
MODELACIÓN ESTRUCTURAL DE LA PROTEÍNA DE LA CÁPSIDE DEL VIRUS A DE LA PAPA (PVA, POTYVIRUS)

STRUCTURAL MODELLING OF POTATO VIRUS A (PVA, POTYVIRUS) COAT PROTEIN

Pablo Gutiérrez^{1, 3}, Sara Bastos-Aristizábal^{1,4}, Mauricio Marín^{2,5}

Resumen

A diferencia de lo que ocurre con diversos virus icosaédricos, la estructura a alta resolución de la cápside de los virus flexuosos de plantas pertenecientes a la familia Potyviridae no ha podido ser determinada aún. Los potyvirus son un grupo de gran importancia económica en la agricultura al afectar cultivos como papa, tomate, tabaco, papaya y caña de azúcar, entre muchos otros; por lo cual la comprensión de su estructura puede arrojar información valiosa para lograr un conocimiento más detallado de sus mecanismos biológicos, con miras al diseño de estrategias de control. En este trabajo se presenta un modelo de la estructura tridimensional de la región central de la proteína de la cápside del virus A de la papa (**PVA**), utilizando una combinación de herramientas de predicción de estructura secundaria y *docking*. El modelo presentado tiene dimensiones compatibles con la estructura de baja resolución obtenida en otros estudios mediante microscopía electrónica y será de gran utilidad en el diseño de experimentos de mutagénesis dirigida, enfocados en el estudio del ensamblaje de la partícula viral y como base para modelar la estructura de otras especies potyvirales de importancia actual en Colombia como el virus Y de la papa (**PVY**), virus de la malformación de las hojas del tomate de árbol (**TaLMV**) y el virus de la mancha anular de la papaya (**PRSV**).

Palabras clave: cápside viral, modelo estructural, potyvirus, PVA

Abstract

In contrast to icosahedric viruses, a high-resolution structure of flexuous viruses of plants belonging to the family Potyviridae has yet to be obtained. Potyviruses are economically important in agriculture due to their ability to infect crops such as potato, tomato, tobacco, papaya, and sugar cane, among many others; thus an understanding of its three-dimensional structure might provide valuable insight into biological mechanisms, with the aim of designing control strategies. In this study, we present a model for the three-dimensional structure of the core region of Potato virus A (**PVA**), using a combination tool to predict the secondary structure and docking. The model presented has dimensions compatible with the low resolution structure in other studies using electron microscopy and will be highly useful in the design of experiments of directed mutagenesis aimed at understanding potyvirus assembly and as a basis for modelling the structure of other important related virus species of importance in Colombia, such as Potato virus Y (**PVY**), Tamarillo leaf malformation virus (**TaLMV**), and Papaya ringspot virus (**PRSV**).

Key words: virus capsid, structural model, Potyvirus, PVA

Recibido: agosto 2010; aceptado: mayo 2011.

¹ Laboratorio de Microbiología Industrial, Facultad de Ciencias. Universidad Nacional de Colombia (Sede Medellín). Medellín (Antioquia), Colombia.

² Laboratorio de Biología Celular y Molecular, Facultad de Ciencias. Universidad Nacional de Colombia (Sede Medellín). Medellín (Antioquia), Colombia.

Correos electrónicos: ³ <paguties@unal.edu.co>; ⁴ <sarabastos111@gmail.com>; ⁵ <mamarinm@unal.edu.co>.

INTRODUCCIÓN

El género *Potyvirus* de la familia Potyviridae es uno de los grupos de virus más numerosos y limitantes de diversos cultivos agrícolas. Los potyvirus se caracterizan por tener estructura de varillas flexuosas de 650 a 900 nm de longitud y 11 a 13 nm de ancho. Su genoma es de ARN de cadena sencilla positiva con tamaño aproximado de 9-10 kpb, poliadenilados en su extremo 3' y asociados a una proteína unida covalentemente al extremo 5' (VPg) (Singh et al. 2008, Van-Regenmortel et al. 2000). El genoma de los miembros del género *Potyvirus* codifica para una poliproteína precursora de 350 kDa (Riechmann et al. 1995), que es subsecuentemente procesada en siete proteínas pequeñas: **P1**, componente asistente (Helper Component); **P3**, de inclusión cilíndrica (**CI**); inclusión nuclear A (**NIa**); inclusión nuclear B (**NIb**); proteína de cápside (**CP**); y dos proteínas putativas conocidas como **6K1** y **6K2**. El clivaje es llevado a cabo por acción de tres proteasas de origen viral: la proteinasa P1 y la proteinasa del componente asistente (HC-pro), catalizan solo reacciones autoproteolíticas en los extremos C terminales (Carrington et al. 1989, Urcuqui-Inchima et al. 2001, Verchot et al. 1991) y los clivajes restantes son catalizados mediante mecanismos trans y autoproteolíticos mediados por la proteína de inclusión nuclear (NIa-Pro), un homólogo de la proteinasa del picornavirus 3C (Carrington y Dougherty 1987, Langenberg y Zhang 1997). El procesamiento y función de todas estas proteínas es aún controvertido pero se cree que muchas de ellas son multifuncionales (Riechmann et al. 1992, Verchot y Carrington 1995).

La principal función de la proteína CP es encapsidar el ARN viral y evitar que sea degradado. Se cree que la interacción específica entre el ARN viral y CP es el principal factor influyente en la encapsidación selectiva del genoma viral en medio de gran cantidad de ARN celular (Duggal y Hall 1993, Lin et al.

