



**UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO**



**INSTITUTO DE INVESTIGACIONES  
QUÍMICO BIOLÓGICAS**

**INSTITUTO DE INVESTIGACIONES QUÍMICO-BIOLÓGICAS**

**MAESTRÍA EN CIENCIAS QUÍMICAS**

**DISEÑO Y SÍNTESIS DE COMPUESTOS POTENCIALMENTE  
TRANSLOCADORES DE GLUT4 A TRAVÉS DEL RECONOCIMIENTO DE  
LA REGIÓN PIF DE LA PDK1**

**Tesis que para obtener el grado de Maestro en Ciencias Químicas**

**Presenta el QFB Abel Suárez Castro.**

**Asesor: Dr. Luis Chacón García**

Morelia Michoacán agosto de 2016

# 1. RESUMEN

El diseño de fármacos asistido por computadora (DIFAC) ha tenido un gran avance y desarrollo durante los últimos años, lo que ha llevado a que los estudios en farmacología no sean vistos sólo de forma experimental en una mesa de trabajo en el laboratorio, sino que, ahora también se contemplan y se incluyen de forma teórica empleando tecnologías como los sistemas de cómputo. En la parte introductoria de esta tesis se habla acerca de los fundamentos y detalles metodológicos que se emplearon en el desarrollo de esta investigación, así como sus ventajas y desventajas además de los retos que enfrentan actualmente, en seguida se presenta un bosquejo general de la diabetes mellitus. En seguida se muestra el desarrollo que fue teniendo la investigación comenzando con derivados de la chalcona por poseer las características similares a los compuestos encontrados como activadores de la enzima PDK1 a la que se le ha adjudicado un papel importante en las vías de señalización de la insulina y que en los últimos años se ha tomado en cuenta como una probable diana farmacológica contra la diabetes mellitus. Después, se describe el trabajo que se desarrolló en el Departamento de Química de la Universidad de Cambridge en el Reino Unido acerca del cribado virtual a partir de la base de datos con más de 20 millones de compuestos. Después de esto los compuestos que fueron seleccionados por el modelado farmacofórico, fueron sometidos a un análisis de acoplamiento molecular obteniendo así un reducido número de compuestos que tienen una probabilidad de ser sobresalientes para cumplir el objetivo de su búsqueda y diseño. Puede parecer complicado el elegir sólo aquellos compuestos que pueden tener una mayor probabilidad de interactuar en la región PIF de la PDK1, para resolver este problema, se utilizaron herramientas de clasificación y agrupamiento de todos los compuestos sobresalientes tomando en cuenta su “huella digital” para poder clasificarlos y tomar en cuenta sólo aquellos que tuvieran alta similitud molecular. De ahí que sólo se seleccionaron nueve compuestos a los que se les predijo una actividad  $pAC_{50}$  frente a la PDK1 utilizando 23 compuestos descritos por Engel *et al.* (2009),

***Palabras clave: DIFAC, diabetes, acoplamiento molecular, QSAR.***

## ABSTRACT

Drug design computer (DIFAC) assisted has had a breakthrough and development in recent years, which has led to studies in pharmacology are not only seen experimentally in a worktable in the laboratory, but, now also contemplated and included theoretically using technologies such as computer systems. In the introductory part of this thesis we present the fundamentals and methodological details that were used in the development of this research as well as its advantages and disadvantages and the challenges they currently face, then a general outline of diabetes mellitus. After justification, objectives and methodology of this work, then the development that was taking the investigation starting with derivatives of chalcone for possessing similar to the compounds found as activators of the enzyme characteristics shown arises PDK1 that it has awarded an important role in the signalling pathways of insulin and in recent years has taken into account as a possible drug target for diabetes mellitus. Then the work developed in the Department of Chemistry at the University of Cambridge in the UK about the virtual screening of zinc-based data (Zinc Data Base) with more than 20 million compounds available to perform it describes a search using the method of filtering Pains is a good tool to rule out the chemical structures that can be considered "promiscuous" or ligands that may interfere in in vitro studies. Noting that the classical rule 5 of Lipinski was not used because the latter focuses largely to predict bioavailability of a compound, however, we know that not all compounds which are already approved drugs meet this requirement. After filtering subjected to a search by pharmacophore modelling based on the structure to select those that meet the minimum requirements necessary to interact in the PIF region of the kinase PDK1. Thereafter which they were selected for pharmacophore modelling, were subjected to molecular analysis anchor thus obtaining a reduced number of outstanding compounds. It may seem complicated to choose only those compounds that may be more likely to interact in the PIF region PDK1, to solve this problem, tools classification and grouping of all outstanding compounds were used taking into account their "fingerprint" for to classify and take into account only those who have high molecular similarity. Thus only nine compounds that are predicted activity against PDK1 using the 23 compounds described by Engel et al,

***Key words: CADD, diabetes, Molecular Docking, QSAR.***

## 2. AGRADECIMIENTOS

Primero quiero agradecer a Dios por darme el regalo de la vida y por darme la fortaleza de continuar mi preparación que cada vez que avanzo me doy cuenta que soy más ignorante por desconocer todo este campo tan extenso y maravilloso como lo es la química. En seguida quisiera dar gracias a mi esposa Paty que ha sido el pilar más fuerte que he tenido en mi vida al compartir la suya conmigo y por estar siempre cuando las cosas van bien pero más cuando van mal. Sabemos que amar es una decisión y lo hemos aprendido juntos, así que este proyecto lleva en gran parte el aporte de su temple enérgico conmigo cuando así lo considero, pero también su temple cariñoso y tierno en los momentos de regocijo a través de estos dos años de travesía, he de confesar que ella ya ha de saber hacer un cálculo de acoplamiento molecular pues para mí cuando habla de estas cosas, siento como que si supiera lo mismo que yo. En seguida quiero agradecer a mis padres que, aunque no han estado conmigo estos últimos 9 años, sé que piden en sus oraciones por mí y que mucho de lo que soy se lo debo a ellos. En seguida quiero agradecer a mi asesor al Dr. Luis Chacón García por mostrarme su confianza plena en todo lo que le propuse, sé que hubiera querido que le demostrara más mi capacidad en síntesis orgánica, pero estamos seguros que fue un gran paso el que me haya transmitido las herramientas necesarias para transitar por este apasionante campo del diseño de fármacos, por todos sus consejos que más que un profesor ha sido un amigo como lo consideramos la mayoría en el laboratorio. A la Dra. Claudia Araceli Contreras Celedón por sus consejos, su amistad y sus sugerencias a mi trabajo, muy valiosos, por cierto, pero sobre todo por ser nuestra guía cuando tenemos alguna duda. También quiero agradecer a la M.C. Maricarmen Ramírez Medel por motivarme a estudiar mi maestría, fue un elemento clave para poder acudir al instituto de investigaciones químico biológicas y que solicitara mi ingreso. Gracias a mis compañeros y amigos que estuvieron presentes durante mi paso por el instituto, Mario, Ale, Silvia, Paco, Yuni, Kari, Fredy, Meli, Gaby, Chava, Fátima, Alejandro, Karen, y especialmente a Ana Lilia que me transmitió todo su conocimiento en cuanto al diseño de fármacos asistido por computadora, espero no haberla defraudado con mi trabajo, a familiares y amigos, especialmente a mi suegra por sus palabras de aliento cuando fueron necesarias, finalmente quiero agradecer al Consejo Nacional de Ciencia y Tecnología (CONACyT), por darme la oportunidad de obtener la beca económica No. 305150, para realizar este trabajo en la UMSNH y por haberme otorgado la Beca Mixta para realizar mi estancia de investigación en la Universidad de Cambridge en el Reino Unido.

