

CARACTERIZACIÓN QUÍMICA IN-SILICO DE INTERACCIONES PROTEÍNA-PROTEÍNA TIPO DOMINIO-MOTIVO PARA EL ESTUDIO DE LAS ACTIVIDADES PATÓGENO-HOSPEDERO EN MODELOS VIRALES.

Materón, L¹, Becerra, A.², Moreno, P².

¹Departamento de Química, Universidad del Valle, Cali, Colombia.

²Escuela de Ingeniería de Sistemas, Universidad del Valle, Cali, Colombia

Corresponsal del artículo: lina.materon@correounivalle.edu.co

RESUMEN.

ANTECEDENTES: Comprender las interacciones entre patógenos y hospederos en las enfermedades de carácter infeccioso son todavía un permitiría la formulación de tratamientos, vacunas, métodos de control y prevención de la enfermedad. Se han catalogado ciertos tipos de interacción, motivos de interacción y motivos lineales. Teniendo en cuenta, que los motivos lineales son cortos y pueden evolucionar rápidamente en proteínas virales. El objetivo de este trabajo es desarrollar, una metodología de caracterización química in silico que permita determinar interacciones tipo proteína-proteína tipo dominio-motivo para el estudio de interacciones hospedero-patógeno.

RESULTADOS: Se desarrolló una metodología para encontrar características estructurales comunes entre interacciones virus-hospedero. Se aplicó la metodología con un conjunto de interacciones entre una proteína viral y una proteína del hospedero, para las que existe el modelo estructural en la base de datos PDB, arrojando los siguientes características: 1) las estructuras que interactúan son “ α -helix”, “ β -sheets” y “loops”, que se pueden verificar por medio de análisis como Área Superficial Accesible de Solvente (SASA), 2) cuando hay presencia de una “ β -strand” esta interacciona con una “ β -strand” formando una “ β -sheet” entre el hospedero y el patógeno, 3) un “loop” en cualquiera de las dos proteínas va enlazado a la interacción con un “ α -helix”, 4) la SASA permite determinar los tipos de enlace encontrados, siendo los que tienen menor valor los que se encuentran más anclados en la interacción, 5) por medio de la medición de distancia del contacto de cada aminoácido del péptido viral con la proteína hospedera se pudo determinar los tipos de enlace químico que favorecen dichas interacciones, 6) los grupos ionizables son determinantes para la estabilidad de la interacción, 7) el pH óptimo para el mínimo de estabilización de cada interacción difiere del valor de punto isoelectrico. Además, se realizó un análisis estadístico del contenido de aminoácidos en las secuencias de las zonas de interacción tanto para hospederos como patógenos virales.

CONCLUSIÓN: Se desarrolló una metodología de caracterización In-Silico de interacciones proteicas hospedero-huésped, por medio de descriptores físico y bioquímicos, con la cual se pueden entender las interacciones proteicas tipo dominio-motivo, desde el punto de vista químico-estructural. Estas ocurren en áreas con alta accesibilidad en las proteínas, sobre todo en su superficie periférica. Estas interacciones de carácter superficiales se da gracias a la presencia de ciertos tipos de enlace como lo son: enlace de hidrógeno, enlaces de puentes salinos, enlaces de Van der Waals. El anclaje que se da para genera las interacciones entre hospedero y huésped de tipo viral, se da gracias a residuos con baja accesibilidad de solvente. El pH es un factor determinante para este tipo de interacciones debido a que dispone todo el ambiente para que la interacción hospedero-patógeno se genere de forma estable. La presencia de ciertos aminoácidos como la fenilalanina, el ácido glutámico y otros, definen la naturaleza química de las secuencias, las estructuras secundarias, en consecuencia, de las interacciones.

PALABRAS CLAVE: Interacciones Hospedero-Patógeno, Virus, PDB, PANDAS, REGULEX, PYMOL, SASA, Propka.

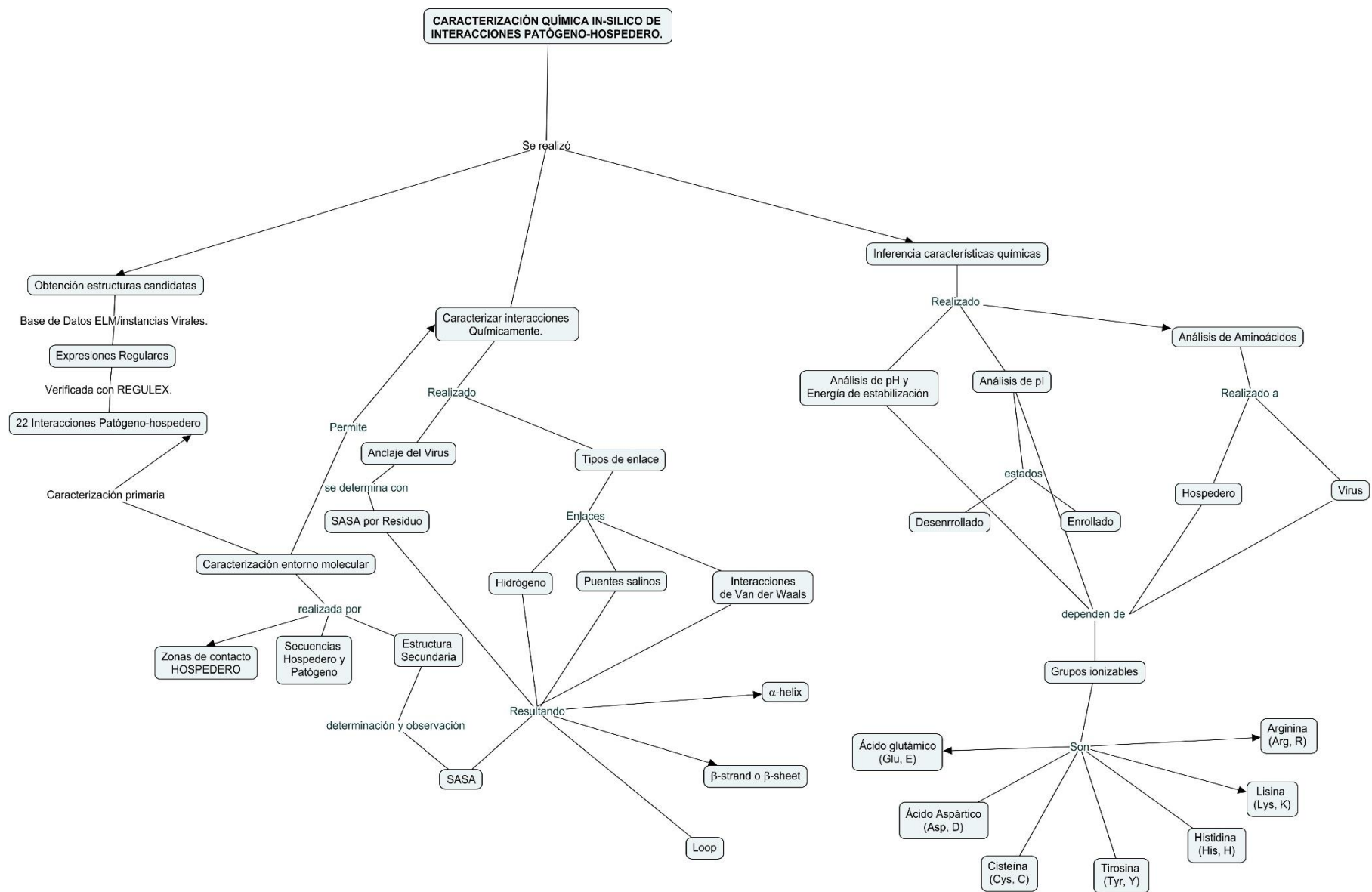


Figura 1. Diagrama resumen de Metodología y resultados obtenidos en la caracterización química in-silico de interacciones patógeno-hospedero.

1. INTRODUCCIÓN.

Las características de las interacciones entre proteínas dentro del mismo hospedero (interacciones endógenas) han sido ampliamente estudiadas. Sin embargo, las interacciones entre hospedero-huésped (interacciones exógenas) aún son objeto de estudio, para ello se han implementado estrategias de carácter experimental y de carácter computacional para hacer el proceso más ágil, sin embargo, siguen siendo un tema de gran interés para la investigación en ciencia actual.¹

Las interacciones hospedero-huésped se han convertido en objetos de total interés dentro de la investigación de sistemas biológicos debido a su importancia en la identificación de los mecanismos subyacentes a la enfermedad.² Los logros en la identificación de estos han permitido que se den avances en el desarrollo de vacunas, fármacos antivirales, diseño de terapias y otros medicamentos para el control, manejo y prevención de las enfermedades.^{3,4} Sin embargo, la escasez de datos experimentales y teóricos sobre las interacciones de los patógenos (huésped) en los hospederos (ser humano) han dificultado la investigación en este campo.

La necesidad de conocer este tipo de mecanismos de interacción se da primordialmente, por la idea de intentar solucionar el problema que ocasionan, enfermedades de tipo infeccioso. La formulación de tratamientos, vacunas u otros métodos de control de una enfermedad generada por patógenos, se basan primero en el descubrimiento del mecanismo de interacción del patógeno con el hospedero. Siguiendo este orden de ideas, se ha identificado que las interacciones motivo-dominio, son las principales agentes causales en la génesis de las enfermedades infecciosas.⁵ Los motivos lineales, son pequeños segmentos conservados, de estructuras secundarias de las proteínas, compuestos por aminoácidos o secuencias de nucleótidos.⁶⁻⁹

Ahora bien, Los virus como los patógenos de los que más de los que es necesario en entender el mecanismo de interacción con las proteínas hospederas sean de humanos o de otros tipos de especies, tienden a copiar los mecanismos celulares intrínsecos responsables de inhibir, señalar y prevenir enfermedades de tipo infeccioso.¹⁰ Estos al copiar las mismas estructuras de las interacciones endógenas, (entre proteínas del mismo hospedero) generan así interacciones exógenas muy similares a las primeras; y traspasando los puntos de control del hospedero y generando así, enfermedad. Sin embargo, hay algo que permite encontrarlos en las secuencias de interacción, los motivos lineales.^{7,11,12}

Los motivos se relacionan directamente con los dominios, que son estructuras terciarias y modulares de las proteínas, haciendo que esta relación motivo-dominio entre huésped y patógeno sea eficaz.^{1,13,14} Es por ello, que la caracterización del comportamiento químico de estas interacciones permite un entendimiento más fino del porque se dan estas. Adicionalmente, la geometría del lugar de reconocimiento y las reacciones involucradas favorecen los mecanismos de producción de enfermedad, en el momento el que el patógeno interacciona con el hospedero. Como objetivos adicionales se pretenden encontrar patrones a nivel

estructural dentro de los sitios de reconocimiento en las proteínas y caracterizar la estructura del motivo y su entorno.^{5,15,16}

La naturaleza química de una interacción proteína-proteína no sólo depende de las estructuras primarias y secundarias encontradas; esta depende de los tipos de enlace que presenta la interacción que, en general, son de tipo enlace de hidrógeno, interacción electrostática, interacción de Van der Waals, entre otras.¹⁷⁻¹⁹ Estas interacciones por tipos de enlace son copiadas por los virus casi de manera idéntica.

Los ambientes donde se desarrollan las interacciones son también muy importantes para generar una estabilidad en esta, el pH es el factor determinante y el más variable del ambiente de interacción dentro de un sistema como lo es el cuerpo de un hospedero, debido a que condiciones como temperatura y presión realmente no son muy variables. Sin embargo, el factor pH según el lugar donde se dé la interacción varía, es por ello que se debe tener en cuenta en la estabilidad energética de la interacción, adicionalmente, el punto isoelectrico también ayuda a determinar cómo se da la interacción y como es el comportamiento tanto de la proteína hospedera como de la proteína viral.²⁰

En este documento se presenta la caracterización química vía computacional o In-Silico de veintidós interacciones proteína-proteína tipo dominio-motivo para el estudio de la actividad patógeno-hospedero en modelos virales. A través de caracterización de estructuras tridimensionales de veintidós interacciones hospedero-patógeno viral, las cuales fueron identificadas, caracterizadas, analizadas y de ellas se pudo inferir un tipo de características químicas en común

Todo esto se realizó haciendo uso de diversas herramientas bioinformáticas como algoritmos, software, bancos de datos y librerías de analítica de datos.

2. METODOLOGÍA.

Para la realización de la caracterización química de diversas interacciones patógeno-hospedero en modelos virales haciendo uso de un método computacional, el proceso fue dividido en tres etapas fundamentalmente; la primera etapa fue la caracterización primaria de las estructuras tridimensionales de las interacciones patógeno-hospedero y su entorno molecular. La segunda etapa, consistió en la identificación y caracterización de las interacciones candidatas teniendo en cuenta la forma de enlace, expresión regular y anclaje del patógeno en el hospedero. La tercera etapa, se basó en la inferencia de características químicas en común presentes en las interacciones previamente estudiadas y caracterizadas.

La primera etapa, caracterización primaria de las estructuras tridimensionales de las interacciones se realizó obteniendo las interacciones candidatas, las cuáles podrían ser muchas, pero el criterio de elección fue la búsqueda de motivos lineales y la presencia de estructura tridimensional de la interacción. El motivo lineal contribuye a la formación de la estructura terciaria, sin necesidad de experimentar cambios, además, que poseen funciones primarias en los dominios globulares de diversas proteínas. Estas son estructuras conservadas, dependen únicamente de su estructura primaria o de aminoácidos y no de su contexto de estructura terciaria.⁶ La base de datos ELM (Eukaryotic Linear Motif Resource)²¹ permite usar el criterio de motivo lineal como la herramienta para encontrar interacciones candidatas y proveer con gran exactitud el motivo interaccionante, además de la estructura tridimensional que representa la interacción. Estas últimas, pueden ser visualizadas como formatos PDB. Los formatos PDB, permite observar las interacciones patógeno hospedero por medio de una interfaz gráfica en un software.

ELM provee las interacciones candidatas y permite filtrarlas tanto por expresión regular como por la presencia de estructura tridimensional, ambos criterios que fueron de vital importancia para esta investigación, siendo este último el más importante, la presencia de una estructura tridimensional (ya modelada), este último será el criterio principal para determinar las interacciones candidatas. El proceso de filtrado se realizó haciendo uso de la librería de Análisis de Datos, PANDAS, escrita en lenguaje Python, que permitió determinar la ausencia o presencia de archivo PDB y así las candidatas analizables. Las interacciones candidatas patógeno-hospedero en modelos virales caracterizadas y filtradas terminaron correspondiendo a 22 (veintidós) estructuras. Las expresiones regulares proporcionadas por la base de datos fueron verificadas por medio de la herramienta virtual, “*regex*” (<https://jex.im/regex/#flags=&re=%5BPSAT%5D%5BQE%5DE>) la cual es una herramienta escrita en dos lenguajes de programación PEARL y Java Script que permiten determinar en un lenguaje simple, las expresiones regulares que poseen los motivos lineales.

Ahora bien, la caracterización primaria de las estructuras candidatas y su entorno molecular se realizó haciendo uso de una interfaz gráfica como lo fue el software PyMol²²⁻²⁴ que permite la interacción entre una consola que permite interaccionar con el ambiente de comandos en lenguaje Python y una interfaz gráfica que muestra todos los resultados y cambios que se van realizando.

Para la localización de las expresiones regulares previamente obtenidas en la base de datos y verificadas con REGULEX, se ejecuta el algoritmo instalado en el programa “*findseq.py*” el cual permite localizar en la secuencia de datos de la interacción señalar los sectores donde se encuentra la expresión regular y conjunto a esto se puede observar en la interfaz gráfica señalada las zonas donde se encuentra el motivo lineal. Este es señalado y delimitado tanto en secuencia como en la estructura tridimensional. Posteriormente, se determinan los

puntos de interacción del motivo lineal del patógeno con la proteína del hospedero, las zonas de contacto y luego se determina la superficie accesible de solvente (SASA) ejecutando el algoritmo, “*SASA.py*”, instalado, también en el software Pymol.

La segunda etapa, consistió en la identificación y caracterización de las interacciones candidatas teniendo en cuenta la forma de enlace, la expresión regular y el anclaje de los aminoácidos a la proteína del hospedero. Para ello, primero se identificaron primero las secuencias de interacción tanto del patógeno o virus y la secuencia donde se ejercía la interacción en el hospedero, señalándola en la interfaz gráfica de Pymol.²² Al determinar esta secuencia, se determina la estructura primaria de estas y se procedió a realizar el análisis de SASA por aminoácido. Este último, sólo se realiza a la estructura primaria del virus o patógeno; este permite determinar los sitios de anclaje del virus y así los puntos donde se debe hacer más hincapié para buscar conexiones y puntos de contacto.