2009). Sin embargo, Merits (1998), reportó que varias proteínas virales; entre ellas CP, también pueden unirse de manera inespecífica al ARN. La CP de los potyvirus se ha dividido en tres dominios principales: el N-terminal variable y C-terminal, expuestos al exterior del virión y sensibles a tratamientos con tripsina y el dominio globular central, el cual es el más conservado. El dominio N-terminal, aunque no está involucrado directamente en la arquitectura del virión, es necesario para la transmisión de los potyvirus por insectos vectores de la familia *Aphididae* y además es el dominio que contiene más epítopes virus-específicos (Lin et al. 2009, Mink et al. 1999, Rajamaki et al. 1998).

Pese a que en las últimas décadas ha habido gran progreso en la determinación estructural de virus icosaédricos, el estudio de los virus filamentosos, con la excepción de los tobamovirus y los bacteriófagos filamentosos, ha tenido poco progreso (Ehrenfeld et al. 2008). Es así como no existe aún un modelo tridimensional que permita explicar los patrones de conservación de la secuencia, ni la manera como las unidades se asocian para formar la estructura viral. Esta información es fundamental para el soporte de estudios sobre la biología del virus y su interacción con hospedantes y vectores biológicos. En este trabajo, se presenta el modelo de la estructura tridimensional de la región central de la proteína de la cápside del PVA (virus A de la papa), un potyvirus que causa mosaicos suaves en diferentes variedades de papa (*Solanum tuberosum*), y para el cual se han propuesto algunos modelos que contrastados con estudios recientes de microscopía electrónica, no concuerdan con las dimensiones esperadas para los componentes de la cápside viral (Kendall et al. 2008). Este modelo servirá de base para la construcción de estructuras por homología de otros potyvirus de importancia económica en Colombia, especialmente de aquellas especies virales que afectan los cul-

tivos de solanáceas como la papa, el tabaco y el tomate de árbol.

MATERIALES Y MÉTODOS

Alineamientos de secuencia y predicción de la estructura secundaria. El alineamiento de secuencias se realizó con el módulo CLUSTAL del programa BIOEDIT (Hall 1999) utilizando la matriz de sustitución BLOSUM62 (Henikoff y Henikoff 1992). Se utilizaron las siguientes secuencias de potyvirus, obtenidas de GenBank: virus de la mancha anular de la papaya (**PRV**), gi: 9629244, virus *Plum pox* (**PPV**), gi: 9626508, virus Y de la papa (**PVY**), gi: 9627728, virus del mosaico del nabo (**TuMV**), gi: 56407093, virus del mosaico de la soya (**SMV**), gi:12018225, virus del grabado del tabaco (**TEV**), gi: 9790340. Para la predicción de estructura secundaria se utilizaron tres métodos diferentes: GORIV (Doolittle et al. 1996), HNN y PredictP (Rost et al. 2004). La predicción de regiones desordenadas se hizo con el programa DisEMBL (Linding et al. 2003a, b) con una ventana de ocho residuos para el algoritmo Savitzky-Golay y el valor umbral de 1,20. La entropía del alineamiento múltiple fue medida con la ecuación de Shannon y se gráfico como un valor de Z utilizando un promedio local de 5 aminoácidos. La hidrofobicidad fue estimada mediante un gráfico de los valores de Hopp y Woods para cada aminoácido con un promedio local de 9 aminoácidos (Hopp y Woods 1981). La confiabilidad de las predicciones fue verificada utilizando la estructura del virus del mosaico del tabaco (**TMV**) como control (PDBid: 2TMV) (Namba y Stubbs 1986).

Modelamiento estructural. La estructura del núcleo central de la proteína de la cápside del potyvirus fue modelada utilizando una estrategia de *docking* secuencial, que consistió en el empaquetamiento gradual de cada una de las hélices utilizando una rutina de *docking* de cuerpo rígido con el programa Autodock 4,0 (Morris et al. 1998).

Para el *docking* se utilizó un algoritmo genético lamarquiano y se realizaron 100 simulaciones con una población inicial de 150 moléculas y un máximo de 2.500.000 evaluaciones de energía. Las tasas de mutación y entrecruzamiento fueron de 0,02 y 0,8, respectivamente. A cada modelo obtenido durante cada paso de *docking* se le adicionaron las regiones conectoras con el programa MODELLER 9v2 (Sali y Blundell 1995) y se realizó minimización de energía con el programa NAMM utilizando el campo de fuerzas CHARMM27 (Phillips et al. 2005). El modelo final fue minimizado mediante un calentamiento gradual de 0 a 310 K en condiciones periódicas de solvente. La calidad del modelo fue evaluada con el programa WHAT IF (Hoofst et al. 1996). Todas las ilustraciones estructurales fueron generadas con PyMOL (DeLano 2002).

RESULTADOS

Validación de las herramientas de predicción de estructura secundaria. Antes de llevar a cabo la predicción de estructura secundaria para la proteína de la cápside de potyvirus, se verificó su precisión y confiabilidad utilizando el virus del TMV como control (datos no mostrados). Varios estudios han establecido que la arquitectura de la cápside de los miembros de la familia Potyviridae es muy similar a la de los virus de varilla rígida como el TMV, del cual se encuentra disponible la estructura cristalográfica (Nemykh et al. 2008, Shukla y Ward 1989). La parte central de la cápside del TMV está compuesta por un agregado de cuatro hélices, LS, RS, RR y LR, que comprende los residuos 19-32, 38-48, 111-135 y 73-87, respectivamente. También existen otros pequeños segmentos helicoidales y una lámina beta ubicados en los extremos N y C terminal. TMV posee un asa interna de estructura desordenada en la forma monomérica que comprende los residuos 97-113 y está involucrada en la unión a los grupos fosfato del ARN.

