

**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACION EN DINAMICA CELULAR

**“Análisis del Desorden Estructural en el Proteoma de
Adenovirus Humano”**

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS

PRESENTA

REINIER FELIPE CÁRDENAS MENA

DIRECTOR DE TESIS

Dr. Carlos Amero



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACION EN DINAMICA CELULAR

Análisis del Desorden Estructural en el Proteoma de Adenovirus Humano

Tesis para optar por el grado de Maestro en Ciencias presenta:

Lic. Reinier Felipe Cárdenas Mena

Asesor

Dr. Carlos Amero Tello

Comité Tutorial

Dra. Carmen Nina Pastor Colón

Centro de Investigación en Dinámica Celular

UAEM

Dr. Ramón Antonio Gonzalez García-Conde

Centro de Investigación en Dinámica Celular

UAEM

Cuernavaca, Morelos 2019

Listado de abreviaturas y acrónimos

AdH	Adenovirus humanos
CD	Dicroísmo Circular
DLS	Dispersión Dinámica de la Luz
DNA	Ácido Desoxirribonucleico
dsDNA	Ácido Desoxirribonucleico de doble cadena
GrB	Granzima B humana
IDPs	Proteínas intrínsecamente desordenadas
IDRs	Regiones intrínsecamente desordenadas
mRNA	Ácido Ribonucleico mensajero
NMR	Resonancia Magnética Nuclear
ORF	Marco abierto de lectura
RNA	Ácido Ribonucleico
RNP	Región de unión a RNA
SAXS	Dispersión de rayos X de ángulo reducido
tRNA	Ácido Ribonucleico transferencia
VIH	Virus de la inmunodeficiencia humana

RESUMEN

Los adenovirus humanos (AdH) son agentes infecciosos ubicuos relacionados con enfermedades respiratorias, conjuntivitis y gastroenteritis; además se ha propuesto que algunos serotipos como AdH 36 son factores etiológicos de obesidad. Por otro lado, la facilidad con la que puede ser manipulado su genoma y la posibilidad de generar altos títulos virales han llevado a que los AdH se empleen como vectores para terapia génica y terapias anti-cáncer. Por lo anterior, conocer cada aspecto del ciclo de replicación viral, así como el funcionamiento de las proteínas virales es de suma importancia. Hasta la fecha se han identificado algunas de las actividades asociadas a la mayoría de las proteínas de los AdH, sin embargo, aún se desconocen los mecanismos moleculares implicados; en la mayoría de los casos, al menos en parte, debido a la falta de información estructural.

Las proteínas intrínsecamente desordenadas (IDPs) y las regiones intrínsecamente desordenadas (IDRs) se definen como proteínas o segmentos polipeptídicos que no poseen una única estructura tridimensional. Las IDPs e IDRs se relacionan con la capacidad de interactuar con diferentes blancos moleculares, lo que le puede explicar la multifuncionalidad de las proteínas con estas características. Se ha propuesto que el nivel de desorden en las proteínas está relacionado con las funciones de esta y con el tipo y número de moléculas con las que interactúa. El análisis de las IDPs e IDRs en el proteoma de AdH resulta útil para comprender la multifuncionalidad de las proteínas virales, así como su mecanismo de acción.

Hasta la fecha el rol de las IDPs/IDRs en el ciclo de replicación de AdH no ha sido evaluado. En el presente trabajo se evaluó la prevalencia de IDPs e IDRs en el proteoma de AdH. Se observó que 20 de las 34 proteínas codificadas en el genoma viral presentan IDRs y que 7 de esas 20 son clasificadas como IDPs. El análisis reveló que en las proteínas estructurales el porcentaje de residuos desordenados está relacionado con la proximidad al DNA en la cápside viral y con el número de interacciones que establece cada proteína. En el caso de las proteínas no estructurales se observó que el porcentaje de aminoácidos desordenados es mayor en las proteínas relacionadas con la replicación del DNA, la transcripción y el procesamiento de los RNA virales, mientras que las proteínas virales con actividades enzimáticas, así como las proteínas encargadas de regular el tráfico vesicular e inhibir la respuesta inmune de las células presentaron bajos niveles de IDRs.

Listado de Figuras

Figura 1. Diagrama del genoma de adenovirus humano.
Figura 2. Proteoma de adenovirus humano.
Figura 3. Diagrama de la partícula viral de adenovirus humano.
Figura 4. Gráfica carga neta media vs hidropatía media.
Figura 5. Predicción de desorden en las proteínas codificadas en E1A y E1B.
Figura 6. Predicción de desorden en las proteínas codificadas en E2.
Figura 7. Predicción de desorden en las proteínas codificadas en E3.
Figura 8. Predicción de desorden en las proteínas codificadas en E4.
Figura 9. Predicción de desorden en las proteínas codificadas en IX y IVa2.
Figura 10. Predicción de desorden en las proteínas codificadas en L1.
Figura 11. Predicción de desorden en las proteínas codificadas en L2.
Figura 12. Predicción de desorden en las proteínas codificadas en L3.
Figura 13. Predicción de desorden en las proteínas codificadas en L4.
Figura 14. Predicción de desorden en las proteínas codificadas en L5.
Figura 15. Distribución del desorden en la partícula viral de AdH.
Figura 16. Caracterización de las IDRs presentes en 22K.
Figura 17. Caracterización de las IDRs presentes en 33K.
Figura 18. Caracterización de las IDRs presentes en DPB.
Figura 19. Caracterización de las IDRs presentes en E1B 19K.
Figura 20. Caracterización de las IDRs presentes en pX.
Figura 21. Caracterización de las IDRs presentes en pVII.
Figura 22. Caracterización de las IDRs presentes en pVI.
Figura 23. Caracterización de las IDRs presentes en pV.
Figura 24. Caracterización de las IDRs presentes en pTP.
Figura 25. Caracterización de las IDRs presentes en AdPol.
Figura 26. Caracterización de las IDRs presentes en pIIIa.
Figura 27. Caracterización de las IDRs presentes en penton.
Figura 28. Caracterización de las IDRs presentes en hexon.
Figura 29. Caracterización de las IDRs presentes en E4orf6/7.
Figura 30. Caracterización de las IDRs presentes en E1A.
Figura 31. Caracterización de las IDRs presentes en 52/55K.
Figura 32. Caracterización de las IDRs presentes en 100K.
Figura 33. Caracterización de las IDRs presentes en E1B 55K.
Figura 34. Coevolución entre las regiones de E1B 55K.
Figura 35. Representación de los pares de aminoácidos con la mayor probabilidad de coevolución.
Figura 36. Caracterización de las IDRs presentes en IVa2.

Listado de Tablas

Tabla 1. Atributos empleados por las redes neurales de PONDR VLXT

Tabla 2. Atributos empleados por la red neural de PONDR XL1_XT
--

Índice

I.1 Adenovirus	1
<i>I.1.1 Clasificación</i>	1
<i>I.1.2 Genoma</i>	1
<i>I.1.3 Proteoma:</i>	2
<i>I.1.4 Estructura:</i>	3
I.2 Proteínas y Regiones Intrínsecamente Desordenadas	4
<i>I.2.1 Predictores de Desorden:</i>	5
I.3 IDPs e IDRs en virus	8
II. Hipótesis	9
III. Objetivos	9
IV. Materiales y Métodos	10
IV.1 Predicción de Desorden	10
IV.2 Caracterización de las regiones desordenadas	10
IV.3 Generación de modelos de las regiones plegadas	11
IV.3 Análisis Co-evolutivo	11
V. Resultados y Discusión	11
V.1 Desorden intrínseco en el proteoma de adenovirus	11
V.2 Distribución del desorden en la partícula viral	24
V.3 Caracterización de las IDRs en las proteínas de adenovirus humano	25
VI. Conclusiones	53
VII. Referencias Bibliográficas	54

I. Introducción

I.1 Adenovirus

I.1.1 Clasificación

La familia *Adenoviridae* se compone de virus desnudos icosaédricos de DNA de doble cadena lineal (dsDNA) que se replican en el núcleo celular [1]. A partir del análisis del genoma, los adenovirus se clasifican en cinco géneros: Mastadenovirus, aislados de mamíferos; Aviadenovirus, aislados de aves; Atadenovirus, aislados de reptiles, rumiantes y aves; Siadenovirus aislados a partir de anfibios, aves y reptiles; e Ichtyadenovirus que solo se ha aislado del Esturión Blanco [2].

Hasta la fecha se han identificado más de 90 serotipos de adenovirus que infectan humanos (AdH). Los AdH se clasifican en siete grupos (A-G) en función de la serología, la hemaglutinación, la oncogenicidad en roedores, la transformación de células en cultivo primario y el análisis de la secuencia de DNA [1].

Los AdH pueden provocar infecciones en las vías respiratorias, conjuntivitis y gastroenteritis [3]. Son responsables del 15 % de las enfermedades respiratorias agudas en humanos y hasta al 10 % de las enfermedades respiratorias en niños menores de cinco años [4]. En pacientes inmunosuprimidos, adenovirus puede causar infecciones fatales que resultan en hepatitis, neumonía y encefalitis [5]. Estos virus son transmitidos por contacto directo, por vía inhalatoria y por vía fecal oral. Los AdH infectan preferentemente las células epiteliales, afectando casi todas las mucosas [3].

I.1.2 Genoma

El genoma de los adenovirus humanos es una molécula de dsDNA de aproximadamente 36 kpb. Esta organizado en cinco unidades tempranas de transcripción (E1A, E1B, E2, E3 y E4), dos unidades intermedias de transcripción (IX y IVa2), y una unidad de transcripción tardía (ML) [1] (Fig. 1). Sin embargo, la clasificación en unidades transcripcionales tempranas o tardías puede no ser la mejor ya que algunas de las unidades transcripcionales tempranas se continúan expresando en la fase tardía y el promotor de ML se encuentra activo durante la fase temprana [6].

El genoma de adenovirus se representa de derecha a izquierda iniciando con la unidad transcripcional de E1A como se muestra en la Figura 1. Las unidades transcripcionales E1A, E1B, IX, ML y E3 están codificadas en la cadena que se lee de izquierda a derecha, mientras que E4, E2 y IVa2 están codificadas en la cadena complementaria.

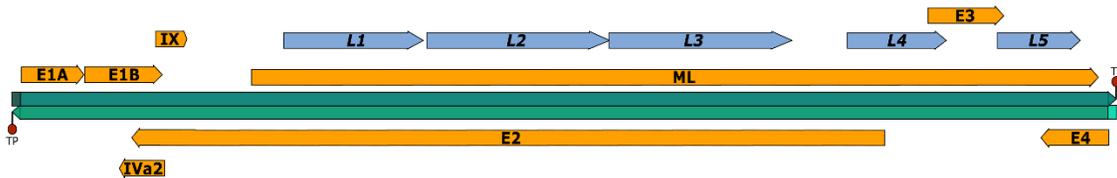


Figura 1. Diagrama del genoma de adenovirus humano. Representación del dsDNA (barras verdes) con la proteína TP (círculos rojos) unida a los extremos 5' de cada cadena. Las flechas de color naranja representan las unidades transcripcionales tempranas (E1A, E1B, E3, E4), intermedias (IX y IVa2) y tardías (ML). Las flechas de color azul claro representan familias de mRNAs virales tardíos (L1-L5) La presentación no se encuentra a escala.

Todos los mRNAs excepto IX son procesados por splicing. El transcrito principal de ML es un pre-mRNA que es procesado por selección de diferentes sitios de poliadenilación generando cinco familias de mRNA (L1, L2, L3, L4 y L5), los cuales posteriormente son procesados por splicing para generar diferentes productos[1, 6].

1.1.3 Proteoma:

La unidad transcripcional E1A codifica para dos proteínas: E1A-12S y E1A-13S, las cuales son nombradas en función del coeficiente de sedimentación de su respectivo mRNA. E1B codifica para al menos 5 producto proteicos los cuales son nombrados en función de su peso molecular (E1B 19K y E1B 55K) y del número de residuos (E1B-84R, E1B-93R y E1B-156R). En E2 se encuentran codificados las proteínas que participan directamente en la replicación del DNA viral y se nombran de acuerdo a su función (AdPol, pTP y DBP). E3 codifica para 7 proteínas las cuales se nombran por su peso molecular (E3 6.7K, E3 11.6K, E3 12.5K, E3 10.4K, E3 14.5K, E3 14.7K y E3 19K). Las proteínas codificadas en E4 se denominan en función del marco abierto de lectura (ORF) correspondiente al mRNA producido por splicing (E4orf1, E4orf3, E4orf4, E4orf6 y E4orf6/7). La unidad transcripcional ML codifica para 14 proteínas que se agrupan en 5 familias de mRNA: L1 (52/55k y pIIIa), L2 (penton, pVII, pV y pX), L3 (pVI, hexon y proteasa), L4 (100k, 22k, 33k y pVIII) y L5 (Fibra).

Por otra parte, las unidades transcripcionales IVa2 y IX codifican para un solo producto proteico respectivamente, los cuales llevan el mismo nombre que la unidad transcripcional[6] (Fig. 2).

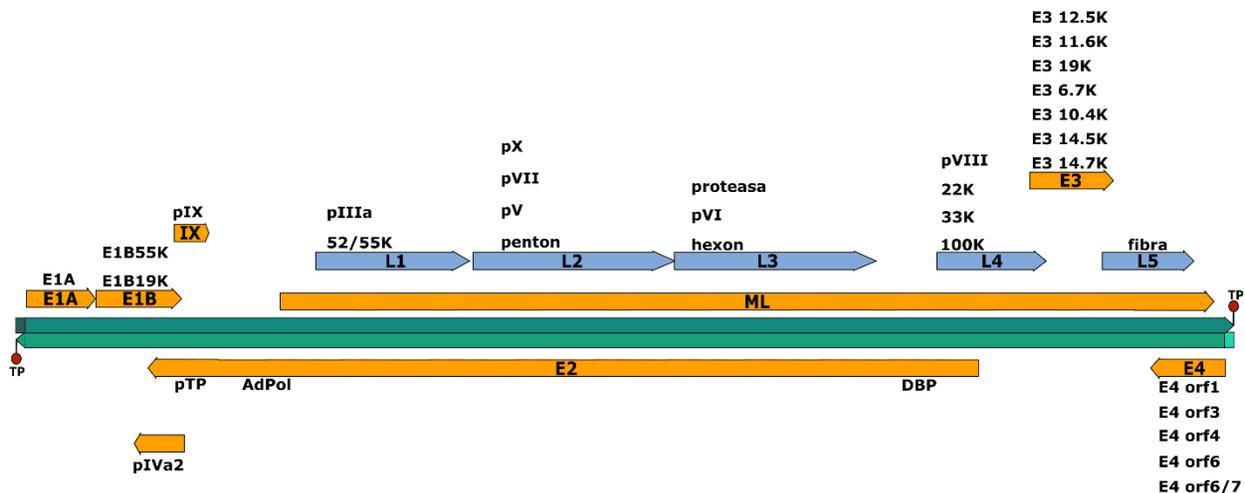


Figura 2. Proteoma de adenovirus humano. Representación del dsDNA (barras verdes) con la proteína TP (círculos rojos) unida a los extremos 5' de cada cadena. Las flechas de color naranja representan las unidades transcripcionales tempranas (E1A, E1B, E3, E4), intermedias (IX y IVa2) y tardías (ML). Las flechas de color azul claro representan familias de mRNAs virales tardíos (L1-L5). El nombre de las proteínas se muestra sobre la unidad transcripcional que las codifica. La representación no está a escala.

1.1.4 Estructura:

Como ya se mencionó, los adenovirus forman partículas desnudas con cápsides icosaédricas y DNA de doble cadena (dsDNA) de aproximadamente 36kbp. La cápside está compuesta por 11 proteínas y tiene un diámetro de aproximadamente 90nm (Fig. 3). La proteína más abundante es hexon, la cual está organizada en 240 trímeros. Penton se encuentra asociado a la proteína fibra, la cual sobresale en cada uno de los 12 vértices de la partícula icosaédrica. Las proteínas IX, VI, VIII y IIIa se encuentran asociadas a la superficie interna de la cápside. El núcleo de la partícula viral está formado por las proteínas V, VII, X y TP, las cuales a excepción de TP interaccionan con el DNA viral. La proteína TP se encuentra unida covalentemente a los extremos 5' del DNA. Además, se encuentran aproximadamente 10 copias de la proteasa viral la cual participa en la maduración y desensamblaje de la partícula viral [7].

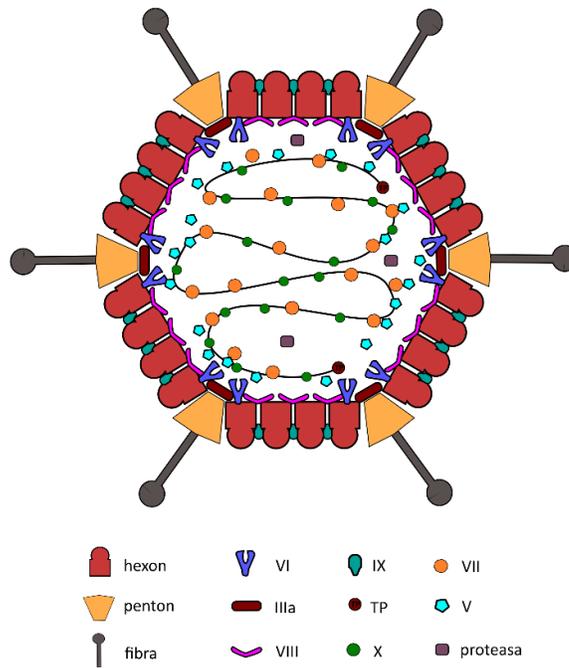


Figura 3. Diagrama de la partícula viral de adenovirus humano. El diagrama representa las proteínas estructurales en el virión. La representación no está a escala.

I.2 Proteínas y Regiones Intrínsecamente Desordenadas

Las proteínas intrínsecamente desordenadas (IDPs) y las regiones intrínsecamente desordenadas (IDRs) se caracterizan por no presentar una única estructura tridimensional en su estado nativo en condiciones fisiológicas, *in vitro* e *in vivo*. Por lo cual es necesario describirlas como un ensamble de distintas conformaciones [8].

La capacidad de las proteínas de plegarse o no en condiciones fisiológicas depende principalmente de la secuencia aminoacídica. Las IDPs en comparación con las proteínas plegadas tienden a presentar menos aminoácidos promotores de orden como Ile, Leu, Val, Trp, Try, Phe, Cys y Asn, y suelen estar enriquecidos en aminoácidos promotores del desorden como Ala, Arg, Gly, Gln, Ser, Glu, Lys y Pro[9]. Existen otros factores como hidropatía, flexibilidad, número de coordinación, volumen y carga neta, entre otros, que también deben tenerse en cuenta para determinar si una proteína es plegada o no.

Como resultado del análisis del rol potencial del desorden intrínseco en las proteínas se propone que: (i) las IDRs brindan las bases estructurales para que una proteína pueda interactuar con varios blancos moleculares; (ii) múltiples IDPs pueden unirse

a la misma proteína y (iii) las IDRs pueden actuar como linkers flexibles entre dominios funcionales facilitando la diversidad de unión [10].

La unión de las IDRs a sus blancos moleculares puede involucrar una transición de desorden a orden antes o durante la interacción. También se propone que puede existir interacción sin plegamiento que involucran complejos altamente dinámicos [10].

Una función ampliamente descrita para las IDPs e IDRs es que actúan como regiones de reconocimiento molecular. Esas interacciones moleculares involucran unión a DNA, mRNA, tRNA, así como a otras proteínas (quinasas, factores de transcripción, inhibidores de la traducción, polimerasas de RNA y de DNA)[11].

Las IDPs e IDRs pueden sufrir modificaciones postraduccionales las cuales pueden modificar su valencia y solubilidad e induce reorganización estructural, regulando así sus funciones. Entre las modificaciones reportadas se encuentran la fosforilación, acetilación, SUMOilación, acilación, metilación, glicosilación, ubiquitinación y ADP-ribosilación. Estudios computacionales han relevado que existe una fuerte correlación entre la predicción de desorden y los sitios donde ocurren las modificaciones postraduccionales. Se cree que el hecho de que las secuencias blanco sean desordenadas podría explicar que las enzimas que llevan a cabo las modificaciones postraduccionales son capaces de unirse a un amplio rango de proteínas blanco[12].

Además de las técnicas experimentales (NMR, CD, SAXS y DLS) estas proteínas pueden ser caracterizadas mediante análisis computacional a partir de la secuencia aminoacídica[10, 13].

1.2.1 Predictores de Desorden:

El primer predictor de desorden intrínseco fue desarrollado en 1997. Este se basaba en solo 67 regiones desordenadas (1,340 residuos) y en regiones ordenadas extraídas manualmente del PDB (16,542 residuos)[14]. En el año 2000 Uversky y colaboradores notaron que las proteínas desordenadas en toda su longitud pueden separarse de las proteínas ordenadas considerando su carga neta promedio e hidropatía[13]. Hasta la fecha se han desarrollado más de 20 predictores de

desorden intrínseco. Dentro de los predictores más empleados se encuentran: PONDR, DISOPRED, IUPred, GlobPlot, DISpro y DisEMBL[15].

1.2.1.1 PONDR:

El predictor de regiones desordenadas naturales (PONDR por sus siglas en inglés) es un conjunto de redes neuronales que usan atributos de secuencias tomadas en ventanas de 9 a 21 aminoácidos. Los atributos más comunes son hidropatía, complejidad de la secuencia y la presencia de determinados aminoácidos.

PONDR presenta varias extensiones cada una de ellas emplea redes neuronales específicas que han sido entrenadas de manera diferente. La primera letra de la extensión corresponde al método por el cual se caracterizaron las proteínas empleadas para el entrenamiento de la red neuronal: X (cristalografía de rayos X), N (resonancia magnética nuclear), C (dicróismo circular) y V (varios métodos). La segunda letra hace referencia a la longitud o a la localización de la región desordenada: S (de 8 a 10 residuos), M (entre 20 y 39 residuos), L (más de 40 residuos), N (extremo N-terminal), C (extremo C-terminal) y T (cualquier extremo).

PONDR VLXT:

Integra tres redes neuronales: VL1[14], XN y CN[16]. La red neuronal de VL1 se entrenó con regiones largas desordenadas caracterizadas por RMN y por cristalografía de rayos X. Las redes neuronales de XN y XC se entrenaron con regiones desordenadas en el N-terminal y C-terminal respectivamente caracterizadas por cristalografía de rayos X. A la combinación de XN con XC se le denomina XT. En la Tabla 1 se muestran los atributos empleados por cada una de estas redes. VL1 predice a partir del aminoácido 11 a cada uno de los extremos, mientras que XT predice a partir de la primera posición hasta el residuo 14 en cada extremo. Para los residuos 11 a 14 en cada extremo se emplea un promedio simple. Además se usan ventanas de 9 aminoácidos para suavizar la predicción a lo largo de toda la secuencia.

Tabla 1. Atributos empleados por las redes neurales de PONDR VLXT

Red	Atributos									
VL1	Número de coordinación	Carga Neta	WFY	W	Y	F	D	E	R	K
XN	Número de coordinación	V	VIYFW	M	N	H	D	PEVK		
XC	Número de coordinación	Hidropatía	VIYFW	M	T	H	PEVK	R		

PONDR XL1_XT:

Esta extensión de PONDR combina la anteriormente mencionada XT con XL1[14]. XL1 está optimizado para regiones de más de 39 residuos y se entrenó con regiones desordenadas caracterizadas por cristalografía de rayos X. XL1 predice a partir del residuo 15 de ambos extremos y se combina con XT forma parecida a VL1. XL1_XT emplea ventanas de 9 residuos para suavizar las predicciones. En la Tabla 2 se muestran los atributos empleados por XL1.