En este orden de ideas, el hallar los puntos de anclaje y contacto más determinantes del virus permitió determinar aquellos aminoácidos que tenían mayor enlace con el hospedero y así determinar la naturaleza química de estos, entonces, ejecutando el algoritmo, instalado en PyMol, “*measurement*” determinó por medio de distancia intermolecular la naturaleza química de los enlaces formados entre el virus y el hospedero. Así también, la clasificación de estos enlaces en tres tipos, enlaces de Van der Waals, enlaces de hidrógeno y enlaces de puente salino.

Cabe aclarar que los algoritmos usados en la primera y segunda etapa de caracterización de las estructuras candidatas se encuentran ambos en PyMol y son proporcionados por las bases de datos del programa, en algunos casos, como el “*findseq.py*” fueron modificados debido a que no compilaban algunos datos.

La tercera etapa, estuvo dada por el estudio de los resultados obtenidos anteriormente en etapas uno y dos, para así generar una inferencia de características químicas comunes de las interacciones; sin embargo, se realizaron otros análisis para así encontrar más evidencia de la relación en común en estas. Para ello se ejecuta el algoritmo “*Propka*”, tanto en el software Pymol, como en el software VMD y en la herramienta en línea “PDB2PQR²⁵ server” (<http://apbs.poissonboltzmann.org/pdb2pqr>) la cual arrojaba al final de la ejecución un archivo “*propka*” el cual al final proporcionaba datos de punto isoelectrico, pH de estabilización y energía de estabilización.

Los archivos pKa, leídos en el software VMD, permiten visualizar gráficas de estabilización energética y la gráfica de punto isoelectrico, obtenida gracias a la optimización del algoritmo “*Propka*” que por medio de la simplificación de algunas interacciones permite obtener los datos previamente mencionados y muchos otros.²⁶ Por último, la determinación de características químicas se basó en un análisis descriptivo estadístico de la presencia o ausencia de los aminoácidos en las estructuras primarias tanto de virus como del hospedero, el análisis permitió por medio de observación básica determinar algunas características químicas importantes para estudiar a futuro otras interacciones virus-hospedero.

3. RESULTADOS

3.1. Clasificación de estructuras candidatas y sus expresiones regulares.

Las interacciones candidatas fueron seleccionadas a partir de la base de datos ELM, teniendo en cuenta dos criterios, la expresión regular y la presencia de una estructura tridimensional de la interacción, como se puede observar en la tabla 1. Las interacciones se filtraron usando la librería de data analítica PANDAS instalada en la terminal de Linux, teniendo en cuenta el criterio de presencia o ausencia de formato PDB, como se evidenciar en la tabla 1.

En la tabla 1, también se puede observar que las interacciones candidatas tienen hospederos diversas proteínas involucradas en diversos procesos y diversos tipos de virus. Adicionalmente, las expresiones regulares, que posteriormente corresponden a motivos lineales, encontradas en la base de datos difieren de los verificados y obtenidos por medio de la herramienta REGULEX, principalmente, a la que la obtención de una expresión regular responde a la probabilidad de encontrar el motivo lineal. Estos, actualmente, aún no son fácilmente predecibles en un contexto de experimentación tanto computacional como in vivo.

Los motivos lineales son elementos cortos y pequeños que se encuentran en grandes proteínas y operan en sitios de regulación, por lo general, se encuentran conservados y aunque intervienen en los procesos no hacen que su presencia genere un cambio significativo en la estructura terciaria tanto de la proteína como de la interacción; en bioinformática existen tres métodos para determinar los patrones de conservación, Los perfiles HMMs, Las redes neuronales artificiales y las expresiones regulares, asociadas a motivos lineales que son fáciles de encontrar.⁶

Las expresiones regulares de los motivos lineales permiten determinar fácilmente la posición de una de las dos partes interaccionantes, en los virus, estas se encuentran conservadas por el modelo de mimica de los virus a las interacciones proteína-proteína del hospedero, por tanto, se puede encontrar el punto de encuentro del virus fácilmente teniendo en cuenta la expresión regular.

Cabe resaltar, que las expresiones obtenidas por REGULEX, se refinaron y también se contrastaron con la literatura existente de cada una de las interacciones y así se obtuvieron con mayor exactitud y precisión para su posterior trabajo con ellas.

Tabla 1. Interacciones patógeno hospedero, candidatas a estudio de caracterización química con su correspondiente expresión regular.²⁷⁻⁵⁰

Complejo de interacción Patógeno-Hospedero	PDB	EXPRESIÓN REGULAR (ELM)	EXPRESIÓN REGULAR (REGULEX)
Calcineurina en complejo con dominio inhibidor de calcineurina de la proteína del African Swine Fever A238L	4F0Z	P.[*P][*P][V][*P]	P.I.IT
			L.V.
Complejo del dominio HAUSP/USP7 con el péptido EBNA1	1YY6	P.E[*P].S[*P]	E.S
14-3-3 Zeta/Fosfopeptido complejoado.	1QJB	R[DE][0,2][DEPG]([ST])([FWYLMV].[*PRIKGN]P) ([*PRIKGN].[2,4][VLMFWYP])	Motivo general: I.R.A
			Motivo Putativo: RS.PS.P
Estructura de Rotavirus NSP1 enlazado con IRF-3	5JER	[VILPF].[1,3]L.(S)	L.IS
			PL.IS
Complejo ALIX-AIP1 CON HIV-1 Dominio tardío YPLTSL	2R02	[LM]YP...[L][*P][*P][L]	LYP...L..L
Complejo ALIX-AIP1 CON HIV-1 Dominio tardío MAGI-1 PDZ-1 / E6CT	2R03	[LM]YP.[L]	[LM]YP.[L]
MAGI-1 PDZ1 enlazado con el péptido C-terminal	2KPL	...[ST].[ACVILF]\$	(ST).(VIL)
Sap97 PDZ2 enlazado con el Péptido C-Terminal	2I04	...[ST].[ACVILF]\$	(ST).(VIL)
Sap97 PDZ3 enlazado con el Péptido C-Terminal	2I0L	...[ST].[ACVILF]\$	(ST).(VIL)
Sap97 PDZ3 enlazado con el Péptido C-Terminal	2I0I	...[ST].[ACVILF]\$	(ST).(VIL)
hDLG/Sap97 PDZ2 en complejo con péptido HPV-	2QOS	...[ST].[ACVILF]\$	(ST).(VIL)
Retinoblastoma pocket en complejo con SV40 Large T Antigen	1GH6	([DEST])*(. {0,4}[L].C.E {1,4}[FLMIVAWPHY]. {0,8}([D EST])\$)	[L].C.[DE]
Proteína no estructural de Coronavirus MHV de tipo salvaje nps15	2GTH	([DEST])*(. {0,4}[L].C.E {1,4}[FLMIVAWPHY]. {0,8}([D EST])\$)	[L].C.[DE]
Retinoblastoma pocket enlazado con E7 motivo LxCxE	1GUX	([DEST])*(. {0,4}[L].C.E {1,4}[FLMIVAWPHY]. {0,8}([D EST])\$)	[L].C.[DE]
Dominio de Retinoblastoma Pocket en complejo	2R7G	...[LMV]...[LM][FY]D.	[L][H].[YF]
Proteína HIV-1 NEF en complejo con Dominio R96I	1EFN	P..P.[KR]	P..P
Estructura del complejo entre el dominio humano	1CYZ	P.Q..D	P.Q.[TSD]
Dominio Mouse Itch 3rd WW enlazado con Péptido derivado de la proteína 2A de la membrana latente	2J09	PP.Y	P.Q..D
Estructura de la glicoproteína de la envoltura del	2HG0	..(N)[*P][ST]..	N.S/T
Dominio C-Terminal de la subunidad PB2 de la polimerasa del virus de la influenza A en complejo	2JDQ	[KR][KR].[7,15][*DE]([K(RK)](RK))([*DE][KR])([KR][*DE])([*DE])	[KR][KR].[KR]
Importina Alfa de ratón en complejo con péptido	1Q1T	[*DE]([K(RK)](RK))[KR][KR][*DE]	[KR][KR].[KR]
Dominios similares a ubiquitina USP7 en	4WPH	K...K	K...K

3.2. Caracterización primaria de las estructuras candidatas por medio del software PyMOL.

3.2.1. Localización de motivo lineales y determinación puntos de interacción en hospedero y en patógeno.

PyMOL es un software escrito en lenguaje Python y Biopython, de licencia libre, que tiene entre sus principales características una interfaz gráfica intuitiva y fácilmente manejable para el usuario y que brinda una facilidad por medio de la ejecución en consola de diferentes algoritmos escritos en lenguaje Python ser ejecutados y así poder estudiar entornos moleculares además de interacciones como las presentadas en esta investigación.²³

La caracterización primaria de las estructuras candidatas se realizó por medio de su estructura tridimensional en formato PDB, en el software con interfaz gráfica Pymol, este a usar algoritmos como *Findseq.py* ejecutado en la consola y agregando la expresión regular verificada por REGULEX en un lenguaje más simple, se pudo determinar el punto de interacción de los virus y así dar

comienzo a la caracterización química de cada una de las veintidós interacciones.

Al correr el algoritmo, el programa permite determinar la zona donde se encuentra la expresión regular, como se puede observar en la figura 2, en la cual se presenta la expresión regular señalada en puntos rojos, al encontrar esta que pertenece a la zona del motivo lineal conservado del virus, y así determinar donde se encuentra este y con qué zona de la proteína hospedera se encuentra interaccionando y así en ambas, tanto en secuencia como en la estructura tridimensional se señalan.

La obtención de las secuencias de interacción del virus y de las zonas de interacción con la proteína hospedera, permitió determinar, por ejemplo, la forma de interacción, las estructuras secundarias que intervenían en cada uno de los procesos de interacción patógeno-hospedero; los tipos de enlace, la geometría de la zona, entre otros.

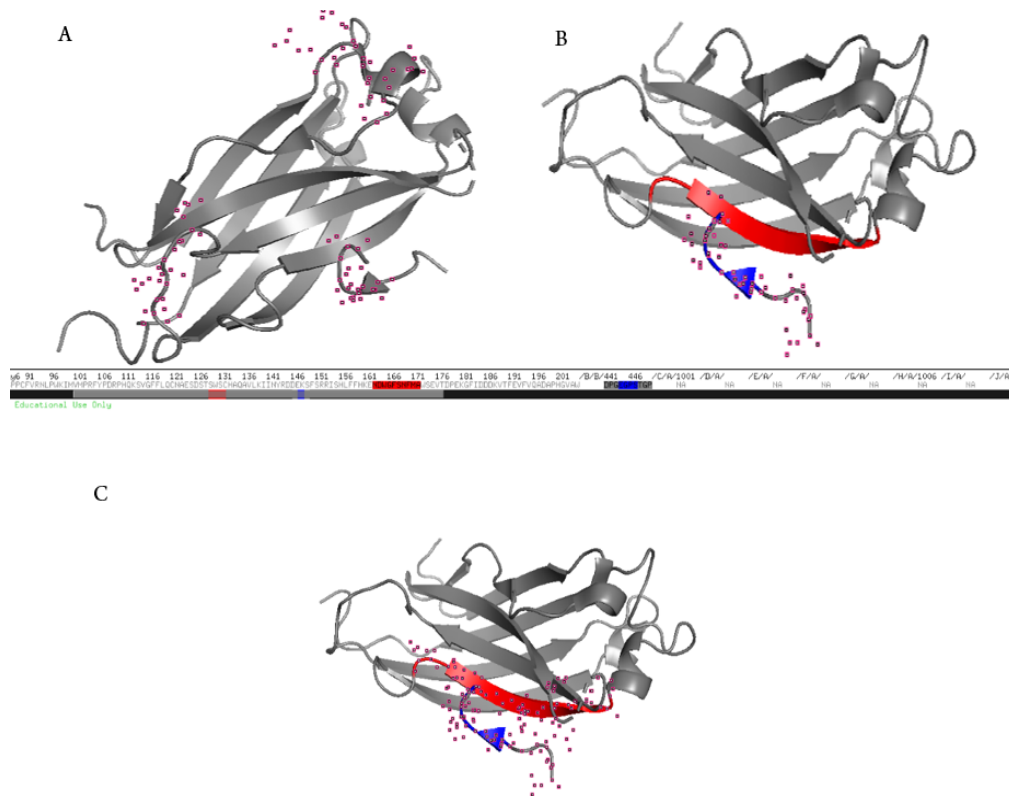


Figura 2. Estructura compleja HAUSP7 con el péptido EBNA1. PDB 1YY6. A. Estructura seleccionando los motivos de expresión regular encontrados. B. Selección de las zonas interaccionantes, virus (azul) proteína hospedera (rojo). C. Determinación de secuencias, en la parte superior de la consola.

3.2.2. Obtención de las estructuras primarias o secuencias de aminoácidos en las interacciones hospedero-patógeno.

Las estructuras primarias de las proteínas corresponden a la secuencia de aminoácidos que dan forma a las estructuras secundarias que son las formaciones que conocemos como hélices alfa, beta hojas y loops, que a su vez en conformación conectada hacen las grandes proteínas. De la estructura primaria o secuencia de aminoácidos primaria se originan todas las formaciones posteriores, esto es debido, a que el ambiente químico de los aminoácidos que conforman esta permite que se haga más sencilla la formación de una estructura secundaria sobre otra.

Para los casos estudiados, después de señalar por medio del uso del algoritmo *findseq.py* en la consola de Pymol, se encontró y señaló la zona donde se encontraba la sección de expresión regular y por consiguiente el motivo lineal que se encontraba en el virus, por tanto, según la teoría del mimicking, también debía tener la proteína que iba a generar la interacción endógena con el hospedero. Esta determinación permitió observar con ayuda de otra de las herramientas del programa la secuencia de aminoácidos que poseía la proteína viral en las 22 interacciones.

Por complementariedad, cercanía y geometría de interacción entre patógeno-hospedero, verificadas en la literatura, se pudo obtener la zona de interacción de la proteína viral-patógena con la

proteína del hospedero. La zona de interacción con la proteína del hospedero fue señalada y se pudo obtener a su vez, la secuencia de aminoácidos que ayudará a determinar cuál sería la estructura secundaria tanto del hospedero como del patógeno.

La tabla 2, presenta las secuencias de aminoácidos tanto de los patógenos virales donde se pueden observar los motivos lineales conservados, como las zonas de interacción de las proteínas hospederas con los patógenos virales. Estas secuencias probablemente darán explicación de las estructuras secundarias que participan en las mencionadas interacciones.

Según la literatura consultada, las zonas de interacción de la proteína hospedera hacen referencia a las zonas de interacción

donde se realizan las interacciones endógenas copiadas por los virus para generar enfermedad, en casos como los virus que tienen una señal de traslocación nuclear, estas son de vital importancia y se encuentran altamente conservadas. Tanto en el virus como en el hospedero así “hackean” el sistema inmune y hacen que el reconocimiento por parte de la célula se haga efectivo para generar la interacción y la posterior enfermedad.⁵¹

Las secuencias de aminoácidos son aquellas que dan pie a la posterior conformación tridimensional que será conocida como estructura secundaria. En este caso de estudio, las interacciones se dan por la superficie o periferia de la proteína hospedera.⁵²

Tabla 2. Secuencias de aminoácidos de la proteína viral y zonas de interacción de la proteína hospedera con el virus.