Las predicciones de estructura secundaria de los programas GOR, HNN y PHD fueron contrastadas con la estructura real del TMV. Los tres programas funcionaron bien en la predicción de las hélices del dominio central mas no en su extensión exacta, lo que es de esperar ya que los métodos para la predicción de estructura secundaria tienen precisión entre el 60 y el 80% (Rost 2001). Se encontró que al establecer un consenso entre las tres herramientas de predicción se puede estimar la posición de las hélices centrales con un margen de error de ± 3 residuos. Sin embargo, los tres programas fallaron en la predicción de las hélices externas, lo cual es un indicativo de que su plegamiento depende de interacciones de tipo terciario que no pueden ser predichas por este tipo de programas. De la misma manera, la predicción de estructuras beta laminares tuvo buena correspondencia con las estructuras extendidas según el gráfico de Ramachandran.

Predicción de estructura secundaria. Las proteínas de la cápside de los potyvirus tiene un tamaño que varía entre 260 y 330 aminoácidos, con un porcentaje de identidad de aproximadamente 75% (figura 1). La región de mayor variabilidad se localiza en el extremo N-terminal que es el elemento determinante de la transmisión por vectores de los potyvirus y el análisis de la estructura secundaria indica que esta región solo tiene tendencia a la formación de una pequeña hélice alfa en el segmento 7-13, pero su baja conservación sugiere que esta no es una característica global de todas las estructuras de cápside (figura 2). Igualmente, varios estudios han demostrado que el dominio N-terminal no está involucrado en la integridad estructural de potyvirus, ni participa de manera significativa en el ensamblaje de la partícula viral (Dolja et al. 1994). Por todo lo anterior, es justificable no incluir este dominio en el modelamiento estructural.

Para el dominio central encontramos cinco regiones con tendencia a la formación de hélices ($\alpha 1a$, 85-102; $\alpha 1b$, 105-114; $\alpha 2$, 157-177; $\alpha 3$, 190-204; y $\alpha 4$, 213-228). La $\alpha 1a$ está compuesta por seis aminoácidos completamente conservados (S83, N84, R86, F92, W95 y V99) y se observa conservación de la carga en la posición 97, donde solo se encuentran aspartatos y glutamatos. La $\alpha 1b$ se caracteriza por la presencia de alto contenido de residuos acídicos y solamente un residuo, M110; está conservado completamente en las secuencias comparadas. La predicción de hélices alfa para otras secuencias de la cápside y su bajo nivel de conservación, indican que la hélice $\alpha 1b$ no siempre está presente (datos no mostrados); sin embargo, basados en los contornos de microscopía electrónica de Kendall et al. (2008), es una hipótesis razonable asumir que $\alpha 1b$ es simplemente una prolongación de $\alpha 1a$. Algo similar ocurrió con la predicción de la hélice larga de TMV que los programas de estructura secundaria tienden a dividirse en dos hélices. Por esta razón se ha modelado $\alpha 1a$ y $\alpha 1b$ como una sola hélice, $\alpha 1$. Las hélices $\alpha 2$, $\alpha 3$ y $\alpha 4$ comprenden algunas de las regiones más conservadas de los potyvirus y no hay duda de que allí se agrupan la mayoría de las interacciones responsables de la estabilidad estructural. El dominio central presenta tres regiones de estructura irregular, que a diferencia de la región N-terminal, son muy conservadas. Las regiones conectoras de las hélices $\alpha 2$ - $\alpha 3$ y $\alpha 4$ - $\alpha 5$ estarían ubicadas en el interior del virus.

El extremo C-terminal es altamente conservado y presenta poca variabilidad en longitud. La predicción de estructura secundaria sugiere la posible presencia de dos láminas beta o estructuras extendidas, $\beta 3$ y $\beta 4$, en las posiciones 234-238, 240-243; y una hélice terminal, $\alpha 6$, que comprende los residuos 255-264. El análisis de hidrofobicidad sugiere que $\beta 4$ es altamente polar y probablemente expuesta; lo cual tiene sentido dado su bajo nivel de conservación. El análisis

