



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**Análisis computacional de la mutación
c.1349_1352delAGTA en el gen fumarato
hidratasa causante de HLRCC: implicaciones
en la estructura proteica y en el metabolismo
celular**

Andrea Yimena Cardona Barreto

Universidad Nacional de Colombia
Facultad de Medicina, Maestría en Genética Humana
Bogotá, Colombia
2018

Análisis computacional de la mutación c.1349_1352delAGTA en el gen fumarato hidratasa causante de HLRCC: implicaciones en la estructura proteica y en el metabolismo celular

Andrea Yimena Cardona Barreto

Tesis presentada como requisito parcial para optar por el título de:

Magister en Genética Humana

Director:

Ph.D., Andrés Mauricio Pinzón Velasco

Codirectora:

MSc. Clara Eugenia Arteaga Díaz

Línea de Investigación:

Biología de sistemas

Grupo de Investigación:

Grupo en Bioinformática y Biología de sistemas (GiBBS)

Universidad Nacional de Colombia

Facultad de Medicina, Maestría en genética Humana

Bogotá, Colombia

2018

A mi mamá quien lucho toda su vida por mi bienestar y a toda mi familia por su apoyo incondicional en todos los momentos de mi vida.

A mi pequeña hija que llevo con su alegría a llenar mi vida de colores y alegría.

Agradecimientos

Agradezco en primer lugar a mi mamá quien a pesar de sus grandes obstáculos dedicó su vida a trabajar para darme una vida hermosa llena de alegrías y aunque ya no está sigue haciendo su trabajo en mi corazón con la semilla que sembró en él.

Agradezco a mi familia porque ayudó a mi mamá en ese proceso y porque creyeron y siguen creyendo en mí. A mis abuelitos Clariza y Antonio, mis tíos Oto y Nana, mi prima María Paula y a mi papa Fernando, mil gracias.

A mi tía nana, mi prima María Paula y a José Elías Moreno por recibirme en su casa, brindarme todas las comodidades, apoyo y amor incondicional para poder realizar esta maestría.

A mi hija Antonia Valencia Cardona porque con su llegada mitigo un poco el dolor de tantas partidas familiares e ilumino la casa con su hermosa sonrisa. Gracias a ti obtuve la fuerza necesaria para terminar este largo y duro proceso.

A la Universidad Nacional de Colombia y a los profesores de la maestría en Genética Humana en especial a Clara Arteaga por su gran amabilidad y preocupación por mi bienestar. Al profesor Mauricio Rey por todos los conocimientos transmitidos y por demostrar su amor hacia su labor y contribuir en mi formación como profesional.

A mi director de tesis el profesor Andrés Mauricio Pinzón Velasco quien confió en mí y me dió la oportunidad de trabajar con él. Le agradezco enormemente por su gran colaboración y su paciencia. Además, por escucharme y aconsejarme en los momentos más difíciles.

A Daniel Camilo Osorio Hurtado quien me ayudó incondicionalmente y con mucha paciencia en el desarrollo de este trabajo.

A Janneth González Santos Doctora en Ciencias Biológicas y al laboratorio en Bioquímica Computacional y Estructural de la Pontificia Universidad Javeriana al que ella hace parte, por permitirme realizar allí la primera fase del modelamiento computacional y docking molecular.

Al Biólogo y estudiante de la maestría en Ciencias Biomédicas de la Universidad del Quindío Johan Villada por su colaboración con los análisis de docking molecular y su disposición para colaborar.

A Oscar Navarro por estar ahí incondicionalmente, por mostrarme lo bonita que puede ser la vida, por darme palabras de aliento en todo momento y por enseñarme a ver la realidad con otros ojos.

A mis compañeros de maestría Ángela Muñoz, Carolina Valencia, Daniel Verbel y Diana Guerrero por su compañía, solidaridad y complicidad en este camino. Mil gracias por todos los momentos gratos, tramos y risas.

A mis amigos incondicionales Natalia Ospina, Andrés Felipe Castro, Rodrigo Restrepo, Mónica Moncada, Cristian Calderón, Lorena García y Diana Fonseca por su acompañamiento y ayuda en todo este proceso y por tratar de alentarme en los momentos de tristeza.

Resumen

Introducción: el gen FH codifica la enzima fumarato hidratasa y mutaciones en éste provocan el síndrome de leiomiomatosis hereditaria y cáncer de células renales (HLRCC).

Objetivo: predecir computacionalmente los posibles cambios estructurales de la proteína fumarato hidratasa relacionados con la mutación c.1349_1352delAGTA en el gen FH causante del síndrome HLRCC, así como la posible implicación de estos cambios en el metabolismo celular.

Metodología: se realizaron análisis estructurales de la proteína fumarato hidratasa con la mutación c.1349_1352delAGTA. Posteriormente, se llevó a cabo una simulación de interacción molecular con sus sustratos, mediante el programa AutoDock Vina. Finalmente, tomando como base la reconstrucción metabólica de humano: Recon 2 y mediante el paquete de análisis COBRA Toolbox, se modelaron computacionalmente las posibles implicaciones, a nivel metabólico, de distintos escenarios de actividad enzimática de la fumarato hidratasa.

Resultados: De acuerdo con los resultados obtenidos mediante la simulación de interacción molecular, en donde se obtuvieron unas energías de -4.5 y -4.4 kcal/mol para la proteína silvestre y mutada con el fumarato respectivamente y de -4.8 y -4.4 kcal/mol para la proteína silvestre y mutada con el malato respectivamente, no se encontraron diferencias significativas entre estas energías. Nuestros resultados sugieren que la mutación c.1349_1352delAGTA en el gen FH podría tener su efecto entonces, al imposibilitar la oligomerización de la proteína durante el establecimiento de puentes de hidrogeno, y de esta manera se inhibiría la formación del sitio activo.

En cuanto a las implicaciones a nivel metabólico de esta posible inhibición de formación del sitio activo, nuestros resultados sugieren un aumento en la reacción de biomasa, lo cual podría sugerir un mecanismo relacionado con cáncer y crecimiento tumoral. Además, se encontró una baja producción del fosfolípido cardiolipina lo que genera alta producción de especies reactivas de oxígeno y se encontró un aumento en el flujo desde el citoplasma hacia la mitocondria de la enzima tioredoxina reductasa, implicada en la disminución en la actividad de la cascada de señalización de apoptosis.

Palabras clave: Fumarato hidratasa, HLRCC, Docking molecular, modelo computacional, tioredoxina, cardiolipina.

Abstract

Introduction: FH gene encodes the fumarate hydratase enzyme. Mutations in this gene can cause the hereditary leiomyomatosis and renal cell cancer (HLRCC) syndrome.

Objective: to predict computationally, the possible structural changes of the fumarate hydratase protein related to the mutation c.1349_1352delAGTA in the FH gene that causes the HLRCC syndrome, as well as the possible implication of these changes in the cellular metabolism.

Methodology: structural analyzes of the fumarate hydratase protein containing the mutation c.1349_1352delAGTA were made. Subsequently, molecular interaction simulation of the protein together with its substrates was carried out by using AutoDock Vina program. Finally, based on the metabolic human reconstruction: Recon 2 and through the COBRA Toolbox analysis package, the possible implications at the metabolic level of different scenarios of enzymatic activity of fumarate hydratase were computationally modeled.

Results: According to the results obtained by the simulation of molecular interaction, were found energies values of -4.5 and -4.4 obtained for the wild and mutated protein with fumarate respectively and -4.8 and -4.4 for the wild protein and mutated with the malate respectively. No significant differences were found between the energy values obtained in none of the cases. In this way, our results suggest that the mutation c.1349_1352delAGTA in the FH gene could have its effect then, by making oligomerization of the protein impossible during the establishment of hydrogen bonds, and in this way the formation of the active site would be inhibited.

Regarding the implications at the metabolic level of the possible inhibition of the active site formation, our results suggest an increase in the biomass reaction, which could suggest a mechanism related to cancer and tumor growth. In addition, a low production of the cardiolipin enzyme was found, which generates high production of reactive oxygen species; furthermore, it was found an increase in the flow of the reduce thioredoxin

enzyme from the cytoplasm to the mitochondria, implied in the decrease in the apoptosis signaling cascade.

Key words: Fumarate hydratase, HLRCC, molecular docking, computational model, cardiolipin and thioredoxin.

Contenido

Resumen	VI
Abstract.....	VII
Contenido.....	IX
Lista de figuras.....	XI
Lista de tablas	XII
Introducción.....	1
1. Marco Teórico.....	4
1.1 Generalidades Síndrome HLRCC.....	4
1.2 Aspectos Clínicos del síndrome HLRCC	4
1.2.1 Leiomiomas cutáneos	4
1.2.2 Leiomiomas uterinos.....	4
1.2.3 Cáncer renal papilar tipo II.....	5
1.3 Aspectos genéticos y moleculares del síndrome HLRCC.....	6
1.3.1 Gen FH.....	6
1.3.2 Proteína fumarato hidratasa.....	6
1.4 Hipótesis acerca de la fisiopatología del síndrome HLRCC.....	9
1.4.1 Aumento en concentraciones de HIF 1-a.....	9
1.4.2 Succinización aberrante	10
1.4.3 Ausencia de reparación de DNA.....	11
1.5 Mutación c.1349_1352delAGTA en el gen FH causante de HLRCC.....	12
1.6 Herramientas computacionales	13
1.6.1 Modelamiento tridimensional de proteínas	14
1.6.2 Docking molecular.....	16
1.6.3 Importancia de las herramientas computacionales en el estudio de la enfermedad HLRCC	17
2. Planteamiento del Problema.....	27
3. Justificación.....	29
4. Objetivos	31
4.1 Objetivo general	31
4.2 Objetivos específico	31

5. Metodología	32
5.1 Modelar y caracterizar las variantes estructurales de la proteína fumarato hidratasa a partir de la mutación c.1349_1352delAGTA en el gen FH	32
5.2 Describir el posible efecto de las variantes halladas de la proteína fumarato hidratasa en la interacción con los ligandos fumarato y malato	35
5.3 Analizar las implicaciones a nivel metabólico de los cambios en la interacción de la proteína fumarato hidratasa mutada con los ligandos fumarato y malato, en un modelo computacional de metabolismo humano	36
6. Resultados y discusión.....	39
6.1 Predecir y caracterizar las variantes estructurales de la proteína fumarato hidratasa a partir de la mutación c.1349_1352delAGTA en el gen FH	39
6.1.1 Estructura cuaternaria de la proteína fumarato hidratasa	44
6.1.2 Propiedades físico-químicas de la proteína fumarato hidratasa silvestre y con la presencia de diferentes mutaciones	50
6.2 Describir el posible efecto de las variantes halladas de la proteína fumarato hidratasa en la interacción con los ligandos fumarato y malato	53
6.3 Analizar las implicaciones de la interacción de la proteína fumarato hidratasa mutada con los ligandos fumarato y malato, en un modelo computacional de metabolismo humano	57
6.3.1 Análisis FBA	57
6.3.2 Análisis FVA	60
7. Conclusiones y recomendaciones	67
7.1 Conclusiones.....	67
7.2 Recomendaciones	68
8. Consideraciones éticas.....	69
9. Bibliografía.....	70

Lista de figuras

<i>Figura 1: Características fenotípicas de HLRCC.</i>	5
<i>Figura 2: Estructura terciaria y cuaternaria de la fumarato hidratasa humana.</i>	7
<i>Figura 3: La enzima fumarato hidratasa cataliza la hidratación de un doble enlace del fumarato para formar malato.</i>	8
<i>Figura 4: Rol del fumarato y otros metabolitos del ciclo del ácido cítrico en la regulación del factor inducible de hipoxia (HIF).</i>	10
<i>Figura 5: Posibles consecuencias de la pérdida de la función de FH en HLRCC (violeta) y el mecanismo de sobre expresión de los genes blanco de Nrf2 (rojo).</i>	11
<i>Figura 6: Pedigree de familia colombiana con integrantes portadores de la mutación c.1349_1352delAGTA en el gen FH causante del síndrome HLRCC.</i>	13
<i>Figura 7: Descripción general del community-driven reconstruction approach para ensamblar Recon 2.</i>	19
<i>Figura 8: Reconstrucción de un modelo metabólico a escala genómica.</i>	21
<i>Figura 9: Base conceptual del modelamiento basado en restricciones y FBA.</i>	24
<i>Figura 10: Alineamiento entre la secuencia de aminoácidos de la proteína fumarato hidratasa silvestre y mutada.</i>	40
<i>Figura 11: Monómero de la proteína Fumarato hidratasa.</i>	41
<i>Figura 12: Dímero que contiene las subunidades B (azul) y D (amarilla).</i>	42
<i>Figura 13: Proteína fumarato hidratasa humana modelada computacionalmente.</i>	45
<i>Figura 14: Plot de Ramachandran que evidencia la conformación estereoquímica de los ángulos phi y psi de la estructura tridimensional de la proteína fumarato hidratasa con la mutación c.1349_1352delATGA.</i>	46
<i>Figura 15: Alineamiento estructural de la proteína fumarato hidratasa humana silvestre y la proteína fumarato hidratasa humana con la mutación c.1349_1352delATGA.</i>	47
<i>Figura 16: Homotetrámeros de la proteína fumarato hidratasa con cada una de las subunidades (Subunidad A: verde, subunidad B: azul, subunidad C: purpura y subunidad D: amarillo).</i>	49
<i>Figura 17: Representación de las hendiduras tanto en la proteína silvestre como mutada.</i>	53
<i>Figura 18: Docking molecular entre la proteína fumarato hidratasa (FH) silvestre y mutada y el ligando malato.</i>	55
<i>Figura 19: Docking molecular entre la proteína fumarato hidratasa (FH) silvestre y mutada y el ligando fumarato.</i>	56
<i>Figura 20: Efecto del aumento del metabolito fumarato por una inactivación de su gen (FH) actuando como un inhibidor del 2-oxoglutarato.</i>	60
<i>Figura 21: Reacción por la cual es sintetizada la cardiolipina en el cuerpo humano.</i>	63
<i>Figura 22: Las células tumorales incorporan menos ácidos grasos poliinsaturados,</i>	65

Lista de tablas

<i>Tabla 1: Enlaces entre la diferentes subunidades tanto de la proteína fumarato silvestre como de la mutada.....</i>	<i>43</i>
<i>Tabla 2: SASA de los aminoácidos que hacen parte del sitio activo de la proteína Fumarato Hidratasa humana silvestre y de la proteína con la mutación c.1349_1352delAGTA.</i>	<i>48</i>
<i>Tabla 3 “Variantes en la proteína fumarato hidratasa humana causantes de la enfermedad HLRCC reportadas en uniProt”.</i>	<i>50</i>
<i>Tabla 4 “Energía de unión y constante de inhibición para los docking entre la proteína fumarato hidratasa silvestre y mutada y los ligandos fumarato y malato”.</i>	<i>54</i>
<i>Tabla 5: Resultados del análisis FBA en la célula control comparada con los otros tres experimentos de la célula con la enzima Fumarato Hidratasa mutada.....</i>	<i>58</i>
<i>Tabla 6: Lower Bound y Upper Bound de reacciones que presentan una diferencia significativa entre el control y los ensayos con las células que tienen deficiencia en la actividad de la enzima Fumarato Hidratasa.....</i>	<i>62</i>

Introducción

El síndrome de leiomiomatosis hereditaria y cáncer de células renales (HLRCC) es un desorden hereditario autosómico dominante caracterizado por una predisposición a desarrollar leiomiomas cutáneos, uterinos y cáncer de células renales. El gen asociado con esta enfermedad es el gen FH, gen que codifica para la enzima fumarato hidratasa la cual cataliza la reacción reversible hidratación/deshidratación de fumarato a L-malato y codifica para dos formas de la enzima que son diferencialmente procesadas (forma mitocondrial y citosólica), pero idénticas en la secuencia. En la mitocondria, la enzima fumarato hidratasa está involucrada en el ciclo del ácido tri-carboxílico o ciclo de Krebs, mientras que la isoforma citoplasmática está involucrada en el metabolismo del fumarato, metabolismo de aminoácidos y en la reparación de rompimiento de doble hebra del DNA (DSBs).

Para el gen FH, a nivel mundial se han reportado alrededor de 50 mutaciones que causan HLRCC, mientras que en Colombia y en Latino América solo ha sido reportada una familia con esta enfermedad (1). En la única familia colombiana conocida con la enfermedad, se reportó la mutación c.1349_1352delAGTA en el año 2016 por Arenas y colaboradores, mutación que no había sido reportada antes para la ciencia. Esta mutación provoca un desplazamiento en el marco de lectura o frameshift en el cual se genera un codón de parada prematuro, lo cual lleva a que solo sean codificados 407 aminoácidos de los 466 que conforman la proteína. En otras palabras, la mutación genera una pérdida de 59 aminoácidos en cada una de las cuatro subunidades que conforman el homotetrámero de la enzima fumarato hidratasa y no se sabe cuáles son las consecuencias estructurales y biológicas que esta mutación genera en la proteína.

La familia que presenta la mutación mencionada, tiene unas características específicas que se basan en alta cantidad de pacientes con presencia de cáncer renal y quistes renales, pocos pacientes con leiomiomas uterinos y ningún paciente con leiomiomas cutáneos, el cual normalmente es el síntoma más común de esta enfermedad.

En cuanto a las causas por las cuales se desarrolla el cáncer renal y los tumores benignos en la piel y en el útero a causa del síndrome HLRCC, durante los últimos años se han desarrollado varias hipótesis, teniendo cada una de ellas argumentos a favor y

argumentos en contra. Sin embargo, en ninguna de estas hipótesis se ha evaluado el posible impacto que tendría en el metabolismo celular la ausencia funcional total o los posibles cambios de afinidad por los sustratos de la enzima fumarato hidratasa, evaluando estas implicaciones desde una perspectiva de biología de sistemas.

Cabe resaltar que los experimentos en laboratorio húmedo son muy valiosos para la investigación científica y que generan grandes conocimientos acerca de las vías metabólicas o de los diferentes procesos que se ven afectados en ciertas situaciones. Sin embargo, en muchas ocasiones realizar este tipo de experimentos es muy difícil ya sea por la dificultad en la obtención de las muestras de los pacientes y/o por los altos costos que estos análisis generan. Por esta razón, es altamente deseable y conveniente implementar técnicas que nos permitan sobrepasar dichas limitaciones y de esta manera realizar acercamientos a los procesos biológicos, bioquímicos, entre otros para entender y visualizar procesos que son difíciles de estudiar debido a las condiciones nombradas. Una de estas técnicas proviene de la biología de sistemas computacional, la cual nos permite realizar predicciones de un sistema biológico dado, mediante la creación de un modelo que se aproxime, lo mejor posible, al fenómeno bajo estudio.

Es importante resaltar que los análisis computacionales tienen una gran ventaja con respecto a los análisis en laboratorio húmedo, y es que permiten tener una mirada más global de las situaciones que ocurren dentro de una célula. Por ejemplo, bajo condiciones específicas o permite analizar complejas cascadas de señalización, cosa que no es posible en el momento de realizar experimentos húmedos ya que estos son más limitados y se requiere de una enorme inversión de tiempo para poder tener una mirada medianamente global. Sin embargo, los análisis computacionales son de carácter predictivo y por lo tanto hipotético, así que finalmente los experimentos en húmedo son fundamentales y tienen un carácter confirmatorio definitivo ya que sin ellos no sería posible entender muchos procesos biológicos. De esta manera, una aproximación ideal es la de acoplar las predicciones computacionales, lo que permite dirigir y enfocar la investigación húmeda, con los análisis de laboratorio convencional que son, como lo mencionamos anteriormente, el sistema de validación de las predicciones y su carácter definitivo.

Por lo anterior, a pesar de que este trabajo se circunscribe exclusivamente al área computacional, mediante el modelado de las posibles implicaciones de la mutación

c.1349_1352delAGTA en el gen FH sobre la estructura de la proteína fumarato hidratasa buscando generar posibles implicaciones de esta mutación en el metabolismo celular, es importante resaltar su importancia y conveniencia, ya que mediante este se han generado hipótesis de trabajo plausible, con alta probabilidad de éxito, que posteriormente pueden ser validadas mediante experimentos en el laboratorio húmedo.

1. Marco Teórico

1.1 Generalidades Síndrome HLRCC

El cáncer renal afectó a más de 403262 individuos globalmente y resultó en al menos 175098 muertes en el año 2018 (2). El 97% de esta patología se presenta en individuos que no tienen una historia familiar de cáncer renal (3). Sin embargo, el 3% restante de los cánceres renales son causados por mutaciones de línea germinal en algunos genes, mutaciones que son heredadas entre los familiares. Estos tipos de cáncer se conocen como cánceres renales hereditarios.

El síndrome de leiomiomatosis hereditaria y cáncer renal hereditario (HLRCC, OMIM 150800) fue descrito por primera vez en el año 2001 por Launonen y colaboradores (4), este es un desorden hereditario autosómico dominante, caracterizado por una predisposición a desarrollar leiomiomas cutáneos y uterinos, cáncer de células renales y leiomiosarcomas uterinos (4).

1.2 Aspectos Clínicos del síndrome HLRCC

1.2.1 Leiomiomas cutáneos

Se ha detectado que los leiomiomas cutáneos (tumores benignos del músculo liso) asociados a HLRCC se presentan en el 75% a 100% de los pacientes afectados (5). Los leiomiomas cutáneos son firmes y de color rojizo, tienen un tamaño entre 0.4 y 2.5 mm (3) y un paciente puede tener entre 0 y más de 100 lesiones (6), las cuales pueden ser dolorosas (5). Estas lesiones se presentan en el tronco y en las extremidades de los pacientes e histológicamente se componen de haces de células de músculo liso con núcleos de extremos romos localizados centralmente (3). La edad media de desarrollo de estas lesiones en la piel es aproximadamente a las 25 años de edad (7) (Figura 1A).

1.2.2 Leiomiomas uterinos

Las mujeres afectadas con HLRCC están en riesgo de desarrollar múltiples fibroides uterinos a edad temprana (28-32 años) (3), estos son detectados en más del 90% de las

mismas (5,6). Además, un alto porcentaje de mujeres afectadas con leiomiomas cutáneos pueden también desarrollar leiomiomas uterinos (6,7). En una paciente se pueden presentar más de 20 leiomiomas uterinos que pueden medir entre 1.5 a 10 cm y los síntomas pueden incluir menstruaciones irregulares, menorragia y dolor (Figura 1B). Histológicamente son caracterizados por células con un único núcleo grande o células multinucleadas con núcleos prominentes orangiofílicos rodeados por un halo peri nuclear.

1.2.3 Cáncer renal papilar tipo II

Los pacientes afectados con HLRCC están en riesgo de desarrollar tempranamente quistes renales bilaterales y/o multifocales y cáncer de células renales papilar tipo II (PRCC) (5,6). A pesar de que HLRCC es una condición dominante, la penetrancia de los tumores renales es incompleta y relativamente baja (14% a 18% de los pacientes con HLRCC presentan tumores renales) (7), razón por la cual muchos casos de HLRCC no son diagnosticados clínicamente (8). Esta patología se puede presentar a una edad muy temprana que va desde los 10 a los 44 años de edad y los tumores se pueden presentar como lesiones solitarias, bilaterales o multifocales (Figura 1C). Estos tumores renales pueden ser muy agresivos y pueden hacer metástasis incluso cuando son muy pequeños (3). Histológicamente el PRCC asociado a HLRCC se caracteriza por citoplasma abundante anfifílico y grandes núcleos con grandes inclusiones eosinofílicas rodeados por un halo claro (3).

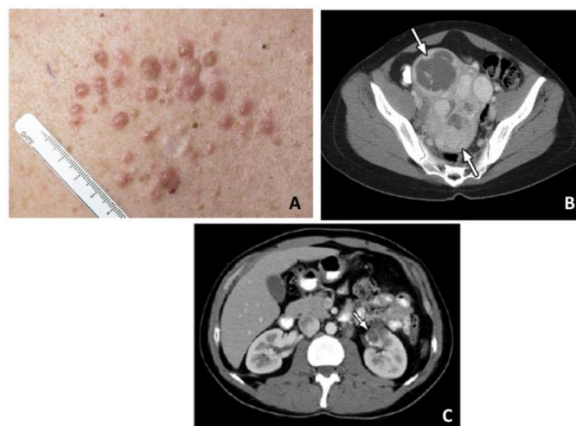


Figura 1: Características fenotípicas de HLRCC.

