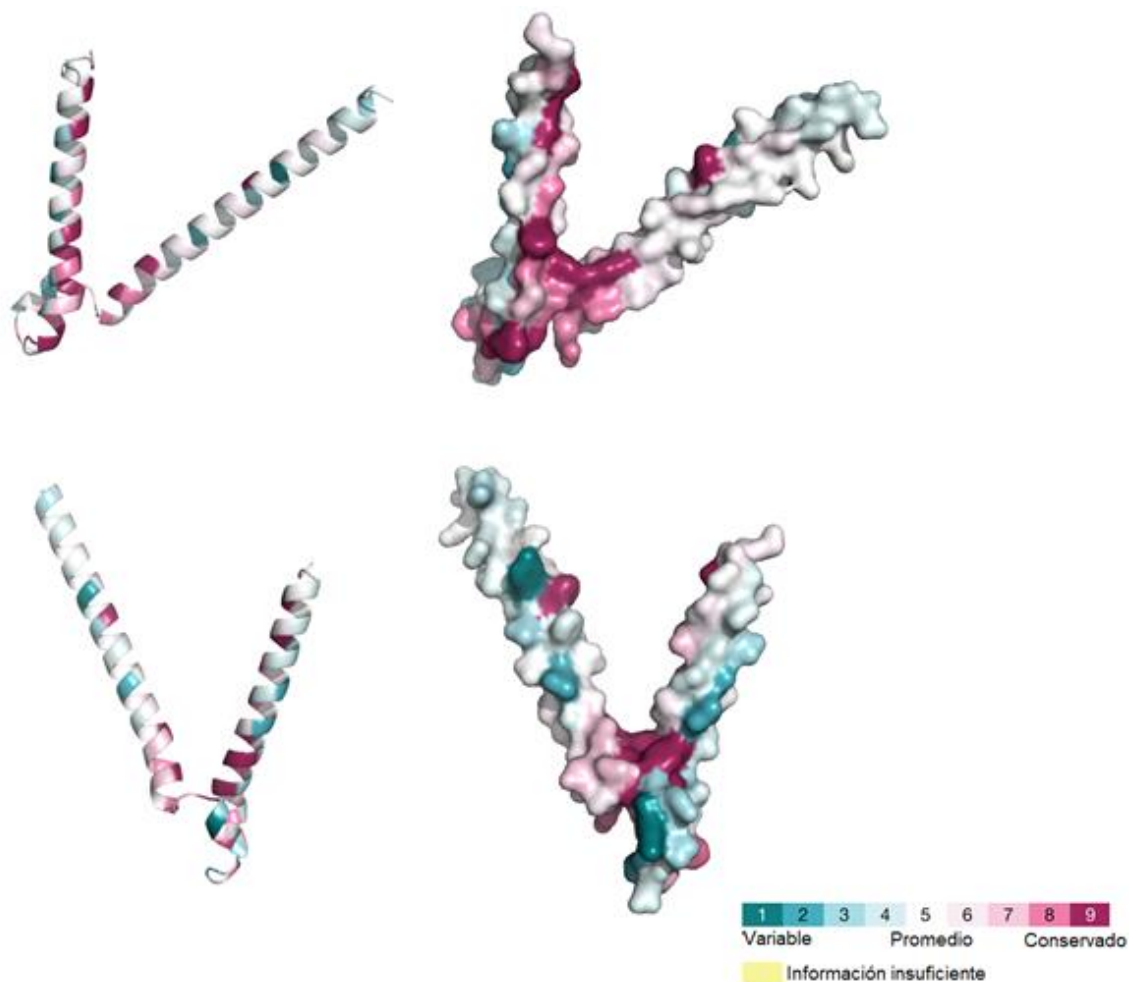
Suplementaria 6



Suplementaria 6. Mapeo de la conservación de la secuencia en la estructura de PIGY. Se muestran distintas orientaciones de la subunidad PIGY. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia).

