

Estudio bioinformático del gen CRN8 en *Phytophthora infestans* causante de la enfermedad tizón tardío en plantas

Sánchez Rebellón Juan Diego ¹, León Rodríguez Daniel Arturo ²,
Gaviria Arias Duverney ²

RESUMEN:

El oomycota *Phytophthora infestans* es un desafío a controlar, debido a que este patógeno posee la capacidad de ocasionarle enfermedades a los cultivos alimenticios ya que se considera un oomiceto fitopatógenos, por lo cual, hoy en día continúa dificultando la producción de cultivos y perjudicando los ecosistemas a escala mundial, esta serie de proceso de infección a los cultivos implica la secreción de proteínas efectoras por parte del patógeno como la CRN8, con la finalidad de manipular los procesos de las células vegetales y la captación de nutrientes para apoyar su crecimiento y proliferación. Por esta razón se llevó a cabo en este estudio el análisis mediante herramientas de biología computacional con el fin de avanzar más sobre la caracterización completa de esta secuencia. Se realizó el alineamiento de diversas secuencias genéticas de CRN8 así como también se realizó una comparación filogenética de estas. En la parte del análisis estructural se realizó el modelamiento de la estructura tridimensional de la proteína producto de este gen mediante el uso de herramientas de predicción de modelos proteicos basados en métodos predictivos por enhebrado y *ab initio*. Se encontró homología de nuestra proteína con modelos de enzimas de tipo transferasa al momento de realizar alineamientos con estructuras contenidas en el Banco de Datos de Proteínas (PDB). De igual manera se logró conseguir un dato importante como lo es el sitio de unión a ligando de la proteína evaluada, encontrando que la molécula S-adenosil-L-homocisteína fue la que mejor encajó en su sitio activo.

PALABRAS CLAVE: *Phytophthora infestans*; CRN8; Proteínas Crinkler; Oomicetos.

1 Estudiante de Microbiología, Facultad de Ciencias de la Salud, Universidad Libre Pereira.

2 Docente Investigador. Programa de Microbiología, Facultad de Ciencias de la Salud, Universidad Libre Pereira

Bioinformatic study of the CRN8 gene in *Phytophthora infestans* that causes late blight disease in plants

ABSTRACT:

The oomycota *Phytophthora infestans* is a challenge to control, because this pathogen has the ability to cause diseases to food crops since it is considered a phytopathogenic oomycete, for which, today it continues to hinder crop production and harming ecosystems. On a global scale, this series of infection processes to cultures involves the secretion of effector proteins by the pathogen such as CRN8, in order to manipulate the processes of plant cells and the uptake of nutrients to support their growth and proliferation. For this reason, the analysis using computational biology tools was carried out in this study in order to further advance the complete characterization of this sequence. The alignment of various genetic sequences of CRN8 was carried out as well as a phylogenetic comparison of these. In the structural analysis part, the three-dimensional structure of the protein product of this gene was modeled using prediction tools of protein models based on predictive methods by threading and *ab initio*. Homology of our protein with models of transferase-type enzymes was found when aligning with structures contained in the Protein Data Bank (PDB). In the same way, it was possible to obtain important data such as the ligand binding site of the evaluated protein, finding that the S-adenosyl-L-homocysteine molecule was the one that best fit into its active site.

KEYWORDS: *Phytophthora infestans*; CRN8; Crinkler proteins; Oomycetes.

I.INTRODUCCIÓN

Entre los principales problemas emergentes en seguridad alimentaria, se encuentran las enfermedades ocasionadas en los cultivos por los oomicetos fitopatógenos, los cuales continúan dificultando la producción de cultivos y perjudicando los ecosistemas a escala mundial (1). Las plantas y los patógenos han venido sosteniendo una relación compleja que por lo general implica la secreción de proteínas efectoras por parte del patógeno para manipular los procesos de las células vegetales y la captación de nutrientes para apoyar su crecimiento y proliferación (2). Aunque, se ha logrado un conocimiento cada vez mayor sobre las proteínas efectoras secretadas que facilitan la colonización del patógeno al huésped (3), se comprende que los organismos patógenos más importantes para los vegetales y animales son los oomicetos (4). Por ello, es tanta la necesidad de comprender los mecanismos que sustentan el parasitismo en este importante grupo de eucariotas, que ha provocado esfuerzos de secuenciación del genoma en varias especies de oomicetos, para traducir los repertorios de genes que están implicados en la infección, la virulencia y los estilos de vida del patógeno (5).

El oomiceto más conocido es *Phytophthora infestans*, agente causal de la enfermedad del tizón tardío en tomate y papa (6). Este microorganismo representa un desafío a controlar, debido a que este patógeno posee un genoma altamente flexible que favorece una

rápida adaptación a las estrategias de control (7), de tal manera, que el estudio de *P. infestans* y otras *Phytophthora spp.*, como por ejemplo; *Phytophthora sojae* y *Phytophthora romarum* han entrado en la era de la genómica, por tanto, se ha evaluado el genoma de estos organismos y se ha determinado que el tamaño del genoma de *P. infestans* es de aproximadamente 240 Mb (Tab.1) (8), con un 51% de secuencias de ADN repetidas (9). Debido a este considerable tamaño y complejidad del genoma, ha conllevado a que los esfuerzos de la genómica de *P. infestans* se centraran en la secuenciación del ADNc para generar colecciones de marcadores de secuencia expresada (EST) de una gran variedad de condiciones de desarrollo, estrés e infección (10). Sin embargo, la anotación precisa de los genomas de *Phytophthora spp.* serán un desafío importante, debido a que la información sobre la estructura de los genes en estas especies se recopila solo de un puñado de ellos, lo cual, resulta en limitaciones significativas para el desarrollo de programas de genes, todo esto es de vital importancia para aumentar el impulso de los estudios de genómica funcional (11).

Además, *P. infestans* es un patógeno hemibiotrófico, el cual, su proceso de infección consiste en dos pasos; uno de ellos consta de una fase biotrófica temprana, seguida de una segunda fase caracterizada por una necrosis extensa en el huésped que permite el crecimiento y esporulación del patógeno (12). Para darse este nivel de infección generado

por el patógeno *P. infestans*, se requiere principalmente una serie de moléculas secretadas por este microorganismo, denominadas efectores, que interfieren con las vías inmunes del huésped y por ello permiten la colonización e infección por parte del patógeno [9], [10]. Por lo general, los oomicetos secretan cientos de proteínas efectoras que se translocan dentro de las células huésped o actúan en el apoplasto (15). Aunque, algunos de los efectores apoplásticos son inhibidores de las hidrolasas vegetales, la mayoría de los efectores citoplásmicos de oomicetos (translocados en el huésped), es decir, los miembros de las familias RXLR y Crinkler (CRN), carecen de similitud con las proteínas conocidas y sus actividades bioquímicas siguen siendo en gran parte desconocidas (16).

No obstante, los efectores citoplasmáticos CRN se identificaron originalmente de *P. infestans*, los cuales se ha demostrado que estos efectores CRN codifican supuestos péptidos que al ser secretados provocan necrosis en la planta, lo cual es una característica de la inmunidad innata de la planta (17,18) termed RXLR (for Arg, any amino acid, Leu, Arg. Por tanto, se ha evidenciado que en los genomas de todas las especies de *Phytophthora spp.* examinadas hasta la fecha albergan grandes repertorios de estos genes efectoras CRN (19). Los cuales, forman la familia más expandida de genes en *P. infestans* (con 196 genes y 254 pseudogenes), esta expansión de los genes CRN en *P. infestans* se debe por su presencia en regiones genómicas con

pocos genes que son ricas en repeticiones y elementos transponibles (20).

De tal manera que las proteínas CRN tienen arquitecturas modulares que incluyen un péptido señal, extremos N conservados y dominios C-terminales muy diversos (21), sin embargo, los extremos CRN N conservados albergan un motivo LXLFLAK distinto, que es seguido por el dominio DWL conservado, adicionalmente, presenta un motivo HVLVXXP altamente conservado que marca el final de los dominios N-terminales de CRN y forman una unión en donde se fusionan diversos dominios C-terminales (20). Por ello, en la actualidad se han identificado distintas proteínas efectoras, entre las que se encuentran las; (CRN2, CRN5, CRN6, CRN8 y CRN12), sin embargo, el CRN8 es el efector más importante para el microorganismo *P. infestans*, ya que requiere de una señal de localización nuclear funcional (NLS) para la acumulación nuclear y la inducción de la muerte celular (22). Esta proteína efectora CRN8 induce la muerte celular cuando se expresa en plantas hospedadoras. Además, actúa como una quinasa y es capaz de autofosforilar, sin embargo, su capacidad para inducir la muerte celular no es un resultado directo de su actividad quinasa, sino una consecuencia del estado fosforilado de los cinco residuos de serina en la proteína CRN8 (21), también está constituida por un dominio D2, el cual, tiene una similitud significativa con las proteínas serina/treonina quinasas (**Fig.1**) de la clase RD (20). Las cua-

les se definen como aquellas quinasas en las que el aspartato catalítico conservado está precedido por un residuo de arginina en el subdominio de quinasa VI (23). Por tanto, al agregar grupos fosfato a las proteínas sustrato, las quinasas dirigen la actividad, la localización y la función general de muchas proteínas, regulando finalmente casi todos los procesos celulares (21).