A. múltiples leiomiomas cutáneos en el torso de un paciente con HLRCC. **B.** imagen de tomografía computada mostrando imágenes de múltiples y grandes leiomiomas uterinos (flechas blancas) en una mujer con HLRCC. **C.** tumor renal izquierdo (flecha blanca) en un paciente con HLRCC” (9).

1.3 Aspectos genéticos y moleculares del síndrome HLRCC

1.3.1 Gen FH

El gen asociado con esta enfermedad es el gen FH (10), gen nuclear que codifica para la enzima fumarato hidratasa tanto citosólica como mitocondrial (EC 4.2.1.2) (11), la cual cataliza la reacción de hidratación de fumarato a L-malato en la mitocondria y la deshidratación de L-malato a fumarato en el citoplasma. Este gen se encuentra localizado en el brazo largo del cromosoma 1 (1q42.1), tiene 10 exones y codifica para dos formas que son diferencialmente procesadas (forma mitocondrial y citosólica), pero idénticas en la secuencia (12). Se ha demostrado que el promotor del gen FH humano contiene dos sitios de inicio de la transcripción a partir de los cuales se transcriben dos tipos de ARNm. El primero es un transcrito que es traducido en una proteína que contiene una secuencia de reconocimiento (47 aminoácidos) por parte de la mitocondria, mientras que el segundo transcrito no tiene dicha secuencia la cual es la enzima citoplasmática (13,14).

1.3.2 Proteína fumarato hidratasa

Cada una de las subunidades que conforma el homotetrámero de la proteína fumarato hidratasa posee 467 aminoácidos y su masa molecular es de 54637 Dalton (Da). La proteína fumarato hidratasa pertenece a la familia de las liasas, específicamente a las hidroliasas encargadas de romper enlaces carbono-oxígeno (15) y es un homotetrámero compuesto de cuatro subunidades idénticas de 50 kilodalton (kDa) cada una. Estructuralmente, cada subunidad de la proteína fumarato hidratasa humana tiene 23 α -helices y ocho hojas β y cada subunidad está construida por tres dominios llamados D1, D2 y D3 (16) (figura 2A). El dominio central, D2, facilita la tetramerización de esta proteína (figura 2B) (17). El dominio 1 está conformado por los residuos 1 a 41, el dominio 2 por los residuos 42 al 415 y el dominio 3 por los residuos 416 al 467 (18).

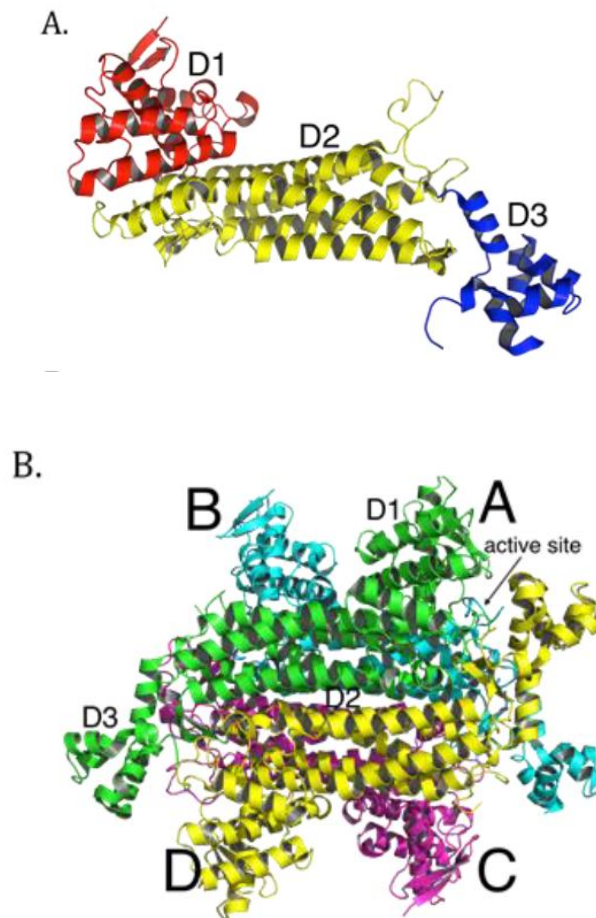


Figura 2: Estructura terciaria y cuaternaria de la fumarato hidratasa humana.

A. Subunidad de la fumarato hidratasa humana que está coloreada por dominios: dominio 1 (D1) rojo, dominio 2 (D2) amarillo y dominio 3 (D3) azul. **B.** Homotetrámero de la proteína fumarato hidratasa humana con cada una de las cuatro subunidades que la conforman marcadas como A, B, C y D. Todo el dominio D2 forma la mayor interfaz entre subunidades después de la oligomerización (17).

La fumarato hidratasa es una enzima encargada de catalizar la adición reversible de una molécula de agua a la unión olefínica del fumarato para formar malato (19). En esta reacción en la dirección de la formación de fumarato, se realiza una eliminación de un hidrogeno como protón de C_3 y un hidroxilo del L-malato para generar la unión olefínica del fumarato (figura 3). Estas dos reacciones ocurren a través de un carbanión intermediario en el cual C_3 es trigonal y C_2 tetraedral. Este aci-carboxilato intermedio tiene una carga negativa localizada principalmente en el grupo carboxilo de C_4 (20) (figura 3).

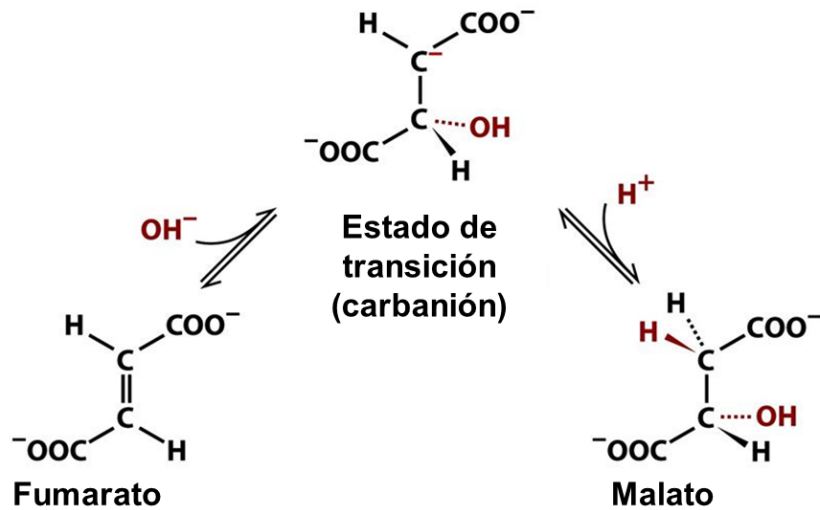


Figura 3: La enzima fumarato hidratasa cataliza la hidratación de un doble enlace del fumarato para formar malato. La reacción de hidratación procede a través de un estado de transición de carbanión. Durante este proceso ocurre primero la adición de un grupo OH^- y posteriormente ocurre la adición de H^+ . Modificado de (21).

La proteína fumarato hidratasa tiene dos sitios de unión al sustrato, el sitio catalítico A y el sitio no catalítico B. En cada una de las proteínas se encuentran cuatro sitios activos independientes (sitio A), los cuales están localizados entre los dominios D1 y D3 en las cuatro esquinas de la enzima. La unión de los dominios D1 y D3 hacen el contorno del punto de entrada a los cuatro sitios activos independientes de la fumarasa (figura 2B) (17). Las regiones que son altamente conservadas en un monómero se pueden observar claramente separadas por largas distancias y de esta manera no pueden constituir un sitio activo como tal (20). Por esta razón, el sitio activo está conformado por tres regiones que aportan tres de las cuatro subunidades y si observamos de forma tridimensional el tetrámero, cada una de estas regiones se unen para formar los sitios catalíticos (figura 2B) (17,20). Las regiones que conforman ese sitio activo son los residuos 133-150, 185-204 y 316-339, la primera y la segunda región construyen la mayoría del sitio de unión al sustrato, mientras que la segunda y tercera región donan los grupos catalíticos H192 (ácido catalítico) y S322 (base catalítica) (17,20,22).

En la mitocondria, la enzima fumarato hidratasa está involucrada en el ciclo del ácido tricarbóxico o ciclo de Krebs (15). Por otra parte, la isoforma citoplasmática está involucrada en el metabolismo del fumarato, el cual es producido en el citosol por

algunas reacciones (13) como el metabolismo de aminoácidos (23) y también está implicado en la reparación de rompimiento de doble hebra del DNA (DSBs) (24).

En los pacientes con HLRCC ocurre una inactivación de un alelo FH que es usualmente heredado de un padre portador de la mutación. La copia funcional restante se pierde o se inactiva a lo largo de la vida llevando a la formación de leiomiomas, quistes o tumores malignos, lo que sugiere que la inactivación completa de FH es la causa de las complicaciones relacionadas con HLRCC (25). El mecanismo por el cual la inactivación de FH conlleva a la formación de tumores renales en pacientes con HLRCC no ha sido bien establecido hasta el momento. Sin embargo, hay tres fuertes hipótesis que no están del todo comprobadas.

1.4 Hipótesis acerca de la fisiopatología del síndrome HLRCC

1.4.1 Aumento en concentraciones de HIF 1-A

La teoría más estudiada propone que la inactivación de FH causa la acumulación de fumarato en la célula aproximadamente 200 veces más comparada con una célula con la enzima fumarato hidratasa activa (26). El fumarato es un inhibidor competitivo de las HIF-prolil hidroxilasas, las cuales en condiciones normales promueven la degradación proteosomal del factor inducible de hipoxia (HIF), el cual es un factor de transcripción que incrementa la expresión de genes que regulan la angiogénesis tales como el factor de crecimiento epidermal vascular (VEGF). Cuando la fumarato hidratasa está inactiva, el fumarato inhibe la HIF-prolil hidroxilasas, evitando la degradación de HIF, el cual al mismo tiempo promueve la transcripción de genes involucrados en supervivencia y proliferación celular (Figura 4) (24,26).

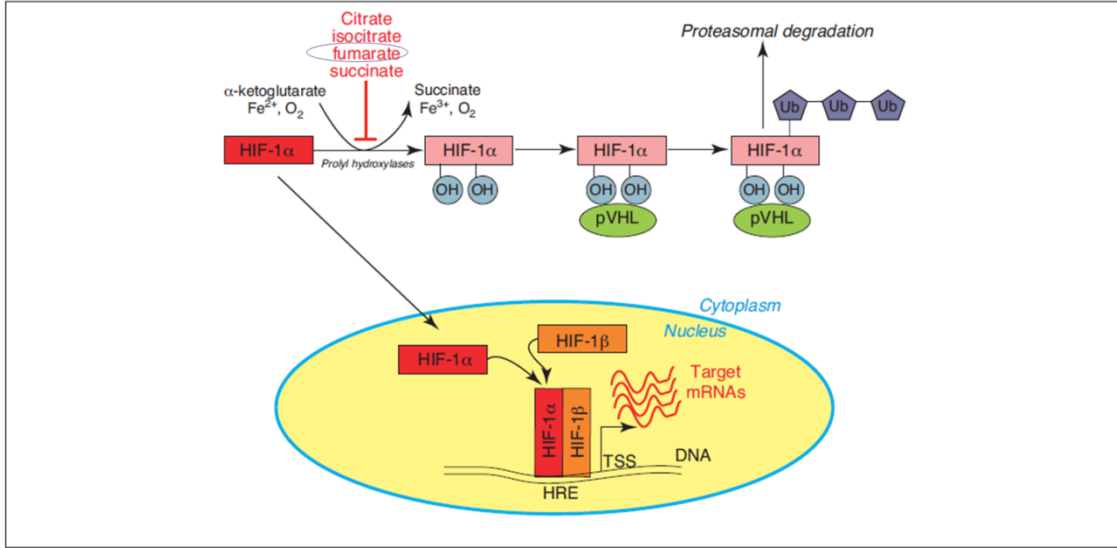


Figura 4: Rol del fumarato y otros metabolitos del ciclo del ácido cítrico en la regulación del factor inducible de hipoxia (HIF).

La acumulación del fumarato inhibe la acción de las prolinas hidroxilasas, lo que evita la degradación de HIF y promueve la translocación de este al núcleo para actuar como factor de transcripción de sus genes blanco” (27).

1.4.2 Succinización aberrante

La segunda hipótesis parte de reconocer a Keap1 como un represor citosólico de la vía Nrf2. Bajo condiciones normales, Keap1 promueve la degradación proteosómica de Nrf2 por interacción con una ubiquitin ligasa. Sin embargo, en momentos de estrés celular (acumulación de especies reactivas de oxígeno), Keap1 funciona como un sensor, lo que permite un cambio conformacional del mismo. Esto impide que se una a Nrf2 e impide que este último sea degradado (28), de esta manera Nrf2 puede cumplir su función de inducir la expresión de genes citoprotectores en respuesta a estrés oxidativo (29). Sin embargo, cuando la proteína FH no está activa se incrementa el fumarato como se dijo anteriormente. El fumarato es un compuesto tiol-reactivo que es capaz de reaccionar con los residuos de cisteína para producir 2-succinil-cisteína en condiciones fisiológicas, proceso llamado succinización. Se sabe que este proceso modifica la actividad de una gran cantidad de proteínas como Keap1, la cual tiene muchos residuos de cisteína expuestos. Tras la modificación covalente de los residuos de cisteína expuestos de Keap1 hay un cambio conformacional de la proteína que bloquea su interacción con Nrf2, razón por la cual la vía Nrf2 es activada constantemente sin existir señales de estrés. Nrf2 es un factor de transcripción que reconoce elementos de respuesta antioxidante (ARE) y sus genes blanco están involucrados en el metabolismo xenobiótico, síntesis de

moléculas antioxidantes y genes involucrados en la vía pentosa fosfato (28). En algunas investigaciones se ha determinado que los genes blanco de transcripción de Nrf2 son coordinadamente sobre expresados en tejidos con la enzima fumarato hidratasa inactiva (25) (Figura 5).

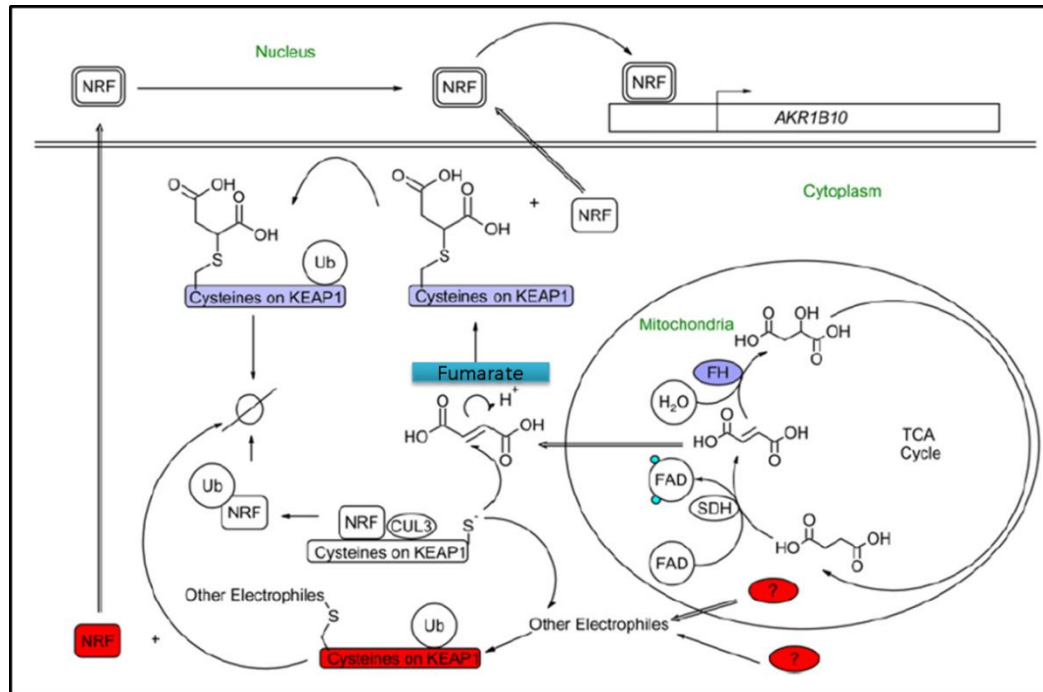


Figura 5: Posibles consecuencias de la pérdida de la función de FH en HLRCC (violeta) y el mecanismo de sobre expresión de los genes blanco de Nrf2 (rojo).

El fumarato que se acumula por la mutación en FH es translocado al citosol, donde reacciona con las cisteínas de Keap1 y altera la conformación de Keap1 liberando a Nrf2 del citoesqueleto. Nrf2 libre puede ser translocado al núcleo, donde se puede unir a los elementos ARE y dirige la expresión de sus genes blanco” (30).

1.4.3 Ausencia de reparación de DNA

La tercera hipótesis, la cual fue planteada en el año 2010, propone que bajo condiciones normales la fumarato hidratasa está localizada tanto en la mitocondria en la que participa en el ciclo de Krebs, como en el citoplasma. La fumarato hidratasa que se encuentra en el citoplasma es reclutada en el núcleo cuando ocurre un rompimiento de doble hebra del DNA. En el núcleo, como consecuencia de la formación de rompimientos de doble hebra del DNA ocurre la fosforilación del aminoácido Thr236 de la enzima fumarato hidratasa por la proteína quinasa dependiente de DNA (DNA-PK). Esta fosforilación es necesaria para el reclutamiento de la fumarato hidratasa en el núcleo. Posteriormente, el fumarato

producido por la proteína fumarato hidratasa a partir de malato, inhibe la lisina desmetilasa 2B dependiente de α -cetoglutarato (KDM2B). La inhibición de KDM2B incrementa la dimetilación de la histona H3 sobre la lisina 36, lo cual permite la acumulación del complejo DNA-PK y la subsecuente reparación del DBS por recombinación no homóloga. Este proceso juega un rol crucial en la detección, regulación y/o estabilidad de la respuesta de la maquinaria de reparación del rompimiento de la doble hebra de DNA (14,24). Por lo tanto, en ausencia de la proteína fumarato hidratasa citosólica no se llevaría a cabo la reparación del DNA de una forma adecuada y se daría paso a la formación de tumores ya que mutaciones tumorigénicas pueden ser establecidas.

1.5 Mutación c.1349_1352delAGTA en el gen FH causante de HLRCC

Se han reportado una gran variedad de mutaciones de línea germinal en el gen FH en individuos con HLRCC. Las mutaciones missense parecen ser las más prevalentes, pero mutaciones frameshift, nonsense y mutaciones en el sitio de corte y empalme (splicing) han sido reportadas también (31). En una familia colombiana, se reportó la mutación c.1349_1352delAGTA en el año 2016 por Arenas y colaboradores (1), la cual no había sido reportada anteriormente en la literatura. Esta mutación provoca un desplazamiento en el marco de lectura o frameshift y posteriormente se genera un codón de parada lo que causa la no codificación de los últimos 59 aminoácidos de cada una de las subunidades de la proteína, razón por la cual cada subunidad de la proteína fumarato hidratasa silvestre tiene una longitud de 466 aminoácidos, mientras que de la proteína con esta mutación tiene una longitud de 407 aminoácidos. La familia que presenta esta mutación, tiene características específicas del Síndrome HLRCC, como son, la presencia de varios miembros afectados por cáncer renal papilar tipo 2 y mujeres con antecedentes de leiomiomas uterinos de inicio temprano; manifestaciones que siguen un claro patrón de herencia autosómica dominante. Así mismo, llama la atención en esta familia que ningún individuo presenta leiomiomas cutáneos, el cual es el síntoma más común para la enfermedad en otras familias (Figura 6).

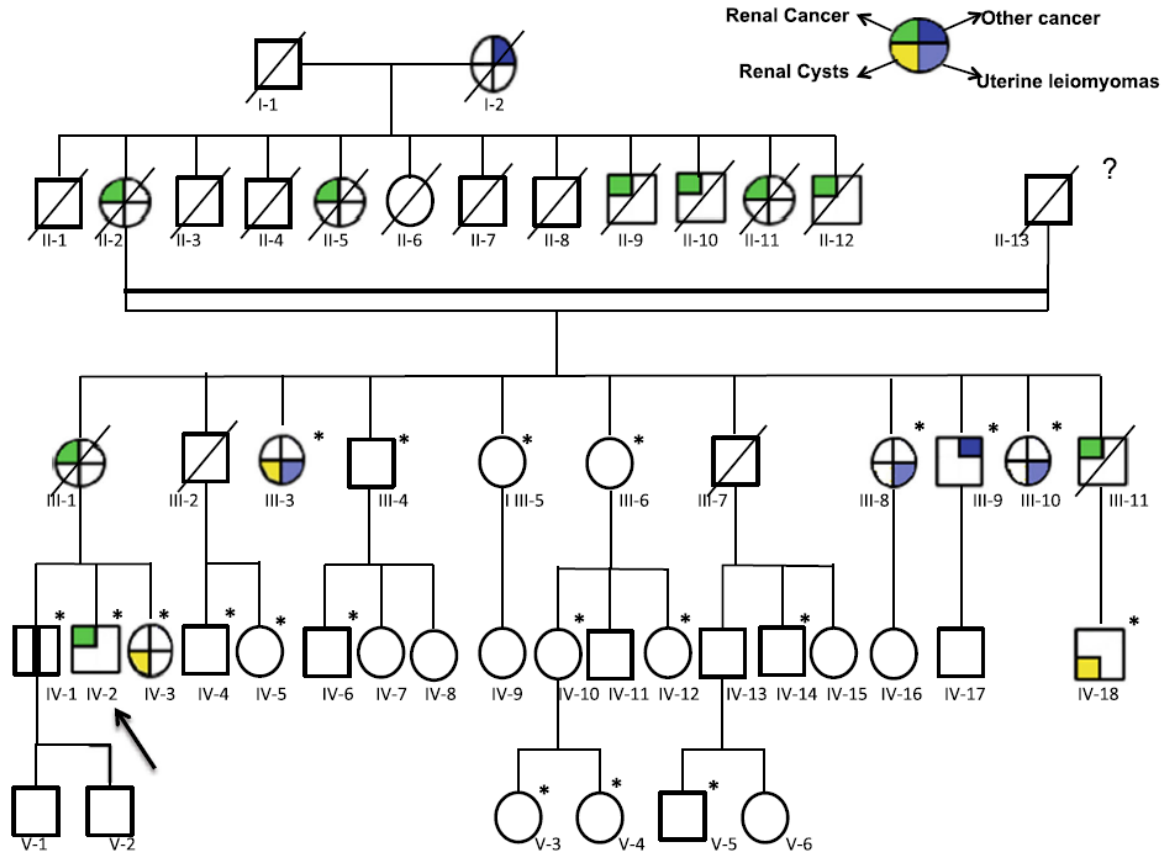


Figura 6: Pedigree de familia colombiana con integrantes portadores de la mutación c.1349_1352delAGTA en el gen FH causante del síndrome HLRCC. Los sujetos II-2 y II-13 eran primos. La información de la familia del sujeto II-13 no está disponible. El sujeto IV-2 es el probando afectado con cáncer de células renales papilar tipo 2. Los asteriscos indican a sujetos se les realizó el análisis para determinar la presencia de la mutación c.1349_1352delAGTA en el gen FH (1).

1.6 Herramientas computacionales

Debido a la diversidad de posibilidades de expresión de la enfermedad, es de gran importancia determinar los cambios estructurales y cambios en la interacción con el ligando que son provocados a partir de mutaciones como la nombrada anteriormente, para lo cual se pueden utilizar herramientas computacionales como modelamiento tridimensional de proteínas, docking molecular y biología de sistemas.