Tabla 2. Atributos empleados por la red neural de PONDR XL1_XT

Red	Atributos									
XL1	Flexibilidad	Hidropatía	C	W	Y	H	D	E	K	S

PONDR VL3-BA:

Esta red neuronal fue entrenada con regiones desordenadas de 152 residuos de longitud caracterizados por varios métodos. Esta red está basada en 20 atributos que incluyen 18 frecuencias de aminoácidos o secuencia de aminoácidos, flexibilidad media y complejidad de la secuencia. Para la predicción se emplean ventanas de 41 aminoácidos y los datos crudos de la predicción son promediados en un output de 31 aminoácidos para obtener la predicción final para cada posición. Los límites entre regiones ordenadas y desordenadas son corregidos utilizando un predictor de límites orden/desorden[17].

PONDR VSL2:

Esta extensión de PONDR combina en una sola red la optimización de regiones desordenadas largas (≥ 30) y cortas (< 30). La red se entrenó con 230 regiones largas desordenadas y 983 regiones cortas desordenadas caracterizadas por varios métodos. Los atributos empleados por esta red son: Complejidad de la secuencia, Relación carga neta/hidrofobicidad, Flexibilidad media, perfil PSI-BLAST y la frecuencia de determinados aminoácidos[17].

1.2.1.2 IUPred:

Este predictor se basa en los principios físicos del orden/desorden en las proteínas. Para ello estima el potencial de los polipéptidos de formar contactos que estabilicen la estructura mediante el uso de un potencial de interacción estadístico. El cálculo del potencial involucra una matriz de predicción de energía de 20x20, la cual esta

parametrizada para acercarse a la energía de los pares en las proteínas globulares. De modo que la probabilidad de un aminoácido de ser desordenado depende de su tipo químico y de su potencial de interacción con sus vecinos. Para cada aminoácido se considera la interacción con los vecinos que están de 2 a 100 residuos en ambas direcciones. La energía de emparejamiento con el resto de la secuencia se transforma a un score probabilístico de 0 a 1, siendo scores de más de 0.5 indicativos de desorden[18].

1.2.1.3 DISOPRED:

Inicialmente se genera un perfil de secuencia para cada proteína utilizando una búsqueda PSI-BLAST en una base de datos de secuencia filtrada. Se construye un vector de entrada para cada residuo a partir del perfil de secuencia en una ventana simétrica de 15 residuos a cada lado. Estos datos son analizados por una máquina de soporte de vectores lineales la cual devuelve el valor de desorden. Este predictor se entrenó con secuencias no redundantes caracterizadas por cristalografía de rayos X[19].

1.3 IDPs e IDRs en virus

Con el descubrimiento de las IDRs e IDPs surgió un nuevo paradigma para el estudio de los virus. El análisis de los proteomas virales ha revelado que el porcentaje de residuos desordenados varía de un 7.3% a un 77.3%. Siendo Coronavirus humano NL63 el de menor porcentaje y el virus del Carcinoma Aviar el mayor porcentaje. Los virus con cinco proteínas o menos generalmente presentan más del 50% de residuos desordenados en su proteoma. Cuando el tamaño del proteoma aumenta, la fracción de residuos desordenados se encuentra entre el 20 y el 40% [20].

El desorden intrínseco en las proteínas virales se ha asociado a la capacidad de realizar más de una función durante el ciclo de replicación viral, así como a la capacidad de estas proteínas de unir diferentes tipos de moléculas como DNA, RNA y/o proteínas. Al igual que las IDRs en células eucariontes, las IDPs virales han sido asociadas con la replicación del DNA, regulación de la transcripción y regulación del ciclo celular [21].

Por otra parte, se ha observado que también puede existir desorden en las partículas virales. Por ejemplo, el N_{Tail} (125 aminoácidos en el extremo C-terminal) de la proteína de la nucleocápside de sarampión se encuentra desordenado, siendo dinámico e intercambiando entre un estado completamente extendido o formando una hélice alfa de 18 residuos. Sobre la base de esa información se establece un marco estructural que permite explicar la dualidad de función del N_{Tail} [22]. Asimismo, la matriz de VIH se predijo como altamente desordenada y en la partícula viral de influenza A se observó que el porcentaje de aminoácidos desordenados aumenta a medida que las proteínas se localizan más cerca del núcleo de RNA [23].

Hasta la fecha no existe un análisis del desorden intrínseco en el proteoma de Adenovirus Humano. Al mismo tiempo, se desconocen los mecanismos de acción de muchas proteínas virales, lo cual se podría deber, en parte, a la ausencia de información estructural de las regiones plegadas y desplegadas. Analizar la distribución y rol de las IDPs e IDRs de AdH proporcionará información relevante para comprender mejor la multifuncionalidad de las proteínas virales, sentando las bases para en un futuro lograr la descripción completa de sus mecanismos de acción.

II. Hipótesis

El grado de desorden en las proteínas de Adenovirus Humano está relacionado con su función durante el ciclo de replicación viral.

III. Objetivos

Objetivo General:

- Analizar la distribución de las IDRs en el proteoma de Adenovirus Humano.

Objetivos Específicos:

1. Identificar las regiones intrínsecamente desordenadas en el proteoma de adenovirus humano.
2. Caracterizar las regiones intrínsecamente desordenadas presentes en el proteoma de adenovirus humano.

3. Determinar el grado de conservación de las regiones desordenadas entre los diferentes serotipos de adenovirus humano.
4. Identificar actividades y funciones putativas asociadas a las regiones intrínsecamente desordenadas.

IV. Materiales y Métodos

IV.1 Predicción de Desorden

Para la predicción de desorden se emplearon las secuencias de las proteínas codificadas en el genoma reportado para AdH5 (Secuencia de referencia en NCBI: AC_000008.1) y los predictores de desorden PONDR® pool [PONDR® VLXT, PONDR® VL1_XT, PONDR® VL3-BA, PONDR® VSL2] (www.pondr.com); IUPRED (<https://iupred2a.elte.hu/>) and DISOPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>). el promedio de la predicción de los seis predictores se empleó para identificar las regiones intrínsecamente desordenadas. Se identificaron como regiones desordenadas relevantes para nuestro análisis segmentos en los cuales 30 o más aminoácidos presentaran valores de predicción de desorden promedio superior a 0.5.

IV.2 Caracterización de las regiones desordenadas

Las regiones predichas como intrínsecamente desordenadas se caracterizaron empleando el servidor CIDER (<http://pappulab.wustl.edu/CIDER/>). CIDER calcula parámetros específicos de las secuencias como la distribución de las cargas opuestas, frecuencia de residuos cargados, carga neta por residuos, hidropatía, proporción de residuos promotores de desorden y genera un diagrama de estado de acuerdo a las cualidades estructurales de la secuencia. Las regiones desordenadas son clasificadas en cinco regiones. *R1*: Secuencia que típicamente son compactas. *R2*: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. *R3*: Secuencias cuya estructura va a depender de la distribución de cargas, si las cargas de la misma polaridad se localizan juntas y no se alternan con las de polaridad opuesta las secuencias se presentarán como estructuras extendidas, de lo contrario serán capaces de plegarse bajo determinadas condiciones al igual que las de tipo

R2. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica.

La predicción de estructura secundaria a partir de la secuencia aminoacídica se realizó con el servidor Psipred (http://bioinf.cs.ucl.ac.uk/psipred_new/).

IV.3 Generación de modelos de las regiones plegadas

Para la generación de modelos, se enviaron las secuencias Fasta de las proteínas y de las regiones predichas como plegadas al servidor I-Tasser (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>). De los modelos generados por I-Tasser, los modelos con más alto C-score fueron elegidos.

IV.3 Análisis Co-evolutivo

Con el objetivo de buscar regiones co-evolutivas con significado estructural potencial empleamos el software GREMLIN [24]. Se empleó el algoritmo HHblits para generar el alineamiento múltiple de secuencia con un valor de $E=1e^{-4}$ y ocho iteraciones. El alineamiento se filtró eliminando las secuencias que presentaban menos de 50 % de identidad y se eliminaron las posiciones que presentaban más del 50 % de gaps. Se realizó un análisis para toda la secuencia de la proteína y un análisis para la región central de la proteína (del aminoácido 147 al 496).

V. Resultados y Discusión

V.1 Desorden intrínseco en el proteoma de adenovirus

Como una primera aproximación para identificar cuáles son las proteínas intrínsecamente desordenadas en el proteoma de AdH se realizó una gráfica de carga neta media vs hidropatía media. Este resultado surgiere que las proteínas E1A, 22K, 33K, pV, pVII y pX son IDPs (Fig. 4). Esta clasificación se basa en el hecho de que las IDPs se suelen agrupar y separan de las proteínas plegadas en este tipo de gráficas [13].

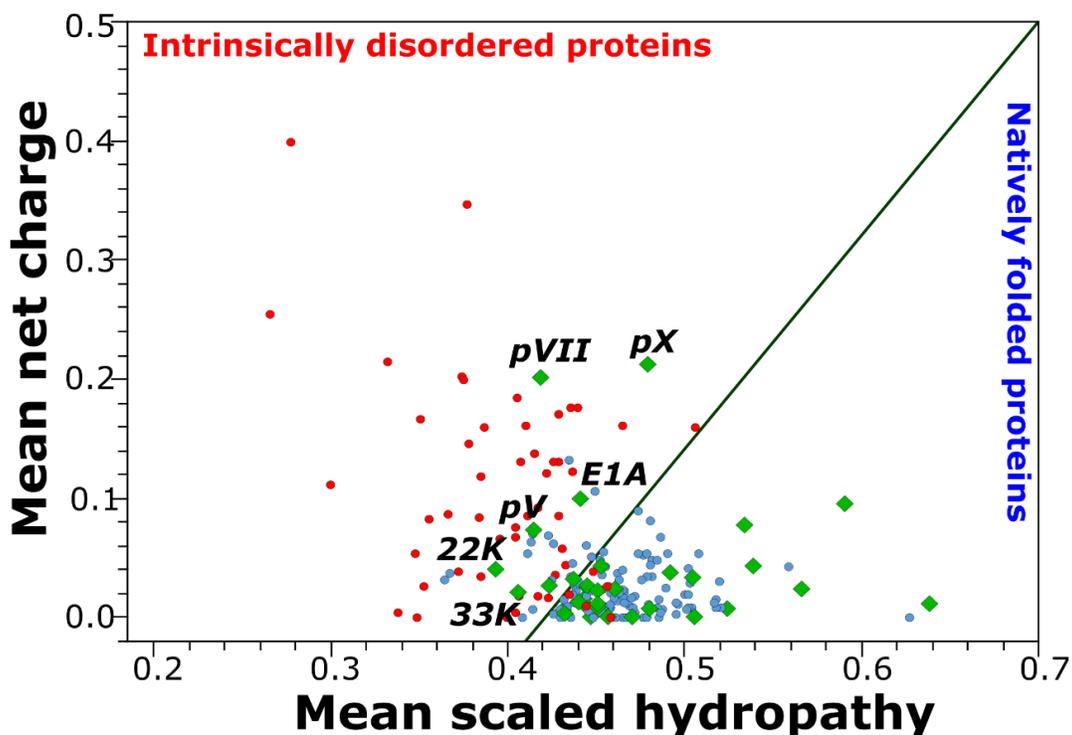


Figura 4. Gráfica carga neta media vs hidropatía media. Los rombos verdes representan las proteínas de adenovirus, se indica el nombre de las proteínas de adenovirus clasificadas como IDPs. La línea representa el límite entre proteínas desordenadas y plegadas. Los círculos azules representan un set de 105 proteínas plegadas y los círculos rojos representan un set de 57 proteínas intrínsecamente desordenadas. Los datos se obtuvieron del programa PONDR.

Para realizar una identificación de las regiones intrínsecamente desordenadas en el proteoma de adenovirus humano se realizó la predicción de desorden intrínseco con el empleo de seis predictores de desorden, el promedio de la predicción de los seis predictores se empleó para identificar las regiones intrínsecamente desordenadas.

La Figura 5 muestra el resultado de la predicción de desorden obtenido para las proteínas E1A, E1B 55K y E1B 19K. La proteína E1A muestra un elevado porcentaje de regiones desordenadas, presentando solo la región 147-183 predicha como ordenada. Este resultado concuerda con lo obtenido en la gráfica de carga neta media vs hidropatía media. Así mismo, existen evidencias experimentales de que la región A53-G91 de E1A es intrínsecamente desordenada. En estos estudios se determinó mediante NMR que esta región se encuentra desplegada y que la interacción con CBP induce la formación de una hélice alfa (E59-I65) y dos hélices 3/10 (L72-V74 y L80-T82) [25]. Durante el ciclo de replicación viral E1A interactúa

con varios blancos moleculares, en ese sentido el hecho de sea una IDP pueden brindarle a la proteína la capacidad de plegarse para interactuar con dichos blancos, así como la flexibilidad para interactuar con más de un blanco la vez.

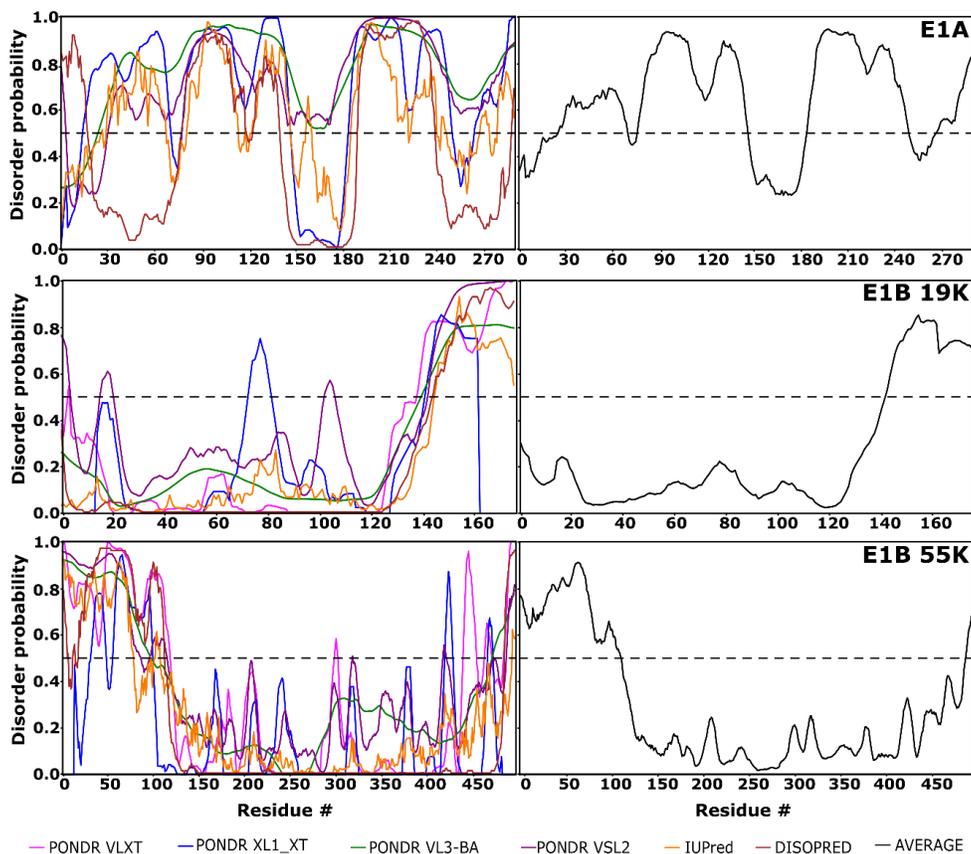


Figura 5. Predicción de desorden en las proteínas codificadas en E1A y E1B. A la izquierda se muestra el resultado de los predictores PONDR VLXT (línea rosa), PONDR VL1_XT (línea azul), PONDR VL3-BA (línea verde), PONDR VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

La predicción de desorden para E1B 19K surge que sus últimos 33 aminoácidos se encuentran desordenados (Fig. 5). E1B 19K está involucrada en la inhibición de la vía apoptótica, al funcionar como una proteína antiapoptótica tipo BCL-2[26]. E1B 19K se une a proteínas proapoptóticas Bax y Bak, para prevenir su oligomerización y por consiguiente la formación de poros en la membrana de la mitocondria[27, 28]. Hasta la fecha no existen reportes de la función de la región desordenada en las actividades de E1B 19K.

La región amino terminal de E1B 55K se predice como una IDR de 113 aminoácidos (Fig. 5). E1B 55K está asociada a varias funciones durante el ciclo de replicación viral: formación de un complejo E3 ubiquitina ligasa [29-31], induce la degradación de la proteína celular Daxx [32], inhibe la activación de la transcripción dependiente de p53 [33], represión de genes inducidos por Interferón [34, 35], la proteína es necesaria para la acumulación de mRNA en el citoplasma y afecta la morfología, organización y composición de los centros de replicación. Para llevar a cabo cada una de esas actividades se propone que la proteína se asocia tanto con proteínas celulares como proteínas virales [36, 37]. Sin embargo, no se conocen los mecanismos moleculares implicados. La presencia de una región desordenada tan extensa le permitiría a la proteína interactuar con diferentes blancos moleculares. Además, se conoce que la proteína es modificada por SUMOilación en esta región, lo cual podría regular las propiedades de la proteína [38]. La SUMOilación de la proteína es necesaria para la localización nuclear, la inhibición de la transactivación de p53 y la degradación de Daxx. Esto puede deberse a que la modificación postraducciona l permite la regulación de las propiedades de la región intrínsecamente desordenada y por ende su interacción con diferentes blancos moleculares. Se demostró por difracción circular y resonancia magnética nuclear que E1B-93R, la cual tiene en común los primeros 79 aminoácidos con E1B 55K es intrínsecamente desordenada. Confirmando así que la región amino terminal de E1B 55k también se presenta como una región intrínsecamente desordenada[39].

Las proteínas codificadas en E2 son las encargadas de llevar a cabo la replicación del genoma viral. Las tres proteínas codificadas en esta unidad transcripcional presentaron al menos una región IDR (Fig. 6). AdPol es la proteína que presenta la actividad DNA polimerasa DNA dependiente. Esta proteína presenta una sola región predicha como IDR, la cual se encuentra en el extremo amino terminal. Por lo general, las enzimas son proteínas plegadas, lo cual se cumple para la mayoría de secuencia de AdPol. La presencia de una región desordenada pudiera deberse a que esta es la región con la cual AdPol interactúa con el DNA, pues se ha reportado que las proteínas de unión a DNA pueden presentar regiones desordenadas que les permiten unirse a los ácidos nucleicos [11].

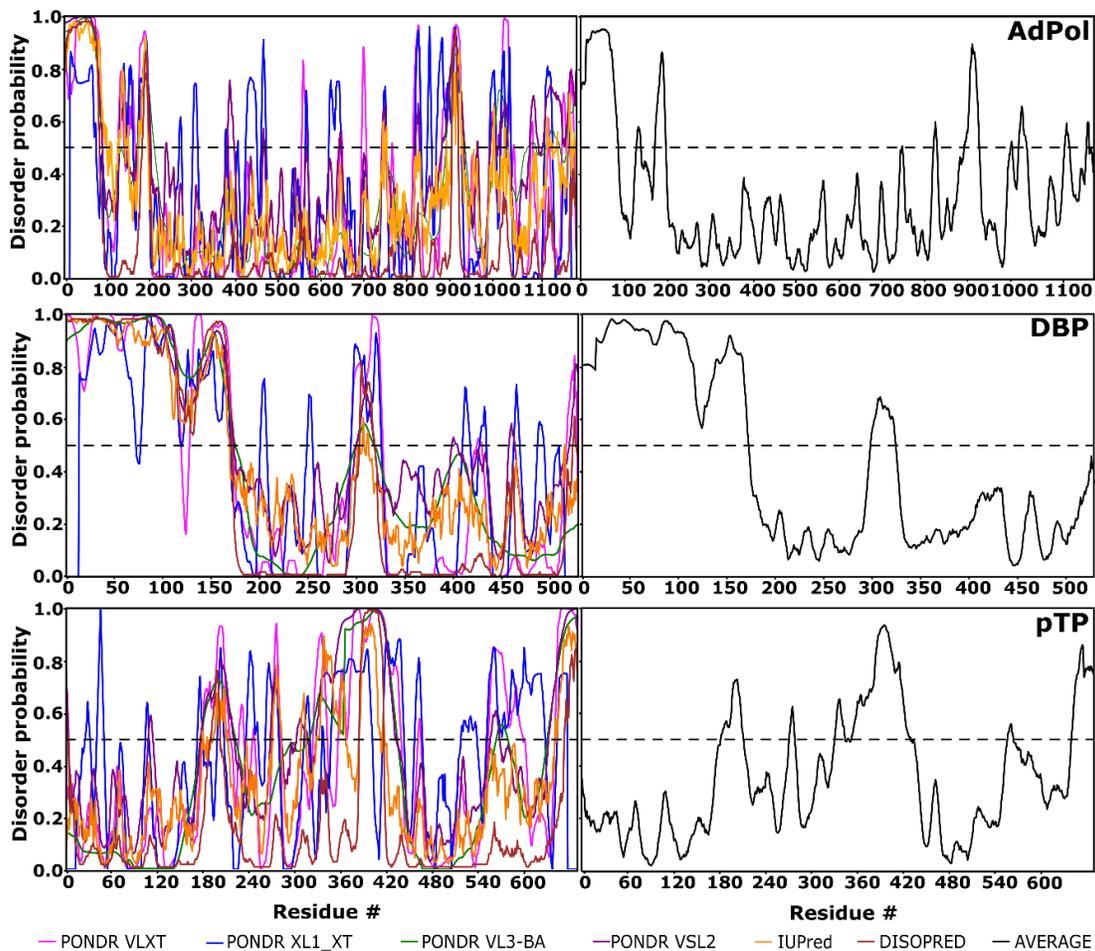


Figura 6. Predicción de desorden en las proteínas codificadas en E2. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL VL3-BA (línea verde), PONDRL VSL2 (línea púrpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

DBP es la proteína que se une a las cadenas simples de DNA viral durante la replicación. DBP presenta dos regiones intrínsecamente desordenadas una en el extremo amino de 173 y otra al centro de la proteína de 30 residuos. La ausencia de estructura de la región desordenada al centro de la proteína la ha sido confirmada por cristalografía de rayos X, ya que esta región presenta pérdida de densidad electrónica, lo cual es indicativo de que esta región se encontraba en diferentes conformaciones en el cristal. El extremo amino terminal de esta proteína no se ha podido cristalizar lo que también puede ser indicativo de que es una región intrínsecamente desordenada [40]. Como se mencionó anteriormente DBP es una

proteína de unión a DNA de simple cadena y el elevado grado de desorden en la misma puede estar relacionado con esta función.

La proteína pTP actúa como primer durante la replicación del genoma viral, pues esta proporciona el extremo 3'-OH necesario para la elongación de la cadena de DNA por AdPol. pTP se encuentra unida a los extremos 5' de cada una de las cadenas del genoma viral. pTP presenta tres regiones desordenadas (183-214, 332-436 y 641-671).

Las proteínas codificadas en E3 participan en la inhibición del sistema inmune y regulan el tráfico vesicular durante la infección. Estas proteínas presentan un bajo porcentaje de aminoácidos desordenados y carecen de regiones intrínsecamente desordenadas (Fig. 7).

Las proteínas codificadas en E4 llevan a cabo diferentes funciones durante el ciclo de replicación viral. Algunas proteínas participan en la regulación del metabolismo celular [41], mientras que otras participan en la degradación de proteínas celulares [42, 43]. Al analizar el desorden estructural en las proteínas de E4 se observó que solo E4 orf6 y E4 orf6/7 presentan IDRs. Cabe señalar que la región desordenada se corresponde a la secuencia que comparten estas proteínas (Fig. 8).

Las proteínas pIX y pIVa2 están codificadas en unidades de transcripción intermedias del mismo nombre. En la Figura 9 se observa la predicción de desorden para estas proteínas. pIX se predice como una proteína ordenada que carecerá de IDRs, mientras que pIVa2 también se predice como una proteína mayoritariamente plegada con una región de 80 residuos desordenados en su extremo amino terminal. pIVa2 es una proteína que se une como dímero al genoma viral y que regula la expresión del promotor de ML [44], además pIVa2 presenta actividad ATPasa *in vitro* [45] y es requerida para la encapsidación del genoma viral [46]. Esta proteína se une a las secuencias de empaquetamiento del DNA viral. Se ha determinado que los residuos 50-100 de la proteína son necesarios para la unión a DNA [47], lo cual concuerda con el hecho de que la mayoría de estos residuos son desordenados y que las proteínas de unión a DNA suelen presentar regiones desplegadas en sus dominios de interacción.