El objetivo de este estudio es proporcionar, mediante herramientas bioinformáticas, una visión general del gen CRN8, su importancia en la regulación de la proteína efectora arrugada 8 en *Phytophthora infestans* y la prospección del gen para el biocontrol de la producción de esta proteína en particular. De igual manera se busca realizar el modelamiento de la estructura proteica producida por este gen, puesto que en la literatura no se encontraron resultados concluyentes que demuestren un diseño tridimensional concreto de la proteína que es codificada por el gen CRN8.

II. METODOLOGÍA

• Búsqueda bibliográfica

Se realizó una búsqueda sistemática de la literatura acerca del tema en concreto en la base de datos *PubMed* (<https://pubmed.ncbi.nlm.nih.gov/>) (24). La búsqueda se restringió a documentos de investigación y revisiones publicados en los últimos 10 años que tuvieran alta relevancia con el tema en específico. Se utilizaron los términos “*Phytophthora*

infestans”, “Bioinformatics” y “CRN8” utilizando el operador booleano AND.

• Búsqueda de la secuencia del gen CRN8 en *Phytophthora infestans*

Se realizó la búsqueda de la secuencia utilizando la base de datos del Centro Nacional de Información Biotecnológica (NCBI) (<https://www.ncbi.nlm.nih.gov/nucleotide/?term=>) (25), más precisamente del GenBank, con el fin de obtener la secuencia de nucleótidos del gen CRN8, luego mediante el alineamiento usando la Herramienta de Búsqueda Básica de Alineación Local (BLAST) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (26), en la cual se buscó que en el genoma del oomycota, *P. infestans* esté presente el gen CRN8, para esto, se tomó como referencia el genoma del oomycota “*Phytophthora infestans* T30-4”.

• Análisis filogenético de secuencias disponibles del gen CRN8

Se realizó un alineamiento múltiple de secuencias con el fin de llevar a cabo una comparación entre las principales secuencias del gen CRN8 pertenecientes a *Phytophthora infestans* disponibles en la base de datos de genes del NCBI. Para esto se utilizó la herramienta CLUSTAL OMEGA (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (27), del instituto europeo de bioinformática. Luego se realizó la construcción y el análisis de las relaciones filogenéticas mediante la herramienta Phylogeny.fr (http://www.phylogeny.fr/simple_phylogeny.cgi) (28), la

cual hace parte del servicio de Métodos y Algoritmos para Bioinformática del Laboratorio de Informática, Robótica y Microelectrónica de Montpellier. Para la visualización de los árboles filogenéticos se utilizó el programa TreeViewX versión 0.5.0 (29).

- **Modelamiento tridimensional de la proteína codificada por el gen CRN8**

El modelo tridimensional de la proteína codificada por el gen en cuestión se obtuvo empleando la herramienta SWISS-MODEL (<https://swissmodel.expasy.org/>) (30), de la universidad de Basilea, a partir de los datos obtenidos de la transcripción realizada a la secuencia de nucleótidos del gen CRN8. Se tuvieron en cuenta los valores arrojados por el análisis de servicio MolProbity, el cual está integrado en SWISS-MODEL. Después fueron empleadas las herramientas Phyre2 (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>) (31) y I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (32) para la predicción de la estructura mediante el modelamiento por *threading* y *ab initio*. Luego se realizó la visualización de los modelos obtenidos empleando el programa RasMol (versión 2.7.5.2) (33).

- **Identificación de los dominios proteicos**

La identificación de los dominios proteicos de la secuencia de la proteína co-

dificada por el gen CRN8 de *P. infestans*, se realizó mediante la página web de búsqueda MOTIF-search (<https://www.genome.jp/tools/motif/>) (34)

III. RESULTADOS Y DISCUSIÓN

Secuencia del gen CRN8

A partir de la secuencia de nucleótidos correspondiente al gen CRN8 de *P. infestans* (1980 bp), se determinó que esta secuencia se encontró dentro del genoma de referencia, el cual correspondía al oomycota *Phytophthora infestans* cepa T30-4 (código de acceso: GCA_000142945.1) (**Fig.2**).

Al observar la (**Fig.2**), se evidenció que al generarse el alineamiento de la secuencia del gen CRN8 con el genoma de *P. infestans* cepa T30-4 (GCA_000142945.1), se obtuvo que la secuencia del gen CRN8 fue la que mayor porcentaje de identidades produjo, además no se presentaron gaps en el alineamiento para esta cepa en específico del oomycota, por lo cual, se decidió elegir esta secuencia para realizar los demás puntos del estudio. Cabe resaltar que en este alineamiento fue posible evidenciar que el gen CRN8 se encuentra entre las posiciones 27 y 1826 bp del genoma de *P. infestans* T30-4. De igual manera que de forma general se ha indicado que este gen se encuentra en esta posición específica del genoma de este microorganismo, debido a que la secuencia del genoma de *P. infestans* T30-4 proporciona una cobertura casi com-

pleta de genes con un 98,2% en cuanto a la alineación del ADN complementario (20)

Analisis filogenético

La secuencia del gen CRN8 correspondió a la referencia AY961456.1 del NCBI, correspondiente a “*Phytophthora infestans* clon PB021B2 ARNm de CRN8 similar a CRN, CDS completo”, además, se determinó que la proteína codificada por el gen CRN8, corresponde a la referencia AAY43402.1 del NCBI, correspondiente a “CRN8 similar al CRN [*Phytophthora infestans*]” cuya secuencia de aminoácidos es de (599 aa). El alineamiento realizado en BLAST de la secuencia de la proteína codificada por el gen CRN8, permitió la identificación de otras especies de *Phytophthora spp.* (**Tab.2**) que presentaron esta misma proteína, de las cuales se seleccionaron 4 secuencias.

Una vez se seleccionaron las secuencias de proteínas, se realizó un alineamiento múltiple de secuencias por medio de la herramienta Clustal omega (**Fig.3**), disponible en el sitio web del Instituto Europeo de Bioinformática, con la finalidad de alinear las diversas secuencias de proteínas (**Tab.2**) correspondiente a la proteína codificada por el gen CRN8, para lograr la mejor coincidencia entre ellas.

Al observar la (**Fig.3**) se obtuvo que las 4 secuencias al alinearse presentaron tres caracteres, uno de ellos es el punto

(.), el cual indica sustituciones menos conservativas; dos puntos (:), indica posiciones en las que se han realizado sustituciones conservativas y el asterisco (*), indica que en dicha posición los residuos son 100% idénticos, es decir, en algunas columnas del alineamiento (35). Por lo tanto, se logró evidenciar una Leucina muy conservada (flecha amarilla) y un Triptófano muy conservado (flecha roja), los cuales se encontraron en todas las secuencias, sin embargo, la secuencia correspondiente a *P. nicotianae* presentó un mayor número de gaps y sitios con una fuerte similitud, es decir, presentó más diferencias con respecto a las demás secuencias. Al comparar estos resultados obtenidos con otros artículos en cuanto al alineamiento de las secuencias de la proteína codificada por el gen CRN8 en distintas especies del género *Phytophthora spp.*, se determinó que no hay un alineamiento perfecto, debido a que las secuencias van evolucionando más rápido que las estructuras o funcionalidades de la proteína (36). Por otra parte, Loytynoja et al. demostró que los alineamientos con más huecos, es decir, menos compacto como se observa en la (**Fig.3**) coinciden mejor con la filogenia y la estructura de las proteínas (35). Por lo general, no es buena idea trabajar con secuencias demasiado parecidas así como no es buena idea trabajar con secuencias muy disímiles, ya que estos alineamientos resultan no ser informativos (37). De manera similar, en el árbol filogenético (**Fig.4**) se observó que la secuencia correspondiente a *P. nicotianae* estuvo excluida

del clado mayor en el que se encontraron todas las demás secuencias, además, cabe resaltar que fue la secuencia más divergente ya que presentó una distancia genética mayor con respecto al resto.

Al observar la **(Fig.4)**, se determinó que el árbol filogenético está compuesto por nodos interiores (puntos de bifurcación) y ramas (líneas de conexión), que conectan con los nodos terminales ocupados por las secuencias de los aminoácidos correspondiente a la proteína CRN8, sin embargo, los nodos son los que representan un ancestro común inferido (38). Por lo tanto, se logró evidenciar distancias entre las secuencias, lo cual se debe a que las distancias entre las especies del género *Phytophthora* spp. se genera en función de las diferencias entre las secuencias de nucleótidos o aminoácidos de las moléculas homólogas. Además, se ha estimado que la longitud de las ramas depende de las distancias, lo cual indica que se unen las ramas de las secuencias con distancias más cortas (39). Por ello, en la **(Fig.4)**, se evidencia que la secuencia correspondiente a *P. infestans* y *P. infestans* T30-4 se encontraron en el mismo nodo, lo cual indica una similitud alta entre las dos secuencias en cuanto la proteína CRN8 (38).

