

Revista BIOCYT es editada en la Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México

HOMOLOGY MODELLING AND MULTIMERIC DOCKING OF THE SMALL OUTER CAPSID PROTEIN FROM THE BACTERIOPHAGE IME08

MODELACIÓN POR COMPARACIÓN Y ACOPLAMIENTO MULTIMÉRICO DE LA PROTEÍNA EXTERIOR PEQUEÑA DE LA CÁPSIDE DEL BACTERIOFAGO IME08

Maicol Ospina-Bedoya

Instituto de Biología, Universidad de Antioquia, Calle 67 N° 53-108, Bloque 7, Apartado Aéreo
1226, Medellín, Antioquia, Colombia.

ABSTRACT

The small outer capsid protein plays a stabilizing role in the viral assembly, adhering to the capsid during the later stages of maturation. This protein acts as glue among adjacent capsomers, protecting the virus against extreme changes. The small outer capsid protein of the bacteriophage IME08 was modelled using structural protein homology. A trimeric protein docking was developed with the best-scored model and important sites of the molecules interfaces were identified. It was used the Swiss Model platform for developing the protein structure. Reliability was assessed by the QMEAN, Verify3D and ERRAT indices. The quality of the whole model was verified by Ramachandran plot and the trimerization model was performed on the platform ClusPro 2.0 Protein-Protein Docking. The structure obtained has a reliability estimator QMEANscore4 of 0.769, rating it as a suitable model. The Z-Score QMEAN value was 0.133, showing that the obtained model is not different from the experimental structures stored in PDB database. The estimators and the Ramachandran plot evaluated positively the model. Finally we identified a loop between two secondary structures as an important site of the interaction of small outer capsid proteins, indicating that from residues 35 to 41 are relevant in the trimerization process.

Key words: Small outer capsid protein, homology modelling, bacteriophage, docking, trimerization

Correspondence to author

m.ospina.bedoya@gmail.com

Manuscrito recibido el 29 de marzo de 2012, aceptado el 05 de junio de 2012.

RESUMEN

La proteína exterior pequeña de cápside cumple una función estabilizadora, adhiriéndose a la cápside durante los últimos estados de maduración. Las moléculas de esta proteína actúan como pegamento entre capsómeros, protegiendo al virus de cambios extremos de pH y temperatura, sin embargo es desconocida la estructura y los mecanismos de acoplamiento de esta proteína. En este trabajo se modeló por comparación la proteína exterior pequeña de la cápside del bacteriófago IME08 y con el modelo se realizó un acople trimérico de la proteína. Se utilizó la plataforma Swiss Model para la elaboración de la estructura. Se evaluó la confiabilidad por los índices QMEAN, Verify3D y ERRAT. La calidad del modelo se verificó por Mapas de Ramachandran. La trimerización del modelo se efectuó en la plataforma ClusProProtein-ProteinDocking 2.0. La estructura obtenida tiene un estimador de confiabilidad QMEANscore4 de 0.769, valorándolo así como un modelo adecuado. El valor QMEAN Z-Score fue de 0.133, sugiriendo que el modelo obtenido no difiere significativamente de las estructuras experimentales de la base de datos PDB. Los estimadores y el mapa de Ramachandran evaluaron positivamente la estructura. Finalmente se identificó que los residuos 35 hasta el 41 son de relevancia en el proceso de trimerización.

Palabras clave: Proteína exterior pequeña, modelación por comparación, cápside, bacteriófago, acoplamiento, trimerización

INTRODUCCIÓN

Los virus son las entidades biológicas más abundantes de toda la tierra, con una trascendencia indiscutible dentro de los procesos biológicos (incluso abióticos) de la vida en el planeta (Forterre, 2005, 2006a, 2006b). Dentro de estos, los bacteriófagos (virus de bacterias) son conocidos no solo como los microorganismos más abundantes sino posiblemente como los más diversos. Se estima que la biomasa de bacteriófagos supera por diez veces la biomasa bacteriana (Labrie y Samson, 2010). Además de su papel en la evolución procariota con la transferencia horizontal de genes y en la ecología de los seres vivos en general, los bacteriófagos tienen una posición privilegiada en los estudios biológicos, ya que han sido parte del desarrollo teórico y práctico de diferentes modelos en biología molecular, celular y evolutiva (Lang y Beatty, 2000; Dennehy, 2009).