1.6.1 Modelamiento tridimensional de proteínas

Determinar la estructura tridimensional de las proteínas aporta información valiosa acerca de la base molecular de la función de las mismas, permitiendo un diseño efectivo de experimentos, como por ejemplo la relación entre diferentes mutaciones con enfermedades (32) como HLRCC. Varios métodos computacionales para modelar proteínas en 3D se han desarrollado. Teniendo en cuenta que la estructura 3D de una proteína es mejor conservada que sus secuencias, es posible identificar una proteína homóloga con su estructura conocida (plantilla) para una secuencia de proteína dada (blanco). En estos casos, el modelamiento por homología ha probado ser un método de escogencia para generar un modelo 3D de confianza desde la secuencia de aminoácidos de la proteína blanco (33). Además, el modelamiento comparativo es el único método que puede de manera segura generar un modelo en 3D de una proteína blanco a partir de su secuencia de aminoácidos. La construcción de estos tipos de modelos requiere al menos una estructura en 3D resuelta experimentalmente o una proteína cristalizada en húmedo (plantilla) que tenga una similitud significativa en la secuencia de aminoácidos con la secuencia de la proteína blanco (32).

1.6.1.1 Pasos para realizar un modelo proteico por homología

Construir un modelo por homología consiste en cuatro pasos principales: 1. Identificación de las plantillas estructurales con herramientas como BLAST y PSI-BLAST, 2. Alineamiento de la secuencia blanco con la estructura plantilla y si se encuentra una identidad entre las secuencias de un 30% o más, las dos proteínas pueden tener una estructura similar si la longitud de la región alineada es suficiente, 3. Una vez se sabe a qué y donde se parece una proteína a una estructura conocida se debe construir el modelo, este es el objetivo de los programas de modelado por homología, los cuales definen en que regiones se parece la proteína modelo a la proteína problema y en estas regiones se transfieren las coordenadas y posteriormente se completan los gaps de las regiones que no son transferibles y por ultimo 4. Evaluar la calidad del modelo. Cada uno de estos pasos requiere software especializado, así como la secuencia de la proteína actualizada y bases de datos de estructuras proteicas (33,34).

1.6.1.2 Evaluación de la calidad del modelo generado

La exactitud de la predicción de la estructura de la proteína está relacionada con el porcentaje de identidad de la secuencia entre la proteína blanco y la plantilla y para eso se usa un valor de distancia el cual es la desviación media cuadrática que por sus siglas en ingles se conoce como RMSD (root mean square deviation). RMSD es una distancia que da idea de la diferencia estructural de dos topologías. De este modo, cuanto menor sea el valor RMSD de comparación de dos estructuras, mayor es la similitud estructural de las mismas. Previo al cálculo de RMSD se deben superponer las dos estructuras que se desean comparar, en otras palabras se debe realizar un alineamiento entre las estructuras tridimensionales de las proteínas que se quieren comparar (35).

1.6.2 Área superficial accesible al solvente

La topología de la superficie de una proteína está íntimamente relacionada con su función; partes de la superficie están directamente involucradas en las interacciones con otras moléculas. La interface proteína-solvente está altamente relacionada con la estructura de la molécula nativa y la reactividad química de varios grupos funcionales dependerá de su posición en esta interface. Una descripción detallada inicial de esa superficie irregular puede ser explicada por una lista de accesibilidad de cada átomo o grupo de átomos en la estructura a una molécula solvente de tamaño definida (36).

El área superficial y accesible al solvente fue originalmente definida y computada por Lee y Richards (36,37) como el área rastreada por el centro de una sonda esférica que representa una molécula de disolvente cuando rueda sobre la superficie de una molécula de interés. Un átomo o grupo de átomos es definido como accesible si una molécula solvente de un tamaño específico puede interactuar con las interacciones Van der Waals. Normalmente, la molécula solvente se asume como una esfera con un radio de 1.4Å y para el cálculo, una esfera se centra en cada posición atómica en la lista de coordenadas y se le asigna un radio igual a la suma del átomo y la de la molécula de disolvente.

1.6.2 Docking molecular

El docking molecular es una técnica computacional que tiene como objetivo predecir si una pequeña molécula en particular se unirá establemente a una proteína blanco a nivel atómico y cómo lo hace (38,39). Esta técnica puede ser aplicada en diferentes campos, uno de ellos está enfocado a un rango de problemas estructurales bioinformáticos tales como la predicción de la función de la proteína o la eficacia de un medicamento al bloquear el sitio activo de una proteína dada. El docking molecular tiene un diseño básico que integra un algoritmo de muestreo, el cual sirve como un motor de búsqueda para explorar la orientación y el espacio conformacional del ligando dado y al menos un método de puntuación (scoring), el cual sirve como un motor de clasificación para evaluar todas las posturas de unión muestreadas (40).

La primera teoría acerca del mecanismo de unión ligando-receptor es la teoría de cerradura y llave propuesta por Fisher (41), en la cual el ligando encaja dentro del receptor como cerradura y llave. Los métodos de docking reportados últimamente fueron basados en esta teoría y tanto el ligando como el receptor fueron tratados como cuerpos rígidos. Posteriormente, fue planteada la teoría de ajuste inducido (induced-fit) (42), indicando que el sitio activo de la proteína es reconfigurado constantemente por interacciones con los ligandos a medida que los ligandos interactúan con la proteína. Esta teoría sugiere que tanto el ligando como el receptor deberían de ser tratados como moléculas flexibles durante el docking (39). Sin embargo, considerando las limitaciones de los recursos computacionales, el docking ha sido desarrollado con un ligando flexible y un receptor rígido, llamado docking semiflexible, el cual es el método más usado.

Programas como AutoDock incorporan métodos de algoritmos genéticos y genéticos lamarkianos para modelar la flexibilidad del ligando mientras mantienen el receptor rígido. La función de puntuación está basada en el Assisted Model Building with Energy Refinement (AMBER) force field, incluyendo fuerzas de van der Waals, uniones de hidrogeno, interacciones electrostáticas, entropía conformacional y términos de desolvatación. Cada término es ponderado usando un factor de escala empírica obtenido de datos experimentales (39).

1.6.3 Importancia de las herramientas computacionales en el estudio de la enfermedad HLRCC

Así como se han determinado diferentes posibilidades para el desarrollo de cáncer renal y tumores benignos por medio de experimentos en laboratorio húmedo, hasta el momento no se han realizado esfuerzos para determinar los mecanismos moleculares que toman lugar luego de la ausencia de actividad catalítica de la proteína fumarato hidratasa en el metabolismo celular, en principio debido a la enorme complejidad que este tipo de aproximación supone desde el punto de vista de herramientas de biología molecular que predominan en la práctica de la biología (43). Sin embargo, la ciencia ha podido en años recientes acercarse más a la aproximación holística de los sistemas biológicos, mediante la creación de modelos computacionales que nos permiten tomar en cuenta mucha de la complejidad biológica, difícil de abalar por dichos medios convencionales. De esta manera, aproximaciones como las redes celulares complejas que pueden generar fenómenos emergentes que serían indetectables por los enfoques reduccionistas (44) son de gran utilidad actualmente.

Ejemplos de este tipo de redes, son las complejas interconexiones celulares en procesos de señalización, interacción proteína-proteína o el conjunto de reacciones bioquímicas en el metabolismo celular. En este último aspecto, el área de la bioquímica durante mucho tiempo se ha ocupado de la determinación y caracterización de la interconexión de las vías metabólicas. Dentro de este tipo de estudios, podemos encontrar el realizado por Yang y colaboradores en el año 2013 (45), en el cual plantean que las células renales cancerosas por causa de la deficiencia de FH, sufren una remodelación metabólica con cambios en la respiración mitocondrial, metabolismo de la glucosa y la glutamina, lo cual soporta la proliferación de células malignas. De esta manera, con el fin de caracterizar las alteraciones en la glicolisis, el ciclo de Krebs y la vía de las pentosas fosfato por medio de espectrometría de masas y resonancia magnética nuclear utilizaron trazadores como [U-¹³C]-glucosa, [U-¹³C,N]-glutamina y [1,2-¹³C₂]-glucosa para trazar las vías de las líneas celulares que tenían en crecimiento. La reprogramación metabólica de las células con ausencia de la actividad catalítica para la enzima fumarato hidratasa fue comparada con la programación de células en las cuales la actividad de la enzima había sido restaurada. Como resultado observaron que las células nulas para FH mostraron una sustancial reorganización de su flujo metabólico para alcanzar su alta demanda de ATP (45).

1.6.3.1 Modelos a escala genómica

En la actualidad, gracias a las capacidades de la secuenciación genómica moderna, la creación de modelos computacionales de estas vías ha incrementado y ha sido integrada con el desarrollo de modelos metabólicos computacionales a escala genómica (46), un proceso mediante el cual, a partir del conjunto de genes codificados en un genoma se infiere el conjunto de reacciones que hacen parte del metabolismo del organismo bajo estudio y se genera un modelo computacional del mismo. Esta metodología hace parte de un número importante de nuevas aproximaciones al estudio de sistemas vivos.

La biología de sistemas es un campo que crece rápidamente y en términos generales parte de su aplicación se encuentra en la reconstrucción y validación de modelos *in silico* de sistemas biológicos. Este campo estudia los fenómenos biológicos a nivel de redes de interacción usando un gran número de datos experimentales. La generación de los modelos *in silico*, se da a partir de la unión de conocimientos bioquímicos, genómicos y genéticos con su respectiva formalización matemática, lo cual permite la generación de, por ejemplo, modelos metabólicos a escala genómica, modelos que permiten predecir rasgos fenotípicos basándose en la composición genética del organismo objetivo (46).

1.6.3.2 Reconstrucciones metabólicas

Una reconstrucción metabólica, es definida como una lista de reacciones bioquímicas que ocurren en un vía de señalización celular particular y las asociaciones entre estas reacciones y genes, transcritos o proteínas relevantes (47,48). Las reconstrucciones de este tipo, incluyen principalmente, pero no exclusivamente, reacciones de las vías metabólicas centrales, tales como la glucólisis, la vía de las pentosas fosfato y el ciclo del ácido cítrico, así como la reacciones que están involucradas en la biosíntesis de aminoácidos y nucleótidos, la síntesis de la pared celular, la biosíntesis de cofactores y utilización alterna de fuentes de nitrógeno y carbono (49). Estos modelos computacionales del metabolismo han permitido el análisis de fenómenos emergentes a través del estudio del conjunto total de interacciones en todas las redes más que en vías o genes individuales y muchas técnicas computacionales han sido desarrolladas para probar sus propiedades. Estos tipos de análisis a nivel de redes son indispensables para entender plenamente las relaciones genotipo-fenotipo en las células. Las reconstrucciones metabólicas a escala genómica contribuyen significativamente al desarrollo de modelos mecanísticos y predictivos a nivel celular, un objetivo poco realista

antes de la revolución genómica (46). Una de las reconstrucciones metabólicas de mayor confiabilidad y en la cual la comunidad científica ha trabajado arduamente los últimos años es la del metabolismo humano, conocida como Recon 2 (50), la cual cuenta con 1789 genes codificantes para enzimas, 7440 reacciones correspondientes a 100 rutas (entre las que se encuentran glicolisis, ciclo de Krebs, pentosas fosfato, síntesis y degradación de nucleótidos, metabolismo de lípidos, entre otras) y 2626 metabolitos únicos distribuidos sobre ocho compartimentos celulares. Para ensamblar Recon 2 se adicionó información metabólica presente en cuatro recursos diferentes a Recon 1 (Figura 7).

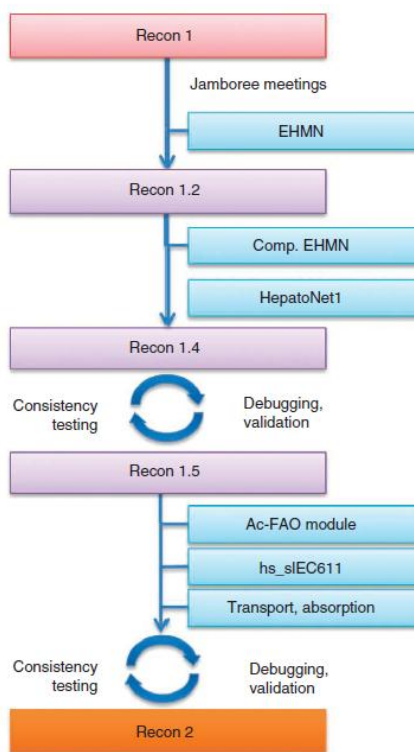


Figura 7: Descripción general del community-driven reconstruction approach para ensamblar Recon 2.

Comp. EHMN, compartmentalized EHMN; and hs_sIEC611, small intestinal enterocyte reconstruction” (50).

Tanto Recon 1 y Recon 2 han probado su utilidad en diversas investigaciones relacionadas con interpretación de datos y experimentos de alto rendimiento para descubrimiento biológico y simulación matemática de fenotipos para guiar experimentos que tienen aplicaciones en estudios de efectos de medicamentos y estudios patológicos(51). Adicional a esto, se han realizado investigaciones de hospedero-

patógeno como por ejemplo, determinar diferencias en los tipos de infecciones de *M. tuberculosis*; para esto se construyó una red combinando Recon 1 derivado del modelo de macrófago alveolar y un modelo existente de *M. tuberculosis* para modelar la infección de *M. tuberculosis* dentro de un fagosoma hospedero (52). De la misma manera en el año 2016 se creó AGORA, la cual es una reconstrucción metabólica a escala genómica para 773 bacterias presentes en el intestino humano; la cual es compatible con Recon 2 y de esta manera se pueden realizar estudios de interacción entre microbioma y hospedero (53). También se han realizado estudios relacionados con la sensibilidad de la insulina y otras aproximaciones farmacológicas, así como predicción de fenotipos en organismos con mutaciones o variantes genéticas (54,55).

1.6.3.3 Proceso para realizar una reconstrucción metabólica a escala genómica

A la fecha, todas las reconstrucciones metabólicas a escala genómica de alta confianza han sido construidas manualmente a través de un proceso de cuatro pasos: (1) una reconstrucción inicial es creada a partir de datos de anotación de genes acoplados con información de bases de datos on line como KEGG (56,57) y ExPasy (58). En otras palabras, la asignación de las reacciones a los genes asociados a algún tipo de actividad enzimática. (2) Curar la reconstrucción inicial, lo cual se realiza a través de una revisión de literatura. Esto consiste en la identificación de las reacciones sin enlazar en la red metabólica y su reconexión mediante búsqueda en la literatura o en análisis experimentales. Posteriormente, la reconstrucción como un conocimiento de base, es convertida a un modelo matemático que puede ser analizado a través de aproximaciones basadas en restricciones. (3) La reconstrucción es validada a través de la comparación de predicciones del modelo para datos fenotípicos. Por ejemplo, evaluando su capacidad de producir ATP a partir de distintas fuentes de glucosa (4). Idealmente, los datos obtenidos de una reconstrucción metabólica es sometida a ciclos de ensayos en laboratorio húmedo y seco continuos, lo cual mejora la exactitud y permite la investigación de hipótesis clave (46) (Figura 8).

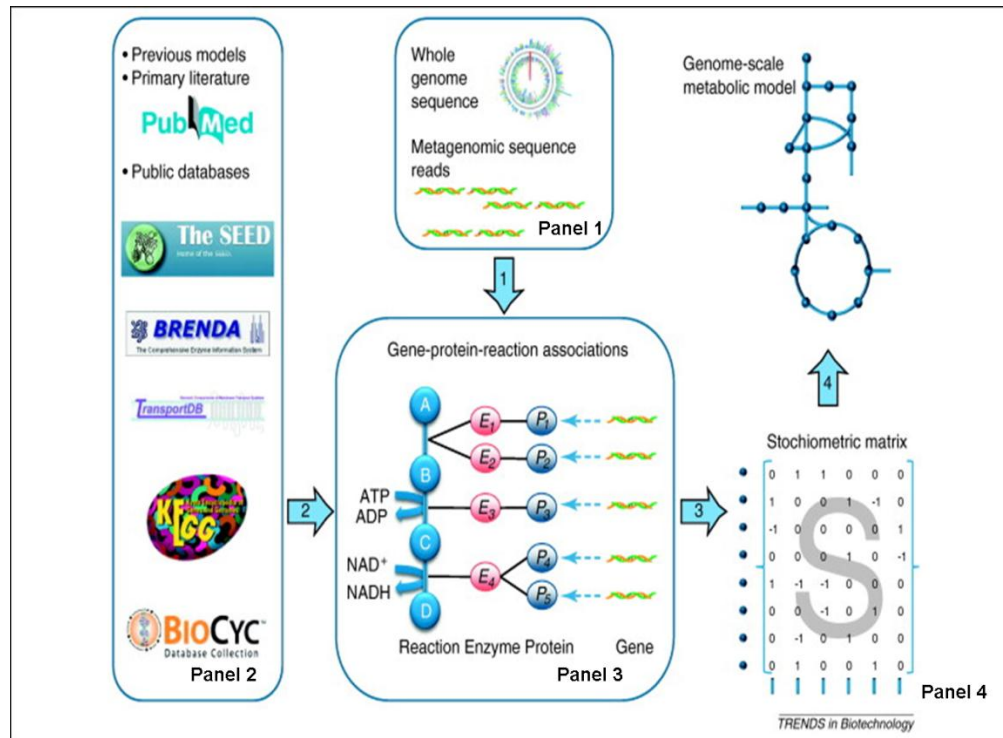


Figura 8: Reconstrucción de un modelo metabólico a escala genómica.

La reconstrucción comienza con un genoma del organismo estudiado (panel 1). Revisión de literatura primaria, modelos previos y bases de datos públicas con el fin de reunir información genómica y metabólica, lo cual facilita el ensamblaje de asociaciones gen-proteína-reacciones (panel 2). La reconstrucción se somete a comparaciones con modelos previos y bases de datos y es curado manualmente con la ayuda de literatura (panel 3). Las reacciones metabólicas son formuladas en una representación matemática en una matriz estequiométrica (panel 4). La matriz estequiométrica en compañía de información acerca de genes, proteínas, metabolitos y reacciones son ensambladas en un modelo metabólico a escala genómica (59).

Las reconstrucciones y sus respectivos modelos permiten la formulación de hipótesis acerca de la presencia de ciertas actividades enzimáticas y la producción de metabolitos a partir de la modificación de ciertos comportamientos dentro del modelo. Para la realización del modelo se hace necesaria la representación matemática del conjunto de reacciones bioquímicas en el mismo. Esto se da mediante la creación de una matriz estequiométrica "S", la cual contiene los coeficientes estequiométricos de cada reacción. De esta manera, cada fila de la matriz representa un compuesto o metabolito único y cada columna representa una reacción, las celdas son los coeficientes estequiométricos de los metabolitos que participan en una reacción (Figura 8, (panel 4)). En la matriz, se coloca un coeficiente negativo por cada metabolito consumido y un coeficiente positivo por cada metabolito producido. Esta estequiometría impone restricciones en el flujo de

metabolitos a través de la red, pues determina cuales y cuantos metabolitos fluyen a través de una reacción particular, asegurando de esta forma que la cantidad total de cualquier compuesto que se produce es igual a la cantidad total que se consume en estado estacionario. De la misma manera, para cada reacción también se establecen límites inferiores y superiores, los cuales definen los flujos máximos y mínimos permitidos en las reacciones. Los balances estequiométricos y los límites inferiores y superiores, definen el espacio de distribuciones de flujo de un sistema, en otras palabras, es la proporción a la cual cada metabolito es consumido o producido por cada reacción (60).

Algunas restricciones también se pueden añadir, como por ejemplo las provenientes de datos de expresión. De esta manera, la construcción de un modelo es en sí mismo el establecimiento de un laboratorio *in silico* al cual se le pueden realizar diferentes preguntas que son respondidas a partir de restricciones dentro del modelo. Existen varios métodos que permiten interrogar o realizar análisis sobre estos modelos, unos de los más usados por la comunidad científica son el Análisis de Balance de Flujo o FBA y el análisis de variabilidad de flujo o FVA por sus siglas en inglés, entre otros.

1.6.3.4 Análisis de balance de flujo (FBA)

El análisis de balance de flujo (FBA), es un método utilizado para el estudio de redes bioquímicas, en particular las reconstrucciones metabólicas como una manera de interrogar una red. Este método calcula el patrón de flujo de metabolitos a través de una red metabólica, lo que hace posible predecir la tasa de crecimiento o la tasa de producción de un metabolito importante biotecnológicamente (60).

El primer paso antes de proceder a realizar un FBA es representar matemáticamente las reacciones metabólicas, la característica fundamental de esta representación es una tabulación (figura 8, (panel 4)). Estas estequiometrias imponen restricciones sobre el flujo de metabolitos a través de la red y estos tipos de ajustes son el corazón del FBA, ya que diferencian a este análisis de modelos basados en parámetros cinéticos de difícil medida (61). El segundo paso consiste en definir un objetivo biológico que es relevante para el problema en estudio. Matemáticamente, el objetivo biológico es representado por una “función objetivo” que indica qué cantidad de cada reacción contribuye al fenotipo (60). Este proceso de optimización produce un set óptimo de valores de flujo metabólico que son consistentes con maximización o minimización del objetivo escogido (46). En un estudio, se pueden establecer una o varias funciones objetivo de acuerdo al problema biológico que quiera ser investigado.

Las representaciones matemáticas de las reacciones metabólicas junto con las representaciones del fenotipo definen un sistema de ecuaciones lineales. En FBA, estas ecuaciones son solucionadas usando programación lineal. Existen muchos algoritmos de programación lineal, los cuales pueden ser muy rápidos identificando soluciones óptimas para grandes sistemas de ecuaciones, un ejemplo de estos son los implementados en el paquete constraint-based reconstruction and analysis (COBRA) COBRA Toolbox (60).

El FBA se refiere al siguiente programa lineal (LP):

$$\begin{aligned} \text{Max}_v Z &= c^T v \\ \text{Sujeto a: } Sv &= 0 \\ v_l &< v < v_u \end{aligned} \quad (1)$$

Donde la matriz **S** es una matriz estequiométrica $m \times n$ con m metabolitos y n reacciones y **c** es el vector que representa la función objetivo lineal. Las variables de decisión **v** representan los flujos, con $V \subseteq \mathbb{R}^n$ y los vectores **v_l** y **v_u** especifican los límites inferiores y superiores respectivamente. Las restricciones **Sv=0** junto con los límites inferior y superior especifican la región factible del problema (62).

La matriz estequiométrica (S) está representada en la figura 8 (panel 3) y es una matriz en la cual la mayoría de reacciones bioquímicas involucra solamente pocos metabolitos diferentes.

Cualquier v que satisface esta ecuación se dice que está en el espacio nulo de S. En cualquier modelo metabólico de gran escala, hay más reacciones que componentes ($n > m$). En otras palabras hay más variables desconocidas que ecuaciones, por lo cual no hay una solución única para este sistema de ecuaciones.

Aunque las restricciones definen un rango de soluciones, es posible identificar y analizar puntos individuales dentro del espacio de solución. FBA es un método para identificar tales puntos óptimos dentro de un espacio restringido (Figura 9).

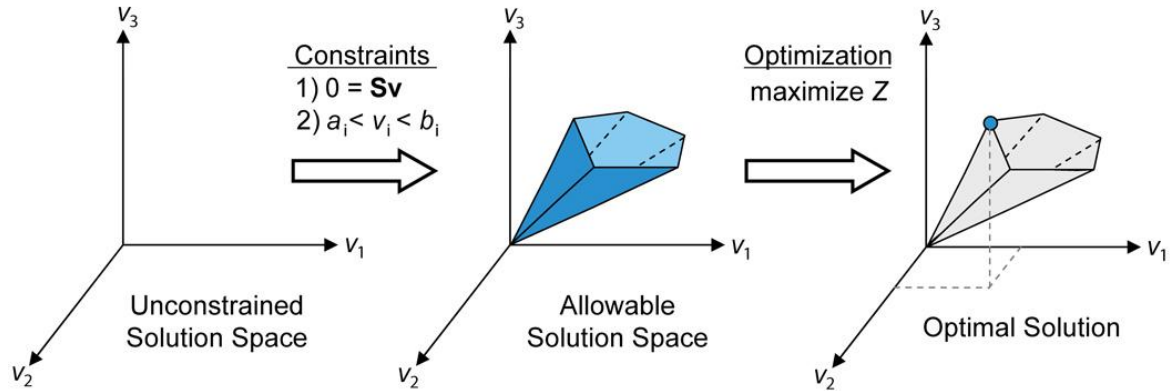


Figura 9: Base conceptual del modelamiento basado en restricciones y FBA.