Modelamiento tridimensional de la proteína codificada por el gen CRN8

Para el modelamiento de la estructura terciaria de la proteína codificada por el gen CRN8 se utilizaron tres herramientas diversas, con la que se obtuvieron

resultados similares en la mayoría de los casos. Mediante el uso de la herramienta SWISS-MODEL, en la cual los modelos se calculan mediante la canalización del modelado por homología basados en el motor del modelado comparativo interno, el cual se enfoca en OpenStructure denominado ProMod3 (40). **(Fig.5)**

Al observar la **(Fig.5)**, se obtuvo el modelo tridimensional de la proteína **(Fig. 5A)**, en el cual se puede apreciar a simple vista una mayor formación de hélices α con respecto a la cantidad de las demás posibles estructuras secundarias, al igual que se observaron los valores de MolProbity (41) 2.24 y de choques estéricos 10.27, siendo los valores ideales para este parámetro de 0 en ambos casos. Al observar el gráfico de Ramachandran **(Fig. 5B)**, en las zonas blancas corresponden a conformaciones donde los átomos del polipéptido están más cercanos que sus radios de van der Waals, y por lo general estas regiones son estéricamente no permitidas para todos los aminoácidos de la proteína, excepto para el aminoácido Glicina, el cual es el único aminoácido que no presenta cadena lateral, en cuanto a las regiones que son de color verde oscuro corresponden a conformaciones donde no hay impedimentos estéricos, es decir estas son zonas permitidas llamadas conformaciones α -hélices y Hojas- β , mientras que las zonas verde claro muestran las regiones permitidas solo si en los cálculos se usan radios van der Waals ligeramente más pequeños, es decir se permi-

te que los átomos puedan estar un poco más cercanos (42). Por lo tanto, se logró evidenciar un mayor agrupamiento de helices α dextrógiras con respecto a las levógiras, así como también es posible evidenciar que hay una gran cantidad de laminas β tanto paralelas como antiparalelas.

Además, se logró determinar la presencia de valores atípicos, que de acuerdo con el análisis integrado de MolProbity, correspondieron al 1,78%, estando este valor por encima del ideal que sería $< 0.2\%$, lo que produjo un número de valores típicos del 89,33%, siendo $> 98\%$ el caso ideal. Cabe mencionar que con respecto al porcentaje de rotámeros atípicos, el modelo de la proteína CRN8 presentó un valor del 1,53%, es decir, fue un resultado que se encontró por encima del valor ideal, el cual es $< 1\%$. Por otro lado, se obtuvo también del modelamiento con SWISS-MODEL la calidad residual (Fig. 5C), que consiste en un análisis de cada uno de los residuos del modelo predicho, los cuales son calificados por un valor de estimación local denominado QMEAN (43), que tiene como valores ideales todos aquellos que se encuentran en $> 0,5$. Por ello, en este gráfico se puede evidenciar que el modelo presenta un QMEAN local (cada residuo) muy bueno, ya que se presentan formaciones de estructuras secundarias típicas, lo que se tradujo en un QMEAN global de -3.57 , es decir, es un modelado de alta calidad, ya que el valor ideal es de > -4.0 .

En otro orden de ideas, se llevó a cabo el modelamiento de la estructura terciaria de la proteína CRN8 usando Phyre2, la cual utiliza herramientas como PSI-BLAST (44), modelos ocultos de Markov (45), 3DLigandSite (46), entre algunas otras más. Todo esto con la finalidad de realizar el alineamiento con estructuras homologas para así lograr el modelamiento de la estructura en cuestión (Fig. 6).

En la (Fig. 6) se puede apreciar el modelo elaborado por Phyre2 de la secuencia de aminoácidos de la proteína CRN8, la cual fue elaborada utilizando como molde la proteína serina/treonina-proteína quinasa vps15 de *Saccharomyces cerevisiae* (código de acceso al PDB: 5DFZ) que resultó ser la molécula con la cual se presentó una mayor cantidad de alineamientos. Además, la identidad de este modelo fue del 16%, y aun así la confianza, el cual es el valor que representa la probabilidad de que la coincidencia entre nuestra secuencia y la plantilla molde representen una verdadera homología, fue del 100%, es decir, es auténtica. Por lo tanto, cabe resaltar que de acuerdo con los diseñadores de esta herramienta, incluso con identidades de secuencia $< 15\%$, los modelos pueden ser muy útiles siempre que la confianza sea alta (31).

Phyre2 también elabora un gráfico en el que se analiza la estructura secundaria y la predicción de desorden del modelo predicho, evidenciando en la (Fig. 7).

Al observar la (**Fig.7**), se determinó que en nuestro modelo se presenta la formación de estructuras secundarias con un nivel aceptable de confianza, sin embargo contrastan algunas regiones en la que se presenta desorden, en los cuales los niveles de confianza son regulares y sugieren que en estos sitios específicos se pueden presentar estructuras atípicas. Lo anterior se supone del hecho de que Phyre2 no predice giros β , curvas β , hélices π o hélices 310, por lo que no se conoce con certeza qué estructuras se generan a partir de estos sitios con niveles bajos de confianza (31)

Esta herramienta, mediante sus algoritmos, identificó a la molécula serina/treonina-proteína quinasa vps15 de *Saccharomyces cerevisiae* como el molde para la predicción de la proteína CRN8, y de acuerdo con el alineamiento de estas dos secuencias (**Fig. 8**).

Al observar la (**Fig.8**), se puede evidenciar la similitud existente entre ellas, en donde se resalta la inserción de nueve aminoácidos en la secuencia de nuestra proteína, así como también 10 pares de deleciones, lo cual generó la no formación de algunas estructuras secundarias que si están presentes en el modelo de la secuencia molde.

Además, esta herramienta permite también la visualización de un ranking de los principales dominios homólogos a nuestra secuencia, por lo que se realizó la superposición de los dos principales homólogos, los cuales fueron los úni-

cos que representaron una confianza del 100%. (**Fig.9**)

En la (Fig.9), se puede resaltar de este alineamiento que tanto la Desviación Cuadrática Media de las Posiciones Atómicas (RMSD) como el Puntaje de Modelado de Plantilla (TM-score) fueron aceptables, ya que se ha definido que la RMSD aceptable para modelo in silico es de 2.5 ángstroms (Å) y de 3 Å para difracción de rayos X (47,48), y en el TM-score dos estructuras son idénticas si presentan un puntaje 1, un puntaje de 0.5 indica el mismo pliegue general para ambas estructuras, y los puntajes < 0.2 indican una similitud no mejor que al azar (49,50).

Finalmente se realizó el modelamiento de la estructura de nuestra proteína utilizando la herramienta I-TASSER (**Fig.10**), la cual integra la predicción de estructuras mediante los métodos de enhebrado y ab initio (32). Esta herramienta genera una valoración del modelo predicho que es denominada puntaje de confianza (C-score), en donde los puntajes > -1.5 indican que el modelo posee una correcta topología global en su estructura (32).

Al observar la (**Fig.10**), Se obtuvo entonces un modelo estructural de la proteína CRN8 (**Fig. 10A**) con unos valores aceptables respecto al C-score (-0,03) y al TM-score (0.71 ± 0.12) que contrastaron con el valor regular de la RMSD (5.3 ± 3.4 Å). Con respecto a lo anterior se puede confirmar que esta herramien-

ta presenta una fuerte correlación entre los puntajes de confianza y de modelado de plantilla, mientras que la correlación con la RMSD no es muy evidente (32). Por otro lado se puede apreciar el Factor B (Fig. 10B), un valor para indicar el alcance de la movilidad térmica inherente de los residuos/átomos en las proteínas, el cual se deduce del enhebrado de proteínas de plantilla del PDB en combinación con los perfiles de secuencia derivados de las bases de datos de secuencias (51). Se observa que mayormente en las hélices y láminas el Factor B fue menor, es decir, estos residuos fueron relativamente más estables en la estructura con respecto a las curvas que presentaron puntajes por encima de cero.

En la (Fig.11) se puede apreciar el alineamiento de la estructura predicha de nuestra proteína en comparación con la más parecida estructuralmente que, similar al alineamiento de Phyre2, fue la serina/treonina-proteína quinasa 24 de *Homo sapiens* (código de acceso al PDB: 3CKX). En este caso se obtuvieron buenos puntajes tanto para la RMSD como para el TM-score, sin embargo fue mejor el TM-score del alineamiento de la estructura predicha por I-TASSER con la estructura modelo, resaltando que ocurrió lo contrario con la RMSD ya que en la comparación realizada por Phyre2 esta fue más cercana a 2. Por otra parte, se obtuvo el modelado de los sitios de unión a ligando de la proteína CRN8, encontrando que el ligando con el mejor C-score (0.21) fue

la S-adenosil-L-homocisteína (Fig.12), la cual es un sulfuro orgánico y un inhibidor competitivo de las metiltransferasas de ADN (52). Se evidenció que en este modelado los residuos del sitio de unión al ligando fueron 129, 131, 135, 152, 153, 154, 172, 173, 174, 175, 190.

Dominios proteicos de la proteína CRN8

Se obtuvieron 3 dominios proteicos dentro de la secuencia de la proteína CRN8, (Fig. 13).

En la (Fig.13), se logró determinar 3 dominios, entre los cuales se encuentran dos dominios de la proteína quinasas de dominio catalítico, y el dominio similar a ubiquitina (Ubl), el primero consta de; la superfamilia de proteínas quinasas, las cuales están compuesta principalmente por los dominios catalíticos de proteínas quinasas específicas de serina / treonina y específicas de tirosina. También incluye RIO quinasas, que son serina proteína quinasas atípicas, aminoglucósido fosfotransferasas y colina quinasas. Estas proteínas catalizan la transferencia del grupo gamma-fosforilo de ATP a grupos hidroxilo en sustratos específicos como serina, treonina o residuos de tirosina de proteínas (53). Mientras que el otro dominio, consta de; que las proteínas de tipo ubiquitina (Ubl) tienen un pliegue de agarre beta de ubiquitina (Ub) similar y se unen a otras proteínas de una manera Ubl pero con funciones bioquímicamente distintas. Las proteínas Ub y Ubl se conjugan

y desconjugan mediante ligasas y peptidasas para modificar covalentemente los polipéptidos diana. Algunos dominios Ubl tienen roles adaptadores en la señalización de Ub al mediar la interacción proteína-proteína (54).