El género de virus "T4-like" (familia Myoviridae) contiene numerosos bacteriófagos líticos que infectan un amplio rango de bacterias, incluyendo varias patogénicas. El genoma de estos bacteriófagos está compuesto por ADN de doble cadena, el cual está en el interior de una cápside icosaédrica, que a su vez está unida a una cola contráctil. A la fecha, menos de 40 secuencias de genomas completos de bacteriófagos "T4-like" se encuentran disponibles y pocas estructuras tridimensionales resueltas experimentalmente se encuentran en las bases de datos macromoleculares. Dentro de este género cabe destacar al bacteriófago T4, ya que ha sido un organismo modelo en biología molecular desde los años 40 y le confiere nombre al género (Huanhua-Jiang et al., 2011).

En este trabajo se pretende aportar a la biología estructural de otro virus del género, el bacteriófago IME08, específicamente a resolver la estructura tridimensional de uno de los componentes de su cápside. Este es un fago lítico enterobacterial, que fue aislado de las aguas residuales de un hospital junto con la cepa 8899 de *Escherichia coli* (Huanhua-Jiang et al., 2011; Xiaofang-Jiang et al., 2011). A pesar de haberse relacionado con *E. coli* y de que eventualmente este fago podría regular las poblaciones bacterianas de ese nicho y a que se encuentra disponible su genoma, nada se conoce acerca de las características estructurales que relacionan la proteína exterior pequeña del bacteriófago, su papel en el ciclo replicación y la maduración de la cápside viral.

Los bacteriófagos con ADN de doble cadena ensamblan su cápside de tal manera que esta se une alrededor de un núcleo que consiste en proteínas de portal dodecaméricas y varias proteínas andamio (Qin et al., 2007). Entonces, parte del núcleo es eliminado y la cápside se expande, iniciando así el empaquetamiento del ADN. Al final el ADN es empacado a una densidad casi cristalina (~500 mg/ml), con una presión aproximada de ~seis MPa. En algunos virus, esta expansión de la procápside expone sitios para la unión de la proteína exterior pequeña (SOC). Es por esto que estructuralmente los virus tienen la necesidad de estabilizar su cápside contra la presión ejercida por el ADN en su interior y sobrevivir en ambientes hostiles. El reforzamiento que provee las proteínas SOC (de nueve kDa) tiene la función de estabilizar y unirse a la cápside durante los últimos estados de la maduración viral. Alrededor de 870 moléculas de SOC se unen para actuar como "pegamento" entre capsómeros hexaméricos adyacentes, formando así una estructura que estabiliza la cápside y la protege contra cambios extremos en el pH o temperatura (Qin et al., 2010).

Análisis con ultracentrifugación mostraron que las proteínas SOC son monoméricas en solución (Qin et al., 2010). Sin embargo una reconstrucción en el bacteriófago RB69 por criomicroscopía electrónica mostró conformaciones triméricas entre proteínas en la superficie de la cápside del virus. Estas conformaciones han sido observadas en otros bacteriófagos no relacionados y virus eucariotas como adenovirus y herpesvirus, posiblemente tienen también un papel determinante en la estabilización estructural de la cápside (Qin et al., 2010).

Es de vital importancia conocer todos los mecanismos implicados en el ciclo de replicación de los bacteriófagos, tanto para entender cómo estos toman ventaja y modifican el ambiente citosólico, como para conocer la manera en que los bacteriófagos controlan poblaciones bacterianas de importancia clínica a gran escala. Por ejemplo, proteínas tóxicas del ántrax unidas a la región N-terminal de la proteína SOC han sido insertadas en la cápside del bacteriófago T4, esto permite una neutralización fuerte de la toxina en ratones y conejos (Qin et al., 2007; Qin et al., 2010).

El objetivo de este trabajo es presentar un modelo por comparación de la proteína SOC del bacteriófago IME08, basado en la estructura tridimensional del bacteriófago RB69 perteneciente al mismo género. Con la estructura obtenida pretendemos un acoplamiento trimérico de la proteína modelada para entender los factores más importantes de dicha interacción.

