

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

PROGRAMA DE DOCTORADO EN CIENCIAS BIOMEDICAS

**ANÁLISIS DE LOS CONTACTOS INTERDOMINIO
DE LA PROTEÍNA NS3 DEL VIRUS DEL DENGUE
Y HEPATITIS C.**

T E S I S

QUE PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

Biol. Luis Rosales León



Director de Tesis: Dra Blanca H. Ruiz Ordaz

México, D.F.

Noviembre 2006



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

En primer lugar quisiera expresar mi más profundo agradecimiento a mi tutor, la Dra. Blanca Ruiz, por todo el apoyo que me brindó de manera incondicional durante todos estos años. De ella he aprendido valiosas lecciones académicas y personales. Es una gran satisfacción para mi poder concluir junto a ella el largo camino que concluimos hace 15 años, en la licenciatura...

También deseo agradecer a los miembros de mi comité tutorial: la Dra. Adela Rodríguez y el Dr. Eduardo Horjales, por su apoyo y valiosas contribuciones al proyecto durante mi permanencia en el Doctorado en Ciencias Biomédicas. De igual modo, son valiosas las aportaciones del resto de los miembros de mi jurado para la obtención de grado: los Drs. Abel Moreno, Roberto Arreguín, Enrique Ángeles y Lorenzo Segovia; gracias a la revisión que de este trabajo hicieron, el presente escrito pudo crecer y madurar a su forma actual.

Agradezco a la Dirección de Computo para la Investigación perteneciente a la Dirección General de Servicios de Cómputo Académico, particularmente al Departamento de Visualización, por los recursos brindados tanto para el desarrollo de este proyecto de investigación, así como para la escritura del artículo resultante. Todas las simulaciones de Dinámica Molecular Interactiva se llevaron a cabo en el Observatorio de Visualización Ixtli, con el apoyo de los departamentos de Visualización y Realidad Virtual.

Finalmente, es importante mencionar el apoyo continuo brindado por Eva María Reyes durante el proceso de escritura del artículo, así las valiosas correcciones de Isabel Perez Monfort a los manuscritos en inglés.

Indudablemente, este ha sido un proceso largo y no hubiese sido posible llevarlo a buen término sin el apoyo de un GRAN número de personas:

- Todos los miembros del laboratorio de la Dra Blanca Ruiz: Giovanito, Charlos, Jorge, Pancho López, Vero, Eneas, Irma e indudablemente Mr. Ramachandran!! Ha sido un placer compartir con ustedes, no sólo el espacio de trabajo, también las pequeñas cosas que día a día forman parte del quehacer científico.
- Todos los miembros del DepVis, especialmente: Salim y Eu, por hacer del cubo un lugar habitable y divertido; Barrón, por compartir sus orgasmos culinarios; Raúl por sus silencios y su elocuencia; Memo, Multti, Liz, Elio y la gran tribu de becarios que pululan por el depvis... Sin olvidar a aquellos que por múltiples razones ya no se hallan entre nosotros: Juan Carlos Espino, Marce y Danespin.
- Concha, Carlos, Abigail, Rolo, Adriana, que aun de lejos siguen siendo una parte importante de mis raíces.
- Anta, por compartir conmigo todo y mucho más...
- Mariana y Rodrigo, habitantes de pepinolandia y sobre todo, de mi corazón.
- Fabiola, gracias por el tiempo que dedicaste a la revisión del escrito. Precio infinitamente tu constante apoyo, amor y comprensión para conmigo, especialmente durante estas febriles semanas de escritura, eres un amor!! 🥰
- A mi familia. Su apoyo ha sido vital para continuar todos los días, especialmente los difíciles... Má, Malú, Eury, Circe; soy afortunado de contar con su amor incondicional.

CONTENIDO

AGRADECIMIENTOS

LISTA DE ABREVIACIONES

PREFACIO

RESUMEN

1. INTRODUCCIÓN

1.1 Simulaciones Moleculares

1.1.1 El Movimiento en las proteínas

1.1.2 Dinámica Molecular

1.1.3 Dinámica Molecular Interactiva

1.1.4 Análisis de Modos Normales

1.2 Predicción de Estructura

1.2.1 Reconocimiento del plegamiento (Threading)

1.2.2 Modelaje Molecular basado en Homología

1.3 Familia Flaviviridae

1.4 NS3

1.5 Helicasas

2. PLANTEAMIENTO DEL PROBLEMA

3. OBJETIVOS

4. MATERIALES Y METODOS

4.1 Modelo de la Proteína NS3 del Virus del Dengue

4.2 Dinámica Molecular de scNS3-NS4a

4.3 Análisis de Modos Normales

4.4 Simulaciones de Dinámica Molecular Interactivas

5. RESULTADOS

5.1 Modelo de scNS3-NS2b

5.2 Simulaciones de Dinámica Molecular para scNS3-NS4a y NS3hel

5.3 Análisis de Modos Normales

5.3.1 Análisis de Modos Normales para scNS3-NS4a

5.3.2 Análisis de Modos Normales para scNS3-NS2b

5.4 Simulaciones de Dinámica Molecular Interactivas

6. DISCUSIÓN

8. BIBLIOGRAFÍA

LISTA DE ABREVIACIONES

DENV	<i>Virus del Dengue</i>
fs	<i>Femtosegundos</i>
HCV	<i>Virus de la Hepatitis C</i>
IMD	<i>Dinámica Molecular Interactiva (por sus siglas en inglés)</i>
NMA	<i>Análisis de Modos Normales (por sus siglas en inglés)</i>
ns	<i>Nanosegundos</i>
NS3hel	<i>Dominio de ARN helicasa de la proteína no estructural 3</i>
NS3pro	<i>Dominio de ARN proteasa serínica de la proteína no estructural 3</i>
NS4a	<i>Proteína no estructural 4a, cofactor del dominio de proteasa serínica de la proteína no estructural 3 en el virus de la hepatitis C</i>
NTP	<i>Nucleótido trifosfatado (por sus siglas en inglés)</i>
ps	<i>Picosegundos</i>
RMSD	<i>Desviación del cuadrado de las distancias atómicas entre dos estructuras superpuestas (por sus siglas en inglés)</i>
RTB	<i>Bloque de rotación/translación (por sus siglas en inglés)</i>
scNS3-NS2b	<i>Modelo teórico construido para la proteína NS3 del virus del Dengue, incorpora la región central del cofactor NS2b en una construcción de una sola cadena.</i>
scNS3-NS4a	<i>Modelo cristalográfico de la proteína NS3 del virus de la Hepatitis C, incorpora la región central del cofactor NS4a en una construcción de una sola cadena.</i>

PREFACIO

El presente trabajo muestra los principales resultados de mi proyecto doctoral desarrollado en el Departamento de Biología Molecular y Biotecnología del Instituto de Investigaciones Biomédicas de la Universidad Nacional Autónoma de México. El contenido se encuentra dividido en dos partes, la primera de las cuales constituye una introducción a los temas relacionados con el desarrollo de mi trabajo de investigación. La segunda parte contiene el desarrollo de mi proyecto doctoral, incluyendo la publicación científica que derivó del mismo.

De manera particular, este trabajo busca responder a la pregunta ¿Cuáles son las bases moleculares que determinan las interacciones entre las diversas unidades estructurales de la proteína NS3? Para ello hice uso de diversas herramientas computacionales para el análisis teórico de estructuras, incluyendo Modelaje por homología, Acoplamiento entre proteínas (Docking), Análisis de Modos Normales, Dinámica Molecular y Dinámica Molecular Interactiva. Las principales aportaciones de estos análisis son las siguientes:

1. Los experimentos de dinámica molecular para scNS3-NS4a indican que el dominio de proteasa serínica de NS3 ejerce un efecto estabilizador sobre el dominio de ARN helicasa. Este efecto es más evidente en las asas de unión a NTPs que se encuentran en la interfase entre los subdominios 1 y 2 de NS3hel y no es corresponde al contacto directo entre estas regiones.
2. Durante la simulación de NS3hel en presencia de solvente, es posible observar que el extremo carboxilo terminal, al encontrarse libre de las interacciones que mantiene con NS3pro es capaz de replegarse e integrarse hacia la última helice α , como fue propuesto por Yao *et al.* basado en comparaciones entre las estructuras cristalográficas disponibles para este dominio.
3. El análisis de modos normales indica que son suficientes 3 modos normales de baja frecuencia para describir los cambios conformacionales de NS3 con implicaciones funcionales. De ellos, los primeros 2 (modos 7 y 8) describen reorientaciones de

NS3pro con respecto a NS3hel, mientras que el modo normal 9 describe reorientaciones entre los subdominios de NS3hel.

4. Asimismo, es posible observar una serie de patrones estructurales comunes a los tres modos estudiados, los cuales incluyen: [1] la presencia de regiones de alta flexibilidad, las cuales actúan a manera de “bisagra”. [2] Los cambios de orientación del dominio 2 de NS3hel, se encuentran acoplados a reorientaciones de menor magnitud en el subdominio 3 de la misma. [3] De la misma manera, cambios en la orientación de NS3pro tienen efectos en el subdominio 3 de NS3hel.
5. Los resultados anteriores, en conjunto con los análisis de IMD, sugieren que es necesario que ocurra una reorientación de NS3pro para que el extremo carboxilo terminal de NS3 pueda ser accesible al solvente y así liberar el sitio activo de la proteasa. De los modos normales analizados, el modo normal 8 es el único que maximiza la exposición al solvente de esta región, aunque es posible que los cambios conformacionales entre NS3pro y NS3hel involucren una mezcla de los modos normales 7 y 8.
6. El presente trabajo constituye el primer análisis del comportamiento dinámico de NS3 mediante experimentos *in silico*. Los resultados aquí mostrados contribuyen con nueva evidencia referente a las bases estructurales de los cambios conformacionales y su posible participación en las funciones de la proteasa-helicasa de diferentes miembros de la Familia *Flaviviridae*.

RESUMEN

La proteína no estructural 3 (NS3) de los miembros de la Familia *Flaviviridae* es una enzima multifuncional, compuesta por un dominio de proteasa serínica (NS3pro) así como por un dominio de ARN helicasa (NS3hel). Diversos reportes han mostrado que las actividades presentes en NS3 son indispensables para la replicación viral, lo cual la ha convertido en un blanco importante para el desarrollo de fármacos antivirales. Aunque se sabe que el ciclo replicativo en *Flaviviridae* requiere de una regulación altamente coordinada de todas las actividades asociadas a NS3, a la fecha se desconoce el mecanismo preciso de las interacciones que las regulan. El presente trabajo busca determinar características estructurales comunes entre las proteínas NS3 del virus del Dengue y Hepatitis C, así como caracterizar a los residuos involucrados en la regulación de los movimientos entre NS3pro y NS3hel. De acuerdo con nuestros análisis, NS3pro incrementa la estabilidad del subdominio 1 de la ARN helicasa. Adicionalmente, el comportamiento dinámico del extremo carboxilo terminal apoya la hipótesis que propone que al ser liberado del sitio activo de NS3pro, los residuos involucrados en esta interacción se pliegan hacia la última α -hélice. Mediante el uso de Análisis de Modos Normales, fue posible evaluar los movimientos colectivos de NS3, y observamos que los 2 modos normales de menor frecuencia son suficientes para describir las reorientaciones de NS3pro con respecto a NS3hel. Estos movimientos en NS3pro favorecen un incremento en la exposición del sitio activo de este dominio, lo cual es de gran importancia durante el procesamiento proteolítico del precursor poliproteínico viral. El tercer modo normal de baja frecuencia se encuentra correlacionado con las reorientaciones de los subdominios de NS3hel, en una manera similar al mecanismo de asociación/disociación a ácidos nucleicos propuesto durante la hidrólisis de NTPs y desenrollamiento de ácidos nucleicos de doble cadena. En conjunto, nuestros datos apoyan un modelo dinámico para NS3, en el cual el movimiento y reorientación entre los dominios de proteasa serínica y ARN helicasa pueden resultar en la regulación de las actividades biológicas presentes en esta proteína.

1. INTRODUCCIÓN

1.1 Simulaciones Moleculares

Los ensayos de dinámica molecular (**MD**) en proteínas y ácidos nucleicos, pueden proveer información importante al respecto de los fenómenos dinámicos de los mismos. La versatilidad de este acercamiento es tan amplia que el principio subyacente a las técnicas de MD ha sido aplicado a muchos otros sistemas compuestos de partículas, incluyendo galaxias. La amplia aplicación de las simulaciones de MD ha sido resultado no sólo del incremento en la disponibilidad de poder de cómputo, si no también del desarrollo de nuevas funciones potenciales, cada vez más exactas.

1.1.1 El Movimiento en las proteínas

De acuerdo con algunas estimaciones, para una proteína promedio de levadura es posible esperar más de 20,000 grados de libertad rotacional, los cuales se conocen como modos normales de vibración. En términos generales, un modo normal integra movimientos de un gran número de

coordenadas internas, incluyendo longitudes y ángulos de enlace (Barth & Zscherp 2002). Actualmente, se acepta que estos modos vibracionales incluyen una gran variedad de rearrreglos estructurales que suceden a lo largo de diversas escalas temporales, desde las vibraciones que suceden en los enlaces, los cuales ocurren en la escala de femtosegundos; rotaciones de las cadenas laterales en la superficie de las proteínas, en la escala de los picosegundos; hasta los grandes cambios conformacionales, los cuales suceden en escalas que van desde nanosegundos hasta microsegundos (Hammes-Schiffer 2002; Hammes-Schiffer & Benkovic 2006).

Del mismo modo, las hipótesis más recientes sobre la función catalítica en proteínas asumen que tanto las enzimas como los sustratos poseen cierta flexibilidad conformacional. En otras palabras, se espera que la movilidad inherente a un plegamiento dado se manifieste de manera continua durante las distintas fases de una reacción catalítica. Estudios recientes sobre la relación entre los cambios conformacionales y la catálisis en proteínas como la deshidrogenasa de alcohol hepática (**LADH**) y la dihidrofolato reductasa (**DHFR**) han permitido evaluar el papel de estos cambios conformacionales en la función de estas proteínas (Hammes-Schiffer & Watney 2006). Los resultados indican que, los movimientos térmicos del complejo enzima+sustrato+cofactor ocurren dentro de las restricciones asociadas al tipo de plegamiento; durante estos movimientos, se muestrea el espacio conformacional disponible al complejo, hacia configuraciones que faciliten la catálisis. Es decir, los movimientos en el complejo componen una “red” de movimientos acoplados, en la cual los movimientos en residuos tanto distales como cercanos al sitio activo, permiten acercar y posicionar favorablemente el sustrato y el cofactor, creando un ambiente electrostático óptimo para que ocurra la reacción química (Hammes-Schiffer & Benkovic 2006).

1.1.2 Dinámica Molecular

El concepto de Dinámica Molecular fue desarrollado originalmente por Alder y Wainwright a finales de la década de los 50's como una técnica para la simulación de un sistema de partículas rígidas en colisión (Alder & Wainwright 1957). Éstas fueron seguidas unos años después por simulaciones de sistemas sencillos, como el estudio de Rahman sobre el comportamiento de Argón líquido, utilizando potenciales de Lennard-Jones (Rahman 1964); así como por simulaciones de agua en estado líquido (Stillinger & Rahman 1974). Sorprendentemente, sólo tres años después, el grupo de Martin Karplus publicó la primera simulación de una macromolécula, el inhibidor de la tripsina pancreática bovina (McCammon, Gelin, & Karplus 1977).

El principio de las simulaciones de MD esta basado en la integración de las ecuaciones de movimiento, es decir, se desea reproducir la trayectoria de un sistema compuesto por N átomos, con masa m ; mediante la resolución de las ecuaciones de movimiento de Newton:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i \quad i = 1..N, \quad (ec. 1)$$

donde $\mathbf{r}_i(t)$ corresponde a la posición de la partícula i . En un momento dado, la fuerza \mathbf{F}_i sobre cada átomo puede ser calculada a partir de las interacciones presentes en el sistema y definida como la derivada de la función de energía potencial V que está determinada por las posiciones de todos los átomos (Lindahl 2001).

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r}_1, \dots, \mathbf{r}_N).$$

Una de las implicaciones de este tipo de aproximación, es que la fuerza sobre cada átomo intentará moverlo hasta alcanzar un estado de baja energía potencial. La aceleración resultante de esta fuerza, es una medida del cambio en las velocidades de las partículas en el sistema, y a partir

de este cambio en la velocidad es posible determinar el cambio en la posición de cada átomo un periodo de tiempo después.

Debido a que la integración de las ecuaciones de movimiento depende de las interacciones calculadas para el sistema, la definición de las funciones potenciales o campos de fuerza (forcefields) es un componente de suma importancia en la precisión de los resultados obtenidos. Sin embargo, un potencial que posea una alta exactitud sería demasiado lento para calcular y por lo tanto de utilidad limitada, de modo que diversas aproximaciones son implementadas: En primer lugar las contribuciones se dividen en interacciones entre átomos enlazados (vgr. energías de enlace, ángulos de enlace y rotaciones de enlaces) e interacciones entre pares de átomos no enlazados (Halgren 1995).

El potencial asociado a las variaciones en la longitud de enlace entre un par de átomos i y j es modelada usualmente como un “resorte” armónico simple:

$$V_{bond}(r_{ij}) = \frac{k_{ij}^b}{2} (r_{ij} - r_{ij}^0)^2$$

donde k_{ij}^b representa una constante de fuerza que describe la resistencia de ese tipo de enlace particular y r_{ij}^0 es la longitud de equilibrio del enlace (Lindahl 2001). De manera similar, la deformación del ángulo de enlace entre los átomos i, j, k es descrita como:

$$V_{angle}(\vartheta_{ijk}) = \frac{k_{ijk}^\vartheta}{2} (\vartheta_{ijk} - \vartheta_{ijk}^0)^2$$

donde k_{ijk}^ϑ representa el ángulo de equilibrio, mientras que la constante ϑ_{ijk}^0 representa la resistencia del ángulo a la deformación. Para moléculas complejas, es necesario tomar en cuenta las variaciones en el potencial debidas a la rotación de enlace, la cual es usualmente modelada como un potencial periódico dihedral:

$$V_{dihedral}(\varphi_{ijkl}) = k_{jk}^\varphi [1 + \cos(n\varphi - \varphi_0)]$$

donde la variable n expresa la multiplicidad de mínimos presentes durante una rotación completa y φ_0 su ubicación (Lindahl 2001).

Debido a que estas interacciones corresponden a aquellas presentes entre átomos enlazados, ocurren de manera local en el espacio, por lo cual su número se incrementa de manera lineal con respecto al tamaño del sistema y su cálculo no representa un gran costo computacional. En contraste, existe un número muy grande de interacciones entre átomos no enlazados y el cómputo de estas fuerzas es la parte más costosa de cualquier simulación, representando cerca del 90 – 95 % de uso del procesador, incluso con las simplificaciones que usualmente se integran a los campos de fuerza (vgr. las interacciones se calculan entre pares de átomos y se ignoran las contribuciones energéticas que ocurren a distancias mayores a un umbral definido) (Wang *et al.* 2001). Usualmente, los componentes repulsivo y dispersivo de las interacciones de no enlace se calculan utilizando una variante de los potenciales de Lennard-Jones:

$$V_{Lennard-Jones}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6}$$

donde $C_{ij}^{(6)}$ y $C_{ij}^{(12)}$ son parámetros que dependen del tipo de átomos involucrados, determinando la cantidad de atracción y repulsión, respectivamente (Lindahl 2001). El potencial resultante es ligeramente atractivo a grandes distancias, pero fuertemente repulsivo cuando existe solapamiento entre átomos. La interacción electrostática entre pares de átomos puede ser calculada por la expresión:

$$V_{Coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

donde q_i y q_j representan las cargas, mientras que ϵ_0 y ϵ_r son expresiones de la permisividad del espacio libre y la permisividad relativa.

