



Universidad de Concepción  
Dirección de Postgrado  
Facultad de Ciencias Biológicas - Programa de Magíster en Bioquímica y Bioinformática

## **Diseño de un Ligando Peptídico con Afinidad por la Hormona Folículo Estimulante Humana**

Tesis para optar al grado de Magíster en Bioquímica y Bioinformática

Jaime Camilo Teneb Lobos  
CONCEPCIÓN-CHILE  
2016

Profesor Guía: Dr. Oliberto Sánchez Ramos  
Profesor Co-guía: Dr. Alexis Salas Burgos  
Dpto. de Farmacología, Facultad de Ciencias Biológicas  
Universidad de Concepción

Esta tesis ha sido realizada en el Departamento de Farmacología de la Facultad de Ciencias Biológicas, Universidad de Concepción.

Profesor Guía

---

Dr. Oliberto Sánchez Ramos  
Facultad de Ciencias Biológicas  
Universidad de Concepción

Profesor Co-Guía

---

Dr. Alexis Salas Burgos  
Facultad de Ciencias Biológicas  
Universidad de Concepción

Comisión Evaluadora:

---

Dr. Jorge Toledo Alonso  
Facultad de Ciencias Biológicas  
Universidad de Concepción

---

Dr. Sergio Onate Betancour  
Facultad de Medicina  
Universidad de Concepción

Director de Programa

---

Dra. Violeta Morín Muñoz  
Facultad de Ciencias Biológicas  
Universidad de Concepción



## INDICE GENERAL

RESUMEN.....	6
ABSTRACT.....	7
I. INTRODUCCIÓN.....	8
1. Relevancia económica de la FSH.....	8
2. Preparaciones de FSH.....	8
3. Análisis de Interfaz proteína-proteína.....	10
4. Estructura del complejo FSH-FSHR.....	11
II. HIPÓTESIS.....	13
III. OBJETIVOS.....	13
OBJETIVO GENERAL.....	13
Objetivos específicos.....	13
IV. METODOLOGÍA.....	14
1. Análisis de interfaz de interacción desde complejos cristalográficos.....	14
2. Modelamiento Molecular.....	14
3. Ensamble termodinámico del sistema para NAMD.....	14
4. Simulaciones de Dinámica Molecular en NAMD.....	15
5. Análisis de las simulaciones de dinámica molecular de NAMD.....	15
6. Ensamblaje termodinámico para de sistema FSH-péptido para Amber.....	15
7. Simulaciones de Dinámica Molecular en AMBER.....	16
8. Cálculo de energía libre de unión del complejo FSH-péptido por MM-PBSA.....	16
V. RESULTADOS.....	17
1. Análisis de la interfaz proteína-proteína del complejo FSH-FSHR <sub>ED</sub> de las estructuras cristalográficas depositadas en la base de datos de proteínas PDB.....	17
2. Análisis del comportamiento del complejo FSH-FSHR y su interfaz proteína-proteína, en un sistema de dinámica molecular con el ectodominio completo del receptor.....	28
2.1 Modelamiento del exodominio completo de FSHR.....	28
2.2 Generación de sistema para simulación de dinámica molecular.....	28
2.3 Simulación de dinámica molecular del complejo FSH-FSHR <sub>ED</sub> .....	30
3. Diseño de un ligando peptídico con afinidad por FSH basado en los análisis de interfaz proteína-proteína del complejo FSH-FSHR <sub>ED</sub> y evaluación de su afinidad mediante un cálculo predictivo de energía libre de unión.....	31
3.1 Diseño de ligando peptídico con afinidad por FSH para uso en cromatografía de afinidad.....	36

3.2 Evaluación del ligando peptídico con afinidad por FSH mediante estudios de simulación de dinámica molecular.....	36
3.3 Predicción de energía libre de unión del complejo FSH-peptido.....	40
<b>VI. DISCUSIÓN.....</b>	<b>42</b>
<b>1. Estudio de la estructura cristalográfica.....</b>	<b>45</b>
<b>2. Modelamiento y simulación de dinámica molecular del complejo FSH-FSHR<sub>ED</sub>.....</b>	<b>45</b>
<b>3. Diseño y evaluación del ligando peptídico con afinidad por FSH.....</b>	<b>46</b>
<b>VII. CONCLUSIONES.....</b>	<b>47</b>
<b>VIII. AGRADECIMIENTOS.....</b>	<b>48</b>
<b>IX. REFERENCIAS.....</b>	<b>49</b>

## INDICE DE FIGURAS

Figura 1. Estructura del complejo FSH-FSHR.....	12
Figura 2. Determinación de Hot-spots de interfaz FSH-FSHR.....	27
Figura 3. Región de FSHR <sub>ED</sub> de estructura indeterminada.....	29
Figura 4. Análisis de la trayectoria de la simulación de dinámica molecular por RMSD.....	32
Figura 5. Validación de SDM por comparación de RMSF y Factor Beta..	33
Figura 6. Energía total durante la SDM del complejo FSH-FSHR <sub>ED</sub> .....	34
Figura 7. Análisis de <i>hotspots</i> de estructura a 46,5 ns de simulación..	37
Figura 8. Trayectorias de las SDM de los sistemas FSH, péptido y complejo FSH-péptido, analizadas por RMSD.....	38
Figura 9. RMSF promedio de residuos del péptido diseñado en condiciones de ausencia y presencia de la proteína FSH.....	39
Figura 10. RMSF promedio de los residuos de interfaz de la FSH con el ligando peptídico en condiciones de ausencia y presencia del péptido.....	41
Figura 11. Equilibración del sistema FSH-péptido.....	43
Figura 12. Parámetros de la simulación productiva de 20 ns de FSH-péptido.....	44

## INDICE DE TABLAS

Tabla 1. Resumen de interfaz entre la cadena $\alpha$ de la FSH y el exodominio del receptor..	19
Tabla 2. Enlaces de Hidrógeno entre la cadena $\alpha$ de la FSH y el exodominio de su receptor en la estructura 4AY9.	20
Tabla 3. Puentes salinos entre la cadena $\alpha$ de la FSH y el exodominio de su receptor en la estructura 4AY9.	20
Tabla 4. Residuos de la interfaz del exodominio de FSHR con la cadena $\alpha$ de la FSH en la estructura 4AY9.	21
Tabla 5. Residuos de la interfaz de la cadena $\alpha$ de la FSH con el exodominio de FSHR en la estructura 4AY9.	22
Tabla 6. Resumen de interfaz entre la cadena $\beta$ de la FSH y el exodominio del receptor.....	23
Tabla 7. Enlaces de Hidrógeno entre la cadena $\beta$ de la FSH y el exodominio de su receptor en la estructura 4AY9.	24
Tabla 8. Puentes salinos entre la cadena $\beta$ de la FSH y el exodominio de su receptor en la estructura 4AY9.	24
Tabla 9. Residuos de la interfaz del exodominio de FSHR con la cadena $\beta$ de la FSH en la estructura 4AY9.	25
Tabla 10. Residuos de la interfaz de la cadena $\beta$ de la FSH con el exodominio de FSHR en la estructura 4AY9.	26
Tabla 11. Análisis de interfaz del complejo FSH-FSHR <sub>ED</sub> en el lazo modelado en los mínimos locales energéticos extraídos de la SDM.....	35

## RESUMEN

La Hormona Folículo Estimulante (FSH) es una glicoproteína que está involucrada en el desarrollo sexual y la función reproductiva en vertebrados. Debido a esto, es utilizada como agente terapéutico para casos de infertilidad de origen femenino desde principio de la década del 60 y posteriormente, también se incorporó en casos de infertilidad de origen masculino. Comercialmente, los fármacos basados en FSH registran ventas de \$1.30 billones de dólares durante el 2013 con un 2% de aumento con respecto al año anterior, presentándose en décimo tercer lugar de las familias de biofármacos más vendidos en el mundo. La principal preparación de FSH humana en la actualidad es Gonal-f con dos tercios del mercado; en conjunto con otros preparados poseen un proceso de purificación basado en cromatografía de afinidad que utiliza anticuerpos monoclonales. La cromatografía de afinidad es un método de separación basado en la propiedad de las moléculas biológicas, principalmente anticuerpos, de unir ligandos específicamente por complementariedad de superficies y fuerzas intermoleculares. A pesar de que el uso de anticuerpos ha demostrado ser muy eficaz en cromatografía de afinidad, actualmente existe una tendencia a reemplazar el uso de anticuerpos por **ligandos peptídicos pequeños** que poseen especificidad y afinidad similar a los anticuerpos, pero presentan una **mayor estabilidad y menor costo de producción**.

En este trabajo se busca diseñar un ligando peptídico que presente alta afinidad por la FSH humana y pueda ser utilizado en un proceso de purificación por cromatografía de afinidad. Para realizar el diseño del péptido, nos hemos basado en el análisis de interfaz proteína-proteína del complejo FSH-FSHR. Se encontró un grupo de residuos con cercanía lineal, conformado por el residuo modificado de tirosina sulfatada sTyr335 y Leu337, que poseen una gran parte de su superficie internalizada en el complejo, y con un cambio de energía de unión ( $\Delta\Delta G$ ) significativo, identificado por barrido de alanina *in silico*, suficientes para ser considerados *hotspots*. Cercano a estos residuos, existe una zona de estructura indeterminada en el receptor.

Con el objetivo de analizar el aporte de los residuos de este lazo en la interacción del complejo, se generó un modelo del exodominio del receptor completo, y así, realizar una simulación de dinámica molecular del complejo por un tiempo de 50 ns, para estudiar su comportamiento en un sistema termodinámico completo. Del análisis de dinámica en estructuras de menor energía, se encontró que el residuo M328, del lazo de estructura desconocida, presenta una superficie internalizada en el complejo considerable, y un  $\Delta\Delta G$  que nos lleva a considerarlo *hotspot*. De esta forma encontramos una secuencia de 10 residuos que posee 3 *hotspots*. Posteriormente, basados en esta secuencia, diseñamos un péptido para utilización en una matriz de purificación de sefarosa N-Hidroxisuccinamida para cromatografía de afinidad y evaluamos la energía de unión del complejo FSH-péptido por MM/PBSA. El valor predictivo de  $\Delta G_{\text{unión}}$  es  $-15,4 \pm 6,9$  kcal/mol lo que sugiere que el péptido diseñado en este trabajo tiene potencial para ser utilizado en la purificación de FSH humana.



## ABSTRACT

The Follicle-Stimulating Hormone (FSH) is a glycoprotein that plays its role in sexual development and reproductive function in vertebrates. Therefore, FSH is used as therapeutic agent for female infertility cases since the beginning of the 60's and later, for male infertility cases. Commercially, FSH based drugs has a sales record of \$1.30 billions of dollars in 2013 with 2% of growing compared to 2012, being the 13<sup>th</sup> most important in the biopharmaceuticals blockbusters worldwide. Currently, the main preparation of human FSH, with two thirds of the market, Gonal-f, and others use an affinity chromatography step in their purification process that involves monoclonal antibodies to capture the protein. Affinity chromatography is a separation method that takes advantage of the specific union of affinity ligands that has biological molecules, especially antibodies. However the use of antibodies has a proved efficacy in affinity chromatography, there is a trend to replace the use of antibodies by **small peptide ligands** that have similar specificity and affinity, but **higher stability and lower production cost**.

This work seeks to design a peptide ligand that has high affinity for the human FSH and could be used in a purification process by affinity chromatography.

We based our design in the information that provides the crystal structure of the complex conformed by the FSH and the exodomain of its receptor. We performed protein-protein interface analysis of the complex in order to find the residues that are crucial for the interaction. We found a group of residues with linear proximity, formed by the modified residue sulphated tyrosine sTyr355 y Leu337, that have buried surface and change of binding energy ( $\Delta\Delta G_{\text{bind}}$ ) in alanine scan for consider it as hotspots. Near to these residues, there is a loop of indeterminate structure in the receptor.

With the aim of analyse the contribution of this loop in the interaction of the complex, we generate a model of the full receptor exodomain to perform a molecular dynamics simulation of the complex of 50 ns, to study its behaviour in dynamics. Through this analysis, we found that the residue M328, from the loop, shows an important buried surface and a  $\Delta\Delta G$  from alanine scanning that lead us to consider it a hotspot. With this methodology we found a 10 residues sequence that contains 3 hotspots. Thus, based in this sequence, we design a peptide for affinity chromatography applications with a N-Hydroxysuccinamide sepharose matrix. We evaluate the binding energy of the FSH-peptide complex by MM/PBSA obtaining a predictive value of  $\Delta G$  of  $-15,4 \pm 6,9$  kcal/mol. This value of  $\Delta G_{\text{bind}}$  suggests that the peptide designed in this work has potential of being used in human FSH purification.

## **I. INTRODUCCIÓN**

La Hormona Folículo Estimulante (FSH) es una glicoproteína que está involucrada en el desarrollo sexual y la función reproductiva en vertebrados. Junto con la Hormona Luteinizante (LH), forman las gonadotrofinas hipofisarias, que son secretadas por la glándula pituitaria anterior en respuesta al factor liberador de gonadotrofina (GnRH) producido por el hipotálamo. Estas hormonas tienen su blanco de acción en las gónadas. Debido a las diferencias sexuales, tanto el blanco como el efecto de la hormona son diferenciales para los géneros masculino y femenino. En el macho, FSH induce y mantiene la espermatogénesis (Pérez y Donoso, 2011) mediante su efecto sobre las células de Sertoli, claves para la espermatogénesis y únicas células masculinas que presentan el receptor de FSH (FSHR) (Simoni y cols., 1997; Heckert y Griswold, 2002). En mujeres, FSH cumple un rol fundamental en la función reproductiva, siendo necesaria y suficiente para inducir la maduración de los folículos ováricos a un estado maduro preovulatorio (Hunzicker-Dunn y Maizels, 2006). Las células granulosas presentes en el folículo ovárico son el blanco de la FSH. El efecto de la interacción FSH-FSHR provoca proliferación, crecimiento y diferenciación en las células granulosas, lo que se traduce en la maduración del folículo. La FSH también provoca un aumento en la expresión del receptor de LH en las células granulosas (Zelevnik y cols., 1974), hormona que finalmente provoca la ovulación en el folículo maduro. Este fenómeno en las mujeres requiere expresión diferencial en el tiempo de las gonadotrofinas hipofisarias, lo que constituye el ciclo menstrual.

Debido a su gran importancia en la fisiología reproductiva, la FSH ha sido ampliamente usada como fármaco en terapia contra la infertilidad, tanto en hombres como en la mujer. Esto a llevado a que los fármacos que contienen FSH se encuentren entre los fármacos proteicos más vendidos a nivel mundial (La Merie Publishing 2014).

### **1. Relevancia económica de la FSH**

En la revisión comercial de *La Merie Publishing* de 2013 (*La Merie Publishing* 2014) muestran a los fármacos basados en FSH como el décimo tercer bloque de biofármacos con mayor ingreso por ventas en el mundo. Estas ventas suman un total de \$1.3 billones de dólares en 2013, con un aumento de 2% con respecto al año anterior. Por otra parte, muestran a Gonal-f, la FSH humana comercializada por Merck Serono con 2 tercios del mercado, mientras que el grupo restante lo lidera Puregon, FSH desarrollada por laboratorios Organon y actualmente adquirida por Merck & Co.

### **2. Preparaciones de FSH**

A principio de la década de 1960 comenzó el uso de gonadotrofinas con objetivo terapéutico para tratamiento de infertilidad en mujeres. Esta preparación de gonadotrofinas contenía actividad LH y FSH, y era obtenida a partir de la orina de mujeres menopáusicas

(Lunenfeld, 2004). Las necesidades particulares de cada paciente llevaron a la búsqueda de preparaciones puras de hormona FSH y LH. Actualmente, existen dos tipos de FSH comercial disponibles en el mercado: las de origen urinario y las generadas por tecnología recombinante en cultivo celular. La primera generación de preparaciones urinarias con actividad FSH y libre de actividad LH se obtuvieron por cromatografía de inmunoafinidad con anticuerpos policlonales anti-LH (urofolitropina), sin embargo, esta preparación continuaba presentando proteínas urinarias que producían reacciones adversas (Giudice y cols., 1994). Luego, gracias al uso de anticuerpos monoclonales, se generó una preparación altamente purificada de FSH (FSH-HP). En este caso, el anticuerpo utilizado en la columna de afinidad une directamente las moléculas de FSH. Esto permitió que el producto purificado de esta forma, aumentara su actividad de 100~150 IU/mg (Unidades Internacionales por miligramo) en la urofolitropina a ~9.000 IU/mg en la FSH-HP. Sin embargo, las preparaciones de hormonas de origen urinario tienen un límite productivo, que no alcanza a cubrir la demanda mundial, dado por las grandes exigencias que requieren los centros de recolección de orina, de los cuales existen muy pocos con las características requeridas para este tipo de preparación.

Por este motivo, y buscando una mejora en la trazabilidad durante la generación del producto, se buscaron soluciones mediante el uso de la expresión de proteínas por vía recombinante. Como productor se escogió la línea celular CHO, debido a los altos niveles de transfección con DNA foráneo descritos y su capacidad de sintetizar glicoproteínas. Además, pueden crecer en cultivos celulares a gran escala. La primera formulación de una FSH recombinante humana (rhFSH, folitropina alfa) para uso clínico fue generada por laboratorios Serono en 1988 y licenciada en el mercado de la Unión Europea en 1995 como Gonal-f<sup>®</sup>. Una FSH recombinante humana similar (folitropina beta, Puregon<sup>®</sup>) fue licenciada por laboratorios Organon en 1996. Para la generación de la cepa productora de folitropina alfa, se realizó una cotransfección de las células CHO con 2 vectores codificantes para cada cadena de la FSH (Howles, 1996), mientras que para la folitropina beta se generó un vector que codificó para ambos genes de la FSH humana. Luego de las transfecciones, se aislaron transfectantes estables con alto nivel de expresión de FSH. Las preparaciones resultantes tienen alta pureza y alta actividad biológica (>10.000 IU/mg). Los pasos de purificación difieren para las dos FSH recombinantes disponibles en el mercado. El proceso de Puregon usa una serie de pasos de cromatografía de intercambio aniónico y catiónico, cromatografía hidrofóbica y cromatografía de exclusión molecular. El proceso de Gonal-f usa una serie similar de 5 pasos cromatográficos y también incluye una cromatografía de inmunoafinidad con un anticuerpo monoclonal.