MATERIALES Y MÉTODOS

Se utilizó la secuencia aminoacídica de la proteína SOC del bacteriófago IME08 (GenBank: ADI55353, 78 residuos). A partir de esta secuencia se identificaron las proteínas homólogas con mayores índices de similitud en la base de datos del Protein Data Bank (PDB) mediante un BLAST (Altschul et al., 1990). El patrón de homología utilizado fue la estructura de la proteína SOC del bacteriófago RB69 (9 kDa) (PDB: 3IG9) con 1,90 Å de resolución. La identificación del molde fue ratificada además por medio de la plataforma SWISS-MODEL Template Identification (Arnold et al., 2006). La estructura seleccionada presentaba un valor de identidad/similitud de 71/76 (93%).

El alineamiento múltiple de secuencias se obtuvo mediante CLUSTAL W2 implementado en la plataforma EMBL-EBI (Larkin et al., 2007). El modelo por comparación fue construido en el servidor SWISS-MODEL AlignmentMode (Schwede, 2003; Arnold et al., 2006) y los ficheros obtenidos fueron editados y almacenados con BIOEDIT (Hall, 1999).

El modelo calculado fue verificado inmediatamente con un análisis en la misma plataforma con los estimadores QMEAN, Anolea y Gromos (Benkert et al., 2011). Posteriormente se utilizó la plataforma SAVE (The Structure Analysis and Verification Server version 4) para evaluar los parámetros Verify3D y ERRAT (Lüthy et al., 1992; Colovos y Yeates, 1993). El parámetro de calidad estereoquímica Ψ (psi) - Φ (phi) del modelo logrado se verificó mediante el uso de Mapas de Ramachandran, realizados en la plataforma Ramachandran PLOT 2.0 Centre Of Excellence in Structural Biology and Biocomputing.

La asignación de la estructura secundaria se realizó a través del servidor JpredStructureprediction server (Cole et al., 2008). Por último, el acoplamiento multimérico proteína-proteína, se realizó en el Cluster en línea ClusProProtein-ProteinDocking de la Universidad de Boston (Comeau et al., 2003; Comeau et al., 2004; Kozakov et al., 2006; Kozakov et al., 2010). Las manipulaciones gráficas fueron hechas en el programa libre Swiss PDB Viewer (DeepView) (Fleishman y Unger, 1999) y PyMol (Delano, 2002).

RESULTADOS

El alineamiento múltiple de secuencia de varios bacteriófagos del género "T4-like" muestra un patrón continuo (fig. 1), y regiones amplias conservadas.

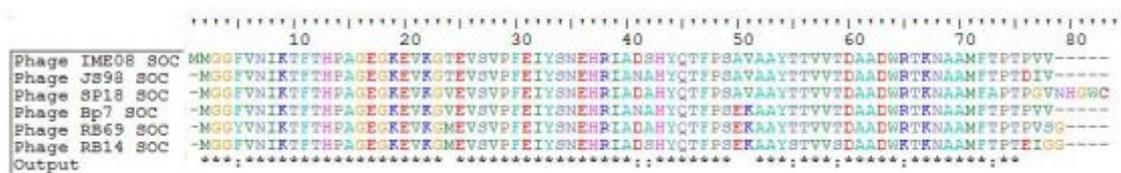


Fig. 1. Alineamiento múltiple de secuencia. Se observan las seis secuencias con los mayores valores de identidad de secuencia. Los colores de los aminoácidos están acorde a sus propiedades químico-físicas. La proteína SOC es bastante conservada en varios bacteriófagos del género

La energía final del modelo por defecto, calculada por GROMOS96 (módulo de DeepView) fue de -2399.377 KJ/mol. El modelo alcanzó un estimador de confiabilidad bastante adecuado, conocido también como QMEANscore4. La puntuación global de todo el modelo que refleja la confiabilidad de la estructura predicha fue de 0.769 (variando de cero a uno). Además, el estimador de fiabilidad QMEAN Z-Score (medida de la calidad absoluta del modelo) tuvo un valor de -0.133, en promedio, estructuras de cristalografía de rayos X de alta resolución tiene Z scores iguales a cero.

