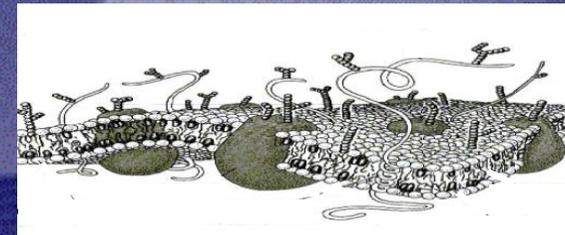


Facultat de Ciències Biològiques  
Departament de Bioquímica i Biologia Molecular  
Universitat de València

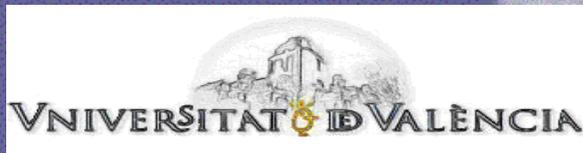


**Determinantes estructurales  
del plegamiento y la interacción  
entre fragmentos transmembrana**



2003 Tesis doctoral Ma del Mar Orzáez Calatayud

Ma del Mar Orzáez Calatayud  
Valencia, 2003



ABREVIATURAS:.....	III
<b>1.INTRODUCCIÓN GENERAL: PROTEÍNAS DE MEMBRANA.....</b>	<b>1</b>
1.1. <i>La membrana</i> .....	12
1.2. <i>Energética del plegamiento de proteínas transmembrana.</i> .....	14
1.3. <i>Biogénesis de proteínas transmembrana.</i> .....	17
1.4. <i>Topología de proteínas transmembrana.</i> .....	18
1.5. <i>Técnicas para la obtención de datos estructurales de proteínas de membrana.</i> .....	19
1.6. <i>Proteínas modelo en el estudio de la asociación de proteínas transmembrana</i> .....	24
Canal de protones M2.....	24
Fosfolamban.....	25
Glicoforina A .....	26
1.7. <i>Tipos de interacciones entre hélices transmembrana.</i> .....	28
Motivos de oligomerización.....	30
1.8. <i>Objetivos.</i> .....	32
<b>2.MATERIALES Y MÉTODOS.....</b>	<b>33</b>
2.1 <i>Manipulación DNA y cepas bacterianas</i> .....	35
2.1.1.Obtención de DNA plasmídico .....	35
2.1.2. Electroforesis en gel de agarosa.....	35
2.1.3. Reacciones de PCR.....	35
2.1.4. Digestión con enzimas de restricción .....	35
2.1.5. Condiciones de crecimiento de <i>E coli.</i> .....	36
2.1.6. Protocolo de transformación de células competentes de <i>E. coli</i> .....	36
2.2. <i>Vectores utilizados y diseño de construcciones.</i> .....	37
2.2.1. Vectores de GpA .....	37
2.2.2. Vectores de Lep.....	38
2.2.3. Preparación de mutantes .....	39
2.2.4. Preparación de construcciones por apareamiento entre oligonucleótidos complementarios.....	42
2.3. <i>Manipulación de proteínas.</i> .....	45
2.3.1. Electroforesis en gel de poliacrilamida .....	45
2.3.2. Expresión y purificación de SN/GpA-6H y mutantes derivados. ....	45
2.3.3. Cuantificación de proteína. ....	46
2.3.4. Cuantificación de los porcentajes de monómero/dímero de SN/GpA y construcciones derivadas. ....	46
2.3.5. Experimentos de competencia proteína/péptido.....	47
2.4. <i>Manipulación de péptidos</i> .....	50
2.4.1. Síntesis y purificación de péptidos.....	50
2.4.2. Espectros de dicroísmo circular.....	55
2.4.3. Electroforesis de Tris-tricina .....	56
2.5. <i>Modelado de estructuras.</i> .....	56
<b>3. INFLUENCIA DEL EXTREMO C-TERMINAL DE GPA EN EL PROCESO DE DIMERIZACIÓN. 59</b>	
3.1. INTRODUCCIÓN .....	61
3.2. RESULTADOS.....	64
3.2.1 <i>Dimerización de los mutantes de delección del extremo C-terminal de GpA.</i> .....	64
3.2.2 <i>Substitución por Leucina de los residuos del extremo C-terminal.</i> .....	66
3.2.3 <i>Influencia de los residuos cargados presentes en el extremo C-terminal en la interacción hélice-hélice.</i> .....	66
3.2.4. <i>Estructura secundaria de los péptidos derivados del extremo C-terminal del fragmento TM de GpA.</i> .....	69
3.2.5. <i>Relevancia del extremo C-terminal en fragmentos transmembrana de mayor longitud.</i> .....	70

3.3. DISCUSIÓN .....	72
<b>4.OLIGOMERIZACIÓN ENTRE FRAGMENTOS TRANSMEMBRANA MODELO.....</b>	<b>77</b>
4.1. INTRODUCCIÓN .....	79
4.2.RESULTADOS .....	81
4.2.1. <i>Contribución de los residuos implicados en el proceso de dimerización.</i> .....	81
4.2.1.1. Inserción secuencial de los residuos del motivo TM de GpA en un fragmento de 18L.....	81
4.2.1.2. Contribución de la leucina 75 al proceso de dimerización.....	86
4.2.1.3. Contribución del residuo en posición 87 al proceso de dimerización.....	87
4.2.2. <i>Efecto de la longitud del fragmento hidrofóbico en la capacidad de dimerización mediada por el motivo GVGVT/18L.</i> .....	88
4.3. DISCUSIÓN .....	90
<b>5. PROLINAS EN FRAGMENTOS TRANSMEMBRANA.....</b>	<b>97</b>
5.1. INTRODUCCIÓN .....	99
5.2. RESULTADOS.....	103
5.2.1. <i>Substitución por prolina de los residuos del segmento TM de GpA.</i> .....	103
5.2.2. <i>Síntesis y purificación del fragmento TM de GpA y del mutante L75P.</i> .....	104
5.2.3. <i>Interfase de interacción proteína/proteína del mutante L75P.</i> .....	105
5.2.3.1. Estudios de competencia proteína /péptido. ....	105
5.2.3.2. Estructura secundaria de los péptidos TM.....	107
5.2.3.3. Estudios de inserción del mutante L75P en membranas naturales.....	110
5.2.4. <i>Estudios de modelado molecular.</i> .....	113
5.3. DISCUSIÓN .....	116
<b>6.CONCLUSIONES.....</b>	<b>121</b>
<b>7.BIBLIOGRAFÍA.....</b>	<b>125</b>

## ABREVIATURAS:

ATP	adenosina trifosfato
BCA	ácido bicincónico
C-terminal	carboxilo terminal
DCM	diclorometano
DIC	diisopropilcarbodiimida
DIEA	N,N-diisopropilmetilamina
DNA	ácido desoxirribonucleico
DNTPs	desoxiribonucleótidos trifosfato
DOPC	dioleoil fosfatidilcolina
DPC	dodecilsulfocolina
EDTA	ácido etilendiaminotetraacético
EPR	resonancia magnética de electrón
F-moc	9-fluorometoxicarbonil
GpA	Glicoforina A
GTP	guanosina trifosfato
HATU	o-(z-azabenzotriazol-1-yl)-N,N,N',N'-(tetrametil-uronium hexafluorofosfato)
HPLC	cromatografía líquida de alta resolución
IPTG	isopropil- $\beta$ -D-tiogalactopiranosido
Lep	peptidasa del péptido señal
LPC	lisofosfatidilcolina
MBP	proteína de unión a la maltosa
MGD	distancia mínima de glicosilación
mRNA	RNA mensajero
N-terminal	amino terminal
PCR	reacción en cadena de la polimerasa
PDB	Protein Data Bank
RMN	resonancia magnética nuclear
RNA	ácido ribonucleico
SDS	dodecil sulfato sódico

## Abreviaturas

---

SDS-PAGE	electroforesis de poliacrilamida en presencia de SDS
SN	nucleasa
SP	péptido señal
SR	receptor de la SRP
SRP	partícula de reconocimiento del péptido señal
TFA	ácido trifluoroacético
TM	transmembrana
UV	ultravioleta

Para nombrar a los aminoácidos se ha utilizado tanto el código de una como el de tres letras.

**1.Introducción general:  
Proteínas de membrana**

Introducción general: Proteínas de membrana \_\_\_\_\_

La membrana celular representa una barrera capaz de independizar el interior de la célula del ambiente externo, permitiéndole generar y mantener un medio completamente diferente al que le rodea. Sin embargo, para generar este orden interno y mantenerlo es necesario que esta frontera permita el intercambio de material e información con el exterior. En este punto desempeñan un papel esencial las proteínas de membrana, que se convierten en el nexo de comunicación entre ambos ambientes. El análisis de secuencias genómicas completas de diversos organismos muestra que, aproximadamente un 20-30% de las pautas de lectura abierta, codifican para proteínas que atraviesan la membrana (Wallin & von Heijne, 1998). Este número aumenta conforme se incrementa la complejidad del organismo, lo que se correlaciona, a su vez, con el incremento de los procesos de comunicación intercelular. De este modo se observa, que el número de proteínas transmembrana es menor en el caso de organismos procariotas que en el de eucariotas (Wallin & von Heijne, 1998). Su enorme importancia se refleja también en el elevado número de procesos básicos para la supervivencia celular en los que este tipo de proteínas están implicadas. La mayoría de componentes tanto de la cadena respiratoria como de la fotosintética son proteínas transmembrana, así como, numerosas proteínas estructurales, transportadores, moléculas implicadas en la transducción de señal (receptores acoplados a proteínas G, receptores del tipo tirosin-quinasa, canales iónicos regulados por ligando o por voltaje, mediadores de la señalización célula-célula), antígenos de superficie, moléculas de adhesión celular, etc. En definitiva, tanto el elevado porcentaje, como las relevantes funciones que desempeñan estas proteínas las convierten en componentes esenciales de la maquinaria celular. De hecho, en la actualidad, la mayoría de las dianas terapéuticas utilizadas para el desarrollo de fármacos son proteínas de membrana (Torres et al., 2003), lo que se ve favorecido por el hecho de que para alcanzar a su molécula objetivo no es necesaria su internalización en la célula.

La identificación de una proteína como transmembrana es relativamente sencilla, gracias a que en los últimos años se han desarrollado numerosos algoritmos (Deber et al., 2001; Martelli et al., 2002; Moller et al., 2001; Pilpel et al., 1999; Sonnhammer et al., 1998) en los que, utilizando diversos criterios, basados fundamentalmente en la existencia de fragmentos hidrofóbicos de una determinada longitud, y usando diversas escalas de hidropatía (Eisenberg et al., 1984; Engelman et al., 1986; Kyte & Doolittle, 1982; von Heijne, 1992); (Wimley & White, 1996); (Deber et al., 1986), se ha conseguido la identificación de proteínas transmembrana, con un alto nivel de fiabilidad. Sin embargo, el medio en el que estas proteínas adoptan su estructura nativa y desarrollan su función, la

membrana, ha dificultado la aplicación directa de las técnicas de estudio habituales para proteínas solubles. Esta dificultad se refleja en el hecho de que en la era de la explosión de obtención de estructuras proteicas de alta resolución, se limita a unas docenas el número de proteínas de membrana resueltas (figura 1.1). Esta carencia es debida a que tanto los procesos de expresión heteróloga, como de purificación, resolución de espectros, cristalización etc. incrementan su dificultad cuando trabajamos con este tipo de proteínas.

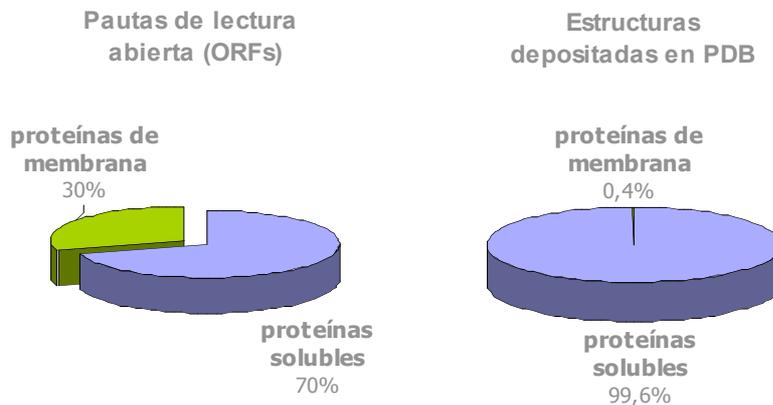


Figura 1.1: En la parte izquierda de la figura se muestra la proporción media de proteínas de membrana vs proteínas solubles en las pautas de lectura abiertas (ORF) de los organismos secuenciados hasta la actualidad. En la parte derecha se representa el porcentaje de este tipo de proteínas cuya estructura está depositada en el PDB (“protein data bank”).

## 1.1. La membrana

La membrana es un medio heterogéneo compuesto principalmente por fosfolípidos y proteínas. El elevado nivel de desorden térmico de estas bicapas fluidas no permite la resolución atómica de imágenes cristalográficas tridimensionales. Sin embargo, el estudio por métodos de difracción de rayos X y de neutrones de bicapas multilamelares, obtenidas mediante la dispersión de fosfolípidos en agua o sobre superficies (cristales líquidos), ha permitido la obtención de imágenes monodimensionales que han aportado información muy relevante sobre la estructura de la membrana (Wiener et al., 1991; Wiener & White, 1992). El estudio de la distribución espacial estadística de los

principales grupos que forman parte de la estructura de los lípidos, respecto a la normal de la bicapa, ha permitido entender que los fosfolípidos se organizan en una bicapa fluida en la que se pueden observar dos zonas diferenciadas por su densidad de carga y por la presencia de agua de hidratación. Las cabezas polares de los fosfolípidos se sitúan preferentemente en una región denominada *interfase* que se extiende unos 15Å a ambos lados de la membrana, mientras que la zona central o *núcleo hidrofóbico* está formado por las cadenas hidrocarbonadas y se extiende por unos 30Å (figura 1.2). La interfase es una zona químicamente heterogénea, donde existe una amplia variedad de posibilidades de interacciones no covalentes con fragmentos proteicos.

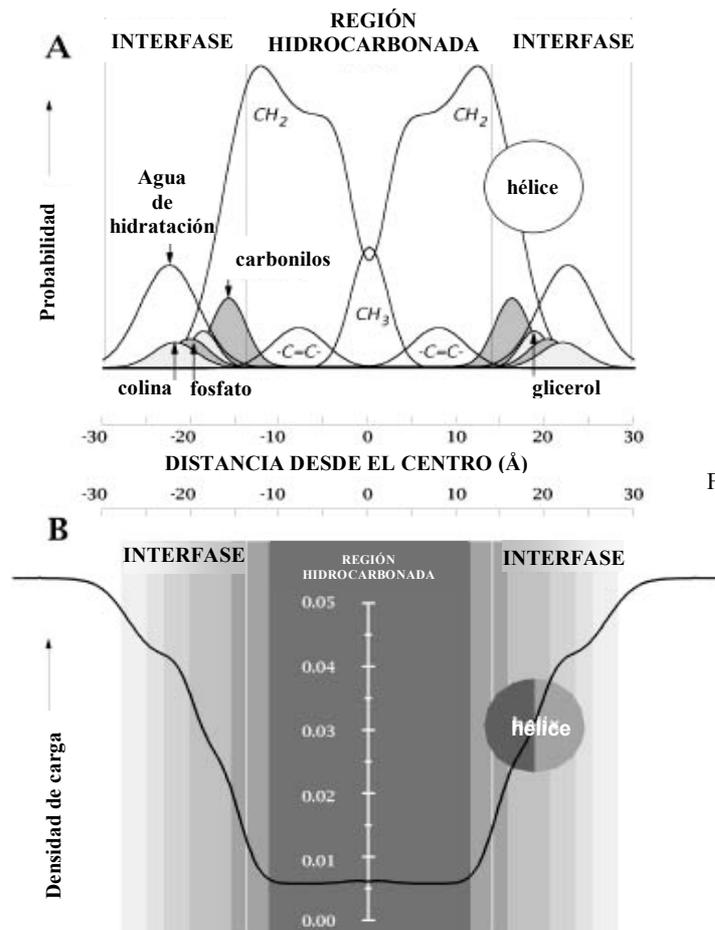


Figura 1.2: (A) Estructura de una bicapa fluida de DOPC (dioleoil fosfatidilcolina) en la que se muestra la probabilidad de encontrar los grupos que forman la estructura del lípido en las diferentes posiciones de la bicapa. En la región de la interfase se muestra la posición que ocuparía una hélice anfipática que se situaría perpendicular a la normal de la bicapa lipídica (B) Perfil de polaridad de esa membrana. Figura adaptada de <http://blanco.biomol.uci.edu> (Dr. Stephen White ).

Es relevante destacar el hecho de que la extensión total de la interfase es equivalente a la de la zona hidrocarbonada, ya que en los primeros modelos se proponía para esta región heterogénea una extensión menor. Además, los 15 Å de esta interfase permiten acomodar una hélice alfa anfipática plegada, perpendicular a la normal de la bicapa, cuyo diámetro es de unos 10 Å (figura 1.2).

## 1.2. Energética del plegamiento de proteínas transmembrana.

En un ambiente tan heterogéneo como el descrito previamente adoptan su estructura nativa las proteínas transmembrana. Los grupos amida y carbonilo del enlace peptídico son grupos polares, de modo que, para mantenerse en el interior de la bicapa, las proteínas deben adoptar plegamientos en los que la exposición al medio de estos grupos quede minimizada. A pesar del número creciente de estructuras de proteínas de membrana que se conoce, sólo se han descrito, hasta el momento, dos tipos de plegamiento, los *haces de hélices alfa* y los *barriles  $\beta$*  (figura 1.3). La mayoría de proteínas conocidas se pliegan *como haces de hélices alfa*, quedando limitado el plegamiento *en barril  $\beta$*  a un pequeño grupo de proteínas, que se encuentran en la membrana externa de bacterias *Gram* negativas,

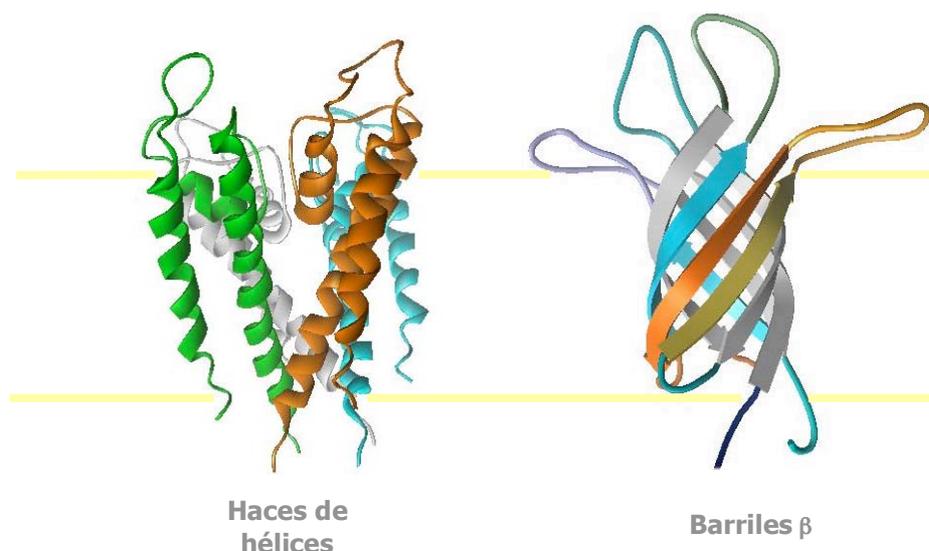


Figura 1.3: Tipos de plegamiento en proteínas transmembrana. En la parte izquierda de la figura se muestra la estructura del canal de potasio KcsA (código PDB, 1BL8) y en la derecha la de la proteína OmpA (código PDB, 1G90).

mitocondrias y cloroplastos, formando rígidos poros, como es el caso de las porinas bacterianas. En cualquier caso, ambos tipos de plegamiento se caracterizan por dar lugar a la formación de estructuras en las que los posibles grupos formadores de puentes de hidrógeno se encuentran satisfechos.

El presente trabajo se centra en el estudio de las proteínas que se pliegan como hélices alfa, por ser el grupo más abundante y sobre el que recaen las funciones más relevantes. Este tipo de proteínas pueden poseer un único fragmento capaz de atravesar la membrana y en tal caso reciben el nombre de *proteínas bitópicas*, o estar formadas por varios de estos fragmentos, denominadas *proteínas politópicas*.

Un paso esencial en la comprensión de cómo se produce el plegamiento de estas proteínas helicoidales transmembrana fue la introducción por Popot y Engelman en 1990 de la hipótesis del *modelo de plegamiento en dos etapas* (Popot & Engelman, 1990; Wimley & White, 1996). Según este modelo la formación de los haces de hélices transmembrana es un proceso que ocurre al menos en dos estadios energéticos independientes (figura 1.4). En primer lugar se forman las hélices de modo individual, proceso que ocurre mediado fundamentalmente por el efecto hidrofóbico. En segundo lugar, las hélices interaccionan unas con otras para dar lugar al plegamiento final de la proteína. Este modelo se propuso, fundamentalmente, basándose en experimentos realizados con bacteriorrodopsina, proteína con siete fragmentos transmembrana, en los que se recupera la estructura activa de la proteína en vesículas lipídicas en presencia del grupo prostético retinal, a partir de dos fragmentos, obtenidos por digestión con Quimotripsina, uno formado por cinco segmentos TM y el otro que contiene los dos restantes (Popot et al., 1987). Este modelo, que ha sido refinado a lo largo de los años, ha constituido la base fundamental para la comprensión del plegamiento de este tipo proteico.

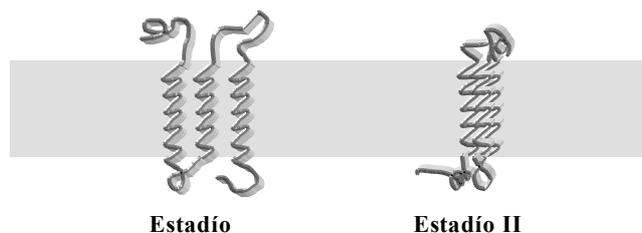


Figura 1.4: Modelo de plegamiento en dos estadios. En el primer estadio se forman las hélices transmembrana de forma independiente, llevándose a cabo la asociación entre ellas en un segundo estadio.

En 1989 Jacobs y White (Jacobs & White, 1989) propusieron un modelo termodinámico en tres etapas para el plegamiento de proteínas de membrana basado en parámetros estructurales y termodinámicos obtenidos del estudio del reparto lípido/agua de pequeños péptidos hidrofóbicos. Dicho modelo consiste en un proceso de reparto en la interfase de la membrana, seguido de una etapa de plegamiento en dicha interfase y su posterior inserción en la membrana. De la combinación de los dos modelos previos surge un modelo en cuatro etapas (figura 1.5) (White & Wimley, 1999): (i) una primera etapa de reparto del polipéptido desplegado en la interfase de la membrana, (ii) la formación de estructura secundaria en ese mismo entorno, (iii) la inserción del polipéptido plegado a través de la membrana, y (iv) la asociación de los elementos de estructura secundaria en el interior de la membrana (figura 1.5). En este punto hay que distinguir entre las proteínas de membrana no constitutivas, como es el caso de muchas toxinas, en las que la ruta de inserción estaría constituida por estas cuatro etapas y las proteínas transmembrana constitutivas en las que el proceso de inserción se inicia directamente asociado a la biosíntesis de las proteínas. En este último caso, partiríamos de la fase en la que los elementos de estructura secundaria se encuentran formados e insertados en la membrana y el proceso en sentido inverso describiría el desplegamiento de la proteína.

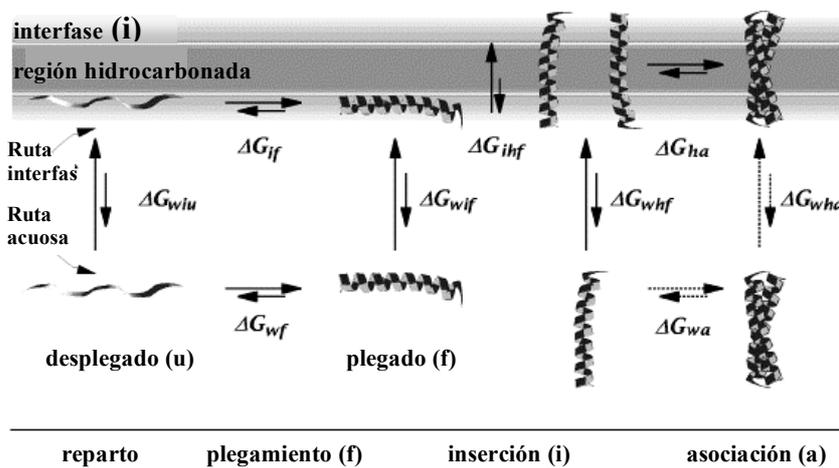


Figura 1.5: Ciclo termodinámico de las cuatro etapas que describen la energética del reparto, plegamiento, inserción y asociación de hélices alfa en una bicapa lipídica. Los símbolos de  $\Delta G$  indican las energías libres estándar de transferencia, para cada etapa del ciclo. i) interfase, w) agua h) núcleo hidrofóbico, u) desplegado, f) plegado, a) asociación. Así, por ejemplo,  $\Delta G_{wif}$ , sería la energía libre de transferencia de un péptido plegado desde la interfase al medio acuoso. Adaptado de <http://blanco.biomol.uci.edu>.

### 1.3. Biogénesis de proteínas transmembrana.

En la célula eucariota, el proceso de inserción de la mayoría de proteínas en la membrana se encuentra acoplado a la síntesis proteica. Las proteínas que van a insertarse en la membrana poseen *secuencias señal* que las dirigen, durante su propia síntesis, a la membrana del retículo endoplasmático. Estas secuencias son bastante heterogéneas, presentando como única característica común la presencia un segmento de entre 6 y 20 residuos hidrofóbicos, seguido de una región de residuos cargados positivamente. Conforme una *secuencia señal* emerge del ribosoma es reconocida por la *partícula de reconocimiento del péptido señal (SRP)*, una ribonucleoproteína formada por 6 polipéptidos y una pequeña molécula de RNA. La unión de la SRP al complejo ribosoma-mRNA-polipéptido naciente, detiene el proceso de traducción. Una de las subunidades de esta proteína (SRP54) posee un surco hidrofóbico, formado por un gran número de residuos de metionina, capaz de unir la secuencia señal. Además, esta subunidad posee también un dominio capaz de unir e hidrolizar GTP. El complejo, mRNA, SRP junto con el ribosoma y la proteína naciente se dirige a la membrana del retículo endoplásmico, para unirse al receptor de la SRP (SR). Este receptor, está compuesto por dos subunidades, SR $\alpha$ , orientada hacia la cara citoplasmática y SR $\beta$ , que se encuentra anclada a la membrana. Estas dos subunidades poseen también capacidad de unir e hidrolizar GTP. La interacción entre la SRP y SR $\alpha$ , produce la liberación de la secuencia señal, la hidrólisis de GTP y la unión del

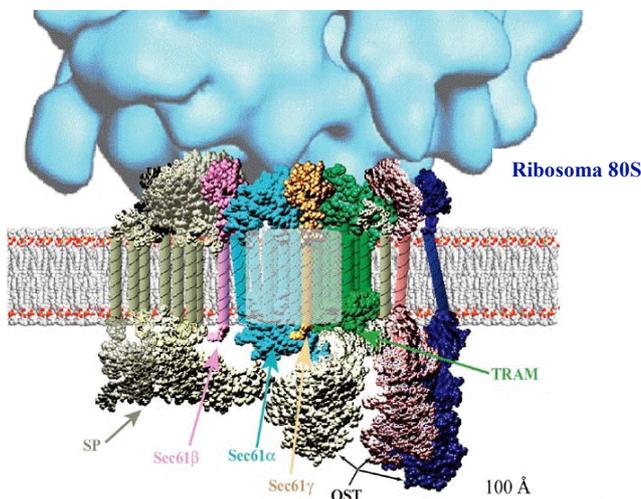


Figura1.6: Traslocón de mamíferos. El traslocón se encuentra formado por 3-4 copias del heterotrímero Sec61 $\alpha\beta\gamma$ . La proteína TRAM interacciona con la secuencia señal y se requiere para una señalización eficiente. En la figura también se muestran la peptidasa del péptido señal (SP) y la oligosacaryltransferasa (OST), proteínas asociadas al traslocón. Figura adaptada de Arthur E. Johnson.

(<http://hscconcord.tamu.edu/AEJohnson/Home>)

ribosoma al traslocón, complejo proteico formado por el heterotrímero Sec1 $\alpha\beta\gamma$ , a través de la proteína Sec61 $\alpha$ . De este modo, se reanuda la traducción, iniciándose el proceso de traslocación de los dominios extracelulares y la difusión lateral (inserción) de los dominios TM a la bicapa. El complejo trimérico formado por las proteínas Sec61 forma un canal acuoso por el que se produce la traslocación, aunque otras proteínas asociadas al traslocón colaboran en este proceso. Así encontramos a la peptidasa del péptido señal (PS) o a la oligosacaril transferasa, que se encargan de la eliminación del segmento que contiene el péptido señal y de la unión covalente de azúcares a las cadenas polipeptídicas que están siendo translocadas, respectivamente (figura 1.6).

Este sistema de traslocación cotraduccional, es muy similar en organismos procariotas, en los que también interviene un canal trimérico formado por proteínas Sec. En todos los casos el canal formado por estas proteínas tiene un comportamiento pasivo, de modo que necesita una fuerza motora adicional para dirigir la traslocación. En el caso de la inserción cotraduccional esta fuerza motora proviene del propio ribosoma, mientras que en la inserción posttraduccional, tenemos que distinguir entre procariotas y eucariotas. En eucariotas intervienen las proteínas integrales de membrana Sec62/Sec63 y la proteína del lumen del retículo BiP, con consumo de ATP. En procariotas la traslocación posttraduccional utiliza la ATPasa SecA citosólica para empujar la cadena polipeptídica a través de la membrana.

#### **1.4. Topología de proteínas transmembrana.**

La topología que adopta una proteína en la membrana depende de factores múltiples entre los que se encuentran la longitud e hidrofobicidad de sus fragmentos hidrofóbicos, la presencia o ausencia de secuencia señal, la extensión de la secuencia hidrofílica en posición N-terminal, las cargas que flanquean sus segmentos TMs, etc. Las proteínas que poseen un único segmento transmembrana se clasifican en tres tipos según su topología y los factores que la determinan (figura 1.7). Las de tipo I, que orientan su extremo C-terminal hacia el citoplasma, poseen una secuencia señal hidrofóbica corta (7-15 residuos) que las dirige a la membrana y que es digerida por proteasas específicas y una señal de parada de transferencia. En las proteínas de tipo II el segmento transmembrana hace de péptido señal y por tanto no se corta, siendo translocado su extremo C-terminal. Las proteínas transmembrana

de tipo III poseen señales de anclaje a la membrana pero reversas, de modo que su extremo C-terminal queda orientado hacia el citoplasma. Existe un último tipo de proteínas cuya inserción es posttraduccional y que no dependen de la maquinaria de secreción, pero que necesitan ATP para la inserción de su fragmento hidrofóbico, localizado comúnmente en el extremo C-terminal de la proteína.

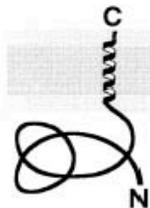
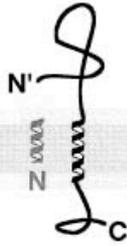
señal	Señales de translocación C-terminales			Translocación N-terminal
determinantes topogénicos	Señal C-terminal	señal de corte + parada de transferencia	señal de anclaje	señal de anclaje reversa
maquinaria	desconocida	SRP/SR/Sec61 +peptidasa señal	SRP/SR/Sec61	SRP/SR/Sec61
exterior topología final citoplasma		 <p><b>Tipo I</b></p>	 <p><b>Tipo II</b></p>	 <p><b>Tipo III</b></p>
ejemplos	Sinaptobrevina; citocromo b <sub>5</sub>	Glicoforina A; receptor LDL	Receptor transferrina; galactosil transferasa	Citocromos P-450

Figura 1.7: Clasificación de proteínas transmembrana bitópicas según su topología y sus determinantes topogénicos.

### 1.5. Técnicas para la obtención de datos estructurales de proteínas de membrana.

La obtención de estructuras de alta resolución de proteínas de membrana, es un trabajo extremadamente complicado, ya que las técnicas de RMN en solución y cristalografía de rayos X, utilizadas habitualmente para la resolución de proteínas solubles no son directamente aplicables. La obtención de cristales de proteínas de membrana en presencia de detergentes o fosfolípidos es en muchos casos imposible. El primer problema que se plantea es la obtención de cantidad suficiente de proteína, puesto que generalmente estas proteínas resultan tóxicas en sistemas heterólogos y su

sobreexpresión no es sencilla. Un problema adicional es encontrar un medio adecuado para disolver la proteína, normalmente se utilizan medios micelares. En cualquier caso, debe tratarse de moléculas capaces de solubilizar la proteína, manteniendo su actividad y de un tamaño lo suficientemente pequeño para permitir el contacto entre las regiones polares de la proteína, de forma que ésta pueda cristalizar en su conformación nativa. Por otra parte, la aplicación de técnicas de RMN en solución a este tipo proteico, se encuentra con la limitación de la necesidad de resolver complejos de gran tamaño proteína-micela. A pesar de todas estas limitaciones, se han obtenido cristales de unas pocas decenas de proteínas de membrana, lo que ha permitido el análisis de su estructura con un elevado nivel de resolución. Una recopilación de proteínas de las que se conoce su estructura 3D se puede encontrar en la base de datos del Dr. Stephen White en [http://blanco.biomol.uci.edu/Membrane\\_Proteins\\_xtal.html](http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Una de las estrategias que se ha mostrado de gran utilidad en la obtención de este tipo de estructuras es la cocrystalización de la proteína en presencia de fragmentos de anticuerpos recombinantes. Esta técnica aumenta las regiones polares de contacto proteína-proteína, favoreciendo la formación de cristales (revisado en (Hunte C, 2002)), lo que resulta de especial utilidad especialmente para aquellas proteínas formadas por regiones transmembrana en las que los lazos citoplasmáticos son muy pequeños.

Entre las alternativas que han surgido para el estudio estructural de proteínas de membrana se encuentran las *técnicas de RMN en estado sólido*, que requieren el marcaje isotópico de la muestra, y en las que las proteínas de membrana son reconstituidas con relativa facilidad en una bicapa fosfolípida. Ya que las interacciones en estado sólido dependen de la orientación, se han desarrollado dos estrategias para el uso de estas técnicas. La primera aproximación es el RMN de ángulo mágico en el que para la obtención del espectro, la muestra gira en torno a un eje con un ángulo de  $54.7^\circ$  respecto al campo magnético, consiguiéndose bandas de resonancia más estrechas y aumentando la resolución. La segunda estrategia consiste en realizar los experimentos de RMN sobre muestra previamente orientadas, bien sea mediante técnicas magnéticas o mecánicas. En cualquier caso, el uso de estas técnicas dista mucho de ser, hoy por hoy, una metodología sencilla que permita la obtención de estructuras con relativa facilidad (*Lipid Bilayers Structure and interactions. J. Katsaras, T. Gutberlet (Eds). Biological physics series. Springer*). Es por ello que se han desarrollado toda una serie de técnicas alternativas que aportan datos estructurales sobre este tipo de proteínas (revisadas en (Torres et al., 2003)). Así han obtenido algunas estructuras tridimensionales de

complejos utilizando técnicas de *difracción de electrones* sobre cristales 2D, sin embargo, la obtención de estos cristales también es complicada. La *crioelectromicroscopía*, puede rendir mejores resultados pero sin alcanzar la elevada resolución que se obtiene por difracción de rayos X. La *microscopía de fuerza atómica* también puede proporcionar información estructural pero sólo de superficies, aunque la mayor ventaja de esta técnica es que las medidas pueden ser tomadas en solución y es posible repetirlas varias veces sobre la misma muestra, de modo que pueden ser observados cambios conformacionales, aunque en cualquier caso, hablamos de resolución laterales de en torno a 10 Å. La *resonancia paramagnética de electron* (EPR) en la que una especie paramagnética (generalmente un grupo nitróxido) es incorporada a la molécula proteica, generalmente por unión a la cadena lateral de uno o más residuos de cisteína también ha sido utilizada para la obtención de datos estructurales de este tipo de proteínas. La señal de absorción EPR depende del ambiente y se ve afectada por la presencia de otras señales paramagnéticas. Esta técnica puede aportar información de orientación rotacional entre hélices alfa, estructura secundaria e incluso tamaño de oligómeros. Sin embargo, su principal problema es el uso de un marcaje muy voluminoso y flexible, lo que puede afectar a la estructura nativa de la proteína, así como a la resolución.