### 3. TABLA DE CONTENIDO

1 RESUMEN .....	I
2 AGRADECIMIENTOS .....	III
3 TABLA DE CONTENIDO .....	IV
4 PRÓLOGO.....	VIII
5 INTRODUCCIÓN AL DISEÑO DE FÁRMACOS ASISTIDO POR COMPUTADORA (DIFAC) .....	1
6. DIABETES MELLITUS .....	8
7. PLANTEAMIENTO DEL PROBLEMA.....	23
8. JUSTIFICACIÓN .....	23
9. HIPÓTESIS.....	23
10. OBJETIVOS.....	24
11. METODOLOGÍA .....	25
12. PRIMEROS ESTUDIOS EN BUSCA DE COMPUESTOS PROBABLEMENTE ACTIVADORES DE LA PDK1 CON DERIVADOS DE CHALCONA.....	26
13. CRIBADO VIRTUAL EN BUSCA DE COMPUESTOS COMO PROBABLES ACTIVADORES DE LA PDK1.....	48
14. ESTUDIO SOBRE UNA PROPUESTA SINTÉTICA DE LOS MEJORES COMPUESTOS OBTENIDOS DEL CRIBADO VIRTUAL.....	83
.....	85
15. CONCLUSIONES Y PERSPECTIVAS.....	97
REFERENCIAS .....	100

## LISTA DE FIGURAS

Figura 1 Prevalencia mundial de diabetes en varones mayores de 18 años durante el año 2014. Tomado de: Global Status Report on noncommunicable diseases 2014) .....	9
Figura 2 Estadística descriptiva que muestra la mortalidad y la distribución de la diabetes mellitus en México clasificada por edades. ....	10
Figura 3 Estructura química de la metformina, una biguanida utilizada como primera opción en el tratamiento de la diabetes mellitus tipo 2 y que acompaña a la mayoría de los esquemas terapéuticos recomendados por la Asociación Americana de la Diabetes. ....	11
Figura 4 Estructura química de la pioglitazona, un hipoglicemiante oral que incrementa la sensibilidad a la insulina especialmente en el tejido adiposo. ....	12
Figura 5 Sitagliptina, un hipoglicemiante oral inhibidor de la dipeptidil peptidasa tipo 4 (DPP4).....	13
Figura 6 Estructura química de la dapagliflozina, un hipoglicemiante oral de reciente uso contra la diabetes mellitus tipo 2, y que es un inhibidor del cotransportador de sodio/glucosa tipo 2 (SGLT2) a nivel tubular renal. ....	13
Figura 7 Papel de la PDK1 en la regulación de varias respuestas biológicas provocadas por la insulina. La PDK1 activa a otras cinasas que llevan a término la señalización de la insulina. PKN: Proteína cinasa serina/treonina; PKB: Proteína cinasa tipo B; PKC: Proteína cinasa de tipo C; PFK2: Fosfofructocinasa 2. Modificado de: K.L.R. Wick. A New Molecular Target of Insulin Action: Regulating the Pivotal PDK1 Current Drug Targets - Immune, Endocrine & Metabolic Disorders, 2001, 1, 209-221.....	16
Figura 8 Proteína cinasa dependiente de inositol trifosfato (PDK1) Código PDB: 4AW0 .....	17
Figura 9 Mecanismo de acción que explica la activación de las algunas cinasas por parte de los compuestos de bajo peso molecular .....	18
Figura 10 Estructura química de los compuestos reportados como activadores de la PDK. El nombre es como aparece en las estructuras cristalográficas depositadas en la base de datos de proteínas (PDB).....	20
Figura 11 Ácido 2 - ((3 - (4-clorofenil)-3-oxo-1-fenil-propil) tio) acético.....	21
Figura 12 Estructura del compuesto 2Z propuesto por Adriana Stroba y cols.....	22
Figura 13 Comparación de las estructuras químicas entre los compuestos conocidos como activadores de la PDK1 (A y C) y la chalcona.(B) .....	26
Figura 14 Compuestos propuestos de novo, derivados de la chalcona.....	27
Figura 22. Estructura química del ácido 2-((3-oxo-1,3-difenilpropil)amino)acético. ....	46

## LISTA DE TABLAS

Tabla 1. Resumen de los principales hipoglicemiantes orales que se utilizan actualmente contra la diabetes mellitus, sólo se presentan uno de cada grupo. ....	14
Tabla 2. Coordenadas de muestreo y dimensiones de la malla generada en Autodock Tools y utilizada en los estudios de acoplamiento molecular. ....	34
Tabla 3. Resultados del cálculo de RMSD teniendo como referencia el ligando de su correspondiente estructura cristalográfica. ....	37
Tabla 4. Resultados del acoplamiento molecular de los 53 ligandos analizados con el receptor 3HRF se incluye en el inicio los resultados del ligando de referencia. ....	38
Tabla 5. Resultados del estudio de Acoplamiento molecular de los 53 ligandos analizados con el receptor 4A06 se incluye en el inicio los resultados del ligando de referencia. ....	39
Tabla 6. Tabla 6. Resultados de los estudios de acoplamiento molecular de los 53 ligandos analizados con el receptor 4AW0 se incluye en el inicio los resultados del ligando de referencia. ....	41
Tabla 7. Resumen de los mejores resultados de energía libre y KI de los ligandos 10, 13, 14, 14a, 20a, 22, 22a, 23, 23a. ....	43
Tabla 8. Condiciones y características de los nueve protocolos empleadas en los estudios de acoplamiento molecular en el programa GOLD. ....	53
Tabla 9. Resultados de los estudios de acoplamiento molecular comparando el desempeño de nueve diferentes protocolos utilizando el teorema bayesiano para calcular sensibilidad y especificidad. ...	68
Tabla 10. Modelos farmacofóricos con sus respectivas características y sus medidas, además de los resultados de sensibilidad y especificidad de cada una de ellas. ....	72
Tabla 11. Resumen de los resultados de los estudios 3DQSAR. ....	79
Tabla 12. Resultados de la predicción de la actividad pAC50 de los 23 compuestos conocidos como activadores de la PDK1. ....	81
Tabla 13. Resultados de la actividad Pac50 calculada para los nueve compuestos obtenidos del cribado virtual. ....	82
Tabla 14. Tabla 1. Resumen de las condiciones e intentos de la tioración del carbonilo de la acetofenona. ....	91
Tabla 15. Resumen de los desplazamientos químicos de los compuestos 6' y 4' . ....	96

## LISTA DE ESQUEMAS

Esquema 1 Primera propuesta de síntesis orgánica para generar un precursor que nos lleve a la obtención de los compuestos derivados de la chalcona que mostraron los mejores resultados de energía libre y de Ki en los estudios de acoplamiento molecular. ....	46
Esquema 2. Propuesta sintética del compuesto 22 y 22a que obtuvieron buenos resultados de energía libre y de Ki en los estudios de acoplamiento molecular. ....	47
Esquema 3. Algoritmo empleado dentro del programa GOLD para llevar a cabo la elección de la mejor pose de cada ligando en el sitio de unión tomando en cuenta los mejores resultados de energía libre. ....	56
Esquema 4. Flujoograma de trabajo para el modelado farmacofórico basado en la estructura. ....	61
Esquema 5. Diagrama de flujo de la metodología empleada en los experimentos de síntesis orgánica. ....	86
Esquema 6. Análisis retrosintético de 1(compuesto ZINC1908976). ....	88

Esquema 7. Análisis retrosintético de 1' (análogo del compuesto ZINC1908976). .....	89
Esquema 8. Ruta general de síntesis del compuesto 1'. .....	90

## 4.PRÓLOGO

La presente tesis, se compone primero de un bosquejo general sobre los métodos que utiliza el DIFAC el cual poya al campo de la química computacional, del cual el enfoque de este trabajo es contra la Diabetes Mellitus que es una enfermedad clasificada como no transmisible, pero no por eso, deja de ser preocupante para la población mundial pues se estima que las cifras incrementen casi al doble de lo que se tiene reportado hasta el tiempo en que se escribe esta tesis. Así pues, deben buscarse medidas para combatir la enfermedad antes mencionada, y uno de los puntos principales además de la prevención, es el mejorar las herramientas de tratamiento para evitar y retardar las complicaciones de la misma. El presente trabajo desde el punto de vista de la química medicinal, pretende contribuir a proponer nuevas formas de estudiar y de buscar nuevos compuestos que puedan ser desarrollados en un futuro a través del desarrollo de fármacos.

Se divide por secciones puesto que no fue un estudio con un solo tipo de compuestos ni tampoco con un cuadro de herramientas y metodologías como es común encontrar en las tesis en el área de las ciencias químicas, pues el DIFAC, más allá de ser un trabajo que muchos clasifican como “apretar botones”, muestra que la química ha dejado desde hace aproximadamente 20 años una ciencia que sólo puede desarrollarse sobre una mesa de un laboratorio.

La parte introductoria pretende abarcar y de esclarecer los fundamentos de cada metodología utilizada, después se presenta la metodología de investigación empleada, desde la justificación, la hipótesis y los objetivos enlistando los métodos que se emplearon, después las siguientes tres secciones muestran una misma estructura que contiene la descripción de la metodología empleada, los resultados obtenidos y al final del trabajo, las conclusiones con su respectiva discusión y al final las conclusiones.

## 5 INTRODUCCIÓN

### 5.1 Introducción al diseño de fármacos asistido por computadora (DIFAC)

El campo de estudio de la farmacología se ha diversificado en años recientes, involucra el conocimiento de la química orgánica, la química inorgánica, la fisiología, la bioquímica, la fisiopatología además del empleo de las herramientas de métodos computacionales en el diseño racional de fármacos llamado DIFAC, o bien llamados estudios in silico<sup>1</sup>.