Suplementaria 7

A

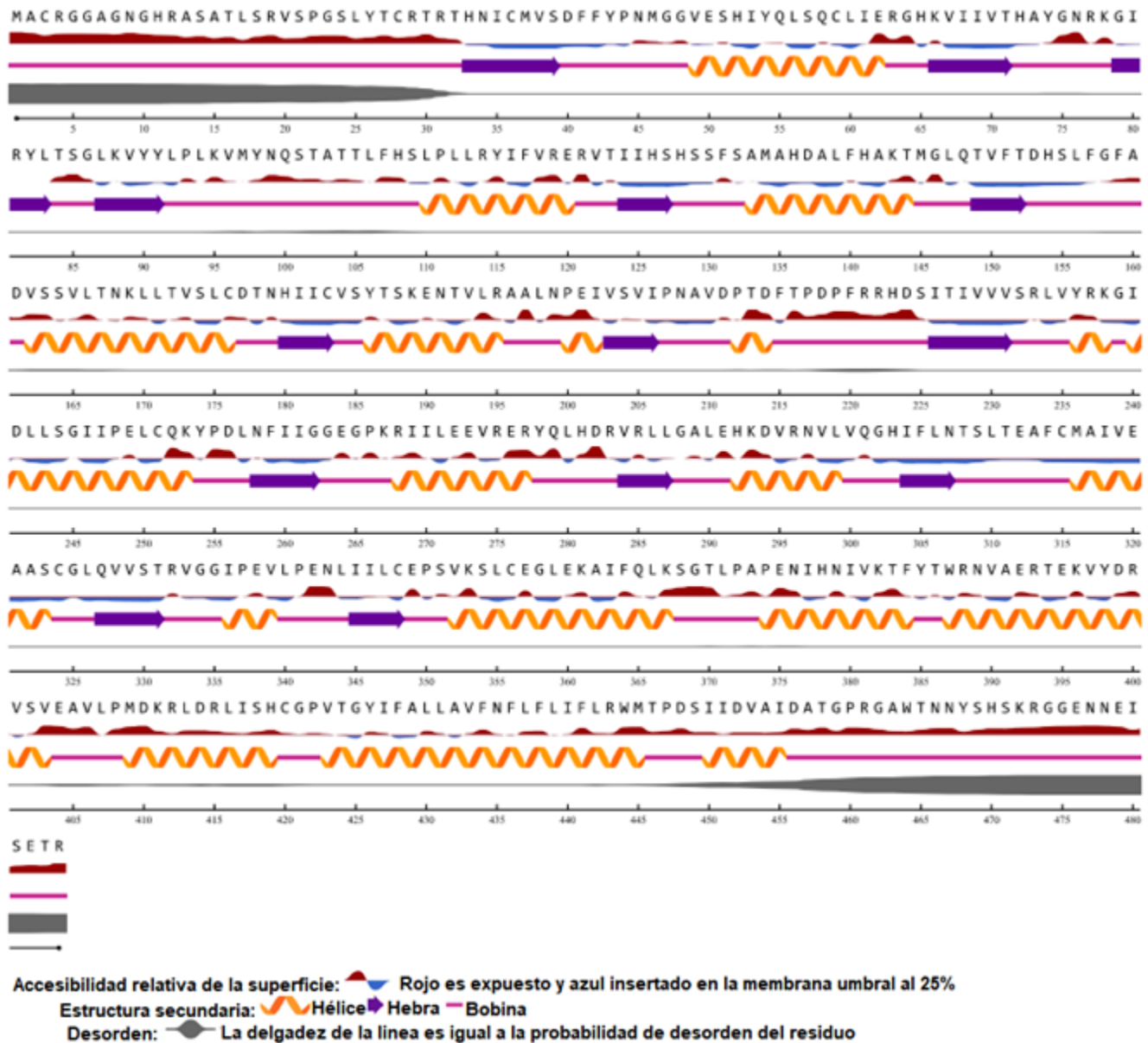


B



Suplementaria 7. Mapeo de la conservación de la secuencia en la estructura de DPM2. Panel A: Se muestran distintas orientaciones de la subunidad DPM2. La estructura se muestra en forma de *cartoon* (izquierda) y su superficie (derecha). El grado de conservación varía según la escala de colores de más variable (verde) a más conservado (fucsia). Panel B: Grupo de residuos aromáticos muy conservados.

Predicción de la estructura secundaria de PIGA realizada con NetSurfP3.0



Suplementaria 8. Predicción de la estructura secundaria de PIGA realizada con NetSurfP3.0. El servidor predice la accesibilidad de la secuencia, la estructura secundaria y el grado de desorden. En el recuadro se indican las referencias para cada una de estas características.

