

Título: Estructura e interacciones de dominios funcionales del ARN del virus de la Hepatitis C

Autor: Jesús Castillo Martínez

Director: José Gallego Sala

Número de páginas: 272

Resumen:

El virus de la hepatitis C (VHC) pertenece al género *Hepacivirus* de la familia Flaviviridae. Caracterizado por primera vez en 1989, actualmente se extiende a nivel global y afecta alrededor de 71 millones de personas. Es una de las principales causas de cirrosis y cáncer hepático y actualmente no existe vacuna eficaz conocida. La partícula viral del VHC contiene una secuencia de ARN genómico de cadena simple y sentido positivo de unos 9600 nucleótidos, que comprende un único marco abierto de lectura (*ORF*, en sus siglas en inglés), el cual codifica para la poliproteína viral que es escindida por proteasas celulares y virales; esta región está flanqueada a ambos lados por dos regiones no traducidas (*UTR*, en sus siglas en inglés) que contienen secuencias estructuradas esenciales para los procesos de traducción, replicación e infectividad. En concreto, el dominio 3'X de 98 nucleótidos, localizado en el extremo 3' del genoma del virus, contiene dos secuencias absolutamente conservadas: *DLS* (*Dimer Linkage Sequence*, en inglés), que permite la dimerización del ARN viral *in vitro*, y *k*, responsable de la interacción con el subdominio 5BSL3.2, situado en el marco abierto de lectura. Estas dos secuencias se encuentran solapadas, característica que hizo pensar durante mucho tiempo, y basados en estudios de modificaciones químicas y enzimáticas, que el dominio 3'X adquiriría dos conformaciones diferentes: una de dos horquillas (SL1' y SL2') donde la región central de la secuencia *DLS* queda libre en el bucle apical de la horquilla SL2' para permitir la dimerización; y una estructura de tres horquillas (SL1, SL2 y SL3), donde se libera la secuencia *k* en el bucle apical de la horquilla SL2 para su interacción con el subdominio 5BSL3.2. De tal forma que este dominio actuaría como un ribointerruptor para regular los procesos de traducción y replicación mediante el intercambio de estas dos conformaciones. El subdominio 5BSL3.2 actúa como elemento de regulación *in cis*, permitiendo guiar las diferentes fases del ciclo viral mediante una red de interacciones ARN-ARN de larga distancia, entre las que se encuentra el complejo 3'X-5BSL3.2, a través de la interacción entre sus secuencias *k* y *k'*, respectivamente; y su interacción con el subdominio IIIId del *IRES* (localizado en la región no traducida 5'), contacto que bloquearía el proceso de traducción.

Resultados previos del grupo indicaron que el dominio 3'X adopta una conformación constituida por dos horquillas (SL1' y SL2') estabilizadas por apilamiento coaxial que expone la secuencia *DLS* en un bucle apical, pero entierra la secuencia *k* en una doble hélice. No se encontraron indicios de cambio conformacional o mezcla de conformaciones. Esta

conformación permite al dominio formar homodímeros mediante la secuencia palindrómica *DLS* y también, sorprendentemente, interaccionar con 5BSL3.2. Los resultados de RMN para esta interacción mostraron que el dominio 3'X mantiene estables las estructuras de la horquilla SL1' y la región basal de SL2' tras el contacto con el subdominio 5BSL3.2, encontrándose únicamente perturbaciones en la región apical de la horquilla SL2', donde se realiza la interacción entre las secuencias k y k'.

