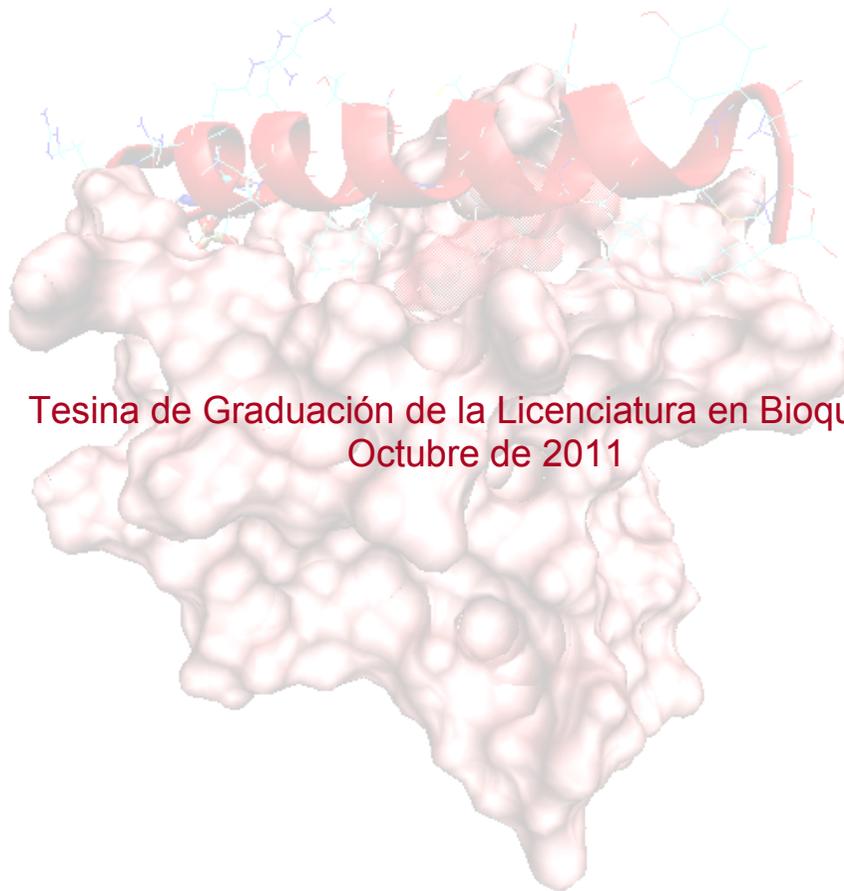


**MODELIZACIÓN ESTRUCTURAL DE LA  
INTERACCIÓN ENTRE PROTEÍNA QUINASA  
DEPENDIENTE DE AMP<sub>c</sub> Y LA PROTEÍNA  
CORE DEL VIRUS DE HEPATITIS C**



Tesina de Graduación de la Licenciatura en Bioquímica  
Octubre de 2011

Astrid Febe Brandner Giménez

Orientador: Dr. Sergio Pantano

Trabajo realizado en el Grupo de Simulaciones Biomoleculares  
Institut Pasteur de Montevideo



## TABLA DE CONTENIDO

RESUMEN.....	4
ABREVIACIONES.....	5
<b>1. INTRODUCCIÓN.....</b>	<b>6-11</b>
<b>1.1 PKA.....</b>	<b>6-7</b>
<b>1.2 AKAPs.....</b>	<b>7-9</b>
<b>1.3 VHC.....</b>	<b>9-11</b>
<b>1.4 Rol de PKA y VHCcore en infecciones de VHC.....</b>	<b>11</b>
<b>2. OBJETIVOS.....</b>	<b>12</b>
<b>3. FUNDAMENTACIÓN METODOLÓGICA.....</b>	<b>13-25</b>
<b>3.1 Simulaciones de dinámica molecular.....</b>	<b>13-23</b>
3.1.1 ¿Por qué usar simulaciones?.....	13
3.1.2 Algunas generalidades.....	13-14
3.1.3 Teoría de las simulaciones de dinámica molecular.....	15-16
3.1.4 Campo de fuerza.....	16-18
3.1.5 Algoritmos utilizados en dinámica molecular.....	19-21
3.1.6 Control térmico y de presión durante las simulaciones.....	21
3.1.7 Algoritmos para minimización energética.....	21-23
<b>3.2 Docking.....</b>	<b>24-25</b>
<b>4. MATERIALES Y MÉTODOS.....</b>	<b>26-28</b>
<b>4.1 Modelos estructurales utilizados.....</b>	<b>26</b>
<b>4.2 Secuencias con numeraciones utilizadas.....</b>	<b>26</b>
<b>4.3 Campo de fuerza y parámetros específicos.....</b>	<b>27</b>
<b>4.4 Potenciales electrostáticos.....</b>	<b>27</b>
<b>4.5 Docking.....</b>	<b>27-28</b>
<b>4.6 Protocolo de simulación.....</b>	<b>28</b>

<b>5. RESULTADOS</b> .....	29-45
<b>5.1 Alineamiento múltiple y estructural</b> .....	29-30
<b>5.2 Potenciales electrostáticos</b> .....	30-31
<b>5.3 Docking</b> .....	32-34
<b>5.4 Dinámica molecular</b> .....	35-45
5.4.1 <i>Análisis general</i> .....	35-38
5.4.1a <i>RMSD</i> .....	35-36
5.4.1b <i>Análisis de cluster de la trayectoria</i> .....	36-38
5.4.2 <i>Análisis estructural</i> .....	38-45
5.4.2a <i>Superficie de interfase</i> .....	38-39
5.4.2b <i>Distancias entre residuos específicos</i> .....	40-42
5.4.2c <i>Interacciones de S116 fosforilada</i> .....	43
5.4.2d <i>Mapa de contactos</i> .....	44-45
<b>6. DISCUSIÓN</b> .....	46-48
<b>7. CONCLUSIONES Y PERSPECTIVAS</b> .....	49
<b>8. REFERENCIAS BIBLIOGRÁFICAS</b> .....	50-52

## RESUMEN

Los mecanismos de señalización celular, tales como la fosforilación, glicosidación, metilación, etc regulan interacciones entre los diversos componentes intracelulares. La proteína quinasa dependiente de AMP cíclico (PKA) es una importante Ser/Thr quinasa que en humanos está representada por cuatro isoformas (tipos I y II, subtipos  $\alpha$  o  $\beta$ ). Está constituida por dos subunidades catalíticas (C), cuya actividad es controlada por sendas subunidades regulatorias (R). Estas subunidades R poseen un dominio de anclaje/dimerización en la región N-terminal seguido por una región flexible que se liga e inhibe a la subunidad C. En el extremo C-terminal de la proteína se encuentran dos dominios de unión a AMPc organizados en tándem. El enlace de AMPc a estos sitios de unión produce un cambio conformacional que resulta en la activación de las subunidades C. Además de servir para la homodimerización, los dominios N-terminal contienen un sitio de reconocimiento específico para proteínas de anclaje molecular, lo que les confiere la posibilidad de ser dirigidas a diferentes compartimientos celulares. Estas proteínas de anclaje pertenecen a la familia de proteínas denominada AKAPs (proteínas de unión a PKA, del inglés: *A Kinase Anchoring Proteins*), que al unirse a PKA posibilitan restringir la fosforilación específica de proteínas a una región celular específica determinada por la proteína de anclaje. Hasta la fecha se han descubierto numerosas AKAPs, usualmente con alta afinidad de unión a RII ( $\alpha$  o  $\beta$ ) o a RI ( $\alpha$  o  $\beta$ ), aunque también se han descrito algunas AKAPs con afinidad por ambas isoformas.

En este estudio se analizó la posibilidad de que la proteína core del virus de la Hepatitis C (VHCcore, específicamente la región comprendida 113-137) pueda actuar como AKAP, uniéndose al dominio de anclaje/dimerización de la isoforma RII $\alpha$  de PKA. Esta hipótesis se basó en las características estructurales de dicha región de VHCcore, que comparte el perfil de otras AKAPs caracterizadas experimentalmente, así como la colocalización de PKA y VHCcore en regiones celulares determinadas. Además, se trabajó con las formas fosforilada y sin fosforilar de la proteína core (VHCcore), con el fin de estudiar el efecto de la fosforilación de Ser116 por PKA en la interacción con PKA y eventualmente plantear un posible mecanismo de regulación mediada por PKA de dicha interacción.

Para esto se utilizaron una serie de métodos teóricos, incluyendo análisis de secuencias, alineamientos estructurales, docking y simulaciones de dinámica molecular. Los resultados de estos estudios sugieren la posibilidad de formación de un complejo de tipo AKAP entre PKA y VHCcore donde la fosforilación parece favorecer la interacción.

## ABREVIACIONES

**AMP:** Adenosín monofosfato

**AMPc:** Adenosín monofosfato cíclico

**AKAP:** Proteína de anclaje a PKA (*A kinase anchoring protein*)

**ATP:** Adenosín trifosfato

**C-PKA:** Subunidad catalítica de PKA

**fVHCcore:** VHCcore fosforilada en Ser116

**PBE:** Ecuación de Poisson-Boltzmann

**PKA:** Proteína quinasa dependiente de AMP cíclico

**PKA-DD:** RII $\alpha$ DD

**PME:** Particle Mesh Ewald

**RE:** Retículo endoplásmico

**RII $\alpha$ DD:** Dominio de docking/dimerización de la subunidad regulatoria II $\alpha$  de PKA

**RII $\alpha$ DD/VHCcore:** Complejo de subunidad regulatoria II $\alpha$  de PKA y VHCcore (aminoácidos 113-137)

**RII $\alpha$ DD/fVHCcore:** RII $\alpha$ DD/VHCcore fosforilado en Ser 116

**R-PKA:** subunidad regulatoria de PKA

**SEP/Sep:** Residuo de serina fosforilada

**VHC:** virus de la hepatitis C

**VHCcore:** Proteína de la cápside del virus de Hepatitis C (*core*) [En la mayor parte del trabajo se refiere a la secuencia aminoacídica modelada 113-137, excepto cuando se especifique que alude a toda la proteína core]

## I. INTRODUCCIÓN

En biología los procesos de señalización cumplen un rol crucial en el mantenimiento del correcto funcionamiento celular. Estos procesos participan en la regulación de la expresión génica, diferenciación celular, respuesta a estímulos, activación e inhibición de enzimas, entre otros.

Existen diversas formas de transmitir señales entre los diferentes protagonistas de las vías de señalización. Comúnmente, se tratan de procesos químicos donde se incorpora un grupo activo a determinado compuesto biológico (generalmente a proteínas, aunque también puede tratarse de ácidos nucleicos) como lo son la fosforilación, glicosilación, miristilación, ubiquitinación o metilación. Este tipo de procesos pueden servir para generar cambios conformacionales o para ser utilizados como señales de reconocimiento para interacciones intermoleculares. El caso de la fosforilación es uno de los más comunes, en cuyo caso el ATP es utilizado para participar directamente de la transferencia de grupos fosfato. Un ejemplo clásico de este tipo de señalización que se da en procesos del metabolismo es la degradación del glucógeno, donde se genera una cascada de modificaciones postrasduccionales, debido a la fosforilación consecutiva de compuestos por medio de proteínas fosforiladas en los pasos previos.

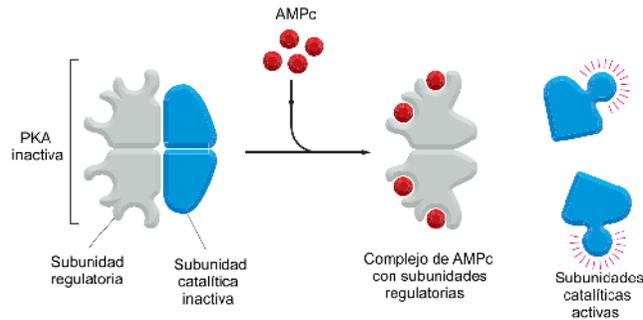
A continuación se presentan los principales protagonistas del sistema biológico estudiado en el presente trabajo.

### 1.1: Proteína quinasa A (PKA)

La proteína quinasa dependiente de AMPc (PKA) es una enzima que cumple un rol importante en numerosas funciones celulares como respuestas inmunitarias, aprendizaje, memoria, metabolismo y secreción de insulina, frecuencia e intensidad del latido cardiaco, etc. [3-11]. PKA es una quinasa que fosforila específicamente serinas o treoninas incluidas en la secuencia consenso RRXS/TY donde X es cualquier residuo pequeño (p.ej: glicina o alanina) e Y es un residuo hidrofóbico grande (p.ej: fenilalanina).

Esta enzima es activada por moléculas de AMP cíclico (AMPc), que es un segundo mensajero producido por la adenilil ciclasa en respuesta a estímulos recibidos de receptores asociados a proteínas G. AMPc tiene como principal efecto la regulación de PKA. En su forma inactiva, PKA existe como un holotetrámero formado por un dímero de subunidades regulatorias (R-PKA) que se ligan e inhiben a dos subunidades catalíticas (C-PKA). Las R-PKA están compuestas por un dominio N-terminal de docking y dimerización (DD), un péptido flexible que funciona como pseudo sustrato de las C-PKA y dos dominios consecutivos que ligan AMPc situados en el extremo carboxilo terminal.

En el ser humano existen cuatro isoformas de la PKA que se diferencian de acuerdo al tipo de subunidad regulatoria que se encuentra en la holoenzima. Éstas pueden ser de tipo I $\alpha$ , I $\beta$ , II $\alpha$  o II $\beta$ . Cuando el AMPc se liga a las R-PKA, estas sufren una transición conformacional que determina la separación de las C-PKA y su consecuente activación.(Figura 1)

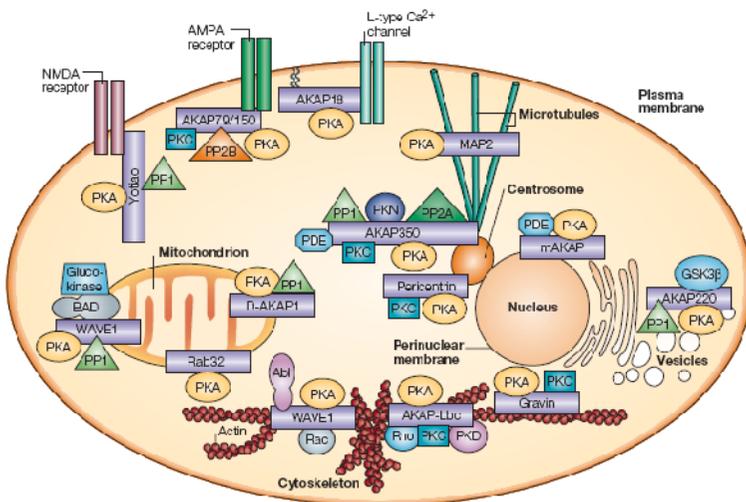


**Figura 1:** Esquema de mecanismo de acción de PKA. [12]

La especificidad de la reacción de fosforilación que se lleva a cabo luego de la activación de la PKA está directamente vinculada a la localización celular de la enzima. Por este motivo es importante que exista una correcta ubicación de PKA en el entorno celular adecuado. Esto se logra a través del anclaje de proteínas de unión a PKA denominadas AKAPs (por su sigla en inglés *A Kinase Anchoring Proteins*) a zonas cercanas al sustrato que fosforilarán.

## 1.2: AKAPs

Estas proteínas funcionan como anclas para las PKA (a través del dominio R-PKA), localizándolas en la vecindad de sus blancos moleculares específicos[2, 13].

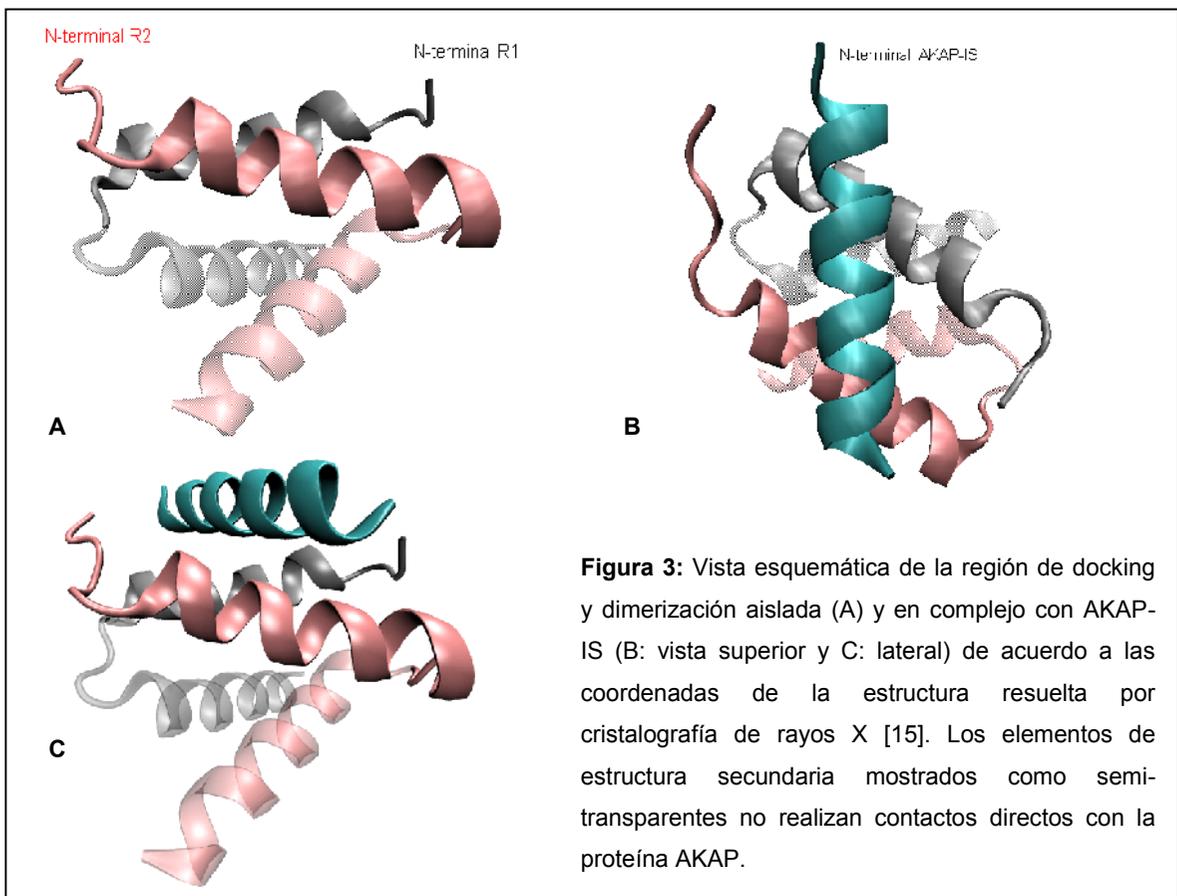


**Figura 2:** Esquema de algunas interacciones AKAP-PKA experimentalmente identificadas. Extraído de [2]

Dentro de las más de 40 proteínas AKAP identificadas se encuentran proteínas que intervienen en numerosos procesos biológicos interactuando con diversos componentes celulares tales como receptores asociados a proteínas G, Adenilil ciclasa, Epacs y sustratos de PKA lo que permite

regular así la vía de señalización del AMPc en numerosos niveles. Dentro de esta gran variedad de destinos de localización de PKA vía AKAPs se han reportado complejos PKA/AKAP ubicados en el núcleo (AKAP95), citoesqueleto (AKAP78, MAP2) en el centrosoma (AKAP350), canales iónicos (AKAP15), peroxisomas (AKAP220), aparato de Golgi (AKAP85) y mitocondria (AKAP84, AKAP149) así como también en membranas (AKAP79/75) y retículo endoplasmático (AKAP100) [14].