INTERACCIONES.	PDB	Secuencia Proteína Viral	Secuencia proteína hospedera
Calcineurina en complejo con dominio inhibidor de calcineurina de la proteína del African Swine Fever A238L	4F0Z	206-211 PKIIT 229-232 LCVK	117-123 MMVGNNL 327-333 NVMIRQ 344-345 PN 353-356 SLPF
Complejo del dominio HAUSP/USP7 con el péptido EBNA1	1YY6	444-447 EGPS	164-170 WGFNFNM
14-3-3 Zeta/Fosfopeptido complejado.	1QJB	ARSH sep YPA	49-60 KNVVGARRSSWR 121-127 MKGDYYR 176-180 VFYYE 222-228 RDNLTLW
Estructura de Rotavirus NSP1 enlazado con IRF-3	5JER	482-491 EEWELLISNS	260-264 YVRHV 285-290 RLGHCH 386-398 SLENTVDLHISNS
Complejo ALIX-AIP1 CON HIV-1 Dominio tardío YPLTSL	2R02	34-44 ELYPLTSLRSL	498-503 VLDKAV 666-681 LVANLKEGTKFYNELT
Complejo ALIX-AIP1 CON HIV-1 Dominio tardío YPDL	2R03	21-28 NYDLSE	487-503 PLRAEGTNFRTVLDKAV 662-679 NFVELVANLKEGTKFYNE
MAGI-1 PDZ-1 / E6CT	2KPL	151-161 RSSRTRRETQV	30-36 TWGGDE 45-47 SLV 76-86 THAQVVKIFQS 112-118 TSLVTSV
MAGI-1 PDZ1 enlazado con el péptido C-terminal de HPV18 E6	2I04	2002-2007 RETQVR	463-469 GFTWGG 510-519 HTHAQVVKIF
Sap97 PDZ2 enlazado con el Péptido C-Terminal de HPV18 E6	2I0L	2000-2006 RRRETQV	330-335 FSIAGG 384-392 EEAVTALKN
Sap97 PDZ3 enlazado con el Péptido C-Terminal de HPV18 E7	2I0I	2000-2006 RRRETQV	475-481 LGFNIVG 524-531 HEQAAAAAL
hDLG/Sap97 PDZ2 en complejo con péptido HPV-18 papilomavirus E6	2QOS	301-306 RRRETQV	14-23 FSIAGGVNQ 65-76 VTHEEAVTALKN

Retinoblastoma pocket en complejo con SV40 Large T Antigen	1GH6	103-107 LFCSE	703-713 IMMCSMYGICKV 719-130 LKFKIIVTAYKD 751-758 SIIVFYNS
Proteína no estructural de Coronavirus MHV de tipo salvaje nps15	2GTH	359-363 LWCNE	259-264 YAFEHV 275-279 GLHLL
Retinoblastoma pocket enlazado con E7 motivo LxCxE	1GUX	1-9 DLYCYEQLN	751-762 SIIVFYNSVFMQ
Dominio de Retinoblastoma Pocket en complejo con Dominio adenovirus E1A CR1	2R7G	40-49 PPTLHELYDL	460-468 MLKSEERL 475-486 KLLNDNIFHMSL 527-534 DFYKVIES 643-646 KSTSL 719-724 LKLFKII 733-740 HAVQETFK 749-764 YDSIIVFYNSVFMQRL
Proteína HIV-1 NEF en complejo con Dominio R96I mutante FYN SH3	1EFN	71-79 RPQVPLRPM	89-95 ALYDYE 133-140 IPSNYVAP
Estructura del complejo entre el dominio humano TRAF2 y un péptido de enlace LMP1	1CYZ	204-210 PQQATDD	393-395 RIY 399-402 DGTG 453-456 SSSF 467-471 SGCPL
Dominio Mouse Itch 3rd WW enlazado con Péptido derivado de la proteína 2A de la membrana latente del virus Epstein-Barr	2J09	101-109 EEPPPYED	21-32 YFVNHNTRITQW
Estructura de la glicoproteína de la envoltura del West Nile Virus	2HG0	Motivo Conservado 153-158 GNYTSQ	Hélice de Glicosilación. 151-162 SHGNYSTQVGAT
Dominio C-Terminal de la subunidad PB2 de la polimerasa del virus de la influenza A en complejo con importina humana alfa5	2JDQ	736-741 KRKRDS 748-756 QTATKRIRM	149-157 WVLTNISG 190-193 VWAL 198-200 GDS 233-236 VWAL 276-280 WALSY 282-286 SDGPN 316-319 ALR 324-327 VTGD 358-365 ACWTISNI 398-403 KEAANA 444-447 LNGL
Importina Alfa de ratón en complejo con péptido SV40 CN no fosforilado	1Q1T	123-133 AAPPKKKRKVE 127-134 KKKRKVEY	105-107 SRE 140-152 SAWALTNISGTS 190-193 AGDG 230-232 TWT 265-270 PEVLAD 280-284 DGPNE 307-309 LPI 313-325 ALRAIGNIVTGT 353-363 KEATWTMSNIT 398-404 AWAITNY
Dominios similares a ubiquitina USP7 en conformación compacta	4WPH	621-625 KCARK	629 S 687 M 680-687 DKDHDVML 756-766 ALDELMDGDII

3.2.3. Estructuras secundarias encontradas en la zona de interacción entre el patógeno viral y la proteína hospedera.

Las estructuras secundarias son las primeras unidades estructurales tridimensionales que hacen parte de las proteínas, entre ellas tenemos: α -helix, β -sheet, β -strand y loops, todas estas son las principales encargadas de generar interacciones con el resto de la proteína y con otras proteínas tanto del mismo hospedero (interacciones endógenas) como con proteínas externas, como las patógenas (Interacciones exógenas). Estas últimas se deben a que son la copia perfecta de las interacciones endógenas y el cuerpo no las detecta y es por ello que los patógenos pueden generar enfermedad.⁵³

Como se puede observar en las tablas 3 y 4, de las estructuras secundarias más comunes que son α -helix, β -sheet, β -strand y loops, los virus y las proteínas hospederas reaccionan como previamente se ha mencionado, en la periferia de las proteínas hospederas donde, en su mayoría, interactúan β -sheet, β -strand y loops. Como se evidencia en la tabla 3, se pueden observar que la mayoría de los virus optan por conformaciones que energéticamente son más estables y sencillas como lo son los loops y las β -strands.⁵⁴

Tabla 3. Estructuras secundarias formadas en las proteínas de los patógenos virales.

PDB/Interacción	Loop	β -Sheet	α -Helix
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)	L.V.	P.I.IT	
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)		E..S	
1QJB (Fosfo péptidos con Murine Polioma Virus)	ARSH		
5JER (IRF-3 con Rotavirus)	PL.IS		
2R02 (Alix-AIP1 con HIV-1)			YP..L
2R03 (alix-AIP1 con YDPL)	LYP.L		
2KPL (MAGI-1 PDZ1/ E6CT)		(ST).(VIL)	
2I04 (MAGI-1 PDZ1/ HPV18-E6)		(ST).(VIL)	
2I0L (PDZ2 CON HPV E6)		(ST).(VIL)	
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)		(ST).(VIL)	
2Q0S (PDZ2 CON HPV-18)		(ST).(VIL)	
1GH6 (Rb con SV40 Large T)	[L].C.[DE]		
2GTH (nsp15 con MHV)	L.C.E		
1GUX (Rb con E7 LxCxE)		[L].C.[DE]	
2R7G (Rb con E1A CR1)	[L].H.L[YF]		
1EFN (HIV-1 Nef con R961 FYN SH3)	P...P		
1CYZ (TRAF2 LMP1)	P.Q.[TSD]		
2J09 (WW domain con Eps-Barr Virus 2A)	PPLP		
2HG0 (West-Nile Virus)			.(N)('P)[ST]
2JDQ (Influenza Virus A con PB2 con importina alpha 5)	[KR][KR].[KR]		
1Q1T (Mouse importin Alpha: SV40 CN)	[KR][KR].[KR]		
4WPH (USP7 con ICP0 Herpes Virus Simple 1)	K...K		

Por otro lado, la tabla 4, donde se pueden observar las estructuras secundarias que participan por el lado de las proteínas hospederas en los casos de estudio. En el caso de los hospederos las estructuras secundarias si se extienden a tener mayor diversidad y no se quedan sólo en los loops y las β -strands sino también en conformaciones de α -helix y β -sheets, que dan pie a posteriores conformaciones tipo V, tipo sándwich que encierran a la proteína viral generando más puntos de contacto con la proteína hospedera y así una interacción más fuerte por varios puntos de la proteína hospedera, lo que también se puede observar con los diferentes fragmentos de secuencia de proteína hospedera presentados en la tabla 3⁵⁵, en interacciones como, por ejemplo, la del dominio C-terminal de la subunidad PB2 de la polimerasa del virus de la influenza A en complejo con la importina humana alfa5. (PDB 2JDQ)

Las interacciones entre hospederos y patógenos de tipo viral se ha evidenciado en diversos estudios, utilizando diferentes enfoques, entre ellos el bioinformático para determinar características en común, hasta ahora lo que se resalta mayoritariamente es el poder del virus de evolucionar y copiar perfectamente el mecanismo de las interacciones endógenas y no ser detectado por los mecanismos de sistema inmune.

Tabla 4. Estructuras secundarias formadas en las proteínas hospederas.

PDB/Interacción	Loop	β -Sheet	α -Helix
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)		PxlIT Contacto	Contacto con LxVx
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)		Contacto E..S	
1QJB (Fosfo péptidos con Murine Polioma Virus)			Contacto con 5,7,9 y 3
5JER (IRF-3 con Rotavirus)			contacto con PLxIS
2R02 (Alix-AIP1 con HIV-1)			Contacto con YP...L
2R03 (alix-AIP1 con YDPL)			Contacto tipo sandwich LYP.L
2KPL (MAGI-1 PDZ1/ E6CT)	Contacto con RSSRTR	Contacto con (ST).(VIL)	Contacto con (ST).(VIL)
2I04 (MAGI-1 PDZ1/ HPV18-E6)		Contacto con (ST).(VIL)	Contacto con (ST).(VIL)
2I0L (PDZ2 CON HPV E6)		Contacto con (ST).(VIL)	Contacto con (ST).(VIL)
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)		Contacto con (ST).(VIL)	Contacto con (ST).(VIL)
2Q0S (PDZ2 CON HPV-18)		Contacto con (ST).(VIL)	Contacto con (ST).(VIL)
1GH6 (Rb con SV40 Large T)			Contacto con [L].C.[DE]
2GTH (nsp15 con MHV)	Contacto con L.C.E		Contacto con L.C.E
1GUX (Rb con E7 LxCxE)			Contacto con [L].C.[DE]
2R7G (Rb con E1A CR1)			Contacto con [L].H.L[YF]
1EFN (HIV-1 Nef con R961 FYN SH3)	Contacto con P..P		Contacto con P..P
1CYZ (TRAF2 LMP1)		Contacto con P.Q.[TSD]	Contacto con P.Q.[TSD]
2J09 (WW domain con Eps-Barr Virus 2A)		Contacto con PP.Y	
2HG0 (West-Nile Virus)	No tiene contacto		
2JDQ (Influenza Virus A con PB2 con importina alpha 5)			Contacto con [KR][KR].[KR]
1Q1T (Mouse importin Alpha: SV40 CN)			Contacto con [KR][KR].[KR]
4WPH (USP7 con ICP0 Herpes Virus Simple 1)		Contacto con K..K	

3.3. Análisis por Área Superficial Accesible de Solvente (ASAS) o SASA.

Una secuencia de aminoácidos en formación de una estructura secundaria, por lo general, se encuentra expuesta a otros factores al conformar una proteína, como tales, la solvatación o efecto del solvente en los aminoácidos que corresponden a esta estructura, los enlaces de hidrógeno, puentes electrostáticos o fuerzas de Van der Waals que se estén ejerciendo dentro de la misma proteína o con el núcleo de esta. Adicionalmente, se debe tener en cuenta que la fuerza impulsora a formar este tipo de estructuras responde a la presencia de aminoácidos de tipo aromático y su capacidad de crear enrollamientos con núcleos hidrofóbicos.

Con este orden de ideas, los aminoácidos de la secuencia forman estructuras secundarias dando respuesta al impulso de creación los núcleos hidrofóbicos; a la medida geométrica del área superficial en el entorno molecular de la exposición de estos aminoácidos que no generan núcleo hidrofóbico se le llama Área Superficial Accesible de Solvente (ASAS o SASA sigla en inglés). El SASA permite predecir de alguna manera la estructura de una proteína. En el caso de estudio, permite corroborar y verificar que las estructuras secundarias encontradas previamente.⁵⁶

Como se evidencia en la tabla 5, los valores calculados gracias al algoritmo SASA.py en la herramienta PyMOL, en los casos donde se presentan interacciones entre loop y β -strand o β -sheet, sea para virus o para hospedero el SASA aumentará, sin embargo, si se presentan interacciones con varias fracciones de secuencia pertenecientes a una α -hélice se presentará el mismo efecto. También, la presencia de α -hélice debido a la conformación enrollada como se puede apreciar en la figura 3, representa una exposición superficial menor de área al solvente lo que disminuiría considerablemente el SASA como se puede observar en los casos de las interacciones codificadas como 4F0Z y 1YY6, donde se puede observar claramente este comportamiento. En casos como el PDB 2JDQ, que tiene varios segmentos de interacción por medio de alfa hélices se encuentra una disminución por este mismo efecto en el hospedero, sin embargo, en el patógeno, se encuentra más expuesto al solvente debido a tener varias zonas de interacción.

Para el PDB 2HG0, hay que destacar que como este no se encuentra interaccionando con alguna proteína hospedera y lo que se realizó fue el estudio de las zonas de interacción del sitio de glicosilación que permite el enlace de este virus a la proteína del hospedero y el cambio conformacional de la Hélice N-Glicosilada también determinante en la interacción del hospedero se encuentra el efecto de que esta Hélice al cambiar su conformación aumenta su exposición al solvente lo que la hace tener un mayor SASA.

Tabla 5. Área Accesible de Solvente (SASA) Calculado para proteína viral y proteína hospedera.

INTERACCIONES	SASA Virus (Å²)	SASA Hospedero (Å²)
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)	642,327	630,472
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)	543,554	205,862
1QJB (Fosfopéptidos con Murine Polioma Virus)	1586,082	1732,626
5JER (IRF-3 con Rotavirus)	3610,881	4673,893
2R02 (Alix-AIP1 con HIV-1)	803,650	891,722
2R03 (alix-AIP1 con YDPL)	768,738	1736,218
2KPL (MAGI-1 PDZ1/ E6CT)	833,003	893,137
2I04 (MAGI-1 PDZ1/ HPV18-E6)	1076,033	1309,058
2I0L (PDZ2 CON HPV E6)	505,381	535,099
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)	1840,353	1193,149
2QOS (PDZ2 CON HPV-18)	659,160	818,982
1GH6 (Rb con SV40 Large T)	233,382	629,304
2GTH (nsp15 con MHV)	832,948	269,361
1GUX (Rb con E7 LxCxE)	730,677	488,073
2R7G (Rb con E1A CR1)	1632,129	4032,591
1EFN (HIV-1 Nef con R961 FYN SH3)	1109,632	1097,140
1CYZ (TRAF2 LMP1)	1019,856	746,069
2J09 (WW domain con Eps-Barr Virus 2A)	1082,426	783,445
2HG0 (West-Nile Virus)	Sitio N-Glicosilado	Hélice N-Glicosilada
	338,066	588,846
2JDQ (Influenza Virus A con PB2 con importina alpha 5)	7416,079	2248,26
1Q1T (Mouse importin Alpha: SV40 CN)	1355,726	1592,645
4WPH (Virus del Herpes con USP7)	734.681	271.030

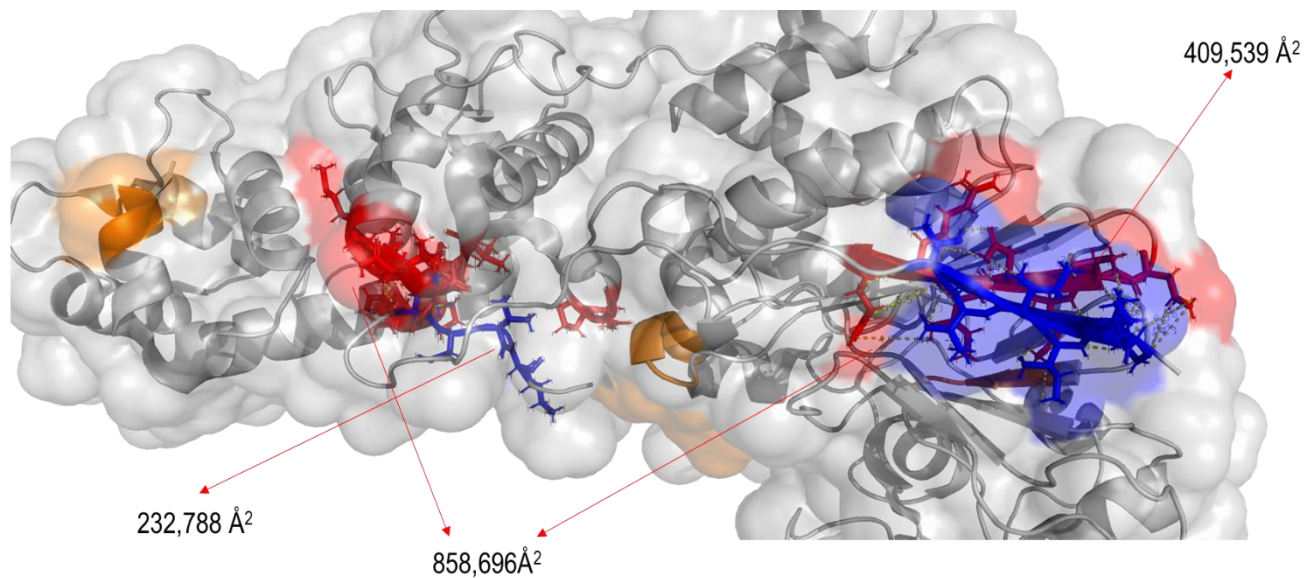


Figura 3. Calcineurina en complejo con el dominio inhibidor de calcineurina del virus de la fiebre porcina africana (African Swine Fever A238L) (PDB 4F0Z)

3.4. ANCLAJE DEL VIRUS.

Las PPIs entre hospederos y patógenos se dan por medio de la copia de la interacción endógena de la proteína por el virus. Esto conlleva a que la proteína viral tenga unos puntos de enlace, unos puntos de anclaje más distinguidos y que en cierta forma son los puntos que copia. Por tanto, para este estudio es relevante analizar los puntos de anclaje del virus, esto se realizó por medio de la medición por residuo del SASA para la proteína viral, encontrando así que los aminoácidos más involucrados en la formación de enlaces dentro de la interacción son aquellos que presentan un porcentaje de SASA más pequeño. El SASA permite dar cuenta del área superficial que tiene disponible una estructura en este caso, residuo de aminoácido, para interactuar con un solvente. Cuando se mide por residuo se mide es por porcentaje y

permite determinar la exposición de cada uno de los residuos o cadenas laterales del aminoácido que están expuestas y las que están formando enlaces con la proteína hospedera.