Una vez que los potenciales para la configuración del sistema a un tiempo t han sido calculados, el siguiente paso consiste en generar una nueva configuración de los átomos a un tiempo $t + \Delta t$ de acuerdo a la dinámica calculada en la ec. 1. Existen diversas aproximaciones para ello, las cuales requieren de una evaluación de la función energética por cada paso (en inglés, timestep). Una de las implementaciones más utilizadas fue desarrollada por Verlet en 1967 (Verlet 1967), basado en la adición y substracción de las expansiones de Taylor para las coordenadas \mathbf{r}_i a los tiempos $t + \Delta t$ y $t - \Delta t$:

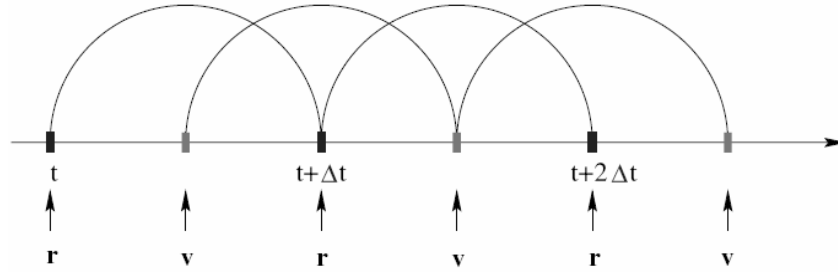
$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \frac{d}{dt} \mathbf{r}_i(t) + \frac{(\Delta t)^2}{2!} \frac{d^2}{dt^2} \mathbf{r}_i(t) - \frac{(\Delta t)^3}{3!} \frac{d^3}{dt^3} \mathbf{r}_i(t) + \mathcal{O}(\Delta t^4)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \frac{d}{dt} \mathbf{r}_i(t) + \frac{(\Delta t)^2}{2!} \frac{d^2}{dt^2} \mathbf{r}_i(t) + \frac{(\Delta t)^3}{3!} \frac{d^3}{dt^3} \mathbf{r}_i(t) + \mathcal{O}(\Delta t^4)$$

lo cual en conjunto con la ec. 1, resulta en:

$$\begin{aligned} \mathbf{r}_i(t + \Delta t) &\approx -\mathbf{r}_i(t - \Delta t) + 2\mathbf{r}_i(t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i \\ \mathbf{v}_i(t) &\approx \frac{1}{2\Delta t} [\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)] \end{aligned}$$

Como resultado de esta implementación, las velocidades son obtenidas como una diferencia de 2 términos de igual orden de magnitud, lo cual hace a los integradores basados en Verlet muy sensibles a la precisión numérica y errores asociados al redondeo (Lindahl 2001). Un segundo integrador derivado de Verlet es el conocido como “Leap-frog”, en el cual las posiciones atómicas y las velocidades son calculadas con diferencia de medio paso o timestep :



(el nombre de este integrador deriva del hecho de que la determinación de posiciones y velocidades difieren por medio paso, como un par de ranas saltando una sobre la otra...) una de las ventajas del esquema “Leap-frog” es que evita las diferencias entre términos de gran magnitud al calcular las velocidades:

$$\begin{aligned} \mathbf{r}_i(t + \Delta t) &\approx \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \Delta t/2) \\ \mathbf{v}_i(t + \Delta t/2) &\approx \mathbf{v}_i(t - \Delta t/2) + \frac{\Delta t}{m_i} \mathbf{F}_i. \end{aligned}$$

Finalmente, algunas de las características observadas en los sistemas moleculares no son atribuibles a los átomos individuales, mas bien son propiedades colectivas del conjunto de átomos, por ejemplo la temperatura y la presión (Mongan & Case 2005). Durante una simulación es importante mantener estas propiedades bajo control, para asegurar que conforme progresa la misma es posible mantenerlas dentro de valores que no desnaturalizan las moléculas bajo estudio.

La temperatura T del sistema es un resultado directo de la energía cinética de todos los átomos y sus grados de libertad N_f :

$$T(t) = \frac{1}{N_f k_B} \sum_{i=1}^N m_i \mathbf{v}_i(t) \cdot \mathbf{v}_i(t).$$

donde k_B representa la constante de Boltzmann (Lindahl 2001; Norberg & Nilsson 2003). Idealmente durante una simulación, la temperatura de un sistema debería permanecer constante, sin embargo debido a los errores ocasionados por el redondeo y las aproximaciones hechas en las

funciones potenciales, es necesario introducir un término λ , el cual permite escalar las velocidades atómicas a cada paso y como resultado, controlar la temperatura del sistema:

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T} - 1 \right)}$$

en la cual T_0 y τ_T corresponden a la temperatura de referencia y a una constante de acoplamiento de la temperatura, respectivamente. De manera similar, la presión del sistema puede ser calculada en base a:

$$\mathbf{p} = 2 \frac{\mathbf{K} - \Xi_{int}}{V}$$

donde \mathbf{K} corresponde a un tensor promedio de la energía cinética y Ξ a la virial del sistema. La presión resultante puede ser controlada mediante un método de acoplamiento similar al empleado para la temperatura:

$$\mu = \sqrt[3]{1 + \frac{\Delta t}{\tau_P} \beta (P - P_0)}$$

donde, para cualquier valor razonable de compresibilidad del sistema β , ocurrirá una relajación exponencial de la presión del sistema (Lindahl 2001).

1.1.3 Dinámica Molecular Interactiva

Como ya observamos en la sección anterior, una de las principales limitantes para la simulación en sistemas moleculares es la restricción de escalas de tiempo reducidas, así como la dificultad de obtener un muestreo estadísticamente representativo del espacio conformacional accesible al sistema de estudio. Con base en lo anterior, en la mayoría de los casos ha sido imposible simular

a nivel molecular muchos de los eventos de importancia biológica. Usualmente las simulaciones de dinámica molecular bajo condiciones de equilibrio se encuentran limitadas al orden de decenas de ns, una escala temporal mucho más corta que muchos de los eventos de interés biológico, tales como cambios conformacionales, permeación de solutos a través de canales, unión y despegado de ligandos a sitios de unión, etc. Como respuesta a estas limitaciones, diversos grupos han implementado herramientas y algoritmos para el desarrollo de una técnica conocida como dinámicas moleculares guiadas. Durante este tipo de simulaciones, el sistema de estudio es guiado hacia un evento particular, mediante la aplicación de una fuerza externa o una función potencial que funciona a manera “de guía” (Grayson, Tajkhorshid, & Schulten 2003). Entre las aproximaciones más comunes a este tipo de simulaciones “fuera del equilibrio”, es posible mencionar los métodos de: Muestreo Sesgado (Umbrella sampling) (McCammon & Harvey 1987), Dinámicas Moleculares Guiadas (steered molecular dynamics, SMD) (Lu *et al.* 1998; Gao, Lu, & Schulten 2002) y Dinámicas Moleculares Interactivas (interactive molecular dynamics, IMD) (Stone *et al.* 2001).

IMD es un método de simulación, el cual permite manipular interactivamente sistemas moleculares, con el fin de probar diversos procesos dinámicos. Metodológicamente, esta aproximación utiliza los mismos campos de fuerza y algoritmos que las aproximaciones clásicas de MD. Sin embargo, en los ensayos de IMD y SMD, se adiciona una fuerza externa al sistema. Esta fuerza es utilizada para investigar las propiedades mecánicas del sistema o para simular procesos que de otra manera requerirían de un tiempo excesivamente largo de simulación mediante un ensayo clásico de MD (Israelewitz, Gao, & Schulten 2001). En contraste con los ensayos de SMD, las fuerzas aplicadas durante un ensayo de IMD son ajustadas continuamente por el usuario, con ayuda de un sistema de visualización de la simulación en tiempo real. La posibilidad de variar la dirección y magnitud de las fuerzas aplicadas, permite al investigador

inducir los cambios conformacionales necesarios para inducir el evento que desea observar (Stone, Gullingsrud, Grayson, & Schulten 2001).

En términos generales, puede decirse que los ensayos de SMD son empleados una vez que el investigador cuenta con una hipótesis que permitirá especificar la posición y dirección de las fuerzas a emplear, las cuales son de magnitud suficiente para permitir al sistema concluir el proceso en un tiempo razonable (normalmente días o semanas de cómputo). El análisis de una trayectoria de SMD permite correlacionar la fuerza aplicada con la evolución del sistema (Israelewitz, Gao, & Schulten 2001). En contraste, los ensayos de IMD tienen como finalidad explorar hipótesis acerca del comportamiento del sistema y los tiempos de simulación suelen ser breves (horas o minutos), por lo cual las fuerzas empleadas tienen la magnitud necesaria para inducir los cambios en dicha escala de tiempo. Por esta razón, las simulaciones de IMD son de carácter cualitativo y las trayectorias resultantes no son adecuadas para análisis numérico.

Como herramienta de investigación, las simulaciones de IMD representan un análisis de gran utilidad, permitiendo al investigador generar y probar hipótesis relativas a mecanismos de reacción, interacción o transporte entre moléculas. Los resultados de este tipo de análisis, aunque básicamente de naturaleza cualitativa, permiten experimentar diversos mecanismos de movimiento e interacción entre moléculas, los cuales por su naturaleza son difíciles de describir únicamente de manera numérica. Una vez establecida la factibilidad de un mecanismo explorado mediante IMD, es posible que el investigador utilice simulaciones no interactivas que permitan caracterizar el mecanismo de manera más fina.

1.1.4 Análisis de Modos Normales

Como se mencionó en las secciones anteriores, debido a la escala temporal en la que ocurren los cambios conformacionales en proteínas, su estudio *in silico* requiere de un conjunto de técnicas computacionales particulares, las cuales permitan acceder a dicha escala de tiempo. En este caso, el uso de las técnicas usuales de dinámica molecular es de poca utilidad, debido a los largos tiempos de cómputo requeridos para muestrear de manera adecuada el espacio conformacional disponible para un sistema de tamaño medio o grande. Por ejemplo, para un sistema cercano a los 100,000 átomos, fue necesaria una semana de tiempo computacional en 32 procesadores de 1.3 GHz, para alcanzar 1 nanosegundo de tiempo de simulación (Tajkhorshid *et al.* 2003). En otro reporte más reciente, la simulación del comportamiento dinámico del virus satélite del mosaico del tabaco, en presencia de solvente explícito, (1,066,628 átomos) se llevó a cabo en 128 nodos Altix duales pertenecientes al NCSA, con un rendimiento de 0.7 ns/día (Freddolino *et al.* 2006).

El análisis de modos normales (o **NMA**, por sus siglas en inglés) es un método teórico para el estudio de los movimientos colectivos en sistemas biológicos. Durante este análisis, el movimiento dinámico de una macromolécula se expresa en términos de una superposición colectiva de variables, conocidas como las coordenadas de los modos normales o vibracionales. En otras palabras, se considera que la energía potencial del sistema se comportará de manera armónica, por lo cual el movimiento dinámico de la molécula puede ser descrito como un conjunto de osciladores armónicos independientes. Los osciladores de baja frecuencia representan movimientos colectivos de gran amplitud que describen las deformaciones más probables para la estructura analizada (Tama 2003).

El objetivo del análisis de modos normales es identificar aquellos movimientos de gran amplitud y baja frecuencia que pueden ser relevantes a la función biológica del sistema estudiado.

En un estudio reciente, el cual incluía más de 3800 cambios conformacionales previamente caracterizados, fue posible observar que más de la mitad de ellos pueden ser modelados aplicando una perturbación en la dirección de 1 o 2 modos normales de baja frecuencia. Más aún, cuando el carácter colectivo de los cambios conformacionales es evidente, un sólo modo normal de baja frecuencia es suficiente para representarlo (Krebs *et al.* 2002).

Las aplicaciones reportadas hasta el momento para esta técnica incluyen: afinamiento de factores B en cristalografía (Kidera & Go 1990;Kidera *et al.* 1992;Kidera & Go 1992;Tirion *et al.* 1995), especificidad enzimática (Ishida, Jochi, & Kidera 1998;Miller & Agard 1999), estudio de cambios conformacionales en proteínas (Harrison 1984;Gibrat & Go 1990;Mouawad & Perahia 1996), complejos macromoleculares (Tama *et al.* 2003) y virus (Tama & Brooks, III 2002), transferencia de electrones (Basu *et al.* 1998a;Basu *et al.* 1998b) y dinámica de ácidos nucleicos (Duong & Zakrzewska 1998;Matsumoto, Tomimoto, & Go 1999).

Inicialmente, los análisis de modos normales estaban restringidos a proteínas de tamaño pequeño, debido a varios factores: Por un lado, los modelos iniciales de NMA, al utilizar potenciales semiempíricos, requerían necesariamente de un paso previo de minimización energética, lo cual resultaba en un incremento en los tiempos de cálculo y deformaciones en la estructura como resultado de que en su mayoría, la minimización energética se llevaba a cabo en el vacío. En 1996 Tirion demostró que el uso de los potenciales semiempíricos puede ser sustituido por una aproximación en la cual las interacciones atómicas son representadas por un modelo simple de red anisotrópica (un potencial hookeano). Para la función energética considerada, la energía potencial del sistema es un mínimo para cualquier configuración de cualquier sistema!! El uso de esta aproximación simplifica los cálculos de modos normales, al tiempo que elimina la necesidad de un paso previo de minimización energética (Tirion 1996).

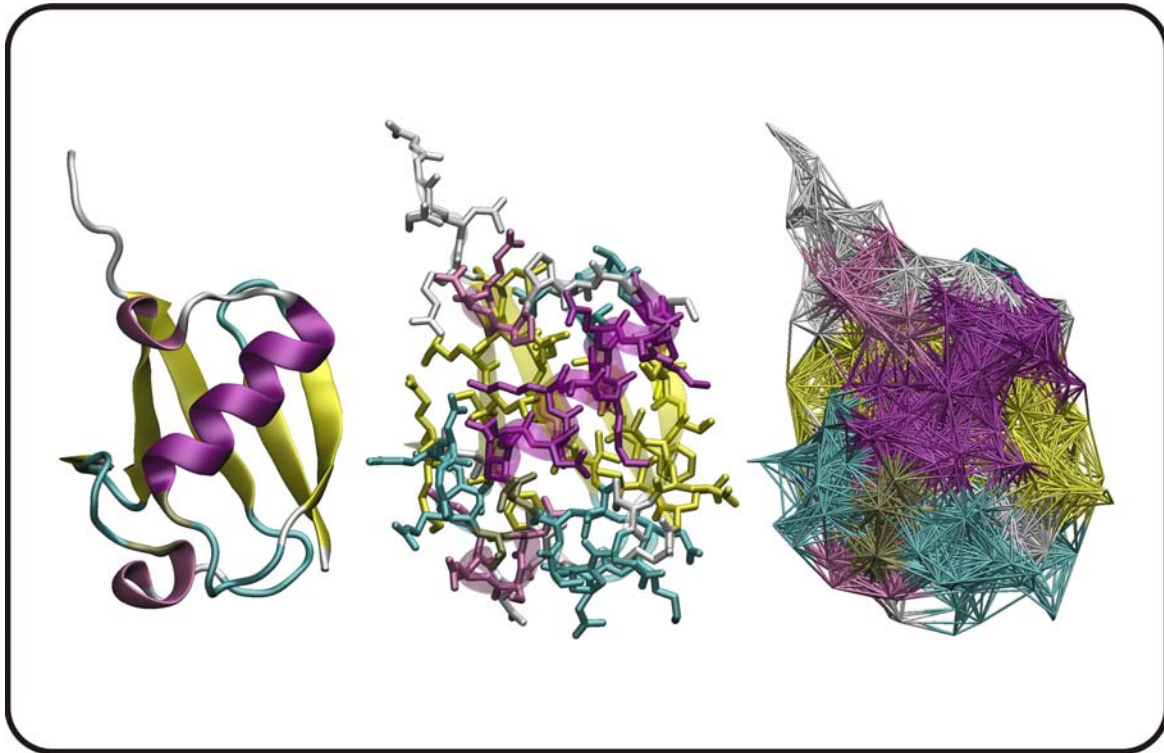


Figura 1 - Modelo de red anisotrópica para la Ubiquitina.

Izquierda: Representación de listón; Centro: Representación de bastón, se ha agregado transparencia al listón, de modo que sea posible identificar a los átomos de la cadena principal; Derecha: Modelo de red anisotrópica, los bastones representan las interacciones atómicas consideradas para un valor umbral de distancia dado (en este caso 5 Å).

Una ventaja adicional en el uso de la implementación de Tirion radica en que permite aplicar las técnicas de NMA a estructuras de baja resolución, a este respecto existen varios reportes en los cuales los modos normales obtenidos para el ribosoma (25 Å) y la ARN polimerasa de *Escherichia coli* (15 Å) son consistentes con la evidencia experimental disponible para dichas estructuras (Frank & Agrawal 2000; Darst *et al.* 2002).

Adicionalmente, en un análisis tradicional de Modos Normales, el paso de minimización energética es seguido por una diagonalización de la matriz Hessiana (la matriz de las segundas derivadas de la función energética), la cual es una matriz de tamaño $3N \times 3N$, donde N representa el número de átomos en el sistema (Tama *et al.* 2000). Debido a que la diagonalización representa otro paso limitante en los ensayos de NMA, se han implementado varias

aproximaciones que buscan simplificar este problema. Perahia *et al.* implementaron una técnica de diagonalización en base mixta, la cual se lleva a cabo una diagonalización iterativa, utilizando soluciones parciales a las derivadas de la función energética (Perahia & Mouawad 1995).

En una aproximación más reciente, un conjunto de átomos o residuos son agrupados en bloques de rotación-translación (RTB). De acuerdo con esta aproximación, los modos normales de baja frecuencia en una proteína pueden ser descritos como movimientos de cuerpos rígidos compuestos por uno o más residuos de aminoácidos (Durand, Trinquier, & Sanejouand 1994). Debido a que en sistemas de gran tamaño, los modos normales de baja frecuencia dependen principalmente de la forma general del sistema (Tama & Sanejouand 2001), esta aproximación tiene un impacto muy pequeño en los modos normales resultantes, al tiempo que permite analizar sistemas biológicos de mayor tamaño, incluyendo a todos los átomos presentes en la estructura de interés (Tama, Gadea, Marques, & Sanejouand 2000).

1.2 Predicción de Estructura

En años recientes, el desarrollo de un gran número de técnicas de alto rendimiento en genómica y proteómica ha permitido la generación de grandes cantidades de información, la cual se encuentra disponible en un creciente número de bases de datos. Un ejemplo de lo anterior se encuentra reflejado en el contenido actual de las principales bases de datos de acceso público: En Noviembre de 2006 el número total de secuencias proteínicas en SwissProt y TrEMBL era de 235,673 y 3,297,108 respectivamente. En contraste con lo anterior, el número de estructuras reportadas para la misma fecha en el Protein Data Bank era de 39,678. En este contexto, nuestra capacidad para obtener información biológicamente relevante a partir de la secuencia de un gen o

una proteína es altamente dependiente de nuestra capacidad para establecer relaciones significativas con las estructuras macromoleculares existentes.

Aunque la solución ideal a este problema sería contar con un modelo o algoritmo capaz de simular el plegamiento de la estructura a partir de una secuencia dada, este acercamiento cuenta con varias limitaciones técnicas que hacen de la predicción de estructura *ab initio* imposible de implementar en un futuro cercano (Finkelstein 1997). Idealmente, un algoritmo para la predicción de estructura debería integrar un modelo energético de gran exactitud, al tiempo que la simplicidad del mismo lo haga computacionalmente factible, una mezcla que aun no ha sido lograda. Este requerimiento para el tratamiento de la energía asociada al plegamiento es de vital importancia, puesto que la función energética debe contar con la suficiente precisión para distinguir al plegamiento nativo como el “mínimo global energético”, entre un gran número de conformeros de baja energía. Estudios realizados por varios grupos han mostrado que las aproximaciones utilizadas en los campos de fuerza actuales no es suficiente para simular el plegamiento de una proteína a su conformación nativa (Mirny & Shakhnovich 1996; Pande & Rokhsar 1999). Sin embargo, una implementación lo suficientemente detallada de las interacciones que producen el estado nativo, incrementaría enormemente los requerimientos computacionales de un algoritmo, volviéndolo impráctico.

Debido a las limitaciones inherentes al problema de predicción de estructura *ab initio*, se han implementado distintas metodologías, las cuales buscan integrar a los modelos la información asociada a las estructuras conocidas. Actualmente existen 2 grandes métodos para la predicción de estructura basada en integración de la información estructural disponible: [1] Reconocimiento del Plegamiento (Threading) y [2] Modelaje Basado en Homología.

1.2.1 Reconocimiento del plegamiento (Threading)

La idea fundamental en los métodos de threading es la de disminuir el número de conformaciones accesibles a la secuencia problema, limitándolo a un subgrupo más pequeño, el cual está integrado por una base de datos de plegamientos conocidos, que funcionan a manera de “molde” y sobre los cuales se superpone esta. Finalmente es necesario encontrar el alineamiento energéticamente óptimo entre este subgrupo (Finkelstein & Reva 1991; Reva, Finkelstein, & Topiol 2002).