La interacción AKAP/R-PKA se realiza a través del reconocimiento específico por parte de una hélice anfipática de 14 a 18 residuos que se inserta en una cavidad hidrofóbica del dominio de docking y dimerización (DD). Experimentalmente, a través de estudios cristalográficos y de resonancia magnética nuclear se ha demostrado que estos dominios DD están constituidos por un homodímero de subunidades regulatorias de PKA que se asocia formando un motivo con cuatro hélices alfa antiparalelas superpuestas en forma de X tal como se muestra debajo, en la figura 3. [15, 16]



Existe además una gran ambigüedad en la definición del motivo AKAP de reconocimiento específico. Del análisis de un alineamiento múltiple de secuencias de todas las proteínas

identificadas como AKAPs se podría definir una secuencia consenso aproximada como **HX(XHHX)<sub>2</sub>XH**, donde **X** representa un aminoácido arbitrario y **H** representa residuos hidrofóbicos que no están absolutamente conservados (*ver columnas resaltadas de figura 7 de la sección 5.1*). Las pequeñas dimensiones del motivo de unión, el número reducido de las secuencias identificadas como AKAPs así como de estructuras resueltas experimentalmente dificultan la predicción e identificación de nuevas proteínas que podrían funcionar como AKAPs. En este campo, la bioinformática estructural ha mostrado ser una herramienta confiable para el diseño de péptidos competidores capaces de ligar selectivamente una de las isoformas de PKA [17-19]. En investigaciones realizadas en el grupo de Simulaciones Biomoleculares del Institut Pasteur de Montevideo se han encontrado resultados análogos para la isoforma II usando técnicas de bioinformática estructural [20], agregando un nuevo nivel de complejidad al reconocimiento proteína-proteína. Por lo tanto, comprender y modelar las diferencias que determinan la interacción específica péptido-proteína constituye el primer paso hacia el diseño racional de pequeñas moléculas, con objetivos terapéuticos, capaces de interrumpir o competir por estas interacciones.

### 1.3: VHC

El virus de la hepatitis C, de la familia *Flaviviridae*, pertenece al género *Hepacivirus*. Su material genético se encuentra en forma de ARN simple hebra positivo, con regiones 5' y 3' UTR (del inglés *Untranslated Region*) que no se traducen. En la región 5' UTR, que corresponde a una secuencia de aproximadamente 340 nucleótidos, se encuentra una región de ingreso interno al ribosoma (IRES) y algunas pequeñas regiones de marcos de lectura abiertos que no son traducidas. El marco de lectura abierto (también conocido como ORF del inglés *Open Reading Frame*) posee aproximadamente 9400 nucleótidos y codifica para una única poliproteína que luego es clivada dando lugar a las diferentes proteínas virales.[21] Las proteínas codificadas por el ARN en el extremo 5' son proteínas estructurales (Core, E1, E2). Luego en la región siguiente del genoma se expresa una pequeña proteína de función desconocida llamada p7, y finalmente, hacia el extremo 3'. las proteínas no estructurales (NS2, NS3, NS4A, NS4B, NS5A y NS5B). Esta organización del genoma se muestra en la Figura 4.

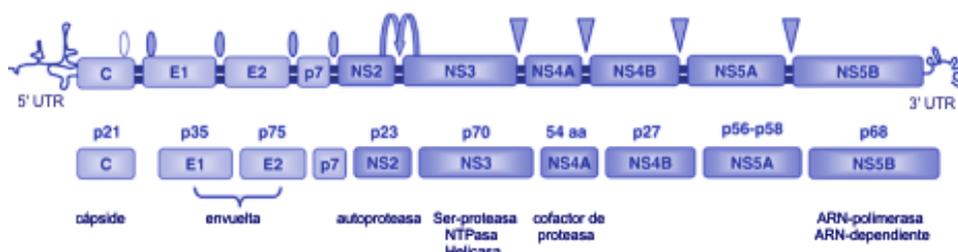


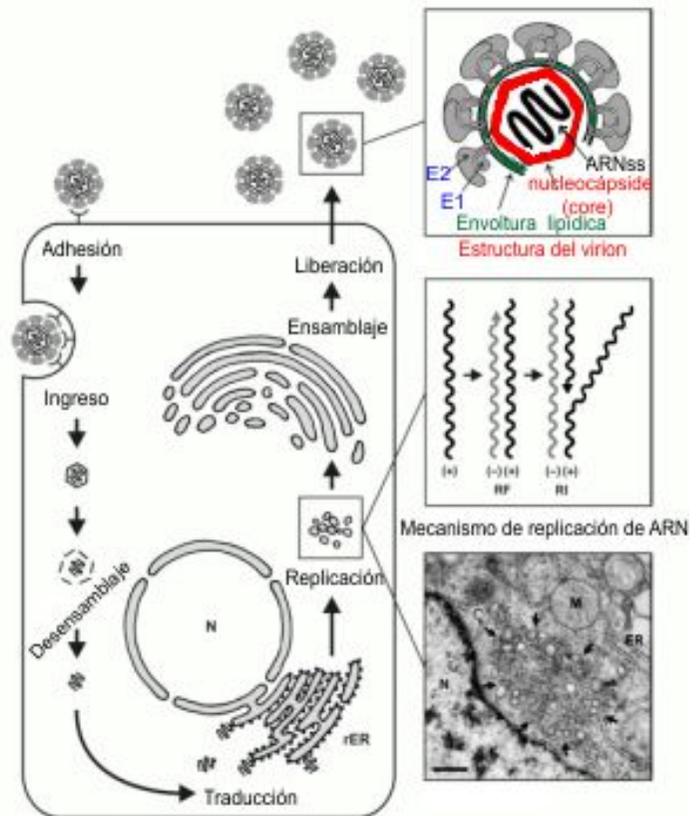
Figura 4: Esquema de proteínas codificadas en el ARN de VHC [22]

El ciclo viral del VHC puede descomponerse en ocho pasos básicos. Primeramente, la adhesión a la membrana celular de la célula que infectará se da mediante proteínas presentes en la envoltura lipídica que se unen a receptores existentes en la membrana celular. En la amplia mayoría se tratan de hepatocitos y la adhesión se da vía interacciones con el receptor SR-BI (*Scavenger receptor class B, type I*) y el complejo CD81. Luego se procede a la internalización vía endocitosis de membrana, entrando

de esta forma la nucleocápside (core con ARN asociado adentro) hacia el interior celular. Posteriormente sucede la liberación del ARN de la envoltura proteica, ya sea durante la penetración por la membrana plasmática o por enzimas que disuelven la cubierta. Una vez liberado el material genético comienza la traducción del ARN viral para así poder producir proteínas virales como la ARN transcriptasa necesaria para la replicación viral. Cuando existe suficiente ARN transcriptasa se produce la síntesis de ARN *antisense* (que servirá de molde para sintetizar las copias de ARN positivo que serán el material genético de la siguiente generación de virus), el que se copiará luego para obtener material genético para los

nuevos virus. El ARN viral dirige la producción de los capsómeros (partículas virales primitivas) así como también los ribosomas participan de la traducción del ARN viral para la síntesis de proteínas. Estos capsómeros se ensamblan con el ARN nuevo formando la cápside. Finalmente las partículas virales viajan por el lado interno de la membrana celular adhiriéndose a ella. La membrana plasmática rodea a la partícula viral para luego liberarla al exterior celular con la correspondiente envoltura lipídica protectora. [23]

Dentro de la numerosa cantidad de procesos en los que interviene el VHC uno de los más importantes a la hora de entender el contexto biológico de la interacción estudiada en el presente trabajo es el de la señalización y junto con ella la modulación de la respuesta inmune en el huésped.



**Figura 5:** Ciclo esquemático del VHC

(Extraído y modificado de [1] )

Las interacciones existentes a lo largo del ciclo de vida de este virus no se conocen en su totalidad todavía. Esto se manifiesta como uno de los principales inconvenientes a la hora de buscar métodos para combatir eficazmente la infección por VHC. Por otro lado, la alta tasa de variabilidad genética que posee por causa de los errores generados por la ARN transcriptasa así como también por los posibles eventos de recombinación entre especies también representa un obstáculo. Por esto es de suma importancia que los posibles blancos de tratamiento sean estructuras con una alta conservación. La proteína core, si bien varía entre los diferentes subtipos de VHC, puede ser un buen candidato para tratamiento ya que tiene un bajo grado de variabilidad dentro de cada genotipo (motivo por el que se la puede utilizar para diferenciar genotipos y no subtipos).

La relevancia biológica de entender mejor los pasos de la infección del virus de la hepatitis C radica en que en presencia de una infección con VHC existen altas chances de generar carcinoma hepatocelular. De hecho, la infección con VHC es uno de los principales causantes de cáncer de hígado en el mundo. Por otro lado, un alto porcentaje de las personas infectadas con VHC acaban teniendo Hepatitis C crónica. En particular, caracterizar las interacciones entre las proteínas de este virus que cumplen un rol importante en el ciclo de vida, ya que el material genético del virus no se inserta en el genoma del portador.

#### **1.4: Rol de PKA y VHCcore en infecciones con VHC**

El ciclo de replicación del virus de la hepatitis C involucra diferentes moléculas, tanto codificadas por el propio virus como las que recluta de la célula infectada para cumplir con las funciones de su ciclo y así poder replicarse. Existen evidencias que relacionan a PKA con un rol dentro del ciclo viral, sugiriendo la activación de PKA en presencia de infección facilitando la entrada del virus a la célula infectada [24]. También se observó experimentalmente que en cultivos celulares infectados por VHC existe fosforilación por PKA de la proteína core en Ser53 y Ser116, siendo la última el blanco mayoritario de esta modificación[25].

La fosforilación no afectaría la formación de partículas NLP (del inglés: *Nucleocapsid-like particles*) pero sí la liberación del anclaje de VHCcore a la membrana reticuloendoplasmática.

Por otro lado VHCcore posee un perfil compatible en su secuencia y estructura con las propiedades compartidas entre todas las AKAPs; ésto es, la presencia de una región de estructura secundaria de hélice alfa anfipática comprendida en la región 113-137 de la proteína core.

A su vez, cuando VHCcore está presente en su forma aislada en un entorno celular, tiene la capacidad de dirigirse a diferentes organelos (ER, mitocondria y en ocasiones también se lo ha ubicado en el núcleo[26]). Se ha reportado, además, que la infección por VHC aumenta los niveles de AMPc y de sustratos fosforilados por PKA en la célula. Esto sugiere que VHCcore podría actuar como AKAP anclando la isoforma II de PKA. Uno de los objetivos de esta tesis es la identificación y caracterización de una región en VHCcore que pueda funcionar como AKAP.

## II. OBJETIVOS

En este proyecto se proponen tres objetivos fundamentales:

- 1) Explorar la posibilidad de que la proteína VHCcore pueda funcionar como AKAP para ligar la subunidad reguladora de PKA II $\alpha$ , e identificar el motivo de enlace.
- 2) Realizar un modelo estructural del dominio de docking y dimerización perteneciente a la subunidad reguladora de PKA II $\alpha$  en complejo con el péptido correspondiente a los residuos identificados en el punto anterior.
- 3) Estudiar el papel que juega la fosforilación de VHCcore y si la presencia de un residuo fosforilado puede actuar como interruptor selectivo de esta interacción.

### III. FUNDAMENTACIÓN METODOLÓGICA

#### 3.1: Simulaciones de dinámica molecular

##### 3.1.1 ¿Por qué usar simulaciones?

La primera pregunta lógica a formularse es ¿Por qué se escogió un método de modelado computacional ante el interés de estudiar un problema biológico? ¿Es un enfoque científico válido? Si bien existe una gran variedad de experimentos que comúnmente se realizan en el laboratorio, se deben dar las condiciones apropiadas para poder estudiar interacciones proteína-proteína. Además existen limitaciones en cuanto a la escala y al detalle en que se obtiene la información, así como también de los costos en tiempo y dinero asociados.

El uso de simulaciones de dinámica molecular, introducido en los años 60, es actualmente cada vez más utilizado. Hoy en día, su utilización conjunta con datos experimentales permite dar mayores detalles y profundizar la comprensión de distintos fenómenos bajo estudio, particularmente en referencia a problemas que se puedan resolver a nivel atómico/molecular. En este campo se han obtenido resultados interesantes de sistemas biológicos cuyos detalles estructurales fueron resueltos con métodos computacionales como las simulaciones de dinámica molecular. Un ejemplo ilustrativo de la utilidad de estos métodos es la descripción atómica del proceso de separación de las dos hebras del ADN mediante simulaciones el cual es difícil de estudiar experimentalmente.[27] Otros sistemas muy estudiados por métodos computacionales son aquellos que involucran proteínas de membrana con la finalidad de elucidar los mecanismos de acción de dichas proteínas a nivel mecánico en presencia de la membrana, así como también el interés en modelar proteínas que actúen como motores moleculares. Este es un campo de gran interés debido a la dificultad de obtener esta información por métodos experimentales. Existen varios ejemplos de este tipo de sistemas, donde uno de los más conocidos es el de la ATP sintetasa (también conocida como  $F_0F_1$ -ATP sintasa) en el que por ejemplo se ha podido analizar los efectos estructurales del complejo en función de la rotación de la subunidad  $\gamma$ . [28]

En los siguientes capítulos de la introducción se pretende dar una noción básica de los principales fundamentos teóricos vinculados con la aproximación computacional elegida para el estudio del sistema de interés.

##### 3.1.2 Algunas generalidades

Existen diferentes métodos teóricos para estudiar estructuras moleculares y todos están fundamentados en algún punto en la ecuación de Schrödinger que describe la vinculación

entre la función de onda de un sistema y su energía. La ecuación de Schrödinger independiente del tiempo, puede escribirse en su forma más desarrollada como:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + \hat{V}(r)\Psi = E\Psi \quad (1.1)$$

Donde  $\hbar = \frac{h}{2\pi}$  es la constante de Planck reducida,  $\hat{V}(r)$  corresponde al término de la energía

potencial de las partículas en función de su posición y  $\Psi$  a la función de onda. Asimismo la ecuación de Schrödinger independiente del tiempo (1.1) puede representarse de forma más reducida de la siguiente manera:

$$\hat{H}\Psi = E\Psi \quad (1.2)$$

Donde  $\hat{H}$  corresponde al operador Hamiltoniano ( $\hat{H} = \frac{\hbar^2}{2m}\nabla^2 + \hat{V}$ ),  $\Psi$  a la función de onda y  $E$  a la energía del sistema (potencial y cinética).

Sin embargo, puntualmente para el caso de simulaciones de dinámica molecular donde el objetivo es describir un sistema molecular en función de su evolución temporal y los sistemas de interés poseen una gran cantidad de átomos (generalmente interesa simular comportamientos de macromoléculas) es que se utiliza un enfoque clásico para modelar las interacciones existentes y consecuentemente poder integrar las ecuaciones de movimiento de Newton. Esta estrategia puede ser implementada como consecuencia de algunas aproximaciones realizadas a detallarse a continuación. La aproximación de Born-Oppenheimer es la primera de ellas. En ella se plantea la posibilidad de disociar los movimientos nucleares de los electrónicos ya que la masa electrónica es mucho menor que la nuclear y las energías asociadas a ambos son considerablemente distintas. Por este motivo, las variaciones en las posiciones nucleares ocurren a una escala temporal mucho mayor comparada con las electrónicas, teniendo entonces los electrones la capacidad de adaptarse rápidamente a estos cambios nucleares, pudiendo considerarse que la adaptación a dichos cambios es casi instantánea. Esta posibilidad de separar la descripción del movimiento electrónico del nuclear es la que permite realizar la segunda aproximación donde se modela el comportamiento de los núcleos de una manera clásica (es decir, cumplen a las leyes clásicas de la física). Lo que nos lleva a introducir la última aproximación utilizada, la base de la obtención de las energías en mecánica clásica: los campos de fuerza, que pueden ser empíricos o con información derivada de la mecánica cuántica, donde se considerarán los efectos electrónicos.

En el caso del presente trabajo se utilizará un enfoque mecánico-clásico de cálculo computacional, dado que se trata de un complejo proteico (de aproximadamente 1700 átomos) e interesa estudiar las características estructurales de interacciones existentes en un sistema macromolecular donde no existen reacciones químicas involucradas en ninguna región a modelar.

### 3.1.3 Teoría de las simulaciones de dinámica molecular

La dinámica molecular es una técnica computacional que se basa en integrar las ecuaciones clásicas de movimiento para los átomos del sistema en cada paso temporal. Es decir que se resuelve la ecuación de movimiento proveniente de la física clásica, que en el caso de un átomo tiene la siguiente expresión:

$$F_i = m_i \cdot a_i = m_i \cdot \ddot{r}_i \quad (1.3)$$

Donde  $F_i$  corresponde a la fuerza actuante sobre un átomo  $i$  como consecuencia de la interacción con el resto de los átomos,  $m_i$  a la masa de dicho átomo y  $a_i$  a la aceleración (o lo que es lo mismo, la segunda derivada de la posición con respecto al tiempo,  $\ddot{r}_i$  ).

