



UNIVERSIDAD AUTÓNOMA DE GUERRERO
UNIDAD ACADÉMICA DE CIENCIAS QUÍMICO BIOLÓGICAS
Unidad Especializada en Microbiología
UNIDAD ACADÉMICA DE MEDICINA

Av. Lázaro Cárdenas S/N. Ciudad Universitaria. Chilpancingo, Gro. C.P. 39090. Tel 01 747 47 2 55 03 y 01 747 47 1 93 10 ext: 4526



DOCTORADO EN CIENCIAS BIOMÉDICAS

**Análisis estructural y funcional de la oncoproteína E6 del VPH16
y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc)
en la interacción con sus proteínas blanco.**

TESIS

Que para obtener el grado de:

DOCTOR EN CIENCIAS BIOMÉDICAS

P R E S E N T A:

M en C. Hugo Alberto Rodríguez Ruiz

Director: Dr. Marco Antonio Leyva Vázquez
Co-dirección: Dra. Olga Lilia Garibay Cerdaneres

Chilpancingo de los Bravo, Gro., julio 2019.

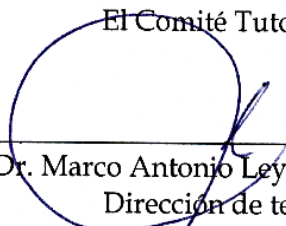


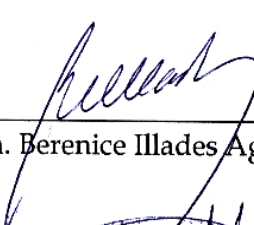
UNIVERSIDAD AUTÓNOMA DE GUERRERO
FACULTAD DE CIENCIAS QUÍMICO BIOLÓGICAS
UNIDAD DE INVESTIGACIÓN ESPECIALIZADA EN MICROBIOLOGÍA
DOCTORADO EN CIENCIAS BIOMÉDICAS


ACTA DE APROBACIÓN DE TESIS

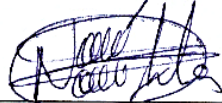
En la ciudad de Chilpancingo, Guerrero, siendo los 05 días del mes de julio del dos mil diecinueve, se reunieron los miembros del Comité Tutorial designado por la Academia de Posgrado del Doctorado en Ciencias Biomédicas, para examinar la tesis titulada "Análisis estructural y funcional de la oncoproteína E6 del VPH 16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc) en la interacción con sus proteína blanco", presentado por el alumno Hugo Alberto Rodríguez Ruíz, para obtener el Grado de Doctor en Ciencias Biomédicas. Después del análisis correspondiente, los miembros del comité manifiestan su aprobación de la tesis, autorizan la impresión final de la misma y aceptan que, cuando se satisfagan los requisitos señalados en el Reglamento General de Estudios de Posgrado e Investigación Vigente, se proceda a la presentación del examen de grado.

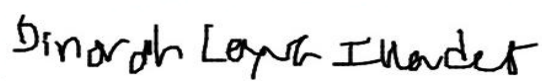
El Comité Tutorial


Dr. Marco Antonio Leyva Vázquez
Dirección de tesis

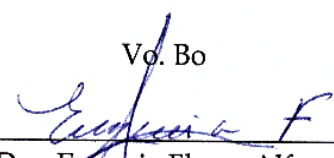

Dra. Berenice Illades Aguiar


Dra. Luz del Carmen Alarcón Romero

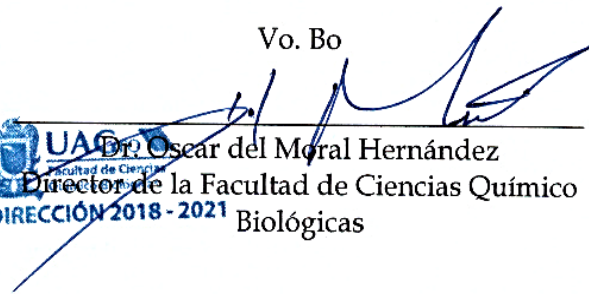

Dr. Napoleón Navarro Tito


Dra. Dinorah Leyva Illades

Vo. Bo


Dra. Eugenia Flores Alfaro
Coordinadora del Doctorado en Ciencias Biomédicas

Vo. Bo


Dr. Oscar del Moral Hernández
Director de la Facultad de Ciencias Químico Biológicas

Este trabajo se realizó en el Laboratorio de Biomedicina Molecular de la Facultad de Ciencias Químico Biológicas de la Universidad Autónoma de Guerrero en Chilpancingo de los Bravo Guerrero y en el Laboratorio de Bioinformática de la Facultad de Ciencias Químico Biológicas de la Universidad Autónoma de Sinaloa.

Bajo la Dirección de:

Dr. Marco Antonio Leyva Vázquez

La Co-dirección de:

Dra. Olga Lilia Garibay Cerdaneres

Dra. Berenice Illades Aguiar

La asesoría de:

Dra. Dinorah Leyva Illades

Dra. Luz del Carmen Alarcón Romero

Dr. Napoleón Navarro Tito

Esta investigación se desarrolló con financiamiento otorgado por CONACYT a través del fondo Ciencia Básica 2016 (288612). Así mismo, con fondo en horas computo por el Laboratorio Nacional de Cómputo de Alto Desempeño (LANCAD 2017-2019, Clúster de Supercómputo Xihcoatl).

Durante el periodo en que curso el Doctorado en Ciencias Biomédicas, el M en C. Hugo Alberto Rodríguez Ruiz, recibió beca del Consejo Nacional de Ciencia y Tecnología con número de registro 557610.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAC), en la interacción con sus proteínas blanco.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc), en la interacción con sus proteínas blanco.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc), en la interacción con sus proteínas blanco.

Índice

Introducción

Capítulo I: Predicción *in silico* de los cambios estructurales de la oncoproteína E6 VPH 16 y sus variantes.

Capítulo II: Dinámica molecular y docking de la interacción de la oncoproteína E6 VPH 16 y sus variantes con p53.

Discusión general.

Conclusión general.

Anexos.

Referencias.

INTRODUCCIÓN

El VPH es un virus de DNA de doble cadena, circular, pequeño de 50 a 55 nanómetros (nm) de diámetro y alrededor de 8 kilobases (Kb) de longitud, presenta un tropismo hacia las células epiteliales, e infecta la piel y las mucosas (Choi and Park, 2016; Lazarczyk et al., 2009). Tiene una cápside icosaédrica compuesta por 72 capsómeros que tienen formas que se asemejan a una estrella de cinco puntos. El genoma del virus está dentro de la cápside, y alberga 8 marcos de lectura que se superponen. Su genoma está dividido en tres regiones: una región temprana (Early: E), una tardía (Late: L) y una región reguladora (Long Control Region: LCR o NonCoding Region: NCR) (Lazarczyk et al., 2009; Zheng and Baker, 2006).

Las tres regiones en las que está dividido el genoma del VPH codifican para diferentes proteínas. La región temprana de aproximadamente 4 Kb, codifica para las proteínas no estructurales: E1, E2, E4, E5, E6 y E7, que son necesarias para los procesos de replicación, transcripción y transformación. Se debe mencionar que las proteínas E6 y E7 son proteínas que juegan un papel importante en el desarrollo del cáncer. Se sabe que la proteína E6 se une al gen supresor de tumor P53, generando así la deficiencia en la reparación del DNA, inhibiendo la apoptosis y generando inestabilidad cromosomal. Por otra parte, la proteína E7 interactúa con varias proteínas celulares, entre ellas la proteína Rb, y también se encuentra involucrada en la inhibición de la apoptosis y la evasión a la respuesta inmune (Doorbar et al., 2015, 2012; Muñoz et al., 2006; Raybould et al., 2011; Woodman et al., 2007).

La región tardía de aproximadamente 3 Kb codifica para dos proteínas estructurales de la cápside, las proteínas L1 y L2; estas dos proteínas son fundamentales para formar los capsómeros y proteger el genoma viral, se expresan al final del ciclo viral y se encuentran en la capa superficial del epitelio. La proteína L1, además de formar los capsómeros junto con la proteína L2 puede unirse ella misma y formar una especie de cápside vacía, que puede generar inmunogenicidad (Chen et al., 2011; Doorbar et al., 2012), mientras que la proteína L2 es necesaria para permitir la entrada del virus a la célula huésped, el transporte de componentes virales dentro

del núcleo y actualmente se sabe que neutraliza anticuerpos contra diferentes tipos de VPH (Pereira et al., 2009).

Finalmente, la región reguladora, la más pequeña de las tres regiones del genoma viral de 1 Kb, no codifica para ninguna proteína viral, sin embargo, contiene elementos reguladores de la expresión y la transcripción viral (Doorbar et al., 2012; Raybould et al., 2011).

Hasta el 2015, el Centro Internacional de Referencia del Papilomavirus Humano (International Human Papillomavirus Reference Center) reportó que existen más de 200 tipos de VPH (www.hpvcenter.net). Los VPH están agrupados en cinco géneros de papilomavirus: Alfa, Beta, Gamma, Mu y Un. El género Alfa-papilomavirus incluye los VPH que infectan mucosa oral y genital, y por lo tanto son los de mayor importancia clínica, a comparación de los otros géneros que están asociados con infecciones cutáneas benignas (Stanley, 2010). Adicionalmente, podemos clasificar a los diferentes tipos de VPH dependiendo de su potencial oncogénico: VPH de Alto Riesgo (VPH-AR): 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, y 59; VPH de Probable Alto Riesgo (VPH-PR): 26, 53, 66, 68, 73, 82; VPH de Bajo Riesgo (VPH-BR): 6, 11, 13, 40, 42, 43, 44, 54, 61, 70, 72, 81 y 89; y VPH de Riesgo Indeterminado (VPH-RI): 2, 3, 7, 10, 27, 28, 29, 30, 32, 34, 55, 57, 62, 67, 69, 71, 74, 77, 83, 84, 85, 86, 87, 90, 91. De la lista de los VPH-AR los VPH 16 y 18 son conocidos como los más carcinogénicos, debido a que están presentes hasta en el 70% de los casos de cáncer de cérvix, vagina, ano, y en el 30-40% de cáncer de vulva y pene (de Villiers et al., 2004; Doorbar et al., 2015, 2012; Hudelist et al., 2004; Muñoz et al., 2006; Stanley, 2010).

El cáncer cervicouterino (CaCU) es el cuarto tipo de cáncer causante de neoplasia maligna en mujeres y uno de los principales causantes de muertes femeninas a nivel mundial, después del cáncer de mama, el cáncer colorrectal y cáncer de pulmón (Negrin et al., 2015; Woodman et al., 2007). La infección persistente causada por los VPH es un evento importante en la progresión y desarrollo del CaCU, esto también asociado a diversos factores de riesgo que permiten la evolución a partir de lesiones intraepiteliales (LEI) hasta cáncer invasor. Sin embargo, la infección causada por los VPH-AR es un evento determinante para el

desarrollo de lesiones precursoras y posteriormente cáncer (Doorbar et al., 2015, 2012; zur Hausen, 2009).

Los VPH-AR están asociados con el cáncer cervicouterino debido a las propiedades biológicas de sus proteínas E6 y E7 que los diferencian de los VPH-BR, ya que estas tienen un alto potencial oncogénico, además de que trabajan sinérgicamente teniendo un papel central en el proceso de transformación maligna (Moody and Laimonis, 2010; Duensing y Münger, 2002; Duensing et al., 2000;). Uno de los VPH-AR mayormente asociado a la carcinogénesis cervical y al CaCU es el VPH-AR tipo 16, el cual se encuentra presente hasta en el 55% de los casos de cáncer cervicouterino (Hudelist et al., 2004; Muñoz et al., 2006). Debido a esto se han desarrollado diversos estudios donde analizan cómo se comporta en comparación con los otros tipos virales. Existen análisis específicos de sus proteínas E6 y E7 donde se ha descubierto su gran capacidad de transformación y utilidad en el desarrollo del cáncer (Duensing et al., 2001, 2000; Duensing y Münger, 2002; Riley et al., 2003).

Las variantes moleculares o variantes intratípicas de los VPH son definidas como aquellos aislados del mismo genotipo que exhiben una divergencia mínima del 10% y menos de un 2% de variación dentro del gen L1. Estas variantes han sido descritas para la mayoría de los genotipos comunes de VPH y cada genotipo puede presentar entre 20 a 60 variantes (Bernard et al., 2005).

Diversos estudios donde han analizado las variantes del VPH16 enfocados en la secuencia de nucleotídica de los genes E6, L1, L2 y LCR han encontrado que el potencial oncogénico de algunas variantes es diferente (Gheit et al., 2011; Zuna et al., 2009). Varios autores han propuesto la unificación de criterios para agrupar las variantes del VPH16 de acuerdo al análisis molecular de la secuencia del gen E6.

Huertas y colaboradores en el 2011, realizaron un análisis para agrupar a las variantes en clases utilizando el análisis del gen E6, proponiendo la siguiente nomenclatura: en el linaje Europeo para hacer referencia a una clase, se coloca la abreviación del linaje al que pertenece, seguida de una letra y un número que hacen referencia al cambio de nucleótido y la posición específica donde se llevó el cambio del nucleótido. En los linajes As, AA, AFR1, AFR2 y NA, para hacer referencia a

una clase, se coloca la abreviatura del linaje seguida de una letra minúscula, así después va la letra que indica el cambio de nucleótido y el número que indica la posición del cambio. Se debe destacar que cada clase está compuesta por muchas variantes las cuales comparten cambios de nucleótidos específicos (Huertas-Salgado et al., 2011).

Por otra parte, Iris Cornet y colaboradores en el 2012, proponen que la nomenclatura debe ser incluyendo las región LCR del genoma de los VPH16, de esta forma para analizar los linajes y sublinajes de las variantes proponen utilizar la región E6 y la LCR (13 y 32 cambios nucleotídicos, respectivamente) (Cornet et al., 2012). El análisis de Iris Cornet y colaboradores es factible y no modifica las propuestas anteriores como la propuesta de Huertas y colaboradores, además la información que se obtiene de la nomenclatura de Iris Cornet se conocen los cambios que pudieran afectar los procesos del ciclo viral debido a que se analiza una región importante del genoma del VPH16 (LCR), y además se incluye a la región que codifica para la proteína E6 que es de gran importancia en la transformación celular.

La proteína E6 es relativamente una proteína pequeña, con una secuencia nucleotídica de 477 nt (83-459), que codifican para un péptido de 158 a.a. (aminoácidos), con un peso molecular aproximado de 16 a 18 kiloDaltones (kDa) (Yue, 2013). La proteína E6 tiene un amplio contenido de α -hélices (6) y láminas β (7) como estructuras secundarias (Figura 1). Estas propiedades, además de su pequeño tamaño hacen a la proteína E6 inestable e insoluble para purificar (Nominé et al., 2001, Nominé et al., 2001), sin embargo, a pesar de ello se sabe ampliamente que tiene dos dedos de zinc (E6N-E6C), con un interdominio de 36 aminoácidos (aa) cada uno, caracterizado por el motivo CXXC (Nominé et al., 2001a); presenta en su extremo carboxilo motivos de unión PDZ que ayudan en la polaridad y adhesión de la proteína (Zanier et al., 2012); además tiene sitios LXXLL denominados motivos LD que median la interacción proteína-proteína, pero además que están involucrados en la regulación celular, adhesión celular y expresión genética (Tumbarello et al., 2002). Todos estos sitios son de posible interacción con proteínas blanco (Cai et al., 2013). Se debe mencionar que la proteína E6 no tiene actividad enzimática y la mayoría de su actividad es debido a

interacciones proteína-proteína (Howie et al., 2009; Ristriani et al., 2001; Yue, 2013).

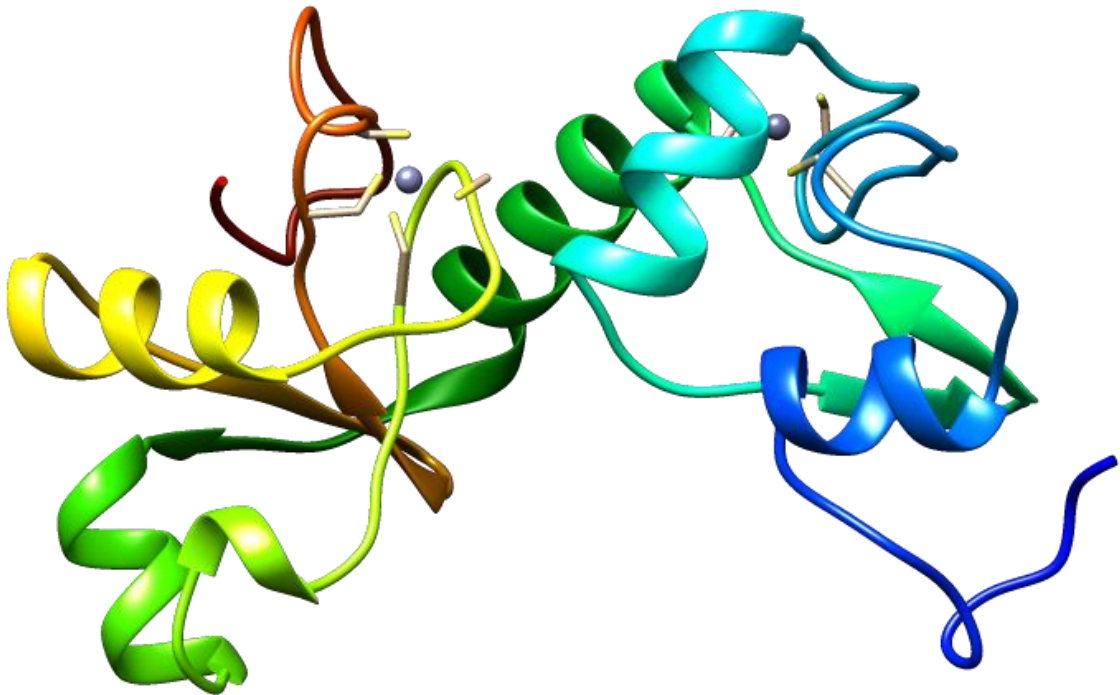


Figura 1. Estructura 3D de la oncoproteína E6 VPH 16 (PDB: 4XR8).

La función clave y específica de la proteína E6 es inhibir la función de p53, una proteína supresora de tumores (Bieging et al., 2014), mediante su degradación por ubiquitinación (Hengstermann et al., 2001; Mantovani y Banks, 2001). Para esto la proteína E6 requiere a la proteína celular asociada a E6 (E6AP). Esta proteína reemplaza a Mdm2 quien en células normales degrada a p53. Esto genera que la vida media de p53 se reduzca de casi tres horas a alrededor de veinte minutos, y el nivel de p53 de esta forma se ve reduciendo en las células infectadas por el virus del VPH en comparación con las células normales.

La proteína E6 también puede retener a p53 en el citoplasma de la célula, bloqueando su translocación al núcleo y así inhibiendo su función independientemente de su degradación (Mantovani y Banks, 2001; Yue, 2013). En consecuencia la proteína E6 inhibe la capacidad de p53 para activar o reprimir la transcripción de sus genes blanco (Figura 2).

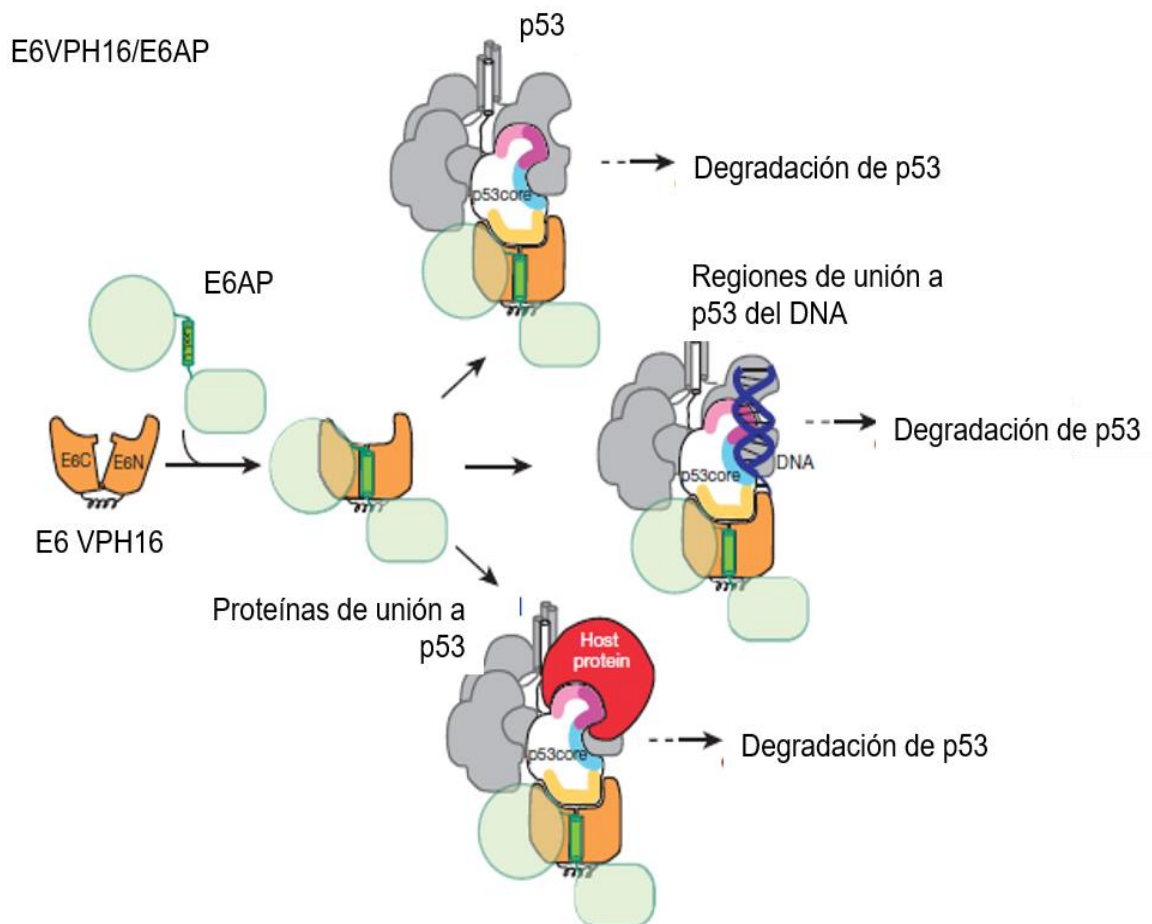


Figura 2. Unión y degradación de E6 a p53. Diferentes formas de degradación de p53 por la unión del complejo E6AP/E6. Aunque p53 se encuentre unido a DNA, a proteínas de unión a p53 o se encuentre libre en la célula el complejo E6AP/E6 se une a p53 y promueve su degradación. Modificada de Martínez-Zapien et al., 2016.