En la tabla 6, se presentan los resultados por SASA en porcentaje por aminoácido según la interacción estudiada, se puede deducir que aquellos que se encuentran “más libres” son aquellos que tienen menos enlaces con la proteína hospedero, y los que se encuentran con porcentajes más pequeños, puede deberse a dos razones, a que presentan varios enlaces con la zona de interacción en proteína hospedera o que se encuentran ocultos y no expuestos en un núcleo o core hidrofóbico.

En la figura 4, se aprecia la forma de interacción del virus Epstein-Barr con la proteína HAUSP/USP7, y como se ancla este virus generando unos puntos específicos, los cuales se observan como esferas de color azul más claro, mientras que aquellas esferas en azul más intenso son aquellas con porcentaje de SASA por residuo mayor.

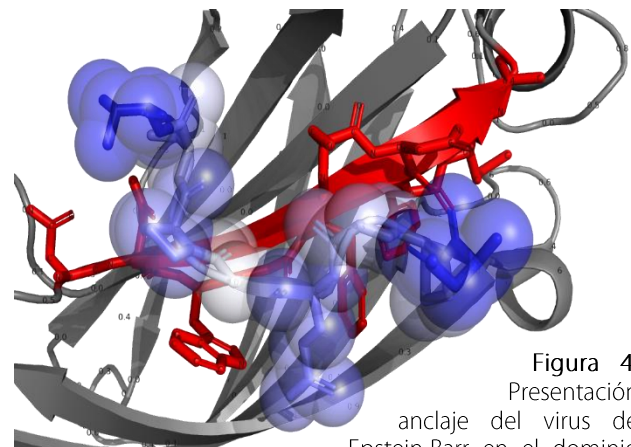


Figura 4. Presentación anclaje del virus de Epstein-Barr en el dominio de la proteína HAUSP/USP7. (PDB 1YY6)

Tabla 6. Datos de SASA por residuo de las secuencias de virus o patógeno .

INTERACCIÓN/Aminoácidos	S	T	Q	N	Y	G	C	A	V	L	I	M	P	F	W	D	E	K	H	R	SECUE NCIA
	Ser	Thr	Gln	Asn	Tyr	Gly	Cys	Ala	Val	Leu	Iso	Met	Pro	Phe	Trp	Asp	Glu	Lys	His	Arg	
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)		47					33		5	33 39 41 4			5					64 84			PKIIT LCVK
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)	16	83				26 5							27 48				50				PGEGPS T
1QJB (Fosfopéptidos con Murine Polioma Virus)	17				31			80 68					25						35	70	ARSH SEP YPA
5JER (IRF-3 con Rotavirus)	45 99			84						24 30	20			16			57 83 68				EEFELLI SNS
2R02 (Alix-AIP1 con HIV-1)	39 77	62			3					34 27 6 49			5				73			72	ELYPLTS LRSL
2R03 (alix-AIP1 con YDPL)	77			78	9					46 36			28			85	87				NLYDLSE
2KPL (MAG-I PDZ1/ E6CT)	52 50	7 0	59						6		21						31			98 36 12 37	RSSRTR RETQV
2I04 (MAG-I PDZ1/ HPV18-E6)		31	64						7								55			46 100	RETQVR
2I0L (PDZ2 CON HPV E6)		19	55						6								66			0 89 46	RRRETQ V
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)		30	78						10								48			0 95 67	RRRETQ V
2QOS (PDZ2 CON HPV-18)		32	64						17								54			75 40	RRRETQ V
1GH6 (Rb con SV40 Large T)	82						3			1				42			23				LFCSE
2GTH (nsp15 con MHV)				71			58			2					48		23				LWCNE
1GUX (Rb con E7 LxCxE)			78	74	36 70		3			6 17 5 10 39						68	16				DLYCYE QLN
2R7G (Rb con E1A CR1)		11			13					5 9 39			30 19			37	34		14		PPTELLH ELYDL x3
		10			14					9 39			49 19			40	34		13		
		39			63					9 83 0			48 100			82	8		39		
1EFN (HIV-1 Nef con R961 FYN SH3)			42						1	56		29	31 11 78							50 14	RPQVPL RPM x2
			46						3	52		22	32 12 83							67 14	
1CYZ (TRAF2 LMP1)		65	57 45					11					34			53 38					PQQATD D x2
		65	49 43					10					41			56 42					
2J09 (WW domain con Eps-Barr Virus 2A)					30								89 63 28 83			78	99 62 84				EEPPPP YED
2HG0 (West-Nile Virus)	70	48	2	40	35	65													7,6,60,52, 40,31		MOTIVO CONSER VADO: GNYSTQ
2JDQ (Influenza Virus A con PB2 con importina alpha 5)	44															24		17 12		5 10	Corto: KRRKDS
		46 37	72					63			43	51						18		37 32	Largo: QTATKRI
1Q1T (Mouse importin Alpha: SV40 CN)					31				34								56	94 33 20 34		14	CORTO: KKRKV EY
								86 51	64				13 52				47	44 21 33 22		31	LARGO: QTATKRI RM
4WPH (USP7 con ICP0 Herpes Virus Simple 1)							41	91										20 22		78	KCARK X2
							42	87										22 17		57	

3.5. Determinación de tipos de enlace entre el patógeno y la zona de contacto con la proteína hospedera.

Las interacciones proteína-proteína (PPIs) sean endógenas o exógenas poseen características en común, las cuales se repiten cuando los modelos se hacen similares. Las PPIs juegan papeles vitales en diferentes procesos celulares, tanta inhibición, traslocación de señales nucleares, replicación e inhibición de ciclos. Para el caso de las interacciones exógenas con patógenos, son copias de las interacciones endógenas naturales del proceso celular del hospedero.

En las PPIs, las fuerzas o enlaces que contribuyen a su formación son: primero, la complementariedad estérica, que permite analizar los contactos entre las interfases superficiales de cada una de las proteínas para determinar su complementariedad; el grado de complementariedad depende del grado de interacción y de la durabilidad de la interacción. Como se puede apreciar en la tabla 7 y en las figuras 5, 6, 7 y 8. La complementariedad de las estructuras secundarias en cuestión de impedimento estérico juega un rol vital en los casos estudiados puesto que en muchos de los casos se observa que se repiten patrones, por ejemplo, interacciones de una fracción de proteína viral en forma de β -strand interacciona con otra β -strand encontrada en el hospedero. en la interacción entre el virus del Herpes con la proteína deubiquitina (PDB 1YY6)⁵⁰. Esta información se puede verificar, además, en la tabla 4 donde se presenta más detalladamente las estructuras secundarias que participan en cada uno de los casos estudiados. Otro caso de complementariedad estérica se presenta en interacciones de forma “*Hélice, Loop, Hélice*”, ejemplos de ello son las interacciones entre el virus causante de la Anemia equina infecciosa (PDB 2R03)³⁶ o el virus de la Influenza A con el complejo con la importina Alpha humana. La interacción tipo “Sandwich” como algunos autores la califican, se presenta comúnmente en factores de transcripción y homeoproteínas.⁵⁷

Segundo, la interacción hidrofóbica o enlaces de Van der Waals^{19,58-60}; estas son particularmente dependientes de la cantidad de aminoácidos de tipo apolar o de tipo aromático presentes en las estructuras primarias o secuencias de las zonas donde se realizan las interacciones patógeno-hospedero, como se puede apreciar en las figuras 5,6,7 y 8; se encuentra que constituyen una porción significativa de los enlaces realizados entre las zonas y que son determinantes para el anclaje de la proteína viral dentro de la proteína hospedera.

Tercero, las interacciones electrostáticas^{17,61-63} o puentes salinos, es otras de las fuerzas que se estudiaron como motivos de interacción entre una proteína hospedera y una proteína patógena viral, en los casos estudiados, los puentes salinos, aunque se encuentran presentes dependen mucho de la cantidad de aminoácidos con carga, sea positiva o negativa que tenga la fracción o motivo lineal interaccionante. Interacciones dependientes de señales de localización nuclear como lo es la del virus de la influenza A con la proteína de la importina Alpha 5 humana. (PDB 2JDQ)³⁸ o la interacción entre el virus del papiloma humano y el “pocket” de la proteína retinoblastoma (PDB 1GUX)³², son específicos para puentes salinos. En ambos casos, se hacen necesarios aminoácidos con cargas positivas como lo son los aminoácidos ácidos como lo son el ácido glutámico y el ácido aspártico o aminoácidos básicos como lo son la arginina, la histidina y la lisina.

Cuarto, los enlaces de hidrógeno o puentes de hidrógeno constituyen la gran mayoría de los enlaces formados entre los aminoácidos de las zonas en contacto entre proteínas hospederas y proteínas virales. Estos en su mayoría se daban por medio de los nitrógenos y oxígenos, encontrados tanto en las cadenas laterales como en los “backbones” o esqueletos de las cadenas de la proteína.⁶⁴⁻⁶⁶

Tabla 7. Relación entre la formación de enlaces y las estructuras secundarias encontradas en las interacciones particulares de las proteínas virales.

PDB/interacción	Virus	Formación de enlaces e interacción entre estructuras secundarias.
4F0Z	African Swine Virus A238L	El motivo P.I.IT que corresponde a una β -strand se contacta con otra β -strand El motivo lineal L.V. que corresponde a un loop se enlaza por medio de un α -
2R03	Equine infectious anemia virus (Isolate Wyoming)	Loop en la proteína viral, entre dos α -helix en el hospedero
4WPH	Herpes Simplex Virus	Loop en el virus entre un β -strand y un α -helix del hospedero
2R7G	Human Adenovirus 5	Interacción entre un loop del virus y α -hélice del hspedero Otra de las fracciones del virus interactua por medio de un α -helix que se enlaza con α -helix del hospedero
2JO9	Human Epstein-Barr Virus Strain B95-8	Enlace por medio de un loop en el virus y un β -strand hospedero. Nota: para la interacción 1CYZ, la interacción es similar, la única diferencia se encuentra en la presencia de un pequeño fragmento de α -hélice interaccionando con la parte final del loop del virus
1YY6	Human Herpes Virus (Strain B95-8)	Enlace por medio del contacto de dos β -Strands, una en el virus, la otra en el hospedero.
1EFN	Human Immunodeficiency virus type 1- (BRU Isolate)	Enlace de un loop en la proteína viral entre α -hélice en el hospedero y un loop también el hospedero.
2R02	Human Immunodeficiency virus type 1- group M subtype j, isolate SE9173	Enlace se da entre tres α -helix, una en el virus, y las otras dos en el hospedero en forma de sandwich.
1GUX	Human Papillomavirus type 16	Contacto entre un β -Strand en el virus con α -helix en el hospedero
2OQS	Human Papillomavirus type 18	Contacto entre un β -Strand del virus con un α -hélice y un β -Strand del hospedero
2JDQ	Influenza Virus A	Interacción entre loops en el virus y α -helix en el hospedero
1QJB	Murine Polyomavirus strain A3	Interacción entre un loop y cuatro α -hélice en el hospedero
2GHT	Murine coronavirus	Enlace de un loop de la proteína viral, α -hélice en el hospedero y un loop también
5JER	Simian rotavirus A-SA11	Interacción entre un loop de la proteína viral y α -hélice en el hospedero
1GH6	Simian Virus 40	Interacción entre un loop de la proteína viral y α -hélice en el hospedero
2GH0	West Nile Virus	α -hélice en el virus

Nota: dentro de la tabla 7, se encuentran 16 de las 22 interacciones que corresponden los 16 diferentes virus estudiados. Y se generaliza, para las otras interacciones.

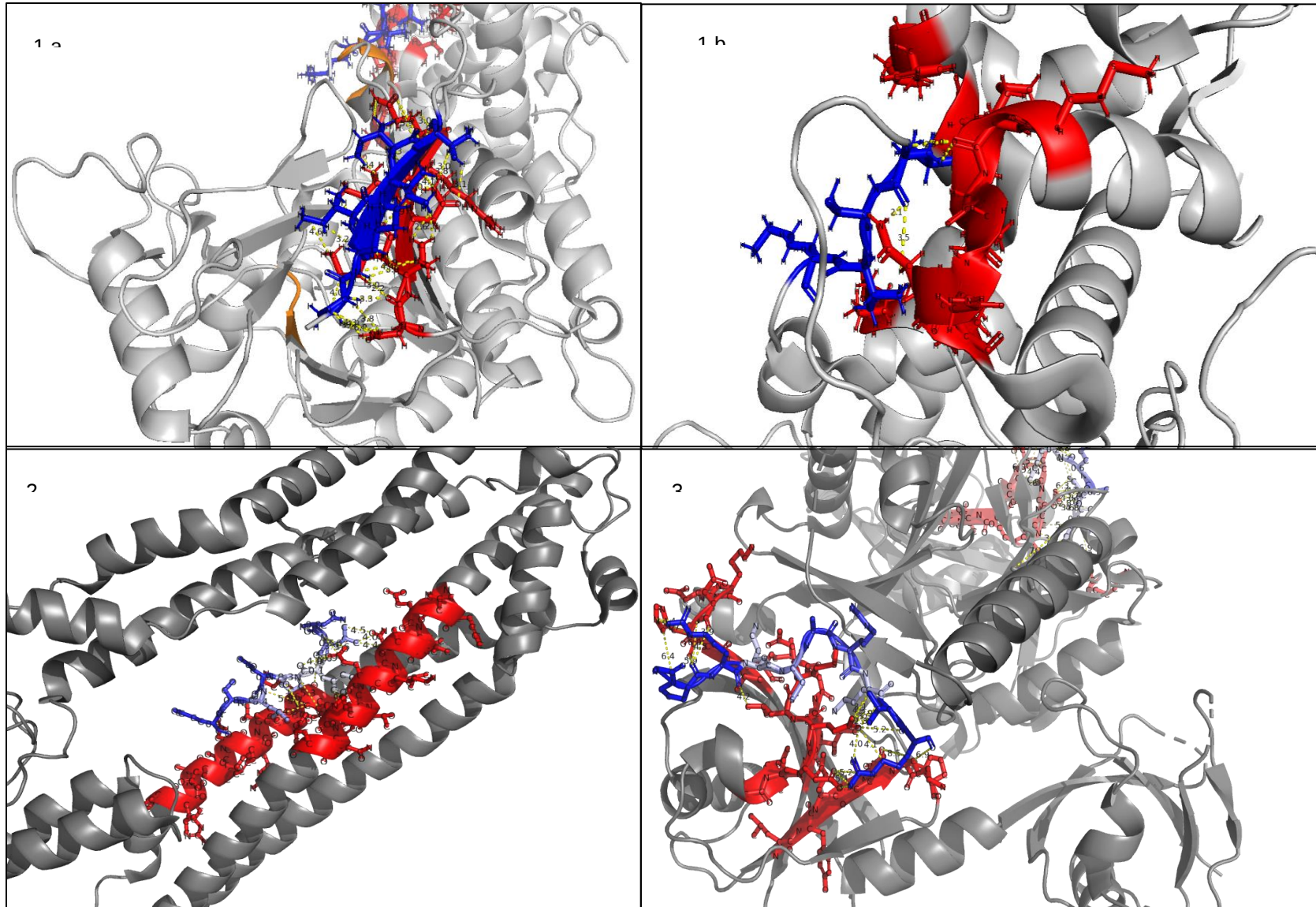


Figura 5 Primera parte de las interacciones entre proteínas hospederas (**rojo**) y proteínas virales (**azul**) con enlaces destacados (**amarillo**). 1) interacción entre Calcineurina y el dominio inhibidor de Calcineurina del virus de la fiebre porcina africana, a) vista de la interacción entre las β strand correspondiente al motivo PxlIT del patógeno y β -strand del hospedero. b) vista de la interacción del loop del virus correspondiente al motivo LxVx con el α -helix del hospedero en enlace. (PDB 4F0Z). 2) Interacción entre la proteína AIP1 con el virus de la anemia equina infecciosa. (PDB 2R03) 3) Interacción entre la proteína USP7 con el virus del herpes simplex. (PDB 4WPH).