La ecuación (1.3) es una ecuación diferencial de segundo orden, que puede reescribirse como dos ecuaciones diferenciales de primer orden, una expresando la derivada del momento y la otra la derivada de las posiciones de los átomos, es decir:

$$\dot{p}_i = \frac{\partial p_i}{\partial t} = F_i \quad (1.4)$$

$$\dot{r}_i = \frac{\partial r_i}{\partial t} = \frac{p_i}{m} \quad (1.5)$$

donde  $p_i$  es el momento de un átomo  $i$  ( $p_i = m_i \cdot v_i$ ),  $r_i$  su posición y  $F_i$  la fuerza total que actúa sobre dicho átomo. A su vez,  $F_i$  puede expresarse como la sumatoria de fuerzas de los pares de átomos no unidos más las fuerzas existentes entre átomos unidos (que pueden involucrar 2, 3 o 4 átomos) más las fuerzas externas o de restricción que puedan existir. Estas fuerzas a su vez, pueden resolverse desde el gradiente de la energía potencial causada por desplazamientos de los átomos:

$$F_i(r_1, \dots, r_N) = -\nabla_{r_i} E_{pot}(r_1, \dots, r_N) \quad (1.6)$$

Finalmente, podemos combinar (1.3) con (1.6) y obtener la siguiente expresión de la ecuación de movimiento para un sistema de  $N$  átomos:

$$-\nabla E_{pot}(r_1, \dots, r_N) = m_i \frac{d^2 r_i}{dt^2} \quad (1.7)$$

Es precisamente en este punto, donde los cálculos están basados en expresiones de la energía potencial  $E_{\text{pot}}(r)$ , que es oportuno introducir el concepto de campo de fuerza.

### 3.1.4 Campos de fuerza (*Force fields*)

En química computacional, un campo de fuerza es un conjunto de parámetros y funciones usados para hallar la energía potencial de un sistema. Estos parámetros pueden obtenerse desde cálculos de mecánica cuántica o en base a datos experimentales. Generalmente se obtienen los datos para moléculas pequeñas o grupos funcionales que luego se transfieren a moléculas mayores.

Si bien existen variaciones en la forma de modelar la energía potencial de un sistema atómico, los campos de fuerza más ampliamente utilizados actualmente han convergido en el uso de un conjunto de términos que representan los modos de interacción interatómicas, como se muestra en la tabla 3.1.

**Tabla 3.1:** Lista de interacciones básicas que se expresan en los campos de fuerza

Tipo de interacción	Expresión matemática para la energía potencial
Cambio en longitud del enlace ("bond stretching")	$\sum_{\text{enlace}} \frac{1}{2} k_r (r - r_0)^2$
Cambio en el ángulo ("angle bending")	$\sum_{\text{ángulo}} \frac{1}{2} k_\theta (\theta - \theta_0)^2$
Cambio en el ángulo del diedro (torsión)	$\sum_{\text{torsión}} \frac{U_n}{2} (1 + \cos(n\phi - \delta))$
Interacción electrostática (Coulomb)	$\sum_{\text{Coulomb}} k_{el} \left( \frac{q_i q_j}{r_{ij}} \right); \quad k_{el} = \frac{1}{4\pi\epsilon}$
Interacción de Van der Waals: potencial de Lennard-Jones	$\sum_{L-J} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right); \quad 4\epsilon\sigma^{12} = A; \quad 4\epsilon\sigma^6 = B$
Expresión de la energía potencial total:	$U = \sum_{\text{enlace}} + \sum_{\text{ángulo}} + \sum_{\text{torsion}} + \sum_{L-J} + \sum_{\text{Coulomb}}$

Estos términos son las expresiones mínimas que posee cualquier campo de fuerza y pueden diferenciarse en dos grandes grupos: las interacciones enlazantes y las no enlazantes. Dentro del primer grupo se encuentran todas aquellas interacciones que se dan como consecuencia de la unión covalente entre átomos. En dicho grupo existen expresiones para las energías aportadas

como consecuencia de la vibración de los enlaces (interacciones que involucran 2 átomos), variación del ángulo entre 3 átomos enlazados entre sí y la variación de ángulos torsionales (4 átomos consecutivos involucrados).

a.1) Potenciales de enlace 1-2: "bond stretching"

El término corresponde al potencial existente como consecuencia de la variación de la longitud del enlace (estiramiento o acortamiento) respecto a la de referencia, siendo la longitud de referencia aquella cuando todos los restantes términos del campo de fuerza son nulos, que no corresponde necesariamente a la longitud de equilibrio del enlace. En mecánica molecular se suele utilizar la expresión escrita en la tabla 3.1, donde se expresa en forma de un simple potencial armónico de Hooke siendo  $k_r$  la constante elástica del resorte,  $r$  la distancia entre los átomos que participan del enlace y  $r_0$  la distancia de equilibrio entre ambos átomos.

a.2) Potenciales de enlace 1-3: "angle bending"

Este tipo de potencial usualmente se formula mediante una expresión armónica al igual que en el caso anterior, con la salvedad de que el potencial se aplica entre dos átomos conectados entre sí mediante un átomo central ligado covalentemente a ambos. En la tabla 3.1  $k_\theta$  corresponde a la constante elástica,  $\theta$  al ángulo definido por los tres átomos y  $\theta_0$  al ángulo de equilibrio entre los 3 átomos.

a.3) Potenciales de enlace 1-4: Diedros torsionales

Describe la posición relativa de dos átomos separados por tres enlaces covalentes, mediante el desarrollo en serie de coseno como se ve en la tabla 3.1, donde  $U_n$  es la altura de la barrera,  $\varphi$  es el ángulo del dihedro definido por los cuatro átomos,  $n$  la multiplicidad y  $\delta$  la fase.

b.1) Potencial electrostático: Potencial de Coulomb

Otro término importante a contemplar son las contribuciones electrostáticas a la energía del sistema a modelar. Este tipo de interacción no enlazante se suele modelar generalmente como potenciales de Coulomb entre pares de átomos. De acuerdo a la expresión de la tabla 3.1  $k_{el}$  es la constante de Coulomb,  $q_i$  y  $q_j$  las cargas de los átomos  $i$  y  $j$  respectivamente y  $r_{ij}$  la distancia entre los átomos  $i$  y  $j$ . Un dato importante a tener en cuenta es el radio de acción de dichas interacciones, que será una de las razones principales en el costo computacional de los cálculos. Es aquí entonces que existen distintas técnicas para considerar la distancia máxima en que dichas interacciones (*cut-off*) así como métodos para evitar problemas de borde (típicamente causados por discontinuidades en la derivada en el punto de borde).[29]

### b.2) Potencial de Lennard-Jones

Las interacciones de Van der Waals son el otro tipo de interacciones no enlazantes que dan estabilidad a la estructura de las moléculas y también se calculan considerando pares de átomos. Estas interacciones incluyen interacciones de tipo dipolo-dipolo, dipolo-dipolo inducido e interacciones de dispersión de London (dipolo inducido instantáneo-dipolo inducido). Una forma de incluirlas dentro de los cálculos de mecánica molecular es como potenciales de tipo Lennard-Jones, donde es posible considerar las dos características importantes de este tipo de interacciones no enlazantes, donde son atractivas a largo alcance y repulsivas a corto alcance. Por ejemplo, expresando el término de Lennard-Jones de acuerdo a la tabla 3.1; donde  $\epsilon$  es la profundidad del pozo y  $\sigma$  el diámetro de colisión (separación interatómica tal que la energía vale cero); el término  $r^{-12}$  modela la repulsión de Pauli, mientras que el  $r^{-6}$  modela las atracciones debida a fuerzas de Van der Waals.

Dependiendo de las propiedades e implementación de los campos de fuerza que se consideren, pueden existir variaciones en las formas de expresar los potenciales así como también en los términos que se puedan agregar. Un ejemplo de esto son aquellos términos extra utilizados históricamente para modelar con mayor detalle los enlaces de hidrógeno (pudiendo agregar un término específico para ello), las interacciones metal-ligando, efectos de solvatación, etc.

Por ejemplo, el campo de fuerza de Amber ff99 (así como su variante ff99SB) es de la forma:

$$U(r) = \sum_{\text{enlace}} \frac{1}{2} k_r (r - r_0)^2 + \sum_{\text{ángulo}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{\text{torsión}} \frac{U_n}{2} (1 + \cos(n\phi - \delta)) + \sum_{\text{Coulomb}} k_{el} \left( \frac{q_i q_j}{r_{ij}} \right) + \sum_{L-J} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$$

Esta expresión es compartida por los campos de fuerza derivados de Amber ff94 (como sucede con ff99 y ff99SB). Otro aspecto interesante a considerar es que para los campos de fuerza de Amber, las cargas se obtienen a partir de una derivación sistemática basándose en cálculos cuánticos, tratándose del método de ajuste restringido de potencial electrostático o RESP (*Restrained ElectroStatic Potential fit*)[30]. En el caso de ff99 los cálculos mecánico-cuánticos se realizan a nivel Hartree-Fock/6-13G\*[31].

### 3.1.5 Algoritmos utilizados en dinámica molecular

Existen varios métodos para integrar las ecuaciones de movimiento newtonianas. En todos ellos se debe considerar un paso temporal (timestep) suficientemente pequeño para asegurar la conservación de la energía, y así poder integrar. Por este motivo los movimientos de mayor frecuencia serán los que limiten el valor de timestep a utilizar. En el caso de las moléculas biológicas éstos corresponden a los movimientos de vibraciones de los enlaces que contengan átomos de hidrógeno, debido a la pequeña masa de dichos átomos. A continuación se presentan brevemente tres de los métodos más comunes para la propagación del movimiento en dinámica molecular y la relación que tienen entre sí[32].

#### Verlet

Este algoritmo de integración es el más simple de entender y se basa en obtener las posiciones a tiempo  $t$  a tiempo  $(t-\delta t)$  y la aceleración a tiempo  $t$ . La aceleración se obtiene directamente a través de la evaluación de la fuerza de la partícula  $i$  considerada del siguiente modo:

$$a_i(t) = \ddot{r}(t) = \frac{d^2 r_i}{dt^2} = \frac{f_i(t)}{m_i} \quad (1.8)$$

Por otro lado, las posiciones se obtienen a través del desarrollo de Taylor de la función que representa la posición en función del tiempo  $t-\delta t$  y  $t+\delta t$ . truncadas en el segundo término, correspondiendo el resto  $O$  al error del método:

$$\left. \begin{array}{l} r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + O \delta t^4 \\ \text{del} \\ r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + O \delta t^4 \end{array} \right\} \begin{array}{l} \text{sumando ambas se obtiene la expresión} \\ \text{algoritmo de Verlet:} \\ r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \end{array} \quad (1.11)$$

Una vez que se obtienen las posiciones se calculan las velocidades en función de ellas, ya que no aparecen explícitamente en la ecuación anterior. Una forma para calcularla consiste simplemente en obtener la velocidad como el cociente entre la diferencia entre las posiciones a tiempo  $t+\delta t$  y  $t-\delta t$  y  $2\delta t$  es decir:

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \quad (1.9)$$

#### Velocity-verlet

Una variante del algoritmo de Verlet es el denominado "Velocity-Verlet", donde las aceleraciones, posiciones y velocidades se calculan todas para el mismo tiempo de acuerdo a las siguientes expresiones:

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{\delta t^2}{2} a(t) \quad (1.10)$$

$$v(t + \delta t) = v(t) + \frac{\delta t}{2} (a(t) + a(t + \delta t)) \quad (1.11)$$

Para continuar con la dinámica es preciso calcular las velocidades a tiempo  $t + \delta t/2$  que se obtienen de la siguiente forma:

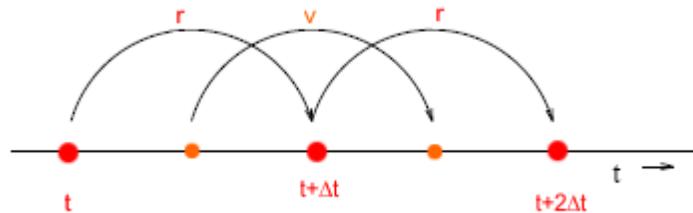
$$v(t + \frac{\delta t}{2}) = v(t) + \frac{\delta t}{2} a(t) \quad (1.12)$$

Luego las nuevas fuerzas se calculan desde las posiciones actuales, dando como resultado  $a(t + \delta t)$  y finalmente, en el último paso se determinan las velocidades para el tiempo  $t + \delta t$  de acuerdo a:

$$v(t + \delta t) = v(t + \frac{\delta t}{2}) + \frac{\delta t}{2} a(t + \delta t) \quad (1.13)$$

### Leap-frog

Este algoritmo lleva el nombre de "leap-frog" (en español: "salto de rana") debido a que se obtienen las velocidades y las posiciones en tiempos diferentes de manera solapada (ver figura 6). La velocidad se calcula cada medio paso de tiempo ( $\delta t/2$ ) mientras que las posiciones lo hacen a cada paso de tiempo ( $\delta t$ ).



**Figura 6:** Esquema de los pasos de integración mediante el algoritmo de leap-frog

Las posiciones y las velocidades se obtienen de los valores a tiempo  $t$  y  $\delta t/2$  respectivamente. Las mismas se actualizan considerando la aceleración  $a$  del mismo modo que sucede en Verlet (ver 1.13).

$$v(t + \frac{\delta t}{2}) = v(t - \frac{\delta t}{2}) + a(t) \delta t \quad (1.14)$$

$$r(t + \delta t) = r(t) + v(t + \frac{\delta t}{2}) \delta t \quad (1.15)$$

Cuando se implementa el algoritmo de leap-frog las velocidades  $v(t + \frac{t}{2})$  se calculan inicialmente de las velocidades a tiempo  $v(t - \frac{\delta t}{2})$  y de las aceleraciones a tiempo  $t$ .

### **3.1.6 Control térmico y de presión durante las simulaciones**

Durante una simulación de dinámica molecular se puede trabajar con diferentes tipos de ensambles termodinámicos, como por ejemplo, el canónico ( $\mathcal{N}, \mathcal{V}, \mathcal{T}$ ), microcanónico ( $\mathcal{N}, \mathcal{V}, \mathcal{E}$ ) o como generalmente sucede en los laboratorios, trabajar con el ensamble isobárico-isotérmico ( $\mathcal{N}, \mathcal{T}, \mathcal{P}$ ). Para asegurarse que se trabaja en el tipo de ensamble deseado es que es necesario utilizar métodos de control de aquellas magnitudes constantes a lo largo de toda la simulación.

Existen diferentes métodos para mantener la presión constante a lo largo de la simulación, uno de ellos es el de Berendsen (escogido en este trabajo), donde se acopla débilmente el sistema a un baño externo usando el principio de mínima perturbación local. Esto se logra agregando un término extra a la ecuación de movimiento que efectúe los cambios necesarios para mantener constante la presión.

$$\left( \frac{\partial p}{\partial t} \right)_{\text{baño}} = \frac{P_0 - P}{\tau_p} \quad (1.16)$$

Siendo  $\tau_p$  la constante del tiempo para el acoplamiento

Análogamente, para asegurar que la temperatura promedio del sistema sea correcta también se agrega un término más a la ecuación de movimiento. Una forma de hacer esto es a través del termóstato de Berendsen donde se eliminan las fluctuaciones en la energía cinética de modo de asegurar que la temperatura se mantenga estable [33]. Las velocidades se escalan entonces a cada paso de modo que la tasa de cambio de la temperatura sea proporcional a la diferencia entre la temperatura del baño acoplado y la del sistema de la siguiente forma:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{\text{baño}} - T(t)) \quad (1.17)$$

Donde  $\tau$  representa la constante de acoplamiento entre el baño y el sistema, a mayor  $\tau$ , mayor será el acoplamiento.

### **3.1.7 Algoritmos para minimización energética**

Una vez obtenida una expresión que permita calcular la energía de un sistema molecular dado, es posible recorrer la superficie de energía potencial que esta expresión describe. Los

mínimos de esta superficie multidimensional representan las conformaciones energéticamente más estables del sistema estudiado.

Para realizar esto existen diversos métodos, donde el más apropiado dependerá del sistema en el que se trabaje, así como del enfoque (cuántico o mecánico clásico), considerando siempre la relación del gasto computacional respecto a la calidad del resultado. Dos de los algoritmos más utilizados son los llamados: "steepest descent" y "conjugate gradient". A modo informativo es conveniente explicar un poco en qué consisten cada uno de estos algoritmos.

### **Descenso por gradiente (*steepest descent*)**

Este método derivativo de primer orden se basa en un algoritmo que explora la superficie de energía potencial moviéndose de un paso a otro en la dirección paralela a la fuerza neta. La fuerza neta se calcula mediante el gradiente de la energía potencial ( $F = -\nabla E_{\text{pot}}$ ) de la estructura inicial de acuerdo con el campo de fuerza elegido. Como se trata de un método iterativo cada nueva posición generada en un paso determinado se obtiene a partir de la posición resultante del paso anterior. Esquemáticamente el algoritmo seguido es el siguiente:

- Dado un punto inicial en la superficie de energía potencial se calcula el gradiente de ella en ese punto, definiendo así la dirección en la que se moverá sobre la superficie:

$$s_k = -\frac{g(x_k)}{|g(x_k)|} \quad \text{donde } s_k \text{ es el vector unitario de dirección de búsqueda, } x_k \text{ las}$$

coordenadas en el paso k y  $g(x_k)$  el gradiente de la energía potencial.

- Con la dirección definida ( $s_k$ ) se calcula la distancia en que se moverá:

$$x_{k+1} = x_k + \lambda_k s_k \quad \text{donde, } \lambda_k \text{ es la magnitud del salto en la dirección de búsqueda.}$$

$\lambda_k$  puede ser un valor arbitrario fijo asignado o obtenerse mediante una búsqueda lineal.

En este último caso el objetivo es identificar el mínimo a lo largo de una dirección específica (de allí el nombre en inglés "*linear search*").

Estos pasos se repiten hasta que se alcance alguna de las condiciones finales dadas por el usuario. Puede ser hasta alcanzar un número máximo de pasos o al alcanzar un valor de gradiente menor a un valor especificado.