Sin restricciones, la distribución de flujo de una red metabólica puede encontrarse en cualquier punto en un espacio de solución. Cuando la restricción del balance de masas es impuesta por la matriz estequiométrica S (panel 1) y se imponen restricciones de capacidad por los límites inferiores y superiores (a_i y b_i) (panel 2), se define el espacio de solución disponible. La red puede adquirir cualquier distribución de flujo dentro de ese espacio, pero los puntos que se encuentran fuera de ese espacio son negados por las restricciones. A través de la optimización de la función objetivo, FBA puede identificar una única distribución de flujo óptimo que se encuentra en el borde del espacio de solución permitido (60).

Sin embargo, es importante mencionar que FBA tiene algunas limitantes, este análisis no usa parámetros cinéticos por lo que no puede predecir concentraciones de metabolitos. De la misma manera, solo es adecuado para la determinación de flujos en estado estacionario. Además, FBA no tiene en cuenta los efectos de regulación tales como la activación de enzimas proteína-quinasa o regulación de la expresión génica, por lo que sus predicciones no siempre son exactas (60).

1.6.3.5 Análisis de variabilidad de flujo (FVA)

Como se dijo anteriormente, por medio de un análisis FBA se obtiene una única solución al problema de optimización y a pesar de que pueden existir varias soluciones al mismo problema, el FBA nos muestra el estado del sistema en un único momento. Sin embargo, pueden existir varias combinaciones de valores de flujos que nos lleven a tener el mismo objetivo máximo o el mismo valor óptimo.

El análisis de variabilidad de flujo (FVA) es una herramienta importante con la cual se extienden los resultados obtenidos por el análisis de balance de flujo (FBA). El FVA determina los valores máximo y mínimo de todos los flujos que satisfarán las restricciones y permitirán el mismo valor objetivo óptimo (63) o el valor máximo y mínimo

posible de flujo a través de una reacción particular para un valor objetivo máximo.. Las aplicaciones de FVA para biología de sistemas molecular, incluye pero no está limitado a la exploración de alternativas óptimas de FBA (62). En otras palabras, el objetivo del FVA es identificar para cada una de las reacciones que participan en el FBA los valores de flujo máximo y mínimo.

Los sistemas biológicos frecuentemente contienen redundancias que contribuyen significativamente a su robustez y esas redundancias en la red metabólica pueden ser analizadas por medio de las distribuciones de flujo optimas alternativas y, por lo tanto, contribuyen a la robustez en el contexto del crecimiento óptimo y la variabilidad de flujo (64,65) y para esto es usado el FVA.

En otras palabras, el FVA es usado para determinar el rango de posibles valores de cada reacción a partir de los múltiples óptimos. La aproximación empieza determinando el valor de la función objetivo. A partir de esta solución, el rango de variabilidad que puede existir en cada flujo en la red debido a soluciones óptimas alternas puede ser calculado a través de una serie de problemas (problema lineal) en los que el valor de la función objetivo original se fija y cada reacción en la red se maximiza y posteriormente se minimiza para determinar el rango factible de valores de flujo para cada reacción (64).

Para algunos flujos, su valor máximo será igual a su valor mínimo, lo que permite una predicción específica. Para otros, puede haber un amplio rango de posibles valores debido a vías alternas (64).

De manera más formal, el FVA se describe de la siguiente manera: sea \mathbf{w} un objetivo biológico como la biomasa o la producción de ATP. Después de resolver (1) con $\mathbf{c}=\mathbf{w}$, FVA resuelve dos problemas de optimización para cada flujo v_i de interés

$$\begin{aligned} \text{Max}_v / \text{Min}_v &= v_i \\ \text{Sujeto a: } \mathbf{S}\mathbf{v} &= \mathbf{0} \\ W^T \mathbf{v} &\geq \pm Z_0 \\ v_l &< v < v_u \end{aligned} \quad (2)$$

Donde $Z_0 = \mathbf{w}^T \mathbf{v}_0$ es una solución óptima para (1), γ es un parámetro, que controla si el análisis se hace w.r.t. estados de red subóptimos ($0 \leq \gamma < 1$) o el estado óptimo ($\gamma=1$).

Asumiendo que todas las reacciones n son de interés, FVA requiere la solución de $2n$ LPs (62).

2. Planteamiento del Problema

El síndrome de leiomiomatosis hereditaria y cáncer renal hereditario (HLRCC, OMIM 150800) es un desorden hereditario autosómico dominante, caracterizado por una predisposición a desarrollar leiomiomas cutáneos y uterinos, cáncer de células renales y leiomiosarcomas uterinos (4). El gen asociado con esta enfermedad es el gen FH, gen nuclear que codifica para la enzima fumarato hidratasa, la cual cataliza la reacción reversible hidratación/deshidratación del fumarato a L-malato. Este gen se encuentra localizado en el brazo largo del cromosoma 1 (1q42.1) y codifica para dos formas que son diferencialmente procesadas (forma mitocondrial y citosólica), pero idénticas en la secuencia (12). En la mitocondria, la enzima fumarato hidratasa está involucrada en el ciclo del ácido tri-carboxílico o ciclo de Krebs (15). Por otra parte, se cree que la isoforma citoplasmática está involucrada en el metabolismo del fumarato (13), metabolismo de aminoácidos (26) y en la reparación de rompimiento de doble hebra del DNA (DSBs) (24).

En una familia colombiana, se reportó la mutación c.1349_1352delAGTA en el año 2016 por Arenas y colaboradores (1). Esta mutación genera un desplazamiento en el marco de lectura (frameshift). Producto de esta modificación se genera un codón de parada y no se codifican los últimos 59 aminoácidos de la proteína. La familia que presenta esta mutación, tiene unas características específicas del síndrome HLRCC, entre ellos, varios pacientes con presencia de cáncer renal papilar tipo 2 y mujeres con leiomiomas uterinos de diagnóstico temprano. No obstante, ningún paciente presenta leiomiomas cutáneos.

Es de gran importancia resaltar que el síndrome HLRCC, como se ha discutido antes, hace parte de las enfermedades que generan cáncer renal hereditario, el cual solo es un 3% de todos los cánceres renales, siendo el 97% restante de tipo esporádico. En este sentido, el HLRCC se considera una patología rara y poco investigada, a pesar de que los pacientes afectados pueden presentar un tipo de cáncer altamente agresivo que puede hacer metástasis incluso en etapas tempranas. Además, es una condición que afecta a varios integrantes de una misma familia, lo que lo hace más impactante para las familias que lo padecen.

Teniendo en cuenta que realizar investigaciones húmedas en células de estos pacientes es difícil, ya que conseguir tejido tumoral es una labor de gran complejidad, los análisis computacionales se convierten en una herramienta de gran ayuda para determinar los cambios estructurales y los comportamientos metabólicos celulares al enfrentarse a las modificaciones de la proteína fumarato hidratasa. Entre estos análisis computacionales se encuentran el análisis de propiedades físico-químicas, el modelamiento de estructuras proteicas en tercera dimensión, el docking molecular y las reconstrucciones metabólicas a escala genómica, lo que permite analizar los impactos de mutaciones sobre las funciones de la proteína y las implicaciones a nivel metabólico de la célula a un nivel de biología de sistemas, permitiendo tener una visión general del comportamiento celular y no solo evaluar unas pocas vías o pocos genes.

Actualmente no hay investigaciones computacionales acerca de la mutación relacionada en este estudio y es de gran importancia dilucidar los impactos de esta mutación sobre la estructura proteica y el metabolismo celular, teniendo en cuenta que los pacientes que son portadores de esta mutación no presentan leiomiomas cutáneos el cual normalmente es el síntoma más común y presentan con mayor frecuencia cáncer papilar renal tipo II el cual es el síntoma menos frecuente es los pacientes con el síndrome HLRCC .

Por todo lo anterior, con este estudio buscamos predecir computacionalmente el posible impacto de la mutación c.1349_1352delAGTA en el gen FH, con respecto a su estructura proteica, así como predecir también los posibles cambios que se pueden presentar en el metabolismo celular frente a estas modificaciones.

3. Justificación

La leiomatosis hereditaria y cáncer renal hereditario (HLRCC) es una patología muy rara y se estima que aproximadamente 200 familias en el mundo la padecen. Este síndrome no ha sido ampliamente estudiado y aunque han surgido un gran número de hipótesis acerca de por qué mutaciones en el gen FH llevan a la formación de leiomiomas y en el peor de los casos llevan al desarrollo de cáncer renal, no se ha identificado con certeza el mecanismo por el cual se desarrollan estas afecciones.

Las mutaciones que se encuentran con más frecuencia en las familias que tienen este síndrome son mutaciones con cambio de sentido (missense), las cuales ocurren justo en algún aminoácido que hace parte del sitio activo impidiendo su acción catalítica, o provocando un cambio estructural que hace que el sitio activo no pueda aceptar de forma adecuada al ligando o incluso mutaciones missense que impiden la oligotetramerización de la proteína y por lo tanto la no formación de los sitios activos.

La mutación c.1349_1352delAGTA en el gen FH, reportada en una familia colombiana tiene una presentación clínica que difiere un poco de la presentación típica del síndrome HLRCC (familia con alto número de familiares con cáncer renal y ningún paciente con leiomatosis cutánea, el cual es el síntoma más común en los pacientes con este síndrome); razón por la cual es de interés evaluar cuál es la modificación estructural que sufre la proteína fumarato hidratasa al presentar esta mutación comparándola con la proteína silvestre.

De la misma manera, a lo largo del tiempo se han estudiado las implicaciones de la ausencia de la actividad de la proteína fumarato hidratasa asociándola con factores de transcripción implicados en la supervivencia celular, pero no se ha evaluado el impacto de la ausencia o posibles alteraciones de la estructura de esta proteína en el metabolismo celular observado desde la biología de sistemas, tema de gran importancia, ya que la proteína fumarato hidratasa está involucrada en el ciclo de krebs, una de las rutas más importantes en el metabolismo celular.

Se debe resaltar que los estudios bioinformáticos y de biología de sistemas son de gran utilidad actualmente, ya que existe una gran cantidad de datos biológicos, químicos, entre otros que deben de ser utilizados y aprovechados al máximo, de esta manera se

pueden realizar estudios de este tipo, en los cuales se pueden hacer aproximaciones que permiten generar hipótesis biológicas acerca de lo que sucede *in vivo* a partir de la simulación computacional, permitiendo obtener resultados plausibles, aprovechando la información que ha sido desarrollada por otros investigadores. Además, es la única manera de tener una visión holística del comportamiento celular frente a ciertas condiciones genéticas o ambientales, cosa que es difícil lograr con experimentos húmedos. Sin embargo, en ningún momento se debe pensar que los experimentos húmedos deben de ser reemplazados completamente por experimentos *in silico*, ya que los primeros son indispensables para llevar a cabo los segundos.

4. Objetivos

4.1 Objetivo general

Describir computacionalmente los posibles cambios estructurales de la proteína fumarato hidratasa relacionados con la mutación c.1349_1352delAGTA en el gen FH causante del síndrome de leiomiomatosis hereditaria y cáncer renal hereditario (HLRCC), así como predecir también las posibles implicaciones de estos cambios en el metabolismo celular humano.

4.2 Objetivos específico

- Modelar y caracterizar las variantes estructurales de la proteína fumarato hidratasa a partir de la mutación c.1349_1352delAGTA en el gen FH.
- Describir el posible efecto de las variantes halladas de la proteína fumarato hidratasa en la interacción con los ligandos fumarato y malato.
- Determinar las posibles implicaciones moleculares de la interacción de la proteína fumarato hidratasa mutada con los ligandos fumarato y malato, en un modelo computacional del metabolismo humano.

5. Metodología

En el año 2016 se llevó a cabo un estudio en el cual se analizó la mutación c.1349_1352delAGTA en el gen FH en una familia colombiana (1). En este estudio, se caracterizó una familia completa determinando los portadores de la mutación y las personas que habían desarrollado los síntomas característicos del síndrome de leiomiomatosis hereditaria y cáncer de células renales (HLRCC) (figura 6). En este estudio fue reportada por primera vez la mutación mencionada en el gen codificante de la proteína fumarato hidratasa.

La metodología del presente estudio se llevó a cabo de acuerdo a cada uno de los objetivos específicos:

5.1 Modelar y caracterizar las variantes estructurales de la proteína fumarato hidratasa a partir de la mutación c.1349_1352delAGTA en el gen FH

La secuencia de la proteína Fumarato Hidratasa humana fue obtenida de UNIPROT con el número de acceso P07954-1 (66). Posteriormente sobre el marco de lectura abierto (ORF) de la proteína Fumarato Hidratasa (NM_000143) (67) humana obtenida del National Center for Biotechnology Information (NCBI), se suprimieron los nucleótidos correspondientes a la mutación c.1349_1352delATGA en el exón 9, mutación encontrada en los pacientes reportados por Arenas y colaboradores con HLRCC (1) y por medio de la herramienta ExpASy- translate tool (68) se obtuvo la secuencia de aminoácidos mutada.

Por medio del software Protein Multiple Sequence Alignment (Muscle) (69) se realizó el alineamiento de la secuencia de aminoácidos nativa contra la secuencia mutada, lo cual fue reportado en formato clustal con el fin de cuantificar el porcentaje de identidad entre dichas secuencias de aminoácidos.

Se visualizó la porción de la proteína que se afecta o se deleciona cuando está presente la mutación en la estructura terciaria (monómero) y se comparó con la estructura terciaria

de la proteína silvestre utilizando el programa UCSF Chimera (70). De la misma manera se señalaron las regiones que hacen parte de los sitios activos de la proteína para poder observar su afectación en la proteína mutada y por último se observó la manera en la cual se forma la estructura tetramérica con el mismo para entender la afectación de esta por la mutación en mención.

Por otro lado, la estructura cuaternaria tridimensional cristalizada de la proteína fumarato hidratasa nativa fue descargada del RCSB Protein Data Bank (RCSB PDB) (71) con el código 3e04, la cual fue obtenida mediante difracción de rayos-X con una resolución de 1.95 Å (72). Mientras que el modelamiento de la estructura cuaternaria tridimensional de la proteína fumarato hidratasa con la mutación c.1349_1352delATGA se llevó a cabo con el programa SWISS-MODEL (33,73,74) usando como plantilla la proteína con código 3e04 depositada en el PDB.

La validación del modelo obtenido se llevó a cabo mediante la herramienta “structure assesment” de SWISS-MODEL (33), determinando la correcta conformación geométrica de la proteína por medio del análisis estereoquímico presentado en el gráfico de Ramachandran (75).

Mediante MultiSeq (76), el cual es un plugin de VMD (77) se llevó a cabo el alineamiento de la estructura cuaternaria tridimensional de la proteína fumarato hidratasa silvestre contra la estructura cuaternaria tridimensional de la proteína fumarato hidratasa mutada y se realizó el análisis de la variación tridimensional de las estructuras calculando el RMSD.

Mediante el programa UCSF ChimeraX (78) se calculó el área de superficie accesible al solvente tanto de la estructura cuaternaria tridimensional silvestre como de la estructura cuaternaria tridimensional mutada para la estructura completa y mediante la herramienta on-line Accessible Surface Area and Accessibility Calculation for Protein (ver. 1.2) de la universidad de Ochanomizu (79) se obtuvo el área de superficie accesible al solvente de cada uno de los residuos que conforman cada una de las proteínas (silvestre y mutada) en cada una de las cadenas.

Por otra parte, se obtuvieron las secuencias de la proteína fumarato hidratasa con las mutaciones reportadas en uniProt (80), esto con el fin de comparar las propiedades físico-químicas de cada una de las proteínas mutadas con la proteína silvestre.

Con la herramienta ExpASy- ProtParam tool (68) se computaron las propiedades físico-químicas de cada una de las secuencias de aminoácidos:

- **Peso molecular (Mw) y punto isoeléctrico (pI):** Esta herramienta calcula estos dos parámetros buscando en las bases de datos Swiss-Prot y TrEMBL proteínas que contengan la composición de aminoácidos más cercana a la proteína blanco (81).
- **Índice alifático:** se define como el volumen relativo ocupado por las cadenas laterales alifáticas de una proteína (alanina, valina, isoleucina y leucina). El índice alifático de una proteína es calculado de acuerdo a la siguiente formula:

$$X_{Ala} + \alpha X_{Val} + \beta(X_{Ile} + X_{Leu})$$

Donde X_{Ala} , X_{Val} , X_{Ile} y X_{Leu} son el porcentaje molar (100 X fracción molar) de alanina, valina, isoleucina y leucina.

Los coeficientes α and β son el volumen relativo de la cadena lateral valina ($\alpha = 2.9$) y de la cadena lateral Leu/Ile ($\beta = 3.9$) a la cadena lateral de alanina (82).

- **GRAVY (promedio de hidrofobicidad):** el valor GRAVY para un péptido o proteína es calculado como la suma de los valores de hidrofobicidad de todos los aminoácidos, dividido por el número de residuos en la secuencia (83).
- **Índice de inestabilidad (II):** provee un estimado de la estabilidad de la proteína en un tubo de ensayo. Un análisis estadístico de 12 proteínas inestables y 32 estables ha revelado que hay ciertos dipéptidos cuya ocurrencia es significativamente diferente en las proteínas inestables comparadas con las estables. Los autores de este método le asignaron un valor de inestabilidad a cada uno de los 400 dipéptidos diferentes (DIWV) (84). Usando este valor es posible computar un índice de inestabilidad el cual se define como:

$$II = (10/L) \sum_{i=1}^{L-1} DIWV_{x_i y_{i+1}}$$

Donde L es la longitud de la secuencia y $DIWV_{x_i y_{i+1}}$ es el valor de inestabilidad para el dipéptido empezando en la posición i .

Una proteína cuyo índice de inestabilidad es más bajo de 40 es catalogada como estable, un valor por encima de 40 predice que la proteína puede ser inestable (84).

Estas propiedades fueron graficadas en diagramas de puntos y se llevó a cabo un análisis estadístico de chi-cuadrado por medio del paquete estadístico R (85). Este análisis estadístico se llevó a cabo con el fin de evaluar la independencia de las variables físico-químicas nombradas anteriormente entre la proteína fumarato hidratasa nativa, la misma proteína con las mutaciones presentes en el gen FH reportadas en uniprot causantes de HLRCC (80) y la proteína con la mutación c.1349_1352delATGA.

5.2 Describir el posible efecto de las variantes halladas de la proteína fumarato hidratasa en la interacción con los ligandos fumarato y malato

Como paso inicial se llevó a cabo un análisis de sitio de unión y análisis de hendidura con el servidor ProFunc (86), el cual sirve para ayudar a identificar la función bioquímica de una proteína a partir de su estructura tridimensional. Esta herramienta combina un número de métodos basados en la secuencia y en la estructura tridimensional utilizando un algoritmo interno y servicios externos a través del uso de servicios web para analizar una estructura cargada en el PDB. Al realizar el análisis de las hendiduras, se asume que el sitio activo de una proteína se encuentra frecuentemente en la hendidura más grande de la proteína y esta hendidura es significativamente más grande que las otras.

ProFunc utiliza el algoritmo SURFNET (87) que detecta las regiones de separación dentro de la proteína mediante el ajuste de esferas de un cierto rango de tamaños entre los átomos de la proteína. Una vez que se identifican las hendiduras y los bolsillos, se pueden comparar con las bases de datos de hendiduras conocidas para identificar similitudes locales.

Para realizar el docking molecular, el metabolito malato fue construido de forma manual por medio del programa PyMOL Viewer versión 1.7.4.5 (88) con base en la estructura

obtenida de la base de datos Zinc Docking (89) con el código de ingreso 895264. Por otro lado, la estructura del fumarato fue obtenida de la base de datos Zinc Docking con el código de ingreso 3860193. Cada uno de los ligandos fue preparado para realizar el docking añadiendo hidrógenos polares y se identificaron enlaces rotables.

La proteína fumarato hidratasa silvestre fue preparada removiendo las moléculas de agua y eliminando moléculas de etilenglicol el cual fue utilizado como ligando en el momento en el que fue cristalizada la proteína. El procedimiento se llevó a cabo con el software AutoDock Vina (90). La malla de muestreo fue centrada en los aminoácidos serina 145 (cadena D), treonina 147 (cadena D), Asparagina 188 (cadena D), Glicina 364 (cadena C), Histidina 235 (cadena A), los cuales hacen parte del sitio activo de la proteína fumarato hidratasa silvestre (91) con una caja de muestreo de 20Å a lo largo de cada eje para realizar los cálculos. Esto se llevó a cabo con la proteína silvestre y mutada con cada uno de los ligandos. El docking molecular fue realizado con metodología semiflexible en la cual el ligando se trató como una molécula flexible, mientras que la proteína fue tratada como una molécula rígida.

Con las energías de unión (Kcal/mol) se comparó la interacción entre cada una de las proteínas con ambos metabolitos y se estableció si la mutación tiene algún efecto en la unión del ligando con la proteína. Posterior a esto y con el software Discovery Studio Visualizer v.17.2.0.16349 (92) se analizaron las interacciones que se establecen entre los residuos de la proteína con los átomos del ligando.

5.3 Determinar las posibles implicaciones moleculares de la interacción de la proteína fumarato hidratasa mutada con los ligandos fumarato y malato, en un modelo computacional del metabolismo humano.

Con el fin de estudiar el posible efecto en el metabolismo humano de la ausencia de actividad de la proteína fumarato hidratasa o la posible alteración de la estructura de la misma y por consiguiente la mala interacción con los ligandos fumarato y malato se utilizó la reconstrucción metabólica humana Recon 2.04 (50).

Para realizar el análisis, se estableció como función objetivo la maximización de biomasa. Esta función fue escogida teniendo en cuenta que la mutación que se presenta en el gen FH provoca cáncer, enfermedad que se caracteriza por una reproducción celular acelerada lo cual se traduce en un incremento de biomasa. Para evaluar las posibles implicaciones a nivel metabólico *in silico* en las células que tiene la mutación c.1349_1352delAGTA en el gen FH humano, se modificaron los límites de la reacción catalizada por la fumarato hidratasa.

Una vez establecidas las restricciones correspondientes, se realizaron cuatro experimentos distintos, cada uno sujeto a su correspondiente análisis de balance de flujo (FBA):

- Análisis de flujo de una célula normal (célula sin ninguna alteración genética, la cual fue considerada como control del estudio).
- Análisis de flujo de una célula con knockout de las reacciones que convierten fumarato en malato y malato en fumarato a nivel mitocondrial y citoplasmático.
- Análisis de flujo de una célula con knockout de la reacción que convierte fumarato en malato a nivel mitocondrial y citoplasmático.
- Análisis de flujo de una célula con knockout de la reacción que convierte malato en fumarato a nivel mitocondrial y citoplasmático.

Además, se llevó a cabo un Análisis de Variabilidad de Flujo (FVA). Estos análisis (FBA y FVA) fueron realizados con el paquete `cobrapy-constraint-based reconstruction and analysis in python` (93).

A partir de los análisis FVA, se obtuvieron los límites de flujo, inferior y superior, en los cuales ocurre cada reacción. Con el fin de identificar cuales reacciones posiblemente eran las que se veían afectadas por la ausencia de la actividad enzimática de la proteína fumarato hidratasa, se compararon los límites en los cuales se daban cada una de las reacciones y se establecieron las diferencias basados en aquellas reacciones en las cuales los rangos entre el control y cualquiera de los experimentos no se solapaban entre sí. Ya que esto quiere decir que los rangos de valores de flujo son totalmente diferentes para alcanzar el mismo valor objetivo. Para llevar a cabo este proceso, se utilizó una función booleana. Todos los análisis se llevaron a cabo comparando el modelo control (proteína fumarato hidratasa activa) contra cada uno de los modelos con la fumarato

hidratasa inactiva. Con los resultados obtenidos, se establecieron las rutas metabólicas afectadas en la presencia de la mutación de la proteína fumarato hidratasa y por medio de literatura se establecieron algunas hipótesis que explican el posible efecto de la mutación de esta proteína que llevaría al desarrollo de cáncer en los pacientes portadores de dicha mutación.