La inclinación de hélices transmembrana incorporadas uniaxialmente en muestras orientadas así como la orientación rotacional interhelicoidal pueden ser estudiadas utilizando técnicas de *infrarrojo de sitio específico*, que monitorizan la absorción infrarroja de ciertos enlaces, como por ejemplo el enlace amida C=O, en un péptido helicoidal. La razón dicroica de esa banda de absorción (relación entre la absorción paralela y perpendicular) se relaciona con la orientación de dicho enlace en el espacio. Para utilizar esta técnica es necesario marcar el enlace a estudiar con un isótopo. Generalmente se utiliza el marcaje doble  $^{13}\text{C}=^{18}\text{O}$  y los datos obtenidos se complementan con simulaciones de dinámica molecular (Torres et al., 2001).

Desde el punto de vista computacional la resolución de estructuras de proteínas de membrana se aborda utilizando diferentes estrategias. Las *simulaciones de dinámica molecular de búsqueda global* consisten en un método computacional que explora las interacciones interhelicoidales entre haces de hélices oligoméricas, los resultados que genera esta técnica suelen ser un grupo de modelos de baja energía, de modo que los datos experimentales deben complementar esta estrategia a la hora de fijar un modelo (Torres et al., 2002). Aunque la incorporación de datos sobre conservación evolutiva de residuos está permitiendo, en algunos casos, obviar la incorporación de resultados

experimentales. También en los últimos años han ido apareciendo las primeras funciones para la predicción de estructura específicas para proteínas de membrana (Fleishman & Ben-Tal, 2002). Por otra parte, se han desarrollado algoritmos que permiten la predicción de la topología de una hélice transmembrana generalmente basados en alineamientos de secuencia (Jayasinghe et al., 2001; Krogh et al., 2001).

Los estudios de equilibrios de sedimentación por *centrifugación analítica* que se han utilizado con mucha frecuencia para el análisis del estado oligomérico de proteínas solubles, han sido también adaptados para el estudio de proteínas de membrana utilizando micelas de detergente (revisado en, (DeGrado et al., 2003). Otra de las técnicas utilizadas para determinar el estado de agregación son experimentos de fluorescencia en los que se mide la *transferencia de energía por resonancia inducida* entre péptidos transmembrana marcados con sondas diferentes de modo que sus espectros de emisión y absorción se solapan (Fisher et al., 1999). También ha sido de gran utilidad para el estudio del estado oligomérico de ciertas proteínas el uso de *electroforesis de poliacrilamida en presencia de SDS*, ya que numerosas proteínas de membrana mantienen su estado oligomérico en presencia de este detergente (Arkin, 2002; Bowie, 1997).

En el proceso de asociación lateral entre hélices transmembrana, segunda etapa del plegamiento de proteínas de membrana según el modelo de plegamiento expuesto previamente, intervienen factores tanto entálpicos como entrópicos que afectan no sólo a las interacciones proteína-proteína ( $\Delta G_{pp}$ ), sino también a las interacciones proteína-lípido ( $\Delta G_{pl}$ ) y lípido-lípido ( $\Delta G_{ll}$ ), ( $\Delta G_{HH} = \Delta G_{pp} + \Delta G_{pl} + \Delta G_{ll}$ ), (Fleming & Engelman, 2001). Sin embargo, se ha demostrado, (Fleming & Engelman, 2001) que la relación entre los datos obtenidos para  $\Delta G_{pp}$  en diferentes medios, miméticos de membrana, es lineal, lo que nos permite el estudio de los factores implicados en la asociación entre hélices en medios miméticos de membrana, esperando un comportamiento cualitativamente comparable al que obtendríamos en ambientes naturales.

En todos los métodos descritos anteriormente, se trabaja con la proteína *in vitro*, sin embargo, en los últimos años, se han desarrollado una serie de sistemas que permiten monitorizar la asociación entre proteínas de membrana *in vivo*. Entre ellos se encuentra el sistema TOXCAT, en el que la homodimerización del factor transcripcional ToxR en la membrana de *E. coli*, mediada por un segmento transmembrana heterólogo, es capaz de activar la expresión del gen marcador CAT (cloramfenicol acetiltransferasa) (Russ & Engelman, 1999). De este modo, sólo presentan resistencia

a cloramfenicol aquellas cepas en las que se produce la dimerización de la proteína. Este sistema incorpora además un marcador extracitoplasmático para determinar la correcta topología, la proteína de unión a lactosa (MBP). MBP sólo es activa en el periplasma, de modo que es capaz de complementar a una cepa con deficiencia en esta proteína únicamente si la inserción es adecuada. Un sistema similar, es el sistema GALLEX (figura 1.8), que permite medir heteroasociación de proteínas de membrana, mediante la fusión de estas a dos dominios diferentes de unión a DNA de la proteína LexA (Schneider & Engelman, 2003). En este método la proteína que permite determinar la topología es, de nuevo, MBP, y el gen marcador Lac Z, responsable de la actividad  $\beta$ -galactosidasa.

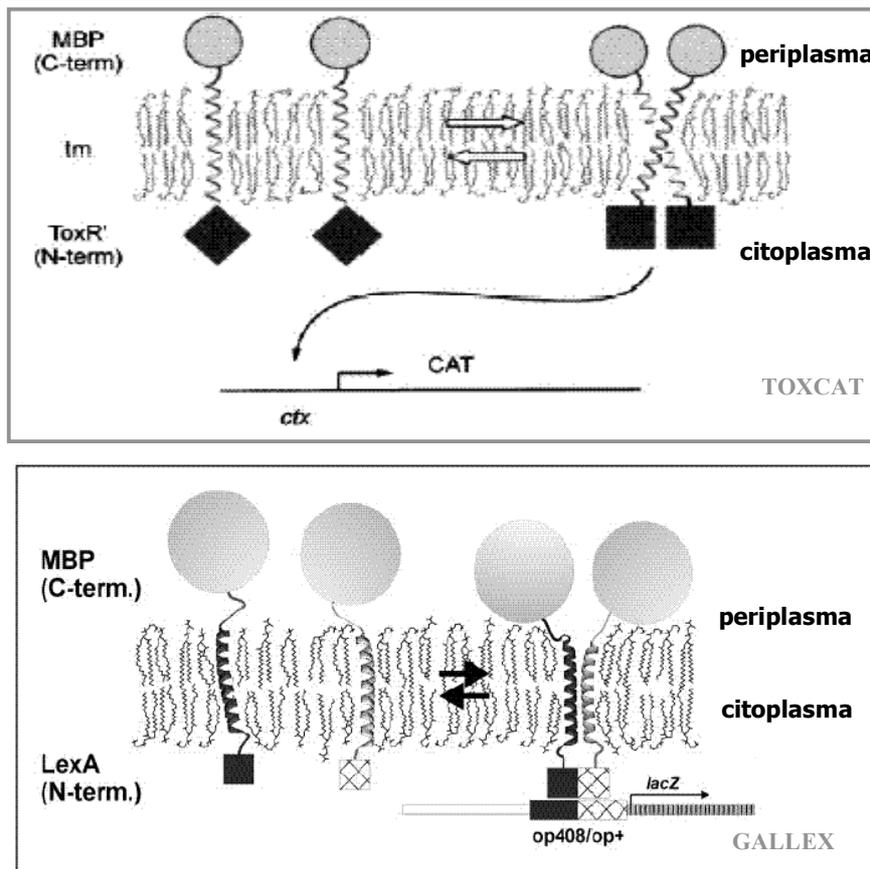


Figura 1.8: Ensayos TOXCAT y GALLEX para la medida de interacciones TM *in vivo*. El dominio transmembrana ancla a la proteína quimera a la membrana interna de *E.coli* con el dominio C-terminal orientado al periplasma (MBP, maltose binding protein) y el dominio de unión al DNA de ToxR o Lex A en N-terminal orientado al citoplasma. La interacción entre los fragmentos TM da lugar a la formación de un dímero con capacidad de unión al operador del gen marcador CAT o LacZ. Adaptado de (Russ & Engelman, 1999), (Schneider & Engelman, 2003).

## 1.6. Proteínas modelo en el estudio de la asociación de proteínas transmembrana.

La mayor parte de los datos estructurales obtenidos hasta la actualidad en el estudio de proteínas de membrana son resultado del uso de proteínas modelo para la obtención de información extrapolable a sistemas más complejos. A continuación, se describen con más detalle, algunos de estos sistemas modelo.

### Canal de protones M2.

La proteína M2 de *Influenza A* (gripe común), es una pequeña proteína integral de membrana (96 residuos) que interviene en el ciclo vital de este virus de RNA, responsable de afecciones respiratorias en humanos. Esta proteína, atraviesa una única vez la membrana y forma tetrámeros unidos por puentes disulfuro. Actúa como un canal iónico durante el proceso de desensamblaje del virus en los endosomas, permitiendo el flujo de protones al interior de las partículas víricas, lo que facilita la desaparición de las interacciones proteína-proteína. Este canal se inhibe de forma específica por amantadina y se activa a pHs bajos (Salom et al., 2000), lo que parece indicar que sólo estará activo en los endosomas y en la red trans Golgi. Las cisteínas no afectan a la actividad de la proteína y se ha visto que péptidos correspondientes al segmento transmembrana de esta proteína son capaces de mimetizar su actividad de canal. El modelo de activación del canal que forma M2 se

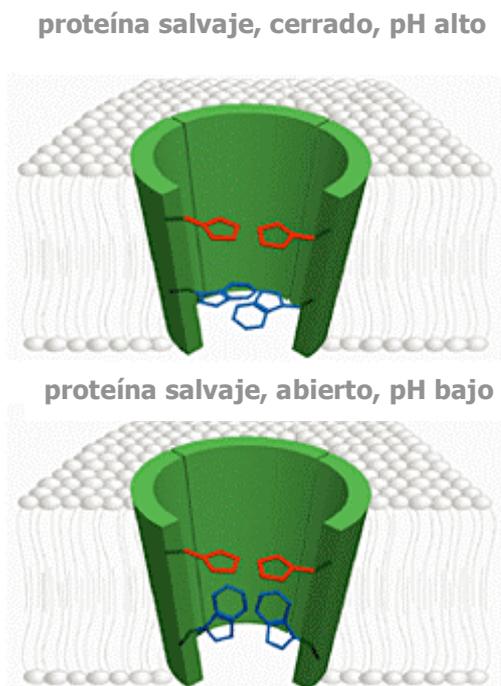


Figura 1.9: Modelo de apertura y cierre del canal de protones M2. (a). Canal cerrado. Cuando el pH es elevado las His37 no tienen carga de modo que los anillos de la cadena lateral del Trp 47, se sitúan en el plano de la bicapa como se observa en la figura y no permiten el paso de protones. (b) Canal abierto. Cuando el pH disminuye las His recuperan la carga de modo que las cadenas laterales de los Trp se disponen perpendicularmente permitiendo el flujo de iones. Figura obtenida del laboratorio de R.A.Lamb. (<http://www.hhmi.org/research/investigators/lamb.html>)

muestra en la figura 1.9, en la que se destacan dos residuos (His37, Trp 47), que parecen ser los responsables de su apertura.

Esta proteína ha sido expresada purificada y reconstituida en liposomas, marcada con ( $^{15}$ ) N, lo que ha permitido realizar experimentos de RMN en estado sólido (Wang et al., 2001),(Tian et al., 2002), que apoyan su estructura tetramérica e indican una inclinación de las hélices con respecto a la perpendicular a la bicapa en torno a  $25^\circ$ .

### Fosfolamban

Fosfolamban es una pequeña proteína de 52 aminoácidos que se encuentra en la membrana del retículo sarcoplásmico, donde se encarga de la inhibición de la  $\text{Ca}^{2+}$ -ATPasa, durante la contracción del músculo cardíaco. La fosforilación de fosfolamban, evita esta inhibición, permitiendo que se restaure la capacidad de establecer un gradiente de  $\text{Ca}^{2+}$  en el retículo sarcoplásmico mediado por dicha bomba. Desde el punto de vista estructural, se ha propuesto que fosfolamban es una proteína homopentamérica (figura 1.10) (Simmerman et al., 1996; Torres et al., 2000) , cuya oligomerización parece estar mediada únicamente por residuos de su segmento transmembrana (Arkin et al., 1994), utilizando un modelo de plegamiento similar a los “*coiled coils*” de proteínas solubles, en el que aparecen los aminoácidos leucina e isoleucina en posiciones  $i$  e  $(i + 3)$  con una periodicidad de 7 residuos ( $\text{L}_{37}\text{xxIxxxLxxI}_{47}$ ), donde la x simboliza cualquier aminoácido hidrofóbico no implicado en la oligomerización. La fosforilación de esta proteína favorece la estabilidad de la estructura pentamérica, lo que podría estar relacionado con su mecanismo de regulación de la  $\text{Ca}^{2+}$ -ATPasa puesto que se ha observado que la unión a esta proteína la lleva a cabo en su forma monomérica.



Figura 1.10: Modelo de la región transmembrana de fosfolamban, en el que se muestra la asociación entre sus hélices transmembrana, en visión lateral (izquierda) y superior (derecha). Adaptado de (Torres et al., 2000)

## Glicoforina A

Glicoforina A es una sialoglicoproteína que se encuentra anclada en la membrana de eritrocitos a través de un único segmento transmembrana. Existen unas 500.000 copias de esta proteína por célula. Se trata de una proteína de tipo I, es decir, su extremo N-terminal es extracitoplasmático. Aunque su función aun no está demasiado clara, a pesar de haber sido la primera proteína transmembrana purificada y secuenciada (Tomita et al., 1978), se conocen algunos datos al respecto, como son, su implicación en la determinación de los subgrupos sanguíneos MN, o su papel como receptor en el proceso de infección mediado por *Plasmodium falciparum* parásito causante de la malaria (Ockenhouse et al., 2001). Además esta proteína se ha utilizado como marcador de la exposición a radiación de los supervivientes de la bomba atómica ya que se demostró una relación lineal entre el nivel de exposición a radiación y la acumulación de mutaciones en Glicoforina A, por lo que también se ha propuesto su uso como marcador tras tratamientos mutagénicos como la quimioterapia (Grant & Bigbee, 1994).

Esta proteína posee la capacidad de formar dímeros a través de su segmento transmembrana (Lemmon et al., 1992a). Desde que se conoció su secuencia se han realizado numerosos estudios con el objetivo de comprender como se produce este proceso de autoasociación entre hélices transmembrana. Experimentos de mutagénesis por saturación (Lemmon et al., 1992b), modelado molecular (Lemmon et al., 1994), inserción de alaninas (Mingarro et al., 1996) y finalmente la resolución estructural por NMR del segmento TM de esta proteína en presencia de micelas de dodecil-fosfocolina (MacKenzie et al., 1997), determinaron un *motivo de dimerización* formado por siete residuos del segmento transmembrana como responsables de la autoasociación de la proteína (<sup>75</sup>L<sup>76</sup>Ixx<sup>79</sup>G<sup>80</sup>Vxx<sup>83</sup>G<sup>84</sup>Vxx<sup>87</sup>T) (figura 1.11).

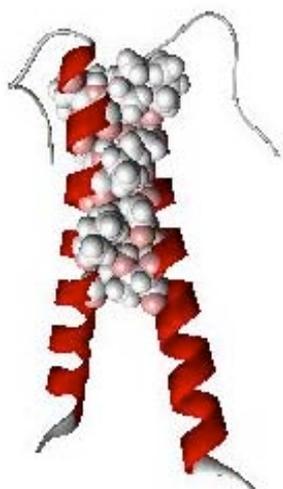


Figura 1.11: Estructura (código PDB,1AFO) del dímero del segmento transmembrana de GpA en micelas. Los residuos implicados en la dimerización (ver texto) se muestran resaltados como bolas con sus radios de Van der Waals.

En esta estructura, en la que las cadenas laterales de algunos aminoácidos se disponen como salientes proteicos que se incorporan a cavidades de la propia hélice, la cadena lateral de la valina 79 de cada uno de los monómeros ocuparía el hueco que deja la glicina 80 del monómero complementario y del mismo modo ocurriría en el caso de la glicina 83 y la valina 84, generando una estructura estrechamente empaquetada en esa región de la proteína. El surco creado por las glicinas es continuado por las cadenas laterales de la leucina 75 en N-terminal y la treonina 87 en C-terminal, mientras que la cadena lateral de la isoleucina 76 se comporta como parte de los salientes. El hidroxilo de la cadena lateral de la treonina 87 formaría un puente de hidrógeno con el carbonilo de la glicina 83 dentro del mismo monómero. De este modo la estructura dimerica de GpA se mantendría únicamente gracias a la existencia de interacciones de tipo *van der Waals*, generándose un ángulo entre hélices de  $-40$  grados.

Posteriormente, se resolvió la estructura de este mismo segmento utilizando NMR de ángulo mágico (Smith et al., 2001), usando membranas de dimiristoil-fosfolina y 1-palmitoil-2-oleoil-fosfolina lo que permitió refinar los tipos de interacciones existentes en la estructura. La mayor aportación de este trabajo consistió en la determinación de una rotación en la cara de interacción entre las hélices que difiere en 25 grados a la determinada por NMR en solución, este hecho provoca que los pares de residuos glicina 79-glicina 79 y glicina 83-glicina 83 queden enfrentados en la estructura, lo que permite interacciones de *van der Waals* entre ellos. Además, sitúa al residuo treonina 87 en disposición de formar puentes de hidrógeno interhelicoidales con la valina 84 de la hélice adyacente, cuya existencia ha sido posteriormente corroborada (Smith et al., 2002). Por otra parte, el ángulo hélice-hélice también queda ligeramente modificado, probablemente como consecuencia de las diferencias en los protocolos experimentales.

El hecho de que GpA tenga un único segmento transmembrana pero sea capaz de dimerizar gracias a él, la convierte en un modelo sencillo para el estudio del proceso de empaquetamiento hélice-hélice. Desde el punto de vista experimental también se puede considerar un modelo apropiado. En general, la sobreexpresión de proteínas transmembrana en sistemas heterólogos resulta muy complicada debido a su evidente toxicidad. Sin embargo, el fragmento transmembrana de GpA, fusionado a una proteína soluble, como es la nucleasa de *Staphylococcus*, se expresa y purifica eficazmente en un sistema heterólogo, como es *Escherichia coli*, manteniendo sus características de dimerización. Por otro lado, la extrapolación en el uso de técnicas habituales para el estudio de la

estructura de proteínas solubles a proteínas transmembrana no es posible, como se ha mencionado anteriormente, en la mayoría de los casos. Sin embargo, la dimerización de GpA es resistente a la presencia de detergentes como el SDS (dodecil sulfato sódico), lo que permite el análisis del grado de dimerización en geles de electroforesis de poliacrilamida desnaturizantes.

Por todos estos motivos GpA, es un modelo completo a la vez que sencillo que ha permitido aportar datos relevantes para la comprensión de los elementos que contribuyen al plegamiento e interacción de proteínas transmembrana. En este trabajo, se ha utilizado dicho sistema modelo para el estudio de algunos de los factores que contribuyen a este proceso.

### **1.7. Tipos de interacciones entre hélices transmembrana.**

La diferencia fundamental en el proceso de plegamiento entre las proteínas solubles y las helicoidales de membrana es consecuencia principalmente del medio en el que se encuentran. La formación de estructura secundaria en ambientes acuosos está desfavorecida, de modo que el proceso de plegamiento y asociación están íntimamente ligados y dirigidos principalmente por el efecto hidrofóbico. Sin embargo, en proteínas de membrana podríamos hablar de un efecto “*solvofóbico*” que dirige las cadenas laterales de los residuos hidrofóbicos hacia la membrana, donde inmediatamente se forma la estructura secundaria en hélice alfa que minimiza la exposición de grupos polares del enlace peptídico al medio (primera etapa del modelo de plegamiento en dos etapas). Después de este proceso de inserción, ya en el interior de la membrana, se produce la asociación entre las diferentes hélices preformadas (segunda etapa del modelo). Entre los factores energéticos que dirigen esta segunda etapa no se encuentra, como es evidente, el efecto hidrofóbico, tan importante en el plegamiento de proteínas solubles y los factores que determinan este proceso son hoy en día objeto de estudio. Se han revelado de gran importancia en esta etapa del plegamiento las interacciones de tipo *van der Waals*, aunque en los últimos años el papel de los *puentes de hidrógeno* en el establecimiento de la estructura final de estas proteínas parece consolidarse y, al tiempo que se incrementa el número de estructuras de alta resolución de este tipo de proteínas, aumenta también el número de este tipo de interacciones descritas. Así se ha definido la existencia de puentes de hidrógeno de dos tipos, aquellos que se establecen entre la cadena lateral de un residuo y el nitrógeno o el oxígeno del esqueleto peptídico y los que se establecen entre cadenas laterales de aminoácidos

diferentes, entre los que cabe destacar principalmente la contribución de los residuos de Ser, Tyr, Thr e His que participan en casi el 50% del número total de este tipo de interacciones observadas en proteínas de membrana (Adamian & Liang, 2002). Por otra parte, también se ha demostrado que los puentes de hidrógeno del tipo  $C_{\alpha}-H\cdots O$  pueden contribuir a la estabilidad de oligómeros en proteínas de membrana (Senes et al., 2001). La introducción de residuos de Asn, Asp, Gln, Glu e His promueve la asociación entre fragmentos transmembrana artificiales probablemente también mediante la formación de puentes de hidrógeno (Gratkowski et al., 2001; Zhou et al., 2000) (Choma et al., 2000; Zhou et al., 2001) y, como se ha mencionado anteriormente, ha sido descrita la formación de una interacción de este tipo que promueve la asociación interhelicoidal en el dímero de GpA entre el grupo hidroxilo de la treonina 87 y el carbonilo de la valina 84 en la cadena complementaria (Smith et al., 2002). No hay que olvidar que la formación de puentes de hidrógeno se puede ver más favorecida en ambientes hidrofóbicos donde no se encuentran solvatados los grupos que van a estar implicados en esta interacción y por tanto su deshidratación previa a la formación del enlace no es necesaria (DeGrado et al., 2003; Zhou et al., 2000).

Sin embargo, no se ha descrito, al menos de forma generalizada, la formación en ambientes de membrana de interacciones del tipo puentes disulfuro o puentes salinos. Este hecho parece estar relacionado con el nivel de promiscuidad y pérdida de la capacidad de modulación que supondría la presencia de este tipo de interacciones fuertes entre hélices TM, que podrían conducir a la asociación no específica y probablemente irreversible entre ellas (Fleming & Engelman, 2001; Zhou et al., 2000; Zhou et al., 2001).

Otra de las diferencias entre los determinantes que dirigen la oligomerización en proteínas solubles y los que intervienen en el caso de proteínas transmembrana es el hecho de que la pérdida de entropía como consecuencia de este proceso de asociación interhelicoidal, es significativamente menor en proteínas de membrana, por diversas razones. En primer lugar, la inserción previa de los monómeros predetermina su orientación, definiendo de antemano las interacciones paralelas *versus* las antiparalelas. Además, se ha observado la presencia de un gran número de residuos  $\beta$ -ramificados en las interfases de interacción (Senes et al., 2000) (Liu et al., 2002), los cuales tienen limitados sus posibles rotámeros en estructuras helicoidales, de modo que, de nuevo, la pérdida de entropía como consecuencia de la oligomerización queda minimizada.

### Motivos de oligomerización

Entre las secuencias proteicas que favorecen la asociación lateral entre hélices transmembrana se han descrito diversos motivos capaces de inducir oligomerización. Entre ellos se encuentran las *cremalleras de serina* (figura 1.12) (Adamian & Liang, 2002) en las que se establecen interacciones del tipo cadena lateral-cadena principal entre dos residuos de serina, con una periodicidad de 7 aminoácidos como se muestra en la figura 1.12. Por otra parte, en experimentos en los que se seleccionan secuencias generadas al azar por su capacidad de oligomerización también ha sido detectados motivos en los que intervienen residuos de Ser y Thr del tipo SxxSSxxT y SxxxSSxxT, que posteriormente han sido localizados en proteínas de membrana naturales (Dawson et al., 2002).

Se han descrito motivos similares a las *cremalleras de leucina* de proteínas solubles, en los que con un patrón típico de siete residuos (*abcdefg-abcdefg*), aparecen aminoácidos hidrofóbicos en posiciones *a* y *d*. Algunos ejemplos de ello están presentes en las proteínas sinaptobrevina y syntaxina (Laage et al., 2000). Del mismo modo que en el caso anterior, también han sido simulados en fragmentos transmembrana artificiales en los que mediante evolución *in vitro* se han detectado

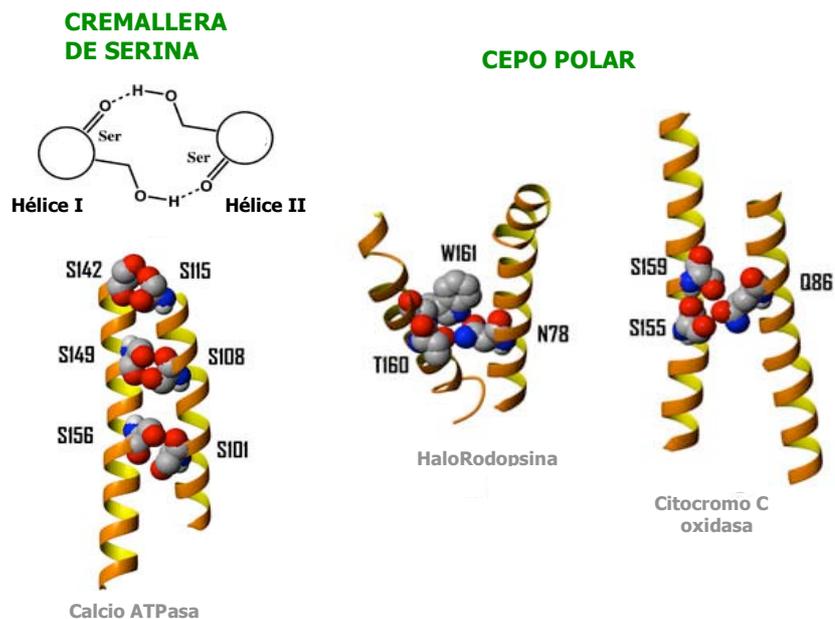


Figura 1.12: Motivos de interacción entre hélices transmembrana.

Como posibles candidatos a ser motivos de interacción generales entre proteínas de membrana (Gurezka & Langosch, 2001).

Otro patrón de interacción de hélices TM es el *cepo polar* (figura 1.12) que se establece entre tres residuos de dos hélices diferentes con dos puentes de hidrógeno interhelicoidales, (Adamian & Liang, 2002). En la mayoría de los casos la cadena lateral de uno de los residuos forma dos puentes de hidrógeno con las cadenas laterales o los átomos de la cadena principal de otros dos residuos (figura 1.12).

Entre los elementos que cabe destacar en el fragmento transmembrana de GpA está la presencia del *motivo de oligomerización* mínimo *GxxxG* (Brosig & Langosch, 1998). Tras la resolución estructural del dímero de GpA el empaquetamiento hélice-hélice que permitía el posicionamiento de las glicinas del motivo se propuso como una posible estrategia más general de oligomerización entre fragmentos transmembrana. De hecho, este motivo se encontró en diversas proteínas transmembrana implicado también en procesos de oligomerización. Así, está presente en la proteína de cubierta del virus M13 (Williams et al., 1995), en el proteoglicano heparan sulfato N-sindecan (Asundi & Carey, 1995), así como en el segmento transmembrana VI del receptor  $\beta_2$ -adrenérgico (Gadbut et al., 1996), en la subunidad  $\epsilon$  de la ATP sintasa mitocondrial de levadura (Arselin et al., 2003). Se han realizado numerosos estudios con el objetivo de comprobar si este motivo se puede considerar secuencia consenso para la interacción hélice-hélice entre fragmentos TM. Se han obtenido desde datos indirectos que relacionan el mayor empaquetamiento observado en proteínas de membrana respecto a sus homólogos solubles, con la abundancia relativa de residuos pequeños en las interfases de interacción (Eilers et al., 2000) (Jiang & Vakser, 2000), datos estadísticos que demuestran la abundancia de las glicinas (Bywater et al., 2001), así como de los pares *GxxxG* en proteínas de membrana (Senes et al., 2000), hasta experimentos de detección de oligomerización con secuencias al azar en las que el motivo *GxxxG* ha sido el que ha aparecido con una mayor frecuencia (Russ & Engelman, 1999). Una de las causas principales de la contribución de este motivo al proceso de empaquetamiento, es el hecho de que la ausencia de cadena lateral de los residuos de glicina favorece el acercamiento entre los esqueletos de las hélices enfrentadas, favoreciendo la formación de puentes de hidrógeno del tipo  $C\alpha-H\cdots O$  (Senes et al., 2001). El carbono alfa es un átomo activado como dador de electrones debido a su unión con los grupos N-H y C=O. Aunque los puentes de hidrógeno del tipo  $C\alpha-H\cdots O$  se han observado en proteínas solubles su

contribución a la estructura no parece ser importante. Sin embargo, el cálculo de la energía de este tipo de enlaces en vacío (Scheiner et al., 2001), otorga un valor de 2,5-3,0 kcal/mol a este enlace, lo que equivale a la mitad de la energía de un puente de hidrógeno convencional N-H $\cdots$ O. Estos valores pueden ser de gran importancia en ambientes hidrofóbicos, ya que se trata de interacciones de tipo electrostático. Este tipo de interacciones se han encontrado no sólo en interfases ricas en residuos de glicina, donde están favorecidas las interacciones C $\alpha$ -H $\cdots$ O=C (esqueleto-esqueleto), sino también ricas en serinas y treoninas donde los grupos hidroxilo de las cadenas laterales permiten la interacción a distancias más largas. En los últimos años el motivo GxxxG también se ha revelado como un motivo de oligomerización entre hélices de proteínas solubles (Kleiger et al., 2002). En definitiva todos estos estudios validan la existencia del motivo GxxxG como un motivo de oligomerización general, al menos para proteínas de membrana. Este hecho contribuye a la consideración de GpA como proteína modelo para el estudio del empaquetamiento entre hélices transmembrana.

### **1.8. Objetivos.**

1. Averiguar el papel que desempeña la región C-terminal del segmento TM de GpA, en el que no se encuentran presentes residuos del motivo de dimerización descrito para esta proteína, en el proceso de empaquetamiento hélice- hélice.
2. Estudiar la contribución individual de los residuos del motivo de dimerización de GpA al proceso global de empaquetamiento hélice-hélice, con el objetivo de obtener un motivo minimizado y comprobar la contribución que la longitud del fragmento hidrofóbico desempeña en dicho proceso.
3. Estudiar las consecuencias que, sobre el empaquetamiento entre hélices transmembrana, provoca la presencia de residuos de prolina.

## **2. Materiales y métodos**



## **2.1 Manipulación DNA y cepas bacterianas**

### **2.1.1. Obtención de DNA plasmídico**

Para la obtención de plásmido a pequeña escala de las diferentes construcciones se ha utilizado el *QIAprep Spin Miniprep Kit* de la casa comercial Qiagen.

### **2.1.2. Electroforesis en gel de agarosa**

Las electroforesis de DNA se han llevado a cabo en un sistema de *Minigeles (Mini Sub Cell GT, Biorad, USA)* utilizando geles de agarosa (*Sigma Aldrich, Germany*) al 1% en tampón TAE (40mM Tris acetato, 2mM Na<sub>2</sub>EDTA, pH 8.5), con 10µg/mL de Bromuro de Etidio (*Sigma Aldrich, Germany*), a 100 voltios durante aproximadamente 30 min.

### **2.1.3. Reacciones de PCR**

Las reacciones de PCR (reacción en cadena de la polimerasa) se realizaron en un termociclador personal Eppendorf Mastercycler. Se utilizó una concentración de oligonucleótidos cebadores de 1µM, unos 5 ng de plásmido molde, 2,5mM de MgCl<sub>2</sub>, 0,2mM de dNTPs y 1 unidad de Taq polimerasa (*Biotaq<sup>TM</sup> DNA Polymerase, Boline, London*) en el tampón proporcionado por la casa comercial. Se realizó un primer paso de desnaturalización 94°C durante 2 min, tras lo que se realizaron 30 ciclos de, 94°C durante 2 min, 2 min a 55°C y 72°C durante el tiempo necesario para completar 2 min por kilobase de DNA a amplificar. Tras estos 30 ciclos se realizó un último paso de amplificación a 72°C durante 2 min. El resultado de la PCR fue purificado utilizando el Kit *QIAquick® PCR Purification (Qiagen, Germany)*. En reacciones en las que se amplificó un fragmento pequeño (100pb) o se observaron bandas adicionales se utilizó la purificación en gel de agarosa utilizando *QIAquick® Gel Extraction kit (Qiagen, Germany)*.

### **2.1.4. Digestión con enzimas de restricción**

Para la digestión del DNA se han utilizado enzimas de restricción de *Boehringer Mannheim, (Alemania)* siguiendo las instrucciones de la casa comercial, salvo para aquellos casos en los que la

diana de restricción estaba en el extremo de la secuencia. En estos experimentos se aumentó la proporción enzima/DNA hasta 10 veces y las digestiones se llevaron a cabo en tiempos de 5-6 h.

#### 2.1.5. Condiciones de crecimiento de *E. coli*.

El crecimiento de las células de *Escherichia coli*, se ha llevado a cabo a 37°C, en medio líquido LB (*Luria Bertani*) que contiene 0.5% de extracto de Levadura (*Pronadisa, Spain*), 1% de Triptona bacteriológica (*Pronadisa, Spain*) y 1% de NaCl o en medio LB sólido con 1,5% de Agar bacteriológico (*Pronadisa, Spain*). El medio fue suplementado en los casos en los que fue necesario con los antibióticos Ampicilina (Roche) a 100µg/mL y cloramfenicol (Roche) 50µg/mL. En el crecimiento de la cepa BL21 pLys (DE3) para la expresión de proteína se utilizó medio TB (*Terrific Broth*) (12g/L triptona, 24g/L extracto de levadura, 4mL/L glicerol, 2.31g KH<sub>2</sub>PO<sub>4</sub>, 12.54 g K<sub>2</sub>HPO<sub>4</sub>).

Todos los medios de cultivo, así como todo el material de trabajo fueron autoclavados durante 20 min a 1 atmósfera (121°C).

#### 2.1.6. Protocolo de transformación de células competentes de *E. coli*

En este trabajo se han utilizado fundamentalmente tres cepas de *E.coli* diferentes MC1061, BL21 pLyS(DE3) y XL1Blue. La obtención de células competentes se ha realizado, en el caso de las cepas MC1061 y BL21pLys DE3, utilizando el método del Cloruro de Rubidio (Current protocols in molecular biology, Jon Wiley & Sons, Inc.). Para la cepa XL1Blue se utilizaron células supercompetentes (*Epicurian Coli® Supercompetent cells, Stratagene, Germany*). El protocolo de transformación utilizado es en todos los casos un protocolo de choque térmico (Hanahan, 1983) pero los tiempos, así como la temperatura de choque térmico varía para cada una de las cepas según los datos que se muestran en la tabla 2.1.