IV. CONCLUSIONES

En los últimos años se han llevado a cabo diversas investigaciones que involucran el estudio del gen CRN8 perteneciente al microorganismo *Phytophthora infestans*. En estos trabajos se lleva a cabo mayormente, la búsqueda de métodos físicos, químicos y biológicos de hacer que este gen regulador de la proteína efectora arrugada 8 sea regulado y así evitar que cultivos de alimentos que son de gran importancia para el consumo lleguen a estar contaminados con la expresión de este efector. El estudio estructural de la proteína codificada por este gen sin duda representa una gran importancia para lograr dilucidar su papel dentro de la síntesis, puesto que no es mucho lo que se conoce al respecto. Se obtuvieron hallazgos importantes como su homología con diversas proteínas de origen fúngico que tienen como cualidad ser enzimas del tipo quinasa. Por otro lado, se logró evidenciar mediante la herramienta I-TASSER el sitio de unión a ligando, encontrando que este sitio encaja perfectamente con la molécula S-adenosil-L-homocisteína. Quedan aún muchas incógnitas sobre la expresión y funcionamiento de esta proteína, por lo que queda abierta la oportunidad a futuras investigaciones

tanto de los genes reguladores como de los demás genes estructurales que hacen parte del clúster encargado de la síntesis de la proteína efectora CRN8 en determinadas condiciones. Por otro lado, se resalta el avance inmenso que se ha realizado en herramientas de predicción de modelos de estructuras terciarias de proteínas, puesto que permiten que investigaciones como esta puedan llevarse a término.

VI. REFERENCIAS

1. Stam R, Jupe J, Howden AJM, Morris JA, Boevink PC, Hedley PE, et al. Identification and Characterisation CRN Effectors in *Phytophthora capsici* Shows Modularity and Functional Diversity. *PLoS One*. 2013;8(3):1–13.
2. Rovenich H, Boshoven JC, Thomma BPHJ. Filamentous pathogen effector functions: Of pathogens, hosts and microbiomes. *Curr Opin Plant Biol*. 2014;20:96–103.
3. Whisson SC, Boevink PC, Wang S, Birch PRJ. The cell biology of late blight disease. *Curr Opin Microbiol*. 2016;34:127–35.
4. Derevnina L, Petre B, Kellner R, Dagdas YF, Sarowar MN, Giannakopoulou A, et al. Emerging oomycete threats to plants and animals. *Philos Trans R Soc B Biol Sci*. 2016;371(1709).

5. Bozkurt TO, Schornack S, Banfield MJ, Kamoun S. Oomycetes, effectors, and all that jazz. *Curr Opin Plant Biol.* 2012;15(4):483–92.
6. Kamoun S, Furzer O, Jones JDG, Judelson HS, Ali GS, Dalio RJD, et al. The Top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol.* 2015;16(4):413–34.
7. Fry WE. *Phytophthora infestans*: New Tools (and Old Ones) Lead to New Understanding and Precision Management. *Annu Rev Phytopathol.* 2016;54:529–47.
8. Tooley PW, Therrien CD. Cytophotometric determination of the nuclear DNA content of 23 Mexican and 18 non-Mexican isolates of *Phytophthora infestans*. *Exp Mycol.* 1987;11(1):19–26.
9. Judelson HS, Randall TA. Families of repeated DNA in the oomycete *Phytophthora infestans* and their distribution within the genus. *Genome.* 1998;41(4):605–15.
10. Randall TA, Dwyer RA, Huitema E, Beyer K, Cvitanich C, Kelkar H, et al. Large-scale gene discovery in the oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. *Mol Plant-Microbe Interact.* 2005;18(3):229–43.
11. Win J, Kanneganti TD, Torto-Alalibo T, Kamoun S. Computational and comparative analyses of 150 full-length cDNA sequences from the oomycete plant pathogen *Phytophthora infestans*. *Fungal Genet Biol.* 2006;43(1):20–33.
12. Pacheco JM. Efectores de oomycetes fitopatogénicos: en la primera línea de ataque. *Fitosanidad.* 2015;19(3):251–7.
13. Chisholm ST, Coaker G, Day B, Staskawicz BJ. Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell.* 2006;124(4):803–14.
14. Hogenhout SA, Van Der Hoorn RAL, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-associated organisms. *Mol Plant-Microbe Interact.* 2009;22(2):115–22.
15. Schornack S, Huitema E, Cano LM, Bozkurt TO, Oliva R, Van Damme M, et al. Ten things to know about oomycete effectors. *Mol Plant Pathol.* 2009;10(6):795–803.
16. Bos JIB, Kanneganti TD, Young C, Cakir C, Huitema E, Win J, et al. The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death

- in *Nicotiana benthamiana*. *Plant J*. 2006;48(2):165–76.
17. Win J, Morgan W, Bos J, Krasileva K V, Cano LM, Chaparro-Garcia A, et al. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell*. 2007;19(8):2349–69.
 18. Torto TA, Li S, Styer A, Huitema E, Testa A, Gow NAR, et al. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res*. 2003;13(7):1675–85.
 19. Kamoun S. Groovy times: filamentous pathogen effectors revealed. *Curr Opin Plant Biol*. 2007;10(4):358–65.
 20. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*. 2009;461(7262):393–8.
 21. Van Damme M, Bozkurt TO, Cakir C, Schornack S, Sklenar J, Jones AME, et al. The Irish Potato Famine Pathogen *Phytophthora infestans* Translocates the CRN8 Kinase into Host Plant Cells. *PLoS Pathog*. 2012;8(8).
 22. Schornack S, Van Damme M, Bozkurt TO, Cano LM, Smoker M, Thines M, et al. Ancient class of translocated oomycete effectors targets the host nucleus. *Proc Natl Acad Sci U S A*. 2010;107(40):17421–6.
 23. Pelech S. of Biology Comment BioMed Central Dimerization in protein kinase signaling. *J Biol*. 2006;5(5):1–7.
 24. Yang H, Lee HJ. Research trend visualization by MeSH terms from PubMed. *Int J Environ Res Public Health*. 2018;15(6).
 25. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank. *Nucleic Acids Res*. 2007;35(SUPPL. 1):21–5.
 26. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
 27. Madeira F, Park Y mi, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res [Internet]*. 2019 Jul 2;47(W1):W636–41. Available from: <https://academic.oup.com/nar/article/47/W1/W636/5446251>
 28. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res [Internet]*. 2008 May 19;36(Web Server):W465–9. Available from: <https://academic.oup.com/nar/article/36/7/W465/2875132>

com/nar/article-lookup/doi/10.1093/nar/gkn180

29. Page RDM. Treeview: An application to display phylogenetic trees on personal computers. *Bioinformatics*. 1996;12(4):357–8.
30. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018 Jul;46(W1):W296–303.
31. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845–58.
32. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9:1–8.
33. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. *Trends Biochem Sci*. 1995;20(9):374–6.
34. Venkataraman A, Chew TH, Hussein ZAM, Shamsir MS. A protein short motif search tool using amino acid sequence and their secondary structure assignment. *Bioinformatics*. 2011;7(6):304–6.
35. Löytynoja A, Goldman N. An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A*. 2005;102(30):10557–62.
36. Santamaría R. Alineamientos de múltiples secuencias. 2015;55.
37. Salcedo R. Introducción a la Bioinformática Práctica 2 : Alineamiento múltiple e Identificación y búsqueda de Motivos . 2012;
38. García Lopez M. Bioinformática en el aula: construcción de árboles filogenéticos a partir de bases de datos moleculares. *Rev Comun y nuevas Tecnol*. 2011;8(14):341–50.
39. Rodríguez Tello E. Conceptos básicos de filogenética molecular. *Conceptos básicos filogenética Mol*. 2013;1–43.
40. Peitsch MC. ProMod and Swiss-model: Internet-based tools for automated comparative protein modelling. *Biochem Soc Trans*. 1996;24(1):274–9.
41. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr*. 2010;66(1):12–21.
42. José Antonio Martínez Oyanedel. Principios de Estructura de Proteínas. Departamento de Biología Molecular . 2012. p. 97–003.

43. Benkert P, Tosatto SCE, Schomburg D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins Struct Funct Genet.* 2008;71(1):261–77.
44. Zhuang Y, Zhou X, Wang S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Plant Syst Evol.* 2012;298(7):3389–402.
45. Eddy SR. Hidden Markov models. *Curr Opin Struct Biol.* 1996;6(3):361–5.
46. Wass MN, Kelley LA, Sternberg MJE. 3DLigandSite: Predicting ligand-binding sites using similar structures. *Nucleic Acids Res.* 2010;38(SUPPL. 2):469–73.
47. Allouche A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. *J Comput Chem.* 2012;32:174–82.
48. Tsai H-HG, Tsai C-J, Ma B, Nussinov R. In silico protein design by combinatorial assembly of protein building blocks. *Protein Sci.* 2009;13(10):2753–65.
49. Zhang Y, Skolnick J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 2005;33(7):2302–9.
50. Xu J, Zhang Y. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics.* 2010;26(7):889–95.
51. Yang J, Wang Y, Zhang Y. ResQ: An Approach to Unified Estimation of B-Factor and Residue-Specific Error in Protein Structure Prediction. *J Mol Biol.* 2016;428(4):693–701.
52. Tehlivets O, Malanovic N, Visram M, Pavkov-Keller T, Keller W. S-adenosyl-L-homocysteine hydrolase and methylation disorders: Yeast as a model system. *Biochim Biophys Acta - Mol Basis Dis.* 2013;1832(1):204–15.
53. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 2017;45(D1):D200–3.
54. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* 2020;48(D1):D265–8.