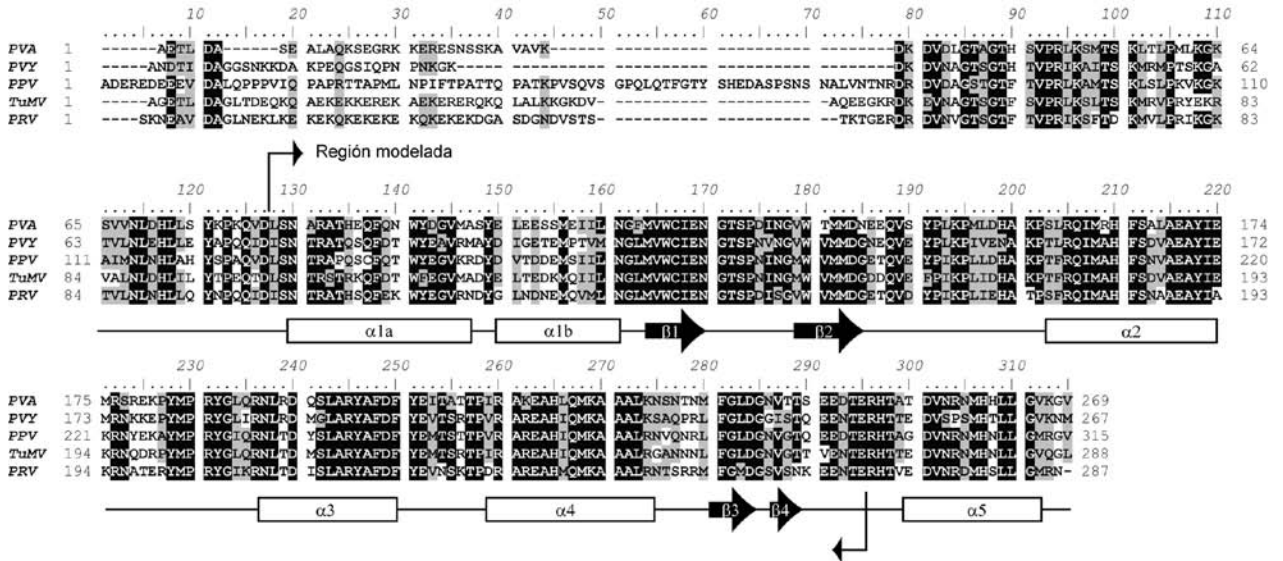


Figura 1. Alineamiento de secuencias de la cápside de potyvirus: virus A de la papa (PVA), virus de Y de la papa (PVY), virus del mosaico de la soya (SMV), virus del grabado del tabaco (TEV), Plum pox virus (PPV), virus del mosaico del nabo (TuMV), virus ringspot de la papaya (PRV). La posición de los elementos de estructura secundaria predichos se encuentran debajo de cada alineamiento

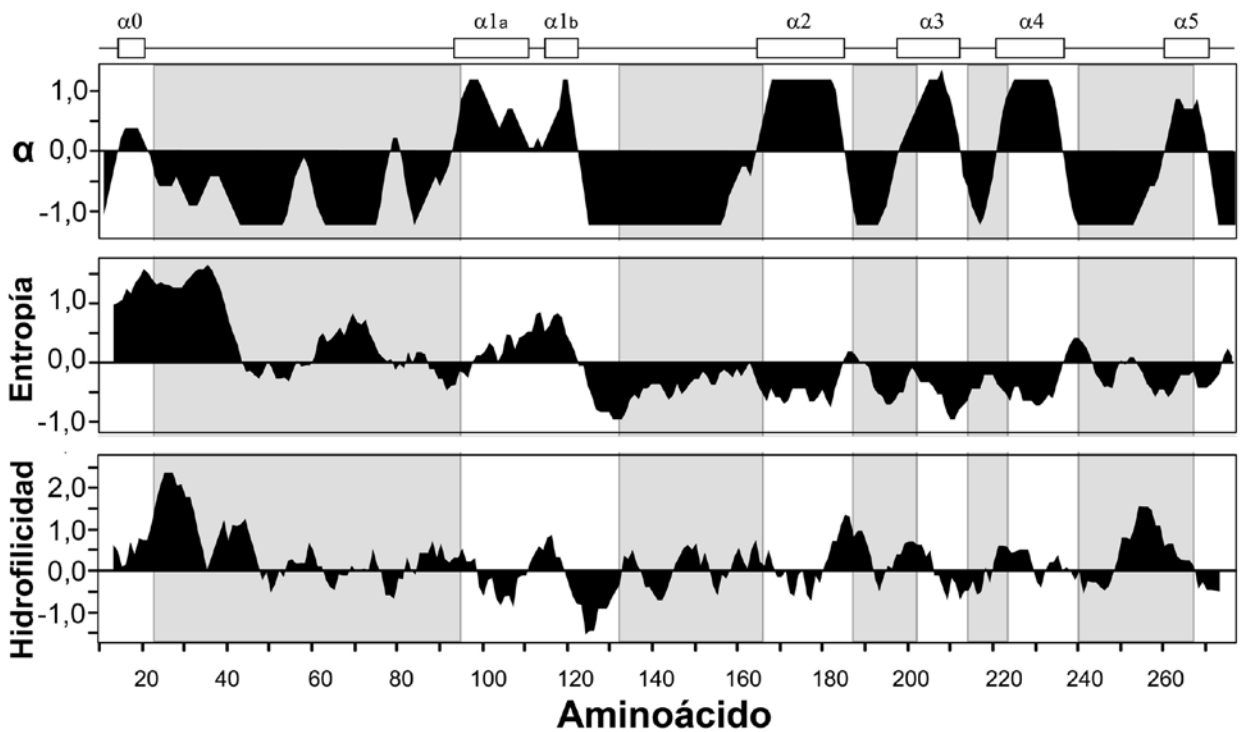


Figura 2. Correlación entre la entropía de las secuencias y el perfil de hidrofiliidad medido con el esquema de Hopp y Woods (Hopp y Woods 1981) para el virus A de la papa. La posición de las hélices alfa es señalada en la parte superior. La región sombreada corresponde a las regiones de estructura irregular predichas con el programa DisEMBL

de regiones no globulares sugiere la presencia de un elemento irregular entre las hélices $\alpha 5$ y $\alpha 6$. La función de la región C-terminal no es del todo clara, pero se ha demostrado que mutaciones y deleciones en esta región del Virus del mosaico de la soya inhibe la interacción entre subunidades de la cápside (Kang et al. 2006).