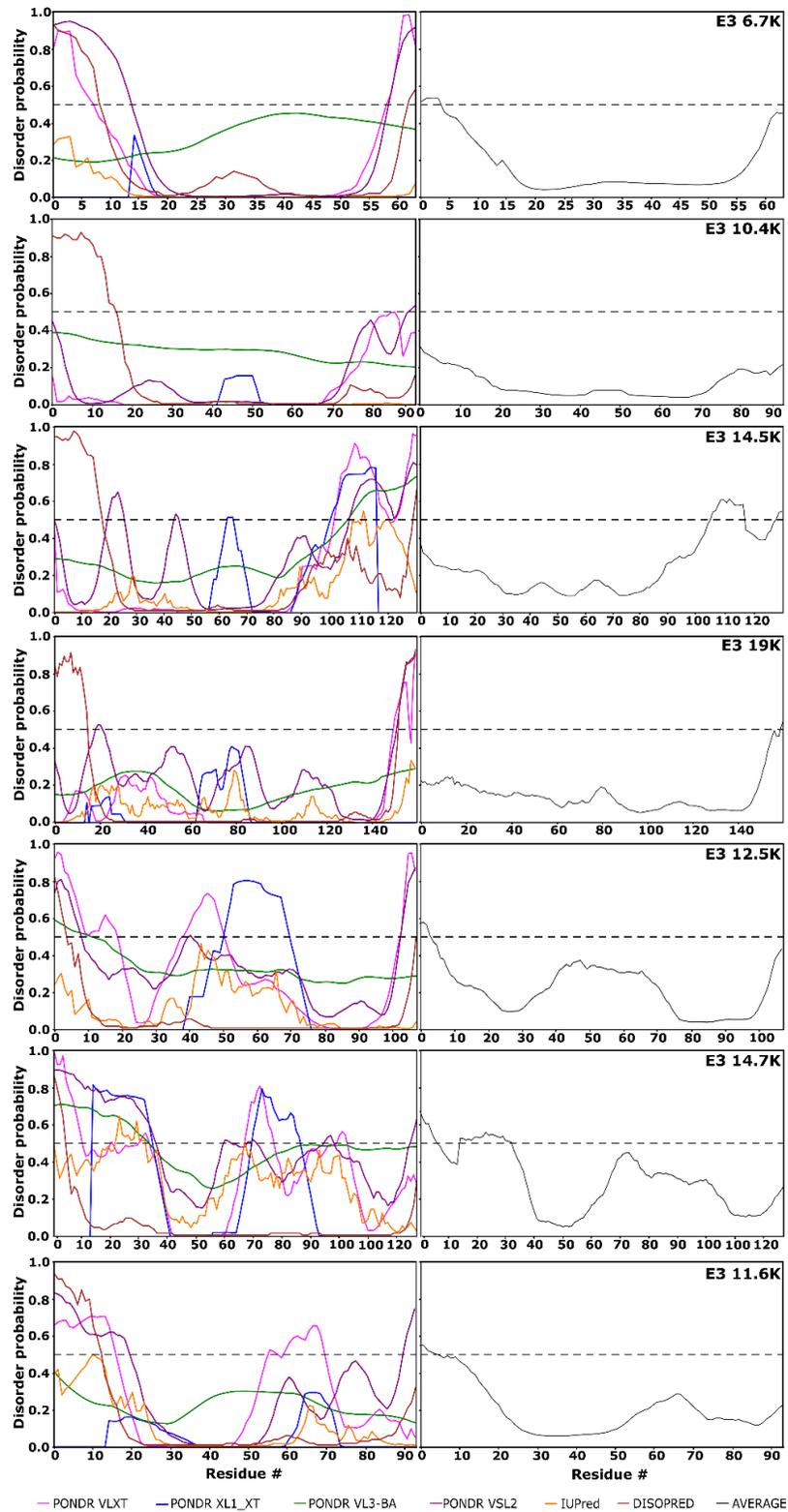


Figura 7. Predicción de desorden en las proteínas codificadas en E3. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL VL3-BA (línea verde), PONDRL VSL2 (línea púrpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

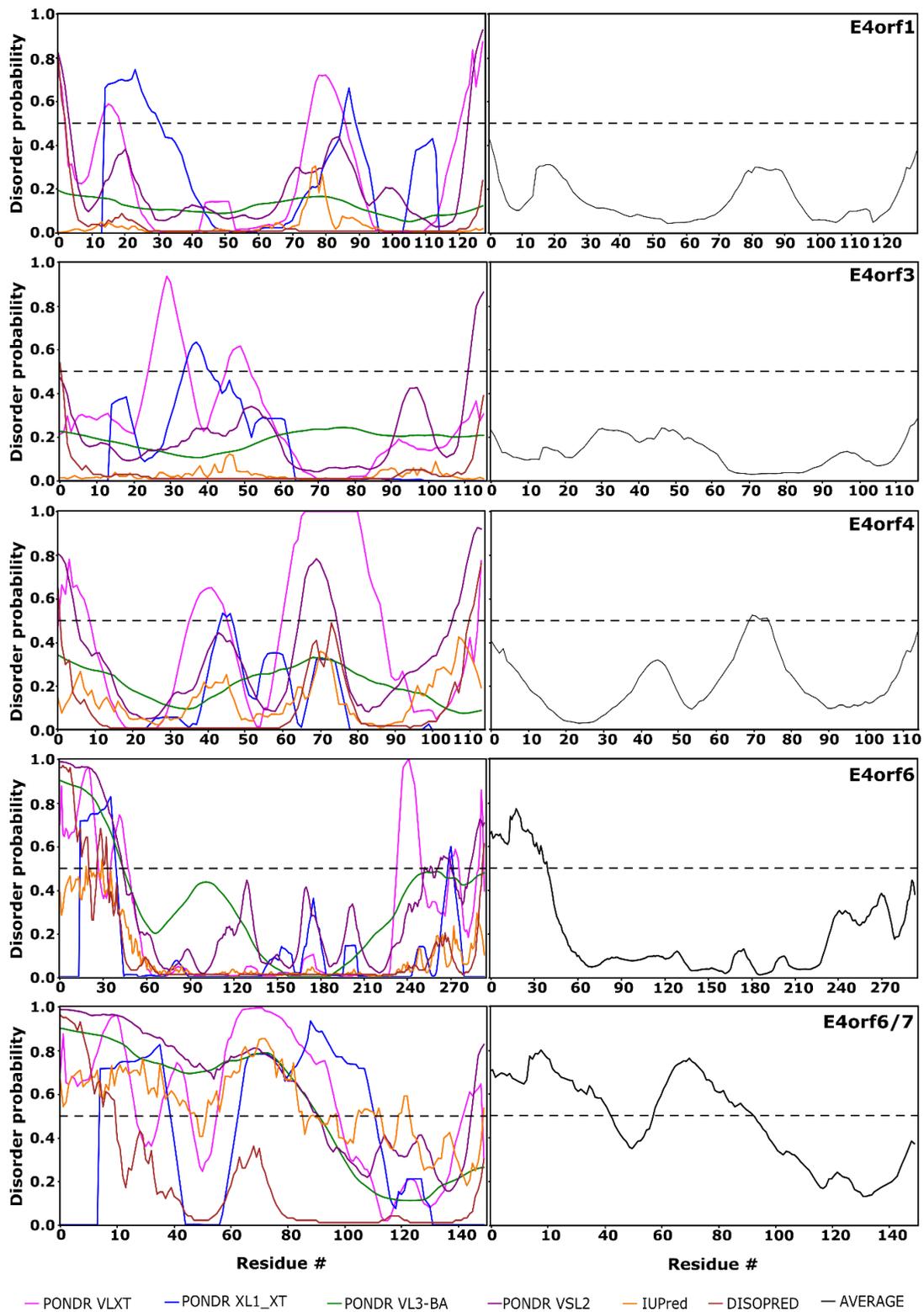


Figura 8. Predicción de desorden en las proteínas codificadas en E4. A la izquierda se muestra el resultado de los predictores PONDR VLXT (línea rosa), PONDR VL1_XT (línea azul), PONDR VL3-BA (línea verde), PONDR VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

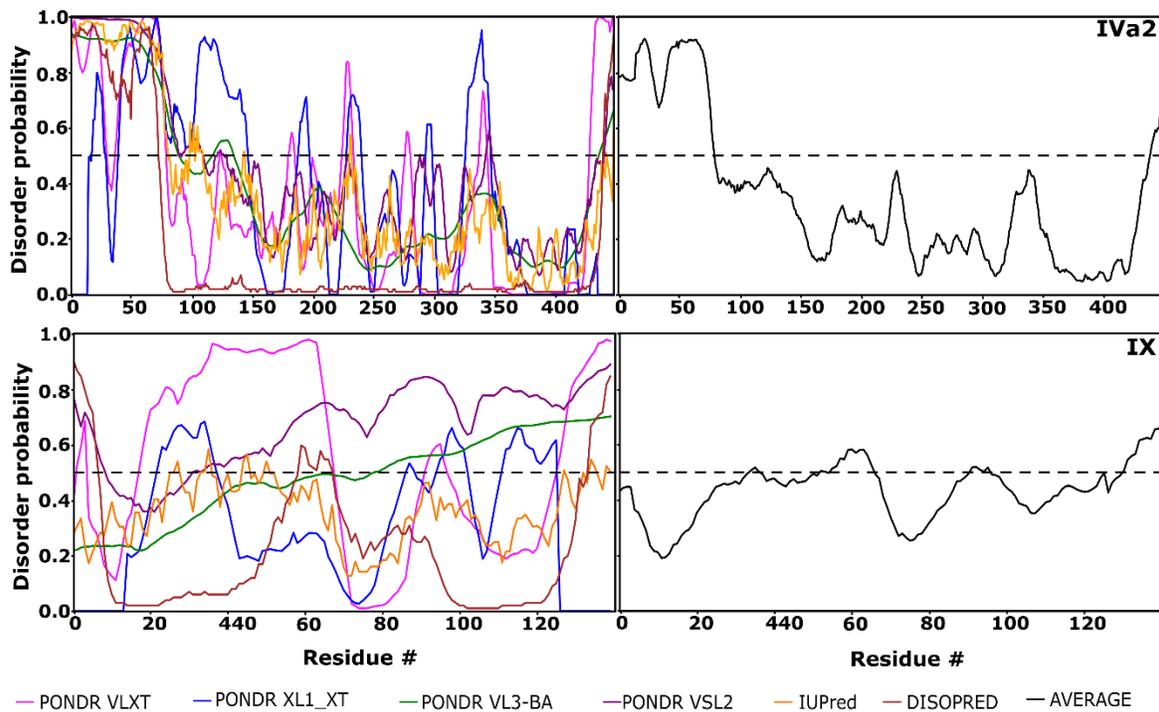


Figura 9. Predicción de desorden en las proteínas codificadas en IX y IVa2. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL VL3-BA (línea verde), PONDRL VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

Las proteínas codificadas en el MLP se producen durante la fase tardía del ciclo de replicación viral, aunque puede existir expresión de algunas de las proteínas en la fase temprana [6]. Estas se agrupan en cinco familias (L1, L2, L3, L4 y L5) de acuerdo al sitio de poliadenilación del RNA.

A la familia L1 pertenecen las proteínas 52/55K y pIIIa. 52/55K participa en la encapsidación serotipo específica de DNA, dicha especificidad es otorgada por la región amino terminal de la proteína [48], la cual se predice como una IDR (Fig. 10). Por otra parte, pIIIa es una proteína estructural que se encuentra asociada a la superficie interna de la cápside.

En familia L2 se encuentran cuatro proteínas, todas ellas presentan regiones intrínsecamente desordenadas (Fig. 11). pV, pVII y pX forman el núcleo de la partícula viral y se asocian directamente con el genoma viral. Estas tres proteínas se clasifican como IDPs (Fig. 4) lo cual concuerda con la predicción de desorden

obtenida para estas proteínas, pues en los tres casos las regiones desordenadas ocupan más de un 60% de la proteína. Llegando a un 80% en el caso de pV. Por otra parte, la proteína penton también es una proteína estructural que se encuentra asociada a la superficie externa de la cápside y que sirve de pase para la proteína fibra. Penton presenta una IDR de 96 aminoácidos localizada al centro de la proteína (A294-R389). La estructura de penton se determinó por cristalografía de rayos X en la cual se observó pérdida de densidad electrónica en la región correspondiente a la IDR [49], confirmando así que este segmento puede adoptar más de una conformación.

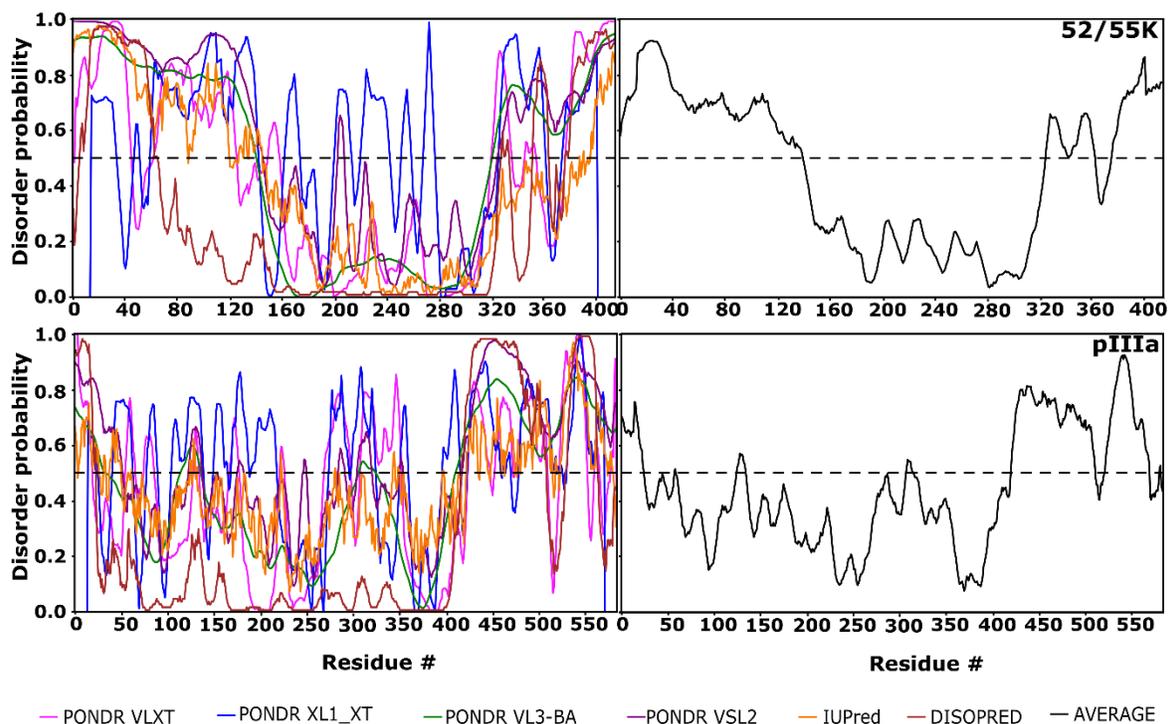


Figura 10. Predicción de desorden en las proteínas codificadas en L1. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL VL3-BA (línea verde), PONDRL VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

Las proteínas tardías pVI, hexon y proteasa pertenecen a la familia L3 de la unidad transcripcional ML. Hexon y pVI son proteínas estructurales, mientras que proteasa participa en el proceso de maduración de la partícula viral y aunque pueden existir hasta diez copias en las partículas virales no se considera una proteína estructural. Proteasa se predice como una proteína plegada (Fig. 12), lo cual concuerda el

hecho de que, por lo general, las enzimas son proteínas con una sola estructura definida. pVI es una proteína estructural que se encuentra asociada a la superficie interna de la cápside. El 55% de pVI está ocupado por una región intrínsecamente desordenada (Fig. 12). Hexon es la proteína más abundante en la partícula viral y se localiza en la superficie externa de la misma. Según la predicción de desorden hexon presenta una IDR de 33 aminoácidos (W135-H167), la cual coincide con una de las regiones hipervariables de esta proteína viral. La estructura de hexon se determinó por cristalografía de rayos X, en la estructura calculada se confirma el hecho de que la proteína esta plegada y se observó ausencia de estructura en la región 137-165.

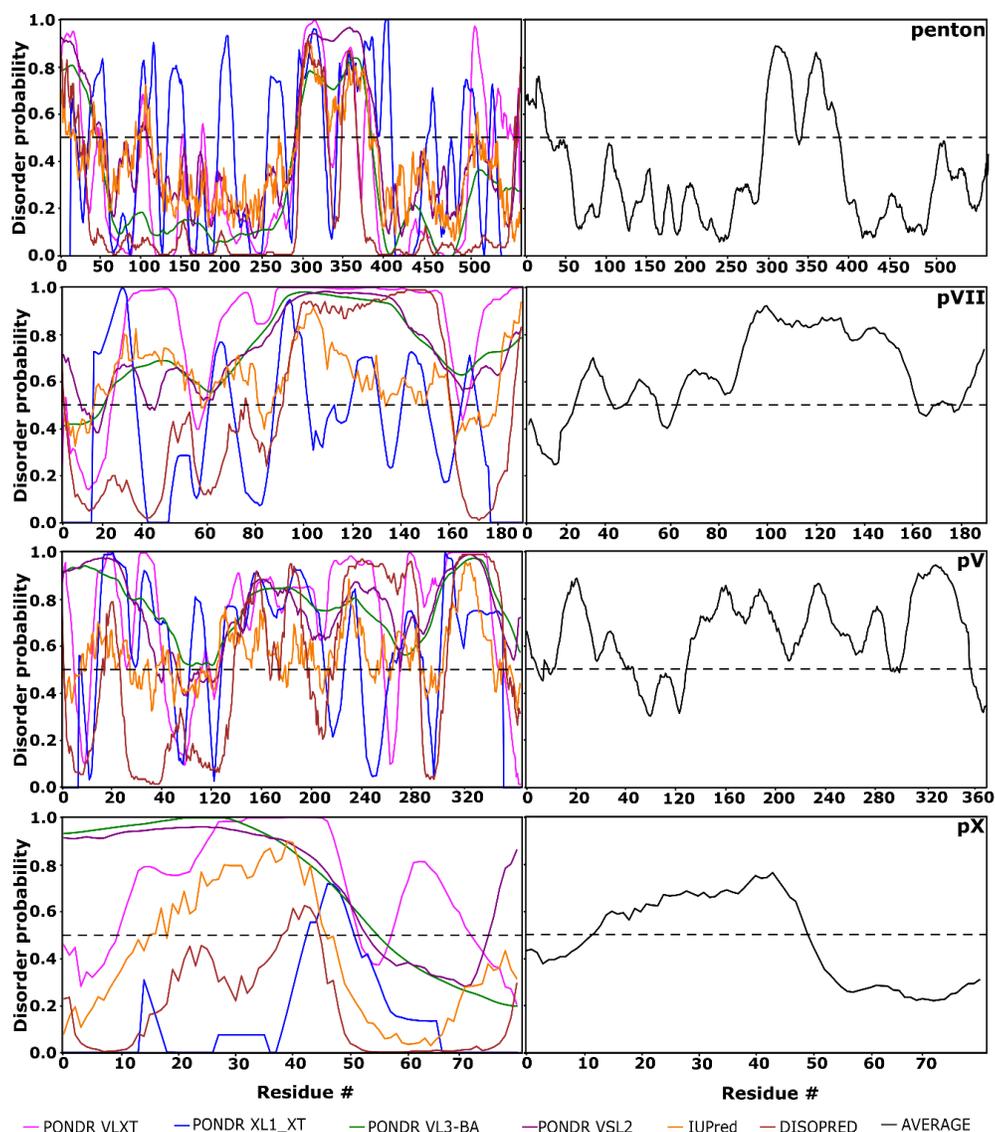


Figura 11. Predicción de desorden en las proteínas codificadas en L2. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL

VL3-BA (línea verde), PONDR VSL2 (línea púrpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

Las proteínas: 100K, 22K, 33K y pVIII pertenecen a la familia L4. La proteína pVIII, a diferencia de las otras proteínas de esta familia, es una proteína estructural que se localiza en la superficie interna de la cápside. pVIII también es la única proteína de esta familia que no presenta IDRs (Fig. 13). 22K y 33K tienen en común los primeros 105 aminoácidos los cuales se predicen como desordenados (Fig. 13). Se ha demostrado que estas proteínas llevan a cabo diferentes funciones durante el ciclo de replicación viral, tanto 22K como 33K se unen a las secuencias de encapsidación en el DNA viral [50]. Por su parte 100K es una proteína de unión a RNA que participa en la expresión selectiva de los mRNAs virales durante la fase tardía de la infección [51]. Esta proteína presenta cuatro regiones intrínsecamente desordenadas distribuidas a lo largo de su secuencia (Fig. 13).

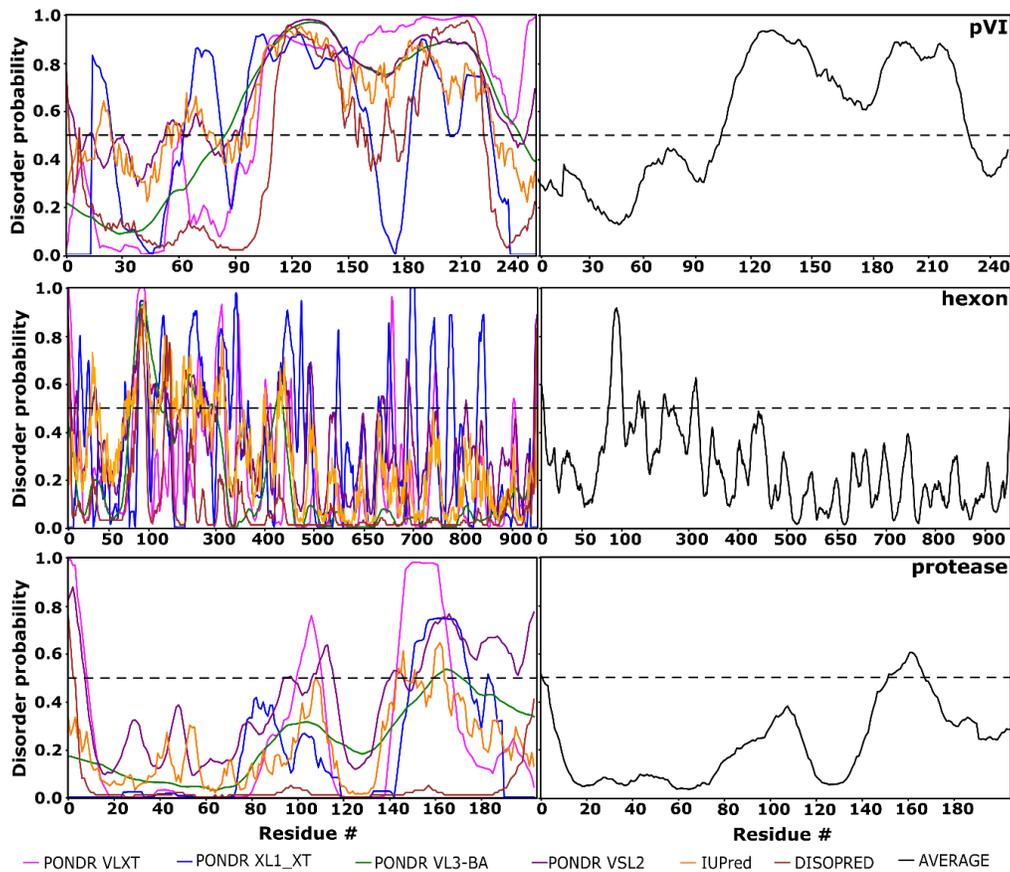


Figura 12. Predicción de desorden en las proteínas codificadas en L3. A la izquierda se muestra el resultado de los predictores PONDRLVXT (línea rosa), PONDRLV1_XT (línea azul), PONDRLV3-BA (línea verde), PONDRLVSL2 (línea púrpura), IUPred (línea naranja) y DISOPRED (línea roja). A la derecha se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

VL3-BA (línea verde), PONDR VSL2 (línea púrpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

La proteína fibra es la única que pertenece a la familia L5 del ML. Fibra es una proteína estructural que se encuentra en la capa externa de la partícula viral, asociada a penton en cada uno de los 12 vértices de la partícula icosaédrica. La estructura de la región carboxilo terminal de fibra se determinó por cristalografía de rayos X, siendo esta región (396-580) plegada, lo cual coincide con la predicción de desorden para esta región de la proteína (Fig. 14).

