



# UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales  
Departamento de Química Biológica

## ESTUDIOS de FUNCIÓN y ESTRUCTURA EN GumK, UNA GLUCURONOSILTRANSFERASA INVOLUCRADA EN LA BIOSÍNTESIS del XANTANO

Tesis presentada para optar al título de Doctor de la Universidad de Buenos  
Aires, área Química Biológica

Lugar de Trabajo: Fundación Instituto Leloir

Lic. Máximo Barreras

Director de Tesis: Dr. Luis Ielpi

2007

## AGRADECIMIENTOS

A Luis Ielpi, por haberme brindado la oportunidad de empezar a aprender Biología Molecular y Biofísica en su Lab, brindándome el apoyo y las herramientas necesarias para hacerlo. También por haberme bancado en todos los momentos, sin excepciones.

A Patri, porque compartimos muchas cosas (el tema de trabajo, todas las horas que estamos en el lab, las dudas y las respuestas en el trabajo de mesada) y lo hicimos de una manera muy honesta y armónica, en un clima de trabajo verdaderamente distendido y alegre. Por los muy buenos recuerdos.

A Mario Bianchet, por la enorme solidaridad de alojarme unos meses en su casa y haberme dado una gran ayuda en la concreción de la parte cristalográfica de esta Tesis.

A Mario Amzel, por permitirme trabajar en su laboratorio en Johns Hopkins.

A Matías, en lo laboral, por la ayuda con las mutantes, la preparación de aceptor y tantas otras cosas. En lo cotidiano por el buen e inteligente humor, una cualidad que aprecio mucho. Y por el regalo que seguro me trae a la vuelta de Europa!

A Silvina, por la ayuda con las complementaciones de las mutantes y por transmitir serenidad y buena onda en el lab, otra cualidad que también aprecio mucho.

A Gastón Mayol, que tempranamente me desasnó en el uso del FPLC y por la ayuda con la purificación de aceptor. Por su calidad como persona, que nos hace sentir siempre cómodos y bienvenidos a todos.

A Patricio Craig, por la ayuda con el Light Scattering.

A Seba Klinke, por el ida y vuelta con las dudas cristalográficas que nos aquejan recurrentemente.

A Benjamín y a Juan, por la ayuda con Linux.

A Vero Iemini, por la ayuda con diferentes técnicas y protocolos, cuando recién había llegado a este lab.

A los pibes: Ramiro, Gastón, Juan, Martín, Mariano porque pudimos armar un grupo distendido y la pasamos muy bien comiendo al mediodía o tomando una cerveza en el Curie y/o cualquier otro piringundín y hablando pavadas. Es una enorme suerte estar en un grupo así.

A los padres de: El Ñato, El Negro, Micu, Lucho, Riki, Seba, Benito, Sarcelo y El Mendu, por haber gestado a éstos, mis -desperdigados por todo el país y también otras latitudes- amigos de siempre.

A Marce, que siempre, siempre, me apoyó para la concreción de esta Tesis.

**A Marce, mi preciosa**

**A mis hermanos, a mis viejos, siempre presentes,  
apoyándome**

# ÍNDICE GENERAL

<b>1. Resumen</b>	1
<b>2. Abstract</b>	2
<b>3. Abreviaturas</b>	3
<b>4. Introducción</b>	5
4.1 Relevancia biológica de las glicosiltransferasas	5
4.2 Clasificación	5
4.3 Mecanismos de reacción	9
4.3.1 Mecanismo de inversión	9
4.3.2 Mecanismo de retención	10
4.4 Descripción estructural	11
4.4.1 Sitios de unión de los sustratos	11
4.4.2 Utilización de iones divalentes	11
4.4.3 Plegamiento y dominios conservados	12
4.4.3.i Superfamilia GT-A: Clan I (inversión) SpsA	13
4.4.3.ii Clan III (retención) LgtC	14
4.4.3.iii Superfamilia GT-B: Clan II (inversión) MurG	15
4.4.3.iv Clan IV PimA	18
4.4.4 Plegamiento GT-C	19
4.5 Función biológica de GumK: Biosíntesis del xantano	19
<b>5. Objetivos</b>	22
5.1 Objetivos específicos	22
<b>6. Materiales y Métodos</b>	23
6.1 Materiales y procedimientos generales	23
6.2 Obtención de suero inmune contra GumK	23
6.3 Cepas y plásmidos	23
6.4 Construcción de plásmidos	24
6.5 Medios y condiciones de cultivo	25
6.6 Introducción de material genético en <i>E. coli</i> y <i>X. campestris</i>	25
6.7 Cuantificación de la producción de xantano	25
6.8 Fraccionamiento celular y detección de las proteínas	26
6.9 Sobreexpresión y purificación de proteínas	26
6.9.1 Sobreexpresión de GumK-s y GumK	26
6.9.2 Solubilización de GumK a partir de membranas	26
6.9.2.i Solubilización por pH: tratamiento con buffer alcalino o con buffer ácido	26
6.9.2.ii Solubilización con fuerza iónica o por desestabilización de la membrana con EDTA	27

6.9.2.iii Solubilización con detergentes no-iónicos	27
6.9.3 Purificación de GumK solubilizada	27
6.9.4 Purificación de GumK sustituida con selenometionina	28
6.10 Actividad enzimática	28
6.10.1 Preparación del sustrato aceptor glicolípido	28
6.10.2 Preparación de diferentes sustratos aceptores marcados radioactivamente	28
6.10.3 Ensayos de actividad de GumK	29
6.10.4 Preparación de oligosacáridos marcados radioactivamente usados como estándares	30
6.10.5 Tratamientos con $\beta$ -glucuronidasa	30
6.11 Cromatografía en capa delgada (TLC) y detección de oligosacáridos radioactivos	31
6.12 Mutagénesis sitio-dirigida	31
6.12.1 Oligonucleótidos mutagénicos	31
6.12.3 Incubaciones radiactivas para la obtención de parámetros cinéticos de GumK y GumK mutada	32
6.13 Cristalización	32
6.13.1 Apo-proteína	32
6.13.2 Complejo GumK-sustrato	33
6.14 Métodos bioinformáticos	33
6.15 Mediciones de tamaño molecular de GumK en solución	33
6.16 Resolución de la estructura cristalina de GumK	33
6.16.1 Caracterización cristalográfica preliminar del cristal nativo	33
6.16.2 Colección de datos de difracción e información de fases	33
6.16.3 GumK en complejo con UDP	34
<b>7. Resultados</b>	<b>35</b>
7.1 Análisis de la función del gen gumK por complementación funcional	35
7.2 GumK se expresa en <i>Xanthomonas campestris</i> a partir de un ORF mas extendido, localizándose en membranas	36
7.3 Producción y purificación de GumK	37
7.3.1 Solubilización	37
7.3.2 Purificación	39
7.4 Análisis de la actividad enzimática de GumK	40
7.4.1 GumK es una glucuronosiltransferasa	40
7.4.2 Caracterización del enlace glucuronosil-glicosídico	41
7.4.3 Utilización de cationes divalentes	42
7.5 Especificidad de aceptor para GumK	43
7.5.1 Cel-Man	43
7.5.2 Pi-Cel-Man y PPI-Cel-Man	43
7.5.3 Fitanyl-PP-Cel-Man	43
7.5.4 Acetilación natural del sustrato aceptor	44
7.6 Ensayos de actividad con diferentes dadores	45
7.7 Estado oligomérico de GumK	45
7.8 Cristalización de GumK	46

7.9 Estructura cristalográfica de la glucuronosiltransferasa GumK	47
7.9.1 Caracterización cristalográfica preeliminar de GumK nativo	47
7.9.2 Búsqueda experimental de la información de fases	48
7.9.2.i Reemplazo Múltiple Isomorfo (MIR)	48
7.9.2.i.a Obtención de derivados	48
7.9.2.i.ii Dispersión anómala múltiple (MAD)	49
7.9.2.i.ii.a Obtención de derivados de seleno-metionina de GumK	49
7.9.2.i.ii.b MAD con derivados con átomos pesados	51
7.10 Características estructurales de GumK	52
7.11 Estudio cristalográfico de la unión de sustratos	53
7.11.1 Cristalización de GumK en complejo con sus sustratos	53
7.11.2 Estudio del complejo GumK-UDP	53
7.12 Unión del aceptor	56
7.12.1 Purificación del sustrato glicolipídico y experimentos de inmersión	56
7.13 Mutagénesis y estudio de la cinética de GumK salvaje y mutantes puntuales	56
7.14 Análisis de las mutantes de unión por complementación in vivo	58
<b>8. Discusión</b>	59
Caracterización de la función y estructura de GumK	59
8.1 Caracterización del gen gumK	59
8.2 Especificidad de sustrato aceptor en GumK	59
8.3 Localización subcelular de GumK	62
8.4 Estado oligomérico	63
8.5 Los cationes divalentes: Cual es su función?	64
8.6 La especificidad por el sustrato dador	64
8.7 El putativo aminoácido catalítico	66
8.8 Paparazeando: La importancia de tener la foto estructural de las GTs	67
<b>9. Perspectivas</b>	68
<b>10. Bibliografía</b>	69

## 1. RESUMEN

GumK es una glucuronosiltransferasa de *Xanthomonas campestris* involucrada en la biosíntesis del exopolisacárido xantano, siendo la enzima encargada de adicionar el cuarto residuo de azúcar en la síntesis de la unidad repetitiva pentasacarídica de este polímero. Específicamente, GumK transfiere un residuo de ácido glucurónico a partir de su sustrato dador UDP-GlcA a su sustrato glicolipídico aceptor, polisoprenil-pirofosfato-glucosa- $\beta$ -1,4-glucosa- $\alpha$ -1,3-manosa. Se caracterizó funcional y estructuralmente a GumK. La caracterización funcional se dio mediante la sobreexpresión, purificado y ensayos de actividad *in vitro* e *in vivo* de la enzima. Mediante estos ensayos de actividad se estudió la especificidad por sus sustratos. Asimismo se estudió su localización subcelular mediante fraccionamiento y detección por western blots. Al determinarse que esta enzima esta asociada a membranas, se realizaron experimentos de disociación para determinar el modo de unión a la membrana. Encontramos que interacciones hidrofóbicas y electrostáticas están involucradas en esta interacción.

La caracterización estructural involucró la cristalización, colección de datos de difracción de rayos-X y resolución de la estructura de la proteína. La estructura fue resuelta con la información de fases proveniente de un experimento de dispersión anómala múltiple sobre un derivado de GumK con átomos de platino. La proteína mostró una estructura consistente de 2 dominios globulares tipo Rossmann, con una hendidura central que es el sitio de unión de sus sustratos y donde ocurre la catálisis. Además, y mediante experimentos de inmersión de cristales, logramos observar y describir el sitio y las interacciones que median la unión de la porción nucleotídica (UDP) del sustrato dador en la enzima. Para confirmar el rol de los residuos involucrados en la unión del UDP, realizamos mutagénesis sitio-dirigida y análisis de la cinética enzimática de las proteínas mutadas. Además realizamos estudios de complementación *in vivo* con estas mutantes. Estos análisis nos permitieron confirmar el rol de al menos media docena de residuos involucrados en la unión del sustrato dador y postular al aspartato 157 como un posible aminoácido catalítico.

Algunos de los resultados obtenidos en esta tesis fueron publicados en:

“Functional characterization of GumK, a membrane-associated  $\beta$ -glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis”. **Barreras, Máximo**; Abdian, Patricia L. and Ielpi, Luis. (2004) **Glycobiology**. 14. pg 233-241.

“Crystallization and preliminary crystallographic characterization of GumK, a membrane-associated glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis”. **Barreras, Máximo**, M.A. Bianchet, and L. Ielpi, **Acta Crystallograph Sect F Struct Biol Cryst Commun** (2006). 62(Pt 9): p. 880-883.

“Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival.” Germán Dunger, Verónica M. Relling, María Laura Tondo, **Máximo Barreras**, Luis Ielpi, Elena G. Orellano, Jorgelina Ottado. **Arch. Microbiol.** (2007).188(2): p. 127-135.

Palabras clave: glicosiltransferasa, glucuronosiltransferasa, GumK, exopolisacárido, *Xanthomonas campestris*, xantano, poliprenol.

## 2. ABSTRACT

GumK is a glucuronosyltransferase from *Xanthomonas campestris* involved in xanthan biosynthesis. This enzyme is responsible for the addition of the fourth sugar residue in xanthan pentasaccharidic subunit. Specifically, GumK transfers a glucuronic acid residue from the donor substrate UDP-GlcA to the acceptor, poliprenil-diphosphate-glucose- $\beta$ -1,4-glucose- $\alpha$ -1,3-mannose. GumK was functional and structurally characterized. Functional characterization of GumK was performed through overexpression, purification and in *vivo/in vitro* enzymatic assays. Substrate specificity was studied with these *in vitro* assays. Subcellular location was studied with cell fractioning, differential centrifugation and detection with western blots. Dissociating treatments from the membrane fraction were carried out to study the specific type of interaction with the membrane. We found that hydrophobic as well as electrostatic interactions were involved in membrane binding.

Structural characterization was performed after crystallization and X-ray data diffraction collection of GumK crystals. The final structure was solved with phase information obtained from a multiple anomalous diffraction experiment carried out on a platinum derivative crystal. The enzyme showed a structure consisting of two Rossmann-fold domains separated by a deep cleft where substrates bind and catalysis occurs. Furthermore, by performing crystal soaking experiments we were able to find and describe the position of the nucleotide portion (UDP) of the donor substrate and the interactions that stabilize the binding of this substrate. To confirm the role of those residues involved in UDP binding we performed site-specific mutagenesis, kinetic characterization and *in vivo* complementation experiments on these mutants. These experiments allowed us to confirm the role of at least half a dozen residues involved in substrate binding and to postulate aspartic acid 157 as the catalytic residue.

### 3. ABREVIATURAS

Las abreviaturas aquí detalladas son aquellas no contempladas en la lista de abreviaturas de la American Society for Microbiology (<http://jb.asm.org/misc/jbitoa.pdf>)

Ac-CoA: acetil coenzima A  
 Ap: Ampicilina  
 C<sub>55</sub>: cadena carbonada de 55 átomos de carbono  
 CAZY: Carbohydrate active enzymes  
 CBM: módulo de unión de carbohidrato  
 Cel: celobiosa  
 Ci: Curie  
 cpm: cuentas por minuto  
 DO: Densidad óptica  
 NDP: nucleótidos difosfato  
 EPS: exopolisacárido  
 Fitanil-PP: fitanil difosfato  
 Gal: galactosa  
 GalA: ácido galacturónico  
 Glc: glucosa  
 GlcA: ácido glucurónico  
 GlcNAc: N-acetil-glucosamina  
 GT: glicosiltransferasa  
 GT-A: superfamilia de GTs con plegamiento tipo A  
 GT-B: superfamilia de GTs con plegamiento tipo B  
 GT-C: superfamilia de GTs con plegamiento tipo C  
 GH: glicosilhidrolasa  
 Gm: gentamicina  
 IR: Índice de refracción  
 kb: kilobase  
 k<sub>cat</sub>: velocidad de recambio enzimático  
 KDO: ácido 2-ceto-3-deoxi octulosónico  
 km: kanamicina  
 K<sub>m</sub>: constante de Michaelis  
 LPS: lipopolisacárido  
 MAD: Dispersión anómala múltiple  
 Man: Manosa  
 MCS: sitio de clonado múltiple  
 MIR: Reemplazo múltiple isomorfo  
 MyM: Materiales y Métodos  
 nt-azúcar: nucleótido azúcar  
 ORF: Marco de lectura abierto  
 PAGE: electroforesis en gel de poliacrilamida  
 PDB: Protein Data Bank  
 PEG: polietilenglicol  
 Pi: fosfato  
 PPI: pirofosfato  
 PMSF: fenil-metil-sulfonil-fluoruro  
 Poliprenil-PP-Cel-Man: polisoprenil-PP-glucosa-β-1,4-glucosa-α-1,3-manosa  
 Poliprenil-PP-Cel-Man-GlcA: polisoprenil-PP-glucosa-β-1,4-glucosa-α-1,3-manosa- β-1,2-ácido glucurónico  
 Poliprenil-PP-Cel-Man-GlcA-Man: polisoprenil-PP-glucosa-β-1,4-glucosa-α-1,3-manosa- β-1,2-ácido glucurónico-β-1,4-manosa  
 PVDF: polivinil difluoruro  
 Rha: ramnosa

## Abreviaturas

Rif: rifampicina

SDS: dodecil sulfato de sodio

TA: temperatura ambiente

TLC: cromatografía en capa fina

$V_{max}$ : velocidad máxima de reacción

xg: factor de fuerza centrífuga

### 4. INTRODUCCIÓN

#### 4.1 Relevancia biológica de las glicosiltransferasas

En términos de masa, las reacciones de glicotransferencia son las más abundantes sobre la tierra. Las reacciones de glicosilación, de gran importancia biológica tanto en procariotas como en eucariotas, son catalizadas por un grupo numeroso y muy diverso de proteínas, las glicosiltransferasas (GTs). Estas enzimas transfieren azúcares desde un dador activado a un aceptor. Las moléculas dadoras del residuo azúcar pueden ser disacáridos, azúcares-fosfato, lípido-azúcares o más comúnmente nucleótidos-azúcares (nt-azúcar). En cambio, las moléculas aceptoras son de naturaleza variada. Basta considerar que todas las categorías de macromoléculas biológicas (oligosacáridos, lípidos, ácidos nucleicos, proteínas) pueden ser glicosiladas, provocando efectos marcados en su función, para comprender la magnitud de esta diversidad. Los glicopéptidos, glicolípidos y otros glicoconjugados están involucrados en funciones como la infección viral, transducción de señales, procesos inflamatorios, interacciones célula-célula, bacteria-huésped y desarrollo embrionario, entre otros. (1-4). Las GTs constituyen blancos importantes de drogas utilizadas para combatir el cáncer (5,6) así como infecciones causadas por virus, bacterias y hongos (7,8). Estas enzimas también representan una importante herramienta biotecnológica en la síntesis químico-enzimática de oligosacáridos de importancia médica o industrial (9-12). Tomando en cuenta el amplio espectro de estructuras que producen y la diversidad de funciones en que están involucradas, se puede pensar en las GTs como uno de los grupos de enzimas más diverso que existe.

A pesar de sus múltiples funciones y aplicaciones, la caracterización funcional y estructural de las GTs tiene aún poco desarrollo. Esto se debe principalmente a las dificultades técnicas encontradas en la expresión y cristalización de GTs, frecuentemente proteínas de membrana o asociadas a ésta. También la caracterización de mecanismos catalíticos suele ser engorrosa, porque debe realizarse con sustratos a veces químicamente complejos y de difícil obtención y/o síntesis. En cuanto a la caracterización estructural, las dificultades se evidencian si pensamos que la primer estructura cristalina de una GT, la  $\beta$ -glicosiltransferasa del fago T4 BGT, apareció hacia 1994 en complejo con su sustrato UDP-Glc (13) y solo luego de 5 años, en 1999, comenzaron a surgir más estructuras (14,15). Los continuos esfuerzos en pos de una caracterización estructural y bioquímica más profunda de este grupo de enzimas han permitido que en la actualidad contemos con la estructura de 59 glicosiltransferasas correspondientes a 27 familias diferentes. Muchas de estas estructuras están en complejo con su sustrato dador y/o aceptor (16-19). Los datos estructurales han provisto información de plegamiento pero también de especificidad de unión de sustratos. Además, también se pueden desprender de estos estudios estructurales inferencias acerca del mecanismo de catálisis y relaciones evolutivas (12,20).

#### 4.2 Clasificación

Con la incesante aparición de nuevas secuencias genómicas el número de secuencias de GTs ha aumentado notablemente. Se estima que aproximadamente el 1 % de los marcos de lectura abierto de cada genoma está involucrado en la tarea de sintetizar enlaces glicosídicos (21). La difícil caracterización funcional y estructural de GTs ha ampliado enormemente la brecha existente entre el pequeño número de GTs de función conocida y miles de secuencias de putativas GTs. Según la base de datos de Entrez-NCBI esto corresponde en la actualidad a poco más de 40.000 genes en todos los organismos (eucariota, procariota, virus) secuenciados hasta el momento. Está claro en este punto el desfase existente, de 3 ordenes de magnitud, entre el número

de secuencias de GTs en bancos de datos y la cantidad de estructuras descriptas. Para paliar este déficit, en la actualidad contamos con métodos de predicción de estructura por modelado (<http://expasy.org/spdbv>) que permiten predecir razonablemente el plegado de la proteína. Sin embargo, la predicción acerca de la especificidad de sustratos no es posible, pues los métodos predictivos son de difícil aplicación en el caso de las GTs a causa de la marcada divergencia de secuencia que se observa y la casi completa ausencia de secuencias consenso.

Si tenemos en cuenta el criterio funcional existen al menos dos grandes grupos de GTs: las enzimas procesivas, como la celulosa sintasa, que transfieren múltiples residuos de azúcar al aceptor; y las enzimas no procesivas, como las que participan en la síntesis de unidades repetitivas de EPS y catalizan la transferencia de un único residuo azúcar al aceptor (22). Por otro lado, puede hacerse otra gran división entre GTs, basándose en la estereoquímica del enlace formado y del dador del residuo azúcar. Así pueden definirse dos grandes grupos: las GTs que actúan con retención y las que actúan con inversión de la configuración del C-anomérico. Las primeras, catalizan la formación de enlaces glicosídicos de estereoquímica idéntica a la del sustrato dador; mientras que las que actúan con inversión producen enlaces glicosídicos de estereoquímica inversa (23). Asimismo, las GTs, al igual que otros grupos de enzimas, han sido clasificadas basándose en la especificidad por sus sustratos y el producto formado, de acuerdo a las recomendaciones de la Unión Internacional de Bioquímica y Biología Molecular (IUBMB) (<http://www.chem.qmul.ac.uk/iubmb>). En el caso de las GTs, este sistema tiene algunas limitantes: en primer lugar necesita de una caracterización detallada de los sustratos y productos de la enzima para poder asignarle un número EC (Enzyme Commission Number), pero además la clasificación se puede volver confusa si tomamos enzimas que actúan sobre varios sustratos distintos, como es el caso de la glucuronosiltransferasas (24), entre tantos otros. Sumado a esto, tampoco refleja características estructurales y mecánicas de las enzimas.

Como habíamos mencionado, una característica muy importante acerca de este grupo de enzimas es su gran divergencia de secuencia aminoacídica: la homología de secuencia entre enzimas emparentadas no suele superar el 10%, y esto solo en el caso de las muy cercanas. Esto impide su agrupamiento utilizando métodos corrientes de alineamiento de secuencias. Por esta razón se idearon varios sistemas de clasificación para las GTs, atendiendo las especiales características de este grupo de enzimas (25,26). En 1997, Campbell y col. proponen un sistema de clasificación para GTs dependientes de nt-azúcares (27), que constituye el cimiento del sistema más completo de clasificación de GTs con que se cuenta actualmente. Este fue ideado a partir de la utilidad del sistema de clasificación de GHs basado en similitud de secuencia. La principal ventaja de este tipo de sistema es que reúne enzimas en familias poli-específicas y permite distinguir relaciones evolutivas entre familias. La validez de estos sistemas de clasificación ha sido luego confirmada al encontrar un mismo tipo de plegamiento y mecanismo catalítico conservado entre miembros de una familia. La clasificación de Campbell comprende secuencias de GTs provenientes de los casi 600 genomas secuenciados hasta el momento. Las secuencias están agrupadas en 90 familias, un número que seguramente aumentará a medida que se incorporen nuevas secuencias genómicas. El criterio de agrupación por familias incluye estereoquímica del enlace formado, identidad del nt-azúcar transferido y homología de secuencia. Asimismo la clasificación ha sido extendida para integrar GTs que utilizan lípido-azúcares y azúcares-fosfato como dadores activados (y no solo nt-azúcar). Todas estas enzimas catalizan reacciones de O-glicosilación, presumiblemente a través de estados de transición con carácter de ión oxocarbonio e involucran la partida de grupos salientes fosforilados generados a partir de los dadores de azúcar activados. Las enzimas que utilizan disacáridos, oligo o polisacáridos como dadores de azúcar (tal es el caso de ciclodextrina-glucanotransferasas, xiloglucan-endotransferasas, dextran-sacarasas, entre otras) no están incluidas en la clasificación

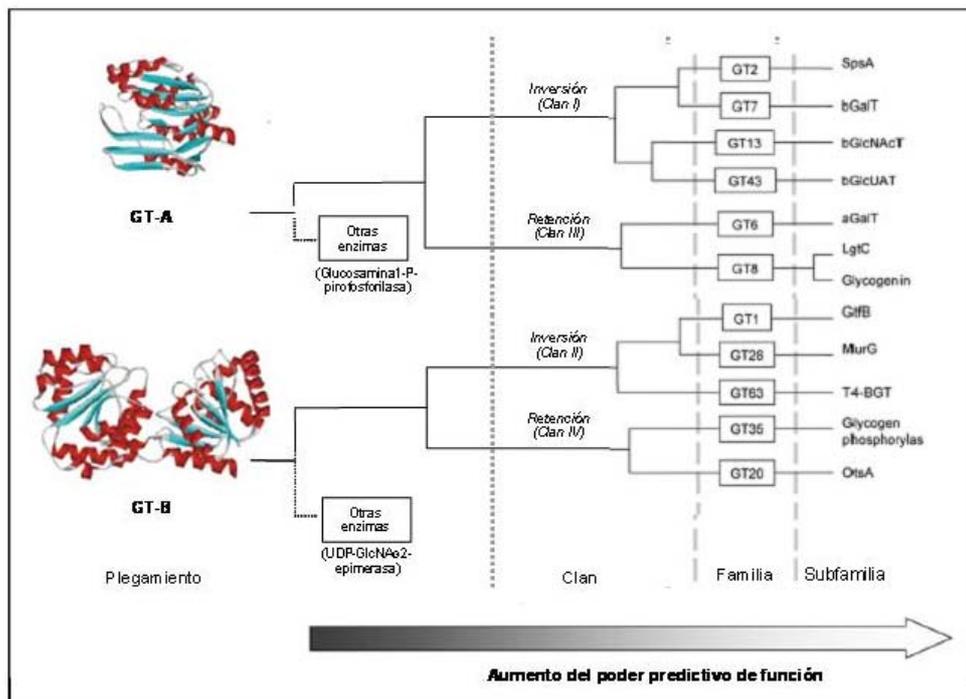
de GTs. Estas enzimas son transglicosidasas, están relacionadas evolutiva y estructuralmente a las GHs y por lo tanto son incluidas en la clasificación de GHs (28,29). Otras enzimas no consideradas GTs son las transferasas de azúcar-fosfato, como las que catalizan la conversión de UDP-azúcar y dolicol-fosfato a UMP y azúcar-difosfato-dolicol. Estas enzimas no son GTs ya que catalizan una reacción de sustitución en el átomo de fósforo del grupo fosfato. Finalmente y dado que la similitud de secuencia es indicativa de similitud estructural, la clasificación de GTs de Campbell (27,30) es capaz de proveer información de la estructura de una proteína. Asimismo, como el mecanismo catalítico es determinado por la posición relativa de ciertos grupos funcionales el mecanismo y la estereoquímica de la reacción catalizada estarían conservados dentro de cada familia. Este hecho tiene sus consecuencias: existen familias como la GT70 de GumK que solo incluyen a esta enzima y sus homólogos en diferentes especies, mientras que otras como GT4 son familias poli-específicas, de una distribución filogenética mucho más amplia.

Por todo esto, a pesar de la importante limitante que implica la baja homología de secuencia, las GTs caracterizadas dan prueba de la validez y capacidad de predicción del sistema de clasificación de Campbell, que surge como un sistema de clasificación predictivo, muy útil si consideramos el volumen de genomas que ya han sido secuenciados, con la consiguiente cantidad de ORFs de putativas GTs que se descubren. Además, un sistema de clasificación de este tipo constituye una de las principales herramientas de la glicobiología y específicamente en proyectos de librerías glicobiológicas (31,32). Una salvedad a tener en cuenta es que la utilización de este sistema para predecir especificidad debe tomarse con cautela, ya que existen casos documentados de GTs en las que el cambio de unos pocos aminoácidos (e inclusive uno solo) alteró la especificidad de sustratos de la enzima (33). Esto es prueba de las dificultades inherentes a la ausencia de secuencias consenso que sean indicativas de la unión de los diferentes sustratos. La clasificación de Campbell de GTs está disponible en la web (<http://www.cazy.org>) y es parte de un sistema de clasificación denominado CAZY (por carbohydrate active enzymes) que incluye también glicosil-hidrolasas, carbohidrato-esterasas, polisacárido-liasas y los módulos de unión a carbohidratos (CBM) asociados a estas enzimas.

Contrariamente a lo que se esperaría si consideramos el gran número de familias de GTs, su enorme variación de secuencia y la amplia diversidad estructural y funcional de sus productos, no se observa traducción en una gran variación a nivel estructural. Todas las estructuras descritas pertenecen a uno de dos tipos de plegamiento: el correspondiente a la superfamilia GT-A o GT-B. Estas 2 superfamilias son variaciones de dominios de plegamiento tipo Rossmann. Los miembros de la superfamilia GT-A muestran un núcleo de 7-8 láminas- $\beta$ , con un motivo DXD (o alguna de sus variantes DD, EXE, DXE), encargado de coordinar la unión de un metal/ribosea en el centro catalítico. Diferencialmente, las enzimas de la superfamilia GT-B no unen iones divalentes y poseen 2 dominios tipo Rossmann bien definidos, con una profunda hendidura entre sí, donde ocurre la unión de sustratos y la catálisis (34,35). Si bien la presencia de cationes divalentes (específicamente  $Mg^{++}$  y  $Mn^{++}$ ) aumenta la tasa de reacción por motivos aun no elucidados, no hay una estricta dependencia de éstos para que la reacción de glicotransferencia ocurra. Aun sin estructura descrita y por ende solamente predicha, se suma una tercera superfamilia a las ya conocidas: GT-C. Esta superfamilia comprende proteínas integrales de membrana con dependencia por sustratos dadores que sean azúcares unidos a poliprenilos. Este plegamiento ha sido descrito por búsquedas en la basic local alignment search tool (BLAST) y por comparaciones estructurales (36). Esta superfamilia parece incluir a 8 familias en CAZy. Todas estas proteínas son fuertemente hidrofóbicas, generalmente con 8 a 13 pasos transmembrana, y se localizan en retículo endoplasmático o en membranas plasmáticas.

Con este panorama estructural, queda de manifiesto la importancia de resolver la estructura de GTs en complejo con sus sustratos. Diferentes aminoácidos son los

encargados de contactar los diferentes sustratos, con la variabilidad en cuanto a la especificidad que esto implica y las consecuencias biotecnológicas que implica poseer esta información (12). Cabe destacar que el tipo de plegamiento no está relacionado con la estereoquímica de la reacción: GTs que actúan con retención de la configuración del C-anomérico claramente presentan estructuras relacionadas con GTs que invierten la configuración. Esto señala un origen evolutivo común para ambos tipos de enzimas. En un intento de amenizar la clasificación estructural con el criterio funcional, se ha propuesto el agrupamiento de familias de GTs que presentan el mismo plegamiento y mecanismo catalítico en clanes (36). Para las GTs que actúan con inversión se han definido dos clanes: clan I, que presenta plegamiento tipo GT-A y clan II con plegamiento tipo GT-B. De manera similar, se definieron dos clanes para GTs que actúan con retención: clan III, con plegamiento GT-A y clan IV, con plegamiento GT-B (Fig. 1)



**Fig 1:** Clasificación jerárquica de GTs. Abarca tipo de plegamiento, clan, familia y subfamilia. Para cada clan, se indica la estereoquímica de la reacción. SpsA,  $\beta$ -GalT,  $\beta$ -1,4-galactosiltransferasa bovina (15);  $\beta$ -GlcNAcT,  $\beta$ -1,2-N-acetilglucosaminiltransferase I de conejo (37);  $\beta$ -1,3-glucuroniltransferasas I humanas (38,39);  $\alpha$ -GalT,  $\alpha$ -1,3-galactosiltransferase bovina (40); LgtC,  $\alpha$ -galactosiltransferase de *N. meningitidis* (41); Glicogenina de conejo (42); GtfB de *Amycolatopsis orientalis* (43); MurG,  $\beta$ -N-acetilglucosaminiltransferase de *E. coli* (44); BGT, DNA  $\beta$ -glucosiltransferasa del bacteriófago T4 (13); GSA, glucógeno sintasa de *A. tumefaciens* (45); OtsA, trealose-6-fosfato sintasa de *E. coli* (42); glucógeno fosforilasa humana (46,47). Adaptado de (21)

Finalmente, una consecuencia importante de la fuerte conservación de estructura en las superfamilias de GTs es que cambiar las especificidades de sustrato puede ser mas sencillo de lo que se pensaba (33,48). Las estructuras de GTs en complejo con sus sustratos nos hacen especular que cambiando pequeñas regiones, incluso un único residuo, se podría relajar e incluso modificar a voluntad la especificidad. Evidentemente esto es de gran utilidad en el campo de síntesis de oligosacáridos y glicoconjugados. Un ejemplo paradigmático de esto último es de la GlcAT-I humana (39,49) en donde el cambio de un residuo de His (His308) a Arg permite que la enzima pueda utilizar muy eficientemente otros sustratos además del UDP-GlcA, su sustrato natural. Entre los nuevos sustratos se encuentran UDP-Glc, UDP-Man o UDP-GlcNAc (48). Como se ve, para un diseño racional de mutantes de especificidades cambiadas la descripción de los contactos entre la enzima y sus sustratos es de vital importancia.

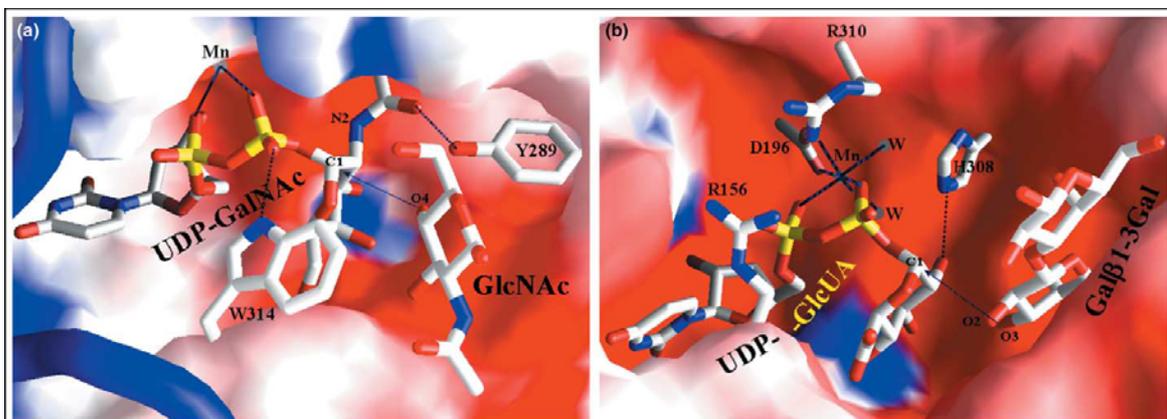
### 4.3 Mecanismos de reacción

#### 4.3.1 Mecanismo de inversión

Las estructuras en anillo de los azúcares poseen un C-anomérico que puede presentar una de dos estereoquímicas:  $\alpha$  ó  $\beta$ . Las GTs que catalizan una sustitución nucleofílica sobre este centro anomérico se agrupan en dos clases según provoquen la inversión o retención de la estereoquímica del centro anomérico. Así, hay dos clases generales de GTs: de inversión o de retención de la configuración anomérica ( $\beta$ -glicosiltransferasas o  $\alpha$ -glicosiltransferasas respectivamente).