Modelo de la proteína de la cápside de PVA.

Debido a la ausencia de estructuras de alta resolución en el *Protein Data Bank* con al menos el 25% de identidad con la secuencia de potyvirus, no fue posible llevar a cabo la modelación por homología de esta proteína. Por esta razón, se implementó un procedimiento que consistió en la obtención de un consenso de estructura secundaria y su posterior ensamblaje terciario mediante un procedimiento iterativo de *docking* y minimización de energía, tal como se describe en la figura 3. En el último paso se adicionaron las regiones no estructuradas y se minimizó el modelo. La calidad de la estructura fue evaluada con el programa WHAT_CHECK (Hoofst et al. 1996). Los valores de Z para el gráfico de Ramachandran y la conformación del esqueleto de la proteína (2,074 y -0,164, respectivamente), están en el rango aceptable. El valor de Z para el empaquetamiento de aminoácidos es de -2,148, que está dentro del rango aceptado para modelos contruidos por homología (-2,0 y -3,0).

El modelo de la región central de la cápside del virus PVA se ilustra en la figura 4. El núcleo central está compuesto por las hélices $\alpha 1$, $\alpha 2$, $\alpha 3$ y $\alpha 4$. La interacción entre estas hélices estaría mediada por interacciones de tipo hidrofóbico donde F92 ($\alpha 1$) se intercala en la hendidura formada por L168 y Y172 de la hélice 2, W95 interactúa con H164, A167 y L168, M100 con I213 y Y103 con I161. Es interesante resaltar que la hélice 1 presenta en sus extremos N y C terminales unos parches de carga positiva y negativa compuestos por los residuos R86, H89, E104, E106, E107, E111 y que el mismo patrón es observado para la hélice $\alpha 2$, donde el parche

básico está conformado por H153, K155 y R159, mientras que el negativo por E170 y E174. Esta distribución de cargas en ambas hélices sugiere una posible interacción intermolecular entre subunidades de la cápside por complementariedad de cargas y potencialmente importante en el ensamblaje de potyvirus (figura 4). La hélice 2 interactuaría con la hélice 3 a través de los residuos L188, R193, S196, L197, Y200 y F204. F165 se intercala entre los residuos Y200 y F204; A169 interactúa con I197, I173-L188, mientras que R193 formaría un par iónico con E170. Las interacciones de $\alpha 3$ con $\alpha 4$ se podrían dar a través de los residuos H219, L220, K223, A224 y L227. El residuo D203 y E206 forma interacciones de carga complementarias con H219, K223 y R199 ($\alpha 3$), mientras que L227 lo haría con el segmento alifático de R199.

La región conectora $\alpha 1$ - $\alpha 2$ está dividida en una región no polar y otra de carácter ácido compuesta por los residuos E123, D129, D138, E140, E141. Todas las evidencias indican que esta región estaría desnaturalizada en la forma monomérica de la cápside y su plegamiento probablemente requiera interacciones cuaternarias con otras subunidades y con el ARN. Algo similar ha sido observado en el TMV donde se ha encontrado que las regiones flexibles contienen gran cantidad de aminoácidos de carga negativa que se repelen entre sí, evitando el ensamblaje del virus. Durante el ensamblaje y en presencia de ARN, estas regiones cambian a una conformación helicoidal que permite la estructuración del virus (Namba 2001).

DISCUSIÓN

Kendall et al. (2008) determinaron recientemente una estructura de baja resolución (22 Å) del potyvirus SMV utilizando difracción de fibras, microscopía crioelectrónica y de transmisión. En esta reconstrucción, las subunidades de la cápside se encuentran bien definidas y presentan una disposición espacial similar a la del TMV,