Aunque los métodos de threading dependen de manera importante de las funciones energéticas implementadas en ellos (de manera similar al problema de predicción de plegamiento *ab initio*), al restringir el espacio conformacional accesible a la secuencia blanco, es posible utilizar funciones energéticas menos precisas para encontrar a la conformación nativa. Sin embargo, este tipo de metodología cuenta con limitaciones importantes. Por ejemplo, si la secuencia problema pertenece a una familia nueva, no representada en la base de datos que integra el grupo de “conformaciones accesibles”, el algoritmo será incapaz de encontrar una estructura energéticamente favorable (Mirny & Shakhnovich 1998).

1.2.2 Modelaje Molecular basado en Homología

Las técnicas de Modelaje Molecular basado en homología parten de la observación de que, si una secuencia de estructura desconocida comparte una identidad de secuencia significativa con una estructura conocida, las dos proteínas compartirán un plegamiento similar, de modo que ésta última puede ser utilizada a manera de molde para construir un modelo de la secuencia problema (Chothia & Lesk 1986). Debido a lo anterior, la calidad del modelo resultante dependerá esencialmente del porcentaje de similitud entre ambas proteínas. Si tenemos en cuenta que para

proteínas con identidades $\geq 50\%$ el RMSD del 90 % de los carbonos α es $\leq 1 \text{ \AA}$, es de esperar que un modelo construido bajo estas condiciones de identidad tenga un alto grado de exactitud en la predicción. Por otro lado, cuando el porcentaje de similitud es cercano al 20 %, el RMSD de las regiones conservadas (aproximadamente 50 % de los residuos) será $\geq 1.8 \text{ \AA}$. Un modelo dentro de este rango de similitud será no sólo técnicamente más complicado de construir, el resultado será de una precisión mucho menor (Tramontano 1998; Tramontano 2003). El modelaje por homología es dependiente de la experiencia previa del investigador, siendo posible en algunos casos obtener predicciones con buen nivel de exactitud con porcentajes de identidad tan bajos como 14% (Venclovas 2003).

Es importante señalar que la exactitud final del modelo no será homogénea (vgr, existiran regiones de mayor/menor exactitud) y esto debe tomarse en cuenta al evaluar los posibles usos y conclusiones que pueden derivarse del mismo. Sin embargo, incluso para modelos construidos en condiciones de baja similitud, es posible evaluar la confiabilidad de sus diferentes partes y tomar esto en cuenta al derivar conclusiones del mismo.

Durante un experimento de Modelaje basado en homología, se debe determinar inicialmente que proteína puede ser utilizada como molde y aunque el factor de mayor importancia será el porcentaje de identidad de secuencia, cuando es posible elegir entre varios moldes con similitud equivalente, es necesario tomar en cuenta otros factores como la resolución, presencia de ligandos, etc.

Una vez determinada la estructura que servirá como molde, se lleva a cabo un alineamiento de secuencia entre ambas proteínas, de modo que posiciones equivalentes se encuentren correctamente alineadas. Dado que sólo existe información estructural para la proteína molde, la asignación de regiones equivalentes debe ser inferida únicamente en base a las

secuencias (D'Alfonso, Tramontano, & Lahm 2001). Durante esta etapa es muy importante revisar el alineamiento para verificar la ausencia de gaps (los cuales representan eventos de inserción/delección de residuos) al interior de elementos de estructura secundaria o en el núcleo de la estructura. Debido a que el alineamiento final determinará el proceso de asignación de coordenadas (vgr. la construcción del modelo) esta etapa reviste de gran importancia y por lo tanto es recomendable utilizar toda la información disponible de diversas fuentes (datos experimentales, información de residuos funcionales, etc) (Tramontano 1998).

Cuando se ha determinado el “alineamiento correcto”, es necesario asignar coordenadas a todos los átomos de la secuencia blanco, para ello es necesario posicionar los átomos de la cadena principal, las cadenas laterales y las zonas de inserciones/delecciones (las cuales son regiones sin correspondencia entre el molde y el modelo). Dependiendo del programa utilizado durante esta etapa, es posible emplear diversas estrategias para la construcción de los loops que presentan indels, incluyendo la construcción basada en reglas heurísticas (Sibanda, Blundell, & Thornton 1989), modelaje *ab initio* (Moult & Unger 1991) y satisfacción de restricciones espaciales (Sali & Blundell 1993). De manera similar el posicionamiento de las cadenas laterales puede estar basado en el uso de diversos acercamientos, como el uso de bibliotecas de rotámeros (Dunbrack, Jr. & Karplus 1993; Dunbrack, Jr. & Karplus 1994), etc.

Finalmente, el modelo construido debe ser afinado para eliminar superposiciones estéricas y optimizar su estereoquímica, usualmente mediante el uso de minimización de energía y dinámica molecular. No hay que olvidar que el proceso de afinamiento no aumentará la calidad final del modelo (la cual depende, como vimos, del porcentaje inicial de similitud y del correcto alineamiento entre ambas proteínas), sólo mejorará la estereoquímica del mismo (Tramontano 1998).

El programa elegido para el modelaje de la proteína scNS3-2b en el presente trabajo fue Modeller, el cual implementa la construcción de modelos moleculares mediante la satisfacción de restricciones espaciales. De manera breve, el proceso inicia a partir del alineamiento de secuencia entre la estructura que servirá de molde y la proteína problema. A partir de este alineamiento, Modeller genera una lista de restricciones compuesta por distancias de enlace y ángulos dihedrales, cuyos valores son obtenidos de una base de datos de alineamientos estructurales. Durante la etapa de afinamiento, las restricciones espaciales, en conjunto con las funciones potenciales derivadas del campo de fuerza CHARMM son tomadas en cuenta para asegurar la correcta estereoquímica del modelo. Ambos conjuntos de datos son combinados en una “función objetiva”, la cual será optimizada, utilizando técnicas de minimización de energía, MD y recocido (simulated annealing) (Sali & Blundell 1993). Entre las ventajas del uso de Modeller, se puede mencionar que minimiza la intervención del usuario durante el afinamiento del modelo, al tiempo que conserva las características estereoquímicas propias del tipo de plegamiento a modelar.

1.3 Familia Flaviviridae

De acuerdo con el Comité Internacional de Taxonomía de virus, la familia *Flaviviridae* está compuesta por más de 80 miembros pertenecientes a 3 géneros (Flavivirus, Pestivirus y Hepacivirus) y un grupo aún no clasificado. Un gran número de miembros de esta familia son importantes patógenos humanos, causantes de enfermedades importantes como fiebres, encefalitis y fiebres hemorrágicas. Algunos de los virus más importantes dentro de este grupo son: Virus de la Fiebre Amarilla, Virus del Dengue (**DENV**), Virus de la Encefalitis Japonesa (**JAENV**) y el Virus de la Hepatitis C (**HCV**).

Las características comunes a esta familia viral incluyen la presencia de una envoltura lipídica rodeando la nucleocápside, un genoma de ARN de polaridad positiva, así como estrategias de replicación similares. Sin embargo, aunque los miembros de cada género pueden ser relacionados entre sí serológicamente (por anticuerpos que presentan reacción cruzada), no es posible detectar esta relación entre los distintos géneros. Algunas de las propiedades más importantes de los viriones pertenecientes a esta familia se ilustran en la Tabla 1.

Los mecanismos de replicación comunes a los miembros de *Flaviviridae* incluyen la unión a receptores y la internalización de los viriones a través de endocitosis mediada por receptores específicos para las proteínas de superficie (Ishak, Tovey, & Howard 1988; Ramos-Castaneda *et al.* 1997; Bielefeldt-Ohmann 1998). La fusión de la envoltura con las membranas celulares depende de un decremento en el pH de las vesículas, lo cual provoca un cambio conformacional en las proteínas de la envoltura viral, permitiendo la exposición de un dominio de fusión a membranas (Rey *et al.* 1995; Bressanelli *et al.* 2004). Una vez fusionada la membrana viral a la vesícula celular, se libera la nucleocápside dentro del citoplasma, donde ocurre la liberación y replicación del genoma viral (Gollins & Porterfield 1985; Gollins & Porterfield 1986; Kimura, Gollins, & Porterfield 1987).

Tabla 1. Características generales de la Familia Flaviviridae.

	Propiedades
Morfología	<p>Virus esféricos envueltos</p> <p>Diámetro de 40-60 nm</p> <p>Una o dos proteínas de envoltura presentes en la membrana, una proteína de membrana.</p> <p>Una proteína formando la nucleocápside</p>
Propiedades fisicoquímicas	$S_{20w} = 140-200$

	<p>$\rho = 1.1-1.23 \text{ g/cm}^3$.</p> <p>Contenido lipídico = 15-20 %</p> <p>Viriones sensibles a calor, pH, solventes orgánicos y detergentes.</p>
Genoma	<p>ARN de polaridad positiva, una sola copia presente por virión.</p> <p>El tamaño del genoma varía entre 9 y 11 kb</p> <p>El genoma es capaz de funcionar como un ARN mensajero con un sólo marco abierto de lectura.</p> <p>El genoma carece de un tracto de poli(A) en su extremo 3'.</p>
Otros	<p>Flaviviridae es la única familia dentro del supergrupo 2 de virus de ARN de cadena sencilla que codifican para una helicasa de ARN.</p>

El ARN genómico es capaz de funcionar como ARN mensajero durante la traducción, las proteínas virales se sintetizan como parte de un precursor poliproteínico de más de 3000 residuos de longitud, compuesto por cerca de 10 proteínas y el cual contiene en su extremo amino terminal las proteínas estructurales, seguido de las proteínas no estructurales (NS, por sus siglas en inglés), las cuales no forman parte de la partícula viral. Para todos los géneros, es posible encontrar firmas características de proteasa serínica, ARN helicasa y ARN polimerasa dependiente de ARN en la región de proteínas NS (Miller & Purcell 1990). La poliproteína es posteriormente procesada por una combinación de proteasas tanto virales, como del hospedero y se cree que las proteínas no estructurales, una vez liberadas del resto del precursor, se integran en un complejo conocido como RNA replicasa. Este tipo de complejos ha sido localizado en compartimientos celulares similares en los diferentes miembros tanto del género *Flavivirus* (NS1, NS2a, NS2b, NS3, NS4a, NS5a, NS5b) (Ishido, Fujita, & Hotta 1998; Cui *et al.* 1998) como para el género *Hepatitis C* (NS2 NS3, NS4a, NS5b) (Chen *et al.* 1997; Sardana *et al.* 1999), lo cual sugiere que dichos complejos se conservan a lo largo de la familia (Kapoor *et al.* 1995). Asimismo, se ha

demostrado que las interacciones de las diversas proteínas no estructurales con proteínas del hospedero pueden jugar un papel importante en la localización de estos complejos en la célula infectada. De esta manera, es posible imaginar que un paso importante durante el establecimiento de la replicación viral incluye la regulación de una o más proteínas a través de la competencia por sitios comunes de unión (Johansson *et al.* 2001b).

Para todos los géneros de la familia Flaviviridae, la replicación ocurre en la región de membranas perinucleares, mediante la síntesis de un intermediario replicativo de polaridad negativa. La morfogénesis viral se lleva a cabo en las vesículas intracelulares, donde ocurre el transporte y la maduración de las glicoproteínas que componen la envoltura viral y las proteínas que más tarde formarán la nucleocápside. La progenie viral es ensamblada en el interior de vesículas intracelulares, las cuales siguen la vía secretoria para fusionarse con la membrana plasmática y liberar los viriones maduros hacia el exterior de la célula.

1.4 NS3.

La proteína no estructural 3 (NS3, por sus siglas en inglés) es una molécula multifuncional involucrada tanto en el procesamiento del precursor poliproteínico como durante la replicación del genoma viral (Sardana *et al.* 2002). Inicialmente, las actividades enzimáticas asociadas a NS3 fueron propuestas con base en los patrones de conservación de residuos y alineamientos múltiples con proteínas celulares homólogas (Gorbalenya *et al.* 1989). Esta proteína cuenta con un dominio de proteasa serínica (**NS3pro**) localizado en su primer tercio, además de un dominio de ARN helicasa/NTPasa (**NS3hel**) ubicado en los dos tercios restantes. Diversos estudios han demostrado que tanto los dominios, como las actividades asociadas a ellos, son indispensables para la

replicación viral en diversos miembros de la familia (Kolykhalov *et al.* 2000); Matusan *et al.* 2001).

NS3pro es responsable del procesamiento del precursor poliproteínico en los límites NS3/NS4a, NS4a/NS4b, NS4b/NS5a y NS5a/NS5b en el género *Hepacivirus* (Bartenschlager *et al.* 1993; Grakoui *et al.* 1993a; Grakoui *et al.* 1993b; Reed, Grakoui, & Rice 1995) y en las uniones virC/anchC, NS2a/NS2b, NS2b/NS3 NS3/NS4a, NS4a/NS4b y NS4b/NS5 en el género *Flavivirus* (Rice, 1996 in Fields; (Stocks & Lobigs 1998). Asimismo existe evidencia experimental que indica que la actividad de NS3pro en ambos géneros se incrementa mediante la asociación con un factor proteínico esencial para la maduración de la proteína viral (Falgout *et al.* 1991; Bartenschlager *et al.* 1995) (NS4a en hepacivirus y NS2b en flavivirus). En ambos géneros, se ha demostrado que este cofactor proteínico se asocia al dominio de proteasa serínica mediante la inserción de una hebra β , convirtiéndose efectivamente en una parte integral del plegamiento de la proteasa viral (Yan *et al.* 1998; (Kim *et al.* 1996).

Debido a que existen varias estructuras reportadas para NS3pro derivado de HCV, ha sido posible establecer una comparación de las diferencias estructurales entre la proteasa, tanto en forma aislada como en complejo con un péptido derivado de NS4a. Las diferencias presentes sugieren que el mecanismo mediante el cual opera la activación dependiente de este cofactor es mediada por una serie de rearrreglos locales de la triada catalítica, hacia una conformación más favorable a la catálisis (Yan *et al.* 1998). Estudios recientes de modelaje molecular en el virus del dengue han mostrado que este mecanismo de activación de NS3pro es posible también en el género *Flavivirus* (Niyomrattanakit *et al.* 2004; Chanprapaph *et al.* 2005).

NS3hel es responsable de la separación de los duplex de ARN en cadenas sencillas durante la replicación viral. A pesar de la conservación de múltiples motivos de secuencia en ambos

dominios, existen varias diferencias importantes entre NS3 de los diversos géneros, incluyendo la presencia de una actividad de ARN 5' trifosfatasa (RTPasa) presente exclusivamente en el dominio de helicasa en el género *Flavivirus* (Bartelma & Padmanabhan 2002), así como diferencias en los mecanismos de activación de NS3pro, especificidad y cofactores asociados.

En 1999 Yao *et al.* determinaron la estructura de una construcción monocatenaria en la cual 13 residuos pertenecientes al cofactor NS4a (Fig. 2, listón rosa) preceden al extremo amino terminal de NS3 (**scNS3-NS4a**). En esta estructura NS3pro (Fig. 2, listones azul y azul claro) muestra un plegamiento de tipo “doble barril β ” característico de proteasas serínicas homólogas a quimotripsina. NS3hel está compuesto por tres subdominios (Fig. 2, listones verde, amarillo y rojo, respectivamente), de los cuales los 2 primeros presentan un plegamiento de tipo β - α - β y se encuentran estructuralmente relacionados entre sí y con otras helicasas; el tercer subdominio de NS3hel está formado principalmente por hélices α y presenta una estructura única a las helicasas de hepacivirus. Una de las características más sorprendentes de la estructura de scNS3-NS4a, es el arreglo espacial entre NS3pro y NS3hel, los cuales están conectados por una cadena interdominio en conformación extendida, mientras que el sitio activo de la proteasa se encuentra orientado hacia la interfase formada por ambos dominios. Por su parte, tanto los sitios de unión a NTP's, como de unión a ARN de cadena sencilla de NS3hel se encuentran orientados lejos de la interfase, expuestos al solvente. Debido a que no existe evidencia del procesamiento proteolítico de NS3 *in vivo*, se ha propuesto que son necesarios cambios locales y globales durante el procesamiento proteolítico del enlace peptídico NS3/NS4a (el cual es un evento intramolecular, también llamado corte en cis), así como durante los subsecuentes cortes mediados por NS3pro (cortes intermoleculares, también llamados en trans) (Yao *et al.* 1999).

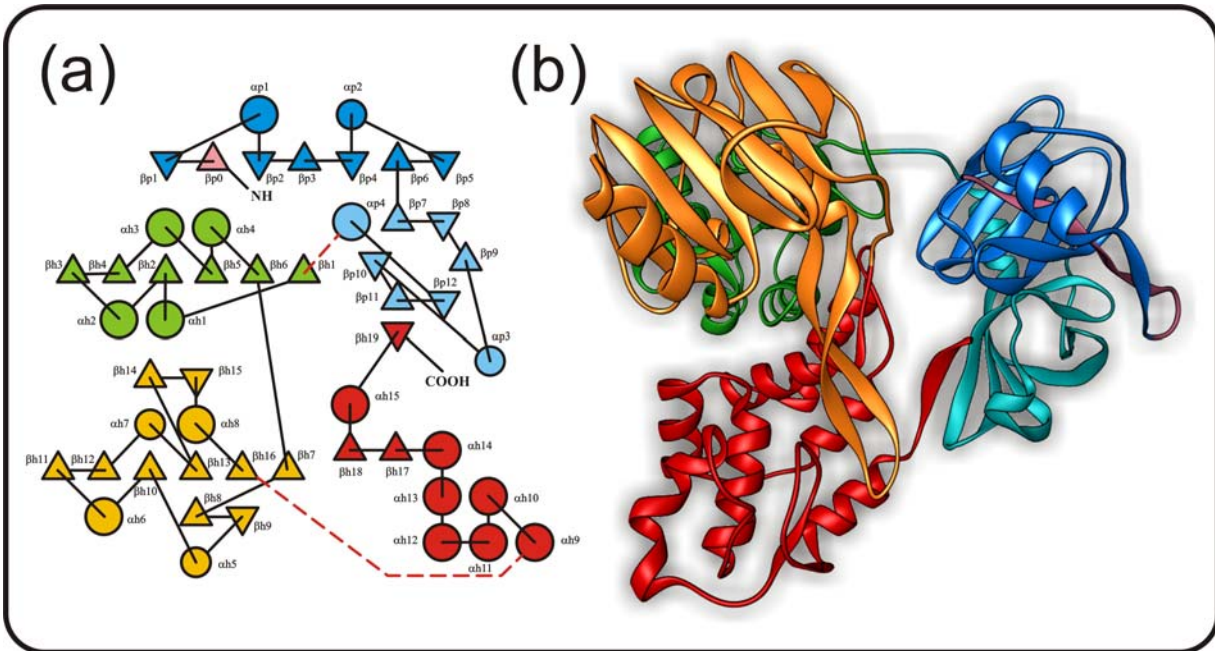


Figura 2 - Organización y Estructura de la proteína scNS3-NS4a de HCV.

(a) Organización de los diversos elementos de la estructura secundaria. Los triángulos representan hebras β y los círculos helices α . (b) Estructura tridimensional de la proteína scNS3-NS4a. Tanto en (a) como en (b) se emplea el mismo código de colores: NS4a (rosa), primer y segundo subdominio de NS3pro (azul y azul claro, respectivamente), NS3hel, primero, segundo y tercer subdominios (verde, oro y rojo, respectivamente).

Aunque inicialmente fue posible demostrar que cada uno de los dominios que componen a NS3 presentan independencia estructural y funcional, actualmente existe un gran número de evidencias que señalan la importancia de las interacciones entre ambos dominios en el polipéptido íntegro *in vivo*. En 1995, un estudio por Kanai *et al.* Reporta que la K_d de unión de poli(U) a scNS3-NS4a íntegra es 10 veces menor que aquella del dominio de helicasa aislado, lo cual sugiere que NS3pro es requerido para una unión óptima de ácidos nucleicos al dominio de ARN helicasa (Kanai, Tanabe, & Kohara 1995). Posteriormente, Morgenstern *et al.* reportaron que, en el complejo NS3-NS4a aislado a partir de células COS transfectadas, la actividad de dominio de proteasa es capaz de ser estimulado 5 veces sobre los controles mediante la adición de poli (U), lo cual sugiere una interdependencia entre las actividades de ARN helicasa y proteasa serínica (Morgenstern *et al.* 1997). En contraste con lo anterior, Gallinari *et al.*

reportaron que para NS3 asociado a un péptido sintético de NS4a, la adición de poli(U)/oligo(U)₁₈ resulta en la inhibición de la actividad de proteasa (Gallinari *et al.* 1998). Debido a que el dominio de proteasa serínica aislado presenta un decremento similar en su actividad, es posible que esta inhibición mediada por ARN fuera causada por una interacción directa con este dominio. Sin embargo, este mismo estudio reporta que la estimulación en la actividad de ATPasa asociada a la presencia de poli(U) es mayor en NS3 íntegra, con respecto al dominio de ARN helicasa aislado.