TABLAS.

Tabla 1. Algunas características del genoma de tres especies de *Phytophthora spp.*

	<i>P. infestans</i>	<i>P. romarum</i>	<i>P. sojae</i>
Tamaño del genoma (Mb)	240	65	95
Total de genes	17797	14451	16988
Genes CRN	196	19	100
Pseudogenes CRN	254	41	102
Total CRN	450	60	202

Tabla 2. Secuencias elegidas de la proteína codificada por gen CRN8 en distintas especies de *Phytophthora spp.* obtenidas de la base de datos del NCBI.

Nombre de la secuencia	Longitud (aa)	Código de acceso (GenBank)
CRN8 similar al CRN [<i>Phytophthora infestans</i>]	599	AAY43402.1
Proteína de la familia Crinkler (CRN) [<i>Phytophthora infestans</i> T30-4]	599	XP_002900721.1
Proteína CRN8 [<i>Phytophthora nicotianae</i>]	446	KUF93016.1
Proteína de la familia Crinkler (CRN) [<i>Phytophthora palmivora</i> var. <i>palmivora</i>]	599	POM77496.1

FIGURAS.

Figura 1. Estructura de la proteína serina/treonina quinasa. A) Fosfoserina, B) Fosfo-treonina. que comparten una similitud con la proteína Clinker CRN8.

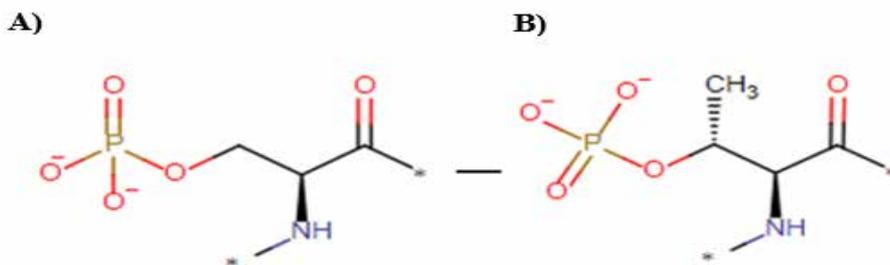


Figura 2. Alineamiento por BLAST del gen CRN8 con el genoma de *Phytophthora infestans* cepa T30-4 (GCA_000142945.1)

Phytophthora infestans T30-4
 Sequence ID: [XM_002900675.1](#) Length: 2427 Number of Matches: 1

Range 1: 8 to 1044 [GenBank](#) [Graphics](#)

Score	Expect	Identities	Gaps	Strand
3395 bits(1638)	0.0	1904/1937(98%)	0/1937(0%)	Plus/Plus
Query 27	ATGGTTACTTTGTTCTGTGCCGTGTTGGTGTGGCTGGGAGCACGTTCCTGTGGACATT	86		
Sbjct 8	ATGGTTACTTTGTTCTGTGCCGTGTTGGTGTGGCTGGGAGCACGTTCCTGTGGACATT	67		
Query 87	AACGAGAACAAAGTCGGTGGGCACTTGAAGAAAGCGATCAAGGACGAGAAAATGTACCAG	146		
Sbjct 68	AACGAGAACAAAGTCGGTGGGCACTTGAAGAAAGCGATCAAGGACGAGAAAATGTACCAG	127		
Query 147	TTTCTGCCGATGAGCTGCAGCTCTTCTGGCAAAGGC GGGGGGCAACGCATGECATACG	206		
Sbjct 128	TTTCTGCCGATGAGCTGCAGCTCTTCTGGCAAAGGC GGGGGGCAACGCATGECATACG	187		
Query 207	AGCCTCACTGAAGATGTGAAGAAAGCTGAAAAGGGCGAGAAGACTGCTCTCGTCAAAATCT	266		
Sbjct 188	AGCCTCACTGAAGATGTGAAGAAAGCTGAAAAGGGCGAGAAGACTGCTCTCGTCAAAATCT	247		
Query 267	CTAACGCAAGAAAGAGAGGAGCTGCAGGGAGAGGATCCGATTTCTGAGTGTCTGGAAAGC	326		
Sbjct 248	CTAACGCAAGAAAGAGAGGAGCTGCAGGGAGAGGATCCGATTTCTGAGTGTCTGGAAAGC	307		
Query 327	ATGGATCCCCGAAAAGTGAACAGATTACGTGCTGGTGGCGCTGCCCTCTGGGACAAGT	386		
Sbjct 308	ATGGATCCCCGAAAAGTGAACAGATTACGTGCTGGTGGCGCTGCCCTCTGGGACAAGT	367		
Query 387	AGTGCCTTATTCTGATGGGACGGACTTGTGGCTGTCGAGATTTCAACATAGTGAAGTT	446		
Sbjct 368	AGTGCCTTATTCTGATGGGACGGACTTATGGCTGTCGAGATTTCAACATAGTGAAGTT	427		
Query 447	GC AAAA CTTACGCTACTGCCGACACGCGGAGATCTAAATGAATTTATCGGCAACCTCTC	506		
Sbjct 428	GC AAAA CTTACGCTACTGCCGACACGCGGAGATCTAAATGAATTTATCGGCAACCTCTC	487		
Query 507	CTGTAAAGATTGGGTTACCGCAAGTCTGTGTTTCAAGCTTGGTCAAGTCCATTGATACTA	566		
Sbjct 488	CTGTAAAGATTGGGTTACCGCAAGTCTGTGTTTCAAGCTTGGTCAAGTCCATTGATACTA	547		
Query 567	GGCCAACTCCTTCGAGATAAAATGTTCCAGCTAAACGACATTTCCGCGTCCGAAATTTTA	626		
Sbjct 548	GGCCAACTCCTTCGAGATAAAATGTTCCAGCTAAACGACATTTCCGCGTCCGAAATTTTG	607		
Query 627	AAAGATTGCGTGTGTTAGCGCTGCGTTCCGTACCCGACAGTGGATGGGGATGCCACAGAG	686		
Sbjct 608	AAAGATTGCGTGTGTTAGCGCTGCGTTCCGTACCCGACAGTGGATGGGGATGCCACAGAG	667		
Query 687	AGTGCATTTCAATATTTCTGGGATCAATATCCGTGTGGTGTCTGGGTTTGTGTTTAGG	746		
Sbjct 668	AGTGCATTTCAATATTTCTGGGATCAATATCCGTGTGGTGTCTGGGTTTGTGTTTAGG	727		
Query 747	CGTGCTTACGTCAACCGGGATTGAGCAGGAAATCCTCGTACGGTTGAAGCGACCCGAC	806		
Sbjct 728	CGTGCTTACGTCAACCGGGATTGAGCAGGAAATCCTCGTACGGTTGAAGCGACCCGAC	787		
Query 807	TTTTTGTGTGCTTTGGATCATATCTGTGTTTTTCTGGTGGAGGAAAAGAACCTCGTACT	866		
Sbjct 788	TTTTTGTGTGCTTTGGATCATATCTGTGTTTTTCTGGTGGAGGAAAAGAACCTCGTACT	847		
Query 867	TCTATCACTGTGCCTCGTGAAGGAGCTCTCAAAAAGCTGGTGGTCTACGGTGGAGTG	926		
Sbjct 848	TCTATCACTGTGCCTCGTGAAGGAGCTCTCAAAAAGCTGGTGGTCTACGGTGGAGTG	907		
Query 927	CCGTATGTGTTGGCTATGCTGCGTGGGCTTCGAGTAGAATTTATGCTATCATCAA	986		
Sbjct 908	CCGTATGTGTTGGCTATGCTGCGTGGGCTTCGAGTAGAATTTATGCTATCATCAA	967		
Query 987	GACGTACGGGCAACGTTAAAACCCACCTTATTGGAGGGTTAACTTGACAGATGCCGCCG	1046		
Sbjct 968	GACGTACGGGCAACGTTAAAACCCACCTTATTGGAGGGTTAACTTGACAGATGCCGCCG	1027		
Query 1047	GAAAGTTTTGCACTTGTACTGGCACTTCTGAATTTGTGTCTGCTTTTCCGGCAATTTGG	1106		
Sbjct 1028	GAAAGTTTTGCACTTGTACTGGCACTTCTGAATTTGTGTCTGCTTTTCCGGCAATTTGG	1087		
Query 1107	CAAAACTGTCCGGCTCGGCAAGGACCGAGTTCATGGACATCCATCGGCCAATGGCGTCT	1166		
Sbjct 1088	CAAAACTGTCCGGCTCGGCAAGGACCGAGTTCATGGACATCCATCGGCCAATGGCGTCT	1147		
Query 1167	AAGGTGCGGTTGAGTCCGATTTTCTCGATAAGATCTTTCATACACAAGAAGAGATTCGT	1226		
Sbjct 1148	AAGGTGCGGTTGAGTCCGATTTTCTCGATAAGATCTTTCATACACAAGAAGAGATTCGT	1207		
Query 1227	CGCGTGAAGCAGATCTACGATTTTGAAGCTTATGGGTTCTTGTGCGCGATGCAAGT	1286		
Sbjct 1208	CGCGTGAAGCAGATCTACGATTTTGAAGCTTATGGGTTCTTGTGCGCGATGCAAGT	1267		
Query 1287	GTACGGTGGATTACAGCAATGAGGTTGACACTTAAGCCCGCGGCTTGAAGTGAAG	1346		
Sbjct 1268	GTACGGTGGATTACAGCAATGAGGTTGACACTTAAGCCCGCGGCTTGAAGTGAAG	1327		
Query 1347	CCATGCAGCTTGAAGTGAAGTGTTCGTGCGCTCGGGAACTTCTCGAGGCCCTGTAGTC	1406		
Sbjct 1328	CCATGCAGCTTGAAGTGAAGTGTTCGTGCGCTCGGGAACTTCTCGAGGCCCTGTAGTC	1387		