Las proteínas E6 y E7 del VPH16 y de todos los VPH-AR ha sido demostrado que tienen interacción específica con la proteína p53 y la proteína Rb, respectivamente. Las proteínas Rb y p53 regulan negativamente al ciclo celular. Por este motivo las

interacciones de estas proteínas juegan un papel importante en la inducción de la inmortalidad celular (Howie et al., 2009).

La importancia de la proteína p53 ha sido reconocida por su papel en mediar la apoptosis y el mantenimiento de la homeostasis celular, así como también la prevención de la transformación neoplásica (Bieging et al., 2014; Zilfou and Lowe, 2009). Una vez que E6 interactúa con p53 mediante la unión con E6AP la degrada, dejando a la célula queda expuesta a cambios moleculares que provocan la transformación maligna (Howie et al., 2009; Mantovani and Banks, 2001; Rivoire et al., 2006).

E6 y E6AP forman un complejo que es estable y capaz de interactuar con diversas proteínas celulares, E6AP es la proteína que en primera estancia interactúa con E6. Después de esta primera interacción, E6/E6AP aumentan el número de interacciones entre estas dos proteínas y diversas proteínas celulares (Martínez-Zapien et al., 2016).

Existe evidencia de que la proteína E6 presenta interacciones con proteínas como: IRF-3 (Factor 3 Regulador de Interferón) y el co-activador notch 1 MAML1 a través de sus motivos LXXLL. Además de estas dos proteínas existe una gran lista de proteínas que se relacionan con diversos procesos celulares importantes para mantener la homeostasis celular. En la siguiente tabla se muestra una lista de las proteínas con las que interactúa la proteína E6 del VHP 16 (Tabla 1).

Tabla1. Principales proteínas que interactúan con la proteína E6 del VPH16.

Proteína	Nombre	Función	Referencia
TP53	Proteína supresora de tumor p53	Ciclo celular/Expresión de genes/Muerte celular programada	Jabin An et al., 2008 Lu Z et al., 2004 Cooper et al., 2003 Scheffner et al., 1992
TP73	Proteína supresora de tumor p73	Expresión de genes/Muerte celular programada	Park JS et al., 2001
RCN2	Reticulocalbina 2	Traducción de señales/Sistema inmune	Chen et al., 1995 Takisawa et al., 2006
SCRIB	Proteína homóloga scribble	Traducción de señales	Thomas et al., 2005 Nakagawa et al., 2000
SIPA1L1	Proteína asociada a proliferación 1 - proteína 1	Sistema neuronal	Lu Z et al., 2004
SIPA1L2	Proteína asociada a proliferación 1 - proteína 2	Función no conocida	Rozenblatt-Rosen et al., 2012
SERTAD1	SERTA proteína 1	Función no conocida	Gupta et al., 2003
TNPO1	Transportina 1	Mantenimiento y biogénesis de orgánulos/Expresión de genes	Roux and Monoianu., 2003
BAK1	proteína antagonista de Bcl-2	Muerte celular programada	Thomas and Banks., 1999
GPS2	proteína supresora de la vía G 2	Organización de la cromatina	Gegenhardt and Silverstein., 2001
BRCA1	proteína susceptible a cáncer de mama 1	Ciclo celular/Metabolismo de proteínas/Reparación de DNA/Expresión de genes	Zhang et al., 2005
CREBBP	proteína de unión a CREB	Homeostasis/Desarrollo/Metabolismo/Traducción de señales/Sistema inmune/Respuesta celular a estrés/Ciclo circadiano/Organización de cromatina/Expresión de genes/Mantenimiento y biogénesis de orgánulos	Patel et al., 1999
CYLD	Ubiquitina carboxilo-terminal hidrolasa CYLD	Traducción de señales/Sistema inmune/Metabolismo de proteínas	An et al., 2008
DLG1	proteína asociada a sinapsis 97	Sistema neuronal/Desarrollo	Thomas et al., 2005 Pim et al., 2002 Thomas et al., 2001 Kiyono et al., 1997
EP300	Acetiltransferasa de histonas p300	Homeostasis/Desarrollo/Metabolismo/Traducción de señales/Ciclo celular/Sistema inmune/Respuesta celular a estrés/Metabolismo de proteínas/Organización de la cromatina/Reparación del DNA/Expresión de genes	Patel et al., 1999
FADD	proteína asociada a muerte FAS	Traducción de señales/Sistema inmune/Muerte celular programada	Filippova et al., 2004
FBLN1	Fibulina 1	Organización de la matriz extracelular	Du et al., 2002 Rozenblatt-Rosen et al., 2012
UBE3A	Ubiquitina ligasa E3A	Sistema inmune	Nakagawa and Huijbregtse., 2000 Vande et al., 1998 Chen et al., 1995
KPNA2	Importina subunidad alfa-1	Sistema inmune/Reparación del DNA	Roux and Monoianu., 2003
KPNB1	Importina subunidad beta-1	Metabolismo/Sistema inmune/Muerte celular programada	Roux and Monoianu., 2003
IRF3	Factor regulador de interferón 3	Sistema inmune	Lu Z et al., 2004 Pim et al., 2002
MAGI1	Guanilato cinasa asociada a la membrana	Comunicación célula - célula	Thomas et al., 2001 Glaunsinger et al., 2000
MCM7	Factor de reparación del DNA MCM7	Ciclo celular/Reparación del DNA	Gupta et al., 2003 Kukimoto et al., 1998
MYC	Proto-oncogen Myc	Traducción de señales/Ciclo celular/Metabolismo de proteínas/Expresión de genes	Veldman et al., 2003
TNFRSF1A	Receptor del factor de necrosis tumoral 1 ^a	Sistema inmune	Filippova et al., 2004
TSC2	Tuberina	Traducción de señales/Sistema inmune/Expresión de genes	Lu Z et al., 2004
TAX1BP3	proteína de unión a Tax1-3	Traducción de señales	Hampson et al., 2004
XRCC1	proteína de reparación de DNA XRCC1	Reparación del DNA	Iftner et al., 2002

La proteína E6 VPH16 interactúa de manera específica con la proteína p53, ésta interacción tiene como resultado la degradación de p53. Mientras que los niveles de la proteína p53 se encuentran regulados de forma común dentro de la célula por la ubiquitina ligasa Mdm2 (Bieging et al., 2014), la proteína E6 VPH16 interactúa con p53 y directamente se ve involucrada en su degradación por el reclutamiento de una ligasa celular diferente, que es la E6AP.

La proteína p53 es ubiquitinada y degradada por E6AP en presencia de la proteína E6 VPH16, promoviendo un mecanismo por el cual las proteínas E6 de los VPH-AR pueden reducir los niveles de p53 y favorecer su replicación en la célula huésped y así generar un estado de carcinogénesis.

Martínez-Zapien y colaboradores en el 2016 dilucidaron el complejo de interacción mediante el cual se lleva a cabo la degradación de la proteína p53 mediante la interacción de la proteína E6 VPH16 y la E6AP. En su investigación demostraron mediante Triple pull-down y cristalografía de difracción de rayos X, que la formación de un trímero en donde se encuentran involucradas las tres proteínas es necesario para la degradación de la proteína p53, y también demostraron que aminoácidos están involucrados en esta interacción (Figura 3 y 4), resaltando la importancia que tiene el péptido de E6AP (12 aminoácidos; ELTLQELLGEER) que se encuentra involucrado en la interacción, en la conservación de la estructura de la proteína E6 VPH16 y su papel en la interacción con la proteína p53 (Martínez-Zapien et al., 2016; Zanier et al., 2013). Hasta el momento es el único reporte que cuenta con la información estructural de la formación del complejo ternario E6 VPH16/E6AP/p53 que está involucrado en la degradación de la proteína p53.

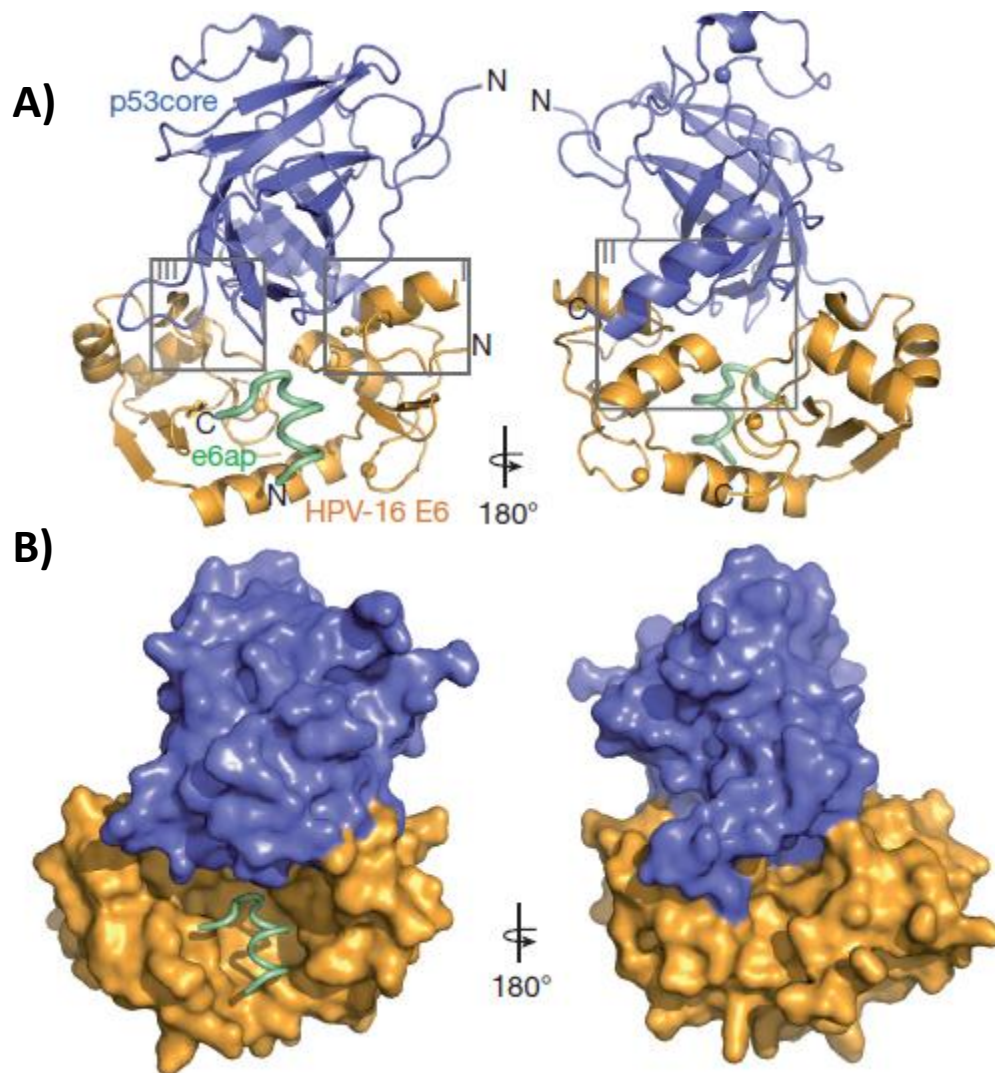


Figura 3. Estructura del complejo ternario E6 VPH16/E6AP/p53. Representación cristalográfica resuelta obtenida mediante cristalografía de rayos X (2.25 Å) del complejo ternario E6 VPH16/E6AP/p53. **A)** Representación en listones del complejo ternario E6 VPH16/E6AP/p53. Verde: péptido de E6AP; Oro: E6 VPH16; Azul: Core de p53. Esferas: Zinc. Cuadros: indican los sitios de interacción en el complejo. **B)** Representación de superficie donde se pueden observar las superficies de interacción en el complejo. Modificado de Martínez-Zapien et al., 2016.

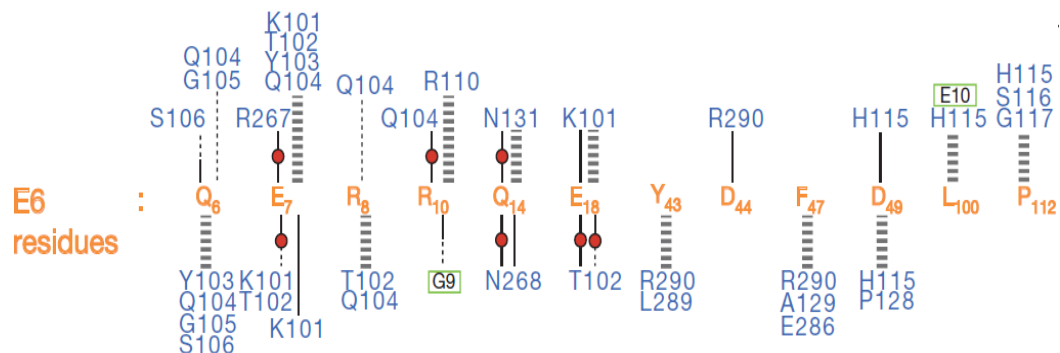


Figura 4. Aminoácidos involucrados en el complejo ternario E6 VPH16/E6AP/p53. Red de interacción del complejo ternario E6 VPH16/E6AP/p53. En la imagen se pueden observar todos los aminoácidos involucrados en la interacción del complejo ternario. Letras oro: aminoácidos E6 VPH16; Letras Azul: aminoácidos de p53; Cuadros verdes: aminoácidos E6AP; Líneas negras gruesas y punteadas: interacciones electroestáticas; Líneas negras: interacciones polares (puentes de hidrogeno); Líneas negras punteadas: interacciones polares (puentes de hidrogeno) involucrados los grupos R de los aminoácidos; Círculos rojos: moléculas de agua. Modificado de Martínez-Zapien et al., 2016.

Un sistema biológico está organizado en diferentes niveles en función del grado de complejidad, no sólo por el número de proteínas (o genes que codifican para ellas), sino también por el número de interacciones combinadas entre ellas a través de splicing, y alteraciones químicas y estructurales de su proteoma (Lander et al., 2001).

La función de una proteína en los diferentes procesos celulares, como expresión de genes, crecimiento celular, proliferación, captación de nutrientes, morfología, movilidad, comunicación intracelular/extracelular y apoptosis, son regulados por diferentes tipos de interacciones entre proteínas (Tuncbag et al., 2008). Debido a la importancia de las interacciones, se han desarrollado diversas técnicas computacionales y experimentales para detectar y predecir los patrones de interacción de proteínas (Planas-Iglesias et al., 2013; Tuncbag et al., 2008).

La proteína E6 está involucrada en la interacción a blancos celulares que interfieren con la transcripción, traducción, remodelación de la cromatina, señalización de citocinas, apoptosis y metástasis. Existe un gran espectro de variación intratípica

de la proteína E6 del VPH16, que se refleja en una alta diversidad en su comportamiento biológico, distribución global y en el riesgo de desarrollar carcinoma cervical.

En nuestro grupo de trabajo, se ha demostrado que las variantes intratípicas de la proteína E6 del VPH16; E-G350, E-A176/G350, E-C188/G350, AAa y AAa son capaces de inducir la expresión diferencial de genes involucrados en el desarrollo del cáncer cervical. Por lo que en este trabajo se propone dilucidar el efecto funcional y estructural de la proteína E6 del VPH16 de referencia con respecto a sus variantes más frecuentes, a través de análisis *in silico* de las secuencias, modelaje molecular (docking y dinámicas moleculares) y la validación experimental a través del análisis de interacciones moleculares.

Capítulo I

Predicción in silico de los cambios estructurales de la oncoproteína E6 VPH 16 y sus variantes.

De: BMC Molecular and Cell Biology Editorial Office <em@editorialmanager.com>

Date: mar., 2 jul. 2019 a las 6:02

Subject: Decision on your Submission to BMC Molecular and Cell Biology - CEBI-D-19-00010R2

To: MARCO ANTONIO LEYVA VAZQUEZ <leyvamarco13@gmail.com>

CEBI-D-19-00010R2

In silico prediction of structural changes in HPV 16 E6 oncoprotein and its variants Hugo Alberto Rodríguez Ruiz, MSc; Olga Lilia Garibay Cerdaneres; Berenice Illades Aguiar; Maricela Sarita Montaña Valdéz; Xiaowei Jiang; MARCO ANTONIO LEYVA VAZQUEZ.

BMC Molecular and Cell Biology

Dear Dr LEYVA VAZQUEZ.

I am pleased to inform you that your manuscript "In silico prediction of structural changes in HPV 16 E6 oncoprotein and its variants" (CEBI-D-19-00010R2) has been accepted for publication in BMC Molecular and Cell Biology.

Before publication, our production team will also check the format of your manuscript to ensure that it conforms to the standards of the journal. They will be in touch shortly to request any necessary changes, or to confirm that none are needed.

Please do not hesitate to contact us if you have any questions regarding your manuscript and I hope that you will consider BMC Molecular and Cell Biology again in the future.

If you wish to co-submit a data note to be published in BMC Research Notes (<https://bmrcresnotes.biomedcentral.com/about/introducing-data-notes>) you can do so by visiting our submission portal <http://www.editorialmanager.com/resn/>. Data notes support open data (<https://www.springernature.com/gp/open-research/open-data>) and help authors to comply with funder policies on data sharing. Please note that this additional service is entirely optional.

Best wishes,

Alexandria Latto
BMC Molecular and Cell Biology
<https://bmcmolcellbiol.biomedcentral.com/>

BMC Molecular and Cell Biology

In silico prediction of structural changes in HPV 16 E6 oncoprotein and its variants

--Manuscript Draft--

Manuscript Number: CEBI-D-19-00010R2

Full Title: In silico prediction of structural changes in HPV 16 E6 oncoprotein and its variants

Article Type: Research article

Section/Category: I don't know, Editor will decide section

Funding Information:

Consejo Nacional de Ciencia y Tecnología (557610) MSc Hugo Alberto Rodríguez Ruiz

Consejo Nacional de Ciencia y Tecnología (288612) Dr Berenice Illades Aguiar

Laboratorio Nacional de Cómputo de Alto Desempeño (2017) Dr Olga Lilia Garibay Cerdaneres

European Research Council (EVOCAN-340560) Dr Xiaowei Jiang

Abstract: Background: HPV16 infection is one of the main risk factors involved in the development of cervical cancer, mainly due to the high oncogenic potential of the viral proteins E6 and E7, which are involved in the different processes of malignant transformation. There is a broad spectrum of intratypical variation of E6, which is reflected in its high diversity, biological behavior, global distribution and risk of causing cervical cancer. Experimental studies have shown that the intratypical variants of the protein E6 from the European variants (E-G350, E-A176/G350, E-C188/G350) and Asian-American variants (AAa and AAc), are capable of inducing the differential expression of genes involved in the development of cervical cancer. Results: An in silico analysis was performed to characterize the molecular effects of these variations using the structure of the HPV16 E6 oncoprotein (PDB: 4XR8; chain H) as a template. In particular, we evaluated the 3D structures of the intratypical variants by structural alignment, ERRAT, Ramachandran plots and prediction of protein disorder, which was further validated by molecular dynamics simulations. Our results, in general, showed no significant changes in the protein 3D structure. However, we observed subtle changes in protein physicochemical features and structural disorder in the N- and C- termini. Conclusions: Our results showed that mutations in the viral oncogene E6 of six high-risk HPV16 variants are effectively neutral and do not cause significant structural changes except slight variations of structural disorder. As structural disorder is involved in rewiring protein-protein interactions, these results suggest a differential pattern of interaction of E6 with the target protein P53 and possibly different patterns of tumor aggressiveness associated with certain types of variants of the E6 oncoprotein.

Corresponding Author:

MARCO ANTONIO LEYVA VAZQUEZ, Ph.D.

Universidad Autónoma de Guerrero

CHILPANCINGO, GUERRERO, Guerrero MEXICO

Corresponding Author E-Mail: leyvamarco13@gmail.com

Corresponding Author Secondary Information:

Corresponding Author's Institution: Universidad Autónoma de Guerrero

Corresponding Author's Secondary Institution:

First Author: Hugo Alberto Rodríguez Ruiz, MSc

First Author Secondary Information:

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAC), en la interacción con sus proteínas blanco.