El estimador Verify3D que determina la compatibilidad del modelo atómico (3D) con la secuencia de aminoácidos (1D) asigna una clase estructural en función de su ubicación y el entorno (alfa, beta, lazo, polar, no polar, etc.), comparando los resultados con estructuras exitosas, dio como resultado que el 96.10% de los residuos tuvieron un valor 3D-1D promediado mayor a 0.2.

La evaluación del modelo con el estimador ERRAT analiza las estadísticas de las interacciones no-enlazantes entre diferentes tipos de átomos y grafica el valor de la función de error en comparación con la posición de un marco deslizante de nueve residuos, calculado mediante una comparación con estadísticas de estructuras altamente refinadas, dando como resultado un factor de calidad total del 97.015%, expresando así el porcentaje de la proteína para la cual el valor de error calculado cae por debajo del límite de rechazo (95%) (Estructuras de alta resolución generalmente producen valores alrededor del 95% o superiores).

El análisis de calidad estereoquímica por el mapa de Ramachandran en el que se visualizan todas las combinaciones posibles de ángulos diédricos Ψ (psi) contra Φ (phi) en los aminoácidos de la proteína SOC mostró que el 83.22% de los residuos se encuentran en regiones totalmente permitidas, el 11.41% se encuentra en regiones adicionalmente permitidas, el 4.70% en regiones generalmente permitidas y solo un 0.67% (un residuo) en regiones fuera de zona (Fig. 2).

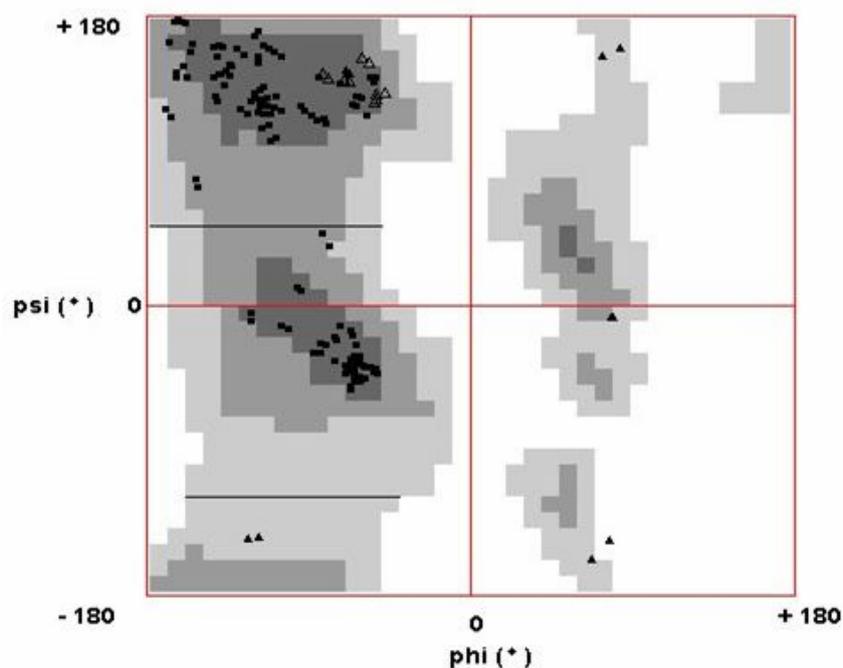


Fig. 2. Mapa de Ramachandran con los ángulos diédricos Ψ (psi) contra Φ (phi) en los aminoácidos de la proteína SOC. Los cuadrados grises oscuros (■) representan regiones totalmente permitidas, los cuadrados grises medios (▒) constituyen regiones adicionalmente permitidas, los cuadrados grises claros (░) simbolizan regiones generalmente permitidas. Los triángulos negros (▲) y blancos (△) representan los residuos de glicina y prolina respectivamente y los cuadrados negros simbolizan los residuos en regiones fuera de zona

La topología de la estructura secundaria total de la proteína SOC modelada consiste en la predominancia de hojas beta hacia el extremo N-terminal mientras que al extremo C-terminal se observa una hélice (Fig. 3).