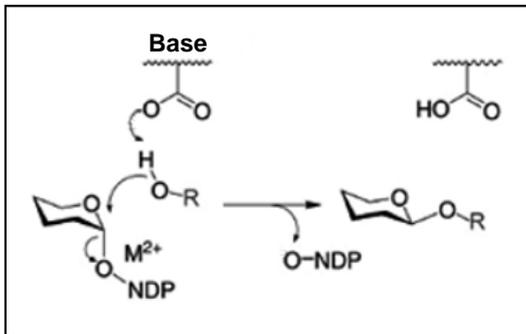
En los nt-azúcares dadores el residuo azúcar se encuentra invariablemente en configuración  $\alpha$ . Por lo tanto las GTs de inversión dan lugar a productos con uniones  $\beta$  (la configuración anomérica ha sido invertida respecto a la del sustrato), mientras que, la retención de la configuración anomérica da lugar a productos con uniones  $\alpha$ , cuya estereoquímica es idéntica a la del residuo azúcar del sustrato dador.

A pesar de que ya está descrito el mecanismo catalítico de las GTs de retención y de inversión, solo podemos considerar que el mecanismo de inversión está bien caracterizado. Las estructuras cristalinas de los complejos con dador y aceptor de las GTs de inversión  $\beta$ -4 Gal-T1, GnT I y GlcAT-I nos muestran la orientación del dador y el aceptor en los bolsillos catalíticos. (Fig. 2) (37,38,49,50). En estas enzimas parte del sustrato dador está cubierto del solvente por un loop flexible, y el extremo no reductor del oligosacárido aceptor se mantiene cercano al azúcar dador por interacciones con residuos específicos, a veces incluidos en este loop flexible, mientras que el resto del oligosacárido sigue expuesto al solvente (lo cual no es tan grave pues no está implicado directamente en la catálisis). El grupo OH reactivo del aceptor se posiciona a una distancia de aproximada de 3 Å del C1 del azúcar dador (Fig. 2).



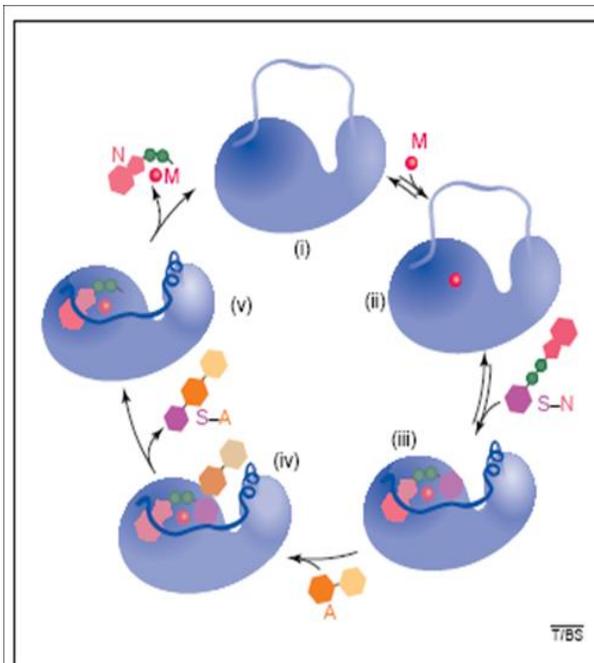
**Fig 2:** Un simple aminoácido determina la especificidad del azúcar dador. (a) Tyr289 en la  $\beta$ -4Gal-T bovina (o Tyr286 en la  $\beta$ -4Gal-T humana) determina la especificidad de la enzima por el dador UDP-Gal. UDP-GalNAc es un sustrato pobre para la  $\beta$ -4Gal-T. El grupo hidroxilo de la Tyr289 interactúa con el grupo N-acetilo del GalNAc del UDP-GalNAc, actuando como un freno molecular y restringiendo su transferencia al aceptor GlcNAc. La mutación del Tyr289 a Leu, Ile o Asp alivia este impedimento estérico, siendo que la mutante Y289L puede transferir GalNAc desde UDP-GalNAc tan eficientemente como la enzima salvaje transfiere galactosa desde UDP-Gal a GlcNAc. (b) En la estructura cristalina de GlcAT-I-UDP-GlcUA-Mn<sup>++</sup> (PDB 1KWS), el N3 del residuo His308 hace un puente de hidrógeno con el hidroxilo-2 de la GlcUA del dador UDP-GlcUA y esto determina la especificidad de GlcAT-1 por el dador. El O3 de la galactosa en el extremo no reductor del aceptor Gal  $\beta$ 1-3Gal (PDB 1FGG) se coloca cercano al C1 del GlcA para el ataque nucleofílico. Cuando His308 se muta a Arg, la enzima GlcAT-I mutante, en contraste a la enzima salvaje, puede transferir manosa, glucosa y GlcNAc desde sus respectivos derivados de UDP con alta eficiencia. El ión Mn<sup>++</sup> en (a), a pesar de presentar una geometría octaédrica, solo se muestra coordinado a 2 átomos de oxígeno del UDP, de los fosfatos  $\alpha$  y  $\beta$  respectivamente. Las superficies moleculares están coloreadas en base a la carga, donde rojo es negativo y azul positivo. Adaptado de (17)

Un residuo de ácido aspártico o glutámico que esté a una distancia de puente de hidrógeno desprotona el grupo OH del aceptor, el cual actúa como base catalítica realizando un ataque nucleofílico sobre el C1 del dador. Los iones divalentes (principalmente  $Mn^{++}$  y  $Mg^{++}$ ) asisten en la partida del UDP, neutralizando la carga negativa que se forma en el fosfato- $\beta$ . El C1 del dador forma un estado de transición tipo ión oxocarbenio, que lleva a la formación del enlace glicosídico entre dador y aceptor con inversión de la configuración anomérica en el C1. Esta secuencia catalítica se correlaciona perfectamente con un mecanismo de reacción de desplazamiento tipo  $SN_2$  (Fig. 3).



**Fig 3:** Mecanismo de reacción para GTs que actúan con inversión. Es una reacción de desplazamiento simple, en la cual el dador de azúcar activado es atacado por el aceptor, resultando en la inversión de la configuración anomérica.

Este estado de transición ha sido atrapado recientemente en  $\beta 4Gal-T1$  durante la transferencia de un residuo de GalNAc desde UDP-GalNAc a glucosa, en presencia de  $\alpha$ -lactoalbumina (51). Luego de la formación del enlace glicosídico, el producto oligosacárido es expulsado de su sitio de unión mientras que el complejo UDP- $Mn^{++}$  todavía permanece en parte escondido debajo del loop flexible. Debido a la pérdida de interacciones con el motivo azúcar del dador, el loop revierte a una conformación abierta, permitiendo que el complejo UDP- $Mn^{++}$  difunda fuera de la enzima (Fig. 4).

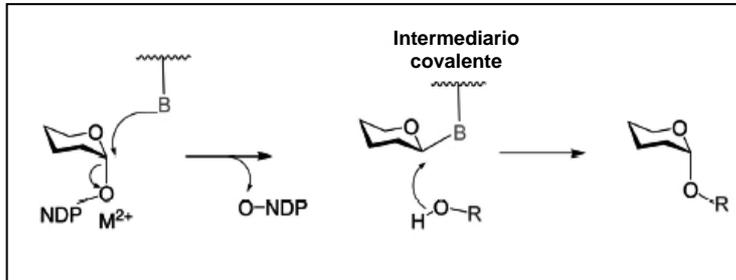


**Fig 4:** Ciclo catalítico de GTs con actividad dependiente de iones. (i) La enzima se muestra como una proteína de 2 dominios en conformación abierta con un loop flexible (en celeste) sin estructura. (ii) El ión metálico (M) se une a la enzima en la conformación abierta. (iii) El nt-azúcar (S-N) se une al complejo enzima-metal y el loop cambia su conformación (en azul oscuro) creando una estructura de tapón o cubierta, que cubre al nt-azúcar cubierto. Se forma una estructura helicoidal en la región C-terminal de este loop que se transforma en parte del sitio de unión del sustrato aceptor. (iv) El aceptor se mantiene en su sitio de unión con su extremo no-reductor cercano al residuo de azúcar del dador, mientras que el resto de la molécula está expuesta al solvente. (v) Luego de la catálisis durante la cual el azúcar es transferido al aceptor, el producto (S-A) es eyectado del sitio de unión, mientras que el complejo UDP-metal permanece cubierto debajo del loop. (i) El loop vuelve a su conformación abierta, facilitando la liberación del complejo UDP-metal y permitiendo que la enzima libre comience un nuevo ciclo catalítico.

### 4.3.2 Mecanismo de retención

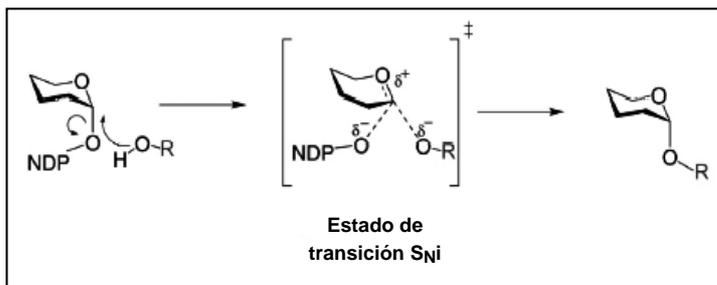
En contraste el mecanismo de inversión, el mecanismo usado por las GTs de retención no está bien definido. Se ha propuesto un mecanismo de doble desplazamiento (similar al descrito para GHs que actúan con retención), (40) en donde el azúcar dador

es primero transferido a un grupo apropiado dentro de la enzima (con inversión de la configuración en C1) produciéndose luego una segunda transferencia donde el C1 revierte a su configuración original (Fig. 5). En la galactosiltransferasa de retención LgtC se han podido identificar intermediarios del azúcar dador con la proteína que podrían atribuirse a un mecanismo de doble desplazamiento (41,52).



**Fig 5:** Mecanismo de reacción propuesto para GTs que actúan con retención. Por analogía con el mecanismo de GHs que retienen la configuración, la reacción ocurriría a través de un intermediario covalente glicosil-enzima.

Sin embargo, las estructuras cristalinas de EXTL2 (53,54),  $\alpha$ -3GT (55) o Fuc-T de *Helicobacter pylori* (18) con el nt-azúcar o UDP unido no estarían apoyando esta hipótesis. De hecho, basados en la evidencia estructural, se ha propuesto un mecanismo de sustitución nucleofílica interna ( $S_{Ni}$ ) donde el ataque nucleofílico ejercido por el aceptor y la partida del grupo saliente ocurren simultáneamente sobre el mismo plano del azúcar (Fig. 6)



**Fig 6:** Mecanismo tipo  $S_{Ni}$  propuesto para GTs que actúan con retención. En el estado de transición, el grupo que ejerce el ataque nucleofílico y el grupo saliente se encuentran simultáneamente en el mismo plano del anillo azúcar.

### 4.4 Descripción estructural

#### 4.4.1 Sitios de unión de los sustratos

Una característica mecanística importante que surge de las estructuras resueltas es la interacción entre los sitios de unión del sustrato dador y del aceptor. Como regla general, se puede afirmar que nt-azúcar se encuentran en el dominio N-terminal de enzimas que presentan plegamiento tipo GT-A, y en el dominio C-terminal de las que poseen plegamiento GT-B, mientras que los sustratos aceptores se unen a sitios ubicados en el otro dominio. Los sustratos aceptores son la parte más diversa de la ecuación catalítica de las GTs. Sus modos de unión son muy variados e implican interacciones electrostáticas y/o hidrofóbicas con la proteína, dependiendo de la naturaleza química del sustrato. Por otro lado, los sustratos dadores (nt-azúcares) tienen un sitio de unión más homogéneo. En líneas generales y para todos los casos descritos, se trata de un bolsillo de carga netamente básica, donde, lógicamente, se acomodan los fosfatos que tienen carga negativa y que son estabilizados mayormente a través de enlaces de hidrógeno. Características y ejemplos de ambos tipos de plegamiento y la forma en que unen a sus sustratos se tratarán en la sección 4.4.3.

#### 4.4.2 Utilización de iones divalentes

Las GTs que presentan plegamiento GT-A se diferencian de las agrupadas como GT-B en su requerimiento de cationes divalentes. Todas las GTs caracterizadas como GT-

A presentan una restricción estructural en la coordinación de un ión metálico por medio de dos átomos de oxígeno de los fosfatos del nt-azúcar y las cadenas laterales de residuos de la enzima. Estas cadenas laterales son aportadas por residuos de un motivo muy conservado que se ha denominado "DXD" (56), aunque existen variaciones del mismo y cualquiera de los Asp puede estar reemplazado por otro residuo o puede ocupar otra posición en la secuencia consenso. Así, variaciones del mismo (NDD, S/TDD, entre otros) se han encontrado en varias GTs cumpliendo la misma función. El motivo DXD cumple una doble función en la catálisis de GTs. Por una parte, coordina la unión de un ión metálico divalente (generalmente  $Mn^{2+}$ ) que es esencial para la catálisis, ayudando a anclar al grupo pirofosfato del nt-azúcar dador en el sitio activo de la enzima. Por otra, asiste la partida del grupo saliente (nucleótido-fosfato), contrarrestando la carga negativa que se desarrolla sobre el fosfato- $\beta$  luego de la ruptura del enlace azúcar-fosfato en el sustrato dador (Fig. 4) (57,58). La mutación del motivo DXD o la ausencia de cationes divalentes en la reacción da lugar a la anulación de la actividad de la enzima. Esta completa dependencia en la presencia de cationes divalentes es un elemento diagnóstico de la bioquímica de las GT-A.

En cambio, las GTs con plegamiento GT-B no presentan este estricto requerimiento y rara vez se ha detectado la presencia de iones metálicos como cofactores en las estructuras resueltas. Uno de los pocos casos descritos es el de la enzima BGT del fago T4, en la que se encuentra un ión metálico unido en cercanía al fosfato- $\beta$  del nt-azúcar (59). A pesar de no presentar una dependencia estricta por cationes divalentes, se ha observado que en presencia de éstos algunos miembros de GT-B muestran una tasa de reacción aumentada, aunque se desconocen las causas (35,60). Finalmente, es necesario aclarar que la presencia del motivo DXD es necesaria pero no suficiente para clasificar a una GT-A. Este motivo se puede encontrar en las otras 2 superfamilias, especialmente en la GT-C, si bien se desconoce que función tendría, si es que presenta alguna(36).

### 4.4.3 Plegamiento y dominios conservados

El plegamiento de GTs, consiste básicamente en una estructura tipo "sandwich"  $\alpha/\beta/\alpha$ , similar al motivo estructural Rossmann (seis hebras- $\beta$  paralelas con topología 321456). Este plegamiento fue descrito originalmente para proteínas que unen nucleótidos (61). El plegamiento GT-A es un sandwich  $\alpha/\beta/\alpha$ , constituido por una lámina- $\beta$  de siete hebras con topología 3214657, con la hebra 6 antiparalela al resto. La lámina- $\beta$  central tiene a su lado otra más pequeña, y la asociación entre ambas contiene al sitio activo. Las enzimas con plegamiento GT-A tienen en común el motivo DXD, cuyos carboxilatos coordinan un ión metálico divalente. Los residuos que preceden al motivo DXD en la secuencia proteica forman una hebra  $\beta$  a través del núcleo de la proteína. Esto explica la observación de que los residuos N-terminales a DXD son de naturaleza no polar. Los miembros de GT-A que actúan con inversión de la configuración presentan un carboxilato (Asp o Glu) equivalente en la estructura, que actúa como base catalítica. En algunos miembros de GT-A un residuo de Asp coordina el N3 de la uridina del nt-azúcar. Más allá de esta limitada similitud, no se han encontrado otras características conservadas en miembros de GT-A.

El plegamiento GT-B, se caracteriza por presentar dos dominios  $\alpha/\beta/\alpha$  tipo Rossmann (dominio 1: lámina- $\beta$  de siete hebras paralelas con topología 3214567; dominio 2: lámina- $\beta$  de seis hebras paralelas con topología 321456) que están asociados más laxamente y están separados por una hendidura profunda. El sitio de unión del nt-azúcar se encuentra sobre la cara C-terminal de la hendidura y este sustrato hace contacto principalmente con residuos del dominio C-terminal. El extremo N-terminal es responsable de la unión del aceptor (62). Los motivos característicos del plegamiento GT-B son, al igual que en GT-A, limitados. Se ha identificado un residuo Glu en una  $\alpha$ -

hélice del dominio C-terminal, que une por medio de puente hidrógeno la ribosa del nt-azúcar; y un loop rico en residuos Gly que sería responsable de la unión de los grupos fosfato del nt-azúcar (35). En algunas GTs, el residuo Glu mencionado forma parte de un motivo EX<sub>7</sub>E, que fue primeramente identificado en un grupo de manosiltransferasas procariotas (63). Este motivo se encuentra en el dominio C-terminal; Asp/Glu/Lys en medio de una hélice- $\alpha$  que ha sido llamada posición 1; mientras que Asp/Glu en el loop anterior a la hélice- $\alpha$ , se ha denominado posición 2 (63). Los residuos Glu del motivo EX<sub>7</sub>E han sido vinculados al mecanismo de doble desplazamiento en GTs que actúan con retención (26,63). Un mecanismo de este tipo requeriría de dos aminoácidos ácidos que actúen como ácido y base respectivamente. Sin embargo, hasta el presente no existen datos experimentales que apoyen esa teoría (64).

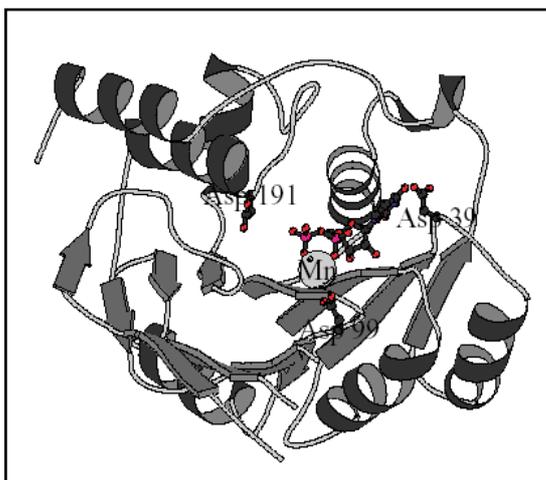
Por otro lado, también se ha predicho la aparición de un nuevo tipo de plegamiento de GTs, que ha sido denominado GT-C y que incluiría a proteínas con múltiples pasos transmembrana y un motivo DXD modificado. Aunque aún no se ha resuelto la estructura de ninguna de estas proteínas, se espera que los miembros de por lo menos 8 familias presenten el mismo tipo de plegamiento (36). Finalmente, no todas las GT-s entran en este esquema de 3 superfamilias: la estructura de la  $\alpha$ -2,3/8-sialiltransferasa CstII de *Campilobacter jejuni*, que utiliza como dador un nucleótido-monofosfato-azúcar (CMP-NeuAc) presenta plegamiento  $\alpha/\beta/\alpha$  con topología diferente a GT-A o GT-B (lámina- $\beta$  de ocho hebras paralelas con topología 54321678) y no contiene motivo DXD (52,65). Este es un ejemplo que muestra que no todas las familias de GTs están incluidas en las categorías de las superfamilias GT-A, B o C.

En lo que sigue, tomaremos como modelo de descripción la estructura de una apo-GT comparandola con la estructura en complejo con sus sustratos para cada uno de los clanes. Las enzimas seleccionadas serán: clan I: SpsA de *Bacillus subtilis* (66), Clan II MurG (16,44), Clan III LgtC (41) y clan IV PimA (19).

### 4.4.3.i Superfamilia GT-A: Clan I (inversión) SpsA

SpsA es una GT involucrada en la producción de la cubierta celular durante el proceso de esporulación de *B. subtilis* (67). La función específica de esta enzima, así como la especificidad por sus sustratos no se conoce, si bien se especula que podría participar en la síntesis de ácido teicoico y colánico, transferencia de Gal y Rha, producción de LPS y síntesis de celulosa y quitina por homología de secuencia con enzimas involucradas en estos procesos (68).

La estructura de SpsA fue resuelta por cristalografía de rayos X a una resolución de 1.5 Å. Se describió la estructura de la apo-enzima y la enzima en complejo con UDP/Mg<sup>++</sup> y UDP/Mn<sup>++</sup> (14). La proteína presenta dos dominios, el N-terminal (residuos 2-100) forma un típico dominio de unión de nucleótido, con cuatro hebras- $\beta$  paralelas flanqueadas a ambos lados por dos hélices- $\alpha$ . El resto de la estructura (residuos 101-256) está compuesta por una lámina- $\beta$  mixta flanqueada por tres hélices- $\alpha$  a un lado y una al otro.



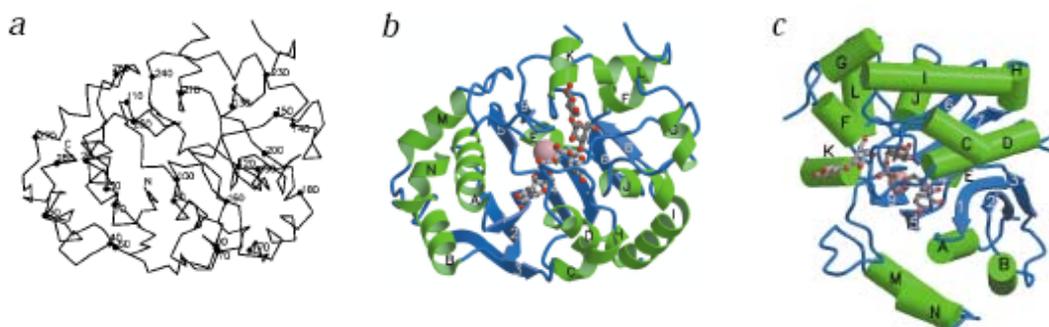
**Fig. 7:** Estructura cristalográfica de SpsA. Se muestra el complejo UDP-Mn<sup>++</sup> de la enzima con los residuos ácidos que participan en la unión del nucleótido y D191, que actuaría como base catalítica.

En la estructura en complejo con UDP, el uracilo se une a un “bolsillo” hidrofóbico (Fig. 7). En este bolsillo, el Asp 39, en el extremo de la hebra-β2, une por puente hidrógeno el N-3 del uracilo. El residuo Asp 39 es uno de los dos Asp estrictamente conservados en miembros de la familia 2. El otro Asp (Asp 99) está involucrado, junto con los grupos fosfato del UDP en la coordinación de Mn<sup>++</sup>. Asp 99 se encuentra en la hebra-β4 y forma un motivo “DD”, variante del DXD que las GT-A presentan y que es responsable de la coordinación del ión divalente involucrado en catálisis. Por su parte, Asp 98 une los hidroxilos de la ribosa. La mayor conservación de secuencia en miembros de la familia 2, se encuentra en el extremo N-terminal. Allí existen tres zonas invariables, donde se encuentran los residuos descritos que hacen contacto con el UDP del sustrato dador (Tyr 11, Asp 39 y Asp 99). El mecanismo de las GTs que proceden con inversión requiere además de un residuo que actúe como base catalítica y provoque la deprotonación del sustrato aceptor. Los resultados de alineamientos de secuencias señalan a Asp 191 como candidato, pero los datos estructurales apuntan también a la tríada formada por Asp 158, His 159 y Cys 160, que se encuentran en una posición adecuada para activar al aceptor, aunque ninguno de estos residuos está conservado. Debido a la ubicación y la conservación de Asp 191, éste fue propuesto como base catalítica en SpsA.

#### 4.4.3.ii Clan III (retención) LgtC

LgtC es una α-1,4-galactosiltransferasa involucrada en la síntesis de lipooligosacáridos de la pared celular de *Neisseria meningitidis*, transfiriendo un residuo de α-d-galactosa desde UDP-Gal a una lactosa (41).

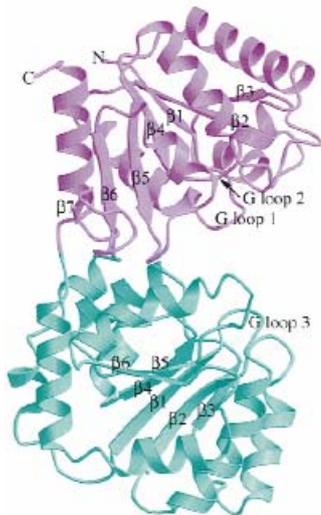
La estructura de LgtC es un monómero que comprende 286 residuos organizados en 2 dominios. El dominio N-terminal es mucho más grande que el C-terminal (residuos 1-247 vs 248-282). El dominio N-terminal tiene un núcleo de siete laminas-β rodeados por α-hélices. Aquí se localiza la hendidura de unión del sustrato aceptor y el sitio de unión del nucleótido-azúcar, que esta formado por el plegamiento de los primeros 100 aminoácidos. El dominio C-terminal es principalmente helicoidal y media la unión a la membrana (Fig. 8)



**Fig. 8:** La arquitectura global de LgtC. **a**, trazado de los Cα en el monómero de LgtC **b**, Estructura de LgtC con sus análogos de sus sustratos unidos. El aceptor se muestra en gris oscuro, el dador en gris claro y el manganeso en rosa **c**, Estructura de LgtC mostrando el dominio de unión de los sustratos N-terminal y el dominio de unión a la membrana C-terminal. Adaptado de (41)

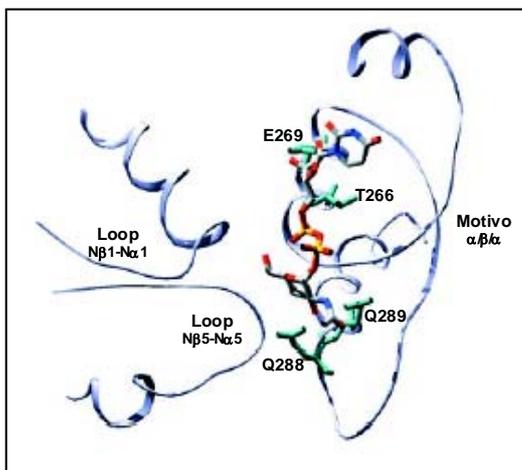
La estructura de esta enzima fue resuelta a 2 Å de resolución y fue cristalizada en presencia de sustratos no hidrolizables tanto del sustrato aceptor como el donante. El análogo del sustrato donante, el UDP-2F-Gal interacciona con residuos que están comprendidos entre el comienzo de la lamina β-1 y la hélice A (residuos 1-28). 2 loops, provenientes de distantes secciones de la molécula, se pliegan a modo de cubierta sobre el dador. Estos loops comprenden los residuos 75-80 y 246-251 y están estabilizados por interacciones con el dador. (Fig. 8 b y c). Específicamente, el uracilo





**Fig.10:** Arquitectura estructural de MurG. El dominio N-terminal se muestra en púrpura y el C-terminal en verde. Adaptado de (44)

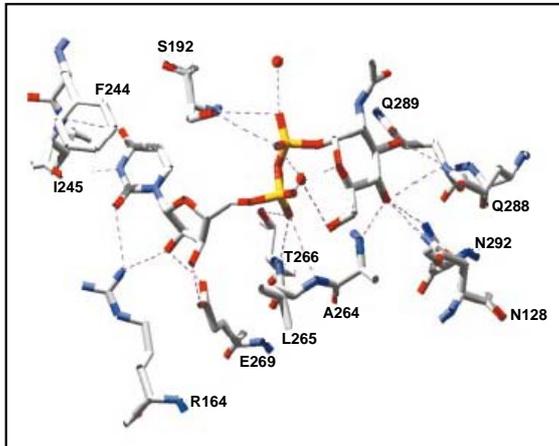
La estructura del complejo MurG/UDP-GlcNAc fue resuelta a 2.5 Å (16). Además el análisis de selectividad por los sustratos se extendió con el estudio cinético de mutantes en residuos de unión a sustratos. La comparación entre ambas estructuras muestra que los dominios N y C-terminales son similares en presencia o ausencia del sustrato dador. Sin embargo, en presencia de UDP-GlcNAc, se observa un cambio en la orientación relativa de ambos dominios que se acercan entre si unos 2.0 Å, siendo el ancho de la hendidura de 16 y no de 18 Å como en la estructura de la apo-enzima. Esto implica que el dominio C-terminal rota unos 10°, funcionando como bisagra la región de unión entre ambos dominios, posiblemente alrededor del residuo N161. Esta es la conformación cerrada de MurG. El nt-azúcar hace contacto con un bolsillo con carga básica de residuos correspondientes a la región C-terminal de la hendidura. Esta región tiene un patrón de de residuos Pro y Gly conservados junto con otros residuos característicos. La región de unión tiene un patrón de plegamiento  $\alpha/\beta/\alpha$ , con las  $\alpha$ -hélices ubicadas cerca de la hendidura entre los dominios N- y C-terminales (Fig. 11). Esta característica estructural, con variaciones, se encuentra conservada en la superfamilia GT-B.



**Fig. 11:** Regiones conservadas en la hendidura central de MurG. Porciones del N- y C-terminal de MurG que conforman el sitio activo. Se observa la unión de UDP-GlcNAc en el C-terminal (motivo  $\alpha/\beta/\alpha$ ) y algunos de los residuos involucrados. En el N-terminal, se muestra el loop rico en Gly  $N\beta 1-N\alpha 1$  y el loop  $N\beta 5-N\alpha 5$  (o loop HEQN) que estaría involucrado en la unión del aceptor.

El nt-azúcar hace contacto con residuos de ambas  $\alpha$ -hélices y de los loops adyacentes. Esta característica estructural, con variaciones, puede ser identificada en la mayoría de los miembros de GT-B. Su función es la de estabilizar al fosfato- $\alpha$  en ausencia de iones metálicos y también es responsable de anclar en su sitio a la ribosa y la hexosa del nt-azúcar. El fosfato- $\beta$  del nt-azúcar está en contacto con el loop denominado GGS que conecta  $C\beta 1$  con  $C\alpha 1$ . La posición de este loop cambia notablemente en la estructura de MurG con el sustrato dador: se mueve hacia abajo,

acercándose al motivo  $\alpha/\beta/\alpha$ . Este movimiento es mediado por el contacto entre el fosfato- $\beta$  y el residuo de Ser (S192 en MurG) (Fig. 11) muy conservado en homólogos de MurG y en algunos miembros de GT-B. Es de notar aquí que la mutación de Ser192 a Ala afecta todos los parámetros cinéticos, incluso la  $K_m$  para el sustrato aceptor. Dado que MurG utiliza un mecanismo secuencial ordenado, en el cual se une primero el nt-azúcar, es probable que cambios conformacionales en el loop GGS estén involucrados en el ajuste requerido para la unión del aceptor.



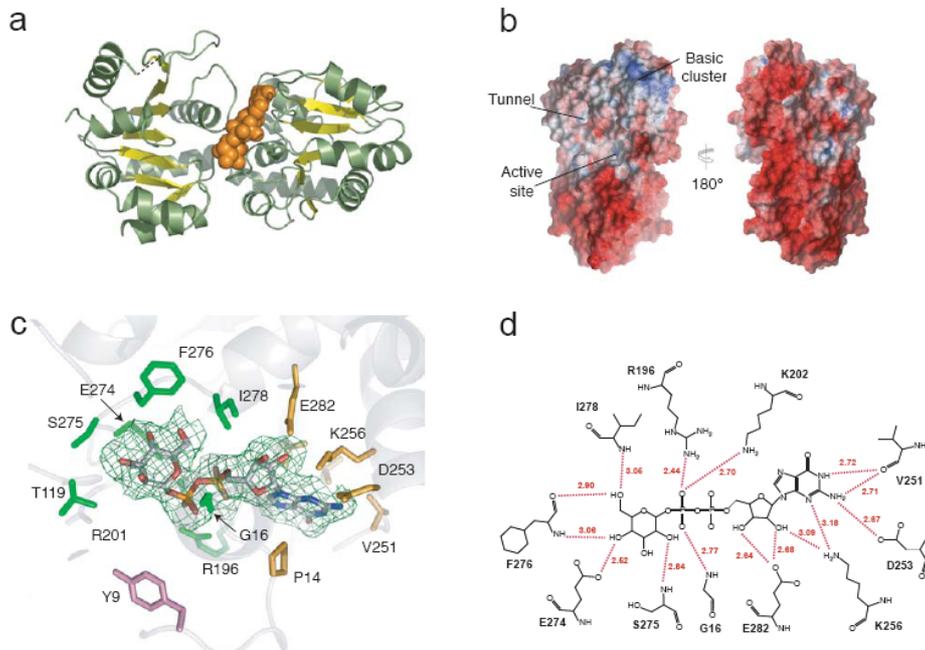
**Fig. 12:** Sitio de unión del sustrato dador. Se muestran los contactos entre UDP-GlcNAc y MurG. Los aminoácidos 264, 265, 269, 288, 289 y 292 se encuentran en el motivo  $\alpha/\beta/\alpha$ . S192 se encuentra en el loop GGS y N128 en el loop HEQN, ambos del dominio N-terminal (16).

El anillo uracilo del nt-azúcar se encuentra en un “bolsillo” delimitado de un lado por  $C\beta_3$  y la hebra que conecta los dominios N y C-terminales, y por otro, por el loop GGS. La base se ancla por interacciones de puente de hidrógeno y por el anillo aromático de F244 que se encuentra rotado respecto a su posición en la enzima libre, ejerciendo interacciones de “stacking” con el anillo uracilo y cerrando el bolsillo (Fig. 12).

Los factores que determinan la especificidad de sustrato se hacen evidentes en la estructura del complejo MurG/UDP-GlcNAc. Por un lado, el bolsillo en que se encuentra el anillo uracilo es demasiado pequeño para acomodar una base purínica; mientras que los puentes de hidrógeno que estabilizan al uracilo, tendrían el efecto contrario sobre citosina. Así la preferencia por UDP parece deberse a una serie de factores estéricos y electrónicos. La cadena lateral de E269 hace contacto con los OH-2' y 3' de la ribosa del UDP. El análisis de una mutante E269A y la utilización de un análogo 2'-deoxi de UDP-GlcNAc muestran que esta interacción es funcionalmente importante y cumple un rol clave en la capacidad de MurG de discriminar entre UDP y TDP. Por otro lado, la estructura del complejo muestra que la presencia de un sustituyente metílico en C5 “chocaría” contra el loop GGS, lo que explica la incapacidad de MurG de utilizar TDP-GlcNAc. Del mismo modo la selectividad por la porción sacarídica se debe a factores estéricos. MurG no es capaz de utilizar UDP-GalNAc como dador, lo que indica que la presencia del OH ecuatorial en C-4 es crítica. Otras regiones conservadas implican al dominio N-terminal. Una de estas regiones (loop HEQN) se encuentra justo frente al sitio de unión de la hexosa y formaría parte del sitio de unión del aceptor. Las otras dos regiones conservadas del dominio N-terminal son dos loops ricos en Gly, semejantes a los loops responsables de la unión de fosfato en proteínas que unen nucleótidos (secuencia consenso GXGXXG). En MurG estos loops estarían involucrados en la unión del di-fosfato del Lípido I, el sustrato aceptor (44). Finalmente, MurG se encuentra asociada a la membrana citoplasmática, pero carece de dominios transmembrana. La estructura revela una región de residuos hidrofóbicos rodeados de residuos básicos en el dominio N-terminal que ha sido propuesto como el sitio de unión a la membrana bacteriana. La asociación a través de esta región permitiría que los loops G del N-terminal se acerquen a la superficie de la membrana donde se encuentra la porción di-fosfato del aceptor. De esta manera, el nt-azúcar dador también podría acceder a la hendidura entre los dominios N- y C-terminales, que aloja al sitio activo.