Tabla 2.1: Condiciones de transformación de las diferentes cepas de *E.coli*

Cepa de <i>E. coli</i>	Temperatura(°C)/ Tiempo (s)	Tiempo en hielo (min)
MC1061	42/90	2
XL1Blue	42/45	2
BL21 pLyS DE3	37/120	5

## 2.2. Vectores utilizados y diseño de construcciones.

### 2.2.1. Vectores de GpA

Para obtener las distintas construcciones de SN/ GpA que se han utilizado en este trabajo se ha partido de un plásmido descrito por (Lemmon et al., 1992a) que contiene los residuos del 73 al 131 de GpA, precedidos por la secuencia de la nucleasa de *Staphylococcus* (SN) y una cola de 6-His en el extremo C-terminal (Mingarro et al., 1996) (figura 2.1). Esta construcción se sitúa entre los sitios *NdeI/ BamHI* y bajo el control del promotor T7 en el plásmido del pET11a de Novagene. A lo largo de este trabajo se nombrará este plásmido como pT7SN/GpA (figura 2.2). Además se dispuso del mutante 5L (Mingarro et al., 1997), que posee 5 residuos de Leucina entre los aminoácidos metionina 81 y alanina 82 del segmento TM de GpA.

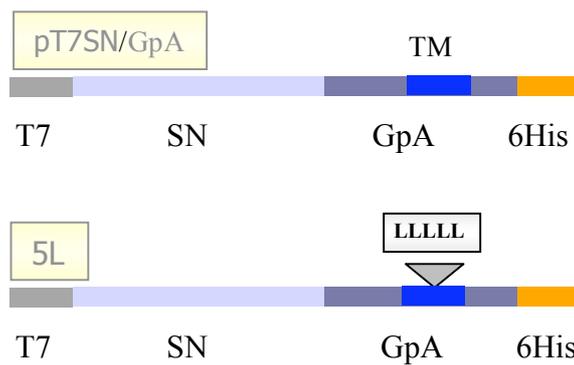


Figura 2.1: Representación esquemática de las construcciones derivadas de pET 11a.

```

1   CATATGACTTCAACTAAAAAATTACATAAAGAACCTGCGACTTTAATTAAGCG
1   H M T S T K K L H K E P A T L I K A

55  ATGATGGTGATACGGTTAAATTAATGTACAAAGGTCAACCAATGACATTCAGA
19  I D G D T V K L M Y K G Q P M T F R

109 CTATTATTGGTTGATACACCTGAAACAAAGCATCCTAAAAAAGGTGTAGAGAAA
37  L L L V D T P E T K H P K K G V E K

163 TATGGTCTGAAGCAAGTGCATTTACGAAAAAATGGTAGAAAAATGCAAAGAAA
55  Y G P E A S A F T K K M V E N A K K

217 ATTGAAGTCGAGTTTGACAAAGGTCAAAGAAGTATAAATATGGACGTGGCTTA
73  I E V E F D K G Q R T D K Y G R G L

271 GCGTATATTTATGGTGATGGAAAAATGGTAAACGAAGCTTTAGTTCGTCAAGGC
91  A Y I Y G D G K M V N E A L V R Q G

325 TTGGCTAAAGTTGCTTATGTTTACAAACCTAACAAATACACATGAACAACATTTA
109 L A K V A Y V Y K P N N T H E Q H L

379 AGAAAAAGTGAAGCACAGCGAAAAAAGAGAAATTAATATTTGGAGCGAAGAC
127 R K S E A Q A K K E K L N I W S E D

433 AACGCTGATTCGGGCCAGAAAGGGTACAACCTGCCCATCATTTCTCTGAACCA
145 N A D S G P E R V Q L A H H F S E P

487 GAGATAACACTCATTATTTTGGGGTGATGGCTGGTGTATTGGAACGATCCTC
163 E I T L I I F G V M A G V I G T I L

541 TTAATTTCTTACGGTATTGCGCGACTGATAAAGAAAAGCCCATCTGATGTAATA
181 L I S Y G I R R L I K K S P S D V K

595 CCTCTCCCTCACCTGACACAGACGTGCCTTTAAGTTCTGTTGAAATTGAGAAC
199 P L P S P D T D V P L S S V E I E N

649 CCCGAAACTTCGGACCAGCACCATCATCATCACCCTAG
217 P E T S D Q H H H H H H H *

```

Figura 2.2: Secuencia de la construcción pT7 SN /GpA. En posición N-terminal se muestra la secuencia de la nucleasa de *Staphylococcus*, en color verde. En azul se muestran los residuos de GpA y en negrita se destacan aquellos correspondientes a la región transmembrana de GpA. La cola de 6 His utilizada para la purificación, se muestra en negrita en la región C-terminal de la proteína.

### 2.2.2. Vectores de Lep

Los vectores que contienen la proteína Lep (*Leader peptidase de E. coli*) proceden del laboratorio del Prof. von Heijne, (Universidad de Estocolmo, Suecia). Son todos derivados del vector pGEM1-Lep (Nilsson et al., 1994) que contiene el gen Lep bajo el control del promotor SP6 (figura 2.3). Este gen se encuentra modificado de forma que se le ha eliminado un sitio natural de glicosilación <sup>214</sup>NET<sup>216</sup>, introduciendo uno nuevo en una posición que optimiza la eficiencia de glicosilación, <sup>90</sup>NST<sup>92</sup>, a 18

residuos del final del segmento TM H2. Este fragmento transmembrana está flanqueado por los sitios de restricción *Bcl I* y *Nde I*, lo que permitió introducir en esa posición el fragmento TM de GpA. Los diferentes vectores de los que se dispuso proporcionados por el mismo laboratorio poseen el sitio de glicosilación NST situado a distancias  $d=9$  y  $d=11$  de H2.

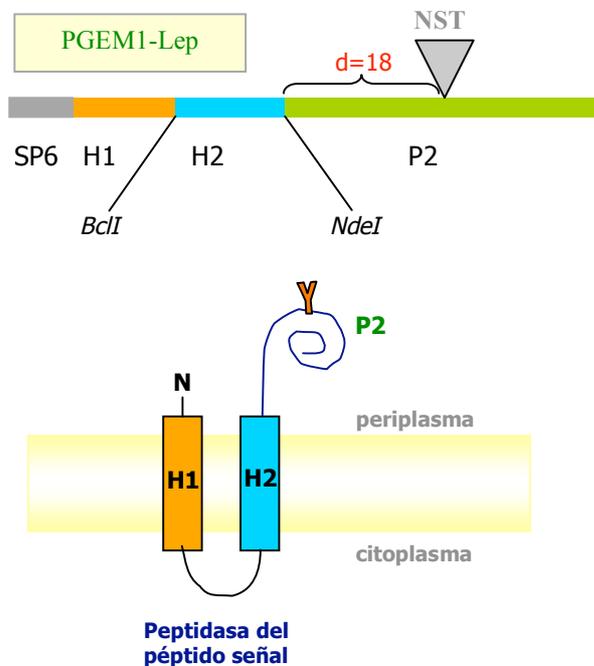


Figura 2.3: Representación esquemática de la construcción pGEM1-Lep que posee dos fragmentos TM, H1 y H2. En azul se muestra el fragmento transmembrana H2, que en las construcciones derivadas se substituye por el segmento TM de GpA. La distancia al sitio de glicosilación NST varía también en los plásmidos derivados. En la parte inferior de la figura se muestra un esquema de cómo se situaría esta proteína una vez insertada en la membrana procariótica.

### 2.2.3. Preparación de mutantes

La preparación de los mutantes de GpA se ha realizado utilizando el kit de mutagénesis (*Quickchange<sup>TM</sup> Site-directed mutagenesis Kit, Stratagene, Germany*). Brevemente, se diseñan oligonucleótidos que aparean en la zona en la que se quiere introducir la mutación (tanto el oligo 5'-3' como el 3'-5') y se realiza una PCR para amplificar el plásmido en el que se desea introducir la mutación utilizando dichos oligos como cebadores de la reacción. El resultado de esta PCR es una

mezcla de plásmido parental (introducido en la reacción como molde) y plásmido mutante. La digestión diferencial del plásmido parental con enzima de restricción *Dpn I* enriquece la muestra en plásmido mutante. Se procede entonces a la transformación en la cepa XL1Blue, utilizando células supercompetentes para obtener una alta eficiencia en la transformación.

El análisis de las colonias obtenidas se lleva a cabo mediante la extracción de plásmido a pequeña escala y su posterior secuenciación (realizada en el servicio de secuenciación del SCSIE, Universitat de València).

Tabla 2.2 Nomenclatura de los diferentes mutantes y secuencia de su segmento TM

Mutante	Secuencia
Wt	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
G79A,V80A	<sup>72</sup> EITLII <b>FAA</b> MAGVIGTILLISYGIRRLIKK <sup>101</sup>
Δ89/93	<sup>72</sup> EITLIIFGVMAGVIGTI-----GIRRLIKK <sup>101</sup>
Δ89/91	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGIRRLIKK <sup>101</sup>
Δ91/93	<sup>72</sup> EITLIIFGVMAGVIGTILL---GIRRLIKK <sup>101</sup>
Δ91	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGIRRLIKK <sup>101</sup>
91-93L	<sup>72</sup> EITLIIFGVMAGVIGTILL <b>LLL</b> GIRRLIKK <sup>101</sup>
91-95L	<sup>72</sup> EITLIIFGVMAGVIGTILL <b>LLLLL</b> RRRLIKK <sup>101</sup>
Δ89/91, Δ96/97	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGI--LIKK <sup>101</sup>
Δ89/91,96/97L	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGI <b>LLL</b> IKK <sup>101</sup>
Δ89/91,96/97L,100/101L	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGI <b>LLLLILL</b> <sup>101</sup>
Δ91, Δ96/97	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGI--LIKK <sup>101</sup>
Δ91,96/97L	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGI <b>LLL</b> IKK <sup>101</sup>
Δ96/97	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGI--LIKK <sup>101</sup>
96/97L	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGI <b>LLL</b> IKK <sup>101</sup>
96/97L,100/101L	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGI <b>LLLLILL</b> <sup>101</sup>
5L	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTILLISYGIRRLIKK <sup>101</sup>
5L, Δ89/93	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTI-----GIRRLIKK <sup>101</sup>
5L, Δ89/91	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTI---SYGIRRLIKK <sup>101</sup>
Mutante	Secuencia

5L, Δ91/93	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTILL---GIRRLIKK <sup>101</sup>
5L, 91/93L	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTILL <b>LLL</b> GIRRLIKK <sup>101</sup>
5L, 91/95L	<sup>72</sup> EITLIIFGVM <b>LLLLL</b> AGVIGTILL <b>LLLLL</b> RRRIKK <sup>101</sup>
E72P	<sup>72</sup> <b>P</b> ITLIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
I73P	<sup>72</sup> <b>EPT</b> LIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
T74P	<sup>72</sup> E <b>IP</b> LIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
L75P	<sup>72</sup> EIT <b>P</b> LIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
I76P	<sup>72</sup> EITL <b>P</b> IFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
I77P	<sup>72</sup> EITLI <b>P</b> FGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
F78P	<sup>72</sup> EITLIIF <b>P</b> GVMAGVIGTILLISYGIRRLIKK <sup>101</sup>
G79P	<sup>72</sup> EITLIIFG <b>V</b> PAGVIGTILLISYGIRRLIKK <sup>101</sup>
V80P	<sup>72</sup> EITLIIFG <b>P</b> MAGVIGTILLISYGIRRLIKK <sup>101</sup>
M81P	<sup>72</sup> EITLIIFG <b>V</b> PAGVIGTILLISYGIRRLIKK <sup>101</sup>
A82P	<sup>72</sup> EITLIIFGVM <b>P</b> GVIGTILLISYGIRRLIKK <sup>101</sup>
G83P	<sup>72</sup> EITLIIFGVM <b>A</b> PVIGTILLISYGIRRLIKK <sup>101</sup>
V84P	<sup>72</sup> EITLIIFGVMAG <b>P</b> IGTILLISYGIRRLIKK <sup>101</sup>
I85P	<sup>72</sup> EITLIIFGVMAG <b>V</b> PGTILLISYGIRRLIKK <sup>101</sup>
G86P	<sup>72</sup> EITLIIFGVMAGVI <b>P</b> TILLISYGIRRLIKK <sup>101</sup>
T87P	<sup>72</sup> EITLIIFGVMAGVIG <b>P</b> ILLISYGIRRLIKK <sup>101</sup>
I88P	<sup>72</sup> EITLIIFGVMAGVIGT <b>P</b> LLISYGIRRLIKK <sup>101</sup>
L89P	<sup>72</sup> EITLIIFGVMAGVIGTI <b>P</b> LISYGIRRLIKK <sup>101</sup>
L90P	<sup>72</sup> EITLIIFGVMAGVIGTIL <b>P</b> ISYGIRRLIKK <sup>101</sup>
I91P	<sup>72</sup> EITLIIFGVMAGVIGTILL <b>P</b> SYGIRRLIKK <sup>101</sup>
S92P	<sup>72</sup> EITLIIFGVMAGVIGTILLI <b>P</b> YGIRRLIKK <sup>101</sup>
Y93P	<sup>72</sup> EITLIIFGVMAGVIGTILLIS <b>P</b> GIRRLIKK <sup>101</sup>
G94P	<sup>72</sup> EITLIIFGVMAGVIGTILLISY <b>P</b> IRRLIKK <sup>101</sup>
I95P	<sup>72</sup> EITLIIFGVMAGVIGTILLISY <b>G</b> PRRLIKK <sup>101</sup>
R96P	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGI <b>P</b> RRLIKK <sup>101</sup>
R97P	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGIR <b>P</b> LIKK <sup>101</sup>
Xho485/KpnI560	<sup>72</sup> LITLIIFGVMAGVIGTILLISYGTRRLIKK <sup>101</sup>
18L	<sup>72</sup> LLLLLLLLLLLLLLLLLLLLLLLLL <b>T</b> RRLIKK <sup>101</sup>
Mutante	Secuencia

G/18L	<sup>72</sup> LLLLLLL <b>G</b> LLLLLLLLLLLLLLLLTRRLIKK <sup>101</sup>
GG/18L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> LLLLLLLLLLLLTRRLIKK <sup>101</sup>
GG/21L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> LLLLLLLLLLLLLLLLTRRLIKK <sup>101</sup>
GVG/18L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> VLLLLLLLLLLLLTRRLIKK <sup>101</sup>
GVG/21L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> VLLLLLLLLLLLLLLLLTRRLIKK <sup>101</sup>
GVGVT/21L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLLLLLLLTRRLIKK <sup>101</sup>
IGVG/18L	<sup>72</sup> LLLL <b>I</b> LL <b>G</b> VLL <b>G</b> VLLLLLLLLLLLLTRRLIKK <sup>101</sup>
IGVG/21L	<sup>72</sup> LLLL <b>I</b> LL <b>G</b> VLL <b>G</b> VLLLLLLLLLLLLLLLLTRRLIKK <sup>101</sup>
GVGVT/15L	<sup>72</sup> LLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
GVGVT/18Lr	<sup>72</sup> LLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLLLLLLLTRRLIKK <sup>101</sup>
GVGVT/24L	<sup>72</sup> LLLLLLLLLLLLLLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
GGI/18L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> LLL <b>I</b> LLTRRLIKK <sup>101</sup>
GGN/18L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> LL <b>N</b> LLTRRLIKK <sup>101</sup>
GVGVT/27L	<sup>72</sup> LLLLLLLLLLLLLLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLLLLLLLTRRLIKK <sup>101</sup>
GVGVT/18L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
AGVGVT/18L	<sup>72</sup> LLL <b>A</b> LLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
IGVGVT/18L	<sup>72</sup> LLL <b>I</b> LLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
GVGVS/18L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> VLL <b>S</b> LLTRRLIKK <sup>101</sup>
VGVGT/18L	<sup>72</sup> LLL <b>V</b> LLL <b>G</b> VLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
GGT/18L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> LL <b>T</b> LLTRRLIKK <sup>101</sup>
GGVT/18L	<sup>72</sup> LLLLLLL <b>G</b> LLL <b>G</b> VLL <b>T</b> LLTRRLIKK <sup>101</sup>
GIGIT/18L	<sup>72</sup> LLLLLLL <b>G</b> ILL <b>G</b> ILL <b>T</b> LLTRRLIKK <sup>101</sup>
GVGT/18L	<sup>72</sup> LLLLLLL <b>G</b> VLL <b>G</b> LL <b>T</b> LLTRRLIKK <sup>101</sup>

#### 2.2.4. Preparación de construcciones por apareamiento entre oligonucleótidos complementarios

Esta estrategia se ha utilizado para preparar las construcciones que se presentan en la tabla 2.3. En primer lugar, para obtener el vector, los plásmidos fueron digeridos con ambas enzimas de restricción, tratados con Fosfatasa Alcalina (*Phosphatase Alcaline Shrimp* de, *Boehringer Mannheim*) y purificados desde gel de agarosa utilizando el *QIAquick® Gel Extraction kit* (*Qiagen, Alemania*).

De modo global, para introducir un inserto de interés entre dos sitios de restricción de un vector se diseñan oligonucleótidos que codifican para la secuencia deseada, flanqueada por extremos colgantes que simulan estar digeridos por dichos enzimas (tabla 2.3).

Por otro lado, los oligonucleótidos complementarios fueron fosforilados por separado utilizando T4 polinucleotido quinasa de Promega, en aquellos casos en los que no se obtuvieron directamente fosforilados en 5' (*Isogen, Holanda*).

Tabla 2.3. Secuencia del segmento TM de las construcciones de polileucinas realizadas por apareamiento entre nucleótidos complementarios.

Mutante	Secuencia TM
18L	<sup>72</sup> LLLLLLLLLLLLLLLLLLLLLLLLL <sup>101</sup> TRRLIKK
G/18L	<sup>72</sup> LLLLLLLL <b>G</b> LLLLLLLLLLLLL <sup>101</sup> TRRLIKK
GG/18L	<sup>72</sup> LLLLLLLL <b>G</b> LLL <b>G</b> LLLLLLLLL <sup>101</sup> TRRLIKK

El apareamiento entre oligonucleótidos complementarios se llevó a cabo incubando 20 µL (30ng/µL) de cada uno de ellos en 20mM Tris-HCl pH 7.4, 2mM MgCl<sub>2</sub>, 50mM NaCl durante 10 min a 75°C y dejando enfriar la muestra lentamente, durante aproximadamente dos horas, hasta alcanzar 30°C.

La reacción de ligación entre los oligonucleótidos apareados y el vector digerido se llevó a cabo en 20µL de volumen de reacción utilizando el enzima T4 DNA ligasa (Promega). Para ello se utilizaron 400ng de oligonucleótidos apareados y aproximadamente 100ng de vector digerido, que fueron incubados durante 5 horas a 16°C. El resultado de la ligación fue utilizado para transformar 100µL de células competentes MC1061. Las colonias resultantes fueron analizadas por secuenciación en el (Servicio de Secuenciación de DNA del SCSIE, Universitat de València).

- Con esta metodología se han realizado dos tipos de construcciones:

### 1. Construcciones de poliLeucinas:

Para realizar este tipo de construcciones se preparó el vector derivado del pT7SN/GpA, al que mediante mutagénesis se le introdujeron los sitios de restricción *XhoI* y *KpnI* (*Xho* 485/*Kpn* 560) en los extremos N y C-terminales del fragmento transmembrana de GpA respectivamente, lo que nos permite la sustitución del fragmento natural por los fragmentos de interés (figura 2.4).

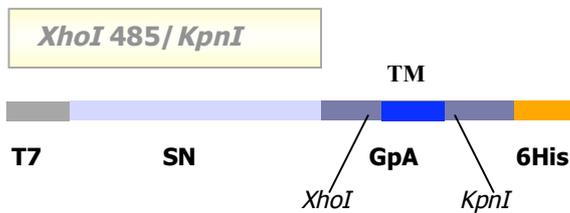


Figura 2.4: Representación esquemática del plásmido derivado de pT7SN/GpA, en el que se han introducido los sitios de restricción *XhoI* y *KpnI* en ambos extremos del fragmento TM .

### 2. Contrucciones GpA-Lep:

Para obtener estos vectores se digirió el vector pGEM1- Lep con los enzimas de restricción *BclI* y *NdeI* para eliminar el fragmento H2 de Lep e introducir en su lugar el fragmento TM de GpA, obteniéndose el vector pGEM1-GpA-Lep (figura 2.5). Para el diseño de esta construcción fue necesario invertir la secuencia natural de GpA, ya que se trata de una proteína de tipo I, es decir, C-terminal citoplasmática y su orientación en la membrana es la contraria a la del fragmento H2 de Lep (figura 2.3).

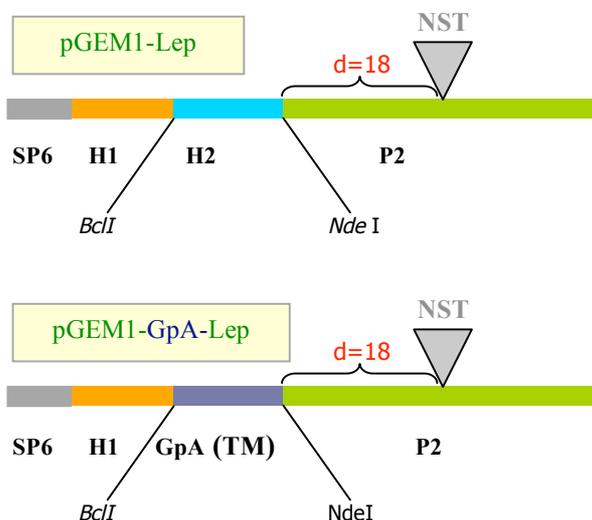


Figura 2.5: Representación esquemática del plásmido pGEM-1-Lep y de su derivado pGEM1-GpA-Lep, en el que la hélice H2 ha sido substituida por el segmento TM de GpA

De este modo la orientación final de GpA es la original de la proteína salvaje. En los plásmidos de Lep con diferentes distancias de glicosilación, la inserción de GpA se realizó por subclonación del fragmento *XbaI/NdeI* en los vectores pertinentes. Los oligonucleótidos utilizados para la subclonación se muestran en la Tabla 2.4.

Tabla 2.4. Oligonucleótidos utilizados para la subclonación de GpA en los diferentes vectores pGEM1-Lep con diferentes distancias de glicosilación . Ambos oligonucleótidos se encuentran fosforilados en posición 5'.

Oligo	Secuencia
H2GpAf	5' GATCAGGCGCATTTGGTTACTCTATTTTACTCATCACGGAATTGTTGG TGCTATGGTGGGGTTTATTATTCTCACAATAGAGCCA-3'
H2GpAr	5' TATGGCTCTATTGTGAGAATAATAAACCCACCATAGCACCAACAATT CCCGTGATGAGTAAAATAGAGTAACCAATGCGCCT-3

## 2.3. Manipulación de proteínas

### 2.3.1. Electroforesis en gel de poliacrilamida.

Las electroforesis de proteínas se realizaron en geles de poliacrilamida en presencia de SDS (Laemmli, 1970), utilizando un sistema de minigeles (*Miniprotean® Electrophoresis cell, Biorad, USA*) y concentraciones de acrilamida variables entre el 10 y el 15%. Las muestras se prepararon añadiendo tampón de muestras para obtener una concentración final de 625mM Tris-HCl pH 6,8, 10% glicerol, 2% SDS, 4% β-mercaptoetanol, 0,025% azul de bromofenol, tras lo que se hirvieron, en la mayoría de los casos, durante 5 min. Las electroforesis se realizaron a intensidad constante de 16mA por mini-gel.

Para la tinción de estos geles se utilizó Azul de Coomassie , tras lo que se sumergieron durante unos 30 min en solución decolorante ( 20% metanol, 5% ácido acético).

### 2.3.2. Expresión y purificación de SN/GpA-6H y mutantes derivados.

Para la expresión de la proteína SN/GpA, así como de los mutantes obtenidos a partir de ella, se prepara una dilución 1/100 en 10 mL de medio TB Ampicilina/ Cloramfenicol a partir de un precultivo crecido durante toda la noche del mutante en LB y se deja crecer el cultivo a 37°C y 200 rpm hasta alcanzar una OD de 2.5. La inducción de la expresión de la proteína se realiza mediante la adición al

medio de IPTG (Isopropil- $\beta$ -D-tiogalactopiranosido) 0,8mM, tras lo que se incuban los cultivos durante 3h a 37°C y 200rpm.

Se recogen las células mediante centrifugación a 3000 rpm 10 minutos y el sedimento celular se resuspende en 500 $\mu$ L de tampón 50mM Tris-HCl pH7.9, 5mM de EDTA, 1mM PMSF. Tras tres rondas de congelación/descongelación, se añade CaCl<sub>2</sub> 10mM a la muestra y se incuba en hielo durante 15 minutos. Se centrifuga la muestra 5 minutos a 12000 rpm y se redisuelve el precipitado en 1mL de TBS 1%SDS. Para romper las células se sonicán en un sonicador Vibra cell 500 (Sonics & Materials Inc, USA) al 50% en posición 4 durante 30 segundos observando que el medio deja de ser viscoso.

Para eliminar los restos celulares se centrifuga 10min a 12000 rpm, desechándose el sedimento. El sobrenadante se incuba en agitación durante una hora a temperatura ambiente con 50 $\mu$ L del volumen de partida de resina Ni-Agarosa (Qiagen). La resina se recoge tras la incubación centrifugando 10 segundos a 12000 rpm y se resuspende en 100 $\mu$ L de TBS 0.5%SDS, tras lo que se introduce la muestra en una minicolumna en la que se realiza un lavado con 1mL de TBS 0.5% SDS 10mM Imidazol. La elución de la muestra se realiza con 55 $\mu$ L de TBS 0.5%SDS 100mM de Imidazol.

### **2.3.3. Cuantificación de proteína.**

La cuantificación de proteínas se realizó utilizando el método del ácido bicinconínico (BCA) (*BCA Protein Assay, Pierce, Perbio, USA*). El principio de este ensayo es similar al del método de Lowry, ya que depende de la reducción de Cu<sup>2+</sup> a Cu<sup>+</sup> que es facilitada al formarse el complejo con el enlace peptídico. El Cu<sup>+</sup> forma un complejo de color púrpura con el BCA que absorbe a  $\lambda$  562nm. La ventaja de este método es que el uso de detergentes como el SDS no afecta a la cuantificación de proteína.

### **2.3.4. Cuantificación de los porcentajes de monómero/dímero de SN/GpA y construcciones derivadas.**

Para la cuantificación del porcentaje de dímero presente en este tipo de muestras se realiza una electroforesis (SDS-PAGE) al 12% de acrilamida en la que se carga un microgramo de muestra a 0.1  $\mu$ g/ $\mu$ L. Las bandas de monómero y dímero que se observan tras tinción con Azul Coomassie se

cuantificaron utilizando el programa Quantity One versión 4.1.1 (Biorad). En todos los casos los porcentajes de dímero son la media del análisis de al menos tres réplicas del mismo mutante.

### **2.3.5. Experimentos de competencia proteína/péptido.**

Los experimentos de competencia proteína/péptido, se realizaron incubando una concentración fija de proteína (4 $\mu$ M) y cantidades crecientes de péptido en tampón TBS (10mM Tris-HCl pH 7.5, 0.9% NaCl), 0,5% de SDS, de modo que la relación proteína/péptido varía en proporción creciente desde 1:1 hasta 1:50. La incubación se llevó a cabo durante 1h a 25°C, habiéndose comprobado previamente que la incubación a tiempos más largos no alteraba el grado de dimerización. Tras esta incubación las muestras son sometidas a electroforesis en gel de poliacrilamida. La tinción del gel con azul de Coomassie permite cuantificar, tanto la desaparición de dímero proteína/proteína como la correspondiente aparición de heterodímero proteína/péptido.

### **2.4. Experimentos de transcripción/traducción *in vitro*.**

Los experimentos de inserción en membranas biológicas se realizaron mediante la transcripción y posterior traducción, en presencia de membranas microsomales, de las construcciones derivadas de pGEM1-Lep. Los microsomas poseen toda la maquinaria celular necesaria para que la proteína se inserte en la membrana. Además también poseen la enzima oligosacaril transferasa (OST) que se encarga de transferir oligosacáridos a las proteínas, allí donde encuentra la secuencia señal N<sub>x</sub>T/S, a medida que ésta emerge del traslocón. (figura 2.6).

En las construcciones derivadas de pGEM1-Lep se ha sustituido el segundo fragmento TM de Lep por el fragmento TM de GpA. En el extremo C-terminal, esta proteína quimérica posee un sitio de glicosilación, de modo que se puede distinguir si el segmento transmembrana se inserta o no dependiendo de su estado de glicosilación. Si existe inserción, el sitio de glicosilación queda accesible al sitio activo de la OST, modificando a la proteína naciente, sin embargo, si no hay inserción la proteína queda sin glicosilar (figura 2.7).

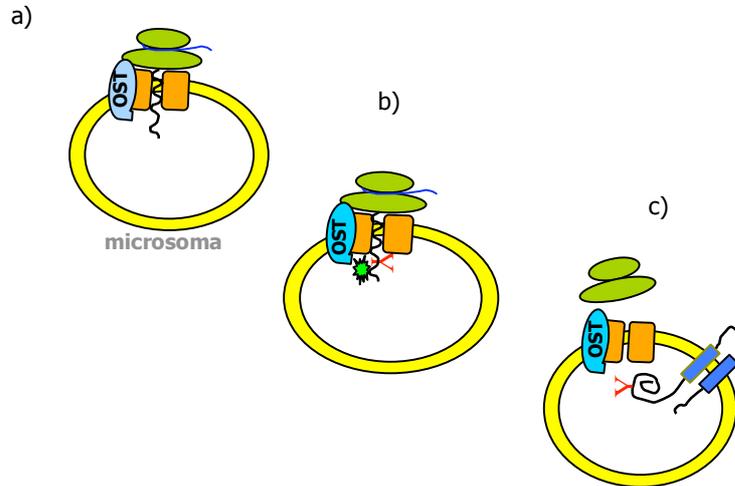


Figura 2.6: Esquema de la traducción *in vitro* en presencia de microsomas.

(a) El ribosoma se une al traslocón (color naranja) situado en la membrana del microsoma y se reanuda la traducción. (b) Cuando la oligosacil transferasa, asociada al traslocón (OST), encuentra una secuencia diana, se produce la glicosilación de la proteína (Y). (c) Se acaba la traducción y el ribosoma se libera del microsoma, quedando la proteína, insertada en la membrana y glicosilada. En este esquema el interior del microsoma corresponde al lumen del retículo y el exterior sería equivalente al citosol celular.

La migración electroforética de una proteína glicosilada sufre un retraso aproximadamente equivalente a un incremento en su peso molecular de 2 kDa, de modo que es posible distinguir en un gel de poliacrilamida, entre ambas opciones (figura 2.7). Para poder visualizar las proteínas traducidas *in vitro*, la traducción se realiza en presencia de L-<sup>35</sup>S-Metionina.

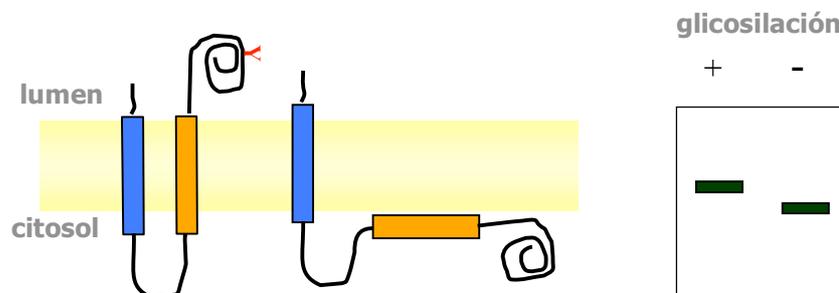


Figura 2.7: Representación esquemática del resultado electroforético que se obtendría para la construcción pGEM1-GpA-Lep dependiendo de si la hélice TM de GpA se insertara (glicosilación +) o no (glicosilación-). En la parte izquierda de la figura se muestra un esquema de lo que estaría ocurriendo con la proteína quimérica en ambas situaciones en la membrana del microsoma.

Para la obtención del RNA se utilizó el sistema Ribomax<sup>TM</sup> Large Scale RNA production System-SP6, (Promega, USA). En un volumen de reacción de 50µL se incubó durante 2h a 37°C una mezcla con aproximadamente 1µg de DNA plasmídico, 10µL de mezcla de rNTPs y 5µL de la mezcla enzimática (RNapolimerasa/fosfatasa) en tampón SP6 1x. El RNA obtenido por transcripción se purificó utilizando RNeasy® Minikit (Qiagen, Germany). La reacción se verificó antes y después de la purificación visualizando una pequeña alícuota en gel de agarosa al 1%.

Para la traducción *in vitro* se utilizó un sistema eucariótico de lisado de reticulocitos, Rabbit Reticulocyte Lysate System (Promega, USA). La reacción se llevó a cabo en un volumen final de 15µL durante 2h a 30°C. La mezcla de reacción se compone de 1µL del RNA obtenido por transcripción, 0,5µL de mezcla de aminoácidos sin metionina, 10µL de lisado de reticulocitos y 1µL de L-<sup>35</sup>S-Metionina de 7.9mCi/mL (Nen<sup>TM</sup> Life Science Products Inc., Boston). En todos los casos las muestras se traducen por duplicado en presencia y ausencia de 1µL de microsomas de pancreas de perro (Promega, USA). En algunas muestras se incluye una reacción control en presencia de péptido aceptor de la glicosilación, el tripéptido NYT que es diana de glicosilación para la glicosiltransferasa compitiendo con la proteína naciente como diana para la transferencia de azúcares a la proteína traducida. Del mismo modo se dispone de un tripéptido, no aceptor, incapaz de ejercer este efecto (QYT). En estos casos los péptidos se añaden a la mezcla de traducción a concentración 60µM.

A los 15 µL de muestra traducida se le añadieron 25 µL de agua y 10 µL de tampón de muestras 5x. Tras hervir durante 5 min se realizó una electroforesis de poliacrilamida al 10%, utilizando 10 µL de cada muestra. El gel se fijó durante 30 min en una solución orgánica (20%metanol, 5%ácido acético), tras lo que se secó en un secador de geles, SG4050 Slab Gel dryer, (Thermosavant, USA) y se expuso en una pantalla intensificadora Bas-IP MP 2025 (Fujifilm, Japan). El análisis de la imagen se llevó a cabo en un FLA-3000 (Fujifilm, Japan), utilizando para procesar las imágenes los programas Image Gaugev3.12 e ImageReader v1.0 (FujiPhotoFilm, Japan)

En algunos experimentos, una vez se dispuso de la mezcla traducida, se realizó la separación de las fracciones soluble y membranosa mediante un protocolo de extracción alcalina. Para ello a los 15µL de muestra traducida se le añadieron 90µL de Carbonato sódico 0,1M y este volumen se dispuso sobre una almohadilla de 50 µL de Carbonato sódico 0,1M, Sacarosa 0,2M. La muestra se ultracentrifugó a 4°C durante 20 min a 40000rpm. Tras separar la fracción soluble a la que se le añade

tampón de muestras, se dejó secar el precipitado, en el que se encuentran las proteínas integrales de membrana, durante 2 min a 95°C y se disolvió en 25µL de tampón de muestras 1x, de este modo ambas fracciones se analizaron en gel de poliacrilamida al 10% como se ha descrito previamente.

## 2.4. Manipulación de péptidos

### 2.4.1. Síntesis y purificación de péptidos

Para la obtención de los péptidos que se manejan en este trabajo se ha utilizado síntesis de péptidos en fase sólida (Erickson & Merrifield, 1976) de tipo F-moc en la que los grupos  $\alpha$ -amino de los aminoácidos (Novagen, Senn Chemicals) están protegidos por el grupo base-lábil N-(9-fluorenil)-metoxicarbonilo (F-moc)(Fields & Noble, 1990), mientras que las cadenas laterales, en aquellos casos en los que es necesario, están protegidas por diferentes grupos ácido lábiles (Tabla 2.5).

Al tratarse en la mayoría de los casos de péptidos con una elevada hidrofobicidad se han utilizado como soporte sólido para la síntesis resinas de baja substitución, como la resina Tentagel (0.22meq/mol), con el fin de evitar la agregación durante el proceso sintético. En los casos en los que se realizó la síntesis de forma manual se utilizaron bolsas de polipropileno semipermeable para la compartimentalización de la resina.