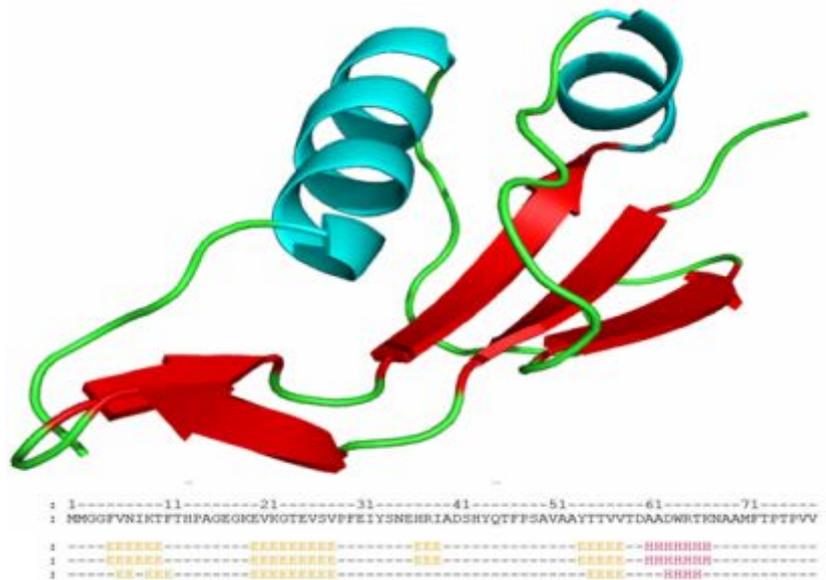


Fig. 3. Estructura del modelo de la proteína SOC del bacteriófago IME08. Resaltando con colores y código de letras las estructuras secundarias más relevantes

El acoplamiento trimérico realizado fue exitoso obteniéndose así una agrupación de modelos con coeficientes balanceados. Los puntajes de energía de los modelos variaron desde -400.5 hasta -569.3 KJ/mol. Los resultados del acoplamiento permiten identificar que la interfase entre las moléculas de la proteína SOC en el trímero está formada por un lazo, conteniendo los residuos 35 hasta el 41 en la región desordenada comprendida entre una hélice y una hoja beta (Fig. 4). Dentro de este lazo los aminoácidos Glu35, Ala39, Asp40, Arg37 poseen las distancias más cercanas a los otros aminoácidos de los monómeros adyacentes ($\approx 1.8-2\text{\AA}$).

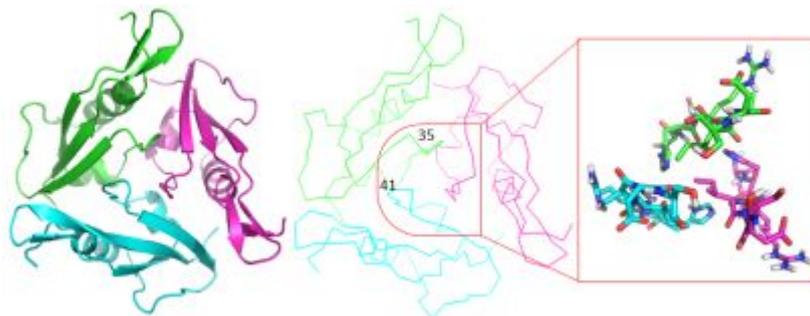


Fig. 4. Acoplamiento multimérico de la proteína SOC. Se resalta un lazo que contiene los aminoácidos 35 a 41. postulándolos como responsables de la trimerización

DISCUSIÓN

La modelación por homología es un sistema de desarrollo paulatino y progresivo, eficiente para obtener información acerca de una proteína de interés (Le et al., 2004; Kaistha et al., 2009). Una vez que se tiene el modelo, puede ser rentable para una variedad de objetivos, por ejemplo para identificar sitios de unión y sitios activos, desarrollo e identificación de ligandos y sustratos específicos, predicción de epítopes antigénicos, refinación de modelos por

rayos X y resonancia magnética nuclear, confirmación de relaciones estructurales distantes y simulaciones de acoplamiento multimérico como en este caso, entre otras muchas opciones (Naranjo et al., 2007; Sánchez et al., 2009).

Este trabajo provee un mayor entendimiento de la interacción de las proteínas de cápside de bacteriófagos que no son modelos biológicos pero que podrían ser usados como herramientas de control, diagnósticas y biotecnológicas.