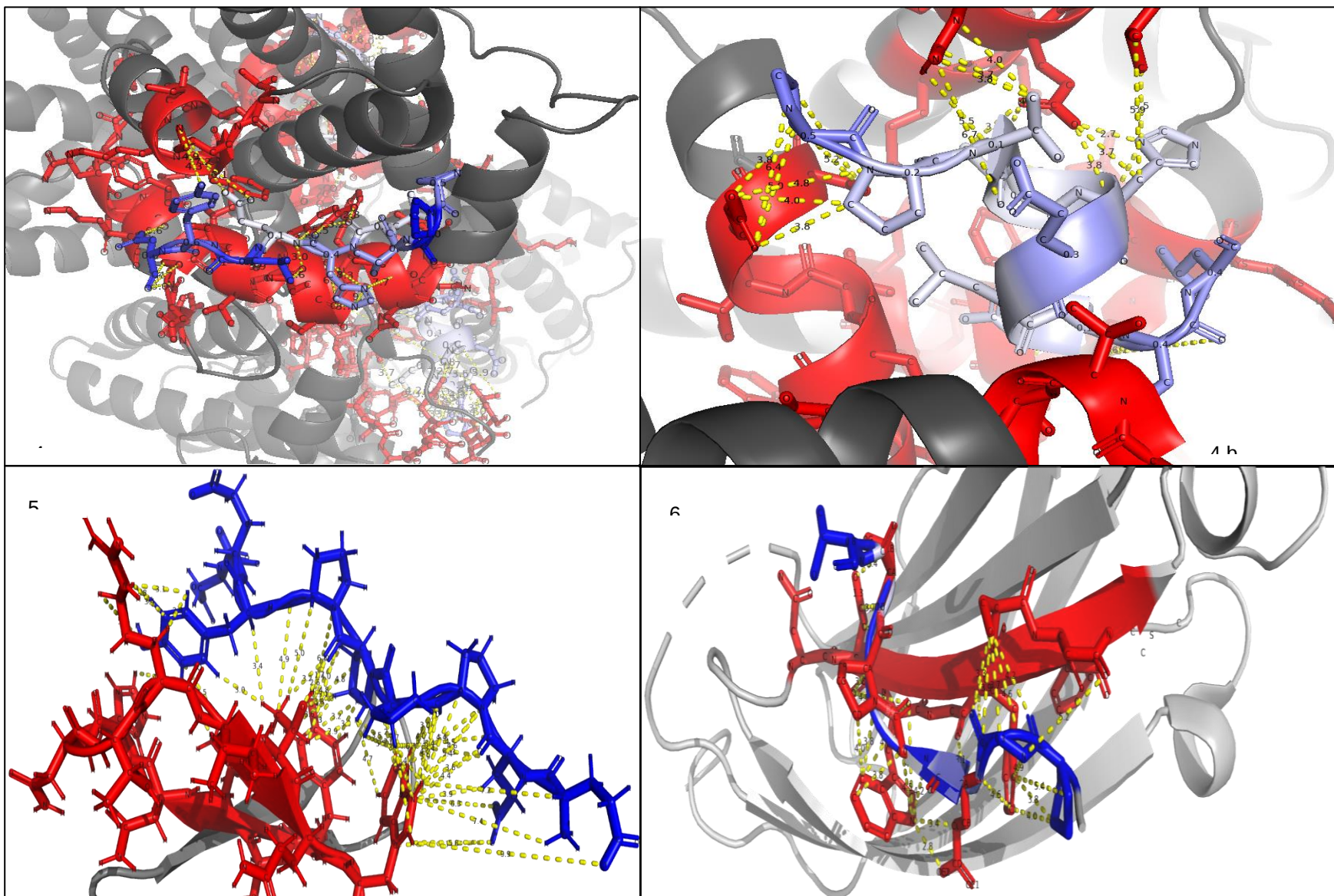


Figura 6. Segunda parte de las interacciones entre proteínas hospederas (rojo) y proteínas virales (azul) con enlaces destacados (amarillo). 4. Interacción entre el “pocket” de la proteína retinoblastoma humana con el Human Adenovirus (PDB 2R7G), 4.a) vista de la interacción el loop del virus y las α -hélices del hospedero. 4.b) vista de la interacción entre el fragmento de α -hélice del Adenovirus humano con α -hélice del hospedero. 5) Interacción entre el virus del Epstein Barr con la proteína del Dominio WW del ratón. (PDB 2JO9) 6) Interacción del virus del herpes con la proteína HAUSP/USP7 (PDB 1YY6)

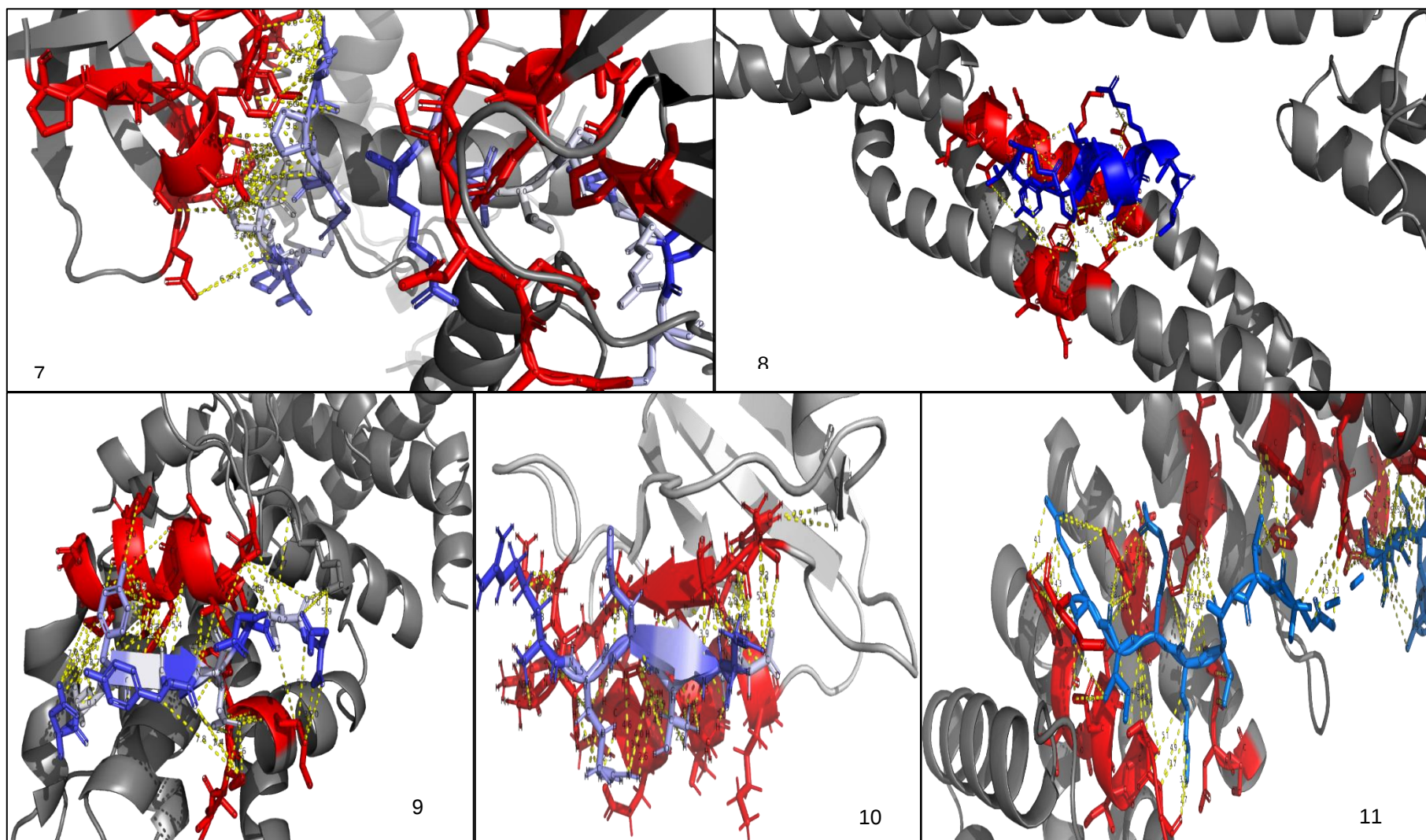


Figura 7. Tercera parte de las interacciones entre proteínas hospederas (rojo) y proteínas virales (azul) con enlaces destacados (amarillo). 7) Interacción entre virus del VIH con el dominio SH3 (PDB 1EFN) 8) Interacción entre el virus del VIH-1 con la proteína del dominio ALIX en formación V por medio del motivo LYPxL. (PDB 2R02) 9) Interacción entre virus del papiloma humano tipo 16 con la proteína "pocket" retinoblastoma. (PDB 1GUX) 10) interacción entre virus Papiloma humano tipo 18 con la proteína Magi-1 PDZ1. (PDB 2QOS) 11) Interacción entre el virus de la influenza A con la proteína Alpha5 importina humana. PDB (2JDQ)

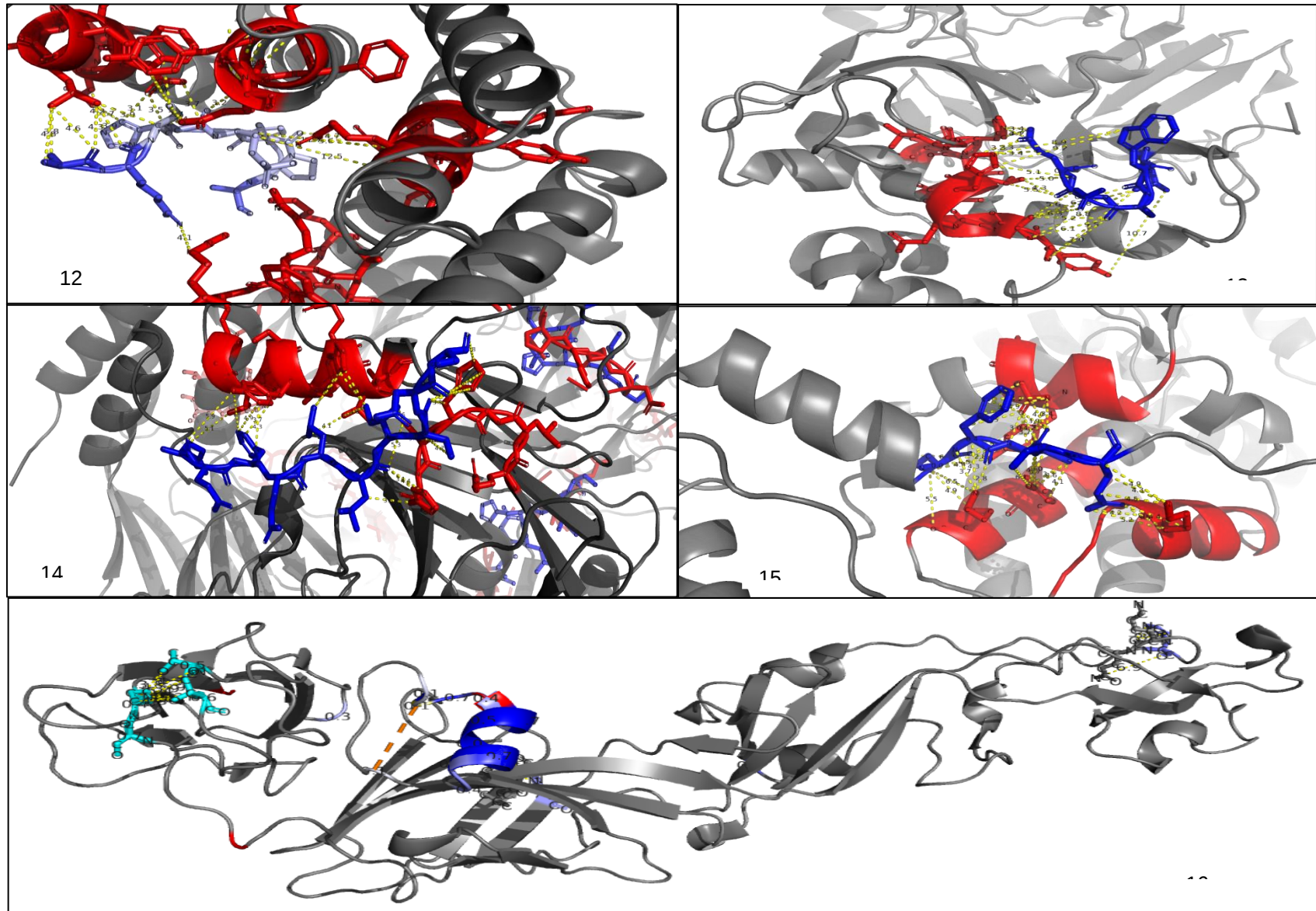


Figura 8. Cuarta parte de las interacciones entre proteínas hospederas (rojo) y proteínas virales (azul) con enlaces destacados (amarillo). 12) Interacción entre Murine Poliovirus Strain A3 y el complejo 14-3-3 Zeta/Phosphopeptide (PDB 1QJB), 13) Interacción entre el Murine Coronavirus y la proteína NSP15 (PDB 2GTH) 14) Interacción entre Simian Rotavirus A-SA11 y la proteína IRF-3 (PDB 5JER), 15) Interacción entre el virus Simian Virus 40 con el "pocket" de retinoblastoma (PDB 1GH6) 16) Virus del Este del Nilo o West Nile Virus señalizando la Hélice de Glicosilación, PDB 2GH0

3.6. Análisis de Energía de estabilización vs. pH de estabilización.

La dependencia del pH y las proteínas se da gracias a la contribución de este a la energía libre de los grupos ionizables de la proteína generando así una desnaturalización de esta. Para las interacciones proteína-proteína en interacciones hospedero-patógeno, sucede de la misma manera, que el pH genera una desnaturalización causando así que ambas proteínas pierdan su conformación natural y así cambien su energía de estabilización. Para calcular los puntos de desnaturalización de cada proteína es necesario calcular primero los pKa de los grupos ionizables dentro de esta y así calcular la dependencia de pH vs. La energía de desnaturalización de la proteína.⁶⁷

La energía de estabilización de una proteína se define como el balance entre las fuerzas que determinan si una proteína está en un equilibrio entre su estado nativo y su estado desnaturalizado⁶⁸. El balance entre las energías libres de estados enrollado y desenrollado. Los estados enrollados estabilizados por los diferentes enlaces dados entre proteínas, por ejemplo, enlaces disulfuro, no covalentes, hidrofóbicos, electrostáticos, enlaces de hidrógeno y enlaces de Van der Waals.⁶⁹

Ahora bien, las interacciones entre proteínas cumplen las mismas leyes pero el que se tengan dos o más proteínas o fragmentos de proteínas en el proceso hace que el cálculo de la energía de estabilización dependiente del pH también sea más complejo, por tanto, se han desarrollado algoritmos como PropKa que por medio de una simplificación de las interacciones permite determinar de forma acertada la energía de estabilización vs el pH de cada interacción y permite su visualización a través de gráficos, como se puede observar en la figura 9 y tablas donde presenta los resultados del análisis.^{26,70,71}

Los archivos obtenidos al correr el algoritmo PropKa, sea en consola nativa del sistema operativo Linux o en la herramienta en línea PDB2PQR server son de carácter “.propka”, como el que se puede observar en la figura 9 el cual es un archivo que toma la interacción de cada proteína, la separa y señala cada grupo ionizable dentro de cada proteína interactuante y determina en sus últimos pasos la energía de estabilización vs. pH y posteriormente,

el pI de la interacción proteína-proteína, del cual se hablará más adelante.⁷¹

Adicionalmente, el pH de máxima estabilización depende de dos factores principales, el primero es la composición de aminoácidos, es decir, la estructura primaria, previamente estudiada y detallada, y la organización de los grupos ionizables en la estructura tridimensional de cada PPI, de la determinación de los grupos ionizables se pueden determinar también diversas fuerzas impulsoras a la formación de las estructuras tridimensionales.⁷²