Tabla 2.5. Aminoácidos utilizados en el proceso de síntesis.

Aminoácido	Código de una letra	Aminoácido protegido	Aminoácido	Código de una letra	Aminoácido protegido
Alanina	A	Fmoc-Ala-OH	Metionina	M	F-moc-Met-OH
Cisteína	C	Fmoc-Cys(Trt)*-OH	Asparagina	N	Fmoc-Asn(Trt)*-OH
Aspártico	D	Fmoc-Asp(tBu)*-OH	Prolina	P	Fmoc-Pro-OH
Glutámico	E	Fmoc-Glu(tBu)*-OH	Glutamina	Q	Fmoc-Gln(Trt)-OH
Fenilalanina	F	Fmoc-Phe-OH	Arginina	R	Fmoc-Arg(Pmc)*-OH
Glicina	G	Fmoc-Gly-OH	Serina	S	Fmoc-Ser(tBu)*-OH
Histidina	H	Fmoc-His(Trt)*-OH	Treonina	T	Fmoc-Thr(tBu)*-OH
Isoleucina	I	Fmoc-Ile-OH	Valina	V	Fmoc-Val-OH
Lisina	K	Fmoc-Lys(BOC)*-OH	Triptófano	W	Fmoc-Trp(BOC)*-OH
Leucina	L	Fmoc-Leu-OH	Tirosina	Y	Fmoc-Tyr(tBu)*-OH

\*Las abreviaturas Trt (trifenilo), t-Bu (terc-butilo), BOC (terc-butoxicarbonilo) y Pmc (2,2,5,7,8-Pentametilcroman-6-sulfonilo corresponden a los grupos protectores de las cadenas laterales de los aminoácidos.

El **proceso sintético manual** consiste en una serie de ciclos de desprotección en medio básico del grupo  $\alpha$ -amino y acople del siguiente residuo que permiten el crecimiento del péptido sobre la resina. En primer lugar se procede a la desprotección de la resina utilizando Piperidina 20% en dimetil formamida (DMF), un medio básico que permite la liberación del grupo F-moc, tras lo que se acopla el aminoácido correspondiente disuelto en DMF. Para el cálculo de la cantidad de aminoácido necesaria se considera un rendimiento de la síntesis del 60% y se utiliza un exceso 6x de la cantidad necesaria. Por cada equivalente de aminoácido se añade un equivalente de 1-hidroxibenzotriazol (HOBt) y dos equivalentes de diisopropilcarbodiimida (DIC), agentes que favorecen el proceso de acople. El proceso de desprotección se suele realizar en dos etapas de 5 y 20 minutos, mientras que el proceso de acople se realiza en 1h 30 minutos. Tanto después del proceso de desprotección como del de acople se realiza la siguiente serie de lavados cortos (1min):

- Tres lavados con DMF para eliminar los restos de piperidina o de aminoácido.
- Tres lavados con isopropanol que favorecen la compresión de la resina eliminando los restos de los huecos intersticiales de la resina.
- Tres lavados con diclorometano (DCM) que favorecen la expansión de nuevo de la resina de forma que tras su secado queda preparada para el siguiente ciclo de acople o desprotección.

El proceso de desprotección y acople a lo largo de la síntesis se ha evaluado en pequeñas muestras de resina utilizando el Test de Kaiser (Kaiser et al., 1970). Este test se basa en la reacción del reactivo ninhidrina con los grupos amino libre, dando lugar a la aparición de un producto de color azul intenso tras un proceso de desprotección, que no aparece en el caso de haber realizado el paso de acople de un nuevo residuo. La desprotección también se cuantificó espectrofotométricamente midiendo la liberación del grupo F-moc a la mezcla de desprotección mediante el seguimiento de la absorbancia a 290nm (coeficiente de extinción molar  $4950 \text{ M}^{-1}\text{cm}^{-1}$ ).

Tras los ciclos necesarios de acople y desprotección para completar la secuencia del péptido, se procede, en algunos casos con objeto de mimetizar el enlace peptídico, a la acetilación de éste. Para ello se desprotege el último residuo de la secuencia y se incuba la resina durante 30 minutos con 20% de anhídrido acético en DMF. Tras repetir los ciclos de lavados se procede a la desprotección completa de las cadenas laterales de los aminoácidos y a la escisión del péptido de la resina mediante

incubación durante 4 horas en una mezcla que contiene 70% de ácido trifluoroacético (TFA), 5% de agua, 2.5% de etanoditiol y 2,5% de clortriisobutilsilano.

La precipitación selectiva del péptido se realiza dejando gotear lentamente la disolución de corte sobre t-butilmetileter frío. El péptido se deja precipitando durante 12h a  $-80^{\circ}\text{C}$ , tras lo que se centrifuga a 3000rpm durante 30 min. El péptido precipitado se redisuelve en ácido acético, se congela y liofiliza, Tras este paso se reconstituye en una solución 90:10 de agua:acetonitrilo, y se procede de nuevo a su congelación y liofilización.

En el caso de la **síntesis automática** se ha utilizado el sintetizador automático 433 A Peptide synthesiser (Applied Biosystems, USA) y los protocolos de síntesis que maneja este aparato para la síntesis “*FastMoc 0.1 monitoring previous peak*”, introduciendo ligeras modificaciones en los tiempos de acople, activadores y tipo de resina utilizada según la necesidad para los diferentes péptidos sintetizados. Para evaluar el proceso de desprotección en este tipo de síntesis el aparato mide la conductividad del medio de desprotección. Tras la síntesis se procede del mismo modo que en el caso anterior a la excisión de la resina y a la reconstitución del péptido mediante liofilización.

La relación de los diferentes péptidos que se han utilizado en este trabajo se muestra en la tabla 2.6.

Tabla 2.6. Péptidos utilizados en el trabajo.y tipo de síntesis

Nombre del péptido	Secuencia	Tipo de síntesis
Ct-wt	Ac-ILLISYGIRRLI-NH <sub>2</sub>	Manual
88-99 96/97K	Ac-ILLISYGIKKLI-NH <sub>2</sub>	Manual
88-99 Δ91	Ac-ILLSYGIRRLI-NH <sub>2</sub>	Manual
88-99 Δ89-91	Ac-ISYGIRRLI-NH <sub>2</sub>	Manual
88-102 Δ89-91	Ac-ISYGIRRLIKKS-NH <sub>2</sub>	Manual
SAT 1	RGGVMAGVIGTGR	Automática
SAT 2	RGLGVLLGVLLTLGR	Automática
SAT 3 (Wt)	KKKKEITLIIFGVMAGVIGTILLISYGKKKK	Automática
SAT4 (L75P)	KKKKEITPIIFGVMAGVIGTILLISYGKKKK	Automática

El proceso de purificación de los péptidos se realizó utilizando cromatografía líquida de alta resolución en fase reversa (RP-HPLC). En primer lugar se analizó el perfil cromatográfico de una pequeña fracción de la muestra (aproximadamente 100µL de concentración 0,2mg/mL disuelta en agua:acetonitrilo (90:10)) en un aparato analítico de la casa Waters compuesto por un inyector Waters 717 plus, un detector fotodiodo Waters 996 y un sistema de bombas Waters 600. Para ello, se utilizó una columna de fase reversa C18 (RP-18) de 5µm de tamaño de partícula, 12,5 cm de longitud y 0,8cm de diámetro (LiChrospher®100, Merck) y una fase móvil compuesta de agua:acetonitrilo, en ambos casos conteniendo 0.1%TFA. El seguimiento del péptido en este tipo de sistemas se realiza midiendo absorbancia a 220nm, longitud de onda a la que absorbe el enlace peptídico. Las muestras se analizaron utilizando inicialmente un gradiente lineal básico de agua:acetonitrilo en el que las proporciones varían de 90:10 a 10:90 en 30 min, que fue optimizado para cada uno de los péptidos, de modo que permitiera su posterior purificación.

Tras obtener un gradiente adecuado en el sistema analítico, se utilizó un sistema preparativo para la purificación, que consiste en un inyector y controlador de bombas Waters Delta Prep. 3000, un detector ultravioleta (UV) L-7400 de Merck y una columna LiChrosorb® RP-18 de 7µm de tamaño de partícula, 27 cm de longitud y 3,5 cm de diámetro. Las fracciones recogidas tras este proceso de purificación, fueron analizadas de nuevo en el HPLC analítico e identificadas por espectrometría de masas MALDI-TOF (*matrix-assisted laser desorption/ionization time of flight*) en el laboratorio de Proteómica, Centro Nacional de Biotecnología, CSIC, Madrid. Las fracciones correspondientes al peso molecular del péptido se agruparon y se liofilizaron, quedando el péptido preparado para su posterior uso. La cuantificación de los péptidos para su uso se realizó mediante espectrofotometría utilizando el máximo de absorbancia a 276 nm del aminoácido tirosina ( $\epsilon=1450 \text{ M}^{-1}\text{cm}^{-1}$ ).

En la síntesis de los péptidos SAT1 y 2 se insertó un cassette RG en ambos extremos con el objetivo de mejorar su solubilidad. Estas secuencias no poseen residuos aromáticos por lo que su cuantificación se realizó por pesada.

La síntesis de fragmentos TM (SAT3 y SAT4) presenta gran dificultad, por un lado, por su condición hidrofóbica y por otro lado, por su gran longitud. Ambas características provocan la agregación del péptido durante el proceso sintético impidiendo que los acoples se lleven a término correctamente y dando lugar a la aparición de péptidos truncados. La purificación de péptidos hidrofóbicos tampoco es un proceso trivial, ya que con los métodos estándar, se obtienen generalmente

perfiles cromatográficos muy anchos que impiden la correcta purificación, por lo que un alto rendimiento de síntesis es esencial para la obtención de un péptido puro. Por todo ello, y tras varios intentos de obtención de estos fragmentos por métodos convencionales de síntesis tanto manual como automática, se adoptó una estrategia descrita recientemente (Melnyk et al., 2001), en la que la adición de cassettes de lisinas en los extremos de un fragmento hidrofóbico facilita tanto el proceso de síntesis, como el posterior manejo y purificación de los péptidos. Para determinar el número de lisinas que se incorpora a ambos extremos de segmento TM, se calcula la hidrofobicidad total del fragmento utilizando la escala de Liu/Deber (Deber et al., 2001; Liu & Deber, 1998), y se añade una lisina en cada extremo del péptido por cada 16 unidades de hidrofobicidad calculada. Se utilizó la síntesis automática aumentando los tiempos de acople mediante la adición de un módulo de espera de 5 minutos y utilizando un exceso 8x en la concentración de aminoácidos. La resina utilizada fue PAL-PEG-PS (*Applied Biosystems, Foster City, CA*) de baja substitución (0.18-0.22mmol/g) para evitar la agregación durante la síntesis. En algunas posiciones se realizaron dobles y triples acoples para evitar la aparición de productos truncados (Fisher & Engelman, 2001) (figura 2.8). En esta síntesis se utilizó el par HATU o-(z-azabenzotriazol-1-yl)-N,N,N',N'-(tetrametil-uronium hexafluorofosfato) / DIEA (N,N-diisopropilmetilamina) como activador potente de la reacción. Los péptidos fueron acetilados, tras lo que se realizó la escisión de la resina con una mezcla de corte que contenía 88% de TFA, 5% fenol, 5% de agua, 2% triisobutilsilano, (adaptado de (Melnyk et al., 2001)).



Figura 2.8: Secuencia del péptido correspondiente a la secuencia salvaje de GpA (SAT 3) y al mutante L75P, (SAT4), con el número de acoples que se realizaron para las diferentes posiciones en el proceso sintético, que se muestra como bolas de color naranja. En azul se muestran los cassettes de residuos de lisina que se añadieron en los extremos N y C-terminales para mejorar la solubilidad de los péptidos

Tras su escisión y posterior liofilización los péptidos fueron purificados por HPLC como se ha descrito previamente y analizados por espectrometría de masas (figura 2.9).

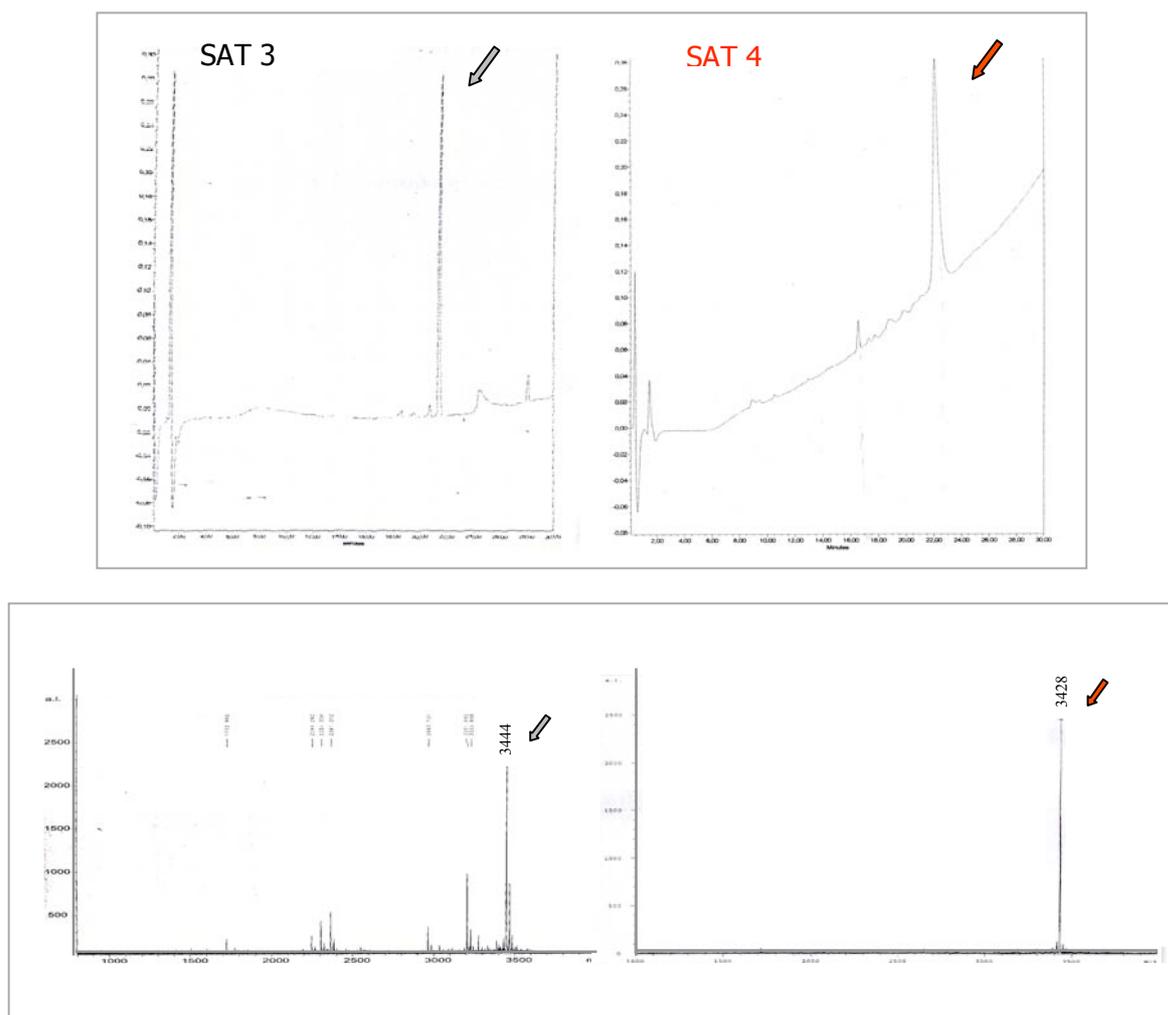


Figura 2.9: En la parte superior se muestran los cromatogramas de HPLC analítico de los péptidos SAT 3 (wt) y SAT 4 (L75P) y en la inferior los espectros de masas obtenidos para estos péptidos.

#### 2.4.2. Espectros de dicroísmo circular.

Los experimentos de dicroísmo circular se realizaron en un espectropolarímetro Jasco J-810. Las muestras se midieron a concentraciones en torno a  $30\mu\text{M}$ . En todos los casos se utilizó una cubeta de 1mm de paso óptico y la temperatura de adquisición de los espectros fue de  $25^{\circ}\text{C}$ , con una velocidad

de barrido de 20 nm/min y un paso de 0.2nm. Los espectros que se presentan son la media de 15 adquisiciones.

Las curvas de desnaturalización térmica se realizaron en el mismo espectropolarímetro con un baño Neslab RT 111 acoplado. En estos experimentos se realizaron medidas de Elipticidad Molar a 222 nm en un rango de temperaturas entre 10 y 90°C, en intervalos de 0.2 grados

### **2.4.3. Electroforesis de Tris-tricina**

Para la visualización de péptidos en electroforesis se utilizaron geles de Tris-tricina en gradiente 10-20% (*Invitrogen, life technologies, Carlsbad*). Las electroforesis se realizaron a 125 voltios y aproximadamente 80mA de intensidad al inicio, durante 90minutos.

### **2.5. Modelado de estructuras.**

Los modelos estructurales tanto de los mutantes de prolinas de GpA, como de las deleciones de regiones C-terminales se han generado *in silico* partiendo de las coordenadas de un fragmento péptidico helicoidal (residuos Ser69-Lys101) que corresponde a la estructura del fragmento transmembrana de esta proteína en presencia de micelas de detergente (PDB 1AFO) (MacKenzie et al., 1997). Los mutantes fueron generados manteniendo la estructura del esqueleto peptídico en el programa Swiss PdB Viewer v3.6b.3 (Guex & Peitsch, 1997; Peitsch et al., 1995; Schwede et al., 2003). Los mutantes de deleción de regiones C-terminales fueron modelados utilizando el servidor de modelado proteico Swissmodel, para secuencias oligoméricas, que utiliza PromodII para generar los modelos (en primer lugar basándose en el alineamiento superpone las estructuras tridimensionales, reconstruye los lazos, recompone el esqueleto peptídico, revisa las cadenas laterales basándose en una librería de rotámeros, comprueba el empaquetamiento y por último revisa la estructura generada) y Gromos96 para minimización de energía.

Los mutantes de sustitución por prolina generados en Swiss PdB Viewer fueron sometidos a un protocolo de minimización de energía y dinámica molecular en vacío utilizando el paquete de programas Gromacs (Berendsen, 1995) (Lindahl, 2001) con precisión doble. Las dinámicas moleculares se realizaron durante 100 picosegundos. Las trayectorias se analizaron con ayuda del

## Materiales y métodos ---

programa VMD (Humphrey, 1996) y aquellas estructuras características fueron seleccionadas y sometidas a minimización de energía.

Las representaciones de las moléculas han sido generadas utilizando el programa WebLab viewerPro3.7(msi).



### **3. Influencia del extremo C-terminal de GpA en el proceso de dimerización.**



### 3.1. INTRODUCCIÓN

La estructura tridimensional de una proteína de membrana depende en gran medida del empaquetamiento que tiene lugar entre los fragmentos que atraviesan la bicapa lipídica. Es por ello necesario estudiar los factores que intervienen en la interacción hélice-hélice para tratar de comprender el proceso de plegamiento. La presencia del motivo GxxxG no es exclusiva de la interfase de dimerización de la proteína GpA, sino que, como se ha explicado previamente, ha sido descrita como un motivo general de interacción entre hélices transmembrana, que facilita el empaquetamiento de este tipo de fragmentos, permitiendo el acercamiento entre hélices. Por ello, averiguar nuevos factores que contribuyen a mantener la estructura del dímero de GpA, puede facilitar la comprensión, no sólo del plegamiento de esta proteína sino también del plegamiento en general de aquellas otras proteínas de membrana que poseen este motivo o motivos similares.

Una de las características generales de las proteínas de membrana es la existencia de residuos cargados positivamente en los extremos de sus fragmentos hidrofóbicos. La presencia de estos residuos puede determinar la topología con la que la proteína se inserta y por tanto influir en el plegamiento final de ésta. En 1988 se propuso una hipótesis para la orientación de las proteínas en la membrana, según la cual uno de los determinantes de la topología de éstas es la distribución de cargas en los extremos del segmento transmembrana, “*positive inside rule*” (von Heijne & Gavel, 1988). Esta “regla” fue descrita, en primer lugar, para organismos procariotas, en los que se observó que los residuos cargados positivamente tenían preferencia por ocupar la cara citoplasmática de la membrana y posteriormente ha sido ampliada también para organismos eucariotas, mitocondrias y cloroplastos (Gavel & von Heijne, 1992; von Heijne & Gavel, 1988). Esta distribución asimétrica de residuos en la membrana ha sido correlacionada con la existencia de un potencial de membrana positivo en el exterior que favorecería la localización de residuos cargados positivamente hacia el interior celular (Andersson & von Heijne, 1994). Sin embargo, esta influencia es evidentemente mayor en el caso de organismos procariotas que en el de eucariotas, donde el potencial de membrana es muy pequeño o prácticamente inexistente. En el caso de estos últimos, parece ejercer un papel más importante la distribución asimétrica de fosfolípidos aniónicos (van Klompenburg et al., 1997a; van Klompenburg et al., 1997b). En las membranas de estos organismos se produce una acumulación de este tipo de fosfolípidos, como la fosfatidilserina, que se localizan casi exclusivamente en la monocapa interna de

la membrana plasmática, transportados de manera activa por la aminofosfolípido tranlocasa. De este modo, la monocapa interna de la membrana queda cargada negativamente favoreciendo la interacción de los residuos cargados positivamente en esta posición.

Si realizamos la comparación de secuencias de Glicoforinas de distintos organismos (figura 3.1) se observa la presencia de residuos cargados positivamente, con un elevado nivel de conservación, en la región C-terminal. Esta región corresponde a la zona de la proteína que queda orientada al citoplasma, de lo que se deduce que estas proteínas se ajustan a la “*positive inside rule*”.

<b>GpA_humana</b>	<b>EPEITLIIFGVMAGVIGTILLISYGIRRLIKK</b>
GLP_cerdo	HAEITGIIFAVMAGLLLIIFLIAYLIRRMICK
GLP_ratón	ALVMILIIILGVMAGIIGTILLISYCSRMTEK
GLP_MACFU	ELVIALIIFGVMAGVIGTILFISYGSRRLLIKK
GLP_caballo	QPVITVVIILGVMAGIIGIILLLAYVSRRLRKR

Figura 3.1: Alineamiento de secuencias del segmento transmembrana de Glicoforinas de distintos organismos. En naranja se encuadran los residuos de la región C-terminal cargados positivamente. En azul se destacan los residuos hidrofóbicos L89-Y93 de GpA humana y sus homólogos en los diferentes organismos.

Por otra parte, en la región C-terminal del fragmento transmembrana de GpA, existe una zona, (L<sup>89</sup>-Y<sup>93</sup>), la cual, según la estructura tridimensional del péptido transmembrana (MacKenzie et al., 1997) se encuentra perfectamente estructurada en  $\alpha$ -hélice, en la que encontramos otras dos características que resultan destacables. Por un lado, la presencia de un residuo aromático como es la tirosina 93. En proteínas de membrana los residuos aromáticos triptófano y tirosina tienden a situar sus cadenas laterales ancladas en la zona de la interfase lípido-agua interaccionando con las cabezas polares de los fosfolípidos y formando lo que se ha denominado un *cinturón aromático* (Killian & von Heijne, 2000) que contribuye a fijar la situación de las hélices con respecto a la perpendicular de la bicapa (figura 3.2).

También en esta región C-terminal encontramos el residuo isoleucina 91, residuo de cadena lateral voluminosa, que se orienta hacia la interfase hélice-hélice según se observa en la estructura del dímero GpA (MacKenzie et al., 1997) (figura 3.3), pudiendo contribuir al mantenimiento del ángulo entre las hélices, lo que podría gobernar el correcto empaquetamiento de la proteína.

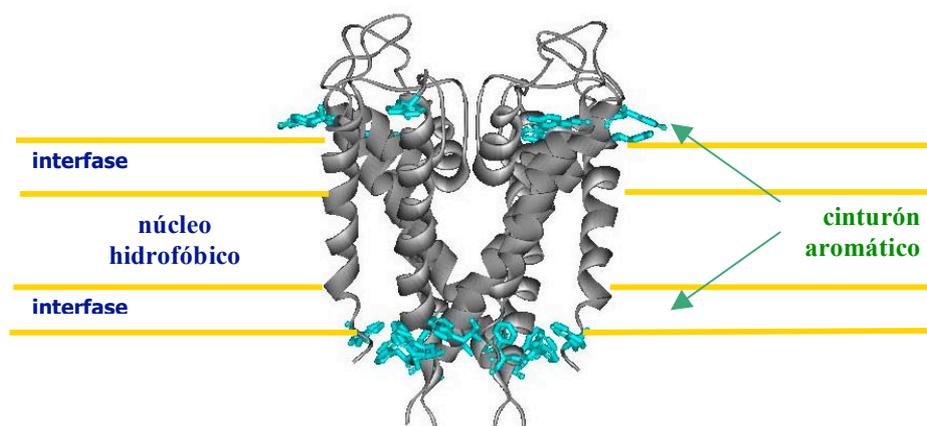


Figura 3.2. Estructura de una proteína de membrana (canal de potasio KcsA, código PDB, 1BL8) en la que se resaltan en azul los residuos aromáticos que se sitúan en la interfase lípido-agua, formando un cinturón aromático que ancla la proteína a la membrana

Sin embargo, estudios de mutagénesis por saturación (Lemmon et al., 1992b) no resaltaron la presencia de residuos cruciales para la formación del dímero de GpA en esta región. En experimentos en los que se realizó la inserción de alaninas a lo largo de todo el segmento transmembrana de GpA tampoco se observó pérdida de dimerización en los mutantes de esta región C-terminal (Mingarro et al., 1996). Con el objetivo de estudiar la importancia de esta zona en el proceso de plegamiento se llevó a cabo su delección obteniéndose el mutante  $\Delta 89/93$  (tabla 3.1) lo que provocó la desaparición del dímero. Este resultado pone de manifiesto cierta repercusión de estos aminoácidos para el empaquetamiento de GpA, lo que sumado al hecho de que el motivo de dimerización de esta proteína puede ser considerado como un modo más general de oligomerización entre fragmentos transmembrana, aumenta el interés por su estudio. Por este motivo esta sección del trabajo se dedicó a tratar de comprender la función que esta región C-terminal de GpA desempeña en la formación del dímero.

## 3.2. RESULTADOS

### 3.2.1 Dimerización de los mutantes de delección del extremo C-terminal de GpA.

Partiendo de los datos que se han presentado y con el objetivo de explicar la presencia de este extremo C-terminal (L<sup>89</sup>-Y<sup>93</sup>) del fragmento TM de GpA y su posible implicación en el correcto empaquetamiento entre las hélices, se diseñó una estrategia basada en la delección sucesiva de los residuos presentes en esta región, que se muestra en color rojo en la figura 3.3.

**<sup>72</sup>EITLIIFGVMAGVIGTILLISYGIRR<sup>97</sup>**

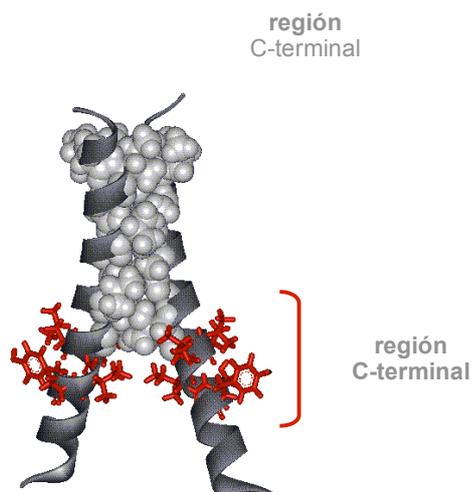


Figura 3.3: Secuencia y estructura del fragmento transmembrana de GpA. En la parte superior de la figura se muestra la secuencia de este fragmento donde se observan en negrita los residuos del motivo de dimerización y en rojo los residuos de la región C-terminal, objeto de estudio. En la parte inferior se muestra la estructura del fragmento TM de GpA (1AFO), en la cual se resaltan en color gris en representación de átomos como bolas con su radio de van der Waals los residuos del motivo de dimerización y en rojo como líneas los residuos de la región C-terminal.

En primer lugar, como se ha mencionado anteriormente, se eliminó la secuencia completa ( $\Delta 89/93$ ), lo que produjo la desaparición del dímero (fig.3.4, carrera 3). Este resultado muestra la necesidad de que esta zona esté presente para que el empaquetamiento hélice-hélice tenga lugar de modo correcto e implica una función no trivial para esta región de la hélice, a pesar de no contener ninguno de los siete residuos del dominio de dimerización de GpA. La delección de estos 5 residuos implica necesariamente el cambio de orientación de los aminoácidos restantes, puesto que eliminamos aproximadamente una vuelta y media de hélice, asumiendo el mantenimiento de la estructura helicoidal canónica que adopta el fragmento TM de GpA en el caso de este mutante de delección. Esta reorientación de los residuos en el mutante  $\Delta 89/93$  podría provocar que quedasen enfrentados los

residuos cargados situados C-terminal, lo que explicaría la pérdida de dimerización observada. Se realizaron, por ello deleciones parciales, en las que se eliminaron sólo tres residuos de esta zona, lo que corresponde a aproximadamente una vuelta de hélice, ya que estas deleciones, al contrario que en el caso anterior, no implican necesariamente un cambio de orientación entre los residuos enfrentados a la interfase de interacción. Sin embargo, los mutantes obtenidos ( $\Delta 89/91$  y  $\Delta 91/93$ ), mostraron también la pérdida del empaquetamiento (fig.3.4, carreras 4 y 5). En ninguna de las construcciones anteriores se observó capacidad de dimerización y puesto que todas tienen en común la deleción de la isoleucina situada en posición 91, se planteó la posibilidad de que pudiera tratarse de un residuo clave para la interacción hélice-hélice, ya que, como se ha mencionado previamente, este aminoácido se encuentra situado en la interfase entre las hélices TM (fig. 3.3 B). Por este motivo se realizó un mutante en el que se eliminaba únicamente este residuo,  $\Delta 91$ . Sin embargo, a diferencia de los casos anteriores, su deleción no afectó de forma significativa al proceso de dimerización (fig. 3.4, carrera 6).

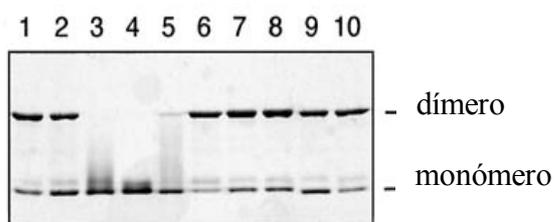


Figura 3.4: Dimerización de los mutantes del extremo C-terminal de GpA. Electroforesis al 12% SDS-PAGE en la que se muestran las extracciones de proteína de los mutantes de deleción y sustitución por leucina de la región C-terminal del fragmento TM de GpA. Carrera 1, proteína salvaje; carrera 2, doble mutante G79A V80A ; carrera 3, mutante  $\Delta 89/93$ ; carrera 4, mutante  $\Delta 89/91$ ; carrera 5, mutante  $\Delta 91/93$ ; carrera 6, mutante  $\Delta 91$ ; carrera 7, mutante 91/93L; carrera 8, mutante 91/95L; carrera 9, deleción de las argininas 96 y 97 ( $\Delta 96/97$ ); carrera 10, sustitución de las argininas 96 y 97 por leucina (96/97L).

En conjunto estos resultados, sugieren que la región C-terminal del fragmento TM de GpA es importante en el proceso de empaquetamiento hélice-hélice de GpA, puesto que los residuos que la constituyen no pueden ser eliminados, aunque su implicación en el proceso de dimerización no resulta evidente.

### **3.2.2 Substitución por Leucina de los residuos del extremo C-terminal.**

Los datos previos, en los que se habían realizado sustituciones individuales de los aminoácidos naturales del fragmento TM de GpA por un conjunto de residuos no cargados, mediante una estrategia de mutagénesis por saturación (Lemmon et al., 1992b), mostraron que la secuencia de la región C-terminal, no contenía ningún residuo que afectase de forma significativa a la dimerización de esta proteína. Con el objetivo de comprobar si sustituciones múltiples justificaban la pérdida del empaquetamiento, de modo que explicasen los resultados obtenidos con los mutantes de delección, se substituyeron grupos de tres y cinco aminoácidos de dicha región por residuos de leucina. La elección de este aminoácido se debe fundamentalmente a dos motivos, por un lado a la alta tendencia que presenta este residuo a formar hélices alfa en ambientes hidrofóbicos (Blondelle et al., 1997; Li & Deber, 1994), lo que permite mantener la estructura secundaria del segmento TM de GpA. Por otra parte, al hecho de que su amplia distribución en fragmentos transmembrana, no se ha visto correlacionada de modo global con su aparición en zonas implicadas en la función de estas proteínas, por lo que, generalmente, se le atribuye un papel meramente estructural.

La sustitución de los residuos I<sup>91</sup>-Y<sup>93</sup> por leucina en el mutante 91-93L (fig 3.4, carrera 7) da lugar a la formación del dímero GpA de modo casi indistinguible al de la proteína salvaje, e incluso en un mutante en el que se realiza la extensión de la sustitución por leucinas hasta el residuo isoleucina 95 (91-95L), el resultado es similar (fig 3.4, carrera 8). Estos resultados descartan una relevancia específica de secuencia de los residuos localizados en esta región del segmento TM.

### **3.2.3 Influencia de los residuos cargados presentes en el extremo C-terminal en la interacción hélice-hélice.**

Como se ha mencionado anteriormente, una de las características conservadas de la región C-terminal de GpA es la existencia de residuos cargados positivamente. Para tratar de comprobar la influencia que estos aminoácidos cargados ejercían sobre la capacidad de dimerizar de los mutantes de delección de esta región, se diseñó un nuevo grupo de construcciones. En primer lugar, se comprobó que ni la delección de las argininas situadas en posición 96 y 97 ( $\Delta$ 96/97), ni su sustitución por leucina

(96/97L), afectaba al proceso de dimerización en la proteína salvaje (fig.3.4, carreras 9 y 10). Del mismo modo, se estudió en el mutante  $\Delta 89/91$ , el efecto que producía la delección o sustitución

Tabla 3.1: Mutantes del extremo C-terminal de GpA y de la construcción 5L

Mutante	Secuencia	%Dímero
Wt	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGIRRLIKK <sup>101</sup>	84
G79A,V80A	<sup>72</sup> EITLIIFAAMAGVIGTILLISYGIRRLIKK <sup>101</sup>	54
$\Delta 89/93$	<sup>72</sup> EITLIIFGVMAGVIGTI-----GIRRLIKK <sup>101</sup>	0
$\Delta 89/91$	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGIRRLIKK <sup>101</sup>	0
$\Delta 91/93$	<sup>72</sup> EITLIIFGVMAGVIGTILL---GIRRLIKK <sup>101</sup>	9
$\Delta 91$	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGIRRLIKK <sup>101</sup>	86
91-93L	<sup>72</sup> EITLIIFGVMAGVIGTILLLLLLGIRRLIKK <sup>101</sup>	78
91-95L	<sup>72</sup> EITLIIFGVMAGVIGTILLLLLLLRRRLIKK <sup>101</sup>	77
$\Delta 89/91, \Delta 96/97$	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGI--LIKK <sup>101</sup>	2
$\Delta 89/91, 96/97L$	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGILLLIKK <sup>101</sup>	2
$\Delta 89/91, 96/97L, 100/101L$	<sup>72</sup> EITLIIFGVMAGVIGTI---SYGILLLILL <sup>101</sup>	73
$\Delta 91, \Delta 96/97$	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGI--LIKK <sup>101</sup>	0
$\Delta 91, 96/97L$	<sup>72</sup> EITLIIFGVMAGVIGTILL-SYGILLLIKK <sup>101</sup>	45
$\Delta 96/97$	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGI--LIKK <sup>101</sup>	69
96/97L	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGILLLIKK <sup>101</sup>	74
96/97L,100/101L	<sup>72</sup> EITLIIFGVMAGVIGTILLISYGILLLILL <sup>101</sup>	55
5L	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTILLISYGIRRLIKK <sup>101</sup>	34
5L, $\Delta 89/93$	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTI-----GIRRLIKK <sup>101</sup>	5
5L, $\Delta 89/91$	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTI---SYGIRRLIKK <sup>101</sup>	1
5L, $\Delta 91/93$	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTILL---GIRRLIKK <sup>101</sup>	8
5L, 91/93L	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTILLLLLLGIRRLIKK <sup>101</sup>	48
5L, 91/95L	<sup>72</sup> EITLIIFGVMLLLLLLAGVIGTILLLLLLLRRRLIKK <sup>101</sup>	50

por leucina de estos residuos cargados. En este caso, ni la sustitución por leucina ( $\Delta 89/91, 96/97L$ ), ni la delección de las argininas ( $\Delta 89/91, \Delta 96/97$ ), que podían interferir en el empaquetamiento, permitieron la dimerización de la proteína (tabla 3.1). Este hecho podría deberse a que en estos mutantes los residuos K<sup>100</sup> y K<sup>101</sup> quedaran orientados a la interfase hélice-hélice, impidiendo del mismo modo que lo harían la argininas el correcto empaquetamiento entre las hélices TM. Para comprobar esta hipótesis se obtuvo el mutante  $\Delta 89/91, 96/97L, 100/101L$  (tabla 3.1) en el que tanto las Arg como las Lys son reemplazadas por residuos de Leu, recuperándose en este caso la capacidad de

dimerización. En este sentido cabe destacar los resultados obtenidos con el mutante  $\Delta 91$  en el que la orientación de los residuos C-terminal cambia, por el hecho de haber eliminado un único residuo. En este mutante, la sustitución de las  $R^{96}$  y  $R^{97}$  por leucina ( $\Delta 91,96/97L$ ), no interfiere en la capacidad de dimerización (tabla 3.2) mientras que su delección ( $\Delta 91,\Delta 96/97$ ) (tabla 3.1), en la que la situación de los residuos  $K^{100}$  y  $K^{101}$  es equivalente a la de las  $R^{96}$  y  $R^{97}$  para los mutantes en los que se eliminan 3 residuos ( $\Delta 89/91$  ó  $\Delta 91/93$ ), puesto que corresponde a deleciones similares en los tres casos, no permite el empaquetamiento.