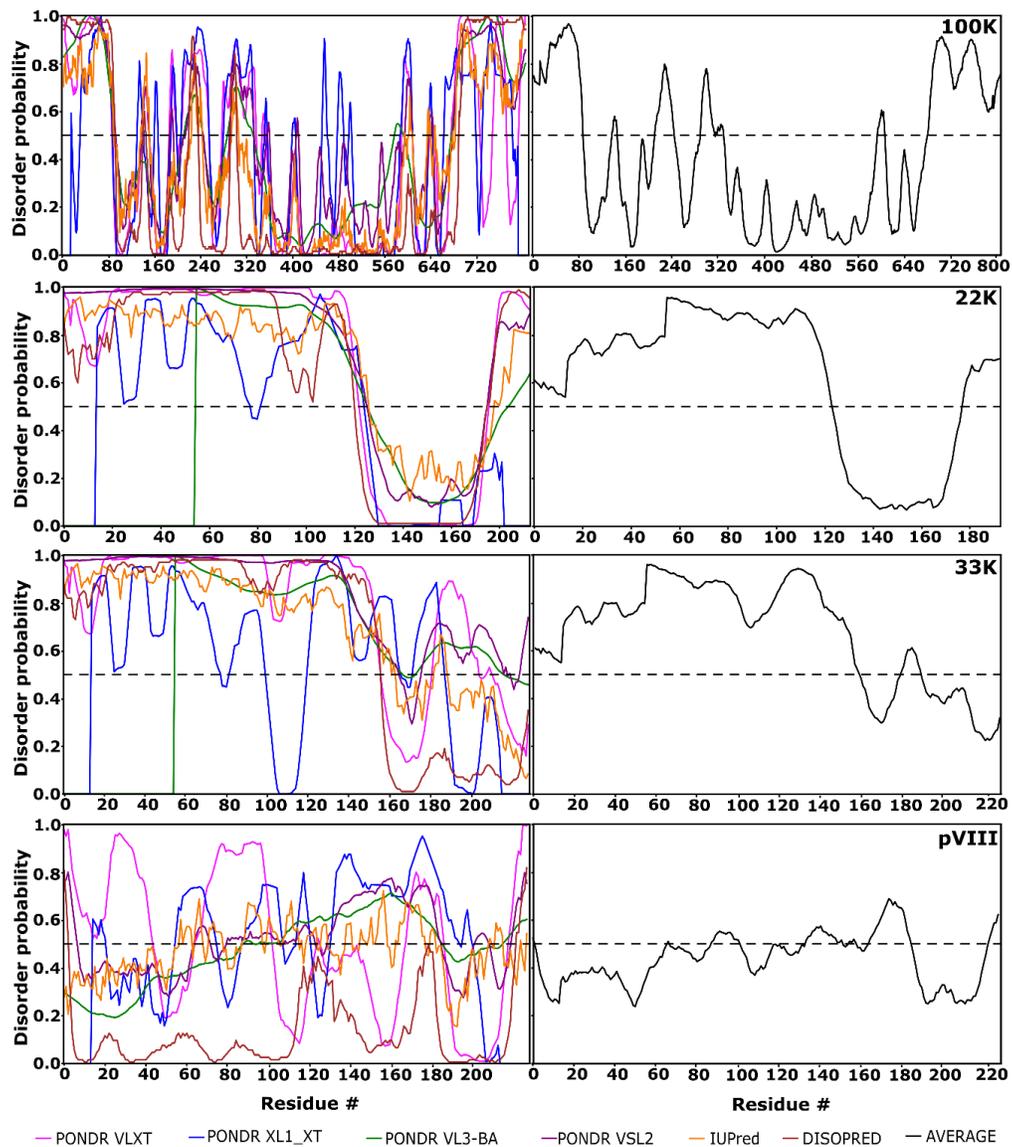


Figura 13. Predicción de desorden en las proteínas codificadas en L4. A la izquierda se muestra el resultado de los predictores PONDR VLXT (línea rosa), PONDR VL1_XT (línea azul), PONDR

VL3-BA (línea verde), PONDR VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

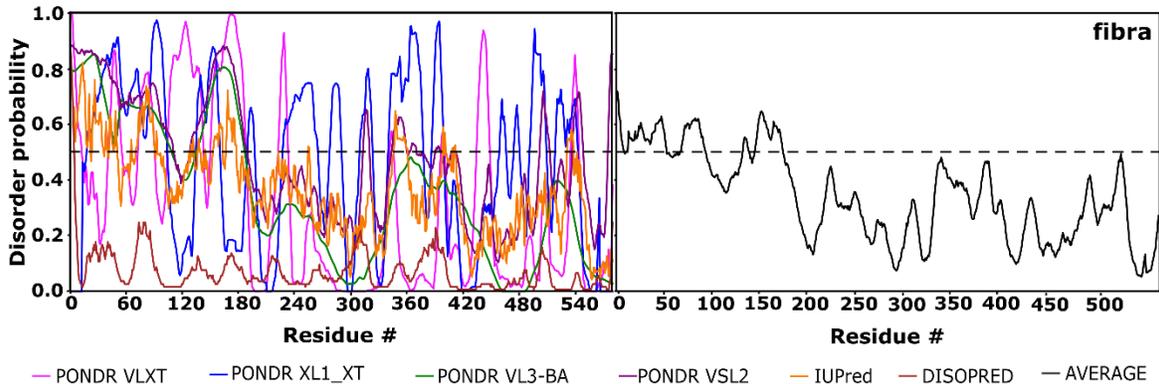


Figura 14. Predicción de desorden en las proteínas codificadas en L5. A la izquierda se muestra el resultado de los predictores PONDRL VLXT (línea rosa), PONDRL VL1_XT (línea azul), PONDRL VL3-BA (línea verde), PONDRL VSL2 (línea purpura), IUPred (línea naranja) y DISOPRED (línea roja). A la izquierda se muestra el promedio de los seis predictores (línea negra). Los valores superiores a 0.5 indican desorden.

V.2 Distribución del desorden en la partícula viral

El análisis del desorden intrínseco en el proteoma de Adenovirus humano revela que existe un elevado porcentaje de secuencias desordenadas a lo largo del proteoma pues 21 de las 34 proteínas analizadas presentaron al menos una región desordenada de 30 aminoácidos o más. Resulta interesante el hecho de la mayoría de las proteínas estructurales contengan IDRs e incluso que algunas sean clasificadas como IDPs (Fig. 4), sin embargo, este no es un comportamiento exclusivo para adenovirus pues se ha reportado que las proteínas en las cápsides de VIH e influenza contienen regiones desordenadas [23]. Si se analiza la distribución de las IDRs en la cápside viral de AdH podemos apreciar que las proteínas con mayor porcentaje de IDRs son las que se unen directamente al DNA viral. Mientras que las proteínas más alejadas del genoma presentaron los menores porcentajes de regiones desordenadas. Este compartimiento también se ha observado en las partículas virales de otros virus. Se plantea que las proteínas de unión al genoma viral presentan un mayor grado de desorden ya que estas deben interactuar con el genoma viral, así como con otras proteínas estructurales [23], de modo que el hecho de ser desordenadas le brinda a la proteína la flexibilidad para

establecer estas interacciones, así como la capacidad de interactuar con múltiples blancos moleculares.

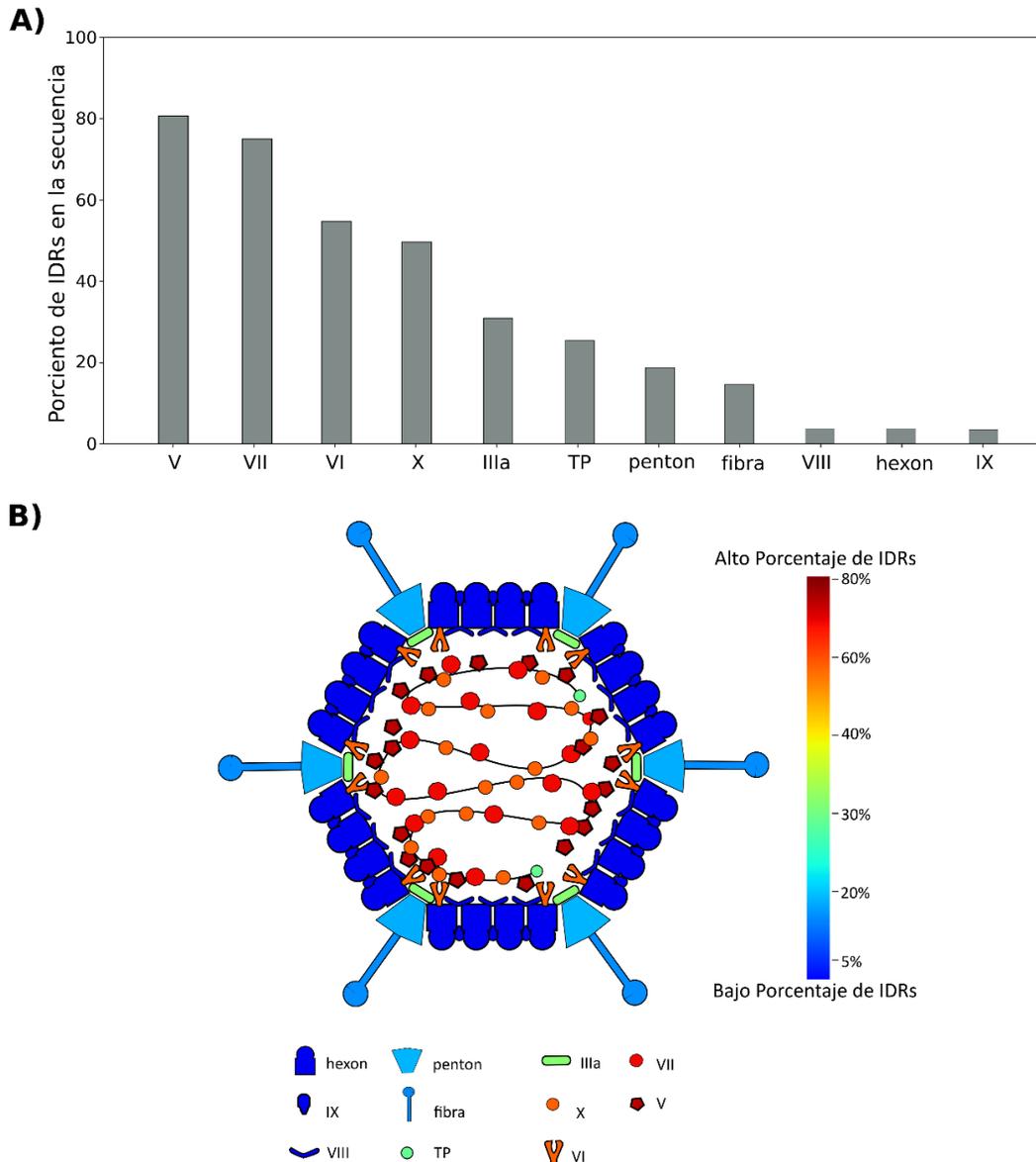


Figura 15. Distribución del desorden en la partícula viral de AdH. (A) Porcentaje de la secuencia ocupado por IDRs en las proteínas estructurales. (B) Representación del desorden intrínseco en la partícula viral. Las proteínas esta coloreadas en función dl porcentaje de IDR. La barra de color indica el porcentaje de IDR en las proteínas. La figura no está en escala.

V.3 Caracterización de las IDRs en las proteínas de adenovirus humano

Con el objetivo de analizar en un mayor grado las regiones desordenadas en las proteínas de adenovirus humano se realizó la caracterización de estas regiones. Además se llevó a cabo un análisis de las actividades y las funciones reportadas

para las proteínas de adenovirus con el fin de determinar las posibles funciones de las IDR en sus mecanismos de acción.

22K:

La proteína viral 22K presenta una IDR de 126 aminoácidos, la cual ocupa el 65% del total de la proteína (Fig. 16A). La caracterización de esta región por CIDER sugiere que esta región es capaz de plegarse en función de su ambiente (Fig. 16B), pudiendo adoptar estructuras secundarias tipo hélices alfa según la predicción de estructura secundaria (Fig. 16C). Aun no se conoce con claridad la función de esta región desordenada de gran tamaño. 22K participa en el empaquetamiento del genoma viral junto a la proteína viral IVa2 [52], por lo cual se pudiera suponer que la presencia de esta IDR está relacionada con la unión de la proteína al DNA, al mismo tiempo que le brinda la flexibilidad para interactuar con IVa2 y con otras proteína durante el proceso de encapsidación.

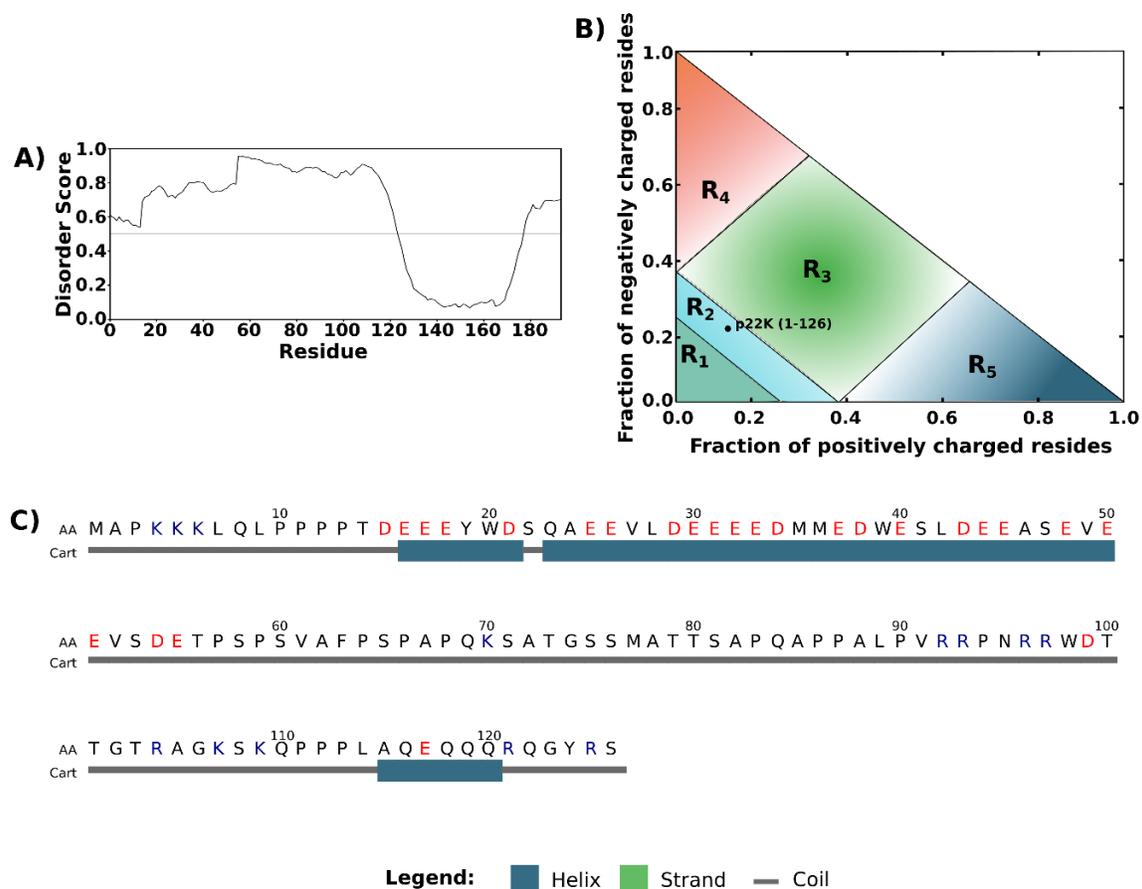


Figura 16. Caracterización de las IDRs presentes en 22K. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia que típicamente

son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDR. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

33K:

La proteína viral 33K presenta una IDR de 163 aminoácidos la cual ocupa el 71% de la misma. 33K al igual que 22K interacciona con DNA y con la proteína IVa2 [45], esto pudiera deberse a que ambas proteína presentan 100% de identidad en el extremo amino terminal (1-105), región que forma parte de las IDRs presentes en ambas proteínas. Por lo que al igual que en 22K esta IDR pudiera estar relacionada con la unión al DNA viral y/o a la interacción con IVa2. A pesar de que tanto 22K como 33K interaccionan con el DNA y IVa2 realizan funciones diferentes en el ciclo de replicación viral [50].

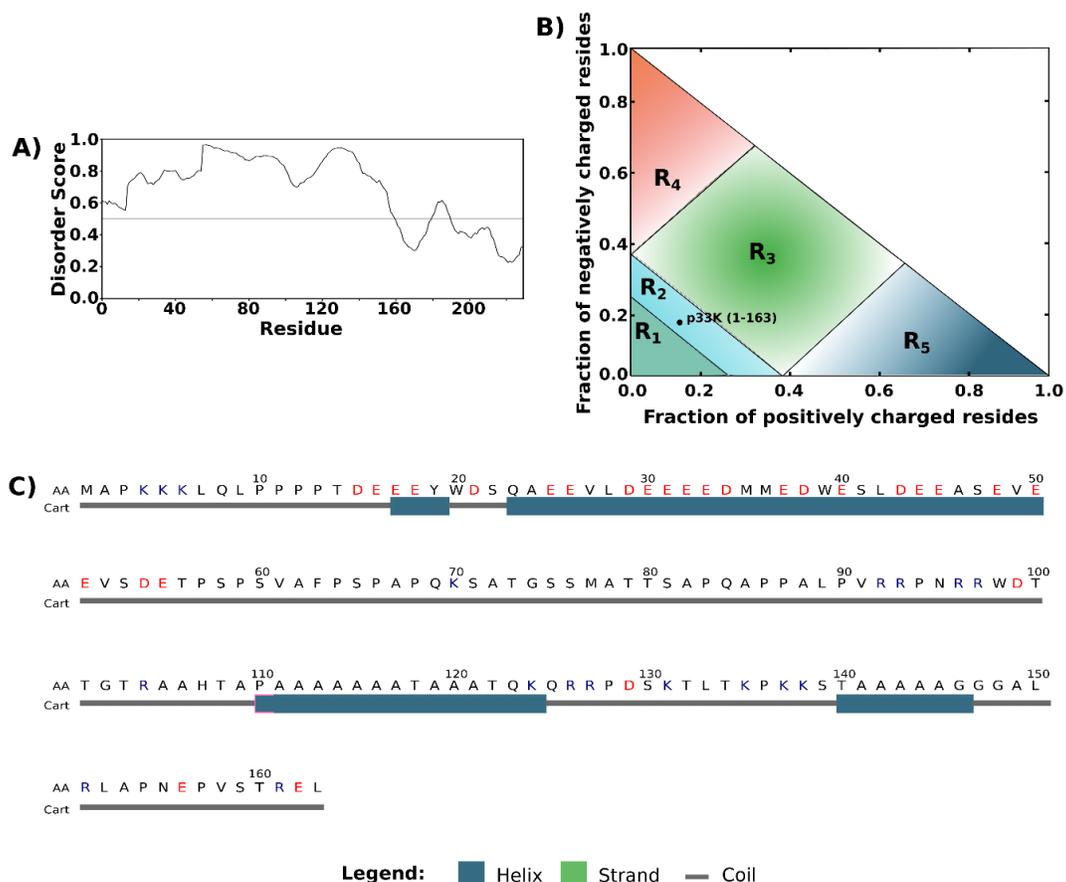


Figura 17. Caracterización de las IDRs presentes en 33K. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2.

IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia que típicamente son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policationica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

DBP:

La proteína de unión a simple cadena de DNA viral (DBP) presenta dos regiones IDRs (Fig. 18A). La primera de ellas ocupa toda la región amino terminal (1-173) y su clasificación por CIDER surge que el posible plegamiento de esta región va a depender de la distribución de cargas de la misma (Fig. 18B). Al analizar la distribución de cargas se observa la presencia de regiones negativas y de regiones positivas las cuales se separan por regiones ricas en prolina. De ahí que en la predicción de estructura secundaria predomine la ausencia de estructura secundaria (Fig. 18C). La segunda región intrínsecamente desordenada se encuentra al centro de la proteína y presenta una longitud de 30 aminoácidos. Esta IDR al igual que la anterior se clasifica como una región cuyo plegamiento depende de la distribución de carga en su secuencia. A diferencia de la región 1-173, la región 297-326 de DBP no presenta parches cargados, si bien se observa que las cargas negativas se encuentran al principio de la secuencia y las positivas al final de la misma, los aminoácidos cargados están separados entre sí por al menos dos aminoácidos no cargados en la mayoría de los casos (Fig. 18C). Aunque se conoce la función de DPB no está claro cuál es el papel de las regiones desordenadas en esta proteína. Se especula que la IDR presente en el extremo amino terminal está implicada en la separación de fases que ocurre durante la formación de los centros de replicación, pues se ha reportado el papel de las IDPs e IDR en este proceso [53].

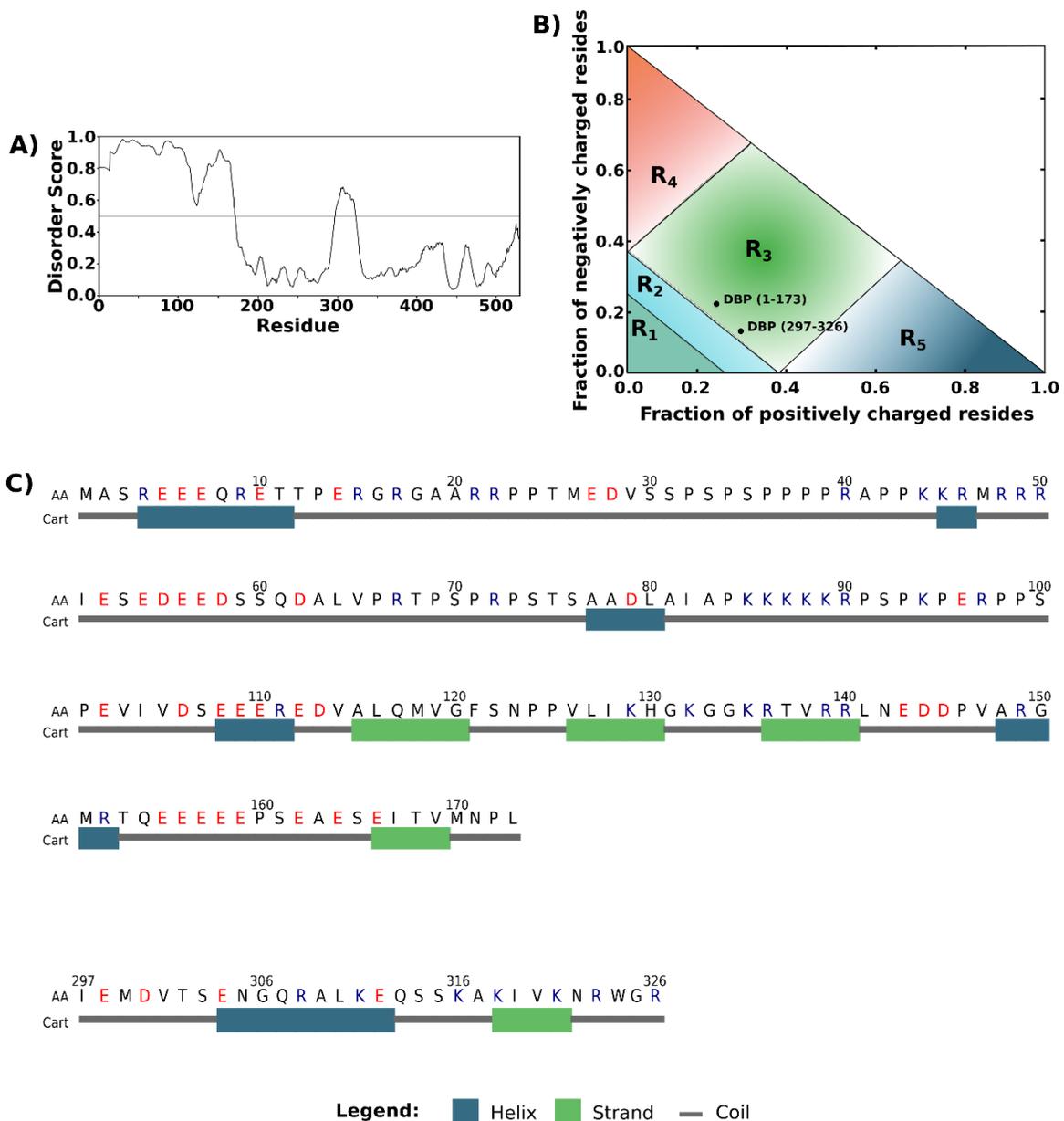


Figura 18. Caracterización de las IDR presentes en DPB. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDR por CIDER. R1: Secuencia que típicamente son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDR. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

E1B 19K:

La proteína E1B 19K presenta una IDR de 34 aminoácidos ubicada en el extremo carboxilo terminal (Fig. 19). Esta IDR se clasifica como una región cuya estructura va a depender de la distribución de cargas (Fig. 19B). Al analizar cómo se distribuyen las cargas en esta secuencia vemos que las cargas negativas se distribuyen a lo largo de la secuencia, mientras que las cargas positivas se localizan mayoritariamente al centro de la región (Fig. 19C), lo cual podría dificultar la formación de estructura secundaria en esta región. Se conoce que la proteína está involucrada en la inhibición de la vía apoptótica, al funcionar como una proteína antiapoptótica tipo BCL-2 [28, 54], al unirse a las proteínas proapoptóticas Bax y Bak, para prevenir su oligomerización y por consiguiente la formación de poros en la membrana de la mitocondria [26, 55]. Sin embargo, no se ha propuesto una posible función para la región intrínsecamente desordenada de esta proteína viral.