Recientemente, Johansson *et al.* encontraron que el dominio de ARN helicasa es capaz de modificar la inhibición mediada por producto en el dominio de proteasa serínica. De manera sorprendente, Johansson reporta que péptidos de menor tamaño son capaces de tener un efecto inhibitorio mayor en NS3 íntegra, comparada con el dominio de proteasa aislado (Johansson *et al.* 2001a). De manera similar, Drouet *et al.* reportaron que tanto el dominio de proteasa serínica aislado, así como el polipéptido completo, son capaces de interactuar con algunos inhibidores de la familia de las serpinas. Sin embargo, sólo NS3 es capaz de dar lugar a la aparición de productos de reacción de alto peso molecular estables (Drouet *et al.* 1999). De manera análoga, el estudio de Gallinari *et al.* muestra que la actividad de ATPasa en el polipéptido completo presenta un incremento en la estimulación causada por la adición de poli(U), cuando el efecto se compara con la estimulación del dominio de helicasa aislado (Gallinari *et al.* 1998).

1.5 Helicasas.

Las helicasas son proteínas cuya función consiste en separar las cadenas de una doble hélice de ácido nucleico mediante un mecanismo dependiente del consumo de NTPs (Soultanas & Wigley 2001). Debido a que una gran parte de la información codificada dentro de los ácidos nucleicos es

accesible únicamente cuando estos se encuentran en forma de cadena sencilla, una gran cantidad de procesos celulares y virales, incluyendo transcripción y recombinación, requieren de la presencia de estas ubicuas proteínas. A manera de ejemplo, en el ser humano se han detectado cinco helicasas homólogas a la ADN helicasa RecQ de *Escherichia coli*, todas ellas relacionadas en eventos de recombinación. De igual manera, se han asociado diferentes enfermedades hereditarias con mutaciones en genes que codifican para helicasas, incluyendo los síndromes de Bloom, Werner y Rothmund-Thomson, los cuales se encuentran caracterizados por inestabilidad genómica y una elevada predisposición al desarrollo de cáncer (Nakayama 2002).

Debido a que las helicasas son una parte fundamental durante el ciclo replicativo de muchos virus, han sido consideradas como un blanco atractivo para el desarrollo de compuestos antivirales. De hecho, para un gran número de familias virales ha sido posible demostrar que la ausencia de una helicasa funcional es capaz de interferir con la producción de progeñie (van Dinten *et al.* 2000;Gu *et al.* 2000;Matusan *et al.* 2001).

Todas las helicasas funcionan *in vivo* como parte integral de maquinarias macromoleculares dedicadas a la manipulación de ácidos nucleicos como parte de eventos que forman parte del ciclo celular, tales como replicación, recombinación, reparación, transcripción y splicing. La función biológica de las helicasas, más allá de la disociación de ácidos nucleicos de doble cadena en cadenas sencillas, se encuentra dirigida a la transferencia de dichas cadenas sencillas hacia la siguiente proteína del complejo macromolecular en el cual actúa (Delagoutte & von Hippel 2002).

El análisis de secuencias de helicasas ha permitido la identificación de una serie de motivos conservados, en base a los cuales se ha propuesto la clasificación de estas proteínas en 3 grandes Superfamilias (Gorbalenya & Koonin 1993). Las helicasas de las superfamilias 1 y 2 contienen 7 motivos conservados, los cuales a pesar de encontrarse dispersos a lo largo de la

secuencia, se encuentran agrupados en la estructura tridimensional de estas proteínas, formando un hueco de unión a NTPs y una parte del sitio de unión a ácidos nucleicos (Hall & Matson 1999).

En términos generales, las helicasas muestran muy poca especificidad hacia la secuencia de su sustrato *in vitro*, los datos derivados de las estructuras determinadas en presencia de ácidos nucleicos, muestran que las interacciones entre ambos involucran principalmente apilamiento de bases con cadenas laterales de residuos aromáticos e interacciones con el esqueleto azúcar-fosfato (Kim *et al.* 1998; Lin & Kim 1999; Caruthers, Johnson, & McKay 2000).

2. PLANTEAMIENTO DEL PROBLEMA

El ciclo replicativo en la Familia *Flaviviridae* requiere de una regulación coordinada de todas las actividades presentes en NS3. Sin embargo, con base a que esta proteína presenta un arreglo estructural único entre sus dominios, la naturaleza exacta de estas interacciones es desconocida. Por lo que la información estructural y funcional precisa de este polipéptido es muy importante para el diseño de nuevos fármacos antivirales.

Con base en lo anterior, el presente trabajo busca evaluar a nivel estructural la presencia de mecanismos comunes para NS3 en DENV y HCV, así como caracterizar las regiones involucradas en la regulación de los movimientos interdominio entre NS3pro y NS3hel y por lo tanto en el comportamiento dinámico de esta proteasa-helicasa en distintos flavivirus.

3. OBJETIVOS

General

- Determinar las características estructurales comunes para la proteína no estructural 3 de los virus de la Hepatitis C (**HCV**) y Dengue (**DENV**).

Particulares

- Caracterizar la influencia del dominio de proteasa serínica en el comportamiento dinámico del dominio de ARN helicasa, mediante simulaciones de Dinámica Molecular.
- Construir un modelo de la proteína NS3 del virus del Dengue, basado en las coordenadas homólogas de la proteína NS3 del virus de la Hepatitis C.
- Caracterizar los cambios conformacionales de NS3 en los virus de la Hepatitis C y Dengue, por medio de Análisis de Modos Normales (**NMA**).
- Determinar, mediante el uso de dinámicas moleculares interactivas, las interacciones entre los dominios de proteasa serínica, ARN helicasa y el extremo carboxilo terminal de NS3.

4. MATERIALES Y METODOS

4.1 Modelo de la Proteína NS3 del Virus del Dengue.

Las diferentes estructuras utilizadas durante la construcción del modelo de la proteína NS3 del virus del Dengue (**scNS3-NS2b**) se obtuvieron del Protein Data Bank, PDB (Berman *et al.* 2000; Berman, Henrick, & Nakamura 2003). A partir de las coordenadas de los dominios aislados de proteasa serínica (**1BEF**, resolución: 2.10 Å) y ARN helicasa (**2BHR**, resolución: 2.80 Å), se determinó inicialmente la orientación relativa de cada dominio, basada en el alineamiento estructural con respecto a la proteína homóloga de HCV reportada por Yao *et al.* (**1CU1**, resolución: 2.50 Å) (Yao et al. 1999). Una vez que ambos dominios estuvieron posicionados, ambas estructuras se utilizaron como moldes para la construcción de un nuevo modelo, el cual incorpora tanto la región central del péptido cofactor NS2b, así como el asa que une ambos dominios. Durante la construcción de este nuevo modelo se utilizó el programa MODELLER (Sali & Blundell 1993).

La estrategia adoptada durante el afinamiento del polipéptido scNS3-NS2b, fue adaptada de Selisko *et al.* (Selisko *et al.* 1999), utilizando el programa X-PLOR (Brünger, Kuriyan, & Karplus 1987) durante todas las fases del siguiente protocolo:

1. Minimización de energía, utilizando el minimizador de Powell (200 pasos, caída inicial de energía $40.0 \text{ kcal/mol}^{-1}$, restricciones armónicas $40.0 \text{ kcal/mol}^{-1} \text{ \AA}^{-2}$).
2. Dinámica molecular (200 pasos, α fijos, timestep 0.5, temperatura inicial 300 °K).
3. 50 ciclos de:
 - a) Minimización de energía de cuerpo rígido (50 pasos, caída inicial de energía $0.1 \text{ kcal mol}^{-1}$), durante esta fase, se consideró como cuerpo rígido a los dominios de proteasa serínica y de ARN helicasa, así como a cada residuo involucrado en el asa interdominio.
 - b) Dinámica molecular (200 pasos, α fijos, timestep 0.2, temperatura inicial 300 °K).

Durante la fase final del afinamiento, se utilizó el programa NAMD2 (Kale *et al.* 1999;Phillips *et al.* 2005), con el conjunto de potenciales CHARMM27 (MacKerell *et al.* 1992;MacKerell *et al.* 1998;MacKerell 2002), así como el método Particle Mesh Edwald (PME) (Darden, York, & Pedersen 1993;Darden & Pedersen 1996) para el cálculo de interacciones electrostáticas a larga distancia. La distancia de corte para las interacciones de van der Waals se fijó en 10 \AA , utilizando un sistema de condiciones periódicas y permitiendo un distancia mínima de 5 \AA entre el modelo y los límites del sistema periódico. Con el fin de permitir un paso (timestep) de 2 fs, todos los enlaces pertenecientes a átomos de hidrógeno se restringieron

utilizando la implementación del algoritmo SHAKE (Ryckaert, Cicciotti, & Berendsen 1977) presente en NAMD2.

Durante la adición de la caja de solvente, una caja de dimensiones 96.0 x 93.0 x 98.5 Å se llenó de moléculas de agua, removiendo posteriormente aquellas moléculas de agua a una distancia menor a 2.4 Å del modelo con ayuda del programa Visual Molecular Dynamics (VMD) (Humphrey, Dalke, & Schulten 1996).

El protocolo utilizado en la última fase de afinamiento consistió en:

- 1) Minimización de Energía, (2500 pasos).
- 2) 2 ciclos de Dinámica Molecular (100 fs) y Minimización de Energía (1000 pasos), con los átomos de la cadena principal fijos.
- 3) 2 ciclos de Dinámica Molecular (100 fs) y Minimización de Energía (1000 pasos), con restricciones armónicas sobre los átomos C α .
- 4) 2 ciclos de Dinámica Molecular (100 fs) y Minimización de Energía (1000 pasos).
- 5) Minimización de Energía, (2500 pasos).

La estabilidad global del sistema proteína/solvente se evaluó durante una simulación de 1ns a 298 °K, muestreando las conformaciones presentes cada 10 ps.

4.2 Dinámica Molecular de scNS3-NS4a.

Con el fin de evaluar las diferencias en la estabilidad y el comportamiento dinámico de scNS3-NS4aa y NS3hel (residuos 201 a 631) para HCV, se llevaron a cabo análisis de dinámica molecular sobre ambos polipéptidos. Las coordenadas de NS3hel se obtuvieron a partir de la

estructura 1CU1, removiendo los residuos 705-720 (NS4a) y 3-200 (NS3pro más la cadena de unión interdominio).

Las simulaciones para todos los modelos se llevaron a cabo con ayuda de NAMD2, utilizando los potenciales CHARMM27, bajo las condiciones mencionadas anteriormente. Inicialmente se generó un sistema consistente en scNS3-NS4a, inmersa en una caja de solvente de dimensiones 74 x 79 x 89 Å. El protocolo de simulación consistió en: 1000 pasos de Minimización de Energía y Dinámica Molecular durante 1 ns a 300 °K, muestreando la trayectoria cada 2 ps. Un protocolo idéntico se llevó a cabo para las coordenadas de NS3hel, ajustando las condiciones periódicas del sistema a una caja de 74 x 79 x 61 Å.

Las variaciones en la estructura se analizaron mediante el alineamiento estructural de las conformaciones muestreadas para ambas trayectorias y calculando el RMSD de los C α . Con la finalidad de obtener una perspectiva óptima de las diferencias en el comportamiento dinámico entre scNS3-NS4a y NS3hel, cada trayectoria fue alineada considerando únicamente los C α del dominio de ARN helicasa, calculando las diferencias en los RMSD para cada residuo con la ayuda de VMD.

4.3 Análisis de Modos Normales.

El análisis de Modos Normales (NMA, por sus siglas en inglés) se llevó a cabo con la ayuda de ElNemo (Suhre & Sanejouand 2004), una interfase de tipo web para el análisis de modos normales mediante el modelo de redes elásticas (Elastic Network Model). Como se ha explicado más ampliamente en la introducción, el uso de las técnicas usuales de dinámica molecular es de poca utilidad, debido a los largos tiempos de cómputo requeridos para muestrear de manera

adecuada el espacio conformacional disponible para un sistema de tamaño medio o grande. En contraste, los cálculos de NMA se basan en la aproximación armónica de la energía potencial de una molécula alrededor de una conformación de mínima energía, utilizando una aproximación energética muy sencilla. ElNemo, está basado en la implementación original de éste modelo hecha por (Tirion 1996), en donde los campos de fuerza empleados usualmente son reemplazados por un parámetro Hookeano único de energía potencial; este tipo de simplificaciones evita la necesidad de llevar a cabo minimizaciones energéticas largas. Otra ventaja inherente al empleo de ElNemo es la utilización de bloques de rotación-translación (RTB, por sus siglas en inglés), permitiendo el tratamiento de modelos a detalle atómico, reduciendo al mismo tiempo los tiempos de cálculo. De acuerdo con los autores, el uso de este tipo de aproximación tiene un impacto muy pequeño en los modos normales de baja frecuencia resultantes.

Durante los análisis que se llevaron a cabo en el presente trabajo, se utilizó una distancia de corte de 9 Å y un tamaño de RTB igual a un residuo, debido a que estos valores dieron la mejor correlación entre los factores B observados y los calculados por ElNemo. La amplitud de las perturbaciones aplicadas para cada modo normal se determinó mediante un análisis de la conservación de la estructura secundaria y la presencia de choques interatómicos, evaluada con PROCHECK (Laskowski *et al.* 1993). Finalmente el análisis de los movimientos relativos de cada región de la proteína se llevó a cabo con la ayuda de DynDom 1.5 (Hayward & Berendsen 1998).

4.4 Simulaciones de Dinámica Molecular Interactivas.

Con el objeto de analizar de una manera más fina las interacciones presentes entre el dominio de proteasa serínica y el extremo carboxilo terminal del dominio de ARN helicasa de scNS3-NS4a, se llevó a cabo una serie de simulaciones de Dinámica Molecular interactivas en la estructura reportada por Yao *et al.* Como se ha descrito en la sección 4.2, la proteína solvatada fue equilibrada durante 1 ns. Para que el sistema pudiese ser simulado a una velocidad razonable para la interacción con el sistema, fue necesario llevar a cabo varias simplificaciones en el mismo: sólo aquellas moléculas de solvente y residuos que se encontraban en un radio de 6 Å alrededor de extremo carboxilo terminal fueron incluidos en la simulación. El número de moléculas de agua aseguró la disponibilidad de las mismas para reemplazar al carboxilo terminal conforme éste era desplazado.

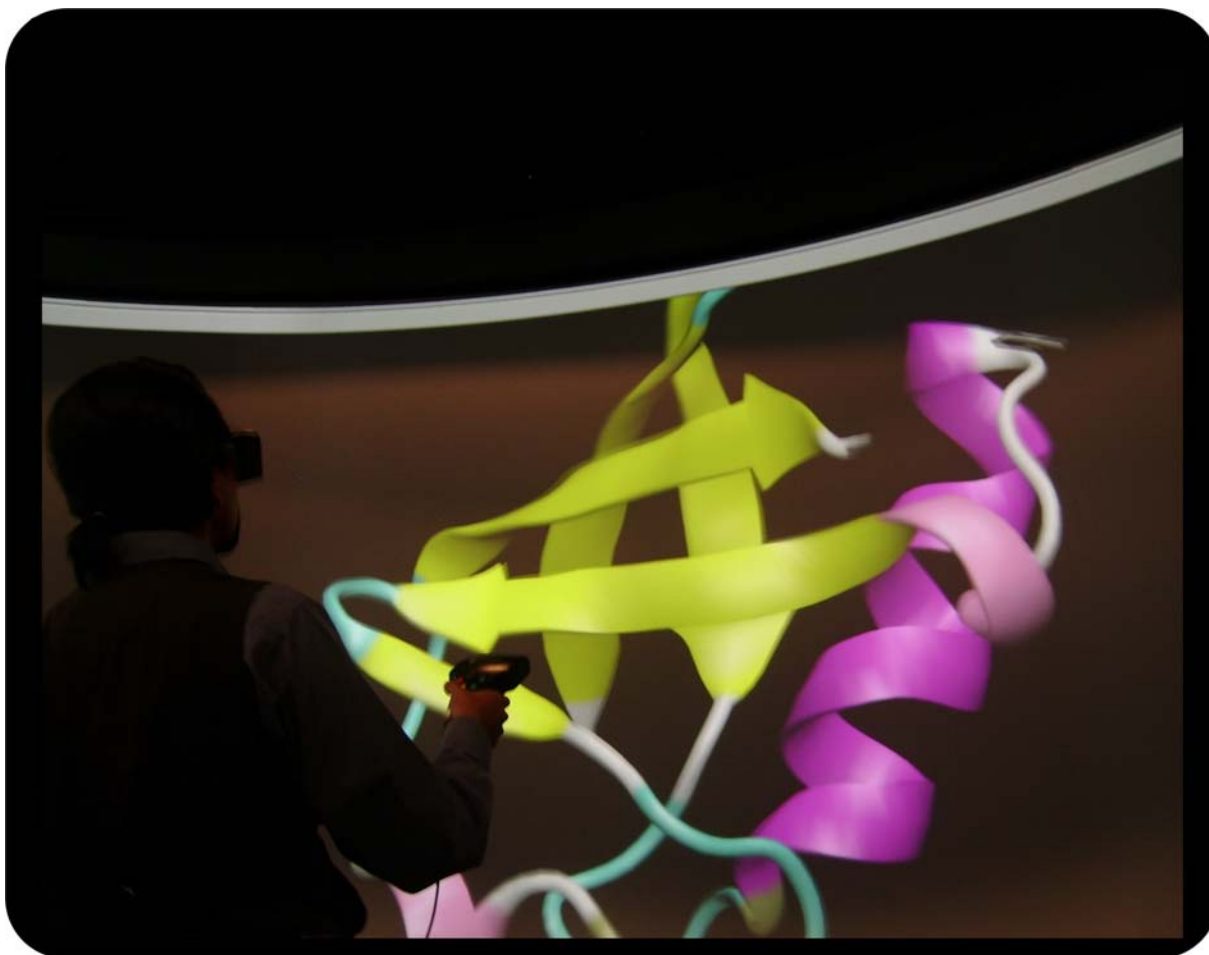


Figura 3 - Sesión de Dinámica Molecular Interactiva.

Durante estos ensayos se aplicó una fuerza sobre el extremo carboxilo de NS3 (representación de bastones), con la finalidad de liberarlo del dominio de proteasa serínica (superficie verde, solo se muestran aquellos átomos a una distancia máxima de 6 Å del extremo carboxilo terminal).

El sistema final contenía 4,617 átomos, 1,950 de los cuales contaban con una completa libertad de movimiento. Durante todas las simulaciones interactivas, se utilizó el conjunto de potenciales CHARMM27, mientras que el sistema se mantuvo a una temperatura constante de 310 °K, mediante control de Langevin, con un coeficiente de 10 ps^{-1} y manteniendo una distancia de corte para las interacciones no covalentes de 10 Å. En total se llevaron a cabo 4 dinámicas interactivas, en las cuales el extremo carboxilo terminal era jalado lentamente a través de la

interfase entre los dominios de ARN helicasa y proteasa serínica, a lo largo de una hora de tiempo real (~ 100 ps de tiempo simulado).

Todas las simulaciones interactivas se llevaron a cabo en el observatorio de realidad virtual Ixtli (Fig. 3), de la Dirección General de Servicios de Cómputo Académico de la UNAM, utilizando 10 procesadores de la supercomputadora Xitle (SGI Onyx, 12 procesadores R14000)

5. RESULTADOS

5.1 Modelo de scNS3-NS2b.

Construimos un modelo para la proteína scNS3-NS2b basado en las estructuras reportadas para el polipéptido completo de HCV. De manera general, nuestro modelo muestra una interfase distinta a la que se encuentra presente en la estructura de referencia (1CU1), lo cual se debe principalmente a las diferencias estructurales presentes en el tercer subdominio de ARN helicasa entre ambos géneros (Figura 4).

La estrategia de afinamiento adoptada nos permitió optimizar la interfase entre ambos dominios, mediante el reacomodo de las cadenas laterales presentes y cambios en la orientación relativa entre ambos dominios funcionales de scNS3-NS2b. Debido a que el modelo inicial, generado con Modeller, contaba con una buena estereoquímica, durante las últimas fases del afinamiento solo fue necesario utilizar dinámicas moleculares de corta duración. El modelo final

comprende 635 residuos, organizados en 5 subdominios. La cadena interdominio tiene una conformación extendida, similar a la presente en la estructura de scNS3-NS4a (Fig 4c y 4d). La interfase entre ambos dominios involucra 12 residuos del dominio de proteasa serínica, así como 15 residuos de NS3hel.