6. Resultados y discusión

6.1 Predecir y caracterizar las variantes estructurales de la proteína fumarato hidratasa a partir de la mutación c.1349_1352delAGTA en el gen FH

Al llevar a cabo la mutación c.1349_1352delATGA sobre el ORF de la proteína fumarato hidratasa y la posterior traducción, se observó que esta mutación genera un desplazamiento en el marco de lectura, seguido por un cambio del aminoácido arginina que se encuentra en la posición 450 en la secuencia de aminoácidos por un aminoácido serina y posteriormente un codón de parada prematuro. Este cambio genera el acortamiento de cada una de las subunidades de la proteína en los últimos 59 aminoácidos, lo que quiere decir que la proteína silvestre sin el péptido señal tiene una longitud de 466 aminoácidos en cada una de sus subunidades, mientras que la proteína que presenta la mutación c.1349_1352delATGA tiene un longitud de 407 aminoácidos en cada una de sus subunidades (Figura 10).

La proteína fumarato hidratasa silvestre y mutada son proteínas con un 99.51% de identidad, ya que desde el aminoácido 1 hasta el 407 no existe ningún tipo de diferencia, la diferencia radica en el final de la secuencia en la cual los últimos aminoácidos de la proteína silvestre no los tiene la proteína mutada. Además, cabe resaltar que las regiones en donde se encuentran los aminoácidos que conforman los sitios activos de la proteína no se encuentran afectados por los cambios generados por la mutación en estudio.

Hablando un poco de la estructura secundaria, se debe recordar que la proteína cuenta con 23 hélices α y 8 hojas- β (17,18) y cuando ocurre la mutación c.1349_1352delATGA se elimina parte de la hélice α 17 y la totalidad de las hélices α 18, 19, 20, 21, 22 y 23. Como se dijo anteriormente, la proteína fumarato hidratasa es un homotetrámero que está conformado por monómeros y que cada monómero está conformada por tres dominios (D1, D2 y D3) y se puede observar que en presencia de la mutación c.1349_1352delATGA la proteína pierde 8 aminoácidos del dominio D2 (parte de la α -hélice 17) y la totalidad del dominio D3 (α -hélices 18, 19, 20 21, 22 y 23) (figura 10 y 11).

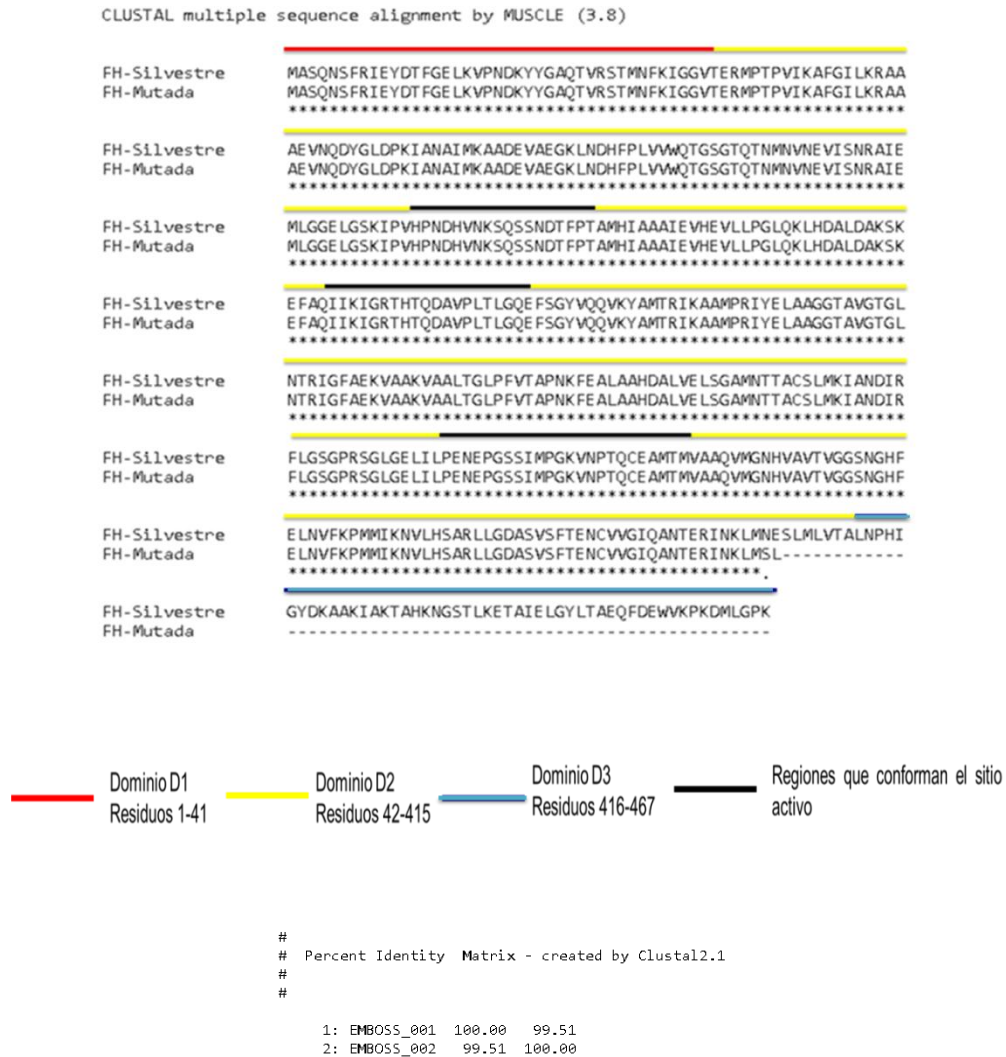


Figura 10: Alineamiento entre la secuencia de aminoácidos de la proteína fumarato hidratasa silvestre y mutada.

En la imagen se puede evidenciar cada uno de los dominios (D1: rojo, D2: amarilla y D3: azul) y las regiones conservadas que hacen parte del sitio activo. En la parte inferior se puede evidenciar que entre ambas proteínas existe un porcentaje de homología del 99.51%.

Teniendo en cuenta lo anterior y tomando en consideración que el sitio activo de la proteína no se ve afectado por la mutación c.1349_1352delATGA, surge la duda de porque la acción catalítica de la proteína se ve afectada si en realidad su sitio activo sigue intacto a pesar de la mutación.

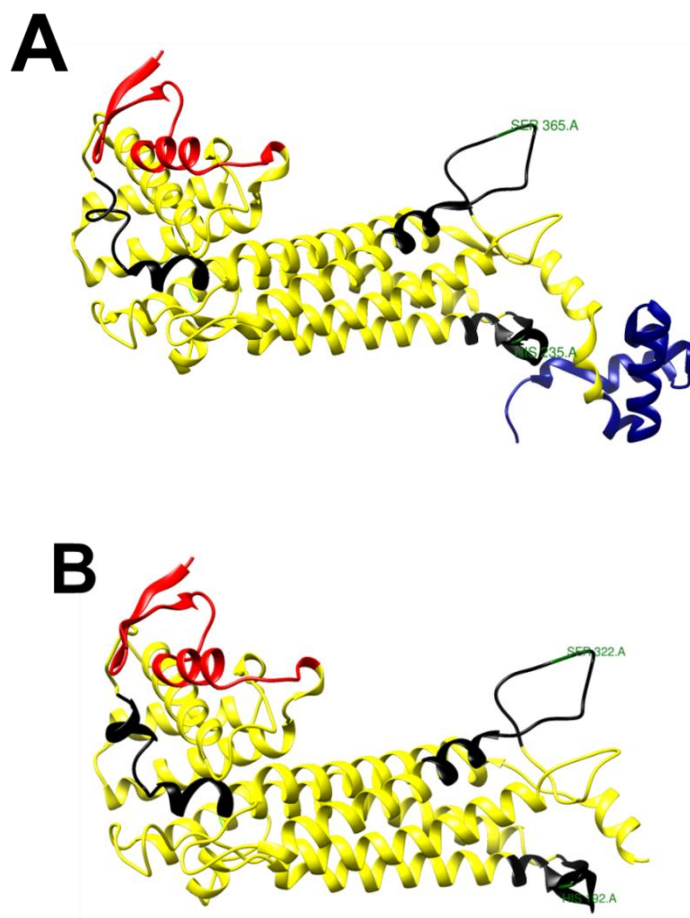


Figura 11: Monómero de la proteína Fumarato hidratasa.

A. Monómero de la proteína fumarato hidratasa humana silvestre con cada uno de los dominios diferenciados. Rojo: dominio 1, amarillo: dominio 2 y azul: dominio 3 y las regiones que conforman los sitios activos resaltadas en negro. **B.** Monómero de la proteína fumarato hidratasa humana mutada con cada uno de los dominios diferenciados. Rojo: dominio 1 y amarillo: dominio 2. También se pueden observar las regiones que conforman los sitios activos resaltadas en negro. Se puede evidenciar que la mutación en mención genera la pérdida de una pequeña porción del dominio 2 y la totalidad del dominio 3 pero en ningún momento afecta las regiones que hacen parte del sitio activo.

Algunos reportes han sugerido que la fumarato hidratasa solo es activa como un tetrámero (15,94) y teniendo en cuenta que al ocurrir la mutación c.1349_1352delATGA se elimina cierta parte del dominio D2 y que este dominio está relacionado con la oligotetramerización de la proteína, se puede inferir que esta mutación puede estar involucrada en la inhibición de la formación adecuada del tetrámero, lo que puede traducirse en una acción catalítica nula de la proteína (17,20,22).

Asimismo, según lo reportado por Weaver y colaboradores en el año 1995 (20), las hélices pertenecientes al dominio D2 forman las principales interacciones entre los monómeros de la proteína fumarato hidratasa para formar el tetrámero y cada subunidad hace contacto con subunidades vecinas a través de ese núcleo helicoidal. Sin embargo, algunas subunidades como A-C y B-D tienen muy poco contacto entre los residuos del dominio D2 y se conectan por medio de los loops en D2 y los aminoácidos que pertenecen al dominio D3 (figura 12). Por esta razón, se puede suponer que la mutación c.1349_1352delATGA puede tener influencia en la no formación del tetrámero y por consiguiente en la acción catalítica deficiente de dicha proteína, ya que según Bulku y colaboradores en el año 2018 (17) un defecto en la interacción entre las subunidades puede causar una formación no adecuada del tetrámero y mal ensamblaje del sitio activo, ya que dicho sitio está conformado por diferentes porciones de las subunidades y se forma cuando estas subunidades se unen de forma adecuada. De la misma manera, la ausencia de una porción tan grande de la proteína puede tener consecuencias estructurales que pueden dañar interacciones indispensables entre átomos que unen diferentes aminoácidos de diferentes subunidades, lo que puede afectar la tetramerización de la proteína.

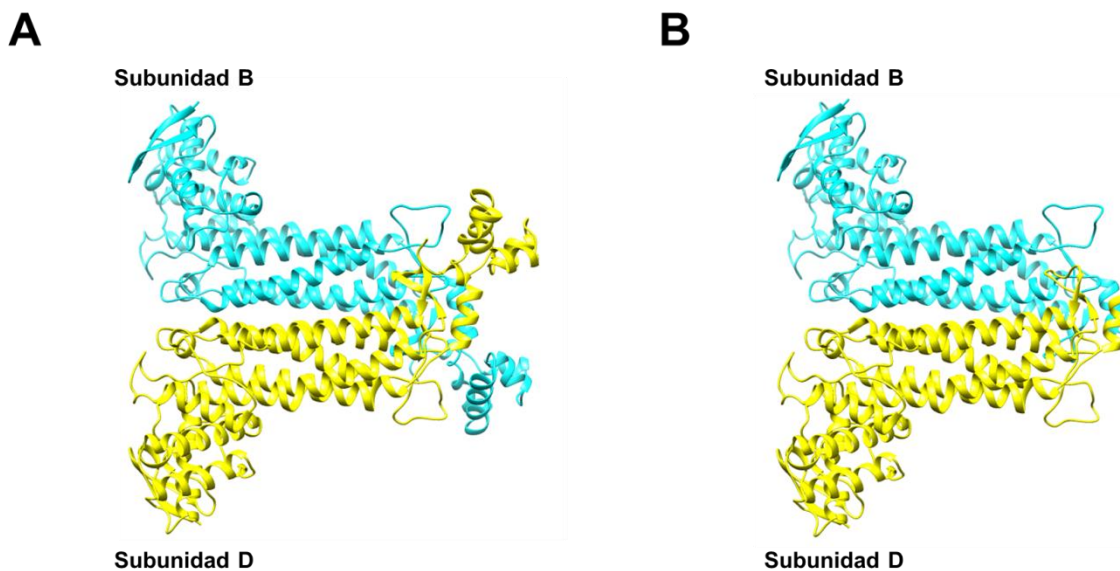


Figura 12: Dímero que contiene las subunidades B (azul) y D (amarilla).

A. Dímero de la proteína fumarato hidratasa silvestre en donde se observa el poco contacto que existe entre las subunidades usando el dominio D2. La mayor parte del contacto ocurre entre las α -hélices del dominio D3 de ambas subunidades. **B.** Dímero de la proteína fumarato hidratasa mutada en el cual se evidencia que se disminuye contacto de las subunidades por la ausencia del dominio D3 y parte del dominio D2.

Teniendo en cuenta lo anterior y al realizar el análisis de interfaz proteína-proteína mediante la herramienta ProFunc el cual brinda una visión de los diferentes tipos de enlaces que se dan entre las subunidades que forman la estructura cuaternaria de la proteína fumarato hidratasa humana, se puede observar que hay grandes diferencias entre el número de enlaces que hay entre ciertas subunidades entre la proteína mutada y silvestre (Tabla 1). Se puede observar que, entre las subunidades A-C en la proteína silvestre se establecen 25 puentes de hidrogeno, mientras que en la proteína mutada se establecen 15. Algo similar ocurre entre las subunidades B-D en las cuales se establecen 28 puentes de hidrogeno entre las subunidades de la proteína silvestre y 15 entre las subunidades de la proteína mutada.

Tabla 1: Enlaces entre la diferentes subunidades tanto de la proteína fumarato silvestre como de la mutada.

Cadena	Proteína silvestre			Proteína mutada		
	Área de interfaz (Å ²)	No. de residuos en la interface	No. de puentes de hidrogeno	Área de interfaz (Å ²)	No. de residuos en la interface	No. de puentes de hidrogeno
A	2413	54	16	2300	48	16
B	2444	52		2300	48	
A	1522	33	25	1178	22	15
C	1574	29		1184	23	
A	3099	63	24	2649	56	22
D	3054	64		2656	56	
B	3055	61	23	2649	56	22
C	3060	58		2656	56	
B	1951	38	28	1184	23	15
D	1893	37		1178	22	
C	2347	51	16	2258	49	16
D	2360	49		2258	49	

Estas diferencias se encuentran entre los números de puentes de hidrogeno y los de las subunidades A-C y B-D (Figura 12), las cuales tienen mayor interacción entre la última porción del dominio D2 y la totalidad del dominio D3. Los puentes de hidrogeno son unos de los mayores determinantes estructurales que controlan las configuraciones al conectar la estructura de la proteína en un equilibrio funcional. La creación o rotura de puentes de hidrogeno afecta profundamente las tasas y los equilibrios dinámicos, que son responsables en gran parte de la actividad biológica de las proteínas (95). En otras palabras los puentes de hidrogeno juegan un rol importante en la estabilización y la función biológica de las macromoléculas (96). Esto permite concluir que la mutación en estudio puede afectar la oligomerización de la proteína y de la misma manera afectar su acción en el ligando ya que el sitio activo se forma solo cuando los residuos de tres de las cadenas se ensamblan perfectamente.

6.1.1 Estructura cuaternaria de la proteína fumarato hidratasa

En la figura 13 se pueden observar tanto la estructura cuaternaria de la proteína fumarato hidratasa silvestre descargada del PDB (71) y la proteína fumarato hidratasa mutada la cual fue modelada por homología utilizando proteínas molde que son derivadas de alineamientos con secuencias de estructura conocida (en este caso la fumarato hidratasa silvestre 3e04) con el programa SWISS-MODEL (33,73,74). Cada una de las cadenas de un color representa una subunidad, lo que indica que esta proteína es un tetramero conformado por cuatro subunidades idénticas como se dijo anteriormente (20,22).

La validación del modelo tridimensional de la proteína fumarato hidratasa mutada fue realizado con la herramienta “structure assesment” de SWISS-MODEL mediante un plot de Ramachandran, el cual traza ángulos *phi* versus *psi* diedros para cada residuo en la proteína (75,97) y se puede usar para hacer una distribución de los pares de ángulos *phi* y *Psi* utilizando un modelo de exclusión de volumen: no se puede superponer dos átomos que no están unidos. Los resultados de esos cálculos muestran una libertad conformacional considerable en los dos ángulos de torsión, pero también un número de restricciones y estas desviaciones pueden ser útiles para medir la calidad de la estructura (98). Los sitios donde hay mayor probabilidad de que los ángulos que conforman la estructura proteica sean posibles corresponden a las regiones más favorables, mientras

que las regiones no permitidas corresponden a sitios con menor probabilidad de acercamiento al ángulo correcto. El actual análisis evidencia que el modelo cumple con los parámetros estereoquímicos de una estructura estable con el 97.12% de los residuos en las regiones más favorables y el 0% de los residuos se encuentran en las regiones no permitidas (Figura 14). Lo anterior quiere decir que el modelo obtenido constituye un buen acercamiento a la estructura cristalizada de la proteína fumarato hidratasa silvestre, lo que era de esperarse, ya que en realidad la diferencia entre las dos proteínas radica en la eliminación de los últimos 59 aminoácidos, lo que no generaría mayor cambio en su estructura tridimensional.

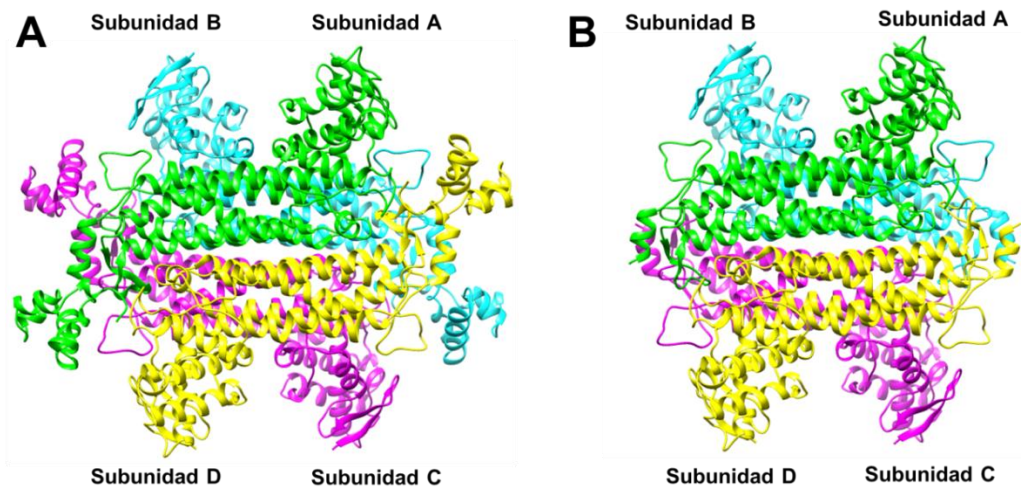


Figura 13: Proteína fumarato hidratasa humana modelada computacionalmente. **A.** Proteína fumarato hidratasa humana silvestre (GMQE: 0.98). **B.** Proteína fumarato hidratasa humana con la mutación c.1349_1352delATGA (GMQE: 0.99)".

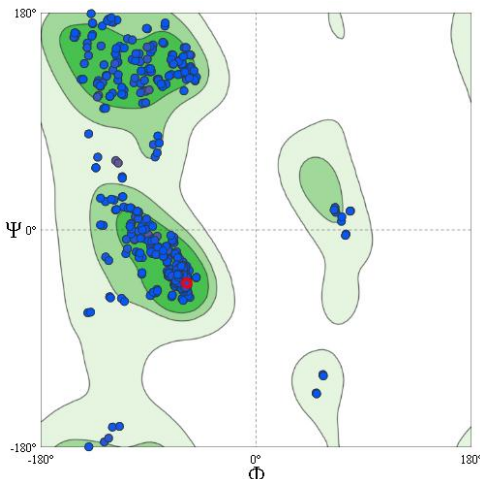


Figura 14: Plot de Ramachandran que evidencia la conformación estereoquímica de los ángulos phi y psi de la estructura tridimensional de la proteína fumarato hidratasa con la mutación c.1349_1352delATGA.

Se muestran las regiones más favorables (verde oscuro), adicionalmente permitidas (verde medio), generosamente permitidas (verde claro) y no permitidas (blanco). Los círculos del color azul indican los residuos de la proteína.

De la misma manera se puede observar que la estimación GMQE (estimación de la calidad del modelo global) es una estimación de la calidad la cual combina propiedades de la alineación de la plantilla con el blanco y el método de búsqueda de la plantilla. Este resultado es expresado de 0 a 1 lo que refleja la precisión esperada del modelo construido con el alineamiento y la plantilla y la cobertura del blanco. Los números más altos indican mayor fiabilidad (73). Con el valor obtenido de la proteína mutada (GMQE: 0.98) (figura 13) se puede asumir que el modelo obtenido es de buena calidad con respecto a su plantilla.

Al realizar el alineamiento entre la estructura tridimensional de la proteína fumarato hidratasa silvestre y la proteína fumarato hidratasa con la mutación en estudio (c.1349_1352delATGA) (figura 15), se determinó la desviación entre ambas (RMSD), encontrando un valor de 0.7814 Å. Cuanto menor sea el valor RMSD al comparar dos estructuras, mayor es la similitud estructural de las mismas, lo cual al mismo tiempo influye en la eficiencia de unión con el sustrato (35).

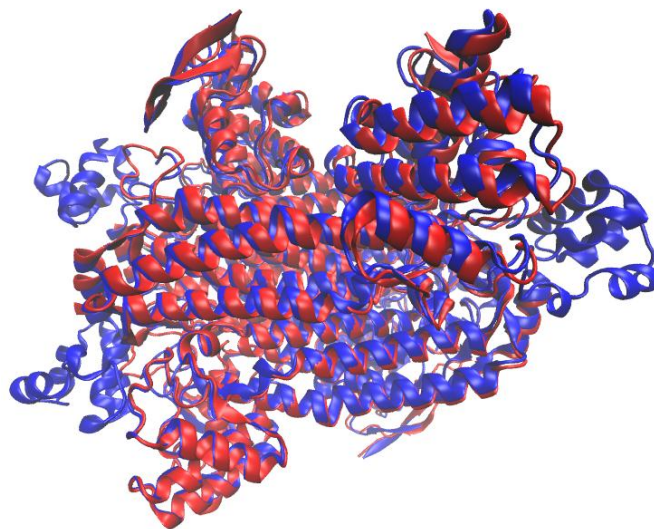


Figura 15: Alineamiento estructural de la proteína fumarato hidratasa humana silvestre y la proteína fumarato hidratasa humana con la mutación c.1349_1352delATGA. La estructura de color azul representa a la proteína fumarato hidratasa silvestre, mientras que la estructura roja representa a la proteína fumarato hidratasa mutada”. RMSD: 0.7814 Å.

Teniendo en cuenta los resultados obtenidos se puede concluir que la divergencia entre la estructura de la proteína silvestre y la mutada es poca en cuanto a su estructura tridimensional, ya que entre más se acerque el valor RMSD a cero es menor la divergencia entre ambas. Esto puede influir en la forma en la que las proteínas se ajustan a los ligandos y realizan su trabajo, lo que quiere decir que la proteína mutada según su estructura puede unirse al ligando de forma similar a como lo hace la proteína silvestre.

Al llevar a cabo la medición de SASA, se encontró que la proteína silvestre presenta un valor SASA de 52198 Å², mientras que el valor SASA obtenido para la proteína mutada fue de 48653 Å². Estos resultados eran esperados, ya que la proteína mutada tiene 59 aminoácidos menos que la proteína silvestre en cada una de las cuatro cadenas y estos aminoácidos que faltan en la proteína mutada se encuentran hacia la periferia de la estructura cuaternaria, razón por la cual el área accesible al solvente de la proteína mutada disminuye notablemente. Por otra parte, cuando analizamos los aminoácidos involucrados en la acción catalítica de la proteína podemos observar que los aminoácidos de la proteína silvestre tienen menos área de superficie accesible al solvente que los aminoácidos de la proteína mutada ya que cada una de las regiones que hacen parte del dominio 3 en cada una de las cadenas o monómero se encarga de proteger la entrada al sitio activo y hace que el solvente tenga menos acceso a estos aminoácidos (Tabla 2). Si

observamos la figura 16B, podemos darnos cuenta que en presencia de la mutación en estudio los aminoácidos que hacen parte del sitio activo quedan totalmente expuestos al solvente y por tal razón el valor SASA es mucho mayor en estos aminoácidos que en los presentes en la proteína silvestre.