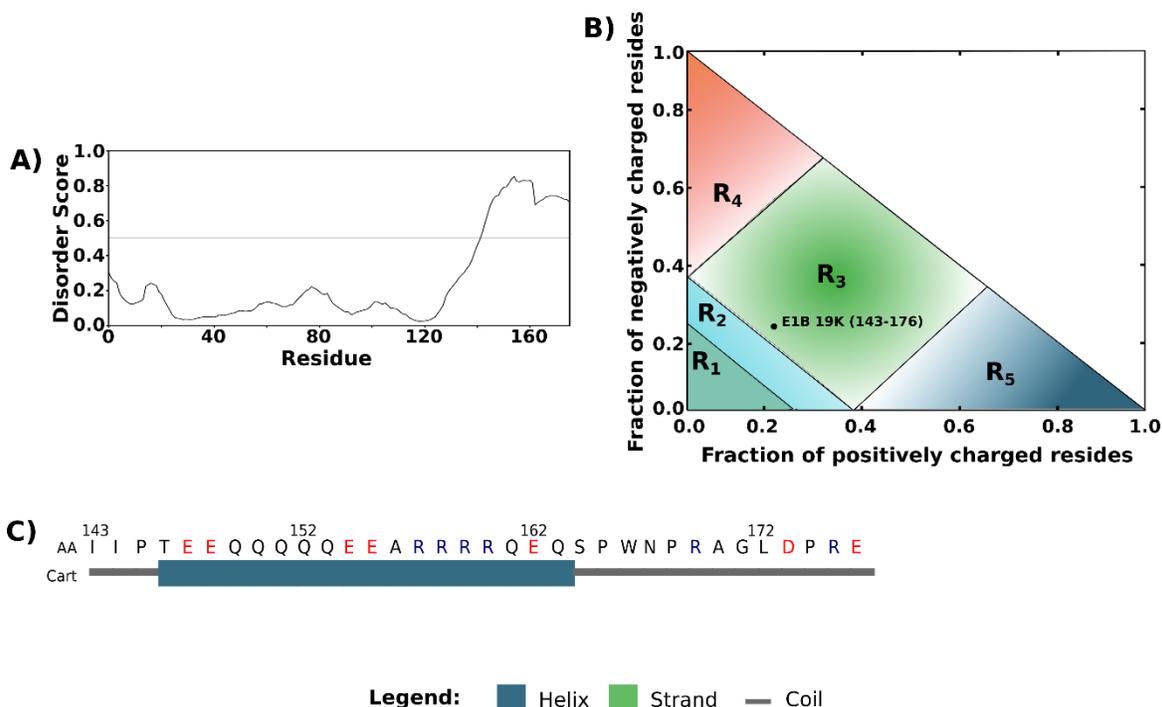


Figura 19. Caracterización de las IDRs presentes en E1B 19K. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policationica. **C)** Predicción de

estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

pX:

La proteína viral pX forma parte del núcleo de la partícula viral e interacciona directamente con el DNA viral [7]. Esta proteína se clasifica como una proteína intrínsecamente desordenada (Fig. 4), presenta una IDR de 51 aminoácidos que ocupa más de la mitad de la proteína (Fig. 20A). la clasificación de la misma por CIDER surgiere que esta región se encuentra compacta (Fig. 20B). Al analizar la secuencia de esta región se observa que es rica en aminoácidos positivos lo cual pudiera estar relacionado con la función de la proteína y su unión al DNA viral.

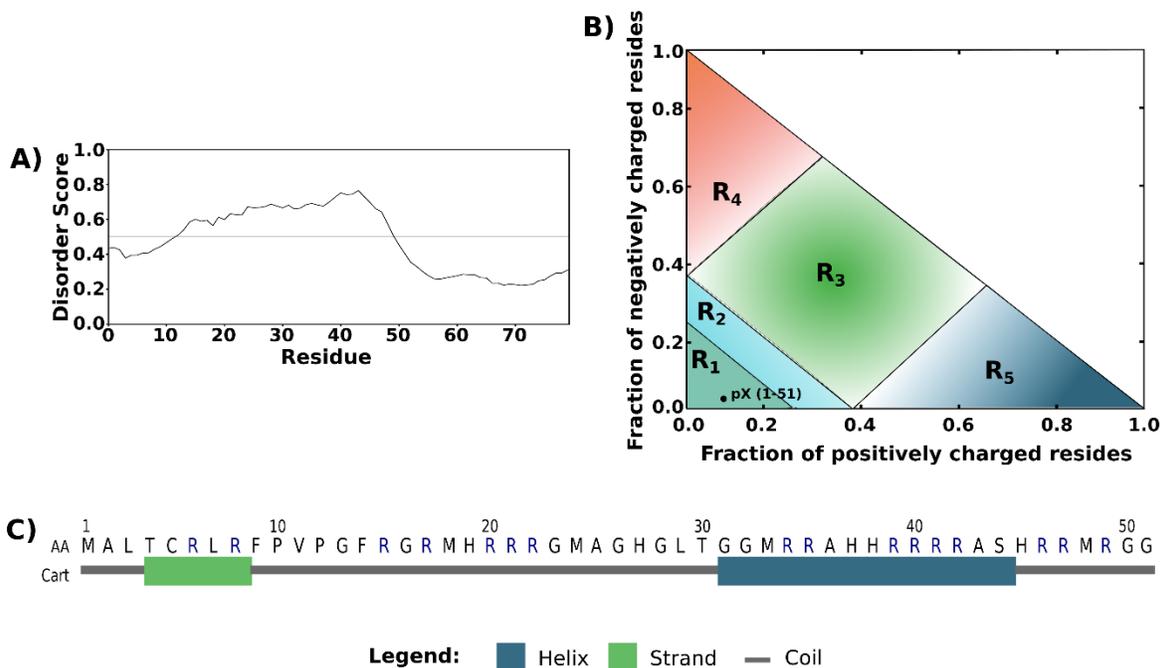


Figura 20. Caracterización de las IDRs presentes en pX. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policationica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

pVII:

La proteína pVII al que pX es una proteína que se una al DNA viral en el núcleo de la partícula viral [7] y se clasifica como una proteína intrínsecamente desordenada (Fig. 4). Esta proteína se encuentra prácticamente desordenada ya que el 89% de la proteína se predice como tal. El análisis de la secuencia de la proteína muestra que es rica en aminoácidos positivos (Fig. 21C) lo cual es de esperar pues interacciona directamente con el DNA. El análisis con CIDER surge que el plegamiento de esta región depende fundamentalmente del ambiente, de modo que las estructuras secundarias predichas pudieran formarse al interactuar con el DNA viral con el resto de las proteínas de la cápside.

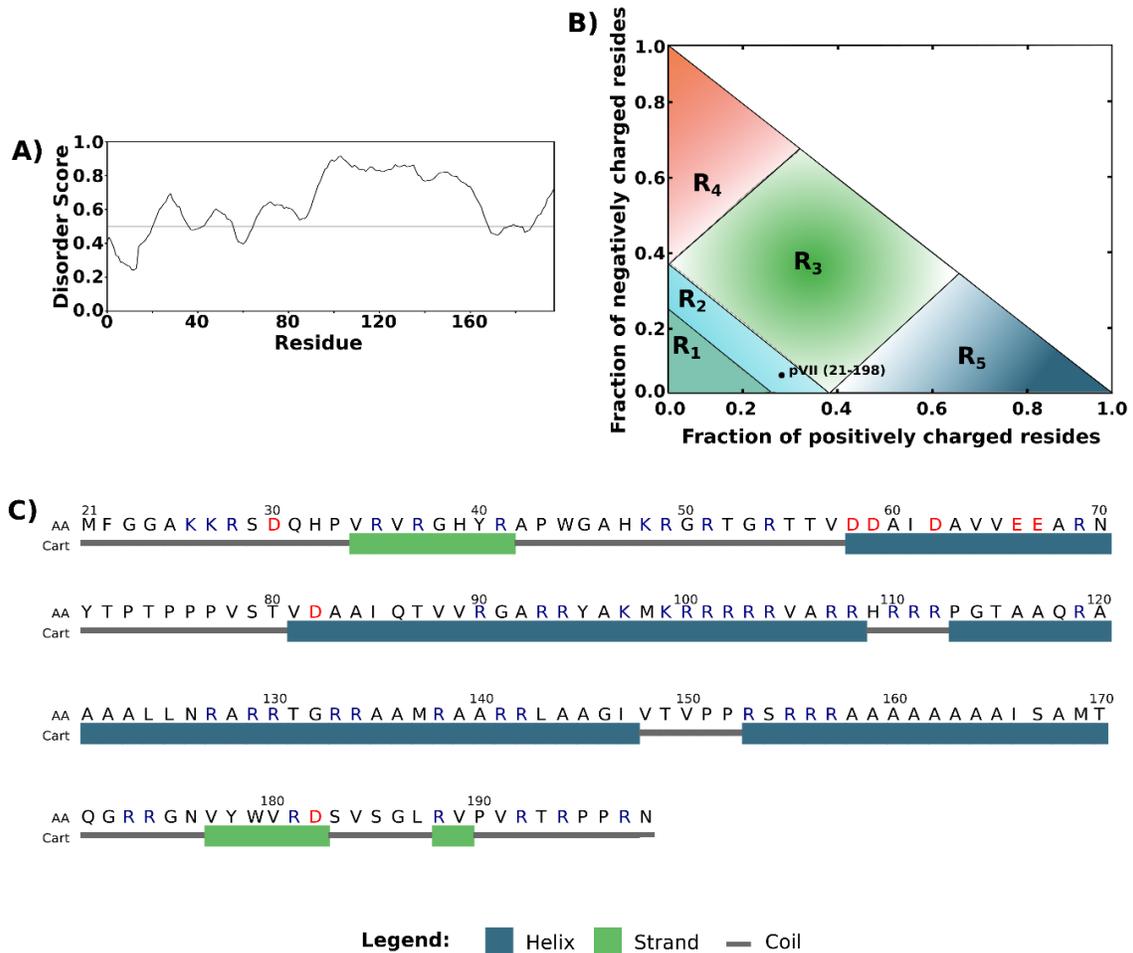


Figura 21. Caracterización de las IDRs presentes en pVII. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R₁: Secuencia típicamente globulares. R₂: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R₃: Secuencias cuya estructura va a depender de la

distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

pVI:

La proteína pVI se encuentra asociada a la superficie interna de la cápside, de manera que interactúa tanto con las proteínas de la superficie externa como con las del núcleo de la partícula viral. Más del 50% de pVI se predice como intrínsecamente desordenado (Fig. 22A). La IDR presente en pVI tiene una longitud de 153 aminoácidos y se clasifica por CIDER como una región típicamente globular y compacta. Se cree que el papel de esta IDR es brindarle a pVI la capacidad de interactuar con varias proteínas de la cápside viral.

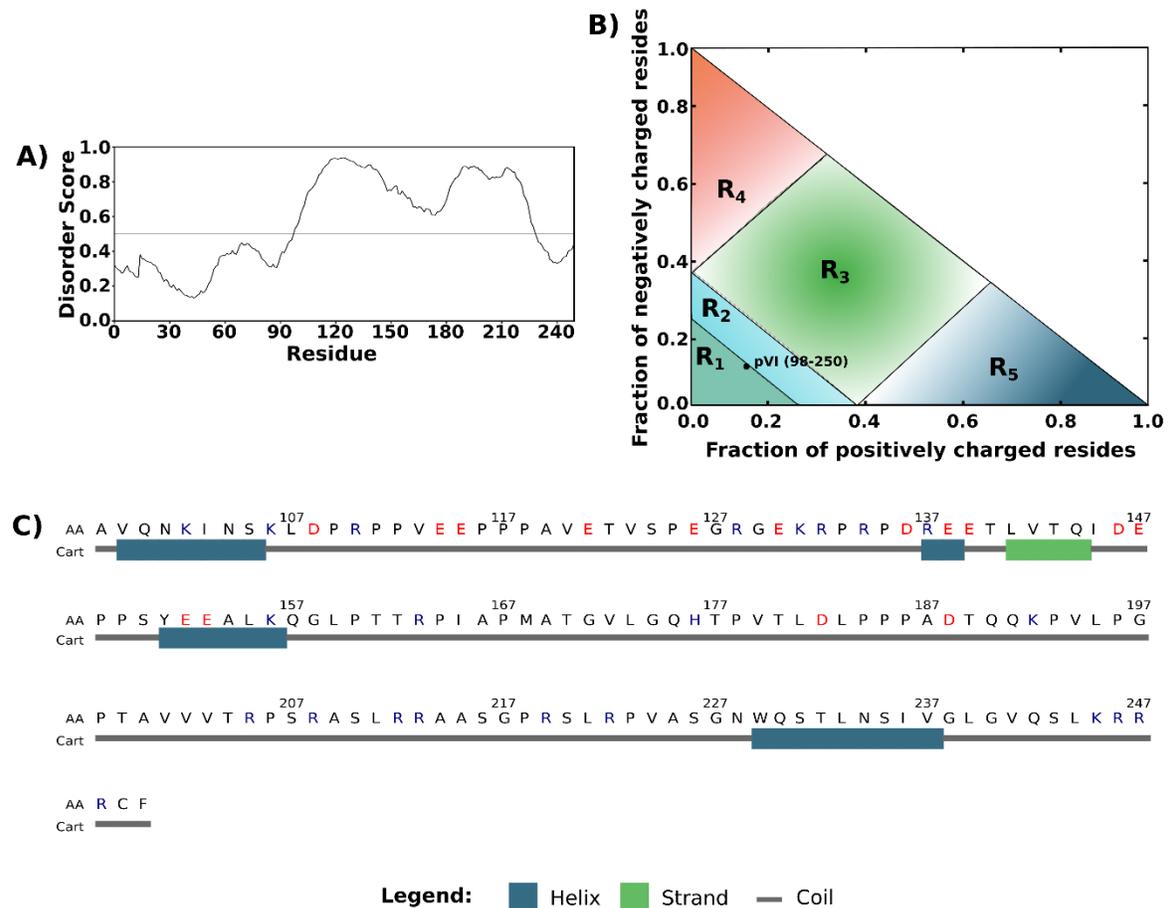


Figura 22. Caracterización de las IDRs presentes en pVI. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración

salina, interacciones en *cís*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

pV:

La proteína viral pV se clasifica como una proteína intrínsecamente desordenada (Fig. 4) que se encuentra interactuando con el genoma viral en el núcleo de la partícula viral. Más de 80% de la proteína se predice como intrínsecamente desordenado y la proteína presenta dos IDRs, una de 87 aminoácidos y otra de 240 aminoácidos (Fig. 23A). Ambas regiones son ricas en aminoácidos positivos, por lo que, la función de las mismas pudiera ser interactuar directamente con el DNA viral.

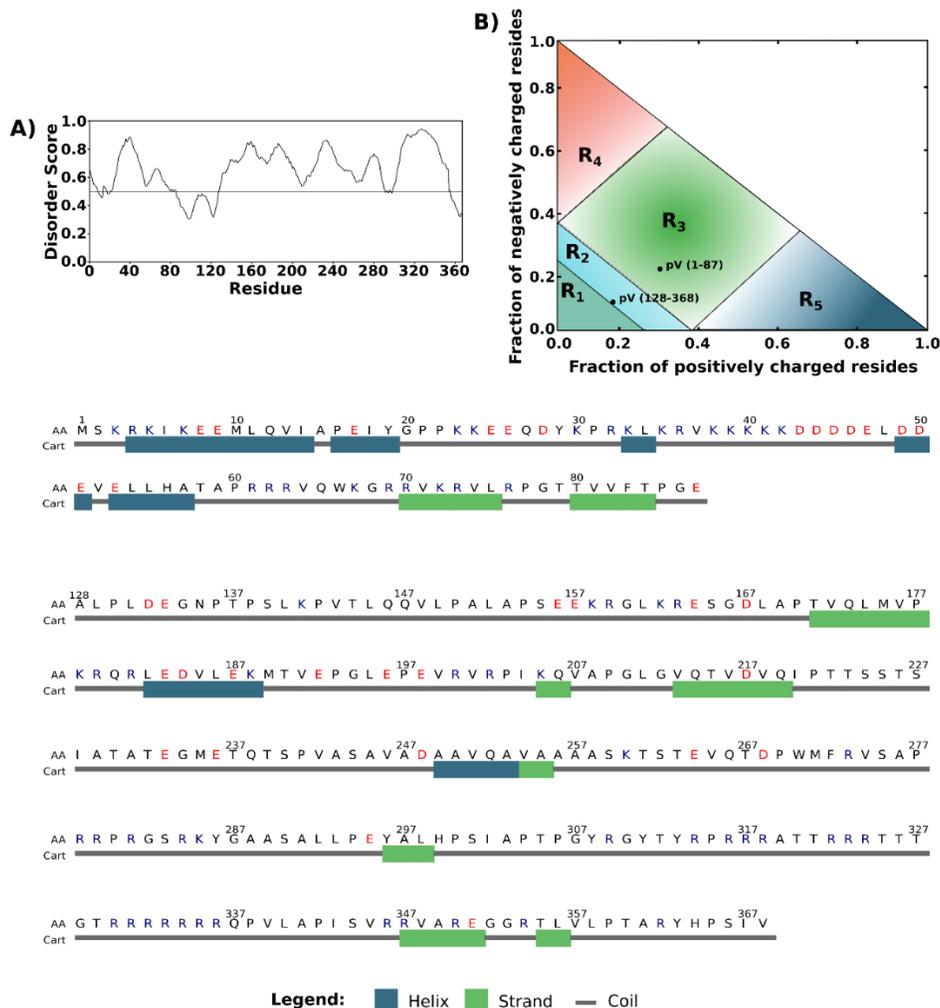


Figura 23. Caracterización de las IDRs presentes en pV. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2.

IUPred y DISOPRED. **B)** Caracterización de las IDR's por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDR's. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

TP:

La proteína terminal TP se encuentra unida covalentemente a los extremos 5' del genoma viral. Su principal función es brindar el extremo 3' OH que la polimerasa viral requiere para la replicación del DNA [1]. TP presenta tres IDR's distribuidas a lo largo de su secuencia (Fig. 24A). La primera de ellas presenta 32 aminoácidos, su posible plegamiento depende de la distribución de cargas en su secuencia, la cual se caracteriza por presentar abundantes cargas negativas hacia un extremo y cargas positivas más dispersas hacia el otro. La segunda IDR presente en TP es la de mayor longitud en esta proteína, esta se localiza en la región central de la misma, desde el residuo 332 al residuo 436 con una extensión de 105 aminoácidos. Al igual que la IDR anterior el plegamiento de esta IDR depende en gran medida de la distribución de las cargas en su secuencia. Al analizar dicha distribución de cargas observamos que estas no se encuentran distribuidas uniformemente y que se observan regiones con alta densidad de aminoácidos positivos como regiones con abundantes aminoácidos negativos, lo cual podría afectar la formación de estructuras secundarias. La tercera región desordenada presente en TP se localiza en el extremo carboxilo terminal de la proteína y se predice como una región carente de estructura secundaria que se encuentra principalmente compacta al ser analizada por CIDER. A pesar de que esta descrita cual es la función principal de la proteína TP, aun no se han propuestos cuales puedan ser las funciones asociadas a las regiones intrínsecamente desordenadas presentes en la misma.

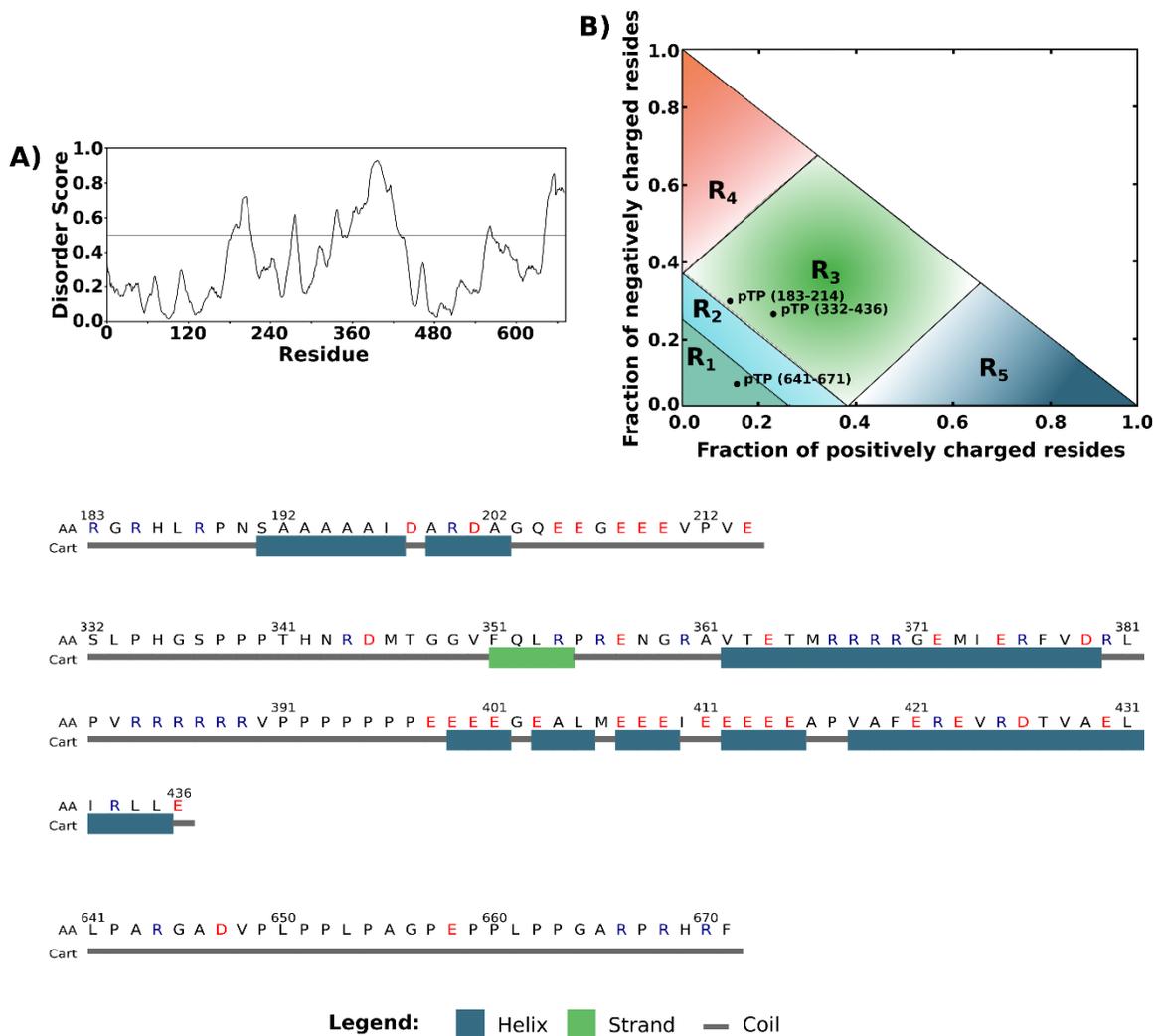


Figura 24. Caracterización de las IDRs presentes en pTP. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policitiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

AdPol:

La polimerasa viral se predice en su mayoría como una proteína plegada, lo cual coincide con el planteamiento de que las enzimas son mayoritariamente estructuradas. A pesar de ello AdPol presenta dos IDRs, una de 91 aminoácidos de longitud y la otra de 41 aminoácidos. La primera de dichas regiones se ubica en el

extremo amino terminal de la proteína y presenta abundantes aminoácidos positivos en su secuencia, por lo que la función de esta región podría ser el reconocimiento del DNA durante la replicación. La segunda IDR presente en AdPol se localiza más hacia el extremo carboxilo de la proteína y su plegamiento se predice que depende de la interacción con factores externos como la concentración salina, pH o interacción con blancos moleculares (Fig. 25).