Los grupos ionizables son aquellos grupos que pueden ser titulables al someter la proteína o el complejo de proteínas a una curva de pH. Estos forman residuos que pueden ser identificados muy fácil químicamente, son los determinantes de la estabilización de una proteína en cuestión energética pues a cierto pH poseen una configuración que permite a la proteína encontrarse en un mínimo de estabilidad. Los grupos ionizables para estudiar en una proteína o en un complejo de interacción son, por ejemplo, los aminoácidos de carga negativa (Ácido Aspártico (ASP o D) y Ácido Glutámico (Glu o E)), aminoácidos de carga positiva (Lisina (Lis o K), Histidina (His o H), Arginina (Arg o R)), aminoácidos polares sólo de dos clases, Tirosina (Tyr o Y) y Cisteína (Cis o C), según la cantidad de grupos ionizables se podría pensar en una explicación para el pH donde se presenta el mínimo de energía de estabilización.^{26,70,73}

En la tabla 8, se pueden evidenciar los resultados obtenidos en cada una de las veintidós PPIs caracterizadas en este estudio, incluyendo la cantidad de grupos ionizables en cada uno de los complejos de interacción, estos podrían explicar en cierta manera, la dependencia del pH para establecer un mínimo de energía en cada complejo. Los complejos, donde la mayoría de los grupos ionizables presenten carácter o carga positivos como lo son; Lisina, Arginina e Histidina, darían pie a que los pH de estabilización mínima posean una naturaleza básica (pH > 7) Y de la misma manera, los complejos donde la mayoría de los grupos ionizables sean de carácter negativo o ácido, como Ácido Aspártico o Ácido Glutámico, tengan un pH de estabilización de naturaleza ácida (pH < 7)^{67,68,72,74}

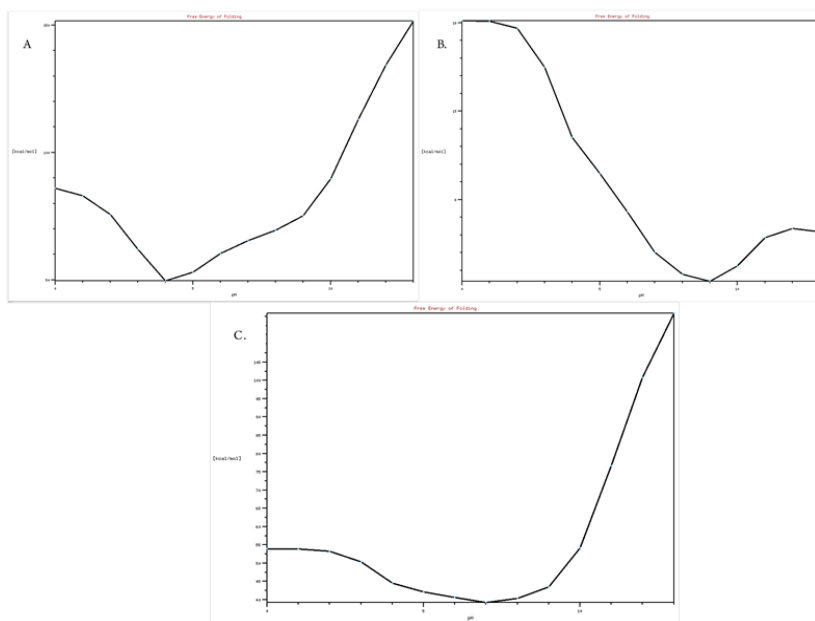


Figura 9. Gráficos correspondientes a la energía libre de desenrollamiento de tres complejos de interacción hospedero-patógeno; A) corresponde a la interacción en el complejo Calcineurina con el virus de la fiebre porcina africana (Interacción con pH de estabilización ácido). B) corresponde a la interacción entre la proteína HAUSP en hospedero humano con el virus EBNA1 (Interacción pH de estabilización básico) C) corresponde a la interacción entre la proteína Retinoblastoma y la oncoproteína causante del cáncer de cuello uterino. E7. (HPV-E7) (interacción con pH de estabilización neutro).

Tabla 8. Resultados del análisis de pH de estabilización vs. Energía de estabilización y grupos ionizables de cada interacción.

INTERACCIÓN	PDB	pH	Energía de estabilización	Grupos Ionizables (cantidad)						
				ASP (D)	GLU (E)	HIS (H)	CYS (C)	TYR (Y)	LYS (K)	ARG (R)
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)	4F0Z	4,1	63,7	41	37	12	15	22	37	24
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)	1YY6	8,8	3,4	10	10	4	3	1	7	8
1QJB (Fosfopéptidos con Murine Polioma Virus)	1QJB	8,6	35,3	28	25	4	6	24	40	24
5JER (IRF-3 con Rotavirus)	5JER	3,9	115,6	48	76	24	20	24	24	56
2R02 (Alix-AIP1 con HIV-1)	2R02	3,8	48,4	42	57	9	10	21	58	58
2R03 (alix-AIP1 con YDPL)	2R03	3,8	46,1	43	57	9	10	21	58	29
2KPL (MAGI-1 PDZ1/ E6CT)	2KPL	4,1	6,5	11	7	3	2	1	9	8
2I04 (MAGI-1 PDZ1/ HPV18-E6)	2I04	2,9	16	14	10	6	4	0	10	10
2I0L (PDZ2 CON HPV E6)	2I0L	10,1	0,5	8	14	6	2	4	18	3
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)	2I0I	12,9	13,8	12	18	6	0	0	9	21
2QOS (PDZ2 CON HPV-18)	2QOS	8,6	5,2	4	7	4	1	2	10	2
1GH6 (Rb con SV40 Large T)	1GH6	6,9	39,3	21	15	3	9	20	36	20
2GTH (nsp15 con MHV)	2GTH	4,3	41,3	28	18	7	9	10	25	14
1GUX (Rb con E7 LxCxE)	1GUX	7,7	24,3	12	24	9	8	18	25	15
2R7G (Rb con E1A CR1)	2R7G	7,2	107,1	29	53	21	14	35	52	35
1EFN (HIV-1 Nef con R961 FYN SH3)	1EFN	3,8	28,4	18	24	14	2	22	12	20
1CYZ (TRAF2 LMP1)	1CYZ	3,8	45,6	40	18	6	9	18	27	33
2J09 (WW domain con Eps-Barr Virus 2A)	2J09	5,6	-6,6	3	5	1	0	2	1	3
2HG0 (West-Nile Virus)	2HG0	8,9	27,9	17	23	11	12	11	23	14
2JDQ (Influenza Virus A con PB2 con importina alpha 5)	2JDQ	3,9	108,7	47	66	8	24	18	54	46
1Q1T (Mouse importin Alpha: SV40 CN)	1Q1T	4,1	38,2	19	27	8	6	7	32	13
4WPH (USP7 con ICP0 Herpes Virus Simple 1)	4WPH	4,1	82,9	58	43	12	12	28	54	44

3.7. Análisis de pI en estados enrollados y desenrollados de cada PPIs.

El punto isoeléctrico o pI es el pH en el cuál una proteína o una interacción entre proteínas tiene una carga electrostática que equivale a cero, es decir, es neutra. Si el pI de una proteína es de tipo neutro, se dan interacciones mayoritariamente entre proteína-proteína, mientras que cuando este se acerca a condiciones máximas, la interacción con los solventes del sistema donde se encuentran las proteínas (agua) generan que las interacciones se den por enlaces tipo hidrógeno e hidrófobos leves.⁷⁵

Al tener un sistema en presencia de una PPI con un pH inferior o superior al de pI las fuerzas electrostáticas generaran repulsión para la generación efectiva de la interacción proteína-proteína, en el caso estudiado hospedero-patógeno viral. El pI genera un balance de cargas positivas y negativas y genera que en este punto la interacción sea óptima entre proteína-proteína.

En los casos de interacción hospedero-patógeno caracterizados, los pI se establecen para la proteína tanto en estado desenrollado como en el estado enrollado, como se presenta en la tabla 9, estos datos permiten entender mejor la naturaleza química del ambiente donde sería más efectiva la interacción, de allí la motivación para realizar este análisis tan específico.⁷⁶

Datos experimentales para calcular el pI de una proteína e incluso de una interacción de proteínas son muy complicados de obtener, esto es debido al diverso número de variables que cambian en el ambiente donde se realizar la medición y calculo de esta, por tanto, se recurre a medios como los algoritmos usados en este estudio. El pI, para estos algoritmos se calcula teniendo en

cuenta la fracción entre residuos ácidos y básicos de los grupos ionizables.⁷²

Al contrastar los datos de la tabla 8 y la tabla 9, se puede observar que los pH de estabilización máxima de las PPIs no coinciden con los pI de estas. Esto se puede deber a que los pI no son necesariamente los puntos donde se genera la máxima estabilidad para que las interacciones se realicen entre proteínas de los hospederos y los patógenos, pero si los ambientes de pH adecuados para que se den interacciones más fuertes entre ellas y puedan dar pie al mecanismo de infección por parte del virus.

En la tabla 9, se presentan los datos de los pI de cada interacción estudiada, como el algoritmo propKa permitió obtenerlos, en estado desenrollado y en estado enrollado. En general, se presentan valores de pI entre 4-7, correspondiente al comportamiento típico de una proteína.⁷⁶

En la figura 10, se pueden observar las gráficas de carga vs. pH de tres de las interacciones estudiadas, los cambios entre las dos gráficas que representan los estados enrollados y desenrollados, son ligeros y como se puede verificar en la tabla 9 estos son menores a 1 unidad de pH.

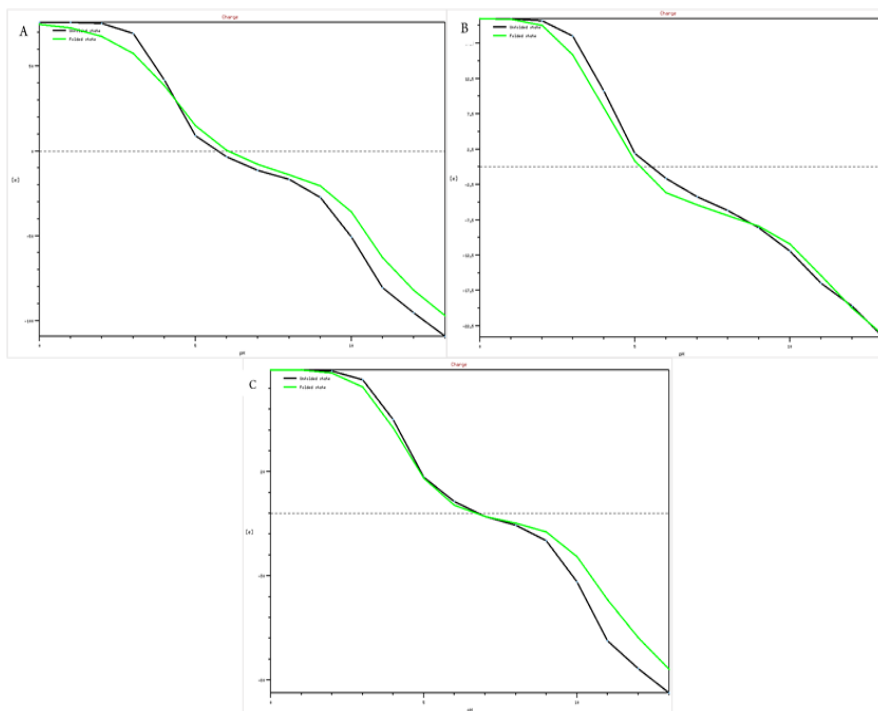


Figura 10. Gráficos correspondientes la carga electrostática vs. pH para determinación del punto isoeléctrico (pI) A) corresponde a la interacción en el complejo Calcineurina con el virus de la fiebre porcina africana. B) corresponde a la interacción entre la proteína HAUSP en hospedero humano con el virus EBNA1 C) corresponde a la interacción entre la proteína Retinoblastoma y la oncoproteína causante del cáncer de cuello uterino. E7. (HPV-E7)

Tabla 9. Datos de punto isoelectrico determinado para diversas interacciones proteína-proteína tipo hospedero-patógeno para estado enrollado y estado desenrollado

INTERACCIÓN	PDB	Enrollado	Desenrollado
4F0Z (CALCINEURINA con African Swine Fever Virus A238L)	4F0Z	5,78	5,34
1YY6 (Deubiquitina con EBNA1, Epstein Barr Virus)	1YY6	5,13	5,42
1QJB (Fosfopéptidos con Murine Polioma Virus)	1QJB	5,11	5,12
5JER (IRF-3 con Rotavirus)	5JER	5,42	5,01
2R02 (Alix-AIP1 con HIV-1)	2R02	5,78	5,76
2R03 (alix-AIP1 con YDPL)	2R03	5,65	5,59
2KPL (MAGI-1 PDZ1/ E6CT)	2KPL	6,62	6,76
2I04 (MAGI-1 PDZ1/ HPV18-E6)	2I04	6,1	6
2I0L (PDZ2 CON HPV E6)	2I0L	6,74	7,03
2I0I (SAP 97 PDZ 3 CON HPV 18 E6)	2I0I	6,99	7,26
2QOS (PDZ2 CON HPV-18)	2QOS	7,55	7,99
1GH6 (Rb con SV40 Large T)	1GH6	6,65	6,75
2GTH (nsp15 con MHV)	2GTH	6,15	5,79
1GUX (Rb con E7 LxCxE)	1GUX	8,76	8,35
2R7G (Rb con E1A CR1)	2R7G	8,35	8,16
1EFN (HIV-1 Nef con R961 FYN SH3)	1EFN	6,2	6,18
1CYZ (TRAF2 LMP1)	1CYZ	7,77	7,73
2J09 (WW domain con Eps-Barr Virus 2A)	2J09	4,31	4,52
2HG0 (West-Nile Virus)	2HG0	6,73	6,93
2JDQ (Influenza Virus A con PB2 con importina alpha 5)	2JDQ	6,22	5,59
1Q1T (Mouse importin Alpha: SV40 CN)	1Q1T	6,93	7,16
4WPH (USP7 con ICP0 Herpes Virus Simple 1)	4WPH	7,41	6,93

3.8. Estadística descriptiva aplicada al contenido de aminoácidos en las secuencias pertenecientes las proteínas hospederas (zonas de interacción con el virus) y proteínas patógenas.

El análisis de estadística descriptiva se realizó creando una matriz que permitía determinar con 0 y 1 la ausencia o la presencia de cierto aminoácido en la secuencia fuera del proteína hospedera o de la proteína patógena.

3.8.1. Contenido de aminoácidos en la secuencia perteneciente a la zona de contacto de la proteína hospedera con la proteína patógena.

La tabla 10, presenta el conteo de variables discretas (aminoácidos) presentes en las secuencias de las zonas de interacción de la proteína hospedera. Estos análisis fueron realizados a través de la herramienta estadística Minitab.

Como se puede observar en la tabla 10, el aminoácido presente en las zonas de contacto de la proteína hospedera que más se encuentra en todas las interacciones es la Fenilalanina (Phe, F), que se encuentra 14 de las 22 interacciones, seguido por Asparagina (Asn, N) encontrándose en 12 de las 22 interacciones. Y Serina (Ser, S) y Ácido Glutámico (Glu o E) ambos encontrándose en 10 de las 22 interacciones.