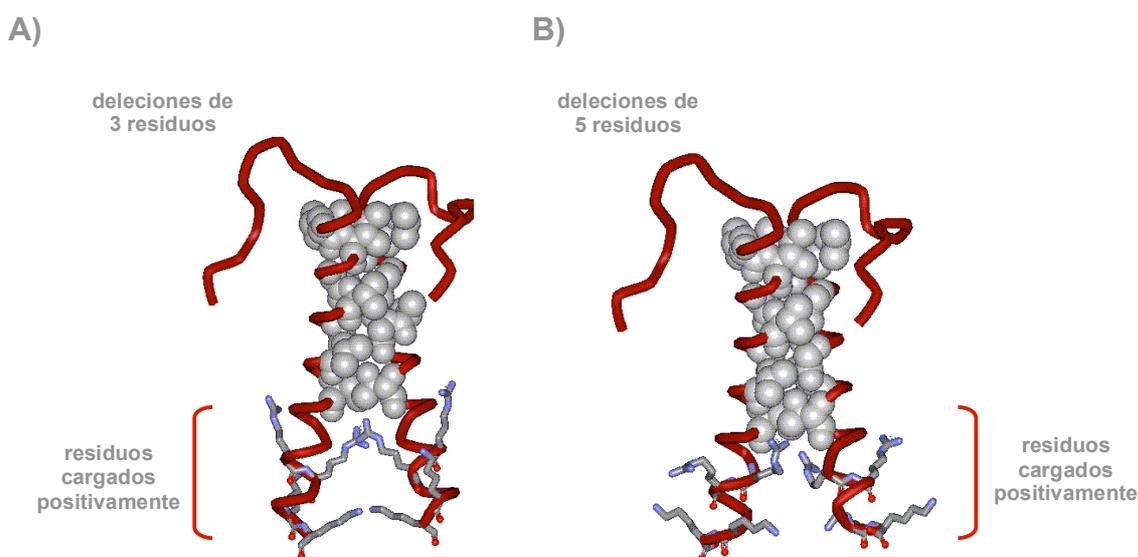


Figura 3.5: Estructura de los mutantes de GpA en los que se delecionan 3 y 5 residuos. Estas estructuras han sido obtenidas mediante modelado por homología respecto de la proteína salvaje, utilizando el programa swissmodel, como se ha descrito en materiales y métodos. A) mutante  $\Delta 89/91$ , B) mutante  $\Delta 89/93$ . Como bolas con su radio de van der Waals se muestran los residuos del motivo de dimerización y como líneas se destacan los residuos cargados positivamente ( $R^{96}$ ,  $R^{97}$ ,  $K^{100}$ ,  $K^{101}$ ) situados en el extremo C-terminal del fragmento.

El modelado por homología de estos mutantes muestra que, tanto en las deleciones de tres, (figura 3.5 A) como en las de cinco residuos (figura 3.5 B), los aminoácidos cargados cambian su situación respecto a la interfase de interacción de tal forma que su posición explicaría la pérdida de la dimerización por repulsiones electrostáticas en todos aquellos mutantes en los que se altere de forma significativa la distancia entre los residuos del motivo de dimerización y la región cargada positivamente.

#### **3.2.4. Estructura secundaria de los péptidos derivados del extremo C-terminal del fragmento TM de GpA.**

La estructura del dominio de dimerización de GpA obtenida por RMN (MacKenzie et al., 1997), muestra que la hélice se extiende desde el residuo E<sup>72</sup> hasta la L<sup>98</sup> (figura 3.3). En todos los estudios presentados hasta el momento se ha considerado que la región C-terminal de los mutantes estudiados conserva su estructura en hélice alfa por encontrarse en la zona transmembrana de la proteína. Sin embargo, con el objetivo de comprobar si en los mutantes de deleción la pérdida de la capacidad de dimerización podría estar relacionada de alguna forma con una pérdida de estructura secundaria en esta zona, se sintetizaron una serie de péptidos (figura 3.6) correspondientes a esta región, tanto de la proteína salvaje como de alguno de los mutantes de deleción.

Los espectros de dicroísmo circular en el ultravioleta lejano de estos péptidos (figura 3.6) realizados en presencia de SDS y a 25°C, muestran dos mínimos, uno a 208nm y otro a 222nm característicos de estructuras plegadas en hélice alfa (Holzwarth & Doty, 1965; Rohl & Baldwin, 1998), incluso en aquellos péptido como 88-99, 89/91 y 88-102, Δ89/91, derivados de mutantes en los que la capacidad de dimerización se había perdido. Estos resultados sugieren que la pérdida del empaquetamiento hélice-hélice de GpA en los mutantes de deleción está relacionada con la posición que ocupan los residuos cargados y no con la pérdida de estructura secundaria en esta zona del fragmento transmembrana.

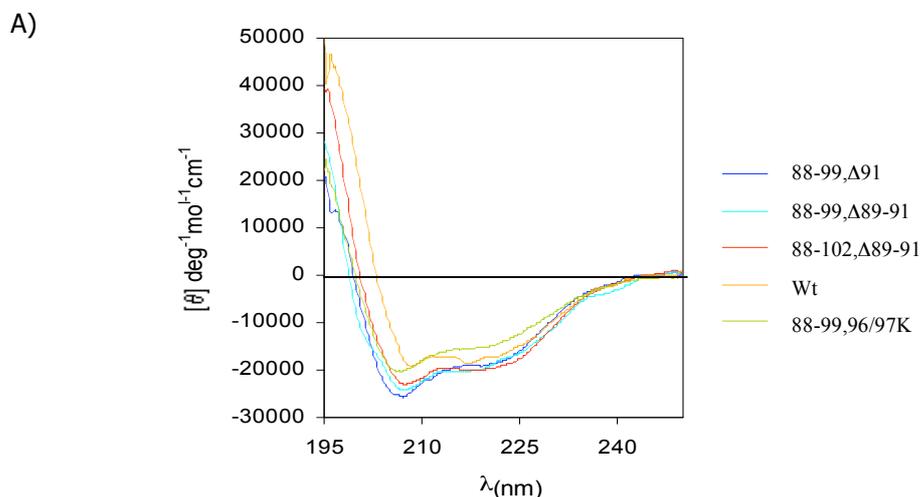


Figura 3.6: A) Espectros de dicroísmo circular de los péptidos derivados del extremo C-terminal de los mutantes de delección en presencia de SDS 3mM. B) Secuencias y valores de elipticidad molar a 222 nm obtenidos para estos péptidos.

### 3.2.5. Relevancia del extremo C-terminal en fragmentos transmembrana de mayor longitud.

Con el objetivo de comprobar si la longitud de hélice hidrofóbica es un factor capaz de intervenir en la dimerización que presentan los mutantes de delección, así como, para estudiar si en una interfase de interacción ligeramente distinta los resultados obtenidos son compatibles con la hipótesis de la función asignada al extremo C-terminal del segmento TM de GpA, se utilizó el mutante 5L, en el que se han insertado cinco residuos de Leu entre la Met que ocupa la posición 81 y la Ala 82 (Mingarro et al., 1997). Este mutante (tabla 3.2.) posee una interfase de dimerización distinta a la de la proteína salvaje (fig. 3.7), en la que el residuo G<sup>79</sup>, no parece ser tan relevante para el empaquetamiento entre hélices,

mientras que los residuos  $G^{83}$ ,  $V^{84}$  y  $T^{87}$  conservan su relevancia, resaltando el papel de la zona C-terminal del motivo de dimerización en el proceso de autoasociación (Mingarro et al., 1997). La estructura obtenida por modelado por homología frente al dímero GpA, muestra la relevancia de esta región C-terminal en el mutante 5L ya que los residuos del motivo de dimerización situados en N-terminal respecto de la inserción ( $L^{75}$ ,  $I^{76}$ ,  $G^{79}$  y  $V^{80}$ ), pierden su localización en la interfase hélice-hélice (fig. 3.7).

<sup>2</sup>EITLIIFGVMLLLLLAGVIGTLLISYGIRR<sup>97</sup>  
región C-terminal

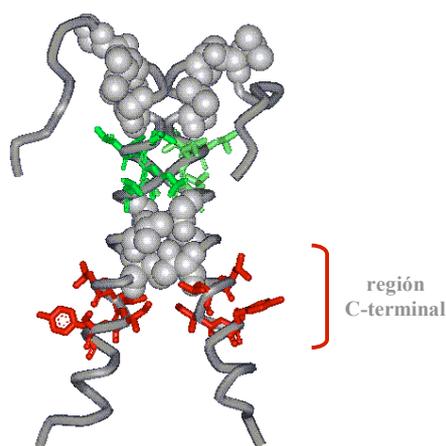


Figura 3.7: Secuencia del segmento transmembrana del mutante 5L y estructura obtenida por modelado por homología respecto a la proteína salvaje. En verde se muestran las 5 leucinas de la inserción y en rojo los residuos de la región C-terminal. Como puede observarse, el empaquetamiento de los residuos  $L^{75}I^{76}G^{79}V^{80}$  de la región N-terminal se ve seriamente dificultado.

La delección completa de la zona C-terminal de este mutante, 5L, $\Delta$ 89/93, así como las delecciones parciales 5L,  $\Delta$ 89/91 y 5L, $\Delta$ 91/93 mostraron resultados paralelos a los obtenidos en el caso de la proteína salvaje ya que en los tres casos el dímero desapareció (tabla 3.2). Es importante resaltar el hecho de que la longitud de hélice hidrofóbica del mutante 5L, $\Delta$ 89/93, es exactamente la misma que la de la proteína GpA salvaje; lo que demuestra que la pérdida de asociación entre las hélices en el mutante  $\Delta$ 89/93 no es simplemente consecuencia del acortamiento del fragmento hidrofóbico. La influencia de la longitud del fragmento hidrofóbico se ha demostrado en otros sistemas como en el estudio de la asociación entre hélices de polileucina, en las que la interacción interhelicoidal es mayor cuando se trata de hélices de 23 leucinas, que en el caso de hélices más pequeñas como las de 19 leucinas (Ren et al., 1999). Es importante señalar que la sustitución de los residuos objeto de estudio por leucina en el mutante 5L, mejora la capacidad de dimerización obtenida (mutantes 5L, 91-93L y

5L, 91-95L, (tabla 3.2), como sería de esperar dada la posición que ocupan en este mutante los residuos de arginina respecto a la interfase de dimerización, descrita para esta construcción y que se centra fundamentalmente en los residuos G<sup>83</sup>, V<sup>84</sup> y T<sup>87</sup>.

### 3.3. DISCUSIÓN

Los trabajos previos realizados sobre GpA se han centrado fundamentalmente en la identificación de residuos esenciales para la dimerización de esta proteína. Así ocurre en los estudios de mutagénesis por saturación realizados por Lemmon y colaboradores (Lemmon et al., 1992b), en los que se pone de manifiesto la existencia de un motivo de dimerización formado por siete residuos, así como, en los estudios estructurales (MacKenzie et al., 1997) en los que se presenta con mayor detalle el tipo de interacciones existentes entre las dos cadenas del dímero que forma GpA. Sin embargo, son numerosos los factores que contribuyen al establecimiento y permanencia del empaquetamiento entre hélices en ambientes de membrana. Este capítulo se ha centrado en el estudio de una región C-terminal del fragmento transmembrana de GpA en la que no parecen existir determinantes de secuencia que, al menos de modo individual, justifiquen su contribución a la estabilidad del dímero. Sin embargo, su delección provoca la desaparición de la interacción. El estudio de la función que el extremo C-terminal realiza en el proceso de dimerización de GpA se ha llevado a cabo utilizando diferentes estrategias. En primer lugar, la pérdida de la capacidad de dimerización en los mutantes de delección, indica que la presencia de esa región podría tener relevancia para la estructura de la proteína. Es por ello, que se aborda el estudio de las necesidades de secuencia global que pueda requerir GpA en la región C-terminal para permitir la oligomerización. En este aspecto, la presencia de dímero en los mutantes de sustitución por leucina (tabla 3.1) indica, que la relevancia de esta región no radica en la secuencia concreta que posee. Tampoco parece ser la clave de este hecho la estructura secundaria que adoptan en esta región los mutantes de delección de tres y cinco residuos, puesto que los estudios de dicroísmo circular, en presencia de SDS, de péptidos derivados de dicha zona, revelan que su estructura se mantiene en hélice alfa como ocurre para la secuencia salvaje (fig.3.6). El hecho de que la capacidad de adoptar estructura secundaria, no vaya ligada a la oligomerización de la proteína, puesto que algunas secuencias aún adoptando una estructura helicoidal no dimerizan, refuerza el modelo de plegamiento en dos estadios para proteínas de membrana y las distingue de los haces de

hélices alfa de proteínas solubles en las que el proceso de plegamiento de las hélices ocurre de forma concomitante a la formación de la estructura terciaria de la proteína. Además, estos resultados están de acuerdo con estudios realizados previamente en los que péptidos correspondientes a la región transmembrana de GpA mostraron que los procesos de adopción de estructura secundaria y oligomerización ocurren de forma independiente en distintos sistemas micelares (Fisher et al., 1999).

Entre las características que intervienen en determinar la topología de una proteína de membrana, se encuentra la presencia de residuos cargados positivamente en su cara citoplasmática. En la figura 3.1 se muestra, como se ha presentado previamente, que en la secuencia de GpA aparecen residuos de este tipo en el extremo C-terminal. Las estructuras obtenidas por modelado por homología (fig 3.5 A y B), apuntan a que el acercamiento y el cambio de orientación en la hélice de los residuos cargados hacia el motivo de dimerización, pueden ser responsables de la pérdida de dimerización de los mutantes de delección, puesto que quedarían enfrentados en la interfase de interacción hélice-hélice impidiendo el empaquetamiento. Este hecho se ve confirmado tras la recuperación del dímero obtenida en los diversos mutantes en los que los aminoácidos cargados positivamente han sido eliminados (tabla 3.1). Además en el caso de las delecciones de tres y cinco residuos, los resultados obtenidos por modelado (fig 3.5.A), muestran una estructura que no sería estable en una membrana donde la cadena polar de la arginina 96 se situaría próxima a la región hidrocarbonada, demasiado hidrofóbica para albergar este residuo polar.

Este conjunto de experimentos revela, en definitiva, que la función de la región C-terminal de GpA consiste en mantener a los residuos cargados positivamente a una distancia y orientación adecuadas respecto a los residuos del motivo de dimerización de esta proteína, de modo que la formación del dímero sea posible. Para tratar de generalizar los resultados obtenidos en este estudio y comprender mejor como estos aminoácidos situados aproximadamente en la interfase lípido-agua son capaces de influir en la estabilidad global de la proteína, sería de gran utilidad poder analizar como influyen estos aminoácidos adicionales al motivo de dimerización de siete residuos de GpA en el ensamblaje de proteínas en las que este motivo ha sido introducido. Sin embargo la mayoría de trabajos realizados utilizando GpA como proteína modelo no han incluido esta región en sus secuencias. Así por ejemplo, los estudios de dimerización *in vivo* en los que se utiliza el segmento TM de esta proteína para dirigir la dimerización de ToxR, incluyen únicamente los residuos 73-89 de GpA (Brosig & Langosch, 1998), o 73-87 (Russ & Engelman, 1999). La misma limitación se encuentra en

el sistema genético basado en el dominio de unión a DNA del represor lambda, donde el último residuo que se incluye es la isoleucina 95 (Leeds et al., 2001), justo el anterior a las argininas 96 y 97.

Entre los factores que se han demostrado capaces de influir en el empaquetamiento entre fragmentos transmembrana se encuentra la longitud de la hélice hidrofóbica (Ren et al., 1999), que puede determinar el ángulo de inserción de un segmento transmembrana en la bicapa lipídica y modificar, de ese modo, la orientación de los residuos que putativamente intervengan en un motivo de oligomerización. Los resultados obtenidos con los mutantes de delección realizados sobre la construcción 5L, son paralelos a los obtenidos para la proteína salvaje, confirmando que, la longitud de la hélice hidrofóbica no es un factor relevante para la pérdida de estructura de dichos mutantes. La construcción 5L, como se ha descrito previamente, muestra una interfase de interacción ligeramente distinta a la de la proteína salvaje, puesto que la inserción de 5 residuos entre los dos pares GV los sitúa necesariamente en caras opuestas de la hélice, habiéndose demostrado que el segundo par GV (G<sup>83</sup>V<sup>84</sup>) es fundamentalmente el responsable de la dimerización observada (Mingarro et al., 1997)(tabla 3.1). Estos resultados han sido corroborados mediante modelado por homología frente a la proteína GpA (figura 3.7) en la que se observa, que los residuos del motivo de dimerización situados en posiciones N-terminal (leucina 75, isoleucina 76, glicina 79 y valina 80), quedan excluidos en esta nueva interfase de interacción entre hélices. La importancia de la situación de los residuos cargados en esta última interfase, se refleja en el comportamiento paralelo que se observa entre los mutantes realizados sobre la proteína salvaje y sobre la construcción 5L. Resultados que apuntan que los datos obtenidos son extrapolables a otras interfases de interacción.

En cuanto al posicionamiento de los residuos respecto a la membrana en el caso del mutante 5LΔ89-93 existen dos posibles situaciones que se esquematizan en la figura 3.8. En la primera de ellas la inserción de 5 residuos de leucina desplazaría a la hélice transmembrana provocando que los aminoácidos situados en posición C-terminal quedasen fuera de la región hidrocarbonada. En este caso la delección de los residuos 89-93 resituaría a los residuos en la posición en la que se encuentran en la proteína salvaje, con la diferencia de que en este mutante la interfase de interacción quedaría alineada con ellos y la dimerización no sería posible. La segunda posibilidad es que la inserción de 5 leucinas desplace la hélice fuera de la región hidrocarbonada, pero hacia su región N-terminal (figura 3.8 derecha). También en este caso los residuos cargados quedarían enfrentados a la interfase de interacción, impidiendo el correcto empaquetamiento hélice-hélice, pero además, a diferencia del caso

anterior, quedarían situados en la región hidrocarbonada donde su localización es poco probable dado su carácter altamente polar. Así pues, en ambas situaciones el proceso de dimerización quedaría dificultado.

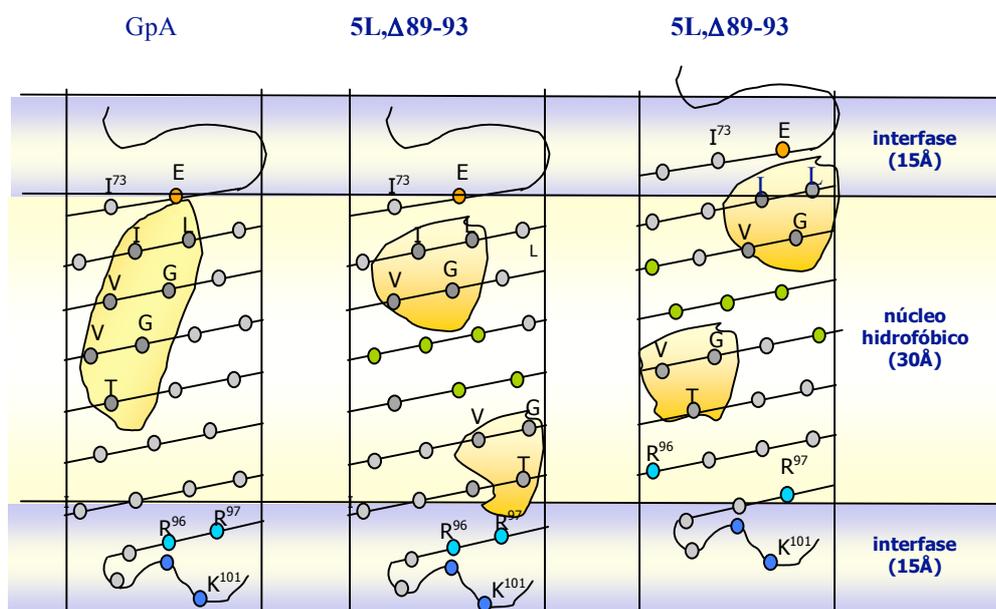


Figura 3.8: Esquema de la posible situación de los residuos de GpA y del mutante 5LΔ89-93 en una bicapa lipídica, en el que se muestran dos posibles modelos para dicho mutante. En negrita y rodeados por un círculo se muestran los residuos del motivo de dimerización de GpA. Como bolas azules se muestran los residuos cargados positivamente y en naranja los cargados negativamente.

La importancia de estos resultados radica en el hecho de que el motivo de dimerización GxxxG, se encuentra presente en numerosas proteínas transmembrana y su implicación en procesos de oligomerización se ha hecho patente mediante estudios diversos. Por ello comprender qué factores pueden estar contribuyendo a que un motivo de este tipo conduzca a un proceso de oligomerización, puede aportar datos que contribuyan al desarrollo de futuros programas de predicción de plegamiento de proteínas transmembrana. La posición que ocupan los residuos cargados situados en posición C-terminal y su orientación respecto a este motivo se convierte así en uno de esos factores a tener en cuenta en este tipo de predicciones. Dada la dificultad que existe en la actualidad para la obtención de

Influencia del extremo C-terminal de GpA en el proceso de dimerización \_\_\_\_\_

estructuras de alta resolución de proteínas de membrana, el desarrollo de programas de predicción de estructura, del mismo modo que ocurre en proteínas solubles, puede convertirse en el futuro en una herramienta esencial para el trabajo con este tipo de proteínas.

## **4. Oligomerización entre fragmentos transmembrana modelo.**



## 4.1. INTRODUCCIÓN

En paralelo al crecimiento exponencial en el número de proteínas conocidas y gracias a la aplicación de aproximaciones genómicas y proteómicas al estudio de los componentes celulares, se han ido desarrollando diversas estrategias para la búsqueda de putativos fragmentos hidrofóbicos que permiten, con éxito, la identificación de proteínas transmembrana. También se conocen numerosos datos del proceso de inserción y la topología que adoptan en la bicapa las proteínas integrales de membrana. De hecho, se han desarrollado algoritmos (revisado en (Chen et al., 2002)) que permiten la predicción teórica de la topología de este tipo de proteínas, como MPEX (S.Jaysinghe, K.Hristova, and S.H.White (2000) <http://blanco.biomol.uci.edu/mpex>) o HMMTOP (Tusnady & Simon, 1998). Sin embargo, mucho queda por entender sobre los factores que contribuyen al proceso de asociación lateral que ocurre entre los distintos fragmentos transmembrana para dar lugar a la estructura final de la proteína. Este proceso de asociación puede tener lugar entre distintos fragmentos dentro de una misma proteína, cuando se trata de proteínas politópicas, o entre fragmentos de cadenas polipeptídicas diferentes, cuando nos referimos a proteínas oligoméricas. Este tipo de interacciones hélice-hélice desempeña un papel esencial no sólo en el plegamiento individual de la proteína sino también en su función, permitiendo la asociación proteína-proteína e interviniendo en procesos de señalización o transporte, esenciales para la supervivencia celular. Numerosos procesos patológicos se producen como consecuencia de una asociación incorrecta entre fragmentos transmembrana, así encontramos que una mutación simple valina-glutámico en el segmento transmembrana del receptor tirosin quinasa codificado por el protooncogen *neu* provoca una activación constitutiva de dicho receptor que desencadena su actividad oncogénica (Bargmann et al., 1986), o alteraciones en el regulador de conductancia de la fibrosis cística, provocadas por la mutación de un residuo de prolina que provoca el plegamiento incorrecto de la proteína (Wigley et al., 2002), con repercusiones patológicas.

El diseño *de novo* de pequeñas proteínas que desempeñan una determinada función se ha convertido en una de las herramientas esenciales en biología, biotecnología y medicina (Martin & Vita, 2000). Sin embargo, para el diseño de estas miniproteínas es necesario, no sólo comprender el plegamiento y la función de las proteínas naturales, sino también minimizar los factores que contribuyen a dichos procesos para tratar de conseguir el máximo de propiedades en proteínas o péptidos de síntesis sencilla

y de función modulable. En este sentido, el estudio de la contribución de los diferentes residuos del motivo de dimerización de GpA al empaquetamiento de hélices transmembrana modelo, contribuiría a minimizar el número de residuos que deben ser incluidos en una secuencia proteica para conseguir que ésta sea capaz de oligomerizar y con ello poder desarrollar en el futuro, estructuras con potencialidad para formar canales en membranas biológicas.

Como se ha explicado anteriormente, el motivo de oligomerización de GpA se asume que está compuesto por estos siete residuos  $L^{75}IxxGVxxGVxxT^{87}$ . La inserción de estos aminoácidos sobre un segmento de polileucinas fue capaz de promover la formación de dímero (Lemmon et al., 1994). Por otra parte, la contribución de los diferentes residuos de dicho motivo de dimerización mediante la utilización de hélices de polivalina y polimetionina también ha sido investigada. En este tipo de esqueletos se observa una especial relevancia del motivo  $G^{79}xxxG^{83}$  en la oligomerización, mientras que la contribución del resto de residuos al proceso de empaquetamiento parece ser menor (Brosig & Langosch, 1998). En el presente capítulo se ha estudiado, la contribución individual de dichos residuos, al proceso global de empaquetamiento hélice-hélice; para lo que se ha procedido a la inserción gradual de los aminoácidos que conforman el *motivo de dimerización* de GpA en un esqueleto homogéneo de polileucinas. La elección de este residuo frente a los esqueletos previamente estudiados se debe, fundamentalmente, a su abundancia en fragmentos transmembrana naturales. Se trata del residuo más abundante en hélices transmembrana, ya que en torno a un 16% del número total de aminoácidos en este tipo de fragmentos son leucinas, mientras que un 8% son valinas y sólo un 4% metioninas (Liu et al., 2002). Además, en estudios sobre la distribución de aminoácidos en proteínas integrales de membrana se observa que los residuos de leucina, no muestran preferencia por encontrarse en posiciones concretas en los fragmentos TM (Ulmschneider & Sansom, 2001), lo que refuerza su contribución fundamentalmente estructural en este tipo de proteínas.

Por otra parte, la longitud media de los fragmentos transmembrana naturales se encuentra en torno a los 21-22 residuos, pero el escenario real es, que podemos encontrar proteínas que atraviesan la bicapa utilizando fragmentos transmembrana de tamaños muy variables. Así pues, se ha demostrado que fragmentos hidrofóbicos de sólo 12 residuos, para los que resulta difícil entender la capacidad de atravesar la bicapa lipídica a menos que las cadenas hidrocarbonadas de los fosfolípidos se acomoden creando una zona en la que su espesor disminuya localmente (figura 4.1.a), son suficientes para dirigir

la inserción en membranas de retículo endoplásmico (Whitley et al., 1996). En contraposición, encontramos también proteínas cuyos fragmentos hidrofóbicos se extienden hasta 27 residuos, tan largos, que es necesario pensar en una orientación de dicho fragmento distinta de la perpendicular a la bicapa (Arkin & Brunger, 1998) que, combinada de nuevo con una adaptación de los fosfolípidos en su entorno inmediato (figura 4.1.b), posibilitaría el que no tenga lugar la exposición de los residuos hidrofóbicos que lo forman al ambiente polar del medio acuoso circundante (Arkin & Brunger, 1998) (figura 4.1.b).

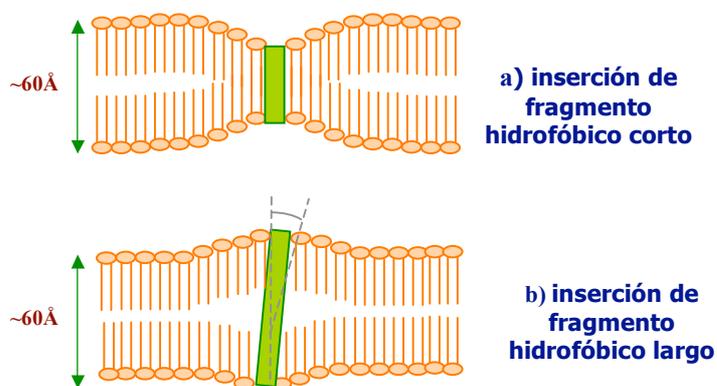


Figura 4.1: Modelos de inserción de fragmentos TM de longitudes extremas (representados como rectángulos) en la bicapa lipídica.

Tanto la hidrofobicidad, como la inclinación que debe adoptar un fragmento transmembrana en la bicapa pueden condicionar su capacidad de asociación. En este capítulo, se ha tratado de averiguar cuáles son los residuos mínimos necesarios para inducir dimerizaciones entre fragmentos TM artificiales y de comprender de que modo la longitud del fragmento hidrofóbico influye en el proceso de interacción entre este tipo de hélices.

## 4.2.RESULTADOS

### 4.2.1. Contribución de los residuos implicados en el proceso de dimerización.

#### 4.2.1.1. Inserción secuencial de los residuos del motivo TM de GpA en un fragmento de 18L.

El primer paso para llevar a cabo el estudio de la influencia que cada uno de los residuos del motivo de GpA ejerce en la capacidad de inducir dimerización, se realizó mediante la introducción gradual de dichos aminoácidos en un fragmento de 18 leucinas. La elección de un molde de 18 residuos se debe

principalmente a que proporciona una longitud de la región hidrofóbica suficiente para atravesar la bicapa lipídica, al tiempo que, la aparición de fragmentos de este tamaño en proteínas transmembrana naturales es relativamente frecuente (Arkin & Brunger, 1998). Para realizar estas construcciones se introdujeron sitios de corte para endonucleasas de restricción en ambos extremos del segmento TM de GpA (ver materiales y métodos tabla 2.2, sección 2.4) y mediante apareamiento entre oligonucleótidos complementarios se introdujeron los fragmentos de polileucina entre dichos sitios de restricción. De esta manera se obtuvieron los mutantes G/L18 y GG/L18 y el resto de construcciones se prepararon a partir de ellas realizando mutagénesis dirigida. Los resultados obtenidos tras la inserción gradual de estos residuos se muestran en la figura 4.2.

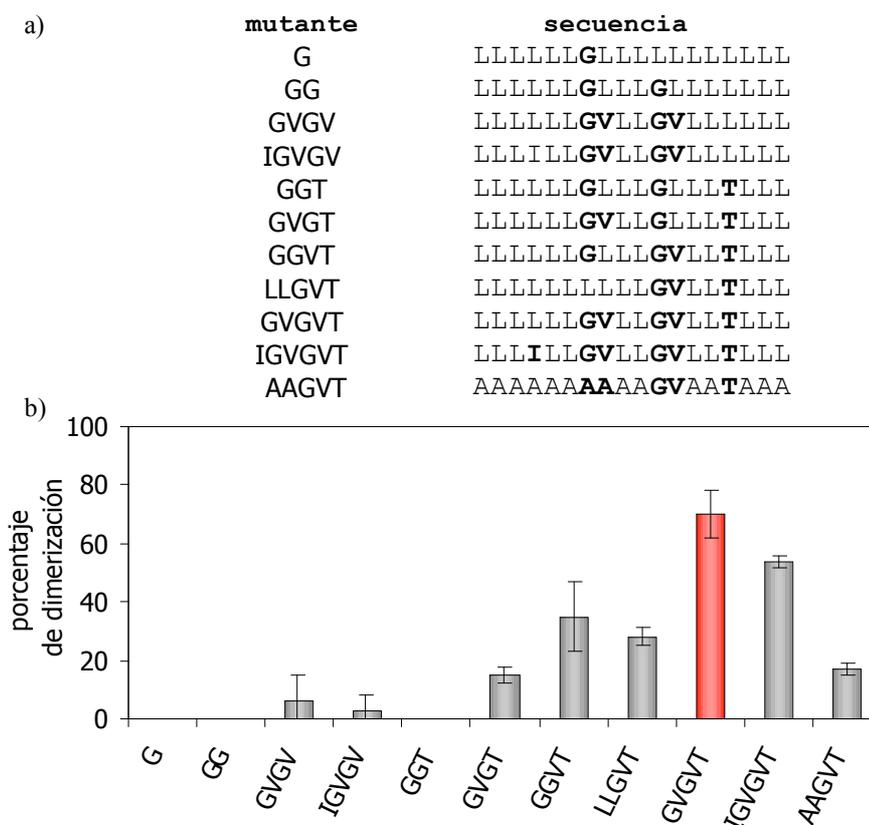


Figura 4.2: Inserción gradual de los residuos del motivo de GpA en un esqueleto de 18 residuos hidrofóbicos.(a) Secuencias de la región transmembrana y (b).porcentajes de dimerización obtenidos para las diferentes construcciones.

Los porcentajes de dimerización indican que la interfase GVGVT es, con diferencia, la más adecuada para la inducción de oligómeros en un fragmento de 18L. También se consiguen niveles de dimerización considerables, aunque nunca tan elevados, en los mutantes GGVT y LLGVT. La presencia de isoleucina en posición 76, (mutantes IGVGVT e IGVGV), no parece ser relevante para el empaquetamiento, ya que su introducción no aumenta los porcentajes de dimerización, a pesar de que el residuo equivalente a éste en el segmento TM de GpA, es uno de los siete aminoácidos componentes del motivo de oligomerización. El porcentaje de dímero obtenido para el mutante LLGVT es del mismo orden que el que presenta la construcción 5L (Tabla 2.2, materiales y métodos) de GpA. En este mutante, como se ha descrito previamente, son estos residuos ( $G^{83}$ ,  $V^{84}$  y  $T^{87}$ ) del extremo C-terminal del motivo de oligomerización, los responsables directos del empaquetamiento de la proteína. Es de destacar, que la secuencia de aminoácidos en la interfase del dímero es similar en estos dos mutantes y acorde con ello también lo es su grado de dimerización (comparar tabla 3.1, capítulo 3, con figura 4.2.b). Para evaluar, la capacidad de inducir dimerización de este motivo en un esqueleto diferente, se introdujeron los residuos GVT sobre un fragmento de polialaninas (mutante AAGVT). En esta nueva construcción también se observó la aparición de dímero, aunque el porcentaje de dimerización obtenido fue ligeramente inferior al que presentaba su homólogo sobre un esqueleto de polileucinas (figura 4.2.b).