### **Gradiente conjugado (*conjugate gradient*)**

El método de gradiente conjugado también es derivativo de primer orden, pero el algoritmo es diferente al del caso anterior. Esto se debe a que se utiliza otro criterio para escoger la dirección

en la que se moverá sobre la superficie de energía potencial. Como lo indica el nombre, la dirección se obtiene de conjugar las dos direcciones de los gradientes de los pasos consecutivos:

$v_k = -g(x_k) + \gamma_k v_{k-1}$  donde  $v_k$  es el vector dirección de búsqueda en el paso k,  $x_k$  las coordenadas en el paso k,  $g(x_k)$  el gradiente de la energía potencial en el paso k y  $\gamma_k$  es un escalar definido en su versión original (algoritmo de Fletcher-Reeves) como:

$$\gamma_k = \frac{g(x_k) \bullet g(x_k)}{g(x_{k-1}) \bullet g(x_{k-1})}$$

Si bien conjugate gradient es computacionalmente más costoso que steepest descent, este método converge más rápidamente que el anterior y comúnmente se lo utiliza a continuación de steepest descent.

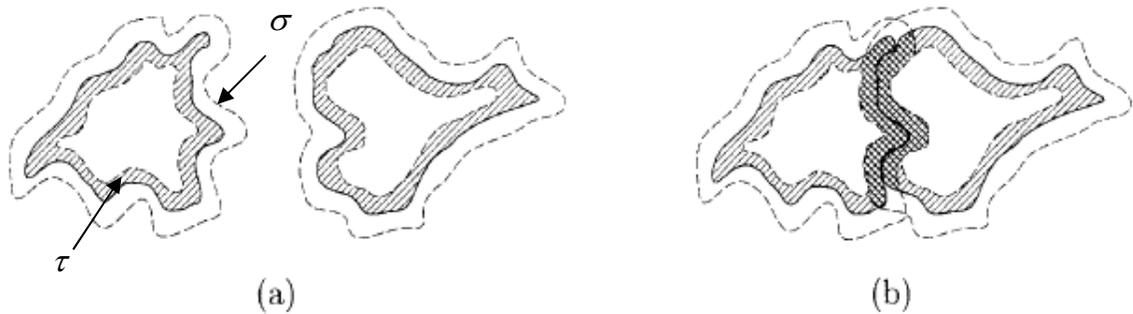
### **3.2 Docking**

Una técnica computacional de amplio uso para estudiar interacciones entre moléculas es el “docking”. Se trata de una herramienta de predicción de complejos receptor-ligando basado en la complementariedad y afinidad entre ellos. Si bien su aplicación estuvo originalmente vinculada al desarrollo de fármacos donde el ligando es generalmente una pequeña molécula orgánica con presuntas propiedades farmacológicas y el receptor corresponde a una macromolécula (generalmente una enzima o porción de la misma donde actúa, como por ejemplo el sitio activo) existe un número creciente de aplicaciones de este tipo de aproximación para estudiar interacciones entre proteínas o partes de las mismas y así tener una primera noción de la viabilidad de dicha interacción. De hecho, existe un creciente interés de la comunidad científica en mejorar los métodos de docking se manifiesta en un proyecto en marcha desde 2001 denominado CAPRI (del inglés: *Critical Assessment of PRediction of Interactions*) donde varios grupos desarrollan algoritmos para predecir complejos proteicos utilizando únicamente las estructuras de los componentes determinados por separado[34].

Genéricamente hablando, un análisis de docking pretende dar una idea de la posibilidad de unión de un complejo, mediante el muestreo de posibles modos de interacción. En el caso de proteínas, debido al gran número de grados de libertad conformacionales, una aproximación ulterior es la de considerar que ambos, el receptor y la “proteína ligando”, son rígidos, es decir considerando sólo seis grados de libertad (tres correspondientes al movimiento de rotación y los otros tres al de traslación). Una vez hecha la búsqueda geométrica existe un algoritmo que permite asignar un puntaje, es decir, una medida de interacción, pudiendo ésta ser tomada como un indicador relacionado a la energía libre de unión. Esto se consigue por medio de la evaluación energética de las diferentes orientaciones relativas entre receptor y ligando que se prueban para poder determinar las mejores de acuerdo al criterio utilizado. Este criterio puede ser simplemente en base a complementariedad estructural o también incluir información de interacciones electrostáticas y de van der Waals. Otras variaciones que se dan a la hora de implementar un análisis de docking están vinculadas a como se modelan el receptor y el ligando. Éstos pueden considerarse a un nivel atómico, donde explícitamente se represente a cada átomo (con sus propiedades espaciales y electrostáticas) o si se representa de manera simplificada (por ejemplo considerando grupos de átomos con características asociadas) a la hora de analizar las interacciones en grupos de átomos.

Por ejemplo, Hex (el programa utilizado en este trabajo para los cálculos de docking) se caracteriza por utilizar correlaciones esféricas de Fourier (*SPF: Spherical Polar Fourier correlations*) para acelerar los cálculos [35]. Utilizan un modelo con dos capas (*double skin*) para representar la superficie que les permite definir dos vectores de coeficientes de expansión denominados  $\tau$  y  $\sigma$ .  $\sigma$  representa la capa más exterior y está asociado al volumen de la superficie molecular y de la superficie accesible al solvente, mientras que  $\tau$  representa la región

interior, es decir los volúmenes de van der Waals de los átomos dentro de la superficie molecular (Figura 7)



**Figura 7a y b:** Esquema de complementariedad de forma de acuerdo al modelo de *double skin*. La línea punteada delimita la capa exterior (volumen de superficie accesible al solvente) mientras que la capa interior está representada por las regiones sombreadas. En b la región central remarcada se señala el volumen accesible al solvente ocluido en la asociación óptima entre ambas proteínas. Modificado de [35].

Como método de scoring utiliza: 
$$S = \int \sigma_A \tau_B dV + \int \tau_A \sigma_B dV - Q \int \tau_A \tau_B dV$$

Donde los primeros dos términos representan las contribuciones favorables mientras que el último término representa la contribución desfavorable (Q es un factor de penalización).

## IV. MÉTODOS

### 4.1 Modelos estructurales utilizados

Como estructuras iniciales de partida para la región de docking y dimerización correspondiente a la subunidad regulatoria de PKA II $\alpha$  (R11 $\alpha$ DD) se tomaron dos estructuras cristalográficas de dicha región tanto en complejo con una proteína de anclaje a PKA, AKAP-IS, (PDBID: 2IZX) como el dímero sólo (PDBID: 2IZY). [15]

Puesto que actualmente no se conoce experimentalmente la estructura de la proteína core del virus de la Hepatitis C, se realizó un modelo basado en la publicación de la predicción de estructura secundaria por el método de predicción auto-optimizado NPSA [26] de la región de interés. Éste predice una estructura de  $\alpha$ -hélice (R113-I137<sup>(\*)</sup>) que luego fue modelada con el programa Swiss-Pdb Viewer v4.0.1[36]. A su vez, como esta proteína tiene una región anfipática que se asimila a la existente en AKAPs, se realizó un análisis múltiple de secuencias contra secuencias AKAPs documentadas. Para obtener un modelo estructural de la interacción AKAP/PKA, se realizó una superposición estructural entre la hélice predicha de VHCcore y el péptido AKAP-IS en presencia del dominio R11 $\alpha$ DD utilizando el algoritmo STAMP implementado en VMD [37]. Esto produjo un modelo del dominio putativo de la interacción entre VHCcore y R11 $\alpha$ DD, que fue utilizado como punto de partida para formar el complejo R11 $\alpha$ DD/VHCcore *in silico*.

### 4.2 Secuencias con numeraciones utilizadas

En negrita se marcan los aminoácidos presentes en los modelos.

#### 1) PKAR11 $\alpha$ DD

El orden de las secuencias corresponde a la primera cadena proteica y la a segunda cadena respectivamente.

(*homo sapiens*):

```
1 SHIQIPPGLT ELLQGYTVEV LRQPPDLVE FAVEYFTRLR EARAPASVLP 50
1 SHIQIPPGLT ELLQGYTVEV LRQPPDLVE FAVEYFTRLR EARAPASVLP 50
```

#### 2) VHCcore

```
1 MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR KTSERSQPRG
61 RRQPIPKARR PEGRTWAQPG YPWPLYGNEG CGWAGWLLSP RGSRPSWGPT DPRRRSRNLG
121 KVIDTLTCGF ADLMGYIPLV GAPLGGAARA LAHGVRVLED GVNyatGNLP GCSFSIFLLA
181 LLSCLTPAS A
```

En rojo se señala la serina fosforilada por PKA

(\*) De ahora en más denominada como VHCcore.

### 4.3 Campo de fuerza y parámetros específicos

El campo de fuerza utilizado fue Amber ff99/SB [31]<sup>(\*)</sup> y el modelo explícito de agua utilizado correspondió a TIP3P. El sistema final se conformó por una caja prismática conteniendo aproximadamente 9800 moléculas de agua e iones de  $K^+$  y  $Cl^-$  necesarios para simular una concentración salina equivalente a 150mM (medio intracelular) trabajando a pH 7.0 (pH en el que los residuos glutamato y aspartato tienen carga negativa; arginina y lisina positiva e histidina es neutro).

Los archivos de topología y coordenadas fueron generados por medio de xLEaP/tLeap[38] y los parámetros de la fosfoserina utilizados correspondieron a aquellos desarrollados por Craft y colaboradores [39].

### 4.4 Potenciales electrostáticos

Se realizaron estudios del potencial electrostático para estudiar el efecto electrostático en la formación de los complejos, esto se realizó para los complejos modelados con los parámetros mencionados en el párrafo anterior (Amber ff99/SB). Para ello, se resolvió la ecuación de Poisson-Boltzman lineal mediante el programa APBS [40]. Las condiciones correspondieron a una temperatura de  $T=300K$ , constante dieléctrica de  $\epsilon=1$  para la región de la proteína y  $\epsilon=78,54$  para la región correspondiente al solvente.

### 4.5 Docking

Se utilizó el programa Hex v.6.1[35] para realizar los estudios de docking rígido entre VHCcore o fVHCcore y PKA  $I\alpha(RII\alpha DD)$ . Hex realiza docking rígido (no se permite la flexibilidad intrínseca de las proteínas, tanto del "receptor" como del "ligando").

Como estructuras de referencia se escogieron los modelos minimizados preparados para la dinámica molecular. Se realizó también docking con AKAP-IS como forma de tener un control. Hex evalúa las energías utilizando los parámetros de Van der Waals de OPLS y las cargas provenientes del campo de fuerza Amber ff99 y las expresa en unidades propias de Hex. Para poder estudiar el sistema formado por PKA y la proteína core fosforilada se utilizaron los valores de las cargas pertenecientes a la serina fosforilada de acuerdo a la parametrización de Craft et al [39]. Dado que el principal interés está enfocado en el efecto provocado por un grupo fosfato se optó por incrementar en un factor de diez el peso de los términos electrostáticos aplicados a la energía

---

<sup>(\*)</sup> Amber ff99SB es una variante de ff99 donde se reparametrizaron los valores de los parámetros torsionales para lograr una mejor descripción de estructuras secundarias tal como se explica en [31].

calculada por el algoritmo de docking, como es sugerido por los autores [35]. Esto se hizo modificando el valor de la variable *factor\_g* adoptando valores de 7000J/mol.

Una vez obtenidos los resultados del docking se visualizaron las 5000 soluciones obtenidas mediante el programa VMD v1.8.7 [41]. Utilizando la herramienta Volmap implementada en la misma versión de VMD[41] se generaron representaciones de ocupancia para caracterizar los modos de unión más frecuentes obtenidos luego del docking. Esto se calculó a un nivel de resolución de 1 Å en función de la ocupancia del C-alfa central del péptido (D14 en AKAP-IS y D123 en VHCcore o fVHCcore) considerando las 5000 soluciones y combinándolas en promedio.

#### 4.6 Protocolo de Simulación

Para las simulaciones de dinámica molecular correspondientes al estudio con solvente explícito se utilizó el programa Gromacs (v4.0.5)[42] con los parámetros correspondientes al campo de fuerza Amber ff99/SB y TIP3P para modelar la porción proteica del sistema y el solvente respectivamente.

Una vez que se terminaron de preparar los modelos se procedió a minimizar y equilibrar el sistema previo al inicio de la dinámica molecular propiamente dicha. Se llevaron a cabo 10000 pasos de minimización combinando los algoritmos de *steepest descent* y *gradiente conjugado* sin restricciones y con un criterio de convergencia <10kJ/mol. Luego se equilibró calentando hasta 300 K a presión constante (1 bar) isotrópica, utilizando el termóstato y baróstato de Berendsen (necesario dado que interesa simular el ensemble NPT).

Se trabajó con condiciones periódicas de contorno, así como con tratamiento de tipo "Particle Mesh Ewald" para las interacciones electrostáticas. El valor de cutoff correspondió a 12 Å para las interacciones de largo alcance (VDW y electrostáticas). Se restringieron los enlaces que involucran hidrógenos por medio del algoritmo Lincs, habilitando utilizar así un paso de integración de 2 fs.

La simulación fue realizada guardando la información cada 8 ps y en fragmentos de a 10 ns hasta totalizar 130 ns de simulación.

Posteriormente se realizaron análisis de cluster de las trayectorias mediante la aplicación *g\_cluster* de Gromacs escogiendo el algoritmo de jarvis-patrick y un cutoff de 2Å de rmsd entre estructuras para definir los vecinos.

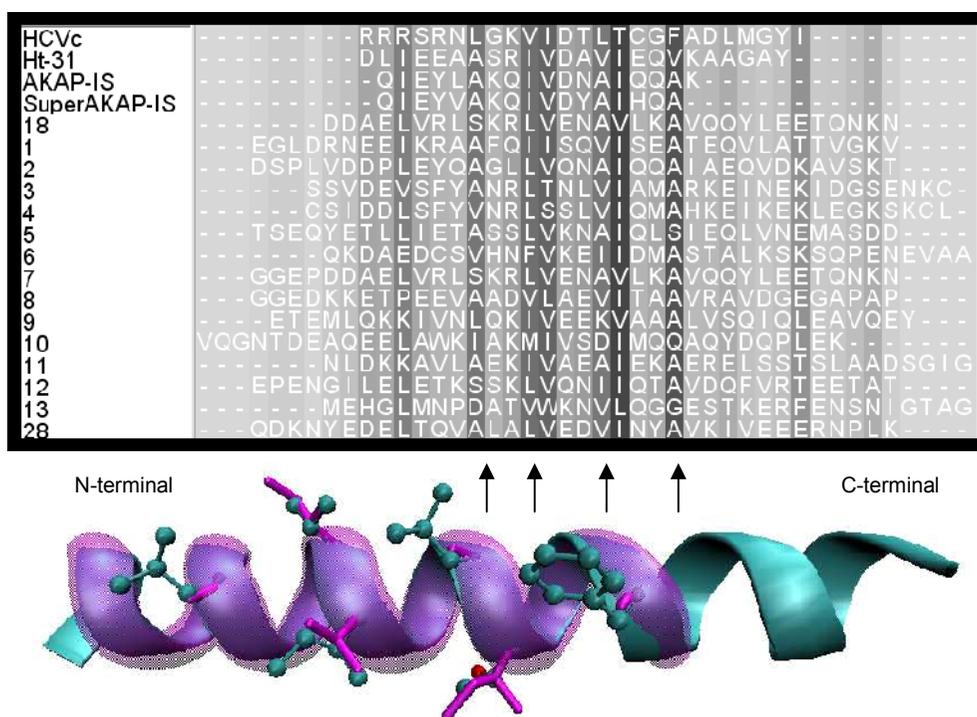
Para calcular el área de interfase se obtuvieron los valores de superficie accesible al solvente mediante la herramienta *g\_sas*, utilizando un radio de solvente de 1,4Å y los valores de radios de Van der Waals correspondientes al campo de fuerza utilizado (ff99SB).

## V. RESULTADOS

### 5.1 Alineamiento múltiple y estructural

Existen numerosos péptidos de unión a PKA que han sido descubiertos en los últimos años. Si bien no existe aún un motivo específico que los permita agrupar (y/o predecir) con exactitud, es sabido que la principal propiedad que éstos conservan es su carácter anfipático y estructura de alfa-hélice.[43]

Una vez identificada la región putativa de interacción se realizó un alineamiento múltiple de secuencias de AKAPs conocidas y la región de interés correspondiente a VHCcore para poder sugerir un modo de unión a PKA (Figura 8).



**Figura 8:** En la parte superior se muestra un alineamiento múltiple VHCcore (aa 113-137) con secuencias AKAPs conocidas, Las columnas oscuras corresponden a regiones de mayor conservación de residuos entre las secuencias estudiadas. En la parte inferior se presentan las estructuras superpuestas acorde al alineamiento de las secuencias y estructural de VHCcore(turquesa) y AKAP-IS(magenta) señalando las cadenas laterales de los residuos hidrofóbicos correspondientes a las cuatro columnas más conservadas del alineamiento (residuos de VHCcore: L119, V122, I123, T126, I127, F130; residuos de AKAP-IS: A9, I12, V13, A16, I17, A20),

De acuerdo al grado de conservación en las secuencias reportadas como AKAPs (Figura 8), se puede apreciar que estas regiones corresponden principalmente a residuos hidrofóbicos

(V,I,L,M,A,F). Las dos columnas centrales (con dos residuos hidrofóbicos cada una) pueden considerarse como regiones complementarias a los bolsillos hidrofóbicos existentes en la superficie de RII $\alpha$ DD.[15, 44] Como puede observarse desde el alineamiento estructural basado en el alineamiento múltiple es posible deducir que el segmento comprendido entre las posiciones 113 a 137 en VHCcore reúne las características de estructura y composición aminoacídica que definen a las proteínas AKAP.

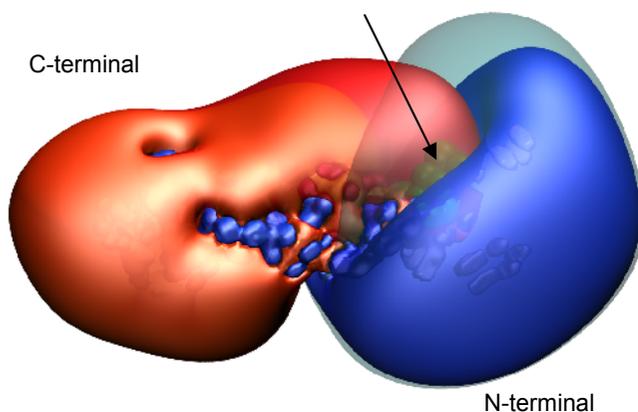
Por este motivo, de acuerdo al método descrito en IV, se realizaron los modelos iniciales tomando como molde la estructura cristalográfica del complejo formado entre PKA-DD y AKAP-IS (PDB ID: 2izx, resolución: 1.8Å) donde se superpuso a VHCcore sobre AKAP-IS de acuerdo al resultado obtenido del alineamiento múltiple y estructural.