Ahora bien, la fenilalanina al ser el aminoácido encontrado en la mayoría de las secuencias de las proteínas hospederas, podría explicar el comportamiento de estas y el tipo de enlaces que se emplean para hacer efectivo el proceso de infección por parte del virus. La figura 11, presenta la estructura de la fenilalanina, la cual posee un gran anillo aromático por el cuál puede realizar muchos enlace de tipo hidrofóbico o enlaces de Van der Waals, característicos de este tipo de interacciones.⁶⁸

La presencia de como Asparagina y Ácido Glutámico dan explicación a tipos de enlace como el de hidrógeno y puentes electrostáticos, también presentes en estas interacciones. Por último, la presencia de un aminoácido como la serina, puede explicarse a través del hecho de que permite cierta flexibilidad dentro de la zona de interacción adicional a que es un aminoácido de carácter polar que permite una interacción por enlaces como los previamente mencionados.⁷¹

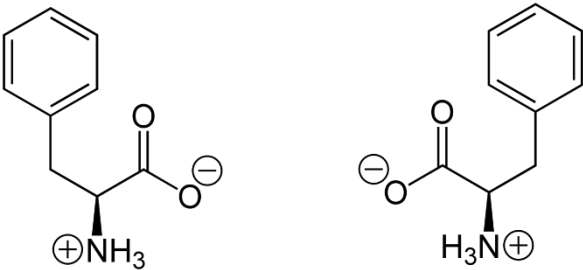


Figura 11. Fenilalanina en configuración de Zwitterion.⁷⁷

Tabla 10. Conteo de variables discretas, aminoácidos contenidos en las secuencias de la zona de interacción en proteínas hospederas

S	Conteo	Porcentaje	T	Conteo	Porcentaje	Q	Conteo	Porcentaje
0	11	50,00	0	16	72,73	0	18	81,82
1	11	50,00	1	6	27,27	1	4	18,18
N=	22		N=	22		N=	22	
N	Conteo	Porcentaje	Y	Conteo	Porcentaje	G	Conteo	Porcentaje
0	10	45,45	0	12	54,55	0	12	54,55
1	12	54,55	1	10	45,45	1	10	45,45
N=	22		N=	22		N=	22	
C	Conteo	Porcentaje	A	Conteo	Porcentaje	V	Conteo	Porcentaje
0	21	95,45	0	14	63,64	0	12	54,55
1	1	4,55	1	8	36,36	1	10	45,45
N=	22		N=	22		N=	22	
L	Conteo	Porcentaje	I	Conteo	Porcentaje	M	Conteo	Porcentaje
0	13	59,09	0	14	63,64	0	18	81,82
1	9	40,91	1	8	36,36	1	4	18,18
N=	22		N=	22		N=	22	
P	Conteo	Porcentaje	F	Conteo	Porcentaje	W	Conteo	Porcentaje
0	18	81,82	0	8	36,36	0	15	68,18
1	4	18,18	1	14	63,64	1	7	31,82
N=	22		N=	22		N=	22	
D	Conteo	Porcentaje	E	Conteo	Porcentaje	K	Conteo	Porcentaje
0	13	59,09	0	11	50,00	0	15	68,18
1	9	40,91	1	11	50,00	1	7	31,82
N=	22		N=	22		N=	22	
H	Conteo	Porcentaje	R	Conteo	Porcentaje			
0	14	63,64	0	15	68,18			

3.8.2. Contenido de aminoácidos en la secuencia perteneciente a la proteína del patógeno viral.

La tabla 11, presenta los resultados del conteo de aminoácidos pertenecientes la secuencia del patógeno viral, al igual que en literal anterior, todo se obtuvo haciendo uso de la herramienta Minitab.

Como se puede observar en la tabla 11, el aminoácido que se encuentra mayoritariamente en las secuencias de las proteínas virales que interactúan en los 22 casos de estudio trabajados, es el Ácido Glutámico (Glu, E) estando presente en 16 de las 22 interacciones, seguido de la Treonina (Thr, T) presente en 13 de las 22 interacciones. Seguido por la Prolina (Pro, P) y por la Arginina (Arg, R).

La existencia de un aminoácido como el Ácido Glutámico, un aminoácido de carácter mayoritariamente ácido en las secuencias de las proteínas de los patógenos, podría deberse al carácter de estas y al tipo de enlace que el patógeno intenta formar con la proteína hospedera tipo enlaces puentes salinos, enlaces de hidrógeno, los cuáles más fácilmente se harán con la proteína hospedera.

Ahora bien, el ácido glutámico puede tener muy buena interacción con la Asparagina de las proteínas hospederas o con la Serina de estas en enlaces de hidrógeno. Por otro lado, se encuentra también la treonina que corresponde a un aminoácido de tipo polar, relativamente pequeño y que brinda flexibilidad para la conformación de la estructura y la formación de enlaces. El grupo OH que este posee puede formar varios enlaces de hidrógeno con la proteína hospedera.

La prolina, sin embargo, es uno de los aminoácidos que, aunque no genera flexibilidad al enlace, permite una unión por medio de enlaces de hidrógeno para las interacciones entre hospederos y patógenos, estos favorecidos por la presencia de un anillo nitrogenado. La presencia de la Arginina contrasta con la presencia de Ácido Glutámico en las proteínas hospederas creando enlaces de carácter electrostático o puentes salinos.^{7370,71}

En la figura 12, se pueden apreciar la estructura de la Prolina y su anillo nitrado que favorece enlaces tipo hidrógeno y del ácido glutámico que es el mayoritario en las proteínas patógenas virales.

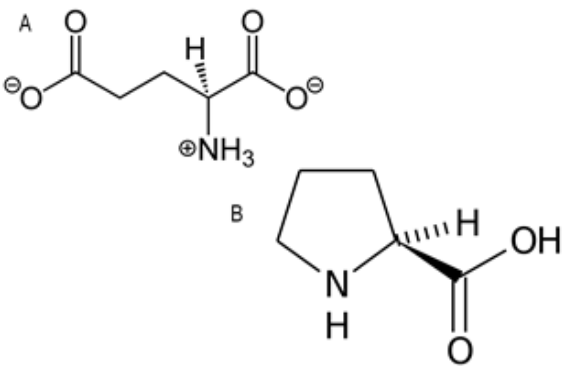


Figura 12. Estructuras de ácido glutámico (A) y Prolina (B)^{78,79}

Tabla 11. Conteo de variables discretas, aminoácidos contenidos en las secuencias de la proteína patógena viral.

S	Conteo	Porcentaje	T	Conteo	Porcentaje	Q	Conteo	Porcentaje
0	13	59,09	0	9	40,91	0	12	54,55
1	9	40,91	1	13	59,09	1	10	45,45
N=	22		N=	22		N=	22	
N	Conteo	Porcentaje	Y	Conteo	Porcentaje	G	Conteo	Porcentaje
0	17	77,27	0	14	63,64	0	20	90,91
1	5	22,73	1	8	36,36	1	2	9,09
N=	22		N=	22		N=	22	
C	Conteo	Porcentaje	A	Conteo	Porcentaje	V	Conteo	Porcentaje
0	17	77,27	0	17	77,27	0	14	63,64
1	5	22,73	1	5	22,73	1	8	36,36
N=	22		N=	22		N=	22	
L	Conteo	Porcentaje	I	Conteo	Porcentaje	M	Conteo	Porcentaje
0	14	63,64	0	17	77,27	0	20	90,91
1	8	36,36	1	5	22,73	1	2	9,09
N=	22		N=	22		N=	22	
P	Conteo	Porcentaje	F	Conteo	Porcentaje	W	Conteo	Porcentaje
0	11	50,00	0	20	90,91	0	20	90,91
1	11	50,00	1	2	9,09	1	2	9,09
N=	22		N=	22		N=	22	
D	Conteo	Porcentaje	E	Conteo	Porcentaje	K	Conteo	Porcentaje
0	16	72,73	0	6	27,27	0	17	77,27
1	6	27,27	1	16	72,73	1	5	22,73
N=	22		N=	22		N=	22	
H	Conteo	Porcentaje	R	Conteo	Porcentaje			
0	19	86,36	0	11	50,00			
1	3	13,64	1	11	50,00			
N=	22		N=	22				

4. CONCLUSIONES.

A partir de los resultados presentados, se puede decir que la caracterización química In-Silico de interacciones proteína-proteína tipo dominio-motivo para el estudio de actividades patógeno-hospedero en modelos virales es posible gracias a diversas al desarrollo de la metodología implementada de para el estudio.

Para resumir, las interacciones proteína-proteína tipo dominio-motivo en actividades patógeno-hospedero en modelos virales, son interacciones que presentan características tales como, motivos lineales repetidos y conservados, que permiten identificar fácilmente a las proteínas virales y así obtener unas secuencias de aminoácidos correspondientes a la estructura primaria de una proteína tanto hospedera como viral. El obtener la estructura primaria de las zonas de interacción entre hospederos y patógenos permite determinar que estructuras secundarias van a interactuar y se verificaría esto con el cálculo de área superficial accesible de solvente (SASA).

Adicionalmente, estos motivos lineales conservados, pueden ser caracterizados y estudiados teniendo en cuenta su anclaje en las zonas de interacción hospedero-patógeno y esto permite determinar cuáles de estos generarán enlaces que serán de tipo enlace de hidrógeno, puente salino y puente salino, además de interacciones hidrófobas como enlaces de Van der Waals. Enlaces; este tipo de enlaces permiten relacionar este tipo específico de interacción entre proteínas (interacción exógena) con el resto de las interacciones (interacciones endógenas).

Por último, los análisis de estabilidad teniendo en cuenta el pH y el punto isoeléctrico dan respuesta a como se hace más estable este tipo de interacciones. Y, el análisis estadístico ayuda a entender como el contenido de aminoácidos de las secuencias en las zonas de interacción de proteínas hospederas y proteínas patógenas tienen una relación directa en como se dan los tipos de enlaces y las estructuras secundarias.

Las interacciones entre proteínas hospederas y patógenos es un campo al que todavía le queda mucha investigación y el uso de la caracterización de las estructuras tridimensionales desde un punto de vista de descripción de factores fisicoquímicos que intervienen en las interacciones, adicionalmente, a estudios de la naturaleza netamente química de las interacciones pueden dar resultados como los presentados anteriormente. Es de vital importancia continuar con este tipo de investigaciones, debido a que pueden abrir puertas para entender más sobre ellas y sus características generales.

5. AGRADECIMIENTOS.

Los autores agradecen a la Universidad del Valle, a la escuela de Ingeniería de Sistemas y Ciencias de la computación y al grupo de Bioinformática.

6. CONFLICTO DE INTERESES.

Los autores declaran no tener conflicto de intereses de ningún tipo de relación a la publicación.

7. REFERENCIAS

- (1) Franzosa, E. A.; Xia, Y. Structural Principles within the Human-Virus Protein-Protein Interaction Network. *Proc. Natl. Acad. Sci.* **2011**, *108* (26), 10538–10543. <https://doi.org/10.1073/pnas.1101440108>.
- (2) Durmus, S.; Cakir, A.; Özgür, A.; Guthke, R. A Review on Computational Systems Biology of Pathogen-Host Interactions. *Front Microbiol* **2015**, *6*, 235.
- (3) Zoraghi, R.; Reiner, N. E. Protein Interaction Networks as Starting Points to Identify Novel Antimicrobial Drug Targets. *Curr. Opin. Microbiol.* **2013**.
- (4) de Chassey, B.; Meyniel-Schicklin, L.; Aublin-Gex, A.; Andre, P.; Lotteau, V. New Horizons for Antiviral Drug Discovery from Virus-Host Protein Interaction Networks. *Curr. Opin. Virol.* **2012**, *2* (5), 606–613.
- (5) Franzosa, E. a; Xia, Y. Structural Models for Host-Pathogen Protein-Protein Interactions: Assessing Coverage and Bias. *Pac. Symp. Biocomput.* **2012**, 287–298. https://doi.org/10.1142/9789814366496_0028.
- (6) Gould, C. M.; Diella, F.; Via, A.; Puntervoll, P.; Gemünd, C.; Chabanis-Davidson, S.; Michael, S.; Sayadi, A.; Bryne, J. C.; Chica, C.; Seiler, M.; Davey, N. E.; Haslam, N.; Weatheritt, R. J.; Budd, A.; Hughes, T.; Paś, J.; Rychlewski, L.; Travé, G.; Aasland, R.; Helmer-Citterich, M.; Linding, R.; Gibson, T. J. ELM: The Status of the 2010 Eukaryotic Linear Motif Resource. *Nucleic Acids Res.* **2009**, *38* (SUPPL.1), 167–180. <https://doi.org/10.1093/nar/gkp1016>.
- (7) Becerra, A.; Bucheli, V. A.; Moreno, P. A. Prediction of Virus-Host Protein-Protein Interactions Mediated by Short Linear Motifs. *BMC Bioinformatics* **2017**, *18* (1), 163. <https://doi.org/10.1186/s12859-017-1570-7>.
- (8) European Bioinformatics Institute. Structural Motifs | <https://www.ebi.ac.uk/training/online/course/biomolecular-structures-introduction-ebi-reso/proteins/structural-motifs> (accessed Nov 8, 2017).
- (9) Kay, B. K.; Williamson, M. P.; Sudol, M. The Importance of Being Proline: The Interaction of Proline-Rich Motifs in Signaling Proteins with Their Cognate Domains. *FASEB J.* **2000**, *14* (2), 231–241.
- (10) Sandoval, A. Characterizing Viruses Mimicry Mechanisms, Universidad del Valle, 2016.
- (11) Siegert, T. R.; Bird, M. J.; Makwana, K. M.; Kritzer, J. A. Analysis of Loops That Mediate Protein-Protein Interactions and Translation into Submicromolar Inhibitors. *J. Am. Chem. Soc.* **2016**, *138* (39), 12876–

12884. <https://doi.org/10.1021/jacs.6b05656>.
- (12) Elde, N. C.; Malik, H. S. The Evolutionary Conundrum of Pathogen Mimicry. *Nature Reviews Microbiology*. 2009. <https://doi.org/10.1038/nrmicro2222>.
 - (13) EUROPEAN BIOINFORMATICS INSTITUTE, E. Domains | EMBL-EBI Train online <https://www.ebi.ac.uk/training/online/course/biomacromolecular-structures-introduction-ebi-reso/proteins/domains> (accessed Nov 8, 2017).
 - (14) Segura-Cabrera, A.; García-Perez, C.; Guo, X.; Rodríguez Perez, M. A.; García-Pérez, C. A.; Guo, X.; Rodríguez-Pérez, M. A. A Viral-Human Interactome Based on Structural Motif-Domain Interactions Captures the Human Infectome. *PLoS One* **2013**, *8* (8). <https://doi.org/10.1371/journal.pone.0071526>.
 - (15) Segura-Cabrera, A.; Bocanegra-García, V.; Lizarazo-Ortega, C.; Guo, X.; Correa-Basurto, J.; Rodríguez-Pérez, M. A. A Computational Analysis of the Binding Mode of Closantel as Inhibitor of the Onchocerca Volvulus Chitinase: Insights on Macrophilicidal Drug Design. *J. Comput. Aided. Mol. Des.* **2011**, *25* (12), 1107–1119. <https://doi.org/10.1007/s10822-011-9489-y>.
 - (16) Jean Beltran, P. M.; Federspiel, J. D.; Sheng, X.; Cristea, I. M. Proteomics and Integrative Omic Approaches for Understanding Host–pathogen Interactions and Infectious Diseases. *Mol. Syst. Biol.* **2017**, *13* (3), 922. <https://doi.org/10.15252/msb.20167062>.
 - (17) Stojanovic, S. D.; Zaric, S. D. Hydrogen Bonds and Hydrophobic Interactions of Porphyrins in Porphyrin-Containing Proteins. *Open Struct. Biol. J.* **2009**, *3* (1), 34–41. <https://doi.org/10.2174/1874199100903010034>.
 - (18) Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114* (13), 6589–6631. <https://doi.org/10.1021/cr400525m>.
 - (19) Roth, C. M.; Neal, B. L.; Lenhoff, A. M. Van Der Waals Interactions Involving Proteins. *Biophys. J.* **1996**, *70* (2 1), 977–987. [https://doi.org/10.1016/S0006-3495\(96\)79641-8](https://doi.org/10.1016/S0006-3495(96)79641-8).
 - (20) Gavenovis, J.; Sheneman, B.; Siegert, T.; Eshelman, M.; Kritzer, J. Comprehensive Analysis of Loops at Protein-Protein Interfaces. *Nat Chem Biol* **2014**, *10* (9), 716–722. <https://doi.org/10.1016/j.gde.2016.03.011>.
 - (21) ELM - Browse/Search instances <http://elm.eu.org/viruses/> (accessed Apr 28, 2020).
 - (22) Schrödinger LCC. The PyMol Molecular Graphics System. 2015.
 - (23) DeLano, W.; Bromberg, S. PyMOL User's Guide (Original). **2004**, 1–66.
 - (24) Makarewicz, T.; Kaźmierkiewicz, R. Improvements in GROMACS Plugin for PyMOL Including Implicit Solvent Simulations and Displaying Results of PCA Analysis. *J. Mol. Model.* **2016**, *22* (5). <https://doi.org/10.1007/s00894-016-2982-4>.
 - (25) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32* (WEB SERVER ISS.), 665–667. <https://doi.org/10.1093/nar/gkh381>.
 - (26) Davies, M. N.; Toseland, C. P.; Moss, D. S.; Flower, D. R. Benchmarking pKa Prediction. *BMC Biochem.* **2006**, *7*, 1–12. <https://doi.org/10.1186/1471-2091-7-18>.
 - (27) Durmuş, S.; Ülgen, K. Comparative Interactomics for Virus–human Protein–protein Interactions: DNA Viruses versus RNA Viruses. *FEBS Open Bio* **2017**, *7* (1), 96–107. <https://doi.org/10.1002/2211-5463.12167>.
 - (28) Nybakken, G. E.; Nelson, C. A.; Chen, B. R.; Diamond, M. S.; Fremont, D. H. Crystal Structure of the West Nile Virus Envelope Glycoprotein. *J. Virol.* **2006**, *80* (23), 11467–11474. <https://doi.org/10.1128/jvi.01125-06>.
 - (29) Liu, X.; Marmorstein, R. Structure of the Retinoblastoma Protein Bound to Adenovirus E1A Reveals the Molecular Basis for Viral Oncoprotein Inactivation of a Tumor Suppressor. *Genes Dev.* **2007**, *21* (21), 2711–2716. <https://doi.org/10.1101/gad.1590607>.
 - (30) Structural Basis for the Inactivation of Retinoblastoma Tumor Suppressor by SV40 Large T Antigen. *EMBO J.* **2001**, *20* (1), 295–304.
 - (31) Xu, X.; Zhai, Y.; Sun, F.; Lou, Z.; Su, D.; Xu, Y.; Zhang, R.; Joachimiak, A.; Zhang, X. C.; Bartlam, M.; Rao, Z. New Antiviral Target Revealed by the Hexameric Structure of Mouse Hepatitis Virus Nonstructural Protein nsp15. *J. Virol.* **2006**, *80* (16), 7909–7917. <https://doi.org/10.1128/jvi.00525-06>.
 - (32) Lee, J.-O.; Russo, A. A.; Pavletich, N. P. Structure of

the Retinoblastoma Tumour-Suppressor Pocket Domain Bound to a Peptide from HPV E7; 1998; Vol. 391.