Para averiguar si la interfase de dimerización que utiliza el mutante GVGVT es similar a la que utiliza GpA, se diseñó la síntesis del péptido SAT 1 (figura 4.3.a). Este péptido, representa la mínima secuencia de GpA que contiene el motivo GVGVT. En ambos extremos del esqueleto peptídico se incluyó un *cassette* RG con el objetivo de mejorar la solubilidad y facilitar la purificación del péptido. Se realizaron experimentos de competencia proteína GVGVT/péptido SAT1. En estos experimentos, como se ha descrito en materiales y métodos, se procedía a incubar la proteína (4 $\mu$ M) en presencia del péptido en tampón TBS 0.5%SDS durante una hora a 25°C, proceso tras el cual las muestras fueron sometidas a electroforesis. La incubación no dio lugar a la aparición de heterodímero proteína-péptido, incluso en presencia de concentraciones elevadas de péptido. Debido a que SAT1 es un péptido pequeño en el que se incluían aminoácidos de la proteína salvaje GpA, la falta de competencia podía deberse a diferencias en la estructura helicoidal de éste, más que a los aminoácidos que participan directamente en la interacción. Por ello, se diseñó un péptido ligeramente más largo, en

el que los aminoácidos no incluidos en el motivo GVGVT fueron substituidos por residuos de leucina. De este modo, se procedió a la síntesis y purificación del péptido SAT2. (figura 4.3.a). Los experimentos de competencia mostraron, del mismo modo que para SAT1, la incapacidad de este péptido para formar heterooligómeros (figura 4.3.b). Se procedió, por ello, a la síntesis del segmento transmembrana completo de GpA (SAT3), con el objetivo de averiguar si la interfase de oligomerización salvaje era distinta de la que encontrábamos en los mutantes de polileucina o por el contrario, la ausencia de interacción con los péptidos anteriores debía atribuirse a su pequeña longitud.

La síntesis de fragmentos TM, como se ha descrito en materiales y métodos, resulta complicada debido a la agregación de los péptidos durante el propio proceso de síntesis. Esto provoca la aparición de péptidos truncados, lo que dificulta en gran medida, no sólo el proceso sintético sino su purificación posterior. Por ello se adaptó un protocolo (Melnik et al., 2001), en el que se utiliza una química ligeramente distinta y se acoplan *cassettes* de lisinas en los extremos del fragmento TM para mejorar su posterior solubilidad, además de utilizarse dobles y triples acoples en diferentes posiciones de la secuencia (Fisher & Engelman, 2001) (ver materiales y métodos, sección 2.4).

Nombre	Secuencia
SAT 1	RG <b>GV</b> MAG <b>VI</b> GTGR
SAT 2	RGL <b>GV</b> LL <b>GV</b> LL <b>TL</b> GR
SAT 3	KKKKEITLIIIF <b>GV</b> MAG <b>VI</b> GT <b>ILL</b> ISYGK

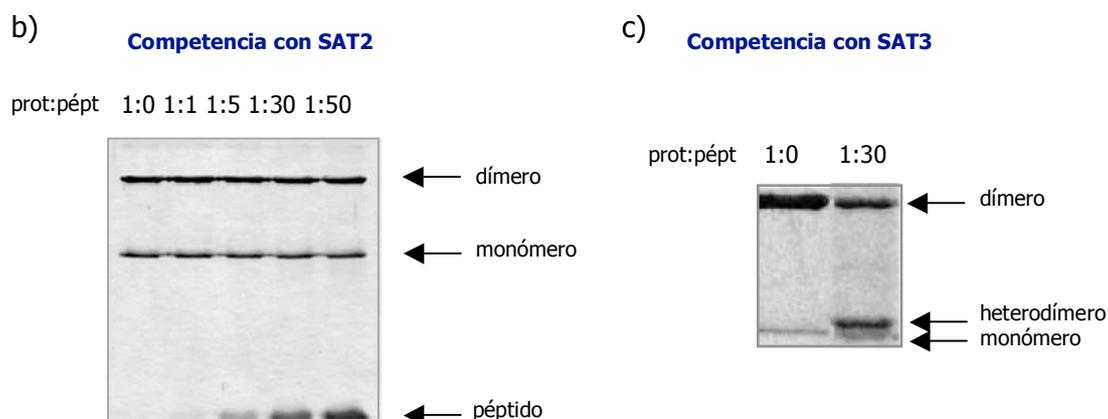


Figura 4.3: Estudios de competencia proteína/ péptido. a) Secuencia de los péptidos hidrofóbicos sintetizados. b) Competencias de los péptidos SAT2 y SAT3 con el mutante GVGVT.

De este modo, se repitieron los experimentos de competencia, con el péptido SAT3. En estos experimentos, como se muestra en la figura 4.3.c, se observó la aparición de una banda de movilidad correspondiente al heterodímero proteína/péptido concomitante a la disminución proporcional de la especie dimérica. Este resultado sugiere que la interfase de dimerización es similar a la de la proteína salvaje para los mutantes GGVV/T (figura 4.3b).

Una vez demostrado que el motivo GVGVT es el mínimo necesario para inducir dimerización y que ésta ocurre utilizando una interfase similar a la de la proteína salvaje, el comportamiento de alguna de estas construcciones fue analizado en una hélice de longitud superior (21L). Una hélice de 18 residuos se extiende por unos 27 Å (1.5 Å de traslación por residuo, en un fragmento estructurado como una hélice alfa canónica), mientras que una hélice de 21 aminoácidos se extiende por 31,5 Å. Como la amplitud de la región hidrocarbonada de la bicapa lipídica se estima del orden de 30Å, se realizaron algunos de los mutantes sobre un esqueleto de 21 leucinas con el objetivo de comprobar si la longitud del segmento hidrofóbico altera la autoasociación de las proteínas. Como se muestra en la figura 4.4 no encontramos diferencias significativas entre los porcentajes de dimerización obtenidos con los dos esqueletos de longitudes diferentes, lo que implica que es la secuencia GVxxGVxxT, la mínima requerida para la autoasociación de fragmentos TM modelo.

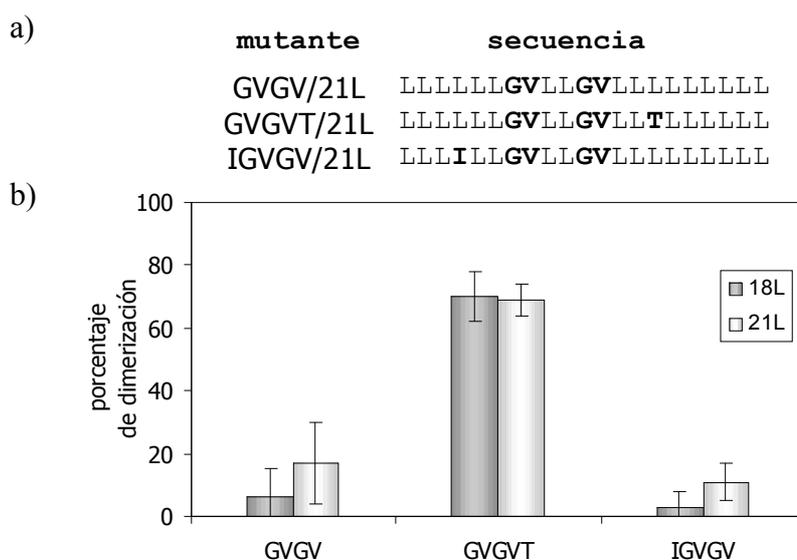


Figura 4.4: Dimerizaciones en esqueletos de 18 y 21 Leu. a) Secuencia del fragmento TM de los mutantes de 21 leucinas. b) Porcentajes de dimerización obtenidos para dichos mutantes frente a sus homólogos de 18L

#### 4.2.1.2. Contribución de la leucina 75 al proceso de dimerización.

Debido a que en la proteína salvaje la leucina 75 del fragmento transmembrana forma parte del motivo de dimerización de 7 aminoácidos (Lemmon et al., 1994) y, ya que en todos los mutantes esta posición está ocupada por un residuo de leucina, por tratarse del esqueleto hidrofóbico elegido, se decidió estudiar el efecto que tenía la sustitución de dicho residuo en la posición equivalente de la construcción GVGVT. Para ello, se diseñaron una serie de mutantes, en los que dicha leucina era substituida por los aminoácidos Ala, Ile o Val, tres residuos muy abundantes en fragmentos TM, de carácter hidrofóbico y en el caso de los residuos Val e Ile con la cadena lateral  $\beta$ -ramificada, lo que podría dificultar la autoasociación. Se estudió el efecto que dicha sustitución provocaba sobre la capacidad de inducir dimerización mediada por el motivo GVGVT. Los resultados obtenidos se muestran en la figura 4.5, donde se puede observar que la presencia de leucina en la posición 75 no parece ser relevante para el empaquetamiento ya que su substitución por otros residuos permite la obtención de un nivel de dimerización muy similar.

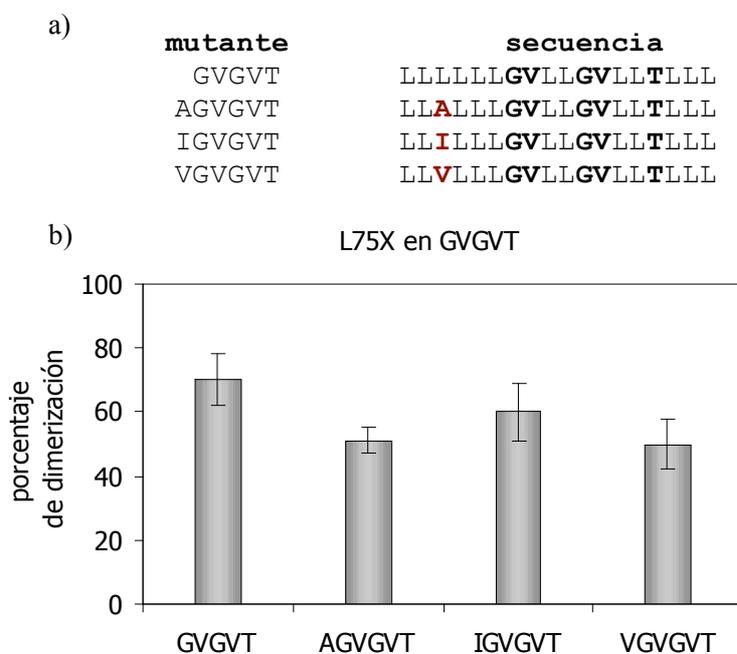


Figura 4.5: Influencia del residuo situado en posición 75. a) Secuencia de los mutantes de GVGVT (18L) para la posición equivalente a la leucina 75 de GpA. b) Porcentajes de dimerización obtenidos para dichos mutantes.

#### 4.2.1.3. Contribución del residuo en posición 87 al proceso de dimerización.

La diferencia encontrada en el grado de dimerización entre los mutantes GVG $\bar{V}$  y GVG $\bar{V}$ T, apunta a un papel muy relevante para el residuo de treonina en este tipo de interacciones. Por otro lado, éste es el único aminoácido polar presente en el motivo minimizado. En el segmento TM de GpA el grupo hidroxilo de este residuo está formando un puente de hidrógeno con el carbonilo del esqueleto péptidico de la valina 84 de la hélice adyacente, contribuyendo de forma significativa a la estabilización de la estructura dimérica (Smith et al., 2002). Además, ha sido descrita la capacidad de inducir la oligomerización de hélices de polileucina mediada por residuos como la asparagina, (Choma et al., 2000; Zhou et al., 2000), mientras que no ocurre lo mismo con el residuo de treonina. Por ello, decidimos sustituir en el mutante GGT, que no dimeriza, el residuo “original” de treonina por el aminoácido asparagina, así como por un residuo  $\beta$ -ramificado como es la isoleucina con objeto de testar si la presencia de estos residuos resultaba ser un determinante estructural suficientemente potente como para inducir homodimerización en nuestras condiciones experimentales, convirtiéndose en un determinante estructural más relevante que la propia treonina. Sin embargo, en ninguno de estos dos casos se indujo la dimerización (figura 4.6).

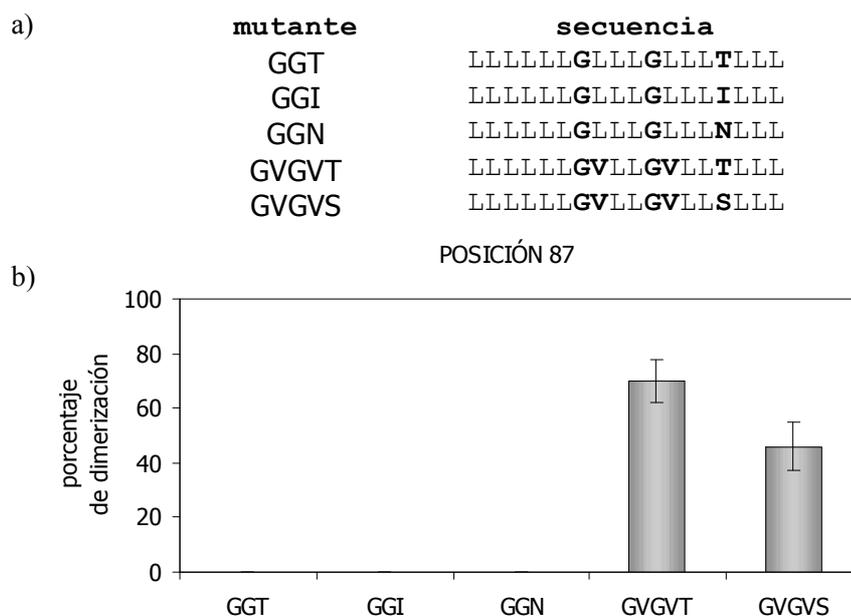


Figura 4.6: Contribución a la dimerización del residuo situado en posición 87 a) Secuencias del segmento TM de los mutantes de la posición equivalente a la treonina 87 de GpA en esqueletos de GG (18L) y GVG(18L). b) Porcentajes de dimerización obtenidos para dichos mutantes.

Por último, se realizó la sustitución por serina, que posee al igual que la treonina un grupo hidroxilo en su cadena lateral sobre el mutante GVGVT, manteniéndose del mismo orden, aunque ligeramente inferiores, los porcentajes de dimerización. Este resultado parece indicar que en estos mutantes, el grupo hidroxilo de la treonina (o la serina) podría estar interaccionando con la valina en posición 84, de modo similar a como ocurre en GpA, lo que por otro lado explicaría el hecho de que esta valina 84 sea más importante para el dímero que la que ocuparía la posición 80 de GpA, como se pone de manifiesto también, en los diferentes porcentajes de dimerización obtenidos para los mutantes GVGT y GGVT (15 y 35 por cien, respectivamente) (figura 4.2).

#### **4.2.2. Efecto de la longitud del fragmento hidrofóbico en la capacidad de dimerización mediada por el motivo GVGVT/18L.**

Con los datos que aportó el estudio de la minimización del motivo de dimerización en los fragmentos de 18 residuos hidrofóbicos, se eligió la interfase GVGVT, como punto de partida para el estudio de la influencia que la longitud del fragmento hidrofóbico puede tener en la capacidad de inducir oligomerizaciones entre hélices TM. Para abordar este estudio, se utilizaron hélices de polileucina de distintos tamaños y se valoró la capacidad de dimerización de los distintos mutantes. La longitud de hélice hidrofóbica se eligió, siempre teniendo en cuenta el rango de tamaños de hélices TM presentes en la naturaleza (Arkin & Brunger, 1998). Los resultados obtenidos se muestran en la figura 4.7, en la que puede observarse que el grado de dimerización obtenido para los distintos mutantes no es constante, observándose una disminución en la cantidad de dímero en los extremos del gráfico.

a)

<b>mutante</b>	<b>secuencia</b>
GVGVT(15L)	LLL <b>GVLLGVLLT</b> LLL
GVGVT(18L)	LLLLLL <b>GVLLGVLLT</b> LLL
GVGVT(21L)	LLLLLL <b>GVLLGVLLT</b> LLLLLL
GVGVT(24L)	LLLLLLLLLL <b>GVLLGVLLT</b> LLLLLL
GVGVT(27L)	LLLLLLLLLL <b>GVLLGVLLT</b> LLLLLLLLLL

b)

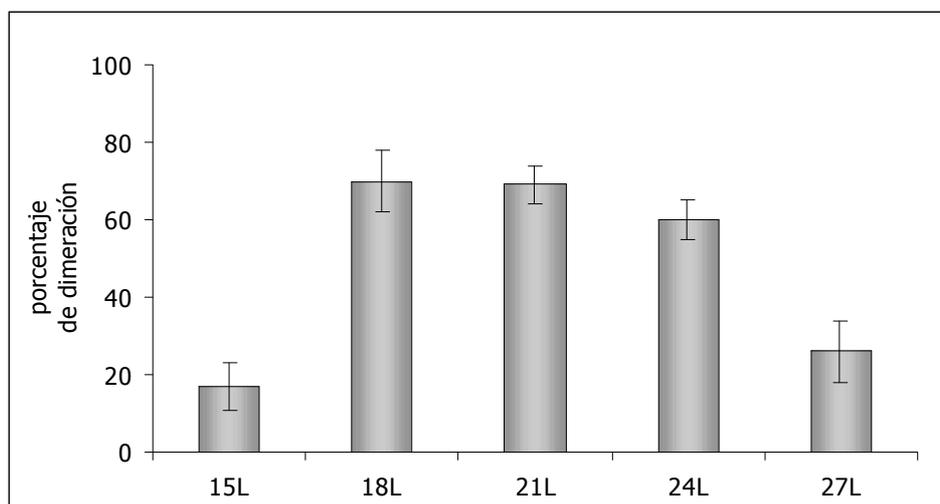


Figura 4.7: Influencia de la longitud del fragmento hidrofóbico en la dimerización del motivo minimizado.

a) Secuencia del fragmento TM de las construcciones con diferente número de leucinas. b)

Porcentajes de dimerización obtenidos para las diferentes longitudes de fragmento TM.

Con la finalidad de verificar si la pérdida de dímero en el mutante GVGVT/15L era consecuencia del acortamiento de la distancia entre el motivo de dimerización y los residuos del extremo N-terminal, nos planteamos el diseño del mutante LLLGVLLGVLLTLLLLLL, (18Lb), cuya longitud total es de 18L y en el que esta distancia en la posición N-terminal es la misma que la que posee el mutante de longitud 15L (figura 4.8). Los resultados muestran que este mutante de 18L pierde en gran medida aunque no totalmente la capacidad de dimerización.

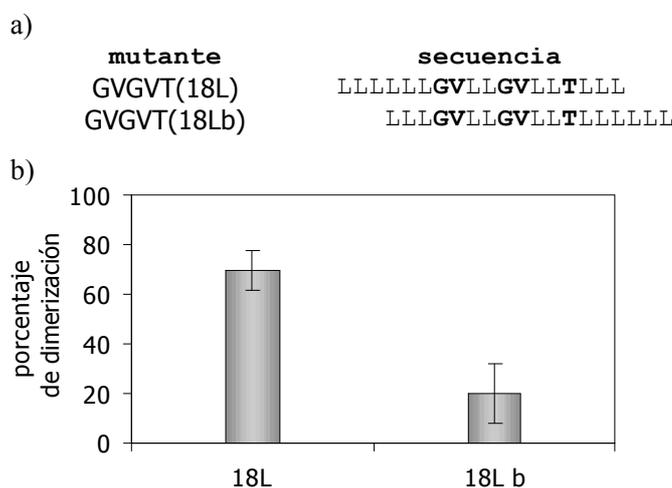


Figura 4.8: Efecto de la posición del motivo GVGVT en fragmentos hidrofóbicos de la misma longitud. a) Secuencia del fragmento TM de los mutantes de 18 leucinas. b) Porcentajes de dimerización obtenidos en función de la distinta longitud hidrofóbica en posición N-terminal.

Este resultado parece indicar que la distancia a los residuos situados en N-terminal puede condicionar el empaquetamiento de las hélices, de forma similar a como ocurría con los residuos situados en C-terminal en la interfase de dimerización de GpA (capítulo anterior). Por otra parte, el menor porcentaje de dimerización obtenido para las contrucciones con fragmentos transmembrana cortos, podría explicar la incapacidad de los péptidos SAT1 y SAT2, de 9 y 11 residuos respectivamente, para heterooligomerizar con la contrucción GVGVT.

### 4.3. DISCUSIÓN

Los resultados obtenidos mediante la adición secuencial de residuos del motivo de dimerización de GpA a un esqueleto de 18 leucinas, nos indican la existencia de una gradación en la aportación de cada uno de los aminoácidos al proceso de empaquetamiento. La presencia de glicinas, no es suficiente para inducir el proceso de oligomerización, sobre el esqueleto de polileucinas. La adición de treonina, tampoco es capaz de inducir formación de dímero, a menos que alguna de las valinas del motivo esté

presente. En este sentido, parece ser más importante la valina situada en posición C-terminal (V84), ya que mientras que con la adición de la primera (mutante GVGT) obtenemos en torno a un 15% de dimerización, el mutante que posee la segunda (GGVT) alcanza porcentajes de dímero del 35%. Sin embargo, la presencia de ambas genera, sin duda, un mejor empaquetamiento, alcanzándose un 70% de dímero, de acuerdo con lo observado en la estructura de RMN de GpA (MacKenzie et al., 1997). En este empaquetamiento, parece ser crucial la presencia de la treonina, puesto que el motivo GVG $\bar{V}$  por sí solo presenta valores de dimerización por debajo del 10%. Dado que, como se ha mencionado previamente, en el segmento TM de GpA, el grupo hidroxilo de la treonina 87 se encuentra formando un puente de hidrógeno con el carbonilo del esqueleto peptídico de la valina 84 de la hélice enfrentada (Smith et al., 2002), podemos esperar un comportamiento similar en el esqueleto de polileucinas. De este modo, el empaquetamiento y la orientación hélice-hélice que genera el motivo GVG $\bar{V}$ , sería necesario para la formación de dicho puente de hidrógeno, lo que explicaría los altos porcentajes de dimerización alcanzados con el motivo GVGVT. De hecho, la sustitución por serina, aminoácido que también posee el grupo hidroxilo, mantiene un grado de dimerización notable, aunque en cualquier caso el empaquetamiento alcanzado es mejor en el caso de la treonina. Es interesante destacar que en ningún caso la dimerización se ve favorecida por la adición de la isoleucina 76 del motivo de GpA, más bien perjudica a este proceso, así pues la adición de dicho residuo sobre el motivo GVGVT provoca una pérdida en la capacidad de autoasociación cercana al 20%. Tampoco la presencia de leucina en posición 75 parece crucial para la inducción del empaquetamiento (figura 3.5). Estos datos apoyan los resultados obtenidos por Lemmon y colaboradores en 1994 (Lemmon et al., 1994), que ya sugerían un papel menos importante para estos dos residuos en el proceso de dimerización. La comparación con resultados previos muestra que en el sistema TOXCAT, (descrito en la Introducción, sección 1.5) (Russ & Engelman, 1999), utilizando esqueletos de 13 residuos de polivalinas o polimetioninas, la adición de sólo dos glicinas del motivo de dimerización es suficiente para inducir la formación del dímero. Aunque resulta difícil comparar resultados entre estos dos estudios, puesto que los sistemas utilizados son distintos, hay que destacar algunas diferencias significativas. En primer lugar, en este sistema, un esqueleto de polivalinas es capaz de autoasociarse de modo que la adición de glicinas, que es realmente la introducción del motivo GVG $\bar{V}$  por la naturaleza del esqueleto peptídico, genera aproximadamente un 11% de dímero más que el propio esqueleto, dato similar al 6% obtenido

para nuestro sistema (Senes et al., 2000). En el caso de polimetioninas los niveles de asociación alcanzados sólo con la introducción de las glicinas son del mismo orden (14%), lo que probablemente tiene que ver con la naturaleza del esqueleto peptídico. La metionina es un residuo más polar, que la leucina y la valina, de modo que la autoasociación, que minimiza la exposición de residuos al medio hidrofóbico, puede estar favorecida en este esqueleto. En nuestros resultados, en el caso de la introducción del motivo GVT en el esqueleto de polialaninas, se obtienen porcentajes de dimerización menores que los observados para el mismo motivo introducido sobre una hélice de polileucina. Sin embargo, debido a que el residuo de alanina es más polar que el aminoácido leucina deberíamos esperar valores de asociación más elevados. Por el contrario, parece que la pérdida de interacciones proteína-lípido, consecuencia de la autoasociación, no es compensada en este caso por la ganancia en interacciones proteína-proteína. Probablemente porque el estrecho empaquetamiento que se consigue en el modelo de entrantes y salientes de GpA no puede ser reproducido en un esqueleto de polialaninas, debido al menor tamaño de su cadena lateral. Si observamos la figura 4.5, aunque la cantidad de dímero del mutante AGVGVT es elevada, se observa una disminución respecto a la construcción GVGVT. Es probable que esa leucina en posición 75, aunque no indispensable, sí contribuya en cierta medida a la interacción, de modo que en el esqueleto de polialaninas la formación de dímero se vea ligeramente disminuida.

Los experimentos de competencia proteína:péptido muestran, por un lado, la incapacidad de péptidos cortos para promover la formación de heterodímeros con el mutante GVGVT. Por otra parte, demuestran que un péptido con la secuencia del fragmento TM de GpA, sí es capaz de inducir la formación de dichos heterodímeros, indicando que el mutante GVGVT utiliza una interfase de dimerización similar a la de la proteína salvaje.

La longitud de fragmentos transmembrana naturales, como se ha comentado anteriormente, es muy variable. Estudios realizados sobre fragmentos TM naturales muestran distribuciones de longitud como las que se presentan en la figura 7.9.

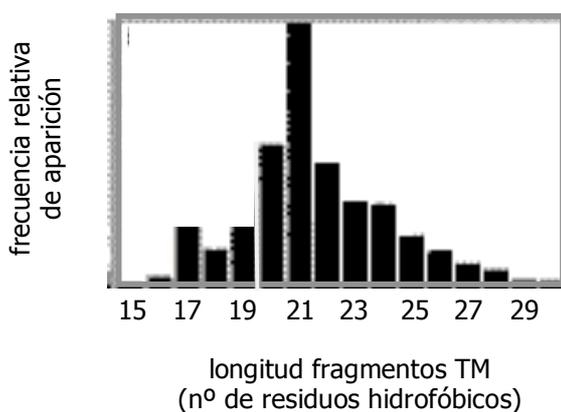


Figura 4.9. Estudio estadístico de la longitud de los segmentos hidrofóbicos en fragmentos TM naturales. Adaptado de (Arkin & Brunger, 1998).

Como puede observarse en la figura 4.9 las longitudes más frecuentes se encuentran en torno a 21 residuos, por lo que se repitieron algunos de los mutantes en un esqueleto de estas dimensiones obteniéndose niveles de dimerización similares a los obtenidos para los mutantes de 18L.

En proteínas de membrana la concordancia entre la longitud de la hélice hidrofóbica y la anchura de la bicapa, es un fenómeno de importancia biológica que se ha visto asociado a procesos tan relevantes para la célula como el direccionamiento de las proteínas desde el Golgi a la membrana plasmática. En la membrana plasmática hay una acumulación de colesterol y esfingolípidos de cadena larga que provocan que la anchura de esta bicapa sea mayor a la de otras membranas celulares. Relacionado con este hecho se ha observado que los fragmentos transmembrana de las proteínas ubicadas en la membrana plasmática son unos 5 residuos como promedio más largos (Munro, 1995). Además también se ha visto relacionada esta concordancia de tamaños con un control de las interacciones hélice-hélice (Ren et al., 1999). Cuando un fragmento TM es más largo que la bicapa en la que se va a insertar, la asociación hélice-hélice se ve favorecida, ya que esta interacción minimiza los contactos hélice-lípido y por tanto los cambios estructurales en la bicapa (Munro, 1995).

Como se ha descrito previamente un fragmento de polileucinas plegado en hélice alfa de un tamaño de 18 residuos se extiende aproximadamente 27Å, mientras que 21 residuos se extenderían 31.5 Å. Ambas longitudes de fragmento podrían ser capaces de cubrir la región hidrocarbonada de la bicapa lipídica. Sin embargo, fragmentos más largos o más cortos probablemente necesitarían reorganizaciones en la bicapa o en su propia orientación para atravesar la membrana. En nuestro sistema de estudio encontramos una situación similar, ya que la región hidrofóbica de los péptidos debería ser cubierta por el diámetro de la micela y una micela de NaSDS posee un diámetro medio en torno a 60Å (Kim et al., 2000), similar a la longitud de la perpendicular de la bicapa.

El comportamiento del motivo GVGVT al ser introducido en fragmentos de polileucina refleja que, en situaciones extremas, la capacidad de oligomerización de dicho motivo se reduce. En el caso de fragmentos pequeños, este hecho parece estar relacionado, con la distancia que presenta el motivo respecto a los residuos cargados, situados en posición N-terminal. Un fragmento de 18 leucinas, en el que esa distancia se reduce a tres residuos, como en el caso del mutante de 15 leucinas, se comporta

del mismo modo que éste, perdiendo la capacidad de dimerización (18Lb). En N-terminal, estos mutantes poseen residuos de glutámico, flanqueando las polileucinas, que podrían fijar la posición del péptido en la bicapa, impidiendo su homodimerización. El ángulo que deben adoptar las hélices entre sí para poder interactuar mantendría, en el caso del mutante 18Lb, el motivo de dimerización demasiado cercano a la región N-terminal, de modo que la situación de los residuos cargados negativamente impediría la oligomerización (figura 2.10 a). Esto no ocurriría en el caso del mutante 18L, a pesar de mantener la misma longitud de región hidrofóbica, puesto que el motivo de dimerización minimizado ocupa una zona más central, aumentando el ángulo de cruce entre las hélices en el extremo N-terminal (figura 2.10 b).

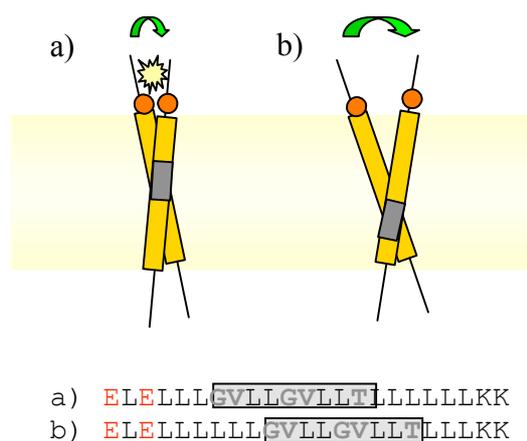


Figura 4.10: Modelo y secuencias de inserción de los dos mutantes de longitud de hélice hidrofóbica 18L.

El resultado obtenido para el mutante 27L, también se podría explicar en términos de ángulo adoptado por la hélice TM con respecto a la bicapa. Mantener residuos hidrofóbicos fuera de la membrana es energéticamente muy desfavorable, de modo que para cubrirlos por la región hidrocarbonada es necesario adoptar un ángulo, que en el caso del mutante 27L, podría no ser compatible con la formación de oligómeros. El balance energético existente entre mantener residuos hidrofóbicos expuestos al medio, favoreciendo la interacción entre monómeros (de modo que las

interacciones proteína-proteína aumenten y disminuyan las de tipo proteína-lípido) y, variar la inclinación del fragmento (de modo que quede minimizada la exposición de residuos al entorno), parece inclinarse en este caso hacia esta segunda opción, impidiendo la formación de dímeros.

En definitiva estos resultados apoyan las hipótesis que barajan una posible contribución de la relación entre la longitud de los fragmentos transmembrana y el espesor de las membranas, como elemento regulador de la interacción proteína-proteína. Es por ello que se podrían encontrar proteínas, que no fuesen capaces de oligomerizar en la membrana del Golgi y sí en la membrana plasmática o viceversa. Del mismo modo, la incorporación de una proteína a islotes de lípidos de composición definida "*rafts*", podría cambiar su capacidad de interacción con otras proteínas.



## **5. Prolinas en fragmentos transmembrana.**



## 5.1. INTRODUCCIÓN

La naturaleza química del residuo prolina le confiere una serie de características que lo diferencian del resto de aminoácidos que componen las proteínas. Su cadena lateral se cicla para formar un enlace con el grupo amino del esqueleto peptídico rindiendo un anillo de pirrolidona y convirtiendo, por lo tanto a la prolina, en un iminoácido (figura 5.1).

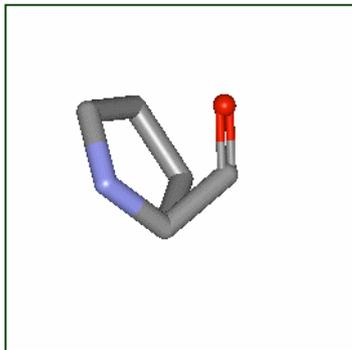


Figura 5.1: Residuo de prolina. En azul se muestra el nitrógeno de la cadena lateral y en rojo el oxígeno del grupo carboxilo.

Por este motivo, la capacidad de formar puentes de hidrógeno mediados por el grupo amino que poseen el resto de residuos, se pierde en el caso de la prolina. Estas características diferenciales, provocan que su tendencia a formar parte de estructuras secundarias canónicas como las hélices alfa o las hojas beta quede limitada. La rigidez estructural que introduce el anillo de pirrolidona provoca que los ángulos que es capaz de adoptar este residuo sean prácticamente incompatibles con estructuras plegadas en hoja  $\beta$  (figura 5.2), generando la aparición de protuberancias en estas estructuras. Mientras que para hojas  $\beta$  el ángulo *phi* suele encontrarse entre -120, -140 grados, en el caso de la prolina dicho ángulo queda restringido por el enlace covalente de su cadena lateral a -60 grados. A esto hay que sumar su incapacidad para formar parte de la red de puentes de hidrógeno que mantiene esta estructura.

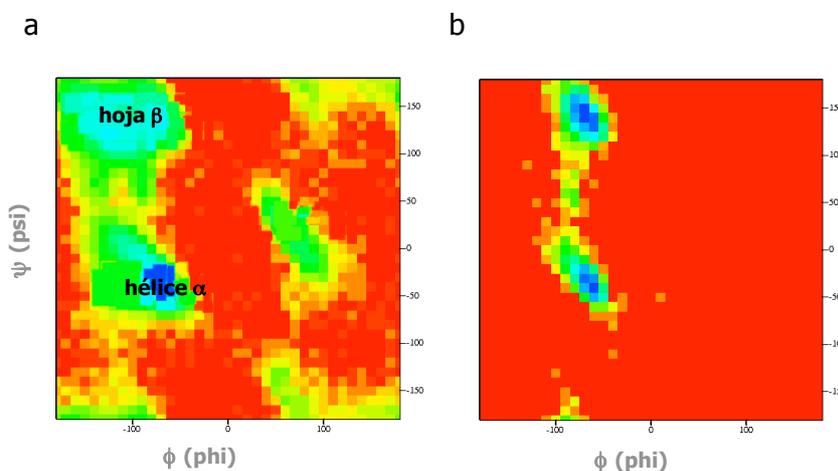


Figura 5.2: Diagrama de Ramachandran de la distribución de los ángulos phi y psi (a) en proteínas y (b) para el residuo de prolina.

En el caso de hélices alfa, la participación del grupo amino del esqueleto peptídico en la formación del ciclo lo incapacita para formar los puentes de hidrógeno de la estructura canónica con el aminoácido situado en posición  $i-4$ . Al mismo tiempo, el volumen de anillo de pirrolidona afecta a las posiciones de los residuos situados en N-terminal, distorsionando el eje de la hélice, afectando también a la formación de puentes de hidrógeno en posición  $i-3$ . Por todos estos motivos, en proteínas globulares, que se encuentran en ambientes acuosos y que tienen por ello la posibilidad de formar puentes de hidrógeno con el entorno, la presencia del residuo prolina se suele asociar a la existencia de zonas de la proteína desestructuradas, lazos, o a ciertas posiciones N-terminales de estructuras helicoidales contribuyendo al “*capping*” o terminación de las hélices (Aurora & Rose, 1998; Serrano & Fersht, 1989). Sin embargo, las proteínas de membrana, como hemos visto, se localizan en un medio hidrofóbico, con el que en general no existe esa posibilidad de formar puentes de hidrógeno. Por ello, todos los grupos polares de las cadenas peptídicas, tienden a formar puentes de hidrógeno intramoleculares, adoptando estructuras regulares tales como las hélices alfa o los barriles  $\beta$ . La aparición en la membrana de un residuo de prolina, que deja libres los carbonilos en posición  $i-3$  e  $i-4$  en un ambiente hidrofóbico, parece, en términos energéticos, más problemática. Sin embargo y pese a estas características la frecuencia de aparición de este residuo en proteínas de membrana es más elevada de lo que cabría esperar (Arkin & Brunger, 1998; Bywater et al., 2001), observándose en

muchos casos que no queda comprometida la estructura helicoidal, a pesar de su presencia (Deisenhofer, 1995).