Actualmente, existe una tendencia a reemplazar el uso de anticuerpos en pasos de cromatografía de afinidad por el uso de ligandos péptidicos pequeños (Kaufman y cols., 2002; Kelley y cols., 2004; Li y cols., 2014; Liu y cols., 2012; Gétaz y cols., 2013; Gétaz y cols., 2013; Heurich y cols., 2013; Makriyannis y Clonis, 1997), debido a que tienen un menor costo de

síntesis y además poseen una mayor estabilidad, lo que lleva a que las columnas basadas en péptidos tengan mayor número de ciclos de uso. En esta tesis buscamos diseñar un péptido que pueda ser utilizado como ligando para cromatografía de afinidad.

### 3. Análisis de Interfaz proteína-proteína.

Las interacciones proteína-proteína tienen un rol crucial en la regulación de los procesos biológicos. El sitio de unión entre proteínas es llamado interfaz. Las propiedades de los residuos de interfaz determinan el reconocimiento proteína-proteína, su unión y afinidad del complejo formado. Las interfaces proteína-proteína (IPP) son generalmente extensas (600 a 1.300 Å<sup>2</sup>) y participan entre 10 y 30 cadenas laterales de residuos de cada proteína (Janin y Chothia, 1990; de Vos, Ultsch y cols., 1992; Davies y cols., 1990). De forma interesante, solo una pequeña proporción de estos residuos aporta la mayoría de la energía libre de unión ( $\Delta G_{\text{unión}}$ ) del complejo proteína-proteína. El primer reporte de este fenómeno es el caso de la hormona de crecimiento humano (hGH) y su receptor hGHbp, donde de  $\approx 30$  residuos de interfaz aportados por cada proteína, tan solo dos triptófanos de hGHbp aportan tres cuartos de la energía de unión total del complejo (Clackson y Wells, 1995). Estos residuos con importante aporte energético a la unión son denominados *hotspots* (Clackson y Wells, 1995; Bogan y Thorn, 1998). De forma general, los *hotspots* son definidos como residuos que al ser mutados por alanina, provocan un aumento de la energía de unión mayor a 2.0 kcal/mol (Clackson y Wells, 1995; Cukuroglu y cols., 2014). Estos residuos no se distribuyen aleatoriamente en la IPP, sino que se presentan agrupados en regiones llamadas regiones calientes (Keskin y cols., 2005). Debido a que la interacción entre regiones calientes en la IPP es fundamental para la función biológica de las proteínas, estas regiones tienden a ser conservadas (Pavelka y cols., 2009). También se determinó que los residuos de mayor tamaño tienen mayor tendencia a comportarse como *hotspots*, siendo Tirosina, Arginina y Triptofano los *hotspot* más comunes (Bogan y Thorn, 1998). Existen bases de datos de *hotspots* determinados experimentalmente, como ASEdb (Thorn y Bogan, 2001) y BID (Fischer y cols., 2003).

Debido a la gran importancia de los *hotspots*, en la IPP y las funciones proteicas, y al alto costo de determinación experimental de estos residuos es que se han invertido grandes esfuerzos en la predicción computacional de estos. Se han utilizado diversos criterios, como la conservación (Pavelka y cols., 2009), el cambio de área superficial accesible al solvente (SASA) (Guharoy y Chakrabarti, 2005; Guney y cols., 2008) y el potencial estadístico (Dehouck y cols., 2013). Otra estrategia utilizada para la predicción de *hotspots* es mediante el uso de simulaciones de dinámica molecular (SDM) de sistemas atómicos integrales (*full-atom*). Esta aproximación permite un análisis detallado de las interacciones en la IPP llegando incluso a nivel atómico. Esta técnica conlleva un alto costo computacional, lo que dificulta su uso a gran escala (Cukuroglu y cols., 2014).

#### 4. Estructura del complejo FSH-FSHR

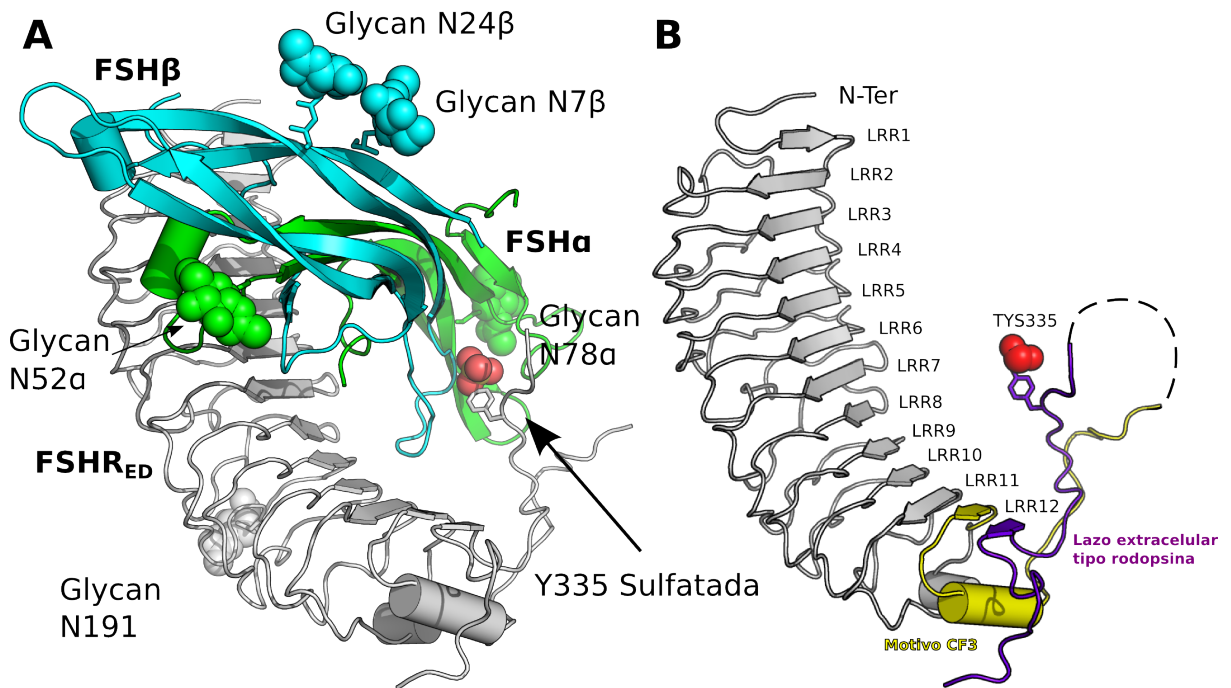
La FSH, al igual que el resto de las hormonas glicoprotéicas (GPHs), LH, gonadotropina coriónica (CG) y hormona estimulante de tiroide (TSH), está formada por dos cadenas. La cadena  $\alpha$  es común para todas las GPHs, mientras que la cadena  $\beta$  es específica de cada hormona. Ambas cadenas tienen plegamientos similares que están formados por motivos centrales de nudos de cisteína desde donde se extienden tres horquillas  $\beta$  (Fox y cols., 2001). Esta estructura se estabiliza con enlaces disulfuro intracatenarios e interacciones no covalentes entre cadenas. Este plegamiento es conservado entre las GPHs, como se observa en el modelo estructural cristalográfico de CG (Lapthorn y cols., 1994). Además, cada cadena de la FSH presenta dos glicosilaciones importantes para la persistencia en circulación y aclaramiento, y su bioactividad (Ulloa-Aguirre y cols., 1995; Ulloa-Aguirre y Timossi, 1998; Szkudlinski y cols., 1995; Ulloa-Aguirre y cols., 1999).

El receptor de la FSH (FSHR) pertenece a la familia de los receptores acoplados a proteína G (GPCR) que poseen 7 dominios transmembrana que participan en la transducción de la señal. A diferencia del común de los GPCRs, que poseen un pequeño ectodominio en el N-terminal, FSHR presenta un gran dominio extracelular, característico de los receptores de glicoproteínas (GPHRs), que unen ligandos de alrededor de 33 kDa.

El complejo de la FSH con el dominio de unión a hormona de su receptor (FSHR<sub>HB</sub>) fue cristalizado y publicado en la base de datos de proteínas (PDB) con el código 1XWD (Fan y Hendrickson, 2005). La interacción FSH/FSHR<sub>HB</sub> tiene forma de apretón de manos y presenta una superficie internalizada de 2.600 Å donde FSHR<sub>HB</sub> interacciona con la cadena  $\alpha$  y  $\beta$  de la FSH.

FSHR<sub>HB</sub> presenta 12 repeticiones de regiones ricas en lisina (LRR) que se muestran en la Figura 1B. Estas regiones forman la “palma” de la mano o FSHR<sub>HB</sub>. Además de la palma, el ectodominio del receptor de FSH (FSHR<sub>ED</sub>) presenta un lazo semejante a un pulgar de una mano derecha (Figura 1). La estructura de parte de este lazo fue descrita recientemente (Morrow y Zhang, 2012; Jiang y cols., 2012). Este lazo presenta 2 motivos distinguibles. El primero está constituido desde Lys264 a Arg298 que fue previamente identificado como motivo CF3 que existe específicamente en proteínas GPCR (Kajava, 1998). El segundo se extiende desde el residuo Tyr322 hasta el residuo Arg366, justo antes del primer dominio transmembrana. Este dominio es comparable con el lazo extracelular de rodopsina (Gromoll y cols., 1996), primer GPCR en tener estructura cristalográfica.

Con el objetivo de diseñar un ligando peptídico con afinidad por FSH, en este trabajo realizaremos un análisis detallado de la IPP presente en el complejo FSH-FSHR<sub>ED</sub> y diseñaremos un péptido derivado del exodominio del FSHR que tenga potencial uso en purificación de FSH por cromatografía de afinidad.



**Figura 1. Estructura del complejo FSH-FSHR.** La hormona folículo estimulante (FSH) está compuesta por una cadena  $\alpha$  (verde), que es común para todas las hormonas glicoproteicas, y una cadena  $\beta$  (cyan) que es exclusiva de cada hormona. Cada una de estas cadenas posee dos sitios de glicosilación (esferas del color de cada cadena) y múltiples enlaces disulfuro intracatenarios, sin embargo, estas cadenas son unidas solo por interacciones no covalentes. La FSH se une a su receptor de la forma mostrada en A. El ectodominio del receptor de FSH (FSHR<sub>ED</sub>) se muestra en B. Este ectodominio posee 12 repeticiones de regiones ricas en lisina (LRR) que forman la estructura similar a una palma que en su base presenta un lazo similar a un pulgar de una mano derecha. Este lazo presenta dos motivos reconocibles, el motivo CF3 (amarillo) y el motivo de lazo extracelular tipo rodopsina (púrpura). Dentro de este último motivo, el lazo presenta una tirosina sulfatada (sY335) como residuo modificado.

## **II. HIPÓTESIS**

El estudio de la interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub> nos permite realizar el diseño dirigido de un ligando peptídico con afinidad por la FSH humana.

## **III. OBJETIVOS**

### **OBJETIVO GENERAL**

**Diseñar un ligando peptídico con afinidad por FSH, utilizando análisis *in silico* de la interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub>, para su aplicación potencial en purificación de FSH por cromatografía de afinidad.**

### **Objetivos específicos**

1. Analizar la interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub> de las estructuras cristalográficas depositadas en la base de datos de proteínas PDB.
2. Analizar el comportamiento del complejo FSH-FSHR y su interfaz proteína-proteína, en un sistema de dinámica molecular con el ectodominio completo del receptor.
3. Diseñar un ligando peptídico con afinidad por FSH basado en los análisis de interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub> y evaluar su afinidad mediante un cálculo predictivo de energía libre de unión.

## IV. METODOLOGÍA

### 1. Análisis de interfaz de interacción desde complejos cristalográficos.

El análisis de interfaz se realizó en dos niveles en este trabajo. El nivel estructural y de superficie de interacción se analizó con el software web PDBePISA (<http://www.ebi.ac.uk/pdbe/pisa/>) disponible en la plataforma del Instituto Europeo de Bioinformática EBI. El nivel energético se analizó con el servidor web BeAtMuSiC (<http://babylone.ulb.ac.be/beatmusic>) que realiza un barrido mutagénico de los residuos de interés por todas las posibilidades de residuos naturales y entrega un valor predictivo del cambio de energía libre de unión basado en potenciales estadísticos.

### 2. Modelamiento Molecular.

La estructura del exodominio del FSHR fue modelada utilizando Modeller 9.12 (Sali y Blundell, 1993). La estructura molde se buscó a través de Blastp contra la base de datos de proteínas, PDB, utilizando la secuencia de FSHR y se seleccionó la mejor por criterios de cobertura y resolución. Posteriormente, se alineó la secuencia del exodominio de FSHR contra la secuencia del molde con el programa de alineamientos múltiples para secuencias de aminoácidos y nucleotídicas, MAFFT v7.149b (Kato y Standley, 2013). Se generaron 5 modelos de FSHR<sub>ED</sub> completo y se escogió el que obtuvo el menor puntaje DOPE (Shen y Sali, 2006).

### 3. Ensamble termodinámico del sistema para NAMD.

Para formar el sistema del complejo FSH con el exodominio completo del receptor se reemplazó la estructura cristalográfica de FSHR<sub>ED</sub> en 4AY9 por el modelo generado mediante un alineamiento estructural realizado en PyMol (Schrödinger, LLC 2010) con el comando *align*. Luego se guardaron en un nuevo archivo PDB las coordenadas de los átomos de FSH y el modelo de FSHR<sub>ED</sub>. Basado en este archivo de coordenadas se generó un archivo de estructura psf para NAMD (Phillips y cols., 2005) utilizando el paquete *psfgen*. Los parámetros y la topología para el residuo modificado sTyr fueron generados en este trabajo. Para esto se tomó como base los archivo de topología y parámetros de CHARMM36 (*Chemistry at HARvard Molecular Mechanics*) (Brooks y cols., 2009). En el caso de la topología para el residuo sulfotirosina sTyr, se hizo una mezcla entre el sulfato encontrado en la topología para lípidos de CHARMM36, mezclado con la tirosina del archivo de topología para proteínas de CHARMM36, de esta forma se creó el nuevo residuo que llamamos TYS. Para obtener los parámetros de TYS, utilizamos el servidor Swissparam (<http://www.swissparam.ch/>) con el archivo estructural de un modelo experimental de TYS obtenido de Expo Ligand (<http://ligand-expo.rcsb.org/>). Luego, los parámetros nuevos fueron agregados al archivo de parámetros de proteínas de CHARMM36. Luego se agregó una caja de agua de 15 Å hacia cada eje desde el complejo con



el *software solvate*, distribuido en el paquete Visual Molecular Dynamics, VMD (Humphrey y cols., 1996). Luego se agregó una concentración de 200 mM de NaCl con *autoionize*, extensión de VMD, neutralizando el sistema y llevándolo a condiciones fisiológicas. Para generar el ensamblaje del sistema FSH-péptido se obtuvieron las coordenadas del complejo FSH-FSHR<sub>ED</sub> en un estado de baja energía obtenido de la SDM generada en NAMD. Se eliminaron los residuos de FSHR que no pertenecen al péptido diseñado, se acetiló el extremo N-terminal y se generaron las mutaciones del diseño mediante el paquete *psfgen*.

#### **4. Simulaciones de Dinámica Molecular en NAMD**

Para las simulaciones de dinámica molecular (SDM) realizadas con NAMD2.9 (Phillips y cols., 2005) se utilizaron los parámetros CHARMM36 (Brooks y cols., 2009) modificados en esta tesis para incluir el residuo modificado de sulfotirosina TYS. Cada paso significó 2 fs. Las condiciones termodinámicas del sistema fueron NPT (isotérmico, isobárico) con presión de 1 atm y temperatura dependiente del ensayo. Cada producción fue precedida por una minimización de 3.000 pasos. Las trayectorias fueron registradas cada 5.000 pasos.

#### **5. Análisis de las simulaciones de dinámica molecular de NAMD**

Los análisis de los datos producidos fueron realizados principalmente mediante cuadernos de IPython (Perez y Granger, 2007) generados en el Laboratorio de Diseño de Drogas de la Universidad de Concepción. Estos cuadernos se basan en paquetes existentes para análisis bioinformático como ProDy (Bakan y cols., 2011) y MDAnalysis (Michaud-Agrawal y cols., 2011). Para generación de gráficos se utilizó el paquete Matplotlib (Hunter, 2007).

#### **6. Ensamblaje termodinámico para de sistema FSH-péptido para Amber.**

Para este sistema se utilizaron las mismas coordenadas utilizadas para el complejo FSH-péptido para NAMD. Para adaptar el PDB de NAMD a Amber (Case y cols., 2015) se utilizó el programa *pdb4amber* con el cual, entre otras cosas, se eliminaron todos los hidrógenos para que sean nombrados nuevamente según la nomenclatura de Amber por *tleap*. El campo de fuerza utilizado fue *ff14SB* de Amber. La parametrización del residuo TYS para Amber se realizó siguiendo el tutorial de parametrización de residuos no tradicionales proporcionado en la página del proveedor del software (<http://ambermd.org/tutorials/basic/tutorial5/>). Manualmente se indicó que el átomo de azufre tomara la conformación *s6* especificada en el Campo de Fuerza General de Amber (GAFF) (Case y cols., 2015), que une 4 átomos en conformación tetraédrica. Las cargas parciales de este residuo fueron designadas manualmente en base al estudio de mecánica cuántica de este residuo realizado en jaguar (Rapp y cols., 2013; Bochevarov y cols., 2013). Luego se generó una caja de agua TIP3P de 12 Å hacia cada eje

desde el complejo con solvatebox. Además se agregaron 2 iones de  $\text{Na}^+$  para neutralizar el sistema con addlons. Se utilizó ionsjc\_tip3p como campo de fuerza para iones.

### **7. Simulaciones de Dinámica Molecular en AMBER**

Previo a la simulación, el sistema fue sometido a una minimización de energía descendente abrupta de 10.000 pasos. Luego el sistema se calentó de 0 a 300 K en 50 ps. Se realizaron 50 ps de equilibración de densidad seguido de 500 ps de equilibración a presión constante a 300 K. Luego se continuó con una producción de 20 ns donde se grabaron las estructuras cada 2 ps.