A principios del siglo XX Paul Erlich propuso el término “receptor” refiriéndose a una estructura macromolecular e donde hoy sabemos que los fármacos llevan a cabo su farmacodinamia al interactuar con el mismo provocando un cambio en algún proceso o función fisiológica. Uno de los métodos en la búsqueda de nuevos fármacos con mayor eficacia y menores efectos adversos ha sido la química computacional a lo que se le llama “Diseño de fármacos in silico o asistido por computadora”. Apoyados de otra herramienta fundamental e imprescindible como es el uso de las estructuras cristalográficas de proteínas almacenadas en una base de datos pública llamada *Protein Data Bank* (Banco de Datos de Proteínas)<sup>2</sup>, en donde se puede visualizar la estructura tridimensional con la ayuda de programas computacionales, permite conocer el sitio o blanco en donde se quiere que el nuevo fármaco interactúe. Para el desarrollo de un nuevo fármaco asistido por computadora, es necesario recopilar toda la información en cuanto a los procesos fisiológicos normales y patológicos alrededor de la enfermedad, luego se proponen nuevas estructuras químicas en las que se puede hacer una predicción teórica en cuanto a sus interacciones y actividad frente al blanco deseado, a través de cálculos computacionales con lo que se han logrado grandes avances en el campo de la farmacología reduciendo el tiempo y los costos invertidos en el desarrollo de nuevos y mejores fármacos<sup>3</sup>.

Hasta hace apenas la década de los ochentas, el estudio de la química dejó de ser visto como una ciencia puramente experimental<sup>4</sup>, es decir, no sólo se puede estudiar el comportamiento de la materia en una mesa de laboratorio con experimentos que involucran reactivos y condiciones específicas de reacción, además del empleo de las herramientas espectroscópicas, ahora, se le puede considerar a la química como el estudio experimental y teórico aplicando simulaciones in silico, que van desde los cálculos de las propiedades básicas de las moléculas que ahora son llamados “sistemas” a través del uso de las matemáticas de la mecánica clásica de Newton en cuanto al movimiento, atracción y repulsión de las moléculas. También ahora se cuenta con el

cálculo de las propiedades fisicoquímicas y de interacción entre diferentes sistemas pequeños por ejemplo una molécula de etanol, hasta un sistema muy grande y complejo como una proteína completa utilizando cálculos matemáticos de la ecuación de Schrödinger aplicando los modelos semi-empíricos, los de Hatree Fock y de la teoría de funcionales de la densidad electrónica (DFT) con un nivel de cálculo mayor<sup>4</sup>. En cuanto al uso de estos modelos matemáticos de la química computacional aplicados a la farmacología, podemos decir que han sido de gran utilidad los dos niveles de cálculo pues los problemas que se pretenden resolver con estos métodos son diversos y cada estudio necesita utilizar uno o varios modelos.

El primer artículo publicado en donde se utilizaron los estudios in silico para el desarrollo de fármacos fue en octubre de 1981 por la revista Fortune con el título: *Uso por primera vez de sistemas computacionales en el desarrollo de fármacos en la farmacéutica Merck* <sup>5</sup>

Durante al menos las últimas cuatro décadas, el desarrollo de los estudios in silico ha tenido grandes avances, dentro de sus objetivos se incluyen la generación de nuevos y mejores compuestos con propiedades aceptables en estudios in vitro e in vivo reduciendo así el número de entidades químicas a ser probadas como candidatos a nuevos fármacos.

Es importante mencionar que los grandes avances en la obtención de la información biológica almacenada en grandes bases de datos, ha contribuido de forma sustancial al DIFAC, sin embargo, han aparecido nuevos retos en esta disciplina debido a las exigencias cada vez mayores en obtener resultados cada vez más confiables, con un menor gasto computacional y con mejores aplicaciones.

El DIFAC lo podemos dividir en tres grandes áreas de investigación: la primera encargada del diseño, la implementación y el mantenimiento de las herramientas necesarias para mejorar los métodos in silico, organizando, analizando y almacenando grandes cantidades de datos de una forma muy rápida, en segundo lugar, tenemos aquellos métodos para identificar, caracterizar y priorizar dianas biológicas estableciendo su relación con la enfermedad que hoy conocemos como bioinformática, y en tercer lugar están los métodos que ayudan a encontrar nuevos y mejores candidatos basados en la estructura o en el ligando y que a su vez ocupan de herramientas y métodos muy conocidos que son los que mayor impacto han tenido a la investigación farmacológica en estudios in silico<sup>6</sup>.

Dentro de los fármacos en donde se han visto involucrados los estudios in silico, tenemos una lista de al menos 12 compuestos de bajo peso molecular entre ellos: Captopril, Dorzolamida,

Saquinavir, Zanamivir, Oseltamivir, Aliskireno, Boceprevir, Nalotrexona, TMI-005, LY-517717, Rupintrivir y NVP-AUY922 que se encuentran en ensayos clínicos o bien que han sido aprobados para su uso terapéutico<sup>7</sup>. El empleo de métodos computacionales mencionados arriba en la tercera categoría, han contribuido en el descubrimiento, optimización y reposicionamiento de compuestos incluso mejores que los ya desarrollados y usados en la terapéutica.

Debe entenderse que el DIFAC no es sólo “apretar botones”. Ciertamente los fármacos no se diseñan en “cajas negras” donde no se sabe que ocurre con los cálculos, en efecto, estos no son estudios aislados pues deben ser complementados con estudios experimentales tras varios ciclos de optimización empleando diferentes métodos llevando esto al análisis extenso y detallado de los resultados que se pueden obtener del DIFAC<sup>1</sup>.

## 5.2 Clasificación de los métodos del DIFAC

Como se comentó arriba, podemos dividir al DIFAC en tres grandes categorías, siendo la tercera de estas, actualmente la más empleada y por lo tanto la más desarrollada, por lo que nos enfocaremos en lo que sigue a la descripción de los métodos sobresalientes dentro del DIFAC. En primer lugar, por lo general, el DIFAC está encaminado al desarrollo de nuevos y mejores compuestos como probables candidatos a ser fármacos, su optimización o bien su reposicionamiento, por lo tanto es necesario conocer con detalle hasta donde sea posible, la información de la diana terapéutica por ejemplo, la función que juega en la fisiopatología de alguna enfermedad, así como la existencia de compuestos aprobados para uso terapéutico y que tengan interacción con la misma, o bien compuestos que hayan sido probados en estudios experimentales, esto, dependiendo del tipo de diana terapéutica que se trate. Además de todo esto, el conocimiento químico en cuanto a comportamiento, reactividad, solubilidad y propiedades de formar interacciones químicas, son elementos clave a considerar para los estudios in silico. Varias herramientas y métodos empleados en la biología celular, la farmacología y la química orgánica han contribuido al mejor desarrollo y diseño de nuevos compuestos o la optimización de algunos de ellos.

### 5.2.1 Métodos basados en la estructura

Este campo de investigación del DIFAC se refiere al aprovechamiento de la información con la que se cuenta en base a la estructura del receptor que puede ser por lo general una proteína, o en algunos otros casos específicos ácidos nucleicos. Estas estructuras han sido obtenidas por medio de difracción de rayos x, también empleando resonancia magnética nuclear, y recientemente, por medio de microscopía electrónica en tres dimensiones. Dentro de la información con la que se cuenta en base a la estructura de los receptores farmacológicos existen algunas bases de datos que se encargan de recibir, analizar y refinar los datos de las coordenadas atómicas obtenidas por los métodos que mencionamos arriba, entre las más conocidas está el Banco de Datos de Proteínas o mejor conocido como PDB (Protein Data Bank por sus siglas en inglés) y que desde 1971 que fue inaugurado cuenta a la fecha con alrededor de 118,587 estructuras accesibles y disponibles para la investigación en todos los niveles<sup>2</sup>. Este enfoque del DIFAC sostiene la hipótesis de que un ligando o compuesto de bajo peso molecular, puede exhibir una interacción específica con alguna parte del receptor y esto a su vez puede dar paso a la predicción de los efectos biológicos que puede presentar dicha interacción. Desde la década de los años ochenta, en la investigación farmacológica se comenzó con el uso y estudio ampliado de este enfoque, y que con los avances de la genómica y proteómica, se ha podido obtener una gran cantidad de receptores que son candidatos a ser estudiados como dianas farmacológicas<sup>7 y 8</sup>. En la actualidad se puede hacer uso de esta herramienta a través del uso del cribado virtual en donde en poco tiempo se puede conocer si es que existe alguna interacción entre una lista de ligandos con múltiples receptores.

Dentro del DIFAC basado en la estructura se cuenta con varios métodos muy accesibles y fáciles de utilizar como: el análisis de acoplamiento molecular mejor conocido como “docking”, el modelado análogo de receptores, el modelado farmacofórico y la dinámica molecular.