Tabla 2: SASA de los aminoácidos que hacen parte del sitio activo de la proteína Fumarato Hidratasa humana silvestre y de la proteína con la mutación c.1349_1352delAGTA.

Proteína silvestre		Proteína mutada	
Cadena A		Cadena A	
Ser 186	4.822	Ser 143	16.609
Ser 187	8.027	Ser 144	57.430
Asn 188	6.872	Asn 145	7.266
Cadena B		Cadena B	
Ser 365	18.340	Ser 322	64.561
Ser 366	25.322	Ser 323	119.088
Lys 371	10.721	Lys 326	47.608
Asn 373	8.481	Asn 330	50.020
Cadena D		Cadena D	
Thr 234	0.154	Thr 191	4.732
His 235	3.279	His 192	8.481

Como se dijo anteriormente, la mutación c.1349_1352delATGA no afecta directamente a los sitios activos de la proteína ni a los residuos catalíticos. Sin embargo, según Bulku y colaboradores cada uno de los cuatro sitios activos independientes están localizados entre los dominios D1 y D3 en las cuatro esquinas de la enzima (figura 17) (17). El dominio central D2 de cada subunidad forma la interfaz principal entre las subunidades y los dominios D2 y D3 se encuentran en las esquinas del homotetrámero y forman el contorno del punto de entrada a los cuatro sitios activos de la enzima (Figura 16).

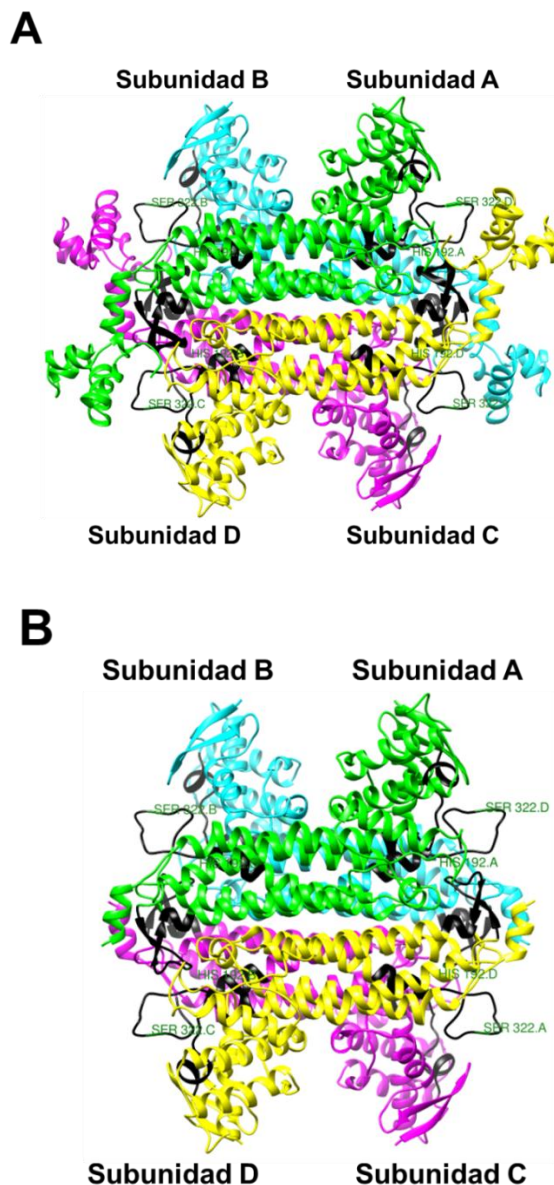


Figura 16: Homotetrámeros de la proteína fumarato hidratasa con cada una de las subunidades (Subunidad A: verde, subunidad B: azul, subunidad C: purpura y subunidad D: amarillo).

A. Homotetrámero de la proteína fumarato hidratasa silvestre con los cuatro sitios activos resaltados en negro y los residuos catalíticos Histidina 192 y serina 322 resaltados en verde oscuro. **B.** Homotetrámero de la proteína fumarato hidratasa con la mutación c.1349_1352delATGA con los cuatro sitios activos resaltados con negro y los residuos catalíticos histidina 192 y serina 322 resaltados en verde oscuro. Se puede observar que la mutación no afecta directamente a las regiones que conforman los sitios activos pero en presencia de la mutación desaparece el dominio D3 y el sitio activo se observa demasiado expuesto hacia el exterior de la proteína.

Esto quiere decir que al ocurrir la ausencia del dominio D3 en cada una de las subunidades que conforman el tetrámero se ve una posible afectación en el sitio catalítico de la misma, lo cual se evidencia por ejemplo en los valores SASA. Estos cambios pueden llevar a que la proteína no tenga buena interacción con el ligando debido a que los residuos catalíticos se pueden ver afectados o modificados por el solvente o a que la proteína cambie conformacionalmente y no se pueda formar el sitio catalítico que depende del ensamblaje perfecto de tres de las cuatro subunidades que conforman la proteína.

6.1.2 Propiedades físico-químicas de la proteína fumarato hidratasa silvestre y con la presencia de diferentes mutaciones

Se obtuvieron las diferentes mutaciones reportadas en uniProt presentes en el gen FH que provocan la enfermedad HLRCC (Tabla 3). Al realizar la comparación de las propiedades fisicoquímicas entre las proteínas con las diferentes mutaciones y la proteína silvestre, solo se encontraron diferencias estadísticamente significativas en la longitud de la secuencia de aminoácidos (p valor: 0.045) y el peso molecular de la proteína con la mutación de interés, respecto a la proteína silvestre (p valor: 2.2e-16) (Figura 17). Esto se debe a que como se dijo anteriormente, al ocurrir la mutación de interés para este estudio se genera un codón de parada temprano y no se traducen 59 aminoácidos, lo que hace que la secuencia de la proteína sea más corta y su peso molecular sea más bajo.

Tabla 3 “Variantes en la proteína fumarato hidratasa humana causantes de la enfermedad HLRCC reportadas en uniProt”.

Identificador	Posición	Descripción
VAR_013497	107-107	N→T
VAR_013498	117-117	A→P
VAR_013499	180-180	H→R
VAR_013500	185-185	Q→R
VAR_002445	230-230	K→R
VAR_013501	233-233	R→H
VAR_013502	282-282	G→V
VAR_013503	328-328	M→R

En cuanto a la inestabilidad se puede observar que la proteína que presenta la mutación en estudio es estable, ya que el valor calculado es menor a 40 (99) al igual que la proteína silvestre y las que presentan las diferentes mutaciones. En cuanto al índice alifático, que es el volumen relativo ocupado por cadenas alifáticas (alanina, valina, isoleucina y leucina), obtuvo un valor alto, incluso más alto que la proteína silvestre y las otras proteínas con mutaciones en el gen FH, lo que indica que la proteína tiene un alto contenido de cadenas alifáticas que la hacen termoestable (99,100). Por último, en cuanto a la hidropatía o valor GRAVY, se determina que la fumarato hidratasa es una proteína hidrofílica pues el valor obtenido es negativo. Sin embargo, la proteína con la mutación en cuestión se acerca al cero, por lo cual se puede inferir que esta mutación puede causar un cambio en la proteína pasando de hidrofílica a hidrofóbica, esto la puede afectar de cierta manera en su función y conformación. La presencia de grupos polares en las cadenas laterales localizadas en la superficie de las proteínas, a diferencia de los grupos hidrofóbicos que se encuentran al interior, es la favorece la solubilidad de las proteínas en el agua (99).

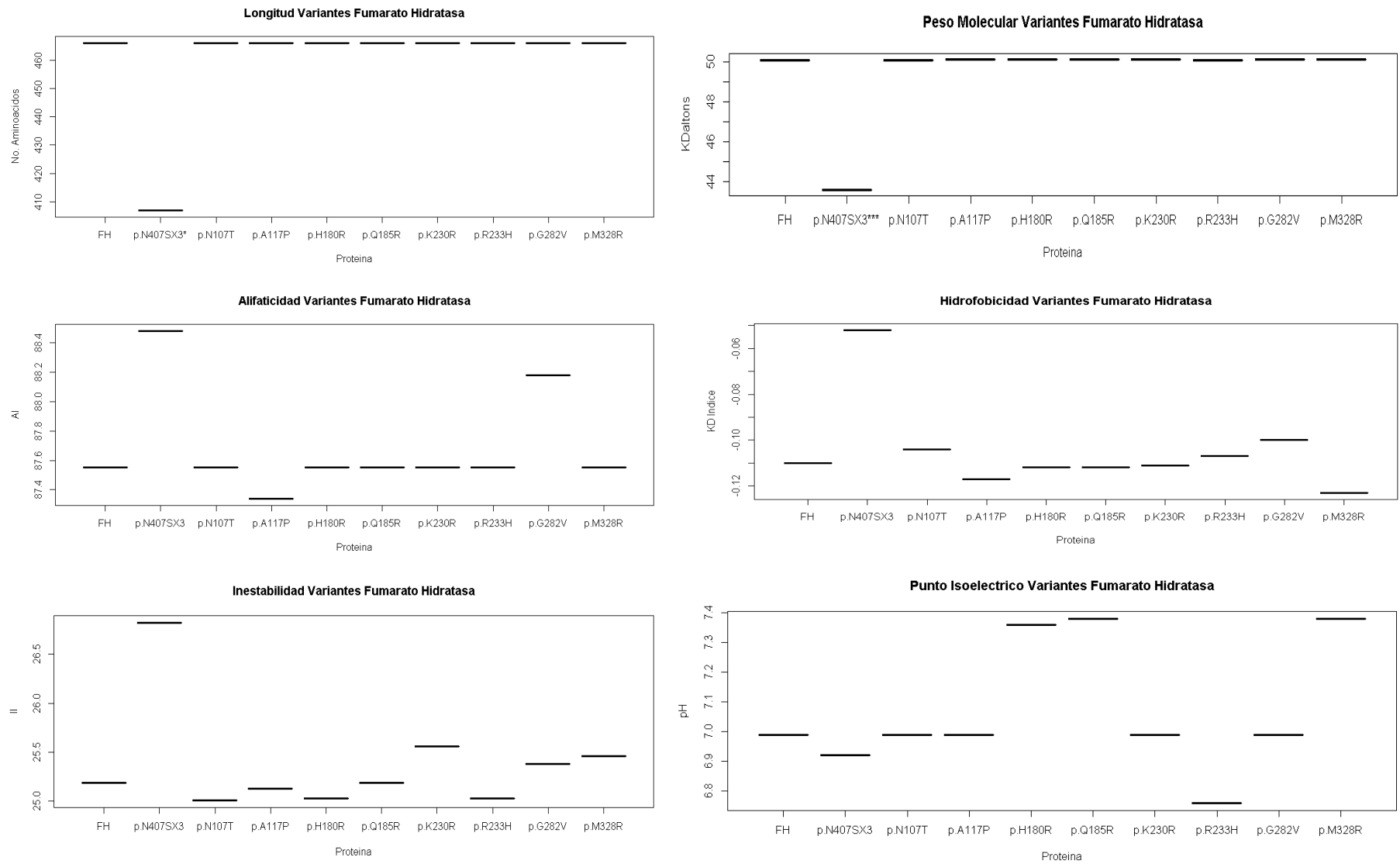


Figura 17: Propiedades fisicoquímicas de las secuencias de aminoácidos de la proteína fumarato hidratasa con mutaciones que provocan HLRCC y de la secuencia de aminoácidos de la proteína fumarato hidratasa silvestre. *p valor menor de 0.05. ***p valor menor de 0.001. FH: fumarato hidratasa silvestre.

6.2 Describir el posible efecto de las variantes halladas de la proteína fumarato hidratasa en la interacción con los ligandos fumarato y malato

Para llevar a cabo este objetivo, primero se realizó un análisis de sitio de unión con la herramienta ProFunc y específicamente un análisis de hendidura (Figura 18). Se puede observar que la proteína silvestre presenta dos hendiduras representativas (color rojo y morado) (Figura 18A), mientras que la proteína mutada presenta una gran hendidura (color rojo) (Figura 18B).

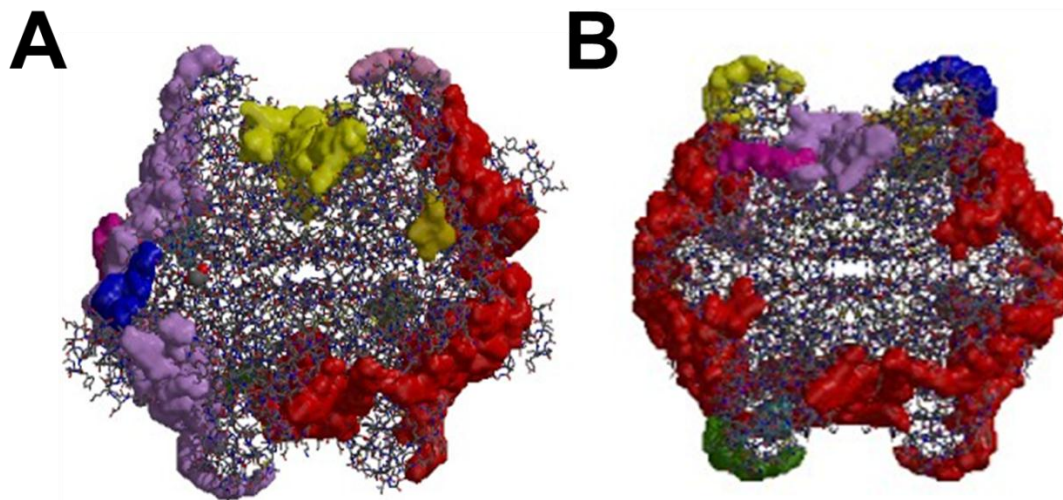


Figura 17: Representación de las hendiduras tanto en la proteína silvestre como mutada. **A.** Proteína fumarato hidratasa silvestre con la representación de las hendiduras. Hendidura mas grande de color rojo y la segunda de color morado. **B.** Proteína fumarato hidratasa con la mutacion c.1349_1352delATGA en la cual se visualiza solo una hendidura de gran tamaño de color rojo.

En la estructura de la proteína fumarato hidratasa humana silvestre se pueden observar dos hendiduras de gran tamaño (color rojo y morado) con un volumen de 21271.36 \AA^3 y 12065.62 \AA^3 , una profundidad de 26.59 \AA y 17.64 \AA . En la estructura de la proteína mutada se puede observar una hendidura de gran tamaño (color rojo) sobre la superficie de la proteína con un volumen de 33770.25 \AA^3 , una profundidad promedio de 36.03 .

Estos sitios en donde se encuentran las hendiduras más representativas coinciden con los sitios en donde se encuentran los cuatro sitios activos de la proteína, tanto de la silvestre como de la mutada. De la misma manera se puede elucidar que los dominios D3 de cada una de las subunidades que conforman la estructura cuaternaria de la proteína cumplen una función protectora del sitio activo de la misma y por esta razón la profundidad de las hendiduras en la proteína mutada es mucho mayor que en la proteína silvestre.

En cuanto a los docking realizados entre la proteína silvestre y mutada con cada uno de los ligandos, se puede observar que las energías de unión resultantes no presentan diferencias (tabla 4). Esto puede ocurrir porque el sitio activo de la proteína no está directamente afectado por la mutación en estudio, razón por la cual en el modelo computacional de la proteína mutada se encuentran los residuos de los aminoácidos disponibles para realizar el docking molecular. Éste hecho puede sugerir que la hipótesis de la no oligomerización de la proteína planteada anteriormente puede ser en realidad la causa del no funcionamiento o funcionamiento inadecuado de la proteína y la causa de la generación del síndrome HLRCC en los pacientes portadores de la mutación.

Tabla 4 “Energía de unión y constante de inhibición para los docking entre la proteína fumarato hidratasa silvestre y mutada y los ligandos fumarato y malato”.

Compuesto	Energías de unión Autodock Vina (kcal/mol)	
	FH silvestre	FH mutada
Fumarato	-4.5	-4.4
Malato	-4.8	-4.4

Las interacciones de enlaces de hidrógeno entre la proteína fumarato hidratasa silvestre con el ligando malato se establecieron entre los aminoácidos Ser145, Lys371 y Asn373, mientras que la fumarato hidratasa mutada tuvo interacciones de puentes de hidrógeno con el malato entre los aminoácidos Asn318, Glu319 y Asn330 (Figura 19).

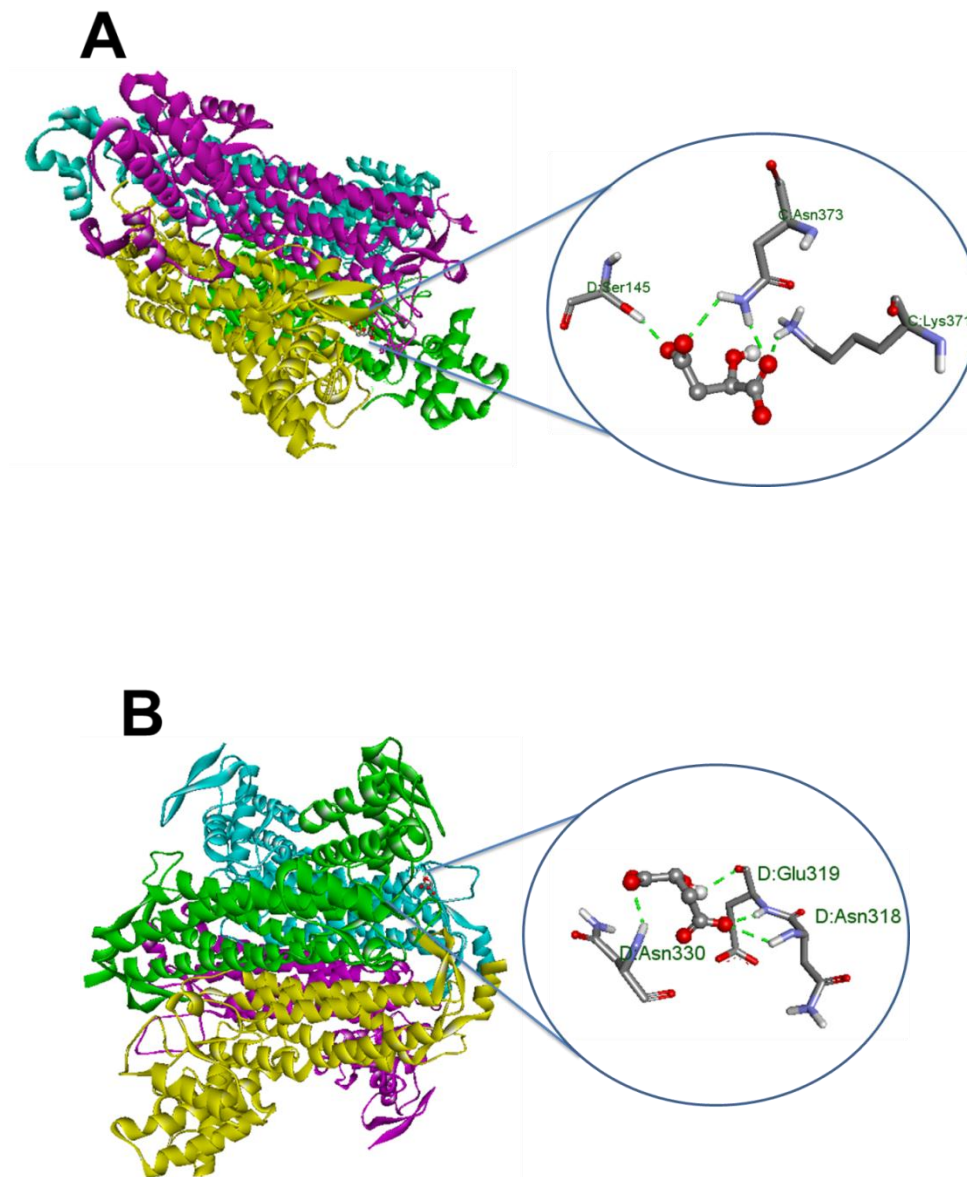


Figura 18: Docking molecular entre la proteína fumarato hidratasa (FH) silvestre y mutada y el ligando malato.

A. Proteína FH silvestre con el ligando malato, energía de unión -4.8 kcal/mol. **B.** Proteína FH mutada con el ligando malato, energía de unión -4.4 kcal/mol. El ligando malato se encuentra representado con bolas y varillas (ball & stick), los aminoácidos con varillas (sticks), los enlaces de hidrogeno están representados por líneas punteadas verdes y cada aminoácido está marcado con su símbolo y posición en la proteína.

Las interacciones de enlaces de hidrogeno entre la proteína fumarato hidratasa silvestre con el ligando fumarato se establecieron entre los aminoácidos Asn361, Glu362, Ser 365, y Asn373, mientras que la fumarato hidratasa mutada tuvo interacciones de enlaces de hidrogeno con el fumarato entre los aminoácidos Thr100, Asn318, Glu319 y Asn330 (Figura 20).

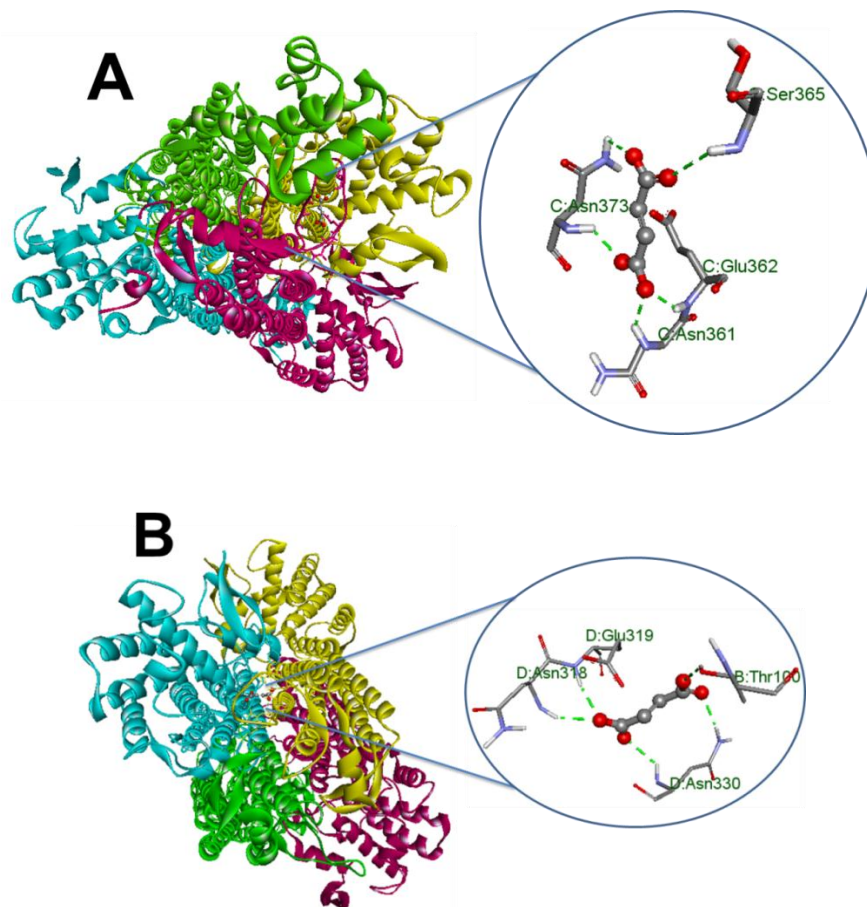


Figura 19: Docking molecular entre la proteína fumarato hidratasa (FH) silvestre y mutada y el ligando fumarato.

A. Proteína FH silvestre con el ligando fumarato, energía de unión -4.5. **B.** Proteína FH mutada con el ligando fumarato, energía de unión -4.4. El ligando fumarato se encuentra representado con bolas y varillas (ball & stick), los aminoácidos con varillas (sticks), los enlaces de hidrogeno entre el ligando y la proteína están representados por líneas punteadas verdes y cada aminoácido está marcado con su símbolo y posición en la proteína.

6.3 Determinar las posibles implicaciones moleculares de la interacción de la proteína fumarato hidratasa mutada con los ligandos fumarato y malato, en un modelo computacional del metabolismo humano.