4.4.3.iv Clan IV: PimA

PimA es una manosiltransferasa de *Mycobacterium smegmatis* encargada de transferir un residuo de manosa a la posición 2 del motivo del fosfatidil-inositol, para producir fosfatidil-inositol-monomanosido. Los fosfatidil-inositol-manosidos son parte del complejo peptidoglicano-micolilarabinogalactano que forma parte de la envoltura celular en micobacterias, y tiene un preponderante papel en la patogenia de estas bacterias.



**Fig. 13:** Estructura de PimA en complejo con el sustrato dador GDP-Man. **A.** Estructura completa de PimA en complejo con GDP-Man (representado como esferas naranjas). La topología de estructura secundaria es  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4-\alpha 3-\beta 5-\alpha 4-\alpha 5-\beta 6-\alpha 6-\beta 7-\alpha 14$  para el dominio N-terminal (arriba) y  $\beta 8-\alpha 7-\beta 9-\alpha 8-\beta 10-\alpha 9-\beta 11-\alpha 10-\beta 12-\alpha 11-\beta 13-\alpha 12-\alpha 13$  para el dominio C-terminal (abajo). **B.** Distribución de carga electrostática en PimA. Las regiones electronegativas se muestran en rojo, azul para las electropositivas y blanco para las neutras. **C.** Mapa final de densidad electrónica ( $2F_o-F_c$ ) para el GDP-Man (contorneado a  $1\sigma$ ). Los residuos de PimA en contacto con la guanosina y el anillo de la ribosa se muestran en naranja y aquellos en contacto con los fosfatos y el anillo de la manosa en verde. Los residuos en contacto con los fosfatos del inositol se muestran en rosa. **D.** Vista esquemática de los puentes de hidrogeno entre GDP-Man y la proteína.

Esta enzima presenta el plegamiento característico de las enzimas de la superfamilia de las GT-B, con dos dominios separados por una hendidura catalítica. El N-terminal comprende los residuos 1-169 y 349-373 y el C-terminal los residuos 170-348. El núcleo de cada dominio esta compuesto de 7 hojas- $\beta$  paralelas y rodeado de 7  $\alpha$ -hélices que las conectan. El dador GDP-Man se une en un bolsillo definido por el loop  $\beta 1-\alpha 1$  del dominio N-terminal, los finales de las hebras  $\beta 8$  y  $\beta 9$  y el loop que conecta  $\beta 8$  y  $\alpha 9$  del dominio C-terminal. Los puentes de hidrogeno que estabilizan al dador se observan en la Fig. 13 c y d. De esta figura es de destacar las interacciones que estabilizan los fosfatos, con los residuos básicos R196 y K202, una forma muy conservada en este clan de estabilizar estas cargas negativas. El sustrato lipídico (fosfatidil-inositol) se une en un tunel hidrofóbico que se observa en el extremo N-terminal. En este, el motivo del inositol es estabilizado por interacciones de stacking con la Y9 (19). Además, del modelo se puede corroborar que la R201 estaría estabilizando los oxígenos de los fosfatos del Pi.

En este terreno solamente se pueden hacer especulaciones, pero se ha sugerido que la superficie del dominio N-terminal, adyacente a la hendidura catalítica y que esta muy conservada en homólogos micobacterianos (Fig. 13 b) podría ser la responsable

de esta asociación, poseyendo un cluster de residuos hidrofóbicos cercanos a un cluster de residuos básicos en la hélice 2 y en el loops que conecta  $\beta$ -3 con  $\alpha$ -2 (residuos 59-70). Estas hipótesis se refuerza con los experimentos de mutagénesis que escinden estas regiones y lo que se observa es una fuerte disminución de la actividad enzimática.

### 4.4.4 Plegamiento GT-C

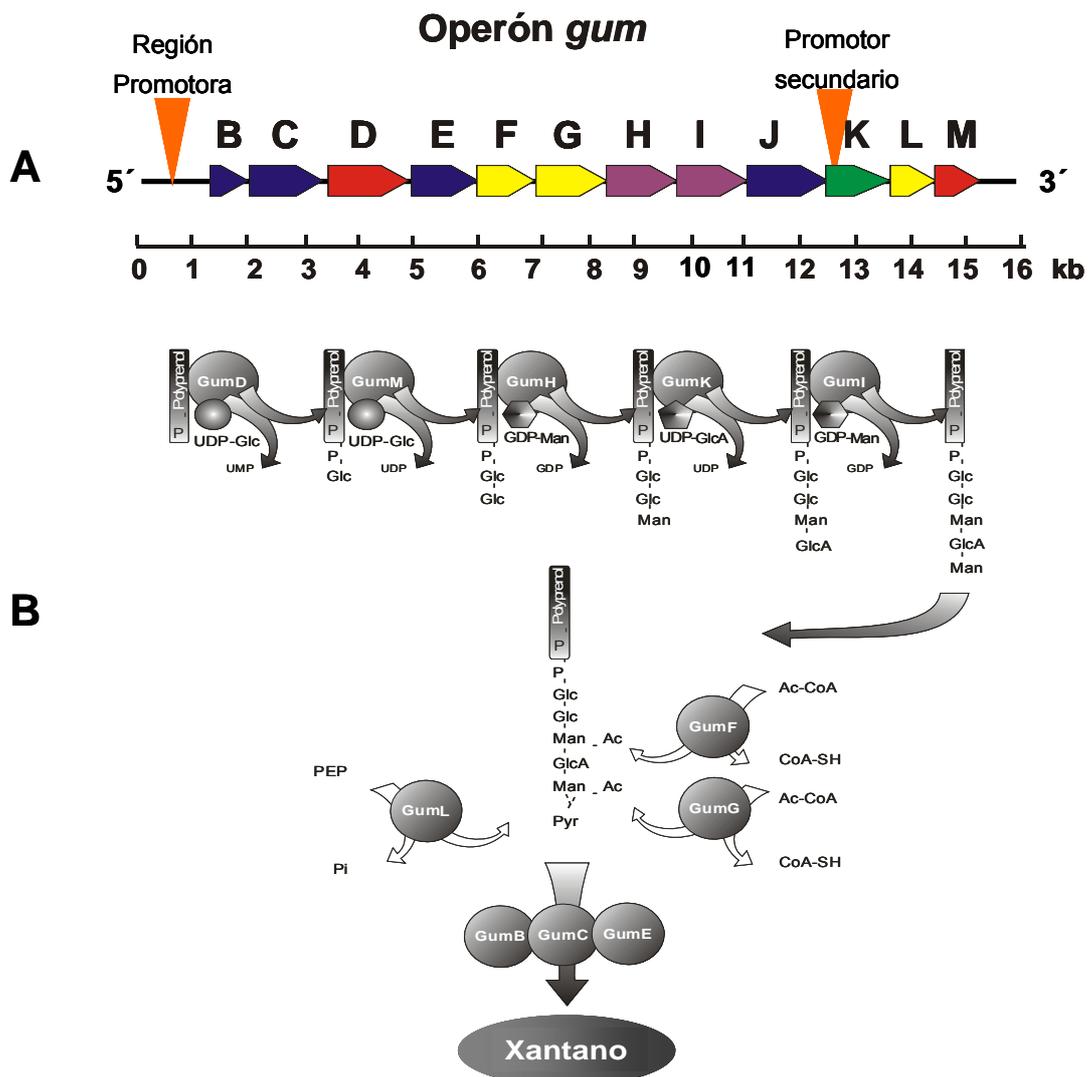
La superfamilia GT-C de reciente descripción, comprende proteínas integrales de membrana con dependencia por sustratos dadores que sean azúcares unidos a poliprenilos. Esta es, de las 3, la superfamilia menos descrita. No existe hasta el momento ninguna estructura cristalina y este plegamiento ha sido descrito en base a alineamientos y comparaciones estructurales (36). Esta superfamilia parece incluir a 8 de las familias de CAZy. Todas son proteínas fuertemente hidrofóbicas, generalmente con 8 a 13 pasos transmembrana. Existe un loop extracitoplasmático con un motivo DXD modificado conservado, presente como EXD, DXE, DDX o DEX. Se ha observado que el rol de este motivo en la manosiltransferasa GPI-MT-I involucrada en la formación de glicosilfosfatidilinositol (GPI) es importante, pues la mutación de los aspartatos a alanina en el motivo anula la actividad. (70). De todos modos, las bases mecánicas de la catálisis en las GT-C no han sido elucidadas, siendo que no se sabe si estos residuos de DXD están involucrados efectivamente en la catálisis teniendo como función la unión de un catión divalente como es el caso de las GT-A. Cerca del extremo carboxiterminal del mismo loop extracitoplasmático existe otro residuo ácido conservado, que también podría estar involucrado en la catálisis, si bien eso resta por definirse en estudios por mutagénesis sitio-dirigida. Un punto interesante acerca de este motivo DXD es que no parece estar relacionado evolutiva y/o filogenéticamente con los de GT-A, si bien podría tener la misma función (como si fuese un caso de evolución divergente). Como características generales, el DXD en GT-C se localiza hacia la porción C-terminal de la primera hélice transmembrana y suele estar seguido de un grupo de aminoácidos hidrofóbicos. Esta forma de arreglo se parece al DXD seguido por una lámina- $\beta$  en el grupo de las GT-A, si bien no existe una similitud estadísticamente significativa entre estas regiones en GT-A y GT-C que se puedan constatar en el contexto de una búsqueda en la base de datos genómicas (36).

La superfamilia GT-C tiene una distribución filética más limitada que la que presentan GT-A o GT-B. GT-C puede encontrarse en muchos de los genomas eucariotas secuenciados, pero esta ausente en archaica y tiene muy poca presencia en procariontes. Además, otra característica distintiva es el tipo de sustratos que utilizan, que en el caso de las GT-C esta limitado a la síntesis de los derivados polisacáridicos del dolicol fosfato, una función más restringida a los eucariotes. Así, las GT-C parecen ser un grupo especializado, recientemente evolucionado, de función y distribución filogenética acotada, que podrían haber adquirido por evolución divergente o por medio de algún proceso recombinatorio el motivo DXD.

### 4.5 Función biológica de GumK: Biosíntesis del xantano

GumK es la 4<sup>o</sup> transferasa involucrada en la biosíntesis del xantano. El xantano es un exopolisacárido (EPS) ácido, producido por *X. campestris*. Este EPS está compuesto de unidades pentasacáridicas formando un esqueleto de celulosa con cadenas laterales trisacáridicas compuestas por Man( $\beta$ -1,4)GlcA( $\beta$ -1,2)Man. Éstas se unen a residuos alternados de Glc del esqueleto de celulosa por enlaces tipo  $\alpha$ -1,3 (71). Aproximadamente la mitad de residuos de Man terminales están sustituidos con piruvato. Los residuos internos de Man, generalmente están O-acetilados en el C-6. La distribución y el contenido de sustituyentes depende de las cepas y de las condiciones

de cultivo (71,72). La vía de síntesis del xantano ha sido caracterizada en detalle y constituye uno de los modelos más completos de biosíntesis de EPS. El estudio de los productos biosintéticos intermediarios producidos por varios mutantes de diferentes ORFs permitió identificar la función de varios genes involucrados en la síntesis de xantano (73-79). Algunos de estos genes están vinculados a la síntesis de los nt-azúcares precursores, mientras que los genes relacionados con las etapas siguientes de la síntesis del xantano están localizados en la región *xpsI* o *gum* (73,74). La región *gum* abarca 16 kb del genoma de *X. campestris* y contiene 12 marcos de lectura abierto (*gumB-gumM*) (GenBank U22511). Las funciones bioquímicas de los productos de los genes *gum* han sido asignadas analizando la formación de intermediarios y polímeros *in vitro* (79). Las etapas en la biosíntesis del xantano y las funciones de los genes *gum* se resumen en la Fig. 14.



**Fig. 14:** **A-** Diagrama de la organización transcripcional del operón *gum*. Las flechas naranjas muestran la posición de las secuencias promotoras de la transcripción. Los ORFs correspondientes a las glucosiltransferasas *gumD* y *gumM* están en rojo. Las manosiltransferasas *gumH* y *gumI* están en violeta y la glucuronosiltransferasa *gumK* en verde (con la posición del promotor secundario solapado en su región N-terminal). Los ORFs correspondientes a proteínas encargadas de la decoración (acetil y piruviltransferasas) están en amarillo y en azul aquellos ORFs para proteínas encargadas de la polimerización o exportación del xantano. **B-** Diagrama de la ruta de síntesis del xantano.

La proteína GumD cataliza la transferencia de Glc-1-P a poliprenil-P que actúa como aceptor. Esta reacción es reversible y constituye el primer paso en la síntesis de la unidad repetitiva del xantano. La proteína GumM cataliza la adición de Glc con

formación de un enlace  $\beta$ -1,4. GumH adiciona un residuo  $\alpha$ -1,3 Man; seguida por GumK, que adiciona un residuo de GlcA en posición  $\beta$ -1,2 y GumI que agrega la Man terminal en posición  $\beta$ -1,4. Las proteínas GumL, GumF y GumG agregan los sustituyentes a partir de acetil-CoA y fosfoenolpiruvato. GumL incorpora piruvato a la Man externa, mientras que los residuos acetato son adicionados por GumF en la Man interna y GumG en la Man externa (79). Mutantes en *gumB*, *gumC* y *gumE* son incapaces de sintetizar el polímero y se ha observado que acumulan unidades pentasacáridicas completas. Debido a esto se ha propuesto que los productos de estos genes actuarían en las etapas de polimerización y exportación del xantano. El rol de GumJ no ha sido aclarado, probablemente actúa en el transporte de la unidad repetitiva ya que ha sido agrupada en la familia PST, que incluye permeasas específicas para polisacáridos, localizadas en la membrana interna (80).

### 5. OBJETIVOS

El estudio estructural-funcional de las GTs se ha convertido en un área muy prolífica donde se están realizando importantes avances en pos de una comprensión mas profunda del mecanismo y función de este importante grupo de enzimas. Como objetivo general de esta Tesis nos planteamos el estudio de la función y estructura de GumK, ejemplo de una GT involucrada en la síntesis de un polisacárido y miembro de la aún no estudiada familia GT70.

#### 5.1 Objetivos específicos

- Estudio de la función bioquímica de GumK: localización subcelular, especificidad de sustratos.
- Estudio de la estructura molecular de GumK: Estructura de la apo-enzima y de complejos enzima-sustrato.
- Estudio de la cinética y función de mutantes de unión del sustrato dador a GumK.

## 6. MATERIALES Y MÉTODOS

### 6.1 Materiales y procedimientos generales

UDP-Glc, GDP-Man, UDP-GlcA, acetil-coenzima A (Ac-CoA) y  $\beta$ -glucuronidasa de hígado bovino se obtuvieron de Sigma. GDP-[ $^{14}\text{C}$ ]Man (280 Ci/mol) y UDP-[ $^{14}\text{C}$ ]GlcA (300 Ci/mol) UDP-[ $^{14}\text{C}$ ]Glc (231 Ci/mol), UDP-[ $^{14}\text{C}$ ]Gal (240 Ci/mol) UDP-[ $^{14}\text{C}$ ]Xyl 300 Ci/mol fueron preparados por Susana Raffo de la Fundación Instituto Leloir como se describe previamente (77,81). Salvo que se indique de otra manera, las proteínas o fracciones celulares fueron solubilizadas en buffer desnaturante 2X (8 M Urea, 3% SDS, 0,4 M ditiotreitól, 100 mM Tris-HCl, pH 8.8) por tratamiento a 100°C 5 min y luego analizadas por SDS-PAGE en geles 10% y visualizadas por tinción con Coomassie blue R250 o por Western blot (82).

### 6.2 Obtención de suero inmune contra GumK

El antisuero contra la proteína GumK fue obtenido a partir de cuerpos de inclusión (CI) recuperados e inyectados en conejo. Los CI resultaron de la sobreexpresión de la proteína en *E. coli* BL21(DE3)/pET29K en cultivos a 37°C, 3h en agitador a 200 rpm. Las células fueron colectadas por centrifugación 15.000 xg 10 min, resuspendidas en buffer 50 mM Tris-HCl pH 7,5 y fragmentadas por 2 tratamientos con prensa de French (20.000 psi). Estos extractos celulares fueron centrifugados a 3.000 xg durante 5 min a 4°C para remover agregados. Los CI fueron recuperados del sobrenadante de 3.000 xg por centrifugación a 13.000 xg durante 15 min a 4°C. El precipitado fue resuspendido en 10 mM Tris-HCl, pH 7,5 y lavado 2 veces en el mismo buffer suplementado con 2 M urea y 0,5 % Tritón X-100. Los CI lavados fueron luego resuspendidos en 50 mM Tris-HCl pH 7,5, 8 M Urea, e incubados con agitación a 4°C durante 24 h para su solubilización. Este sobrenadante se centrifugó a 12.000 xg, 10 min, 4°C y el sobrenadante fue diluido hasta una concentración 0,3 M de urea. La urea fue removida por diálisis en membranas de tamaño de poro entre 12 y 14 kDa. La proteína obtenida fue liofilizada y luego utilizada para inmunizar ratones. La inmunización de los ratones fue realizada por protocolos estándares. A saber: inmunización con 0,1 mg de proteína por ratón, en emulsión con coadyuvante de Freund completo para la primera aplicación e incompleto para las siguientes aplicaciones, una cada 10 días. El antisuero policlonal anti-GumK obtenido fue titulado y utilizado en diluciones que van de 1/100 a 1/1.000 en reacciones de western blot. Anticuerpos anti IgG de ratón, conjugados a fosfatasa alcalina (Sigma) fueron usados como anticuerpos secundarios y la detección en membranas de polivinildifluoruro (PVDF) fue realizada con los sustratos BCIP/NBT (Promega).

### 6.3 Cepas y plásmidos

**Tabla 1:** Cepas bacterianas y plásmidos utilizados

Cepa	Características relevantes	Origen/referencia
<b><i>Xantomonas campestris</i></b>		
FC2	Mutante Rif <sup>r</sup> de la cepa silvestre NRRL B-1459.	(77)
XcK	FC2 llevando el plásmido pGum52-18S integrado en el genoma. Km <sup>r</sup>	(79)
Xc1231	Rif <sup>r</sup> $\Delta$ <i>gum</i>	(73)

### *Escherichia coli*

JM109	F' <i>traD36proA+B+lac<sup>q</sup> Δ(lacA)M15/Δ(lac-proAB) glnV44e14-gyrA96 recA1 relA1 endA1 thi hsdR17</i>	New England Biolabs
DH5α	F <sup>-</sup> <i>endA1 hsdR17 supE44 thi-1 recA1 gyrA relA1 ΔlacU169 (φ80dlacZΔM15)</i>	Invitrogen
BL21(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm (DE3)</i>	Novagen
B834(DE3)	F- <i>ompT gal hsdSB (rB-mB-) met dcm lon λDE3 pLysS</i>	Novagen

---

### Vectores y plásmidos

pCHC3	Derivado pRK293 Sall-Sall 12.4 kb del operon <i>gum</i>	(74)
pBBRprom	Derivado de pBBR1MCS-5 con promotor <i>gum</i> en 5' del MCS Gm <sup>r</sup>	(83)
pET29a	Ribonucleasa S en el N-terminal Km <sup>r</sup>	Novagen
pET22b	Histag en el C-terminal Ap <sup>r</sup>	Novagen
pET29K-s	Clonado <i>Nde I-Bam HI gumk-s</i> en pET29a	esta Tesis
pET29K	Clonado <i>Nde I-Bam HI gumk</i> en pET29a	esta Tesis
pETHisKC	Clonado <i>Nde I-Xho I gumk</i> en pET22b, His Tag C-term	esta Tesis
pBBRK-s	Clonado <i>Nde I-Bam HI gumk-s</i> en pBBRprom	esta Tesis
pBBRK	Clonado <i>Nde I-Bam HI gumk</i> en pBBRprom	esta Tesis

---

## 6.4 Construcción de plásmidos

Los plásmidos fueron construidos como se describe a continuación: (i) pBBRK-s. La versión original del gen *gumK* (Nº GeneBank. U22511) que de aquí en mas denominaremos *gumK-s*, fue amplificada por PCR a partir de pCHC3 (74) usando los oligonucleótidos FGKCS (5'-GGAAGACCATATGTTCCGCTGGTATG-3' [sitio *NdeI* subrayado]) y BGKC (5'-CGCGGATCCCTCCTCAATGTGAGAGCGCTGCC-3' [sitio *BamHI* subrayado]), y digerido con las enzimas de restricción *NdeI* y *BamHI*. El fragmento de 902 pb fue ligado en el vector de amplio rango de hospedador pBBRprom (pBBR1MCS-5 con la región del promotor del operón *gum* insertada en el sitio de clonado múltiple) (83) produciendo el plásmido pBBRK-s. (ii) pBBRK. El ORF extendido para *gumK* (GeneBank Nº AY170889), para el cual mantenemos el nombre *gumK*, que se extiende 315 pb río arriba de la posición +1 de *gumK-s*, fue también amplificado a partir de pCHC3 usando los oligonucleótidos FGKLS (5'-GACACGGCATATGAGCGTCTCTC-3'[sitio *NdeI* subrayado]) y BGKC. El producto de PCR de 1217 pb fue digerido con *NdeI* y *BamHI* y ligado en pBBRprom para producir el plásmido pBBRK. (iii) pET22HisKC. El ORF *gumK* fue amplificado de pCHC3 con los oligonucleótidos FGKL (5'-CACGGCCCATGGGCGTCTCTC-3' [sitio *NcoI* subrayado]) y BHisKC (5'-CCGCTCGAGATGTGAGAGCGCTGC 3' [sitio *XhoI* subrayado]). El producto de PCR de 1216 pb fue digerido con *NcoI* y *XhoI* y ligado en los sitios correspondientes del vector pET22b(+) para producir el plásmido

pET22HisKC. Este plásmido codifica para GumK con una secuencia C-terminal LE(H)<sub>6</sub> para su posterior purificación. (iv) pET29K y pET29K-s. Los fragmentos *NdeI-BamHI* correspondientes a los ORFs para *gumK-s* y *gumK* fueron clonados en pET29a(+) resultando en los plásmidos pET29K-s y pET29K, respectivamente. Todos los fragmentos clonados fueron secuenciados para verificar que la secuencia de los genes era la correcta. Para todos los clonados realizados, los vectores e insertos digeridos fueron mezclados en proporción 10/1 y ligados con 10 U de T4-ADN ligasa (New England Biolabs) durante 2h a TA.

### 6.5 Medios y condiciones de cultivo.

La cepa *E. coli* DH5 $\alpha$  fue usada para el clonado de genes. *E. coli* BL21(DE3) o *E. coli* B834(DE3) fueron usadas en todos los casos para la expresión proteica. Las células de *E. coli* fueron cultivadas en medio Luria Broth (82) en agitador a 200 rpm, 37°C o en medio DLM para la producción de GumK sustituida con selenometioninas (84) 100 mg/l de un set de los 20 L-aminoácidos, 25 g/l glucosa, 1g/l NH<sub>4</sub>Cl, 3 g/l PO<sub>4</sub>K<sub>2</sub>H, 0,5 g/l MgCl<sub>2</sub>.6H<sub>2</sub>O, 0,5 g/l Citrato de sodio, 30 mg/l Se-Met, 10N NaOH pH 7,5. Este medio fue suplementado con: 500 mg/l adenina, 500 mg/l uracilo, 800 mg/l cisteína, 800 mg/l guanosina, 500 mg/l timidina, 5 mg/l tiamina, 4 mg/l d-biotina, llamándose a este medio DLMM (por DLM modificado) Los cultivos se realizaron en agitador a 200 rpm, 21°C *X. campestris* FC2 (cepa salvaje), XcK (*gumK*<sup>-</sup>) y Xc1231 ( $\Delta$ *gum*) fueron cultivadas en medio TY (5 g/l Triptona, 3 g/l extracto de levadura y 0,7 g/l CaCl<sub>2</sub>), en medio YM (5 g/l Triptona, 3 g/l extracto de levadura, 3 g/l extracto de malta y 10 g/l glucosa) o en medio SFFM (XOL modificado) (10 g/l K<sub>2</sub>HPO<sub>4</sub>, 0,2 g/l KH<sub>2</sub>PO<sub>4</sub>, 1 g/l (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,001 g/l MnCl<sub>2</sub>, 0,1 g/l MgCl<sub>2</sub>, 0,125 g/l Triptona, 0,125 g/l extracto de levadura, 0,125 g/l extracto de malta, 0,005 g/l FeSO<sub>4</sub> y 40 g/l glucosa, pH 7,5). Los antibióticos fueron usados a las siguientes concentraciones cuando correspondía: para *E. coli*: kanamicina, 30  $\mu$ g/ml; gentamicina, 10  $\mu$ g/ml; ampicilina 200  $\mu$ g/ml; para *X. campestris*: kanamicina, 50  $\mu$ g/ml; gentamicina, 30  $\mu$ g/ml.

### 6.6 Introducción de material genético en *E. coli* y *X. campestris*

La introducción de plásmidos en *E. coli* y *X. campestris* fue llevada a cabo por electroporación utilizando un electroporador Gene Pulser Transfection System (Bio-Rad). La preparación de células electrocompetentes y las condiciones de electroporación fueron las especificadas por el fabricante.

### 6.7 Cuantificación de la producción de xantano

Las cepas salvaje y mutantes (portando o no los plásmidos con los ORFs *gumK* y *gumK-s*) de *X. campestris* fueron inoculadas e incubadas por 72 h en medio SFFM con los antibióticos correspondientes. El xantano producido es precipitado y cuantificado por el método del cloruro de cetilpiridinio (CPC) (85) con algunas modificaciones. Brevemente, 20  $\mu$ l de sobrenadante de cultivo, conteniendo xantano en suspensión y soluciones acuosas de estándares de xantano de 0,75, 1,5, 2,25, 3 y 3,75 % fueron mezcladas con 3 ml de agua desionizada. Se agregaron 3 ml de una solución 0,36 % (p/v) de CPC y la mezcla fue incubada a 37°C por 1 h. Esta mezcla fue centrifugada a 5.000 xg 20 min para separar el precipitado de CPC/xantano. Se trazó una curva de calibración midiendo la absorbancia a 260 nm de estos estándares de xantano. La cantidad de polisacárido producida en las muestras fue estimada de su A<sub>260</sub> en relación a la curva de calibración.

### 6.8 Fraccionamiento celular y detección de las proteínas

Cultivos de *X. campestris* fueron inoculados e incubados en medio TY hasta fase estacionaria. Las bacterias fueron colectadas por centrifugación, lavadas 2 veces con agua y resuspendidas en buffer de resuspensión (70 mM Tris-HCl, 10 mM MgCl<sub>2</sub> pH 8,2) según la relación 50 equivalentes de DO/ml. Las bacterias fueron fragmentadas por 3 pasajes a través de una prensa de French a 20.000 psi. Las células no fragmentadas fueron removidas por centrifugación a 2.000 xg, 4°C, 10 min. El sobrenadante fue ultracentrifugado a 100.000 xg, 4°C, por 2 h para separar la fracción de membranas totales de la fracción soluble. La fracción de membranas totales fue solubilizada en buffer de resuspensión con el agregado de 4% SDS en un volumen equivalente al de la fracción soluble. Se tomaron alícuotas de cada fracción y fueron analizadas por SDS-PAGE y transferidas a membranas de PVDF para detección por Western blot con anticuerpos policlonales anti-GumK. Como control de la separación de fracciones subcelulares se midió el marcador de membrana ácido 2-ceto-3-deoxioctanoico (KDO) (86) y fueron analizadas por SDS-PAGE o inmunoblot. Los sobrenadantes de 100.000 xg de *E. coli* fueron utilizados para la purificación de proteína, como se describe más adelante

### 6.9 Sobreexpresión y purificación de proteínas

#### 6.9.1 Sobreexpresión de GumK-s y GumK

Para la expresión de GumK y GumK-s 200 ml de medio Luria Broth fueron inoculados con *E. coli* BL21(DE3) con los plásmidos pET29K o pET29K-s e incubados en un agitador giratorio a 200 rpm, 37°C. Al alcanzar una OD<sub>600</sub> de 0,8, se adicionó 0,5 mM IPTG y las incubaciones continuaron a 21°C por espacio de 15-18 h. Luego de esto las bacterias fueron centrifugadas, lavadas 2 veces con 70 mM Tris-HCl pH 8,2 y resuspendidos en este buffer a 50 DO equivalentes/ml. Las bacterias en suspensión fueron fragmentadas con 2 tratamientos con una Prensa de French a 20.000 psi. Los extractos celulares resultantes fueron centrifugados a 2.000 y 5.000 xg para separar células intactas de debris celular, y finalmente a 100.000 xg para separar membranas totales de la fracción soluble (citoplasma y periplasma). Los precipitados conteniendo la fracción de membranas fueron resuspendidos en 70 mM Tris-HCl pH 8,2. Estos precipitados fueron usados como fuente de enzima para ensayos de actividad previamente a contar con la enzima purificada. Alícuotas de los cultivos inducidos y controles fueron analizadas por SDS-PAGE. Para la inmunodetección del epítipo His-tag se utilizó anticuerpo monoclonal de ratón (Sigma) en dilución 1/2.000.

#### 6.9.2 Solubilización de GumK a partir de membranas

Membranas totales de *X. campestris* XcK/pBBRK o *E. coli* BL21(DE3)/pETHisKC fueron obtenidas del precipitado de 100.000 xg como se describió en la sección 6.8. Para estudiar el tipo de interacción de GumK con la membrana y además contar con proteína purificada, esta fracción fue sometida a distintos tratamientos, como se describe a continuación

##### 6.9.2.i Solubilización por pH: tratamiento con buffer alcalino o con buffer ácido

La fracción de membranas de *X. campestris* XcK/pBBRK o de *E. coli* BL21(DE3)/pETHisKC fue resuspendida en 100 mM Na<sub>2</sub>CO<sub>3</sub> pH 11,4 o con 100 mM Buffer citrato pH 3,4 para obtener una concentración final de proteínas de 5 mg/ml.

La solución fue incubada en hielo durante 30 min con agitación. Luego de la incubación se separaron la fracción solubilizada y el precipitado por centrifugación a 100.000 xg durante 1 h. Las proteínas presentes en ambas fracciones fueron detectadas por SDS-PAGE e inmunoblot. Para la inmunodetección del epítipo His-tag se utilizó anticuerpo monoclonal de ratón (Sigma) en dilución 1/2.000. GumK fue cuantificada por densitometría utilizando el software Gelworks 1D Analysis (Nonlinear Dynamics Ltd.).

### 6.9.2.ii Solubilización con fuerza iónica o por desestabilización de la membrana con EDTA

La fracción de membranas de *X. campestris* XcK/pBBRK o de *E. coli* BL21(DE3)/pETHisKC fue resuspendida en 10mM EDTA pH 8,2 o bien 1M NaCl (87) para obtener una concentración final de proteínas de 5 mg/ml. La solución fue incubada en hielo durante 30 min con agitación. Luego de la incubación se separaron la fracción solubilizada y el precipitado por centrifugación a 100.000 xg durante 1 h. El tratamiento prosiguió como se describe en el punto anterior.

### 6.9.2.iii Solubilización con detergentes no-iónicos

Se probaron diferentes concentraciones de 2 detergentes: Tritón X-100 y Tween 20. Utilizamos una concentración de proteínas totales de 10 mg/ml y concentraciones variables de los detergentes (0,1, 0,5, 1, 2, 4%) buscando la concentración mínima que nos permitiese extraer la mayor cantidad posible de enzima. La concentración final que satisfizo estas condiciones fue del 1%. Las condiciones finales de extracción fueron: 1% Tritón X-100 en agitación suave a 4°C. Bajo estas condiciones se recupero alrededor de un 60% de la enzima GumK presente en la membrana, solubilizada por el detergente.

### 6.9.3 Purificación de GumK solubilizada

Para la purificación de GumK, 1 l de bacterias *E. coli* BL21(DE3)/pET22HisKC a una DO<sub>600</sub> de 0,7, fue inducida con 0,5 mM IPTG e incubadas a 21°C por 15-18 h en agitador a 200 rpm. Las bacterias fueron cosechadas, lavadas 2 veces con 70 mM Tris-HCl pH 8,2, y resuspendidas en el mismo buffer a 20 mg/ml y fragmentadas con el uso de la prensa de French a 20.000 psi 4°C. Este extracto celular es diluido hasta una concentración de proteínas totales de 5 mg/ml, con buffer 70 mM Tris-HCl pH 8,2, 1% Tritón X-100 y es incubado con leve agitación a 4°C por 2 h, para luego ser centrifugado a 100.000 xg por 1h. El sobrenadante fue recuperado y filtrado con filtros de PVDF de 0,22 µm de poro. El sobrenadante filtrado, conteniendo a GumK solubilizada fue aplicado a columnas Ni-NTA columns (GE healthcare) conectada a un equipo de HPLC y preequilibrada en 0,5 M NaCl, 10 mM imidazol, 0,05% Tritón X-100, 20 mM Tris-HCl pH 8.0 (buffer imac A) a 0,5 ml/min. La columna fue lavada con el mismo buffer y luego se realizó elución competitiva con un gradiente de imidazol (20-400 mM) en 60 min a 1 ml/min. GumK eluyó en una concentración aproximada de imidazol de 250 mM. La proteína fue entonces concentrada por ultrafiltración y se aplicó en alícuotas de 10 mg a una columna de filtración por gel Superdex 200 HR 10/30 (GE healthcare) preequilibrada en 0,4 M NaCl, 0,05 % Tritón X-100, 50 mM Tris-HCl pH 8,0 para separar los últimos contaminantes. La elución de proteínas en cada etapa se siguió por A<sub>280</sub>. La enzima eluyó como un pico único, mostrando en geles SDS-PAGE una única banda de alrededor de 45 kDa, el peso molecular esperado. Por densitometría se estimó la pureza de la proteína como mayor al 95%. La producción de enzima fue de aproximadamente 20 mg/l de cultivo

bacteriano. Esta proteína concentrada es la utilizada para los ensayos de actividad y cristalización.

### 6.9.4 Purificación de GumK sustituida con selenometionina (Se-Met GumK)

Para la expresión de Se-Met GumK se utilizó la cepa auxotrofa B834(DE3)/pETHisKC en medio DLMM. Se observó máxima expresión de Se-Met GumK a 21°C, 20 h, 0,5 mM de IPTG. Se purificó esta proteína de la misma manera que GumK nativa y agregándose a los buffers de lavado y elución TCEP (Tris 2-carboxi-etil fosfina) 2 mM como agente reductor, evitando otros agentes con mayor poder reductor como DTT o  $\beta$ -mercaptoetanol (incompatibles con columnas IMAC). La proteína obtenida como el sobrenadante de una ultracentrifugación de 100.000 xg fue utilizada para la purificación, que se realizó de la misma manera que para GumK nativa. El rendimiento de la proteína sustituida fue de aproximadamente 12 mg/l de cultivo bacteriano.