En esta tesis doctoral hemos analizado, en primer lugar, la estructura y la función del dominio 3'X, mediante experimentos de espectroscopía de resonancia magnética nuclear (RMN), ensayos de desnaturalización térmica medida por espectroscopía de ultravioleta (UV) y electroforesis, combinados con mutagénesis y ensayos funcionales. Se diseñaron un set de seis mutantes capaces de estabilizar (U_3C , $U_{55}C$ y $U_3C/U_{55}C$) o desestabilizar (U_3G , $G_{50}C/C_{52}G$ y $U_3G/G_{50}C/C_{52}G$) la conformación de dos horquillas del dominio salvaje. Los mutantes estabilizadores, así como el mutante simple desestabilizador U_3G , mostraron un comportamiento en estado monomérico y dimérico similar al encontrado a la secuencia salvaje, al contrario que los mutantes desestabilizadores, indicando que estos podrían adquirir una conformación similar al dominio 3'X. Estos resultados se corroboraron mediante espectroscopía de RMN, demostrando que los mutantes estabilizadores y U_3G adquieren una conformación de dos horquillas, mientras que los mutantes desestabilizadores estaban capacitados para formar la horquilla SL1, pero no se encontraron indicios de la formación de la horquilla SL2'. Basados en ensayos de desnaturalización térmica medida por espectroscopía de UV, se observó una temperatura de fusión correspondiente a la horquilla SL1' para todos los mutantes y una segunda temperatura de fusión correspondiente a la horquilla SL2' únicamente para los mutantes estabilizadores y U_3G , siendo esta mayor a medida que se incrementaba la estabilidad de la región basal de la horquilla SL2' mediante las mutaciones introducidas. Además, los mutantes estabilizadores, así como U_3G , estaban capacitados para interaccionar con el subdominio 5BSL3.2 mediante un proceso similar al encontrado en la secuencia salvaje, ya que solo se hallaron perturbaciones en la región apical de la secuencia SL2' tras la formación del complejo en los ensayos de RMN. Sin embargo, los mutantes desestabilizadores no estaban capacitados para interaccionar con el subdominio 5BSL3.2. Finalmente, los mutantes estabilizadores, así como el mutante simple U_3G , dieron lugar a mejores niveles de traducción y se observó una posible relación entre la estabilidad de la región basal de la horquilla SL2' con el proceso de replicación.

Tras determinar la estructura y función del dominio 3'X y comprobar lo inusual que era su interacción con el subdominio 5BSL3.2 y su importancia en el ciclo viral, se decidió estudiar

la conformación tridimensional de este complejo, así como la conformación del subdominio 5BSL3.2, ya que aún no existían datos estructurales para los mismo. Estos se analizaron mediante experimentos de espectroscopía de resonancia magnética nuclear (RMN), dispersión de rayos X de ángulo pequeño (SAXS, en sus siglas en inglés) y electroforesis. Se diseñaron dos secuencias mutantes correspondientes al dominio 3'X y la horquilla SL2' capaces de interactuar con el subdominio 5BSL3.2 pero deficientes en homodimerización, insertando dos mutaciones en la región apical de la horquilla SL2', dentro de la secuencia *DLS*, sin afectar a la secuencia *k*. Los ensayos se realizaron en ausencia y presencia de 6 mM magnesio, ya que fueron las mejores condiciones para establecer el complejo 3'X-5BSL3.2

Los resultados indicaron que el dominio 3'X adquiere una conformación tridimensional significativamente estable, formada por dos horquillas (SL1' y SL2') apiladas coaxialmente. Esta estructura expone la región central de la secuencia *DLS* en el bucle apical de SL2', así como los tres nucleótidos terminales del genoma del virus desapareados en la intersección entre SL1' y SL2'. Esta conformación es compatible con replicación, ya que se han observado que la presencia de al menos tres nucleótidos desapareados tras una estructura estable favorece este proceso. El magnesio es uno de los principales factores que estabiliza las estructuras de ARN y permite las interacciones ARN-ARN, por lo que se le consideró uno de los responsables de permitir el cambio conformacional del dominio 3'X. Mediante los resultados de dispersión de rayos X de ángulo pequeño hemos demostrado que este catión no provoca dicho cambio conformacional, generando únicamente una reducción del tamaño y la flexibilidad del dominio.

El subdominio 5BSL3.2, por otra parte, adoptó una estructura tridimensional muy estable con forma de "L", similar a la de dominios de ARN conservados de *Flavivirus*, importantes para el proceso de replicación. Esta conformación estaba constituida por dos regiones de doble hélice separadas por una protuberancia de 8 nucleótidos, más un bucle apical de 12 nucleótidos. Esta gran protuberancia fue la responsable de generar la curvatura entre las dos hélices que componen el sistema, siendo más aguda a mayor fuerza iónica. Además, en el bucle apical se localizó el segmento 5'-CACAG-3', dentro de la secuencia *k'*, presente en otros dominios de *Flavivirus* e imprescindible para su interacción con las polimerasas virales. Esta estructura exhibió baja flexibilidad en ausencia y presencia de magnesio, lo que parece indicar que tanto la protuberancia como el bucle apical podrían estar estructurados.