6.3.1 Análisis FBA

Teniendo en cuenta que las energías de unión entre las simulaciones de interacción molecular no se encontraron diferencias entre las proteínas silvestre y mutada y los dos ligandos, se realizaron los análisis con el bloqueo total de las reacciones:

- Reacción que convierte malato a fumarato en mitocondria y citoplasma.
- Reacción que convierte fumarato a malato en mitocondria y citoplasma.
- Knockout total de la enzima fumarato hidratasa (ambas reacciones) en mitocondria y citoplasma.
- Célula sin ningún tipo de modificación en la proteína fumarato hidratasa como control.

Después de llevar a cabo el análisis FBA para los cuatro ensayos planteados, se encontró que al bloquear la acción de la proteína Fumarato Hidratasa en el metabolismo celular sea total o parcialmente, ocurren cambios importantes que reflejan algunas condiciones características de tejidos tumorales (tabla 5).

Tabla 5: Resultados del análisis FBA en la célula control comparada con los otros tres experimentos de la célula con la enzima Fumarato Hidratasa mutada.

Reacción	Nombre de la reacción	Célula control	Célula con reacción bloqueada F→M	Célula con reacción bloqueada M→F	Célula con enzima FH con knock-out total
o2 [e]	Oxígeno	28.2	10.2	20.9	20.9
akg[e]	Intercambio de 2-oxoglutarato	1	No	No	No
lnlnca[e]	Intercambio de ácido alfa-linoleico	1	1	No	No
strdnc[e]	Intercambio de ácido estearidónico	1	No	No	No
crvnc[e]	Intercambio de ácido docosaheptaenoico	1	No	No	No

6.3.1.1 Oxígeno

En el tejido control el cual no presenta ninguna mutación, se observa una entrada de oxígeno superior a la que se presenta en los análisis realizados con el bloqueo de la enzima Fumarato Hidratasa (tabla 5).

El oxígeno es una molécula requerida para el metabolismo aerobio, para mantener la bioenergética intracelular y sirve como aceptor de electrones en muchas reacciones orgánicas e inorgánicas. Sin embargo, cuando un tejido presenta niveles bajos de oxígeno se conoce como hipoxia, la cual ocurre en una gran cantidad de condiciones patológicas incluyendo isquemias, inflamación y el crecimiento de tumores sólidos. La hipoxia surge en los tumores a través de la proliferación incontrolada de células cancerígenas en la ausencia de una base vascular sólida. Debido a la rápida proliferación de las células cancerígenas, el tumor es incapaz de desarrollar una vasculatura adecuada provocando poca entrada de nutrientes y de oxígeno por lo tanto el tejido tumoral es normalmente un tejido hipoxico. Esa hipoxia, lleva a la producción de factores angiogénicos desde los sitios tumorales hipóxicos lo cual desencadena en la vascularización de la masa tumoral (101). Además, la exposición a hipoxia por un largo tiempo está asociado con alta frecuencia de rompimiento del ADN y acumulación de errores de replicación lo que lleva a una inestabilidad genética y mutagénesis (102).

De la misma manera, la hipoxia induce una serie de vías de señalización intracelulares complejas, como el factor inducible de hipoxia (HIF) (103), P13K/AKT/mTOR (104), vía ERK y NFκB (103), las cuales están involucradas en proliferación celular, supervivencia, apoptosis, metabolismo, migración e inflamación. Sullivan y colaboradores en el año 2013, reportaron que las células de una línea celular establecida a partir de una metástasis de un paciente con HLRCC (UOK262) eran dependientes de la glicolisis para su supervivencia y exhibían un consumo de oxígeno mitocondrial casi que indetectable (105). La señalización que mejor caracteriza el cambio en las células cancerígenas humanas deficientes de la enzima fumarato hidratasa es la inducción del estado pseudohipoxico. A pesar de la presencia amplia de oxígeno y una pérdida del consumo del mismo, las células deficientes de la enzima fumarato hidratasa tienen activación crónica del factor inducible de hipoxia (HIF). De esta manera, es probable que nuestros resultados reflejen esta condición de bajo intercambio de oxígeno desde el medio extracelular al citoplasma.

6.3.1.2 Intercambio de 2-oxoglutarato

Se observa una diferencia en el intercambio de 2-oxoglutarato, intercambio que ocurre en el control pero no ocurre en los ensayos con modificaciones en la acción de la enzima Fumarato Hidratasa. Isaacs y colaboradores en el año 2005 demostraron que un aumento modesto (2 veces) en las concentraciones de fumarato celular causado por la pérdida de expresión del gen FH es suficiente para promover la acumulación del factor inducible de hipoxia inhibiendo competitivamente el co-sustrato 2-oxoglutarato de las enzimas HIF prolin hidroxilasas, lo que demuestra una unión directa entre el aumento del fumarato y la desregulación de la inactivación del factor inducible de hipoxia (106) (Figura 21). Estos datos fueron corroborados por Koivunen y colaboradores en el año 2016, quienes propusieron al 3-oxoglutarato el cual es un análogo de 2-oxoglutarato como un potencial terapéutico del cáncer dirigido a HIF- α in vitro e in vivo (107).

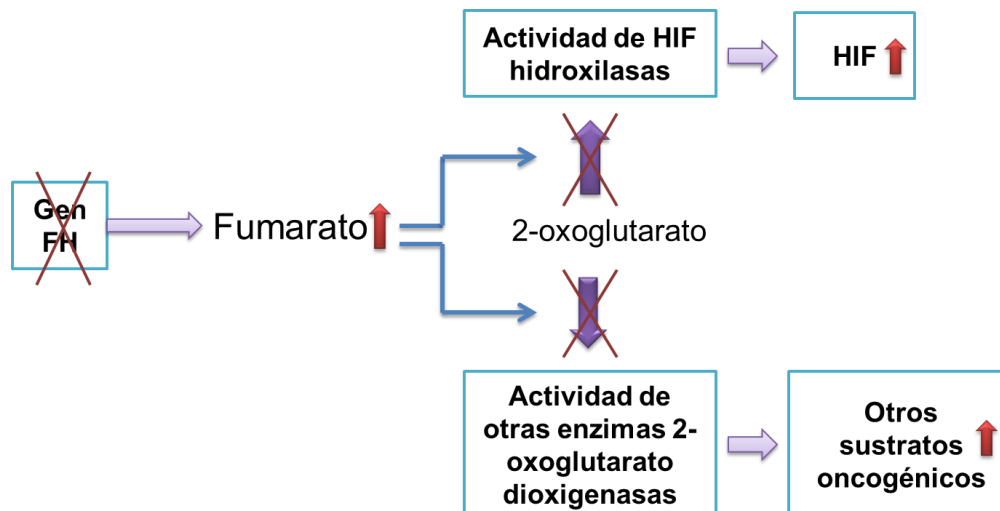


Figura 20: Efecto del aumento del metabolito fumarato por una inactivación de su gen (FH) actuando como un inhibidor del 2-oxoglutarato.

Se puede observar que en ausencia de 2-oxoglutarato, se inhibe la actividad de las enzimas HIF hidroxilasas y 2-oxoglutarato dioxigenasas y esta inactivación lleva al aumento de sustratos oncogénicos como HIF.

6.3.1.3 InInca[e], strdnc[e] y crvnc[e]

Por otro lado, las reacciones que están involucradas con el ingreso de ácidos grasos a la célula también se ven afectadas por la mutación en el gen FH. Se puede observar que en las células con bloqueo de las reacciones de la enzima fumarato hidratasa no hay entrada de estos ácidos grasos como alfa-linoleico, ácido estearidónico y ácido docosahexaenoico a diferencia de la célula control, en la cual los ácidos grasos nombrados anteriormente sí ingresan a la célula (tabla 5). Los ácidos grasos son esenciales para el buen funcionamiento corporal y su ausencia está relacionada con diferentes patologías en el ser humano, entre ellas el cáncer.

6.3.2 Análisis FVA

Al llevar a cabo los análisis de FVA y realizar el análisis booleano, se logró identificar cuales reacciones posiblemente eran las que se veían afectadas por la ausencia de la actividad enzimática de la proteína fumarato hidratasa y se llegó a la conclusión que hay siete reacciones que están posiblemente afectadas cuando la reacción que sintetiza

fumarato a partir de malato está bloqueada o cuando se presenta un knock-out total de la proteína y solo una reacción que se ve afectada cuando la actividad enzimática es deficiente al convertir fumarato en malato (tabla 6).

6.3.2.1 Biomass_reaction

Lo primero que se puede identificar es que cuando se encuentra bloqueada la reacción por la cual la enzima Fumarato Hidratasa convierte malato en fumarato o se realiza el knockout de la enzima se observa un incremento en la reacción de biomasa que fue la función objetivo planteado en este estudio.

En modelos metabólicos, la reacción de biomasa es incluida para permitir la simulación del crecimiento y la división celular a través de la producción simultánea de todos los bloques de construcción de moléculas pequeñas de biomasa. Lo anterior, quiere decir que la célula encuentra la manera de generar biomasa en mayor cantidad a pesar de la ausencia de la actividad de la enzima Fumarato Hidratasa en el ciclo del ácido tricarbónico el cual es fundamental en la generación de ATP para que la célula lleve a cabo las funciones necesarias para su supervivencia. Esto de alguna manera puede indicar que la inactivación de esta enzima lleva a que la célula obtenga energía de forma diferente y el posterior desarrollo de cáncer, ya que el cáncer se puede traducir en aumento de biomasa para la formación tumoral.

Tabla 6: Lower Bound y Upper Bound de reacciones que presentan una diferencia significativa entre el control y los ensayos con las células que tienen deficiencia en la actividad de la enzima Fumarato Hidratasa.

*Datos que en realidad presentan diferencias significativas respecto al control.

Reacción	Célula Control		Célula con reacción bloqueada F→M		Célula con reacción bloqueada M→F		Célula con enzima FH con knock-out total	
	Lb	Ub	Lb	Ub	Lb	Ub	Lb	Ub
CLS_hs	0.037282932	0.037282932	0.037282932	0.037282932	0.037282931*	0.037282931*	0.037282931*	0.037282931*
EX_na1(e)	2.02363E-11	1.02318E-12	-1.81899E-12	4.91127E-11	0*	-1.81899E-12*	0*	-1.81899E-12*
FADtm	1.06581E-14	4.61853E-14	0	9.54969E-12	0*	-9.09495E-13*	0*	-9.09495E-13*
NADK	9.09495E-13	0	-5.29243E-13	5.82645E-13	1.25056E-12*	1.31877E-11*	1.25056E-12*	1.31877E-11*
NADtm	-3.41061E-13	9.10637E-13	6.25278E-13	7.38964E-12	1.13687E-12*	2.19913E-12*	1.13687E-12*	2.19913E-12*
Biomass_reaction	3.198055582	3.198055582	3.198055582	3.198055582	3.19805558220602*	3.19805558220602*	3.19805558220602*	3.19805558220602*
R1441	0	0	1.26104E-11*	2.27374E-13**	-2.13163E-14*	-5.68434E-13*	-2.13163E-14*	-5.68434E-13*

Composición de cada una de las reacciones**CLS_hs:** [c] : cdpdag_hs + pglyc_hs --> clpn_hs + cmp + h**FADtm:** fad[c] <=> fad[m]: Transporte de FAD dentro de la mitocondria**NADK:** [c] : atp + nad --> adp + h + nadp: metabolismo de NAD**NADtm:** nad[c] <=> nad[m]: Transporte de NAD dentro de la mitocondria**R1441:** trdrd[c] -> trdrd[m]: transporte active de tioredoxina reductasa hacia el interior de la mitocondria

6.3.2.2 CLS_hs

Esta es una reacción en la cual se sintetiza cardiolípidina (CL) en el ser humano (figura 22). CL es considerado el fosfolípido más importante de la mitocondria y se encuentra casi que exclusivamente en la membrana mitocondrial interna donde es esencial para el funcionamiento óptimo de numerosas enzimas que están involucradas en el metabolismo energético mitocondrial (108) y es requerido para la estabilización de los supercomplejos de la cadena de transporte y el control respiratorio óptimo, lo que sugiere el importante rol de CL en el proceso bioenergético mitocondrial (109). Adicional a esto, tiene una habilidad única para interactuar con proteínas y tiene una función importante al mantener la fluidez y la estabilidad osmótica de la membrana interna mitocondrial. Por último, la cardiolípidina ancla electrostáticamente el citocromo c a la membrana interna mitocondrial y por tal razón puede jugar un papel crucial en la regulación de la liberación de citocromo c lo cual genera procesos apoptóticos (108).

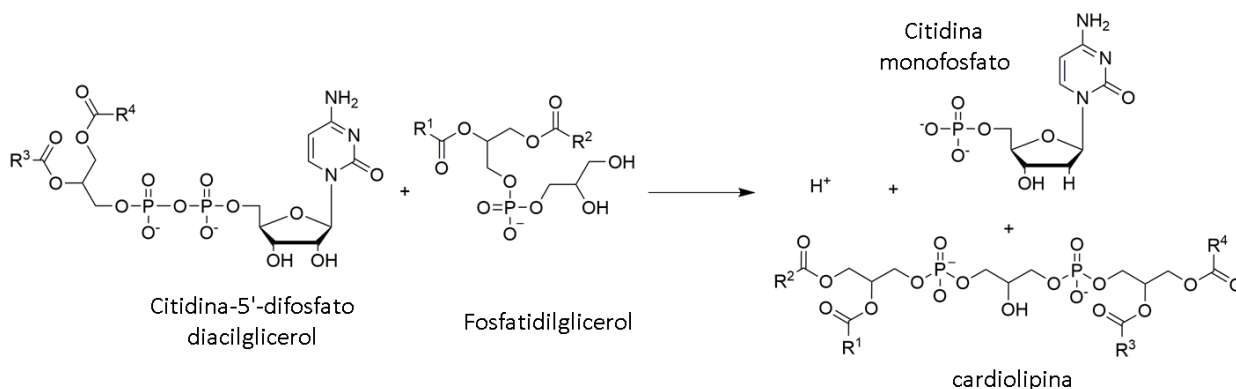


Figura 21: Reacción por la cual es sintetizada la cardiolípidina en el cuerpo humano.

La composición de la cadena del grupo acilo de CL es altamente específica, ya que está formada principalmente por cadenas de acilo insaturadas de 18 carbonos, de los cuales la mayoría es ácido linoleico en gran parte de los tejidos de mamíferos (110). En el presente estudio se observa una disminución estadísticamente significativa en el flujo de la reacción CLS_hs, lo que quiere decir que la síntesis de CL disminuye en la ausencia de la función de la enzima FH. Esto puede ocurrir porque la cardiolípidina necesita gran cantidad de ácido linoleico para poder ser sintetizada y como se vio en el análisis FBA, el ácido

linoleico no tiene flujo hacia el interior de la célula en la presencia de la mutación del gen FH, por lo cual un insumo importante para producir CL está ausente.

En la mayoría de los tejidos de mamíferos, TLCL (tetralinoleoil cardiolipina) es la principal cardiolipina, la cual está conformada por cuatro cadenas de ácido linoleico (111) y algunos estudios han indicado que la peroxidación de CL está involucrada en la apoptosis a través de la liberación de factores proapoptóticos tales como el citocromo c desde la mitocondria hacia el citosol (112). Uno de los productos de la peroxidación de CL es 4-hidroxi-2-nonenal (4-HNE) el cual puede alcanzar varias vías de señalización para provocar la muerte celular principalmente formando aductos covalentes con proteínas funcionales, ácidos nucleicos y membranas lipídicas, lo cual puede sugerir que 4-HNE derivado de la peroxidación de CL puede tener efectos profundos en la función mitocondrial y apoptosis (111).

En un estudio realizado por Zhong y colaboradores en el año 2016 (111), plantean que los lípidos de la mitocondria juegan algún papel crucial en la resistencia a la apoptosis en la células cancerígenas. En ese estudio, con el fin de investigar los cambios en el perfil de la CL durante la progresión del cáncer, se utilizaron células HCC (líneas celulares de carcinoma hepatocelular) como modelo *in vivo* y el análisis lipídico mostró un decrecimiento significativo del contenido total de CL y TLCL en tejidos tumorales comparados con los tejidos adyacentes y se hipotetiza que las células cancerígenas prefieren ácidos grasos saturados, sobre los ácidos grasos poli-insaturados como el ácido linoleico para evadir la apoptosis a través de la peroxidación de la CL en la mitocondria y al mismo tiempo establecieron que la producción de 4-HNE a través de la oxidación de TLCL decreció en los tejidos tumorales (figura 23).

Sin embargo, Kiebish y colaboradores en el año 2008 plantearon que las células cancerígenas tienen anomalías en el contenido y la composición de CL, lo cual está asociado con irregularidades en el transporte de electrones (113). Adicional a esto, anomalías en la composición y contenido de CL puede prevenir la oxidación de la coenzima Q lo que produce especies reactivas de oxígeno durante la progresión tumoral (114), esto puede desencadenar una desestabilización genómica, mal funcionamiento de los genes supresores de tumores y descontrol en la proliferación celular (115). Además, se ha encontrado que la CL interactúa con una gran cantidad de proteínas que se encuentran en la membrana interna de la mitocondria incluyendo los complejos de la

cadena de transporte de electrones (116) y afectando la fosforilación oxidativa. Teniendo en cuenta esto y la hipótesis postulada por Warburg, en donde plantea que hay una disfunción mitocondrial en células cancerígenas, se puede asumir que las alteraciones de las funciones mitocondriales mediadas por CL pueden ser esenciales para la generación de tumores y la progresión tumoral.

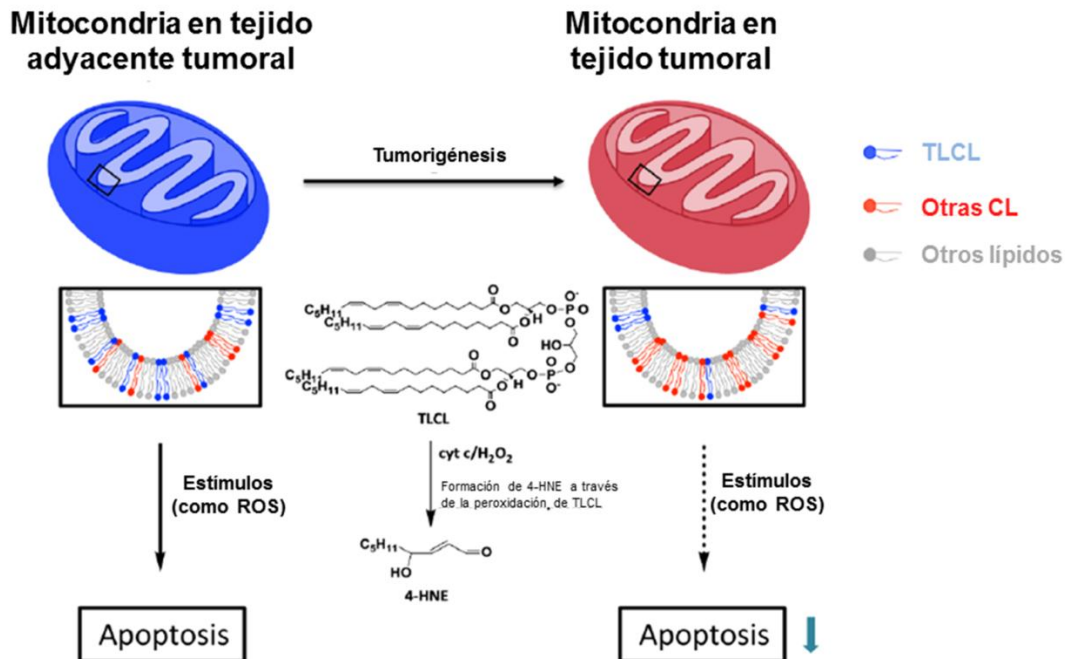


Figura 22: Las células tumorales incorporan menos ácidos grasos poliinsaturados, como el ácido linoleico dentro de su grupo de CL mitocondriales para evadir la muerte apoptótica a través de la regulación a la baja de la oxidación de la CL y la subsecuente formación de 4-HNE. Tomado y modificado de Zhong y colaboradores 2016 (111).

6.3.2.3 R1441

Esta reacción está relacionada con el transporte activo de la proteína tiorredoxina reductasa desde el citoplasma hacia la mitocondria. En los resultados obtenidos en el presente estudio, se puede observar que cuando la enzima FH está inhabilitada para convertir malato a fumarato o se encuentra knockout total de la enzima, aumenta el flujo de tiorredoxina reducida desde el citoplasma hacia la mitocondria como se puede observar en la tabla 6.

La enzima Tiorredoxina (Trx) funciona como un regulador clave en el proceso redox asociado al estrés oxidativo. El rol fundamental del sistema redox es proveer un ambiente reducido y proteger la célula de diferentes daños provocados por los efectos del estrés oxidativo que en última instancia conduce a la apoptosis. Por lo tanto, cualquier cambio importante en la actividad de Trx y la actividad de las enzimas reductoras que dependen de Trx para regenerar sus capacidades reductoras eventualmente conducirán a una mayor cantidad de proteínas oxidadas y nitrosiladas con el subsiguiente mal funcionamiento de la proteína. Trx interactúa con moléculas de señalización sensibles a redox. Hay un gran número de factores de transcripción que son reguladas por redox y contienen cisteínas sensibles a redox en su dominio de unión de ADN. Estas incluyen NF- κ B, proteína 21 (p21), proteína 53 (p53), el factor inducible de hipoxia 1-alfa (HIF-1 α), el factor nuclear eritroide-2 (Nrf2) y la quinasa 1 reguladora de la señal de apoptosis (ASK1). La tiorredoxina oxidada con un enlace disulfuro intramolecular entre C32 y C35 está inactiva y es incapaz de unirse a ASK1. Sin embargo, Trx reducida puede unirse a ASK1. Al unirse, Trx inactiva a ASK1 al dirigirlo hacia ubiquitinación y degradación. La degradación de ASK1 resulta en la disminución de la cascada de señalización de apoptosis con las proteínas quinasas c-Jun N-terminal (JNK) y la proteína 38 (p38), lo que inhibe la apoptosis (117).

En un estudio realizado por Mahmood y colaboradores en el año 2013, se reporta que Trx-2 (proteína Tiorredoxina mitocondrial) es codificada por un gen nuclear (TRX2) con un péptido señal de dirección mitocondrial. Esta proteína es expresada en las células de todos los órganos y tejidos y la sobreexpresión de Trx-2 puede facilitar el desarrollo de resistencia a la apoptosis inducida por ROS. Entonces, el rol dual de la tiorredoxina reductasa en cáncer, prevención/promoción puede depender del estado del desarrollo del cáncer (118).

Todo lo anterior, indica que la inhibición de la reacción malato \rightarrow fumarato por parte de la enzima FH, lleva a un aumento en el flujo de Trx reducida desde el citoplasma hasta la mitocondria y la enzima Trx reducida da como resultado una inhibición de la apoptosis, una característica muy importante en el desarrollo y progresión del cáncer.

7. Conclusiones y recomendaciones

7.1 Conclusiones

Nuestros resultados sugieren que la mutación c.1349_1352delATGA genera un cambio de una arginina por una serina, lo que provoca un codón de parada prematuro. Este cambio tiene como consecuencia el acortamiento de cada subunidad de la proteína en los últimos 59 aminoácidos, lo que quiere decir que cada subunidad de la proteína silvestre tiene una longitud de 466 aminoácidos, mientras que cada subunidad de la proteína que presenta la mutación tiene una longitud de 407 aminoácidos.

Al realizar los análisis estructurales se evidencia que la mutación c.1349_1352delATGA no tiene ninguna afectación sobre el sitio activo de la proteína y que su mal funcionamiento puede ser causado por un ensamblaje inadecuado de las cuatro subunidades, esto producto de la delección de una porción del dominio D2 y la totalidad del dominio D3 como se evidencia en los análisis de la interacción entre las subunidades. Además, también se puede inferir que la ausencia del dominio D3 puede llevar a la exposición al solvente de los residuos que conforman el sitio activo y de esta manera afectar su conformación y su unión con el ligando.