## 6.10 Actividad enzimática

### 6.10.1 Preparación del sustrato aceptor glicolípido

El sustrato aceptor de GumK no está disponible comercialmente, por lo tanto debía ser sintetizado químicamente o enzimáticamente para luego ser purificado. Por esto, intentamos hacer una síntesis quimio-enzimática a partir de fitanil-pirofosfocelobiosa (Fitanil-PP-Cel) incubado con GDP-Man en presencia de AceA, una manosiltransferasa de *Acetobacter xylinum* purificada y disponible en nuestro laboratorio. (6). Esta síntesis no fue posible, principalmente debido a problemas con la preparación de Fitanil-PP-Cel. Como alternativa proseguimos con la extracción y purificación del sustrato natural polisoprenil-PP-glucosa- $\beta$ -1,4-glucosa- $\alpha$ -1,3-manosa (Poliprenil-PP-Cel-Man) a partir de su fuente natural, la membrana de *X. campestris*. Para aquellas incubaciones que incluían el sustrato aceptor sin marca radioactiva purificamos 2  $\mu$ moles de sustrato por el método de las células permeabilizadas (16). Para ello se prepararon 20 litros de cultivo saturado de *X. campestris* FC2 en medio SFFM, sin glucosa para evitar la síntesis de xantano. Este medio fue elegido por sobre el medio YM o TY por su mejor rendimiento en unidades de DO por unidad de volumen. Los cultivos bacterianos, al llegar a fase estacionaria (en aproximadamente 72hs) fueron colectados por centrifugación 5.000 xg 30 min y resuspendidos en 200 ml de 10 mM EDTA. Las células fueron permeabilizadas con EDTA como se describió anteriormente (77). Estas células así permeabilizadas fueron utilizadas en una incubación con los precursores nucleotídicos 3 mM UDP-Glc y 1,5 mM GDP-Man en buffer 0,2 M Tris-HCl pH 8,2 suplementado con 75 mM MgCl<sub>2</sub>. El producto de esta incubación, junto con impurezas de la preparación, fue recuperado por extracción con solventes orgánicos: brevemente, se lavaron los precursores nucleotídicos con 250 ml de agua 3 veces y se extraen los glicolípidos con un volumen de 200 ml de cloroformo:metanol:agua 1:2:0.3 3 veces. Este extracto que contiene lípidos azúcares intermediarios (Poliprenil-PP-Cel-Man, entre otros) será referido como extracto 1203. Este sustrato es almacenado a -20°C para ser usado posteriormente. El sustrato fue finalmente concentrado bajo corriente de N<sub>2</sub> hasta 20 mM.

### 6.10.2 Preparación de diferentes sustratos aceptores marcados radioactivamente

Para la obtención del sustrato natural en cantidades suficientes solo para 3-4 reacciones se utilizaron 150  $\mu$ l de células permeabilizadas (3-4 mg de proteína total) que fueron incubadas con 8 mM  $MgCl_2$ , 70 mM Tris-HCl pH 8,2, UDP-Glc y GDP-[ $^{14}C$ ]Man. Las reacciones se llevaron a cabo en un volumen final de 350  $\mu$ l a 20°C durante 30 min. La reacción se detuvo con el agregado de 1 ml de 10 mM EDTA, 70 mM Tris-HCl pH 8,2. Las células fueron recuperadas por centrifugación y lavadas dos veces con el mismo buffer. El aceptor presente en la membrana de las células se extrajo con tres lavados de 150  $\mu$ l de cloroformo-metanol-agua (1:2:0.3 v/v/v). El extracto fue concentrado bajo corriente de nitrógeno a un volumen aproximado de 20  $\mu$ l para ser utilizado inmediatamente en la reacción de glicosiltransferencia de GumK.

La obtención de Cel-Man-PPi se hizo de acuerdo a García y col. (81). Se realizó una incubación con células permeabilizadas en presencia de UDP-Glc y GDP-[ $^{14}C$ ]Man, para favorecer la formación Poliprenil-PP-Cel-Man y se extrajo la fracción de glicolípidos (extracto 1203) que fue reducida hasta 20  $\mu$ l bajo corriente de  $N_2$ . Se agregaron 160  $\mu$ l de 50 % fenol y se incubó a 70°C durante 30 min. Luego, la mezcla fue enfriada a TA y centrifugada a 13.000 xg, obteniendo una fase acuosa y otra fenólica. Esta última fue lavada tres veces con 60  $\mu$ l de agua. Las fases acuosas de cada lavado fueron juntadas y lavadas tres veces con éter etílico. La muestra obtenida fue utilizada como aceptor en la reacción de GumK.

El aceptor Poliprenil-PP-Cel-(Ac)Man fue preparado mediante la adición de 0,7 mM Ac-CoA a la mezcla de reacción. Bajo estas condiciones, aproximadamente el 60% de Poliprenil-PP-Cel-Man se transforma a su forma acetilada (y por ende queda un 40% no acetilado). Este sustrato mixto se denominará "aceptor acetilado mixto". Los extractos conteniendo el aceptor acetilado mixto fueron reducidos bajo corriente de  $N_2$  hasta un volumen de aproximadamente 15  $\mu$ l y usados en ensayos de actividad.

El sustrato sintético Fitanyl-PP-Cel-Man fue obtenido por incubación con Fitanyl-PP-Cel y GDP-Man con la manosiltransferasa de *A. xylinum* AceA como se reporta en (88). Brevemente, la mezcla de incubación para preparar Fitanyl-PP-Cel-Man contenía 0,4  $\mu$ g de AceA purificada, 1% Tritón X-100, 6 mM  $MgCl_2$ , 20  $\mu$ M de aceptor sintético Fitanyl-Cel-PP y 0,25 mM de GDP-Man en un volumen final de 100  $\mu$ l. Luego de la incubación a 37°C por 10 min, la reacción fue detenida por la adición de 200  $\mu$ l de cloroformo:metanol (1:1) y 120  $\mu$ l de agua. El aceptor glicolípido es recuperado en la fase orgánica luego de centrifugar a 6.000 xg por 5 min. El volumen del aceptor recuperado fue reducido bajo flujo de  $N_2$  hasta aproximadamente 15  $\mu$ l.

### 6.10.3 Ensayos de actividad de GumK

La actividad de GumK se midió mediante la transferencia de un residuo de GlcA desde el nucleótido-azúcar UDP-GlcA hacia su aceptor natural Poliprenil-PP-Cel-Man o sintético Fitanyl-PP-Cel-Man. La mezcla de reacción contenía 15  $\mu$ l del aceptor glicolípido, 0,5 mM  $MgCl_2$ , 3% Tritón X-100 y 0,15  $\mu$ Ci UDP-[ $^{14}C$ ]GlcA, 70 mM Tris-HCl pH 8,2, en un volumen final de 100  $\mu$ l. La reacción comienza con la adición de 1  $\mu$ g de GumK purificada o 50  $\mu$ g del precipitado de 100.000 xg conteniendo GumK o GumK-s (las construcciones sin tags). Las incubaciones se realizaron a 20°C por 30 min y son finalizadas con la adición de 200  $\mu$ l de cloroformo:metanol (1:1) y 120  $\mu$ l de agua. Esta mezcla fue centrifugada 6.000 xg 10 min, produciéndose una partición en 2 fases. La fase orgánica, conteniendo los

productos liposacáridicos de reacción, fue recuperada y la fase soluble fue desechada. La fracción orgánica fue sometida a hidrólisis ácida suave del enlace fosfato (0,01 N HCl 10 min a 100°C) para liberar la fracción oligosacáridica del aceptor. La fracción lipídica fue extraída con 0,6 ml de cloroformo:metanol (2:1). La fase acuosa que contiene los oligosacáridos fue secada por completo y resuspendida en 0,5 ml de piridina. Las sales remanentes fueron removidas por centrifugación y el sobrenadante conteniendo los oligosacáridos marcados fue recuperado. Se cuantificó la radiactividad presente con un contador de centelleo, midiendo el 10% de este sobrenadante. El resto de la muestra fue secada, resuspendida en agua y analizado por TLC. Se utilizó como fase móvil una mezcla de 1-propanol:nitrometano:agua (5:2:2). Las marcas radioactivas correspondientes a los oligosacáridos liberados fueron visualizadas y cuantificadas usando el software ImageQuant software en un scanner Storm 820 (Molecular Dynamics).

Para los ensayos en los que se probaron sustratos aceptores solubles en agua, la reacción procedió como se detalla arriba, pero en este caso, luego de la partición orgánica, se recuperó la fracción soluble (donde se hallan los putativos productos de esta reacción). Se eliminaron las sales de estas fracciones acuosas por extracción con piridina y los oligosacáridos producidos en estas reacciones fueron analizados por TLC.

Para determinar la importancia de iones metálicos en la reacción catalizada por GumK, se omitió la presencia de MgCl<sub>2</sub> en la mezcla de incubación. El efecto del EDTA se evaluó preincubando la enzima con 10 mM del agente quelante, 4°C durante 1 h.

Para determinar la especificidad de dador, se realizaron las incubaciones como se describe pero reemplazando el dador radiactivo con ADP-[<sup>14</sup>C]Glc, UDP-[<sup>14</sup>C]Glc, UDP-[<sup>14</sup>C]Gal, UDP-[<sup>14</sup>C]GlcA, UDP-[<sup>14</sup>C]Xyl o UDP-GalA, UDP-GlcNac, GDP-[<sup>14</sup>C]Man o UDP-[<sup>14</sup>C]Xyl. Los glicolípidos fueron aislados de la fase orgánica y se continuaron procesando como se indicó anteriormente.

### 6.10.4 Preparación de oligosacáridos marcados radioactivamente usados como estándares

Cel-Man, Cel-(Ac)Man y Glc-β-1,4-Glc-α-1,3-Man-β-1,2-GlcA-β-1,4-Man (Cel-Man-GlcA-Man) (estándares de trisacárido, trisacárido acetilado y pentasacárido respectivamente) marcados radioactivamente fueron preparados por tratamiento ácido suave a partir de los correspondientes poliprenil-difosfo-oligosacáridos, obtenidos mediante incubaciones con células permeabilizadas como se detalla en la sección anterior. El tetrasacárido GlcA-β-1,2-Man-α-1,3-Glc-β-1,4-Glc (Cel-Man-GlcA) y su derivado acetilado (Cel-(Ac)Man-GlcA) fueron preparados por un procedimiento similar, reemplazando células permeabilizadas de la cepa salvaje FC2 por células de la mutante de *X. campestris* Xcl. Esta cepa es incapaz de transferir el último residuo de manosa al tetrasacárido intermediario (79).

### 6.10.5 Tratamientos con β-glucuronidasa

Los tratamientos con β-glucuronidasa fueron realizados como se describe anteriormente (60). La mezcla de reacción contiene: 100 mM citrato de sodio pH 5,0, 30.000 cpm de oligosacáridos marcado radioactivamente, obtenido por hidrólisis ácida suave de Poliprenil-PP-Cel-Man-GlcA, y 0,4 U de β-glucuronidasa de hígado bovino en un volumen final de 50 μl (Sigma). Las incubaciones se realizaron por 8 h a 37°C y las reacciones finalizaron con la adición de 1 volumen de etanol. Los sobrenadantes fueron desalados como se indica para los ensayos de actividad de GumK y analizados por TLC usando butanol:ácido acético:agua (2:1:1) como fase móvil.

**6.11 Cromatografía en capa delgada (TLC) y detección de oligosacáridos radioactivos.**

Las muestras de oligosacáridos obtenidas de las distintas incubaciones fueron sembradas junto a estándares radiactivos en placas de TLC (silica gel 60, 0,25 mm, SIGMA). Las cromatografías se desarrollaron dos veces en 1-propanol-nitrometano-agua (5:2:2 v/v). Los oligosacáridos radioactivos fueron visualizados con un Phosphoimager (Storm 840, Amersham-Pharmacia) luego de una exposición de 24 h para 10.000 cpm por muestra/calle.

**6.12 Mutagénesis sitio-dirigida**

Todas las mutantes de GumK en residuos específicos fueron obtenidas con el QuickChange Site directed mutagenesis Kit. Sobre pETHisKC (Qiagen), utilizando los oligonucleótidos mutagénicos que figuran en la Tabla 2.

**Tabla. 2:** Oligonucleótidos mutagénicos

Secuencia de los primers	Cambio introducido
FD157A: 5'-CTATCGCGCGTGCGCCGGGCTGAGCACC -3'	D157A
BD157A: 5'-GGTGCTCAGCCC <u>GGCC</u> GACGCGGATAG -3'	D157A
FD157E: 5'-CTATCGCGCGTGCGAAGGGCTGAGCACC -3'	D157E
BD157E: 5'-GGTGCTCAGCCC <u>TTCC</u> GACGCGGATAG -3'	D157E
FD157N: 5'-CTATCGCGCGTGCAACGGGCTGAGCACC -3'	D157N
BD157N: 5'-GGTGCTCAGCCC <u>GTTCC</u> GACGCGGATAG -3'	D157N
FD207A: 5'-GTAGGCCACGGCGTG <u>GCCC</u> ACAACCTCGATC -3'	D207A
BD207A: 5'-GATCGAGGTTGTGGG <u>CC</u> CACGCGTGCCCTAC -3'	D207A
FD234A: 5'-GGGTGCGATGCTGTTT <u>GCCC</u> CTGAATTTTTCGTCG -3'	D234A
BD234A: 5'-CGACGAAAAATTCAGGGGCAACAGCATCGACCC -3'	D234A
FE192A: 5'-GCGATGGCCGCAGCAGTAGTGAGCCGC -3'	E192A
BE192E: 5'-GCGGCTCACTACT <u>GCT</u> GCGGCCATCGC -3'	E192A
FE272A: 5'-GTCATTGTCTATGGC <u>GCA</u> ATGAAGCACGCGCAG -3'	E272A
BE272A: 5'-CTGCGCGTGCTTCA <u>TTC</u> GCGCCATAGACAATGAC -3'	E272A
FR29A: 5'-TTCCGCACGCCAG <u>CT</u> CGGGCTAACATC -3'	R29A
RR29A: 5'-GATGTTAGCCCAG <u>CT</u> TGGCGTGCGGAA -3'	R29A
FM231A: 5'-GCAGTCGGGT <u>CGGC</u> GCTGTTTGATCCTG -3'	M231A
RM231A: 5'-CAGGATCAAACAG <u>CGCC</u> GACCCGACTGC -3'	M231A
FE272D: 5'- CATTGTCTATGGC <u>GAC</u> ATGAAGCACGCG -3'	E272D
RE272D: 5'-CGCGTGCTTCAT <u>GTC</u> GCCATAGACAATG -3'	E272D
FY292A: 5'-GGCATTGCGCCT <u>GCC</u> GCGTGCAGCAG -3'	Y292A
RY292A: 5'-CTGCTCGGACGCG <u>GCC</u> AGGCGCAATGCC -3'	Y292A
FL301A: 5'-GTGCCGGTGTAT <u>GCC</u> GCAGACAGCTC -3'	L301A
RL301A: 5'-GAGCTGTCTG <u>CCGC</u> ATACACCGGCAC -3'	L301A
S304A: 5'-GTATCTGGCAGAC <u>GCCT</u> CAATGAAATTG -3'	S304A
RS304A: 5'-CAATTTTATTGAGGCGTCTGCCAGATAC -3'	S304A
FK307A: 5'-GACAGCTCAAT <u>GGC</u> ATTGCTGCAATAC -3'	K307A
RK307A: 5'-GTATTGCAGCAAT <u>GCC</u> ATTGAGCTGTC -3'	K307A
FQ310A: 5'-CAATGAAATTGCTG <u>GCA</u> TACGACTTTTTTC -3'	Q310A
RQ310A: 5'-GAAAAAGTCGTAT <u>GCC</u> AGCAATTTTCATTG -3'	Q310A

La fidelidad de las reacciones de mutagénesis fue confirmada por secuenciación de los ORF completos. La expresión y purificación de las proteínas mutadas fue

realizada de la misma manera que GumK salvaje. Las mutantes purificadas fueron almacenadas a 4°C.

### 6.12.3 Incubaciones radiactivas para la obtención de parámetros cinéticos de GumK y GumK mutada

Los parámetros cinéticos fueron determinados utilizando varias concentraciones de un sustrato (12,5-1.000  $\mu\text{M}$  UDP-GlcA, 25-1.000  $\mu\text{M}$  Poliprenil-PP-Cel-Man) y manteniendo el otro a una concentración fija (2 mM UDP-GlcA, 0,5 mM Poliprenil-PP-Cel-Man). Los parámetros cinéticos de GumK salvaje o mutantes fueron medidos utilizando los ensayos de actividad con sustrato radiactivo (UDP-[ $^{14}\text{C}$ ]GlcA en diferentes concentraciones y actividad específica). Las incubaciones se realizaron agregando 0,5  $\mu\text{g}$  de GumK en un volumen de 100  $\mu\text{l}$ . Uno de los 2 sustratos fue agregado a concentraciones saturantes (aproximadamente 10 veces la  $K_m$  para UDP-GlcA y aproximadamente 3 veces la  $K_m$  del aceptor, pues la cantidad y solubilidad de esta molécula era limitada) y el restante a concentraciones variables para graficar los puntos de las curvas. La incubación incluía 2% Tritón X-100 y se realizaba en buffer 20 mM Tris-HCl pH 8.0. El tiempo de incubación es de 2 min. La reacción se detenía mediante el agregado de 200  $\mu\text{l}$  de 1:1 cloroformo:metanol y 120  $\mu\text{l}$  de agua. La extracción de los productos radiactivos se realizó como se describe en la sección 6.10.3. Cada incubación se realizó por triplicado y los resultados mostrados son el resultado promedio de estas 3. Los parámetros cinéticos fueron obtenidos de gráficos de Lineweaver-Burk y utilizando el programa de calculo de constantes cinéticas Prism4. En todos los casos, el error estándar no superó el 5%.

### 6.13 Cristalización

#### 6.13.1 Apo-proteína

Las condiciones iniciales de cristalogénesis fueron obtenidas utilizando el método de la gota colgante en placas utilizando los buffers de cristalización Crystal Screen kits (Hampton Research). La búsqueda de las condiciones iniciales de cristalización se llevo a cabo utilizando volúmenes de reservorio de 500  $\mu\text{l}$ . GumK, luego de ser purificada, fue concentrada mediante ultracentrifugación hasta 20-23 mg/ml. Los cristales fueron obtenidos mezclando 2  $\mu\text{l}$  de proteína con 2  $\mu\text{l}$  de los buffers de cristalización consistentes en 30% PEG 4000, 0,1 M Tris-HCl, pH 8,5 y como co-precipitante A) 0,2 M  $\text{Li}_2\text{SO}_4$ , B) 0,2 M  $\text{MgCl}_2$  o C) 0,2 M acetato de sodio. Los cristales obtenidos mostraron una forma romboidal, creciendo hasta un tamaño final de 0,2 x 0,1 mm en 3 semanas a 20°C. La optimización fue lograda mediante suaves variaciones alrededor de las condiciones iniciales, variando la concentración de las sales ( $\text{Li}_2\text{SO}_4$ ,  $\text{MgCl}_2$ , o acetato de sodio) o el PEG. El pH óptimo de cristalización fue ajustado mediante el uso buffers Tris-HCl y fosfato en el rango de 6-9. Asimismo, se logro una mayor optimización con el uso de Hampton Research Additive y Detergent Screens 1, 2 y 3. Finalmente, los cristales de mejor forma y tamaño y que mejor difractaban se obtuvieron mezclando 3  $\mu\text{l}$  de proteína con 1  $\mu\text{l}$  del buffer de cristalización optimizado (35% PEG 3350, 0,1 M Tris-HCl; 0,2 M  $\text{Li}_2\text{SO}_4$ , 0,1 M CsCl, pH 8,2). Con esto se obtuvieron cristales de 0,3 x 0,15 mm que polarizaban fuertemente la luz. Previo a la colección de datos, los cristales fueron transferidos a solución crioprotectora 35% (w/v) PEG 3350; 0,1 M Tris-HCl; 0,2 M  $\text{Li}_2\text{SO}_4$ , pH 8,2 por 1 min y congelados en nitrógeno líquido para su posterior análisis.

### 6.13.2 Complejo GumK-sustrato

Para la obtención de este complejo se realizaron 2 series de experimentos: en primer lugar se intentó co-cristalizar a GumK en presencia de 10 mM UDP-GlcA. Esta aproximación no funcionó, pues en estos experimentos solo se obtuvieron cristales de apariencia heterogénea y que no difractaban los rayos-X. Como segunda opción se realizaron experimentos de inmersión de los cristales en buffer de cristalización suplementados con 10 o 100 mM UDP-GlcA por espacio de 30 min a 8 hs. Luego de la inmersión estos cristales fueron congelados en nitrógeno líquido y se colectaron grupos de datos de difracción en el sincrotrón LNLS, Campinas, SP, Brasil. También se realizaron experimentos de co-cristalización en presencia del sustrato aceptor natural Poliprenil-PP-Cel-Man purificado de membranas de *X. campestris*. Los tiempos de inmersión fueron de 1, 3 y 8 hs en este caso con una concentración final de 10 mM.

### 6.14 Métodos bioinformáticos

Para la predicción de semejanza estructural, estructura secundaria o alineamientos de secuencia se utilizaron los programas o servidores Dali, BLAST, Jpred, disponibles en el server ExPASy (<http://expasy.org/>). Para la predicción de dominios transmembrana se utilizó el programa DAS (Dense alignment surface method) que se encuentra disponible en (<http://www.sbc.su.se/~miklos/DAS/>) (89). La predicción de coiled-coils se realizó con el server de predicción disponible en ([http://www.ch.embnet.org/software/COILS\\_form.html](http://www.ch.embnet.org/software/COILS_form.html)) (90). Los cálculos de potencial electrostático de superficie de la proteína se realizaron según el algoritmo de Poisson-Boltzmann y se visualizaron por Pymol (DeLano Scientific LLC).

### 6.15 Mediciones de tamaño molecular de GumK en solución

Las mediciones de dispersión de luz para tamaño de partícula en GumK se realizaron en HPLC LKB 2248, en columnas de filtración por gel S-200 (GE healthcare), utilizando un detector de luz UV variable Waters 486 y un refractómetro diferencial LKB para las mediciones de índice de refracción. La concentración de GumK era de 20 mg/ml y la masa total de proteína fue de 400 µg.

### 6.16 Resolución de la estructura cristalina de GumK

#### 6.16.1 Caracterización cristalográfica preliminar del cristal nativo

Los sets de datos de difracción de rayos X fueron colectados a partir de cristales únicos a 110 K en los sincrotrones de NSLS, Brookhaven National Laboratories, NY, EEUU, en la línea X12C ( $\lambda=1,100$  Å) equipada con un detector del tipo ADSC Q210. Los datos de difracción fueron procesados utilizando los programas MOSFLM, SCALA y TRUNCATE del pack de programas CCP4 v5.0.2 (91). Los datos fueron integrados y procesados hasta una resolución de 1,9 Å. Para este grupo de datos de GumK nativo en total se procesaron 240 imágenes de difracción de un ángulo de oscilación de 0,5°.

#### 6.16.2 Colección de datos de difracción e información de fases

Se tomaron cristales de derivados isomorfos de Platino luego de inmersión de 2h en buffer de cristalización optimizado suplementado con 10 mM de  $K_2PtCl_4$ . Estos cristales fueron congelados en corriente de  $N_2$  utilizando el buffer de cristalización

como crioprotector. Se realizaron escaneos de fluorescencia sobre el cristal derivado para determinar correctamente los valores de  $f'$  y  $f''$  correspondientes y se determinaron los valores de pico e inflexión de la curva de fluorescencia, correspondiendo a longitudes de onda: 1,0718 y 1,0722 Å respectivamente. Se colectaron sets de datos de difracción por rayos-X en ambas longitudes de onda a una temperatura de 100 K usando radiación de sincrotrón (línea X12C, NSLS, BNL, EEUU) en un detector ADSC Q210 modificado. Cada set de datos constó de 240 imágenes de difracción de 0,5° de oscilación. Las intensidades de las reflexiones fueron integradas usando MOSFLM v6.2.3, agrupadas con SCALA v3.2.5 y reducidas con Truncate (91). La estructura fue resuelta con un experimento de MAD usando 2 longitudes de onda ( $\lambda=1,0718$  y 1,0722 correspondientes a los puntos del máxima e inflexión del derivado de Pt). Por cálculos de ocupancia y medición por métodos directos según el programa SHELX se encontraron 4 o más átomos de platino unidos a la estructura. Las posiciones, factores B de temperatura y ocupancias de estos átomos fueron refinadas usando Sharp (92). Luego de 6 rondas de refinamiento e inspección de mapas de Fourier residuales, se identificaron 4 posiciones de platino en la estructura. La modificación de la densidad de solvente se realizó usando los programas DM y Solomon (91). Todo este procedimiento resultó en un mapa de densidad electrónica interpretable. Parámetros estadísticos como el phasing power anómalo mostraban que se podía obtener información de fases a una resolución de 2,5 Å. La extensión de fases y el refinamiento final de la estructura se realizó con los datos procesados para el cristal nativo descrito anteriormente. Una fracción del 5% de los datos de difracción fueron dejados aparte para los cálculos de  $R_{\text{free}}$ .

### 6.16.3 GumK en complejo con UDP

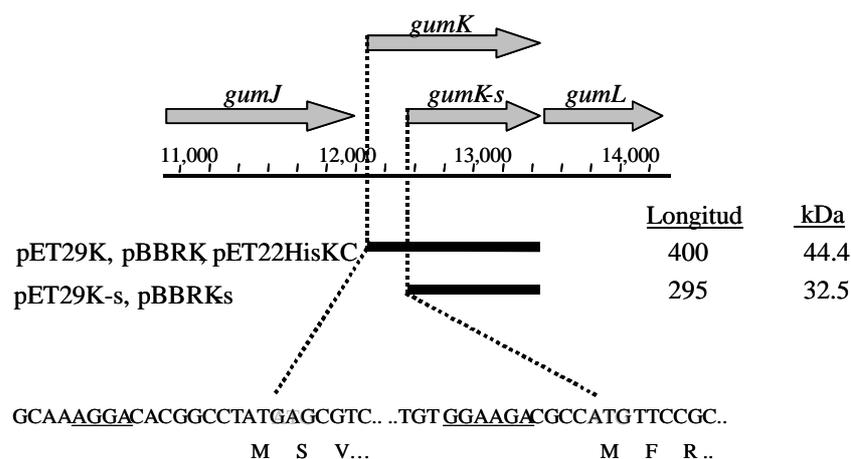
Los datos de difracción fueron colectados en la línea DO3B-MX1, LNLS, Campinas, SP, Brasil, a una longitud de onda  $\lambda= 1,427$  Å, a 100 K en un detector CCD (MarResearch). Los grupos de datos fueron integrados y procesados con los programas MOSFLM, SCALA y TRUNCATE del set de programas CCP4. Para la construcción de modelos y mapas de densidad se utilizó el programa ARP/wARP con las fases de la apo-enzima como modelo para el trazado de la proteína. Como resultado de este procedimiento se obtuvieron mapas de densidad atómica y coordenadas para la gran mayoría de las posiciones atómicas de la proteína. Estas coordenadas fueron refinadas por Refmac o a mano en Coot (93,94). Luego, en mapas de densidad atómica 2Fo-Fc se buscó la densidad atómica correspondiente a la posición del sustrato.

## 7. RESULTADOS

### 7.1 Análisis de la función del gen *gumK* por complementación funcional

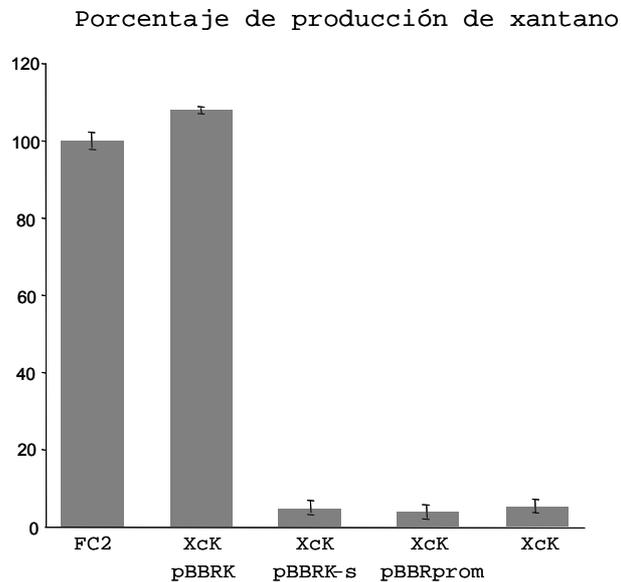
Una de las líneas de trabajo de nuestro laboratorio ha sido el estudio de la función y estructura de las proteínas encargadas de la síntesis del xantano. El análisis de los intermediarios biosintéticos producidos por mutantes genómicos de diferentes ORFs del operón *gum* fue el punto de partida para estos estudios. El análisis de los intermediarios producidos por XcK (mutante *gumK* de *X. campestris*) (79) mostraban que esta mutante acumulaba Poliprenil-PP-Cel-Man, indicando una deficiencia en la actividad de transferencia del residuo de ácido glucurónico a la subunidad pentasacárida del xantano, Poliprenil-PP-Cel-Man-GlcA-Man. Para confirmar que el ORF *gumK* se correspondía con la proteína GumK (y por ende con la actividad de  $\beta$ -glucuronosiltransferasa), se llevó a cabo un experimento de complementación con la mutante XcK portando el plásmido pBBRK-s (que contiene el ORF de 890 pb postulado para *gumK*, Fig. 15). El plásmido pBBRK-s deriva del vector modificado pBBRprom, que es a su vez un derivado de pBBRMCS5 (95) con el agregado de toda la región promotora del operón *gum* hacia el 5' del sitio de clonado múltiple. Este vector está especialmente diseñado para emular lo mejor posible la situación en el genoma de la bacteria. El objetivo es evaluar lo más fehacientemente posible los resultados de complementación *in vivo* de la mutante XcK. Así, el exopolisacárido xantano, en caso de ser producido, fue aislado y cuantificado de cultivos de fase estacionaria de las cepas FC2 (salvaje) y la mutante XcK portando los plásmidos pBBRprom o pBBRSK-s. Sorprendentemente la cuantificación del polisacárido producido por XcK/pBBRK-s fue un 95% menor que la de la cepa salvaje, mostrando que la complementación había sido nula, si bien habíamos verificado previamente por western blot la expresión de la proteína a niveles satisfactorios (Fig. 17)

Al inspeccionar mas detenidamente la secuencia nucleotídica del operón *gum*, se hizo inmediatamente evidente que *gumK* se podía extender 315 pb hacia el extremo 5' usando un nuevo codón de iniciación ATG que respeta el marco de lectura. Para este nuevo ORF extendido se mantuvo el nombre *gumK* (número de acceso GeneBank AY170889). Consecuentemente el nombre para el ORF mas corto de *gumK* fue cambiado a *gumK-s*. Se encontró una secuencia del tipo Shine–Dalgarno 8 nucleótidos río arriba de este nuevo codón de iniciación (Fig. 15)



**Fig 15:** Mapa parcial y organización genética del operón *gum* de *X. campestris*. Las flechas indican los ORFs. Los números de nucleótidos se refieren al operón *gum* (GenBank U22511). También se presentan los ORFs *gumJ* y *gumL* orientados en la misma dirección. Las regiones 5' de los ORFs *gumK* y *gumK-s* están representadas mostrando los primeros residuos de cada proteína. Los potenciales sitios de unión a ribosomas están subrayados. Ambos ORFs fueron clonados produciendo los plásmidos que se indican.

El gen *gumK* de 1.203 pb fue clonado en pBBRprom y el plásmido resultante pBBRK fue introducido en la mutante XcK. En este caso se observa que la síntesis de xantano se reestablece a niveles comparables a la cepa salvaje (Fig. 16). Este resultado fue la primera confirmación de que es el gen de 1.203 pb el que restaura en la mutante la actividad glucuronosiltransferasa.



**Fig. 16:** Producción relativa porcentual de xantano en la mutante XcK portando el plásmido pBBRK, pBBRK-s o pBBRprom. Los resultados presentados representan la media de al menos 3 mediciones independientes.

### 7.2 GumK se expresa en *Xanthomonas campestris* a partir de un ORF mas extendido, localizándose en membranas

La secuencia de GumK-s es 105 residuos mas corta que la de GumK (Fig. 15, 295 residuos vs 400 residuos). Con el objetivo de establecer cual de estas proteínas se expresa *in vivo*, se realizaron ensayos de western blot usando suero de ratones con anticuerpos anti-GumK. Una proteína con peso molecular de 44 kDa fue detectada en extractos de membrana de *X. campestris* salvaje FC2 (Fig. 17, calle 1). Este valor se corresponde con el valor de peso molecular de GumK, lo cual indica que la proteína de 44 kDa representa la proteína expresada por la cepa salvaje. En este mismo ensayo, se separaron las fracciones soluble y de membrana (MyM sección 6.8) en cada una de las cepas analizadas. De esta manera se pudo establecer la localización subcelular de GumK tanto en la cepa salvaje, mutante y mutante complementada. En la fracción de membranas de la cepa salvaje se observó una única banda correspondiente a GumK, que estaba ausente en la cepa con delección de todo el operón *gum* (Xc1231) y en la mutante XcK, indicando que efectivamente la señal de 44 kDa corresponde a GumK (Fig. 17 calles 9-12). La enzima también fue detectada en la mutante complementada XcK/pBBRK y en XcK/pBBRK-s, así como en Xc1231/pBBRK. Interesantemente, en la mutante Xc1231/pBBRSK la enzima se localizó en membranas de la bacteria, indicando que la unión a membranas ocurre aun en ausencia de las demás proteínas del operón *gum*, tal vez descartando una asociación funcional entre GTs, flipasas y demás componentes de la maquinaria de biosíntesis y exportación del xantano.

Estos resultados claramente muestran que la proteína de 44 kDa se sintetiza en la cepa salvaje a partir del gen *gumK* de 1.203 pb. Además la proteína activa se encuentra en la membrana de la bacteria. Lo primero que notamos (y que luego confirmamos en la estructura cristalográfica) es que esta proteína no presentaba hélices transmembrana. Por ende la proteína sería periférica de membrana, posiblemente unida mediante interacciones del tipo electrostáticas/hidrofóbicas.

También existía la posibilidad que hubiera interacción con la membrana a través de hélices anfipáticas, posibilidad que fue descartada por estudios de predicción por secuencia y análisis de la estructura cristalina. La localización en membranas también se observa en otras GTs involucradas en la vía biosintética del xantano (96) (datos no publicados de nuestro laboratorio) o de acetano (88), indicando que GTs involucradas en la síntesis de EPS pueden unirse a membrana de manera intrínseca, es decir sin asociarse a otros componentes de la maquinaria biosintética.



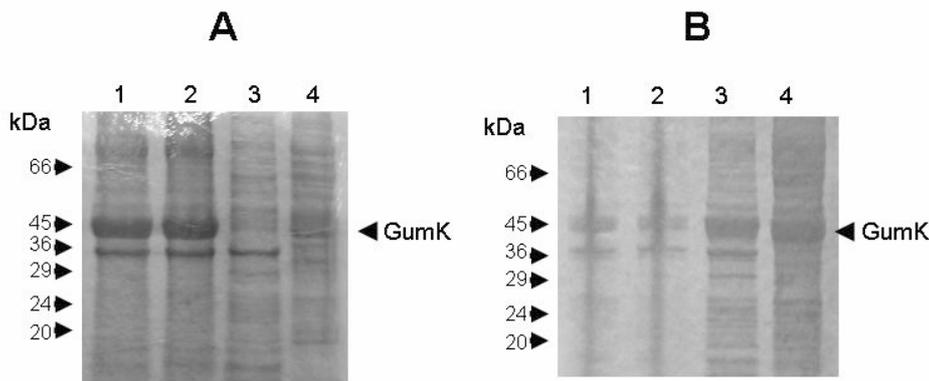
**Fig. 17:** Detección por western blot y localización subcelular de GumK en la cepa salvaje FC2 de *X. campestris* y en distintas cepas mutantes. Las fracciones de membrana (M) y soluble (S) fueron preparadas como se describe en MyM sección 6.8. Calles 1 y 2 correspondientes a la cepa salvaje FC2, Calles 3 y 4 cepa Xc 1231/pBBRK, calles 5 y 6 cepa XcK/pBBRK, calles 7 y 8 cepa XcK/pBBRK-s. Las calles 9-10 (cepa mutante XcK) y 11-12 (mutante Xc 1231) son controles negativos.