#### 5.2.1.1 Estudios de acoplamiento molecular

Los estudios de acoplamiento molecular o mejor conocidos por su palabra inglesa *docking*, son una técnica computacional automatizada que consta de algoritmos que determinan la predicción de cómo interactúa un ligando en un sitio específico de un receptor<sup>9</sup>. Por lo general, este método es casi siempre ligando-receptor, pero en algunas ocasiones puede ser receptor-receptor, o la evaluación puede ser en orden inverso, es decir receptor-ligando a lo que se le conoce como análisis de acoplamiento molecular inverso<sup>6</sup>. El método general común, puede

predecir la orientación del ligando, su conformación geométrica y el puntaje que alcanza en los cálculos según el algoritmo empleado, es decir, el puntaje viene dado por los resultados obtenidos y que son diferentes en cada programa computacional que puede hacer cálculos de acoplamiento molecular. Este tipo de cálculos ayuda a predecir la conformación y pose entre el ligando y su receptor, que, para estudios biológicos, el reconocimiento molecular es esencial para conocer la forma en cómo se lleva a cabo la interacción entre ambos a nivel atómico, es decir, las interacciones débiles como los puentes de hidrógeno, las atracciones y repulsiones electrostáticas, así como las hidrofóbicas<sup>10</sup>. Por lo general, en el análisis de acoplamiento molecular, se utilizan ecuaciones matemáticas aplicando la mecánica clásica de Newton, tomando en cuenta parámetros electrónicos, el impedimento estérico, las fuerzas de atracción y de repulsión con el empleo de campos de fuerza para sistemas atómicos de bajo peso y alto peso molecular, es decir, ligandos o bien macromoléculas respectivamente<sup>11</sup>. Para poder hacer esas aproximaciones, el acoplamiento molecular dispone de varias etapas de cálculo, cada una enfocada a obtener los mejores resultados a través de un algoritmo, en donde el mejor resultado es la mínima energía libre total del sistema (ligando-receptor) aplicando algoritmos de búsqueda, potenciales de campos de fuerza y funciones de puntaje<sup>10</sup>.

#### *Algoritmos de búsqueda*

Existen dos componentes básicos en un programa para realizar estudios de acoplamiento molecular, el algoritmo de búsqueda y la función de puntaje<sup>12</sup>. El algoritmo de búsqueda, posiciona al ligando en diferentes conformaciones, localizaciones y orientaciones basado en los giros libres de torsión que vienen dados por el estado de hibridación en los enlaces covalentes. Los campos de fuerza apoyan al cálculo tomando en cuenta los parámetros arriba mencionados y sitúa cada vez al ligando en algún sitio específico sobre el receptor. La elección de un algoritmo de búsqueda determina cómo se puede alcanzar una predicción aproximada de la pose más favorable según los valores de energía libre y en algunos casos, el cálculo de la constante de afinidad  $K_i$ . Cada estudio y proyecto, debe en lo particular tomar los métodos y procesos generales como primera instancia para luego refinarlo y elegir uno de forma específica según los objetivos que se pretenden alcanzar. Cada software contiene sus propias formas de búsqueda con distintas variables, por ejemplo, el software Gold para el acoplamiento molecular, es de mucha ayuda sus algoritmos de búsqueda si se está trabajando con cinasas pues cuenta con la parametrización necesaria para los cambios conformacionales que sufren ciertos aminoácidos en su sitio activo. Los algoritmos genéticos los más utilizados debido primero a la

calidad de los resultados que se pueden obtener realizando de forma adecuada la preparación de las estructuras químicas de partida para el cálculo. Dentro de los algoritmos de búsqueda, los más utilizados son el método Monte Carlo que realiza un cálculo global aleatorio<sup>6</sup> que en algunas ocasiones se encarga de llevar a cabo la rotación de enlaces simples además de elegir las posiciones de densidad electrónica de forma aleatoria llevando a cambios estructurales y en otros casos modificando la estereoquímica lo cual lo convierte en una desventaja.

### *Funciones de puntaje*

Las funciones de puntaje son las responsables de hacer el cálculo de la mejor energía de interacción entre el ligando y el receptor que es la más favorable en la cual posiblemente tenga interacción el ligando con su receptor

Los términos comúnmente utilizados en el análisis docking en cuanto a la predicción de la interacción de un ligando con el receptor se pueden estudiar con respecto de : su conformación, que se define como la forma tridimensional en la cual se establecen las interacciones más estables entre los dos sistemas. En los cálculos de acoplamiento molecular la conformación predicha no necesariamente es la conformación más estable del ligando pues la función de puntaje se encarga de encontrar la mejor posición independientemente de las fuerzas intramoleculares, más bien, influyen de una forma mayoritaria las fuerzas intermoleculares. En algunas ocasiones algunos investigadores discuten sobre la necesidad de que antes de realizar el cálculo del acoplamiento molecular se debe obtener un análisis conformacional del ligando para partir de su mínima conformación energética, siendo esto innecesario por lo que se ha comentado arriba. La translación que viene representada por la posición espacial con coordenadas específicas en el sitio de unión con el receptor siendo esta característica el resultado de predicción que conduce a una posición, resultados de energía libre y aquellos que algún programa de anclaje molecular pueda ofrecer utilizando su función de puntaje, por eso es tan necesario evaluar este parámetro de un cálculo de anclaje molecular antes de explorar grandes bases de datos por ejemplo en un cribado virtual.

### *Potenciales y campos de fuerza*

Al igual que en la dinámica molecular, se pueden utilizar campos de fuerza para realizar los cálculos de atracción y repulsión entre los átomos involucrados en el acoplamiento molecular tomando en cuenta algunas variables físicas y químicas que prevalecen en las interacciones ligando-proteína. Existen algunos campos de fuerza comunes y bien conocidos con diferentes aplicaciones al diseño de fármacos asistido por computadora donde las principales diferencias son los términos que se incluyen son en base a la energía libre entre el ligando y el receptor, de forma intra e intermolecular. Dentro de los campos de fuerza más sobresalientes está el AMBER<sup>13</sup> y el campo de fuerza CHARMM<sup>14</sup>

#### 5.2.1.2 Modelado farmacofórico

Dentro de las primeras etapas de un proyecto de investigación sobre la búsqueda de nuevos compuestos con probable actividad biológica, los investigadores eligen algunas opciones sobre cómo conseguir encontrar estructuras químicas que cumplan los requisitos necesarios para tener interacción con el sitio de interés de una diana farmacológica, para estos propósitos, se tiene la disponibilidad de la información estructural. El farmacóforo se define como el conjunto de interacciones electrónicas y estéricas que tiene un ligando en el sitio de interés en el receptor<sup>15</sup>. Las características comunes que definen a un farmacóforo son, por ejemplo, las interacciones puentes de hidrógeno a través de grupos aceptores o donadores de los mismos, además de grupos hidrofóbicos, grupos ácidos, grupos básicos, cargas parciales, y regiones hidrofóbicas aromáticas. Estas características han sido bien utilizadas en el cribado virtual, diseño *de novo*, y optimización de compuestos líderes a través de su modificación del auxóforo también llamada región haptopofórica<sup>16</sup>. En la actualidad el uso del modelado farmacofórico ha tenido un gran impacto casi a la par del empleo de estudios de acoplamiento molecular, pues ofrece entre otras ventajas, el diseño de nuevos compuestos basado en la estructura o bien basados en el ligando<sup>17</sup>.

## 6. DIABETES MELLITUS

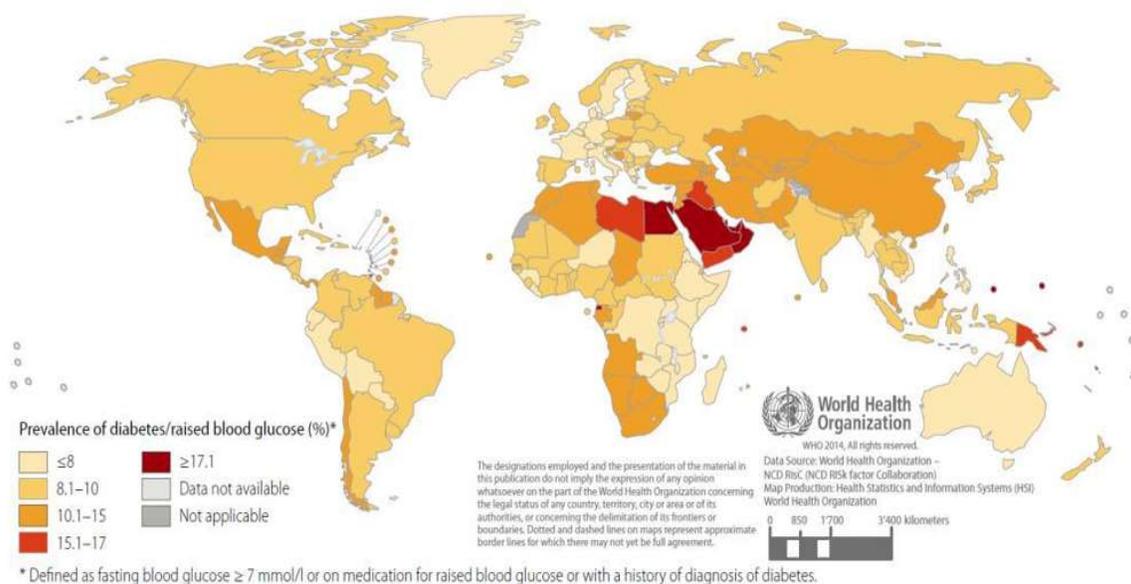
Es importante hablar de la diabetes mellitus (DM) una enfermedad de tipo no transmisible (ENT) que es un problema mundial por el que pasa nuestro planeta, y que la Organización Mundial de la Salud junto con otros organismos encargados de combatirla, han llevado a cabo campañas de prevención, estudios clínicos y farmacológicos bien definidos para evitar que esta progrese y que en un futuro no nos encontremos en el lugar que según nos dicen los investigadores para el 2030<sup>18</sup>.