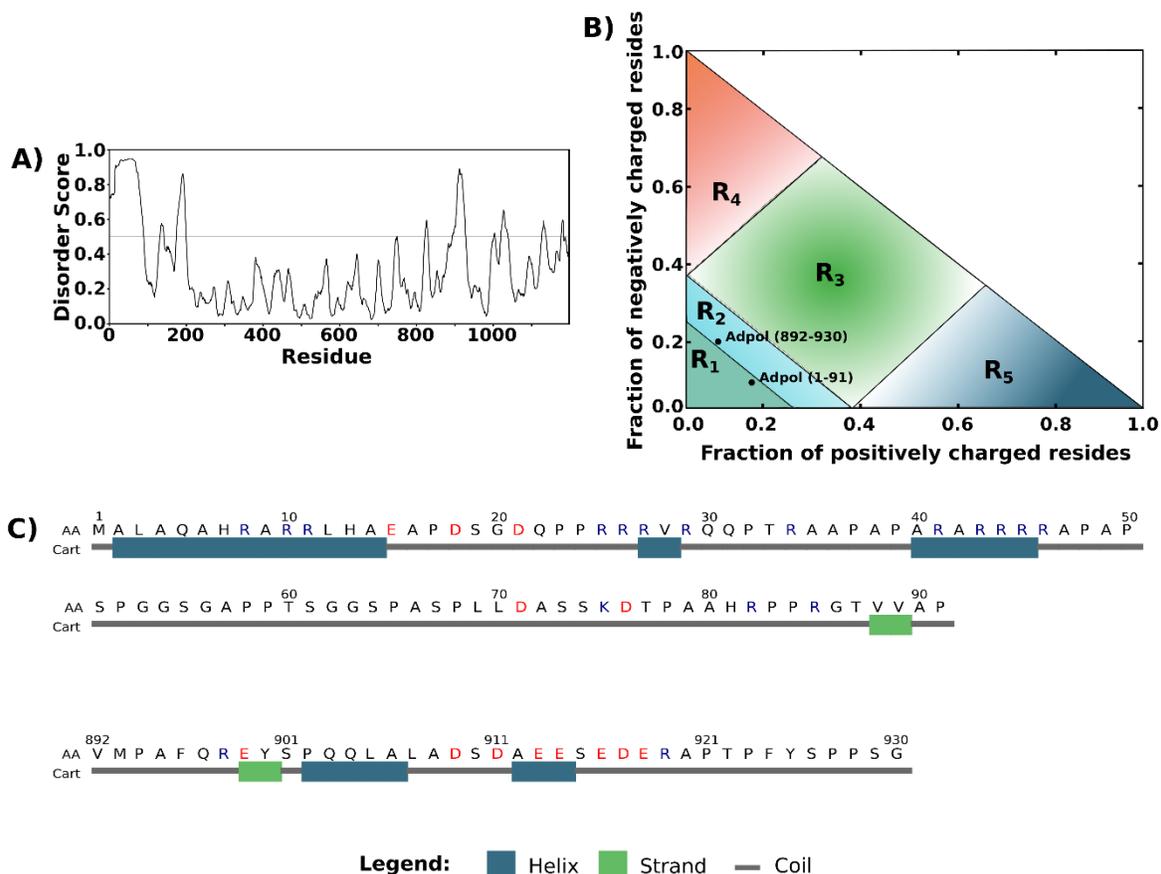


Figura 25. Caracterización de las IDRs presentes en AdPol. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

Illa:

La proteína estructural Illa forma parte de la cápside viral y se encuentra asociada a la superficie interna de la cápside, por lo que esta proteína debe interactuar a la vez con proteína de la superficie externa de la partícula viral, como con proteínas del núcleo de la partícula viral. En este sentido la presencia de una IDR hacia su extremo carboxilo terminal le podría brindar a la proteína la capacidad de interactuar con más de un blanco molecular. La IDR presente en Illa, al ser clasificada por CIDER (Fig. 26B) se predice como una secuencia cuyo plegamiento depende de la interacción con sus blancos moleculares lo que está de acuerdo con lo planteado anteriormente en relación a su posible función.

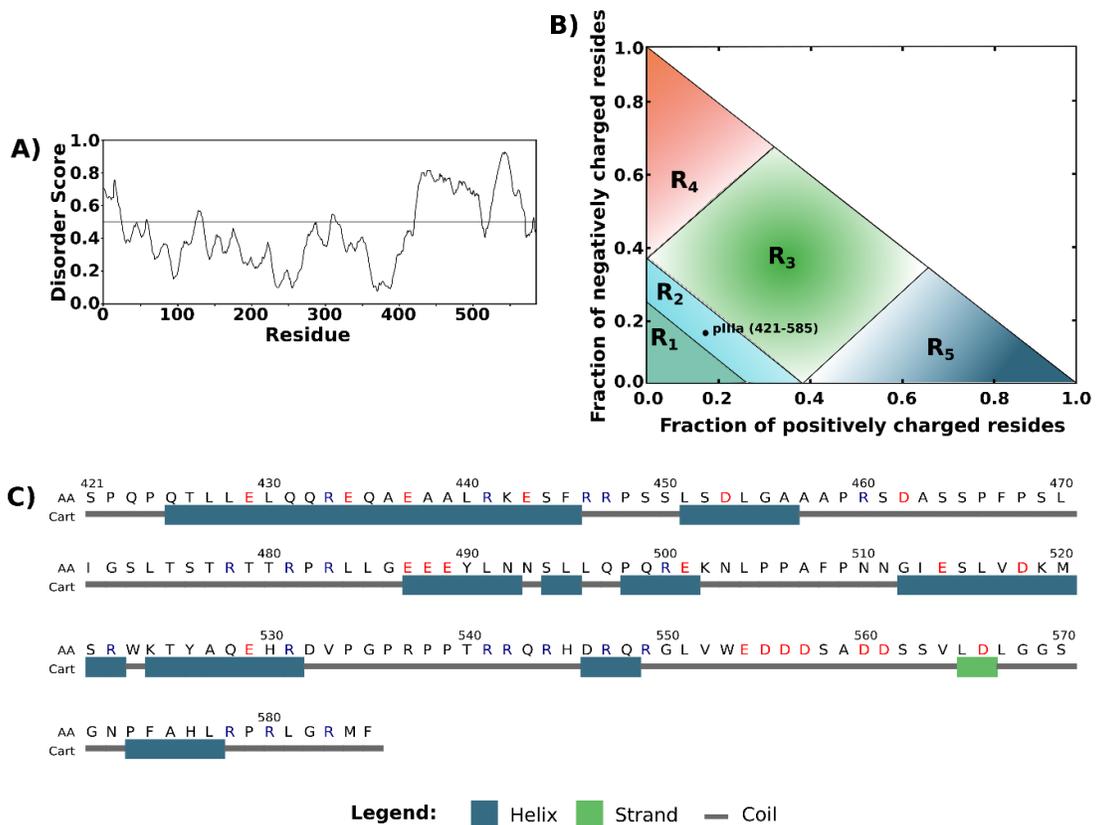


Figura 26. Caracterización de las IDRs presentes en pIlla. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policitiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

Penton:

Penton es una proteína estructural que se localiza en la superficie externa de la partícula viral y se localiza en cada uno de los vértices del icosaedro, actuando como base para la proteína fibra. Esta proteína viral presenta dos IDR, una de 30 aminoácidos localizada en el extremo amino terminal y otra de 96 aminoácidos localizada al centro de la proteína. La estructura cristalográfica obtenida para esta proteína viral confirma la existencia de estas regiones desordenadas, pues dichas regiones presentan pérdida de densidad electrónica, lo cual se corresponden con regiones que no presentaron una única conformación en el cristal.

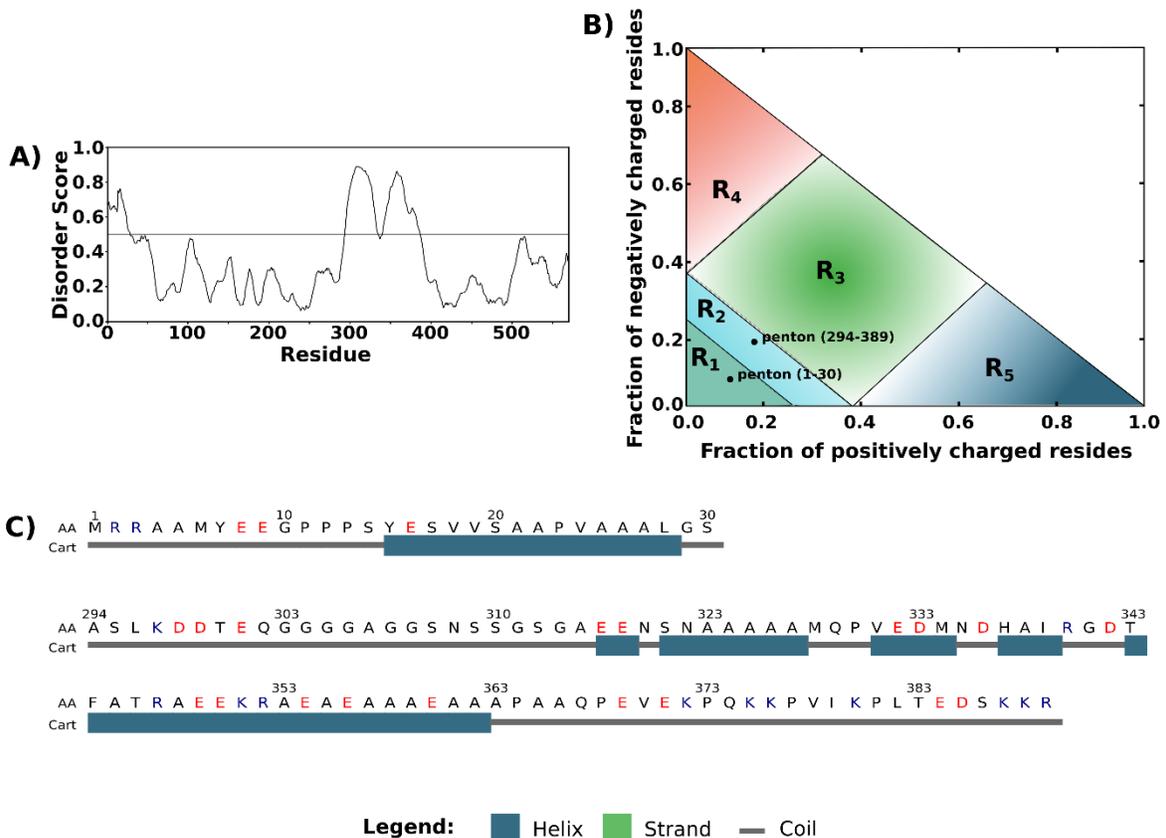


Figura 27. Caracterización de las IDRs presentes en penton. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policationica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

Hexon:

Hexon es la proteína más abundante en la partícula viral y juega un papel importante en el transporte de la partícula viral hacia el núcleo celular para la liberación del genoma nuclear, ya que es la única proteína de la cápside que interacciona con las Dineínas celulares [56]. Hexon presenta siete regiones hipervariables a lo largo de su secuencia [57], las cuales están relacionadas con la evasión del sistema inmune por parte de la partícula viral [58]. Hexon presenta una única IDR de 33 aminoácidos, la cual coincide con la región Hipervariable 1. Esta región se caracteriza por presentar abundantes aminoácidos negativos lo que la clasifica como una secuencia que se encuentra mayoritariamente extendida (Fig. 28B).

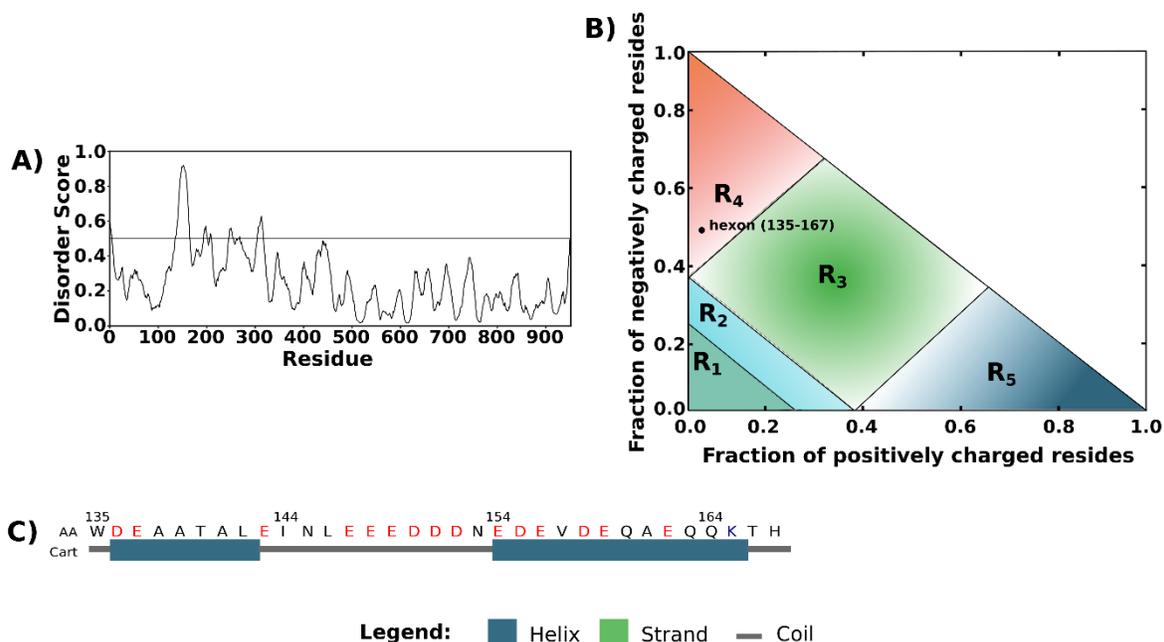


Figura 28. Caracterización de las IDRs presentes en hexon. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

E4orf6/7:

La proteína temprana E4orf6/7 presenta una única IDR de 93 aminoácidos que ocupa el 62% de la proteína (Fig. 29A). Esta IDR se predice como una secuencia globular (Fig. 29B) que prácticamente carece de estructura secundaria (Fig. 29C).

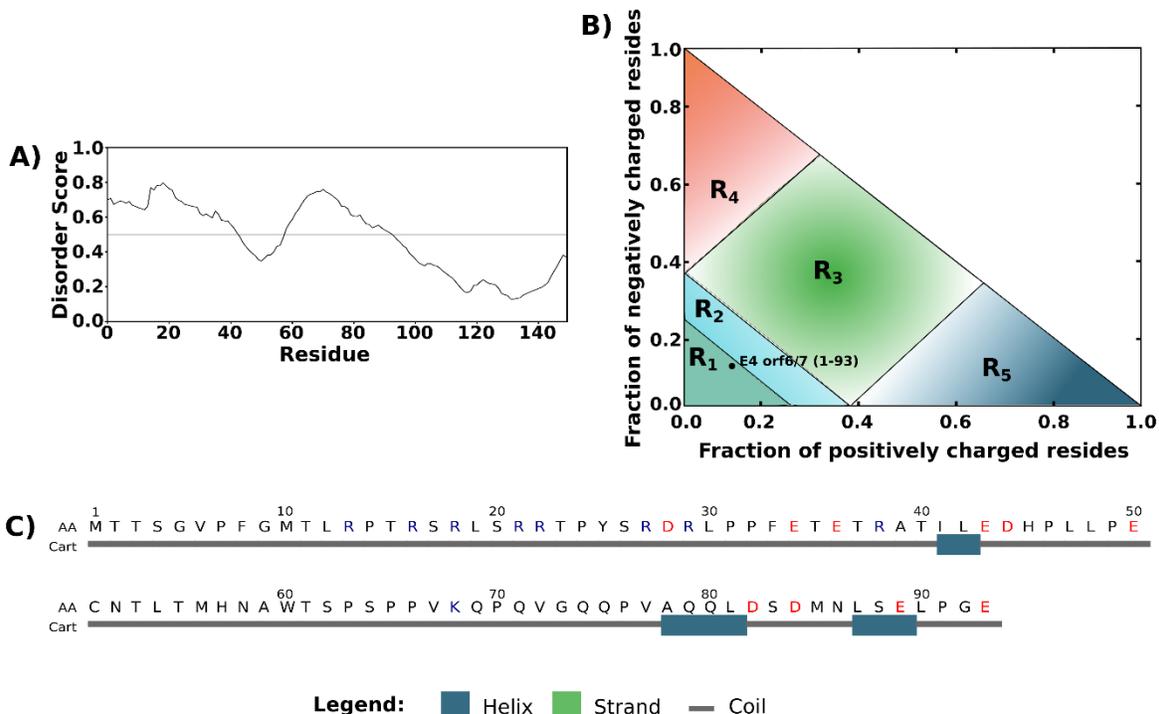


Figura 29. Caracterización de las IDRs presentes en E4orf6/7. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globular. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policitiónica. **C)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

E1A:

E1A es una proteína intrínsecamente desordenada (Fig. 4), en la cual se predice que más del 80% de sus residuos carecen de estructura (Fig. 28B). Esta proteína actúa como regulador de la transcripción durante el ciclo de replicación viral e induce la entrada a la fase S del ciclo de replicación celular al interactuar con la proteína Retinoblastoma e inducir su disociación de E2F. Estudios por RNM han demostrado que la proteína se encuentra desplegada y que la interacción con CBP induce la

formación de una hélice alfa (E59-I65) [25], la cual también se predice por Psipred (Fig. 28C). Durante el ciclo de replicación viral E1A interactúa con varios blancos moleculares, en ese sentido el hecho de sea una IDP pueden brindarle a la proteína la capacidad de plegarse para interactuar con dichos blancos, así como la flexibilidad para interactuar con más de un blanco a la vez.

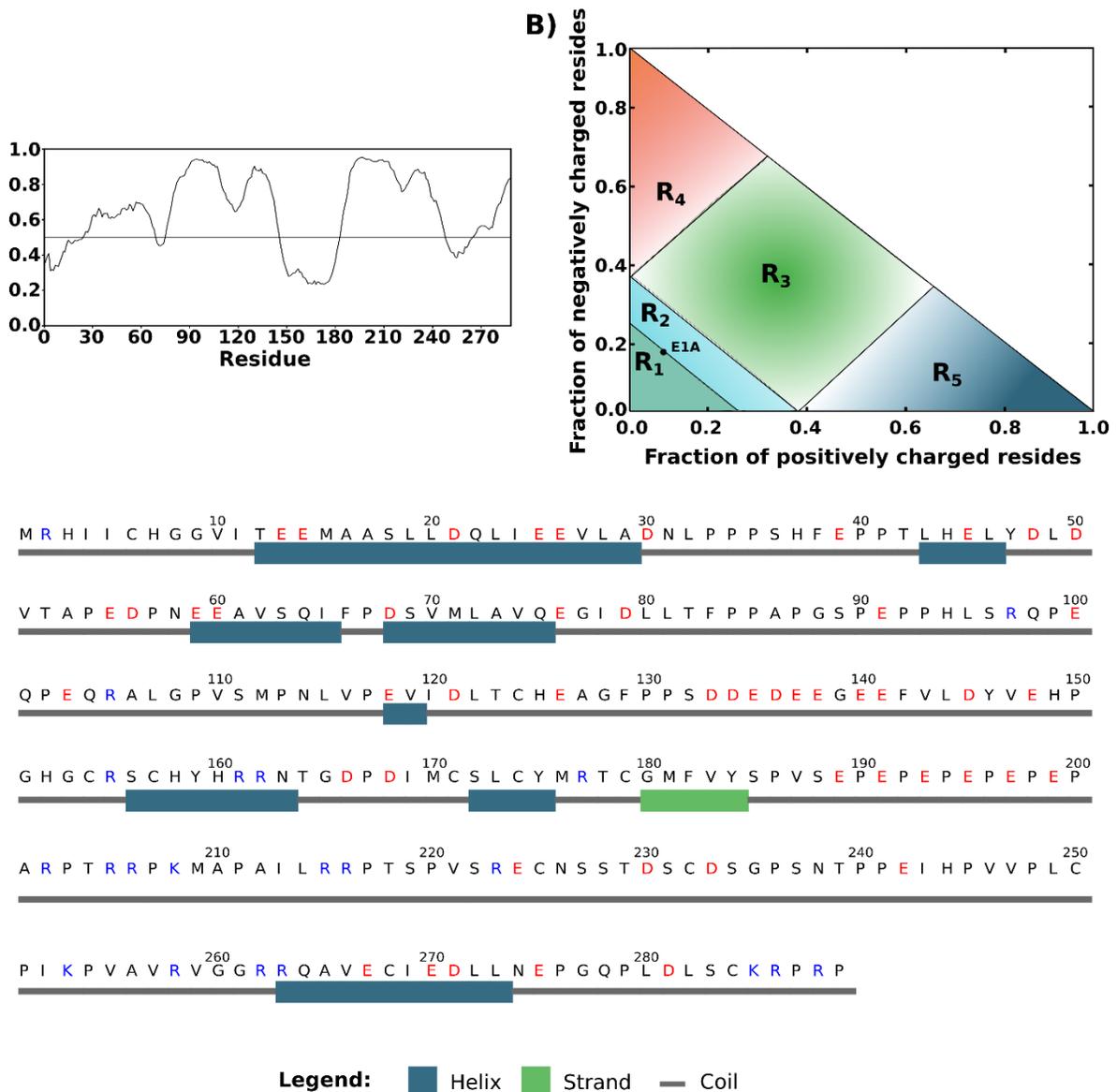


Figura 30. Caracterización de las IDRs presentes en E1A. A) Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia típicamente globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatónica. **C)** Predicción de

estructura secundaria en las IDR. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

52/55K:

Estas proteínas virales son derivados fosforilados de 52 y 55 kDa obtenidos a partir de un polipéptido de 415 residuos con una talla molecular de 48kDa. Los estados fosforilados de esta proteína no cambian durante la infección. 52/55K presenta un sitio de corte para la proteasa viral entre los residuos T351-G352, lo cual origina un producto de 40 kDa. Durante la infección esta proteína viral existe como un homodímero unido por puente disulfuro. Se han observado los dímeros 52/55K-52/55K, 52/55K-40kDa y 40kDa-40kDa [59]. Tanto la predicción de desorden como la predicción de estructura secundaria (Fig. 31A y 31D) surgieron que la región donde ocurre el corte proteolítico carece de estructura. Se ha reportado la presencia de IDRs en las regiones donde ocurren los cortes proteolíticos lo cual permite que la proteasa sea capaz de reconocer diferentes blancos moleculares [10]. El desorden en las regiones que sufren cortes proteolítico también se ha observado en la poliproteína de Chikungunya virus [21]. La caracterización de las regiones desordenadas (Fig. 31B) surge que la región en la cual se encuentra el corte proteolítico es capaz de plegarse bajo determinadas condiciones ambientales. Bajo estos supuestos, la formación de las hélices que se predicen que flanquean el sitio de corte (Fig. 31D) puede inducirse en presencia de la proteasa viral, permitiendo el reconocimiento y la interacción molecular.