En el BLAST realizado con varias secuencias de bacteriófagos del género "T4-like" se encontró que el gen de la proteína exterior pequeña de cápside está bien conservado y específicamente la sección que se encontró está asociada con la trimerización (aminoácidos 35 al 40), no tiene cambio alguno, esto establece la relevancia de esta región para la función estabilizadora de la proteína.

Trabajos anteriores sugieren que los valores de estructuras cristalográficas con homología mayores al 65% e identidad superiores al 50%, se consideran suficientes para realizar un modelo por similitud (Lin et al., 2003; Naranjo et al., 2007). El porcentaje de identidad de la proteína empleada en el presente estudio fue del 93%. Asimismo, la energía final del modelo fue de -2399.377 KJ/mol, se demuestra así su estabilidad.

La confiabilidad de la estructura tuvo una puntuación global de 0.769 (variando de cero a uno). El estimador de fiabilidad QMEAN Z-Score tuvo un valor de -0.133. Se muestra así que la estructura obtenida no difiere mucho de los valores QMEAN esperados para estructuras experimentales con un tamaño equivalente en la base de datos PDB. El estimador Verify3D dio como resultado que el 96.10% de los residuos tuvieron un valor 3D-1D promediado mayor a 0.2, se revela así la afinidad de la estructura con la secuencia primaria. El estimador ERRAT evaluó el modelo con un factor de calidad total del 97.015%, similar a valores de estructuras de alta resolución.

El análisis de calidad estereoquímica por el mapa de Ramachandran de la proteína SOC indicó que el 83.22% de los residuos se encuentran en regiones totalmente permitidas, el 11.41% se encuentra en regiones adicionalmente permitidas, el 4.70% en regiones generalmente permitidas y solo un 0.67% (un residuo) en regiones fuera de zona. El aminoácido fuera de zona es una glicina en la región N-terminal, esta zona es desordenada y dada su posición no interviene en la estructura secundaria de la proteína, ni en las regiones relevantes para la estabilidad o función. Se revela con esto que en general la conformación total de la estructura es apropiada y este modelo se puede considerar como válido.

El acoplamiento multimérico fue correcto y permitió identificar que la región implicada en la trimerización es un lazo de siete aminoácidos que se encuentra entre una hélice y una hoja beta. Particularmente los residuos Glu35, Ala39, Asp40, Arg37 presentan las condiciones para formar puentes de hidrogeno y están a distancias menores de dos angstrom.

En conclusión, con este trabajo se logró obtener un modelo por homología exitoso de la proteína SOC del bacteriófago IME08, con el cual se realizó un acoplamiento trimérico identificando así las regiones que pueden estar implicadas en la trimerización de las moléculas. La relevancia de los trabajos sobre las interacciones triméricas de la proteína SOC son de gran importancia ya que permite entender los procesos de estabilización de la cápside en sucesos de estrés, cambios de temperatura y pH.

AGRADECIMIENTOS

El autor desea agradecer a la Universidad de Antioquia por proveer los espacios adecuados para esta investigación. A los profesores asociados al curso de bioinformática, así como a mis compañeros de clase. Agradecimientos especiales a Sara Isabel Garcés, David Vásquez y Sergio Muñoz por aportar con sugerencias y comentarios al manuscrito.