### 6.1 Epidemiología actual y el futuro de la diabetes mellitus

Las enfermedades no transmisibles (ENT) son los padecimientos que están preocupando en todo el mundo, la Dra. Margaret Chan, directora general de la Organización Mundial de la Salud (OMS) ha enviado un mensaje dentro del Reporte de Status Global de las ENT al respecto de este problema de salud a nivel mundial: *el mundo ha alcanzado un decisivo punto de la historia en las enfermedades no transmisibles, y tiene una sin precedente oportunidad para alterar su curso*. La mayoría de los países se han puesto de acuerdo para establecer un tiempo límite para poder frenar nueve objetivos que son un problema de salud para el 2025, entre los principales tenemos reducir el alcohol, la insuficiente actividad física, consumo de sodio y tabaquismo, los cuales se han visto involucrados como factores de riesgo que incrementan la posibilidad de padecer enfermedades metabólicas como la diabetes mellitus y obesidad. El último mensaje fue: *el más importante mensaje del Reporte de Status Global de las ENT es que actualmente la comunidad global tiene la oportunidad de cambiar el curso de la epidemia de las ENT*, así, todos los países que participan activamente dentro de la OMS, tienen una agenda global para la prevención y el control de las ENT, en donde la responsabilidad es compartida entre todos los países basadas en objetivos concretos<sup>19</sup>. Según las predicciones realizadas por los expertos en las ENT que durante los años 2011 al 2025, las perdidas acumulativas debido a las ENT, han sido estimadas en aproximadamente \$7 trillones de dólares, lo que representa una suma mucho mayor a los \$11.2 millones de dólares para implementar intervenciones de alto impacto para reducir las ENT. La obesidad, incrementa el riesgo de padecer diabetes mellitus, hipertensión, enfermedad coronaria, la prevalencia mundial de obesidad ha incrementado casi al doble desde 1980, para el 2014 (**Figura 1**) se encontró un 11% en el sexo masculino y un 15% para el sexo femenino mayores de 18 años ambos. Para el 2013, Se reportó más de 42 millones de niños

menores de 5 años con obesidad. La prevalencia global de diabetes en el 2014 fue estimada de un 9%<sup>20</sup>.

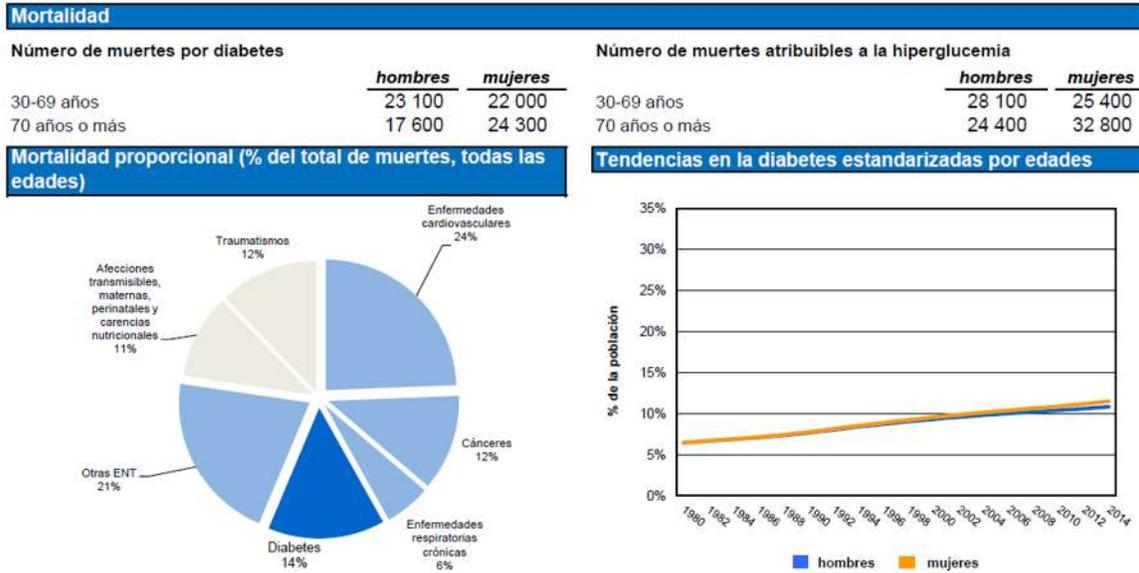
La Diabetes Mellitus es un síndrome clínico que resulta de la secreción deficiente de insulina, o resistencia a la misma. Esta deficiencia puede ser absoluta o relativa; es absoluta cuando su secreción es menor de lo normal y relativa cuando la acción de la hormona está disminuida<sup>21</sup>. Se estima que existen aproximadamente 6.4 millones de personas en México que la padecen lo cual genera grandes gastos en salud que pueden ser prevenibles<sup>22</sup>.



**Figura 1** Prevalencia mundial de diabetes en varones mayores de 18 años durante el año 2014. Tomado de: *Global Status Report on noncommunicable diseases 2014*

Los reportes más recientes de la OMS en un resumen sobre el problema que se enfrenta refieren lo siguiente: en los países de ingresos bajos y medianos con frecuencia no se dispone de los medicamentos esenciales que son fundamentales para controlar la diabetes. Los gobiernos, los profesionales de la salud, las personas con diabetes, la sociedad civil, los productores de alimentos y fabricantes de productos alimenticios y los proveedores de medicamentos: todos son partes interesadas. En conjunto pueden hacer una contribución importante para detener el aumento de la diabetes y mejorar la calidad de vida de las personas que tienen la enfermedad.

El reporte de las estadísticas de la DM en México (**Figura 2.**) indica que ha habido un incremento de la tasa de mortalidad siendo el sexo masculino en edad juvenil quien la padece más<sup>23</sup>.



**Figura 2** Estadística descriptiva que muestra la mortalidad y la distribución de la diabetes mellitus en México clasificada por edades.

El problema del incremento de las ENT como la DM va en aumento en América de un 54%, es decir más de 54.9 millones de personas que padecerán DM, pronóstico obtenido por el modelo de Rowley et al. para el año 2030<sup>18</sup>. Que consiste en incluir los censos poblacionales disponibles en América desde al año 1990, dividiendo a la población total por segmentos de edad y estimando a través de un modelo predictivo (lineal) la prevalencia y los casos diagnosticados con DM en diferentes tiempos con circunstancias ambientales o externas similares teniendo en cuenta las publicaciones sobresalientes sobre la cantidad de personas afectadas con la enfermedad, información proporcionada por el Centro Nacional de Estadística en Salud de los Estados Unidos y el Centro de Control de Enfermedades (CDC).

Sobre las razones importantes para el desarrollo de este trabajo, se decidió desde un inicio abordar el problema de la enfermedad ya que es un área de estudio bastante amplia en donde podemos incursionar con estudios de DIFAC enfocados a nuevas propuestas terapéuticas sin dejar de lado el lugar que tiene la prevención de la misma enfermedad y que este último tema es algo en lo que deberían apostar más la sociedad en general, no sólo los dirigentes

gubernamentales y aquellos en los que en ocasiones se considera que tienen el poder de cambiar la situación.

## 6.2 Tratamiento farmacológico de la diabetes mellitus tipo 2, fármacos actuales, y nuevas perspectivas en el tratamiento de esta enfermedad

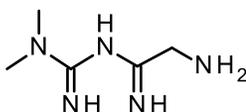
El tratamiento de la diabetes mellitus es tan antiguo como la misma enfermedad. Las biguanidas se identificaron en la planta *Galega officinalis*, cuyos extractos se utilizaban en la edad media para tratar esta enfermedad<sup>21</sup>.