La proteína 52/55K participa en la encapsidación serotipo específica del DNA viral [50]. Esta proteína se une de manera indirecta a las secuencias de empaquetamiento e interactúa directamente con la proteína viral IVa2. Aunque su interacción con IVa2 no es necesaria para la unión al DNA si lo es para la encapsidación [60, 61]. Los primeros 173 residuos del extremo amino terminal son esenciales para la unión con IVa2 [48], mientras que los primeros 191 son los que permiten la encapsidación serotipo específica del DNA viral [50]. El análisis del desorden en la región amino terminal muestra que esta región es una IDR (Fig. 31A). Lo cual le da la posibilidad de interactuar con más de un blanco molecular y la flexibilidad para interactuar con más de uno a la vez pues la proteína además de IVa2 debe interactuar con otra u otras proteínas de unión a DNA que garanticen su

unión indirecta a las secuencias de empaquetamiento. La caracterización de la IDR en el extremo amino por CIDER surge que esta región es capaz de plegarse para interactuar con blancos moleculares pues se encuentra en la región R3 (Fig. 31B). La formación o no de estructura en las secuencias tipo R3 va a depender de la distribución de cargas. Si analizamos la distribución de cargas en el extremo amino observamos que estas están dispersas a lo largo de la secuencia por lo que no deben existir repulsiones electrostáticas que impidan la formación de las estructuras secundarias predichas por IUPred (Fig. 31D). Todo lo discutido anteriormente evidencia que las IDRs en la proteína viral 52/55K son de vital importancia para el funcionamiento de la proteína, pues se localizan en las regiones de interacción molecular descritas, así como en las regiones de reconocimiento de la proteasa viral.

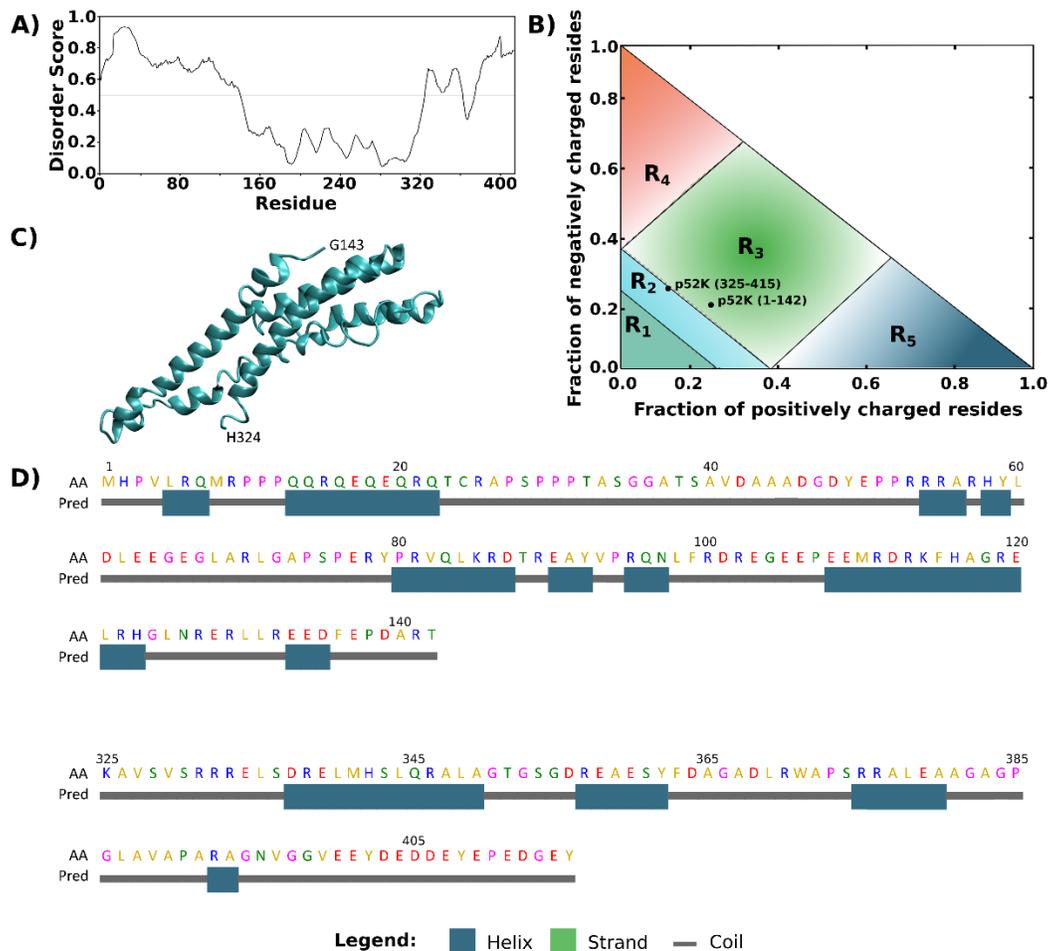


Figura 31. Caracterización de las IDRs presentes en 52/55K. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2. IUPred y DISOPRED. **B)** Caracterización de las IDRs por CIDER. R1: Secuencia que típicamente

son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policationica. **C)** Modelo de la región G143-H324 obtenido con I-TASER (C-score=-2.67) **D)** Predicción de estructura secundaria en las IDRs. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

100K:

100K es una proteína viral de 807 aminoácidos codificada en la región L4 de la unidad transcripcional de ML. Esta proteína promueve la traducción selectiva de los mRNAs virales durante la fase tardía de la infección [62] y promueve la trimerización e importación de la proteína viral Hexon al núcleo celular [63]. 100K favorece la traducción de los mRNAs virales bloqueando la traducción de los mRNAs celulares al promover un mecanismo alternativo para la traducción (Ribosome Shunting), para ello la proteína se une de manera selectiva a la región 5' no codificante de los mRNAs virales tardíos (5'NCR) e interacciona con el factor de traducción celular eIF4G [64]. La mutación de los residuos 329-RRK-331 a Ala evita la interacción de 100K con eIF4G [64], lo cual pudiera indicar que la interacción de estas proteínas se debe a interacciones electrostáticas. Por otra parte, la fosforilación de las Y365 y Y682 es necesaria para la interacción de 100K con 5'NCR. El efecto de la fosforilación de las tirosinas es aditivo y se necesita de ambas fosforilaciones para expresión selectiva de los mRNAs virales [65]. La fosforilación puede afectar las propiedades de la región en la que se encuentra (Fig. 32D) ya que las modificaciones postraduccionales afectan las propiedades de la IDRs y con ello las funciones de estas. Las modificaciones postraduccionales en las IDRs pueden inducir o evitar el plegamiento, afectar la distribución de carga, etc., regulando así la interacción de estas con sus blancos moleculares. En el caso de 100K la fosforilación de las tirosinas presentes en las IDRs permite la unión de la proteína a los mRNAs virales.

La metilación de las argininas 727, 736 y 741 es necesaria para la unión de la proteína 100K a Hexon [51]. La metilación de las argininas disminuye la carga positiva, cambiando así la distribución de cargas, la hidropatía y la valencia de los polipéptidos, afectando así sus propiedades de unión.

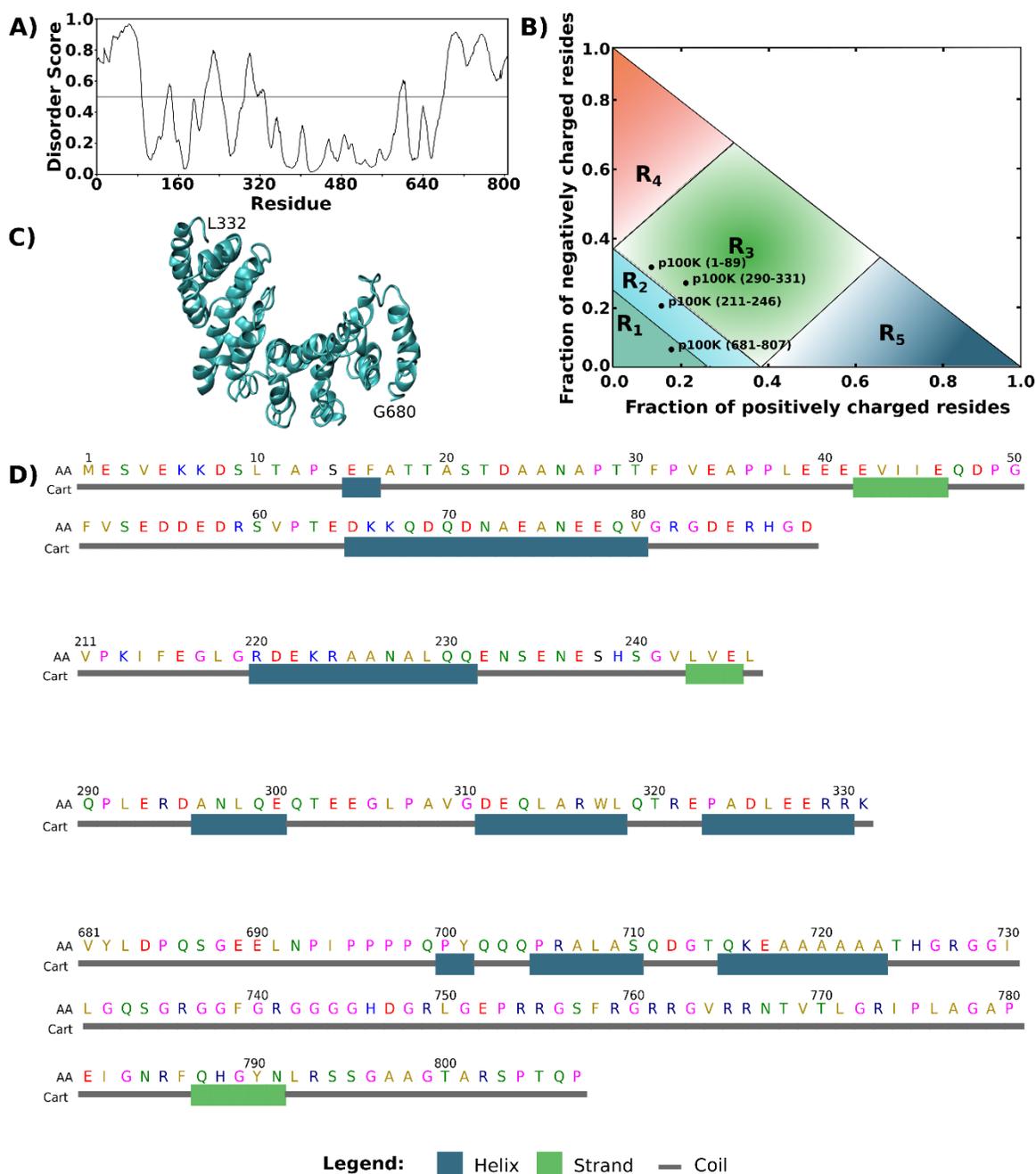


Figura 32. Caracterización de las IDRs presentes en 100K. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de la IDR por CIDER. R1: Secuencia que típicamente son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Modelo de la región L332-G680 obtenido con I-TASER (C-score=-3.83) **D)** Predicción de estructura secundaria en la IDR.

La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

La proteína 100K inhibe la actividad de la Granzima B humana (GrB). 100K actúa como un inhibidor competitivo interactuando directamente con el sitio activo. Los residuos D48 y P49 son los implicados en dicha interacción y los residuos del G688 al E781 son necesarios para la formación del complejo inhibidor-enzima [66]. El hecho de que los residuos implicados en la interacción con el sitio activo de la GrB sean desordenados coincide con el planteamiento de que las proteasas son capaces de reconocer varios blancos moleculares por la ausencia de estructura en los mismos [10].

E1B 55K:

E1B 55K es una proteína temprana de 496 aminoácidos que es modificada por SUMOilación en la Lys 104 y por fosforilación en Ser 490, Ser 491 y Thr 495. E1B 55K lleva a cabo múltiples funciones durante el ciclo de replicación viral. Entre las funciones que se han identificado para la proteína viral se encuentra la formación de un complejo E3 ubiquitina ligasa junto con la proteína viral E4 Orf6 y las proteínas celulares Elonginas B y C, el factor Rbx1 y Cullin 5. Este complejo se encarga de la ubiquitilación de proteínas celulares, marcándolas para ser degradadas por vía proteosomal [29]. Los sustratos identificados para el complejo incluyen el supresor tumoral p53 [67], proteínas de la maquinaria de reparación celular de daño a DNA (Mre11, Rad50, NBS1)[30, 42], la DNA ligasa IV, la integrina alfa-3 [43], la helicasa Bloom [68] y la proteínas de señalización celular ALCAM y EPHA2 [31]. La sustitución de los aminoácidos que son fosforilados por Ala resulta en la deficiente degradación de las proteínas celulares antes mencionadas [69]. La región en la que se encuentran los sitios de fosforilación presenta un elevado nivel de desorden estructural (Fig. 33A). En este sentido, la fosforilación de estos residuos podría cambiar las propiedades de interacción de la proteína con los blancos del complejo E3 Ubiquitina Ligasa o con las proteínas que forman dicho complejo. E1B 55K es capaz de regular la expresión de genes celulares y virales. Se conoce que inhibe la activación de la transcripción dependiente de p53 por interacción con el dominio de activación de este factor celular [33] y está implicada en la represión de la expresión de genes inducidos por interferón [34, 35]. Por otra parte, estudios recientes en

nuestro laboratorio han demostrado que E1B 55K participa en la regulación de la expresión de promotores virales en el contexto de la infección. Sin embargo, no se conoce si la regulación de la expresión de estos genes implica la unión directa de la proteína al DNA. Resultados obtenidos en el laboratorio indican E1B 55K impacta en la morfología, organización y composición de los centros de replicación viral en el núcleo de las células infectadas. Se ha observado que las mutaciones que simulan el estado no fosforilado de la proteína afectan la funcionalidad de la proteína resultando en una baja eficiencia la producción de progenie, así como la degradación de p53 [69]. Por otro lado, las mutaciones que simulan el estado trifosforilado de la proteína traen consigo un aumento en la producción de progenie viral. Separación de fases. Además, E1B 55K está involucrada en el procesamiento selectivo de los RNA virales en el núcleo celular. Se ha demostrado que la presencia de la proteína es necesaria para la acumulación de mRNA en el citoplasma así como para la entrada eficiente a la fase tardía de la infección [36, 37]. El análisis de la secuencia primaria de E1B 55K indica que la proteína presenta un dominio de unión a RNA (RNP). Los resultados obtenidos en el laboratorio indican que E1B 55K es capaz de interactuar con RNA *in vitro*, así como en el contexto de la infección en células humanas no transformadas y que las sustituciones en el RNP afectan la interacción de la proteína con el RNA viral, la replicación de DNA viral y la producción de progenie sin afectar la actividad E3 ubiquitina ligasa ni la distribución intranuclear de la proteína [70].

La región amino terminal (1-113) de E1B 55K se predice como intrínsecamente desordenada (Fig. 33A). Su caracterización por CIDER sugiere que se presenta mayoritariamente compacta (Fig. 33B) a pesar de no presentar abundantes estructuras secundarias predichas (Fig. 33D). Al igual que la predicción de desorden, la modelación por homología presenta el extremo amino terminal desplegado (Fig. 33C). La presencia de una región desordenada en la proteína E1B 55K pudiera estar relacionada con la capacidad de la proteína de realizar varias funciones durante el ciclo de replicación viral e interactuar con múltiples blancos moleculares.

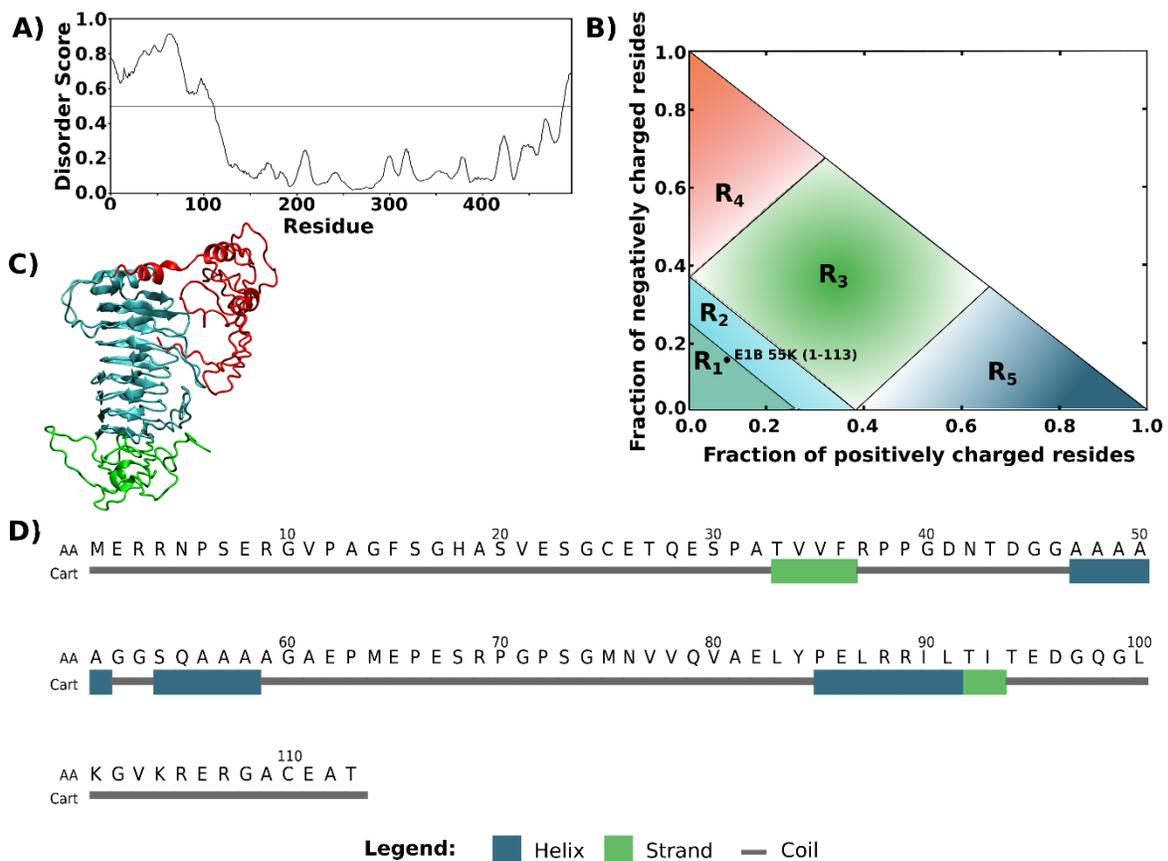


Figura 33. Caracterización de las IDRs presentes en E1B 55K. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de la IDR por CIDER. R1: Secuencia que típicamente son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Modelo de la proteína obtenido con I-TASER (C-score=-0.85) En rojo se representa el extremo amino terminal (1-146), en cyan la región central (147-385) y en verde la región carboxilo terminal (386-496) **D)** Predicción de estructura secundaria en la IDR. La línea gris indica ausencia de estructura secundaria, los cuadros verdes hojas beta.

Con el fin de analizar si existían interacciones entre las diferentes regiones predichas en la proteína se llevó a cabo un análisis coevolutivo. Se observó que la mayoría de las evidencias coevolutivas se observaron entre aminoácidos cercanos en secuencia y que las pocas interacciones observadas entre regiones distantes presentan bajo nivel de significación. No se observaron evidencias de coevolución

entre aminoácidos presentes en diferentes regiones de E1B 55K (Fig. 34A). Este resultado junto con la predicción de desorden sugiere que E1B 55K puede estar constituida por tres dominios independientes. Cabe señalar que la ausencia de evidencias coevolutivas entre la región central y carboxilo terminal no implica la ausencia de interacción entre estos dominios en la proteína, debido a que este análisis se realiza a partir de la secuencia primaria de la proteína. Los mayores scores de coevolución se observaron entre aminoácidos que se encuentran en la región central de la proteína (Fig. 34B), lo que reafirma la idea de que esta región se encuentra más estructurada que el resto de la proteína.

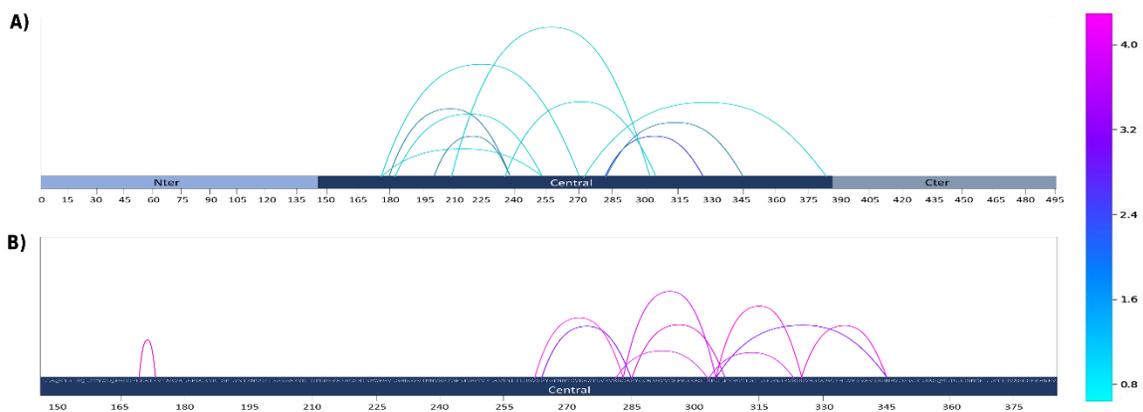


Figura 34. Coevolución entre las regiones de E1B 55K. El color de las líneas indica el score de coevolución A) Las líneas conectan los aminoácidos que coevolucionan y presentan una distancia en secuencia superior a los 40 aminoácidos. B) Las líneas conectan los aminoácidos que coevolucionan con los mayores scores.

Con el fin de analizar la distribución en la estructura tridimensional de los aminoácidos que coevolucionan con mayor score, estos se representaron en los modelos obtenidos en I-Tasser. Se observó que los aminoácidos que coevolucionan con mayor probabilidad son aminoácidos que se encuentran aproximadamente a 20 aminoácidos de distancia en secuencia y que interaccionan entre sí en el modelo de E1B 55K (Fig. 35). El análisis coevolutivo muestra que existe fuerte coevolución entre las fenilalaninas 264, 285 y 307, así como entre las asparaginas 325 y 345 las cuales pueden estar implicadas en la estabilización de la estructura solenoidal de la proteína. Escaleras similares de fenilalaninas se observaron estabilizando la estructura de la proteína LH3 de adenovirus tipo I de serpiente, proteína con mayor

identidad con E1B 55K [71]. Estos resultados junto a la modelación por homología refuerzan la idea que la región central de E1B 55 se pliega en un solenoide beta.