Query 987 GACGTCACGGGCAACGTTAAAACCCACCTTATTGGAGGGTTAACTTGCAGCATGCGCCG 1046
 |||
 Sbjct 968 GACGTCACGGGCAACGTTAAAACCCACCTTATTGGAGGGTTAACTTGCAGCATGCGCCG 1027

Query 1047 GAACGTTTTGACTTGTACTGGCACTTCTGAATTTGTGTCTGCTTTCCCGCAATTGTG 1106
 |||
 Sbjct 1028 GAACGTTTTGACTTGTACTGGCACTTCTGAATTTGTGTCTGCTTTCCCGCAATTGCG 1087

Query 1107 CAAAACGTCCGGCTCGGCAGGGACCGAGTTCATGGACATCCATCGGCCAATGGCGTC 1166
 |||
 Sbjct 1088 CAAAACGTCCGGCTCGGCAGGGACCGAGTTCATGGACATCCATCGGCCAATGGCGTT 1147

Query 1167 AAGGTGCGGTTGAGTCCGATTTTCGTGATAAGATCTTTCACACACAAGAAGAGTATCGT 1226
 |||
 Sbjct 1148 AAGGTGCGGTTGAGTCCGATTTTCGTGATAAGATCTTTCACACACAAGAAGAGTATCGT 1207

Query 1227 CGCGTGAAGCAGATCTACGATTCTTTGAAAGCTTATGGGGTTCCTTGTGCCGATGCAGTC 1286
 |||
 Sbjct 1208 CGCGTGAAGCAGATCTACGATTCTTTGAAAGCTTATAGGATTCCTTGTGCCGATGCAGTT 1267

Query 1287 GTCACGGTGGATTGAGCAATGAGGTTGACACTTAAGCCCGCGGATTGAGATGAAG 1346
 |||
 Sbjct 1268 GTCACGGTGGATTGAGCAATGAGGTTGACACTTAAGCCCGCGGATTGAGATGAAG 1327

Query 1347 CCATGCAGCTTGAAGTGTTCGTGCGCTCGGGAACGTTCTCGAGGCCCTGTAGTC 1406
 |||
 Sbjct 1328 CCATGCAGCTTGAAGTGTTCGTGCGCTCGGGAACGTTCTCGAGGCCCTGTAGTC 1387

Query 1407 CTCCATCGCAACGGGTGGATGCACAGAGACATCCGTTGGTCTAACGTGATCAAGCATATC 1466
 |||
 Sbjct 1388 CTCCATCGCAACGGGTGGATGCACAGAGACATCCGTTGGTCTAACGTGATCAAGCATATC 1447

Query 1467 GACCGGTTGAATGGTTTTGATCGACTTCGCTGATGCAGCTCAAAGTCCGCAGAAATAT 1526
 |||
 Sbjct 1448 GACCGGTTGAATGGTTTTGATCGACTTCGCTGATGCAGCTCAAAGTCCGCAGAAATAT 1507

Query 1527 CCCAGCGGCGATCACTTGACGACGACGAACACGCTTCAGATATTTTATGGAAGGTGGA 1586
 |||
 Sbjct 1508 CCCAGCGGCGATCACTTGACGACGACGAACACGCTTCAGATATTTTATGGAAGGTGGA 1567

Query 1587 TCCCACACGACTGCGGTTGACTTATGGCCGTGGGATACCTGGTGAAGACAAGTAAGATT 1646
 |||
 Sbjct 1568 TCCCACACGACTGCGGTTGACTTATGGCCGTGGGATACCTGGTGAAGACAAGTAAGATT 1627

Query 1647 GAACGGGAATGGACTGCGGAGCCCGAAAAGAGCGCTTTTCTAGATCGGTTGATGAACCCA 1706
 |||
 Sbjct 1628 GAACGGGAATGGACTGCGGAGCCCGAAAAGAGCGCTTTTCTAGATCGGTTGATGAACACA 1687

1707 GACCCGAGTGCTCGGCCGACGGCAGACGAAGCCTTACAGCTTCTGTACGTTTTGAACGT 1766
 |||
 1688 GACCCGAGTGCTCGGCCGACGGCAGACGAAGCCTTACAGCTTCTGTACGTTTTGAACGT 1747

1767 GAAGCTGCTGAGCAGGAATCACAGGGCAAGGGTGTGCGCAAGAAAGCACAGACGTGCCTGA 1826
 |||
 1748 GAAGCTGCTGAGCAGGAATCACAGGGCAAGGGTGTGCGCAAGAAAGCAAAGACGTGCGTAA 1807

1827 ATTTTTACATCGCTATGGCACTACCATTACGCGCTGGGACGAAGTACTATGGTACTAACA 1886
 |||
 1808 ATTTTTGCATCGCTATGGCACTACCATTACGCGCTGGGACAAAGTACTATGGTACTAACA 1867

1887 ACAGGAGTAGTCGATGAATGAGCAACACAGCTTTACTTCGAATGATGAGAAATCTCGGC 1946
 |||
 1868 ACAGGAGTAGTCGATGAATGAGCAACACAGCTTTACTTCGAATGATGAGAAATCTCGGT 1927

1947 TGACGAGAGAAACTCGT 1963
 |||
 1928 TGACGAGAGAAACTCGT 1944

Figura 3. Alineamiento múltiple de la secuencia de los aminoácidos pertenecientes a cada una de las secuencias correspondiente a la proteína codificada por el gen CRN8 obtenidas en la base de datos del NCBI realizado mediante Clustal Omega.



Figura 4. Árbol filogenético obtenido del alineamiento de las secuencias de aminoácidos de la proteína CRN8 (proteína efectora arrugada 8) mediante la herramienta Phylogeny.ft.

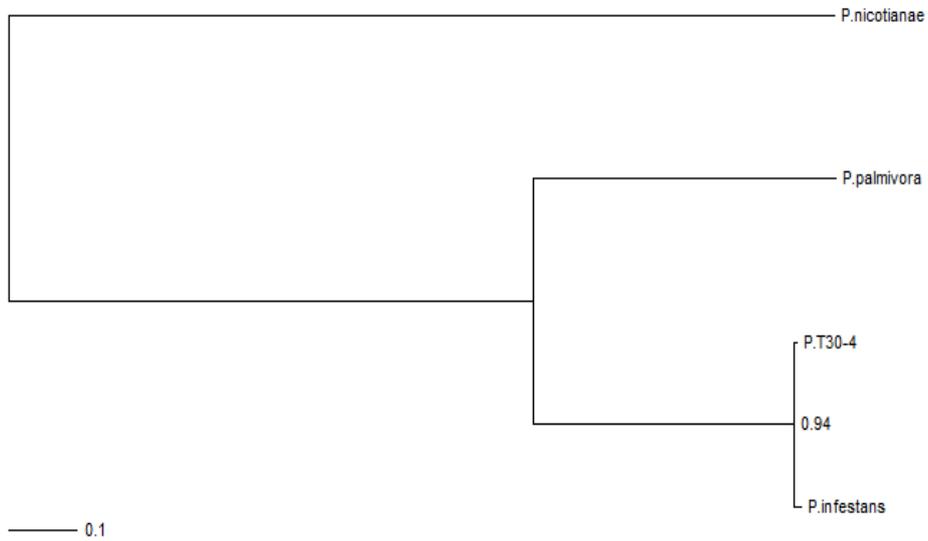
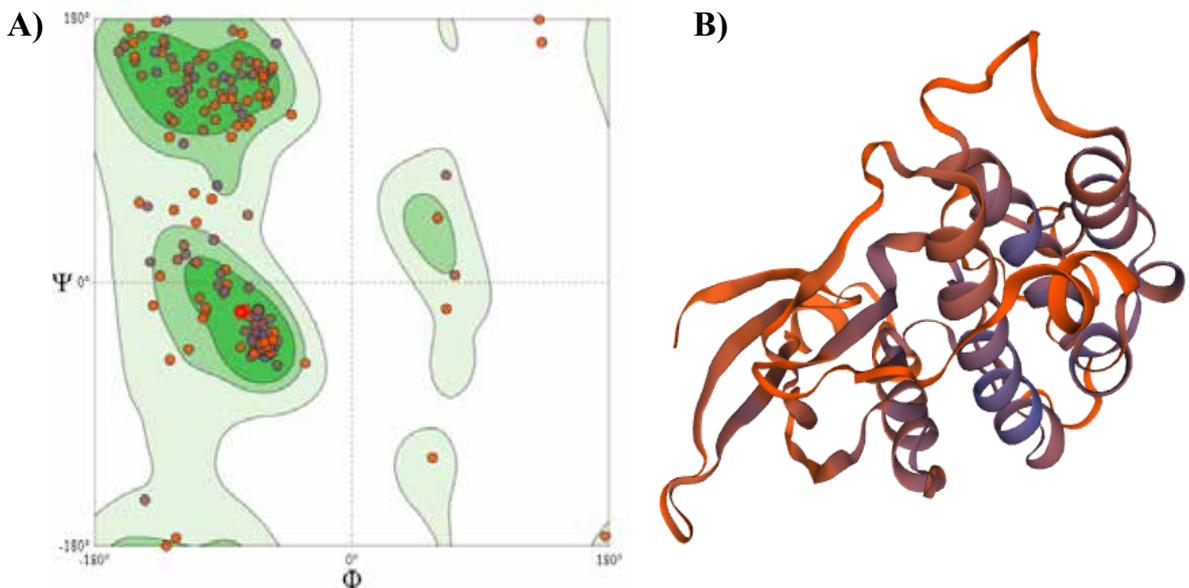


Figura 5. Modelado de la estructura terciaria de la proteína CRN8 mediante la herramienta SWISS-MODEL. **(A)** Modelo tridimensional predicho de la proteína CRN8. **(B)** Gráfico de Ramachandran general. **(C)** Calidad residual del modelo estructural. B: residuo en puente β aislado; C: bucle o irregular, E: cadena extendida, participa en escalera & β ; G: hélices 3_{10} ; H: hélice α ; I: hélice π ; T: giro unido por hidrógeno; S: curva.