Con los adelantos científicos y tecnológicos se logró identificar la fisiopatología de la enfermedad y con ella los componentes bioquímicos que intervienen en la resistencia a la insulina o bien cuando no se encuentra en un organismo vivo. Así se han diseñado varios compuestos útiles en el retardo de la absorción de carbohidratos como la acarbosa, secretagogos de insulina como las sulfonilureas o las gliptinas análogos de GLP-1<sup>24</sup>. Sin embargo, no se había tenido reporte de la enzima PDK1 y su papel en la vía de señalización de la insulina.

Así, existen dos vías de administración de tratamiento contra la diabetes mellitus, tratamiento oral y los parenterales.

### *Metformina.*

Junto con la tolbutamida fueron los primeros sensibilizadores de insulina en el tratamiento de la diabetes mellitus. Es una biguanida (**Figura 3**) cuyo mecanismo de acción es activar a la AMPK disminuyendo así la concentración de ATP intracelular provocando la sensibilización a la insulina en las células de músculo esquelético<sup>25</sup>.

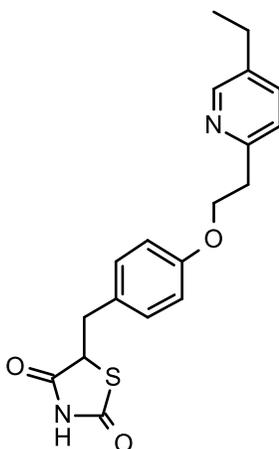


**Figura 3** Estructura química de la metformina, una biguanida utilizada como primera opción en el tratamiento de la diabetes mellitus tipo 2 y que acompaña a la mayoría de los esquemas terapéuticos recomendados por la Asociación Americana de la Diabetes.

Es el fármaco oral de primera elección en el tratamiento de la diabetes mellitus tipo 2, reconocido en todas las guías y recomendaciones clínicas actuales. El estudio: UKPDS (United Kingdom Prospective Diabetes Study) y DCCT (Diabetes Control and Complications Trial) manifestaron los beneficios de utilizar metformina incluso como profiláctico en diabetes mellitus tipo 2<sup>26</sup>

#### *Nuevos sensibilizadores de insulina.*

Las tiazolidinedionas como la pioglitazona (**Figura 4**) son compuestos que activan al receptor activador de la proliferación del peroxisoma de tipo gamma (PPAR gamma) provocando así la sobre expresión de receptores de insulina en músculo esquelético y tejido adiposo aumentando la sensibilidad a la misma <sup>25</sup>.

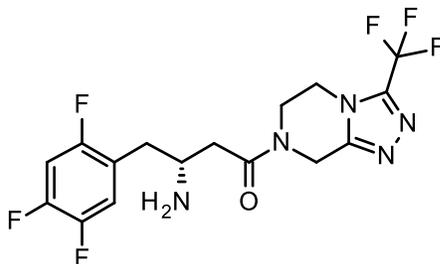


*Figura 4 Estructura química de la pioglitazona, un hipoglicemiante oral que incrementa la sensibilidad a la insulina especialmente en el tejido adiposo.*

#### *Activadores de secreción de insulina.*

Las gliptinas como la sitagliptina (**Figura 5**) ocupan el primer lugar del tratamiento con antidiabéticos orales porque no afectan al peso corporal, están desprovistas de efecto hipoglicemiante y se pueden combinar con cualquier otro fármaco usado actualmente. La gliptina se define como un fármaco que produce el efecto metabólico de una hormona natural de origen intestinal secretada en respuesta a la ingestión de alimentos, el péptido glucanoide

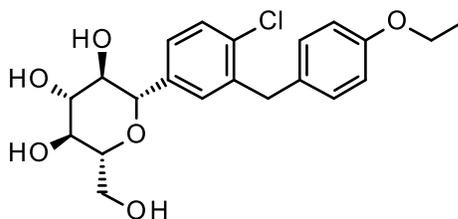
de tipo 1 (GLP-1) que induce la secreción de insulina y simultáneamente inhibe la del glucagón<sup>25</sup>.



**Figura 5** Sitagliptina, un hipoglicemiante oral inhibidor de la dipeptidil peptidasa tipo 4 (DPP4).

#### *Inhibidores del cotransportador de sodio/glucosa SGLT2*

Este grupo de fármacos representan la farmacoterapia oral más reciente contra DM, poseen algunas ventajas sobre la terapia convencional ya conocida y utilizada desde hace décadas, pues se pueden utilizar en el tratamiento de la DM tipo 2 y por primera ocasión un antidiabético oral se puede emplear contra la DM tipo 1. Actúan inhibiendo la reabsorción de glucosa en el túbulo contorneado proximal impidiendo el regreso de iones sodio acompañados con el soluto glucosa provocando glucosuria y por consiguiente se conocen bien los efectos adversos esperados como lo es el caso de las infecciones de las vías urinarias<sup>27</sup>. En recientes fechas, la FDA de los Estados Unidos de América ha emitido un comunicado sobre la preocupación que existe en la fase IV de estos fármacos pues ya hay otros tres congéneres de la dapagliflozina (**Figura 6**) en el mercado de los cuales se han reportado varios casos de cetoacidosis llevando a los pacientes a los servicios de urgencias de forma inmediata para recibir el tratamiento adecuado, siendo así nuevos fármacos de los que se tiene una recuperación temprana sobre los efectos adversos antes de lo previsto.



**Figura 6** Estructura química de la dapagliflozina, un hipoglicemiante oral de reciente uso contra la diabetes mellitus tipo 2, y que es un inhibidor del cotransportador de sodio/glucosa tipo 2 (SGLT2) a nivel tubular renal

En la **Tabla 1** se resume el tratamiento farmacológico de la DM tipo 2 resaltando el fármaco, la diana farmacológica y la ruta metabólica asociada

**Tabla 1** Resumen de los principales hipoglucemiantes orales que se utilizan actualmente contra la diabetes mellitus, sólo se presentan uno de cada grupo.

Fármaco	Diana farmacológica	Ruta metabólica asociada
Promotores del metabolismo de glucosa (biguanidas)		
Metformina	NADH deshidrogenasa-ubiquinona	Fosforilación oxidativa
Sensibilizadores de insulina		
Pioglitazona, rosiglitazona	PPAR-gamma	Transcripción de DNA en la síntesis de receptores insulino-tróficos en tejido adiposo
Análogos de incretinas		
Sitagliptina, vidagliptina	Receptor GLP-1	Estimulación de la secreción de insulina
Retardadores de la absorción de glucosa		
Acarbosa	Glicosidasa alfa	Hidrolisis de disacáridos en intestino delgado
Inhibidores DPP4		
Exenatida, liraglutida	Dipeptidil peptidasa tipo 4	Proteólisis de incretinas
Inhibidores de los canales de potasio en la célula beta pancreática (sulfonilureas)		
Glibenclamida	Receptor de sulfonil ureas (SUR)	Secreción de insulina por las células beta pancreáticas
Inhibidores del cotransportador sodio/glucosa SGLT2		
Dapagliflozina	Transportador SGLT2	Reabsorción renal de glucosa

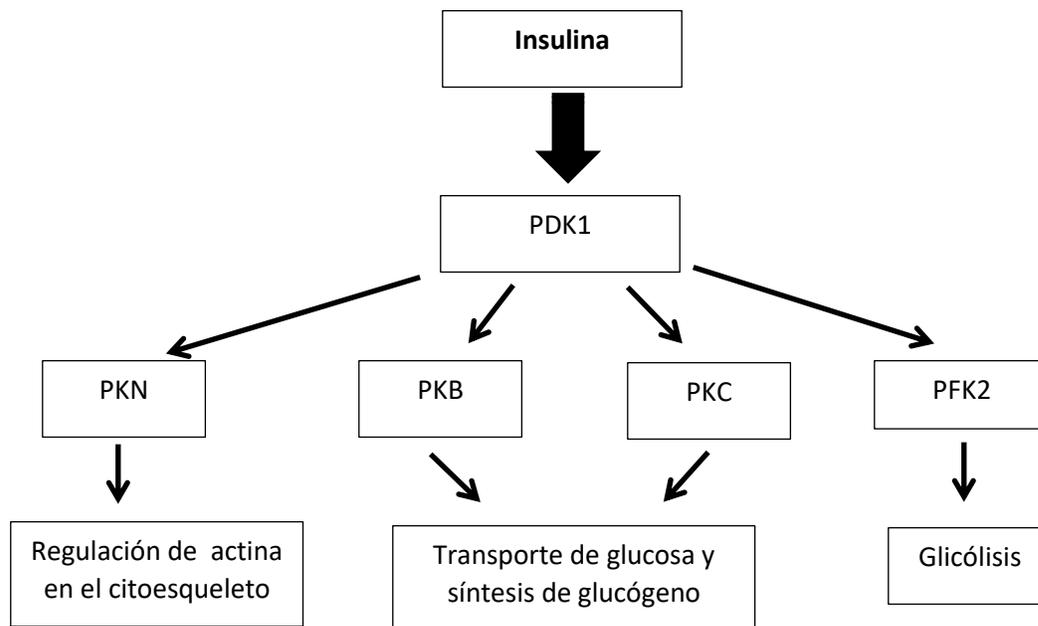
Recientemente los reportes sobre la monoterapia o politerapia farmacológica oral contra la DM tipo 2 y 1 es muy diferente, lo que comparten ambas es que se ha optado por el uso de la insulina basal dentro de los primeros estadios después de que se ha realizado el diagnóstico de la enfermedad<sup>24</sup>, además de que los pacientes que padecen DM tipo 2 tardan más tiempo en