Order of Authors:

Hugo Alberto Rodríguez Ruiz

Olga Lilia Garibay Cerdenares

Berenice Illades Aguiar

Maricela Sarita Montaña Valdéz

Xiaowei Jiang

MARCO ANTONIO LEYVA VAZQUEZ

***In silico* prediction of structural changes in HPV16 E6 oncoprotein and its variants**

Hugo Alberto Rodríguez-Ruiz¹, Olga Lilia Garibay-Cerdenares^{1,2}, Berenice Illades-Aguilar¹, Sarita Montaña³, Xiaowei Jiang^{4,5,6*} and Marco Antonio Leyva-Vázquez^{1*}

¹ Molecular Biomedicine Laboratory. Faculty of Chemical-Biological Sciences. Autonomous University of Guerrero, Mexico; hugordzrz@uagro.mx (HARR); olgaribay@uagro.mx (OLGC), billades@uagro.mx (BIA)

² CONACyT-Autonomous University of Guerrero, Mexico; olgaribayce@conacyt.mx (OLGC)

³ Faculty of Chemical-Biological Sciences. Autonomous University of Sinaloa, Mexico; mmontano@uas.edu.mx (SM)

⁴ Division of Birth Cohort Study, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

⁵ Guangzhou-Birmingham Joint Research Center for Birth Cohorts and Disease Cohorts, Guangzhou Medical University, Guangzhou, China

⁶ Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

* Correspondence: leyvamarco13@gmail.com; Tel.: +52 (747) 4710901 or jiangx@tcd.ie

ABSTRACT

Background: HPV16 infection is one of the main risk factors involved in the development of cervical cancer, mainly due to the high oncogenic potential of the viral proteins E6 and E7, which are involved in the different processes of malignant transformation. There is a broad spectrum of intratypical variation of E6, which is reflected in its high diversity, biological behavior, global distribution and risk of causing cervical cancer. Experimental studies have shown that the intratypical variants of the protein E6 from the European variants (E-G350, E-A176/G350 and E-C188/G350) and Asian-American variants (AAa and AAc), are capable of inducing the differential expression of genes involved in the development of cervical cancer.

Results: An *in silico* analysis was performed to characterize the molecular effects of these variations using the structure of the HPV16 E6 oncoprotein (PDB: 4XR8; chain H) as a template. In particular, we evaluated the 3D structures of the intratypical variants by structural alignment, ERRAT, Ramachandran plots and prediction of protein disorder, which was further validated by molecular dynamics simulations. Our results, in general, showed no significant changes in the protein 3D structure. However, we observed subtle changes in protein physicochemical features and structural disorder in the N- and C- termini.

Conclusions: Our results showed that mutations in the viral oncogene E6 of six high-risk HPV16 variants are effectively neutral and do not cause significant structural changes except slight variations of structural disorder. As structural disorder is involved in rewiring protein-protein interactions, these results suggest a differential pattern of interaction of E6 with the target protein P53 and possibly different patterns of tumor aggressiveness associated with certain types of variants of the E6 oncoprotein.

Keywords: HPV16; E6 oncoprotein variants; 3D structure; molecular modeling; protein disorder; molecular dynamics simulations.

Background

Persistent infection with high-risk human papilloma virus (HR-HPV) is necessary but not sufficient for the development of cervical cancer (CC). High-risk HPV type 16 (HR-HPV16) is the

causal agent of more than half of the CC in the world [1]. Its high oncogenic potential is mainly due to the E6 and E7 oncoproteins, as they are key regulators of the cell cycle [2]. Like other HR-HPV, HPV16 has well-preserved distinctive intratypic variants by geographical origin [3], and their global distribution and risk of cervical cancer appear to be dependent on the population [4, 5].

In our group, it has been reported that the intratypical variants of HPV16 E6, namely, E-G350, E-A176/G350, E-C188/G350, AAa and AAc, are the most common and have the most oncogenic potential in the development of CC in southern Mexico, in comparison with the HPV16 E6 reference [6]. Moreover, we analyzed the effects of the expression of HPV16 E6 variants (E-G350, E-A176/G350, E-C188/G350, AAa and AAc) and the E6 reference on global gene expression profiles through an *in vitro* model, showing that HPV-16 variants are capable of inducing differential expression of host genes involved in the development of CC, such as genes involved in adhesion, angiogenesis, apoptosis, differentiation, cell cycle, proliferation, transcription and protein translation [7].

In protein evolution, a mutation that changes an amino acid is non-synonymous, while a mutation does not change an amino acid is synonymous. Non-synonymous mutations can be detrimental, beneficial or neutral to viral fitness in the host and could often be explained by subtle changes at the protein structural level [8, 9]. In this study, we adopted an *in silico* approach to evaluate the E6 structural changes. We generated the 3D structures of the five intratypical E6 variants using the crystallized structure of the mutated HPV16 E6 as a template (PDB: 4XR8, chain H from crystal structure of the HPV16 E6/E6AP/p53 ternary complex at 2.25 Å resolution) [10]. We further predicted the structural disorder of the six variants using IUPRED2A [11] and performed molecular dynamics simulations. Our results show that mutations observed in different E6 variants do not significantly alter their 3D structures. However, these non-synonymous mutations slightly modify the structural disorder tendency in the amino- and carboxyl-terminus of HPV16 E6, with the amino-terminus being most affected, which is further supported by the molecular dynamics simulation analysis. These changes may lead to differential binding to host P53 proteins and potentially other proteins, which likely affect the oncogenic potential of the different HPV16 strains investigated here.

Results

Multiple alignment of HPV16 E6 sequences and its variants

The protein sequences of each HPV16 E6 variant, reference E6, E-G350, E-A176/G350, E-C188/G350, AAa and AAc, were obtained through a literature search [12]. To compare each variant with the E6 reference, these sequences were subsequently aligned by Clustal W [13].

Figure 1 shows the alignment of protein sequences of all intratypical variants. In Figure 1A, the primary structures of the E6 reference and its variants are shown. Six non-synonymous sites of amino acid changes are observed among the variants: L83V (E-G350), D25N/L83V (E-A176/G350), E29Q/L83V (E-C188/G350), Q14H/H78Y/L83V (AAa) and Q14H/I27R/H78Y/L83V (AAc). Although there are amino acid changes in every variant in comparison with the reference, all mutated amino acids remain hydrophilic (green), except the change in I27R (red) located in E6 AAc. This mutation changed isoleucine (I), one of the largest aliphatic amino acids whose structure is often relegated to the hydrophobic core of a protein fold, to a basic amino acid, arginine (R), whose lateral chain is able to have a permanent positive charge in living systems and thus is capable of generating exquisite molecular interactions. In Figure 1B, the secondary structures of variants were also aligned, as we can see that their secondary structures are maintained with no changes.

3D structures

To understand the structural role of the six non-synonymous mutations, we used the template 3D experimental structure PDB: 4XR8, which contains the 3D structure of a mutated E6 (chain H). To obtain the E6 reference 3D structure, the SCRWL4 program [14] was used to revert all amino acid changes to the E6 reference in the PDB structure. Subsequently, the 3D structures of all variants were obtained in a similar way. In Figure 2, the structures from the E6 variants and the reference proteins are shown from two different angles: reference (blue), E-G350 (yellow), E-A176/G350 (green), E-C188/G350 (cyan), AAa (pink) and AAc (orange).

In Figure 3, the 3D structural alignment of each variant to the reference is shown. We highlighted mutated amino acids among the models (licorice). The structural overlap shows different orientations in the axes (X: red, Y: green; and Z: blue). Two amino acid mutations, H78Y (AAa and AAc) and L83V (all variants), are shown in Figure 3A. Additionally, it is possible to identify changes in the amino

acids Q14H (AAa and AAc), D25N (E-A176/G350), I27R (AAc) and E29Q (E-C188/G350) in Figure 3B. Finally, in Figure 3C, we structurally aligned all six protein structures from the E6 reference and all five variants (E-G350, E-A176/G350, E-C188/G350, AAa and AAc). Interestingly, all models have the same general structural conformation, and the non-synonymous mutations of the variants do not have a visible effect on the 3D structure of E6, which clearly demonstrates that the 3D structure of the protein seems to be preserved despite variations in amino acids.

In Figure 4, from the above aligned protein structures (Figure 3), we isolated the aligned amino acid changes in every variant in comparison with the E6 reference. The subtle modifications in the orientation of the side chains are due to a particular amino acid change whose functional groups belong to the same biochemical group; however, a drastic change occurs in the lateral chain in the change of I27R in E6 AAc due amino acids of different biochemical behavior.

3D structure evaluation

To evaluate and validate the accuracy of the predicted 3D structures of the HPV16 E6 reference and its variants, Rampage and ERRAT servers were used. Ramachandran plot analysis of the HPV16 reference and its variants are also shown. In Figure 5, the disallowed regions are colored pale-yellow. Red color indicates low-energy regions. Brown color indicates allowed regions, and yellow indicates the so-called generously allowed regions. This analysis showed 92.6% amino acids in the highly favored region, 7.4% in the favored region and no atypical amino acids in the unfavored region, which suggest that the generated models are reliable. Evaluation of the quality of the models generated using the ERRAT server showed an overall quality factor that was different for each one of the variants (Figure 6), with a value greater than 50 considered favorable to the E6 reference. E-C188/G350 presented a quality factor of 90.909, while for E6 E-G350, the quality factor was 90.210; however, E6 E-A176/G350 presented the lowest value of quality factor of all variants at 89.510. Finally, for the variants AAa and AAc, the quality factors were 90.210 and 92.308, respectively. Despite the fact that all the variants obtained a favorable value, the variant AAc obtained the highest overall quality factor.

Physical and chemical properties

To analyze the possible modifications in the physicochemical properties of the HPV16 E6 reference and its variants, we submitted the primary sequences to the ExPASy's ProtParam Tool server. This server allows the theoretical prediction of several physical and chemical parameters from a given primary protein sequence, such as molecular weight (MW), theoretical isoelectric point (pI), amino acid composition, atomic composition, extinction coefficient (EC), estimated half-life, instability index (II), aliphatic index (AI) and grand average of hydropathicity (GRAVY).

As a result, the theoretical value of the isoelectric point (pI) calculated for the reference HPV16 E6 reference as well as the different variants were in the range of 9.01 to 9.10, indicating that the E6 reference and variants have a basic isoelectric point. The EC indicates how much light a protein absorbs at a certain wavelength but is also a value used to determine the concentrations of proteins in solution to perform processes of purification. This property was different between the variants, introducing an EC value of $21275 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm for the E6 reference and the variants E-G350, E-A176/G350, E-C188/G350, and an EC value of $22765 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm for the variants AAa and AAc. The II provides an estimate of the stability of a protein in a test tube; a protein whose II is smaller than 40 is predicted to be stable, while a value above 40 predicts a highly unstable protein. All variants analyzed, including the reference, presented a high instability index: E6 reference: 73.25; E-G350: 72.75; E-A176/G350: 77.23; and E-C188/G350: 71.50. However, the AAa and AAc variants showed the lowest rates of instability at 67.61 and 66.34, respectively. The AI, which is defined as the relative volume occupied by the aliphatic side chains (alanine, valine, leucine, and isoleucine), is considered a positive factor for the increase in the heat stability of globular proteins. The AI value of the E6 reference was the largest at 70.99, followed by the AI of the variants E-G350, E-A176/G350, E-C188/G350 and AAa, which was 70.33. Finally, the AI of the AAc variant was the lowest, with a value of 67.75. The GRAVY value for a peptide or protein is the sum of the hydropathic values of all the amino acids divided by the number of residues in the sequence, indicating the feature hydrophobic (positive values) or hydrophilic (negative values) of a protein while taking into account the length of the sequence of amino acids. This value for the proteins analyzed was -0.734 for the reference, -0.732 for the variants E-G350, E-A176/G350, E-C188/G350, -0.717 for AAa variant and -0.777 for variant AAc, with the latter variant being the most hydrophilic.

Structural disorder prediction of E6 reference and variants

Proteins are known to have different levels of structural disorder either for the whole protein or within their protein domains. There are specific amino acids contributing to structural disorder at different levels, which can lead to differential interactions with target proteins [15]. Here, we investigated the structural disorder in the E6 reference and its variants using IUPRED2A [11]. Figure 7A shows that the structural disorder for each variant does not deviate significantly from each other. However, there are several noticeable differences due to non-synonymous mutations. Around site 14, the E6 reference shares similar disorder scores with variants E-G350, E-A176/G350 and E-C188/G350, which have higher disorder tendencies than variants AAa and AAc. AAc has a higher disorder tendency than variant AAa. Around site 29, variant AAc has the highest disorder tendency of any other variant with the same disorder tendency. From sites 78 to 90, the reference has similar disorder tendencies as variants E-G350, E-A176/G350 and E-C188/G350, which have higher disorder scores than AAa and AAc variants. Even when the prediction analysis of IUPRED2A showed no significant variations related to protein disorder, it is noticeable that all variations are located in the amino acid modified in each variant, and most disorder prediction sites could explain subtle structural changes that could affect the interaction of E6 with host proteins (Figure 7B).

Molecular dynamics simulation analysis

As the SCRWL program only optimizes sidechain conformations when building homology models, to further understand the effect of the six non-synonymous mutations of E6 oncoprotein on side chain and backbone conformations, a MD simulation of 10 ns of the reference E6 oncoprotein and variants E-G350, E-A176/G350, E-C188/G350, AAa and AAc was performed to analyze the structural impact of the intratypical variations, and structure overlapping are shown (Figure 8). The frames of 0, 5 and 10 ns were obtained to visualize the perturbations on backbones (Figure 8A-C) and side chains (Figure 8D-E) during the simulations, where most variations are located in the side chains of the non-synonymous mutations (Figure 8D-E). The figure 8A shows certain instability in amino and carboxyl regions after a structural overlap of all variants. Particularly, there is a notable increase of flexibility (disorder) in amino- and carboxyl- regions of all variants through the simulation consistent with the disorder prediction, being more visible in the Asian American variants (Figure 8C) when compared with the European variants (EUR; E-G350, E-A176/G350, E-C188/G350) (Figure 8B). Moreover, the most representative side chain variations are shown in Figure 8D for EUR variants (D25N, E29Q and L83V), and in Figure 8E for AA variants (Q14H, I27R and H78Y).

Discussion

High-risk mucosal HPV infections are responsible for the majority of cervical, anal, rectal, and penile cancers, as well as an increasingly high proportion of oropharyngeal cancers. The two main viral HPV oncogenes required to establish and maintain the tumorigenic phenotype encode two early expressed oncoproteins, called E6 and E7 [16]. It is known that the HPV16 E6 reference is involved in several biological processes of malignant transformation, such as cell cycle, apoptosis, DNA repair, immune response, organization of chromatin, and cell communication [17, 18].

Our initial structural analysis based on homology modeling did not find any significant change of the protein structures of the variant E6 proteins. We subsequently used ExPASy's ProtParam Tool server to predict the physicochemical properties of these proteins. It was possible to predict some physical and chemical parameters from the E6 reference and their variants, which showed minimal differences in the isoelectric point, but when the high index of instability was analyzed, AA variants (a and c) showed the highest values of instability. The instability index could be used as an indirect measure of the *in vivo* half-life of a protein [19]. Proteins that have an *in vivo* half-life of less than 5 h have been shown to have an instability index of more than 40, whereas those that have an *in vivo* half-life of more than 16 h have an instability index of less than 40 [20]. The values obtained for the proteins in this work are in the range of 66-69, and these theoretical predictions suggest a long half-life at an average of 5 h for variants of the E6 oncoprotein. Finally, the hydrophobicity was determined using the GRAVY index (the sum of hydropathic values of all amino acids divided by the protein length), whose values could be related to the solubility and aggregation of a protein, which can greatly affect the recovery efficiency in the recovery process [21, 22]. In all E6 analyzed, the negative values are associated with their hydrophilic behavior and could be related to the cytoplasmic localization or localization in the lumen of certain organelles.

In the UniProt database there are 9 PDB structures corresponding to HPV16 E6 (1VZN, 2FK4, 2KPL, 2LJX, 2LJY, 2LJZ, 4GIZ, 4JOP and 4XR8), where we used 4XR8 [10]. Unlike the rest of the crystallized structures, which correspond to the amino or carboxyl ends, or to binding domains, the 4XR8 structure corresponds to the structure of a quaternary complex HPV16 E6 / E6AP / p53, in which the domains associated with the interaction with E6AP and P53 are shown. This is the only complex allows us to get the complete structure of E6 with their respective interactions with P53 and perform homology modeling, which further supported our analysis of the molecular dynamics

simulations. Our study reveals that these non-synonymous mutations on E6 from these highly oncogenic variants cause subtle structural and potential functional changes, which may play an important role in mediating how they interact with P53. These differential interactions may subsequently determine the level of susceptibility of developing cervical cancer or other related cancers. Particularly, it can be clearly seen that Q14H mutation is right at the binding interface of E6-P53core, which is located at the sub-interfaces I of the binding interface (Supplementary Figure S1) [10]. Previous functional studies demonstrate this site (Q14A) plays an important role in P53 binding and subsequent degradation [10]. Other mutations located around this and other active sites may allosterically contribute to P53 binding and degradation [10, 23]. Further docking analysis shows that although the site of interaction with E6AP-E6 does not change, the interaction pattern between E6-E6AP and P53 does change, potentially due to the subtle structural effect of these pathogenic mutations at both the N and C regions of the viral E6 protein (unpublished data). However, we caution that due to limited number of E6 structures currently available, our homology model may not be entirely accurate. This can be improved in the future when more full-length E6 structures are available.

There are proteins that have 3D structural conformation highly heterogeneous and are known to have different levels of structural disorder, from slightly flexible to intrinsically disordered [24]. Their structural conformations depend on different conditions triggered by pH, temperature, redox potential, mechanical force, light exposure and various types of interactions [25]. The biological function of these proteins is directly linked to their foldability during interaction with their targets. HPV-16 E6 is a protein with regions of different structural disorder levels, particularly the N and C termini as evidenced by the molecular dynamics simulations. In this *in silico* analysis, the amino acid changes of variants are predominantly located in the N-terminal region, we suggest that a subtle change in structural disorder tendency could be enough to confer variations in HPV16's oncogenic potential, which should aid future functional studies of these types of mutations.

The present study provides a framework to evaluate the mechanism by which HPV16 E6 structural disorder could be related to differential molecular interactions to key host tumour suppressive and/or oncogenic pathways and, most importantly, to establish a new point of view about the design of new therapeutic strategies based on oncoproteins as molecular targets of viral oncoprotein plasticity. However, we caution that due to limited number of E6 structures currently

available, our homology modelling may not be entirely accurate. This can be improved in the future when more full-length E6 structures are available.

Conclusions

In silico analysis shows that the primary sequences of all HPV E6 variants analyzed have variations in the first 83 amino acids; however, their general 3D structure does not change. The molecular modeling and molecular dynamics simulations of these protein variants showed minimal changes in general structure, but broad changes in their physicochemical parameters, which are possibly involved in the differential pattern of interactions with protein targets.

Methods

Obtaining the target sequence and its variants

The primary sequence of amino acids in the HPV16 E6 reference was obtained from the UniProt database (<http://www.uniprot.org/>) with accession number P03126. The E6 oncoprotein consists of 477 nucleotides that encode 151 amino acids (aa), presents two zinc fingers and has a PDZ domain in its carboxyl-terminal [26]. The changes in amino acids in each variant analyzed are shown in table 1. The amino acid sequences of the E6 variants E-G350, E-A176/G350, E-C188/G350, AAa and AAc were reported by Huertas-Salgado et al. [12]. To submit the sequences to alignment, an amino acid substitution from the reference sequence of HPV16 E6 was performed.

Generation of 3D structures.

The complete HPV16 E6 oncoprotein three-dimensional (3D) structure is found in the Protein Data Bank (PDB) with the access code 4XR8, chain H (<http://www.rcsb.org/pdb/>) [10]. To obtain the secondary structure, the primary sequence (P03126) and its respective variants were submitted to PDBsum server [27]. The sequence of the E6 oncoprotein in this PDB is called E6 4C/4S and has four point mutations (Ser80Cys, Ser97Cys, Ser111Cys, and Ser140Cys) [10]. To analyze the mutations of the E6 oncoprotein variants, E-G350, E-A176/G350, E-C188/G350, AAa, and AAc, and the E6 reference, the SCWRL4 program was used [14].

Validation of modeled structures

The accuracy of the predicted models was evaluated by Ramachandran plot using the RAMPAGE server (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) [28] and ERRAT server (<http://services.mbi.ucla.edu/ERRAT/>) [29] to check the quality of these models.

Calculation of Physiochemical Properties

To calculate the physical and chemical parameters of proteins, such as molecular weight, theoretical isoelectric point, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity, we used the ExPASy's ProtParam Tool server (<https://web.expasy.org/protparam/>) [30].

Visualization

All 3D structures were visualized in the VMD 1.9.1 program (Visual Molecular Dynamics), which is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3D graphics and built-in scripting (<http://www.ks.uiuc.edu/Research/vmd/>) [31].

Structural disorder prediction

The IUPRED2A is based on the calculation of the pairwise amino acid interaction energies in a given length. So, a single amino acid change (e.g., Q14H mutation) could potentially change the disorder tendency in that region [32]. This server takes amino acid sequence in FASTA format as input. The results are returned in either text or graphical format, specifying the disorder tendency of each residue along the sequence. This score can take a value between 0 and 1. Residues with a predicted score above 0.4 are considered disordered [33].

Molecular dynamics simulation

We used the NAMD 2.8 program [34] to perform the molecular dynamics (MD) simulations of the reference E6 oncoprotein and variants E-G350, E-A176/G350, E-C188/G350, AAa and AAc. We collaborated with the Laboratory of Molecular Modeling and Bioinformatics of the Facultad de Ciencias Químico Biológicas de la Universidad Autónoma de Sinaloa, using the Hybrid Cluster Xihcoatl (<http://clusterhibrido.cinvestav.mx>) LANCAD and GPU-CUDA with video cards graphics NVIDIA Tesla C2070/Tesla C2075. The force fields CHARMM22 and CHARMM27 [35] were used for topologies and lipids. The TIP3 model was used for water molecules. The system was solvated using the psfgen plugin in the VMD program [31]. To add water molecules and ions to neutralize the system, we added 9,725 water molecules and 15 Cl⁻ for the E6 reference. For variant E-G350, we added 7924 water molecules and 15 Cl⁻. For variant E-A176/G350, we added 7924 water molecules and 16 Cl⁻. For variant E-C188/G350, we added 7926 water molecules and 15 Cl⁻. Finally, for variant AAa we added 7924 water molecules and 15 Cl⁻. For variant AAc, we added 7860 water molecules and 12 Cl⁻. Moreover, for all structures two Zn⁺ were added. The system was submitted to minimization energy for 1000 steps followed by equilibration for 1 ns under constant temperature and pressure (NPT) with protein and lipid atoms restrained. Molecular dynamics simulations were run for 10 ns using the NTV ensemble, considering E6 and its variants as soluble proteins. The frames from 0, 5 and 10ns were obtained with the Carma software [36]. The 3D structures visualization and structural alignment were performed by VMD software [31].