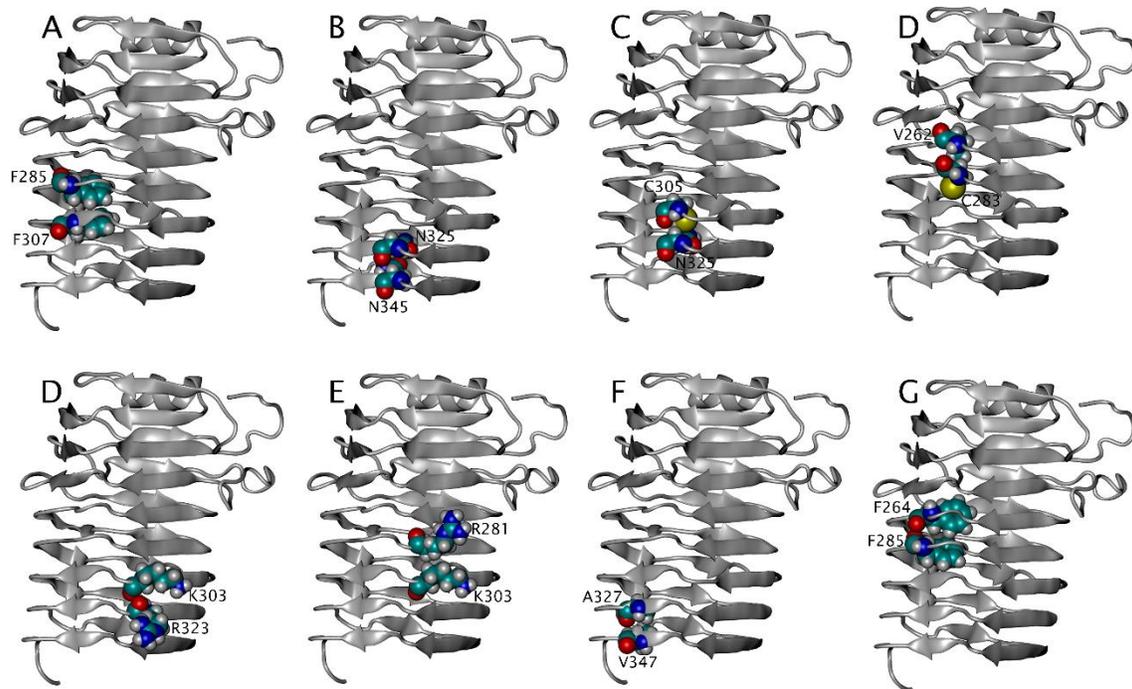


Figura 35. Representación de los pares de aminoácidos con la mayor probabilidad de coevolución. Modelo tridimensional obtenido por I-Tasser con la representación en VDW de los aminoácidos que co-evolucionan con mayor probabilidad. A) F285-F307. B) N325-N345. C) C305-N325. D) V262-C283. E) K303-R323. F) R281-K303. G) A327-V347. H) F264-F285.

IVa2:

La proteína viral IVa2 está formada por 449 residuos y se expresa durante la transición de la etapa temprana a la etapa tardía del ciclo de replicación viral de AdH. IVa2 se une como dímero al promotor de ML, regulando así la transcripción de sus genes. Los residuos del D50 a la S100, al igual que de la T359 a la L449 son necesarios para la interacción de la proteína con el genoma viral [44, 72]. Estos residuos deben estar implicados en la dimerización de la proteína y/o en la unión directa al genoma. El extremo amino terminal (80 residuos) se predice como intrínsecamente desordenado (Fig. 36A) y la predicción de estructura secundaria sugiere que en esta región solo se forma una hélice de tres aminoácidos (P22-Q24)

(Fig. 36D). Por su parte, la caracterización por CIDER en cuanto a la distribución de las cargas surge que el comportamiento del amino terminal depende de factores externos como la concentración salina, interacciones en *cis* o unión de ligandos (Fig. 36B). Al analizar la secuencia de los primeros 80 residuos de IVa2 podemos observar que esta región presenta una cantidad elevada de residuos positivos, lo que le permitiría la interacción con las cargas negativas del DNA (Fig. 36D). La predicción de la estructura secundaria del resto de la proteína [72], así como la modelación de la estructura por homología (Fig. 36C) suponen la formación de dos hélices anfipáticas en las regiones imprescindibles para la unión con DNA (D79-L98 y M413-L432), siendo estas las que pueden interactuar directamente con el genoma viral.

También se conoce que IVa2 interactúa directamente con 52/55K [60] y que se une de manera serotipo específica a las secuencias de empaquetamiento del genoma viral [47, 73]. IVa2 es una proteína esencial para el empaquetamiento, en su ausencia solo se forma un bajo número de cápsidas vacías [74]. IVa2 también interactúa con la proteína viral de unión a DNA 22K, formando un complejo que se une a las secuencias de empaquetamiento [52, 74]. Se propone que la función de IVa2 durante la encapsidación es actuar como la ATPasa del Motor de Vértice Único para la encapsidación de DNA de doble cadena [45]. La presencia de este mecanismo de encapsidación se ha observado en Herpes virus y Phagos, sin embargo, no existen evidencias contundentes que permitan afirmar que este es el mecanismo de encapsidación empleado por Adenovirus. Aunque si se ha demostrado que IVa2 es capaz de unir ATP ya que posee dos regiones por las cuales ocurre dicha interacción: Walker A (G181-S188) y Walker B (I275-E281) [50]. Además, se demostró que *in vitro* IVa2 en presencia de la proteína viral 33K presenta actividad ATPasa [45]. Las regiones de unión a ATP se encuentran en la región predicha como plegada lo cual concuerda con el supuesto que las enzimas generalmente son proteínas bien plegadas. El caso de IVa2 es muy interesante pues se observa en una misma proteína como el hecho de que una proteína este o no plegada está relacionada con su función y el tipo y número de interacciones, pues la región desordenada aparentemente está implicada en la unión a DNA, mientras que la plegada lleva a cabo la hidrólisis de moléculas de ATP.

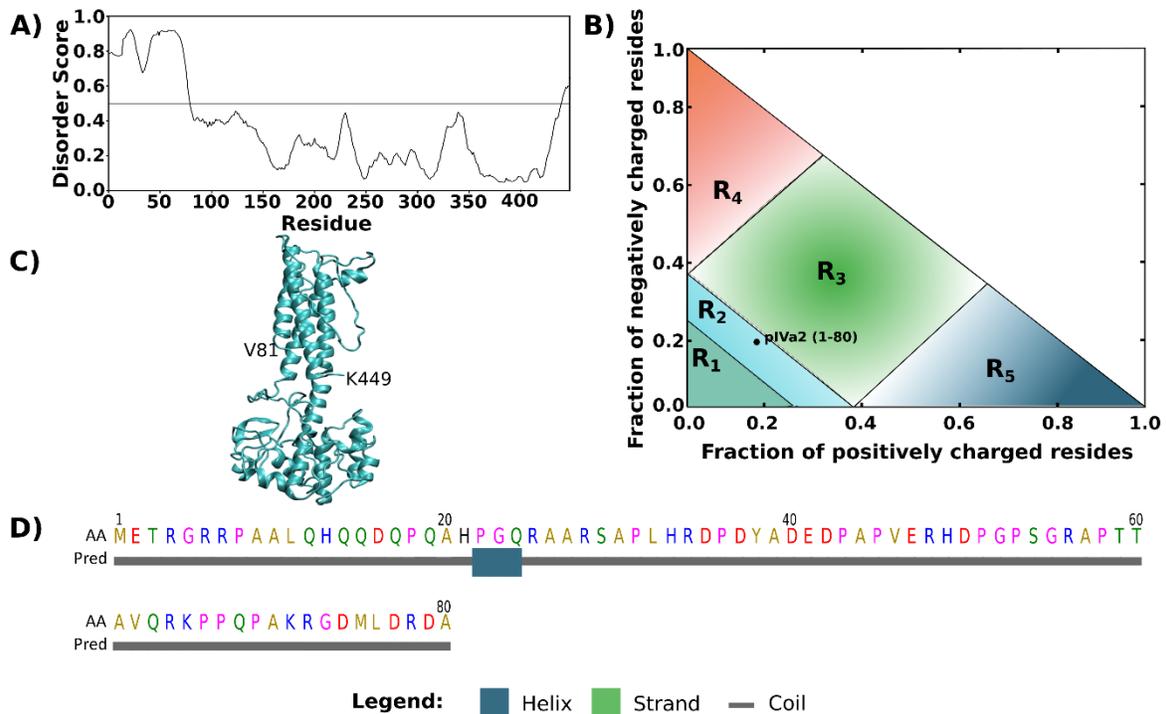


Figura 36. Caracterización de las IDRs presentes en IVa2. **A)** Promedio de la predicción de desorden con los predictores PONDR VLXT, PONDR XL1_XT, PONDR VL3-BA, PONDR VSL2, IUPred y DISOPRED. **B)** Caracterización de la IDR por CIDER. R1: Secuencia que típicamente son globulares. R2: Secuencias cuyo comportamiento depende de factores externos como concentración salina, interacciones en *cis*, unión de ligandos. R3: Secuencias cuya estructura va a depender de la distribución de cargas. R4: Secuencias con estructura extendida debido a la repulsión polianiónica. R5: Secuencias con estructura extendida debido a la repulsión policatiónica. **C)** Modelo de la región V81-K449 obtenido con I-TASER (C-score=-2.79) **D)** Predicción de estructura secundaria en la IDR. La línea gris indica ausencia de estructura secundaria, los cuadros azules hélices y los cuadros verdes hojas beta.

VI. Conclusiones

- El proteoma de adenovirus presenta elevados niveles de desorden estructural, 20 de las 34 proteínas analizadas presentaron regiones intrínsecamente desordenadas.
- En la cápside viral de adenovirus humano el porcentaje de IDRs en las secuencias está relacionado con la proximidad de las proteínas al DNA y el tipo de interacciones que establece cada proteína.

- Las proteínas que participan en la replicación del DNA, la transcripción y el procesamiento de los mRNA presentaron mayores porcentajes de aminoácidos desordenados que las proteínas virales con actividades enzimáticas, así como las proteínas encargadas de regular el tráfico vesicular e inhibir la respuesta inmune de las células.

VII. Referencias Bibliográficas

1. Fields, B., D. Knipe, and P. Howley, *Fields virology 6th ed.* Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health, 2013: p. 1704-1731.
2. Davison, A.J., M. Benkő, and B. Harrach, *Genetic content and evolution of adenoviruses.* Journal of General Virology, 2003. **84**(11): p. 2895-2908.
3. Lichtenstein, D.L. and W.S. Wold, *Experimental infections of humans with wild-type adenoviruses and with replication-competent adenovirus vectors: replication, safety, and transmission.* Cancer gene therapy, 2004. **11**(12): p. 819-829.
4. Parthasarathy, A., *Textbook of pediatric infectious diseases.* 2013: JP Medical Ltd.
5. Lion, T., *Adenovirus infections in immunocompetent and immunocompromised patients.* Clinical microbiology reviews, 2014. **27**(3): p. 441-462.
6. Weinberg, J. and K. Spindler, *Adenoviruses.* 2014.
7. Nemerow, G., et al., *Insights into adenovirus host cell interactions from structural studies.* Virology, 2009. **384**(2): p. 380-388.
8. Dunker, A.K., et al., *Intrinsic protein disorder in complete genomes.* Genome Informatics, 2000. **11**: p. 161-171.
9. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nature reviews Molecular cell biology, 2005. **6**(3): p. 197.
10. Uversky, V.N., *Intrinsically disordered proteins from A to Z.* The international journal of biochemistry & cell biology, 2011. **43**(8): p. 1090-1103.
11. Dunker, A.K., et al., *Intrinsic disorder and protein function.* Biochemistry, 2002. **41**(21): p. 6573-6582.
12. Dunker, A.K., et al., *Function and structure of inherently disordered proteins.* Current opinion in structural biology, 2008. **18**(6): p. 756-764.
13. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins: structure, function, and bioinformatics, 2000. **41**(3): p. 415-427.
14. Romero, P., et al. *Identifying disordered regions in proteins from amino acid sequence.* in *Proceedings of International Conference on Neural Networks (ICNN'97).* 1997. IEEE.
15. He, B., et al., *Predicting intrinsic disorder in proteins: an overview.* Cell research, 2009. **19**(8): p. 929.
16. Li, X., et al., *Predicting protein disorder for N-, C-and internal regions.* Genome Informatics, 1999. **10**: p. 30-40.
17. Obradovic, Z., et al., *Predicting intrinsic disorder from amino acid sequence.* Proteins: Structure, Function, and Bioinformatics, 2003. **53**(S6): p. 566-572.
18. Dosztányi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.* Bioinformatics, 2005. **21**(16): p. 3433-3434.
19. Jones, D.T. and D. Cozzetto, *DISOPRED3: precise disordered region predictions with annotated protein-binding activity.* Bioinformatics, 2014. **31**(6): p. 857-863.

20. Xue, B., A.K. Dunker, and V.N. Uversky, *Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life*. Journal of Biomolecular Structure and Dynamics, 2012. **30**(2): p. 137-149.
21. Singh, A., et al., *Deciphering the dark proteome of Chikungunya virus*. Scientific reports, 2018. **8**(1): p. 5822.
22. Jensen, M.R., et al., *Intrinsic disorder in measles virus nucleocapsids*. Proceedings of the National Academy of Sciences, 2011. **108**(24): p. 9839-9844.
23. Goh, G.K.-M., A.K. Dunker, and V.N. Uversky, *Protein intrinsic disorder toolbox for comparative analysis of viral proteins*. BMC genomics, 2008. **9**(2): p. S4.
24. Kamisetty, H., S. Ovchinnikov, and D. Baker, *Assessing the utility of coevolution-based residue-residue contact predictions in a sequence-and structure-rich era*. Proceedings of the National Academy of Sciences, 2013. **110**(39): p. 15674-15679.
25. Ferreon, J.C., et al., *Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein*. Proceedings of the National Academy of Sciences, 2009. **106**(32): p. 13260-13265.
26. Cuconati, A. and E. White, *Viral homologs of BCL-2: role of apoptosis in the regulation of virus infection*. Genes & development, 2002. **16**(19): p. 2465-2478.
27. Mun, K., *Study on the alternative splicing in the human adenovirus serotype 5 E1B region*. 2014.
28. Rao, L., et al., *The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bcl-2 proteins*. Proceedings of the National Academy of Sciences, 1992. **89**(16): p. 7742-7746.
29. Harada, J.N., et al., *Analysis of the adenovirus E1B-55K-anchored proteome reveals its link to ubiquitination machinery*. Journal of virology, 2002. **76**(18): p. 9194-9206.
30. Stracker, T.H., C.T. Carson, and M.D. Weitzman, *Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex*. Nature, 2002. **418**(6895): p. 348-352.
31. Fu, Y.R., et al., *Comparison of protein expression during wild-type, and E1B-55k-deletion, adenovirus infection using quantitative time-course proteomics*. Journal of General Virology, 2017. **98**(6): p. 1377-1388.
32. Schreiner, S., et al., *Proteasome-dependent degradation of Daxx by the viral E1B-55K protein in human adenovirus-infected cells*. Journal of virology, 2010. **84**(14): p. 7029-7038.
33. Martin, M.E. and A.J. Berk, *Adenovirus E1B 55K represses p53 activation in vitro*. Journal of virology, 1998. **72**(4): p. 3146-3154.
34. Chahal, J.S., J. Qi, and S. Flint, *The human adenovirus type 5 E1B 55 kDa protein obstructs inhibition of viral replication by type I interferon in normal human cells*. PLoS pathogens, 2012. **8**(8): p. e1002853.
35. Chahal, J.S., et al., *The repression domain of the E1B 55 kDa protein participates in countering interferon-induced inhibition of adenovirus replication*. Journal of virology, 2013: p. JVI. 03387-12.
36. Gonzalez, R.A. and S. Flint, *Effects of mutations in the adenoviral E1B 55-kilodalton protein coding sequence on viral late mRNA metabolism*. Journal of virology, 2002. **76**(9): p. 4507-4519.
37. Gonzalez, R., et al., *Adenovirus E1B 55-kilodalton protein is required for both regulation of mRNA export and efficient entry into the late phase of infection in normal human fibroblasts*. Journal of virology, 2006. **80**(2): p. 964-974.
38. Blackford, A.N. and R.J. Grand, *Adenovirus E1B 55-kilodalton protein: multiple roles in viral infection and cell transformation*. Journal of virology, 2009. **83**(9): p. 4000-4012.
39. Sieber, T., et al., *Intrinsic disorder in the common N-terminus of human adenovirus 5 E1B-55K and its related E1BN proteins indicated by studies on E1B-93R*. Virology, 2011. **418**(2): p. 133-143.

40. Kanellopoulos, P.N., et al., *Conformational Change of the Adenovirus DNA-Binding Protein Induced by Soaking Crystals with K3UO2F5 Solutions*. Acta Crystallographica Section D: Biological Crystallography, 1996. **52**(5): p. 942-945.
41. Sapunar, J., et al., *Adenovirus 36 seropositivity is related to obesity risk, glycemic control, and leptin levels in Chilean subjects*. International Journal of Obesity, 2019: p. 1.
42. Cheng, C.Y., et al., *Role of E1B55K in E4orf6/E1B55K E3 ligase complexes formed by different human adenovirus serotypes*. Journal of virology, 2013. **87**(11): p. 6232-6245.
43. Cheng, C.Y., et al., *The E4orf6/E1B55K E3 ubiquitin ligase complexes of human adenoviruses exhibit heterogeneity in composition and substrate specificity*. Journal of virology, 2011. **85**(2): p. 765-775.
44. Lutz, P., et al., *Nucleoplasmic and nucleolar distribution of the adenovirus IVa2 gene product*. Journal of virology, 1996. **70**(6): p. 3449-3460.
45. Ahi, Y.S., et al., *Adenoviral L4 33K forms ring-like oligomers and stimulates ATPase activity of IVa2: implications in viral genome packaging*. Frontiers in microbiology, 2015. **6**: p. 318.
46. Zhang, W. and M.J. Imperiale, *Requirement of the adenovirus IVa2 protein for virus assembly*. Journal of virology, 2003. **77**(6): p. 3586-3594.
47. Zhang, W., et al., *Role for the adenovirus IVa2 protein in packaging of viral DNA*. Journal of virology, 2001. **75**(21): p. 10446-10454.
48. Perez-Romero, P., K.E. Gustin, and M.J. Imperiale, *Dependence of the encapsidation function of the adenovirus L1 52/55-kilodalton protein on its ability to bind the packaging sequence*. Journal of virology, 2006. **80**(4): p. 1965-1971.
49. Zubieta, C., et al., *The structure of the human adenovirus 2 penton*. Molecular cell, 2005. **17**(1): p. 121-135.
50. Wohl, B.P. and P. Hearing, *Role for the L1-52/55K protein in the serotype specificity of adenovirus DNA packaging*. Journal of virology, 2008. **82**(10): p. 5089-5092.
51. Koyuncu, O.Ö. and T. Dobner, *Arginine methylation of human adenovirus type 5 L4 100-kilodalton protein is required for efficient virus production*. Journal of virology, 2009. **83**(10): p. 4778-4790.
52. Ewing, S.G., et al., *Ternary complex formation on the adenovirus packaging sequence by the IVa2 and L4 22-kilodalton proteins*. Journal of virology, 2007. **81**(22): p. 12450-12457.
53. Boeynaems, S., et al., *Protein phase separation: a new phase in cell biology*. Trends in cell biology, 2018. **28**(6): p. 420-435.
54. Boyd, J.M., et al., *Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set of cellular proteins*. Cell, 1994. **79**(2): p. 341-351.
55. White, E., *Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus*. Oncogene, 2001. **20**(54).
56. Bremner, K.H., et al., *Adenovirus transport via direct interaction of cytoplasmic dynein with the viral capsid hexon subunit*. Cell host & microbe, 2009. **6**(6): p. 523-535.
57. Crawford-Miksza, L. and D.P. Schnurr, *Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues*. Journal of virology, 1996. **70**(3): p. 1836-1844.
58. Khare, R., et al., *Identification of adenovirus serotype 5 hexon regions that interact with scavenger receptors*. Journal of virology, 2012. **86**(4): p. 2293-2301.
59. Hasson, T.B., D. Ornelles, and T. Shenk, *Adenovirus L1 52-and 55-kilodalton proteins are present within assembling virions and colocalize with nuclear structures distinct from replication centers*. Journal of virology, 1992. **66**(10): p. 6133-6142.
60. Perez-Romero, P., et al., *Analysis of the interaction of the adenovirus L1 52/55-kilodalton and IVa2 proteins with the packaging sequence in vivo and in vitro*. Journal of virology, 2005. **79**(4): p. 2366-2374.

61. Gustin, K.E., P. Lutz, and M.J. Imperiale, *Interaction of the adenovirus L1 52/55-kilodalton protein with the IVa2 gene product during infection*. Journal of virology, 1996. **70**(9): p. 6463-6467.
62. Hayes, B., et al., *The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species*. Journal of virology, 1990. **64**(6): p. 2732-2742.
63. Hong, S.S., et al., *The 100K-chaperone protein from adenovirus serotype 2 (Subgroup C) assists in trimerization and nuclear localization of hexons from subgroups C and B adenoviruses*. Journal of molecular biology, 2005. **352**(1): p. 125-138.
64. Xi, Q., R. Cuesta, and R.J. Schneider, *Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting*. Genes & development, 2004. **18**(16): p. 1997-2009.
65. Xi, Q., R. Cuesta, and R.J. Schneider, *Regulation of translation by ribosome shunting through phosphotyrosine-dependent coupling of adenovirus protein 100k to viral mRNAs*. Journal of virology, 2005. **79**(9): p. 5676-5683.
66. Andrade, F., L.A. Casciola-Rosen, and A. Rosen, *A novel domain in adenovirus L4-100K is required for stable binding and efficient inhibition of human granzyme B: possible interaction with a species-specific exosite*. Molecular and cellular biology, 2003. **23**(17): p. 6315-6326.
67. Querido, E., et al., *Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex*. Genes & development, 2001. **15**(23): p. 3104-3117.
68. Orazio, N.I., et al., *The adenovirus E1b55K/E4orf6 complex induces degradation of the Bloom helicase during infection*. Journal of virology, 2011. **85**(4): p. 1887-1892.
69. Schwartz, R.A., et al., *Distinct requirements of adenovirus E1b55K protein for degradation of cellular substrates*. Journal of virology, 2008. **82**(18): p. 9043-9055.
70. Tejera, B., et al., *The human adenovirus type 5 E1B 55kDa protein interacts with RNA promoting timely DNA replication and viral late mRNA metabolism*. PloS one, 2019. **14**(4): p. e0214882.
71. Menéndez-Conejero, R., et al., *Structure of a reptilian adenovirus reveals a phage tailspike fold stabilizing a vertebrate virus capsid*. Structure, 2017. **25**(10): p. 1562-1573. e5.
72. Lutz, P. and C. Keding, *Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter*. Journal of virology, 1996. **70**(3): p. 1396-1405.
73. Zhang, W. and M.J. Imperiale, *Interaction of the adenovirus IVa2 protein with viral packaging sequences*. Journal of virology, 2000. **74**(6): p. 2687-2693.
74. Ostapchuk, P., M. Almond, and P. Hearing, *Characterization of empty adenovirus particles assembled in the absence of a functional adenovirus IVa2 protein*. Journal of virology, 2011. **85**(11): p. 5524-5531.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Coordinación de Programas Educativos

Posgrado en Ciencias



DR. VICTOR BARBA LÓPEZ
COORDINADOR DEL POSGRADO EN CIENCIAS
PRESENTE

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada “Análisis del Desorden Estructural en el Proteoma de Adenovirus Humano” que presenta el alumno Reinier Felipe Cárdenas Mena (10009537) para obtener el título de **Maestro en Ciencias**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Enrique Rudiño Piñera IBT-UNAM	Aprobado	
Dra. Carmen Nina Pastor Colón CIDC-UAEM	Aprobado	
Dra. Lina Andrea Rivillas Acevedo CIDC-UAEM	Aprobado	Lina Rivillas
Dr. Ramón A. Gonzalez García-Conde CIDC-UAEM	APROBADO	
Dr. Carlos Amero Tello CIQ-UAEM	Aprobado	