### **8. Cálculo de energía libre de unión del complejo FSH-péptido por MM-PBSA.**

Este cálculo se realizó siguiendo el tutorial de MM-PBSA proporcionado en la página de Amber (<http://ambermd.org/tutorials/advanced/tutorial3/index.htm>), con algunas modificaciones. Se utilizaron los últimos 3 ns de una simulación de 20 ns. Se realizó un muestreo de 75 estructuras cada 40 ps cada una según el algoritmo de MMPBSA.py proporcionado por Amber (Case y cols., 2015).

## V. RESULTADOS

En esta tesis se buscó el diseño de ligandos peptídicos con afinidad por FSH con la proyección de generar un sistema de purificación de FSH por afinidad basado en los ligandos encontrados. Para ello se analizó la IPP del complejo FSH-FSHR<sub>ED</sub> con el fin de identificar los residuos de FSHR de mayor importancia en dicha interacción. También, para estudiar la participación en la IPP del lazo entre los residuos Ser295 y Thr331 sin estructura determinada, se generó un modelo del FSHR con la secuencia completa del ectodominio, se generó el complejo entre este modelo y la FSH y se realizó una simulación de dinámica molecular de 50 ns en condiciones fisiológicas. Finalmente, se diseñó un ligando con potencial afinidad por FSH, basado en la información generada en este trabajo y se evaluó su afinidad por cálculo predictivo de la energía de unión del complejo FSH-peptido. Los resultados obtenidos se describen a continuación.

### 1. Análisis de la interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub> de las estructuras cristalográficas depositadas en la base de datos de proteínas PDB.

El estudio de la interfaz del complejo FSH-FSHR busca la identificación de los residuos del FSHR que presentan mayor importancia en la interacción con el ligando. Para esto se analizaron características como superficie internalizada en la interfaz, potencialidad de formación de puentes de hidrógeno, potencialidad de formación de puentes salinos y cambio de energía libre de unión del complejo ( $\Delta\Delta G_{unión}$ ) utilizado barrido de alaninas *in silico*.

Existen tres archivos con coordenadas para el complejo FSH-FSHR depositados en la base de datos de estructura de proteínas, Protein Data Bank (PDB). Por orden de publicación, el primero es 1XWD (Fan y Hendrickson, 2005), que solo muestra la estructura del FSHR<sub>HB</sub>, o palma, que corresponde a la zona más rígida del ectodominio del receptor, y posee una resolución de 2,92 Å. El segundo es 4AY9 (Jiang y cols., 2012), que presenta parte del lazo extracelular del FSHR y tiene una resolución de 2,5 Å. Finalmente, el más reciente es 4MQW (Jiang y cols., 2014), que entrega información similar a 4AY9, pero que en su publicación presentan evidencia que muestra que FSHR actúa como un trímero. Esta última estructura posee una resolución de 2,9 Å. Debido a que entrega mayor información y posee la resolución más fina, los análisis los realizamos en base a la estructura 4AY9.

El estudio de la IPP del complejo lo realizamos en el servidor de proteínas, interfaces, estructuras y ensamblajes, **PDBePISA** (<http://www.ebi.ac.uk/pdbe/pisa/>) del Instituto Bioinformático Europeo (EBI). Este servidor entrega información de los residuos del complejo y sobre como participan en la interfaz, en particular, el servidor predice si el residuo está formando algún tipo de enlace tipo disulfuro o de hidrógeno, puente salino o enlace covalente (HSDC), además calcula el área accesible al solvente (ASA), el área de superficie internalizada (BSA), el porcentaje de área internalizada. Este estudio analiza la IPP entre pares de

moléculas. Por esta razón, se analizaron las interfaces entre la cadena  $\alpha$  de la FSH (FSH $_{\alpha}$ ) y FSHR $_{ED}$ , y la cadena  $\beta$  de la FSH (FSH $_{\beta}$ ) y FSHR $_{ED}$ . Debido a que el objetivo de este análisis es conocer los residuos de FSHR $_{ED}$  que participan en la unión de FSH, nos enfocaremos en los resultados referentes al receptor.

Los resultados del análisis de interfaz entre FSH $_{\alpha}$  y FSHR $_{ED}$  se encuentran en las tablas 1-5. El resumen mostrado en la tabla 1 indica que FSHR $_{ED}$  posee 141 átomos en la interfaz con FSH $_{\alpha}$ , presentes en 43 residuos que están detallados en la tabla 4. La predicción de enlaces de hidrógeno y puentes salinos se muestran en las tablas 2 y 3, respectivamente. El resultado del software predice la formación de 15 enlaces de hidrógeno y 6 puentes salinos entre FSH $_{\alpha}$  y FSHR $_{ED}$ .

Los resultados del análisis de interfaz entre FSH $_{\beta}$  y FSHR $_{ED}$  se encuentran en las tablas 6-10. El resumen mostrado en la tabla 6 indica que FSHR $_{ED}$  posee 99 átomos en la interfaz con FSH $_{\beta}$ , presentes en 36 residuos que están detallados en la tabla 9. La predicción de enlaces de hidrógeno y puentes salinos se muestran en las tablas 7 y 8, respectivamente. En este caso, el software predice 9 potenciales puentes de hidrógeno y 9 puentes salinos para FSH $_{\beta}$  y FSHR $_{ED}$ . Según el análisis de PDBePISA, la interfaz de FSHR $_{ED}$  y FSH está compuesta por 79 residuos pertenecientes al receptor, sin embargo, es sabido que en las interacciones proteína-proteína, es una fracción minoritaria de los residuos de la interfaz los que son responsables de la mayoría de la afinidad de unión (Moreira y cols., 2007; Clackson y Wells, 1995; Morrow y Zhang, 2012). Estos residuos críticos, comúnmente conocidos como *hotspots*, son usualmente definidos como posiciones en que una mutación causaría un aumento en la energía libre de unión de al menos 2,0 kcal/mol. Tradicionalmente, se ha utilizado el barrido de mutagénesis de alanina para caracterizar experimentalmente la interfaz proteína-proteína e identificar sus *hotspots*, que constituyen un blanco importante para la modulación de interacciones proteína-proteína (Wells, 1991; Bogan y Thorn, 1998). Para caracterizar la IPP de FSHR $_{ED}$  y FSH, y encontrar sus *hotspots*, utilizamos una aproximación de barrido de mutagénesis de alanina *in silico*, mediante el servidor bioinformático **BeAtMuSiC** (<http://babylone.ulb.ac.be/beatmusic>) (Thorn y Bogan, 2001; Dehouck y cols., 2013). Para este trabajo, definimos como *hotspots* los residuos que al ser modificados por alanina aumentan la energía libre de unión,  $\Delta\Delta G_{unión}$ , en más de 1,5 kcal/mol. La Figura 2 muestra los *hotspots* identificados para cada cadena del complejo, en la interacción FSH-FSHR $_{ED}$ . Generalmente, los *hotspots* se presentan en claros grupos diferenciados. Para este complejo se pueden distinguir 3 grupos de *hotspots* mostrados en Fig. 2A-C. En cada grupo de *hotspots*, el receptor posee un residuo con  $\Delta\Delta G_{unión}$  mayor a 2,0 kcal/mol, estos son Leu55, Tyr124 y el residuo modificado sTyr335 con 2,44; 2,55 y 3,26 kcal/mol, respectivamente.

Aunque estos grupos de *hotspots* son definidos por cercanía espacial, es decir, cercanos en el espacio de la estructura tridimensional; una característica que puede facilitar el diseño del ligando peptídico con afinidad por FSH es la cercanía de secuencia de los *hotspots*. En la Figura 2 podemos observar que los únicos *hotspots* con cercanía de secuencia de FSHR $_{ED}$  son

**Tabla 1. Resumen de interfaz entre la cadena  $\alpha$  (A) de la FSH y el exodominio del receptor (X).** Esta tabla muestra un resumen del análisis de interfaz realizado por PDBePISA, para la interacción entre la cadena  $\alpha$  de la FSH, denominada A, y el exodominio del receptor, denominada X, en el complejo 4ay9 de PDB.

Resumen de Interfaz				
	Estructura 1		Estructura 2	
<b>Rango de selección</b>	X		A	
Clase	Proteína		Proteína	
Operación de simetría	x,y,z		x,y,z	
ID de simetría	1_555		0_555	
<b>Número de átomos</b>				
interfaz	141	5.70%	96	14.20%
superficie	1388	56.40%	493	73.00%
total	2462	100.00%	675	100.00%
<b>Número de residuos</b>				
interfaz	43	14.00%	28	31.80%
superficie	275	89.60%	88	100.00%
total	307	100.00%	88	100.00%
<b>Area accesible al solvente (Å<sup>2</sup>)</b>				
interfaz	1009.7	6.40%	1089.5	15.30%
total	15885.2	100.00%	7110	100.00%
<b>Energía de solvatación (kcal/mol)</b>				
estructura aislada	-280.8	100.00%	-62.6	100.00%
ganancia en formación de complejo	-0.4	0.10%	-2.7	4.30%
ganancia promedio	-2.2	0.80%	-5.8	9.20%
P-value	0.672		0.806	

**Tabla 2. Enlaces de Hidrógeno entre la cadena  $\alpha$  (A) de la FSH y el exodominio de su receptor (X) en la estructura 4AY9.**

##	Estructura 1	Dist. [Å]	Estructura 2
1	X:ASN 106[ N ]	3.69	A:THR 46[ O ]
2	X:ASN 129[ ND2]	3.19	A:VAL 49[ O ]
3	X:LYS 74[ NZ ]	3.51	A:THR 86[ OG1]
4	X:TYS 335[ O3 ]	3.16	A:ASN 15[ ND2]
5	X:TYS 335[ O2 ]	3.01	A:GLN 27[ NE2]
6	X:ASP 81[ OD2]	2.95	A:ARG 42[ NH1]
7	X:ASP 81[ OD1]	3.82	A:ARG 42[ NH2]
8	X:GLN 79[ OE1]	2.74	A:SER 43[ OG ]
9	X:ASN 129[ OD1]	3.18	A:LEU 48[ N ]
10	X:ASN 129[ OD1]	2.87	A:VAL 49[ N ]
11	X:ASP 153[ OD2]	2.90	A:LYS 51[ NZ ]
12	X:GLN 145[ OE1]	3.66	A:ARG 67[ NH2]
13	X:GLU 99[ OE2]	3.01	A:THR 86[ OG1]
14	X:GLU 99[ OE1]	2.82	A:TYR 88[ OH ]
15	X:ASP 150[ OD2]	3.19	A:LYS 91[ NZ ]

**Tabla 3. Puentes salinos entre la cadena  $\alpha$  (A) de la FSH y el exodominio de su receptor (X) en la estructura 4AY9.**

##	Estructura 1	Dist. [Å]	Estructura 2
1	X:ASP 81[ OD2]	2.95	A:ARG 42[ NH1]
2	X:ASP 81[ OD1]	3.82	A:ARG 42[ NH2]
3	X:ASP 81[ OD2]	3.44	A:ARG 42[ NH2]
4	X:ASP 153[ OD1]	3.39	A:LYS 51[ NZ ]
5	X:ASP 153[ OD2]	2.90	A:LYS 51[ NZ ]
6	X:ASP 150[ OD2]	3.19	A:LYS 91[ NZ ]

**Tabla 4. Residuos de la interfaz del exodominio de FSHR (X) con la cadena  $\alpha$  (A) de la FSH en la estructura 4AY9.**

Estructura 1	HSDC	ASA	BSA	$\Delta iG$
X:GLU 34		89.88	1.11	-0.01
X:SER 35		62.17	5.28	-0.06
X:LEU 55		72.84	44.27	0.67
X:LYS 57		78.98	8.49	-0.05
X:GLU 73		66.67	11.57	-0.14
X:LYS 74	H	61.75	31.40	-0.09
X:GLN 79	H	80.16	39.46	-0.15
X:ASN 80		0.76	0.59	0.00
X:ASP 81	HS	59.96	40.17	-0.22
X:HIS 98		38.34	5.95	0.10
X:GLU 99	H	29.38	26.92	-0.38
X:ARG 101		59.20	13.90	0.01
X:LYS 104		75.27	37.49	0.52
X:ASN 106	H	95.74	63.52	0.36
X:GLN 123		76.51	13.81	0.12
X:TYR 124		52.50	49.30	0.36
X:LEU 126		19.38	7.59	0.12
X:ASN 129	H	80.45	74.73	-0.89
X:THR 130		8.73	8.61	0.00
X:GLY 131		4.66	4.50	0.07
X:LEU 144		106.12	1.48	-0.02
X:GLN 145	H	60.72	29.56	-0.46
X:LEU 148		36.61	30.97	0.50
X:ASP 150	HS	9.42	9.18	-0.10
X:GLN 152		34.44	16.64	-0.18
X:ASP 153	HS	56.93	40.35	-0.20
X:ILE 155		126.32	50.02	0.80
X:ASN 156		43.01	5.38	-0.06
X:GLU 171		67.10	7.23	-0.03
X:ILE 174		49.23	17.31	0.28
X:TRP 176		72.35	41.14	0.66
X:ASN 178		13.97	10.33	-0.12
X:LYS 179		89.40	14.36	-0.27
X:ASN 199		26.63	5.08	-0.06
X:HIS 289		69.35	3.20	-0.79
X:PRO 290		131.82	67.61	0.82
X:ILE 291		118.53	3.56	-0.04
X:CYS 292		33.45	5.31	0.15
X:LYS 294		134.76	23.11	-0.60
X:ASP 334		113.51	5.60	-0.08
X:TYR 335	H	256.65	105.40	-0.52
X:ASP 336		52.11	3.35	0.05
X:LEU 337		116.48	24.84	0.34

**Tabla 5. Residuos de la interfaz de la cadena  $\alpha$  (A) de la FSH con el exodominio de FSHR (X) en la estructura 4AY9.**

Estructura 2	HSDC	ASA	BSA	$\Delta iG$
A:ASN 15	H	19.78	14.87	-0.17
A:PRO 16		117.58	2.85	0.05
A:LEU 17		127.51	65.93	1.05
A:PHE 18		53.73	20.79	0.33
A:GLN 27	H	63.81	17.89	-0.20
A:ARG 42	HS	176.01	94.13	-0.59
A:SER 43	H	43.22	23.66	0.05
A:LYS 45		157.28	20.91	-0.11
A:THR 46	H	98.27	98.27	0.43
A:MET 47		44.94	31.96	0.81
A:LEU 48	H	184.00	114.10	1.55
A:VAL 49	H	101.32	63.17	0.88
A:LYS 51	HS	119.68	45.04	-0.72
A:LYS 63		132.09	12.05	-0.45
A:TYR 65		81.03	8.89	-0.05
A:ARG 67	H	184.74	29.71	-0.70
A:MET 71		110.27	33.34	0.23
A:GLY 72		81.53	8.78	0.09
A:PHE 74		115.80	57.99	0.93
A:HIS 83		65.36	3.93	-0.15
A:SER 85		35.24	19.08	-0.22
A:THR 86	H	85.71	54.81	0.27
A:CYS 87		12.70	4.42	-0.05
A:TYR 88	H	157.00	113.36	0.33
A:TYR 89		198.98	52.06	0.46
A:HIS 90		92.73	0.15	0.00
A:LYS 91	HS	159.36	56.45	-1.12
A:SER 92		165.52	20.86	-0.23



**Tabla 6. Resumen de interfaz entre la cadena  $\beta$  (B) de la FSH y el exodominio del receptor (X).** Esta tabla muestra un resumen del análisis de interfaz realizado por PDBePISA, para la interacción entre la cadena  $\beta$  de la FSH, denominada B, y el exodominio del receptor, denominada X, en el complejo 4ay9 de PDB.

Resumen de Interfaz				
	Estructura 1		Estructura 2	
<b>Rango de selección</b>	X		B	
Clase	Proteína		Proteína	
Operación de simetría	x,y,z		x,y,z	
ID de simetría	1_555		0_555	
<b>Número de átomos</b>				
interfaz	99	4.00%	88	10.60%
superficie	1388	56.40%	619	74.50%
total	2462	100.00%	831	100.00%
<b>Número de residuos</b>				
interfaz	36	11.70%	23	21.50%
superficie	275	89.60%	107	100.00%
total	307	100.00%	107	100.00%
<b>Area accesible al solvente (Å<sup>2</sup>)</b>				
interfaz	864.8	5.40%	906.6	11.00%
total	15885.2	100.00%	8267.9	100.00%
<b>Energía de solvatación (kcal/mol)</b>				
estructura aislada	-280.8	100.00%	-91.1	100.00%
ganancia en formación de complejo	1.4	-0.50%	-1.1	1.20%
ganancia promedio	-1.5	0.50%	-2.2	2.40%
P-value	0.818		0.643	

**Tabla 7. Enlaces de Hidrógeno entre la cadena  $\beta$  (B) de la FSH y el exodominio de su receptor (X) en la estructura 4AY9.**

##	Estructura 1	Dist. [Å]	Estructura 2
1	X:LYS 243[ NZ ]	2.60	B:ALA 43[ O ]
2	X:LYS 179[ NZ ]	3.86	B:SER 89[ O ]
3	X:LYS 104[ NZ ]	3.10	B:ASP 93[ OD1 ]
4	X:TYS 335[ O2 ]	2.71	B:VAL 38[ N ]
5	X:TYS 335[ O2 ]	3.79	B:TYR 39[ N ]
6	X:TYS 335[ OH ]	3.73	B:TYR 39[ N ]
7	X:GLU 197[ OE1 ]	2.53	B:ARG 44[ NH1 ]
8	X:GLU 76[ OE1 ]	2.92	B:ARG 97[ NH2 ]
9	X:GLU 34[ OE2 ]	2.63	B:TYR 103[ OH ]

**Tabla 8. Puentes salinos entre la cadena  $\beta$  (B) de la FSH y el exodominio de su receptor (X) en la estructura 4AY9.**