### 7.3 Producción y purificación de GumK

#### 7.3.1 Solubilización

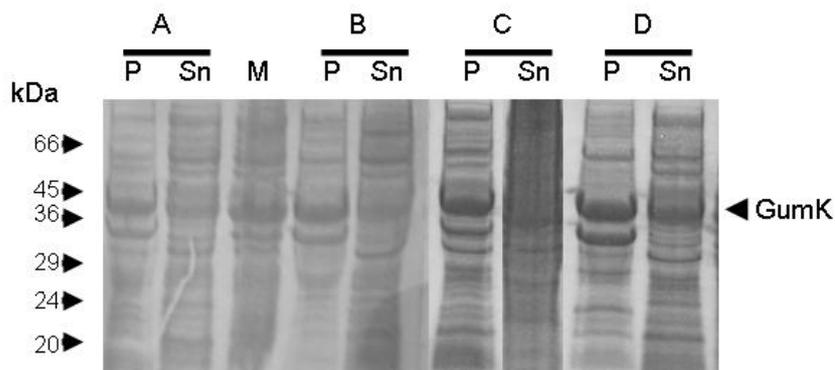
Para su purificación GumK fue subclonada en el vector de expresión procariótico pET22b, codificando para una proteína con una cola de polihistidina ( $\text{His}_6$ ) en su extremo C-terminal. El plásmido que codifica esta proteína recombinante, pETHisKC fue introducido en la cepa de *E. coli* BL21(DE3) y la proteína fue sobrepresada según las condiciones descritas en MyM sección 6.9.3.

Una prominente banda de alrededor de 45 kDa para GumK pudo ser observada en extractos celulares de cultivos de *E. coli* BL21(DE3)/pET22HisKC inducidos con 0,5 mM IPTG (Fig. 18A calle 1) a una temperatura de 37°C. GumK inducida en estas condiciones precipitaba mayormente como CI, sin actividad enzimática, y por ende no aptos para estudios bioquímicos o estructurales (Fig. 18A calle 2). Así, como primer paso necesitábamos inducir la expresión de la proteína de manera tal que no se produjeran CI. Para ello hubo que buscar condiciones más suaves de inducción. Se probaron combinaciones de diferentes condiciones, variando concentración de inductor, tiempo y sobre todo temperatura de inducción para obtener la mayor cantidad de proteína activa. La combinación óptima fue: inducción a una  $\text{DO}_{600}$  de 0,7, 0,5 mM IPTG, 15-18 hs a 21°C y 200 rpm en agitador rotatorio. Bajo estas condiciones obtuvimos la mayor proporción de proteína activa (Fig. 18B).



**Fig 18:** Sobreexpresión de GumK a 37°C (A) y a 18°C (B): Calle 1: extracto crudo del cultivo de *E. coli* BL21(DE3)/pET22HisKC . Calle 2: precipitado de 5.000 xg (fracción en la que precipitan los cuerpos de inclusión). Calle 3: precipitado de 100.000 xg (fracción de membranas. Calle 4 sobrenadante de 100.000 xg (fracción soluble).

En segundo lugar necesitábamos solubilizar la enzima de la membrana bacteriana para realizar ensayos bioquímicos y estudios estructurales. Ya habíamos observado que GumK era una proteína localizada en membranas. Como la naturaleza exacta de esta interacción no estaba en claro, iniciamos una serie de tratamientos para solubilizar a GumK de las membranas. Una de las posibilidades más factibles era que la interacción fuera del tipo hidrofóbica y/o electrostática. Los primeros tratamientos que probamos para solubilizar GumK fueron de extracción con alta fuerza iónica o con pHs extremos. En la Fig. 19 observamos que los tratamientos con carbonato de sodio, buffer citrato o EDTA no tienen efecto y no solubilizan a la enzima. Por otro lado, se observa que el tratamiento con NaCl produce la solubilización de proteína. El resultado se observa en la Fig. 19A donde se puede apreciar que una fracción del 20-30% (calculado por densitometría) de proteína parece haberse solubilizado, pero la mayor parte se mantiene en membrana. De todos modos, el mayor problema se observó a las pocas horas de haber obtenido esta fracción solubilizada, donde se constató la total precipitación de la enzima (datos no mostrados).

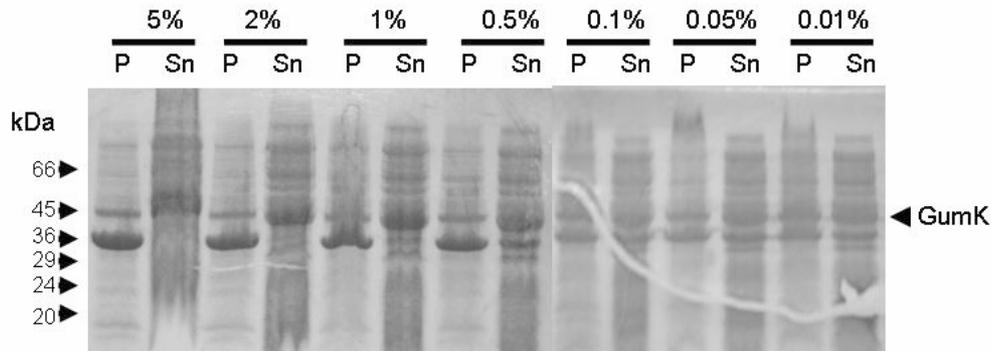


**Fig 19:** Solubilización de GumK por tratamiento con fuerza iónica o pHs extremos: A) 1M NaCl B) 10 mM EDTA pH 8,0 C) 100 mM Buffer citrato pH 3,4 D) 100 mM Carbonato de Sodio pH 11,4. P y Sn se refiere a precipitado y sobrenadante de 100.000 xg (fracciones de membrana y soluble respectivamente). M es la fracción de membranas de *E. coli* BL21(DE3)/pET22HisKC conteniendo GumK.

La insistencia con usar sales en la solubilización tenía una razón: en caso de ser necesario recurrir a técnicas espectroscópicas con GumK, las sales pueden ser eliminadas, intercambiadas por otras y en general no interfieren a la hora de hacer mediciones. En cambio los detergentes en general tienen señal propia (en mediciones

de dicroísmo circular, al UV o fluorescencia) y esto hace que estos estudios sean más difíciles de realizar.

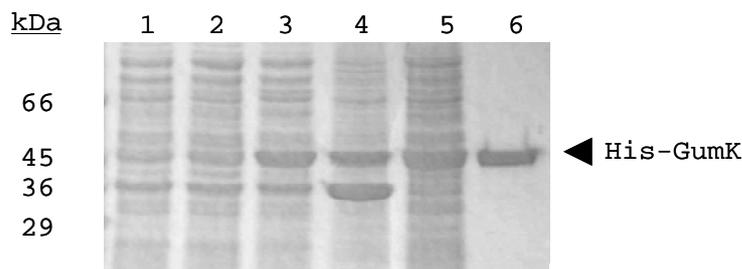
De todos modos, la extracción con sales no había sido satisfactoria, específicamente porque no era capaz de mantener a la proteína en solución. Por ende utilizamos detergentes no iónicos como Tritón X-100 o Tween 20 para extraer y solubilizar efectivamente a GumK de membranas. La Fig. 20 muestra que las condiciones finales de extracción fueron: 1% Tritón X-100 en agitación suave a 4°C. Bajo estas condiciones se recuperó alrededor de un 60% de proteína soluble y activa.



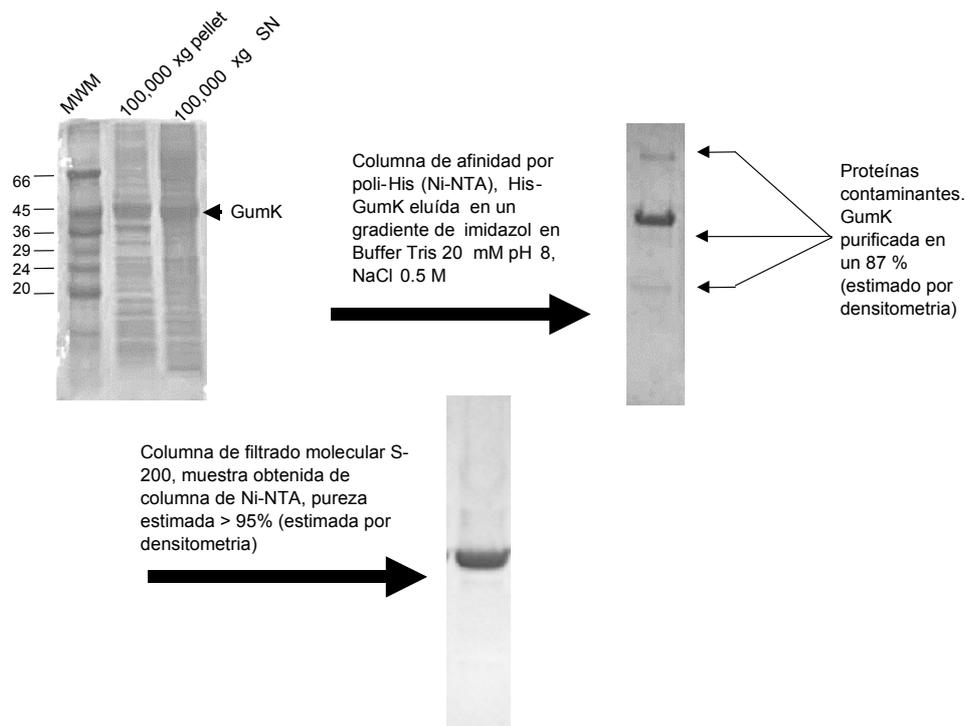
**Fig 20:** Solubilización de GumK a partir de extractos crudos de *E. coli* BL21(DE3)/pET22HisKC. Se utilizó una concentración de proteína total de 10 mg/ml. Se muestra el precipitado (P) y sobrenadante (Sn) de la ultracentrifugación a 100.000 xg para tratamientos con diferentes concentraciones de Tritón X-100.

### 7.3.2 Purificación

Las células de *E. coli* BL21(DE3)/pET22HisKC fueron cultivadas e inducidas como se describe en la sección anterior. La proteína fue purificada por interacción con una matriz de Ni-NTA a través de su cola de poli-His y eluída con un gradiente de 20 a 400 mM de imidazol. La proteína así obtenida fue concentrada y se pudo observar que eluía junto con algunos contaminantes (Fig. 22). Estos contaminantes fueron eliminados por filtración en columna de exclusión por volumen Supedex S-200. La pureza final fue mayor al 95% (Fig. 21 calle 6 y Fig. 22). En teoría esta pureza es suficiente para comenzar ensayos de cristalización para el estudio de la estructura de esta GT. El rendimiento final fue de 20 mg/ml.



**Fig. 21:** Expresión de GumK. La calle 1 muestra las proteínas solubles de la cepa *E. coli* BL21(DE3)/pET22b luego de inducción con IPTG (control negativo). Calle 2: Proteínas solubles de *E. coli* BL21(DE3)/pET22HisKC no inducido (control negativo). La calle 3 muestra la fracción de proteínas totales en *E. coli* BL21(DE3)/pET22HisKC luego de inducción según las condiciones ya descritas. Calles 4 y 5: precipitado y sobrenadante de 100.000 xg luego de fragmentación con prensa de French de *E. coli* BL21(DE3)/pET22HisKC y tratamiento de proteínas totales con 1% Tritón X-100. La calle 6 muestra a GumK purificada. La flecha indica la posición de His-GumK con un peso molecular de ~45 kDa.



**Fig. 22:** Purificación de GumK-HisTag. Se muestran las fracciones separadas por ultracentrifugación de 100.000 xg (precipitado y sobrenadante). GumK es purificada a partir del sobrenadante de 100.000 xg que contiene la proteína soluble y activa.

## 7.4 Análisis de la actividad enzimática de GumK

### 7.4.1 GumK es una glucuronosiltransferasa

Hasta ahora teníamos pruebas genéticas de la actividad enzimática de GumK. A fin de analizar mas directamente esta actividad, desarrollamos un protocolo de incubaciones *in vitro* con la proteína purificada y sus putativos sustratos (75,77). Como sustrato hemos utilizado un extracto glicolipídico de *X. campestris* enriquecido en el aceptor. Células de *X. campestris* permeabilizadas con EDTA incubadas con UDP-Glc y GDP-Man como azúcares donantes acumulan principalmente Poliprenil-PP-Cel-Man en el extracto orgánico. GumK fue incubada con UDP-[<sup>14</sup>C]GlcA en presencia o ausencia del glicolípido aceptor. Seguidamente se determinó la incorporación de [<sup>14</sup>C]GlcA en la fase orgánica. No se observo incorporación de radioactividad en la fase orgánica al realizarse incubaciones en ausencia de enzima (controles negativos). Por el contrario se observó la incorporación de radioactividad en la fase orgánica en incubaciones conteniendo GumK y el aceptor glicolipídico. Esto ratifica que GumK es una glucuronosiltransferasa. Los oligosacáridos fueron separados de la fracción lipídica solubilizandose en la fracción acuosa mediante hidrólisis ácida suave usando 0,01 N HCl, indicando que los azúcares están unidos a la fracción lipídica a través de un puente pirofosfato. Este comportamiento es similar al del autentico Poliprenil-PP-Cel-Man-GlcA producido por *X. campestris*. Los oligosacáridos liberados, marcados radiactivamente fueron analizados por TLC. Se detectó un compuesto radioactivo que comigra con autentico Cel-Man-GlcA (Fig. 23, calle 1), indicando que el oligosacárido era un tetrasacárido conteniendo GlcA.

Para confirmar que Poliprenil-PP-Cel-Man era el aceptor del residuo de GlcA se preparó Poliprenil-PP-Cel-[<sup>14</sup>C]Man utilizando células permeabilizadas de *X. campestris*. El aceptor marcado se incubó en presencia de UDP-GlcA y GumK. El material radioactivo fue tratado con hidrólisis ácida suave y analizado por TLC. En incubaciones sin enzima se observó la presencia de Cel-Man no modificada (Fig. 23,

calle 3). Por el contrario, un tetrasacárido que migra junto con Cel-Man-GlcA fue observado cuando se realizó una incubación con GumK (Fig. 23, calle 2). Estos resultados confirman que GumK cataliza la transferencia de un residuo de ácido glucurónico de UDP-GlcA a Poliprenil-PP-Cel-Man dando como producto final Poliprenil-PP-Cel-Man-GlcA.

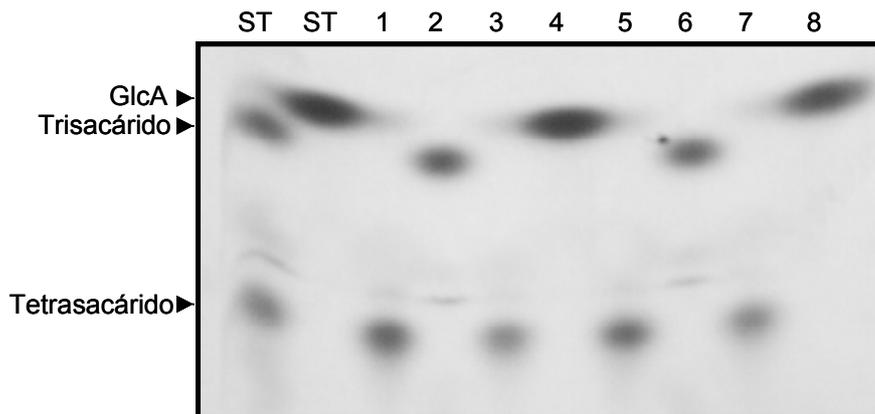


**Fig. 23:** Análisis por TLC de los oligosacáridos liberados luego de los ensayos de actividad glucuronosiltransferasa. Las incubaciones se realizaron en presencia de GumK, Poliprenil-P-P-Cel-Man y UDP<sup>[14C]</sup>GlcA (calle 1); GumK Poliprenil-PP-Cel-<sup>[14C]</sup>Man y UDP-GlcA (calle 2) o incubaciones en ausencia de enzima, Poliprenil-PP-Cel-<sup>[14C]</sup>Man y UDP-GlcA (calle 3). Trisacárido, tetrasacárido y pentasacárido se refieren a Cel-<sup>[14C]</sup>Man, Cel-<sup>[14C]</sup>Man-GlcA y Cel-<sup>[14C]</sup>Man-GlcA-Man usados como estándares (ST).

Finalmente, tampoco se observó actividad enzimática cuando se realizaron incubaciones conteniendo GumK-s (proteína de 295 residuos), el aceptor glicolípido y UDP-<sup>[14C]</sup>GlcA, indicando que es la proteína GumK (de 400 residuos) la que lleva a cabo la actividad glucuronosiltransferasa.

#### 7.4.2 Caracterización del enlace glucuronosil-glicosídico

Se utilizó β-glucuronidasa de hígado bovino para determinar la estereoquímica del enlace Man-GlcA producido por GumK. El tetrasacárido (Cel-Man-GlcA) obtenido marcado en <sup>[14C]</sup>GlcA o en <sup>[14C]</sup>Man fue digerido con β-glucuronidasa. Como se muestra en la Fig. 24, calle 2, la digestión del tetrasacárido marcado con <sup>[14C]</sup>Man produjo un trisacárido <sup>[14C]</sup>Man que migra junto con Cel-Man, mientras que el tetrasacárido marcado con <sup>[14C]</sup>GlcA libera <sup>[14C]</sup>GlcA libre (calle 4). El auténtico Poliprenil-PP-Cel-Man-GlcA obtenido en incubaciones con células permeabilizadas es también sensible a la β-glucuronidasa (calles 6 y 8). Estos resultados son consistentes con un enlace β-glucuronil para la unión Man-GlcA. Estos resultados indican que la estereoquímica del enlace formado es β, siendo GumK una β-glucuronosiltransferasa.

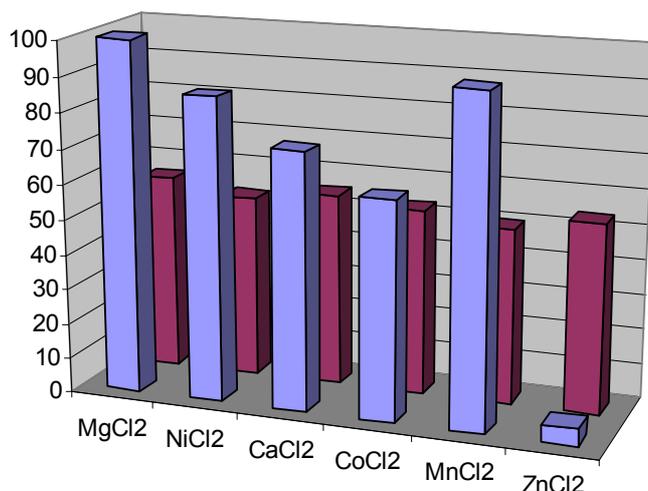


**Fig 24:** Análisis por TLC de los oligosacáridos liberados luego del tratamiento con  $\beta$ -glucuronidasa. El Trisacárido (Cel-[ $^{14}$ C]Man), tetrasacárido (Cel-[ $^{14}$ C]Man-GlcA) y [ $^{14}$ C]-GlcA (calle ST) se utilizaron como estándares. Se muestra el tetrasacárido (Cel-[ $^{14}$ C] Man-GlcA) obtenido en una incubación con GumK antes (calle 1) y después (calle 2) del tratamiento con  $\beta$ -glucuronidasa. Se muestra el tetrasacárido (Cel-Man-[ $^{14}$ C] GlcA) obtenido en una incubación con GumK antes (calle 3) y después (calle 4) del tratamiento con  $\beta$ -glucuronidasa. Tetrasacárido (Cel-[ $^{14}$ C]Man-GlcA) obtenido en una incubación usando células de la cepa salvaje FC2 permeabilizadas con EDTA, antes (calle 5) y después (calle 6) del tratamiento con  $\beta$ -glucuronidasa. Tetrasacárido (Cel-Man-[ $^{14}$ C]GlcA) obtenido en una incubación utilizando células de la cepa FC2 permeabilizadas con EDTA, antes (calle 7) y después (calle 8) del tratamiento con  $\beta$ -glucuronidasa.

#### 7.4.3 Utilización de cationes divalentes

En las incubaciones *in vitro* que realizamos con GumK utilizamos cationes divalentes, específicamente 5 mM  $MgCl_2$ . GumK es una GT-B, y por ende no necesita de la estricta presencia de cationes divalentes en la mezcla de reacción, si bien la actividad de esta enzima esta afectada por la presencia de éstos. Por esta razón decidimos estudiar el comportamiento de GumK en presencia de otros cationes además del  $Mg^{++}$  o bien en ausencia de éstos. Las incubaciones se realizaron como se indica en MyM, sección 6.10.3. Los resultados se muestran en la Fig. 25 y muestran que el catión óptimo para la función de GumK es el  $Mg^{++}$ , si bien con otros cationes divalentes como el Ni, Ca, Co o Mn tiene una actividad que ronda el 60-90% de lo observado con  $Mg^{++}$ . Interesantemente, en incubaciones en ausencia de cationes (y con 10 mM EDTA como agente quelante), la actividad se reduce aproximadamente un 50% en todos los casos, pudiendo considerarse a esta actividad como el valor "basal". Esta no estricta dependencia de la presencia de cationes divalentes es una evidencia bioquímica directa de que GumK es una GT-B, y era esperable este comportamiento, también observado en otras GT-B. También se observa que el  $Zn^{++}$  anula la actividad de la enzima por razones que desconocemos. De estos resultados se desprende que bioquímicamente, GumK tiene un comportamiento esperable para una GT-B, con actividad aumentada en presencia de divalentes (especialmente para  $Mg^{++}$  y  $Mn^{++}$ ), pero reteniendo un porcentaje de esta actividad aun en ausencia de cationes. Coincidentemente, en la estructura de GumK no hemos encontrado un ión  $Mg^{++}$  que ayude a coordinar los fosfatos del dador o alguna indicación de otra función como cofactor en la catálisis, por lo que no podemos adjudicarle a estos iones ninguna función evidente.

## Porcentaje de reacción



**Fig 25:** Porcentaje de actividad respecto a la presencia del catión  $Mg^{++}$ . Se muestran los porcentajes para diferentes cationes utilizados en las incubaciones *in vitro* (en azul). También se muestran los porcentajes en presencia de EDTA (en rojo).

## 7.5 Especificidad de aceptor para GumK

### 7.5.1 Cel-Man

En primer lugar, analizamos si GumK es capaz de lograr la transferencia de un residuo de ácido glucurónico en un aceptor trisacárido sin lípido. Cel-[<sup>14</sup>C]Man se obtuvo por tratamiento con hidrólisis ácida suave y fue incubada con GumK y analizada por TLC. La Fig. 26A presenta evidencia de que el puente pirofosfato y/o la fracción lipídica son indispensables para el reconocimiento del sustrato y/o la actividad catalítica, pues la enzima no pudo transformar este trisacárido en el producto tetrasacárido.

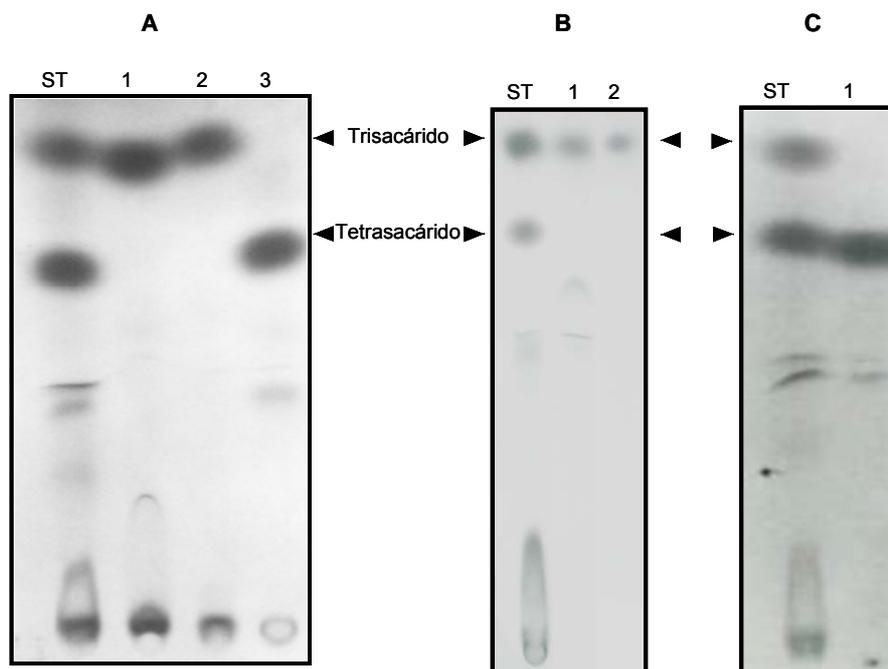
### 7.5.2 P<sub>i</sub>-Cel-Man y PP<sub>i</sub>-Cel-Man

En la Fig. 26B observamos que la porción sacarídica del aceptor no era reconocida por la enzima. Decidimos entonces hacer un análisis más pormenorizado, seccionando la molécula aceptora para entender la importancia de cada uno de sus componentes. Por esto decidimos obtener P<sub>i</sub>-Cel-Man y PP<sub>i</sub>-Cel-Man para probar si GumK es capaz de transferir GlcA al trisacárido fosforilado. La obtención de este sustrato truncado se basa en un protocolo reportado anteriormente (MyM sección 6.10.2). Ambas formas del sustrato fosforilado fueron incubadas en presencia de GumK y UDP-GlcA. Luego de la incubación las muestras fueron analizadas por TLC. La Fig. 26B muestra que en ninguno de los dos casos hubo formación de tetrasacárido como si ocurría en el caso del aceptor con la porción lipídica intacta. Se desprende de estos resultados que el pirofosfato y el trisacárido del aceptor no son suficientes para el reconocimiento por parte de la enzima.

### 7.5.3 Fitanyl-PP-Cel-Man

De los resultados anteriores surge una pregunta: si el aceptor es un poliprenilo de cadena insaturada, y la porción lipídica es imprescindible para el reconocimiento del aceptor, podría por ejemplo un fitanilo de cadena saturada suplantarlos de manera satisfactoria? Para elucidar la contribución de la fracción lipídica en el proceso de reconocimiento, y basados en reportes que indicaban que el fitanol puede reemplazar la porción lipídica de la molécula en otras glicosiltransferasas (88), estudiamos la capacidad de GumK soluble de transferir un residuo de ácido glucurónico a Fitanyl-PP-Cel-Man. El aceptor Fitanyl-PP-Cel-Man fue preparado a partir de la transferencia de una manosa desde GDP-Man a Fitanyl-PP-Cel, catalizado por la enzima AceA de *Acetobacter xylinum*. La Fig. 26C muestra que el compuesto con fitanol puede aceptar

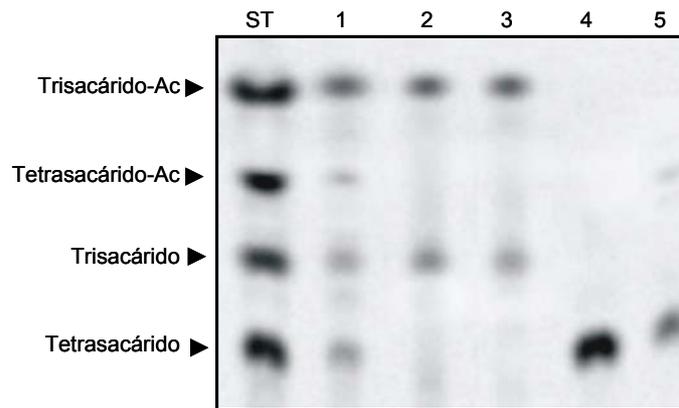
un residuo de ácido glucurónico. Esto indica que la longitud o el grado de saturación de la cadena hidrocarbonada no es un requisito tan estricto.



**Fig. 26:** Análisis por TLC de los oligosacáridos liberados luego de incubación con aceptor delipidificado (A)  $P_i$ -Cel-Man y  $PP_i$ -Cel-Man (B) o fitanil-PP-Cel-Man (C). Las incubaciones se realizaron en presencia de la enzima GumK, Cel- $[^{14}C]$ Man y UDP-GlcA (A, calle 1), Incubación sin enzima y con Poliprenil-PP-Cel- $[^{14}C]$ Man y UDP-GlcA (A, calle 2, control negativo), y GumK, Poliprenil-PP-Cel- $[^{14}C]$ Man y UDP-GlcA (A, calle 3, control positivo); (B) Incubación en presencia de GumK,  $P_i$ -Cel- $[^{14}C]$ Man y UDP-GlcA (B, calle 1), GumK,  $[^{14}C]$ Man-Cel-PPi y UDP-GlcA (B, calle 2). (C) La incubación se realizó en presencia GumK, Fitanil-PP-Cel-Man y UDP- $[^{14}C]$ GlcA (C, calle 1). Trisacárido y tetrasacárido se refieren a Cel- $[^{14}C]$ Man y Cel- $[^{14}C]$ Man-GlcA usados como estándares (ST).

#### 7.5.4 Acetilación natural del sustrato aceptor

Las manosas internas y externas del xantano están sustituidas con proporciones variables de residuos O-acetilo (71). Previamente se ha descrito que Ac-CoA es el donante y que el Poliprenil-PP-Cel-Man así como el Poliprenil-PP-Cel-Man-GlcA-Man funcionan como aceptores formando Poliprenil-PP-Cel-(Ac)Man y Poliprenil-PP-Cel-(Ac)Man-GlcA-Man. Estas reacciones son catalizadas por la acetiltransferasa I, codificada por el gen *gumF* (79). Investigamos la influencia del residuo acetilo en la especificidad de GumK. Cuando el Poliprenil-PP-Cel-(Ac)Man- $[^{14}C]$  o Poliprenil-PP-Cel-(Ac)Man fueron incubados con GumK en presencia de UDP-GlcA marcado o no radioactivamente, se observó muy baja actividad de transferencia del residuo de GlcA, de solamente el 10% de la cantidad total de sustrato (Fig. 27 calles 1 y 5). Esto estaría indicando que la presencia del grupo acetilo bloquea la reacción y entonces el reconocimiento del aceptor se ha alterado por parte de GumK. Podemos decir entonces que el sustrato acetilado no es un buen aceptor para GumK, pues su transformación en producto final es muy baja (si bien no es nula) comparada con el sustrato no acetilado.



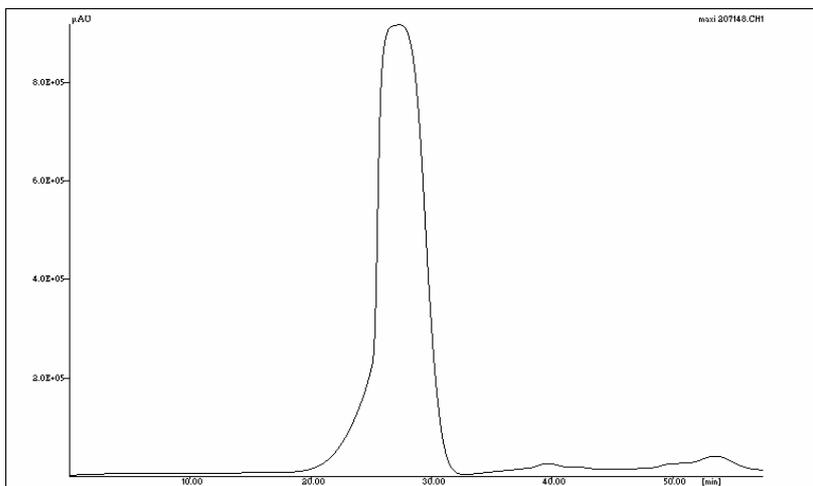
**Fig. 27:** Efecto de la O-acetilación del sustrato sobre la actividad de GumK. El aceptor acetilado Poliprenil-PP-Cel-(Ac)Man fue sintetizado como se describe en MyM. La calle 1 muestra los productos de reacción resultantes de incubar el sustrato acetilado Poliprenil-PP-Cel-(Ac)Man<sup>[14C]</sup> mezclado con el sustrato radiactivo sin acetilar (sustrato mixto). La calle 2 muestra las posiciones correspondientes a esta mezcla de sustrato y la calle 3 muestra al sustrato incubado en ausencia de GumK como control negativo. Calle 4: Incubación con sustrato marcado sin acetilar, GumK y UDP-GlcA como control positivo y calle 5: sustrato acetilado mixto no marcado radiactivamente incubado en presencia de UDP-[<sup>14</sup>C]GlcA, y GumK. Trisacárido, trisacárido-Ac, tetrasacárido y tetrasacárido-Ac se refieren a Cel- [<sup>14</sup>C] Man, Cel- [<sup>14</sup>C] Man-Ac, Cel- [<sup>14</sup>C] Man-GlcA, Cel- [<sup>14</sup>C] Man-Ac-GlcA usados como estándares (calle ST).

### 7.6 Ensayos de actividad con diferentes dadores

Como sustrato dador se probaron diferentes nt-az: UDP-[<sup>14</sup>C]Glc, GDP-[<sup>14</sup>C]Man, UDP-[<sup>14</sup>C]Gal, UDP-[<sup>14</sup>C]Xyl, UDP-Rha, UDP-GlcNac, UDP-GalNac y UDP-GalA. Las reacciones se llevaron a cabo como se describe en MyM sección 6.10.3. En ningún caso se observó actividad detectable. Esto implica que GumK posee una estricta especificidad por su aceptor, un punto que será desarrollado en la Discusión.

### 7.7 Estado oligomérico de GumK

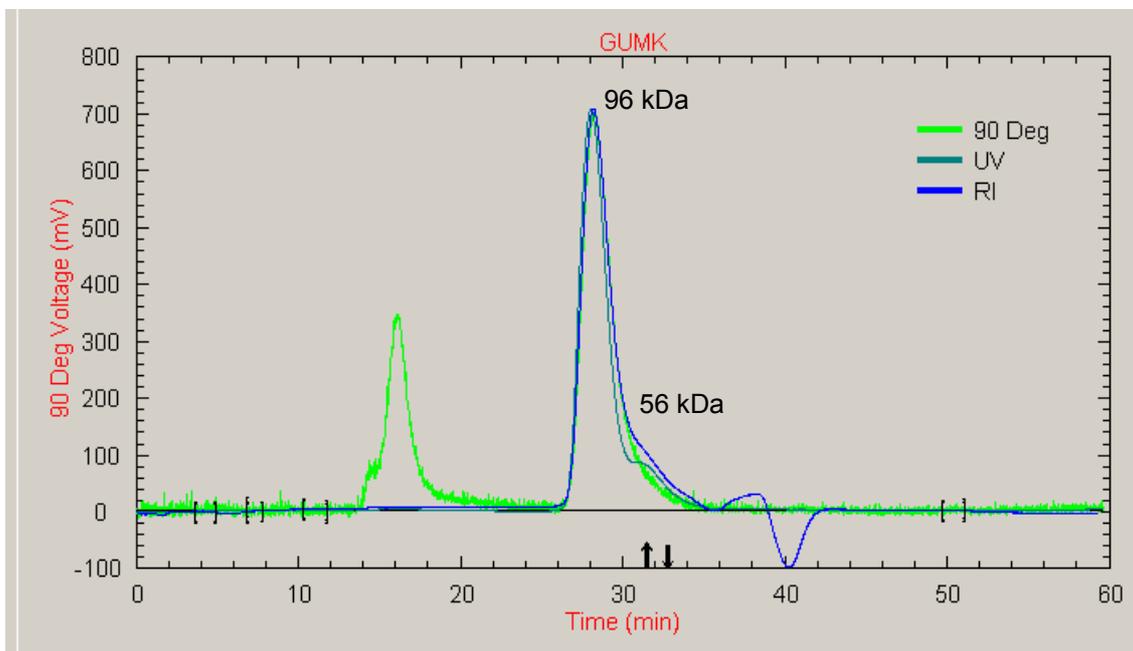
Durante la purificación de GumK, existía un fenómeno que generaba controversia y era el perfil de elución de la proteína en cromatografía líquida de filtración en gel (Fig. 28).