Abbreviations

HR-HPV	High Risk Human Papilloma Virus
CC	Cervical Cancer
HPV16	Human Papilloma Virus Type 16
PDB	Protein Data Bank
3D	Three Dimensional

Declarations

Ethics approval and consent to participate. Not applicable.

Consent for publication. Not applicable.

Availability of data and material: Not applicable.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was partially supported by grants from Basic Science 2016 (288612) from CONACyT and Laboratorio Nacional de Cómputo de Alto Desempeño (LANCAD 2017, Clúster de Supercómputo Xihucóatl). XJ was partially supported by an advanced ERC grant (EVOCAN-340560) awarded to Ian PM Tomlinson.

Acknowledgments: Rodríguez-Ruíz HA was a recipient of a CONACyT fellowship (No. 557610).

Author's contributions: This work was carried out through the collaboration of all the authors. MALV, MSM, OLGC, HARR and XJ were responsible of the research design. MSM and HARR were involved in the generation of 3D structures and molecular dynamics. XJ carried out the protein disorder and structural analysis. HARR, MSM and OLGC carried out the *in silico* analysis, and together with MALV, BIA and XJ participated in the results interpretation and drafting of the manuscript. XJ and OLGC wrote the final version of the manuscript. All authors had read and approved the final manuscript.

Ethics approval and consent to participate. Not applicable

Consent for publication. Not applicable

Availability of data and materials. Not applicable

References

1. Zur Hausen H: **Papillomaviruses in the causation of human cancers - a brief historical account.** *Virology* 2009, **384**(2):260-265.
2. Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I: **Human papillomavirus molecular biology and disease association.** *Rev Med Virol* 2015, **25**(S1):2-23.
3. Calleja-Macias IE, Kalantari M, Huh J, Ortiz-Lopez R, Rojas-Martinez A, Gonzalez-Guerrero JF, Williamson A-L, Hagmar B, Wiley DJ, Villarreal L *et al*: **Genomic diversity of human papillomavirus-16, 18, 31, and 35 isolates in a Mexican population and relationship to European, African, and Native American variants.** *Virology* 2004, **319**(2):315-323.
4. Cornet I, Gheit T, Iannacone MR, Vignat J, Sylla BS, Del Mistro A, Franceschi S, Tommasino M, Clifford GM: **HPV16 genetic variation and the development of cervical cancer worldwide.** *Br J Cancer* 2013, **108**(1):240-244.
5. Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Greggi S, Buonaguro L, Buonaguro FM: **Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix.** *Gynecol Oncol* 2011, **121**(1):32-42.
6. Ortiz-Ortiz J, Alarcón-Romero LdC, Jiménez-López MA, Garzón-Barrientos VH, Calleja-Macías I, Barrera-Saldaña HA, Leyva-Vázquez MA, Illades-Aguiar B: **Association of human papillomavirus 16 E6 variants with cervical carcinoma and precursor lesions in women from Southern Mexico.** *Virol J* 2015, **12**(1):29.
7. Zacapala-Gómez AE, Del Moral-Hernández O, Villegas-Sepúlveda N, Hidalgo-Miranda A, Romero-Córdoba SL, Beltrán-Anaya FO, Leyva-Vázquez MA, Alarcón-Romero LdC, Illades-Aguiar B: **Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells.** *Virology* 2016, **488**:187-195.
8. Olabode AS, Jiang X, Robertson DL, Lovell SC: **Ebolavirus is evolving but not changing: No evidence for functional change in EBOV from 1976 to the 2014 outbreak.** *Virology* 2015, **482**:202-207.
9. Studer Romain A, Dessailly Benoit H, Orengo Christine A: **Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes.** *Biochem J* 2013, **449**(3):581.
10. Martinez-Zapian D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, Forster A, Cousido-Siah A, Masson M, Pol SV, Podjarny A *et al*: **Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53.** *Nature* 2016, **529**(7587):541-545.
11. Meszaros B, Erdos G, Dosztanyi Z: **IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding.** *Nucleic Acids Res* 2018, **46**(W1):W329-W337.
12. Huertas-Salgado A, Martín-Gómez DC, Moreno P, Murillo R, Bravo MM, Villa L, Molano M: **E6 molecular variants of human papillomavirus (HPV) type 16: An updated and unified criterion for clustering and nomenclature.** *Virology* 2011, **410**(1):201-215.
13. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al*: **Clustal W and Clustal X version 2.0.** *Bioinformatics* 2007, **23**(21):2947-2948.

14. Krivov Georgii G, Shapovalov Maxim V, Dunbrack Roland L: **Improved prediction of protein side-chain conformations with SCWRL4.** *Proteins* 2009, **77**(4):778-795.
15. Vladimir NU, Christopher JO, Dunker AK: **Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept.** *Annu Rev Biophys* 2008, **37**(1):215-246.
16. McBride AA: **Oncogenic human papillomaviruses.** *Philos Trans R Soc Lond B Biol Sci* 2017, **372**(1732):20160273.
17. Howie HL, Katzenellenbogen RA, Galloway DA: **Papillomavirus E6 proteins.** *Virology* 2009, **384**(2):324-334.
18. Mittal S, Banks L: **Molecular mechanisms underlying human papillomavirus E6 and E7 oncoprotein-induced cell transformation.** *Mutat Res Rev Mutat Res* 2017, **772**:23-35.
19. Guruprasad K, Reddy BVB, Pandit MW: **Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence.** *Protein Eng Des Sel* 1990, **4**(2):155-161.
20. Rogers S, Wells R, Rechsteiner M: **Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis.** *Science* 1986, **234**(4774):364.
21. Huang H-J, Chen W-Y, Wu J-H: **Total Protein Extraction for Metaproteomics Analysis of Methane Producing Biofilm: The Effects of Detergents.** *Int J Mol Sci* 2014, **15**(6):10169-10184.
22. Kyte J, Doolittle RF: **A simple method for displaying the hydropathic character of a protein.** *J Mol Biol* 1982, **157**(1):105-132.
23. Dalal S, Gao Q, Androphy EJ, Band V: **Mutational analysis of human papillomavirus type 16 E6 demonstrates that p53 degradation is necessary for immortalization of mammary epithelial cells.** *J Virol* 1996, **70**(2):683-688.
24. Babu MM, Barbar E, Blackledge M, Bondos SE, Dosztányi Z, Dyson HJ, Forman-Kay J, Fuxreiter M, Gsponer J, Han K-H *et al*: **What's in a name? Why these proteins are intrinsically disordered AU - Dunker, A. Keith.** *Intrinsically Disord Proteins* 2013, **1**(1):e24157.
25. Zhang Y, Launay H, Schramm A, Lebrun R, Gontero B: **Exploring intrinsically disordered proteins in Chlamydomonas reinhardtii.** *Sci Rep* 2018, **8**(1):6805.
26. Mantovani F, Banks L: **The Human Papillomavirus E6 protein and its contribution to malignant progression.** *Oncogene* 2001, **20**(54):7874-7887.
27. Laskowski RA, Chistyakov VV, Thornton JM: **PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids.** *Nucleic Acids Res* 2004, **33**(Database issue):D266-D268.
28. Lovell Simon C, Davis Ian W, Arendall WB, de Bakker Paul IW, Word JM, Prisant Michael G, Richardson Jane S, Richardson David C: **Structure validation by C α geometry: ϕ , ψ and C β deviation.** *Proteins* 2003, **50**(3):437-450.
29. Colovos C, Yeates Todd O: **Verification of protein structures: Patterns of nonbonded atomic interactions.** *Protein Sci* 1993, **2**(9):1511-1519.

30. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, Bairoch A: **Protein Identification and Analysis Tools on the ExPASy Server**. In: *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; **2005**: 571-607.
31. Humphrey W, Dalke A, Schulten K: **VMD: Visual molecular dynamics**. *J Mol Graphics* 1996, **14**(1):33-38.
32. Dosztányi Z, Csizmok V, Tompa P, Simon I: **IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content**. *Bioinformatics* 2005, **21**(16):3433-3434.
33. Dosztányi Z: **Prediction of protein disorder based on IUPred**. *Protein Sci* 2018, **27**(1):331-340.
34. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K: **Scalable molecular dynamics with NAMD**. *J Comput Chem* 2005, **26**(16):1781-1802.
35. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S *et al*: **All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins**. *J Phys Chem B* 1998, **102**(18):3586-3616.
36. Koukos PI, Glykos NM: **Grcarma: A fully automated task-oriented interface for the analysis of molecular dynamics trajectories**. *J Comput Chem* 2013, **34**(26):2310-2312.

Tables

Table 1. Amino acids differentials between the sequence of the E6 reference and variants of HPV16.

E6	14	25	27	29	78	83
Reference	Q	D	I	E	H	L
151 aa						
E-G350						V
E-A176/G350		N				V
E-C188/G350				Q		V
AaA	H				Y	V
AAc	H		R		Y	V

Figures

Figure 1. Alignment of the protein sequences and their secondary structures from E6 oncoprotein and variants. (A) Shows the alignment of amino acid sequences of E6 oncoprotein reference and its intratypical variants. Amino acidic changes of every variant are highlighted in colors. (B) Shows the structural alignment of their secondary structures.

Figure 2. Structures of the reference E6 oncoprotein and five variants. (A) Structure of the E6 reference oncoprotein: blue and its variants: (B) E6-G350: yellow, (C) E6-A176/G350: green, (D) E6-C188/G350: cyan, (E) E6 AAa: pink and (F) E6 AAc: orange. The silver spheres indicate zinc molecules. The orientation of the proteins is indicated by the axes, X: red, Y: green; Z: blue.

Figure 3. 3D Structural alignment of the reference E6 oncoprotein and its variants with amino acid changes highlighted (licorice). (A) Visualization of amino acid changes: H78C and L83V; (B) Visualization of amino acid changes: Q14H, D25N, I27R and E29Q; (C) Visualization of global overlapping structures. Reference E6 oncoprotein (blue) and its variants: E6-G350 (yellow); E6-A176/G350 (green); E6-C188/G350 (cyan); E6 AAa (pink) and E6 AAc (orange). The silver spheres indicate zinc molecules. The orientation of the proteins is defined by the axes (X: red, Y: green; Z: blue).

Figure 4. Amino acids from the non-synonymous sites among the E6 reference and its variants. (A) Two different amino acids at position 14 (Q14) in the sequence alignment between the reference E6 oncoprotein and AAa and AAc variants; (B) Two different amino acids at position 25 (D25) between the reference E6 oncoprotein and alternative E6-A-176/G350 are shown; (C) Amino acid at position 27 (I27) of the E6 oncoprotein reference differs in the variant AAc by R27; (D) The amino acid at position 29 (E29) differs in the variant E6-C188/G350; (E) shows the amino acid at position 78 (H78) that differs in the AAa and AAc variants; (F) shows the amino acid at position 83 (L83) which is mutated in all variants. The same colors are used across panels for different amino acids from the same variant. Different colors are used as follow: E6 reference oncoprotein (blue), E6-G350 (yellow), E6-A176/G350 (green), E6-C188/G350 (cyan), E6 AAa (pink) and E6 AAc (orange). The silver spheres indicate zinc molecules. The orientation of the amino acids is defined by the axes shown (X: red, Y: green; Z: blue).

Figure 5. Ramachandran plot analysis. The analysis showed 92.6% of amino acids in all variations were located in the region highly favored, 7.4% in the favored region and no atypical amino acid in the not favored region. (A) E6 Reference; (B) E6-G350; (C) E6-A176/G350; (D) E6-C188/G350; (E) E6 AAa and (F) E6 AAc.

Figure 6. ERRAT server analysis. Graphics show the overall quality factor of the reference E6 oncoprotein and its variants. (A) E6 oncoprotein reference; (B) E6-G350; (C) E6-A176/G350; (D) E6-C188/G350; (E) E6 AAa and (F) E6 AAc. *On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value. Overall quality factor, is expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit.

Figure 7. Protein disorder analysis. (A) Graphic shows the disorder tendency for each amino acid predicted by UPRED2A of the E6 reference and its variants E6-G350; E6-A176/G350; E6-C188/G350; E6 AAa and E6 AAc. (B) The primary sequences of E6 reference and its variants are showed, the white boxes highlighting the amino acid changes on each variant in comparison with E6 reference.

Figure 8. Molecular dynamics simulations of E6 oncoproteins. (A) Shows a structural overlap of all variants at 0, 5 and 10 ns. (B, C) The structural overlap of reference E6 oncoprotein with EUR variants and AA variants are shown. (D) Visualization of amino acids conformational changes: D25N, E29Q and L83V. (E) Visualization of amino acids conformational changes: Q14H, I27R and H78Y. Same colors of Figure 2 are used for each variant.

Supplementary Figure

Figure S1. Illustration of the non-synonymous mutations on E6 structure (colored grey) in complex with the p53core (colored wheat). Mutations Q14H, D25N, I27R and E29Q are colored red for the N-terminus mutations, while mutations H78Y and L83V are colored green for C-terminus mutations. (A) The side view of the six non-synonymous mutations. (B) Top view of the six non-synonymous mutations. PDB 4XR8 and PyMOL 2.3.0 were used to map these mutations.