Enfermedades/síndromes asociados a las subunidades del complejo GPI-GnT

(Información extraída del servidor Uniprot (<https://www.uniprot.org/>))

Enfermedades/síndromes asociados a PIGA:

Síndrome de anomalías congénitas múltiples – hipotonía – convulsiones tipo 2 (MCAHS 2): Trastorno recesivo del desarrollo ligado al cromosoma X caracterizado por rasgos dismórficos, hipotonía neonatal, convulsiones mioclónicas y anomalías congénitas variables que afectan a los sistemas nervioso central, cardíaco y urinario. La mayoría de los individuos afectados mueren en la infancia (MIM:300818)

Hemoglobinuria Paroxística Nocturna (HPN): Trastorno caracterizado por anemia hemolítica con hemoglobinuria, trombosis y deficiencia de la hematopoyesis. La descomposición de los glóbulos rojos con liberación de hemoglobina en la orina se manifiesta sobre todo por una orina de color oscuro por la mañana (MIM:300868).

Trastorno del Neurodesarrollo con Epilepsia y Hemocromatosis (NEDEPH): Trastorno recesivo ligado al cromosoma X caracterizado por retraso grave del desarrollo, discapacidad intelectual, epilepsia de aparición precoz y sobrecarga sistémica precoz de hierro que da lugar a hemocromatosis de aparición juvenil. Otros rasgos adicionales variables pueden incluir contracturas articulares, discapacidad visual o auditiva y anomalías cutáneas (MIM:301072).

Enfermedades/síndromes asociados a PIGC:

Defecto en la biosíntesis de GPI 16: Trastorno autosómico recesivo caracterizado por retraso en el desarrollo psicomotor, discapacidad intelectual y convulsiones (MIM:617816).

Enfermedades/síndromes asociados a PIGH:

Defecto en la biosíntesis de GPI 17: Trastorno autosómico recesivo caracterizado por déficits neurológicos variables que se manifiestan en la infancia o en la niñez temprana. Los rasgos clínicos incluyen dificultades de aprendizaje, retraso del desarrollo de leve a moderado, convulsiones de gravedad variable, comportamiento agresivo o excesivamente amistoso y rasgos autistas (MIM:618010).

Enfermedades/síndromes asociados a PIGP:

Encefalopatía Epiléptica Infantil Temprana 55 (EIEE 55): Forma de encefalopatía epiléptica, un grupo heterogéneo de epilepsias graves de aparición temprana caracterizadas por crisis refractarias, alteraciones del neurodesarrollo y mal pronóstico. El desarrollo es normal antes de la aparición de las crisis, tras lo cual se hacen evidentes los retrasos cognitivos y motores. La EIEE 55 es una enfermedad autosómica recesiva (MIM:617599).

Enfermedades/síndromes asociados a PIGQ:

Encefalopatía Epiléptica Infantil Temprana 77 (EIEE 77): Su nombre alternativo es Síndrome de anomalías congénitas múltiples-hipotonía-convulsiones 4 (MCAHS 4). Es un síndrome autosómico recesivo caracterizado por la aparición de convulsiones refractarias en los primeros meses de vida. Otros rasgos clínicos incluyen un retraso global grave del desarrollo, rasgos faciales dismórficos y anomalías esqueléticas, renales y oftálmicas. A nivel celular, el trastorno está causado por un defecto en la síntesis de GPI (MIM:618548).

Enfermedades/síndromes asociados a PIGY:

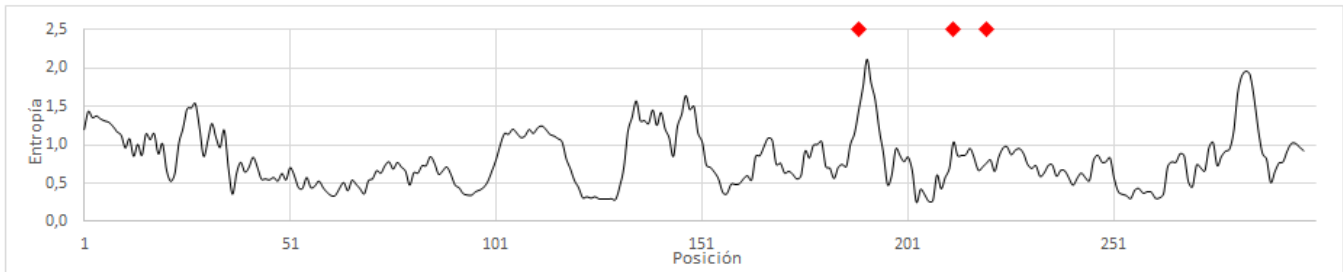
Síndrome de Hiperfosfatasa con Retraso Mental 6 (HPMRS 6): Trastorno multisistémico autosómico recesivo caracterizado por un retraso grave del desarrollo, dismorfismo, convulsiones, cataratas y muerte prematura en algunos pacientes (MIM:616809).

Enfermedades/síndromes asociados a DPM2:

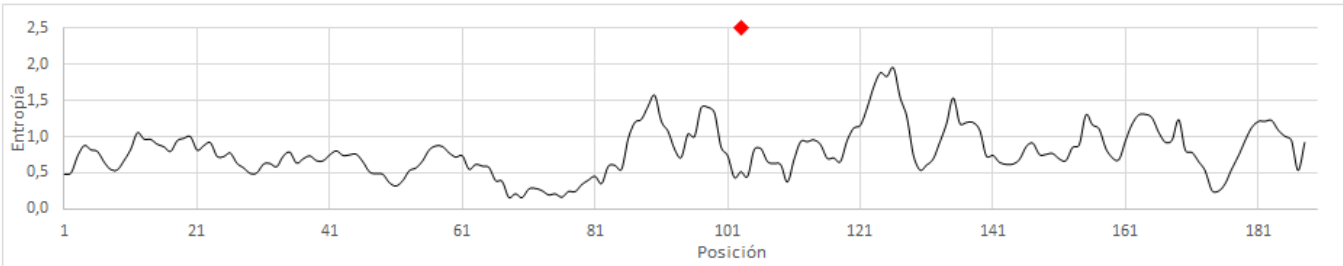
Trastorno Complejo de la Glicosilación Tipo Ie: Una forma de trastorno congénito de la glicosilación, un trastorno multisistémico causado por un defecto en la biosíntesis de glicoproteínas y caracterizado por glicoproteínas séricas infraglicosiladas. Los trastornos congénitos de la glicosilación dan lugar a una amplia variedad de características clínicas, como defectos en el desarrollo del sistema nervioso, retraso psicomotor, rasgos dismórficos, hipotonía, trastornos de la coagulación e inmunodeficiencia. El amplio espectro de características refleja el papel crítico de las N-glicoproteínas durante el desarrollo embrionario, la diferenciación y el mantenimiento de las funciones celulares. Algunos pacientes CDG1U presentan cambios distróficos observados en la biopsia muscular y una reducción de O-manosilglicanos en el alfa-distroglicano (MIM:615042).