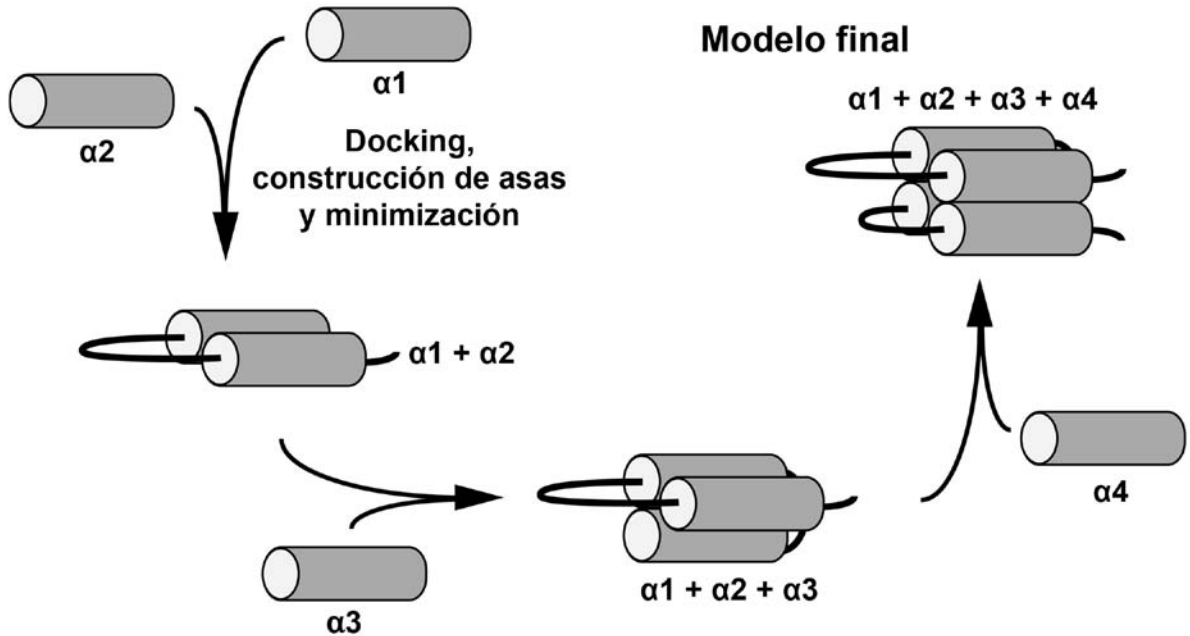


Figura 3. Ilustración esquemática del proceso de *docking* iterativo utilizado para la modelación de la cápside de potyvirus

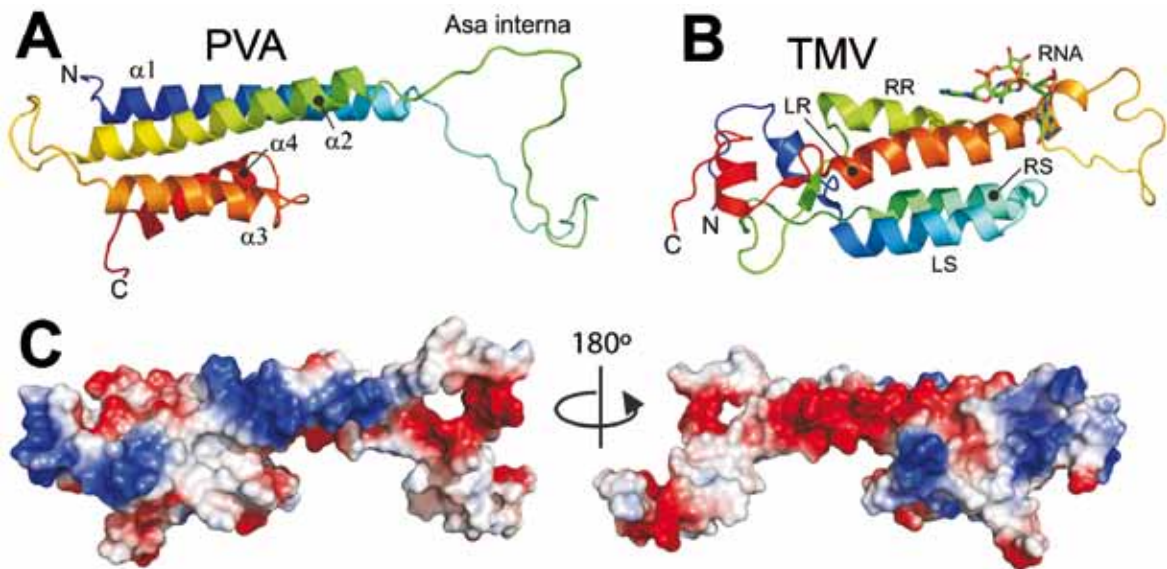


Figura 4. Modelo para la estructura tridimensional de región central del virus A de la papa. **A.** Modelo propuesto para la cápside del virus A de la papa. **B.** Estructura del virus del mosaico del tabaco. **C.** Electroestática del modelo de la cápside

pero con un ancho mayor y longitud menor. Las subunidades del SMV tienen 55 Å de longitud y un ancho máximo de 35 Å. La estructura del SMV tiene un paso de rosca de 33 Å, 8,8 subunidades por vuelta y un diámetro de 140 Å. La densidad

radial del SMV sugiere que el virión tiene un hueco central con un radio de 15 Å y radio máximo de 70 Å. Como puede apreciarse en la figura 5, el modelo se ajusta de manera satisfactoria a la superficie experimental y proponemos

que las regiones faltantes corresponden a los extremos N y C-terminales no modelados en este trabajo. El dominio C-terminal parece ser fundamental en determinar las interacciones verticales con las hélices $\alpha 1$ y $\alpha 2$. En la estructura por microscopía electrónica se observa en la vista superior que cada subunidad presenta dos segmentos que forman un ángulo de aproximadamente 40° . En nuestro modelo este quiebre de estructura fue obtenido de manera natural con el procedimiento de *docking*, lo cual sugiere que es una consecuencia del empaquetamiento de las hélices del núcleo estructural.

Los dos segmentos lineales estarían compuestos por $\alpha 1$ - $\alpha 2$ y luego $\alpha 3$ - $\alpha 4$, el dominio no modelado correspondería a una protuberancia hacia la derecha y que proponemos que estaría formado por la unión de las regiones conservadas

N y C-terminales. El modelo también encaja de manera satisfactoria con la vista lateral de la estructura del virus. En esta imagen se observa que las subunidades de los potyvirus presentan una estructura escalonada, la parte superior corresponde a $\alpha 1$ - $\alpha 2$, luego $\alpha 3$ - $\alpha 4$ y por último la región no modelada que determina las interacciones verticales entre subunidades. Si se asume que la región no modelada está conformada por los residuos conservados de las regiones N y C terminales y que la forma de este dominio es aproximadamente esférico, se estima que el diámetro de dicha esfera debe ser del orden de 28 \AA , que es un valor comparable con la regiones señaladas con círculos en la figura 5. Nemykh et al. (2008) publicaron un modelo de la estructura de la cápside de PVA; sin embargo, esta estructura excede las dimensiones de potyvirus obtenidas por microscopía electrónica.