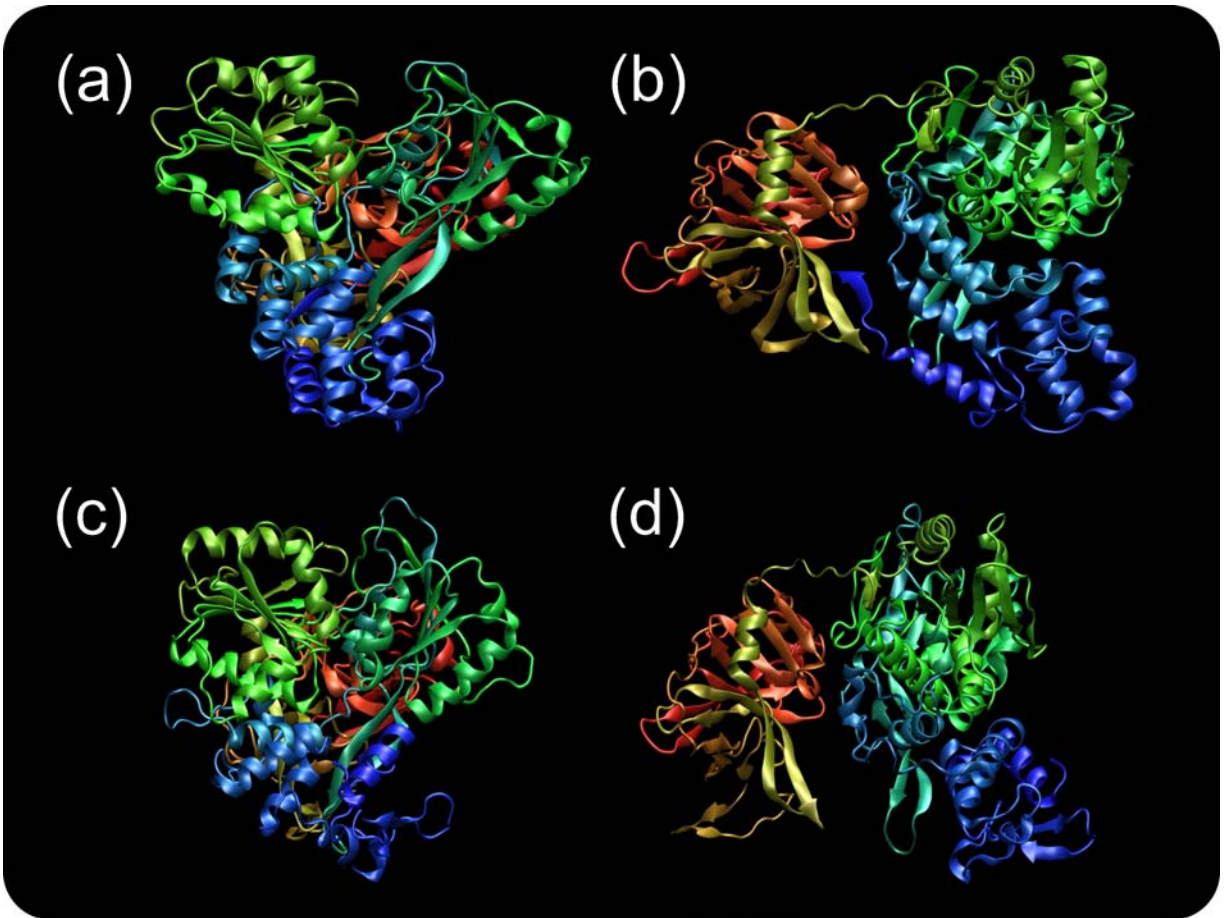


Figura 4 - Estructura de las proteínas NS3 de HCV y DENV.

(a) Representación esquemática de listón de la proteína scNS3-NS4a de HCV, el dominio de ARN helicasa se encuentra al frente. (b) scNS3-NS4a, vista lateral; la proteína ha sido rotada 90° en el eje y con respecto a la vista en (a). (c) Representación esquemática de listón de la proteína scNS3-NS2b de DENV, la orientación de la proteína es similar a la vista en (a). (d) scNS3-NS2b, vista lateral, la orientación es similar a (b).

Las interacciones locales en la interfase están dadas principalmente por contactos hidrofóbicos, con la presencia de 6 puentes de hidrógeno entre residuos de los dominios adyacentes.

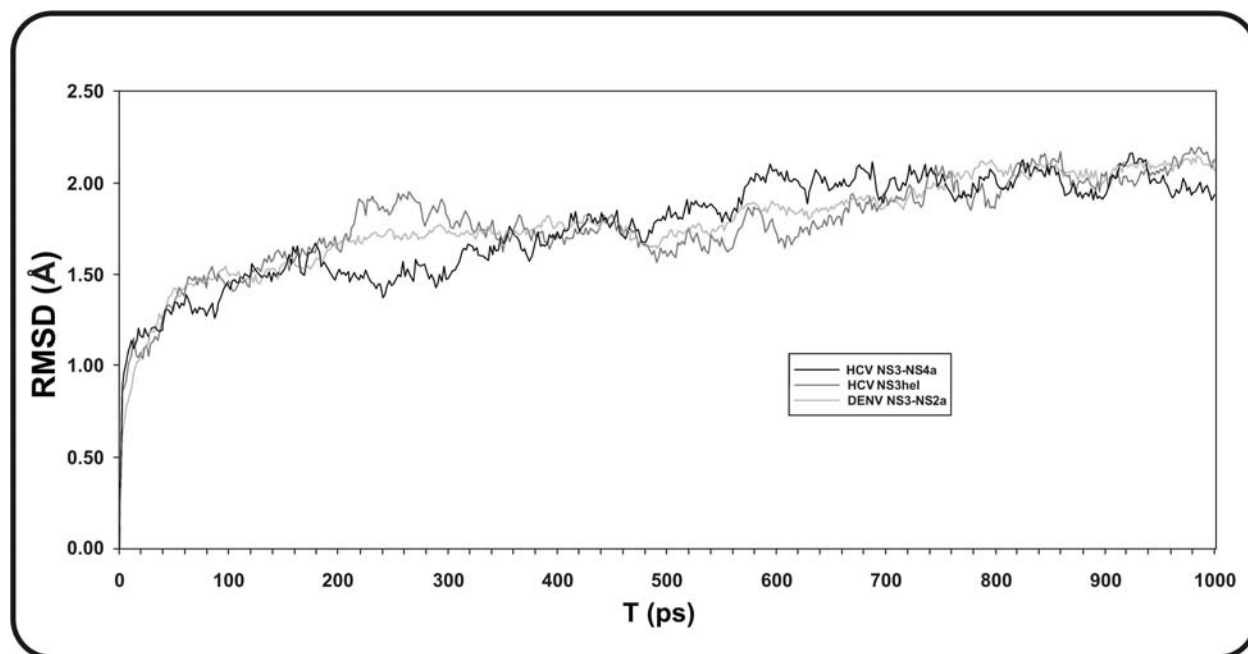


Figura 5 - Análisis del comportamiento dinámico de NS3 mediante Dinámica Molecular.

RMSD para los C α de distintos fragmentos de NS3 en Flaviviridae a lo largo de una simulación de DM durante 1 ns. En negro, scNS3-NS4a de Hepatitis C; en gris medio NS3hel, modelo derivado de las coordenadas de HCV; en gris claro modelo teórico de scNS3-NS2b para DENV.

Posteriormente, evaluamos la estabilidad del modelo a lo largo de una dinámica molecular larga (1 ns). Como es posible observar en la Figura 5, las variaciones en el RMSD alcanzan una meseta después de 100 ps, indicando que la simulación ha alcanzado el equilibrio. El valor global de RMSD para las 500 conformaciones muestreadas es de 1.49 Å, como es de esperarse, las regiones de mayor movilidad corresponden a aquellos residuos presentes en las asas, indicando la estabilidad presente en las regiones con estructura secundaria repetitiva (Figura 6).

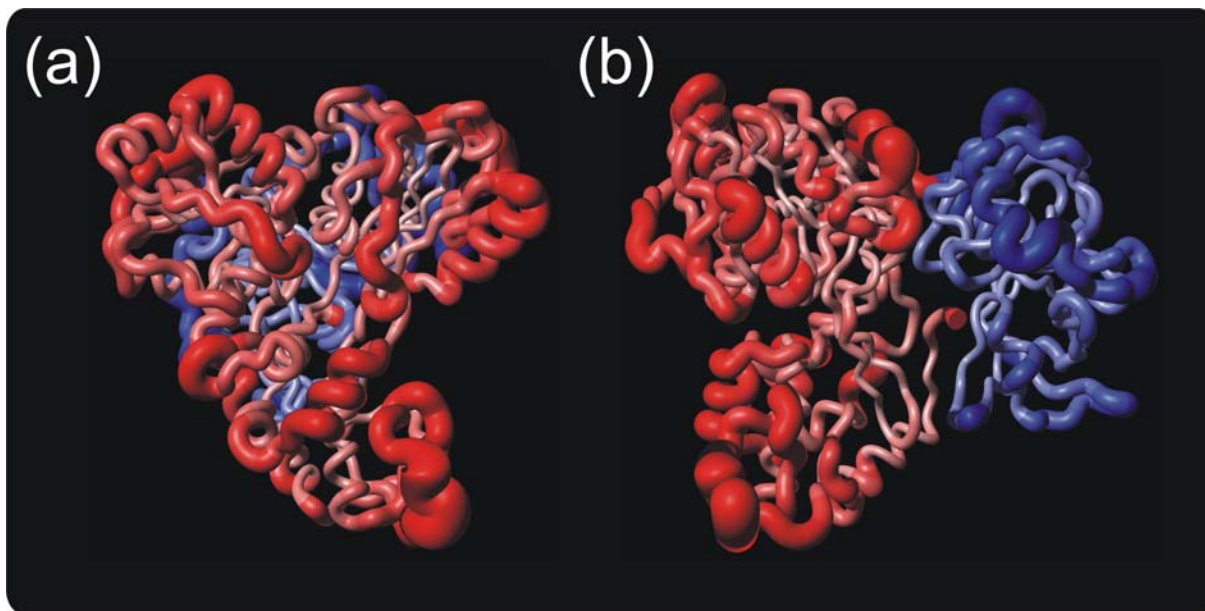


Figura 6 - Análisis del comportamiento dinámico de NS3 mediante Dinámica Molecular.

(a) Representación de listón de la proteína scNS3-NS4a. de HCV. El grosor del listón y la intensidad del color representan el tamaño relativo de las variaciones en RMS para la traza de C α . NS3pro y NS3hel se muestran en azul y rojo, respectivamente. (b) Vista lateral de la proteína, la cual representa una rotación de 90° en el eje vertical con respecto a la vista en (a).

5.2 Simulaciones de Dinámica Molecular para scNS3-NS4a y NS3hel.

Durante este análisis llevamos a cabo diferentes simulaciones de dinámica molecular de 1 ns a 300 °K en la presencia de solvente, tanto para la estructura reportada por Yao *et al.*, como para la estructura del dominio de ARN helicasa aislado. Este análisis tuvo como objetivo evaluar si existían diferencias en el comportamiento dinámico de ambos modelos, las cuales podrían ser causadas por las interacciones entre los dominios de proteasa serínica y ARN helicasa. Para ello, se muestrearon 500 conformaciones y se calcularon las variaciones en el RMSD a lo largo de cada simulación. En la Figura 7 se muestran las variaciones de RMSD para los C α de scNS3-NS4a y NS3hel. En cada caso las variaciones han sido reportadas con respecto a la conformación inicial (T = 0 ns). Es posible observar que para scNS3-NS4a, la simulación alcanza el equilibrio

cerca de los 100 ps, mientras que NS3hel tarda un tiempo más largo para alcanzar un grupo de conformaciones estables.

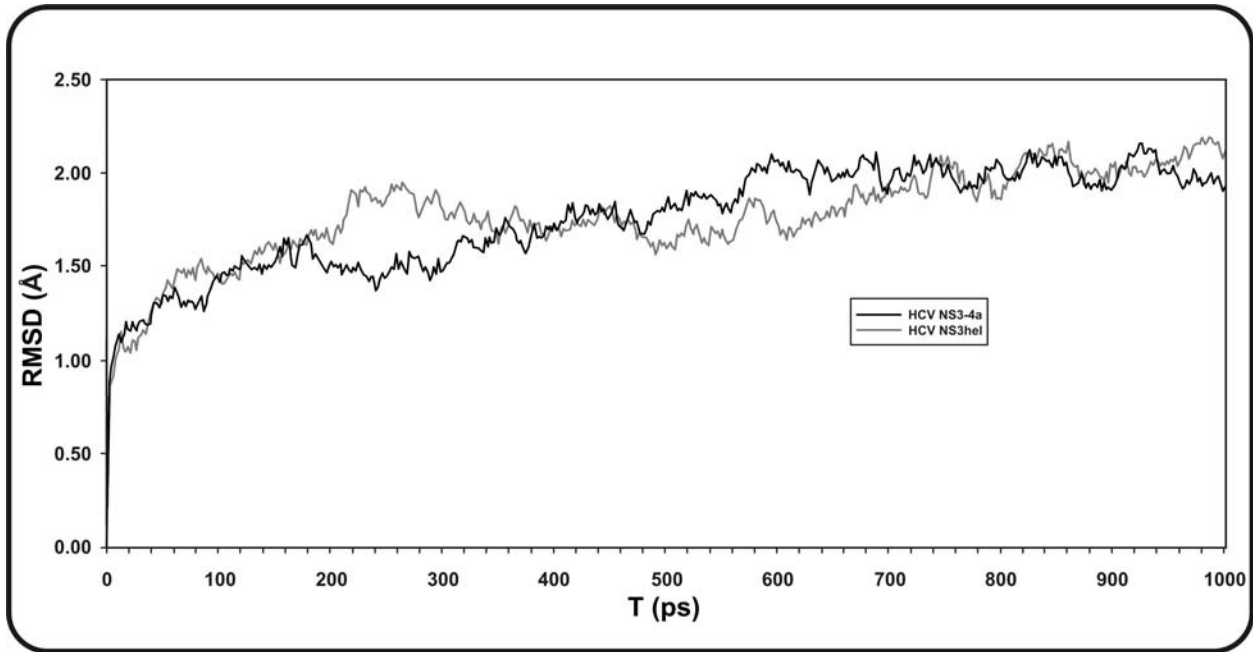


Figura 7 - Análisis del comportamiento dinámico de scNS3-NS4a mediante Dinámica Molecular.

RMSD para los C α de distintos fragmentos de HCV NS3 a lo largo de una simulación de DM durante 1 ns. En negro, scNS3-NS4a de Hepatitis C; en gris NS3hel.

Debido a que ambos sistemas partieron de estructuras idénticas para el dominio de ARN helicasa, es posible que estas diferencias puedan ser atribuidas a un efecto estabilizador de NS3pro sobre los residuos involucrados en la interfase con NS3hel, principalmente en el extremo carboxilo terminal, el cual se encuentra en contacto directo con el sitio activo de la proteasa.

Para determinar con mayor detalle aquellos residuos involucrados en las diferencias de movilidad observadas, se calculó el RMSD de cada uno de los C α para NS3hel, para ambas trayectorias. Posteriormente, se comparó el promedio de las mediciones de RMSD entre residuos equivalentes en ambos dominios de ARN helicasa, normalizando los valores. La Figura 8a muestra la distribución de los cambios normalizados de RMSD a lo largo de la secuencia de NS3hel. Es posible observar que cerca del 90 % de los residuos cuentan con diferencias menores

a 0.19, sin embargo, la distribución del 10 % restante no se encuentra aleatoriamente distribuida. En la Figura 8b, se muestran aquellos residuos con diferencias mayores a 0.19, los cuales han sido mapeados en la estructura de referencia, 1CU1 (el radio del listón representa las desviaciones en RMSD calculados para la traza de $C\alpha$).

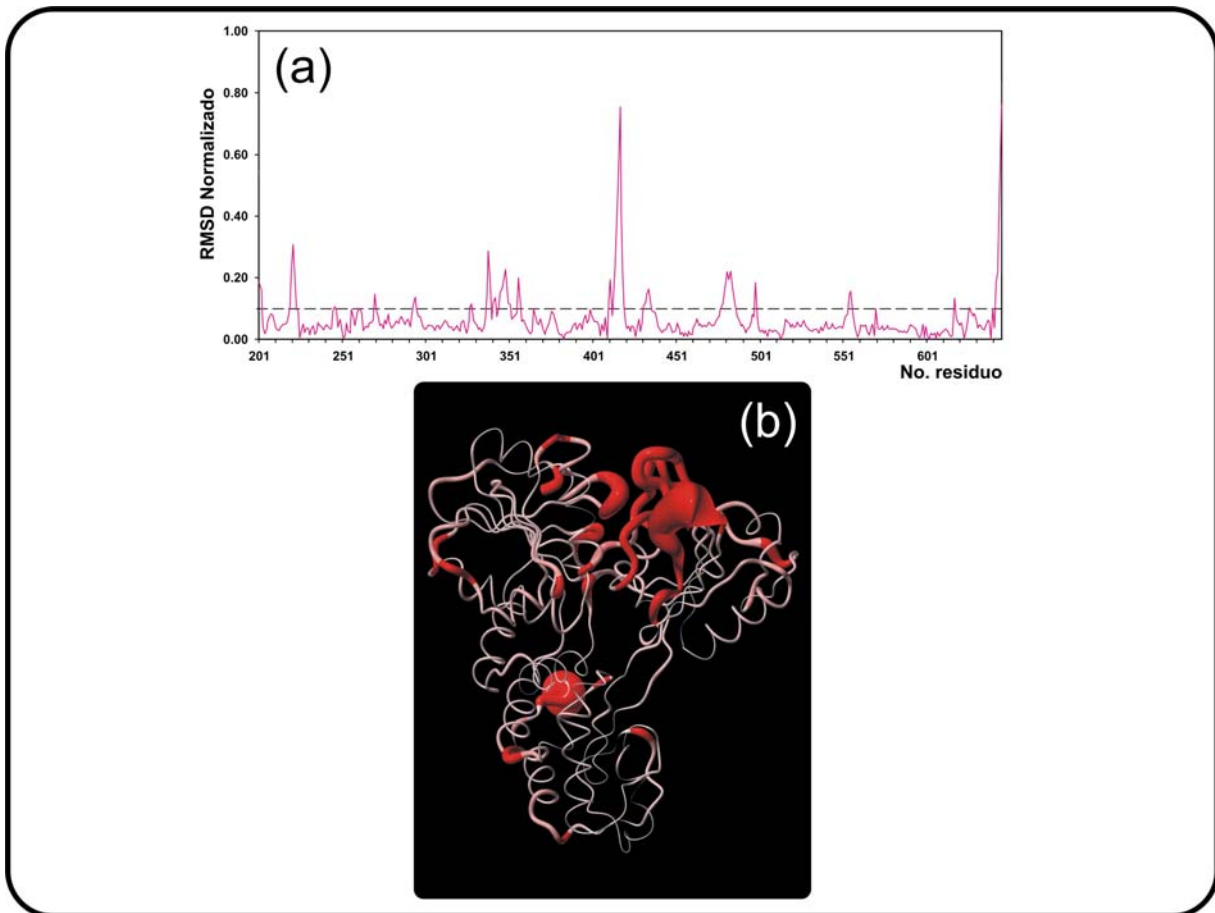


Figura 8 - Diferencias en el comportamiento dinámico de scNS3-NS4a.

(a) Diferencias en la movilidad por residuo entre scNS3-NS4a y NS3hel. Cada punto representa la diferencia, expresada como valores absolutos, entre el RMSD obtenido para NS3hel y el dominio de ARN helicasa de scNS3-NS4a. La línea punteada representa la desviación estándar de todas las diferencias consideradas. (b) Representación de listón de NS3hel, el grosor representa el tamaño relativo de las diferencias en RMSD entre NS3hel y el dominio de ARN helicasa de scNS3-NS4a. Los residuos marcados en rojo presentan diferencias más grandes a 1 desviación estándar.

Es posible observar que la mayor parte de los residuos con gran movilidad se encuentra localizada en 2 asas presentes en los subdominios 2 y 3 de NS3hel. Es importante notar que en

algunos sitios el grosor del listón es mínimo, lo cual permite apreciar la estabilidad de la estructura secundaria repetitiva en los modelos a lo largo de las simulaciones. En conjunto, estos datos sugieren que la presencia de NS3pro ejerce un efecto importante en la estabilidad del subdominio 1 de la helicasa, lo cual incluye a la mayor parte de los residuos involucrados en la unión e hidrólisis de NTPs.