Gráficos de entropía en función de la posición de las mutaciones en la secuencia

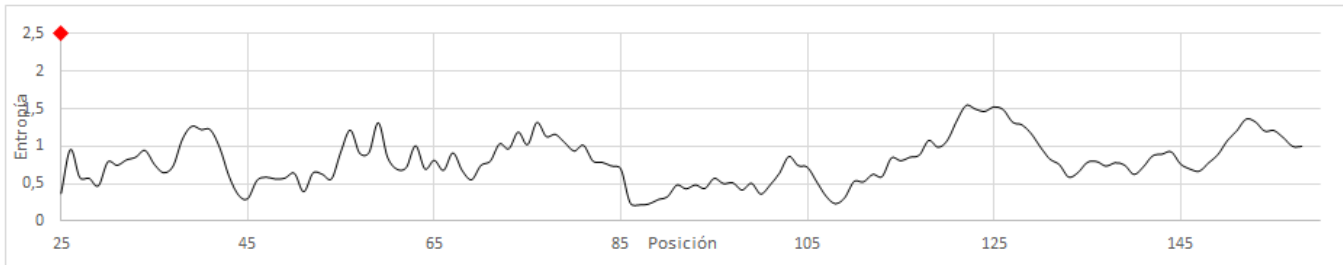
PIGC



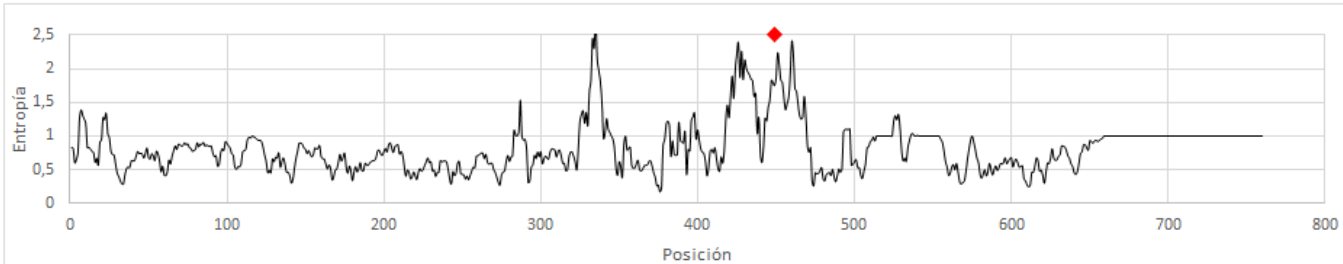
PIGH



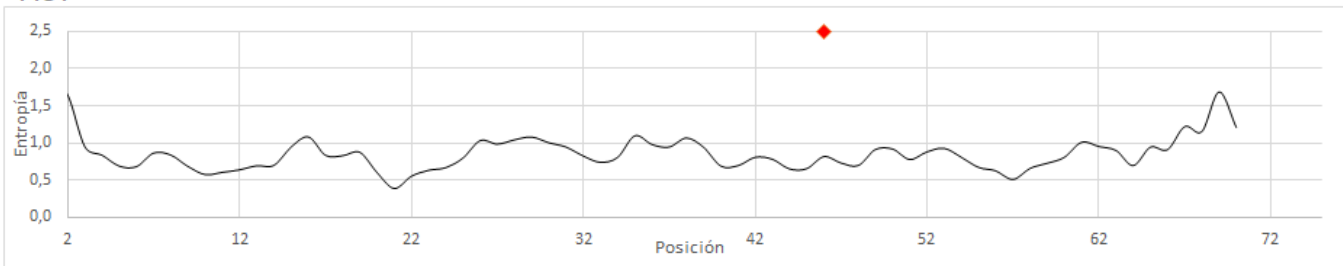
PIGP



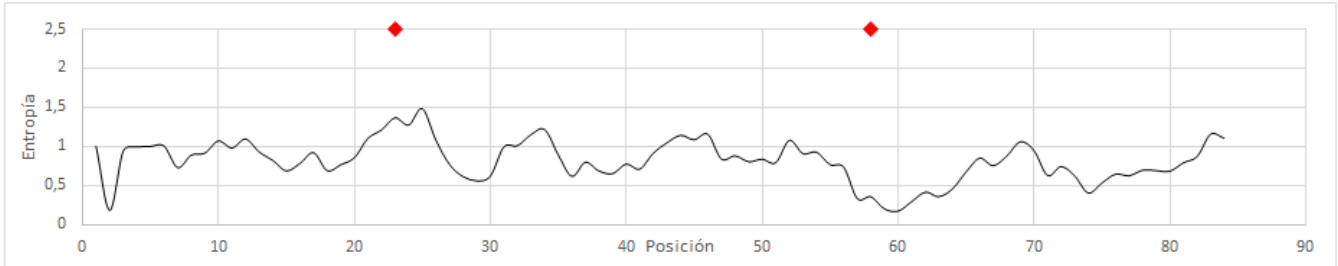
PIGQ



PIGY



DPM2



Suplementaria 9. Gráfico de entropía en función de la posición de las mutaciones (comienza en la página anterior). Para cada subunidad se graficó la entropía (relacionada con el grado de conservación de la secuencia) en función de la posición de las mutaciones halladas en las bases de datos. Los puntos (en rojo) en la parte superior del gráfico representan la posición de las mutaciones. Se graficaron solo las mutaciones patogénicas y probablemente patogénicas.