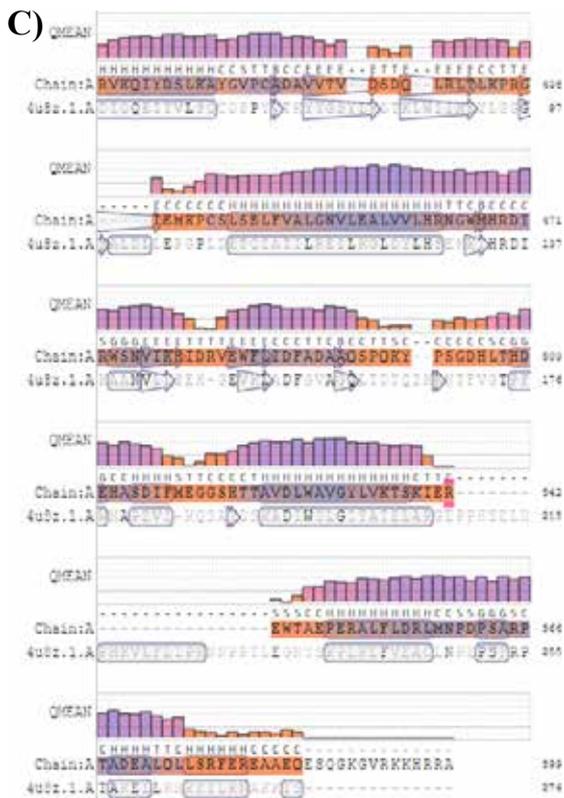


Figura 6. Modelo tridimensional predicho por Phyre2 de la proteína CRN8 basado en el modelo de la molécula serina/treonina-proteína quinasa vps15 de *Saccharomyces cerevisiae* (código de acceso al PDB: 5DFZ)

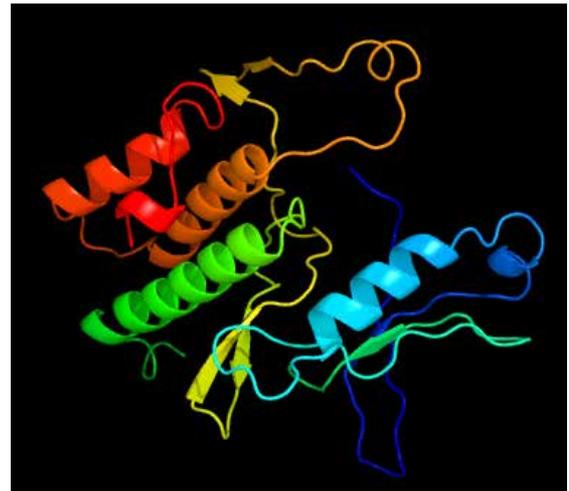
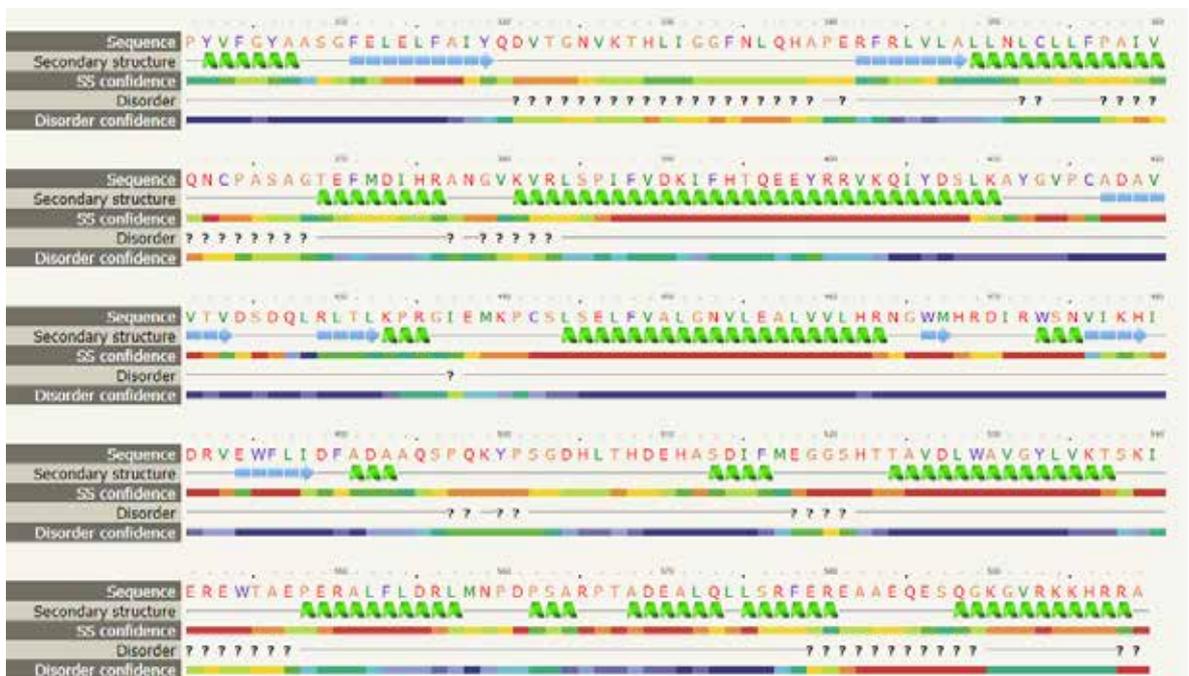


Figura 7. Análisis de la estructura secundaria y predicción de desorden por Phyre2 sobre la estructura proteica modelada de CRN8.



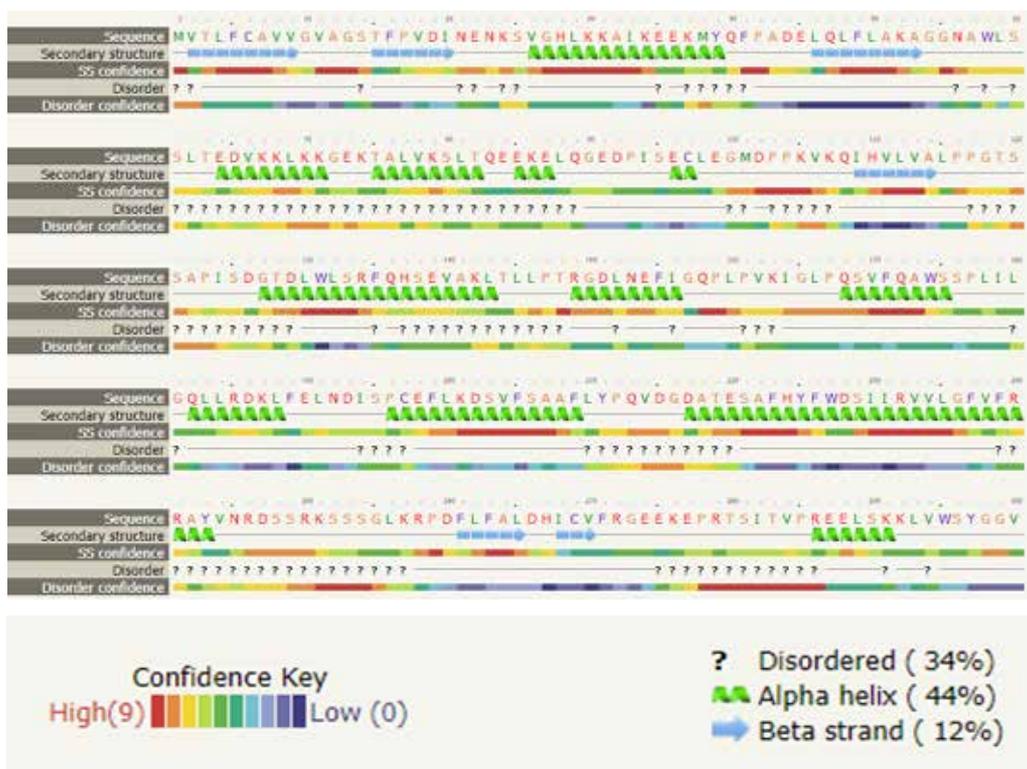
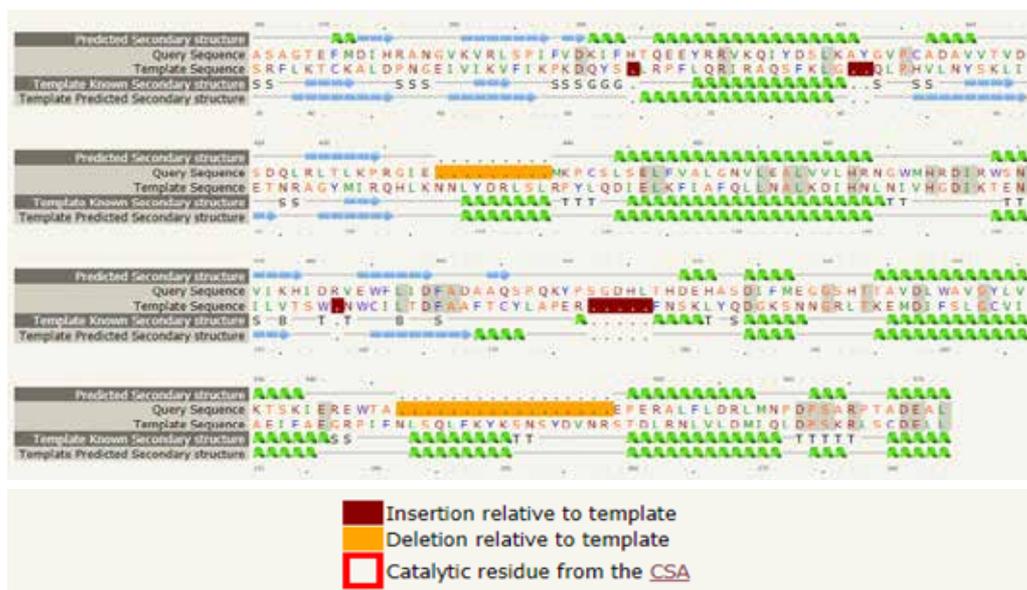