Figure 1

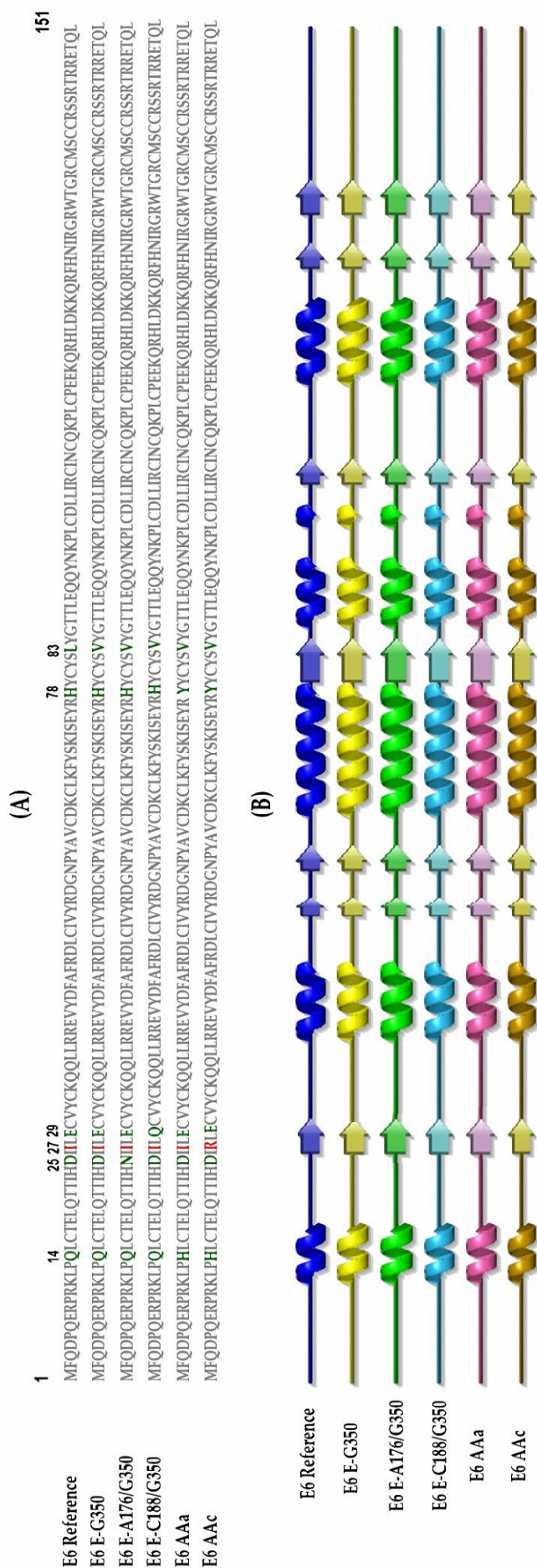


Figure 2

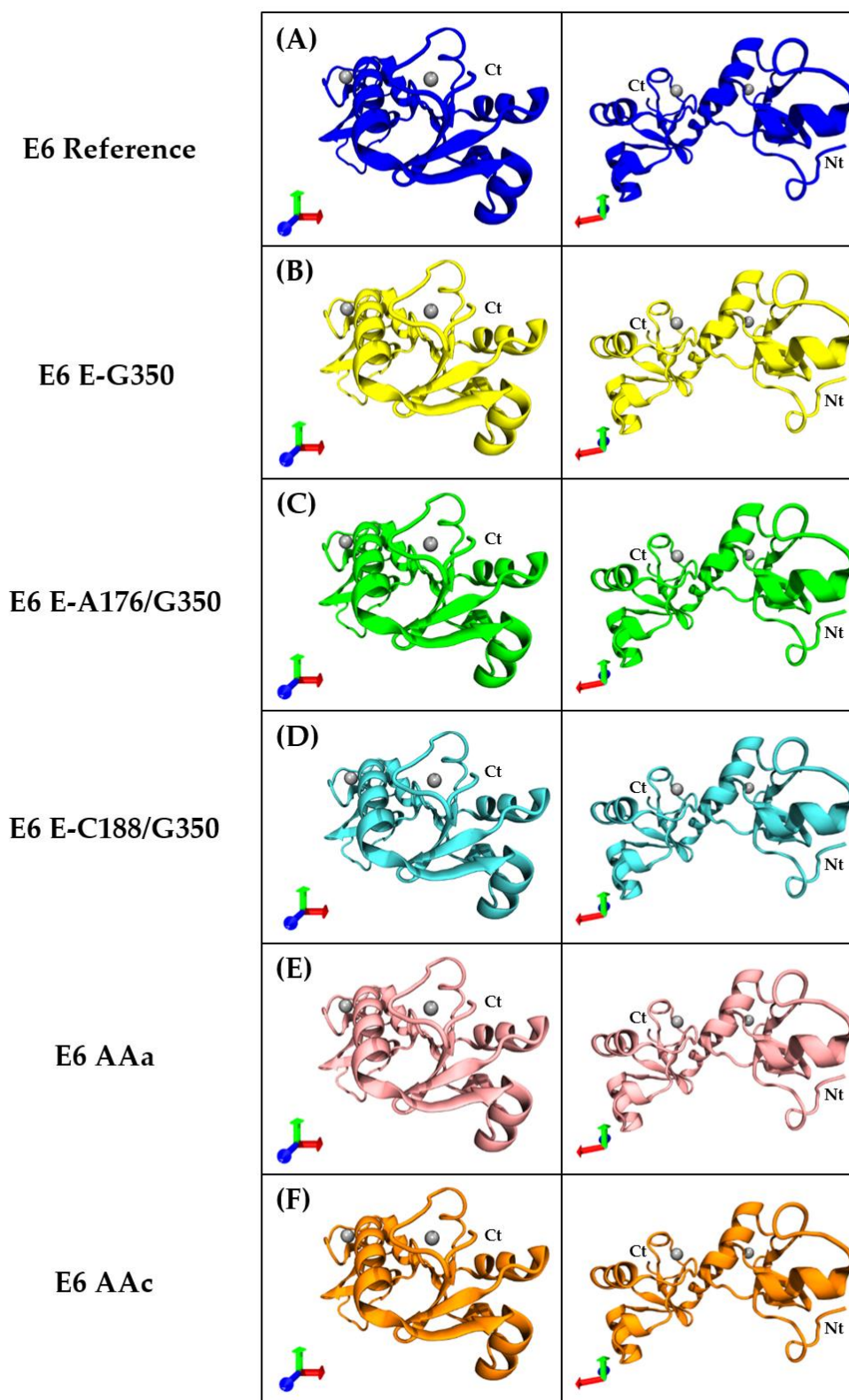


Figure 3

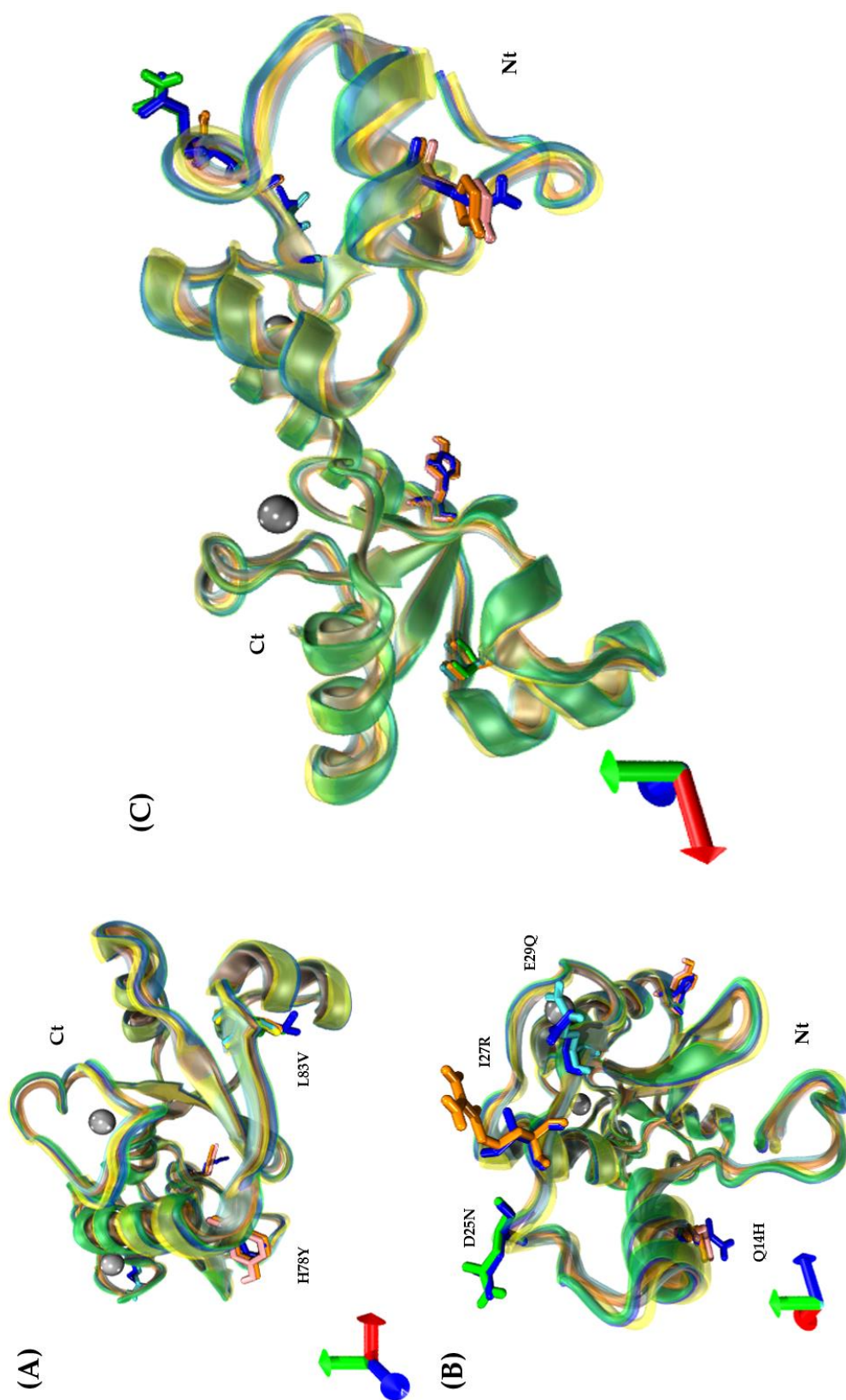


Figure 4

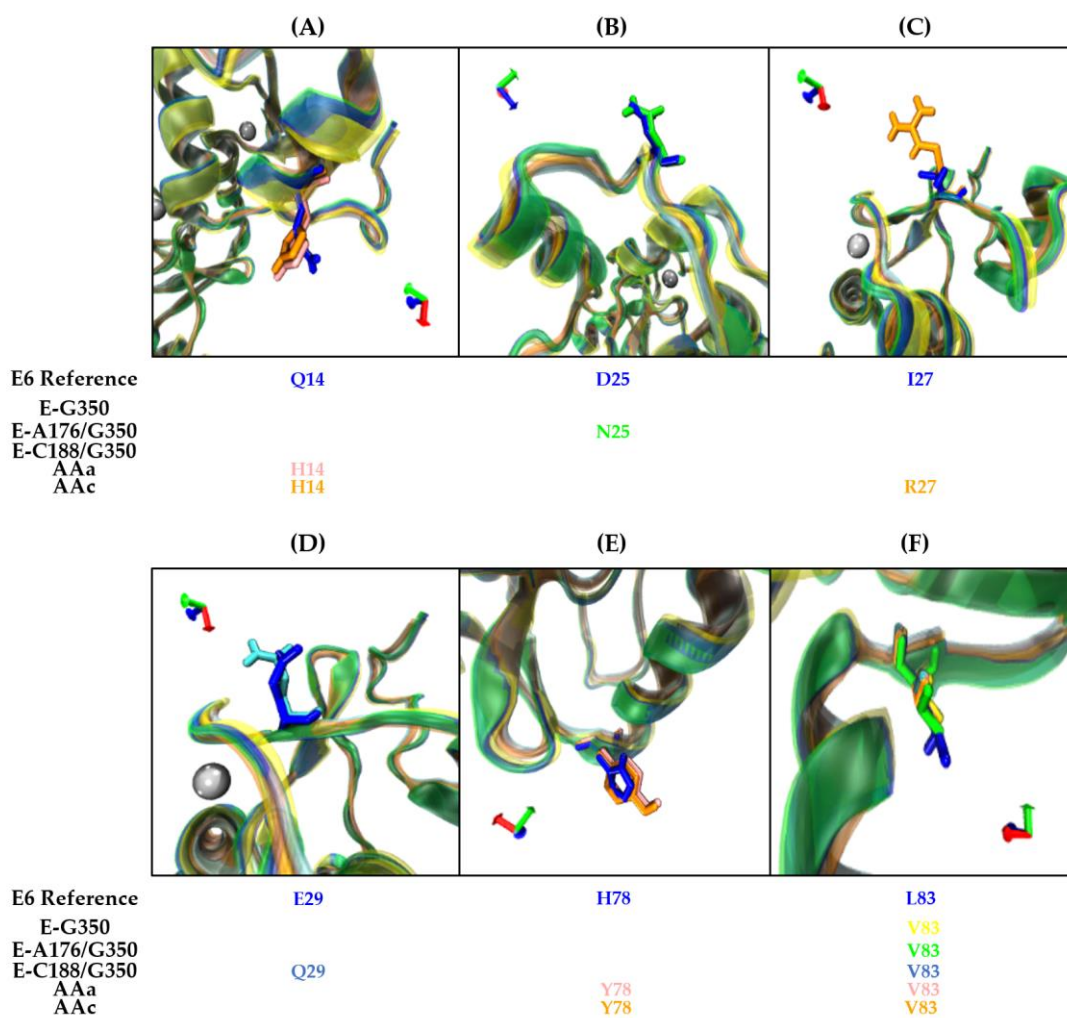


Figure 5

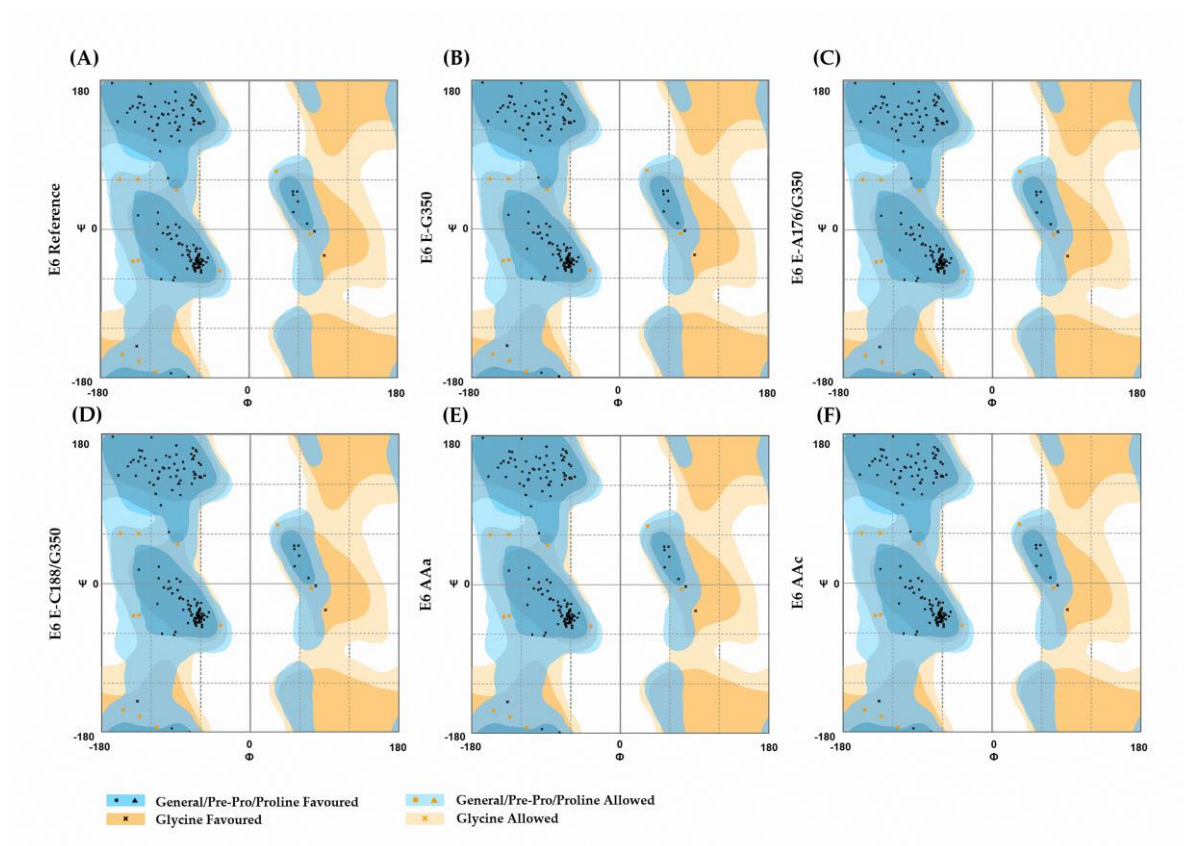


Figure 6

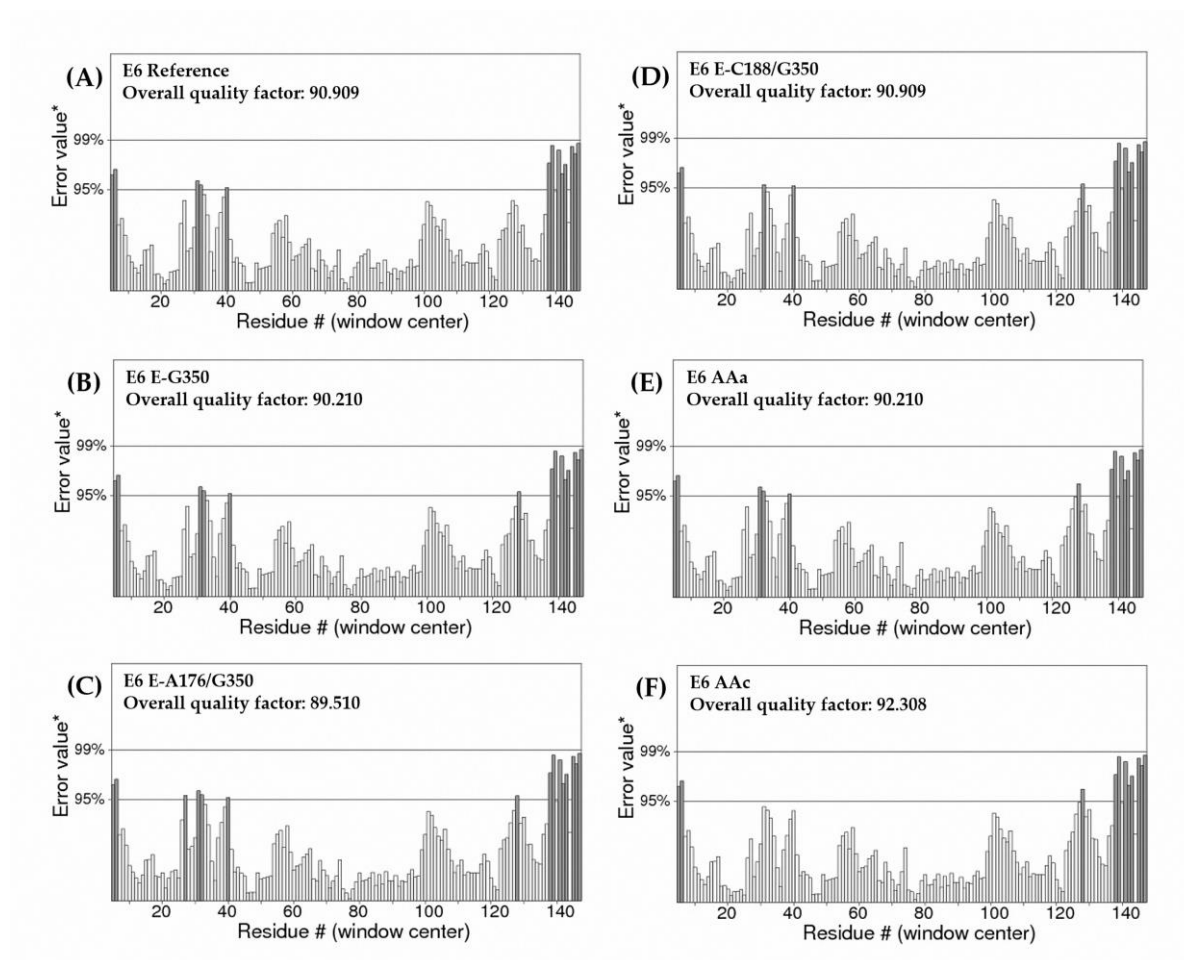


Figure 7

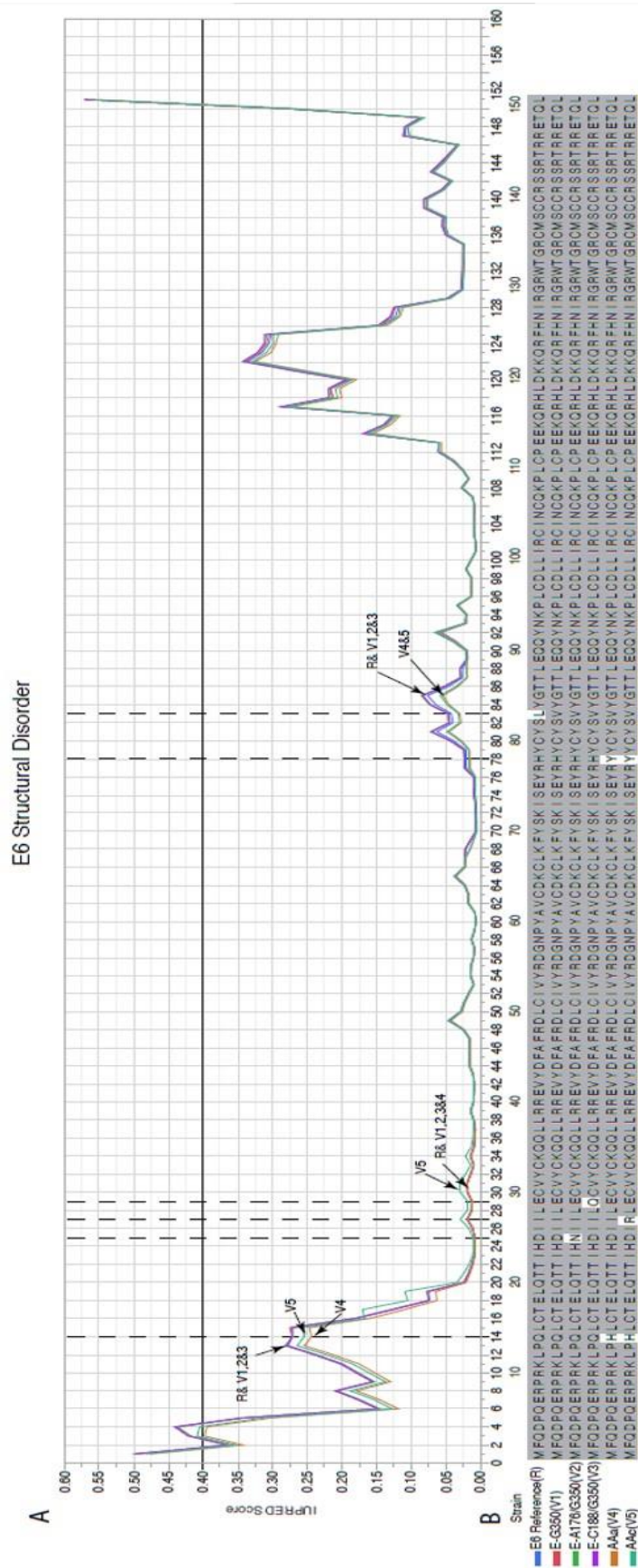
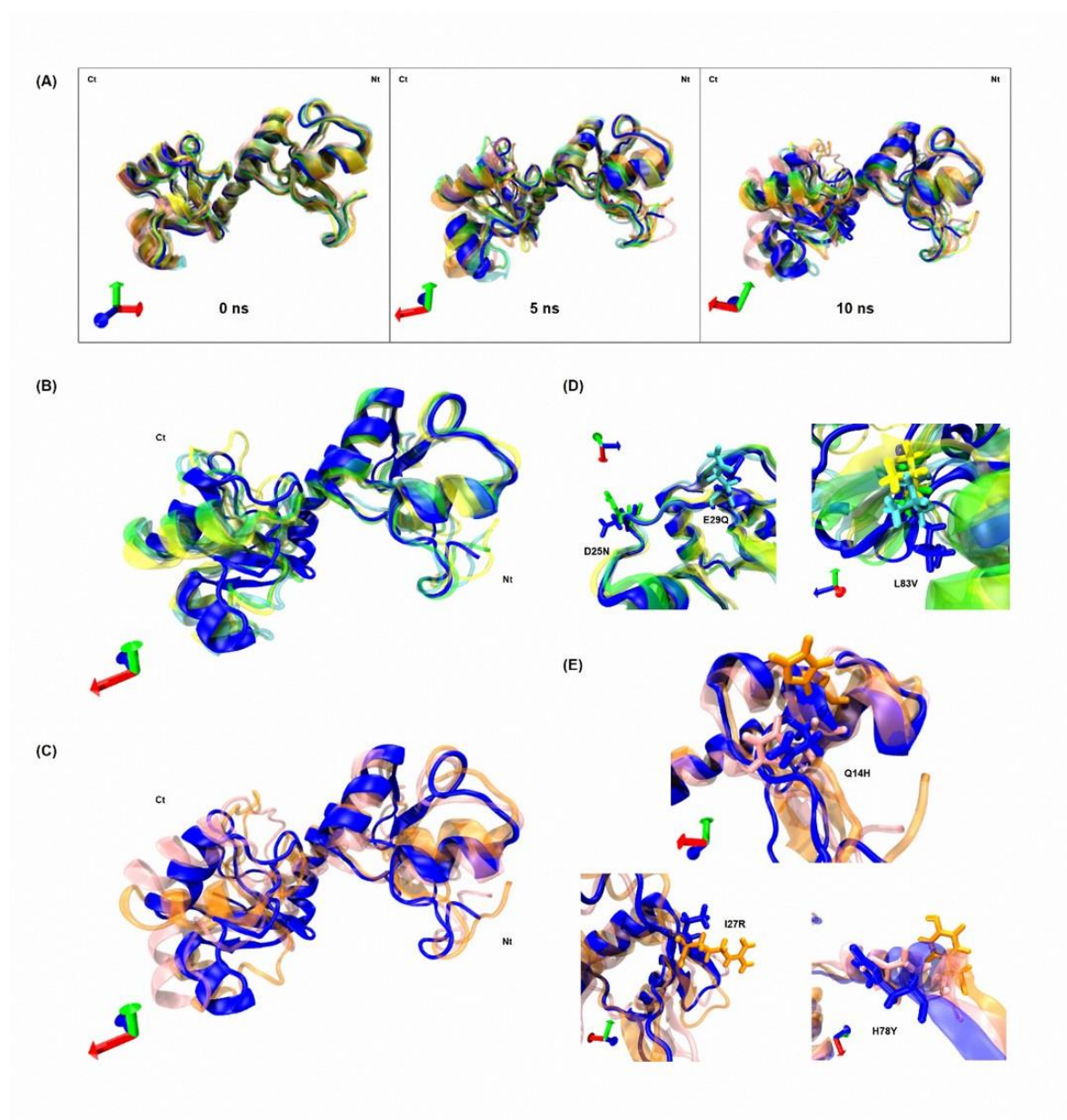
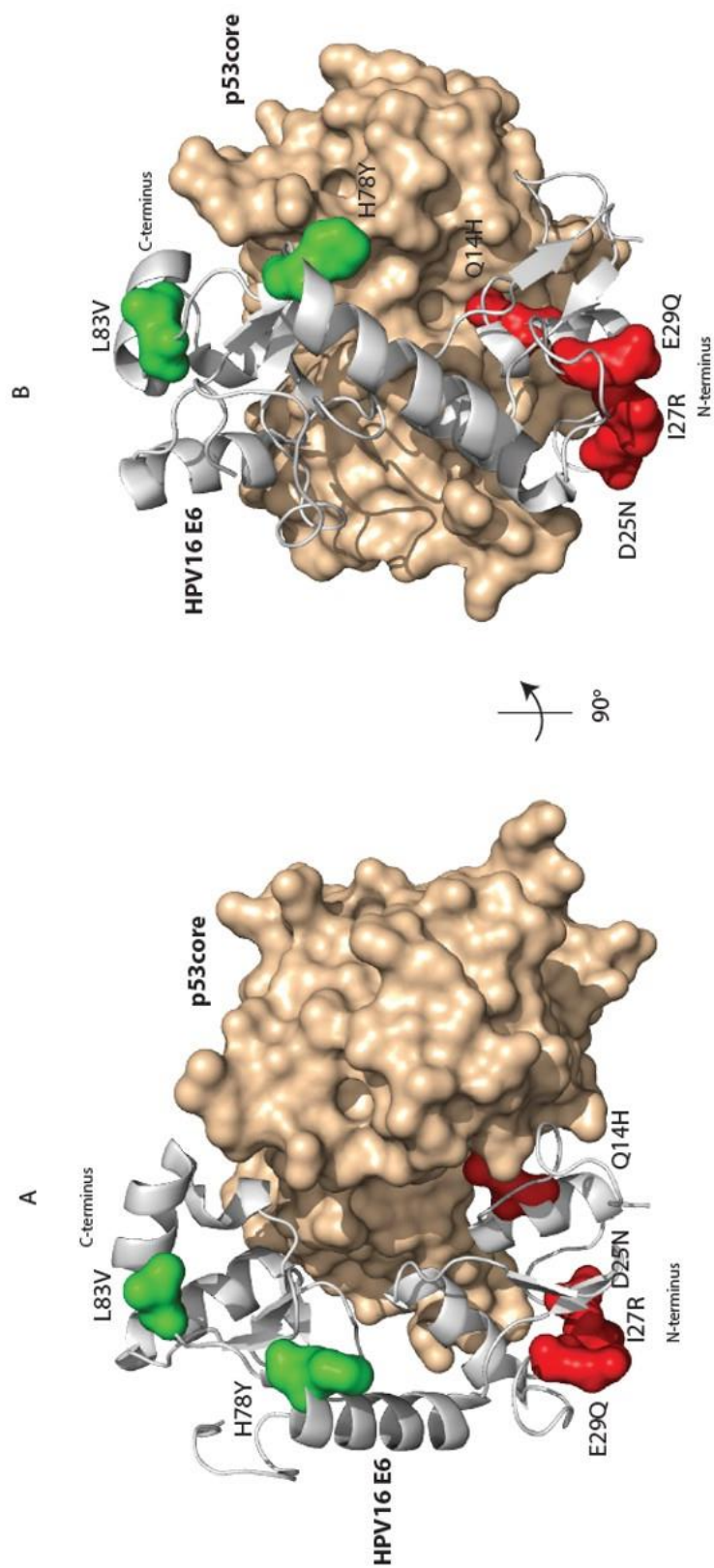


Figure 8



Supplementary Figure



Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAC), en la interacción con sus proteínas blanco.