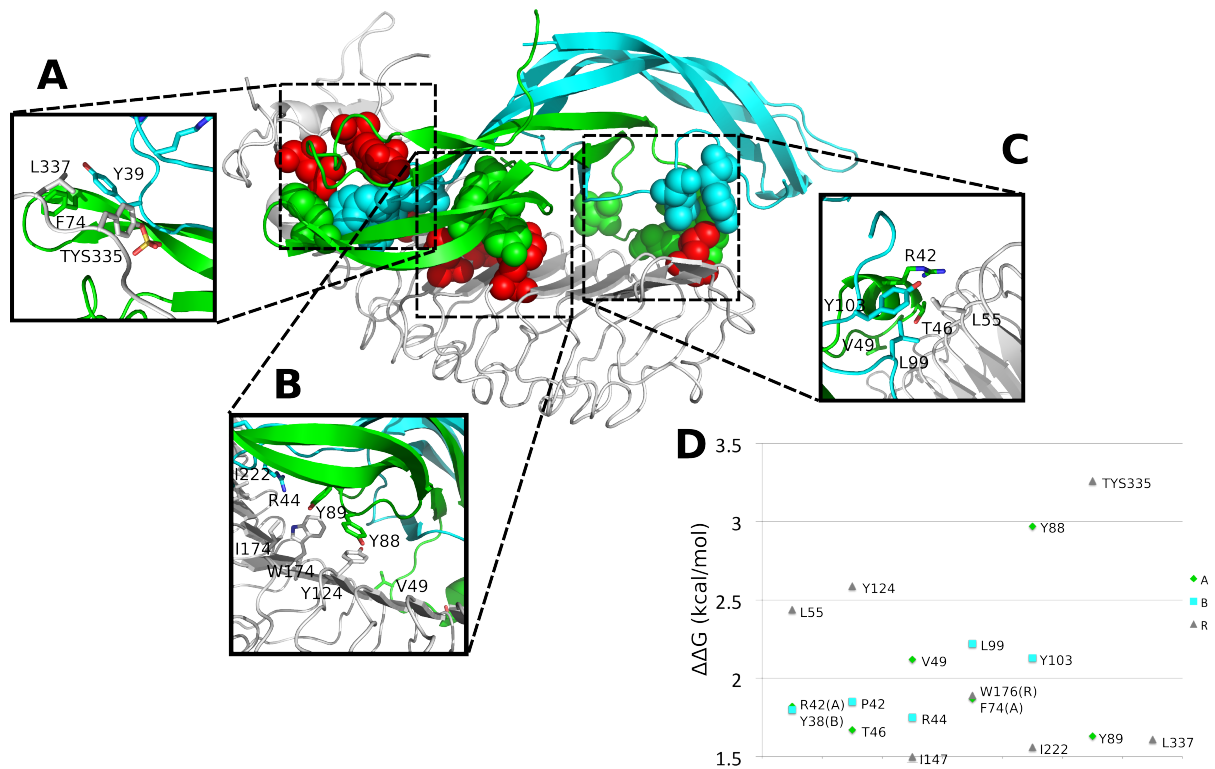
##	Estructura 1	Dist. [Å]	Estructura 2
1	X:LYS 179[ NZ ]	3.97	B:ASP 90[ OD1 ]
2	X:LYS 104[ NZ ]	3.10	B:ASP 93[ OD1 ]
3	X:LYS 104[ NZ ]	3.21	B:ASP 93[ OD2 ]
4	X:GLU 197[ OE1 ]	3.92	B:ARG 44[ NE ]
5	X:GLU 197[ OE1 ]	2.53	B:ARG 44[ NH1 ]
6	X:GLU 197[ OE2 ]	3.17	B:ARG 44[ NH1 ]
7	X:GLU 197[ OE1 ]	3.90	B:ARG 44[ NH2 ]
8	X:GLU 76[ OE1 ]	2.92	B:ARG 97[ NH2 ]
9	X:GLU 76[ OE2 ]	3.86	B:ARG 97[ NH2 ]

**Tabla 9. Residuos de la interfaz del exodominio de FSHR (X) con la cadena  $\beta$  (B) de la FSH en la estructura 4AY9.**

Estructura 1	HSDC	ASA	BSA	$\Delta iG$
X:GLN 33		65.88	16.02	-0.17
X:GLU 34	H	89.88	23.04	-0.27
X:GLU 50		23.98	4.42	-0.05
X:ARG 52		82.90	39.32	-0.91
X:VAL 54		22.70	7.45	0.12
X:LEU 55		72.84	28.56	0.46
X:GLU 76	HS	30.14	23.80	-0.34
X:SER 78		3.92	0.50	-0.01
X:GLN 79		80.16	40.69	-0.55
X:GLU 99		29.38	0.99	-0.01
X:ARG 101		59.20	45.30	-0.90
X:GLU 103		23.49	19.39	-0.17
X:LYS 104	HS	75.27	37.78	-0.41
X:LEU 126		19.38	2.29	0.04
X:SER 128		21.51	0.25	0.00
X:ASN 129		80.45	3.20	-0.04
X:LYS 146		116.96	37.24	0.50
X:GLN 152		34.44	7.71	-0.13
X:ASP 153		56.93	9.57	-0.16
X:ILE 174		49.23	26.11	0.42
X:TRP 176		72.35	7.35	0.12
X:ASN 178		13.97	0.29	0.00
X:LYS 179	HS	89.40	54.10	0.33
X:ASP 196		48.27	35.79	0.04
X:GLU 197	HS	34.42	24.40	-0.26
X:ASP 202		34.20	5.24	0.00
X:VAL 221		48.79	26.95	0.43
X:ILE 222		40.84	36.15	0.58
X:LYS 242		79.30	16.86	0.13
X:LYS 243	H	60.96	45.50	-1.25
X:ARG 245		66.13	7.28	-0.21
X:MET 265		79.18	3.64	0.15
X:THR 331		197.09	14.19	0.18
X:ASP 334		113.51	23.72	-0.11
X:TYS 335	H	256.65	143.27	0.33
X:LEU 337		116.48	46.45	0.74

**Tabla 10. Residuos de la interfaz de la cadena  $\beta$  (B) de la FSH con el exodominio de FSHR (X) en la estructura 4AY9.**

Estructura 2	HSDC	ASA	BSA	$\Delta iG$
B:TYR 33		146.32	0.47	0.01
B:ARG 35		177.38	39.23	-0.18
B:ASP 36		118.49	0.12	0.00
B:LEU 37		61.05	31.45	0.50
B:VAL 38	H	148.96	12.83	0.09
B:TYR 39	H	172.53	54.97	0.55
B:LYS 40		124.62	27.47	0.17
B:ASP 41		52.96	25.47	-0.23
B:PRO 42		121.95	83.04	0.89
B:ALA 43	H	84.51	75.42	0.78
B:ARG 44	HS	136.93	90.01	-1.17
B:PRO 45		88.81	18.40	0.29
B:LYS 46		94.21	25.11	-0.65
B:SER 89	H	87.11	38.71	-0.02
B:ASP 90	S	140.66	28.51	-0.27
B:ASP 93	HS	117.09	39.29	-0.36
B:CYS 94		66.64	4.49	-0.05
B:THR 95		79.36	30.62	0.28
B:VAL 96		125.82	33.59	0.51
B:ARG 97	HS	197.15	100.47	-1.51
B:GLY 98		58.48	10.21	0.16
B:LEU 99		142.79	84.70	1.34
B:TYR 103	H	112.98	51.96	-0.05



**Figura 2. Determinación de Hot-spots de interfaz FSH-FSHR.** La determinación de *hotspots* se realizó mediante un barrido de mutagénesis de alanina *in silico* realizado por el servidor BeAtMuSiC. Se denominó *hotspot* a un residuo que modificado por alanina aumenta la energía libre de unión,  $\Delta\Delta G$ , en más de 1.5 kcal/mol. El resultado muestra la identificación de 3 grupos de *hotspots*. Los *hotspots* del receptor son mostrados como esferas rojas, mientras que los de las cadenas de la FSH son mostrados como esferas del color de la cadena (A: verde y B: cyan). Los distintos grupos se muestran en detalle en A, B y C, donde se muestran los residuos participantes. En D se muestra el cambio de energía libre de unión,  $\Delta\Delta G$ , de cada *hotspot* (A: cadena  $\alpha$ , B: cadena  $\beta$  y R: receptor). Se observa que el mayor  $\Delta\Delta G$  es presentado por el residuo TYS335 del receptor y que es significativamente mayor con respecto al resto de los *hotspots* del receptor.

sTyr335 y Leu337.

Estos resultados llevan nuestra atención hacia el residuo modificado de FSHR, sTyr335, debido a que es el residuo con mayor superficie enterrada, tanto en FSH<sub>α</sub> como en FSH<sub>β</sub>, además posee cercanía de secuencia con el *hotspot* Leu337 y genera el mayor aumento en la energía libre de unión al ser modificado por alanina, mostrándose como el *hotspot* de FSHR<sub>ED</sub> más importante en esta interacción proteína-proteína. De esta forma, FSHR<sub>sTyr335</sub> se convierte en un punto de partida para el diseño de un ligando peptídico con afinidad por FSH.

## **2. Análisis del comportamiento del complejo FSH-FSHR y su interfaz proteína-proteína, en un sistema de dinámica molecular con el ectodominio completo del receptor.**

En el objetivo anterior encontramos que FSHR<sub>sTyr335</sub> es el *hotspot* más importante de la interacción FSH-FSHR<sub>ED</sub>. Sin embargo, este residuo se encuentra a cuatro residuos, hacia el N-Terminal, de una región de estructura indeterminada experimentalmente. Debido a que esta región es rica en residuos hidrofóbicos (Figura 3), creemos que puede aportar residuos importantes en esta interacción.

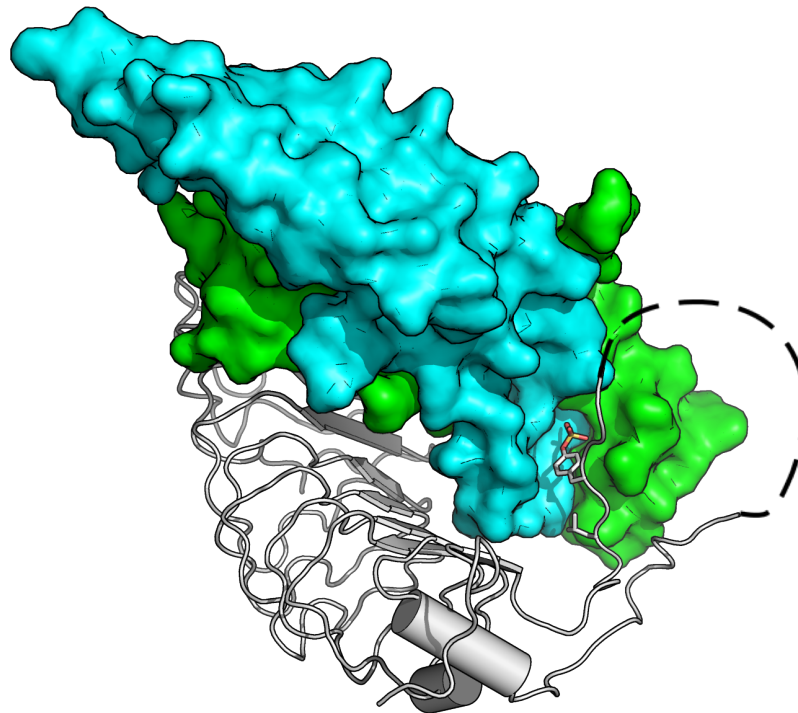
### **2.1 Modelamiento del exodominio completo de FSHR**

Para estudiar este lazo de estructura indeterminada decidimos generar un modelo estructural del exodominio completo de FSHR con el programa de modelamiento estructural por homología, Modeller versión 9.12 (Webb y Sali, 2014; Sali y Blundell, 1993). Como molde utilizamos la estructura 4AY9, que posee la mayor cobertura de la secuencia del exodominio y la menor resolución, de las encontradas a través de Blastp contra la base de datos de PDB utilizando la secuencia de FSHR. Posteriormente, se alineó la secuencia del exodominio de FSHR contra la secuencia del exodominio de FSHR encontrada en 4AY9, que carece la secuencia del lazo de estructura indeterminada. Se generaron 5 modelos de FSHR<sub>ED</sub> completo y se escogió el que obtuvo el menor puntaje DOPE. El puntaje DOPE (Discrete Optimized Protein Energy) es una medida relacionada a energía derivada del potencial estadístico de distancia de los átomos, tomando como referencia las distancias atómicas encontradas las estructuras depositadas en PDB. Esto conlleva a que el puntaje DOPE no dependa de parámetros ajustables (Shen y Sali, 2006). Para formar el complejo FSH-FSHR<sub>ED</sub> completo, se alineó el modelo de FSHR<sub>ED</sub> con FSHR<sub>ED</sub> de la estructura 4AY9, para posteriormente, eliminar los átomos de FSHR<sub>ED</sub> original y ser reemplazados por el modelo que contiene el lazo.

### **2.2 Generación de sistema para simulación de dinámica molecular**

Para estudiar el comportamiento de este lazo en una situación fisiológica decidimos realizar una simulación de dinámica molecular del complejo modelado FSH-FSHR<sub>ED</sub>. Debido a que los archivos de topología y parámetros necesarios para generar el sistema de dinámica molecular

SELHPICNKSILRQEVDYMTQARGQRSSLAEDNESSYRSGFDMTYTEFD(sY)DLCNEVVDVTCSP



**Figura 3. Región de FSHR<sub>ED</sub> de estructura indeterminada.** En la parte superior de la figura se muestra la secuencia de la región de estructura indeterminada en gris, rodeada por la secuencia de la estructura conocida en negro. En rojo se muestran los *hotspots* encontrados en esta región donde sY es la sulfotirosina TYS335. En la parte inferior de la figura se muestra la ubicación espacial del lazo de estructura indeterminada en forma de línea punteada y su cercanía con los *hotspots* de FSHR, TYS335 y L337, para quienes se muestra su cadena lateral.

y ejecutar la simulación, contienen solo la información de los residuos tradicionales, y que el complejo FSH-FSHR<sub>ED</sub> posee un residuo modificado, FSHR<sub>sTyr335</sub>, que es crucial para la interacción de estas proteínas, fue necesario generar un archivo de topología y de parámetros específico para nuestro sistema. Para esto se tomó como base los archivo de topología y parámetros de CHARMM36 (*Chemistry at HARvard Molecular Mechanics*) (Brooks y cols., 2009). En el caso de la topología para el residuo sulfotirosina sTyr, se hizo una mezcla entre el sulfato encontrado en la topología para lípidos de CHARMM36, mezclado con la tirosina del archivo de topología para proteínas de CHARMM36, de esta forma se creó el nuevo residuo que llamamos TYS, su nombre en nomenclatura de tres letras. Para obtener los parámetros de TYS, utilizamos el servidor Swissparam (<http://www.swissparam.ch/>) con el archivo estructural de un modelo experimental de TYS obtenido de Expo Ligand (<http://ligand-expo.rcsb.org/>). Luego, los parámetros nuevos fueron agregados al archivo de parámetros de proteínas de CHARMM36.

El archivo de extensión PDB del modelo del complejo FSH-FSHR<sub>ED</sub> generado en 2.1 contiene información de coordenadas atómicas, sin embargo, para realizar una simulación de dinámica molecular es necesario tener un archivo de información estructural de extensión psf. Para generar el psf del complejo se utilizó el paquete psfgen que es una herramienta de construcción de estructuras. Al generar el archivo de estructura con psfgen, se debe contar con un archivo de topología CHARMM que contenga todos los residuos o grupos atómicos presentes en el pdb de origen. En esta etapa utilizamos el archivo de topología CHARMM36 generado anteriormente que contiene la topología para el residuo TYS además de los otros residuos comunes. También, al generar el archivo de estructura, se debe tener en cuenta aspectos como la protonación de residuos y la formación de puentes disulfuro. Para determinar la protonación de los residuos del complejo FSH-FSHR<sub>ED</sub> a pH fisiológico 7.4, se utilizó el servidor propka ([www.propka.org](http://www.propka.org)), el que no indica cambio de estado de protonación a los residuos del complejo. Los puentes disulfuro fueron obtenidos de las estructuras cristalizadas para el complejo y se establecieron como enlaces disulfuro en la generación del archivo de información estructural psf. Luego se solvató este complejo generando una caja de agua con 20 Å hacia cada eje del complejo. Finalmente, se le agregaron iones de NaCl a una concentración fisiológica de 200 mM con la extensión del programa Visual Molecular Dynamics (VMD), autoionize. Esto generó un sistema de 97.401 átomos.

### **2.3 Simulación de dinámica molecular del complejo FSH-FSHR<sub>ED</sub>**

La simulación de dinámica molecular (SDM) se realizó utilizando el programa NAMD, para dinámicas moleculares escalables (Webb y Sali, 2014; Phillips y cols., 2005). Para acelerar la obtención de resultados utilizamos la tecnología de cálculo GPU, CUDA, que paraleliza el cálculo en los múltiples núcleos de tarjetas gráficas especializadas desarrolladas por NVIDIA. Esta tecnología nos permitió obtener el resultado de 50 ns de simulación en 7 días de cálculo



que hubiera tardado 4,6 meses en el *cluster* CPU con que disponíamos en la Universidad. La estación de trabajo GPU que utilizamos fue facilitada por Exxact Corporation (<http://www.exxactcorp.com/>), compañía proveedora californiana líder en computo de alto rendimiento y soluciones de visualización.

Para observar si la SDM logró llegar a una estabilización en 50 ns se generó un gráfico de RMSD de cada cadena del complejo FSH-FSHR<sub>ED</sub> durante toda la trayectoria contra  $t_0$  (Figura 4). El gráfico muestra que ya a los 15 ns de simulación, las cadenas del complejo tienen un comportamiento estable, lo que indica que la SDM llegó a un equilibrio en 15 ns y desde este tiempo hasta 50 ns se realizarán los análisis posteriores.