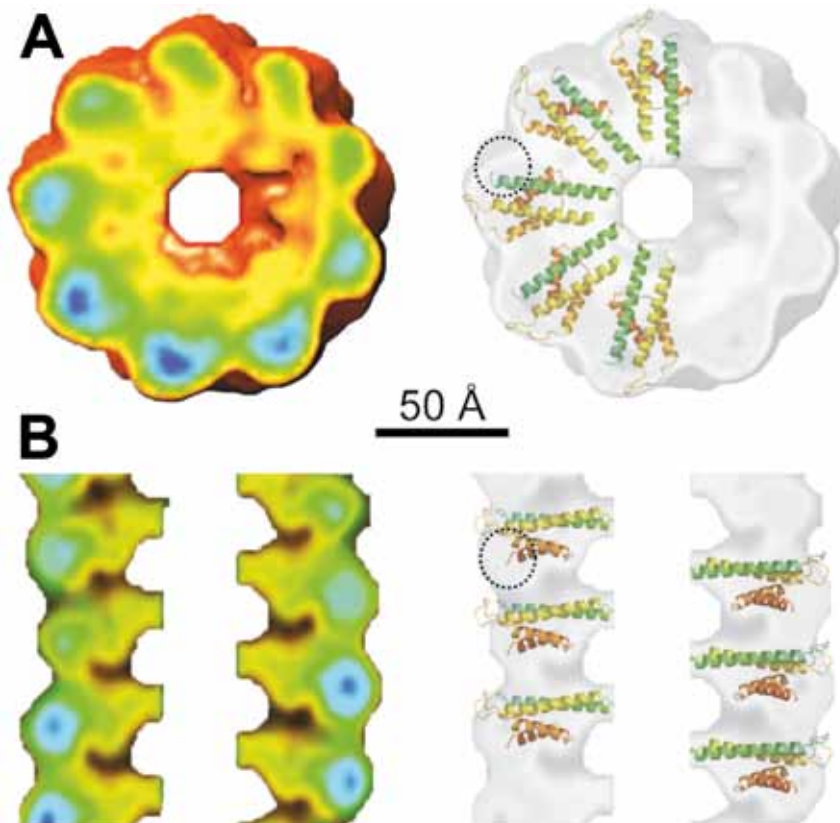


Figura 5. Figura adaptada de Kendal et al. (2008), donde se ilustra la superficie obtenida por microscopía electrónica de un potyvirus y su comparación con el modelo. La región propuesta para el dominio C-terminal se encuentra señalada con el círculo

En nuestra opinión, el error de dicha estructura consiste en asumir que la región C-terminal es colineal al resto de la estructura. En la propuesta aquí planteada, se sugiere que dicha región está empaquetada con las hélices $\alpha 3$ y $\alpha 4$ y tiene contactos intermoleculares con las $\alpha 1$ y $\alpha 2$ de la subunidad inmediatamente inferior.

Desafortunadamente, existen muy pocos estudios donde se evalué el efecto de mutaciones de la proteína de la cápside sobre el ensamblaje de potyvirus, los cuales podrían ser de gran utilidad en el refinamiento del modelo presentado (Lin et al. 2009). Sin embargo, consideramos que la estructura presentada puede servir como una herramienta guía en el diseño de experimentos de mutagénesis dirigida que permitan la mejor comprensión de los patrones de conservación de la proteína de la cápside, así como el mecanismo de ensamblaje y reconocimiento del ARN viral. Se espera que los procedimientos desarrollados en este trabajo, puedan ser empleados en el futuro próximo para modelar la estructura tridimensional de potyvirus de importancia económica en Colombia como PVY, TaLMV, PRSV, entre otros, como una estrategia para el diseño de herramientas de diagnóstico (ej. diseño de anticuerpos con base en las características estructurales de la partícula viral) y de métodos de control (e. g., generación de variedades mejoradas con tolerancia/resistencia a virus) de las enfermedades asociadas en dichos virus en los agroecosistemas tropicales.

AGRADECIMIENTOS

Esta investigación se realizó gracias al apoyo económico del proyecto Colciencias: “Desarrollo de métodos de detección serológica y molecular del complejo de virus asociado al mosaico del tomate de árbol en Colombia”, Contrato 408-2007 y de la Dirección de Investigaciones de la Universidad Nacional de Colombia sede Medellín (DIME).