## 5.2 Potenciales electrostáticos

Si bien el tipo de interacciones características que se dan entre las AKAPS y PKA son mayormente hidrofóbicas (por eso se tratan de hélices  $\alpha$  anfipáticas), una peculiaridad observada en algunas AKAPs es la influencia de interacciones entre residuos cargados de AKAPs y de PKA en las regiones que flanquean el típico motivo de unión a PKA.

Es muy importante notar que la secuencia del péptido identificado (residuos 113 a 137) contiene el motivo blanco de fosforilación de PKA (Arg-X-X-Ser). De hecho, ha sido demostrado que la Serina 116 es uno de los blancos mayoritarios de fosforilación en VHCcore por la isoforma II de PKA [25, 45]. Con el propósito de investigar el efecto de la fosforilación en la interacción propuesta se realizaron cálculos de potenciales electrostáticos de ambos complejos modelados así como también del complejo cristalográfico.

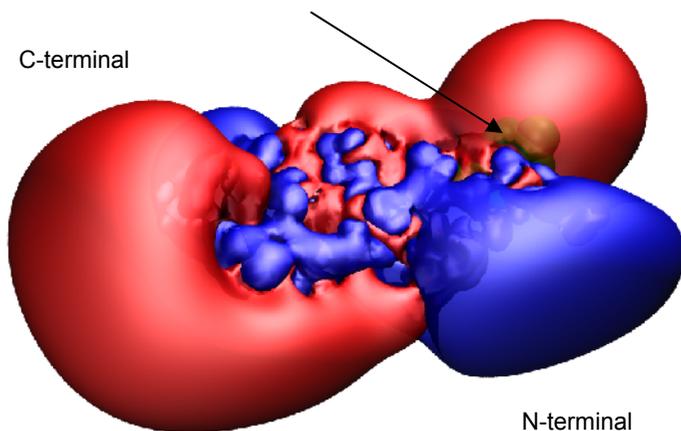
En la figura 9 se muestran los potenciales electrostáticos superpuestos de VHCcore fosforilado y sin fosforilar. Como es de esperarse, se puede apreciar claramente el efecto de la fosforilación en el potencial correspondiente a la región amino terminal del péptido donde se



encuentra la serina fosforilada (con carga neta = -2). El potencial electrostático en N-terminal no solamente presenta menor valor positivo sino que éste habilita una prolongación de la región con potencial negativo hacia dicha región.

**Figura 9:** Potenciales electrostáticos de fVHCcore (rojo= -5 mV; azul= 5 mV) y VHCcore (naranja= -5 mV; turquesa= 5 mV). La flecha señala la ubicación correspondiente a S116 (representada como esferas de VDW verdes).

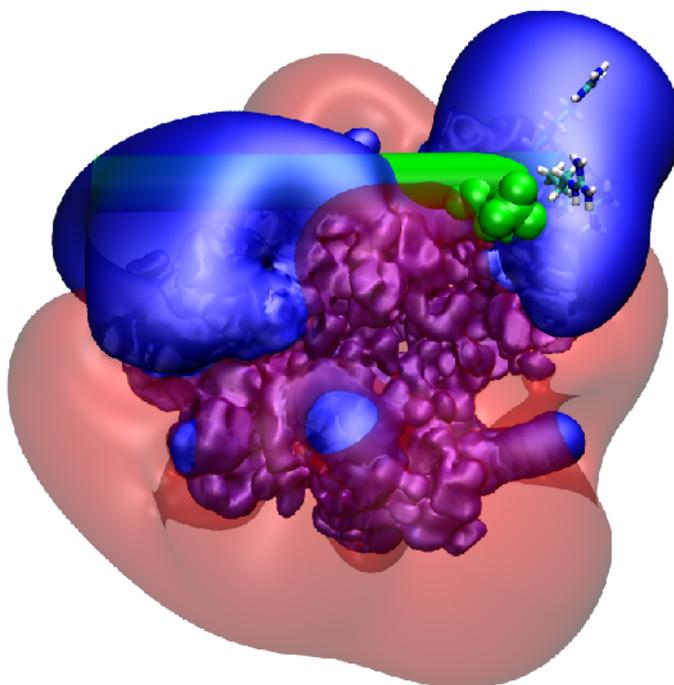
Sin embargo, resulta más interesante observar esto en comparación con AKAP-IS, debido a que



**Figura 10:** Potenciales electrostáticos de AKAP-IS (rojo= -5 mV y azul =5 mV). La flecha señala la ubicación correspondiente a E6 (representada como esferas de VDW verdes).

tiene una interacción conocida con PKA. En la figura 10 Se puede apreciar la isosuperficie correspondiente a los valores de potencial de 5 mV (5 mV=azul, -5 mV=rojo). Aquí se marcó en color naranja (señalado con una flecha) al E6 de AKAP-IS equivalente a la S116 de acuerdo al alineamiento propuesto. Como se trata de un glutamato, se aprecia en esa región un foco de potencial negativo en el extremo amino-terminal de AKAP-IS. El hecho de que exista esta región negativa en el extremo N-terminal apoya la idea

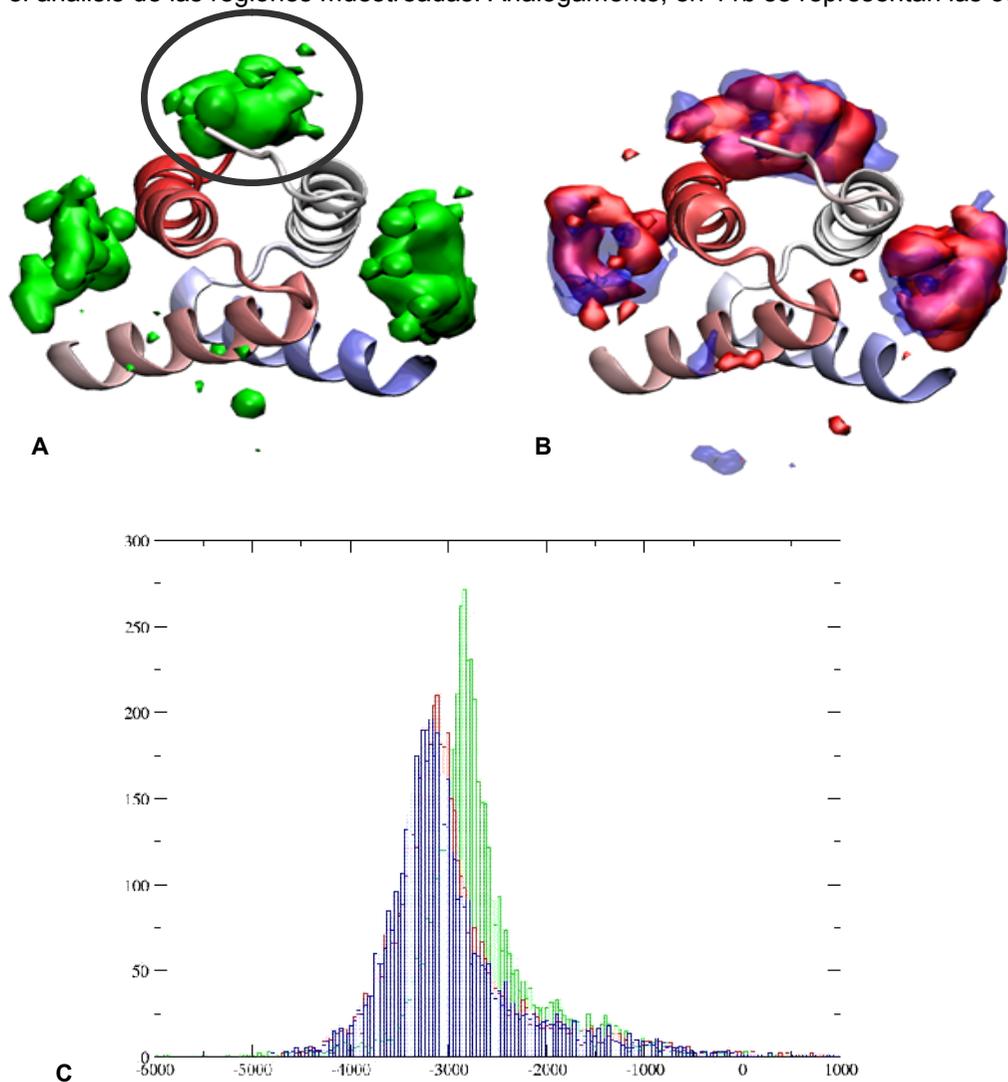
que la fosforilación de VHCcore pueda favorecer la interacción con PKA, ya que reduce la carga positiva de las tres argininas en esa región. Esto podría ser importante, ya que la importante carga positiva en la región N-terminal podría repeler la interacción con PKA, puesto que en PKA, en la región complementaria existe también una zona de potencial positivo (ver figura 11).



**Figura 11:** Potenciales electrostáticos de RII $\alpha$ DD (rojo= -5 mV y azul =5 mV). Se muestra además en complejo con fVHCcore (cilindro verde), donde se señala también a S116 fosforilada (representada en verde como superficie de VDW). En la región cercana al extremo N-terminal de fVHCcore se observan las argininas del péptido ubicadas en la zona de potencial positivo de PKA.

### 5.3 Docking

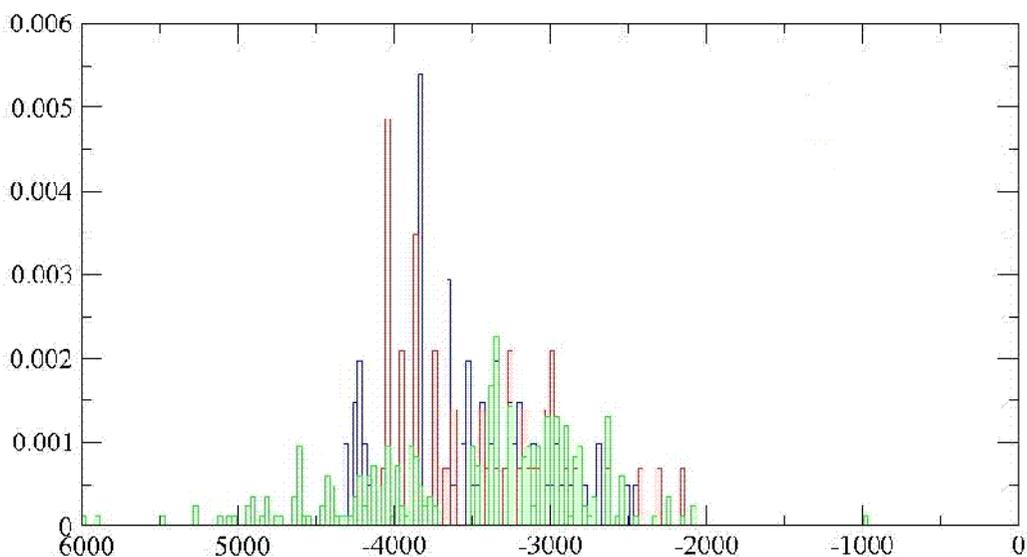
Como siguiente paso se realizó un estudio de docking para observar los principales modos de unión de los péptidos VHCcore y fVHCcore al dominio RII $\alpha$ DD. Como sistema de control positivo se utilizó el complejo RII $\alpha$ DD /AKAP-IS (ver sección 4.5). El resultado del estudio de docking para el caso del complejo RII $\alpha$ DD /AKAP-IS se ilustra en la figura 12a donde se representa la ocupancia del átomo de carbono alfa correspondiente al residuo central de los péptidos modelados (D124 en VHCcore o fVHCcore y D14 en AKAP-IS). Este criterio fue escogido para facilitar el análisis de las regiones muestreadas. Análogamente, en 11b se representan las ocupan-



**Figura 12:** Representación de la ocupancia de las soluciones de docking de los péptidos de estudio. En (A) se muestra la ocupancia correspondiente a las soluciones de AKAP-IS (regiones verdes) resaltando el motivo de unión AKAP y en (B) las azules corresponden a VHCcore y las rojas a fVHCcore. En (C) se observan las energías de docking obtenidas con Hex (v.6.1) de las 5000 soluciones finales para los complejos RII $\alpha$ DD/AKAP-IS (verde), RII $\alpha$ DD/fVHCcore (rojo) y RII $\alpha$ DD/VHCcore (azul).

cias de fVHCcore y VHCcore. En los tres casos se pueden identificar claramente tres regiones mayoritarias de poblaciones, donde dos de ellas corresponden a interacciones inespecíficas en los surcos de cada protómero de RII $\alpha$ DD. Únicamente una región se correlaciona con el motivo de unión AKAP específico (región central resaltada con el círculo), que se puede discriminar por la información obtenida experimentalmente de elucidación estructural de complejos AKAP-PKA. Es posible entonces decir que el algoritmo de docking hizo un correcto mapeo de la superficie de RII $\alpha$ DD lo que se correlaciona con los perfiles energéticos obtenidos de todas las soluciones de docking donde se observa una típica distribución de forma de campana donde la distribución de las soluciones totales de fVHCcore y VHCcore aparentan estar levemente más sesgadas a regiones más favorables (menores energías).

De cualquier manera, para hacer un análisis más meticuloso es conveniente depurar las soluciones considerando la información estructural que concierne específicamente al modo de unión AKAP-PKA. Para ello se realizó un análisis más minucioso observando únicamente aquellas soluciones de docking correspondientes a las orientaciones de cada uno de los tres casos de estudio cuyos rmsd distan hasta 2Å del complejo cristalográfico. La figura 13 muestra un histograma normalizado para cada uno de los casos. Haciendo un primer examen visual de los puntajes de estas soluciones depuradas es posible observar que en el caso de AKAP-IS (verde) existe una amplia distribución de puntajes de docking registrados (variando desde la más favorable -6000 hasta -1000), mientras que en los complejos problema existe una leve tendencia a localizar soluciones con mayor probabilidad en las regiones correspondientes al intervalo (-4000,-2000) aproximadamente.



**Figura 13:** Gráfica mostrando los histogramas normalizados correspondientes a las diferentes energías de docking de las orientaciones que distan hasta 2Å de la estructura de referencia. RII $\alpha$ DD/AKAP-IS (verde), RII $\alpha$ DD/fVHCcore (rojo) y RII $\alpha$ DD/VHCcore (azul). Las unidades energéticas corresponden a unidades de Hex.

Por otro lado, se realizó un test estadístico para comparar las distribuciones de dichas energías correspondientes a los complejos depurados. Se utilizó el test de suma de rangos de Wilcoxon para muestras independientes no pareadas a un nivel de confianza de 0.95 bajo la hipótesis nula que ambas distribuciones comparadas son iguales. Los resultados de este test arrojaron que los perfiles energéticos obtenidos entre el complejo fosforilado y el no fosforilado no son estadísticamente diferenciables así como tampoco los son en comparación con AKAP-IS (Tabla 5.1)

**Tabla 5.1:** Resultados del Test de suma de rangos de Wilcoxon

<b>muestras</b>	<b>W</b>	<b>p-valor</b>
VHCcore/fVHCcore	1510	0.4958
VHCcore/AKAP-IS	9193	0.6608
AKAP-IS/fVHCcore	6333.5	0.5248

De cualquier manera, aunque no se pueda rechazar la hipótesis nula y por lo tanto no poder diferenciar estadísticamente las poblaciones de energías de docking obtenidas para cada complejo de estudio (considerado la restricción de que la estructura no se desvíe en más de 2Å de la cristalográfica tomada como referencia), se observó que ambos complejos modelados (R11 $\alpha$ DD/VHCcore y R11 $\alpha$ DD/fVHCcore) presentan soluciones en el sitio conocido donde se anclan las AKAPs presentando inclusive un perfil similar al observado en AKAP-IS. Esto implica que las soluciones obtenidas son cuanti y cualitativamente comparables a la estructura cristalográfica de referencia (complejo con AKAP-IS), de acuerdo a la hipótesis que VHCcore puede establecer interacciones proteína-proteína del tipo AKAP con R11 $\alpha$ DD. Por todos estos motivos, y como interesa estudiar el efecto de la fosforilación en S116 de VHCcore, se buscó un enfoque que permite estudiar con el detalle suficiente el efecto de la fosforilación en la interacción planteada, eligiendo un método más sensible al efecto de la fosforilación aunque computacionalmente más costoso, como son las simulaciones de dinámica molecular.