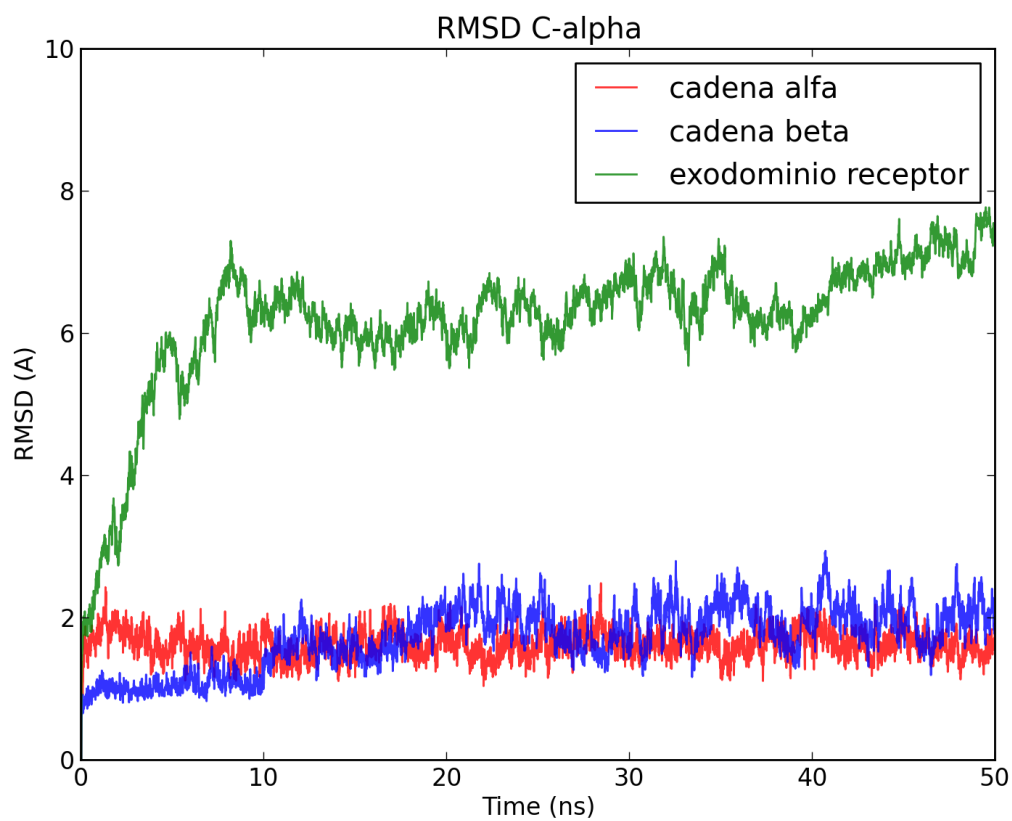
Una forma de validar los resultados de la SDM es comparando los residuos con alta movilidad en la trayectoria de la SDM con el factor beta de los carbonos  $\alpha$  (C $\alpha$ ) de los residuos, que es una medida experimental de la movilidad del C $\alpha$  en el cristal. Para ello se grafica el RMSF (root mean square fluctuations) que indica el movimiento de los residuos durante la SDM contra el factor beta para los residuos de cada cadena del complejo (Figura 5). En cada caso se observa una estrecha relación entre RMSF y factor beta, lo que valida nuestros resultados. De forma interesante, para el lazo modelado de FSHR<sub>ED</sub> se observan valores muy altos de RMSF y no se cuenta con datos de factor beta debido a que estos residuos no poseen determinación experimental. Esta alta movilidad mostrada en la SDM puede ser causal de la dificultad para determinar la estructura de este segmento.

Para encontrar estructuras representativas dentro de la SDM, analizamos la energía total de la porción proteica del sistema en la trayectoria total (Figura 6). La energía se analizó utilizando la extensión del programa visualizador VMD, NAMD Energy. Escogimos las estructuras correspondientes a 14,7; 16,4 y 46,5 ns porque son mínimos locales energéticos importantes en la trayectoria de la SDM.

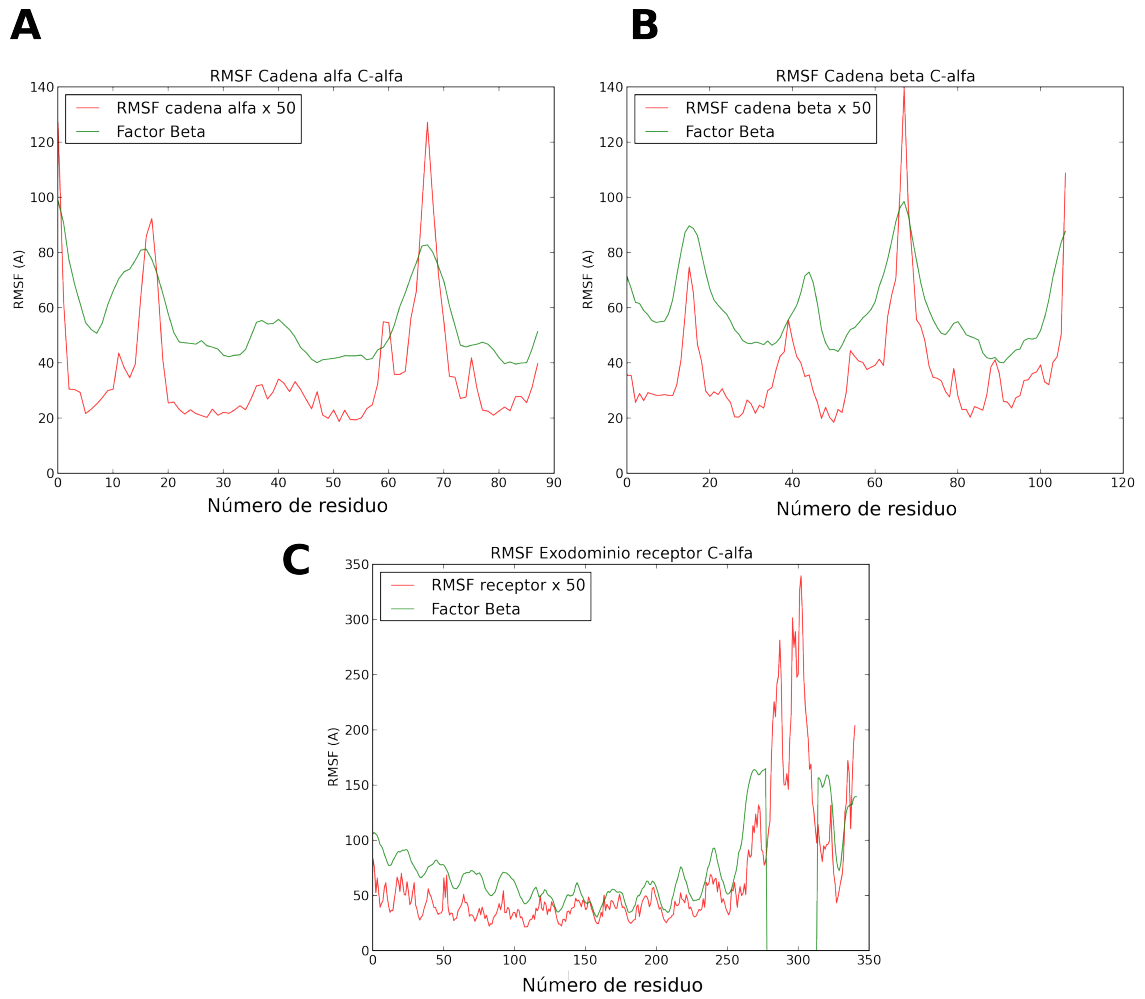
Para cada una de estas estructuras fueron analizadas en PDBePISA para conocer el comportamiento del lazo modelado (Tabla 11). Sorprendentemente, el residuo Met328 de FSHR<sub>ED</sub> muestra, de forma consistente, una alta área de superficie internalizada, similar a lo mostrado por los *hotspots* cercanos sTyr335 y Leu337. Posteriormente se realizó un análisis de barrido de mutagenesis de alanina *in silico* para encontrar los *hotspots*, de forma similar a lo realizado en el objetivo específico 1, en la estructura obtenida de la SDM (Figura 7). Este resultado muestra a Met328 como un nuevo *hotspot*, obtenido mediante el modelamiento del lazo de estructura indeterminada y la posterior SDM del complejo completo de 50 ns.

### **3. Diseño de un ligando peptídico con afinidad por FSH basado en los análisis de interfaz proteína-proteína del complejo FSH-FSHR<sub>ED</sub> y evaluación de su afinidad mediante un cálculo predictivo de energía libre de unión.**

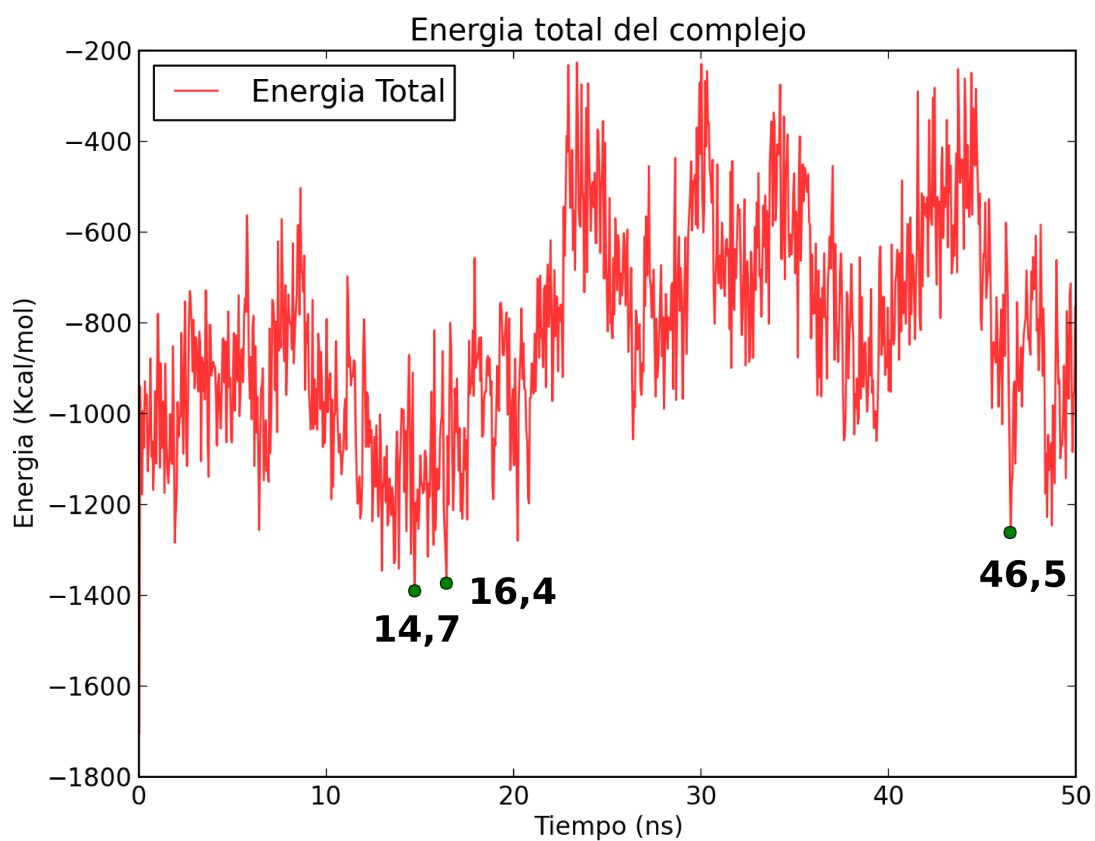
En este objetivo se diseña un ligando peptídico para la FSH para utilización en purificación por cromatografía de afinidad y se evalúa la energía de unión del complejo FSH-peptido mediante un cálculo predictivo de energía de unión basado en SDM.



**Figura 4. Análisis de la trayectoria de la simulación de dinámica molecular por RMSD.** Para cada cadena del complejo FSH-FSHR<sub>ED</sub> se calcula el RMSD durante toda la trayectoria contra  $t_0$ . En rojo se muestra FSH<sub>α</sub>, en azul FSH<sub>β</sub> y en verde, FSHR<sub>ED</sub>. Se observa que todas las cadenas del complejo, luego de movimientos importantes, llegan a una estabilización.



**Figura 5. Validación de SDM por comparación de RMSF y Factor Beta.** En esta figura se muestra el gráfico de correlación entre el RMSF y el factor beta para cada cadena del complejo FSH-FSHR<sub>ED</sub>. El cálculo de RMSF se realizó con los últimos 10 ns de la simulación. Para que la correlación fuera visible en un gráfico, se multiplicó el RMSF por un factor de 50.



**Figura 6. Energía total durante la SDM del complejo FSH-FSHR<sub>ED</sub>.** Se graficó la energía total de la porción proteica del sistema, obtenida con NAMD energy, durante el trayecto completo de la simulación. En verde se muestran los tres mínimos locales más importantes. Se escogió analizar el marco correspondiente a 46,5 ns porque involucra un alto grado de avance en la simulación y porque es un mínimo local importante.

**Tabla 11. Análisis de interfaz del complejo FSH-FSHR<sub>ED</sub> en el lazo modelado en los mínimos locales energéticos extraídos de la SDM.** HSDC: Residuos formando enlaces Hidrógeno/Disulfuro, puentes Salinos o unión Covalente. ASA: área de superficie accesible, Å<sup>2</sup>. BSA: área de superficie internalizada, Å<sup>2</sup>. IIII porcentaje de área internalizada, una barra por 10%.

Estructura	16.4 ns			46.5 ns		
	HSDC	ASA	BSA	HSDC	ASA	BSA
GLY 325		40.46	0.00		56.71	0.00
PHE 326		185.71	0.00		116.46	0.00
ASP 327		86.90	0.00		125.24	0.00
MET 328		118.97	62.40 IIIII		130.67	69.77 IIIII
THR 329		130.32	16.28 II		84.38	3.90 I
TYR 330		53.62	0.00		31.23	0.00
THR 331		14.31	0.00		38.44	0.00
GLU 332		116.95	0.00		135.46	0.00
PHE 333		78.58	14.05 II		87.23	5.96 I
ASP 334		91.06	11.22 II		108.50	25.59 III
TYS 335	H	254.69	206.45 IIIIIIIII	H	270.49	214.48 IIIIIIIII
ASP 336		33.59	3.55 II		33.43	6.49 II
LEU 337		194.73	151.60 IIIIIIIII		168.50	98.39 IIIIIII
CYS 338		23.75	0.00		37.96	0.00
ASN 339		75.35	40.74 IIIII		91.29	36.91 IIIII
GLU 340		51.51	0.00		83.61	0.00
VAL 341		76.24	0.00		82.67	0.00
VAL 342		21.26	0.00		86.53	0.00

### **3.1 Diseño de ligando peptídico con afinidad por FSH para uso en cromatografía de afinidad.**

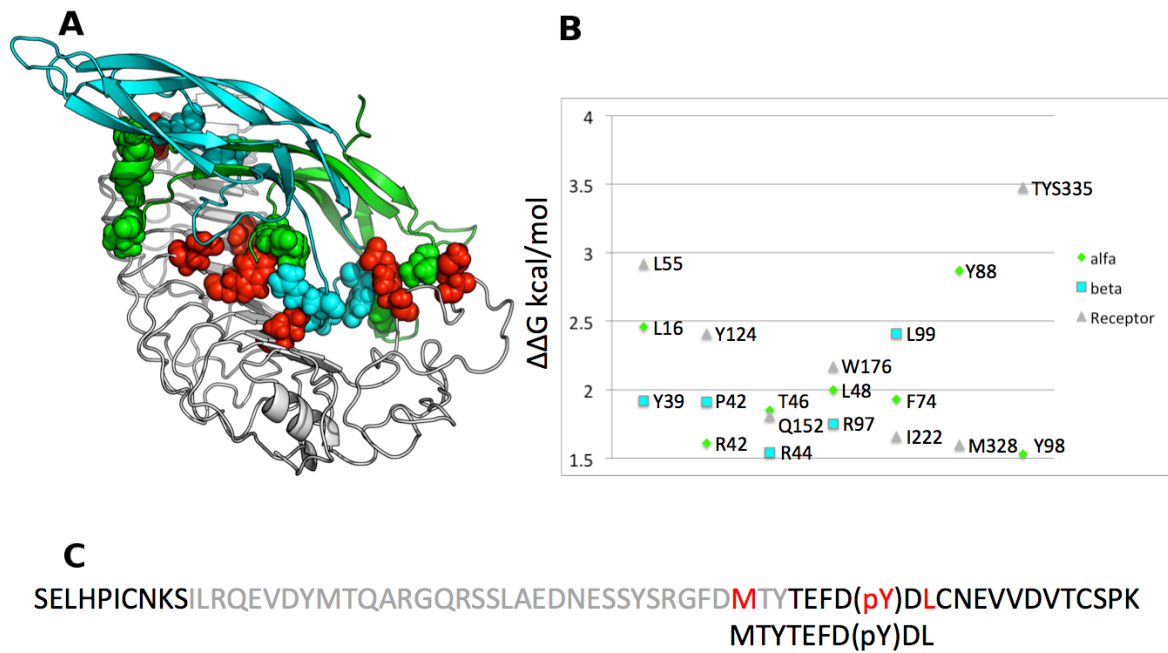
El diseño del ligando peptídico con afinidad por FSH se realizó en base a los residuos del lazo exoplasmático de FSHR. Específicamente, se utilizó la secuencia MTYTEFD(sY)DLCN mostrada en la Figura 7C. Se escogió esta secuencia ya que fue la única que encontramos que posee *hotspots* con cercanía en la secuencia (M328, sY335 y L337). Para orientar el ligando peptídico de forma similar a como se presenta en FSHR decidimos que debe unirse a la matriz por el extremo C-terminal del péptido. Para esto, diseñamos los extremos del péptido de forma que se una por el extremo C-terminal a una matriz de N-Hidroxisuccinimidil-sefarosa, reactivo que une aminas primarias. Para lograr la orientación correcta, acetilaremos la amina primaria del extremo N-terminal y el residuo asparragina (N) del C-terminal lo mutaremos los lisina (K), residuo que donará la amina primaria de su grupo radical para lograr la unión a la matriz. Finalmente, para evitar reacciones conducentes a enlaces covalentes no deseados, decidimos mutar la cisteína (C) por serina (S), residuo con semejanza estructural, pero que no genera puentes disulfuro. De esta forma, el ligando peptídico diseñado queda de la siguiente forma.