**Fig. 28:** Perfil de elución de GumK en una columna de filtración por gel Superdex 200 para un rango de 10-600 kDa. La proteína eluye en un volumen de aproximadamente 12 ml (24 min en el grafico a un flujo de 0,5 ml/min), y correspondiente a un peso molecular de unos 250-300 kDa si tomamos como referencia estándares de peso de proteínas solubles.

Por lo que se observa en el cromatograma, y tomando en cuenta una curva de estándares de peso de proteínas solubles, GumK estaría eluyendo con un peso aproximado de 250-300 kDa. Esto evidentemente se correspondería con un multímero,

cuya valencia no es clara, sobre todo si tomamos en cuenta que la presencia de detergente en la solución bien podría estar afectando el radio hidrodinámico de la proteína. Entonces, ¿cuál es el estado oligomérico de esta proteína activa en solución? Para estudiar este aspecto se utilizaron ensayos de dispersión de luz. La proteína purificada fue analizada por IR, UV y dispersión de luz estática. Este análisis arrojó como resultado una partícula mayoritaria de 96 kDa (>95%), correspondiendo a un dímero en solución (Fig. 29). Este resultado nos estaría indicando entonces que la proteína funcional sería un dímero. Perfiles de elución que se corresponden a partículas de alto peso molecular, generalmente mayores de 200 kDa son observados tanto en GumK como otras GTs de membrana (97). Esto abre el interrogante de cuál sería el significado biológico de tener un dímero en solución, pues hay que tener en cuenta que las hipótesis acerca de cómo GumK se une a la membrana proponen la existencia de un monómero y no un dímero. A esto se agrega que la unidad asimétrica en la estructura cristalina es un monómero. Este punto se desarrollará más extensamente en la Discusión.



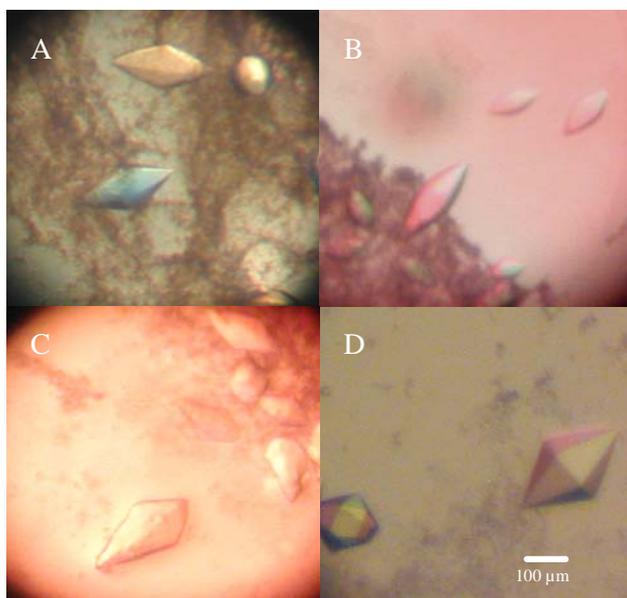
**Fig. 29:** Diagrama de elución de GumK en columnas de filtración por peso molecular (Superdex 200, GE healthcare). Se observa un pico mayoritario que corresponde a un dímero (aproximadamente 90 kDa, 95% de la proteína presente en la muestra), y un pico menor (56 kDa) que podría corresponder a un monómero.

### 7.8 Cristalización de GumK

GumK fue purificada a un grado mayor al 95% y concentrada hasta 20 mg/ml en buffer 0,4 M NaCl 50 mM Tris-HCl 0,05% Tritón X-100 pH 8,0. Con esta proteína se determinaron las condiciones iniciales de cristalización utilizando el Crystal Screen I Kit, (Hampton, Inc). En este experimento se hallaron 3 condiciones de cristalización, con formación de cristales de aspecto romboidal y bordes poco definidos: 1) 0,2 M  $\text{Li}_2\text{SO}_4$ , 0,1 M Tris HCl pH 8,5, 30% PEG 4.000. 2) 0,2 M  $\text{MgCl}_2$ , 0,1 M Tris HCl pH 8,5, 30% PEG 4.000 y 3) 0,2 M Acetato de Sodio, 0,1 M Tris HCl pH 8,5, 30% PEG 4.000 (Fig. 30A, B, C). Para optimizar estas condiciones iniciales se variaron levemente las condiciones de pH, concentración de precipitantes y relación de volumen entre solución de proteína concentrada y buffer de cristalización en ensayos de gota colgante. Este proceso culminó con 3 condiciones diferentes de cristalización: 1) 0,2 M  $\text{Li}_2\text{SO}_4$ , 35% PEG 3350, 0,1 M ClCs, 0,1 M Tris HCl pH 8,2. 2) 0,2 M  $\text{MgCl}_2$ , 30% PEG

## Resultados

3350, 0,1 M ClCs, 0,1 M Tris HCl pH 8,0. 3) 0,2 M Acetato de Sodio, 30% PEG 3350, 0,1 M ClCs, 0,1 M Tris HCl pH 8,5, todos en una relación 3:1 v/v de proteína/buffer de cristalización, con formación de cristales en aproximadamente 2-3 semanas. La condición de cristalización con 0,2 M  $\text{Li}_2\text{SO}_4$ , 35% PEG 3350, 0,1 M ClCs, 0,1 M Tris HCl pH 8,2 fue la que logró los cristales mas definidos y de mayor tamaño, con un tamaño máximo de 0,4 x 0,2 mm (Fig. 30D). Estos últimos cristales polarizaban la luz fuertemente, indicando ordenamiento molecular.



**Fig. 30:** Cristales de GumK. Se observa claramente su forma bipiramidal con bordes definidos. El tamaño promedio es de 0,4 x 0,2 mm.

### 7.9 Estructura cristalográfica de la glucuronosiltransferasa GumK

#### 7.9.1 Caracterización cristalográfica preliminar de GumK nativo

Para los cristales de GumK obtuvimos datos de difracción a 6 Å de resolución en un difractor de ánodo rotatorio disponible en nuestro instituto (ánodo rotatorio de Cu Siemens M18X-02-014,  $\lambda = 1,54$  Å. Detector de Área Multiwire Siemens X1000). Este resultado nos alentó a buscar una fuente de rayos-X de alta energía como es el caso de un difractor de mayor potencia o una línea de rayos-X en un sincrotrón para obtener datos de mayor resolución. Con este fin se estableció una colaboración con el Dr. Mario Amzel del Department of Biophysics & Biophysical Chemistry, Johns Hopkins University, Baltimore, Maryland, USA En el laboratorio del Dr. Amzel purifiqué y cristalicé GumK con condiciones optimizadas. Una vez obtenidos los cristales se colectaron datos en difractores de ánodo rotatorio con detectores de área del tipo R-axis (Rigaku Corp.). Los datos obtenidos para el cristal nativo tenían una resolución de 2,7 Å. La resolución y parámetros estadísticos en el procesamiento de los datos fueron mejorados notablemente colectando sets de datos en una línea de alta potencia, específicamente en el sincrotrón del NSLS, BNL, NY, EEUU (Fig. 31A y B). Las colecciones de datos se realizaron a 110 K en la línea X12C a una longitud de onda de 1,100 Å. La solución de cristalización 0,2 M  $\text{Li}_2\text{SO}_4$ , 35% PEG 3350, 0,1 M Tris HCl, 0,1 M ClCs funcionó exitosamente como crioprotector, por lo que se utilizaron cristales obtenidos en esta solución para obtener datos de difracción. La línea X12C esta equipada con un detector ADSC Q210. Los cristales difractaron a 1,9 Å. Los datos fueron procesados usando los programas MOSFLM, SCALA y TRUNCATE (Tabla 3) del paquete CCP4(91). Los cristales mostraron tener una simetría  $P6_522$  con un volumen de celda de 2.307.219,25 Å<sup>3</sup>. Hay una molécula de GumK por unidad

asimétrica y el contenido de solvente asciende a un 58.8%. En los mapas de Patterson nativos no se observaron picos significativos fuera del origen.

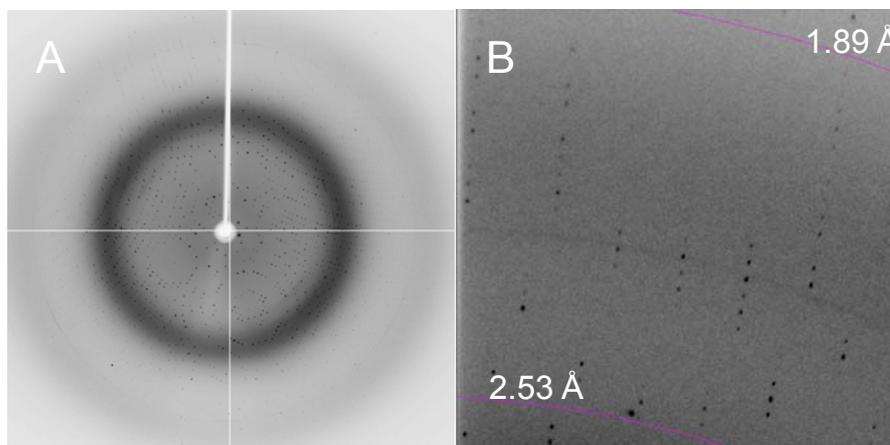
La primera caracterización cristalográfica de GumK se realizó con un set de datos de 240 imágenes de difracción de 0,5° de oscilación, integrando los datos hasta una resolución de 1,9 Å. Los primeros intentos para resolver la estructura por reemplazo molecular usando modelos como la glucógeno sintasa (PDB 1rzu) o la N-acetilglucosaminiltransferasa MurG (PDB 1f0k) en el programa Amore no fueron exitosos, principalmente debido a la baja homología de secuencia existente entre estas enzimas y GumK.

**Tabla 3:** Estadísticas de colección de datos de difracción de rayos-X para GumK nativo

Grupo espacial	P6 <sub>5</sub> 22
Sistema cristalino	Primitivo Hexagonal
Parámetros de la celad unidad (Å, °)	a=123,63, b=123,63, c=174,30 α=90, β=90, γ=120
Rango de resolución (Å)	35,5-1,9 (2,0-1,9)
Reflexiones únicas	62.401
Rsym (%) <sup>§</sup>	6,9 (32,3)
Porcentual de reflexiones medidas (%)	100 (100)
[I/σ(I)]	9,9 (2,2)
Mosaicidad(°)	0,18
Observaciones totales	879.919
Multiplicidad	14,1

§  $R_{sym} = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]|}{\sum_{hkl} \sum_i I_{hkl,i}}$  donde  $[I_{hkl}]$  es la media de la intensidad de las múltiples observaciones  $I_{hkl,i}$  en el caso de reflexiones relacionadas por simetría.

Números en paréntesis corresponden a la capa de mayor resolución.



**Fig. 31:** (A) Patrón de difracción de GumK colectado en la línea X12C del sincrotrón NSLS-BNL. (B) Detalle de (A). El ángulo de oscilación es de 0,5°. Se observa difracción hasta 1,4 Å.

### 7.9.2 Búsqueda experimental de la información de fases

#### 7.9.2.i Reemplazo Múltiple Isomorfo (MIR)

##### 7.9.2.i.a Obtención de derivados

Los intentos de realizar reemplazos moleculares con modelos de baja homología de secuencia no resultaron exitosos. Paralelamente a esto, buscamos obtener derivados de la proteína con metales de átomos pesados. Con este fin, realizamos experimentos

de inmersión de los cristales en solución de cristalización suplementada con átomos pesados para obtener conjugados que, al ser expuestos a los rayos-X, cambien las intensidades de las difracciones en puntos específicos, como primera medida para obtener la información del ángulo de fase de las difracciones. Esta información es esencial para la determinación de la estructura molecular de la proteína. Se obtuvieron cristales derivados isomorfos por inmersión de cristales nativos en buffer de cristalización con el agregado de diferentes sales de átomos pesados (normalmente sales de Hg, Pt, Au, Pb, Br, Mn, Ta, Gd, Sm, Re, Tb, Ag, entre otros). Para el caso particular de los experimentos de MIR, se obtuvieron sets de datos de cristales derivados isomorfos por inmersión en soluciones de átomos pesados en buffer de cristalización. Los compuestos utilizados en inmersiones que produjeron cristales que conservaban señal de difracción (en general los cristales adquirirían un aspecto más deteriorado, resultando en una muy pobre o nula difracción) fueron: 0,1 mM Acetato de Gadolinio, 6 h de inmersión; 50 mM Bromuro de Sodio 10 min y 0,1 mM HgCl<sub>2</sub> 2 h y 1mM K<sub>2</sub>PtCl<sub>4</sub> 6 h. Las resoluciones de estos sets de datos son de 2,7, 3,0, 3,1 y 2,8 Å respectivamente. La integración y reducción de los datos se realizó como se indica en MyM sección 6.16.1. Para la localización de átomos pesados y cálculo de fases se utilizaron los programas SHELX-97, Patterson search y MLPHARE (91). Los parámetros de celda de estos derivados son esencialmente idénticos a los del cristal de la enzima nativa. Parámetros como el Phasing power y la Diferencia isomorfa media de los derivados versus el cristal nativo se muestran en la Tabla 4. Estos datos son una medida de la información de fases en la difracción de los derivados y como se observa muestra que en éstos la información de fases era escasa. Principalmente el Phasing power era muy bajo, suponemos que esencialmente debido a una señal anómala también muy baja, producto de sets de datos que mostraban un R<sub>merge</sub> muy alto (generalmente más de 10%), lo cual dificultaba la identificación de la auténtica información de fases por sobre el “ruido”, o error asociado a la colecta de datos del set de datos.

**Tabla 4:** Estadísticas del cálculo de fases por Reemplazo isomorfo múltiple. Los valores en paréntesis corresponden a la capa más alta de resolución.

Análisis MIR-derivados	HgCl <sub>2</sub>	GdCl <sub>3</sub>	BrNa	K <sub>2</sub> PtCl <sub>4</sub>
Diferencia media isomorfa	0,104	0,093	0,041	0,152
Phasing power	1,21 (0,68)	1,02 (0,50)	0,73 (0,50)	2,58 (0,92)
R <sub>cullis</sub>	0,89 (0,92)	0,95 (0,98)	0,97 (0,99)	0,98 (0,99)

Diferencia media isomorfa =  $\Sigma[F_{PH} - F_P] / \Sigma F_{PH}$ , donde F<sub>PH</sub> y F<sub>P</sub> son los factores de estructura del derivado y el nativo respectivamente.

El Phasing power es el coeficiente entre la media del factor de estructura del derivado y la media del error de la diferencia vectorial de fases

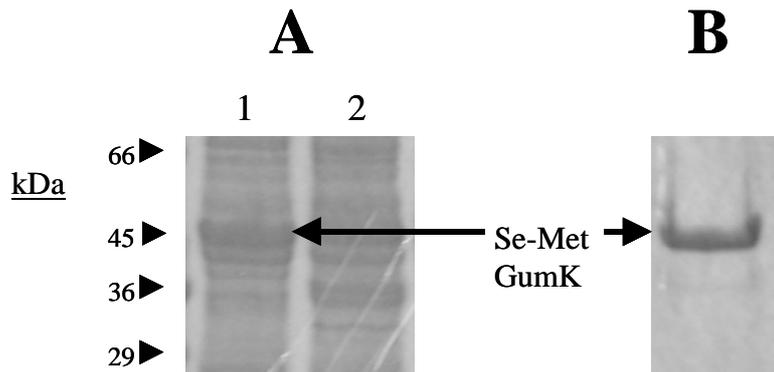
R<sub>cullis</sub> es la media de la diferencia vectorial de fases dividida por la diferencia dispersiva o anómala.

### 7.9.2.ii Dispersión anómala múltiple (MAD)

#### 7.9.2.ii.a Obtención de derivados de seleno-metionina de GumK

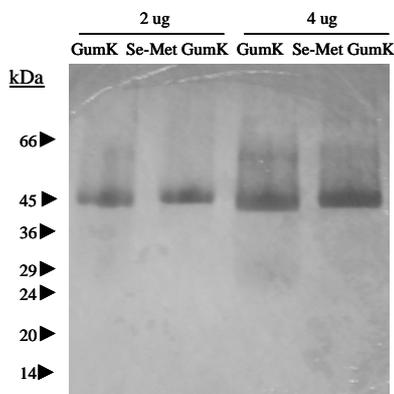
Los experimentos con derivados no se limitaron a MIR. De hecho, estos experimentos se realizaron a raíz de que sólo contábamos con una línea de rayos-X de longitud de onda fija (DO3B-MX1,  $\lambda=1,427$ ). Cuando pudimos contar con una línea de longitud de onda variable (NSLS, X12C), llevamos a cabo experimentos de MAD. Para ello necesitábamos incorporar en la proteína átomos pesados que tuvieran una dispersión atómica anómala a longitudes de onda sintonizables y medibles en un sincrotrón y utilizar esta información para determinar las posiciones atómicas en la red cristalina. El átomo de elección para estos experimentos es el selenio, en reemplazo del azufre de la metionina. Este selenio se incorpora como Selenio-Metionina (Se-Met) en la

proteína de interés. Para obtener Se-Met GumK primero se constató la expresión de GumK en la cepa de *E. coli* auxotrofa para metionina B834(DE3). Se comprobó primero que esta cepa podía expresar GumK en medio LB y una vez constatado esto, se comenzó con la expresión de GumK en medio mínimo. Las pruebas de crecimiento mostraron que el medio ideal para estas cepas era el DLMM (MyM sección 6.5). Se observó máxima expresión de Se-Met GumK a 21°C, 20 h, 0,5 mM IPTG. Se purificó esta proteína de la misma manera que GumK nativa y agregándose a los buffers de lavado y elución TCEP (Tris 2-carboxy etil fosfina) 2 mM como agente reductor, evitando otros agentes con mayor poder reductor como DTT o β-mercaptoetanol (incompatibles con columnas IMAC). La proteína obtenida como el sobrenadante de una ultracentrifugación de 100.000 xg fue utilizada para la purificación (Fig. 32).



**Fig.32:** Purificación de Se-Met GumK. (A) Células de *E. coli* B834(DE3) pETHisKC fueron fragmentadas con prensa de French y ultracentrifugadas a 100.000 xg. Se analizó sobrenadante (1) y precipitado (2) a través de SDS-PAGE en gels 10%. (B) La proteína purificada mostró una pureza superior al 95%.

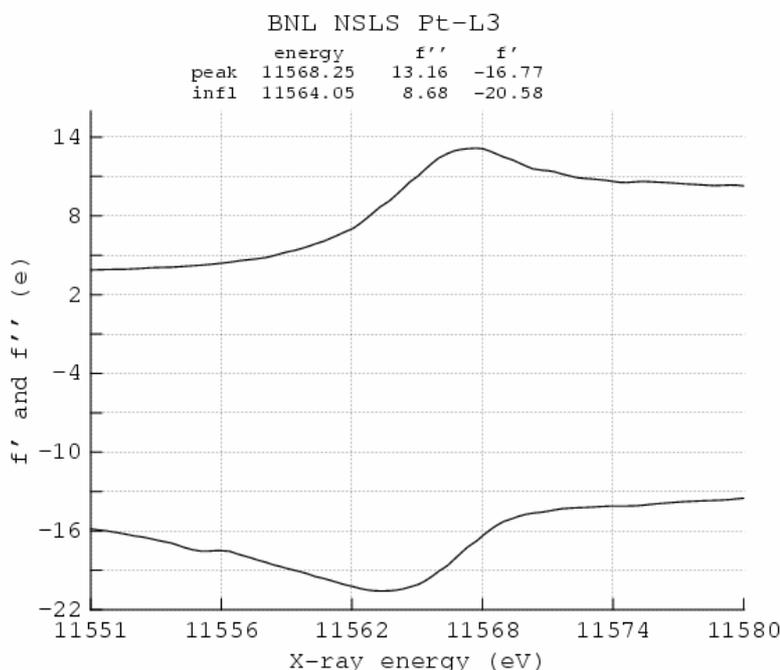
GumK posee 10 metioninas en su secuencia; la incorporación de 10 Se-Met produce una diferencia de masa molecular de 0,7 kDa, observable en un gel 15% de poliacrilamida. Utilizamos esta forma de verificación de la producción de Se-Met GumK y no, por ejemplo, un espectro de masa porque la presencia de detergentes dificulta las mediciones y la interpretación de los espectros. Se observó una diferencia de peso entre la proteína nativa y la Se-Met proteína en una corrida electroforética en SDS-PAGE (Fig. 33). Tras su purificación se establecieron las condiciones de cristalización para Se-Met GumK. Se observó que la proteína sustituida cristaliza bajo condiciones similares a la nativa (solo se aumentó en un 10% la concentración de los precipitantes PEG y Li<sub>2</sub>SO<sub>4</sub>), pero los cristales que se formaban eran de apariencia heterogenea (no mostrado) y no difractaban al ser expuestas a los rayos-X. Intentamos entonces buscar condiciones de cristalización que nos permitieran optimizar estos cristales utilizando kits de aditivos y detergentes para optimizar los cristales. Lamentablemente, no pudimos sortear este obstáculo técnico y no obtuvimos cristales de GumK sustituido con Se-Met para realizar experimentos de MAD.



**Fig. 33:** SDS-PAGE acrilamida 15%. Se compara la movilidad electroforética de GumK vs Se-Met GumK.

## 7.9.2.ii.b MAD con derivados con átomos pesados

Como tercera (y exitosa) aproximación para resolver el problema de las fases se buscó la información de fases a partir de derivados de átomos pesados obtenidos por experimentos de inmersión. La información de fases para la resolución de la estructura de GumK se obtuvo con un experimento de dispersión anómala múltiple (MAD) a partir de un derivado de  $K_2PtCl_4$  (Tabla 5). El derivado se obtuvo por inmersión de cristales en solución de cristalización con el agregado de 10 mM  $K_2PtCl_4$  por 2h. Este cristal difractó satisfactoriamente hasta 2,0 Å. Los grupos de datos para el experimento de MAD fueron colectados a las longitudes de onda  $\lambda=1,0718$  Å y  $\lambda=1,0722$  Å, correspondientes a los niveles energéticos máximo y de inflexión de absorbancia en un espectro de fluorescencia para el platino (Fig. 34). Los datos de difracción fueron colectados en la línea X12C del NSLS, BNL, NY, EEUU, a una temperatura de 110 K. Los datos fueron primeramente procesados usando MOSFLM, SCALA y TRUNCATE del compendio de programas para cristalografía CCP4 (7). El derivado analizado mostró tener una simetría del tipo  $p6_522$ , con parámetros de celda idénticos a los del cristal nativo. El phasing power anómalo, calculado usando el programa SHARP llegaba a 6,0 en las capas de difracción de menor resolución, llegando hasta 1,0 a 2,5 Å. Con estos valores de información de fases fue posible resolver la estructura de la proteína. La combinación de los grupos de datos del cristal nativo y derivados fue hecha con los programas CAD y MLPHARE. Las posiciones de los 4 átomos de platino unidos a la estructura de la proteína fueron resueltas con SHELX y refinadas con SHARP. El porcentaje de solvente y las regiones de menor densidad en el cristal fueron determinadas utilizando los programas SOLOMON y DM de CCP4. El trazado de la cadena peptídica se realizó en primera instancia con el programa ARP/wARP(98) y luego fue refinado a mano utilizando el programa Coot. Las estadísticas finales de  $R_{free}$  y restricciones en el posicionamiento de los residuos y sus cadenas laterales (plot de Ramachadran) y moléculas de solvente en la estructura fueron hechas con Refmac y Procheck. La estructura se encuentra depositada en el banco de estructuras de la RCSB con el número de PDB 2HY7. El modelo final muestra muy buena estereoquímica, con un gráfico de Ramachadran que muestra al 91,0% de los residuos en las regiones más favorables, un 8,3% en regiones adicionalmente permitidas y solo un 0,6% en regiones solo generosamente permitidas, con 0% de residuos en regiones no permitidas.



**Fig. 34:** Diagrama de fluorescencia para platino. Esta medición se realizó sobre el derivado con Pt unido a su estructura. En el diagrama se muestran la magnitud de la contribución de los componentes real ( $f'$ ) e imaginario ( $f''$ ) para la dispersión anómala.

**Tabla 5:** Estadísticas de la colección de datos para derivados con átomos pesados. Los valores en paréntesis corresponden a la capa más alta de resolución.

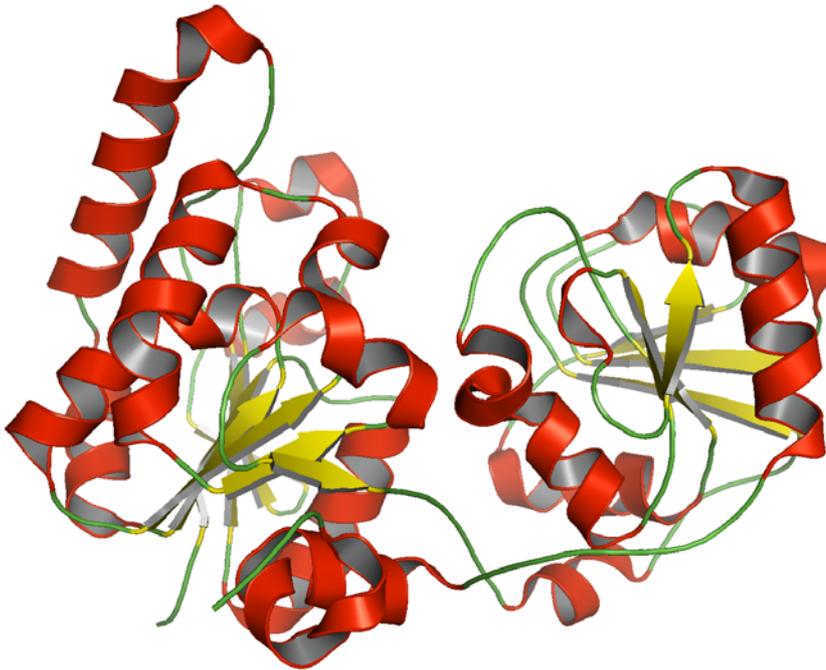
Derivado K <sub>2</sub> PtCl <sub>4</sub>		
Longitudes de onda (Å)	1,0718	1,0722
Rango de resolución (Å)	33,15-2,0 (2,1-2,0)	33,33-2,0 (2,1-2,0)
Reflexiones únicas	51.441	51.635
Rsym (%) <sup>§</sup>	7,3 (38,9)	7,3 (38,9)
Porcentual de reflexiones medidas (%)	95,5 (75,1)	95,5 (74,8)
Porcentual de reflexiones anómalas medidas (%)	91,8 (63,9)	92,1 (74,8)
[I/σ(I)]	17,3 (2,8)	17,4 (2,8)
Mosaicidad(°)	0,21	0,21
Observaciones totales	325.071	330.711
Multiplicidad	6,3	6,3
Multiplicidad de observaciones anómalas	3,4	3,4
<b>Estadísticas de obtención de fases por MAD</b>		
Phasing power isomorfo/[anómalo]	- [2,094]	0,543 [1,267]
Figura de mérito total	0,4811	
Rcullis isomorfo/[anómalo]	- [0,726]	0,683 [0,739]

$$\S \text{Rsym} = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]|}{\sum_{hkl} \sum_i I_{hkl}}$$

## 7.10 Características estructurales de GumK

La enzima posee dimensiones totales de 50 x 50 x 65 Å (Fig. 35). En los mapas de densidad electrónica se pudo observar y trazar desde el residuo 13 hasta el 385 (sobre 406). Los 12 primeros residuos y los residuos 386-406 (que incluyen los 6 residuos de la cola de poli-histidina con que se purificó la proteína) no son visibles, presumiblemente por estar muy desordenados o ser móviles en el cristal. El parámetro estadístico  $R_{\text{free}}$ , que mide la veracidad del modelo estructural, dio un valor de 0,20, indicando un muy buen ajuste entre los datos y el modelo que se muestra. Se observó que el plegamiento de GumK se corresponde con el de la superfamilia GT-B dentro de las glicosiltransferasas. GumK presenta 2 dominios, uno amino y otro carboxi-terminal, con plegamiento tipo Rossmann. Estos dominios muestran una alta homología estructural entre sí (rmsd= 2,90 Å sobre 88 Cα), mostrando que a pesar de la enorme variación de secuencia entre uno y otro dominio, el mismo tipo de plegamiento es adoptado. Estos 2 dominios están compuestos de un núcleo de hojas β casi paralelas, rodeadas de α-hélices, la típica disposición de los dominios tipo Rossmann. En la estructura se observa que GumK está compuesta de 16 α-hélices y 14 hojas β. El dominio N-terminal va desde la Arg13 hasta la His201 e incluye además la hélice Cα7 (residuos 362-380) y el último loop visible en la estructura (residuos 381-385). Posee 10 α-hélices interconectadas con loops de distinto tamaño a 8 hojas-β. La His201 es el residuo final de la hoja Nβ8 y también marca el comienzo de un loop de 8 aminoácidos que se extiende hasta la Asn209 y que interconecta ambos dominios, enlazando a la hoja Nβ8 con la hélice Cα10. Estos dominios están separados por una hendidura de 25 Å de ancho máximo por 15 Å de profundidad (Fig. 35). El loop que los conecta ambos dominios y el que une las hélices Cα6 con Cα7 forman el piso o fondo de esta hendidura. Tanto desde el dominio N como C-terminal protruyen hacia el interior de esta hendidura loops de longitud variable que interconectan α-hélices con hojas β. Es en estas regiones altamente variables donde se encontrarían los aminoácidos catalíticos y de unión a los sustratos, tanto del dador como del aceptor. El sitio de unión del sustrato dador y las interacciones que unen al UDP a la proteína se discutirán más adelante. El sitio de unión del aceptor glicolipídico no pudo ser

observado por cristalografía y las hipótesis que intentan explicar donde se uniría el aceptor serán detalladas en la Discusión.



**Fig. 35** : Representación de la estructura de GumK. En rojo se muestran las  $\alpha$ -hélices, las hojas  $\beta$  en amarillo y los loops que las interconectan en verde. El dominio globular a la izquierda corresponde al extremo N-terminal.

Por experimentos de localización subcelular ya habíamos descrito que GumK se encuentra asociada a membranas (60). Se observa que en el dominio N-terminal se encuentra un parche de aminoácidos básicos intercalados con aminoácidos hidrofóbicos (Fig. 41, Discusión). Ante la evidente ausencia de regiones transmembrana, se especula que GumK podría asociarse a membrana mediante una combinación de factores electrostáticos e hidrofóbicos. Este modelo ya ha sido propuesto para otras GTs (19,44). Además, la localización de este parche es consistente con la hipótesis que muestra una región o canal hidrofóbico sobre la cara N-terminal de la hendidura como el sitio de unión del aceptor.

### 7.11 Estudio cristalográfico de la unión de sustratos

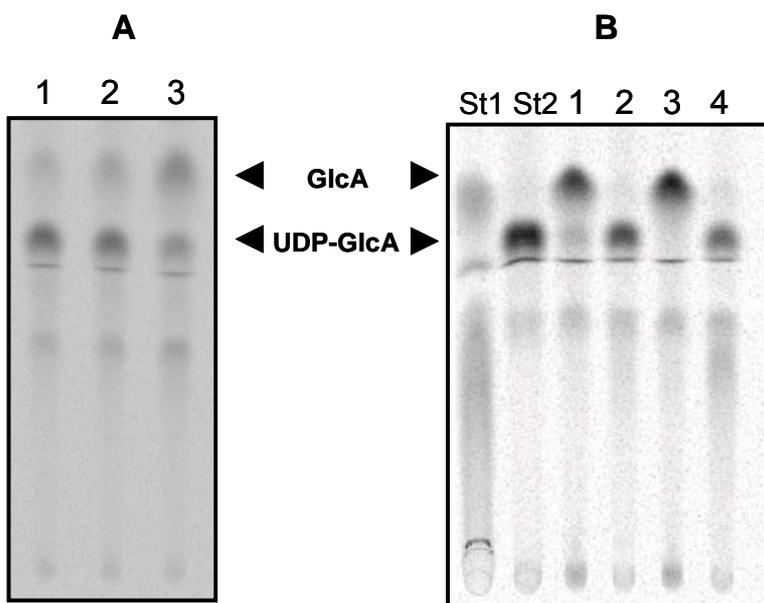
#### 7.11.1 Cristalización de GumK en complejo con sus sustratos

Además de obtener el modelo estructural de la enzima GumK, otro de nuestros objetivos es describir las interacciones de la enzima con sus sustratos. Con este fin, intentamos co-cristalizar GumK en presencia de UDP-GlcA. Los ensayos de co-cristalización se realizaron con la técnica de la gota colgante, en el mismo buffer de cristalización y mismas condiciones que para la apo-enzima. Lo que obtuvimos son cristales de apariencia heterogénea que carecen de un orden molecular y no difractan al ser expuestos a los rayos-X. Seguidamente realizamos experimentos de inmersión de los cristales nativos en soluciones 10 o 100 mM de su sustrato dador UDP-GlcA por lapsos de entre 30 min a 20 hs. También se realizaron experimentos de co-cristalización en presencia del sustrato aceptor natural Poliprenil-PP-Cel-Man purificado de membranas de *X. campestris*. Los tiempos de inmersión fueron de 1, 3 y 8 hs en este caso. Estos cristales no cambiaron su apariencia al ser sumergidos en la solución conteniendo su sustrato aceptor o dador.

#### 7.11.2 Estudio del complejo GumK-UDP

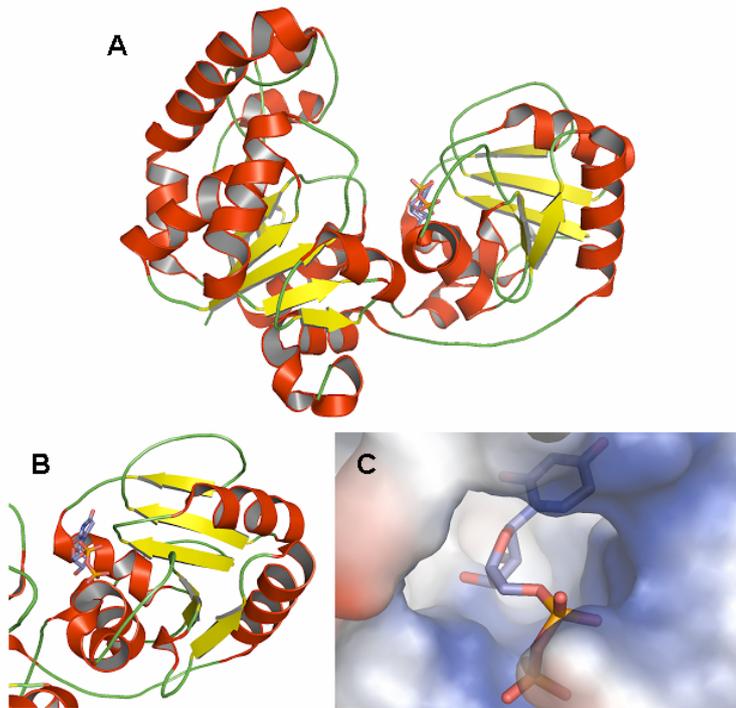
## Resultados

Luego de los experimentos de inmersión los cristales fueron congelados en nitrógeno líquido y se colectaron grupos de datos de difracción en el sincrotrón del LNLS, Campinas, SP, Brasil. Todas las mediciones de datos de difracción fueron hechas a una longitud de onda  $\lambda = 1,427 \text{ \AA}$ , a 100 K, en la línea DO3B-MX1. Los datos de difracción fueron colectados en un detector CCD (MarResearch). Los grupos de datos fueron integrados y procesados con los programas MOSFLM, SCALA y TRUNCATE. Los datos de difracción con mejor resolución y estadísticas de integración se obtuvieron de un experimento de inmersión de 2h en buffer de cristalización suplementado con 100 mM UDP-GlcA. Los parámetros de celda se mantuvieron sin cambios, con una simetría  $P6_522$  y  $a = 121,21 \text{ \AA}$ ,  $b = 121,21 \text{ \AA}$ ,  $c = 170,72 \text{ \AA}$  y  $\alpha = 90^\circ$   $\beta = 90^\circ$   $\gamma = 120^\circ$ . Para la construcción de modelos y mapas de densidad de GumK en complejo con su sustrato soluble se utilizó el programa ARP/wARP intercalando este procedimiento con refinamientos utilizando Refmac. El modelo inicial fue apo-GumK. Se pudo asignar en los mapas de densidad aproximadamente el 90% de la secuencia de GumK con un  $R_{\text{free}}$  inicial de 0,293. Las regiones restantes fueron construidas manualmente usando Coot. Luego, en mapas de densidad atómica  $2Fo-Fc$  se buscó la densidad atómica correspondiente al UDP-GlcA. Se pudo observar la presencia del UDP en 3 cristales, siempre en la misma posición, pero el GlcA no fue detectado. Existía la posibilidad que la enzima hidrolizase la unión fosfato entre el azúcar y el nucleótido, aun en estado cristalino, posibilidad que fue estudiada mediante la incubación de la enzima en buffer de cristalización con UDP- $^{14}\text{C}$ GlcA y posterior análisis de los productos resultantes por TLC (Fig. 36). Estos análisis indican que el UDP-GlcA comienza a hidrolizarse tras 1 h de incubación. Así, es probable que la enzima hidrolize la unión entre este azúcar y los fosfatos del nucleótido en el tiempo del experimento de inmersión.