conseguir las metas terapéuticas deseadas. Los dos grandes estudios únicos en su género y en su duración de casi 20 años, el UKPDS y el DCCT<sup>28</sup> han traído grandes avances sobre la detección, el diagnóstico, tratamiento y seguimiento de las complicaciones de la enfermedad. Estos estudios fueron llevados a cabo el primero en el Reino Unido y el segundo en los Estados Unidos de América en donde se probaron la mayoría de las combinaciones terapéuticas conocidas contra DM2 y el manejo de los diferentes tipos de insulina en la DM1, en donde se observó el comportamiento fisiológico y fisiopatológico de diferentes pacientes con diferentes complicaciones y cómo el tratamiento combinado ayudaba a alcanzar los niveles deseados de glucosa haciéndose evidente a través de la medición de la hemoglobina glucosilada HbA1c la cual se le considera ahora como un marcador diagnóstico de DM en los países desarrollados que pueden determinarla por medio de HPLC<sup>29</sup>. Sin embargo, desde la década pasada han aparecido algunos inconvenientes en cuanto al manejo terapéutico de los fármacos orales y existen algunos artículos de revisión especialmente enfocados a los nuevos tratamientos en contra de la enfermedad como los inhibidores GSKT2. En un estudio, se evaluó a la dapagliflozina, inhibidores DPP4 y análogos GLP-1 en combinación cada uno con sulfonilureas y se encontró que la combinación estos últimos con inhibidores SGLT2 llevaban a un control temprano de los niveles glicémicos comparados con la monoterapia con sulfonilureas, pero no con los otros fármacos mencionados. En otro estudio se evaluó la eficacia y seguridad del uso de la Metformina, sulfonilureas, tiazolidinedionas, inhibidores DPP4, análogos a GLP-1, inhibidores de alfa glicosilasa y algunos hipolipemiantes en donde las conclusiones importantes fueron que: la Metformina presenta una excelente eficacia favorable evitando el incremento de peso, y el control de lípidos siendo la primera recomendación como monoterapia, los inhibidores DPP4 mostraron poca eficacia comparados con la Metformina pero demostraron un bajo riesgo de causar hipoglicemia y los demás fármacos mostraron tener un desempeño bajo combinados con los demás fármacos incluso sólo con Metformina incrementando el riesgo de sufrir hipoglicemia<sup>30</sup>.

### 6.3 Proteína cinasa dependiente de fosfatidil inositol 1 (PDK1) y su papel en la diabetes mellitus.

La PDK1 (**Figura 8**) juega un papel muy importante en la respuesta posterior a la estimulación del receptor de insulina y que a su vez provoca la activación de la PKB terminando con la translocación de GLUT 4 en la membrana celular. Se conoce bien el papel que tienen los segundos mensajeros derivados de fosfatidilinositol, trifosfato (PIP3) y difosfato (PIP2) en las

vías de señalización de factores de crecimiento e insulina. Cuando se producen por estimulación de insulina, activan a una serie de serina/treonina proteínas cinasas como la PDK1 que a su vez fosforila a la PKB (Protein kinase B) <sup>31</sup>, concluyendo en la translocación del transportador de glucosa GLUT4 en tejido adiposo y músculo esquelético (**Figura 7**).



**Figura 7** Papel de la PDK1 en la regulación de varias respuestas biológicas provocadas por la insulina. La PDK1 activa a otras cinasas que llevan a término la señalización de la insulina. PKN: Proteína cinasa serina/treonina; PKB: Proteína cinasa tipo B; PKC: Proteína cinasa de tipo C; PFK2: Fosfofructocinasa 2. Modificado de: K.L.R. Wick. *A New Molecular Target of Insulin Action: Regulating the Pivotal PDK1 Current Drug Targets - Immune, Endocrine & Metabolic Disorders*, 2001, 1, 209-221

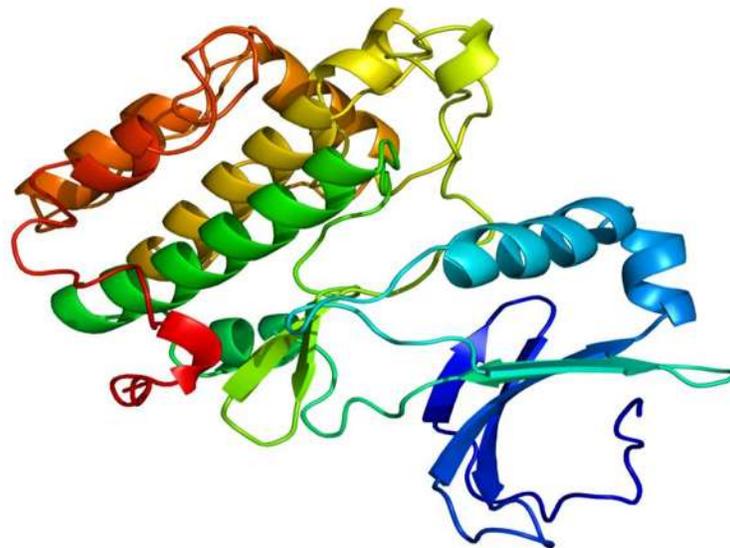
En 1997 se identificó a un nuevo miembro de las cinasas dependientes de fosfatidilinositol que mostró activar a la PKB en respuesta a la estimulación de la insulina. Esta cinasa de serina/treonina fue llamada proteína cinasa dependiente de fosfatidilinositol trifosfato, su gen está localizado en el cromosoma 13.3 16p. Es una proteína de 63 KDa que tiene un dominio catalítico N-terminal y un dominio C-terminal no catalítico llamado dominio PH (pleckstrin homology) al cual se une el fosfatidilinositol trifosfato y difosfato. La regulación de la fosforilación de la PDK1 con otras cinasas se da en el fragmento de interacción con la PDK1 o también llamado PIF (PDK1 Interacting fragment) <sup>32</sup>, el cual ha sido objeto de estudio como

diana terapéutica para varios desordenes que tienen relación con la señalización de la insulina, entre ellos la diabetes mellitus.

### 6.3.1 Relación de la PDK1 con el transportador de glucosa GLUT4

La señalización de la insulina se da a través de una serie de cinasas en particular la PDK1 (de la cual se habla más adelante), en donde el efecto final sobre los carbohidratos es el transporte de glucosa por la translocación de GLUT4 (Transportador de glucosa) perteneciente a una familia de transportadores de la que se han descrito 13 genes que expresan moléculas diferentes, de todas ellas GLUT4 es la mejor conocida, y según los datos actuales la más importante, pues representa el 90% de los transportadores de glucosa en el adipocito y el 75% de los transportadores que se encuentran en la membrana plasmática tras el estímulo de la misma<sup>25</sup>.

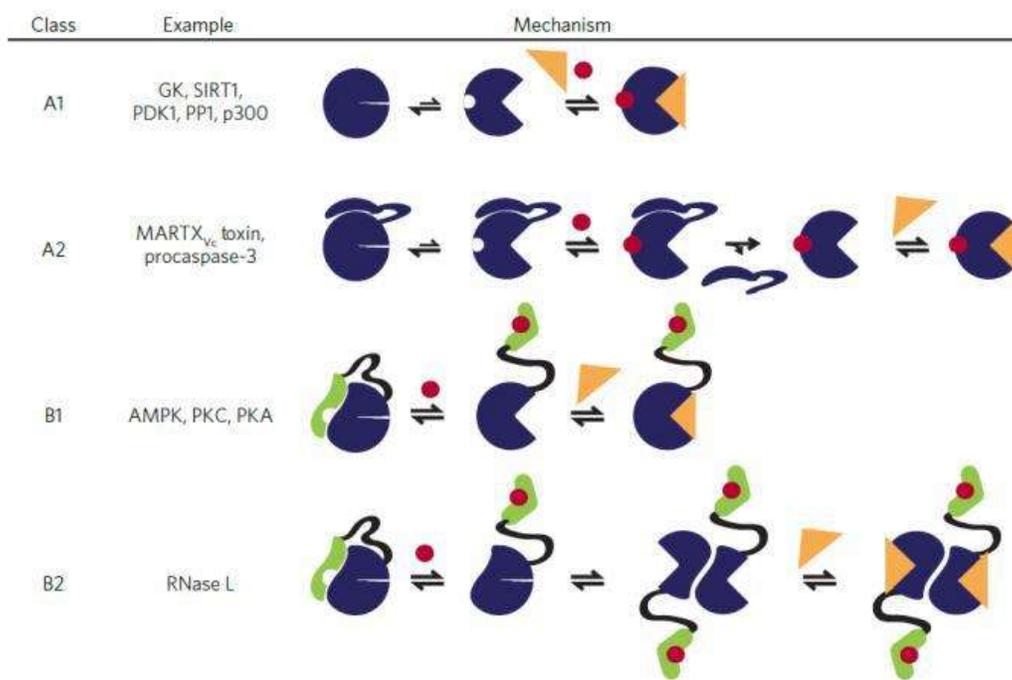
Los estudios previos de Phillip Cohen y cols. (1997) sobre el papel que juegan las cinasas en las enfermedades como el cáncer y la diabetes mellitus entre otras, abrieron la puerta al estudio de posibles planteamientos farmacológicos a través del diseño de compuestos basados en las interacciones con la región PIF que probablemente activen a la PDK1<sup>33</sup>.