REFERENCIAS

- Carrington JC, Freed DD, Sanders TC. 1989. Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and in vitro. *Journal of Virology* 63: 4459- 4463.
- Carrington JC, Dougherty WG. 1987. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. *Journal of Virology*, 61: 2540-2548.
- DeLano WL. 2002. The PyMOL Molecular Graphics System. DeLano Scientific Homepage. Fecha de acceso: 10 de noviembre de 2010. Disponible en: <<http://www.pymol.org>>.
- Dolja VV, Haldeman R, Robertson NL, Dougherty WG, Carrington JC. 1994. Distinct functions of capsid protein in assembly and movement of *Tobacco etch potyvirus* in plants. *EMBO Journal*, 13: 1482-1491.
- Doolittle RF, Garnier J, Gibrat JF, Robson B. 1996. GOR secondary structure prediction method version IV. *Methods in Enzymology*, 266: 540-553.
- Duggal R, Hall TC. 1993. Identification of domains in Brome mosaic virus RNA-1 and coat protein necessary for specific interaction and encapsidation. *Journal of Virology*, 67: 6406-6412.
- Ehrenfeld N, González A, Canon P, Medina C, Pérez-Acle T, Arce-Johnson P. 2008. Structure-function relationship between the tobamovirus TMV-Cg coat protein and the HR-like response. *Journal of General Virology*, 89: 809-817.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41: 95-98.
- Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, 89: 10915-10919.
- Hooft RWW, Vriend G, Sander C, Abola EE. 1996. Errors in protein structures. *Nature*, 381: 272.
- Hopp TP, Woods KR. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 78: 3824-3828.
- Kang SH, Lim WS, Hwang SH, Park JW, Choi HS, Kim KH. 2006. Importance of the C-terminal domain of Soybean mosaic virus coat protein for subunit interactions. *Journal of General Virology*, 87: 225-229.
- Kendall A, McDonald M, Bian W, Bowles T, Baumgarten SC, Shi J, Stewart PL, Bullitt E, Gore D, Irving TC, Havens WM, Ghabrial SA, Wall JS, Stubbs G. 2008. Structure of flexible filamentous plant viruses. *Journal of Virology*, 82: 9546-9554.
- Langenberg WG, Zhang L. 1997. Immunocytology shows the presence of Tobacco etch virus P3 protein in nuclear inclusions. *Journal of Structural Biology*, 118: 243-247.

- Lin L, Shi Y, Luo Z, Lu Z, Zheng H, Yan F, Chen J, Chen J, Adams MJ, Wu Y. 2009. Protein-protein interactions in two potyviruses using the yeast two-hybrid system. *Virus Research*, 142: 36-40.
- Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. 2003a. Protein disorder prediction: implications for structural proteomics. *Structure*, 11: 1453-1459.
- Linding R, Russell RB, Neduva V, Gibson TJ. 2003b. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Research*, 31(13): 3701-3708.
- Merits A, Guo D, Saarma M. 1998. VPg, coat protein and five non-structural proteins of *Potato A potyvirus* bind RNA in a sequence-unspecific manner. *Journal of General Virology*, 79: 3123-3127.
- Mink GI, Vetten HJ, Wyatt SD, Berger PH, Silbernagel MJ. 1999. Three epitopes located on the coat protein amino terminus of viruses in the Bean common mosaic potyvirus subgroup. *Archives of Virology*, 144: 1174-1189.
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. 1998. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *Journal of Computational Chemistry*, 19: 1639-1662.
- Namba K. 2001. Roles of partly unfolded conformations in macromolecular self-assembly. *Genes Cells*, 6: 1-12.
- Namba K, Stubbs G. 1986. Structure of Tobacco mosaic virus at 3.6 Å resolution: implications for assembly. *Science*, 231: 1401-1406.
- Nemykh MA, Efimov AV, Novikov VK, Orlov VN, Arutyunyan AM, Drachev VA, Lukashina EV, Baratova LA, Dobrov EN. 2008. One more probable structural transition in Potato virus X virions and a revised model of the virus coat protein structure. *Virology*, 373: 61-71.
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26: 1781-1802.
- Rajamaki M, Merits A, Rabenstein F, Andrejeva J, Paulin L, Kekarainen T, Kreuze JF, Forster RL, Valkonen JP. 1998. Biological, serological and molecular differences among isolates of a Potato A potyvirus. *Phytopathology*, 88: 311-321.
- Riechmann JL, Laín S, García JA. 1992. Highlights and prospects of potyvirus molecular biology. *Journal of General Virology*, 73: 1-16.
- Riechmann JL, Cervera MT, García JA. 1995. Processing of the *Plum pox* virus polyprotein at the P3-6K1 junction is not required for virus viability. *Journal of General Virology*, 76: 951-956.
- Rost B. 2001. Review: Protein secondary structure prediction continues to rise. *Journal of Structural Biology*, 134: 204-218.
- Rost B, Yachdav G, Liu J. 2004. The predict protein Server. *Nucleic Acids Research*, 32: 321-326.
- Sali A, Blundell TL. 1995. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234: 779-815.
- Shukla DD, Ward CW. 1989. Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. *Advances in Virus Research*, 36: 273-314.
- Singh RP, Valkonen JPT, Gray SM, Boonham N, Jones RAC, Kerlan C, Schubert J. 2008. Brief review: The naming of potato virus Y strains infecting potato. *Archives of Virology* 153: 1-13.
- Urcuqui-Inchima S, Haenni AL, Bernardi F. 2001. Potyvirus proteins: a wealth of functions. *Virus Research*, 74 (1-2): 157-75.
- Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB. 2000. *Virus Taxonomy - Seventh Report of the International Committee on Taxonomy of Viruses*. Academic Press. New York. 1162 p.
- Verchot J, Koonin EV, Carrington JC. 1991. The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology*, 185: 527-35.
- Verchot J, Carrington JC. 1995. Evidence that the potyvirus P1 proteinase functions in *trans* as an accessory factor for genome amplification. *Journal of Virology*, 69: 3668-3674.