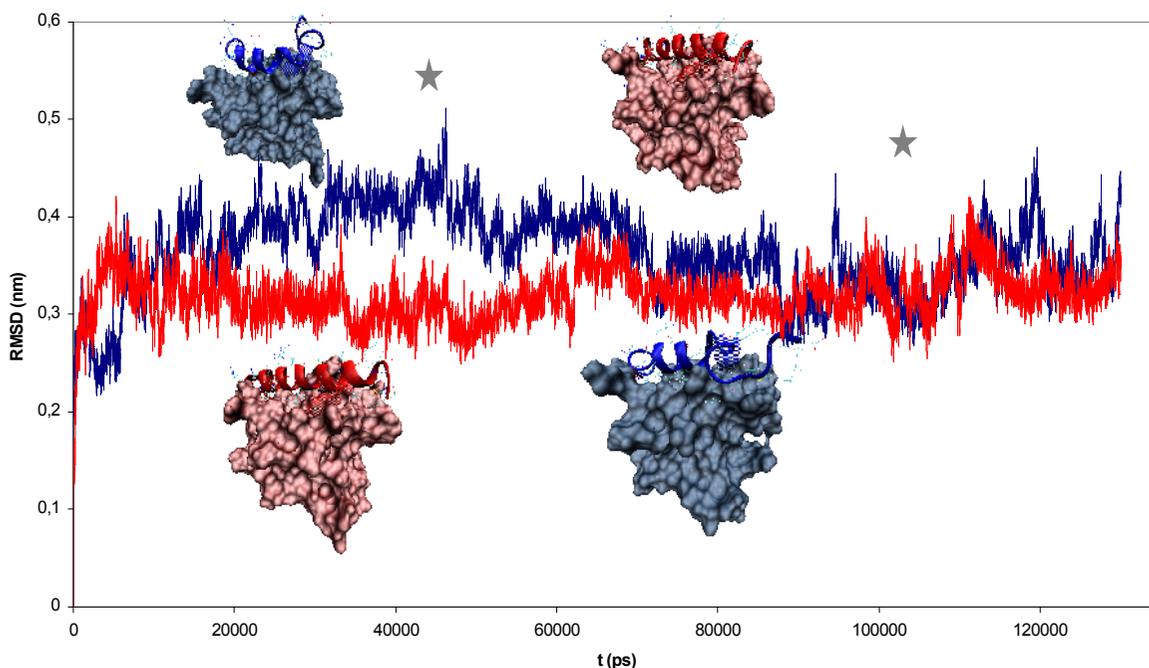
## 5.4 Dinámica molecular

### 5.4.1 Análisis general

Con el fin de estudiar el efecto de la fosforilación sobre la estructura del complejo RII $\alpha$ DD/VHCcore se realizaron simulaciones de dinámica molecular. A partir del modelo estructural descrito en la sección 5.1, se simularon dos sistemas, fosforilado y no fosforilado en el ensemble NPT (número de partículas, presión y temperatura constante). Se decidió partir desde el modelo basado en el alineamiento estructural dado que por este método se obtiene una única solución cercana a la información cristalográfica disponible para un ligando similar (AKAP-IS), mientras que por el método de docking se obtiene una cantidad arbitrariamente grande de soluciones refinadas con la misma información de la estructura cristalográfica. Luego de las primeras fases de minimización y equilibración donde se controló que se alcanzaran las condiciones de estabilidad en la temperatura, presión y energía se procedió con la fase de producción. A continuación se muestran los resultados de ambas simulaciones en términos comparativos.

### **RMSD**

Para analizar el desarrollo de la simulación de una manera global se graficó la evolución de la desviación cuadrática media (en inglés: *RMSD*) respecto al conformero inicial considerando todos los átomos del complejo luego de haber hecho una superposición basada sólo en los átomos correspondientes a RII $\alpha$ DD (Figura 14).

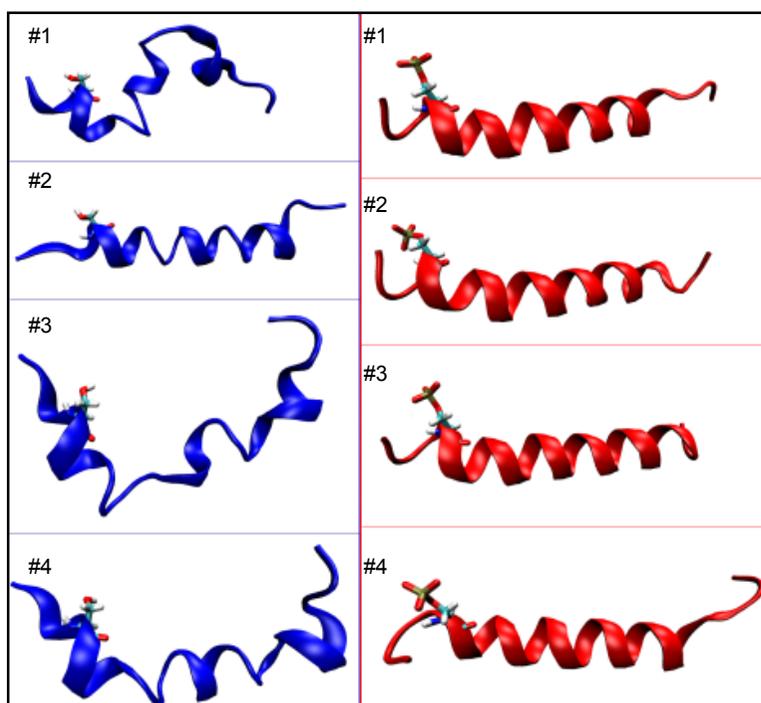


**Figura 14 (pág anterior):** Gráfica de RMSD(t) para el complejo correspondiente de RII $\alpha$ DD formado con VHCcore (azul) o con la proteína core fosforilada (rojo), luego de realizado el ajuste lineal considerando únicamente RII $\alpha$ DD con respecto a la estructura inicial luego de la minimización. Las estrellas señalan puntos de referencia de donde se tomaron las estructuras de los insertos.

Se puede apreciar que los valores de RMSD se mantienen con menos fluctuaciones en el caso del complejo RII $\alpha$ DD/fVHCcore, con una variación máxima de alrededor de 4 Å de la estructura de referencia (estructura inicial luego de la equilibración), mientras que el máximo valor que adopta el complejo RII $\alpha$ DD/VHCcore es del entorno de 5 Å.

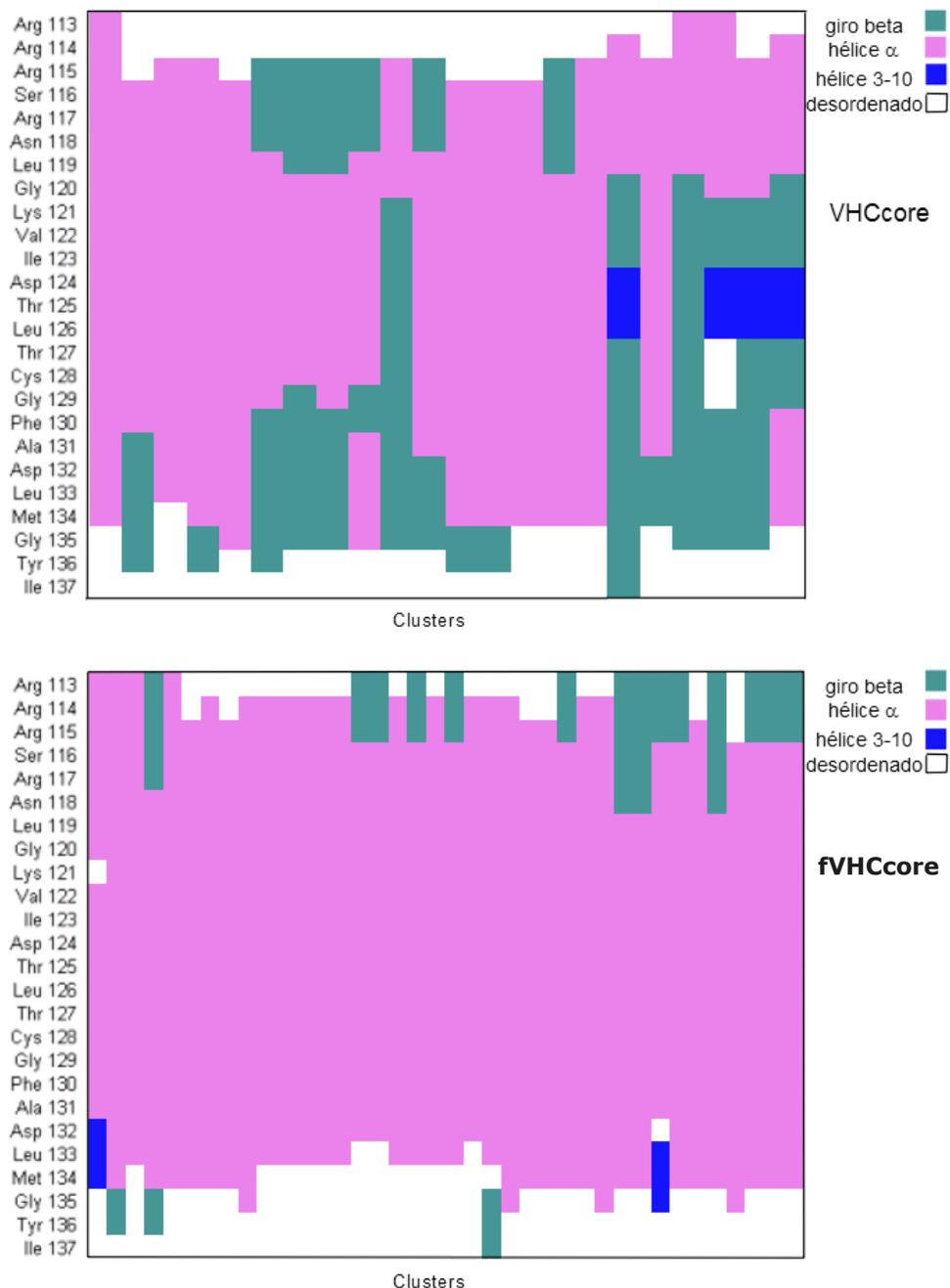
### Análisis de cluster de la trayectoria

La clusterización de las trayectorias de los modelos fue realizada con la finalidad de identificar las estructuras más representativas que adoptan los péptidos estudiados a lo largo de la evolución de la dinámica. En la figura 15 se muestran las conformaciones más representativas de los clusters más poblados en ambos casos de estudio.



**Figura 15:** Estructuras más representativas de cada uno de los primeros 4 clusters encontrados en las trayectorias correspondientes a RII $\alpha$ DD/VHCcore(azul) y RII $\alpha$ DD/fVHCcore (rojo). El aminoácido señalado es S116 (en hélice azul) y S116 fosforilada (en hélice roja).

Otra forma de visualizar rápidamente la variación en la estructura secundaria de los péptidos estudiados se presenta a continuación. En este caso se clasifica la estructura secundaria, como está definida en DSSP[46] que adopta cada aminoácido en cada una de las estructuras más representativas de cada cluster encontrado a lo largo de la trayectoria (figura 16). Se puede apreciar claramente que en la imagen superior, que corresponde a VHCcore existen mayores variaciones de la estructura helicoidal inicial (menos residuos se encuentran en la forma de hélice  $\alpha$ , que corresponde al color rosado).



**Figura 16:** Diagramas que muestran la estructura secundaria de la que está formando parte cada aminoácido que componen al péptido para cada una de las estructuras más representativas de cada cluster.

A raíz de estos cambios conformacionales extremos fácilmente visibles al observar la fluctuación en la estructura secundaria de VHCcore, se decidió realizar los análisis estructurales que estén basados en la formación de un complejo estable con RII $\alpha$ DD/fVHCcore.

### **5.4.2 Análisis estructural**

#### **Superficie accesible al solvente: superficie de interfase**

Para analizar los cambios en la superficie de la interacción durante el transcurso de la simulación, se procedió a estudiar la superficie de interfase entre el fVHCcore y RII $\alpha$ DD. Se utilizó para ello la aplicación *g\_sas* presente en el paquete de herramientas de análisis de Gromacs que calcula la superficie accesible al solvente (ver sección 4.6). Se consideró como área de interfase el área de RII $\alpha$ DD en contacto con fVHCcore más el área de fVHCcore en contacto con RII $\alpha$ DD. Para ello se definió la superficie de RII $\alpha$ DD que participa en la interfase PKA/fVHCcore como:

$$Int_{PKA} = Int_{PKA}^{PKA} - Int_{complejo}^{PKA} \quad (I)$$

donde los subíndices corresponden a la porción considerada para el cálculo y los superíndices a la región considerada para el resultado. Es decir, el área de interfase correspondiente a la porción RII $\alpha$ DD se calculó mediante la diferencia de la superficie accesible al solvente de RII $\alpha$ DD sin estar formando complejo ( $Int_{PKA}^{PKA}$ ) y la superficie accesible al solvente de RII $\alpha$ DD formando el complejo con fVHCcore ( $Int_{complejo}^{PKA}$ ).

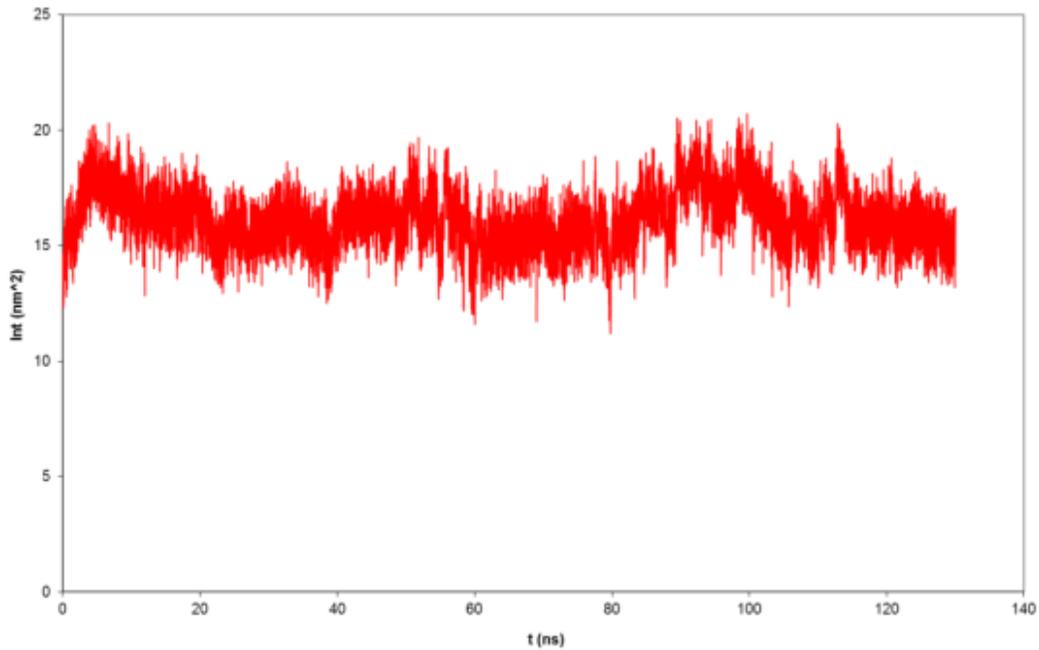
Análogamente, se realizó el mismo procedimiento para obtener el área correspondiente a la porción de fVHCcore en la interfase como se muestra en la ecuación (II):

$$Int_{fVHC} = Int_{fVHC}^{fVHC} - Int_{complejo}^{fVHC} \quad (II)$$

Finalmente, se definió:

$$Int_{PKA/fVHC} = Int_{PKA} + Int_{fVHC} \quad (III)$$

Este análisis se hizo para considerando todos los frames de la simulación, resultando así la gráfica de  $Int_{PKA/fVHC}(t)$  (Figura 17, página siguiente).



**Figura 17:** Gráfica de la superficie de interfase entre RII $\alpha$ DD y fVHCcore en función del tiempo de simulación transcurrido (ns)

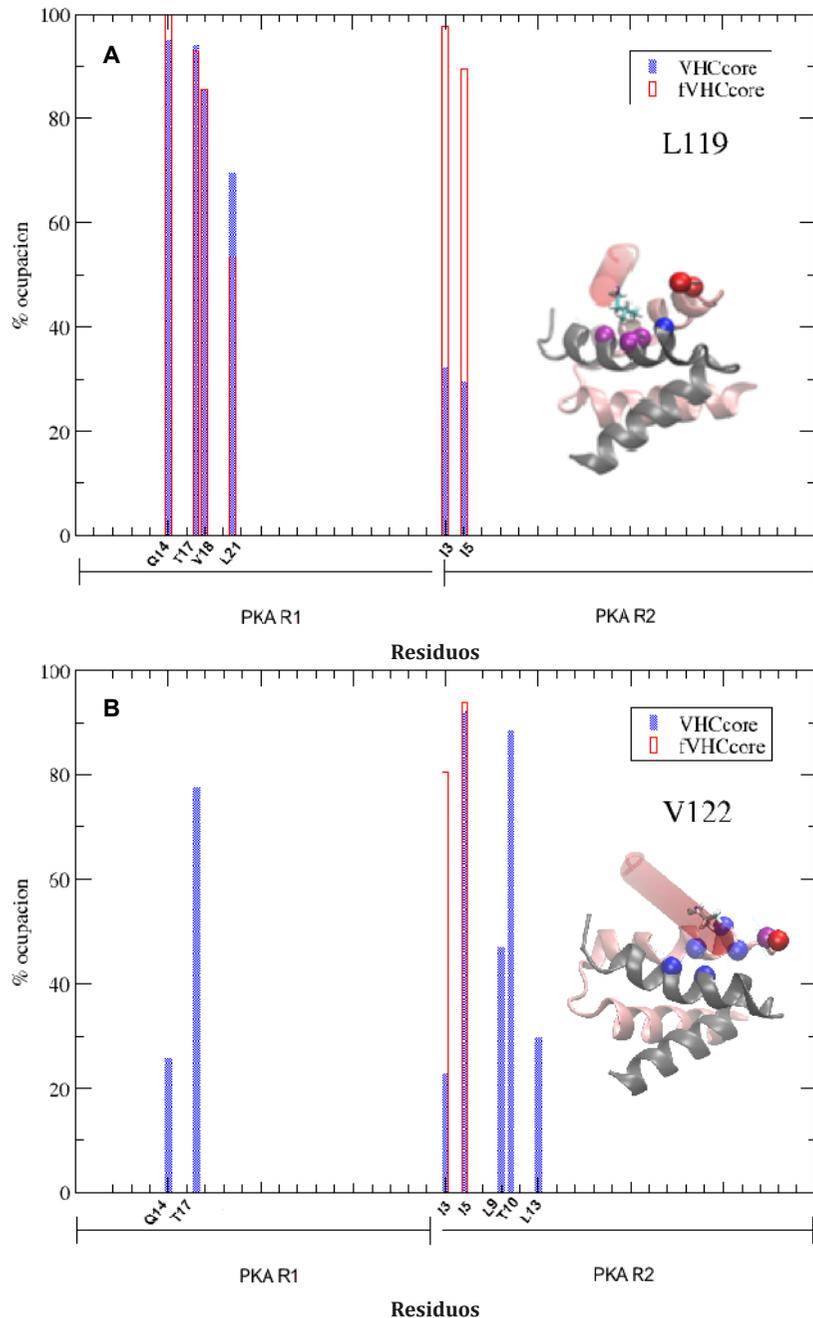
Para tener una idea más cuantitativa del valor de la superficie se calculó el área promedio de la Interfase tomando los datos a partir de los 50ns de simulación. Ésta valió  $\overline{Int}_{PKA/fVHC} = 16,105 \text{ nm}^2$  que está en el mismo orden de magnitud que el valor de  $14,33 \text{ nm}^2$  de área de interfase entre RII $\alpha$ DD y AKAP-IS reportada por Gold et al.[15].