Para el estudio del complejo formado entre el dominio 3'X y el subdominio 5BSL3.2 se siguió una estrategia simplificada de espectroscopía de RMN basada en la inserción de una base inosina en una posición específica de la horquilla SL2'. Esta base posee la capacidad de interactuar con uracilo y citosina, pero su desplazamiento químico en los ensayos de RMN se encuentra muy alejado de los desplazamientos químicos de los pares de bases canónicos y, además, es diferente dependiendo de si aparea con uracilo o con citosina. Estas características nos permitieron distinguir la estructura adquirida por la horquilla SL2' tras su interacción con 5BSL3.2. Finalmente, se analizó el complejo 3'X-5BSL3.2 mediante ensayos de dispersión de rayos X de ángulo pequeño, para corroborar los resultados obtenidos en RMN.

Este complejo formó una estructura inusual, consistente en una intersección de besos con forma de "Y" cuyo eje inferior correspondía al subdominio SL1' y la doble hélice basal de SL2', y cuyos dos ejes superiores estaban compuestos por el subdominio 5BSL3.2 y una nueva horquilla de 17 nucleótidos formada por nucleótidos *DLS* liberados por la interacción y donde se localizaba en nuevo par de bases formado por la inosina tras la formación del complejo. Las secuencias *k* y *k'* formaron una doble hélice intermolecular de 7 nucleótidos que se estabilizó mediante apilamiento coaxial con las dobles hélices que la flanqueaban, correspondientes a la horquilla apical de 5BSL3.2 y una nueva doble hélice intermedia de la horquilla SL2', surgida tras el reordenamiento de esta región.

Guiados por las evidencias experimentales de la interacción entre 5BSL3.2 y NS5B, y por las similitudes observadas con los flavivirus, construimos un modelo de la estructura en "L" del subdominio 5BSL3.2 unida al dominio del pulgar de la polimerasa NS5B. En este modelo, la complementariedad de tamaño entre la estructura proteica y el ARN resultó ser excelente y, en línea con lo observado en *Flavivirus*, el bucle apical de 5BSL3.2 conteniendo la secuencia CACAG se colocó cerca de una región rica en argininas del pulgar, conservada en distintos virus del género *Hepacivirus*. La protuberancia de 5BSL3.2 quedó orientada hacia los dedos de la polimerasa, mientras que la doble hélice inferior se localizó cerca de la palma. La formación de este complejo no bloquea ni el canal de unión de la hebra molde de la polimerasa NS5B y ni el canal de salida de la doble hebra de ARN. Por tanto, este modelo de interacción NS5B-5BSL3.2 es compatible con la entrada de los tres nucleótidos terminales desapareados del dominio 3'X en el sitio activo de la polimerasa, donde actuarían como molde para iniciar la síntesis de la hebra negativa del ARN viral.

Estos modelos explicarían por qué el complejo desfavorecería la replicación. Cuando el complejo no está formado, la estructura en L del subdominio 5BSL3.2 se uniría a la

polimerasa NS5B, estado con la suficiente flexibilidad para que el dominio 3'X alcance el canal de entrada de la hebra molde para iniciar la síntesis de la hebra negativa. Cuando se forma el complejo 3'X-5BSL3.2 relativamente estable, una hipotética unión de la polimerasa a la estructura en Y del complejo no permitiría la entrada de los nucleótidos terminales a este canal y, viceversa, la unión de la polimerasa a los nucleótidos terminales estaría desfavorecida por la voluminosa región superior de la estructura en Y del complejo. Esta situación tendría dos consecuencias: el subdominio 5BSL3.2 no interactuaría con el *IRES*, liberándolo y permitiendo la traducción de la poliproteína viral; y el bloqueo de la secuencia palindrómica *DLS*.

Aún se desconoce como la homodimerización realizada a través de la secuencia *DLS* afecta realmente al ciclo viral, pero esta investigación y la de otros autores proponen que puede regular la transición entre traducción y replicación. A altas concentraciones de ARN viral, la formación de homodímeros impediría la formación del complejo 3'X-5BSL3.2 al bloquear la secuencia *k*, favoreciendo la unión de la polimerasa a 5BSL3.2 y el acceso de los nucleótidos terminales de 3'X al canal de entrada, estabilizando el proceso de síntesis de la hebra negativa. Por el contrario, el complejo bloquearía la formación de homodímeros desfavoreciendo la replicación. Recientemente, se han detectado otras secuencias palindrómicas en el dominio SLA de *Flavivirus* que promueven la dimerización, apuntando a un posible mecanismo de regulación común de la familia Flaviviridae.