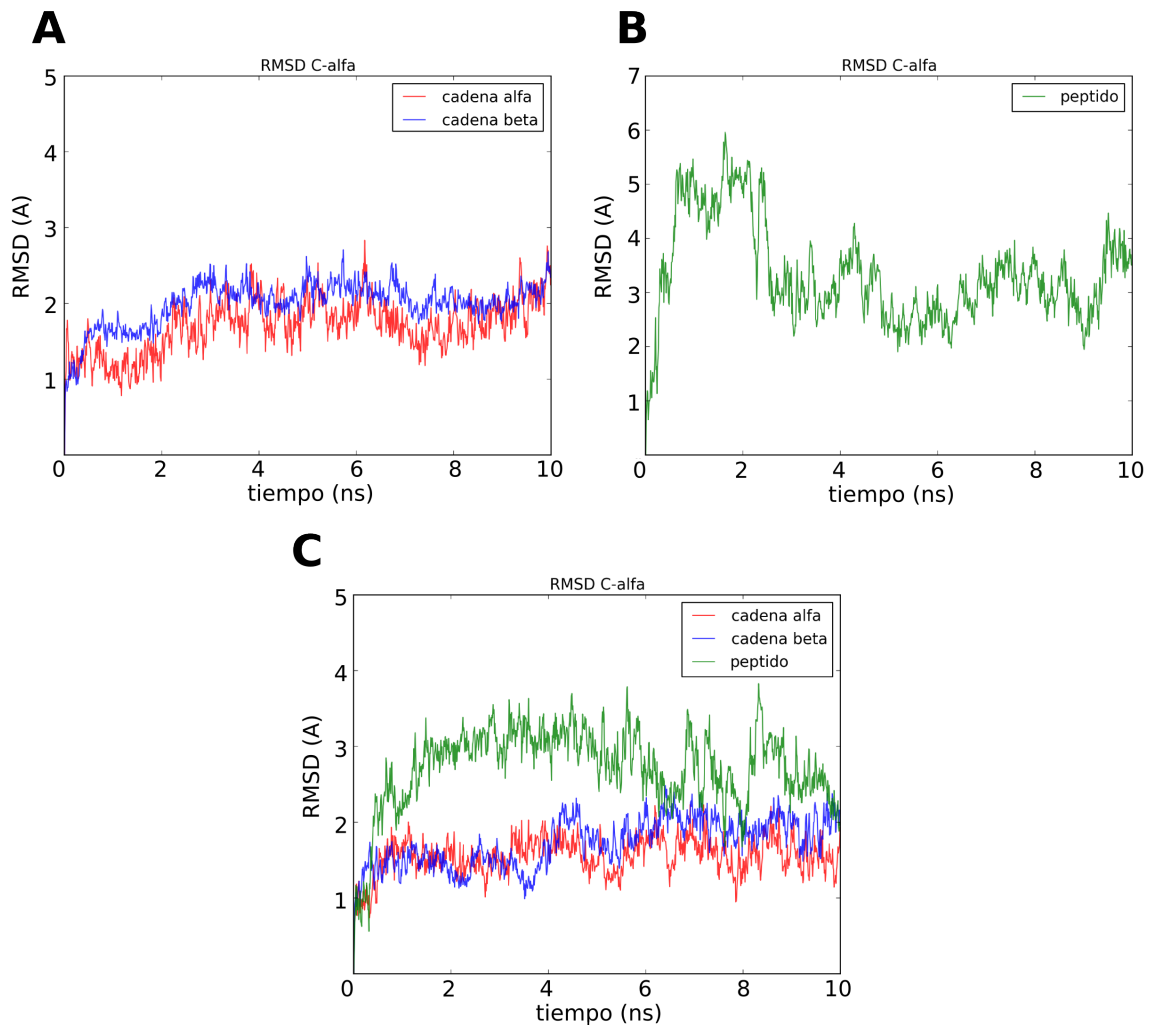
### **3.2 Evaluación del ligando peptídico con afinidad por FSH mediante estudios de simulación de dinámica molecular.**

El péptido definido anteriormente fue generado bioinformáticamente, en base al marco de tiempo 46,5 ns de la SDM del complejo FSH-FSHR<sub>ED</sub>. Se extrajo el péptido original de lazo de FSHR<sub>ED</sub>, se realizaron las mutaciones C338S y N339K, y se acetiló el extremo N-terminal con el paquete psfgen. Para realizar un análisis del comportamiento del complejo FSH-péptido se realizaron 3 SDM, la proteína FSH, el péptido y el complejo FSH-péptido, todas en agua con 200 mM de NaCl. Para simular las condiciones de un proceso de purificación se utilizaron condiciones de temperatura estándar, es decir 25° C o 298 K. Se generaron 10 ns de simulación por cada sistema. Al analizar las simulaciones por RMSD durante la trayectoria encontramos que todos los sistemas se encuentran estabilizados en 5 ns (Figura 8). Para hacer cálculos posteriores se utilizarán los últimos 3 ns de simulación de cada sistema.

Para hacer un análisis de la estabilización del sistema en la formación del complejo FSH-péptido realizamos una comparación de los RMSF promedio por residuo en los sistemas independientes con el sistema del complejo FSH-péptido. Para el péptido, se observa que la formación del complejo disminuye el movimiento de todos los residuos que lo componen (Figura 9). En particular, se presentan disminuciones de RMSF importantes para los residuos (enumerados con respecto al péptido) Met1 y Thr2, ambos residuos del extremo N-terminal de péptido y modelados en esta tesis. Para los residuos Tyr3, Phe6, sTyr8, Asp9 y Leu10 se

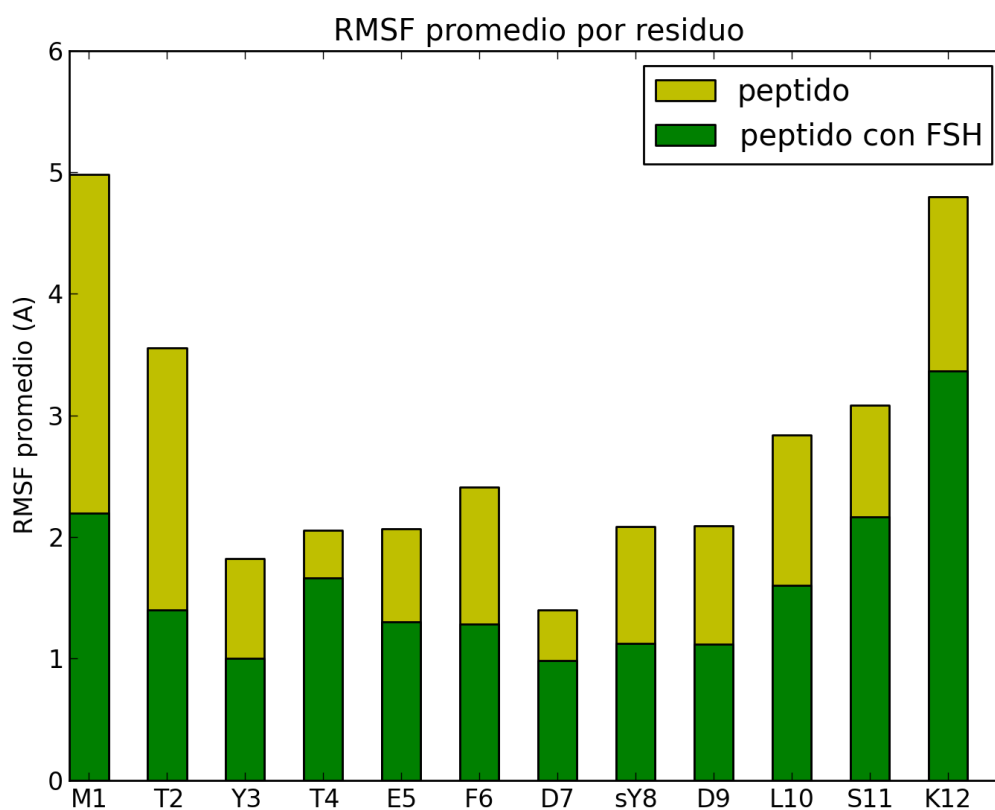


**Figura 7. Análisis de hotspots de estructura a 46,5 ns de simulación.** En A se observa la distribución espacial de los grupos de hotspots. En B se grafican los hotspots con respecto al  $\Delta\Delta G$  producido al ser mutado cada residuo por alanina. Se puede apreciar una gran conservación con respecto a los hotspots obtenidos de la estructura cristalográfica mostrados en la Figura 2. De lazo modelado, es considerado hotspot el residuo M328. En C se muestra en negro la secuencia de FSH<sub>ED</sub> con estructura determinada, en gris, la secuencia del lazo modelado, y en rojo, los hotspots encontrados en esta zona. Si bien es cierto L337 no alcanza un valor de  $\Delta\Delta G$  para ser considerado hotspot en este análisis, está muy cerca del umbral y en el análisis de la estructura cristalográfica si obtiene un valor de  $\Delta\Delta G$  para ser considerado hotspot. Bajo esta secuencia mostramos el diseño del péptido que creemos que tiene altas posibilidades de comportarse como un ligando de afinidad para la FSH.



**Figura 8. Trayectorias de las SDM de los sistemas FSH, péptido y complejo FSH-péptido, analizadas por RMSD.** Se graficaron los RMSD durante la trayectoria de cada cadena peptídica en las distintas SDM, FSH (A), péptido (B) y complejo FSH-péptido.





**Figura 9. RMSF promedio de residuos del péptido diseñado en condiciones de ausencia y presencia de la proteína FSH.** El gráfico muestra el RMSF promedio de cada residuo del ligando peptídico diseñado. En amarillo se muestra el RMSF resultado de la simulación con sólo el péptido. En verde se muestra el RMSF del péptido en el complejo FSH-péptido. Se observa que para todos los casos, el RMSF disminuye cuando el péptido se encuentra acomplexado con la proteína FSH.

observa una disminución del RMSF promedio por residuo cercana a la mitad en la condición en complejo.

Para analizar el RMSF promedio por residuo de la proteína FSH, seleccionamos previamente los residuos de la FSH que participan en la interfaz con el ligando peptídico. Esto lo realizamos mediante el servidor PDBePISA. Estos resultados fueron graficados por cadena al igual que se hizo para el péptido (Figura 10). Estos resultados muestran diferencias más conservadoras entre ambas condiciones. Esto es esperable debido a que una proteína globular como la FSH posee una estructura tridimensional estabilizada intrínsecamente. Sin embargo, hay 3 residuos en la cadena  $\alpha$ , Met71, Gly73 y Phe74, y un residuo en la cadena  $\beta$ , Asp41, que presentan disminuciones en el RMSF en la condición acomplejada. También en la cadena  $\beta$  se observa que los residuos Leu37 y Val38 aumentan su RMSD en la situación de complejo. Aún así, la estabilización en los residuos analizados es mayor en la situación en complejo FSH-péptido.

### 3.3 Predicción de energía libre de unión del complejo FSH-péptido

Para estimar la energía libre de unión ( $\Delta G_{\text{unión}}$ ) del complejo FSH-péptido se utilizó una aproximación consistente en un muestreo de estados del complejo derivados de una SDM equilibrada y cálculo de  $\Delta G_{\text{unión}}$  basado en Mecánica Molecular con electrostática de Poisson-Boltzmann y solvatación de área superficial (MM/PBSA).

MM/PBSA es un método de punto final para cálculo de  $\Delta G_{\text{unión}}$  originalmente desarrollado para el software de SDM AMBER (Kollman y cols., 2000; Miller y cols., 2012) y posteriormente implementado para el software GROMACS (Kumari y cols., 2014). Este método utiliza fotografías de dinámicas estabilizadas para calcular el  $\Delta G_{\text{unión}}$  mediante la fórmula:

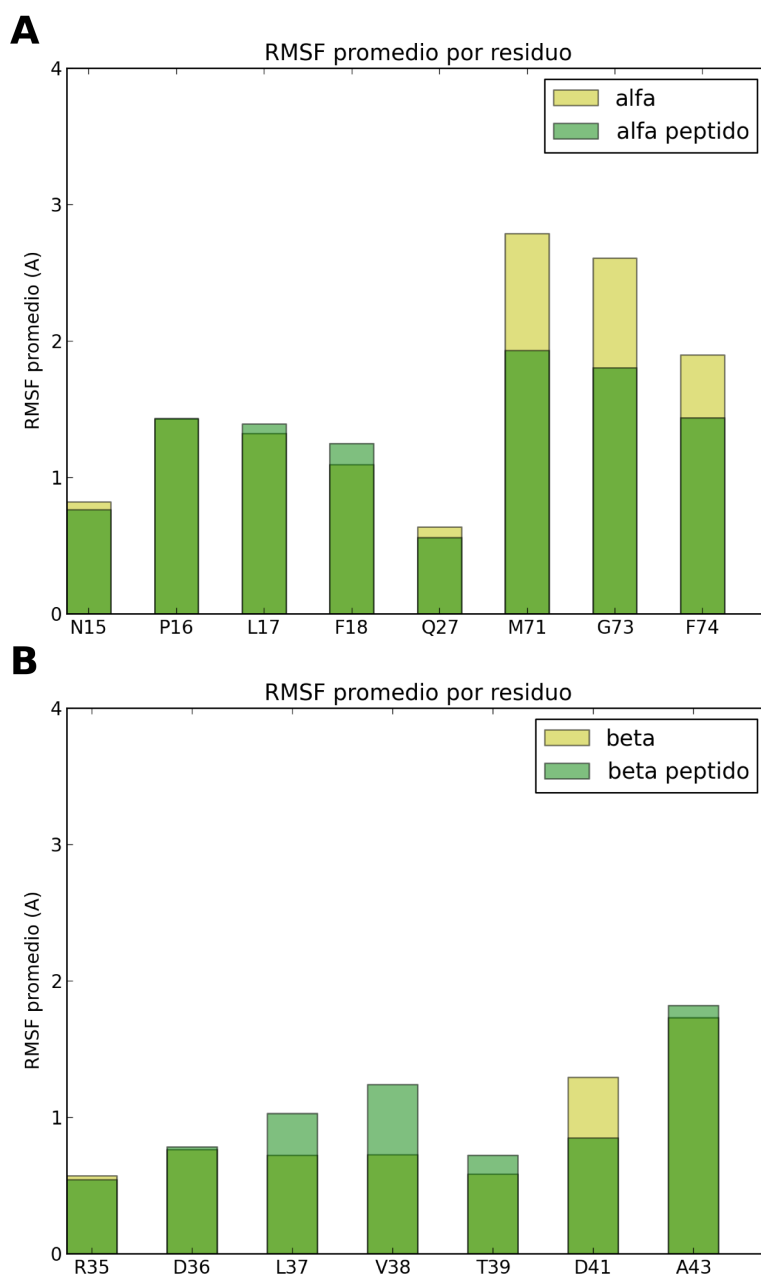
$$\Delta G_{\text{unión}} = G_{\text{PL}} - G_{\text{P}} - G_{\text{L}}$$

Donde PL es complejo proteína-ligando, P es la proteína, L el ligando y G es la energía libre de cada uno de ellos. Para MM/PBSA G es estimada según la siguiente suma:

$$G = E_{\text{enl}} + E_{\text{el}} + E_{\text{vdW}} + G_{\text{pol}} + G_{\text{np}} - TS$$

Donde los tres primeros términos corresponden a los términos de energía de MM estándar de enlace (enlace, ángulo y dihédrico), electrostática e interacciones de van der Waals.  $G_{\text{pol}}$  y  $G_{\text{np}}$  son las contribuciones polares y no polares de la energía libre de solvatación.  $G_{\text{pol}}$  es obtenido resolviendo la ecuación de PB mientras que el término no polar es estimado de una relación lineal al área superficial accesible al solvente (SASA). El último término de la ecuación es la temperatura absoluta,  $T$ , por la entropía,  $S$ , estimada por un análisis de modo normal de las frecuencias vibracionales.

Debido a que este cálculo está mejor implementado en SDM realizadas con el paquete de SDM Amber, es que trasladamos nuestro sistema FSH-péptido generado en NAMD a un sistema de Amber. Esto significó una nueva parametrización del residuo TYS que está especificada en la metodología (VI-6). Previo a la SDM, se realizó una minimización de energía descendente abrupta de 10.000 pasos. Luego se realizó una etapa de calentamiento del sistema de 0 a 300



**Figura 10. RMSF promedio de los residuos de interfaz de la FSH con el ligando peptídico en condiciones de ausencia y presencia del péptido.** En cada gráfico se muestran los residuos de interfaz de la FSH con el péptido, determinados mediante el servidor PDBePISA, de la cadena  $\alpha$  (A) y la cadena  $\beta$  (B) de la FSH. En amarillo se muestra la condición en ausencia de péptido mientras que en verde se muestra la condición en complejo FSH-péptido. Para la cadena  $\alpha$  se observan 3 residuos que disminuyen su RMSF, M71, G73 y F74, mientras que el resto de los residuos no muestran mayores diferencias en comportamiento. Por otra parte, para la cadena  $\beta$  hay un comportamiento más difuso que muestra al residuo D41 más estabilizado en condición de complejo FSH-péptido, mientras el residuo V38 se muestra más estable en la condición sin péptido. El resto de los residuos de interfaz de la cadena  $\beta$  muestra ligeros aumentos de RMSF en la condición con péptido (L37, T38 y A43) o no muestra diferencias entre ambas condiciones.