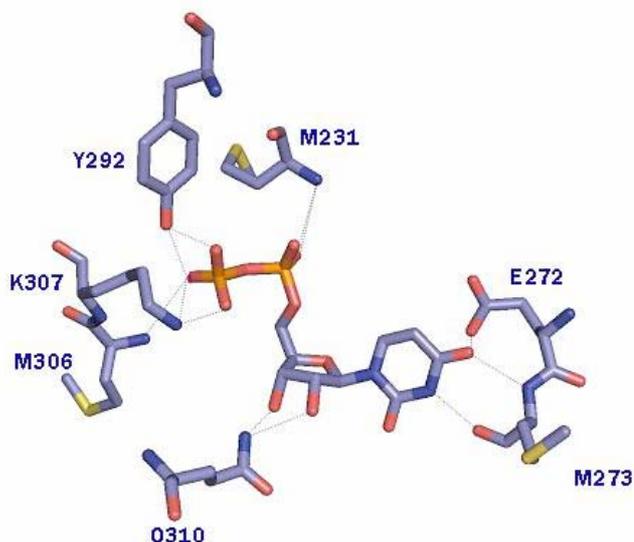
**Fig. 36:** Hidrólisis del sustrato dador. A) Análisis cromatográfico del sustrato UDP- $^{14}\text{C}$ GlcA luego de incubaciones en buffer de cristalización en presencia de GumK. Calle 1: 1 h de incubación, calle 2: 2h, calle 3: 5 h. B) Calle 1: 6h de incubación, calle 2: control negativo (incubación sin GumK) de 6h de incubación, calle 3: 20 h de incubación, calle 4: control negativo de 20 h de incubación. Las calles St1 y St2 corresponden a ácido glucurónico y UDP-GlcA respectivamente, utilizados como estándares.

Es de notar que esto no es un comportamiento inesperado, pues existen varios ejemplos en la literatura de hidrólisis de la unión entre el azúcar y el nucleótido para experimentos de inmersión con nt-azúcar (13,18). Lamentablemente, en caso de reducir el tiempo de inmersión (por ejemplo 30 min o 1h), no pudimos observar la presencia del sustrato unido a la proteína, por lo que no nos fue posible subsanar este problema.



**Fig. 37:** Sitio de unión del UDP en la estructura de GumK. A) Estructura de GumK con el UDP unido sobre la cara C-terminal de la hendidura catalítica. B) Detalle de la unión, se muestran 2 loops y una  $\alpha$ -hélice que están interaccionando con el sustrato. C) Representación de la carga de superficie en GumK, mostrando específicamente el bolsillo de unión del UDP. En rojo se muestran las superficies con carga negativa, blanco neutro y azul carga positiva.

El modelo final del complejo GumK-UDP mostró una estereoquímica muy satisfactoria, con el 91,9% de los residuos en las regiones más favorecidas y el 8,1% restante en regiones adicionalmente favorecidas. El  $R_{free}$  fue de 0,216 en la estructura final. La estructura fue depositada en la base de datos del RCSB, número de PDB 2Q6V. La Fig. 37 muestra la localización del sustrato. La unión del UDP se observa en lo que denominamos el bolsillo de unión del sustrato. Este bolsillo se encuentra en el lado C-terminal de la hendidura, en una superficie con carga positiva. Las interacciones que unen al UDP a la enzima pueden describirse de la siguiente manera: el grupo NH imídico de la cadena principal de Met231 están interaccionando por enlace de hidrogeno con el O2 del fosfato  $\alpha$  del UDP. Lo mismo ocurre entre el grupo  $NH_2$  de la cadena lateral de la Lys307 los O2 y O3 del fosfato  $\beta$ . Los O1 y O3 del fosfato  $\beta$  también son estabilizados por enlaces de hidrogeno con el NH imídico de Met306. Finalmente, el OH del anillo aromático de Tyr292 interacciona con el O2 $\beta$ . La ribosa está estabilizada por contactos entre sus O3'' y O2'' con el  $NH_2$  de la cadena lateral de la Gln310 y entre el O2' y el carbonilo de la Gln310 a través de una molécula de  $H_2O$ . Finalmente la uridina esta estabilizada por puentes de hidrogeno entre su O4' y el grupo NH imidico de Glu272 y el N3' interacciona con el carbonilo de la cadena principal de Met273. Estas interacciones estan graficadas en la Fig. 38. Las distancias entre átomos se detallan en la Tabla 6.



**Fig. 38:** Esquema de los enlaces de hidrógeno que estabilizan al UDP en su unión a GumK. Nótese que la posición de los residuos M306 y K307 ha sido levemente modificada para permitir una mejor visualización.

**Tabla 6:** Puentes de Hidrógeno entre UDP y GumK

Residuo de GumK	Distancia al átomo de UDP(Å)
M231/N	fosfato- $\alpha$ O2A (2.92)
Y292/OH	fosfato- $\beta$ O2B (2.42)
M306/N	fosfato- $\beta$ O1B (3.06)
M306/N	fosfato- $\beta$ O1B (3.26)
K307/N	fosfato- $\alpha$ O2A (2.85)
K307/N	fosfato- $\beta$ O2B (2.35)
K307/N	fosfato- $\beta$ O3B (2.51)
Q310/NE2	ribosa O3'' (2.50)
Q310/NE2	ribosa O2'' (2.88)
M273/N	Uracilo O4' (2.66)
M273/O	Uracilo N3' (2.71)
E272/OE1	Uracilo O4' (3.36)

## 7.12 Unión del aceptor

### 7.12.1 Purificación del sustrato glicolípido y experimentos de inmersión

El estudio de la unión del sustrato aceptor a GumK presentaba 2 grandes dificultades técnicas: 1- El sustrato no está disponible comercialmente, por lo tanto debía ser sintetizado química o enzimáticamente para luego ser purificado. 2- Se deben hacer experimentos de inmersión con cristales nativos, siendo que no sabemos si primero se une el sustrato dador o el aceptor. Por ello no sabemos, *a priori*, si el sustrato va a ser aceptado por la enzima. Para paliar el primero de los problemas, intentamos hacer una síntesis quimio-enzimática a partir de Fitanil-PP-Cel incubado con GDP-manosa en presencia de AceA, una manosiltransferasa de *A. xylinum* purificada y disponible en nuestro laboratorio. Esta síntesis no fue posible, principalmente debido a problemas con la preparación de Fitanil-PP-Cel. Como alternativa, entonces, proseguimos con la extracción y purificación del sustrato natural Poliprenil-PP-Cel-Man a partir de su fuente natural, la membrana de *Xanthomonas campestris*. Así, purificamos 2  $\mu$ moles de sustrato por el método de las células permeabilizadas como se detalla en MyM sección 6.10.1. Este sustrato glicolípido fue utilizado en experimentos de inmersión 10 mM 2 a 8 hs solo o en presencia de UDP. Los cristales fueron difractados en el sincrotrón LNLS, Campinas, SP, Brasil. Todas las mediciones de datos de difracción fueron hechas a una longitud de onda  $\lambda = 1,427$  Å. Los grupos de datos fueron integrados y procesados de la misma manera que para los complejos GumK-UDP. En este caso, analizando los mapas de densidad  $2Fo-Fc$  no se observó la presencia del sustrato aceptor, ni completa ni parcialmente. Las razones para esto pueden estar en que la enzima, en su conformación cristalina no acepta al sustrato aceptor, o bien que el sustrato no está lo suficientemente inmóvil como para aparecer nítidamente en los mapas de densidad.

### 7.13 Mutagénesis y estudio de la cinética de GumK salvaje y mutantes puntuales

La información que nos mostró la estructura de GumK nos sirvió para definir, en primer lugar el tipo de plegamiento de la proteína y además definir que aminoácidos y que tipo de interacciones están promoviendo la unión entre la proteína y sus sustratos. Con esta información en mano decidimos mutagenizar los aminoácidos involucrados en estas interacciones con el objetivo final de estudiar los parámetros cinéticos de estas mutantes. La mutagénesis sitio-dirigida fue realizada con el kit de mutagénesis QuickChange 2 (Stratagene). Se realizaron las siguientes mutaciones: D157A, D157E, D157N, E192A, D207A, M231A, D234A, E272A, E272D, Y292A, L301A, S304A,

K307A y Q310A. Todos estos residuos están involucrados en la unión del UDP o son candidatos a ser el aminoácido catalítico, pues están apuntando su cadena lateral hacia la hendidura central, en las inmediaciones del sitio de unión de los sustratos. Las mutaciones se verificaron por secuenciación y las proteínas mutadas fueron purificadas por el mismo método que se había descrito anteriormente (MyM sección 6.9.3) Para la medición de los parámetros cinéticos se utilizó un método con sustrato marcado radiactivamente, como se describe en MyM sección 6.12.3.

**Tabla 7:** Caracterización cinética de las mutantes de unión al UDP. ND= No detectada.  $K_m$ : Constante de Michaelis.  $V_{max}$ : velocidad máxima de reacción.  $K_{cat}$ : constante catalítica (número de recambio)

MUTANTE	$K_m$ UDPGlcA mM	$V_{max}$ pmoles/ $\mu$ g min	$K_m$ acceptor mM	$V_{max}$ pmoles/ $\mu$ g min	$K_{cat}$ min <sup>-1</sup>
Salvaje	0,062 +/- 0,005	116,2 +/- 2,937	0,197 +/- 0,016	115,9 +/- 2,923	5,23
D157A/E/N	ND	ND	ND	ND	ND
E192A	0,061 +/- 0,006	115 +/- 3,170	0,199 +/- 0,019	118,9 +/- 3,393	5,18
D207A	0,060 +/- 0,005	111,8 +/- 2,655	0,194 +/- 0,017	115 +/- 3,057	5,04
M231A	0,063 +/- 0,008	116,2 +/- 4,195	0,218 +/- 0,016	119,1 +/- 2,697	5,23
D234A	0,063 +/- 0,008	117 +/- 4,259	0,205 +/- 0,018	120 +/- 3,248	5,27
E272A	0,125 +/- 0,009	50,47 +/- 1,182	0,291 +/- 0,021	52,22 +/- 1,222	2,27
E272D	0,100 +/- 0,004	95,02 +/- 1,092	0,267 +/- 0,013	99,48 +/- 1,577	4,28
Y292A	0,212 +/- 0,025	27,52 +/- 1,004	0,274 +/- 0,030	23,46 +/- 0,836	1,23
L301A	0,068 +/- 0,007	120,9 +/- 3,867	0,203 +/- 0,019	121,0 +/- 3,380	5,44
S304A	0,063 +/- 0,009	117,3 +/- 4,749	0,237 +/- 0,014	111,2 +/- 2,021	5,28
K307A	0,435 +/- 0,033	15,22 +/- 0,298	0,385 +/- 0,064	10,82 +/- 0,636	0,68
Q310A	1,186 +/- 0,099	99,85 +/- 2,338	0,246 +/- 0,020	98,12 +/- 2,517	4,50

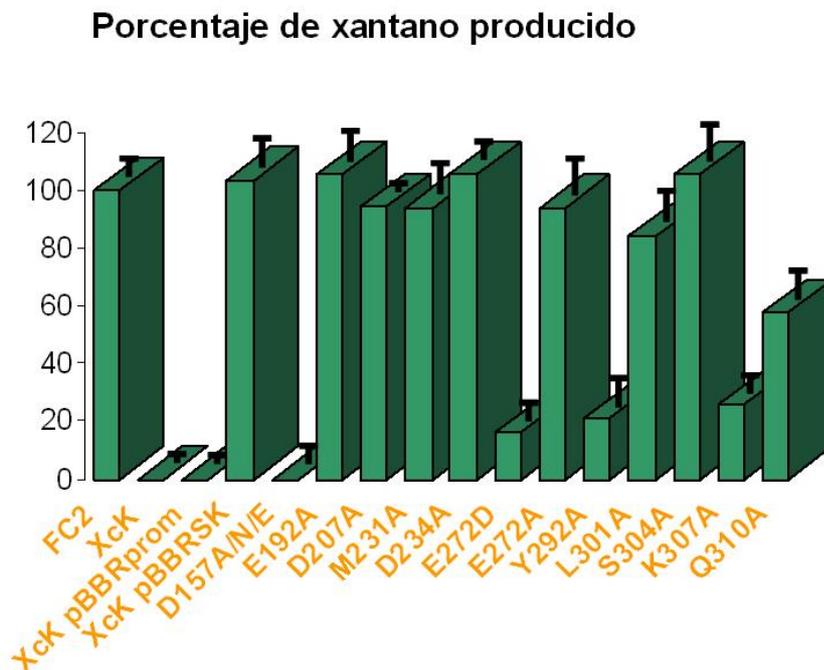
Mediante el análisis de los parámetros cinéticos, se observa que las mutantes que tienen afectada su función son 5: D157A/E/N, E272A o D, Y292A, K307A y Q310A. Esto coincide con aquellas mutantes que tienen función de unión del sustrato. Estos resultados entonces están confirmando la importancia de estos contactos en la unión del sustrato dador. Un caso especial es el de la mutante D157A/N/E. Estas mutantes son las únicas que muestran actividad nula, inclusive considerando las mutaciones más conservadoras del Asp original a Asn o Glu, algo que no se observa para los otros residuos ácidos que apuntan hacia el sitio activo. Esto abre lugar a especulaciones acerca de la función de este residuo, pues su posición dentro de la estructura de la proteína lo hace un candidato ideal para ser el aminoácido catalítico (Fig. 43 Discusión), sobre todo si tomamos en cuenta la evidente analogía que se da con otros residuos ácidos en los cuales se ha demostrado su rol en el mecanismo catalítico; como ejemplos podemos nombrar a los residuos D100 de la  $\beta$ -glucosiltransferasa BGT o E95 de la fucosiltransferasa FucT (18,99), cuya función es realizar el ataque nucleofílico sobre el aceptor, generando un intermediario sobre el que ocurrirá la glicosiltransferencia. Este punto se ampliara con más detalle en la Discusión.

Para el resto de las mutaciones, E272A/D, Y292A, K307A y Q310A, también involucradas en la unión de diferentes átomos del UDP (Fig. 38), el grado de afectación de la función es variable. Si bien pudimos observar un cambio apreciable en los parámetros cinéticos de E272A/D o Y292A, las variaciones más fuertes están centradas en las mutantes K307A y Q310A. Estas 2 mutantes están encargadas de unir oxígenos de los fosfatos o interaccionan con la ribosa, respectivamente. Era particularmente llamativo el caso de Q310A, donde la  $K_m$  sube unas 20 veces, pero aún así la  $V_{max}$  se modifica solo levemente, razón por la cual la constante de intercambio  $K_{cat}$  se mantiene sin mayores alteraciones. Esto puede señalar una interacción importante al momento de unir el dador. Diferente es el caso de K307A. En este caso se observa que la  $K_m$  sube unas 6 veces, acompañada de un descenso de aproximadamente el 90% en la velocidad de reacción, lo que indica que la ausencia de

esta interacción con los fosfatos disminuye la velocidad de la reacción, lo cual queda en evidencia al observar que el valor de  $K_{cat}$  es de un 10-15% para esta mutante respecto a la enzima salvaje. Este mismo comportamiento se observa para Y292A y D272A, donde las disminuciones de la constante de recambio  $K_{cat}$  respecto a la enzima salvaje son de alrededor de un 75 y 50% respectivamente, mientras que la  $K_m$  tiene un aumento de entre 2 y 3 veces.

### 7.14 Análisis de las mutantes de unión por complementación *in vivo*

Nuestra intención al realizar las complementaciones *in vivo* con plásmidos portando *gumK* mutada es analizar magnitud de la afectación en su función en un ensayo mucho mas sensible que la actividad *in vitro*. Para esto clonamos todas las mutantes en el vector pBBRprom y cuantificamos la producción de xantano en la mutante de *X. campestris* XcK portando el plásmido con las diferentes mutaciones. La Fig. 39 muestra el gráfico de la producción relativa porcentual de xantano en las cepas complementadas con las mutantes. Es de destacar el resultado de la complementación con las mutantes D157A/Q/N, que muestran actividad nula, confirmando los resultados de las curvas de cinética enzimática. Este resultado es otra indicación en el sentido que D157 puede tener un rol central en la actividad catalítica de GumK, es decir que podría ser el residuo encargado de realizar el ataque nucleofílico sobre el sustrato en el que se va a transferir el GlcA. Además, con respecto a las demás mutaciones, en general se confirma que la actividad de la enzima esta alterada, en grados que son proporcionales a la afectación que muestran sus parámetros cinéticos, solo que en este caso afectando un proceso que no solo involucra a GumK sino a todas las enzimas del operón *gum*. Esto vuelve a confirmar el rol de estos residuos en la unión y especificidad por UDP-GlcA por parte de GumK.



**Fig. 39:** Producción relativa porcentual de xantano para las mutantes de unión del UDP respecto a la bacteria salvaje. XcK y XcK/pBBRprom son controles negativos. XcK/pBBRSK es un control positivo. Los datos son la media de al menos 3 mediciones independientes.

## 8. DISCUSIÓN

### Caracterización de la función y estructura de GumK

#### 8.1 Caracterización del gen *gumK*

GumK es la primera GT del operón *gum* de *X. campestris* cuyas propiedades bioquímicas han sido estudiadas y caracterizadas en ensayos *in vitro*, involucrando la sobreexpresión y purificación de esta enzima. Según nuestros reportes previos (79), la actividad  $\beta$ -glucuronosiltransferasa en la biosíntesis de la subunidad repetitiva del xantano estaba codificada por un ORF de 888 pb, que codificaba para una proteína de 295 residuos. Los resultados de los experimentos de complementación *in vivo*, y análisis funcional indican que este mismo ORF extendiéndose 315 pb hacia el 5' y codificando para una proteína extendida de 400 aminoácidos es la encargada de esta actividad en la bacteria. Este ORF extendido también se encuentra presente en especies muy emparentadas cuyo genoma ha sido secuenciado, como *Xylella fastidiosa* o *Xanthomonas axonopodis* pv. *citri*, entre otras del género *Xanthomonas* (100). En trabajos previos de nuestro laboratorio, se observó que la transcripción de los genes *gum* se produce a partir de un promotor único que se encuentra río arriba del gen *gumB*. Además se postuló un segundo promotor que se encontraría río arriba del ORF *gumK-s* (101). Como es evidente, este segundo promotor estaría localizado dentro de la secuencia del gen *gumK*, en su región 5' (Fig. 14 Introducción). Para estudiar la posibilidad de que este promotor alternativo tuviese algún rol en la regulación de la síntesis del xantano, expresamos las 2 proteínas, GumK y GumK-s en la cepa salvaje *X. campestris* FC2. Esta cepa sobreexpresando GumK completa o troncada pudo producir xantano al mismo nivel que la cepa salvaje (datos no mostrados). Este resultado sugiere que bajo nuestras condiciones la presencia de GumK-s no interfiere en la producción de xantano, aunque no se descarta una función regulatoria para este segundo promotor.

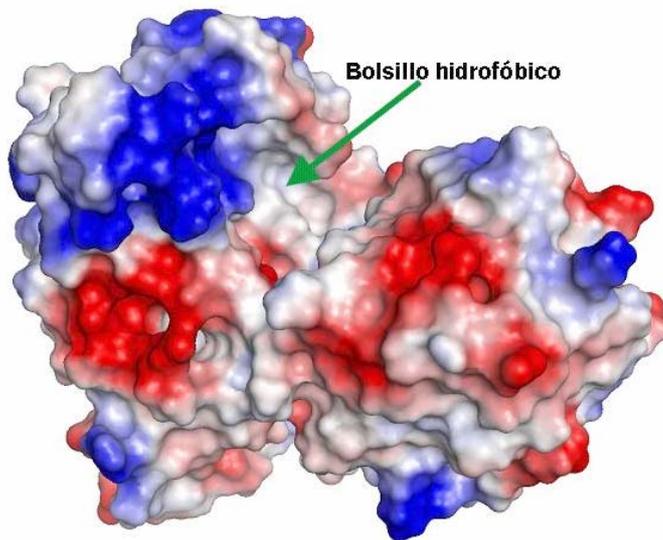
#### 8.2 Especificidad de sustrato aceptor en GumK

Para estudiar funcionalmente a GumK, realizamos ensayos de actividad y cinética enzimática con un extracto de glicolípidos obtenido de células de *X. campestris*, enriquecidos en el aceptor Poliprenil-PP-Cel-Man. Los resultados obtenidos en las incubaciones *in vitro* con la proteína recombinante purificada nos permitieron confirmar que GumK es una glucuronosiltransferasa. La identificación del producto por TLC dió el resultado esperado: celobiosil-manosa-glucuronil es la porción sacarídica de este producto. Además, el análisis del producto de reacción indica que la unión de la porción oligosacarídica al lípido es a través de un enlace pirofosfato. Por otra parte confirmamos que el enlace formado en la reacción catalizada por GumK tiene configuración beta.

En el estudio de la especificidad por su sustrato, observamos que GumK es incapaz de transferir un residuo de GlcA a Cel-Man (aceptor sin los fosfatos y delipidificado) o a aceptores con la porción fosfatídica intacta como el Pi-Cel-Man o PPI-Cel-Man. Este resultado coincide con lo reportado para otras GTs (102,103) e indica la importancia de esta porción de la molécula aceptora en el reconocimiento por parte de la enzima.

A pesar de nuestros esfuerzos, no pudimos co-cristalizar a GumK en presencia de su aceptor, y solo podemos hacer especulaciones acerca de la forma de unión de este sustrato a GumK. En nuestro caso, y basado en los reportes estructurales y estudios de docking para otras GTs con aceptores glicolípidicos (19), postulamos la hipótesis de que el aceptor estaría uniéndose en un bolsillo o canal hidrofóbico sobre la cara N-terminal de la hendidura catalítica, con el extremo no-reductor hacia el piso de esta abertura. Esta situación está graficada en la Fig. 40. Este canal, que posee

aproximadamente  $560 \text{ \AA}^2$  de superficie sería lo suficientemente amplio como para acomodar el trisacárido y los fosfatos del aceptor. Además, las paredes de esta cavidad están rodeadas de varios residuos básicos como argininas, lisinas, 3 histidinas y residuos aromáticos como 2 tirosinas y 1 fenilalanina, en concordancia con las características generales de los sitios de unión de carbohidratos (104). Para WaaG una GT de *Streptomyces viridochromogenes* también se ha postulado un canal hidrofóbico de similares características (105). El aceptor es un glicolípido intermediario en la síntesis del lipopolisacárido de esta bacteria, de naturaleza química similar a nuestro aceptor. Los dominios N-terminal de GumK y WaaG están estructuralmente muy relacionados, con un rmsd=3,4 Å sobre 123 residuos alineados. También se han postulado otras formas de unión de un aceptor glicolipídico a GTs de membrana. En el caso de MurG, una proteína con plegamiento tipo GT-B, se ha propuesto (en base a lo observado en pirofosforilasas) que 2 loops ricos en glicina en el N-terminal, variantes de la secuencia consenso GXGXXG encontrada en otras proteínas que unen dinucleótidos, sería la responsable del reconocimiento del puente difosfato en la porción Poliprenil-PP del sustrato aceptor (44).



**Fig. 40:** Diagrama mostrando la carga de superficie en GumK. En azul la carga positiva, rojo negativa y en blanco carga neutra. Se observa claramente un bolsillo hidrofóbico sobre la cara N-terminal de la hendidura catalítica donde se podría acomodar el sustrato aceptor.

Lo cierto es que en el caso de GTs con un sustrato de membrana como es el glicolípido Poliprenil-PP-Cel-Man o aceptores similares hasta el momento no se ha podido cristalizar la enzima en presencia de este sustrato o en presencia de aceptor y dador. En definitiva, la cuestión de como los fosfatos, la porción lipídica y la sacarídica del aceptor son reconocidos por GumK y/o otras GTs que unan glicolípidos está aún sin constatación experimental.

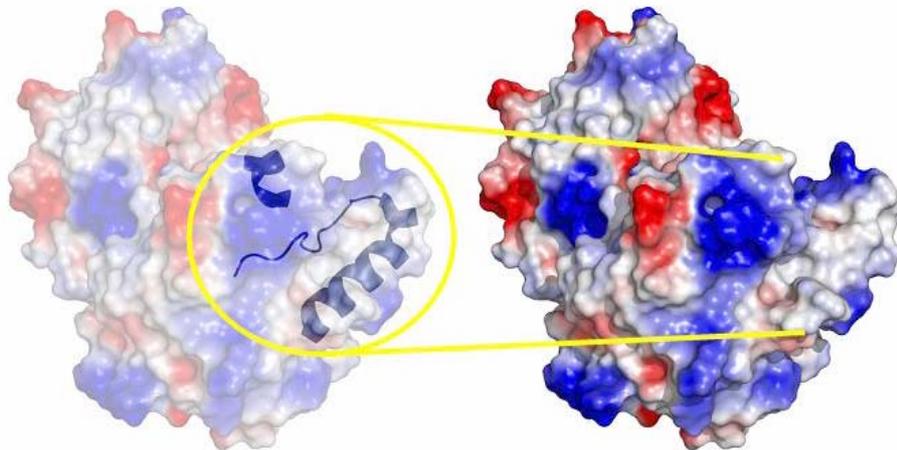
En relación a este punto, muchos exopolisacáridos de bacterias gram-negativas son ensamblados a partir de unidades oligosacarídicas unidas a poliisoprenil-pirofosfatos (106). La porción lipídica del aceptor, en los pocos casos que ha sido estudiada, resultó ser un poliisoprenol, específicamente un undecaprenol, una molécula insaturada de  $C_{55}$  (107). En el caso del lípido transportador de la subunidad pentasacáridica del xantano, la longitud exacta de la cadena del poliisoprenol no se conoce, si bien esta característica no parece ser un estricto requerimiento en el reconocimiento del sustrato aceptor. La sustitución del lípido natural por el fitanilo, lípido saturado ( $C_{20}$ ), resulta en el aceptor Fitanyl-PP-Cel-Man, aceptor reconocido por la enzima. Así, el grado de saturación del lípido y el largo de cadena no son factores críticos. En casos mejor documentados, como el análisis cinético de MurG usando diferentes cadenas lipídicas en análogos de su aceptor glicolipídico, mostraron que la enzima reconoce una cadena de hasta 10 átomos de carbono, y una gran variedad de

cadena, tanto saturadas como insaturadas, si bien prefiere que contengan un doble enlace cis-álíico (108). Este es, hasta el momento, el único trabajo completo publicado en este sentido, lo cual habla de las dificultades técnicas que implica la síntesis y/o purificación de sustrato aceptor glicolipídico y análogos sintéticos. La preferencia de las GTs bacterianas por unidades isoprenoide insaturadas en el aceptor ha sido observada anteriormente (108-110) y podría relacionarse con la proximidad entre la porción lipídica y el punto de acción de la enzima sobre el sustrato aceptor.

Basándose en estos resultados y comparando los requerimientos por el aceptor de MurG Chen y col. (97,108) proponen que las proteínas integrales de membrana que operan en la interfase requieren aceptores con cadenas lipídicas largas, ya que imitan el ambiente de membrana, donde los sustratos se encuentran próximos. Por el contrario, MurG, que también actúa en la interfase, pero carece de dominios transmembrana, puede utilizar sustratos sintéticos con cadenas lipídicas más cortas. Esta predicción sería extensible a todas las GTs que utilizan un sustrato soluble y otro embebido en la membrana. Nuestros resultados apoyan esta teoría, aunque sólo la disponibilidad de nuevos aceptores purificados para GumK nos permitirá hacer un análisis más detallado en cuanto a las características de la porción lipídica. Además, la constatación experimental sobre cuestiones concernientes a la bioquímica del sustrato aceptor como el porqué de la importancia e interacciones del puente pirofosfato y la cadena hidrocarbonada será estudiada en este laboratorio en el futuro, cuando dispongamos de análogos del sustrato natural (que actualmente son de dificultosa síntesis química). Finalmente, la disponibilidad de GumK purificada nos permitió estudiar en mayor detalle la incorporación de grupos acetilos en la biosíntesis del xantano. Por resultados anteriores, sabíamos que ocurría la síntesis de Poliprenil-PP-Cel-Man(Ac) además de su variante no acetilada en células permeabilizadas (77), sugiriendo que esta molécula podía ser reconocida por GumK. En tal caso, el producto de esta reacción, el Poliprenil-PP-Cel-Man(Ac) (que es posible producir en la cepa mutante XcK) sería transformado al pentasacárido acetilado Poliprenil-PP-Cel-Man(Ac)-GlcA. Lo que encontramos es que, *in vitro*, GumK es apenas capaz de utilizar Poliprenil-PP-Cel-Man(Ac) como aceptor. Como se muestra en Resultados (Fig. 27) la cantidad de Poliprenil-PP-Cel-Man(Ac)-GlcA producida representa solo el 10 % de la cantidad de aceptor acetilado presente en la reacción. Esta tasa de transformación esta muy por debajo de lo observado cuando usamos el aceptor no acetilado, donde la transformación es del 100%. Además, la presencia del aceptor acetilado, parece ejercer un efecto negativo sobre la actividad de GumK, ya que solo el 60% del Poliprenil-PP-Cel-Man se transforma a Poliprenil-PP-Cel-Man-GlcA. Estructuralmente puede especularse que la presencia de este grupo acetilo dificulta la correcta unión del aceptor a GumK, interfiriendo directa o indirectamente en el ataque nucleofílico de la manosa, explicando entonces que la reacción no ocurra o lo haga lentamente en presencia del aceptor acetilado. Estos resultados sugieren que en la bacteria el intermediario tetrasacarídico Poliprenil-PP-Cel-Man-GlcA debe ser sintetizado por GumK antes de ser decorado por la acetiltransferasa GumF (Fig. 14 Introducción). Una evidencia a favor de esta hipótesis es que el producto de GumK Poliprenil-PP-Cel-Man-GlcA funciona satisfactoriamente como sustrato de la acetiltransferasa GumF (datos no publicados de nuestro laboratorio).

### 8.3 Localización subcelular de GumK

Anteriormente se había sugerido que las GTs relacionadas con la biosíntesis de succinoglicano y xantano estarían localizadas sobre la cara interna de la membrana plasmática (111). Lo mismo se ha sugerido para varias GTs de estructura conocida (19,44). Para determinar la localización subcelular de GumK hemos usado western blots. Los resultados muestran que la proteína esta asociada a membranas en *X. campestris*. Se ha visto que muchas GTs bacterianas son proteínas integrales de membrana. Otras, como GumK, están asociadas a la membrana a pesar de carecer de pasos transmembrana, algo que es evidente desde la estructura. En estos casos, la naturaleza de la interacción proteína-membrana no esta claramente establecida. Nuestra hipótesis de trabajo para futuros estudios en el laboratorio es que la interacción membrana-GumK es del tipo hidrofóbico/electroestático, a través de una región o parche en el dominio N-terminal. Este parche se ubicaría hacia la región más distal de este dominio, exponiendo residuos las cadenas laterales de argininas, lisinas, metioninas y leucinas presentes en las hélices Na2, Na4 y el loop que precede a esta última. Esto se ve claramente en la Fig. 41 donde se resalta la posición y carga electrostática del parche que mediaría la interacción con la membrana.



**Fig. 41:** Diagrama mostrando los componentes estructurales y la carga de superficie del parche hidrofóbico presente en el dominio N-terminal de GumK. Esta región presenta características hidrofóbicas/básicas.

Otra posibilidad es la formación de hélices anfipáticas. Existen casos como por ejemplo la  $\alpha$ -fucosiltransferasa FucT de *H. pylori* donde se observan claramente cierres de leucina que proveen anclaje a la membrana (18). También se ha demostrado que interacciones electrostáticas e hidrofóbicas están involucradas en la unión y activación de la  $\alpha$ -monoglucosil-diacilglicerol sintasa de *Acholeplasma laidlawii*, alMGS, en un proceso en dos etapas. La enzima se une a la membrana de manera irreversible, y tanto la asociación como la actividad enzimática estarían fuertemente influenciadas por la composición local de la membrana. En especial la carga de los lípidos de membrana podría estar determinando cambios conformacionales, específicamente la formación de una hélice anfipática, que estaría interaccionando con la membrana (112,113). Este comportamiento, por homología estructural, podría también verificarse en otras GTs que no poseen segmentos transmembrana detectables por los medios de predicción, y sin embargo están asociadas a la membrana. En el caso de GumK, la predicción de hélices anfipáticas arroja resultados nulos, por lo que descartamos esta segunda hipótesis acerca del modo de unión a membranas.

Otras propiedades que presenta GumK es que la proteína sólo es solubilizada completamente de la fracción de membranas mediante el uso de detergente; y que su

pl teórico es elevado (8,65) y además se observa una marcada diferencia en el pl para el dominio C-terminal (6,2), comparándolo con el N-terminal (9,97). Esto es una característica común a las proteínas asociadas a la membrana. Respecto a esto último, Edman y col. (114) observaron una distribución de carga distintiva en varias GTs (mayor concentración de cargas positivas en el N-terminal) que se localizan en la membrana y lo relacionaron a la interacción con la misma. Estos autores obtuvieron evidencia experimental de la participación del N-terminal de aMGS en la unión a membrana, estudiando la localización *in vivo* de híbridos de los dominios N- y C-terminal con la proteína verde fluorescente. También el N-terminal de la glucuroniltransferasa GelK de *Sphingomonas paucilomobis*, que presenta un pl elevado, fue expresado separadamente del dominio C-terminal y se lo encontró asociado a la fracción de membranas (115). Específicamente en este caso también se observó que la proteína mantenía actividad si bien era residual.

Resumiendo, pareciera que las GTs que utilizan sustratos lipídicos dependen, al menos en parte, de interacciones electrostáticas mediadas por aminoácidos básicos con fosfolípidos para su unión a la membrana. En el caso de GumK, el tratamiento de la fracción de membranas con buffer alcalino o ácido, resultó en una pobre solubilización de la proteína (5% aproximadamente). Con el uso de NaCl 1M, la proporción de proteína soluble aumentó (20%), si bien precipitó inmediatamente después de la extracción. Contrariamente, la utilización de detergentes no iónicos del tipo polioxietilenglicol (Tritón X-100 y Tween 20) extraen eficientemente a GumK de la fracción de membranas y producen su solubilización completa (y no precipita a los pocos min como en el caso del tratamiento con NaCl). Estos resultados sugieren que las interacciones electrostáticas con las cargas negativas de la membrana son de importancia en el acercamiento de la proteína a la superficie de la bicapa lipídica, donde se encuentra su sustrato aceptor; pero que son las interacciones hidrofóbicas las determinantes en la asociación de GumK con la membrana y de su estabilidad en solución.