### Distancias entre residuos específicos

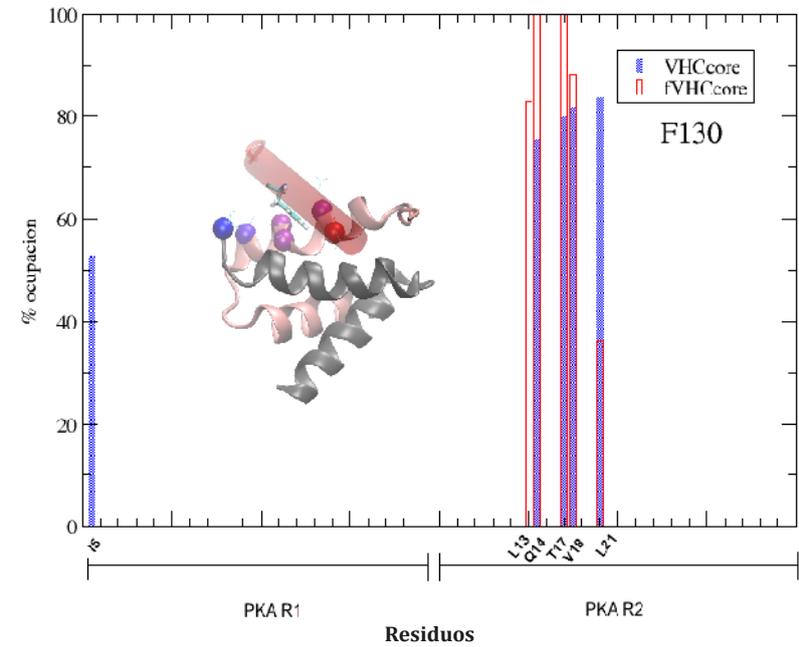
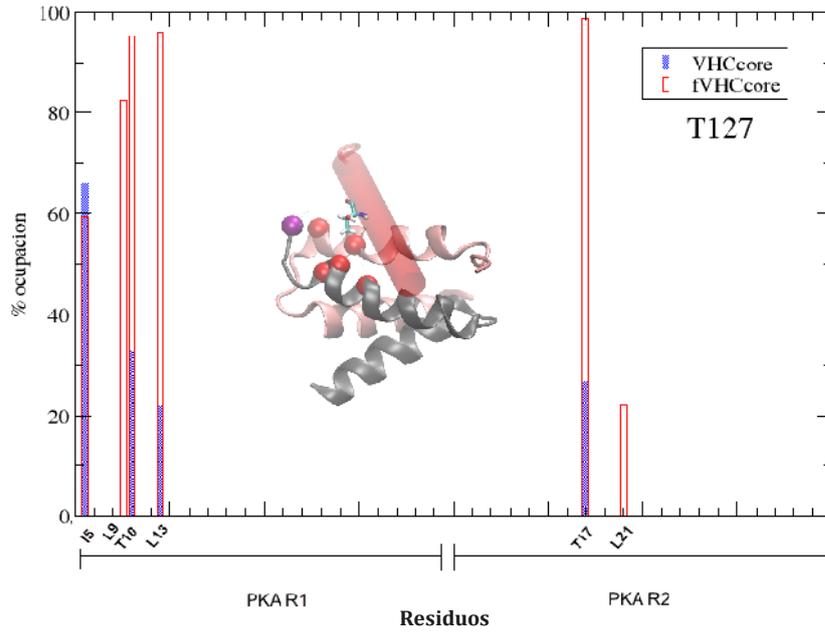
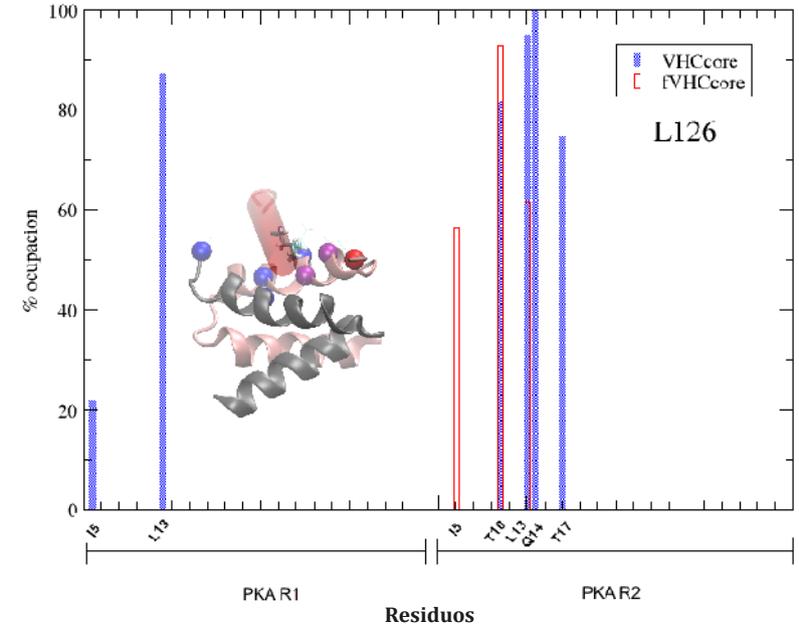
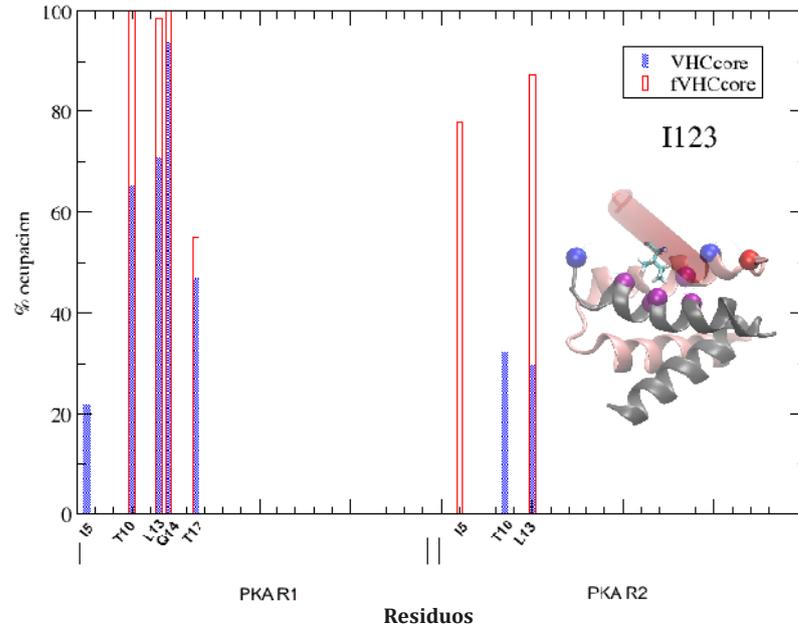
Se analizaron las distancias entre los residuos pertenecientes a las regiones más conservadas de VHCcore y fVHCcore (de acuerdo al alineamiento múltiple en el que se basó el alineamiento estructural, ver sección 5.1) con la finalidad de caracterizar mejor las interacciones existentes entre las secuencias peptídicas estudiadas y PKA. El criterio asumido para poder identificar regiones de interacciones

conservadas a lo largo de las simulaciones contempló la presencia de átomos ubicados a una distancia máxima de 6Å desde el centro de masa de los aminoácidos seleccionados (L119, V122, I123, T126, I127, F130).

Las Figuras 18 A-F representan los los valores de porcentaje de ocupación de las interacciones halladas entre pares de residuos (VHCcore-PKA y fVHCcore-PKA) a lo largo de toda la trayectoria simulada para cada uno de los seis aminoácidos que corresponden a las regiones más conservadas del alineamiento. Para simplificar el análisis, se graficaron sólo



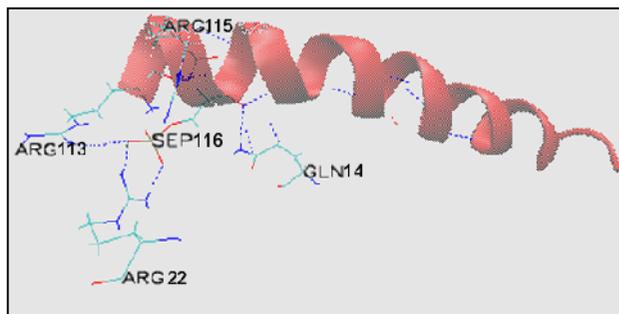
**Figura 18 A-F (continua en página siguiente):** Gráficas de ocupación de interacciones entre pares de residuos fVHCcore//PKA (rojo) o VHCcore/PKA (azul). Los insertos señalan la ubicación de los C-alfa de los residuos de PKA como esferas de VDW siguiendo el mismo criterio de colores. Las esferas púrpuras corresponden a residuos compartidos entre fVHCcore y VHCcore. PKA R1 y PKA R2 señalan la primera y segunda cadena de PKA respectivamente a la que pertenecen los residuos señalados sobre el eje X.



aquellos pares de residuos que aparecen formando interacciones en más del 20% de la trayectoria. Observando la totalidad de las gráficas es posible apreciar que en términos generales las interacciones que involucran a los residuos de fVHCcore (barras con contorno rojo) se encuentran más localizados y con mayor nivel de ocupación a lo largo de la trayectoria que las interacciones entre residuos VHCcore y PKA (barras azules). En el caso de VHCcore se puede interpretar que se pierden aquellos contactos iniciales por interacciones más inespecíficas que se dan más dispersas e inconstantes sobre la superficie de PKA. Esto se ve claramente en el caso de los puntos de contacto entre V122 o L126 en VHCcore y PKA, donde aparecen pares nuevos de interacciones en regiones más alejadas de la zona inicial. Particularmente, es interesante observar la presencia de interacciones que involucran al extremo amino terminal de la segunda cadena de PKA (denotada en las gráficas como PKA R2) ya que este tipo de interacciones están caracterizadas experimentalmente como importantes en la interacción PKA-AKAP[15]. Por otro lado, T10 y Q14 (tanto de R1 como de R2) también participan en interacciones de tipo no hidrofóbicas que han sido identificadas en la estructura cristalográfica de referencia y aparecen en las interacciones RII $\alpha$ /VHCcore y particularmente en RII $\alpha$ /fVHCcore.

### Interacciones de S116 fosforilada

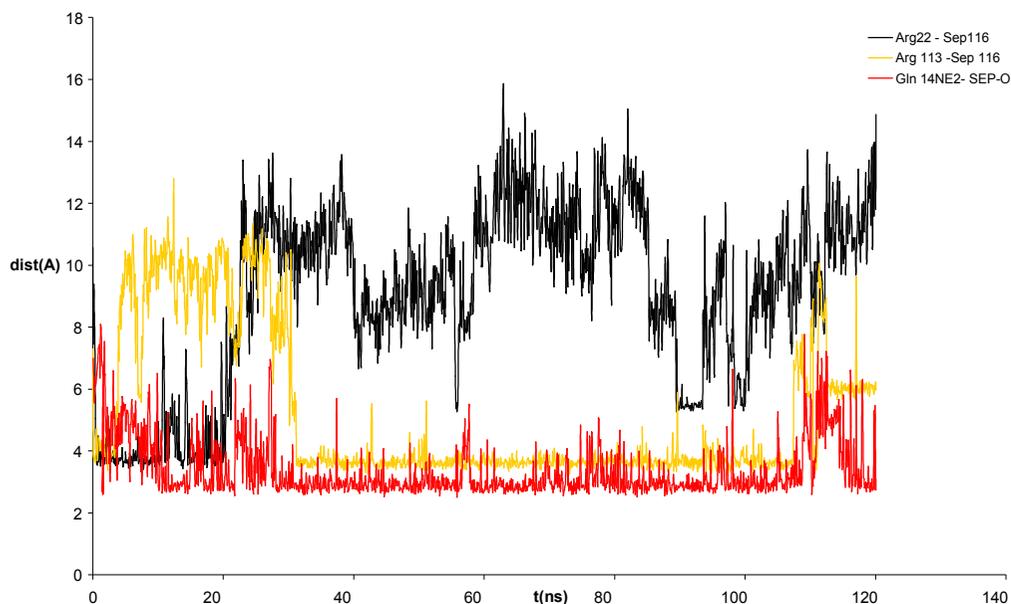
Otro aspecto interesante de analizar en el complejo fosforilado es la estabilidad de las interacciones en las que interviene S116 fosforilada como por ejemplo enlaces de hidrógeno y puentes salinos, poniendo especial atención a aquellos existentes entre el péptido y RII $\alpha$ DD, si bien los existentes intra-AKAP y entre las cadenas de RII $\alpha$ DD son indicadores de estabilidad estructural.



**Figura 19:** Se observan los residuos involucrados en la interacción específica con el grupo fosforilo de la serina fosforilada, así como la interacción entre Q14 de RII $\alpha$ DD y el oxígeno carbonílico del enlace peptídico correspondiente a dicha serina.

En la figura 19 se muestra un ejemplo de la disposición de las interacciones directamente relacionadas. Una vez identificados estos residuos se hizo un seguimiento de las distancias entre los mismos a lo largo de la simulación. Para esto se eligieron seguir las distancias entre el Nitrógeno  $\epsilon$  de la Q14 y el oxígeno carbonílico de S116 fosforilada, para así considerar la alternancia entre los hidrógenos de la glutamina que establecen el puente de hidrógeno.

Del mismo modo las distancias graficadas en la figura 20 siguen el mismo criterio para hacer un seguimiento considerando esa movilidad con las argininas.



**Figura 20:** Gráfica de distancias (Å) entre Sep116 y residuos de interacción vs tiempo de simulación (ns). Las distancias graficadas corresponden a los pares de aminoácidos de acuerdo a la leyenda.

## Mapa de contactos

Para poder tener una noción más concreta de los residuos que participan específicamente en la interacción en estudio se realizó un mapa de contacto para RII $\alpha$ DD/fVHCcore<sup>(\*)</sup> así como para RII $\alpha$ DD/AKAP-IS. El último se usó como patrón de guía a la hora de comparar, ya que AKAP-IS es una proteína AKAP bien conocida. Para ello se utilizó la aplicación g\_mdmat de Gromacs escogiendo como distancia de interacción máxima 8Å. Particularmente interesante es comparar la forma fosforilada con respecto a RII $\alpha$ DD/AKAP-IS como se muestra en la figura 21a. No existen grandes cambios en los contactos que se dan únicamente dentro del dominio RII $\alpha$ DD. Los únicos cambios observados en esa región corresponden a los extremos de las alfa hélice, regiones más móviles de la proteína. Esto se observa en particular con la región amino terminal del segundo fragmento proteico de RII $\alpha$ DD en el mapa de resta de contactos entre fVHCcore y AKAP-IS, se aprecian cambios más localizados que involucran a I123 y L126 de fVHCcore.

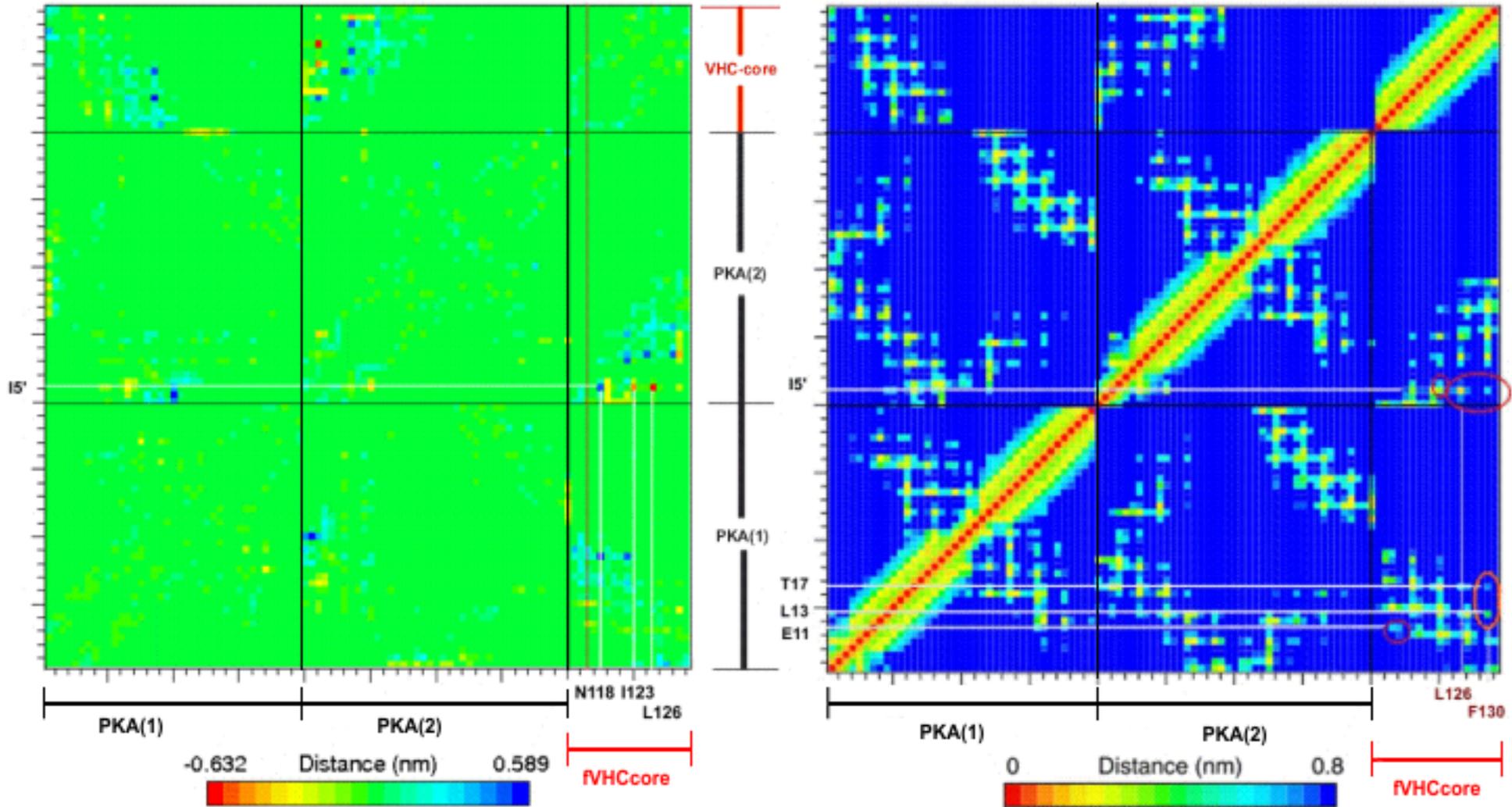
Otro cambio que se observa a nivel de la región amino terminal de la segunda cadena peptídica de RII $\alpha$  se da en la zona de interacción con la primera cadena de PKA, quizás ocasionado como consecuencia de la interacción más estrecha de esta porción con los residuos I123 y L126 de VHCcore y el alejamiento de los contactos equivalentes que se dan naturalmente en la interfaz RII $\alpha$ DD/AKAP-IS entre dicha región de PKA y AKAP-IS.