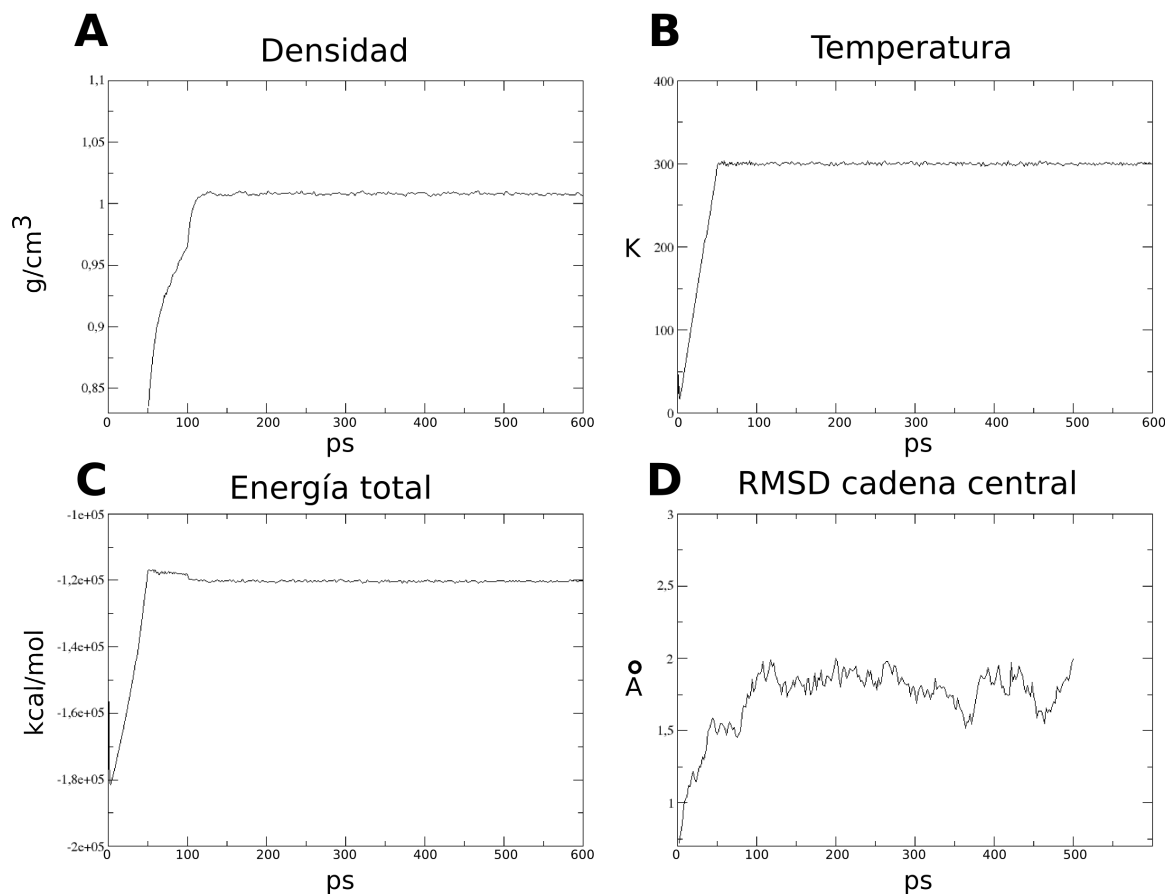
K en 50 ps (Figura 11B). Posteriormente, se llevó a densidad estable de 1 g/cm<sup>3</sup> en un proceso de otros 50 ps (Figura 11A). Finalmente, se realizó una equilibración del sistema por 500 ps. Esta equilibración fue evaluada por los parámetros de energía total del sistema (Figura 11C) y RMSD de la cadena central (Figura 11D).

A partir de este sistema equilibrado se procedió a realizar una SDM productiva de 20 ns rescatando, cada 2 ps, los parámetros evaluados anteriormente y las estructuras generadas. La estabilidad de la producción se puede observar en la Figura 12A-C. El RMSD de la cadena central durante la SDM productiva muestra cierta estabilidad, sin embargo, también muestra un cambio importante cerca de los 16 ns. Es por esto que para el cálculo de energía libre de unión con MM/PBSA, que requiere una trayectoria estable, se utilizaron los últimos 3 ns. El valor predictivo de  $\Delta G_{\text{unión}}$  para el complejo FSH-péptido obtenido por esta metodología es de  $-15,4 \pm 6,9$  kcal/mol. Este valor entrega soporte a la utilización del péptido diseñado en este trabajo para utilización en purificación de FSH por cromatografía de afinidad, mostrando de forma predictiva que la unión de la proteína FSH al péptido diseñado lleva a la formación de un complejo estable.

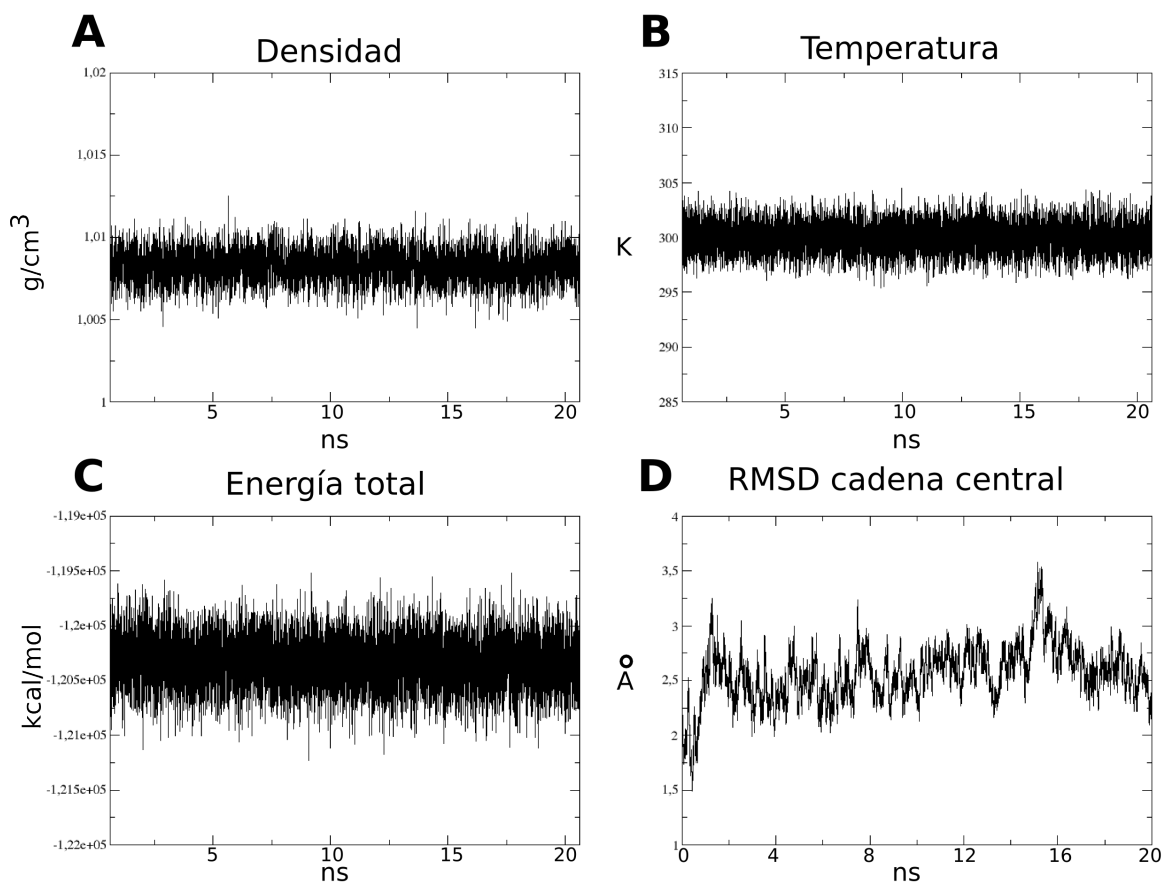
## VI. DISCUSIÓN

En este trabajo se ha mostrado el diseño y la evaluación de un ligando peptídico con afinidad por la proteína FSH para potencial uso en purificación por cromatografía de afinidad. El abordaje del diseño fue basado en el estudio de la IPP del complejo FSH-FSHR, de estructura recientemente publicada (Jiang y cols., 2012). Esta aproximación fue validada durante el transcurso de nuestro trabajo por la publicación en que diseñan ligandos peptídicos con afinidad por capsómeros de partículas tipo virus (Li y cols., 2014). Estas partículas poseen importancia en las áreas de la vacunología, terapia génica, entrega de fármacos y ciencia de materiales. Ellos utilizaron la interacción conocida entre el capsomero Cap y el extremo C-terminal de la proteína menor de envoltura, VP2-C, depositada como estructura cristalográfica en la base de datos de proteínas, PDB, con el código 1CN3. Ellos utilizaron análisis energéticos para determinar los residuos más importantes en la interacción y establecer una biblioteca de diversidad de 6859 péptidos de 8 residuos de largo con residuos invariables (*hotspots*) y residuos variables. Posteriormente, mediante progresivos tipos de barrido virtual se fue filtrando los mejores ligandos, para con los últimos mejores 10 candidatos, realizar estudios de SDM y análisis energético de  $\Delta G$  mediante Mecánica Molecular con Poisson-Boltzmann y solvatación de área superficial (MM/PBSA).

El motivo por el cual el método de MM/PBSA es cada vez más usado (Genheden y Ryde, 2015) es por que es un intermedio razonable entre exactitud del cálculo y requerimientos computacionales siendo los más exacto que los métodos tradicionales de puntaje de docking molecular y no requiere tanta capacidad computacional como los cálculos de perturbación



**Figura 11. Equilibración del sistema FSH-péptido. Cada gráfico muestra distintos parámetros evaluados durante el proceso de equilibración.** En (A) se muestra que la densidad llega a 1 g/cm<sup>3</sup> al finalizar los primeros 100 ps de simulación. (B) muestra la etapa de calentamiento que lleva el sistema de 0 K a 300 K durante los primeros 50 ps y luego se mantiene constante a 300 K. La energía total del sistema (C) llega a un equilibrio en  $-1,2 \times 10^5$  kcal/mol cerca de los primeros 150 ps. Finalmente, el RMSD de la cadena central contra la estructura inicial se muestra en (D). Este RMSD se calculó solo para los últimos 500 ps de equilibración que corresponde al tiempo de simulación sin restricciones. Todos estos parámetros muestran estar estables lo que significa una equilibración exitosa.



**Figura 12. Parámetros de la simulación productiva de 20 ns de FSH-péptido.** Cada gráfico muestra distintos parámetros evaluados durante el proceso de simulación productiva del sistema FSH-péptido. Los parámetros de densidad (A), temperatura (B) y energía total del sistema (C) se observan constantes durante la producción. El RMSD de la cadena central contra la estructura inicial se muestra en (D). Este RMSD se calculó para los 20 ns de simulación productiva. Para el cálculo de energía libre de unión por MM/PBSA se utilizó desde 10 a 13 ns debido que muestran un comportamiento estable en el RMSD.

alquímica. Finalmente, el trabajo de Li *et al.*, demuestra experimentalmente que el péptido mejor calificado *in silico*, es capaz de capturar la proteína de interés en condiciones de purificación por cromatografía de afinidad.

### 1. Estudio de la estructura cristalográfica

Nuestro punto de partida fue el estudio de la estructura cristalográfica del complejo FSH-FSHR<sub>ED</sub>. En la base de datos de proteínas, PDB, existen hasta la fecha 3 estructuras para el complejo FSH-FSHR. Nosotros decidimos basarnos en la estructura de código 4AY9 ya que posee la mejor resolución y una mayor cobertura del exodominio del receptor. Para estudiar el área superficial de interacción, global y por residuo, utilizamos el servidor web PDBePISA (Krissinel y Henrick, 2007; Krissinel, 2010). Este servidor está enfocado en la determinación de la presencia de complejos proteicos de relevancia biológica en las estructuras cristalográficas determinadas experimentalmente. Sin embargo, nosotros lo utilizamos para obtener información sobre los residuos de interfaz del complejo FSH-FSHR<sub>ED</sub>, y para determinar los residuos del receptor que más aportan a la interacción en términos de superficie. Luego, para evaluar la importancia de los residuos de interfaz en la energía libre de unión,  $\Delta G_{\text{unión}}$ , para el complejo FSH-FSHR<sub>ED</sub>, utilizamos el servidor web BeAtMuSiC (Dehouck y cols., 2013), que realiza predicciones del cambio de energía libre de unión,  $\Delta\Delta G_{\text{unión}}$ , de un barrido de mutaciones puntuales. Nosotros consideramos las mutaciones de cada residuo de interfaz por alanina para de esta forma establecer la presencia de *hotspots*. Consideramos como hotspot al residuo que al ser mutado por alanina genera un  $\Delta\Delta G_{\text{unión}}$  mayor a 1,5 kcal/mol. Este valor fue escogido debido a que nos permitió observar las 3 agrupaciones de los *hotspots* presentes en el complejo, que con valores de corte de  $\Delta\Delta G_{\text{unión}}$  mayores, como 2,0 kcal/mol, no podríamos haber detectado. Los dos análisis realizados en esta parte del estudio coinciden, mediante diferentes parámetros, que el residuo modificado sTyr335 de FSHR, es el residuo que más aporta a la interacción del complejo por parte del receptor.

Además, la agrupación de *hotspots* donde se presenta sTyr335, es la única que muestra residuos cercanos en la secuencia proteica en el receptor (sTyr335, Leu337). Esto último lo consideramos como un hecho que facilita el diseño del ligando de afinidad por FSH.

La importancia de esta región se ha demostrado para el receptor de la hormona luteinizante, proteína estrechamente relacionada a la FSH, para la cual se sabe que los residuos Tyr331 y Asp330 del receptor son claves para la activación del mismo (Bruysters y cols., 2008). Por estas razones, nos enfocamos en esta región de FSHR<sub>ED</sub> para realizar el diseño del ligando peptídico de afinidad por FSH.

### 2. Modelamiento y simulación de dinámica molecular del complejo FSH-FSHR<sub>ED</sub>

La región en que se encuentran los *hotspots*, pTyr335 y Leu337, es en el lazo exoplasmático de FSHR. La estructura de este lazo sólo se conoce en parte, razón por la cual, completamos la estructura exoplásmica mediante modelamiento por Modeller. Si bien es cierto, Modeller realiza

modelamiento por homología y no poseemos un molde estructural para el lazo, utilizamos esta metodología esperando observar el comportamiento del lazo en una SDM. Previamente a realizar las dinámicas, un paso crítico para obtener resultados útiles fue la parametrización del residuo de tirosina sulfatada, TYS, en campos de fuerza CHARMM36. Esto se realizó usando información del modelo experimental de TYS en la base de datos de ligandos, Ligand Expo (Feng y cols., 2004) e información de la topología del grupo sulfato unido a un carbono obtenida de la topología CHARMM36 para lípidos (Klauda y cols., 2010). Podemos comparar el resultado de nuestra parametrización del residuo TYS con lo obtenido por Rapp et al., quienes parametrizaron el residuo mediante el software de cálculos de mecánica cuántica, Jaguar (Rapp y cols., 2013). La mayor diferencia en cargas parciales entre nuestra parametrización y la realizada por Jaguar es de 0,04 lo que indica diferencias menores y que los resultados obtenidos serán representativos de la realidad.

La SDM del complejo FSH-FSHR<sub>ED</sub> generada comprendió 50 ns en condiciones fisiológicas, es decir, 37° C y una concentración de sales de 200 mM. Además, se realizó en una caja de agua con condiciones periódicas de borde, lo que simula un sistema infinito en que una molécula que sale por una cara del sistema, aparece por la opuesta. Consideramos que el sistema está estabilizado a los 15 ns, lo que se muestra en la figura 4. La correlación entre el factor beta y el RMSF de las distintas cadenas, que fue calculado con los últimos 10 ns de la simulación del sistema, valida nuestra simulación y parametrización del residuo TYS. Para analizar el resultado de la dinámica, utilizamos marcos de baja energía evaluados por energía total de la porción proteica del sistema mediante NAMDEnergy. Los marcos evaluados arrojaron resultados similares en residuos de interfaz por PDBePISA (Tabla 11). Para búsqueda de *hotspots* se analizó solo el marco de trayectoria más avanzada, que corresponde a 46,5 ns. Este análisis muestra un *hotspot* nuevo, M328, ubicado en el lazo modelado de FSHR. Este residuo también muestra un área de superficie considerable participando en la interacción con la FSH, lo que nos sugiere la importancia de este residuo y que puede ser utilizado en el diseño del ligando de afinidad por FSH.

### **3. Diseño y evaluación del ligando peptídico con afinidad por FSH**

Para el diseño tomamos en cuenta la secuencia de FSHR que contiene los 3 *hotspots* del lazo extracelular, Met328, pTyr335 y Leu337. La cisteína presente en la secuencia fue reemplazada por serina para evitar la formación de enlaces covalentes indeseados. También, para lograr una adecuada presentación del péptido, lo diseñamos de tal forma que una el C-terminal a la matriz de sefrosa N-hidroxisuccinamida activada, que une aminas primarias. Para ello bloqueamos el extremo N-terminal con una acetilación y en el extremo C-terminal, presentamos una lisina, como lo realizado por Kelley et al., para la unión a la matriz del ligando peptídico de afinidad por Factor VIII (Kelley y cols., 2004).

Luego de generar *in silico* el péptido con las modificaciones anteriormente señaladas, se generó un sistema con el péptido acoplado a FSH según el marco en 46,5 ns de la dinámica



del complejo FSH-FSHR<sub>ED</sub>. Este sistema FSH-péptido se llevó a dinámica por 10 ns en condiciones de temperatura ambiente, para simular un proceso de purificación, y con concentración de sales de 200 mM. Además, se generaron simulaciones del péptido y la FSH individualmente en las mismas condiciones. Todas estas dinámicas se encontraban estabilizadas a los 5 ns de simulación. Como una aproximación a la estabilización provocada a por la interacción FSH-péptido, analizamos el RMSF promedio por residuo en el péptido y en los residuos de interfaz de la FSH. Para el péptido es notoria una mayor estabilización en el complejo, mientras que para la FSH es más sutil pero aún así se puede considerar una estabilización. Nuestra suposición es que un complejo más estable, con menor movimiento, se traduce en una menor energía lo que significa una tendencia a la formación del complejo. Para corroborar esta supuesta estabilidad, recurrimos a otro paquete de SDM, Amber, que tiene implementado el cálculo de  $\Delta G_{\text{unión}}$  predictivo de MM/PBSA. Con esta metodología, obtuvimos un valor de  $\Delta G_{\text{unión}}$  de  $-15,4 \pm 6,9$  kcal/mol. Este valor negativo de  $\Delta G_{\text{unión}}$  significa que el complejo FSH-péptido se comporta de manera estable. Sin embargo, al compararlo al valor de  $-76,31$  kcal/mol obtenido en el trabajo de Li et al. (Li y cols., 2014), podríamos decir que es necesario someter nuestro péptido a un proceso de optimización similar al trabajo de Li, en que modifican todos los residuos que no son *hotspots*, por toda la combinatoria de residuos naturales para luego realizar distintas clases de filtros y llegar al péptido con mejor afinidad por la proteína.

Nuestros resultados sugieren que el péptido diseñado en este trabajo, Ac-MTYTEFD(sY)DLSK, es capaz de unir establemente a la proteína FSH y tiene potencial de, 1) ser utilizado como ligando de afinidad para FSH, o 2) ser optimizado para lograr mejores afinidades.

## VII. CONCLUSIONES

Mediante este trabajo, identificamos la región del lazo exoplasmático de FSHR como una región con alto potencial para diseño de ligandos peptídicos con afinidad por FSH e identificamos una secuencia de 12 residuos que posee afinidad por FSH *in silico* por predicción de  $\Delta G_{\text{unión}}$ .