Finalmente, otra posibilidad es la formación de complejos entre proteínas unidas a la membrana y citoplasmáticas involucradas en la biosíntesis de polisacáridos. Esta posibilidad ha sido sugerida anteriormente (116-118). En definitiva, nuestros resultados sugieren que GumK no necesitaría de la presencia del resto de las proteínas de la región *gum* para estar asociada a membranas. Esto se demuestra por el hecho de que GumK también se localiza en membranas en la mutante de *X. campestris* del operón *gum* Xc1231 (Fig. 17, Resultados). Así concluimos que esta interacción es una característica intrínseca de la proteína, si bien no descartamos asociaciones del tipo funcional, donde complejos transitorios estarían maximizando la eficiencia en las reacciones catalizadas por las GTs del operón *gum*.

### 8.4 Estado oligomérico

Un punto controvertido en la caracterización de GumK es su perfil de elución en una cromatografía de filtrado en gel. Según este primer parámetro, el PM de GumK corresponde a unos 250-300 kDa. Cálculos más específicos, hechos mediante un experimento de dispersión de luz estático, muestran que el PM de la partícula mayoritaria presente es de aproximadamente 92 kDa. Como el PM del monómero de GumK corresponde a 44,5 kDa, este resultado corresponde a un dímero en solución. Este dímero es funcional, como se verifica por los análisis de la actividad *in vitro*, por lo que se abre el interrogante de cual es la región de la proteína encargada de la interacción proteína-proteína. Ciertamente, en otras GTs donde se observa un comportamiento similar, se han presentado 2 explicaciones: por un lado se especula que superficies hidrofóbicas son las responsables por la asociación inespecífica. En otros casos se ha demostrado una funcionalidad para esta asociación, en un tipo de procesividad que derivaría de esta disposición (119). Lo cierto es que en el caso de GumK necesitamos saber cual es la disposición oligomérica *in vivo*, pues no sabemos

si es esta una condición normal (fisiológica) o un simple artefacto que se produce al tener proteína purificada y concentrada a aproximadamente 20 mg/ml. Asimismo, en estos mismos ensayos de dispersión de luz se observa un pico de 56 kDa (podría ser el monómero), representando no más de un 5% del total de proteína. Estructuralmente, observamos que en la red cristalina la unidad asimétrica es un monómero y no hay asociaciones de mayor valencia, si bien lo que ocurre en la disposición cristalina no tiene necesariamente correlato fisiológico.

### 8.5 Los cationes divalentes:Cuál es su función?

Las estructuras de las GTs resueltas hasta el momento muestran que la característica distintiva entre GT-A y GT-B acerca de la utilización o no de iones metálicos como cofactores de reacción, se debe a que la unión del nt-azúcar es distinta en estas dos superfamilias. Las GTs pertenecientes a GT-B, estabilizan las cargas negativas de los grupos fosfato del nt-azúcar a través de una estructura que presenta un patrón característico tipo  $\alpha/\beta/\alpha$  y que se ha convertido en una especie de marca para la identificación rápida de miembros del clan GT-B. Nuestros resultados muestran que GumK contacta al UDP a partir de residuos presentes en este motivo. En otros casos documentados, las interacciones de enlace de hidrogeno y Van der Waals de este motivo conservado con átomos de los nt-azucars es también la solución al problema de como unir al nt-azúcar en GT-B (17,120). La otra solución es la que utilizan los miembros del grupo GT-A, que consiste en la coordinación de un ión metálico por las cargas negativas del grupo pirofosfato del nt-azúcar. En este grupo, la presencia de iones divalentes es necesaria en el mecanismo catalítico, ya que el ión metálico, coordinado por el motivo DXD, asiste la partida del grupo saliente, estabilizando la carga generada sobre el fosfato- $\alpha$  (17). Hoy en día esta más claro que el mecanismo catalítico de las GT-B de inversión es del tipo ataque nucleofílico  $SN_2$  del sustrato aceptor activado (por desprotonación de un grupo OH) sobre el C1 del dador (Fig. 3 Introducción). No esta claro en este esquema cual sería la función de un catión divalente, y es importante recordar que no hay estructura de una GT-B en la que se observe la presencia de este putativo cofactor. De todos modos, nuestros estudios del requerimiento cationes divalentes como cofactores para la actividad enzimática de GumK mostró que, excepto para el  $Zn^{2+}$ , en general la enzima puede utilizar diferentes cationes con diferentes eficiencias, siendo  $Mg^{++}$  el que produce una mayor actividad catalítica, pero en ningún caso estos cationes son esenciales para la catálisis, por motivos que aun no están claros.

### 8.6 La especificidad por el sustrato dador

Las interacciones entre el UDP y GumK se pueden dividir en 3 grupos: los puentes de hidrogeno con el uracilo, con la ribosa y con los fosfatos. De estos contactos, ninguno parece afectar definitivamente la actividad de la enzima, si bien las mutaciones Q310A y K307A, responsables de parte de los contactos con la ribosa y los fosfatos tiene un marcado efecto de aumento sobre la  $K_m$  y sobre todo que la eficiencia catalítica  $k_{cat}$  disminuya notablemente. Un comportamiento similar se registra para los cambios Y292A y E272A/D, involucrados respectivamente en los contactos entre la proteína y los fosfatos  $\beta$  y el uracilo del dador, respectivamente. En estos casos la  $K_m$  muestra un aumento no tan pronunciado y la  $k_{cat}$  se altera mas levemente, mostrando que en todos los casos estos contactos son relevantes pero no imprescindibles para el reconocimiento y unión del dador (Tabla 7, Resultados). Interesantemente, el descenso en la eficiencia catalítica de estas mutantes purificadas se verifica como una menor producción de xantano *in vivo* en los ensayos de complementación, indicando que afectando la actividad de uno de los componentes de la maquinaria biosintética del xantano se puede afectar todo el proceso y su eficiencia final. Esta observación es

importante desde el punto de vista evolutivo, pues implica una fuerte presión de selección en cuanto a mantener la eficiencia y la especificidad por el dador.

El análisis de la especificidad por el sustrato dador, mostró que GumK sólo es capaz de transferir el residuo azúcar a partir de UDP-GlcA. En la sección 7.6 de Resultados se muestra la lista de hipotéticos dadores utilizados en incubaciones *in vitro* con GumK. De esta lista, hay ciertos sustratos en los cuales era esperable que no hubiera reacción, específicamente los que poseen GDP o ADP como la porción nucleotídica. Esta claro por la estructura de estas moléculas, que por su doble anillo purínico tendrían claros impedimentos estéricos para encajar en el bolsillo de unión del dador, que esta preparando para recibir al UDP (Fig. 37 y 38 Resultados). En otros casos los fundamentos de la especificidad son más sutiles, por ejemplo para UDP-Glc, UDP-GalA o UDP-GalNAc. En ninguno de estos casos se ha observado actividad, si bien los impedimentos estéricos en este caso no son tan evidentes. Evidentemente, al no poseer la estructura de la enzima con el dador completo solo podemos especular acerca de cuales serian las interacciones que determinan la especificidad por la porción sacarídica. Por modelado de la posición del GlcA, este azúcar debería estar interaccionando con las cadenas laterales de los aminoácidos Ser304 y Leu301, que estarían a distancia de enlace de hidrógeno. Por nuestros resultados de incubaciones *in vitro* y complementaciones *in vivo*, vemos que la mutación de estos residuos no tiene mayor efecto sobre la actividad de GumK (exceptuando la complementación de la mutante L301A, cuya producción de xantano disminuye un 25%). Estos resultados no nos permiten descartar una función de estos residuos en la especificidad por la porción sacarídica, pero para aclarar mejor este punto será necesario obtener un complejo cristalino con UDP-GlcA. Por otra parte, estos resultados muestran que GumK no es una enzima con selectividad relajada, sino más bien con una selectividad muy estricta, característica muy extendida en estas enzimas. Funcionalmente, esto tiene su razón de ser: si GumK transfiriese al aceptor un residuo de azúcar tan parecido como la glucosa a partir de UDP-Glc (ampliamente disponible en el citoplasma de la célula), modificaría la estructura del xantano, y sobre todo sus propiedades reológicas, tan importantes a la virulencia de la bacteria.

En términos estructurales, observamos que la coordinación del dador en GT-B esta enmarcada en un bolsillo de carga básica, promoviendo de esta manera la unión y estabilización de los fosfatos. A manera de ejemplo, para la transferasa MurG, hay un patrón  $\alpha/\beta/\alpha$  en la cual las 2 hélices están localizadas cerca de la hendidura catalítica y el dador hace contacto con residuos ubicados en estas hélices y en los loops que las conectan (Fig. 11 Introducción). En GumK, esta subunidad  $\alpha/\beta/\alpha$  también se verifica, con contactos que estabilizan al dador concentrados sobre las hélices Ca3 y Ca4 (residuos 274-284 y 300-314 respectivamente) o los loops adyacentes a éstas (Fig. 42). Esto demuestra, nuevamente, la sorprendente conservación estructural en las GT-B, no solo a nivel de plegamiento general, sino además de subunidades estructurales que unen los sustratos. Como corolario, si calculamos la homología estructural (<http://www.ebi.ac.uk/dali>) del dominio C-terminal de GumK con el de otras GTs se encuentra una muy marcada superposición del sitio activo de por ejemplo la glucógeno sintasa de *Agrobacterium tumefaciens* (rmsd=3,8 Å, PDB 1rzu) o la glicosiltransferasa WaaG (rmsd=3,3 Å, PDB 2iv7), sin importar que estas GTs utilicen sustratos diferentes. Esta similitud en la arquitectura del sitio activo refuerza la idea de un origen evolutivo común para las superfamilias de GTs.

Para fines biotecnológicos, una especificidad “relajada” e inclusive cambiada seria de interés. Podemos señalar que vemos factible un cambio en la especificidad a partir de UDP-azucres, dada la constricción estructural existente para acomodar un dador purínico más voluminoso. De la Fig. 42 se desprende que una factible manera de modificar la especificidad por el dador es mutar residuos de la subunidad  $\alpha/\beta/\alpha$ , en especial los de Ca4 donde parecen estar los residuos responsables de unir el ácido glucurónico.

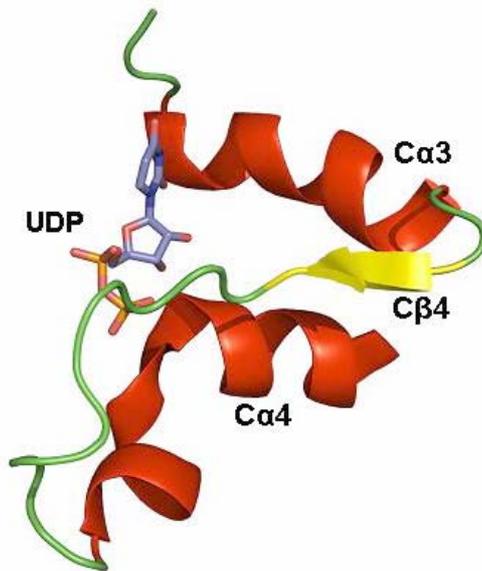
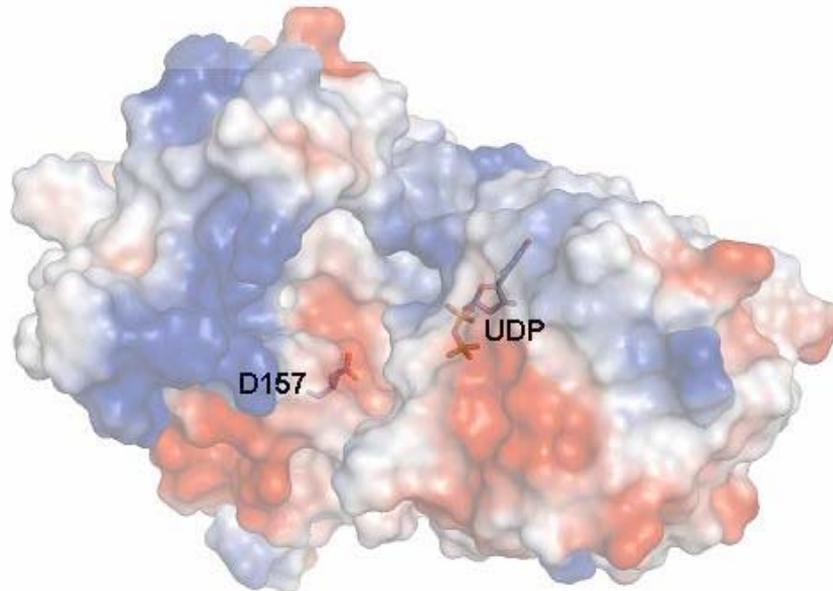


Fig. 42: Motivo  $\alpha/\beta/\alpha$  de unión a sustrato dador en GT-B. Este motivo se encuentra sobre la cara C-terminal de la hendidura central de GumK.

### 8.7 El putativo aminoácido catalítico

Una discusión interesante se suscita en torno del residuo D157. En la comparación entre las estructuras de la enzima unida a UDP y apo-GumK la principal diferencia consiste en la formación de un loop conteniendo a D157. En el cristal GumK en complejo con UDP, la hélice C $\alpha$ 7 se reordena como un loop (residuos 156-164). Este loop, que contiene a D157, cambia de posición relativa y se extiende, con D157 en su extremo, alrededor de 2 Å en dirección a la hendidura central, acercándose al sitio de unión de sustratos (Fig. 43). El cambio D157A/N/E anula totalmente la actividad de GumK, y esto también se verifica en la complementación *in vivo*. Además, como se señala en resultados, GumK mostró actividad hidrolítica sobre UDP-GlcA (Fig. 36 Resultados) en ausencia del sustrato aceptor y tras 1h de incubación, pero la mutante D157A no fue capaz de hidrolizar al UDP-GlcA (datos no mostrados) aún después de una incubación de 20 h. Esta evidencia estructural y bioquímica muestra a este residuo como un fuerte candidato a ser el aminoácido catalítico (responsable de la desprotonación sobre el aceptor para el posterior ataque nucleofílico sobre el sustrato dador). Esto se ve claramente en la Fig. 43 donde se muestra la posición relativa de este putativo aminoácido catalítico. En una posición similar dentro de la hendidura catalítica, se encuentran también los residuos D100 de la glucosiltransferasa BGT del fago T4 y el E95 de la fucosiltransferasa de *H. pylori* (18,99). En el caso de FucT (como en el caso de GumK también), el entorno básico del sitio de unión del aceptor subiría el pK<sub>a</sub> del aceptor, haciendo más factible el ataque nucleofílico de E95. Para confirmar esta hipótesis sería necesario estudiar la posición definitiva del aceptor en un complejo ternario dador-GumK-aceptor. La posición de D157 en tal complejo nos mostraría su factibilidad para realizar la desprotonación que inicia la glicosiltransferencia.



**Fig. 43:** Posición relativa del residuo D157. La cadena lateral de este residuo ácido se muestra como carga negativa (rojo) rodeada de la carga básica (azul) del bolsillo de unión del dador o carga neutra (blanco) en el putativo bolsillo de unión del aceptor.

### 8.8 Paparazeando: La importancia de tener la foto estructural de las GTs

GumK posee las características estructurales de la superfamilia GT-B y presenta, como todas las GTs, la dual característica de una bajísima homología de secuencia con un alto grado de conservación estructural. La comparación estructural con otras GTs muestra que GumK no difiere demasiado de otras GT-B, con un patrón de plegamiento tipo Rossmann que se constata en todas las estructuras publicadas. Esta similitud también se extiende a los 2 dominios globulares N y C-terminal que comparados entre sí muestran un rmsd de 2,9 Å para 88 C- $\alpha$  alineados. Si a esto se agrega la casi absoluta ausencia de secuencias consenso, excepto quizás por el DXD de la superfamilia GT-A (y sin embargo ésta no es una marca exclusiva de superfamilia) o EX7E de las transferasas que actúan con retención es fácil entender que las diferencias más importantes están cifradas en los determinantes estructurales implicados en la unión de los sustratos y en la localización subcelular (en este caso, unión a la membrana). Por caso, el motivo  $\alpha/\beta/\alpha$  es de conservación netamente estructural (no hay una secuencia consenso establecida) y se hace evidente con la estructura cristalográfica en mano, y no tanto con los métodos de predicción por secuencia disponibles. Información mas detallada como la naturaleza de las interacciones responsables de la especificidad por los sustratos o la identificación del residuo catalítico son también puntos que se ven mucho mas esclarecidos con una estructura en mano que con una mera secuencia. Esta información sobre los determinantes estructurales que permiten la unión y discriminación de los diferentes sustratos de las GTs es lo que en el futuro nos permitirá modificar racionalmente la función de este universalmente distribuido, funcionalmente muy diverso y biológicamente esencial grupo de enzimas. Esto redundaría en la producción de nuevos compuestos (oligo y polisacáridos, entre otros) y nuevas funciones biológicas, con toda la potencialidad que esto implica.

## 9. PERSPECTIVAS

A partir de los resultados de esta Tesis se generan nuevos interrogantes que responder. En el futuro podríamos:

- ❖ Co-cristalizar GumK con sus sustratos, bien sea en complejos binarios o ternarios, utilizando mutantes puntuales de actividad nula o análogos no hidrolizables de los sustratos. La información que se desprenda de estudios estructurales de este tipo (inducción de cambios conformacionales, lugar y forma de unión del sustrato aceptor, posición de residuos catalíticamente importantes) nos facilitará una mejor comprensión de la función de GumK.
- ❖ Intentar modificar racionalmente la especificidad por el sustrato dador a partir de la información disponible sobre residuos involucrados en la unión de este sustrato.
- ❖ Determinar fehacientemente la forma de interacción de GumK con la membrana. Investigar el rol de los lípidos de membrana en la unión de GumK y en su actividad catalítica.
- ❖ Realizar estudios cinéticos en presencia de aceptores con cadenas lipídicas de diferente grado de saturación y longitud.
- ❖ Investigar, mediante técnicas para determinar interacciones proteína-proteína (FRET, IP, entre otros) la aparición de complejos (permanentes o transitorios) de la maquinaria biosintética del xantano en la membrana bacteriana.

## 10 BIBLIOGRAFÍA

1. Doores, K. J., Gamblin, D. P., and Davis, B. G. (2006) *Chemistry* 12, 656-665
2. Dube, D. H., and Bertozzi, C. R. (2005) *Nat Rev Drug Discov* 4, 477-488
3. Feizi, T., and Mulloy, B. (2003) *Curr Opin Struct Biol* 13, 602-604
4. Ratner, D. M., Adams, E. W., Disney, M. D., and Seeberger, P. H. (2004) *ChemBiochem* 5, 1375-1383
5. Brockhausen, I., Yang, X., and Harrison, M. (2006) *Methods Mol Biol* 347, 211-236
6. Miyamoto, S. (2006) *Curr Opin Mol Ther* 8, 507-513
7. Sugimoto, Y., Sakoh, H., and Yamada, K. (2004) *Curr Drug Targets Infect Disord* 4, 311-322
8. Van Dyken, S. J., Green, R. S., and Marth, J. D. (2007) *Mol Cell Biol* 27, 1096-1111
9. Tracy, B. S., Avci, F. Y., Linhardt, R. J., and DeAngelis, P. L. (2007) *J Biol Chem* 282, 337-344
10. Tanaka, E., Nakahara, Y., Kuroda, Y., Takano, Y., Kojima, N., and Hojo, H. (2006) *Biosci Biotechnol Biochem* 70, 2515-2522
11. Blixt, O., and Razi, N. (2006) *Methods Enzymol* 415, 137-153
12. Hancock, S. M., Vaughan, M. D., and Withers, S. G. (2006) *Curr Opin Chem Biol* 10, 509-519
13. Vrieland, A., Ruger, W., Driessen, H. P., and Freemont, P. S. (1994) *Embo J* 13, 3413-3422
14. Charnock, S. J., and Davies, G. J. (1999) *Biochemistry* 38, 6380-6385
15. Gastinel, L. N., Cambillau, C., and Bourne, Y. (1999) *EMBO J.* 18, 3546-3557
16. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) *Proc Natl Acad Sci U S A* 100, 845-849
17. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) *Trends Biochem Sci* 30, 53-62
18. Sun, H. Y., Lin, S. W., Ko, T. P., Pan, J. F., Liu, C. L., Lin, C. N., Wang, A. H., and Lin, C. H. (2007) *J Biol Chem* 282, 9973-9982
19. Guerin, M. E., Kordulakova, J., Schaeffer, F., Svetlikova, Z., Buschiazzo, A., Giganti, D., Gicquel, B., Mikusova, K., Jackson, M., and Alzari, P. M. (2007) *J Biol Chem* 282, 20705-20714
20. Kikuchi, N., and Narimatsu, H. (2006) *Biochim Biophys Acta* 1760, 578-583
21. Coutinho, P., M., E. Deleury, G. J. Davies and B. Henrissat. (2003) *J Mol Biol* 328, 307-317
22. Saxena, I. M., and Brown, R. M., Jr. (2000) *Curr Opin Plant Biol* 3, 523-531
23. Davies, G. J. (2001) *Nat Struct Biol* 8, 98-100
24. Iyanagi, T. (2007) *Int Rev Cytol* 260, 35-112
25. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) *Nucleic Acids Res* 32, D138-141
26. Kapitonov, D., and R. K. Yu. (1999) *Glycobiology* 9, 961-978

27. Campbell, J. A., G. J. Davies, V. Bulone, and B. Henrissat. (1997) *Biochem. J.* 326, 929-939
28. Henrissat, B. (1998) *Biochem Soc Trans* 26, 153-156
29. Henrissat, B., and Bairoch, A. (1996) *Biochem J* 316 ( Pt 2), 695-696
30. Campbell, J. A., G. J. Davies, V. Bulone, and B. Henrissat. (1998) *Biochem. J.* 329, 719
31. Lockhoff, O., and Frappa, I. (2002) *Comb Chem High Throughput Screen* 5, 361-372
32. Plante, O. J. (2005) *Comb Chem High Throughput Screen* 8, 153-159
33. Marcus, S. L., Polakowski, R., Seto, N. O., Leinala, E., Borisova, S., Blancher, A., Roubinet, F., Evans, S. V., and Palcic, M. M. (2003) *J Biol Chem* 278, 12403-12405
34. Bourne, Y. a. H., B. (2001) *Curr Opin Struct Biol* 11, 593-600
35. Hu, Y., and S. Walker. (2002) *Chemistry & Biology* 9, 1287-1296
36. Liu, J., and Mushegian, A. (2003) *Protein Sci* 12, 1418-1431
37. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) *Embo J* 19, 5269-5280
38. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J Biol Chem* 275, 34580-34585
39. Kakuda, S., Shiba, T., Ishiguro, M., Tagawa, H., Oka, S., Kajihara, Y., Kawasaki, T., Wakatsuki, S., and Kato, R. (2004) *J Biol Chem* 279, 22693-22703
40. Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H., and Joziase, D. H. (2001) *Embo J* 20, 638-649
41. Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) *Nat Struct Biol* 8, 166-175
42. Gibbons, B. J., Roach, P. J., and Hurley, T. D. (2002) *J Mol Biol* 319, 463-477
43. Mulichak, A. M., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C. T., and Garavito, R. M. (2003) *Proc Natl Acad Sci U S A* 100, 9238-9243
44. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) *Protein Sci* 9, 1045-1052
45. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) *Embo J* 23, 3196-3205
46. Rath, V. L., Ammirati, M., LeMotte, P. K., Fennell, K. F., Mansour, M. N., Danley, D. E., Hynes, T. R., Schulte, G. K., Wasilko, D. J., and Pandit, J. (2000) *Mol Cell* 6, 139-148
47. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) *Chem Biol* 9, 1337-1346
48. Ouzzine, M., Gulberti, S., Levoin, N., Netter, P., Magdalou, J., and Fournel-Gigleux, S. (2002) *J Biol Chem* 277, 25439-25445
49. Pedersen, L. C., T. A. Darden and M. Negishi. (2002) *J Biol Chem* 277, 21869-21873
50. Ramakrishnan, B., and Qasba, P. K. (2001) *J Mol Biol* 310, 205-218
51. Ramakrishnan, B., Ramasamy, V., and Qasba, P. K. (2006) *J Mol Biol* 357, 1619-1633
52. Lairson, L. L., Chiu, C. P., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C., and Withers, S. G. (2004) *J Biol Chem* 279, 28339-28344
53. Negishi, M., Dong, J., Darden, T. A., Pedersen, L. G., and Pedersen, L. C. (2003) *Biochem Biophys Res Commun* 303, 393-398

54. Pedersen, L. C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J. M., Pedersen, L. G., Sugahara, K., and Negishi, M. (2003) *J Biol Chem* 278, 14420-14428
55. Boix, E., Zhang, Y., Swaminathan, G. J., Brew, K., and Acharya, K. R. (2002) *J Biol Chem* 277, 28310-28318
56. Wiggins, C. A., and Munro, S. (1998) *Proc Natl Acad Sci U S A* 95, 7945-7950
57. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) *Glycobiology* 16, 29R-37R
58. Unligil, U. M., and Rini, J. M. (2000) *Curr Opin Struct Biol* 10, 510-517
59. Morera, S., Lariviere, L., Kurzeck, J., Aschke-Sonnenborn, U., Freemont, P. S., Janin, J., and Ruger, W. (2001) *J Mol Biol* 311, 569-577
60. Barreras, M., Abdian, P. L., and Ielpi, L. (2004) *Glycobiology* 14, 233-241
61. Lesk, A. M. (1995) *Curr Opin Struct Biol* 5, 775-783
62. Ni, L., Chokhawala, H. A., Cao, H., Henning, R., Ng, L., Huang, S., Yu, H., Chen, X., and Fisher, A. J. (2007) *Biochemistry* 46, 6288-6298
63. Geremia, R. A., Petroni, E. A., Ielpi, L., and Henrissat, B. (1996) *Biochem J* 318 ( Pt 1), 133-138
64. Kostova, Z., Yan, B. C., Vainauskas, S., Schwartz, R., Menon, A. K., and Orlean, P. (2003) *Eur J Biochem* 270, 4507-4514
65. Chiu, C. P., Watts, A. G., Lairson, L. L., Gilbert, M., Lim, D., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2004) *Nat Struct Mol Biol* 11, 163-170
66. Tarbouriech, N., Charnock, S. J., and Davies, G. J. (2001) *J Mol Biol* 314, 655-661
67. Stragier, P., and Losick, R. (1996) *Annu Rev Genet* 30, 297-241
68. Davies, C., White, S. W., and Nicholas, R. A. (2001) *J Biol Chem* 276, 616-623
69. Ramakrishnan, B., Balaji, P. V., and Qasba, P. K. (2002) *J Mol Biol* 318, 491-502
70. Maeda, Y., Watanabe, R., Harris, C. L., Hong, Y., Ohishi, K., Kinoshita, K., and Kinoshita, T. (2001) *Embo J* 20, 250-261
71. Jansson, P. E., L. Kenne, and B. Lindberg. (1975) *Carbohydr. Res.* 45, 275-282
72. Cadmus, M. C., Rogovin, S. P., Burton, K. A., Pittsley, J. E., Knutson, C. A., and Jeanes, A. (1976) *Can J Microbiol* 22, 942-948
73. Capage, M. A., D. H. Doherty, M.R. Betlach, and R. W. Vanderslice. (1987), International patent.
74. Harding, N. E., J. M. Cleary, D. K. Cabanas, I. G. Rosen, and K. S. Kang. (1987) *J. Bacteriol.* 169, 2854-2861
75. Harding, N. E., S. Raffo, A. Raimondi, J. M. Cleary, and L. Ielpi. (1993) *J. Gen. Microbiol.* 139, 447-457
76. Hotte, B., Rath-Arnold, I., Puhler, A., and Simon, R. (1990) *J Bacteriol* 172, 2804-2807
77. Ielpi, L., R. O. Couso, and M. A. Dankert. (1993) *J. Bacteriol.* 175, 2490-2500
78. Ferreiro, D. U. (1997) in *Departamento de Biología*, pp. 39, Universidad de Buenos Aires, Buenos Aires
79. Katzen, F., D. U. Ferreiro, C. G. Oddo, M. V. Ielmini, A. Becker, A. Pühler, and L. Ielpi. (1998) *J. Bacteriol.* 180, 1607-1617

80. Becker, A., Katzen, F., Puhler, A., and Ielpi, L. (1998) *Appl Microbiol Biotechnol* 50, 145-152
81. Garcia, R. C., Recondo, E., and Dankert, M. (1974) *Eur J Biochem* 43, 93-105
82. Sambrook, J., E. F. Fritsch and T. Maniatis. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
83. Ielmini, M. V. (2007) in *Fundación Instituto Leloir*, Universidad de Buenos Aires, Buenos Aires
84. Colligan, J. E., Dunn, B. E., Ploegh, H. L., Speicher, D. W., and Wingfield, T. W. (2004) *Current Protocols in Protein Science*. Current Protocols (Chanda, V. B., Ed.), 1. 3 vols., John Wiley & Sons, Inc., New York
85. Scott, J. E. (1965) in *General polysaccharides* (Whistler, R. L., ed) Vol. V, pp. 38-44, Academic Press, New York
86. Murray, R. G. E., Wood, W. A., and Krieg, N. R. (1994) *Methods for General and Molecular Bacteriology* (Gerhardt, P., Ed.), ASM, Washington
87. Aitken, A., Learmonth, M., Walker, J. M., Waterborg, J. H., H.R., M., Kruger, N. J., Root, D. D., and Wang, K. (1996) *The Protein Protocols Handbook* (Walker, J. M., Ed.), Humana Press, New Jersey
88. Abdian, P. L., A. C. Lellouch, C. Gautier, L. Ielpi, and R. A. Geremia. (2000) *J. Biol. Chem.* 275, 40568-40575
89. Cserzo, M., Wallin, E., Simon, I., von Heijne, G., and Elofsson, A. (1997) *Protein Eng* 10, 673-676
90. Lupas, A., Van Dyke, M., and Stock, J. (1991) *Science* 252, 1162-1164
91. (1994) *Acta Crystallogr D Biol Crystallogr* 50, 760-763
92. Vonnrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) *Methods Mol Biol* 364, 215-230
93. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2184-2195
94. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132
95. Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop 2nd, and K. M. Peterson. (1995) *Gene* 166, 175-176
96. Camicia, G. L. (2005) in *Fundación Instituto Leloir*, Universidad de Buenos Aires, Buenos Aires
97. Ha, S., Chang, E., Lo, M., Park, P., Ge, M., and Walker, S. (1999) *J Amer Chem Soc* 121, 8415-8426
98. Evrard, G. X., Langer, G. G., Perrakis, A., and Lamzin, V. S. (2007) *Acta Crystallogr D Biol Crystallogr* 63, 108-117
99. Lariviere, L., Gueguen-Chaignon, V., and Morera, S. (2003) *J Mol Biol* 330, 1077-1086
100. Simpson, A. J., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bove, M. R. Briones, M. R. Bueno, A. A. Camargo, L. E. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. Costa, C. M. Costa-Neto, L. L. Coutinho, M. Cristofani, E. Dias-Neto, C. Docena, H. El-Dorry, A. P. Facincani, A. J. Ferreira, V. C. Ferreira, J. A. Ferro, J. S. Fraga, S. C.

- Franca, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. Goldman, S. L. Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. Leite, E. G. Lemos, M. V. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, M. A. Machado, A. M. Madeira, H. M. Madeira, C. L. Marino, M. V. Marques, E. A. Martins, E. M. Martins, A. Y. Matsukuma, C. F. Menck, E. C. Miracca, C. Y. Miyaki, C. B. Monteriro-Vitorello, D. H. Moon, M. A. Nagai, A. L. Nascimento, L. E. Netto, A. Nhani Jr, F. G. Nobrega, L. R. Nunes, M. A. Oliveira, M. C. de Oliveira, R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. Pereira, H. A. Pereira Jr, J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, A. J. de Rosa, V. E. de Rosa Jr, R. G. de Sa, R. V. Santelli, H. E. Sawasaki, A. C. da Silva, A. M. da Silva, F. R. da Silva, W. A. da Silva Jr, J. F. da Silveira, M. L. Silvestri, W. J. Siqueira, A. A. de Souza, A. P. de Souza, M. F. Terenzi, D. Truffi, S. M. Tsai, M. H. Tshako, H. Vallada, M. A. Van Sluys, S. Verjovski-Almeida, A. L. Vettore, M. A. Zago, M. Zatz, J. Meidanis, and J. C. Setubal. (2000) *Nature* 406, 151-157
101. Katzen, F., A. Becker, A. Zorreguieta, A. Puhler, and L. Ielpi. (1996) *J. Bacteriol.* 178, 4313-4318
  102. Geremia, R. A., Roux, M., Ferreiro, D. U., Dauphin-Dubois, R., Lellouch, A. C., and Ielpi, L. (1999) *Mol Gen Genet* 261, 933-940
  103. Lellouch, A. C., and R. A. Geremia. (1999) *J. Bacteriol.* 181, 1141-1148
  104. Davies, G. J., Gloster, T. M., and Henrissat, B. (2005) *Curr Opin Struct Biol* 15, 637-645
  105. Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2006) *Chem Biol* 13, 1143-1152
  106. Sutherland, I. W., and M. I. Tait. (1992) in *Encyclopedia of microbiology*, 1st Ed., pp. 339-349, Academic press, San Diego, Calif.
  107. Johnson, J. G., and Wilson, D. B. (1977) *J Bacteriol* 129, 225-236
  108. Chen, L., H. Men, S. Ha, X. Y. Ye, L. Brunner, Y. Hu, and S. Walker. (2002) *Biochemistry* 41, 6824-6833
  109. Rush, J. S., van Leyen, K., Ouerfelli, O., Wolucka, B., and Waechter, C. J. (1998) *Glycobiology* 8, 1195-1205
  110. Mankowski, T., Sasak, W., Janczura, E., and Chojnacki, T. (1977) *Arch Biochem Biophys* 181, 393-401
  111. Harding, N. E., J. M. Cleary and L. Ielpi. (1995) in *Food Biotechnology: Microorganisms* (Hui, Y. H., and G. G. Khachatourians, ed), pp. 495-514, VCH publishers., New York.
  112. Berg, S., Edman, M., Li, L., Wikstrom, M., and Wieslander, A. (2001) *J Biol Chem* 276, 22056-22063
  113. Lind, J., Ramo, T., Klement, M. L., Barany-Wallje, E., Epand, R. M., Epand, R. F., Maler, L., and Wieslander, A. (2007) *Biochemistry* 46, 5664-5677
  114. Edman, M., Berg, S., Storm, P., Wikstrom, M., Vikstrom, S., Ohman, A., and Wieslander, A. (2003) *J Biol Chem* 278, 8420-8428
  115. Videira, P., A. Fialho, R. A. Geremia, C. Breton, and I. Sa-Correia. (2001) *Biochem. J.* 358, 457-464
  116. Hodson, N., G. Griffiths, N. Cook, M. Pourhossein, E. Gottfridson, T. Lind, K. Lidholt, and I. S. Roberts. (2000) *J. Biol. Chem.* 275, 27311-27315

117. Rigg, G. P., B. Barrett, and I. S. Roberts. (1998) *Microbiology* 144, 2905-2914
118. Whitfield, C., and I. S. Roberts. (1999) *Mol. Microbiol.* 31, 1307-1319
119. Yuan, Y., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., and Walker, S. (2007) *Proc Natl Acad Sci U S A* 104, 5348-5353
120. Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) *Nature* 250, 194-199