También se realizó un estudio más minucioso de las interacciones específicas que se dan en el complejo fosforilado con respecto a RII $\alpha$ DD/AKAP-IS. En la figura 21b se muestra el mapa de contacto del complejo RII $\alpha$ DD/fVHCcore. Se marcaron las regiones de RII $\alpha$ DD que intervienen en la interacción con fVHCcore y que no participan con la misma intensidad en la interacción entre RII $\alpha$ DD y AKAP-IS: en rojo se observan las diferencias (i.e: zonas "nuevas" de interacción) y en naranja los aminoácidos que presentan diferencias en cuanto a distancias de interacción del complejo de estudio con respecto a RII $\alpha$ DD/AKAP-IS (es decir aquellas interacciones características en PKA-fVHCcore que no se observan en AKAP-IS).

Por otro lado, en RII $\alpha$ DD/fVHCcore existe también una zona "nueva" de interacciones que involucran al quinto residuo de la segunda cadena de RII $\alpha$ DD (I5') con la fenilalanina del péptido (F130), como se observa en el globo rojo de la figura 21b.

---

(\*) La estructura escogida para realizar el mapa de contacto para RII $\alpha$ DD/fVHCcore consistió en la estructura promedio minimizada considerando la trayectoria desde los 50ns hasta el final.



**Figura 21 a y b:** Resta de mapa de contactos entre RII $\alpha$ DD/fVHCcore y RII $\alpha$ DD/AKAP-IS (izq) donde se señalan los residuos que tienen una interacción más estrecha en el caso de RII $\alpha$ DD/fVHCcore (líneas grises verticales): I123, L126. A la derecha se ve el mapa de contacto que corresponde al complejo RII $\alpha$ DD/fVHCcore donde se muestran las interacciones características entre VHCcore y PKA, que no se conservan en PKA-AKAP-IS. En ambos mapas se trazaron líneas de referencia para diferenciar las zonas que corresponden a la subunidad reguladora (líneas horizontales negras) así como para diferenciar cuando comienza el péptido (línea vertical negra).

## VI. DISCUSIÓN

El mayor desafío en la predicción de la capacidad de funcionar como AKAP en una proteína reside en la amplia variedad a nivel de secuencia de este tipo de moléculas. Esto se debe a que la propiedad más compartida entre ellas radica en la presencia de una estructura de hélice- $\alpha$  anfipática de aproximadamente 18 aminoácidos, un criterio poco específico y tan sensible como para captar las soluciones correctas pero junto con ellas muchos falsos positivos. De cualquier manera, dado el conocimiento estructural de la región de docking y dimerización, se ha mostrado recientemente que un enfoque bioinformático puede ser de utilidad a la hora de identificar un posible dominio de unión a PKA dentro de un candidato a AKAP [20, 47]. Sin embargo, también fueron identificadas AKAPs no canónicas que no interactúan con PKA vía su dominio de docking/dimerización [48] lo que dificulta aún más la búsqueda de este tipo de proteínas, motivo que genera el interés en poseer también información experimental que complemente los resultados de la modelización. En este trabajo se modelaron los complejos basados en los alineamientos múltiple y estructural a partir de una estructura cristalográfica reportada entre un péptido de unión a R-PKA de alta afinidad, por lo que se consideró una interacción AKAP-específica como punto de partida.

Los potenciales electrostáticos calculados pretenden dar una idea más cualitativa del efecto de que genera la fosforilación de la S116. Como se explicó anteriormente, el cambio en el potencial observado en el complejo con el fVHCcore se debe a la carga negativa presente en el fosfato, que neutralizan en gran medida la zona de mayor valor potencial positivo en la región amino-terminal del péptido donde se ubican los residuos arginina (región con la secuencia 113 RRRSR 117). La mayor estabilidad estructural del complejo fosforilado a lo largo de la simulación con respecto al no fosforilado se explica en gran medida por esta disminución de la carga local en esa región del péptido junto con la mayor posibilidad de formar interacciones estabilizantes con los residuos de PKA que se da al estar la serina fosforilada.

Los resultados obtenidos por el enfoque de docking se realizaron con la finalidad de averiguar si la fosforilación favorecía algún tipo de unión con PKA-DD. Si bien es cierto que las soluciones de docking para proteínas no deberían utilizarse ciegamente sin considerar información biológica complementaria, siendo conscientes de las limitaciones del método escogido es posible interpretar cuidadosamente algunos elementos dadas las condiciones escogidas. Aún habiendo ponderado la electrostática más fuertemente a la hora del cálculo no existe evidencia para indicar que VHCcore y fVHCcore interactúen de manera diferente como posible AKAP con PKA-DD ya que no se observan diferencias energéticas concluyentes por este método de estudio. Por este motivo se consideró que el método que no es lo suficientemente sensible para resolver este último

problema planteado. De cualquier manera, el hecho de que se hayan encontrado soluciones que indicaron la posibilidad de una interacción de VHCcore (o su versión fosforilada) con el surco hidrofóbico de PKA, así como la validación con las soluciones obtenidas de AKAP-IS fue decisiva a la hora de resolver cambiar por un método más costoso pero que pudiera permitir un estudio más profundo del efecto de la fosforilación de Ser 116 en la interacción.

Con respecto a la información obtenida de las simulaciones de dinámica molecular, dentro de los análisis generales se puede apreciar que existen mayores fluctuaciones del RMSD en el caso de RII $\alpha$ DD/VHCcore que en RII $\alpha$ DD/fVHCcore. Estas fluctuaciones resultan particularmente notorias al considerar los resultados de la clusterización de la trayectoria (figuras 15 y 16), donde queda en evidencia la mayor estabilidad estructural del péptido fosforilado.

Por otro lado, la información obtenida de la superficie de interfase entre RII $\alpha$ DD y fVHCcore oscila entorno a un valor promedio de  $\overline{Int}_{PKA/fVHC} = 16,11 \text{ nm}^2$  (calculada como el área promedio de la Interfase considerando los datos a partir de los 50ns de simulación). Si bien existen diferencias en la evolución temporal del área de interfase están en el mismo orden que el valor experimental presentado por Gold et al,  $14,33 \text{ nm}^2$ , por lo que el área es similar a la observada en el complejo formado con una AKAP, hecho que es coherente con la hipótesis inicial de que el fragmento estudiado de VHCcore pudiera actuar como AKAP.

Cuando se comparan los resultados obtenidos de las simulaciones con los datos estructurales obtenidos experimentalmente del complejo formado entre la PKA y el péptido de alta afinidad desarrollado por Gold et al[15] llamado AKAP-IS se pueden distinguir interacciones conservadas entre residuos claves de PKA y la proteína de unión a PKA. Estas están constituidas por los residuos Q14 y T10 de PKA donde Q14 interactúa principalmente con L119 e I123 de fVHCcore (y en menor grado también con las argininas o la serina fosforilada) mientras que T10' (T10 de la segunda cadena proteica del dímero RII $\alpha$ ) lo hace con T127 y F130 de VHCcore. Es importante señalar que la primera de estas interacciones involucra al residuo de PKA conocido por tener interacciones dependientes de secuencias AKAP-específicas. Esto explica que este residuo de la PKA también se aprecie como integrante de una interacción AKAP-PKA en la estructura cristalográfica de referencia (PDBID 2IZX) así como en la estructura obtenida por RMN de RII $\alpha$ DD/Ht31 (PDBID: 2DRN) [16] aunque en el caso de estudio la interacción entre Q14 no se da con la cadena lateral expuesta del residuo correspondiente de la AKAP, como sí sucede con AKAP-IS. Como dato adicional, a través de informaciones preliminares obtenidas del grupo del Dr. McLauchlan de la Universidad de Glasgow, se ha sabido que una mutación F130E de VHCcore inhabilita la formación de partículas virales, hecho que resulta interesante ya que en nuestro modelo F130 forma parte de la región de interfase proteína-proteína.

Considerando los resultados obtenidos y contrastándolos con la bibliografía, donde se plantea que la infección con VHC tiende a aumentar la producción de AMPc y de sustratos fosforilados por PKA generando así la activación de PKA (AMPc –dependiente)[24] es posible aventurar una posible interacción de tipo AKAP con la proteína core (p21) de VHC en su conformación fosforilada. Esto podría plantearse por ejemplo en el marco de la localización de PKA en el retículo endoplásmico vía VHCcore, que posee dominios de unión al RE.[49].

## **VII. CONCLUSIONES Y PERSPECTIVAS**

Como primer resultado del trabajo realizado se pudo cumplir con el objetivo inicial de diseñar un modelo del complejo PKA-DD/VHCcore. Un argumento para afirmar esto radica en que las soluciones obtenidas por el método de docking para los complejos son comparables a nivel cualitativo y cuantitativo con la estructura cristalográfica de referencia lo que concuerda con la hipótesis inicial que VHCcore pueda unirse a PKA como AKAP.

Asimismo, los resultados arrojados por los potenciales electrostáticos y las simulaciones señalan una mayor estabilidad en el complejo formado por la forma fosforilada (fVHCcore). Esto permite avanzar un paso más en el estudio propuesto y sugerir que el efecto de fosforilación influiría en la modulación de la interacción de manera de favorecer la formación del complejo.

Teniendo todos los resultados obtenidos en consideración, es posible sugerir una hipótesis donde la fosforilación de VHCcore mediada por PKA pudiera actuar como interruptor en la formación del complejo y así ser un mecanismo novedoso de autorregulación en interacciones PKA-AKAPs. Los resultados del presente estudio serán utilizados para el diseño de nuevos experimentos por parte de colaboradores de nuestro grupo tendientes a esclarecer el rol de la interacción VHCcore-PKA en la infección viral.

## VIII. REFERENCIAS BIBLIOGRÁFICAS

1. Seng-Lai Tan, P., ed. *Hepatitis C Viruses Genomes and Molecular Biology*. 2006: Lilly Research Laboratories, Indianapolis, USA Norfolk (UK): Horizon Bioscience; 2006.
2. Wong, W. and J.D. Scott, *AKAP signalling complexes: focal points in space and time*. *Nat Rev Mol Cell Biol*, 2004. **5**(12): p. 959-70.
3. Mosenden, R. and K. Taskén, *Cyclic AMP-mediated immune regulation -- Overview of mechanisms of action in T cells*. *Cellular Signalling*. **23**(6): p. 1009-1016.
4. Sanderson, J.L. and M.L. Dell'Acqua, *AKAP signaling complexes in regulation of excitatory synaptic plasticity*. *Neuroscientist*. **17**(3): p. 321-36.
5. Rangarajan, P.N., K. Umesono, and R.M. Evans, *Modulation of glucocorticoid receptor function by protein kinase A*. *Mol Endocrinol*, 1992. **6**(9): p. 1451-7.
6. Diviani, D. and J.D. Scott, *AKAP signaling complexes at the cytoskeleton*. *J Cell Sci*, 2001. **114**(Pt 8): p. 1431-7.
7. Torgersen, K.M., et al., *Molecular mechanisms for protein kinase A-mediated modulation of immune function*. *Cellular Signalling*, 2002. **14**(1): p. 1-9.
8. Abel, T. and P.V. Nguyen, *Regulation of hippocampus-dependent memory by cyclic AMP-dependent protein kinase*. *Prog Brain Res*, 2008. **169**: p. 97-115.
9. Schillace, R.V., et al., *A-kinase anchoring in dendritic cells is required for antigen presentation*. *PLoS One*, 2009. **4**(3): p. e4807.
10. Gao, T., et al., *cAMP-Dependent Regulation of Cardiac L-Type Ca<sup>2+</sup> Channels Requires Membrane Targeting of PKA and Phosphorylation of Channel Subunits*. *Neuron*, 1997. **19**(1): p. 185-196.
11. Marx, S.O., et al., *PKA Phosphorylation Dissociates FKBP12.6 from the Calcium Release Channel (Ryanodine Receptor): Defective Regulation in Failing Hearts*. *Cell*, 2000. **101**(4): p. 365-376.
12. Alberts B, J.A., Lewis J, Rafi M, Roberts K and Walter P, *Molecular Biology of the Cell*. 2008, Garland Science.
13. Langeberg, L.K. and J.D. Scott, *A-kinase-anchoring proteins*. *Journal of Cell Science*, 2005. **118**(15): p. 3217-3220.
14. Skalhegg, B.S. and K. Tasken, *Specificity in the cAMP/PKA signaling pathway. differential expression, regulation, and subcellular localization of subunits of PKA*. *Front Biosci*, 1997. **2**: p. d331-42.
15. Gold, M.G., et al., *Molecular basis of AKAP specificity for PKA regulatory subunits*. *Mol Cell*, 2006. **24**(3): p. 383-95.
16. Newlon, M.G., et al., *A novel mechanism of PKA anchoring revealed by solution structures of anchoring complexes*. *EMBO J*, 2001. **20**(7): p. 1651-62.
17. Alto, N.M., et al., *Bioinformatic design of A-kinase anchoring protein-in silico: a potent and selective peptide antagonist of type II protein kinase A anchoring*. *Proc Natl Acad Sci U S A*, 2003. **100**(8): p. 4445-50.
18. Hundsrucker, C., et al., *High-affinity AKAP7delta-protein kinase A interaction yields novel protein kinase A-anchoring disruptor peptides*. *Biochem J*, 2006. **396**(2): p. 297-306.

19. Chang, C.E., et al., *Entropic contributions and the influence of the hydrophobic environment in promiscuous protein-protein association*. Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7456-61.
20. Berrera., Zaccolo., Pantano., *en preparación*.
21. Allan Granoff, R.G.W., ed. *Encyclopedia of Virology*. 2nd ed. Hepatitis C Virus, ed. R.H. Purcell. Vol. 1. 1999, Academic Press. 657-663.
22. Levrero, M., *Viral hepatitis and liver cancer: the case of hepatitis C*. Oncogene, 2006. **25**(27): p. 3834-47.
23. *Hepatitis C Virus Database (HCVdB)*. 2004-2010 Available from: [http://www.hcvdb.org/hcv\\_info.asp](http://www.hcvdb.org/hcv_info.asp).
24. Farquhar, M.J., et al., *Protein kinase A-dependent step(s) in hepatitis C virus entry and infectivity*. J Virol, 2008. **82**(17): p. 8797-811.
25. Shih, C.M., et al., *Modulation of the trans-suppression activity of hepatitis C virus core protein by phosphorylation*. J Virol, 1995. **69**(2): p. 1160-71.
26. Suzuki, R., et al., *Molecular determinants for subcellular localization of hepatitis C virus core protein*. J Virol, 2005. **79**(2): p. 1271-81.
27. Perez, A. and M. Orozco, *Real-Time Atomistic Description of DNA Unfolding*. Angewandte Chemie International Edition, 2010. **49**(28): p. 4805-4808.
28. Karplus, M. and J. Kuriyan, *Molecular dynamics and protein function*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(19): p. 6679-6685.
29. Schlick, T., *Molecular Modeling and Simulation, An Interdisciplinary Guide*. 2nd ed. 2010.
30. Bayly, C.I., et al., *A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model*. The Journal of Physical Chemistry, 1993. **97**(40): p. 10269-10280.
31. Hornak, V., et al., *Comparison of multiple Amber force fields and development of improved protein backbone parameters*. Proteins, 2006. **65**(3): p. 712-25.
32. Leach, A.R., ed. *Molecular modelling principles and applications* 2nd ed. 2001, Pearson Education Limited.
33. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath*. The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
34. *CAPRI: Critical Assessment of PRediction of Interactions*. 2011; Available from: <http://www.ebi.ac.uk/msd-srv/capri>.
35. Ritchie, D.W. and G.J. Kemp, *Protein docking using spherical polar Fourier correlations*. Proteins, 2000. **39**(2): p. 178-94.
36. Guex, N. and M.C. Peitsch, *SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling*. Electrophoresis, 1997. **18**(15): p. 2714-2723.
37. Russell, R.B. and G.J. Barton, *Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels*. Proteins, 1992. **14**(2): p. 309-23.
38. D.A. Case, T.A.D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo., R. C. W. M. Crowley, W. Zhang, K.M. Merz, B.Wang, S. Hayik, A. Roitberg, G. Seabra, I., et al. , *AMBER 10*, S.F. University of California, Editor. 2008.
39. Craft, J.W., Jr. and G.B. Legge, *An AMBER/DYANA/MOLMOL phosphorylated amino acid library set and incorporation into NMR structure calculations*. J Biomol NMR, 2005. **33**(1): p. 15-24.

40. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome*. Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10037-41.
41. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*. J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
42. Hess, B., et al., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*. Journal of Chemical Theory and Computation, 2008. **4**(3): p. 435-447.
43. Herberg, F.W., et al., *Analysis of A-kinase anchoring protein (AKAP) interaction with protein kinase A (PKA) regulatory subunits: PKA isoform specificity in AKAP binding*. J Mol Biol, 2000. **298**(2): p. 329-39.
44. Sarma, G.N., et al., *Structure of D-AKAP2:PKA RI complex: insights into AKAP specificity and selectivity*. Structure. **18**(2): p. 155-66.
45. Lu, W. and J.H. Ou, *Phosphorylation of hepatitis C virus core protein by protein kinase A and protein kinase C*. Virology, 2002. **300**(1): p. 20-30.
46. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features*. Biopolymers, 1983. **22**(12): p. 2577-637.
47. Hou, T., Y. Li, and W. Wang, *Prediction of peptides binding to the PKA RII{alpha} subunit using a hierarchical strategy*. Bioinformatics. **27**(13): p. 1814-21.
48. Diviani, D., et al., *Pericentrin anchors protein kinase A at the centrosome through a newly identified RII-binding domain*. Curr Biol, 2000. **10**(7): p. 417-20.
49. McLauchlan, J., et al., *Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets*. EMBO J, 2002. **21**(15): p. 3980-8.