- (33) Charbonnier, S.; Nominé, Y.; Ramírez, J.; Luck, K.; Chapelle, A.; Stote, R. H.; Travé, G.; Kieffer, B.; Atkinson, R. A. The Structural and Dynamic Response of MAGI-1 PDZ1 with Noncanonical Domain Boundaries to the Binding of Human Papillomavirus E6. *J. Mol. Biol.* **2011**, *406* (5), 745–763. <https://doi.org/10.1016/j.jmb.2011.01.015>.
- (34) Liu, Y.; Henry, G. D.; Hegde, R. S.; Baleja, J. D. Solution Structure of the hDlg/SAP97 PDZ2 Domain and Its Mechanism of Interaction with HPV-18 Papillomavirus E6 Protein. *Biochemistry* **2007**, *46* (38), 10864–10874. <https://doi.org/10.1021/bi700879k>.
- (35) FONTES, M. R. M.; TEH, T.; TOTH, G.; JOHN, A.; PAVO, I.; JANS, D. A.; KOBE, B. Role of Flanking Sequences and Phosphorylation in the Recognition of the Simian-Virus-40 Large T-Antigen Nuclear Localization Sequences by Importin- α . *Biochem. J.* **2003**, *375* (2), 339–349. <https://doi.org/10.1042/bj20030510>.
- (36) Zhai, Q.; Fisher, R. D.; Chung, H. Y.; Myszk, D. G.; Sundquist, W. I.; Hill, C. P. Structural and Functional Studies of ALIX Interactions with YPX N L Late Domains of HIV-1 and EIAV. *Nat. Struct. Mol. Biol.* **2008**, *15* (1), 43–49. <https://doi.org/10.1038/nsmb1319>.
- (37) Rittinger, K.; Budman, J.; Xu, J.; Volinia, S.; Cantley, L. C.; Smerdon, S. J.; Gamblin, S. J.; Yaffe, M. B. Structural Analysis of 14-3-3 Phosphopeptide Complexes Identifies a Dual Role for the Nuclear Export Signal of 14-3-3 in Ligand Binding. *Mol. Cell* **1999**, *4* (2), 153–166. [https://doi.org/10.1016/S1097-2765\(00\)80363-9](https://doi.org/10.1016/S1097-2765(00)80363-9).
- (38) Tarendeau, F.; Boudet, J.; Guilligay, D.; Mas, P. J.; Bougault, C. M.; Boulo, S.; Baudin, F.; Ruigrok, R. W. H.; Daigle, N.; Ellenberg, J.; Cusack, S.; Simorre, J. P.; Hart, D. J. Structure and Nuclear Import Function of the C-Terminal Domain of Influenza Virus Polymerase PB2 Subunit. *Nat. Struct. Mol. Biol.* **2007**, *14* (3), 229–233. <https://doi.org/10.1038/nsmb1212>.
- (39) Ciferri, C.; Pasqualato, S.; Screpanti, E.; Varet, G.; Santaguida, S.; Dos Reis, G.; Maiolica, A.; Polka, J.; De Luca, J. G.; De Wulf, P.; Salek, M.; Rappsilber, J.; Moores, C. A.; Salmon, E. D.; Musacchio, A. Implications for Kinetochore-Microtubule Attachment from the Structure of an Engineered Ndc80 Complex. *Cell* **2008**, *133* (3), 427–439. <https://doi.org/10.1016/j.cell.2008.03.020>.
- (40) Zhao, B.; Gao, X.; Du, F.; Shu, C.; Sankaran, B.; Li, P.; Shelton, C. L.; Herr, A. B.; Ji, J.-Y.; Gao, X.; Sankaran, B.; Du, F.; Shelton, C. L.; Herr, A. B.; Ji, J.-Y.; Li, P. Structural Basis for Concerted Recruitment and Activation of IRF-3 by Innate Immune Adaptor Proteins. *Proc. Natl. Acad. Sci.* **2016**, *113* (24), E3403–E3412. <https://doi.org/10.1073/pnas.1603269113>.
- (41) Ye, H.; Park, Y. C.; Kreishman, M.; Kieff, E.; Wu, H. *The Structural Basis of the Recognition of Diverse Receptor Sequences by TRAF2*; New York, New York, 1999; Vol. 4.
- (42) Grigoriu, S.; Bond, R.; Cossio, P.; Chen, J. A.; Ly, N.; Hummer, G.; Page, R.; Cyert, M. S.; Peti, W. The Molecular Mechanism of Substrate Engagement and Immunosuppressant Inhibition of Calcineurin. *PLoS Biol.* **2013**, *11* (2). <https://doi.org/10.1371/journal.pbio.1001492>.
- (43) Morales, B.; Ramirez-Espain, X.; Shaw, A. Z.; Martin-Malpartida, P.; Yraola, F.; Sánchez-Tilló, E.; Farrera, C.; Celada, A.; Royo, M.; Macias, M. J. NMR Structural Studies of the ItchWW3 Domain Reveal That Phosphorylation at T30 Inhibits the Interaction with PPxY-Containing Ligands. *Structure* **2007**, *15* (4), 473–483. <https://doi.org/10.1016/j.str.2007.03.005>.
- (44) Zhang, Y.; Dasgupta, J.; Ma, R. Z.; Banks, L.; Thomas, M.; Chen, X. S. Structures of a Human Papillomavirus (HPV) E6 Polypeptide Bound to MAGUK Proteins: Mechanisms of Targeting Tumor Suppressors by a High-Risk HPV Oncoprotein. *J. Virol.* **2007**, *81* (7), 3618–3626. <https://doi.org/10.1128/jvi.02044-06>.
- (45) Lee, C.-H.; Saksela, K.; Mirza, U. A.; Chait, B. T.; Kuriyan, J. *Crystal Structure of the Conserved Core of HIV-1 NEF Complexed with a SRC Family SH3 Domain*; New York, New York, 1996; Vol. 85.
- (46) Zhang, R. G.; Lee, W.; Shire, K.; Holowaty, M. N.; Sarkari, F.; Arrowsmith, C. H.; Liao, J.; Edwards, A. M.; Nguyen, T.; Frappier, L.; Sheng, Y.; Saridakis, V.; Sheng, Y.; Sarkari, F.; Holowaty, M. N.; Shire, K.; Nguyen, T.; Zhang, R. G.; Liao, J.; Lee, W.; Edwards, A. M.; Arrowsmith, C. H.; Frappier, L. Structure of the p53 Binding Domain of HAUSP/USP7 Bound to Epstein-Barr Nuclear Antigen 1. *Mol. Cell* **2005**, *18* (1), 25–36. <https://doi.org/10.1016/j.molcel.2005.02.029>.
- (47) Shakin-Eshleman, S. H.; Spitalnik, S. L.; Kasturi, L. The Amino Acid at the X Position of an Asn-X-Ser Sequon Is an Important Determinant of N-Linked Core-Glycosylation Efficiency. *J. Biol. Chem.* **1996**, *271* (11), 6363–6366. <https://doi.org/10.1074/jbc.271.11.6363>.

- (48) Honda, A.; Mizumoto, K.; Ishihama, A. Two Separate Sequences of PB2 Subunit Constitute the RNA Cap-Binding Site of Influenza Virus RNA Polymerase. *Genes to Cells* **1999**, 4 (8), 475–485. <https://doi.org/10.1046/j.1365-2443.1999.00275.x>.
- (49) Mukaigawa, J.; Nayak, D. P. Two Signals Mediate Nuclear Localization of Influenza Virus (A/WSN/33) Polymerase Basic Protein 2. *J. Virol.* **1991**, 65 (1), 245–253.
- (50) Pfoh, R.; Lacdao, I. K.; Georges, A. A.; Capar, A.; Zheng, H.; Frappier, L.; Saridakis, V. Crystal Structure of USP7 Ubiquitin-like Domains with an ICP0 Peptide Reveals a Novel Mechanism Used by Viral and Cellular Proteins to Target USP7. *PLoS Pathog.* **2015**, 11 (6), 1–23. <https://doi.org/10.1371/journal.ppat.1004950>.
- (51) Franzosa, E. A.; Xia, Y.; Michael Levitt, B. Structural Principles within the Human-Virus Protein-Protein Interaction Network. *Proc. Natl. Acad. Sci.* **2011**, 108 (26).
- (52) Hagai, T.; Azia, A.; Babu, M. M.; Andino, R. Use of Host-like Peptide Motifs in Viral Proteins Is a Prevalent Strategy in Host-Virus Interactions. *Cell Rep.* **2014**, 7 (5), 1729–1739. <https://doi.org/10.1016/j.celrep.2014.04.052>.
- (53) Milroy, L.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. Modulators of Protein – Protein Interactions. **2014**. <https://doi.org/10.1021/cr400698c>.
- (54) Remaut, H.; Waksman, G. Protein-Protein Interaction through β -Strand Addition. *Trends in Biochemical Sciences* 2006, pp 436–444. <https://doi.org/10.1016/j.tibs.2006.06.007>.
- (55) Finkel, T.; Duc, J.; Fearon, E. R.; Dang, C. V.; Tomaselli, G. F. Detection and Modulation in Vivo of Helix-Loop-Helix Protein-Protein Interactions. *J. Biol. Chem.* **1993**, 268 (1), 5–8.
- (56) Durham, E.; Dorr, B.; Woetzel, N.; Staritzbichler, R.; Meiler, J. Solvent Accessible Surface Area Approximations for Rapid and Accurate Protein Structure Prediction. *J. Mol. Model.* **2009**, 15 (9), 1093–1108. <https://doi.org/10.1007/s00894-009-0454-9>.
- (57) Poulin, G.; Lebel, M.; Chamberland, M.; Paradis, F. W.; Drouin, J. Specific Protein-Protein Interaction between Basic Helix-Loop-Helix Transcription Factors and Homeoproteins of the Pitx Family. *Mol. Cell. Biol.* **2000**, 20 (13), 4826–4837. <https://doi.org/10.1128/mcb.20.13.4826-4837.2000>.
- (58) Roth, C. M.; Neal, B. L.; Lenhoff, A. M. Van Der Waals Interactions Involving Proteins. *Biophys. J.* **1996**, 70 (February). [https://doi.org/10.1016/S0006-3495\(96\)79641-8](https://doi.org/10.1016/S0006-3495(96)79641-8).
- (59) Sung, S. S. Peptide Folding Driven by Van Der Waals Interactions. *Protein Sci.* **2015**, 24 (9), 1383–1388. <https://doi.org/10.1002/pro.2710>.
- (60) Sung, S.-S. Peptide Folding Driven by Van Der Waals Interactions. **2015**. <https://doi.org/10.1002/pro.2710>.
- (61) Casasola, R. Electrospinning of Poly (Lactic) Acid for Biomedical Applications: Analysis of Solution Properties and Process Parameters, Drug Encapsulation and Release. **2016**.
- (62) Waugh, D. F. Protein-Protein Interactions. *Adv. Protein Chem.* **1954**, 9, 325–437. [https://doi.org/10.1016/S0065-3233\(08\)60210-7](https://doi.org/10.1016/S0065-3233(08)60210-7).
- (63) Mariano, R.; Wuchty, S. Structure-Based Prediction of Host-???pathogen Protein Interactions. *Curr. Opin. Struct. Biol.* **2017**, 44, 119–124. <https://doi.org/10.1016/j.sbi.2017.02.007>.
- (64) Vinogradov, S. Hydrogen Bonds in Crystal Structures of Amino Acids, Peptides and Related Molecules. *Int. J. Pept. Protein Res* **1979**, 14 (1959), 281–289.
- (65) Vinogradov, S. N. *HYDROGEN BONDS IN CRYSTAL S AMINO ACIDS, PEPTIDES A N D RELA*; 1979; Vol. 14.
- (66) Scheiner, S.; Kar, T.; Gu, Y. Strength of the CaH...O Hydrogen Bond of Amino Acid Residues. *J. Biol. Chem.* **2001**, 276 (13), 9832–9837. <https://doi.org/10.1074/jbc.M010770200>.
- (67) Yang, A. S.; Honig, B. On the pH Dependence of Protein Stability. *Journal of Molecular Biology.* 1993, pp 459–474. <https://doi.org/10.1006/jmbi.1993.1294>.
- (68) Gromiha, M. M. Protein Stability. In *Protein Bioinformatics from sequence to function*; Elsevier: Tokio, Japan., 2010; pp 209–245.
- (69) Gromiha, M. M.; Nagarajan, R.; Selvaraj, S. Protein Structural Bioinformatics: An Overview. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **2019**, 2, 445–459. <https://doi.org/10.1016/B978-0-12-809633-8.20278-1>.
- (70) Rostkowski, M.; Olsson, M. H.; Søndergaard, C. R.; Jensen, J. H. Graphical Analysis of pH-Dependent Properties of Proteins Predicted Using PROPKA. *BMC Struct. Biol.* **2011**, 11 (Cc). <https://doi.org/10.1186/1472-6807-11-6>.
- (71) Kazlauskas, R. Engineering More Stable Proteins. *Chem. Soc. Rev.* **2018**, 47 (24), 9026–9045.

<https://doi.org/10.1039/c8cs00014j>.

- (72) Alexov, E. Numerical Calculations of the pH of Maximal Protein Stability: The Effect of the Sequence Composition and Three-Dimensional Structure. *Eur. J. Biochem.* **2004**, 271 (1), 173–185. <https://doi.org/10.1046/j.1432-1033.2003.03917.x>.
- (73) Demuru, S.; Fanelli, A. Politecnico Di Torino Peptides Analysis Using VMD and Its PropKa Extension. **2015**, No. September. <https://doi.org/10.13140/RG.2.1.3524.9121>.
- (74) Obenauer, J. C.; Yaffe, M. B. Computational Prediction of Protein-Protein Interactions. In *Protein-Protein interactions methods and applications*; Fu, H., Ed.; Humana Press, 2004; Vol. 261, pp 445–468.
- (75) Gehring, C.; Davenport, M.; Jaczynski, J. Functional and Nutritional Quality of Protein and Lipid Recovered from Fish Processing by-Products and Underutilized Aquatic Species Using Isoelectric Solubilization / Precipitation. *Curr. Nutr. Food Sci.* **2009**, 5 (1), 17–39. <https://doi.org/10.2174/157340109787314703>.
- (76) Ciborowski, P.; Silberring, J. PROTEOMIC PROFILING AND ANALYTICAL The Crossroads. 2016, p 279.
- (77) La Fenilalanina, ácido, Zwitterion imagen png - imagen transparente descarga gratuita <https://www.freepng.es/png-uof9r1/> (accessed May 6, 2020).
- (78) Descargar libre | Msg ácido glutámico glutamato sal umami, gránulo PNG Clipart | PNGOcean <https://www.pngocean.com/gratis-png-clipart-ewrmg/descargar> (accessed May 6, 2020).
- (79) L-PROLINA | Quimica Alkano <http://quimicaalkano.com/product/l-prolina/> (accessed May 6, 2020).