Capítulo II

Dinámica molecular y docking de la interacción entre la oncoproteína E6 VPH 16 y sus variantes con p53

Introducción

Interacción de la oncoproteína E6 y p53

Se ha demostrado que la proteína E6 VPH16 interactúa de manera específica con la proteína p53 y que esta interacción tiene como resultado la degradación de p53 (Hengstermann et al., 2001; Mantovani y Banks, 2001). Mientras que los niveles de la proteína p53 se encuentran regulados de forma común dentro de la célula por la ubiquitina ligasa Mdm2 (Bieging et al., 2014), la proteína E6 VPH16 interactúa con p53 y directamente se ve involucrada en su degradación por el reclutamiento de una ligasa celular diferente, que es la E6AP.

La proteína p53 es ubiquitinada y degradada por E6AP en presencia de la proteína E6 del VPH16, promoviendo un mecanismo por el cual las proteínas E6 de los VPH-AR pueden reducir los niveles de p53 y favorecer su replicación en la célula huésped y así generar un estado de carcinogénesis.

Martínez-Zapien y colaboradores en el 2016 dilucidaron el complejo de interacción mediante el cual se lleva a cabo la degradación de la proteína p53 mediante la interacción de la proteína E6 del VPH16 y la E6AP. En su investigación demostraron mediante Triple pull-down y cristalografía de difracción de rayos X, que la formación de un trímero en donde se encuentran involucradas las tres proteínas es necesario para la degradación de la proteína p53, y también demostraron cuales aminoácidos por las tres partes son los que se ven involucrados en esta interacción. Además, describen la importancia que tiene el péptido de E6AP (12 aminoácidos; ELTLQELLGEER) que se encuentra involucrado en la interacción, en la conservación de la estructura de la proteína E6 VPH16 y como ayuda en la interacción con la proteína p53, aunque no se vea involucrado de forma directa en esta (Martínez-Zapien et al., 2016; Zanier et al., 2013).

Hasta el momento es el único reporte que cuenta con la información estructural de la formación del complejo ternario E6 VPH16/E6AP/p53 que está involucrado en la degradación de la proteína p53.

En nuestro grupo de trabajo, se ha demostrado que las variantes intratípicas de la proteína E6 del VPH16; variantes Europeas: E-G350, E-A176/G350, E-C188/G350;

variantes Asiático Americanas: AAa y AAa, son capaces de inducir la expresión diferencial de genes involucrados en el desarrollo del cáncer cervical (Zacapala-Gómez et al., 2016). Además, recientemente en el 2019, Rodríguez-Ruiz y colaboradores han dilucidado que los cambios en estas variantes de la oncoproteína E6 VPH 16 modifican la estructura 3D de la proteína, y que estas modificaciones afectan el movimiento de los extremos amino y carboxilo, además de aumentar los niveles del desorden de la proteína en los puntos donde se generan los cambios de aminoácidos en cada una de las variantes. Siendo esto un punto de partida importante para evaluar cómo afectan estos cambios de aminoácidos la interacción con otras proteínas blanco, se ha propuesto en este trabajo dilucidar el efecto estructural de la oncoproteína E6 del VPH 16 y sus variantes en la interacción con p53, una de sus principales proteínas blanco, esto mediante análisis *in silico* que involucra docking y dinámicas moleculares.

Materiales y métodos

Estructuras 3D

Para realizar los diferentes análisis *in silico* se usaron las estructuras 3D de la oncoproteína E6 VPH 16 y variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAa) generadas mediante modelado por homología utilizando el programa SCWRL4 (Krivov et al., 2009). La estructura 3D de la proteína p53 que se usó para los análisis fue el PDB: 4XR8 (Martínez-Zapien et al., 2016), utilizando solo el core de unión al DNA, que es también el sitio de interacción con la oncoproteína E6.

Es importante señalar que para que ocurra la degradación de p53 es necesario la unión de la ubiquitina ligasa E6AP, por tanto se utilizó el péptido ELTLQELLGEER (Martínez-Zapien et al., 2016) correspondiente al sitio de interacción con E6.

Dinámicas moleculares, simulación de E6 VPH 16 y sus variantes

Las simulaciones de dinámicas moleculares (DM) de las oncoproteínas E6 VPH16, y las variantes: E-G350, E-A176/G350, E-C188/G350, AAa y AAa se realizaron en el Cluster Híbrido Xihcoatl (<http://clusterhibrido.cinvestav.mx/>) usando los programas NAMD 2.8 y CHARMM27. Los átomos de hidrogeno se agregaron utilizando *psfgen* del programa VMD. El sistema fue llevado a una minimización de energía por 1000 tiempos seguido por un equilibrio de 1 ns y se continuó con la simulación. La simulación final corrió por 100 ns usando el ensamble NTV.

Análisis de docking

Para realizar el docking de la oncoproteína E6 VPH 16 y sus variantes con p53 se utilizó el programa ClusPro Server (Kozakov et al., 2013).

Se seleccionó como referencia el modelo ternario de interacción descrito por Martínez-Zapien et al., 2016 de E6/E6AP/p53. Se seleccionó este modelo por la información específica que se tiene de la interacción de E6 con estas dos proteínas, además de la importancia biológica que tiene la interacción de E6 VPH16 con E6AP y la proteína p53.

Para el docking se utilizaron los resultados del último frame de las dinámicas moleculares de 100 ns de cada una de las variantes de la oncoproteína E6, incluida la proteína E6 VPH 16, esto para obtener una mejor conformación de las proteínas con respecto al modelo de partida generado a través de modelos de homología.

Resultados

Dinámicas moleculares de E6 VPH 16 y variantes

Las conformaciones estructurales de la oncoproteína E6 VPH 16 y sus variantes después de las DM de 100 ns se muestran en una superposición en la Figura 1. Como se puede apreciar en los resultados de las DM al realizar la superposición de las estructuras de las diferentes variantes; E-G350, E-A176/G350, E-C188/G350, AAa y AAa de la oncoproteína E6 VPH 16 con la estructura de la E6 de referencia se observa el efecto de los cambios de aminoácidos de cada una de las variantes, siendo más evidentes los cambios presentes en los extremos amino y carboxilo de las diferentes proteínas.

Evaluación de la interacción de la oncoproteína E6 VPH 16 y variantes con p53

Utilizando las estructuras de la oncoproteína E6 VPH 16 y sus variantes obtenidas al final de las dinámicas moleculares, se procedió a realizar los análisis de interacciones. Primeramente, se realizó un docking de las oncoproteínas E6 versus un péptido de E6AP (12 aminoácidos; Figura 2), esto para obtener el complejo E6/E6AP, posteriormente utilizando este complejo se procedió a realizar los análisis de interacción contra el sitio Core de p53. Se obtuvieron los diferentes patrones de unión entre E6/E6AP versus el Core de p53.

En la Figura 3 se muestra el resultado solo de la interacción de la oncoproteína E6 VPH16/E6AP versus p53, y se resaltan los sitios donde se localizan cada uno de los cambios presentes en las diferentes variantes de E6.

La proteína E6 VPH 16 y las variantes E-G350, E-A176/G350, E-C188/G350, AAa y AAa, tienen diferentes patrones de interacción tanto con el péptido de E6AP, así como este complejo (E6/E6AP) con respecto al Core de p53. Las variantes

pertenecientes al Linaje Europeo (E6 VPH16 y las variantes E-G350, E-A176/G350, E-C188/G350), tienen similitudes en cuanto a la formación de los patrones de interacción en el trímero con el péptido de E6AP y el Core de p53.

En el trímero de interacción de la oncoproteína E6 VPH16 con E6AP los aminoácidos involucrados por parte de la proteína E6 VPH16 son; para la interacción de la proteína E6 con el péptido de E6AP: C51, V53, V62, V131, R55, T32, T70, R77, **H78**, P5, R8, R10, I128, R129, S74, R131, Q107, L50 Y L67, y por parte del péptido de E6AP: E382, R383, L373, L378, E377, T374, L375, L379, E381 y E372. Para la interacción de la proteína E6 VPH16 con la proteína p53 los aminoácidos involucrados son: Q20, R10, K94, Q6, R8, E7, P9, E18, **Q14**, T17, P13, F47, D98, L99, L100, C97 y L115, y por parte de la proteína p53: Q144, D228, T118, G117, A119, S116, H115, L114, G105, T107, G108, S106, D148, T148, Q104, N268, R110, L101, T102, Q100, R290, A129 y N131.

Las variantes Europeas E-G350, E-A176/G350, E-C188/G350, difieren de la oncoproteína E6 VPH 16 en distintos aminoácidos, específicamente en la interacción con p53. Para E-G350 los aminoácidos que difieren por parte de E6; P5, Q3, D4, T22, C16, L19 y T21., por p53 son; F270, S269, K164, P128, S227, K120 y L290. Para E-A176/G350 por E6; P95, P112, T43, I23, H24, C111, R8, P9 y M1, por p53; T146, S227, G226, C229, P128, T150, S99, S149, S269 y R267. Finalmente, para E-C188/G350 por E6; D4, F2, P112, A46, T43, H24 y K11, por p53; K164, L130, S99, K292, K291 y P128.

De igual forma, las variantes E6 VPH16 AAa y AAc se comportan de forma distinta, aunque pertenezcan al mismo linaje. En la interacción dentro del trímero, no sólo difieren en los patrones de interacción y aminoácidos con las variantes Europeas, sino que también entre las dos variantes, E6 VPH16 AAa y AAc, ya que presentan modificaciones en los patrones de interacción en algunos aminoácidos.

En el trímero E6 VPH16 AAa, E6AP y la proteína p53, los aminoácidos que se encuentran involucrados por parte de la proteína E6 VPH 16 AAa en la interacción con el péptido de E6AP son: V62, F45, C51, V53, R55, Y32, L50, L67, Y70, Q107, R102, L100, S74, R131, **Y78** y R129; y por parte del péptido de E6AP: L378, E377, T374, L375, L379, E381, E372 y Q376. Los aminoácidos involucrados de proteína

E6 VPH16 AAa en la interacción con p53 los aminoácidos son: Q20, R10, K94, Q6, R8, E7, P9, E18, **H14**, T17, D49, P13, F47, I23, H24, Y43, D44, R40 y P5; y por parte de la proteína p53 los aminoácidos son: S166, D148, G108, F109, R110, Q104, G105, S106, Y103, K101, R267, T102, S269, N131, P128, D288, N228, L130, A129, K292, Q100, L289, R290 y E286.

Por otro lado, en el trímero E6 VPH16 AAc, E6AP y la proteína p53, los aminoácidos involucrados en la interacción y los patrones de interacción son distintos. En este trímero para la interacción de la proteína E6 VPH16 AAc y el péptido de E6AP los aminoácidos involucrados por parte de la proteína E6 y del péptido E6AP son los mismos que en la proteína E6 VPH 16 AAa y el péptido, sin embargo, en la interacción de E6 VPH 16 AAc y la proteína p53 cambian los aminoácidos y los patrones de interacción. En la interacción de la proteína E6 VPH16 AAc con la proteína p53, los aminoácidos involucrados son los siguientes: Q20, R10, K94, Q6, R8, E7, P9, E18, **H14**, T17, D49, P13, F47, I23, H24, Y43, D44, R40 y P5; y por parte de la proteína p53 los aminoácidos involucrados son: N131, S269, D228, R267, K101, T102, Y103, S106, G105, Q104, D148, G108, R110, Q100, S166, P128, L130, A129, R290, E286, L289, K292 y D288. Estos valores son representados en la Tabla 1.

Tabla1. Redes de interacción resultados del Docking. Tabla del tipo y número de interacciones entre la variante E6 VPH16 y las proteínas E6AP y p53.

E6 VPH16/E6AP vs P53	Residuos en interacción	Enlaces iónicos	Puentes de Hidrogeno	Interacciones electroestáticas
E6AP	19:10	4	3	146
P53	17:23	3	12	157
E6 E-G350/E6AP vs P53				
E6AP	21:11	6	7	158
P53	17:20	2	12	131
E6 E-A176-G350/E6AP vs P53				
E6AP	19:10	5	5	100
P53	22:21	1	9	170
E6 E-C188-G350/E6AP vs P53				
E6AP	18:12	5	2	143
P53	20:21	2	6	206
E6 AAa/E6AP vs P53				
E6AP	15:12	5	14	105
P53	34:30	1	16	380
E6 AAc/E6AP vs P53				
E6AP	11:11	4	2	131
P53	23:25	2	7	148

Discusión

En este trabajo, se han aplicado técnicas *in silico* actuales de alta calidad, para conocer las diferencias generadas por cambios de aminoácidos en estructura e interacción de la oncoproteína E6 del VPH16 y sus variantes; E-G350, E-A176/G350, E-C188/G350, AAa y AAa con p53. Se ha demostrado que en las diferentes variantes genotípicas de E6 VPH16 generan diferentes perfiles de expresión de genes (Zacapala-gomez) asociados a diferentes patrones de degradación de p53 (Dalal et al., 1996; Martínez-Zapien et al., 2016), además de favorecer procesos celulares malignos en la generación de cáncer cervical (Moody and Laimonis, 2010; Duensing y Münger, 2002; Duensing et al., 2000).

Al analizar la interacción de las dos moléculas principales involucradas en la desregulación de procesos celulares que están involucrados en los cambios patológicos en las células, se ha dilucidado teóricamente, mediante análisis *in silico* el efecto de mutaciones de aminoácidos en la interacción de dos proteínas, la oncoproteína E6 y sus variantes versus p53.

Inicialmente se obtuvieron estructuras 3D de la oncoproteína E6 VPH16 y sus variantes; E-G350, E-A176/G350, E-C188/G350, AAa y AAa, esto mediante modelamiento por homología, usando el software SCWRL4 (Krivov et al., 2019) que es capaz de predecir cambios particulares de aminoácidos por mutaciones puntuales. Posterior a la generación de las estructuras se realizaron dinámicas moleculares de cada una de las variantes para ver los efectos estructurales de los cambios de aminoácidos, observando que todas las variantes presentan diversos

niveles de flexibilidad en sus cadenas laterales dado por estas mutaciones.

Las variantes analizadas pertenecientes al linaje Europeo que fueron; E-G350, E-A176/G350 y E-C188/G350, incluyendo a la E6 VPH16 presentan una flexibilidad estructural similar en su extremo amino y carboxilo. Mientras que las variantes analizadas pertenecientes al linaje Asiático Americano; AAa y AAc se comportan de manera distinta a las variantes Europeas. Estas dos variantes presentan un mayor cambio estructural en su extremo carboxilo, pero también tiene un ligero efecto en el extremo amino, esto en comparación con la E6 VPH16. Tamarozzi y Giuliatti en el 2018 han evaluado la capacidad intrínseca del desorden la oncoproteína E6 de diferentes tipos de VPH's, incluyendo VPH de Bajo Riesgo, como VPH6 y de Alto Riesgo como los VPH16 y VPH18, donde observaron que los efectos de los cambios en las secuencias de las proteínas E6 se ven principalmente en los extremos amino y carboxilo, concluyen que las mutaciones puntuales y los cambios de la secuencia en las regiones amino y carboxilo favorecen el desorden de esta proteína permitiéndole mayor movilidad en sus extremos, y aumentando su capacidad de interactuar con otras proteínas blanco.

Para evaluar las diferencias en la capacidad de interactuar de la oncoproteína E6 y sus variantes con p53, usando estas estructuras realizamos experimentos de docking. Primero, realizamos docking de las estructuras sin dinámicas moleculares, de las oncoproteínas E6 con un péptido de E6AP y estas dos con p53 (Figura suplementaria 1), esto debido a que para que ocurra la degradación de P53 es necesaria la unión primera con E6AP (Hengstermann et al., 2001; Mantovani y Banks, 2001), sin embargo, se debe señalar que para la interacción de

la proteína E6 con P53 no es necesario la unión primera de E6AP, pueden ocurrir como eventos independientes (Martínez-Zapien et al., 2016). Posteriormente, se generaron interacciones moleculares entre proteínas generadas a partir de las dinámicas moleculares de las oncoproteínas E6.

El docking que se generó sin dinámicas moleculares fue para tener un punto de partida al momento de evaluar las interacciones usando dinámicas moleculares, y poder observar a los aminoácidos que permanecían incluidos en las interacciones, a los aminoácidos que se perdían en las interacciones y el número de las interacciones generadas o perdidas (Tabla1).

Observamos que con estos modelos generados las capacidades de generar interacciones entre la oncoproteína E6 y p53 son diferentes. En comparación con el primer análisis sin dinámicas moleculares que realizamos donde no se observan diferencias entre las variantes Europeas, pero si entre las variantes Asiático Americanas.

Con las dinámicas moleculares el comportamiento de las variantes Europeas; E-G350, E-A176/G350, E-C188/G350 en comparación con la E6 VPH16 (Figura Suplementaria 1), muestra el efecto de los cambios de aminoácidos en la interacción con P53, donde podemos observar un mayor número de interacciones en la variante E-C188/G350 con respecto a las variantes Europeas (Tabla1). Es interesante señalar que esta variante presenta mutaciones en el extremo amino (E29Q), que es donde se observó un mayor desorden en las variantes Europeas. Mientras que en las variantes Asiático Americanas la E6 AAa presente el mayor número de interacciones con p53, que involucran no sólo el número de interacciones generadas

(iónicas, puentes de hidrogeno, electroestáticas), sino también el número de aminoácidos involucrados en el sitio de interacción, esto en comparación con la E6 AAc e incluso con las variantes Europeas.

Todos estos datos son de importancia para evaluar la capacidad de la oncoproteína E6 para interactuar con p53. En un estudio experimental e *in silico* White y colaboradores en el 2012, evaluaron la capacidad de interactuar de las oncoproteínas E6 de diferentes tipos de VPH con diversas proteínas blancos celulares, y observaron que cambios de un aminoácido o la pérdida de una pequeña región de la proteína E6 pueden afectar en gran forma la interacción con sus blancos, mostrando principalmente las diferencias en la capacidad de interactuar con proteínas importantes para la regulación de señales celulares involucradas en procesos como apoptosis y ciclo celular (White et al., 2012).

Este estudio provee información importante no sólo del número de interacciones diferencial entre las variantes de la oncoproteína E6, sino que permite ver el efecto de como la presencia de mutaciones en una misma proteína puede afectar la interacción con sus proteínas blanco.

Conclusión

El modelo molecular y las simulaciones de dinámica molecular de estas variantes de la oncoproteína E6 permiten ver el efecto de las mutaciones puntuales de cada variante.

Aunque las variantes de la proteína E6 no muestran cambios estructurales en lo general, su capacidad de interacción con otras proteínas como p53 fue diferente, ganando o perdiendo interacciones en comparación con la proteína E6 VPH16. Dichas interacciones

favorecen el desorden de las variantes Asiático Americanas en comparación con las Europeas.

Aminoácidos como la Histidina en la posición 14 podrían ser considerados como futuro blanco para el desarrollo de fármacos que beneficien el efecto dinámico de oncoproteínas como E6 en la gran plasticidad que poseen para interaccionar con múltiples proteínas celulares íntimamente relacionadas con procesos como el desarrollo y progresión del cáncer cervical.

Figuras

Figura 1. Dinámicas moleculares de la oncoproteína E6. En la figura se observa una superposición del último frame resultado de las dinámicas moleculares de 100 ns la oncoproteína E6 y sus variantes. Código de colores: oncoproteína E6 VPH 16 (azul); E6-G350 (amarillo); E6-A176/G350 (verde); E6-C188/G350 (purpura); E6 AAa (naranja) y E6 AAc (rojo).

Figura 2. Interacción de la oncoproteína E6 VPH16 ref vs E6AP. En la figura se observa el resultado del docking entre la oncoproteína E6 VPH16 versus E6AP. Código de colores: oncoproteína E6 VPH 16 (azul) y péptido E6AP (naranja).

Figura 3. Interacción de la oncoproteína E6 VPH16 ref/E6AP versus p53. En la figura se observa el resultado del docking entre la oncoproteína E6 VPH16/E6AP versus el Core de p53. Código de colores: oncoproteína E6 VPH 16 (azul); péptido E6AP (naranja) y p53 Core (verde).

Figuras 1.