Al realizar el docking molecular, se observó que no hay diferencias al comparar las energías de unión que se generan entre las proteínas fumarato hidratasa silvestre y mutada con cada uno de los ligandos, esto se debe posiblemente a que los sitios activos de la proteína mutada no se ven afectados en ningún momento por la mutación en estudio y al realizar el modelamiento de la estructura tridimensional de la proteína mutada los sitios activos quedan intactos. Por este motivo, la proteína mutada puede interactuar de forma normal con el ligando cuando se realizan los análisis de docking molecular. Sin embargo, esto puede darnos indicios de que la ausencia de la acción catalítica de esta proteína en los pacientes con HLRCC provocada por la mutación c.1349_1352delATGA puede ser causada por una oligotetramerización inadecuada de la proteína y la formación de un sitio activo que no es funcional.

En cuanto a los resultados encontrados al realizar la simulación en el metabolismo celular humano, se observó que la proteína fumarato hidratasa mutada cuando es incapaz de convertir malato en fumarato presenta las mayores afectaciones en el metabolismo de la

celula. Inicialmente se observa que la funcion objetivo que en este caso es biomasa se ve aumentada significativamente, lo que se puede traducir en crecimiento celular acelerado, una caracterisitca visible en procesos cancerigenos. Uno de los hallazgos mas importantes es que se presenta una disminucion en el flujo de la reacción CLS_hs, la cual da como resultado final la síntesis de la enzima cardiolipina, el cual es el fosfolípido más importante de la membrana mitocondrial y una disminución de esta enzima se traduce en un funcionamiento inadecuado del mecanismo de apoptosis ante estímulos. .

También se puede observar un cambio en la reacción r441, la cual se basa en el la translocación de la enzima tiorredoxina reducida desde el citoplasma hacia la mitocondria. Esta es una enzima que está encargada de la regulación del proceso redox asociado al estrés oxidativo y su función es prevenir a la célula de diferentes daños causados por este ambiente. Sin embargo, cuando se presenta sobre-expresión de esta enzima, la misma es capaz de unirse a la enzima ASK1 que la dirige hacia ubiquitinación y degradación. La degradación de ASK1 resulta en la disminución de la cascada de señalización de apoptosis con las proteínas quinasas c-Jun N-terminal (JNK) y la proteína 38 (p38), lo que inhibe la apoptosis, característica fundamental en el desarrollo tumoral. Entonces los hechos de que la cardiolipina produzca ROS en exceso y que se sobre-exprese la tiorredoxina reducida y se dirija hacia la mitocondria pueden estar unidos y ser importantes en el desarrollo del cáncer en los pacientes que presentan mutaciones en el gen FH.

7.2 Recomendaciones

Se recomienda llevar a cabo los ensayos experimentales en laboratorio húmedo relacionados con la acción del fosfolípido cardiolipina y la enzima tiorredoxina reductasa en la enfermedad HLRCC, y de esta manera verificar si las mismas están relacionadas con el desarrollo de cáncer en células renales y si estos mecanismos llevan a la progresión acelerada de este cáncer por su influencia sobre el mecanismo de apoptosis. .

8. Consideraciones éticas

El presente proyecto de investigación tiene en cuenta las consideraciones de varios estamentos internacionales y nacionales. Entre ellas se encuentra la Declaración de Helsinki, la cual declara las pautas para la realización de estudios que están relacionados con seres humanos y resalta los derechos de los sujetos de investigación, derechos que están por encima de la ciencia y la sociedad (Asociación médica mundial, 2008).

Respecto a las normas establecidas por el estado colombiano, el presente proyecto está dentro de los parámetros establecidos en la resolución 8430 de 1993, por la cual se establecen las normas para realizar investigaciones en salud en el país (Republica de Colombia, 1993).

Teniendo en cuenta lo anterior y de acuerdo con lo expuesto en la resolución 8430, el presente es un estudio sin riesgo para los pacientes que van a participar en el mismo. Se definen estudios sin riesgo como aquellos que emplean técnicas y métodos de investigación documental retrospectivos y aquellos en los que no se realiza ninguna intervención o modificación intencionada de las variables biológicas, fisiológicas, psicológicas o sociales de los individuos que participan en el estudio, entre los que se consideran: revisión de historias clínicas, entrevistas, cuestionarios y otros en los que no se le identifique ni se traten aspectos sensitivos de su conducta (Republica de Colombia, 1993).

Se tuvieron en cuenta otras consideraciones éticas para la realización del estudio:

- En el estudio realizado por Arenas y colaboradores en el cual se estudió la mutación, se obtuvo un consentimiento informado, dando a conocer de una manera comprensible los objetivos, duración, diseño metodológico, riesgos y beneficios del estudio y a cada uno de los participantes se le informó acerca de sus derechos como sujeto de investigación.
- La intimidad, anonimato y datos de los sujetos que participen en el estudio serán protegidas.

Los investigadores declaran que no tienen conflictos de interés para la realización del presente trabajo de investigación.

9. Bibliografía

1. Arenas Valencia C, Rodríguez López ML, Cardona Barreto AY, Garavito Rodríguez E, Arteaga Díaz CE. Hereditary leiomyomatosis and renal cell cancer syndrome: identification and clinical characterization of a novel mutation in the FH gene in a Colombian family. *Fam Cancer*. 2017 ;16(1): 117–22.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018; 68(6): 394–424.
3. Schmidt LS, Marston Linehan W. Hereditary leiomyomatosis and renal cell carcinoma. *Int J Nephrol Renovasc Dis*. 2014; 7: 253–60.
4. Launonen V, Vierimaa O, Kiuru M, Isola J, Roth S, Pukkala E, et al. Inherited susceptibility to uterine leiomyomas and renal cell cancer. *Proc Natl Acad Sci*. 2001; 98(6): 3387–92.
5. Linehan WM, Rouault TA. Molecular pathways: Fumarate Hydratase-deficient kidney cancer - Targeting the warburg effect in cancer. *Clin Cancer Res*. 2013; 19(13): 3345–52.
6. Wei MH, Toure O, Glenn GM, Pithukpakorn M, Neckers L, Stolle C, et al. Novel mutations in FH and expansion of the spectrum of phenotypes expressed in families with hereditary leiomyomatosis and renal cell cancer. *J Med Genet*. 2006; 43(1): 18–27.
7. Toro JR, Nickerson ML, Wei M-H, Warren MB, Glenn GM, Turner ML, et al. Mutations in the Fumarate Hydratase Gene Cause Hereditary Leiomyomatosis and Renal Cell Cancer in Families in North America. *Am J Hum Genet*. 2003; 73(1): 95–106.
8. Bardella C, El-Bahrawy M, Frizzell N, Adam J, Ternette N, Hatipoglu E, et al. Aberrant succination of proteins in fumarate hydratase-deficient mice and HLRCC patients is a robust biomarker of mutation status. *J Pathol*. 2011; 225(1): 4–11.
9. Grubb RL, Franks ME, Toro J, Middleton L, Choyke L, Fowler S, et al. Hereditary Leiomyomatosis and Renal Cell Cancer: A Syndrome Associated With an Aggressive Form of Inherited Renal Cancer. *J Urol*. 2007; 177(6): 2074–80.
10. Tomlinson IPM, Alam NA, Rowan AJ, Barclay E, Jaeger EEM, Kelsell D, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin

- leiomyomata and papillary renal cell cancer the multiple leiomyoma consortium. *Nat Genet.* 2002; 30(4): 406–10.
11. Hayman S, Alberty RA. THE ISOLATION AND KINETICS OF TWO FORMS OF FUMARASE FROM TORULA YEAST*. *Ann N Y Acad Sci.* 2006; 94(3): 812–24.
 12. Picaud S, Kavanagh KL, Yue WW, Lee WH, Muller-Knapp S, Gileadi O, et al. Structural basis of fumarate hydratase deficiency. *J Inher Metab Dis [Internet].* 2011; 34(3): 671–6.
 13. Alam NA, Olpin S, Rowan A, Kelsell D, Leigh IM, Tomlinson IPM, et al. Missense mutations in fumarate hydratase in multiple cutaneous and uterine leiomyomatosis and renal cell cancer. *J Mol Diagnostics.* 2005; 7(4): 437–43.
 14. Azem A, Herrmann J, Pines O, Leshets M, Silas YBH, Lehming N. Fumarase: From the TCA Cycle to DNA Damage Response and Tumor Suppression. *Front Mol Biosci | www.frontiersin.org .* 2018; 1:68.
 15. Lorenzato A, Olivero M, Perro M, Brière JJ, Rustin P, Di Renzo MF. A cancer-predisposing “hot spot” mutation of the fumarase gene creates a dominant negative protein. *Int J Cancer.* 2008; 122(4): 947–51.
 16. Puthan Veetil V, Fibriansah G, Raj H, Thunnissen AMWH, Poelarends GJ. Aspartase/fumarase superfamily: A common catalytic strategy involving general base-catalyzed formation of a highly stabilized aci - Carboxylate intermediate. *Biochemistry.* 2012; 51(21): 4237–43.
 17. Bulku A, Weaver TM, Berkmen MB. Biochemical Characterization of Two Clinically-Relevant Human Fumarase Variants Defective for Oligomerization. *Open Biochem J.* 2018; 12(1): 1–15.
 18. Pereira de Pádua RA, Nonato MC. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of recombinant human fumarase. *Acta Crystallogr Sect F Struct Biol Commun.* 2014; 70(1): 120–2.
 19. Hill RL, Teipel JW. 19 Fumarase and Crotonase. *Enzymes.* 1971; 5(C): 539–71.
 20. Weaver TM, Levitt DG, Donnelly MI, Wilkens Stevens PP, Banaszak LJ. The multisubunit active site of fumarase c from escherichia coli. *Nat Struct Biol.* 1995; 2(8): 654–62.
 21. Voet D, Voet JG, Pratt CW. *Fundamentals of Biochemistry: Life at the Molecular Level.* 2012. 1200 p.
 22. Weaver T, Banaszak L. Crystallographic studies of the catalytic and a second site in fumarase C from Escherichia coli. *Biochemistry.* 1996; 35(44): 13955–65.

23. Alam NA, Rowan AJ, Wortham NC, Pollard PJ, Mitchell M, Tyrer JP, et al. Genetic and functional analyses of FH mutations in multiple cutaneous and uterine leiomyomatosis, hereditary leiomyomatosis and renal cancer, and fumarate hydratase deficiency. *Hum Mol Genet.* 2003; 12(11): 1241–52.
24. Yogev O, Yogev O, Singer E, Shaulian E, Goldberg M, Fox TD, et al. Fumarase: A mitochondrial metabolic enzyme and a cytosolic/nuclear component of the dna damage response. *PLoS Biol.* 2010; 8(3).
25. Ooi A, Furge KA. Fumarate hydratase inactivation in renal tumors: HIF1??, NRF2 and “cryptic targets” of transcription factors. *Chin J Cancer.* 2012; 31(9): 413–20.
26. Pollard PJ, Brière JJ, Alam NA, Barwell J, Barclay E, Wortham NC, et al. Accumulation of Krebs cycle intermediates and over-expression of HIF1 α in tumours which result from germline FH and SDH mutations. *Hum Mol Genet.* 2005; 14(15): 2231–9.
27. Raimundo N, Baysal BE, Shadel GS, Raimundo. Revisiting the TCA cycle: Signaling to tumor formation. *Trends Mol Med.* 2012; 17(11): 641–9.
28. Malhotra D, Portales-Casamar E, Singh A, Srivastava S, Arenillas D, Happel C, et al. Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through chip-seq profiling and network analysis. *Nucleic Acids Res.* 2010; 38(17): 5718–34.
29. Mitsuishi Y, Motohashi H, Yamamoto M. The Keap1–Nrf2 system in cancers: stress response and anabolic metabolism. *Front Oncol.* 2012; 2:1–13.
30. Ooi A, Wong JC, Petillo D, Roossien D, Perrier-Trudova V, Whitten D, et al. An Antioxidant Response Phenotype Shared between Hereditary and Sporadic Type 2 Papillary Renal Cell Carcinoma. *Cancer Cell.* 2011; 20(4): 511–23.
31. Reyes C, Karamurzin Y, Frizzell N, Garg K, Nonaka D, Chen Y, et al. *HHS Public Access.* 2015; 27(7): 1020–7.
32. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* 2003; 31(13): 3381–5.
33. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics.* 2006; 22(2): 195–201.
34. Edwards YJK. *Bioinformatics and Functional Genomics. Briefings Funct Genomics Proteomics.* 2004; 3(2): 187–90.
35. Mas Benavente JM, Manuel J. Comparación computacional de estructuras de

- proteínas aplicación al estudio de un inhibidor de carboxipeptidasa como agente antitumoral. Universitat Autònoma de Barcelona; 2001.
36. Lee B, Richards FM. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol.* 1971; 55(3): 379-404.
 37. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. *Science* [Internet]. 1983; 221(4612): 709-13.
 38. Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. 2012; 26(9): 1169-75.
 39. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr Comput Aided-Drug Des.* 2011; 7(2): 146-57.
 40. Li X, Li Y, Cheng T, Liu Z, Wang R. Evaluation of the performance of four molecular docking programs on a diverse set of protein-ligand complexes. *J Comput Chem* [Internet]. 2010; 31(11): 2109-25.
 41. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch Chem Gesellschaft.* 1894; 27(3): 2985-93.
 42. Koshland DE. Correlation of Structure and Function in Enzyme Action. *Science* (80-). 1963; 142(3599):1533-41.
 43. Singh RS. Darwin to DNA, molecules to morphology: the end of classical population genetics and the road ahead. *Genome.* 2003; 46(6): 938-42.
 44. Westerhoff H V., Palsson BO. The evolution of molecular biology into systems biology. *Nat Biotechnol.* 2004; 22(10): 1249-52.
 45. Yang Y, Lane AN, Ricketts CJ, Sourbier C, Wei MH, Shuch B, et al. Metabolic Reprogramming for Producing Energy and Reducing Power in Fumarate Hydratase Null Cells from Hereditary Leiomyomatosis Renal Cell Carcinoma. *PLoS One.* 2013; 8(8): 1-14.
 46. Oberhardt M a, Palsson BØ, Papin J a. Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* [Internet]. 2009; 5(320): 320.
 47. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc.* 2007; 2(3): 727-38.
 48. Copeland WB, Bartley BA, Chandran D, Galdzicki M, Kim KH, Sleight SC, et al. Computational tools for metabolic engineering. *Metab Eng.* 2012; 14(3): 270-80.
 49. Borodina I, Nielsen J. From genomes to in silico cells via metabolic networks. *Curr*

- Opin Biotechnol. 2005; 16(3 SPEC. ISS.): 350–5.
50. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. *Nat Biotechnol.* 2013; 31(5): 419–25.
 51. Bordbar A, Palsson BO. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. *J Intern Med.* 2012; 271(2): 131–41.
 52. Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, Jamshidi N. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol.* 2010; 6: 422.
 53. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* [Internet]. 2016; 35(1): 81–9.
 54. Dé F, Capel R, Klimčá E, Viguerie N, Roussel B, Vítková M, et al. Macrophages and Adipocytes in Human Obesity Adipose Tissue Gene Expression and Insulin Sensitivity During Calorie Restriction and Weight Stabilization. 2009 [cited 2018 Nov 19]; Available from: <http://diabetes.diabetesjournals.org>
 55. Zelezniak A, Pers TH, Soares SO, Patti ME, Patil KR. Metabolic Network Topology Reveals Transcriptional Regulatory Signatures of Type 2 Diabetes. *PLoS Comput Biol.* 2010; 6(4):1000729.
 56. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 2016; 44(D1) :D457–62.
 57. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 1999; 27(1): 29–34.
 58. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003; 31(13): 3784–8.
 59. Karlsson FH, Nookaew I, Petranovic D, Nielsen J. Prospects for systems biology and modeling of the gut microbiome. *Trends Biotechnol.* 2011; 29(6): 251–8.
 60. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotechnol.* 2010; 28(3): 245–8.
 61. Edwards JS, Covert M, Palsson B. Metabolic modelling of microbes: The flux-balance approach. *Environ Microbiol.* 2002; 4(3): 133–40.
 62. Gudmundsson S, Thiele I. Computationally efficient flux variability analysis. *BMC*

- Bioinformatics. 2010; 11: 489.
63. Müller AC, Bockmayr A. Flux modules in metabolic networks. *J Math Biol* . 2014; 69(5): 1151–79.
 64. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng*. 2003; 5(4): 264–76.
 65. Brien O', Monk EJ, Palsson JM, Bernhard, Palsson &. Using Genome-scale Models to Predict Biological Capabilities. *Cell*. 2018; 161(5): 971–87.
 66. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. *Nucleic Acids Res*. 2015; 43(D1): D204–12.
 67. Pithukpakorn M, Toro JR. Hereditary Leiomyomatosis and Renal Cell Cancer [Internet]. GeneReviews®. University of Washington, Seattle; 1993.
 68. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, De Castro E, et al. ExpASY: SIB bioinformatics resource portal. *Nucleic Acids Res*. 2012; 40(W1): 597–603.
 69. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32(5): 1792–7.
 70. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* [Internet]. 2004; 25(13): 1605–12.
 71. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res*. 2000; 28(1): 235–42.
 72. Kavanagh K., Picaud S, Pilka E., Yue W., Roos A., Filippakopoulos P, et al. Crystal structure of human fumarate hydratase. *To Be Publ*. 2011; 671–6.
 73. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014; 42(W1): 252–8.
 74. Kiefer F, Arnold K, K??nzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res*. 2009; 37(SUPPL. 1): 387–92.
 75. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol* [Internet]. 1963; 7(1): 95–9.
 76. Roberts E, Eargle J, Wright D, Luthey-Schulten Z. MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. 2006; 7: 1–11.
 77. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*

- [Internet]. 1996; 14(1): 33–8.
78. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* 2018; 27(1): 14–25.
 79. Center for Informational Biology OU. Accessible Surface Area and Accessibility Calculation for Protein. 2012.
 80. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. *Nucleic Acids Res.* 2015; 43(D1): D204–12.
 81. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez J ???C, et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis.* 1993; 14(1): 1023–31.
 82. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem.* 1980; 88(6): 1895–8.
 83. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982; 157(1): 105–32.
 84. 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): 155–61.
 85. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vol. 1, R Foundation for Statistical Computing. 2017; 1-2630 p.
 86. Laskowski RA, Watson JD, Thornton JM. ProFunc: a server for predicting protein function from 3D structure.
 87. Laskowski RA. SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph.* 1995; 13(5): 323–30.
 88. Schrödinger L. : The PyMOL Molecular Graphics System Schrödinger. 2003.
 89. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A free tool to discover chemistry for biology. *J Chem Inf Model.* 2012; 52(7): 1757–68.
 90. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2009 .
 91. Subasri S, Chaudhary SK, Sekar K, Keshewani M, Velmurugan D. Molecular docking and molecular dynamics simulations of fumarate hydratase and its mutant H235N complexed with pyromellitic acid and citrate. *J Bioinform Comput Biol.* 2017; 15(06): 1750026.

92. Dassault Systèmes BIOVIA. Discovery Studio Visualizar. San Diego: Dassault Systèmes; 2016.
93. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: CONstraints-Based Reconstruction and Analysis for Python. *BMC Syst Biol.* 2013; 7.
94. HILL RL, KANAREK L. THE SUBUNITS OF FUMARASE. *Brookhaven Symp Biol.* 1964; 17: 80–97.
95. Sheu S-Y, Yang D-Y, Selzle HL, Schlag EW. Energetics of hydrogen bonds in peptides. 2003.
96. Jiang L, Lai L. CHO Hydrogen Bonds at Protein-Protein Interfaces* □ S Downloaded from. *J Biol Chem.* 2002; 277(40): 37732–40.
97. Saddala MS, Kandimalla R, Adi PJ, Bhashyam SS, Asupatri UR. Novel 1, 4-dihydropyridines for L-type calcium channel as antagonists for cadmium toxicity. *Sci Rep.* 2017; 7(1): 45211.
98. Hooft RWW, Sander C, Vriend G. Objectively judging the quality of a protein structure from a Ramachandran plot. Vol. 13, *CABIOS.* 1997.
99. Gasteiger E, Hoogland C, Gattiker a, Duvaud S, Wilkins M, Appel R, et al. Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protoc Handb.* 2005; 571–607.
100. Cozzzone AJ. *Proteins : Fundamental Chemical Properties.* Life Sci. 2002; 1–10.
101. Eales KL, Hollinshead KER, Tennant DA. Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis.* 2016; 5(1): e190.
102. Chan N, Koritzinsky M, Zhao H, Bindra R, Glazer PM, Powell S, et al. Chronic hypoxia decreases synthesis of homologous recombination proteins to offset chemoresistance and radioresistance. *Cancer Res.* 2008; 68(2): 605–14.
103. Muz B, Puente P de la, Azab F, Azab AK. The role of hypoxia in cancer progression angiogenesis metastasis and resistane to therapy. *Hypoxia.* 2015; 3: 83–92.
104. Courtney R, Ngo DC, Malik N, Ververis K, Tortorella SM, Karagiannis TC. Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. *Mol Biol Rep.* 2015; 42(4): 841–51.
105. Sullivan LB, Martinez-Garcia E, Nguyen H, Mullen AR, Dufour E, Sudarshan S, et al. The Proto-oncometabolite Fumarate Binds Glutathione to Amplify ROS-Dependent Signaling. *Mol Cell .* 2013; 51(2): 236–48.
106. Isaacs JS, Yun JJ, Mole DR, Lee S, Torres-Cabala C, Chung YL, et al. HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer:

- Novel role of fumarate in regulation of HIF stability. *Cancer Cell*. 2005; 8(2): 143–53.
107. Koivunen P, Fell S, Lu W, Rabinowitz J, Kung A, Schlisio S. The 2-oxoglutarate analog 3-oxoglutarate decreases normoxic hypoxia-inducible factor-1α in cancer cells, induces cell death, and reduces tumor xenograft growth. *Hypoxia*. 2016; 4: 15.
 108. Chicco AJ, Sparagna GC. Role of cardiolipin alterations in mitochondrial dysfunction and disease. *AJP Cell Physiol*. 2006; 292(1): C33–44.
 109. Paradies G, Paradies V, De Benedictis V, Ruggiero FM, Petrosillo G. Functional role of cardiolipin in mitochondrial bioenergetics. *Biochim Biophys Acta - Bioenerg*. 2014; 1837(4): 408–17.
 110. Lee H-J, Mayette J, Rapoport S, Bazinet R. Selective remodeling of cardiolipin fatty acids in the aged rat heart. *Lipids Health Dis*. 2006; 5(1): 2.
 111. Zhong H, Xiao M, Zarkovic K, Zhu M, Sa R, Lu J, et al. Mitochondrial control of apoptosis through modulation of cardiolipin oxidation in hepatocellular carcinoma: A novel link between oxidative stress and cancer. *Free Radic Biol Med*. 2017; 102: 67–76.
 112. Kagan VE, Tyurina YY, Tyurin VA, Mohammadyani D, Pedro J, Angeli F, et al. Cardiolipin Signaling Mechanisms: Collapse of Asymmetry and Oxidation.
 113. Kiebish MA, Han X, Cheng H, Seyfried TN. *In Vitro* Growth Environment Produces Lipidomic and Electron Transport Chain Abnormalities in Mitochondria from Non-Tumorigenic Astrocytes and Brain Tumours. *ASN Neuro*. 2009; 1(3): AN20090011.
 114. Raja V, Greenberg ML. The functions of cardiolipin in cellular metabolism—potential modifiers of the Barth syndrome phenotype. *Chem Phys Lipids*. 2014; 179: 49–56.
 115. Trachootham D, Alexandre J, Huang P. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? *Nat Rev Drug Discov [Internet]*. 2009; 8(7): 579–91.
 116. Zhong H, Yin H. Role of lipid peroxidation derived 4-hydroxynonenal (4-HNE) in cancer: Focusing on mitochondria. *Redox Biol*. 2015; 4: 193–9.
 117. Lee S, Kim SM, Lee RT. Thioredoxin and Thioredoxin Target Proteins: From Molecular Mechanisms to Functional Significance. *Antioxid Redox Signal*. 2013; 18(10): 1165–207.
 118. Mahmood DFD, Abderrazak A, El Hadri K, Simmet T, Rouis M. The Thioredoxin System as a Therapeutic Target in Human Health and Disease. *Antioxid Redox Signal*. 2013; 19(11): 1266–303.