Adicionalmente, la simulación de NS3hel brindó información importante referente al comportamiento del extremo carboxilo de NS3 en solución (es decir, cuando no se encuentra interaccionando directamente con el sitio activo de la proteasa). A partir de las conformaciones muestreadas y analizadas es posible observar que la longitud de la última hélice α se incrementa conforme la simulación progresa (Fig. 9, Material Suplementario 1). Estos cambios, han sido propuestos anteriormente por Yao *et al.*, basados en la comparación entre sus coordenadas atómicas y las estructuras previamente reportadas por otros grupos. Sin embargo, todas las estructuras reportadas a la fecha para NS3hel, carecen de varios residuos en el extremo carboxilo terminal, debido a la gran movilidad de esta región. Nuestros datos apoyan la hipótesis de que, una vez liberado del sitio activo de la proteasa, una parte de los residuos que lo conforman son capaces de replegarse hacia la última hélice α ; sin embargo, los últimos 5 residuos permanecen en una conformación extendida, lo cual les permitiría formar nuevamente la interfase con NS3pro, una vez que ésta ha llevado a cabo los cortes necesarios en la poliproteína viral.

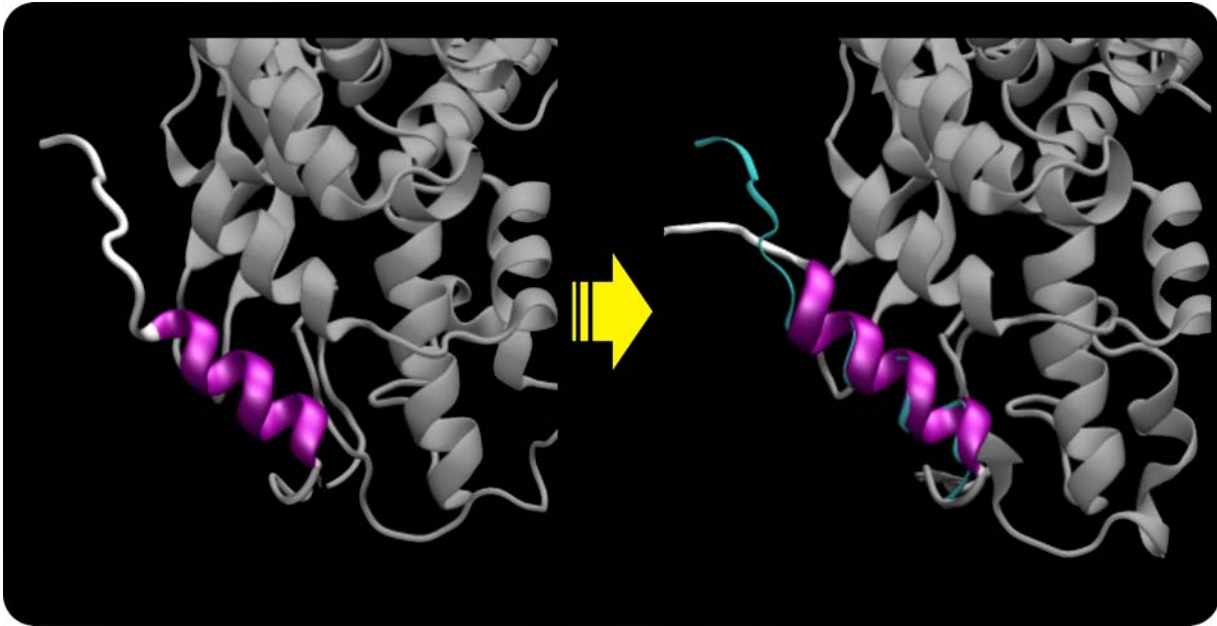


Figura 9 - Cambios Estructurales en la región carboxilo terminal de NS3 para HCV.

A la izquierda se muestra el estado inicial, correspondiente a las coordenadas obtenidas de 1CU1. A la derecha, estado final después de 1 ps. Para mayor claridad, las coordenadas iniciales del extremo carboxilo terminal (listón azul claro) se han alineado al estado final.

En apoyo a ésta hipótesis, Wang *et al.* proponen que la presencia de un residuo conservado de Thr en el sitio de corte del carboxilo terminal de NS3, en lugar de Cys (el cual representa el residuo óptimo para el corte en esta región) es necesario para permitir la asociación/disociación de la interfase entre NS3pro y NS3hel, previniendo la inhibición de la proteasa causada por el producto de corte en esta región (Wang *et al.* 2004).

5.3.1 Análisis de Modos Normales para scNS3-NS4a.

Con la finalidad de caracterizar las propiedades dinámicas de NS3, se llevó a cabo un análisis de modos normales para el modelo de HCV. El uso de Elnemo nos permitió realizar la caracterización sobre un modelo que incluía todos los átomos presentes en la estructura, utilizando un residuo por cada RTB. Reportes previos han permitido concluir que el uso este tipo

de aproximación tiene un impacto menor en los modos resultantes (Tama & Sanejouand 2001). Del análisis de modos normales para scNS3-NS4a, se escogieron y analizaron los tres modos de menor frecuencia (modos normales 7 a 9, pues los primeros seis representan modos triviales.) Las Tabla 2 a 4 muestran las características de todas las reorientaciones entre dominios asociadas a estos modos.

Nota: Para los capítulos referentes al análisis de modos normales, el término dominio hace referencia a las propiedades dinámicas de un conjunto de átomos. Así, dominio es definido como un agrupamiento semi-rígido de residuos que cuenta con propiedades rotacionales distintas del resto del polipéptido. En el resto del escrito, el término dominio hace referencia a la definición más aceptada desde el punto de vista estructural.

Modo Normal 7 Este modo incluye un movimiento rotacional de la proteasa con respecto a los subdominios 1 y 2 de la helicasa, incluyendo la participación de una región que funciona a manera de bisagra (residuos 181 a 188, listón púrpura; Fig. 10a. Material Suplementario 2); en este caso el eje de rotación es casi perpendicular a la interfase entre ambas regiones. Acoplado al movimiento anterior, existe una rotación de 15.7 ° del subdominio 3 de la helicasa, con respecto al resto de la misma. Esta reorientación incluye residuos de 5 regiones flexibles (residuos 431-435, 446-451, 480-487, 509-520, 523-525; Fig. 10b, listón verde), mientras que el eje de rotación es casi perpendicular al anterior.

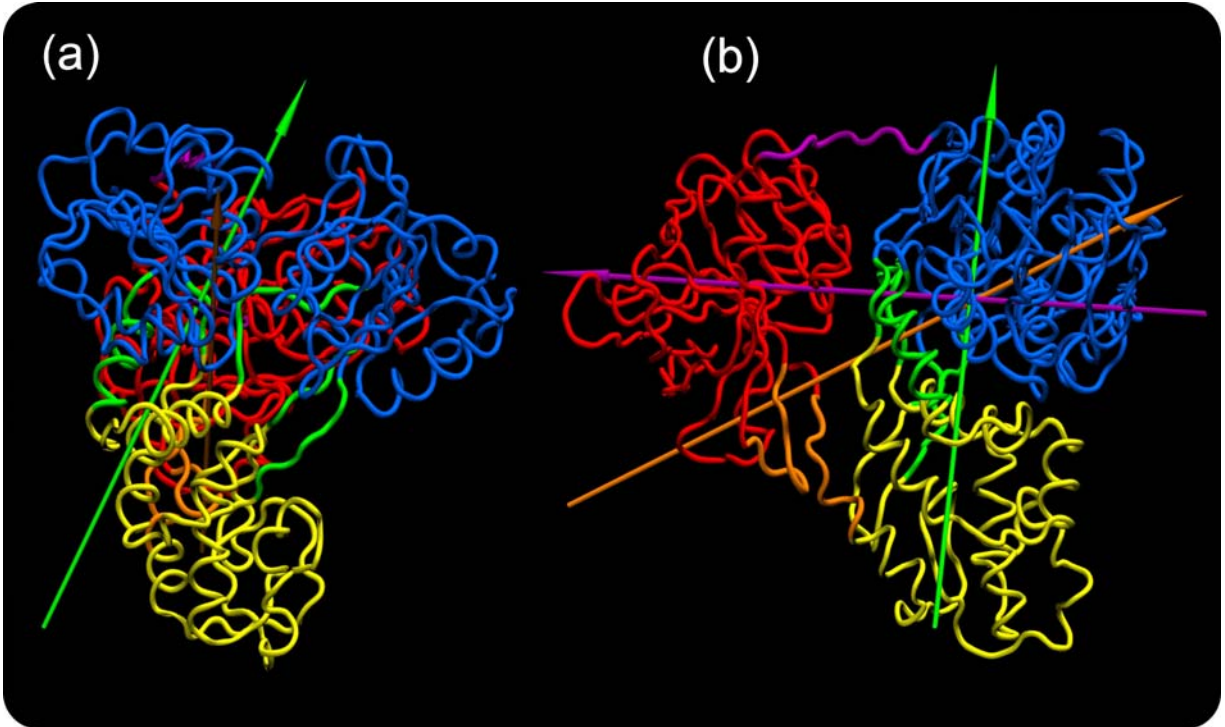


Figura 10 - Modo Normal 7.

(a) Vista frontal a NS3hel. (b) Vista de la interfase, la cual representa una rotación de 90° con respecto a la vista en (a). En ambos casos el dominio fijo (listón rojo) involucra a NS3pro, mientras que los dos dominios móviles están formados por los subdominios 1-2 (listón azul) y 3 (listón amarillo). Se indican las regiones de alta flexibilidad entre el dominio fijo y los subdominios 1-2 (listón púrpura) y 3 (listón naranja), así como entre los subdominios 1-2 y 3 (listón verde). Las flechas representan el eje de rotación entre el dominio fijo y los dominios móviles, el color de cada una corresponde al color de las regiones flexibles.

Tabla 2 Modo Normal 7, Caracterización de Dominios Dinámicos para scNS3-NS4a.

Modo Normal	Dominio Fijo* [RMSD]	Dominio Móvil** [RMSD]	Regiones de conexión	Ángulo de Rotación	% de Cerrado	Ángulo***
Modo 7	Dominio 1 (297 residuos) [0.642 Å]	Dominio 2 (193 residuos) [0.796 Å]	181-188	30.6°	4.44	12.0°
		Dominio 3 (151 residuos) [0.754 Å]	431-435 446-451 480-487 509-520 523-525	15.7°	5.99	14.0°
	Dominio 2 (193 residuos) [0.796 Å]	Dominio 3 (151 residuos) [0.754 Å]	157-165 622-629	33.8°	83.30	65.0°

* Dominio es definido como un agrupamiento semi-rígido de residuos que cuenta con propiedades rotacionales distintas del resto del polipéptido.

** El Dominio Móvil es considerado en base a su movimiento relativo al Dominio Fijo.

*** Definido como el ángulo entre el eje de giro y una línea que une los centros de masa entre los dominios Móvil y Fijo.

Es importante mencionar que, para las amplitudes seleccionadas, la integridad de la interfase entre el dominio de proteasa serínica y el subdominio 3 de la RNA helicasa no es afectada, sin embargo, la mayor parte de esta interfase se encuentra involucrada en las regiones flexibles que permiten acoplar las reorientaciones entre ambos dominios.

Modo Normal 8 Este modo describe un movimiento de apertura/cerrado entre la RNA helicasa y la proteasa (Fig. 11 & Material Suplementario 3), en este caso, los subdominios 1 y 2 de la helicasa forman una unidad semi-rígida, mientras que el eje de rotación asociado a esta reorientación es paralelo a la interfase entre ambos dominios.

De manera similar al modo 7, el movimiento de la proteasa se encuentra acoplado a una reorientación del subdominio 3 de NS3hel; sin embargo, en este caso el movimiento de este subdominio es más limitado que el observado en el modo anterior. Asimismo, la mayor parte de los residuos que participan en regiones de alta flexibilidad pertenecen a la interfase entre ambos dominios, por lo cual la misma no se disgrega completamente para las amplitudes seleccionadas.

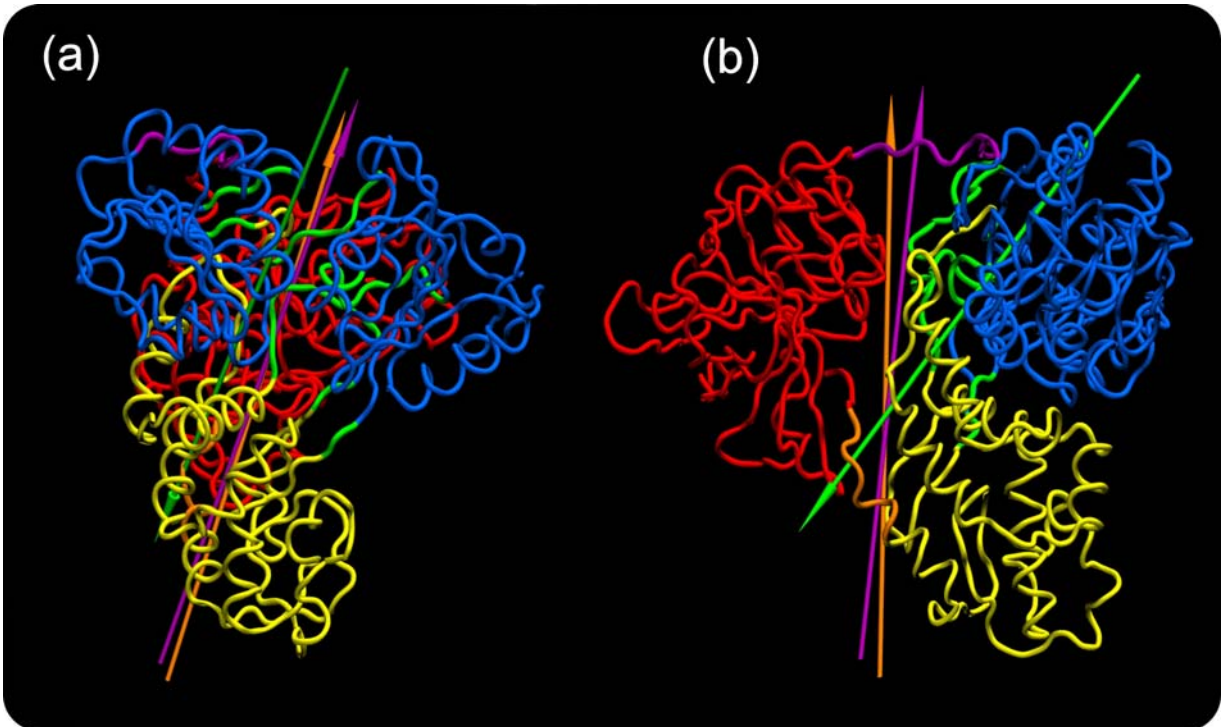


Figura 11 - Modo Normal 8.

(a) Vista frontal a NS3hel. (b) Vista de la interfase, la cual representa una rotación de 90° con respecto a la vista en (a). En ambos casos el dominio fijo (listón rojo) involucra a NS3pro, mientras que los dos dominios móviles están formados por los subdominios 1-2 (listón azul) y 3 (listón amarillo). Se indican las regiones de alta flexibilidad entre el dominio fijo y los subdominios 1-2 (listón púrpura) y 3 (listón naranja), así como entre los subdominios 1-2 y 3 (listón verde). Las flechas representan el eje de rotación entre el dominio fijo y los dominios móviles, el color de cada una corresponde al color de las regiones flexibles.

Tabla 3 Modo Normal 8, Caracterización de Dominios Dinámicos para scNS3-NS4a.

Modo Normal	Dominio Fijo* [RMSD]	Dominio Móvil** [RMSD]	Regiones de conexión	Ángulo de Rotación	% de Cerrado	Ángulo***
Modo 8	Dominio 1 (200 residuos) [0.836 Å]	Dominio 2 (265 residuos) [0.748 Å]	184-193	35.6°	96.2	78.7°
		Dominio 3 (173 residuos) [0.521 Å]	622-628	31.0°	66.5	54.6°
	Dominio 2 (265 residuos) [0.748 Å]	Dominio 3 (173 residuos) [0.521 Å]	202-205 321-322 327-333 433-434 448-449 453-458 477-487	33.8°	83.30	65°

* Dominio es definido como un agrupamiento semi-rígido de residuos que cuenta con propiedades rotacionales distintas del resto del polipéptido.

** El Dominio Móvil es considerado en base a su movimiento relativo al Dominio Fijo.

*** Definido como el ángulo entre el eje de giro y una línea que une los centros de masa entre los dominios Móvil y Fijo.

Modo Normal 9 En este modo, es posible observar una unidad semi-rígida formada por el dominio de proteasa y el subdominio 1 de NS3hel (listón rojo, Fig. 12 & Material Suplementario 4), mientras que las reorientaciones restantes describen un movimiento de apertura/cerrado entre los subdominios 2 y 3 de la helicasa.

En este modo normal, existen dos regiones flexibles que funcionan como bisagras para el desplazamiento de los subdominios 1 (454-457, 480-483) y 2 (532, 623-624).

De acuerdo con los datos anteriores, los movimientos de la proteasa presentes en los modos normales 7 y 8 se encuentran acoplados a los movimientos en la helicasa debido a la participación de residuos en la interfase. De manera particular, encontramos tres regiones (56, 78, 79), (326-327, 524) y (483-485, 520-522) que se hallan muy cercanas a los ejes de rotación entre ambos dominios y que funcionan como “regiones pivote” o de giro. Asimismo, existen elementos bien delimitados de la estructura secundaria (residuos 613-623, hélice α ; residuos 430-452, giro beta) que presentan gran importancia, acoplando los movimientos del dominio de proteasa a las reorientaciones del subdominio 3 de la helicasa.

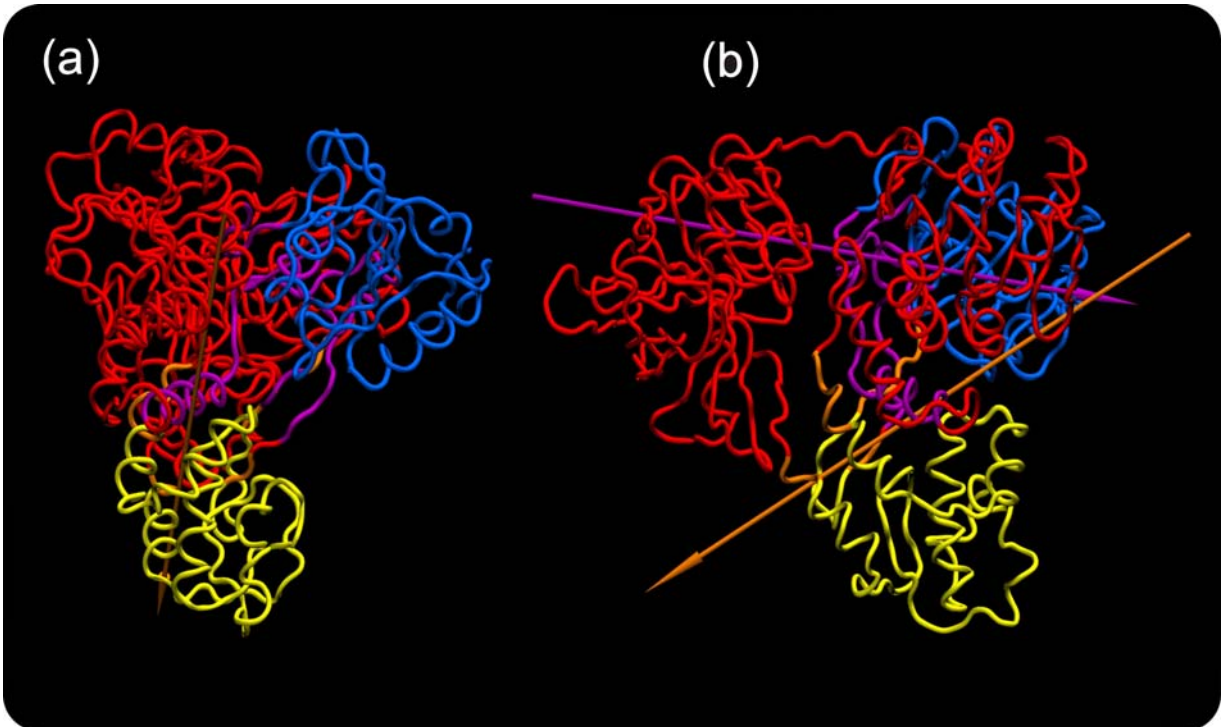


Figura 12 - Modo Normal 9.