Figura 8. Alineamiento de los modelos estructurales de la proteína de CRN8 generada por Phyre2 y el modelo de serina/treonina-proteína quinasa vps15 de *S. cerevisiae* (código de acceso al PDB: 5DFZ). G: hélice de 3 vueltas (hélices 310); I: hélice de 5 vueltas (hélice π); T: giro unido por hidrógeno; B: residuo en puente β aislado; S: curva.



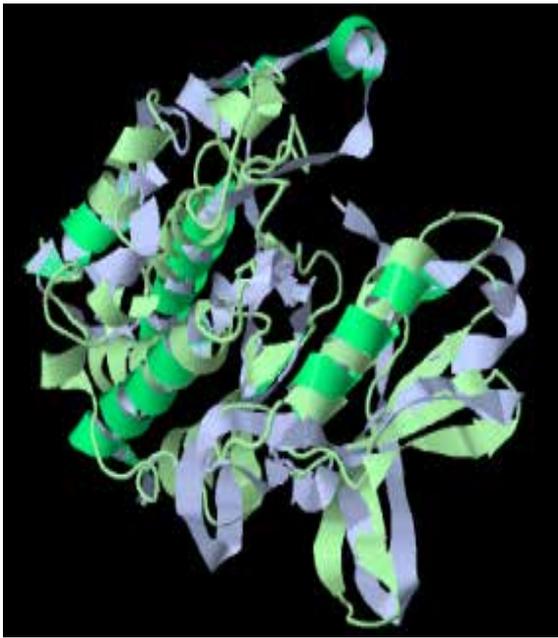


Figura 9. Superposición realizada por Pyre2 del modelo de CRN8 basado en el modelo de serina/treonina-proteína quinasa vps15 de *S. cerevisiae* (código de acceso al PDB: 5DFZ) y serina/treonina-proteína quinasa 24 de *Homo sapiens* (código de acceso al PDB: 3CKX).

Figura 10. Modelado de la estructura terciaria de la proteína CRN8 mediante la herramienta I-TASSER. **(A)** Modelo tridimensional predicho de la proteína CRN8. **(B)** Factor B normalizado predicho obtenido del enhebrado de proteínas de plantilla del PDB en combinación con los perfiles de secuencia derivados de las bases de datos de secuencias

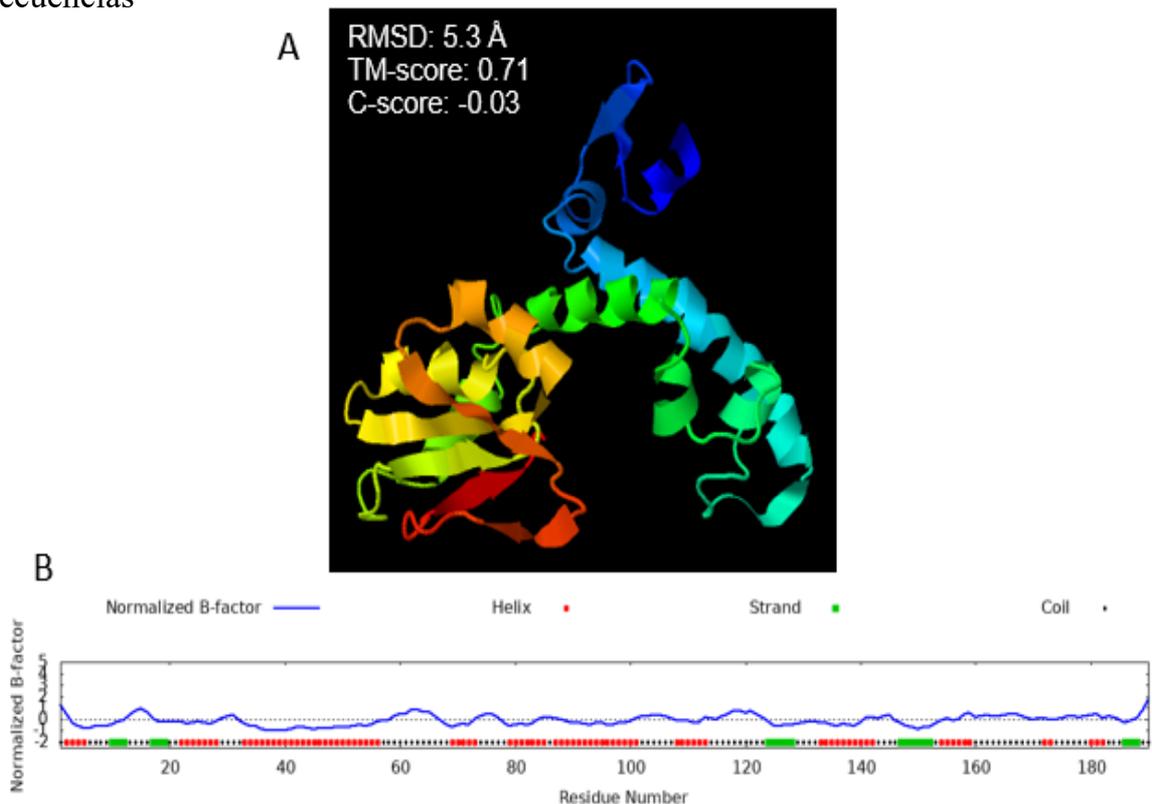


Figura 11. Alineamiento de la estructura entre el modelo de CRN8 realizado por I-TASSER (mostrado en dibujo) y el modelo de serina/treonina-proteína quinasa 24 de *Homo sapiens* (código de acceso al PDB: 3CKX) (mostrado en esqueleto)

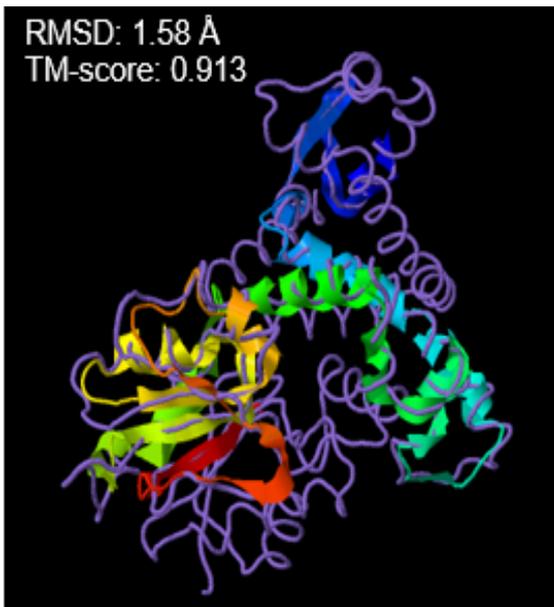


Figura 12. Modelado de los sitios de unión a ligando de la estructura tri-dimensional de CRN8 predicha por I-TASSER.

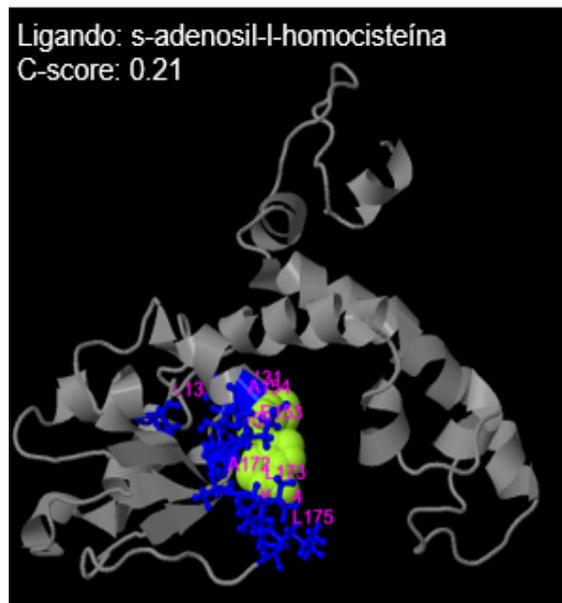


Figura 13. Dominios proteicos identificados en la secuencia de la proteína CRN8