REFERENCIAS

1. Altschul, S., Gish, W., y Miller, W., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3): 403-410.
2. Arnold, K., Bordoli, L., Kopp, J., y Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England)*, 22(2): 195-201.
3. Benkert, P., Biasini, M., y Schwede, T., 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)*, 27(3): 343-50.
4. Cole, C., Barber, J. D., y Barton, G. J., 2008. The Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, 36: W197-201.
5. Colovos, C., y Yeates, T., 1993. Verification of protein structures: Patterns of non-bonded atomic interactions. *Molecular Biology*, 2: 1511-1519.
6. Comeau, S. R., Gatchell, D. W., Vajda, S., y Camacho, C. J., 2003. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 20(1): 45-50.
7. Comeau, Stephen R, Gatchell, D. W., Vajda, S., y Camacho, C. J., 2004. ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic acids research*, 32: W96-9.
8. Delano, W. The PyMOL Molecular Graphics System. 2002. <http://www.pymol.org/> (accesado febrero 10, 2012).
9. Dennehy, J. J., 2009. Bacteriophages as model organisms for virus emergence research. *Trends in Microbiology*, 17(10): 450-7.
10. Fleishman, S., y Unger, V., 1999. Protein modeling for all. *Trends in biochemical sciences*, 24(9): 364-36.
11. Forterre, P., 2005. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie*, 87: 793-803.
12. Forterre, P., 2006a. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Research*, 117: 5-16.
13. Forterre, P., 2006b. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: A hypothesis for the origin of cellular domain. *PNAS*, 103(10).
14. Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41: 95-98.
15. Huanhua, Jiang, X., Wang, S., Li, C., Chen, B., An, X. y Mi, Z., 2011. The complete genome sequence of a novel T4-like bacteriophage, IME08. *Archives of Virology*, 156(8): 1489-92.
16. Kaistha, S. D., Sinha, R., y Model, S., 2009. Homology modeling of phosphorylthymidine kinase of enterohemorrhagic *Escherichia coli* OH:157. *Bioinformation* 3(6): 240-243.
17. Kozakov, D., Brenke, R., Comeau, S. R., y Vajda, S., 2006. PIPER: An FFT-Based Protein Docking Program with Pair wise Potentials. *Bioinformatics*, 406(August): 392-406.

18. Kozakov, D., Hall, D. R., Beglov, D., Brenke, R., Comeau, S. R., Shen, Y. y Li, K., 2010. Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. *Proteins*, 78(15): 3124-30.
19. Labrie, S. y Samson, J., 2010. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology*, 8: 317-327.
20. Lang, a S., y Beatty, J. T., 2000. Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2): 859-64.
21. Larkin, M. A, Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. a, McWilliam y H., Valentin, F., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, 23(21), 2947-8.
22. Le, D. T., Yoon, M.-Y., Kim, Y. T., y Choi, J.-D., 2004. Homology modeling of the structure of tobacco acetohydroxy acid synthase and examination of the active site by site-directed mutagenesis. *Biochemical and Biophysical Research Communications*, 317(3): 930-8.
23. Lin, T. H., Tsai, K.-C., y Lo, T.-C., 2003. Homology modeling of the central catalytic domain of insertion sequence ISLC3 isolated from *Lactobacillus casei* ATCC 393. *Protein Engineering Design and Selection*, 16(11): 819-829.
24. Lüthy, R., Bowie, J. U., y Eisenberg, D., 1992. Assessment of protein models with three-dimensional profiles. *Nature*, 356(6364): 83-5.
25. Naranjo, D., Arce, H. D. D., Pérez, E., Valiente, P. A., Carrasco, R., y Martínez, S., 2007. Modelación por homología de la catepsina B de *Fasciola hepatica*. *Revista de Salud Animal*, 29(1): 58-64.
26. Qin, L., Shivachandra, S. B., Zhang, Z., y Rao, V. B., 2007. Assembly of the small outer capsid protein, Soc, on bacteriophage T4: a novel system for high density display of multiple large anthrax toxins and foreign proteins on phage capsid. *Journal of Molecular Biology*, 370(5): 1006-19.
27. Qin, L., Fokine, A., O'Donnell, E., Rao, V. B., y Rossmann, M. G., 2010. Structure of the small outer capsid protein, Soc: a clamp for stabilizing capsids of T4-like phages. *Journal of Molecular Biology*, 395(4): 728-41.
28. Schwede, T., 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31(13): 3381-3385.
29. Sánchez, A., Fierro, L., y Castaño, R. S., 2009. Aproximación *in silico* a la estructura 3D de la proteína antiveneno DM64 de la zarigüeya (*Mammalia: Marsupialia: Didelphidae*). *Revista Académica Colombiana de Ciencias*, 23(126): 104-124.
30. Xiaofang, Jiang., Jiang, H., Li, C., Wang, S., Mi, Z., An, X. y Chen, J., 2011. Sequence characteristics of T4-like bacteriophage IME08 genome termini I revealed by high through put sequencing. *Virology Journal*, 8(1): 194.