(a) Vista frontal a NS3hel. (b) Vista de la interfase, la cual representa una rotación de 90° con respecto a la vista en (a). En ambos casos el dominio fijo (listón rojo) involucra a NS3pro y el primer subdominio de NS3hel, mientras que los dos dominios móviles están formados por los subdominios 2 (listón azul) y 3 (listón amarillo). Se indican las regiones de alta flexibilidad entre el dominio fijo y los subdominios 2 (listón púrpura) y 3 (listón naranja). Las flechas representan el eje de rotación entre el dominio fijo y los dominios móviles, el color de cada una corresponde a el color de las regiones flexibles.

Tabla 4 Modo Normal 9, Caracterización de Dominios Dinámicos para scNS3-NS4a.

Modo Normal	Dominio Fijo* [RMSD]	Dominio Móvil** [RMSD]	Regiones de conexión	Ángulo de Rotación	% de Cerrado	Ángulo***
Modo 9	Dominio 1 (408 residuos) [0.964 Å]	Dominio 2 (131 residuos) [0.475 Å]	324-329 429-435 448-458 477-499	16.9°	73.3	58.9°
		Dominio 3 (102 residuos) [0.715 Å]	432-436 445-446 527-532 623-625	18.0°	99.1	84.5°

* Dominio es definido como un agrupamiento semi-rígido de residuos que cuenta con propiedades rotacionales distintas del resto del polipéptido.

** El Dominio Móvil es considerado en base a su movimiento relativo al Dominio Fijo.

*** Definido como el ángulo entre el eje de giro y una línea que une los centros de masa entre los dominios Móvil y Fijo.

5.3.2 Análisis de Modos Normales para scNS3-NS2b.

Con el fin de comparar las propiedades dinámicas entre los modelos de NS3 para HCV y DENV, analizamos el comportamiento dinámico de este último. La Tabla 5 muestra las características de todos los movimientos asociados a los modos normales analizados. Como era de esperarse, scNS3-NS2b presenta un elevado número de modos normales de baja frecuencia asociados a valores altos de colectividad. Debido a que nuestro modelo de NS3 carece de los contactos presentes en el extremo carboxilo terminal de NS3-HCV, es de esperar que se presente un número mayor de movimientos con alta colectividad, debido a que la proteasa cuenta con una gran cantidad de grados de libertad para moverse.

De acuerdo a nuestro análisis con DynDom, es necesario tomar en cuenta los primeros 5 modos normales de más baja frecuencia para poder incluir las reorientaciones de los dominios que son equivalentes a los modos normales 7 a 9 de NS3-HCV.

Tabla 5 Análisis de Modos Normales, Caracterización de Dominios Dinámicos en los 3 modos de menor frecuencia para scNS3-NS2b.

Modo Normal	Dominio Fijo* [RMSD]	Dominio Móvil** [RMSD]	Regiones de conexión	Ángulo de Rotación	% de Cerrado	Ángulo***
Modo 7	Dominio 1 (197 residuos) [1.428 Å]	Dominio 2 (434 residuos) [0.831 Å]	194-204	30.0°	44.73	41.9°
Modo 8	Dominio 4 (309 residuos) [0.943 Å]	Dominio 3 (103 residuos) [0.619 Å]	187-204	29.5°	8.7	17.1°
		Dominio 1 (89 residuos) [0.713 Å]	335-341 491-500	26.2°	81.5	64.5°
		Dominio 2 (117 residuos) [0.782 Å]	445-451 459-463 507-513 532-534	13.6°	62.24	52.0°
	Dominio 1 (89 residuos) [0.713 Å]	Dominio 3 (103 residuos) [0.619 Å]	42-44 55-56 64-65 97-98	9.7°	99.4	85.7°
Modo 9	Dominio 2 (114 residuos) [1.359 Å]	Dominio 3 (84 residuos) [0.848 Å]	101-103 184-187	15.8°	12.2	20.5°
		Dominio 1 (433 residuos) [1.473 Å]	191-204 338-340 494-499	24.8°	95.8	77.7°
Modo 10	Dominio 3 (328 residuos) [2.000 Å]	Dominio 1 (189 residuos) [0.878 Å]	185-203 338-340	24.4°	88.2	69.9°
		Dominio 2 (100 residuos) [1.743 Å]	446-455 458-462 541-547	37.2°	35.3	36.4°
Modo 11	Dominio 2 (352 residuos) [2.754 Å]	Dominio 2 (178 residuos) [1.428 Å]	205-208 212-225 240-247 301-302 305-306 331-332 446-447 460-463 502-503	19.7°	92.8	74.5°
			Dominio 3 (101 residuos) [1.573 Å]	449-450 459-461 542-544	22.4°	99.4

* Dominio es definido como un agrupamiento semi-rígido de residuos que cuenta con propiedades rotacionales distintas al resto del polipéptido.

** El Dominio Móvil es considerado en base a su movimiento relativo al Dominio Fijo.

*** Definido como el ángulo entre el eje de giro y una línea que une los centros de masa entre los dominios Móvil y Fijo.

5.4 Simulaciones de Dinámica Molecular Interactivas.

Con la finalidad de evaluar los posibles mecanismos que regulan la exposición del sitio activo de NS3pro y la ruptura de la interfase entre éste y NS3hel, se llevaron a cabo varios análisis de IMD. Durante los ensayos previos de dinámica molecular no interactiva, el extremo carboxilo terminal de NS3 se encontraba posicionado sobre el sitio activo de la proteasa. Sin embargo, los movimientos atómicos asociados a la temperatura no son suficientes para liberar los átomos del extremo carboxilo terminal. De esta manera, para poder acelerar el proceso de liberación del carboxilo, una fuerza fue aplicada en esta región durante estos ensayos. Nuestros resultados muestran que, para el modelo derivado de la estructura experimental (1CU1), existe un gran número de impedimentos estéricos que mantienen en posición la región carboxilo terminal, impidiendo el acceso a las moléculas de agua y por consiguiente previniendo que las mismas puedan competir con esta región de la interfase por la formación de puentes de hidrógeno. Estas observaciones están de acuerdo con una propuesta previa hecha por el grupo de Yao *et al.*, de acuerdo a la cual es necesario que exista una apretura previa en la interfase, la cual permitiría la exposición del extremo carboxilo terminal y su interacción con el sustrato y/o solvente.

Para evaluar esta hipótesis, se llevaron a cabo nuevas simulaciones, tomando como conformación inicial la estructura derivada del modo normal 8 que presentaba una máxima exposición al solvente del extremo carboxilo terminal. Nuestras simulaciones de IMD indican que es necesario un movimiento de apertura similar al observado en los modos normales 7 y 8, con el fin de permitir el acceso del solvente al sitio activo de la proteasa. Una vez que las moléculas de agua han alcanzado el sitio activo, es posible que faciliten la liberación del extremo carboxilo terminal al competir con el mismo por la formación de puentes de hidrógeno con la proteasa.

6. DISCUSIÓN

Con el fin de complementar nuestros datos estructurales y las observaciones derivadas de los ensayos de MD, llevamos a cabo una serie de análisis de modos normales, para los cuales se utilizaron modelos explícitos, tomando en cuenta las interacciones atómicas individuales. Nuestros análisis predicen la presencia de varias regiones de alta flexibilidad entre NS3pro y NS3hel. Estas regiones “bisagra” se encuentran asociadas al menos a un par de reorientaciones entre “dominios” semi-rígidos: un movimiento de giro (modo normal 7) y un movimiento de apertura/cierre (modo normal 8). Los cambios asociados a estos movimientos incluyen la presencia de una región que permite el acoplamiento de los movimientos de la proteasa a una reorientación limitada del subdominio 3 de la helicasa. Asimismo existe una “región pivote” adicional, formada por los residuos de la región interdominio.

Los análisis de modos normales realizados en el presente trabajo han permitido determinar la presencia de regiones flexibles presentes en la interfase entre NS3pro y NS3hel. Ambos dominios tienen codificada la plasticidad estructural necesaria para desempeñar sus

funciones. De manera particular, el subdominio 3 de la helicasa presenta una alta flexibilidad, la cual es de vital importancia para acoplar los movimientos entre NS3pro y NS3hel.

Nuestros datos son consistentes con un modelo en el cual NS3 es capaz de regular las actividades presentes en sus dominios mediante la modificación de las interacciones entre ambos (las reorientaciones entre dominios/subdominios los cuales representan movimientos de baja frecuencia y una gran colectividad; vgr. cambios conformacionales), igualmente, éstas reorientaciones están acopladas a movimientos de menor magnitud en las regiones de mayor flexibilidad, así como en las regiones de los sitios activos (movimientos localizados, de mayor frecuencia y menor colectividad). Estudios previos han mostrado que la proteasa es capaz de contribuir con sitios adicionales para la unión a ARN en el polipéptido íntegro (Gallinari, 1998); de esta manera es posible imaginar que la posición de la proteasa con respecto a la helicasa puede jugar un papel importante en la regulación de la actividad de éste dominio.

Es muy probable que las reorientaciones de “dominio rígidos” asociadas a los modos normales 7 y 8 representen todos los cambios conformacionales necesarios en la interfase para permitir la exposición del sitio activo de NS3pro, permitiendo la entrada del sustrato y/o solvente. A este respecto, nuestros análisis de IMD muestran que la reorientación entre dominios es un paso importante en la liberación del extremo carboxilo terminal de NS3, al permitir la exposición de esta región al solvente. Reportes previos sugieren que la presencia del dominio de proteasa, asociado a la región central de NS4a incrementa la actividad de RNA helicasa presente en NS3-HCV (Howe *et al.* 1999). De manera similar, Lam *et al.* han reportado que la presencia de un dominio intacto de proteasa se encuentra asociada a una mayor procesividad en el desenrollamiento de dsARN, cuando la actividad de NS3hel es comparada a la presente en el polipéptido completo (Lam, Rypma, & Frick 2004). Estos mismos autores sugieren que la presencia de NS3pro permite un uso más eficiente de la hidrólisis de ATP durante el

desenrollamiento de dsARN. Estos datos sugieren que NS3pro participa de manera importante, incrementando las actividades de ARN helicasa/NTPasa de NS3. Debido a que el dominio de proteasa serínica se encuentra a una gran distancia de los residuos involucrados en estas actividades, es muy probable que NS3pro participe de manera indirecta, estabilizando y promoviendo una conformación más activa de los residuos involucrados en la unión de ARN e hidrólisis de NTPs, tal como fue posible observar en nuestros análisis de Dinámica Molecular. La participación de residuos distales a los sitios de unión o activos ha sido reportado con anterioridad para otras enzimas, tales como la DHFR, donde la mutación de algunos residuos distales al sitio activo, como la Gly 121 disminuye la tasa de reacción en un factor de 163 (Agarwal *et al.* 2002). De este modo, es muy probable que algunos residuos presentes en el dominio de proteasa participen de manera activa en la red de interacciones que posibilitan los cambios conformacionales asociados al desenrollamiento de ácidos nucleicos y la hidrólisis de NTPs por parte de NS3hel.

Con respecto al modelo de scNS3-NS2b, nuestros análisis muestran que la asociación propuesta para los dominios de proteasa y helicasa es estable a lo largo de una simulación de 1 ns. En este momento, aún no es clara la importancia biológica de los eventos de proteólisis *in vitro* reportados para scNS3-NS2b. Es probable, que de manera similar a lo que se ha observado para HCV, el dominio de proteasa participe en la unión de NS3 a ácidos nucleicos, tal como fue propuesto por Xu *et al.* a partir de estudios de modelaje molecular (Xu *et al.* 2005).

Una de las características más notables del nuestro modelo, es la ausencia de las interacciones entre el extremo carboxilo terminal y el dominio de proteasa serínica que se hallan presentes en NS3-HCV. Como resultado de estas diferencias, nuestro análisis de modos normales predice que las reorientaciones de NS3pro para el virus del dengue se encuentran asociadas a un

mayor número de modos normales de baja frecuencia. Sin embargo, nuestros análisis muestran varias características comunes al comportamiento dinámico de NS3 en ambos virus. De manera particular, es claro que las regiones flexibles presentes en el subdominio 2 de la helicasa juegan un papel de gran importancia en los cambios conformacionales de este dominio. Adicionalmente, es claro que los movimientos del segundo subdominio se encuentran asociados a movimientos de menor magnitud en el subdominio 3, todo lo cual indica la presencia de un mecanismo estructural común que puede ser asociado a la función biológica de esta familia de enzimas.

Nuestro trabajo contribuye con nueva evidencia estructural con respecto al papel biológico de los cambios conformacionales entre los dominios de NS3 y la participación que estos cambios pueden tener en la regulación de las diversas actividades presentes en la proteína completa. De manera particular, el complejo NS3pro-NS4a de HCV juega un papel de gran importancia en la regulación temporal de las actividades de unión a ARN, NTPasa y desenrollamiento de dsARN (Gallinari *et al.* 1999;Pang *et al.* 2002). De esta manera, los cambios en la orientación relativa de la proteasa pueden ser un factor determinante en la regulación de su efecto sobre el dominio de helicasa en los miembros de *Flaviviridae*.

8. BIBLIOGRAFÍA

1. Agarwal, P. K., Billeter, S. R., Rajagopalan, P. T., Benkovic, S. J., & Hammes-Schiffer, S. 2002, "Network of coupled promoting motions in enzyme catalysis", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 99, no. 5, pp. 2794-2799.
2. Alder, B. J. & Wainwright, T. E. 1957, "Phase Transition for a Hard Sphere System", *J.Chem.Phys.*, vol. 27, p. 1208.
3. Bartelma, G. & Padmanabhan, R. 2002, "Expression, purification, and characterization of the RNA 5'-triphosphatase activity of dengue virus type 2 nonstructural protein 3", *Virology*, vol. 299, no. 1, pp. 122-132.
4. Bartenschlager, R., Ilbhorn-Laake, L., Mous, J., & Jacobsen, H. 1993, "Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions", *J.Virol.*, vol. 67, no. 7, pp. 3835-3844.
5. Bartenschlager, R., Lohmann, V., Wilkinson, T., & Koch, J. O. 1995, "Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation", *J.Virol.*, vol. 69, no. 12, pp. 7519-7528.
6. Barth, A. & Zscherp, C. 2002, "What vibrations tell us about proteins", *Q.Rev.Biophys.*, vol. 35, no. 4, pp. 369-430.
7. Basu, G., Kitao, A., Kuki, A., & Go, N. 1998a, "Protein electron transfer reorganization energy spectrum from normal mode analysis. 1. Theory", *Journal of Physical Chemistry B*, vol. 102, no. 11, pp. 2076-2084.
8. Basu, G., Kitao, A., Kuki, A., & Go, N. 1998b, "Protein electron transfer reorganization energy spectrum from normal mode analysis. 2. Application to Ru-modified cytochrome c", *Journal of Physical Chemistry B*, vol. 102, no. 11, pp. 2085-2094.
9. Berman, H., Henrick, K., & Nakamura, H. 2003, "Announcing the worldwide Protein Data Bank", *Nat.Struct.Biol.*, vol. 10, no. 12, p. 980.
10. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. 2000, "The Protein Data Bank", *Nucleic Acids Res.*, vol. 28, no. 1, pp. 235-242.
11. Bielefeldt-Ohmann, H. 1998, "Analysis of antibody-independent binding of dengue viruses and dengue virus envelope protein to human myelomonocytic cells and B lymphocytes", *Virus Res.*, vol. 57, no. 1, pp. 63-79.
12. Bressanelli, S., Stiasny, K., Allison, S. L., Stura, E. A., Duquerroy, S., Lescar, J., Heinz, F. X., & Rey, F. A. 2004, "Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation", *EMBO J.*, vol. 23, no. 4, pp. 728-738.

13. Brünger, A. T., Kuriyan, J., & Karplus, M. 1987, "Crystallographic R-factor refinement by molecular dynamics", *Science* no. 235, pp. 458-460.
14. Caruthers, J. M., Johnson, E. R., & McKay, D. B. 2000, "Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 97, no. 24, pp. 13080-13085.
15. Chanprapaph, S., Saparpakorn, P., Sangma, C., Niyomrattanakit, P., Hannongbua, S., Angsuthanasombat, C., & Katzenmeier, G. 2005, "Competitive inhibition of the dengue virus NS3 serine protease by synthetic peptides representing polyprotein cleavage sites", *Biochem.Biophys.Res.Commun.*, vol. 330, no. 4, pp. 1237-1246.
16. Chen, C. J., Kuo, M. D., Chien, L. J., Hsu, S. L., Wang, Y. M., & Lin, J. H. 1997, "RNA-protein interactions: involvement of NS3, NS5, and 3' noncoding regions of Japanese encephalitis virus genomic RNA", *J.Virol.*, vol. 71, no. 5, pp. 3466-3473.
17. Chothia, C. & Lesk, A. M. 1986, "The relation between the divergence of sequence and structure in proteins", *EMBO J.*, vol. 5, no. 4, pp. 823-826.
18. Cui, T., Sugrue, R. J., Xu, Q., Lee, A. K., Chan, Y. C., & Fu, J. 1998, "Recombinant dengue virus type 1 NS3 protein exhibits specific viral RNA binding and NTPase activity regulated by the NS5 protein", *Virology*, vol. 246, no. 2, pp. 409-417.
19. D'Alfonso, G., Tramontano, A., & Lahm, A. 2001, "Structural conservation in single-domain proteins: implications for homology modeling", *J.Struct.Biol.*, vol. 134, no. 2-3, pp. 246-256.
20. Darden, T. & Pedersen, L. 1996, "Use of fast Ewald summation in molecular dynamics simulations", *Abstracts of Papers of the American Chemical Society*, vol. 212, p. 6-COMP.
21. Darden, T., York, D. M., & Pedersen, L. 1993, "Particle mesh Ewald. An $N \cdot \log(N)$ method for Ewald sums in large systems", *J.Chem.Phys.*, vol. 98, pp. 10089-10092.
22. Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G., & Wriggers, W. 2002, "Conformational flexibility of bacterial RNA polymerase", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 99, no. 7, pp. 4296-4301.
23. Delagoutte, E. & von Hippel, P. H. 2002, "Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases", *Q.Rev.Biophys.*, vol. 35, no. 4, pp. 431-478.
24. Drouet, C., Bouillet, L., Csopaki, F., & Colomb, M. G. 1999, "Hepatitis C virus NS3 serine protease interacts with the serpin C1 inhibitor", *FEBS Lett.*, vol. 458, no. 3, pp. 415-418.
25. Dunbrack, R. L., Jr. & Karplus, M. 1993, "Backbone-dependent rotamer library for proteins. Application to side-chain prediction", *J.Mol.Biol.*, vol. 230, no. 2, pp. 543-574.
26. Dunbrack, R. L., Jr. & Karplus, M. 1994, "Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains", *Nat.Struct.Biol.*, vol. 1, no. 5, pp. 334-340.
27. Duong, T. H. & Zakrzewska, K. 1998, "Sequence specificity of bacteriophage 434 repressor-operator complexation", *J.Mol.Biol.*, vol. 280, no. 1, pp. 31-39.
28. Durand, P., Trinquier, G., & Sanejouand, Y. H. 1994, "A new approach for determining low-frequency normal modes in macromolecules", *Biopolymers*, vol. 34, pp. 759-771.
29. Falgout, B., Pethel, M., Zhang, Y. M., & Lai, C. J. 1991, "Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins", *J.Virol.*, vol. 65, no. 5, pp. 2467-2475.
30. Finkelstein, A. V. 1997, "Protein structure: what is it possible to predict now?", *Curr.Opin.Struct.Biol.*, vol. 7, no. 1, pp. 60-71.
31. Finkelstein, A. V. & Reva, B. A. 1991, "A search for the most stable folds of protein chains", *Nature*, vol. 351, no. 6326, pp. 497-499.
32. Frank, J. & Agrawal, R. K. 2000, "A ratchet-like inter-subunit reorganization of the ribosome during translocation", *Nature*, vol. 406, no. 6793, pp. 318-322.
33. Freddolino, P. L., Arkhipov, A. S., Larson, S. B., McPherson, A., & Schulten, K. 2006, "Molecular dynamics simulations of the complete satellite tobacco mosaic virus", *Structure.*, vol. 14, no. 3, pp. 437-449.
34. Gallinari, P., Brennan, D., Nardi, C., Brunetti, M., Tomei, L., Steinkuhler, C., & De, F. R. 1998, "Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus", *J.Virol.*, vol. 72, no. 8, pp. 6758-6769.