## VIII. AGRADECIMIENTOS

Se agradece al financiamiento otorgado por Becas Conicyt para estudios de Magíster en Chile e INNOVA BIOBIO con el financiamiento de apoyo a la realización de tesis de postgrado, proyecto 13.1279-EM.TES.

También se agradece a las instituciones que facilitaron sus instalaciones para realizar los cálculos de simulación presentes en este trabajo. Estas instituciones son Exxact Corporation, California, Estados Unidos (<http://exxactcorp.com>) y el Laboratorio Nacional de Cómputo de Alto Rendimiento, NLHPC, ubicado en el Centro de Modelamiento Matemático (CMM), Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile (<http://www.nlhpc.cl>).



## IX. REFERENCIAS

- Bakan, A, L M Meireles, and I Bahar. 2011. "ProDy: Protein Dynamics Inferred From Theory and Experiments." *Bioinformatics (Oxford, England)* 27 (11) (May 25): 1575–1577. doi:10.1093/bioinformatics/btr168.
- Bochevarov, Art D, Edward Harder, Thomas F Hughes, Jeremy R Greenwood, Dale A Braden, Dean M Philipp, David Rinaldo, Mathew D Halls, Jing Zhang, and Richard A Friesner. 2013. "Jaguar: a High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences." *International Journal of Quantum Chemistry* 113 (18) (July 4): 2110–2142. doi:10.1002/qua.24481.
- Bogan, A A, and K S Thorn. 1998. "Anatomy of Hot Spots in Protein Interfaces.." *Journal of Molecular Biology* 280 (1) (July 3): 1–9. doi:10.1006/jmbi.1998.1843.
- Brooks, B R, C L Brooks, A D Mackerell, L Nilsson, R J Petrella, B Roux, Y Won, et al. 2009. "CHARMM: the Biomolecular Simulation Program.." *Journal of Computational Chemistry* 30 (10) (July 30): 1545–1614. doi:10.1002/jcc.21287.
- Bruysters, Martijn, Miriam Verhoef-Post, and Axel P N Themmen. 2008. "Asp330 and Tyr331 in the C-Terminal Cysteine-Rich Region of the Luteinizing Hormone Receptor Are Key Residues in Hormone-Induced Receptor Activation.." *The Journal of Biological Chemistry* 283 (38) (September 19): 25821–25828. doi:10.1074/jbc.M804395200.
- Case, D A, J Berryman, R M Betz, D S Cerutti, T E Cheatham, III, T A Darden, et al. 2015. "Amber 2015."
- Clackson, T, and J A Wells. 1995. "A Hot Spot of Binding Energy in a Hormone-Receptor Interface.." *Science (New York, N.Y.)* 267 (5196) (January 20): 383–386.
- Cukuroglu, Engin, H Billur Engin, Attila Gursoy, and Ozlem Keskin. 2014. "Hot Spots in Protein-Protein Interfaces: Towards Drug Discovery." *Progress in Biophysics and Molecular Biology* 116 (2-3) (November 12): 165–173. doi:10.1016/j.pbiomolbio.2014.06.003.
- Davies, D R, E A Padlan, and S Sheriff. 1990. "Antibody-Antigen Complexes.." *Annual Review of Biochemistry* 59: 439–473. doi:10.1146/annurev.bi.59.070190.002255.
- de Vos, A M, M Ultsch, and A A Kossiakoff. 1992. "Human Growth Hormone and Extracellular Domain of Its Receptor: Crystal Structure of the Complex.." *Science (New York, N.Y.)* 255 (5042) (January 17): 306–312.
- Dehouck, Y, J M Kwasigroch, M Rooman, and D Gilis. 2013. "BeAtMuSiC: Prediction of Changes in Protein-Protein Binding Affinity on Mutations." *Nucleic Acids Research* 41 (W1) (June 22): W333–W339. doi:10.1093/nar/gkt450.
- Fan, Qing R, and Wayne A Hendrickson. 2005. "Structure of Human Follicle-Stimulating Hormone in Complex with Its Receptor.." *Nature* 433 (7023) (January 20): 269–277. doi:10.1038/nature03206.
- Feng, Zukang, Li Chen, Himabindu Maddula, Ozgur Akcan, Rose Oughtred, Helen M Berman, and John Westbrook. 2004. "Ligand Depot: a Data Warehouse for Ligands Bound to Macromolecules.." *Bioinformatics (Oxford, England)* 20 (13) (September 1): 2153–2155. doi:10.1093/bioinformatics/bth214.
- Fischer, T B, K V Arunachalam, D Bailey, V Mangual, S Bakhru, R Russo, D Huang, et al. 2003. "The Binding Interface Database (BID): a Compilation of Amino Acid Hot Spots in Protein Interfaces.." *Bioinformatics (Oxford, England)* 19 (11) (July 22): 1453–1454.
- Fox, K M, J A Dias, and P Van Roey. 2001. "Three-Dimensional Structure of Human Follicle-Stimulating Hormone.." *Molecular Endocrinology (Baltimore, Md.)* 15 (3) (March): 378–389. doi:10.1210/mend.15.3.0603.
- Genheden, Samuel, and Ulf Ryde. 2015. "The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities." *Expert Opinion on Drug Discovery* 10 (5) (May): 449–461. doi:10.1517/17460441.2015.1032936.
- Gétaz, David, Alessandro Butté, and Massimo Morbidelli. 2013. "Model-Based Design Space Determination of Peptide Chromatographic Purification Processes." *Journal of Chromatography A* 1284 (April 5): 80–87. doi:10.1016/j.chroma.2013.01.117.
- Gétaz, David, Guido Stroehlein, Alessandro Butté, and Massimo Morbidelli. 2013. "Model-Based Design of Peptide Chromatographic Purification Processes." *Journal of Chromatography A* 1284 (April 5): 69–79. doi:10.1016/j.chroma.2013.01.118.
- Gromoll, J, E Pekel, and E Nieschlag. 1996. "The Structure and Organization of the Human Follicle-Stimulating Hormone Receptor (FSHR) Gene.." *Genomics* 35 (2) (July 15): 308–

311. doi:10.1006/geno.1996.0361.
- Guharoy, Mainak, and Pinak Chakrabarti. 2005. "Conservation and Relative Importance of Residues Across Protein-Protein Interfaces.." *Proceedings of the National Academy of Sciences of the United States of America* 102 (43) (October 25): 15447–15452. doi:10.1073/pnas.0505425102.
- Guney, Emre, Nurcan Tuncbag, Ozlem Keskin, and Attila Gursoy. 2008. "HotSprint: Database of Computational Hot Spots in Protein Interfaces.." *Nucleic Acids Research* 36 (Database issue) (January): D662–6. doi:10.1093/nar/gkm813.
- Heckert, Leslie L, and Michael D Griswold. 2002. "The Expression of the Follicle-Stimulating Hormone Receptor in Spermatogenesis.." *Recent Progress in Hormone Research* 57: 129–148.
- Heurich, Meike, Zeynep Altintas, and Ibtisam Tothill. 2013. "Computational Design of Peptide Ligands for Ochratoxin A." *Toxins* 5 (6) (June): 1202–1218. doi:10.3390/toxins5061202.
- Humphrey, William, Andrew Dalke, and Klaus Schulten. 1996. "VMD -- Visual Molecular Dynamics." *Journal of Molecular Graphics* 14: 33–38.
- Hunter, J. D. 2007. "Matplotlib: a 2D Graphics Environment." *Computing in Science & Engineering* 9 (3): 90–95.
- Hunzicker-Dunn, Mary, and Evelyn T Maizels. 2006. "FSH Signaling Pathways in Immature Granulosa Cells That Regulate Target Gene Expression: Branching Out From Protein Kinase a.." *Cellular Signalling* 18 (9) (September): 1351–1359. doi:10.1016/j.cellsig.2006.02.011.
- Janin, J, and C Chothia. 1990. "The Structure of Protein-Protein Recognition Sites.." *The Journal of Biological Chemistry* 265 (27) (September 25): 16027–16030.
- Jiang, X, D Fischer, X Chen, S D McKenna, H Liu, V Sriraman, H N Yu, A Goutopoulos, S Arkinstall, and X He. 2014. "Evidence for Follicle-Stimulating Hormone Receptor as a Functional Trimer." *Journal of Biological Chemistry* 289 (20) (May 16): 14273–14282. doi:10.1074/jbc.M114.549592.
- Jiang, Xuliang, Heli Liu, Xiaoyan Chen, Po-Han Chen, David Fischer, Venkataraman Sriraman, Henry N Yu, Steve Arkinstall, and Xiaolin He. 2012. "Structure of Follicle-Stimulating Hormone in Complex with the Entire Ectodomain of Its Receptor.." *Proceedings of the National Academy of Sciences of the United States of America* 109 (31) (July 31): 12491–12496. doi:10.1073/pnas.1206643109.
- Kajava, A V. 1998. "Structural Diversity of Leucine-Rich Repeat Proteins.." *Journal of Molecular Biology* 277 (3) (April 3): 519–527. doi:10.1006/jmbi.1998.1643.
- Katoh, Kazutaka, and Daron M Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.." *Molecular Biology and Evolution* 30 (4) (April): 772–780. doi:10.1093/molbev/mst010.
- Kaufman, Deborah B, Marc E Hentsch, George A Baumbach, Joseph A Buettner, Christopher A Dadd, Ping Y Huang, David J Hammond, and Ruben G Carbonell. 2002. "Affinity Purification of Fibrinogen Using a Ligand From a Peptide Library." *Biotechnology and Bioengineering* 77 (3) (February 5): 278–289. doi:10.1002/bit.10120.
- Kelley, Brian D, James Booth, Molly Tannatt, Qi-Long Wu, Robert Ladner, Jinan Yu, Daniel Potter, and Arthur Ley. 2004. "Isolation of a Peptide Ligand for Affinity Purification of Factor VIII Using Phage Display." *Journal of Chromatography A* 1038 (1-2) (June): 121–130. doi:10.1016/j.chroma.2004.03.041.
- Keskin, Ozlem, Buyong Ma, and Ruth Nussinov. 2005. "Hot Regions in Protein--Protein Interactions: the Organization and Contribution of Structurally Conserved Hot Spot Residues.." *Journal of Molecular Biology* 345 (5) (February 4): 1281–1294. doi:10.1016/j.jmb.2004.10.077.
- Klauda, Jeffery B, Richard M Venable, J Alfredo Freites, Joseph W O'Connor, Douglas J Tobias, Carlos Mondragon-Ramirez, Igor Vorobyov, Alexander D MacKerell, and Richard W Pastor. 2010. "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types.." *The Journal of Physical Chemistry B* 114 (23) (June 17): 7830–7843. doi:10.1021/jp101759q.
- Kollman, P A, I Massova, C Reyes, B Kuhn, S Huo, L Chong, M Lee, et al. 2000. "Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models.." *Accounts of Chemical Research* 33 (12) (December): 889–897.
- Krissinel, Evgeny. 2010. "Crystal Contacts as Nature's Docking Solutions." *Journal of*

- Computational Chemistry* 31 (1) (January 15): 133–143. doi:10.1002/jcc.21303.
- Krissinel, Evgeny, and Kim Henrick. 2007. "Inference of Macromolecular Assemblies From Crystalline State.." *Journal of Molecular Biology* 372 (3) (September 21): 774–797. doi:10.1016/j.jmb.2007.05.022.
- Kumari, Rashmi, Rajendra Kumar, Open Source Drug Discovery Consortium, and Andrew Lynn. 2014. "G\_Mmpbsa—a GROMACS Tool for High-Throughput MM-PBSA Calculations." *Journal of Chemical Information and Modeling* 54 (7) (July 28): 1951–1962. doi:10.1021/ci500020m.
- La Merie Publishing. 2014. "2013 Sales of Recombinant Therapeutic Antibodies & Proteins" (March 1): 1–37.
- Lapthorn, A J, D C Harris, A Littlejohn, J W Lustbader, R E Canfield, K J Machin, F J Morgan, and N W Isaacs. 1994. "Crystal Structure of Human Chorionic Gonadotropin.." *Nature* 369 (6480) (June 9): 455–461. doi:10.1038/369455a0.
- Li, Yanying, Xiaodan Liu, Xiaoyan Dong, Lin Zhang, and Yan Sun. 2014. "Biomimetic Design of Affinity Peptide Ligand for Capsomere of Virus-Like Particle." *Langmuir* 30 (28) (July 22): 8500–8508. doi:10.1021/la5017438.
- Liu, Zhuo, Patrick V Gurgel, and Ruben G Carbonell. 2012. "Purification of Human Immunoglobulins a, G and M From Cohn Fraction II/III by Small Peptide Affinity Chromatography." *Journal of Chromatography A* 1262 (November 2): 169–179. doi:10.1016/j.chroma.2012.09.026.
- Makriyannis, T, and Y D Clonis. 1997. "Design and Study of Peptide-Ligand Affinity Chromatography Adsorbents: Application to the Case of Trypsin Purification From Bovine Pancreas.." *Biotechnology and Bioengineering* 53 (1) (January 5): 49–57. doi:10.1002/(SICI)1097-0290(19970105)53:1<49::AID-BIT8>3.0.CO;2-Z.
- Michaud-Agrawal, Naveen, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. 2011. "MDAnalysis: a Toolkit for the Analysis of Molecular Dynamics Simulations.." *Journal of Computational Chemistry* 32 (10) (July 30): 2319–2327. doi:10.1002/jcc.21787.
- Miller, Bill R, III, T Dwight McGee Jr., Jason M Swails, Nadine Homeyer, Holger Gohlke, and Adrian E Roitberg. 2012. "MMPBSA.Py: an Efficient Program for End-State Free Energy Calculations." *Journal of Chemical Theory and Computation* 8 (9) (September 11): 3314–3321. doi:10.1021/ct300418h.
- Moreira, Irina S, Pedro A Fernandes, and Maria J Ramos. 2007. "Hot Spots--a Review of the Protein-Protein Interface Determinant Amino-Acid Residues.." *Proteins: Structure, Function, and Bioinformatics* 68 (4) (September 1): 803–812. doi:10.1002/prot.21396.
- Morrow, John Kenneth, and Shuxing Zhang. 2012. "Computational Prediction of Protein Hot Spot Residues.." *Current Pharmaceutical Design* 18 (9): 1255–1265.
- Pavelka, Antonin, Eva Chovancova, and Jiri Damborsky. 2009. "HotSpot Wizard: a Web Server for Identification of Hot Spots in Protein Engineering.." *Nucleic Acids Research* 37 (Web Server issue) (July): W376–83. doi:10.1093/nar/gkp410.
- Perez, Fernando, and Brian E. Granger. 2007. "iPython: a System for Interactive Scientific Computing." *Computing in Science & Engineering* {9} (3) (MAY-JUN): {21–29}. doi:10.1109/MCSE.2007.53.
- Pérez, Alfredo, and Enrique Donoso. 2011. "Gametogénesis." In *Obtetricia*. Mediterraneo.
- Phillips, James C, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kalé, and Klaus Schulten. 2005. "Scalable Molecular Dynamics with NAMD." *Journal of Computational Chemistry* 26 (16): 1781–1802. doi:10.1002/jcc.20289.
- Rapp, Chaya, Hadassa Klerman, Emily Levine, and Christopher L McClendonChandra Verma to. 2013. "Hydrogen Bond Strengths in Phosphorylated and Sulfated Amino Acid Residues." Chandra VermaEd. *PLoS ONE* 8 (3) (March 5): e57804. doi:10.1371/journal.pone.0057804.s002.
- Sali, A, and T L Blundell. 1993. "Comparative Protein Modelling by Satisfaction of Spatial Restraints.." *Journal of Molecular Biology* 234 (3) (December 5): 779–815. doi:10.1006/jmbi.1993.1626.
- Schrödinger, LLC. 2010. "The PyMOL Molecular Graphics System, Version~1.3r1."
- Shen, Min-Yi, and Andrej Sali. 2006. "Statistical Potential for Assessment and Prediction of Protein Structures." *Protein Science* 15 (11) (November): 2507–2524. doi:10.1110/ps.062416606.

- Simoni, M, J Gromoll, and E Nieschlag. 1997. "The Follicle-Stimulating Hormone Receptor: Biochemistry, Molecular Biology, Physiology, and Pathophysiology.." *Endocrine Reviews* 18 (6) (December): 739–773. doi:10.1210/edrv.18.6.0320.
- Szkudlinski, M W, N R Thotakura, J E Tropea, M Grossmann, and B D Weintraub. 1995. "Asparagine-Linked Oligosaccharide Structures Determine Clearance and Organ Distribution of Pituitary and Recombinant Thyrotropin.." *Endocrinology* 136 (8) (August): 3325–3330. doi:10.1210/endo.136.8.7628367.
- Thorn, K S, and A A Bogan. 2001. "ASEdb: a Database of Alanine Mutations and Their Effects on the Free Energy of Binding in Protein Interactions.." *Bioinformatics (Oxford, England)* 17 (3) (March): 284–285.
- Ulloa-Aguirre, A, A R Midgley, I Z Beitins, and V Padmanabhan. 1995. "Follicle-Stimulating Isohormones: Characterization and Physiological Relevance.." *Endocrine Reviews* 16 (6) (December): 765–787. doi:10.1210/edrv-16-6-765.
- Ulloa-Aguirre, A, and C Timossi. 1998. "Structure-Function Relationship of Follicle-Stimulating Hormone and Its Receptor.." *Human Reproduction Update* 4 (3) (April): 260–283.
- Ulloa-Aguirre, A, C Timossi, P Damián-Matsumura, and J A Dias. 1999. "Role of Glycosylation in Function of Follicle-Stimulating Hormone.." *Endocrine* 11 (3) (December): 205–215. doi:10.1385/ENDO:11:3:205.
- Webb, Benjamin, and Andrej Sali. 2014. "Protein Structure Modeling with MODELLER.." *Methods in Molecular Biology (Clifton, N.J.)* 1137: 1–15. doi:10.1007/978-1-4939-0366-5\_1.
- Wells, J A. 1991. "Systematic Mutational Analyses of Protein-Protein Interfaces.." *Methods in Enzymology* 202: 390–411.
- Zeleznik, A J, A R Midgley, and L E Reichert. 1974. "Granulosa Cell Maturation in the Rat: Increased Binding of Human Chorionic Gonadotropin Following Treatment with Follicle-Stimulating Hormone in Vivo.." *Endocrinology* 95 (3) (September): 818–825. doi:10.1210/endo-95-3-818.