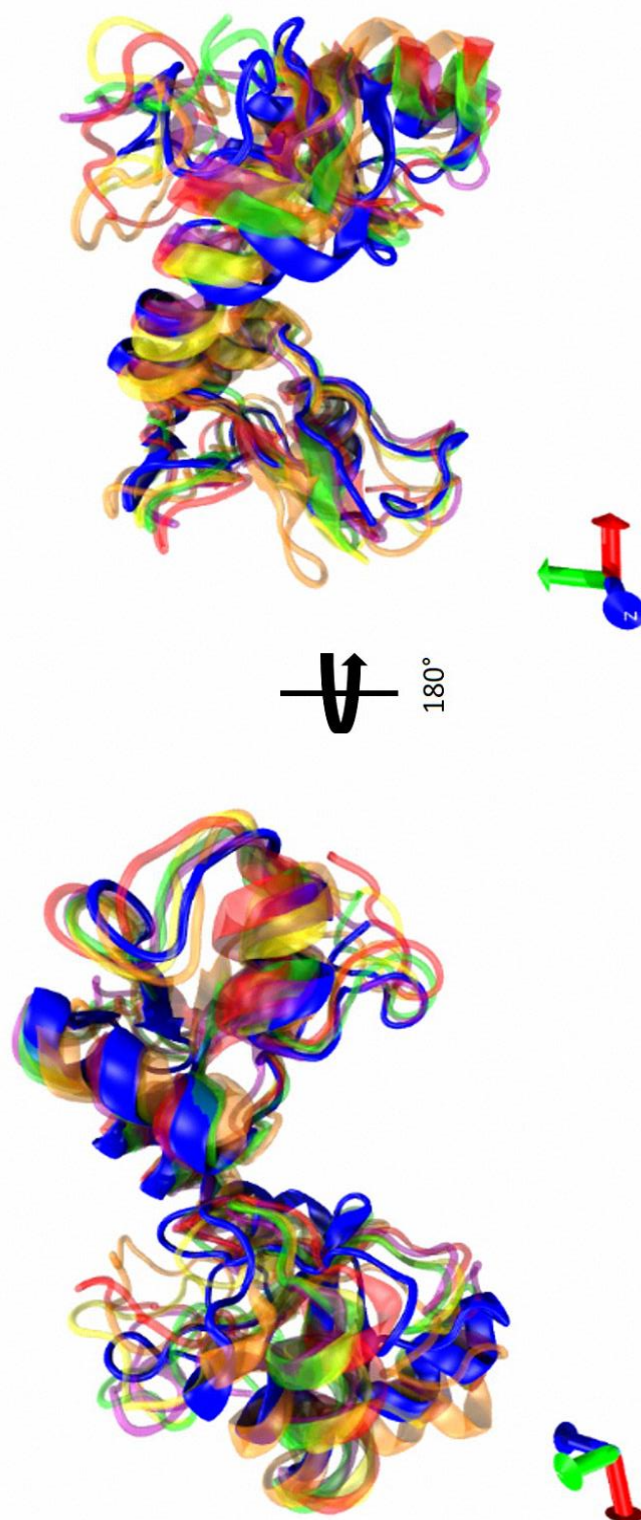


Figura 2.

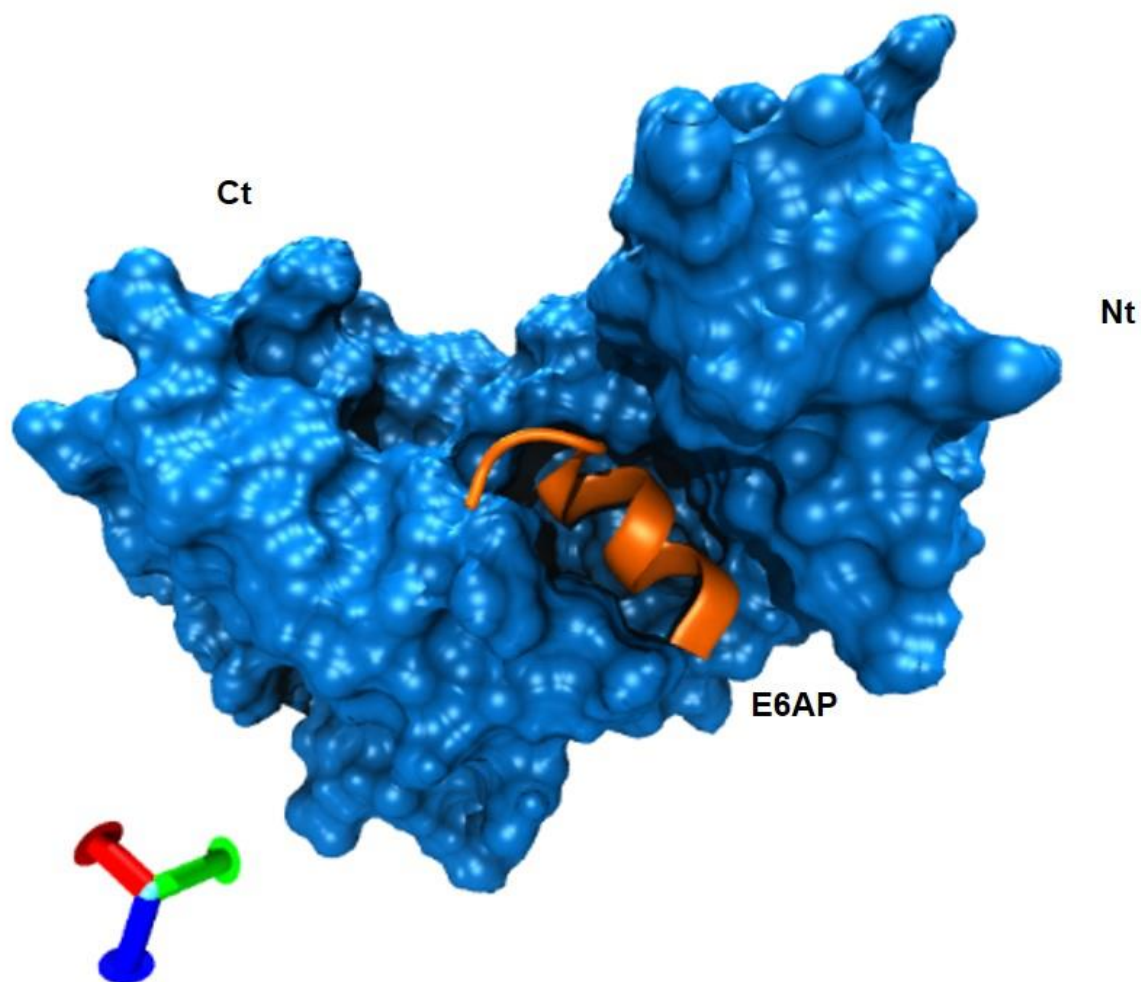
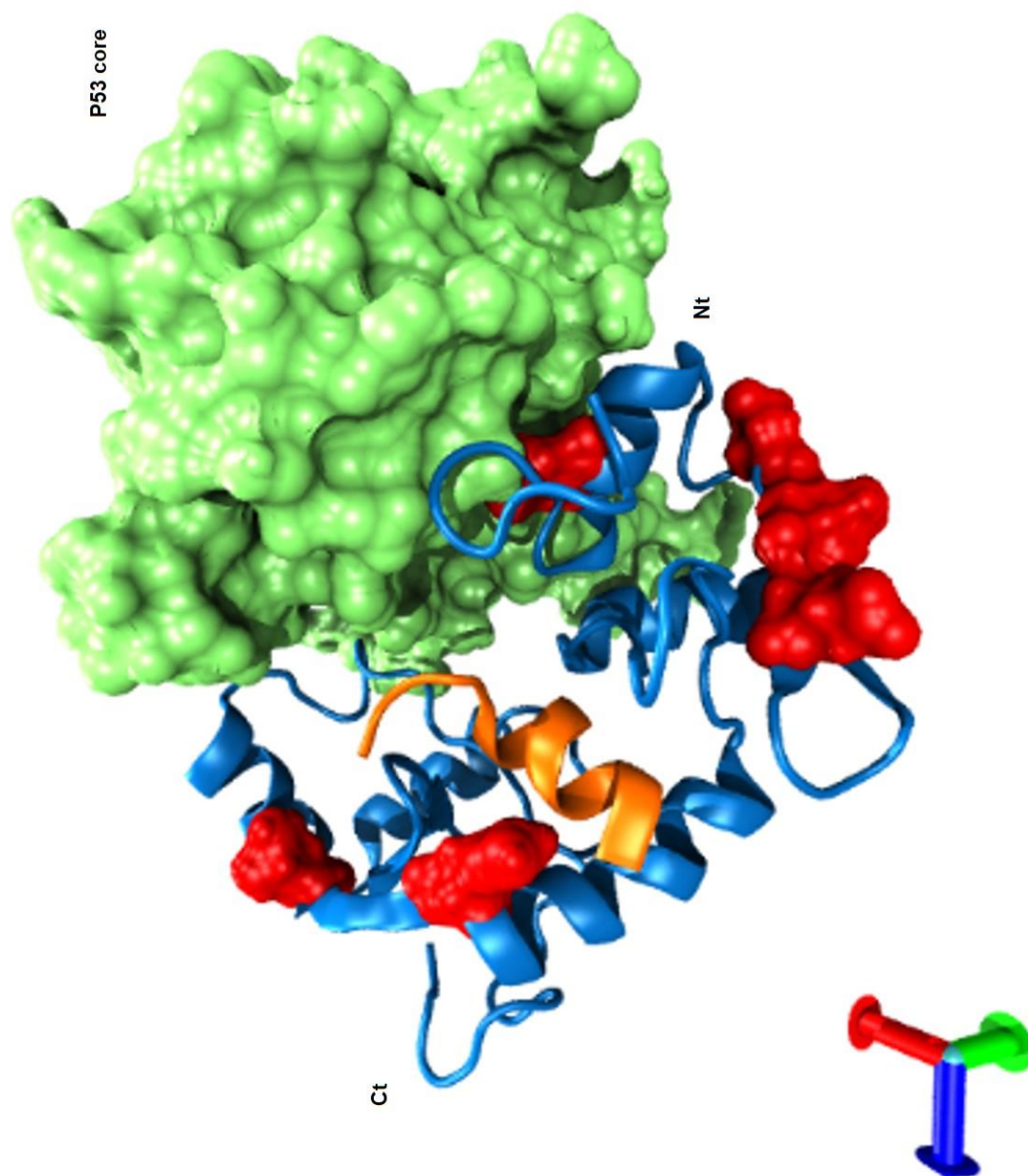


Figura 3.



Anexos

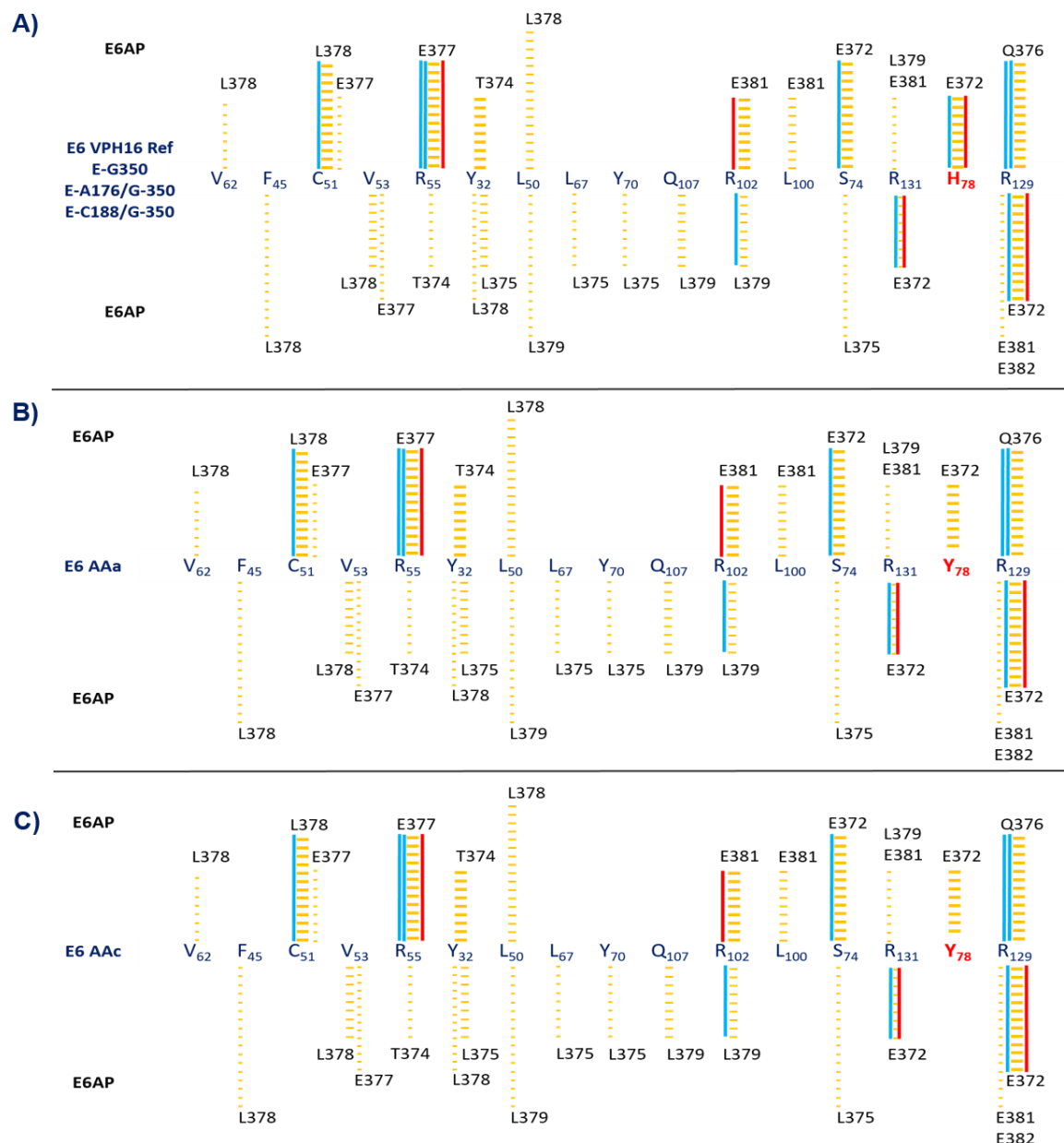


Figura Suplementaria 1. Redes de interacción resultados del Docking sin dinámicas moleculares.

Redes de interacción resultados del Docking de E6 VPH16 WT y variantes: Europeas (EUR); E-G350, E-A176/G-350, E-C188/G-350 y Asiático Americanas: AAa y AAC., versus las proteínas E6AP y p53. **A)** Redes de interacción entre E6 VPH16 EUR y E6AP. **B)** Redes de interacción entre E6 VPH16 AAa y E6AP. **C)** Redes de interacción entre E6 VPH16 AAC y E6AP. **Código de colores:** Letras azul oscuro: aminoácidos E6 VPH16 y variantes; Letras rojas: aminoácidos E6 VPH16 de interés; Letras negras: aminoácidos de E6AP; Letras rosas: aminoácidos de la proteína p53; Letras tachadas: pérdida de interacción; Letras subrayadas: nueva interacción; Líneas azules: puentes de hidrogeno; Líneas rojas: enlaces iónicos; Líneas amarillas gruesas punteadas: interacciones

electroestáticas fuertes; Líneas amarillas punteadas: interacciones electroestáticas débiles; Sombras negras: pérdida del enlace.

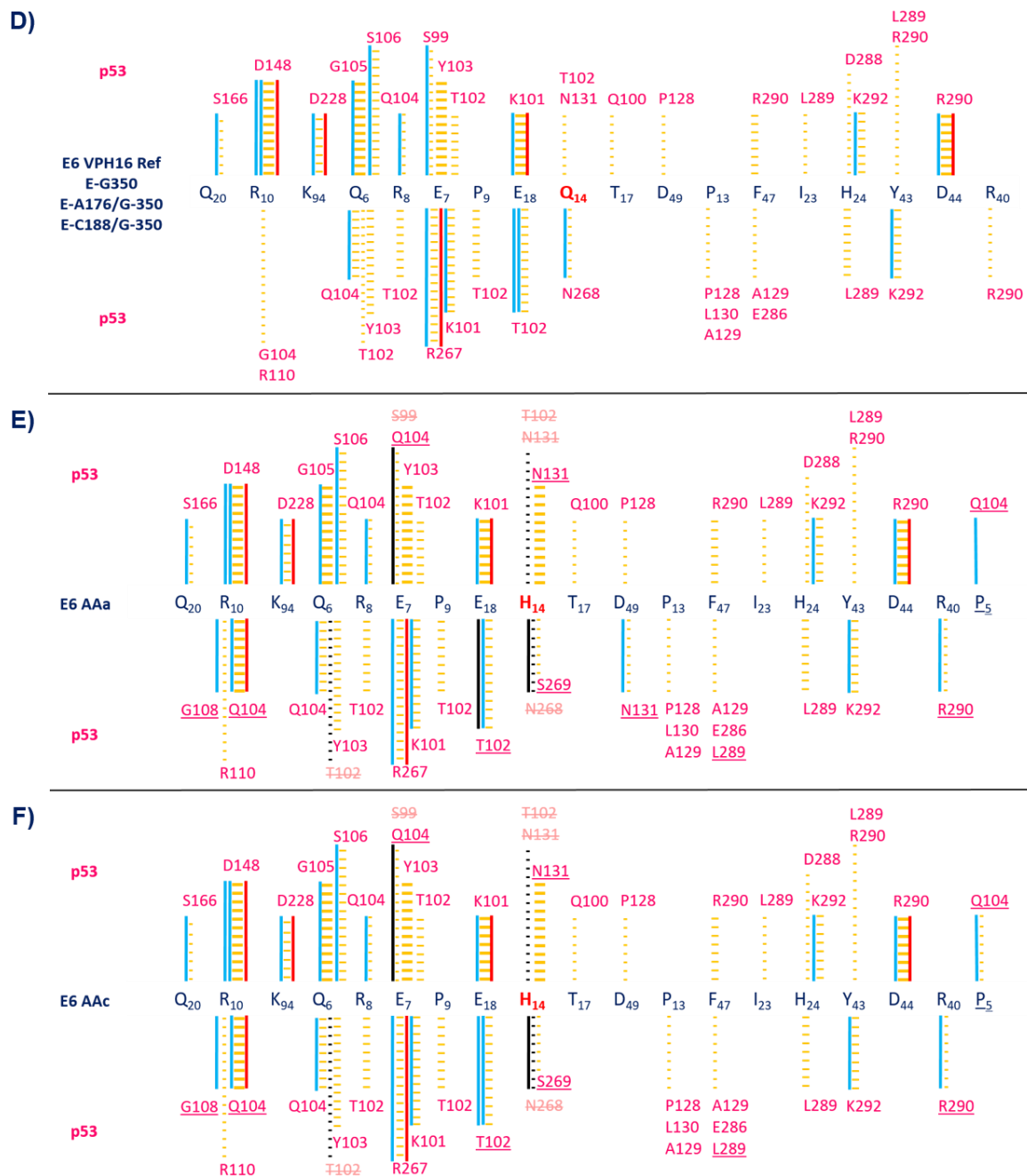


Figura Suplementaria 1. Redes de interacción resultados del Docking sin dinámicas moleculares.

D) Redes de interacción entre E6 VPH16 EUR y la proteína p53. **E)** Redes de interacción entre E6 VPH16 AAa y la proteína p53. **F)** Redes de interacción entre E6 VPH16 AAc y la proteína p53.

Código de colores: Letras azul oscuro: aminoácidos E6 VPH16 y variantes; Letras rojas: aminoácidos E6 VPH16 de interés; Letras negras: aminoácidos de E6AP; Letras rosas: aminoácidos de la proteína p53; Letras tachadas: pérdida de interacción; Letras subrayadas: nueva interacción; Líneas azules: puentes de hidrogeno; Líneas rojas: enlaces iónicos; Líneas amarillas gruesas punteadas: interacciones electroestáticas fuertes; Líneas amarillas punteadas: interacciones electroestáticas débiles; Sombras negras: pérdida del enlace.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc), en la interacción con sus proteínas blanco.

G)	E6 VPH16 EUR vs	No. De residuos en interacción	No. De enlaces iónicos	No. De enlaces disulfuro	No. De puentes de hidrogeno	No. Interacciones electroestáticas
	E6AP	16:9	5	-	10	87
	p53	18:23	5	-	18	174

H)	E6 VPH16 AAa vs	No. De residuos en interacción	No. De enlaces iónicos	No. De enlaces disulfuro	No. De puentes de hidrogeno	No. Interacciones electroestáticas
	E6AP	16:9	4	-	9	92
	p53	19:24	5	-	21	204

I)	E6 VPH16 AAc vs	No. De residuos en interacción	No. De enlaces iónicos	No. De enlaces disulfuro	No. De puentes de hidrogeno	No. Interacciones electroestáticas
	E6AP	16:9	4	-	9	91
	p53	19:23	5	-	21	202

Figura Suplementaria 1. Redes de interacción resultados del Docking sin dinámicas moleculares.

G) Tabla del tipo y número de interacciones entre la variante E6 VPH16 EUR y las proteínas E6AP y p53. **H)** Tabla del tipo y número de interacciones entre la variante E6 VPH16 AAa y las proteínas E6AP y p53. **I)** Tabla del tipo y número de interacciones entre la variante E6 VPH16 AAc y las proteínas E6AP y p53.

Referencias

Bieging, K.T., Mello, S.S., Attardi, L.D., 2014. Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* 14, 359–370. doi:10.1038/nrc3711

Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38. <http://www.ks.uiuc.edu/Research/vmd/>

James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26:1781-1802, 2005.

Martinez-Zapien, D., Ruiz, F.X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Pol, S.V., Podjarny, A., Travé, G., Zanier, K. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* 529, 541–545. doi:10.1038/nature16481

Zacapala-Gómez, A.E., et al., Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells. *Virology*, 2016. 488: p. 187-195 DOI: 10.1016/j.virol.2015.11.017.