35. Gallinari, P., Paolini, C., Brennan, D., Nardi, C., Steinkuhler, C., & De, F. R. 1999, "Modulation of hepatitis C virus NS3 protease and helicase activities through the interaction with NS4A", *Biochemistry*, vol. 38, no. 17, pp. 5620-5632.
36. Gao, M., Lu, H., & Schulten, K. 2002, "Unfolding of titin domains studied by molecular dynamics simulations", *J.Muscle Res.Cell Motil.*, vol. 23, no. 5-6, pp. 513-521.
37. Gibrat, J. F. & Go, N. 1990, "Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion", *Proteins*, vol. 8, no. 3, pp. 258-279.
38. Gollins, S. W. & Porterfield, J. S. 1985, "Flavivirus infection enhancement in macrophages: an electron microscopic study of viral cellular entry", *J.Gen.Virol.*, vol. 66 (Pt 9), pp. 1969-1982.
39. Gollins, S. W. & Porterfield, J. S. 1986, "pH-dependent fusion between the flavivirus West Nile and liposomal model membranes", *J.Gen.Virol.*, vol. 67 (Pt 1), pp. 157-166.
40. Gorbalenya, A. E., Donchenko, A. P., Koonin, E. V., & Blinov, V. M. 1989, "N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases", *Nucleic Acids Res.*, vol. 17, no. 10, pp. 3889-3897.
41. Gorbalenya, A. E. & Koonin, E. V. 1993, "Helicases: amino acid sequence comparisons and structure-function relationship", *Curr.Opin.Struct.Biol.*, vol. 3, no. 3, pp. 419-429.
42. Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., & Rice, C. M. 1993a, "Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites", *J.Virol.*, vol. 67, no. 5, pp. 2832-2843.
43. Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M., & Rice, C. M. 1993b, "Expression and identification of hepatitis C virus polyprotein cleavage products", *J.Virol.*, vol. 67, no. 3, pp. 1385-1395.
44. Grayson, P., Tajkhorshid, E., & Schulten, K. 2003, "Mechanisms of selectivity in channels and enzymes studied with interactive molecular dynamics", *Biophys.J.*, vol. 85, no. 1, pp. 36-48.
45. Gu, B., Liu, C., Lin-Goerke, J., Maley, D. R., Gutshall, L. L., Feltenberger, C. A., & Del Vecchio, A. M. 2000, "The RNA helicase and nucleotide triphosphatase activities of the bovine viral diarrhea virus NS3 protein are essential for viral replication", *J.Virol.*, vol. 74, no. 4, pp. 1794-1800.
46. Halgren, T. A. 1995, "Potential energy functions", *Curr.Opin.Struct.Biol.*, vol. 5, no. 2, pp. 205-210.
47. Hall, M. C. & Matson, S. W. 1999, "Helicase motifs: the engine that powers DNA unwinding", *Mol.Microbiol.*, vol. 34, no. 5, pp. 867-877.
48. Hammes-Schiffer, S. 2002, "Impact of enzyme motion on activity", *Biochemistry*, vol. 41, no. 45, pp. 13335-13343.
49. Hammes-Schiffer, S. & Benkovic, S. J. 2006, "Relating protein motion to catalysis", *Annu.Rev.Biochem.*, vol. 75, pp. 519-541.
50. Hammes-Schiffer, S. & Watney, J. B. 2006, "Hydride transfer catalysed by Escherichia coli and Bacillus subtilis dihydrofolate reductase: coupled motions and distal mutations", *Philos.Trans.R.Soc.Lond B Biol.Sci.*, vol. 361, no. 1472, pp. 1365-1373.
51. Harrison, R. W. 1984, "Variational calculation of the normal modes of a large macromolecule: methods and some initial results", *Biopolymers*, vol. 23, no. 12, pp. 2943-2949.
52. Hayward, S. & Berendsen, H. J. 1998, "Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme", *Proteins*, vol. 30, no. 2, pp. 144-154.
53. Howe, A. Y., Chase, R., Taremi, S. S., Risano, C., Beyer, B., Malcolm, B., & Lau, J. Y. 1999, "A novel recombinant single-chain hepatitis C virus NS3-NS4A protein with improved helicase activity", *Protein Sci.*, vol. 8, no. 6, pp. 1332-1341.
54. Humphrey, W., Dalke, A., & Schulten, K. 1996, "VMD: visual molecular dynamics", *J.Mol.Graph.*, vol. 14, no. 1, pp. 33-38.
55. Ishak, R., Tovey, D. G., & Howard, C. R. 1988, "Morphogenesis of yellow fever virus 17D in infected cell cultures", *J.Gen.Virol.*, vol. 69 (Pt 2), pp. 325-335.
56. Ishida, H., Jochi, Y., & Kidera, A. 1998, "Dynamic structure of subtilisin-eglin c complex studied by normal mode analysis", *Proteins*, vol. 32, no. 3, pp. 324-333.
57. Ishido, S., Fujita, T., & Hotta, H. 1998, "Complex formation of NS5B with NS3 and NS4A proteins

- of hepatitis C virus", *Biochem.Biophys.Res.Commun.*, vol. 244, no. 1, pp. 35-40.
58. Isralewitz, B., Gao, M., & Schulten, K. 2001, "Steered molecular dynamics and mechanical functions of proteins", *Curr.Opin.Struct.Biol.*, vol. 11, no. 2, pp. 224-230.
 59. Johansson, A., Hubatsch, I., Akerblom, E., Lindeberg, G., Winiwarter, S., Danielson, U. H., & Hallberg, A. 2001a, "Inhibition of hepatitis C virus NS3 protease activity by product-based peptides is dependent on helicase domain", *Bioorg.Med.Chem.Lett.*, vol. 11, no. 2, pp. 203-206.
 60. Johansson, M., Brooks, A. J., Jans, D. A., & Vasudevan, S. G. 2001b, "A small region of the dengue virus-encoded RNA-dependent RNA polymerase, NS5, confers interaction with both the nuclear transport receptor importin-beta and the viral helicase, NS3", *J.Gen.Virol.*, vol. 82, no. Pt 4, pp. 735-745.
 61. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., & Schulten, K. 1999, "NAMD2: Greater scalability for parallel molecular dynamics", *Journal of Computational Physics*, vol. 151, no. 1, pp. 283-312.
 62. Kanai, A., Tanabe, K., & Kohara, M. 1995, "Poly(U) binding activity of hepatitis C virus NS3 protein, a putative RNA helicase", *FEBS Lett.*, vol. 376, no. 3, pp. 221-224.
 63. Kapoor, M., Zhang, L., Ramachandra, M., Kusukawa, J., Ebner, K. E., & Padmanabhan, R. 1995, "Association between NS3 and NS5 proteins of dengue virus type 2 in the putative RNA replicase is linked to differential phosphorylation of NS5", *J.Biol.Chem.*, vol. 270, no. 32, pp. 19100-19106.
 64. Kidera, A. & Go, N. 1990, "Refinement of protein dynamic structure: normal mode refinement", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 87, no. 10, pp. 3718-3722.
 65. Kidera, A. & Go, N. 1992, "Normal mode refinement: crystallographic refinement of protein dynamic structure. I. Theory and test by simulated diffraction data", *J.Mol.Biol.*, vol. 225, no. 2, pp. 457-475.
 66. Kidera, A., Inaka, K., Matsushima, M., & Go, N. 1992, "Normal mode refinement: crystallographic refinement of protein dynamic structure. II. Application to human lysozyme", *J.Mol.Biol.*, vol. 225, no. 2, pp. 477-486.
 67. Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C., & Caron, P. R. 1998, "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding", *Structure.*, vol. 6, no. 1, pp. 89-100.
 68. Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., & Thomson, J. A. 1996, "Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide", *Cell*, vol. 87, no. 2, pp. 343-355.
 69. Kimura, T., Gollins, S. W., & Porterfield, J. S. 1987, "[Early interaction of the flavivirus West Nile virus with mouse macrophage cell line P 388D1: pH-dependent fusion between the virus particles and cell membranes]", *Uirusu*, vol. 37, no. 1, pp. 103-109.
 70. Kolykhalov, A. A., Mihalik, K., Feinstone, S. M., & Rice, C. M. 2000, "Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo", *J.Virol.*, vol. 74, no. 4, pp. 2046-2051.
 71. Krebs, W. G., Alexandrov, V., Wilson, C. A., Echols, N., Yu, H., & Gerstein, M. 2002, "Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic", *Proteins*, vol. 48, no. 4, pp. 682-695.
 72. Lam, A. M., Rypma, R. S., & Frick, D. N. 2004, "Enhanced nucleic acid binding to ATP-bound hepatitis C virus NS3 helicase at low pH activates RNA unwinding", *Nucleic Acids Res.*, vol. 32, no. 13, pp. 4060-4070.
 73. Laskowski, R. A., Mac Arthur, M. W., Moss, D. S., & Thornton, J. M. 1993, "PROCHECK: a program to check the stereochemical quality of protein structures", *J.Appl.Cryst.*, vol. 26, pp. 283-291.
 74. Lin, C. & Kim, J. L. 1999, "Structure-based mutagenesis study of hepatitis C virus NS3 helicase", *J.Virol.*, vol. 73, no. 10, pp. 8798-8807.
 75. Lindahl, E. 2001, *Computational Modeling of Biological Membrane and Interface Dynamics*, Doctoral Dissertation, Royal Institute of Technology, Department of Physics.

76. Lu, H., Isralewitz, B., Krammer, A., Vogel, V., & Schulten, K. 1998, "Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation", *Biophys.J.*, vol. 75, no. 2, pp. 662-671.
77. MacKerell, A. D. 2002, "CHARMM biomolecular force field: Recent developments and future directions", *Abstracts of Papers of the American Chemical Society*, vol. 224, p. U471.
78. MacKerell, A. D., Bashford, D., Bellot, M., Dunbrack, L., Evanseck, J., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, I. W. E., Roux, B., Schlenkrich, M., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. 1998, "All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field", *J.Phys.Chem.B*, vol. 102, pp. 3586-3616.
79. MacKerell, A. D., Bashford, D., Bellot, M., Dunbrack, L., Evanseck, J., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Roux, B., Schlenkrich, M., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. 1992, "Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations", *FASEB*, vol. 6, p. 143.
80. Matsumoto, A., Tomimoto, M., & Go, N. 1999, "Dynamical structure of transfer RNA studied by normal mode analysis", *Eur.Biophys.J.*, vol. 28, no. 5, pp. 369-379.
81. Matusan, A. E., Kelley, P. G., Pryor, M. J., Whisstock, J. C., Davidson, A. D., & Wright, P. J. 2001, "Mutagenesis of the dengue virus type 2 NS3 proteinase and the production of growth-restricted virus", *J.Gen.Virol.*, vol. 82, no. Pt 7, pp. 1647-1656.
82. McCammon, J. A., Gelin, B. R., & Karplus, M. 1977, "Dynamics of folded proteins", *Nature*, vol. 267, no. 5612, pp. 585-590.
83. McCammon, J. A. & Harvey, S. C. 1987, *Dynamics of Proteins and Nucleic Acids* Cambridge University Press, Cambridge, UK.
84. Miller, D. W. & Agard, D. A. 1999, "Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease", *J.Mol.Biol.*, vol. 286, no. 1, pp. 267-278.
85. Miller, R. H. & Purcell, R. H. 1990, "Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups", *Proc.Natl.Acad.Sci.U.S A*, vol. 87, no. 6, pp. 2057-2061.
86. Mirny, L. A. & Shakhnovich, E. I. 1996, "How to derive a protein folding potential? A new approach to an old problem", *J.Mol.Biol.*, vol. 264, no. 5, pp. 1164-1179.
87. Mirny, L. A. & Shakhnovich, E. I. 1998, "Protein structure prediction by threading. Why it works and why it does not", *J.Mol.Biol.*, vol. 283, no. 2, pp. 507-526.
88. Mongan, J. & Case, D. A. 2005, "Biomolecular simulations at constant pH", *Curr.Opin.Struct.Biol.*, vol. 15, no. 2, pp. 157-163.
89. Morgenstern, K. A., Landro, J. A., Hsiao, K., Lin, C., Gu, Y., Su, M. S., & Thomson, J. A. 1997, "Polynucleotide modulation of the protease, nucleoside triphosphatase, and helicase activities of a hepatitis C virus NS3-NS4A complex isolated from transfected COS cells", *J.Virol.*, vol. 71, no. 5, pp. 3767-3775.
90. Mouawad, L. & Perahia, D. 1996, "Motions in hemoglobin studied by normal mode analysis and energy minimization: evidence for the existence of tertiary T-like, quaternary R-like intermediate structures", *J.Mol.Biol.*, vol. 258, no. 2, pp. 393-410.
91. Moult, J. & Unger, R. 1991, "An analysis of protein folding pathways", *Biochemistry*, vol. 30, no. 16, pp. 3816-3824.
92. Nakayama, H. 2002, "RecQ family helicases: roles as tumor suppressor proteins", *Oncogene*, vol. 21, no. 58, pp. 9008-9021.
93. Niyomrattanakit, P., Winoyanuwattikun, P., Chanprapaph, S., Angsuthanasombat, C., Panyim, S., & Katzenmeier, G. 2004, "Identification of residues in the dengue virus type 2 NS2B cofactor that are critical for NS3 protease activation", *J.Virol.*, vol. 78, no. 24, pp. 13708-13716.
94. Norberg, J. & Nilsson, L. 2003, "Advances in biomolecular simulations: methodology and recent applications", *Q.Rev.Biophys.*, vol. 36, no. 3, pp. 257-306.
95. Pande, V. S. & Rokhsar, D. S. 1999, "Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G",

- Proc.Natl.Acad.Sci.U.S.A.*, vol. 96, no. 16, pp. 9062-9067.
96. Pang, P. S., Jankowsky, E., Planet, P. J., & Pyle, A. M. 2002, "The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding", *EMBO J.*, vol. 21, no. 5, pp. 1168-1176.
 97. Perahia, D. & Mouawad, L. 1995, "Computation of low-frequency normal modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin", *Comput.Chem.*, vol. 19, no. 3, pp. 241-246.
 98. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., & Schulten, K. 2005, "Scalable molecular dynamics with NAMD", *J.Comput.Chem.*, vol. 26, no. 16, pp. 1781-1802.
 99. Rahman, A. 1964, "Correlations in the Motion of Atoms in Liquid Argon", *Phys.Rev.*, vol. 136, p. A405-A411.
 100. Ramos-Castaneda, J., Imbert, J. L., Barron, B. L., & Ramos, C. 1997, "A 65-kDa trypsin-sensible membrane cell protein as a possible receptor for dengue virus in cultured neuroblastoma cells", *J.Neurovirol.*, vol. 3, no. 6, pp. 435-440.
 101. Reed, K. E., Grakoui, A., & Rice, C. M. 1995, "Hepatitis C virus-encoded NS2-3 protease: cleavage-site mutagenesis and requirements for bimolecular cleavage", *J.Virol.*, vol. 69, no. 7, pp. 4127-4136.
 102. Reva, B., Finkelstein, A., & Topiol, S. 2002, "Threading with chemostructural restrictions method for predicting fold and functionally significant residues: application to dipeptidylpeptidase IV (DPP-IV)", *Proteins*, vol. 47, no. 2, pp. 180-193.
 103. Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C., & Harrison, S. C. 1995, "The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution", *Nature*, vol. 375, no. 6529, pp. 291-298.
 104. Ryckaert, J. P., Cicciotti, G., & Berendsen, H. J. C. 1977, "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes", *J.Comput.Phys.*, vol. 23, pp. 327-341.
 105. Sali, A. & Blundell, T. L. 1993, "Comparative protein modelling by satisfaction of spatial restraints", *J.Mol.Biol.*, vol. 234, no. 3, pp. 779-815.
 106. Sardana, V. V., Blue, J. T., Zugay-Murphy, J., Sardana, M. K., & Kuo, L. C. 1999, "A uniquely purified HCV NS3 protease and NS4A(21-34) peptide form a highly active serine protease complex in peptide hydrolysis", *Protein Expr.Purif.*, vol. 16, no. 3, pp. 440-447.
 107. Selisko, B., Licea, A. F., Becerril, B., Zamudio, F., Possani, L. D., & Horjales, E. 1999, "Antibody BCF2 against scorpion toxin Cn2 from *Centruroides noxius* Hoffmann: primary structure and three-dimensional model as free Fv fragment and complexed with its antigen", *Proteins*, vol. 37, no. 1, pp. 130-143.
 108. Sibanda, B. L., Blundell, T. L., & Thornton, J. M. 1989, "Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering", *J.Mol.Biol.*, vol. 206, no. 4, pp. 759-777.
 109. Soutanas, P. & Wigley, D. B. 2001, "Unwinding the 'Gordian knot' of helicase action", *Trends Biochem.Sci.*, vol. 26, no. 1, pp. 47-54.
 110. Stocks, C. E. & Lobigs, M. 1998, "Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM", *J.Virol.*, vol. 72, no. 3, pp. 2141-2149.
 111. Stone, J., Gullingsrud, J., Grayson, P., & Schulten, K. 2001, "A system for interactive molecular dynamics simulation," in *2001 ACM Symposium on Interactive 3D Graphics*, ACM SIGGRAPH, ed., New York, pp. 191-194.
 112. Suhre, K. & Sanejouand, Y. H. 2004, "ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement", *Nucleic Acids Res.*, vol. 32, no. Web Server issue, p. W610-W614.
 113. Tajkhorshid, E., Aksimentiev, A., Balabin, I., Gao, M., Isralewitz, B., Phillips, J. C., Zhu, F., & Schulten, K. 2003, "Large scale simulation of protein mechanics and function", *Adv.Protein Chem.*, vol. 66, pp. 195-247.
 114. Tama, F. 2003, "Normal mode analysis with simplified models to investigate the global dynamics of biological systems", *Protein Pept.Lett.*, vol. 10, no. 2, pp. 119-132.
 115. Tama, F. & Brooks, C. L., III 2002, "The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus", *J.Mol.Biol.*, vol. 318, no. 3, pp. 733-747.

116. Tama, F., Gadea, F. X., Marques, O., & Sanejouand, Y. H. 2000, "Building-block approach for determining low-frequency normal modes of macromolecules", *Proteins*, vol. 41, no. 1, pp. 1-7.
117. Tama, F. & Sanejouand, Y. H. 2001, "Conformational change of proteins arising from normal mode calculations", *Protein Eng*, vol. 14, no. 1, pp. 1-6.
118. Tama, F., Valle, M., Frank, J., & Brooks, C. L., III 2003, "Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 100, no. 16, pp. 9319-9323.
119. Tirion, M. M. 1996, "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis", *PHYSICAL REVIEW LETTERS*, vol. 77, no. 9, pp. 1905-1908.
120. Tirion, M. M., ben-Avraham, D., Lorenz, M., & Holmes, K. C. 1995, "Normal modes as refinement parameters for the F-actin model", *Biophys.J.*, vol. 68, no. 1, pp. 5-12.
121. Tramontano, A. 1998, "Homology modeling with low sequence identity", *Methods*, vol. 14, no. 3, pp. 293-300.
122. Tramontano, A. 2003, "Of men and machines", *Nat.Struct.Biol.*, vol. 10, no. 2, pp. 87-90.
123. van Dinten, L. C., van, T. H., Gorbalenya, A. E., & Snijder, E. J. 2000, "The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis", *J.Virol.*, vol. 74, no. 11, pp. 5213-5223.
124. Venclovas, C. 2003, "Comparative modeling in CASP5: progress is evident, but alignment errors remain a significant hindrance", *Proteins*, vol. 53 Suppl 6, pp. 380-388.
125. Verlet, L. 1967, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", *Phys.Rev.*, vol. 159, no. 1, pp. 98-103.
126. Wang, W., Donini, O., Reyes, C. M., & Kollman, P. A. 2001, "Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions", *Annu.Rev.Biophys.Biomol.Struct.*, vol. 30, pp. 211-243.
127. Wang, W., Lahser, F. C., Yi, M., Wright-Minogue, J., Xia, E., Weber, P. C., Lemon, S. M., & Malcolm, B. A. 2004, "Conserved C-terminal threonine of hepatitis C virus NS3 regulates autoproteolysis and prevents product inhibition", *J.Virol.*, vol. 78, no. 2, pp. 700-709.
128. Xu, T., Sampath, A., Chao, A., Wen, D., Nanao, M., Chene, P., Vasudevan, S. G., & Lescar, J. 2005, "Structure of the Dengue virus helicase/nucleoside triphosphatase catalytic domain at a resolution of 2.4 Å", *J.Virol.*, vol. 79, no. 16, pp. 10278-10288.
129. Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J. L., Sardana, M., Steinkuehler, C., Tomei, L., De, F. R., Kuo, L. C., & Chen, Z. 1998, "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2 Å resolution structure in a hexagonal crystal form", *Protein Sci.*, vol. 7, no. 4, pp. 837-847.
130. Yao, N., Reichert, P., Taremi, S. S., Prorise, W. W., & Weber, P. C. 1999, "Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase", *Structure*, vol. 7, no. 11, pp. 1353-1363.