Desde un punto de vista estructural la presencia de este residuo en segmentos transmembrana se ha visto a menudo asociada a la aparición de deformaciones en el eje de la hélice denominadas acodamientos (*kinks*) (Bywater et al., 2001; Sansom & Weinstein, 2000; von Heijne, 1991). Un acodamiento se caracteriza por provocar un cambio en la dirección del eje de la hélice, de modo que podríamos definir un ángulo  $\theta$  entre el fragmento de hélice anterior a la prolina y el posterior (figura 5.3)

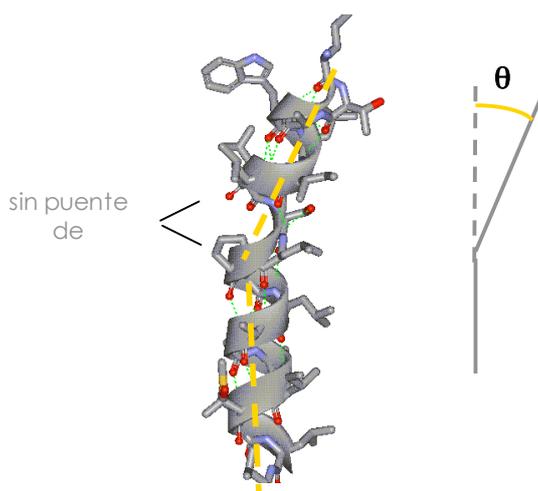


Figura 5.3. Acodamiento provocado por la presencia de un residuo de prolina en una hélice alfa canónica.

La flexibilidad conformacional intrínseca que los residuos de prolina confieren a las hélices, sugiere la posible existencia de un papel dinámico para estos acodamientos pudiendo intervenir en la interconversión entre diferentes estados proteicos como respuesta a estímulos externos. Este papel ha sido puesto de manifiesto para diversos canales, como por ejemplo, en la apertura de las uniones Gap Cx32 reguladas por voltaje (Ri et al., 1999). En la secuencia primaria de este tipo de canales de agua, existe una prolina muy conservada que se encuentra formando un *acodamiento*. La apertura diferencial del ángulo de éste, que se produce como consecuencia de un cambio en la red de puentes de

hidrógeno que mantienen la estructura proteica, podría explicar la transición del canal desde su posición cerrada a su posición abierta. Un papel similar se le ha atribuido a la aparición de este residuo en el caso de los canales de potasio regulados por voltaje (Sansom & Weinstein, 2000). Así como, la presencia de residuos de prolina también parece tener un papel crucial en el funcionamiento de receptores, como es el caso de los receptores acoplados a proteínas G (GPCRs).

Por otro lado recientemente, se le ha atribuido al residuo de prolina un papel relevante en el proceso de plegamiento correcto e inserción de proteínas transmembrana. Los segmentos transmembrana son ricos en residuos  $\beta$ -ramificados que, aunque en la bicapa lipídica se encuentran formando parte de hélices  $\alpha$ , en ambientes solubles, favorecen la formación de estructuras en hoja  $\beta$  con alta tendencia para la agregación (Arkin & Brunger, 1998). En el regulador de conductancia de la fibrosis cística se ha encontrado un residuo de prolina que parece estar implicado en evitar intermediarios de plegamiento incorrectos de la proteína que conduzcan a su inactivación (Wigley et al., 2002), relacionándose este hecho con su incapacidad para formar parte de estructuras en hoja  $\beta$ . En la estabilización de la estructura correcta de una proteína, intervienen numerosos aspectos, entre los que se encuentra, la desestabilización de estructuras alternativas que conduzcan a la obtención de una proteína inactiva. Durante el proceso de inserción y plegamiento, una proteína transmembrana tiene que atravesar ambientes diversos como son el poro acuoso del traslocón y el ambiente proteico en el camino de salida de éste para incorporarse a la membrana. El residuo prolina, al menos en este caso, parece jugar un papel esencial en evitar que en este camino hacia la inserción y plegamiento de las proteínas en la membrana, se produzcan intermediarios de plegamiento no deseados, en concreto estructuras en hoja  $\beta$  que pueden conducir a la agregación de la proteína naciente.

Aunque la pérdida del patrón de puentes de hidrógeno que provoca la presencia de un residuo como la prolina en una hélice alfa, se puede considerar un problema en el proceso de inserción de los fragmentos transmembrana, en el escenario real de la membrana biológica, la energía de transferencia desde el medio acuoso a la bicapa de una hélice transmembrana compuesta únicamente de aminoácidos hidrofóbicos, supera el mínimo necesario para alcanzar la inserción estable (Popot & Engelman, 2000). Así pues, la introducción de residuos que generan diversidad en estas hélices, como la prolina, podría ser interesante para, como se deduce de los ejemplos anteriores, introducir tanto funcionalidad como flexibilidad en este tipo de fragmentos proteicos.

## 5.2. RESULTADOS

### 5.2.1. Substitución por prolina de los residuos del segmento TM de GpA.

La aproximación que se utilizó para comprender el efecto que un residuo como la prolina podía tener en el empaquetamiento entre hélices TM, fue la sustitución de cada uno de los residuos del fragmento TM de GpA por dicho aminoácido y el estudio de sus consecuencias sobre la capacidad de dimerización de esta proteína. Para ello, mediante mutagénesis dirigida, se cambiaron los residuos naturales desde el E<sub>72</sub> hasta la R<sub>97</sub> por prolina y tras la expresión y purificación de las proteínas mutantes, se evaluó en gel de poliacrilamida el porcentaje de dimerización de cada muestra, obteniéndose los resultados que se muestran en la figura 5.4

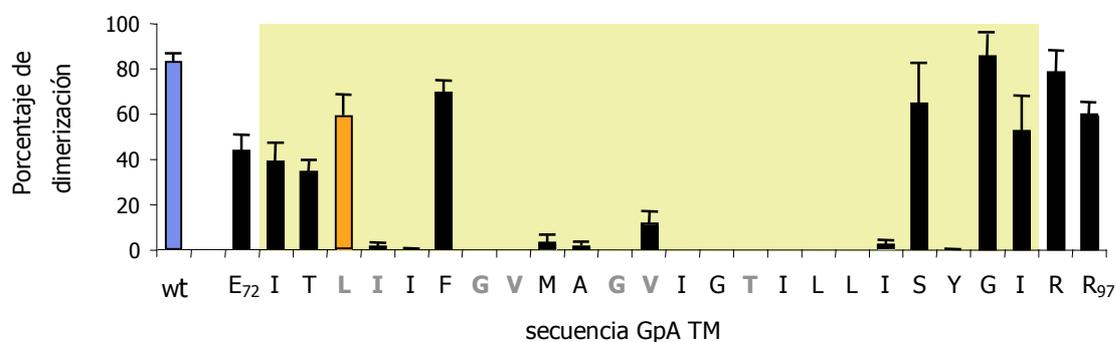


Figura 5.4: Porcentajes de dimerización obtenidos para las sustituciones por prolina de los residuos del segmento TM de GpA. Cada punto representa el resultado del análisis de tres muestras independientes  $\pm$  su desviación estándar. Los residuos del motivo de dimerización de GpA se muestran en gris. En color azul se muestra el porcentaje de dimerización de la proteína salvaje y en naranja se destaca el del mutante L75P. La banda amarilla corresponde a la localización aproximada de la región hidrofóbica de la membrana (~30Å).

Como se puede observar, los datos muestran que, de modo global existe una mayor permisividad a la presencia de prolinas en los extremos del fragmento transmembrana, mientras que, su presencia en la zona central provoca la desaparición del dímero. Entre estos resultados cabe



mutación introducida. Sin embargo, este hecho no parece deberse a un modo de interacción distinto entre monómeros, ya que mezclas equimoleculares de ambos péptidos muestran una única banda de movilidad electroforética intermedia (figura 5.6, carrera 4). Es importante destacar el hecho de que la presencia de residuos de lisina en los extremos del fragmento, facilita la solubilidad de estos sin afectar al proceso de dimerización.

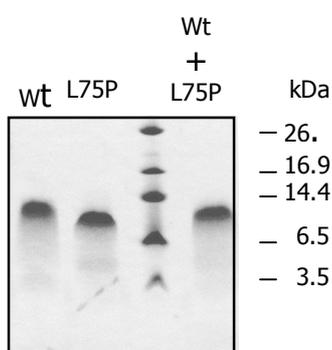


Figura 5.6: Electroforesis de Tris-tricina de los péptidos wt y L75P. En las dos primeras carreras se aplicaron los péptidos por separado y en la cuarta aparece una mezcla equimolar de los dos péptidos.

### 5.2.3. Interfase de interacción proteína/proteína del mutante L75P.

#### 5.2.3.1. Estudios de competencia proteína /péptido.

Con la finalidad de estudiar la especificidad de la interacción entre la interfase que presenta el mutante L75P y la de la secuencia salvaje, se realizaron una serie de experimentos de competencia proteína/péptido. En estos experimentos, como se ha explicado en materiales y métodos (sección 2.3.5), se realiza una incubación a 25°C durante 1h entre la proteína y el péptido correspondientes. De este modo, podemos averiguar si el péptido es capaz de competir con la proteína por la formación del dímero observando, de forma concomitante, la aparición de una banda de heterodímero proteína-péptido de movilidad diferente. Estos ensayos se realizaron a diferentes concentraciones de péptido. El análisis de los resultados obtenidos en los geles de electroforesis permitió la cuantificación de cada una de las especies presentes. Se realizó el ajuste de los datos experimentales de porcentajes de heterodimerización a una ecuación hiperbólica para cada una de las competencias (figura 5.7). De este modo, se calculó el valor de la concentración de péptido a la que el porcentaje de heterodímero es la mitad del porcentaje de heterodímero máximo (IC50) lo que permitió la comparación de este

parámetro en las diferentes interfases de interacción. Los resultados de IC50 se muestran en la tabla 5.1.

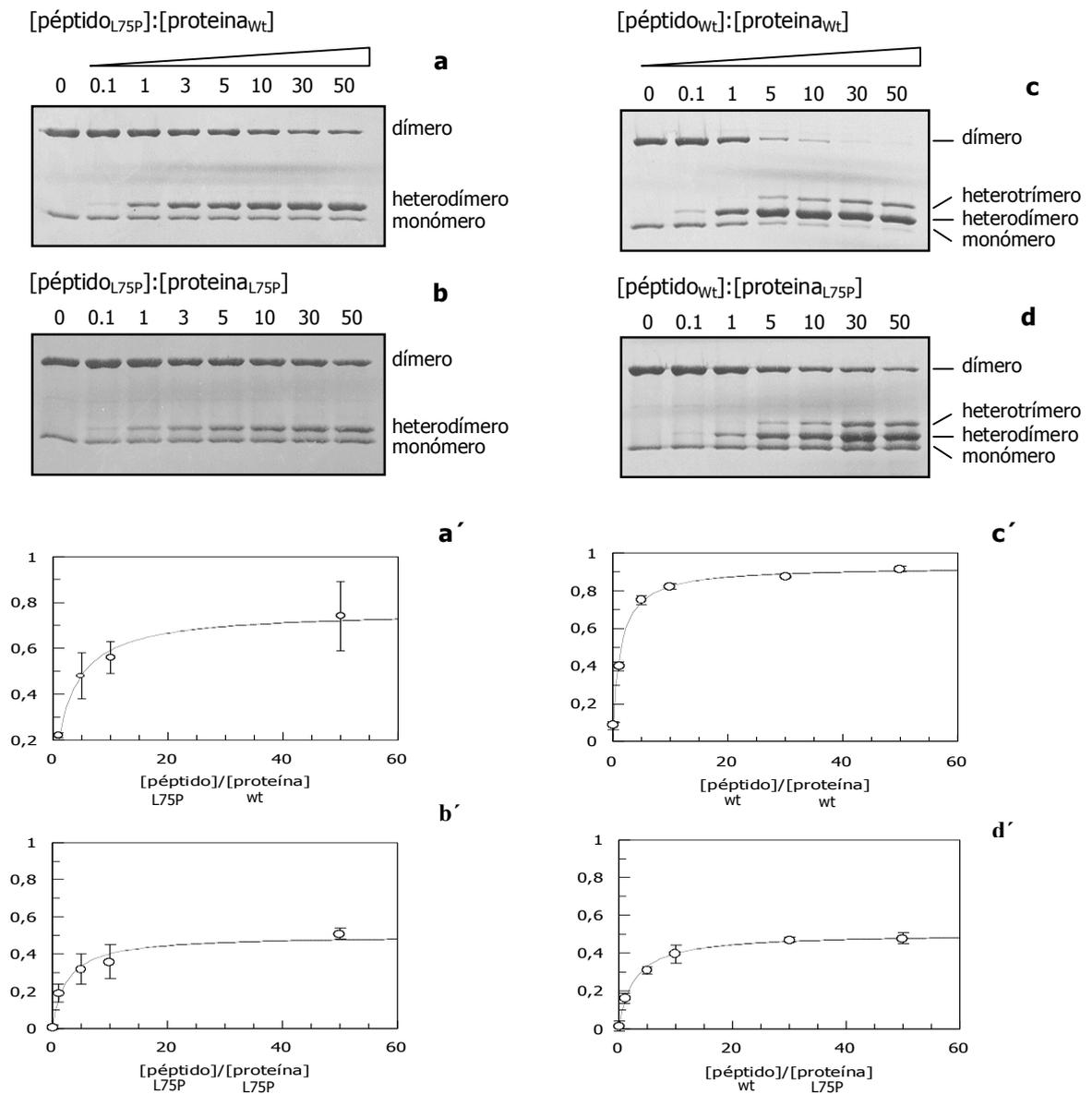


Figura 5.7.: Experimentos de competencia proteína/péptido para las distintas interfases de interacción y los correspondientes ajustes de la aparición de heterodímero a curvas hiperbólicas. Las barras de error resultan del análisis de tres muestras independientes.

Como se puede observar (figura 5.7 c y d), en aquellos casos en los que se utiliza el péptido salvaje a concentraciones muy elevadas aparece una banda que por su movilidad electroforética corresponde a la presencia de heterotrímeros formados por una molécula de proteína y dos de péptido. Este hecho complica los cálculos de IC50, por lo que para la obtención de estos valores se han tenido en cuenta únicamente aquellas concentraciones en las que la cantidad de heterotrímero no supera el 10% del total de proteína. De los resultados obtenidos para las diferentes interfases de interacción, los valores de IC50 son aproximadamente del mismo orden, si bien menores para el homodímero salvaje (wt/wt) (tabla 5.1), de forma análoga a lo que ocurre con los porcentajes de dimerización obtenidos para la proteína salvaje frente al mutante L75P.

Tabla 5.1. Interfases de interacción péptido/proteína. IC50 obtenida para las diferentes combinaciones.

<b>Interfase de dimerización péptido/proteína</b>	<b>Ic50 (<math>\mu\text{M}</math>)</b>
L75P/wt	12 $\pm$ 2
L75P/L75P	10 $\pm$ 4
wt/wt	5 $\pm$ 1
wt/L75P	11 $\pm$ 2

### 5.2.3.2. Estructura secundaria de los péptidos TM.

La presencia de residuos de prolina en proteínas solubles, como se ha mencionado anteriormente, resulta en general incompatible con la formación de estructura secundaria. Por este motivo se considera al residuo prolina como un “perturbador de hélices”. A pesar de ello, el alto porcentaje de dimerización encontrado para el mutante L75P sugiere el mantenimiento de un elevado contenido en hélice alfa. Con el objetivo de averiguar hasta que punto el mutante L75P conserva la estructura secundaria en hélice alfa similar a la de la proteína salvaje, se realizaron espectros de dicroísmo circular para ambos péptidos a 25°C en presencia de tampón Tris 10mM pH 7.5, con o sin SDS, detergente mimético de ambientes de membrana (Gierasch et al., 1982), que corresponde al utilizado en los geles de electroforesis.

Un primer análisis de la elipticidad molar a 222 nm que presentan ambos péptidos (figura 5.8), indica que la estructura secundaria en el péptido que contiene la secuencia mutante L75P, se mantiene. El comportamiento que observamos en agua, es típico de un péptido de membrana en el que la hélice alfa desaparece, ya que en este ambiente se ve favorecida la formación de puentes de hidrógeno con el medio acuoso. Sin embargo, la adición de SDS proporciona un ambiente hidrofóbico adecuado, en el que la formación de puentes de hidrógeno intramoleculares se ve favorecida y por tanto las estructuras en hélice alfa quedan estabilizadas. En este entorno los algoritmos de predicción de estructura secundaria (Yang et al., 1986) muestran un contenido helicoidal en torno a un 80% para ambos péptidos.

En espectroscopía de dicroísmo circular, la relación de elipticidad molar  $\Theta_{220}/\Theta_{208}$ , se utiliza para determinar el estado de agregación de un péptido (Lau et al., 1984). Cuando esta relación se encuentra en torno a 0.8 se propone la existencia del péptido en su forma monomérica, mientras que un valor en torno a 1 se relaciona con la existencia de formas oligoméricas. En nuestro caso, tanto el péptido mutante como el salvaje muestran una relación  $\Theta_{220}/\Theta_{208}$  de 1, lo que sugiere la existencia del péptido en su forma dimérica también en las condiciones de los experimentos de espectroscopía, en concordancia con los resultados obtenidos en electroforesis (figura 5.6).

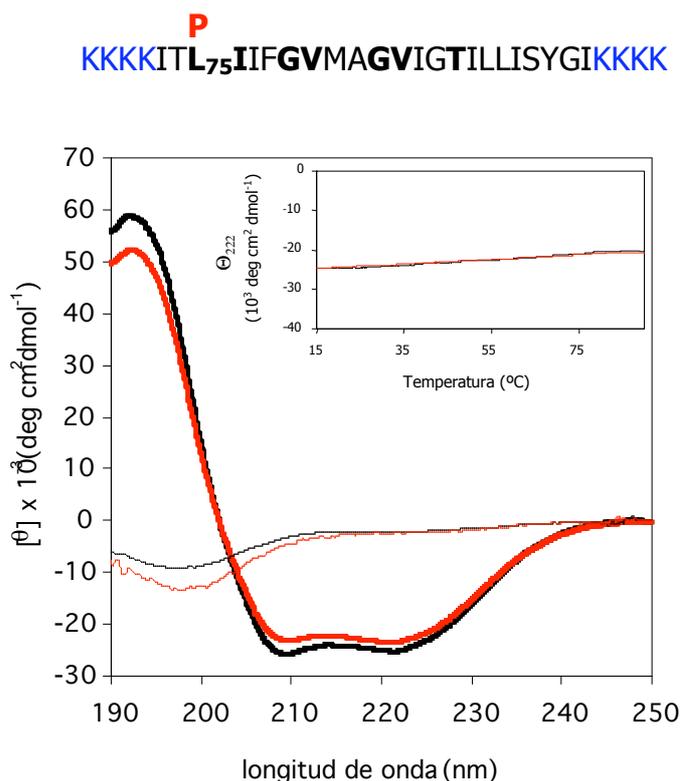


Figura 5.8: Espectros de dicroísmo circular de los péptidos, wt (línea negra) y L75P (línea roja), realizados a 25°C, en medio acuoso (línea fina) o en presencia de SDS 10mM (línea gruesa). En el inserto se muestran los resultados de los experimentos de desnaturalización térmica realizados en presencia de SDS 10 mM, para la proteína wt (negro) y la mutante (línea roja).

Por otro lado, los resultados de los experimentos de estabilidad térmica (sección 2.4.2, materiales y métodos), en los que las diferentes muestras son sometidas a un aumento de temperatura lineal registrándose la elipticidad molar a 222nm, indican que el proceso de desnaturalización es muy similar para las dos muestras (figura 3.8, inserto). La presencia de un residuo de prolina en la hélice del mutante L75P, podría haber provocado que en la curva de desnaturalización apareciera una etapa adicional, correspondiente a la desnaturalización térmica de la región N-terminal. Sin embargo, en ninguno de los dos casos se observa una pérdida total de la estructura secundaria, probablemente porque la exposición al medio hidrofóbico de los grupos polares está muy desfavorecida desde el punto de vista energético. Por otra parte, la relación  $\Theta_{220}/\Theta_{208}$  parece mantenerse alrededor de la unidad a lo largo del proceso de desnaturalización, lo que sugiere que el dímero también es estable a temperaturas elevadas, de nuevo en concordancia con los datos de electroforesis, en los que las muestras son calentadas a 95°C durante 5 minutos.

Además de los estudios de estructura secundaria en presencia de SDS, también se estudió el comportamiento de estos péptidos, en micelas de lípidos, como la lisofosfatidilcolina (LPC) y la dodecilsfosfocolina (DPC) (figura 5.9), medio este último en el que se obtuvieron los espectros de RMN con los que se resolvió la estructura del péptido transmembrana de GpA (MacKenzie et al., 1997).



Figura 5.9: Estructuras químicas de los lípidos utilizados para realizar los experimentos de dicroísmo circular. Donde R' representa una cadena cualquiera de ácido graso.

Los resultados obtenidos en ambos tipos de micelas, muestran un comportamiento similar, en el que las diferencias estructurales entre los dos péptidos no parecen ser significativas (figura 5.10).

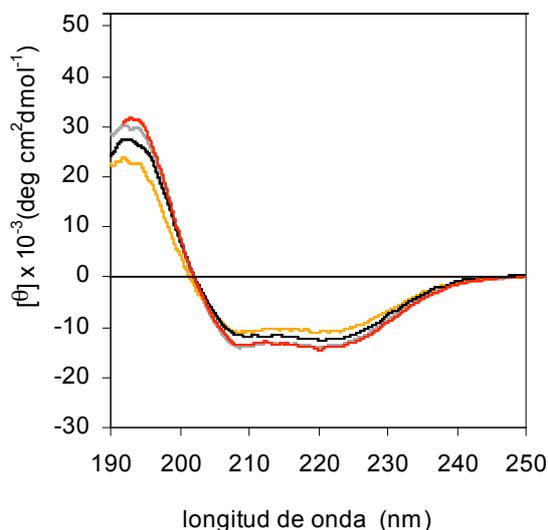


Figura 5.10: Espectros de dicroísmo circular en presencia de micelas lipídicas. Línea negra, péptido wt en micelas de LPC. 4mM Línea roja, péptido L75P en micelas de LPC 10mM. Línea gris, péptido wt en presencia de micelas de DPC 4mM. Línea naranja, péptido L75P en presencia de micelas de DPC 10 mM.

### 5.2.3.3. Estudios de inserción del mutante L75P en membranas naturales.

Los estudios de dicroísmo circular nos dan una idea del comportamiento global del péptido TM de GpA y del mutante L75P en ambientes miméticos de membrana. Sin embargo, para intentar obtener información estructural del extremo la hélice TM en el que se sitúa el residuo 75, en membranas biológicas, se ha utilizado otra aproximación, basada en la técnica de mapeo por glicosilación (Nilsson & von Heijne, 1993). Esta técnica consiste en la realización de experimentos de transcripción/traducción *in vitro* en presencia de membranas microsomales. Estos microsomas poseen no sólo la maquinaria de translocación, sino también, la enzima oligosacaryltransferasa (OST) encargada de los procesos de N-glicosilación en el retículo endoplasmático (ver figura 1.6, en la introducción). La OST reconoce las secuencias Asn-X-Thr/Ser proteicas, en las que la posición central puede estar ocupada por cualquier aminoácido excepto prolina, transfiriendo oligosacáridos de manosa obtenidos de un transportador de dolicol a estas secuencias receptoras, de forma concomitante con el proceso de traducción o inmediatamente después de que ésta se produzca. Como consecuencia de la adición de estos azúcares, la movilidad electroforética de la proteína queda alterada, retrasándose su migración de forma equivalente a un incremento en su masa molecular de aproximadamente 2 kDa en geles SDS-PAGE. El hecho de que el sitio activo de la OST se encuentra a una distancia fija de la

bicapa lipídica, proporciona una técnica eficaz para el estudio de longitudes de fragmentos transmembrana (revisado en van Geest & Lolkema, 2000).

Como proteína modelo para este tipo de estudios se utiliza la proteína Lep (peptidasa del péptido señal). Se trata de una proteína de la membrana interna de *E. coli*, muy bien caracterizada que posee dos segmentos TM (H1 y H2) unidos por un lazo citoplasmático, un pequeño dominio N terminal periplásmico y una larga región C-terminal polar (P2) también localizada en el periplasma (figura 5.11).

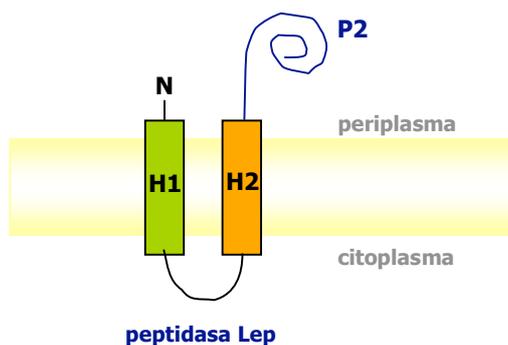


Figura 5.11: Esquema de la disposición en la membrana interna de *E. coli* de la proteína Lep.

Para llevar a cabo estos estudios se han realizado una serie de construcciones, (ver materiales y métodos, sección 2.4), en las que se ha introducido el segmento TM de la proteína GpA salvaje y el del mutante L75P, sustituyendo la hélice H2 de Lep, en los distintos vectores pGEM1-Lep y obteniéndose mutantes en los que la distancia de glicosilación, respecto al segmento transmembrana, varía de forma controlada. Estas construcciones, en las que la diana de glicosilación se encuentra en diferentes posiciones nos permiten averiguar la distancia mínima de glicosilación (MGD) para el segmento TM de GpA, así como para el mutante L75P (figura 5.12).

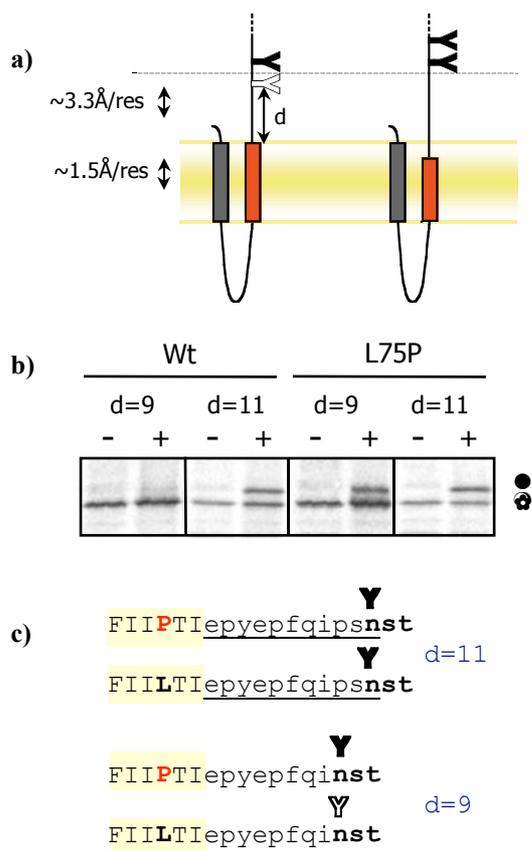


Figura 5.12.: Distancia mínima de glicosilación del fragmento TM de GpA wt frente al mutante L75P. ( a) Modelo de las quimeras de Lepen el que se muestra la relación entre la glicosilación y la longitud de la cadena polipeptídica a partir del segmento TM (b) Eficiencia de glicosilación para los sitios aceptores a distancias  $d=9$  y  $d=11$ . La migración de la banda glicosilada es la correspondiente al círculo negro. El círculo vacío corresponde a la banda de proteína sin glicosilar (c) Esquema de los resultados obtenidos con las secuencias salvaje y mutante, para las diferentes distancias de glicosilación.

Para la interpretación de estos resultados debe tenerse en cuenta que esta distancia corresponde a la que separa el sitio activo de la OST del principio del fragmento TM, de tal modo que esta enzima sea capaz de acceder a la diana de glicosilación de la proteína de membrana. Si el extremo de la hélice TM del mutante L75P, se encuentra plegado del mismo modo que en la proteína salvaje, la MGD no debería variar entre ambas proteínas. Sin embargo, como se observa en la figura 5.12, para la proteína salvaje la MGD es de 11 residuos, mientras que en el caso de la proteína mutante L75P el número de residuos necesarios para posicionar la diana de glicosilación en el sitio activo de la OST se reduce a 9. El hecho de que a distancia de 9 residuos la proteína mutante sea capaz de glicosilarse, sugiere que en la zona donde se encuentra el residuo de prolina esta proteína adopta una conformación más extendida, de modo que, para cubrir la misma distancia hasta alcanzar el centro

activo de la glicosiltransferasa, necesita un número menor de residuos. Esto se debe a que los aminoácidos cubren una distancia mayor en su conformación extendida ( $\sim 3.3 \text{ \AA}$ ) que en conformación helicoidal ( $\sim 1.5 \text{ \AA}$ ).

#### **5.2.4. Estudios de modelado molecular.**

Con el propósito de evaluar los efectos de la presencia de prolina en las distintas posiciones de la hélice transmembrana de GpA, tanto en la propia estructura de la hélice, como en el empaquetamiento hélice-hélice se realizaron estudios de minimización de energía y dinámica molecular de los distintos mutantes en colaboración con el Dr Jesús Salgado (Universidad de Valencia), (ver materiales y métodos sección 2.5).

Las mutaciones que afectan a la estructura secundaria de la primera vuelta de la región N-terminal de una hélice provocan una desestabilización de esta región, que se manifiesta mediante la extensión de esta zona de la proteína (figura 5.13). Sin embargo, el efecto que esto provoca sobre la interfase de dimerización de estos mutantes es diferente. Mientras que para el mutante L75P la interfase de interacción permanece prácticamente intacta, ya que no observamos un cambio drástico en el empaquetamiento entre el residuo en posición 75 y las isoleucinas en posición 76, en el caso de los mutantes I76P e I77P esta interfase queda claramente afectada (figura 5.14). Estas diferencias podrían explicar la pérdida de dímero observada para estas últimas construcciones al tiempo que se mantiene la dimerización en el mutante L75P. Con un efecto similar, la substitución por prolina de la posición 78, a pesar de encontrarse muy cercana al residuo crucial G79P, no provoca grandes cambios en la interfase de interacción entre hélices, de nuevo en concordancia con los resultados obtenidos en los estudios de dimerización (figura 5.13 c y 5.14).

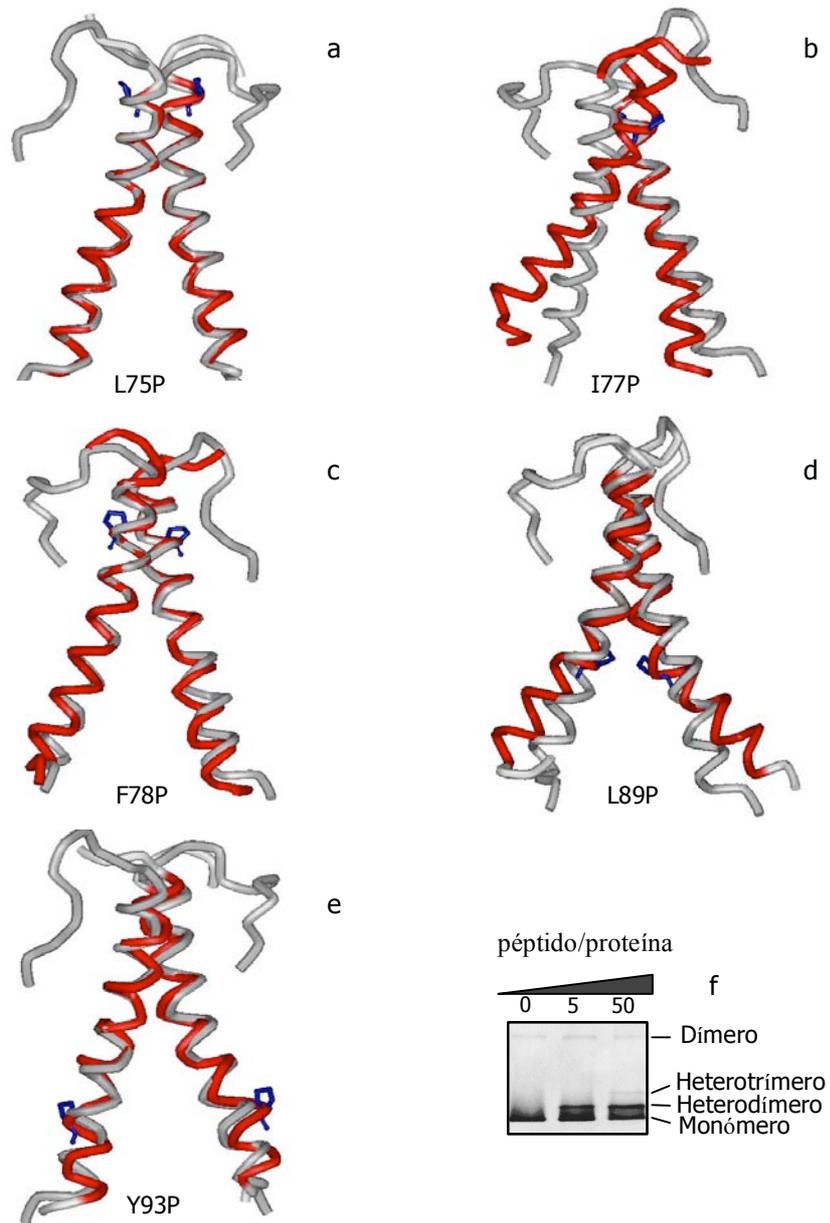


Figura 5.13: Modelos estructurales de los mutantes de prolina del fragmento TM de GpA (color rojo), comparado con la estructura del fragmento TM de la proteína salvaje (color gris). (a),(b) y (c) corresponden a la región N-terminal mientras que (d) y (e) son mutantes del extremo C-terminal. (f) heterooligómeros formados entre el mutante Y93P y un péptido correspondiente a la secuencia TM de la proteína wt.

Cuando la sustitución de residuos se realiza en la región C-terminal, lejos de la interfase de interacción (Ej. L89P, figura 5.13, d), en la mayoría de los casos se observa la aparición de un *acodamiento* en la hélice, que al distorsionar la dirección de ésta afecta a los residuos implicados en el empaquetamiento. Este comportamiento se extiende hasta que alcanzamos posiciones más alejadas del motivo de dimerización en las que la aparición de estas deformaciones en la hélice no afecta a la posición relativa de los residuos en la región de interacción. Aun así, encontramos un mutante, el Y93P, en el que los resultados de la dinámica molecular no muestran grandes cambios respecto a la proteína salvaje y, sin embargo, existe una pérdida de dimerización evidente. Para todos aquellas construcciones que no presentaron capacidad de dimerización, se realizaron experimentos de competencia con el péptido salvaje, con el objetivo de evaluar hasta que punto se encontraba dañada la interfase de interacción en estos mutantes. Sin embargo, únicamente se observó la aparición de heterodímeros proteína mutante/péptido salvaje para la construcción Y93P, lo que indica que la perturbación en la estructura en este mutante no debe ser muy grande, ya que permite la heterooligomerización con la interfase de la proteína salvaje (figura 5.13 (e) y(f)).