Zanier, K., Charbonnier, S., Sidi, A. O. M. O., McEwen, A. G., Ferrario, M. G., Poussin, P., Travé, G. (2013). Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* (New York, N.Y.), 339(6120), 694–698. <http://doi.org/10.1126/science.1229934>

Zanier, K., Charbonnier, S., Sidi, A. O. M. O., McEwen, A. G., Ferrario, M. G., Poussin, P., ... Travé, G. (2013). Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* (New York, N.Y.), 339(6120), 694–698. <http://doi.org/10.1126/science.1229934>

G. G. Krivov, M. V. Shapovalov, and R. L. Dunbrack, Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* (2009).

Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38. <http://www.ks.uiuc.edu/Research/vmd/>

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAc), en la interacción con sus proteínas blanco.

UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. J Comput Chem. 2004 Oct;25(13):1605-12. <http://www.rbvi.ucsf.edu/chimera>

James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26:1781-1802, 2005.

Discusión general

Se sabe que la oncoproteína E6 HPV16 está involucrada en varios procesos biológicos que llevan a transformación maligna, como el ciclo celular, la apoptosis, la reparación del ADN, la respuesta inmune, la organización de la cromatina y la comunicación celular (Howie et al., 2009; Mittal y Banks 2017).

El uso de la bioinformática para conocer las diferentes condiciones de una patología, han logrado grandes desarrollos técnicos y metodológicos. Esto debido a la precisión de los análisis y a su alta confiabilidad, ya que se basan en los conocimientos biológicos ya obtenidos para la generación y desarrollo de nuevo conocimiento.

En este trabajo, se han aplicado técnicas *in silico* actuales de alta calidad, para conocer las diferencias generadas por cambios de aminoácidos en estructura e interacción de la oncoproteína E6 del VPH16 y sus variantes; E-G350, E-A176/G350, E-C188/G350, AAa y AAc. Esto debido a que se ha observado que estos cambios favorecen diferentes funciones por parte de estas oncoproteínas en expresión de genes y generación de procesos malignos.

En primera estancia, se generaron estructuras 3D mediante modelado por homología de las diferentes oncoproteínas E6, esto utilizando un software que permite el análisis estructural de mutaciones de aminoácidos en una proteína, al hacer el análisis de estas estructuras no se encontró ningún cambio significativo de la oncoproteína E6 ni de sus variantes. Esto debido a que los modelos generados solo modificaban las cadenas laterales de los aminoácidos mutados.

Las simulaciones de dinámicas moleculares son análisis realizados de manera *in silico* utilizando cómputo de alto desempeño, estas permiten el estudio de moléculas y su comportamiento dentro de un sistema. Nosotros realizamos dinámicas moleculares para evaluar el comportamiento estructural de la oncoproteína E6 y sus variantes. Lo que queríamos, era analizar el efecto de las mutaciones en un contexto simulado.

Este análisis reveló que estas mutaciones no sinónimas en la oncoproteína E6 causan cambios estructurales sutiles y potencialmente funcionales, que pueden desempeñar un papel importante en la mediación de cómo interactúan con otras proteínas, como lo es P53.

Para comprobar esto y evaluar las diferencias en la capacidad de interactuar de la oncoproteína E6 y sus variantes con p53, usando estas estructuras realizamos experimentos de docking, que permiten evaluar la interacción entre moléculas. Primero, realizamos docking de las estructuras sin dinámicas moleculares para tener un punto de partida y conocer que aminoácidos se encontraban involucrados en esta interacción. El proceso de docking fue, la oncoproteína E6 con un péptido de E6AP y estas dos contra p53 (Figura suplementaria 1, capítulo II), esto debido a que para que ocurra la degradación de p53 es necesaria la unión primera con E6AP (Hengstermann et al., 2001; Mantovani y Banks, 2001), sin embargo, se debe señalar que para la interacción de la proteína E6 con p53 no es necesario la unión primera de E6AP, pueden ocurrir como eventos independientes (Martínez-Zapien et al., 2016). Posteriormente, repetimos el análisis pero ahora utilizando las estructuras obtenidas después de las dinámicas moleculares de las oncoproteínas E6. Con estos modelos de docking

obtenidos observamos que las capacidades de generar interacciones entre la oncoproteína E6 y sus variantes con p53 son diferentes, y estas interacciones diferenciales pueden determinar posteriormente otros procesos.

En particular, como resultado de los análisis de interacción pudimos observar que la mutación presente en las variantes AA (a y c) Q14H, está justo en la interfaz de interacción de E6 y P53, y se ha visto en estudios funcionales previos este sitio (Q14) desempeña un papel importante en la unión de P53 y su posterior degradación. Otras mutaciones ubicadas alrededor de este y otros sitios activos pueden contribuir de forma alostérica a la unión y degradación de P53 (Dalal et al., 1996; Martínez-Zapien et al., 2016), como lo serían las mutaciones en las demás variantes.

El presente estudio proporciona un marco para evaluar el mecanismo mediante el cual el efecto estructural de las variantes; E-G350, E-A176/G350, E-C188/G350, AAa y AAa, con respecto a la oncoproteína E6 VPH16 podría estar relacionado con interacciones moleculares diferenciales en procesos malignos clave dentro de las células huésped, y lo más importante, establecer un nuevo punto de vista sobre el diseño de nuevas estrategias terapéuticas basadas en oncoproteínas como dianas moleculares.

Conclusiones generales

El análisis *in silico* muestra que las secuencias primarias de todas las variantes de la oncoproteínas analizadas tienen variaciones en los primeros 83 aminoácidos; Sin embargo, su estructura general 3D no cambia.

Las simulaciones de dinámica molecular de estas variantes de la oncoproteína E6 mostraron cambios sutiles en la estructura general de la proteína, pero cambios amplios en sus extremos amino y carboxilo, además de modificaciones en parámetros fisicoquímicos, que posiblemente estén involucrados en el patrón diferencial de interacciones con otras moléculas.

Aunque las variantes de la proteína E6 no muestran cambios estructurales en lo general, los cambios de aminoácidos afectan su capacidad de interacción con otras proteínas como lo fue con p53.

Conocer los sitios de interacción y los aminoácidos involucrados, entre la proteína E6 y sus variantes con p53, favorece a futuro el desarrollo de alguna estrategia terapéutica (como fármacos) que evite la interacción de estas proteínas.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAC), en la interacción con sus proteínas blanco.

Análisis estructural y funcional de la oncoproteína E6 del VPH16 y sus variantes (E-G350, E-A176/G350, E-C188/G350, AAa y AAC), en la interacción con sus proteínas blanco.

ANEXOS

TERMOFORESIS DE MICROESCALA: ANÁLISIS DE LA INTERACCIÓN DE LA ONCOPROTEINA E6 VPH 16 Y SUS VARIANTES CON P53

Introducción

Debido a la importancia de las interacciones de proteínas, y la caracterización de estas interacciones se han desarrollado diversas técnicas computacionales y experimentales para detectar y predecir los patrones de interacción de proteínas (Planas-Iglesias et al., 2013; Tuncbag et al., 2008). La técnica de Termoforesis de Microescala (MST), es un nuevo y revolucionario enfoque biofísico para medir interacciones moleculares. MST detecta los cambios en la capa de hidratación de las moléculas y las medidas de las interacciones moleculares en condiciones fisiológicas y no se requiere la inmovilización del sustrato que se utiliza. El principio de la MST es elevar el flujo de las moléculas en un gradiente de temperatura, así la entropía de solvatación y la capa de hidratación de las moléculas proporcionan la fuerza motriz del flujo, y cualquier cambio en la capa de hidratación de las biomoléculas debido a cambios estructurales de estas afectara la movilidad termoforética, este dato se utilizará para determinar afinidades de unión y las actividades enzimáticas con alta precisión y sensibilidad (Jerabek-Willemsen et al., 2014).

OBJETIVO 4: Validar las interacciones moleculares entre la oncoproteína E6 VPH16 y sus variantes (E-G350, E-A176/G-350, E-C188/G-350, AAa, AAc) con las proteínas blanco (p53).

- Construcciones de E6 VPH16 WT y variantes
- Validación de las construcciones
- Producción de las proteínas recombinantes
- Análisis de la interacción proteína-proteína por Termoforesis de Microescala (Monolith NT.115)

En relación a la obtención de las proteínas recombinantes:

Los resultados que se tienen de este objetivo son:

1. **Construcciones** de las variantes en el vector de expresión: pQE30-E6AAc, pQE30-E6188 y pQE30-E6350.
2. Generación de **proteínas recombinantes**: Se están haciendo ensayos de inducción con la construcción pQE30-AAc a diferentes tiempos y concentraciones de IPTG. Esto debido a que el primer intento de inducción como se propone en la literatura (Illiano et al., 2016) no fue satisfactorio. Actualmente, se ha logrado obtener proteína solo cambiando las concentraciones de IPTG y aumentando los tiempos de inducción. En la **Figura 1** se muestra un Western Blot resultado preliminar de la inducción de la oncoproteína E6 VPH 16 AAc.

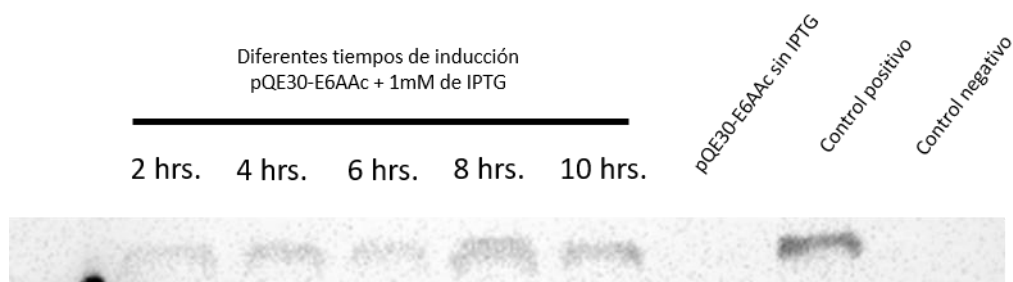


Figura 1. Tiempos de inducción. En la figura se muestran los resultados de western blot de diferentes tiempos de inducción con 1mM de IPTG utilizando la construcción pQE30-E6AAc. Como se puede observar se está produciendo la proteína E6.

Referencias generales

Bernard, H.-U., Burk, R.D., Chen, Z., van Doorslaer, K., Hausen, H. zur, de Villiers, E.-M., 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401, 70–79. doi:10.1016/j.virol.2010.02.002

Biegging, K.T., Mello, S.S., Attardi, L.D., 2014. Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* 14, 359–370. doi:10.1038/nrc3711

Cai, Q., Lv, L., Shao, Q., Li, X., Dian, A., 2013. Human papillomavirus early proteins and apoptosis. *Arch. Gynecol. Obstet.* 287, 541–548. doi:10.1007/s00404-012-2665-z

Chen, J., Ni, G., Liu, X.S., 2011. Papillomavirus virus like particle-based therapeutic vaccine against human papillomavirus infection related diseases: Immunological problems and future directions. *Cell. Immunol.* 269, 5–9. doi:10.1016/j.cellimm.2011.03.003

Choi, Y.J., Park, J.S., 2016. Clinical significance of human papillomavirus genotyping. *J. Gynecol. Oncol.* 27. doi:10.3802/jgo.2016.27.e21

Cornet, I., Gheit, T., Franceschi, S., Vignat, J., Burk, R.D., Sylla, B.S., Tommasino, M., Clifford, G.M., the IARC HPV Variant Study Group, 2012. Human Papillomavirus Type 16 Genetic Variants: Phylogeny and Classification Based on E6 and LCR. *J. Virol.* 86, 6855–6861. doi:10.1128/JVI.00483-12

de Villiers, E.-M., Fauquet, C., Broker, T.R., Bernard, H.-U., zur Hausen, H., 2004. Classification of papillomaviruses. *Virology* 324, 17–27. doi:10.1016/j.virol.2004.03.033

Doorbar, J., Egawa, N., Griffin, H., Kranjec, C., Murakami, I., 2015. Human papillomavirus molecular biology and disease association: Human papillomavirus. *Rev. Med. Virol.* 25, 2–23. doi:10.1002/rmv.1822

Doorbar, J., Quint, W., Banks, L., Bravo, I.G., Stoler, M., Broker, T.R., Stanley, M.A., 2012. The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine* 30, F55–F70. doi:10.1016/j.vaccine.2012.06.083

Duensing, S., Duensing, A., Crum, C.P., Münger, K., 2001. Human papillomavirus type 16 E7 oncoprotein-induced abnormal centrosome synthesis is an early event in the evolving malignant phenotype. *Cancer Res.* 61, 2356–2360.

Duensing, S., Lee, L.Y., Duensing, A., Basile, J., Piboonniyom, S., Gonzalez, S., Crum, C.P., Munger, K., 2000. The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proc. Natl. Acad. Sci. U. S. A.* 97, 10002–10007. doi:10.1073/pnas.170093297

Duensing, S., Münger, K., 2002. The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. *Cancer Res.* 62, 7075–7082.

G Capote Negrin, L., 2015. Epidemiology of cervical cancer in Latin America. *ecancermedicalscience* 9. doi:10.3332/ecancer.2015.577

Gheit, T., Cornet, I., Clifford, G.M., Iftner, T., Munk, C., Tommasino, M., Kjaer, S.K., 2011. Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 20, 1315–1321. doi:10.1158/1055-9965.EPI-10-1187

Hengstermann, A., Linares, L.K., Ciechanover, A., Whitaker, N.J., Scheffner, M., 2001. Complete switch from Mdm2 to human papillomavirus E6-mediated degradation of p53 in cervical cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 98, 1218–1223. doi:10.1073/pnas.031470698

Ho, L., Chan, S.Y., Burk, R.D., Das, B.C., Fujinaga, K., Icenogle, J.P., Kahn, T., Kiviat, N., Lancaster, W., Mavromara-Nazos, P., 1993. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J. Virol.* 67, 6413–6423.

Ho, L., Chan, S.Y., Chow, V., Chong, T., Tay, S.K., Villa, L.L., Bernard, H.U., 1991. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J. Clin. Microbiol.* 29, 1765–1772.

Howie, H.L., Katzenellenbogen, R.A., Galloway, D.A., 2009. Papillomavirus E6 proteins. *Virology* 384, 324–334. doi:10.1016/j.virol.2008.11.017

Hudelist, G., Manavi, M., Pischinger, K.I., Watkins-Riedel, T., Singer, C.F., Kubista, E., Czerwenka, K.F., 2004. Physical state and expression

of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol. Oncol.* 92, 873–880. doi:10.1016/j.ygyno.2003.11.035

Huertas-Salgado, A., Martín-Gámez, D.C., Moreno, P., Murillo, R., Bravo, M.M., Villa, L., Molano, M., 2011. E6 molecular variants of human papillomavirus (HPV) type 16: an updated and unified criterion for clustering and nomenclature. *Virology* 410, 201–215. doi:10.1016/j.virol.2010.10.039

Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P., Breitsprecher, D., 2014. MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* 1077, 101–113. doi:10.1016/j.molstruc.2014.03.009

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., et al., 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi:10.1038/35057062

Lazarczyk, M., Cassonnet, P., Pons, C., Jacob, Y., Favre, M., 2009. The EVER Proteins as a Natural Barrier against Papillomaviruses: a New Insight into the Pathogenesis of Human Papillomavirus Infections. *Microbiol. Mol. Biol. Rev.* 73, 348–370. doi:10.1128/MMBR.00033-08

Mantovani, F., Banks, L., 2001. The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene* 20, 7874–7887. doi:10.1038/sj.onc.1204869

Martinez-Zapien, D., Ruiz, F.X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Pol, S.V., Podjarny, A., Travé, G., Zanier, K., 2016. Structure of the E6/E6AP/p53 complex

required for HPV-mediated degradation of p53. *Nature* 529, 541–545.
doi:10.1038/nature16481

Muñoz, N., Castellsagué, X., de González, A.B., Gissmann, L., 2006. Chapter 1: HPV in the etiology of human cancer. *Vaccine* 24, S1–S10.
doi:10.1016/j.vaccine.2006.05.115

Nominé, Y., Ristriani, T., Laurent, C., Lefèvre, J.F., Weiss E, null, Travé, G., 2001a. Formation of soluble inclusion bodies by hpv e6 oncoprotein fused to maltose-binding protein. *Protein Expr. Purif.* 23, 22–32.
doi:10.1006/prep.2001.1451

Nominé, Y., Ristriani, T., Laurent, C., Lefèvre, J.F., Weiss E, null, Travé, G., 2001b. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng.* 14, 297–305.

Pereira, R., Hitzeroth, I.I., Rybicki, E.P., 2009. Insights into the role and function of L2, the minor capsid protein of papillomaviruses. *Arch. Virol.* 154, 187–197. doi:10.1007/s00705-009-0310-3

Planas-Iglesias, J., Bonet, J., García-García, J., Marín-López, M.A., Feliu, E., Oliva, B., 2013. Understanding Protein–Protein Interactions Using Local Structural Features. *J. Mol. Biol.* 425, 1210–1224.
doi:10.1016/j.jmb.2013.01.014

Raybould, R., Fiander, A., Hibbitts, S., 2011. Human Papillomavirus Integration and its Role in Cervical Malignant Progression. *Open Clin. Cancer J.* 5, 1–7. doi:10.2174/1874189401105010001

Riley, R.R., Duensing, S., Brake, T., Münger, K., Lambert, P.F., Arbeit, J.M., 2003. Dissection of human papillomavirus E6 and E7 function in

transgenic mouse models of cervical carcinogenesis. *Cancer Res.* 63, 4862–4871.

Ristriani, T., Nominé, Y., Masson, M., Weiss, É., Travé, G., 2001. Specific recognition of four-way DNA junctions by the C-terminal zinc-binding domain of HPV oncoprotein E6. *J. Mol. Biol.* 305, 729–739. doi:10.1006/jmbi.2000.4330

Rivoire, W.A., Corleta, H.V.E., Brum, I.S., Capp, E., 2006. Biologia molecular do câncer cervical. *Rev. Bras. Saúde Materno Infant.* 6, 447–451. doi:10.1590/S1519-38292006000400012

Seedorf, K., Krämmer, G., Dürst, M., Suhai, S., Röwekamp, W.G., 1985. Human papillomavirus type 16 DNA sequence. *Virology* 145, 181–185.

Stanley, M., 2010. Pathology and epidemiology of HPV infection in females. *Gynecol. Oncol.* 117, S5–S10. doi:10.1016/j.ygyno.2010.01.024

Tumbarello, D.A., Brown, M.C., Turner, C.E., 2002. The paxillin LD motifs. *FEBS Lett.* 513, 114–118.

Tuncbag, N., Kar, G., Keskin, O., Gursoy, A., Nussinov, R., 2008. A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. *Brief. Bioinform.* 10, 217–232. doi:10.1093/bib/bbp001

Wheeler, C.M., Yamada, T., Hildesheim, A., Jenison, S.A., 1997. Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. *J. Clin. Microbiol.* 35, 11–19.

Wiederschain, G.Y., 2006. Protein-protein interactions. A molecular cloning manual: 2nd Edition, Golemis, E. A., and Adams, P. D. (eds.)

Cold Spring Harbor Laboratory Press, New York, 938 p., 2005, Biochem. Mosc. 71, 697–697. doi:10.1134/S0006297906060162

Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D., Duhr, S., 2010. Protein-binding assays in biological liquids using microscale thermophoresis. Nat. Commun. 1, 100. doi:10.1038/ncomms1093

Woodman, C.B.J., Collins, S.I., Young, L.S., 2007. The natural history of cervical HPV infection: unresolved issues. Nat. Rev. Cancer 7, 11–22. doi:10.1038/nrc2050

Yamada, T., Manos, M.M., Peto, J., Greer, C.E., Munoz, N., Bosch, F.X., Wheeler, C.M., 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J. Virol. 71, 2463–2472.

Yamada, T., Wheeler, C.M., Halpern, A.L., Stewart, A.C., Hildesheim, A., Jenison, S.A., 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. J. Virol. 69, 7743–7753.

Yue, Y., 2013. Human papillomavirus oncoproteins and apoptosis (Review). Exp. Ther. Med. doi:10.3892/etm.2013.1374

Zanier, K., ould M'hamed ould Sidi, A., Boulade-Ladame, C., Rybin, V., Chappelle, A., Atkinson, A., Kieffer, B., Travé, G., 2012. Solution Structure Analysis of the HPV16 E6 Oncoprotein Reveals a Self-Association Mechanism Required for E6-Mediated Degradation of p53. Structure 20, 604–617. doi:10.1016/j.str.2012.02.001

Zheng, Z.-M., Baker, C.C., 2006. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front. Biosci. J. Virtual Libr.* 11, 2286–2302.

Zilfou, J.T., Lowe, S.W., 2009. Tumor Suppressive Functions of p53. *Cold Spring Harb. Perspect. Biol.* 1, a001883–a001883. doi:10.1101/cshperspect.a001883

Zuna, R.E., Moore, W.E., Shanesmith, R.P., Dunn, S.T., Wang, S.S., Schiffman, M., Blakey, G.L., Teel, T., 2009. Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. *Int. J. Cancer* 125, 2609–2613. doi:10.1002/ijc.24706

zur Hausen, H., 2009. Papillomaviruses in the causation of human cancers — a brief historical account. *Virology* 384, 260–265. doi:10.1016/j.virol.2008.11.046

Martinez-Zapien, D., Ruiz, F.X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Pol, S.V., Podjarny, A., Travé, G., Zanier, K. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* 529, 541–545. doi:10.1038/nature16481

Zanier, K., Charbonnier, S., Sidi, A. O. M. O., McEwen, A. G., Ferrario, M. G., Poussin, P., ... Travé, G. (2013). Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science (New York, N.Y.)*, 339(6120), 694–698. <http://doi.org/10.1126/science.1229934>

G. G. Krivov, M. V. Shapovalov, and R. L. Dunbrack, Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* (2009).

Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38.
<http://www.ks.uiuc.edu/Research/vmd/>

UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem*. 2004 Oct;25(13):1605-12.
<http://www.rbvi.ucsf.edu/chimera>

James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26:1781-1802, 2005.