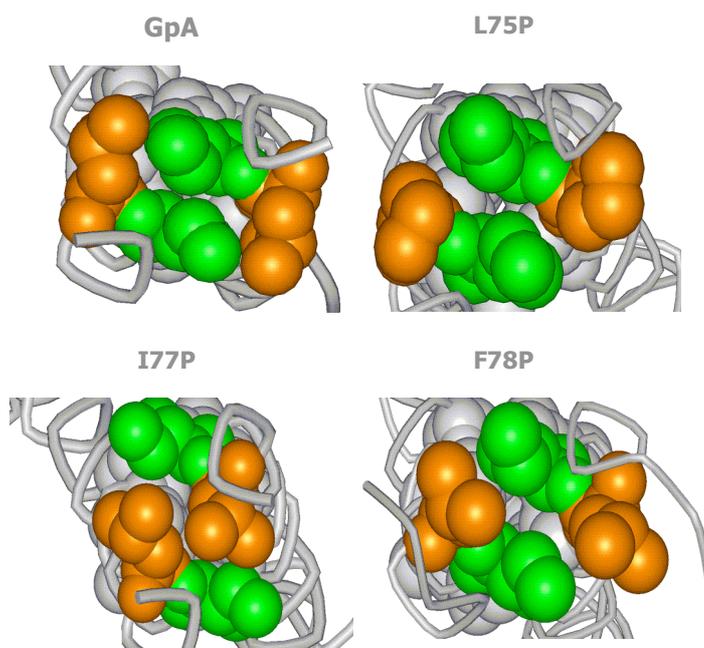


Figura 5.14: Vista superior de la interfase de oligomerización de la proteína salvaje y diversos mutantes. El esqueleto de la proteína se representa como un tubo y los residuos implicados en la dimerización como bolas, con su radio de van der Waals. En color naranja se destaca el residuo en posición 75, mientras que en verde se muestra la isoleucina 76.

### 5.3. DISCUSIÓN

De los resultados obtenidos en este estudio se deduce que, la presencia de prolinas en posiciones centrales del fragmento transmembrana afecta en gran medida a la capacidad de interacción hélice-hélice, siendo mejor tolerada su presencia en las regiones terminales (figura 5.15). Estos resultados, están de acuerdo con la distribución de residuos de prolina encontrada en fragmentos TM de proteínas naturales. La distribución que muestran estos residuos ha sido estudiada sobre fragmentos transmembrana de los que se conoce su estructura con una resolución mayor de 4Å (Ulmschneider & Sansom, 2001). En este estudio, se observa la aparición frecuente de residuos de prolina en los extremos de las hélices transmembrana, acompañada de la aparición de un pequeño pico en regiones centrales de la hélice (figura 5.15). De hecho, en nuestros experimentos, a pesar de estar afectando a un residuo crucial para la oligomerización, el mutante V84P, presenta cierto grado de dimerización, encontrándose, también, en una posición central de la hélice TM.

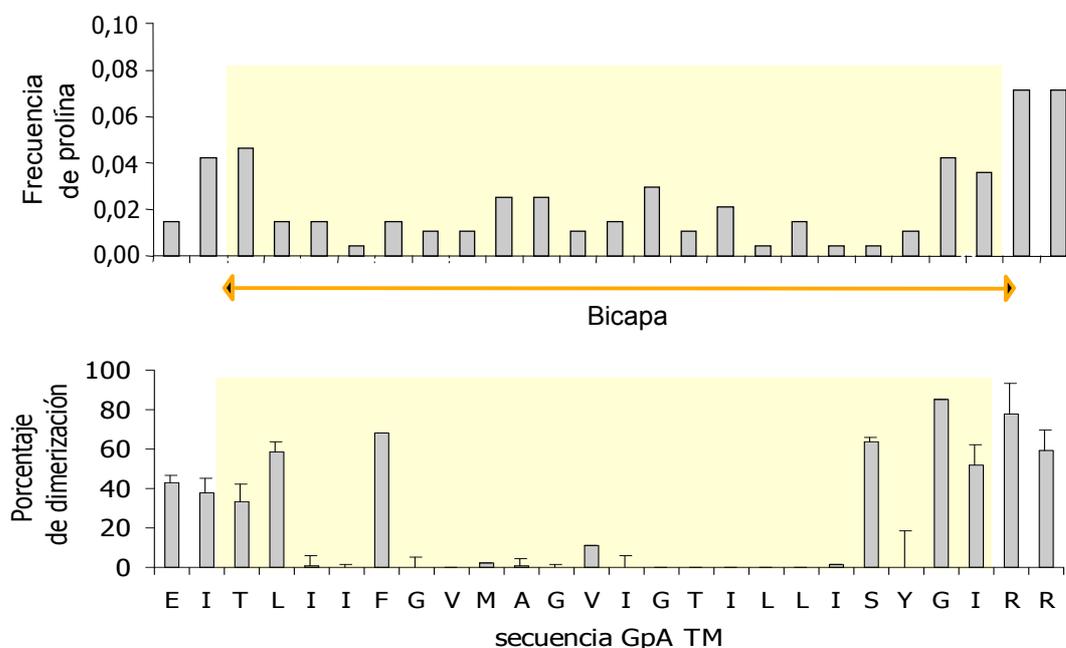


Figura 5.15: Frecuencia de aparición de prolinas de fragmentos TM naturales (Ulmschneider & Sansom, 2001). Comparación con los datos de dimerización obtenidos para los mutantes de prolina del segmento TM de GpA en el presente trabajo .

La aparición de este pico central no se observó, sin embargo, en estudios previos realizados sobre proteínas con un único segmento TM (Landolt-Marticorena et al., 1993). Esta aparente discrepancia entre ambos trabajos teóricos se puede explicar si tenemos en cuenta que las diferencias entre estos dos tipos proteicos se deben principalmente a la inexistencia de interfases de interacción hélice-hélice en aquellas proteínas que poseen un único segmento TM. GpA debido a su capacidad de dimerización, se comporta en muchos aspectos como una proteína politópica, por lo que para este tipo de estudios se podría englobar en este tipo proteico. Por otra parte la aparición de este pico central, es más evidente en estudios recientes realizados por el grupo de investigación del profesor Sansom (Cordes et al., 2002). Desde el punto de vista de empaquetamiento entre hélices, la aparición de este pico central podría verse sesgada debido a la elevada presencia en este análisis de proteínas del tipo receptores acoplados a proteínas G y canales en las que la aparición de ese acodamiento central provocado por la presencia de prolina, parece desempeñar un papel funcional.

En proteínas solubles, cuando una prolina aparece en una hélice, lo hace con mayor frecuencia en la segunda posición de la primera vuelta de hélice, afectando así, en menor medida, su incapacidad de formar puentes de hidrógeno, a la estructura global (Serrano & Fersht, 1989). La preferencia que presenta el residuo prolina por ocupar regiones terminales de los fragmentos TM, puede deberse a la mayor diversidad química que existe en la interfase de la membrana, donde la presencia de los grupos polares de los fosfolípidos constituyentes de las membranas biológicas aumenta las posibilidades de establecer interacciones que compensen la incapacidad de este iminoácido para formar puentes de hidrógeno en posiciones canónicas. Esta mayor diversidad química parece favorecer también la tolerancia de estos residuos en regiones terminales cuando nos referimos al empaquetamiento hélice-hélice. Su presencia en zonas centrales restringe las posibilidades conformacionales, afectando en gran medida a la interacción entre fragmentos. En cambio, cuando un residuo de prolina se sitúa en regiones terminales la flexibilidad conformacional aumenta, tolerando que el empaquetamiento entre hélices se mantenga.

Por otra parte, si nos fijamos en la posición que el motivo de dimerización de GpA presenta respecto al fragmento TM, observamos una distribución asimétrica, de modo que este motivo se sitúa más cercano al extremo N-terminal de la hélice TM. Según esta situación, cabría esperar que la presencia de prolinas fuera mejor tolerada en la región C-terminal, lejana a la interfase de

dimerización. Sin embargo, los resultados experimentales no parecen ajustarse a esta distribución asimétrica. De hecho, en posiciones N-terminales, tan cercanas al motivo de dimerización como es el caso de la F78, que se sitúa justo antes de un residuo crucial para la formación del dímero como es la G79, encontramos porcentajes de dimerización muy elevados. Esta falta de asimetría puede tener dos explicaciones. Por una parte, puede ser debida al hecho de que los residuos de prolina afectan fundamentalmente a la estructuración de las posiciones situadas en N-terminal respecto a su situación, de modo que las sustituciones en la región C-terminal afectan en realidad a regiones anteriores de la secuencia (aproximadamente una vuelta de hélice). Por otra parte, se podría explicar por la propia asimetría del motivo de dimerización, en el que, como previamente ha sido demostrado, parecen jugar un papel más crucial los residuos situados en C-terminal ( $G^{83}$ ,  $V^{84}$  y  $T^{87}$ ), de modo que las sustituciones en posiciones N-terminales serían mejor toleradas.

Los resultados obtenidos para algunos mutantes concretos, como F78, que por su situación en la hélice, presenta un grado de dimerización más elevado de lo que cabría esperar, encuentran cierta explicación en los experimentos de modelado y dinámica molecular que muestran claramente como la interfase de interacción se mantiene prácticamente intacta con respecto a la de la proteína salvaje (figura 5.14). Por otro lado, en el mutante Y93P (figura 5.13 (e)), que se encuentra a más de una vuelta de hélice de distancia del último residuo del motivo de dimerización (Thr87), parece que los cambios observados en la dinámica molecular no justifican la desaparición del dímero. Sin embargo, los experimentos de heterodimerización indican que los cambios estructurales de esta proteína no deben afectar de forma drástica a la interfase de dimerización puesto que este mutante es capaz de heterooligomerizar con péptidos derivados de la secuencia de la proteína salvaje (figura 5.13 (f)). Probablemente la imposibilidad de formar dímero de este mutante Y93P sea debida a un impedimento estructural en la región C-terminal al motivo de dimerización provocado por la presencia de la prolina, que puede ser eludido en el caso de la formación de heterodímero con el péptido salvaje. En resumen, el hecho de que la distribución del residuo prolina en fragmentos TM naturales coincida con las posiciones en que la capacidad de oligomerizar se mantiene parece indicar que entre estos dos hechos existe una relación directa. De este modo la distribución que observamos para este residuo en fragmentos TM naturales podría ser consecuencia de que la presencia de prolina en esas posiciones permite la interacción entre fragmentos transmembrana.

El caso del mutante L75P, en el que el porcentaje de dimerización obtenido es más elevado de lo esperado, dada su descripción como uno de los residuos del motivo de dimerización de GpA, justifican un estudio más detallado de este mutante. De los experimentos de diroísmo circular parece deducirse que no existen diferencias importantes en cuanto a la estructura secundaria global de estos dos péptidos transmembrana (figuras 5.9 y 5.10). El estudio local de la estructura de la región N-terminal utilizando un sistema de mapeo por glicosilación, demuestra que esta zona en la proteína transmembrana se encuentra más extendida, disminuyendo por ello la MGD (figura 5.16).

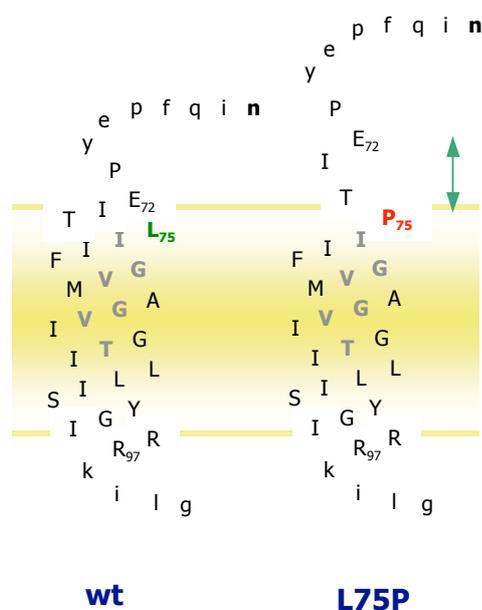


Figura 5.16: Modelo de la situación en membrana de la región N-terminal del fragmento TM de GpA comparada con la del mutante L75P en las contrucciones derivadas de Lep. En mayúsculas se muestran los residuos del fragmento TM de GpA y en minúsculas los aminoácidos que aparecen como resultado de la clonación, incluyendo el residuo aceptor Asn, que se muestra en negrita.

Los estudios de competencia proteína-péptido muestran que en todos los casos en los que en la interfase de interacción hélice-hélice al menos uno de los dos monómeros corresponde a la proteína mutante, la IC50 aumenta (tabla 5.1). Es decir, que la capacidad de dimerizar disminuye de forma equivalente en una interfase mixta wt-L75P y en aquellos casos en los que los dos monómeros presentan la secuencia la proteína mutante, indicando que la posible pérdida de interacciones entre los monómeros debe ser equivalente en ambos casos.

Sin embargo, hay que destacar que en estudios previos se demostró que la sustitución de la leucina 75 por residuos como alanina, valina, isoleucina, triptófano, glicina o treonina provocaban la desaparición del dímero (Lemmon et al., 1992b). Pese a ello, algunas sustituciones, como el cambio por cisteína, metionina o fenilalanina, permitieron la detección de cantidades de dímero significativas. Del mismo modo que, como se ha descrito en el capítulo anterior, en la transferencia del motivo de dimerización de GpA a un esqueleto de polileucinas la posición equivalente a la que ocupa la leucina 75 parece ser relativamente permisiva (figura 4.5). El hecho de que la sustitución por prolina mantenga la capacidad de oligomerización indica que la modificación de la zona de interacción que provoca el cambio por prolina es menor que la provocada por otros residuos y que, probablemente la importancia del residuo leucina 75, no radica en las interacciones que establece sino en que permite, como algunos otros aminoácidos también pueden hacerlo, la formación de la interfase de oligomerización.

En conjunto los resultados obtenidos sobre el empaquetamiento entre hélices TM podrían explicar el hecho de que en fragmentos TM naturales, la situación preferencial de residuos de prolina sea su aparición en regiones terminales. En estas posiciones además de poder jugar un papel funcional no interferirían en el empaquetamiento entre hélices. En definitiva, muestran que el empaquetamiento entre hélices transmembrana puede ser uno de los factores que determinan la frecuencia de aparición de residuos de prolina en proteínas de membrana. Lo que sugiere que además del papel funcional que se le ha atribuido en procesos de transducción de señal y regulación de la apertura de canales, este residuo ejerce también un papel significativo en el plegamiento de este tipo de proteínas.

Conclusiones

---

## **6. Conclusiones**

Conclusiones

---

## CONCLUSIONES.

1. La importancia del extremo C-terminal del segmento transmembrana de GpA en el plegamiento de esta proteína, se debe a su capacidad para mantener los residuos cargados positivamente a una distancia y orientación adecuadas, de modo que el empaquetamiento de la proteína sea posible. El posicionamiento de estos residuos, muy abundantes en los extremos de fragmentos TM, debe ser tenido en cuenta en las predicciones de plegamiento para este tipo de proteínas.
2. Sobre un esqueleto de polileucinas se puede reproducir el estado de oligomerización de GpA, introduciendo parte de los residuos del motivo de dimerización de esta proteína ( $G^{79}V_{xx}GV_{xx}T^{87}$ ). Los residuos  $L^{75}$  y  $I^{76}$  no parecen ser relevantes para la interacción, mientras que los residuos  $^{83}GV_{xx}T^{87}$  parecen indispensables. La  $T^{87}$ , contribuye en gran medida a la formación del dímero, probablemente mediante el establecimiento de un puente de hidrógeno interhelicoidal.
3. La capacidad de inducir dimerización mediada por el motivo ( $G^{79}V_{xx}GV_{xx}T^{87}$ ), se puede ver modificada en función de la longitud del fragmento hidrofóbico sobre el que se introduzca, probablemente porque el ángulo de interacción entre los monómeros, varía según la longitud de la hélice.
4. Respecto del empaquetamiento entre hélices TM, los residuos de prolina son mejor tolerados en los extremos de este tipo de fragmentos, donde parecen afectar en menor medida al empaquetamiento hélice-hélice, como queda patente en los experimentos de modelado y dinámica molecular. Además, la mayor frecuencia de aparición en segmentos transmembrana naturales, coincide con aquellas posiciones en las que no se altera el empaquetamiento interhelicoidal.
5. En el mutante L75P, el residuo prolina provoca que el extremo N-terminal de la proteína adopte una estructura más extendida, sin apenas distorsionar la interfase de dimerización de GpA lo que permite la formación de homodímeros.



## **7. Bibliografía**



## **BIBLIOGRAFÍA:**

- Adamian, L. & Liang, J. (2002). Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. *Proteins* **47**(2), 209-18.
- Andersson, H. & von Heijne, G. (1994). Membrane protein topology: effects of  $\Delta\mu_{H^+}$  on the translocation of charged residues explain the 'positive inside' rule. *Embo J* **13**(10), 2267-72.
- Arkin, I. T. (2002). Structural aspects of oligomerization taking place between the transmembrane alpha-helices of bitopic membrane proteins. *Biochim Biophys Acta* **1565**(2), 347-63.
- Arkin, I. T., Adams, P. D., MacKenzie, K. R., Lemmon, M. A., Brunger, A. T. & Engelman, D. M. (1994). Structural organization of the pentameric transmembrane alpha-helices of phospholamban, a cardiac ion channel. *Embo J* **13**(20), 4757-64.
- Arkin, I. T. & Brunger, A. T. (1998). Statistical analysis of predicted transmembrane alpha-helices. *Biochim Biophys Acta* **1429**(1), 113-28.
- Arselin, G., Giraud, M. F., Dautant, A., Vaillier, J., Brethes, D., Couлары-Salin, B., Schaeffer, J. & Velours, J. (2003). The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *Eur J Biochem* **270**(8), 1875-84.
- Asundi, V. K. & Carey, D. J. (1995). Self-association of N-syndecan (syndecan-3) core protein is mediated by a novel structural motif in the transmembrane domain and ectodomain flanking region. *J Biol Chem* **270**(44), 26404-26410.
- Aurora, R. & Rose, G. D. (1998). Helix capping. *Protein Sci* **7**(1), 21-38.
- Bargmann, C. I., Hung, M. C. & Weinberg, R. A. (1986). The neu oncogene encodes an epidermal growth factor receptor-related protein. *Nature* **319**(6050), 226-30.
- Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Comp Phys Comm* **91**, 43-56.
- Blondelle, S. E., Forood, B., Houghten, R. A. & PerezPaya, E. (1997). Secondary structure induction in aqueous vs membrane-like environments. *Biopolymers* **42**(4), 489-498.
- Bowie, J. U. (1997). Helix packing in membrane proteins. *J Mol Biol* **272**(5), 780-9.

- Brosig, B. & Langosch, D. (1998). The dimerization motif of the glycophorin A transmembrane segment in membranes: Importance of glycine residues. *Protein Sci* **7**, 1052-1056.
- Bywater, R. P., Thomas, D. & Vriend, G. (2001). A sequence and structural study of transmembrane helices. *J Comput Aided Mol Des* **15**(6), 533-52.
- Chen, C. P., Kernytsky, A. & Rost, B. (2002). Transmembrane helix predictions revisited. *Protein Sci* **11**(12), 2774-91.
- Choma, C., Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2000). Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol* **7**(2), 161-6.
- Cordes, F. S., Bright, J. N. & Sansom, M. S. (2002). Proline-induced distortions of transmembrane helices. *J Mol Biol* **323**(5), 951-60.
- Dawson, J. P., Weinger, J. S. & Engelman, D. M. (2002). Motifs of serine and threonine can drive association of transmembrane helices. *J Mol Biol* **316**(3), 799-805.
- Deber, C. M., Brandl, C. J., Deber, R. B., Hsu, L. C. & Young, X. K. (1986). Amino acid composition of the membrane and aqueous domains of integral membrane proteins. *Arch Biochem Biophys* **251**(1), 68-76.
- Deber, C. M., Wang, C., Liu, L. P., Prior, A. S., Agrawal, S., Muskat, B. L. & Cuticchia, A. J. (2001). TM Finder: a prediction program for transmembrane protein segments using a combination of hydrophobicity and nonpolar phase helicity scales. *Protein Sci* **10**(1), 212-9.
- DeGrado, W. F., Gratkowski, H. & Lear, J. D. (2003). How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Sci* **12**(4), 647-65.
- Deisenhofer, J., Epp, O., Sinning, I. & Michel, H. (1995). Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodospseudomonas viridis*. *J Mol Biol* **246**, 429-457.
- Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O. & Fleming, P. J. (2000). Internal packing of helical membrane proteins. *Proc Natl Acad Sci U S A* **97**(11), 5796-801.
- Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol* **179**(1), 125-42.
- Engelman, D. M., Steitz, T. A. & Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Biophys Chem* **15**, 321-53.

- Erickson, B. W. & Merrifield, R. B. (1976). Solid phase peptide synthesis. In *The Proteins* (Hill, R. L. & Neurath, H., eds.), Vol. 2, pp. 255-527. Academic Press, London.
- Fields, G. B. & Noble, R. L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Prot. Res.* **35**, 161-214.
- Fisher, L. E. & Engelman, D. M. (2001). High-yield synthesis and purification of an alpha-helical transmembrane domain. *Anal Biochem* **293**(1), 102-8.
- Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (1999). Detergents modulate dimerization, but not helicity, of the glycoporphin A transmembrane domain. *J Mol Biol* **293**(3), 639-51.
- Fleishman, S. J. & Ben-Tal, N. (2002). A novel scoring function for predicting the conformations of tightly packed pairs of transmembrane alpha-helices. *J Mol Biol* **321**(2), 363-78.
- Fleming, K. G. & Engelman, D. M. (2001). Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc Natl Acad Sci U S A* **98**(25), 14340-4.
- Gadbut, A. P., Riccardi, D., Wu, L., Hebert, S. C. & Galper, J. B. (1996). Specificity of coupling of muscarinic receptor isoforms to a novel chick inward-rectifying acetylcholine-sensitive K<sup>+</sup> channel. *J Biol Chem* **271**(11), 6398-402.
- Gavel, Y. & von Heijne, G. (1992). The distribution of charged amino acids in mitochondrial inner-membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. *Eur J Biochem* **205**(3), 1207-15.
- Gierasch, L. M., Lacy, J. E., Thompson, K. F., Rockwell, A. L. & Watnick, P. I. (1982). Conformations of model peptides in membrane-mimetic environments. *Biophys J* **37**(1), 275-84.
- Grant, S. G. & Bigbee, W. L. (1994). Bone marrow somatic mutation after genotoxic cancer therapy. *Lancet* **343**(8911), 1507-8.
- Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc Natl Acad Sci U S A* **98**(3), 880-5.
- Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**(15), 2714-23.
- Gurezka, R. & Langosch, D. (2001). In vitro selection of membrane-spanning leucine zipper protein-protein interaction motifs using POSSYCCAT. *J Biol Chem* **276**(49), 45580-7.

- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**(4), 557-80.
- Holzwarth, G. & Doty, P. (1965). The ultraviolet circular dichroism of polypeptides. *J Am Chem Soc* **87**, 218-228.
- Humphrey, W., Dalke, A., Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38.
- Hunte C, M. H. (2002). Crystallisation of membrane proteins mediated by antibody fragments. *Current Opinion in Structural biology* **12**, 503-508.
- Jacobs, R. E. & White, S. H. (1989). The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices. *Biochemistry* **28**(8), 3421-37.
- Jayasinghe, S., Hristova, K. & White, S. H. (2001). MPtopo: A database of membrane protein topology. *Protein Sci* **10**(2), 455-8.
- Jiang, S. & Vakser, I. A. (2000). Side chains in transmembrane helices are shorter at helix-helix interfaces. *Proteins* **40**(3), 429-35.
- Kaiser, E., Colescott, R. L., Bossinger, C. D. & Cook, P. I. (1970). Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem* **34**(2), 595-8.
- Killian, J. A. & von Heijne, G. (2000). How proteins adapt to a membrane-water interface. *Trends-in-Biochemical-Sciences* **25**(9), 429-434.
- Kim, B.-J., Oh, S.-G., Han, M.-G. & Im, S.-S. (2000). Preparation of polyaniline nanoparticles in micellar solutions as polymerization medium. *Langmuir* **16**, 5841-5845.
- Kleiger, G., Grothe, R., Mallick, P. & Eisenberg, D. (2002). GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* **41**(19), 5990-7.
- Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**(3), 567-80.
- Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**(1), 105-32.

- Laage, R., Rohde, J., Brosig, B. & Langosch, D. (2000). A conserved membrane-spanning amino acid motif drives homomeric and supports heteromeric assembly of presynaptic SNARE proteins. *J Biol Chem* **275**(23), 17481-7.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(259), 680-5.
- Landolt-Marticorena, C., Williams, K. A., Deber, C. M. & Reithmeier, R. A. (1993). Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. *J Mol Biol* **229**(3), 602-8.
- Lau, S. Y., Taneja, A. K. & Hodges, R. S. (1984). Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J Biol Chem* **259**(21), 13253-61.
- Leeds, J. A., Boyd, D., Huber, D. R., Sonoda, G. K., Luu, H. T., Engelman, D. M. & Beckwith, J. (2001). Genetic selection for and molecular dynamic modeling of a protein transmembrane domain multimerization motif from a random Escherichia coli genomic library. *J Mol Biol* **313**(1), 181-95.
- Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B.-J., Dempsey, C. E. & Engelman, D. M. (1992a). Glycophorin A dimerization is driven by specific interactions between transmembrane  $\alpha$ -helices. *J Biol Chem* **267**, 7683-7689.
- Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J. & Engelman, D. M. (1992b). Sequence specificity in the dimerization of transmembrane  $\alpha$ -helices. *Biochemistry* **31**(51), 12719-12725.
- Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brünger, A. T. & Engelman, D. M. (1994). A dimerization motif for transmembrane  $\alpha$ -helices. *Nature Struct Biol* **1**, 157-163.
- Li, S.-C. & Deber, C. M. (1994). A measure of helical propensity for amino acids in membrane environments. *Nature Struct Biol* **1**, 368-373.
- Lindahl, E., Hess, B., van der Spoel, D. (2001). GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* **7**, 306-317.
- Liu, L. P. & Deber, C. M. (1998). Uncoupling hydrophobicity and helicity in transmembrane segments. Alpha-helical propensities of the amino acids in non-polar environments. *J Biol Chem* **273**(37), 23645-8.

- Liu, Y., Engelman, D. M. & Gerstein, M. (2002). Genomic analysis of membrane protein families: abundance and conserved motifs. *Genome Biol* **3**(10), RESEARCH0054.
- MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science* **276**(5309), 131-3.
- Martelli, P. L., Fariselli, P., Krogh, A. & Casadio, R. (2002). A sequence-profile-based HMM for predicting and discriminating beta barrel membrane proteins. *Bioinformatics* **18**(Suppl 1), S46-53.
- Martin, L. & Vita, C. (2000). Engineering novel bioactive mini-proteins from small size natural and de novo designed scaffolds. *Curr Protein Pept Sci* **1**(4), 403-30.
- Melnyk, R. A., Partridge, A. W. & Deber, C. M. (2001). Retention of native-like oligomerization states in transmembrane segment peptides: application to the Escherichia coli aspartate receptor. *Biochemistry* **40**(37), 11106-13.
- Mingarro, I., Elofsson, A. & von Heijne, G. (1997). Helix-helix packing in a membrane-like environment. *J Mol Biol* **272**(4), 633-641.
- Mingarro, I., Whitley, P., Lemmon, M. A. & von Heijne, G. (1996). Ala-insertion scanning mutagenesis of the glycoporphin A transmembrane helix. A rapid way to map helix-helix interactions in integral membrane proteins. *Protein Sci* **5**, 1339-1341.
- Moller, S., Croning, M. D. & Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**(7), 646-53.
- Munro, S. (1995). An investigation of the role of transmembrane domains in Golgi protein retention. *Embo J* **14**(19), 4695-704.
- Nilsson, I. & von Heijne, G. (1993). Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J Biol Chem* **268**(8), 5798-5801.
- Nilsson, I., Whitley, P. & von Heijne, G. (1994). The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J Cell Biol* **126**(5), 1127-32.
- Ockenhouse, C. F., Barbosa, A., Blackall, D. P., Murphy, C. I., Kashala, O., Dutta, S., Lanar, D. E. & Daugherty, J. R. (2001). Sialic acid-dependent binding of baculovirus-expressed recombinant

- antigens from Plasmodium falciparum EBA-175 to Glycophorin A. *Mol Biochem Parasitol* **113**(1), 9-21.
- Peitsch, M. C., Wells, T. N., Stampf, D. R. & Sussman, J. L. (1995). The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. *Trends Biochem Sci* **20**(2), 82-4.
- Pilpel, Y., Ben-Tal, N. & Lancet, D. (1999). kPROT: a knowledge-based scale for the propensity of residue orientation in transmembrane segments. Application to membrane protein structure prediction. *J Mol Biol* **294**(4), 921-35.
- Popot, J. L. & Engelman, D. M. (1990). Membrane protein folding and oligomerization - The 2-stage model. *Biochemistry* **29**(17), 4031-4037.
- Popot, J. L. & Engelman, D. M. (2000). Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem* **69**, 881-922.
- Popot, J. L., Gerchman, S. E. & Engelman, D. M. (1987). Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J Mol Biol* **198**(4), 655-76.
- Ren, J., Lew, S., Wang, J. & London, E. (1999). Control of the Transmembrane Orientation and Interhelical Interactions within Membranes by Hydrophobic Helix Length. *Biochemistry* **38**, 5905-5912.
- Ri, Y., Ballesteros, J. A., Abrams, C. K., Oh, S., Verselis, V. K., Weinstein, H. & Bargiello, T. A. (1999). The role of a conserved proline residue in mediating conformational changes associated with voltage gating of Cx32 gap junctions. *Biophys J* **76**(6), 2887-98.
- Rohl, C. A. & Baldwin, R. L. (1998). Deciphering rules of helix stability in peptides. *Methods Enzymol* **295**, 1-26.
- Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci USA* **96**, 863-868.
- Salom, D., Hill, B. R., Lear, J. D. & DeGrado, W. F. (2000). pH-dependent tetramerization and amantadine binding of the transmembrane helix of M2 from the influenza A virus. *Biochemistry* **39**(46), 14160-70.
- Sansom, M. S. & Weinstein, H. (2000). Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices. *Trends Pharmacol Sci* **21**(11), 445-51.
- Scheiner, S., Kar, T. & Gu, Y. (2001). Strength of the C $\alpha$  H..O hydrogen bond of amino acid residues. *J Biol Chem* **276**(13), 9832-7.

- Schneider, D. & Engelman, D. M. (2003). GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. *J Biol Chem* **278**(5), 3105-11.
- Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* **31**(13), 3381-5.
- Senes, A., Gerstein, M. & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* **296**(3), 921-36.
- Senes, A., Ubarretxena-Belandia, I. & Engelman, D. M. (2001). The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* **98**(16), 9056-61.
- Serrano, L. & Fersht, A. R. (1989). Capping and alpha-helix stability. *Nature* **342**(6247), 296-9.
- Simmerman, H. K., Kobayashi, Y. M., Autry, J. M. & Jones, L. R. (1996). A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a coiled-coil pore structure. *J Biol Chem* **271**(10), 5941-6.
- Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W. W., Groesbeek, M., Metz, G., Ziliox, M. & Aimoto, S. (2002). Implications of threonine hydrogen bonding in the glycoporphin A transmembrane helix dimer. *Biophysical Journal* **82**(5), 2476-2486.
- Smith, S. O., Song, D., Shekar, S., Groesbeek, M., Ziliox, M. & Aimoto, S. (2001). Structure of the transmembrane dimer interface of glycoporphin A in membrane bilayers. *Biochemistry* **40**(22), 6553-8.
- Sonnhammer, E. L., von Heijne, G. & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**, 175-82.
- Tian, C., Tobler, K., Lamb, R. A., Pinto, L. H. & Cross, T. A. (2002). Expression and initial structural insights from solid-state NMR of the M2 proton channel from influenza A virus. *Biochemistry* **41**(37), 11294-300.
- Tomita, M., Furthmayr, H. & Marchesi, V. T. (1978). Primary structure of human erythrocyte glycoporphin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry* **17**(22), 4756-70.

- Torres, J., Adams, P. D. & Arkin, I. T. (2000). Use of a new label,  $(^{13}\text{C})\text{C}=\text{O}$ , in the determination of a structural model of phospholamban in a lipid bilayer. Spatial restraints resolve the ambiguity arising from interpretations of mutagenesis data. *J Mol Biol* **300**(4), 677-85.
- Torres, J., Briggs, J. A. & Arkin, I. T. (2002). Contribution of energy values to the analysis of global searching molecular dynamics simulations of transmembrane helical bundles. *Biophys J* **82**(6), 3063-71.
- Torres, J., Kukol, A., Goodman, J. M. & Arkin, I. T. (2001). Site-specific examination of secondary structure and orientation determination in membrane proteins: the peptidic  $(^{13}\text{C})\text{C}=\text{O}$  group as a novel infrared probe. *Biopolymers* **59**(6), 396-401.
- Torres, J., Stevens, T. J. & Samsó, M. (2003). Membrane proteins: the 'Wild West' of structural biology. *Trends Biochem Sci* **28**(3), 137-44.
- Tusnady, G. E. & Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283**(2), 489-506.
- Ulmschneider, M. B. & Sansom, M. S. (2001). Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta* **1512**(1), 1-14.
- van Geest, M. & Lolkema, J. S. (2000). Membrane topology and insertion of membrane proteins: Search for topogenic signals. *Microbiology and Molecular Biology Reviews* **64**(1), 13-33,CP1.
- van Klompenburg, W., Nilsson, I., von Heijne, G. & de Kruijff, B. (1997a). Anionic phospholipids are determinants of membrane protein topology. *Embo J* **16**(14), 4261-6.
- van Klompenburg, W., Ridder, A. N., van Raalte, A. L., Killian, A. J., von Heijne, G. & de Kruijff, B. (1997b). In vitro membrane integration of leader peptidase depends on the Sec machinery and anionic phospholipids and can occur post-translationally. *FEBS Lett* **413**(1), 109-14.
- von Heijne, G. (1991). Proline kinks in transmembrane  $\alpha$ -helices. *J. Mol. Biol.* **218**(3), 499-503.
- von Heijne, G. (1992). Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225**(2), 487-94.
- von Heijne, G. & Gavel, Y. (1988). Topogenic signals in integral membrane proteins. *Eur J Biochem* **174**(4), 671-8.
- Wallin, E. & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**(4), 1029-38.

- Wang, J., Kim, S., Kovacs, F. & Cross, T. A. (2001). Structure of the transmembrane region of the M2 protein H(+) channel. *Protein Sci* **10**(11), 2241-50.
- White, S. H. & Wimley, W. C. (1999). Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **28**, 319-65.
- Whitley, P., Grahn, E., Kutay, U., Rapoport, T. & von Heijne, G. (1996). A 12 residues long polyleucine tail is sufficient to anchor syntrophin to the ER membrane. *J Biol Chem* **271**, 7583-7586.
- Wiener, M. C., King, G. I. & White, S. H. (1991). Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. I. Scaling of neutron data and the distributions of double bonds and water. *Biophys J* **60**(3), 568-76.
- Wiener, M. C. & White, S. H. (1992). Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. II. Distribution and packing of terminal methyl groups. *Biophys J* **61**(2), 428-33.
- Wigley, W. C., Corboy, M. J., Cutler, T. D., Thibodeau, P. H., Oldan, J., Lee, M. G., Rizo, J., Hunt, J. F. & Thomas, P. J. (2002). A protein sequence that can encode native structure by disfavoring alternate conformations. *Nat Struct Biol* **9**(5), 381-8.
- Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A. & Deber, C. M. (1995). Packing of coat protein amphipathic and transmembrane helices in filamentous bacteriophage M13: Role of small residues in protein oligomerization. *J Mol Biol* **252**, 6-14.
- Wimley, W. C. & White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* **3**(10), 842-8.
- Yang, J. T., Wu, C. S. & Martinez, H. M. (1986). Calculation of protein conformation from circular dichroism. *Methods Enzymol* **130**, 208-69.
- Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. (2000). Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat Struct Biol* **7**(2), 154-60.
- Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc Natl Acad Sci U S A* **98**(5), 2250-2255.