**Figura 8** Proteína cinasa dependiente de inositol trifosfato (PDK1) Código PDB: 4AW0

### 6.3.2 Activando enzimas

Dentro de la investigación del desarrollo y búsqueda de compuestos que puedan interactuar con enzimas, específicamente con las cinasas se enfoca a buscar inhibidores, no obstante los compuestos que se comportan como activadores, proveen de una mejor comprensión de los mecanismos de interacción alostérica, siendo las cinasas las enzimas que poseen un mayor número de sitios alostéricos para poder tener una regulación adecuada, ya sea en la forma activa o en la forma inactiva en respuesta a una vía de señalización celular. Julie A Zorn *et al.* reportaron una revisión acerca de los compuestos que se comportan como activadores de cinasas (**Figura 9**), y subrayaron que existe sólo una docena de compuestos caracterizados cabalmente en cuanto a la forma de interactuar con las cinasas, y ellos clasificaron los mecanismos por los cuales compuestos de bajo peso molecular activan a las cinasas siendo la PDK1 una de las cinasas en las que se conocen algunos activadores de bajo peso molecular<sup>34</sup>.



**Figura 9** Mecanismo de acción que explica la activación de las algunas cinasas por parte de los compuestos de bajo peso molecular

Una de las ventajas de los compuestos activadores de cinasas de bajo peso molecular comparados con inhibidores para las mismas enzimas, es que interactúan de forma alostérica con dichas enzimas, lo cual nos conduce a tener una menor probabilidad de que estos sitios bien conocidos se modifiquen en su secuencia de aminoácidos, ofreciendo así mejores oportunidades para generar compuestos con alta especificidad.

### 6.3.3 PDK1 como una posible diana farmacológica contra diabetes mellitus.

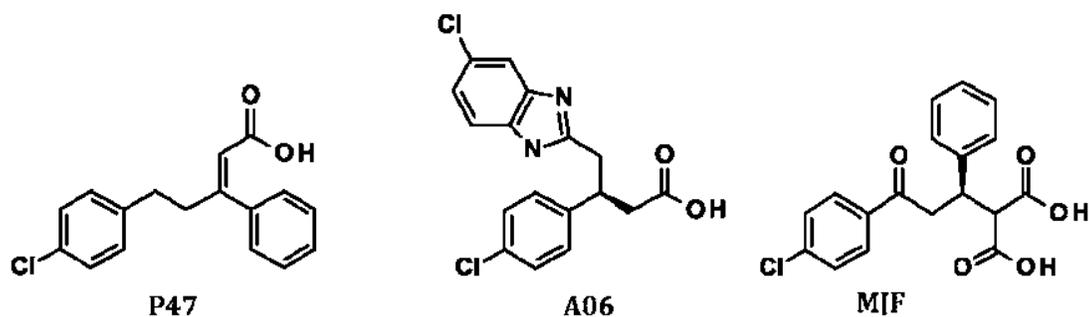
Lawlor et al. (2002) hicieron los primeros intentos de estudiar las funciones fisiológicas de PDK1 en respuesta a las señales de insulina en el organismo<sup>35</sup>. La mayoría de los tejidos de mamíferos llevan a cabo el control del gasto energético de acuerdo a la capacidad de producción de ATP con cualquiera de los sustratos disponibles o en las condiciones hormonales en estado fisiológico y en un tiempo específico. Las hormonas como la insulina juegan un papel importante en hidratos de carbono y ácidos grasos, especialmente a nivel de la glicólisis y la lipólisis. La glucólisis es mediada hasta un punto específico para la acción de la insulina a través de la translocación de glucosa por el GLUT4 y la activación de la PFK2. El primer informe de la evidencia genética que PDK1 media la activación de insulina de 6-fosfofructoquinasa-2 fue reportado por Alfonso Mora et al. (2003), donde se evaluaron el papel PDK1 en ratones con y sin PDK1, la inyección de la insulina y los diferentes tiempos y, al mismo tiempo que se extrae el corazón y medición de los niveles y la actividad PFK2 y fructosa 2,6-bisfosfatasa. En ratones con PDK1, PFK-2 actividad fue estimulada casi 2 veces dentro de 5 min de la inyección de insulina y la acumulación de fructosa-2,6-bisfosfato se detectó después de 10 min, alcanzando 5 veces los niveles más altos que en mPDK1 no estimulada (+ / +) corazones después de 20 min. En los corazones de ratones sin PDK1, ningún aumento significativo en la actividad PFK-2 o fructosa niveles de 2,6-bisfosfato se detectó después de una inyección de insulina<sup>36</sup>.

Bayascas Jose R et al. estudiaron la influencia de PDK1 en resistencia a la insulina señalando que la vía de señalización PDK1 juega un papel crucial en el control de la homeostasis de la glucosa. Generaron un modelo de ratones mutados en la PDK1 (K465E). Esto conduce a un daño de señalización específica, reduciendo la actividad de las isoformas de PKB, lo que resulta en la aparición de resistencia a la insulina en ratones modificados para producir la enzima presentando deterioro en la respuesta a la insulina conduciendo a la intolerancia a la glucosa que normalmente preceden a la aparición de la diabetes tipo 2. Sin embargo, la activación de la

misma vía de señalización por factores de crecimiento extracelulares conduce a la desregulación del crecimiento, la supervivencia, la proliferación, la apoptosis<sup>37</sup>.

#### Primeros compuestos activadores de la PDK1

Tres compuestos de bajo peso molecular se han desarrollado dirigidos a la región PIF de la PDK1 (**Figura 10**). Estos compuestos tienen la capacidad de activar específicamente PDK1 mediante la imitación de la transición conformacional que se produce normalmente tras la interacción de PDK1, sin afectar directamente a la actividad catalítica intrínseca de otras quinasas AGC relacionadas<sup>38</sup>.



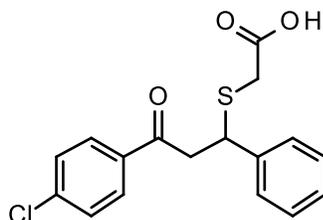
**Figura 10** Estructura química de los compuestos reportados como activadores de la PDK. El nombre es como aparece en las estructuras cristalográficas depositadas en la base de datos de proteínas (PDB).

Matthias Engel y cols. (2006) concluyen en su trabajo que la esencia de los mecanismos que involucran la unión de un ligando en el sitio regulatorio PIF la posibilidad de diseñar nuevos compuestos que activen a la PDK1 y que en un futuro pudieran desarrollarse fármacos en base a los resultados obtenidos. Tomaron como base de datos a 60, 000 modelos de compuestos de bajo peso molecular desarrollados en May bridge (Tintagel) y fueron optimizados por el método computacional Unity 4.3 (Tripos), entre 250 y 2,500 compuestos mostraron predicción de poseer características de unión a la región PIF. La selección de los compuestos fue evaluada y sólo 220 fueron probados in vitro. Los resultados revelaron que sólo dos

compuestos incrementaban significativamente la actividad de la PDK1 frente a un sustrato polipeptídico similar a los residuos que se fosforilan en la PKB. Así pues, seleccionaron al compuesto denominado “**Compuesto 1**” (**Figura 11**), mostrando una activación máxima de la PDK1<sup>39</sup>.

Stroba et al. 2009 informaron los resultados de describen la síntesis, potencia la activación y el análisis de unión calorimétrico de los compuestos clasificados como moduladores alostéricos del PDK1, con base en los estudios de estructura-actividad. Los resultados termodinámicos demuestran una confirmación de que la región PIF de la PDK1 es un sitio accesible para el desarrollo de fármacos y gracias a esto se han podido definir las características químicas estructurales mínimas para los candidatos a activadores para la PDK1<sup>38</sup>.

#### *Primer diseño de compuestos potencialmente activadores de la PDK1*

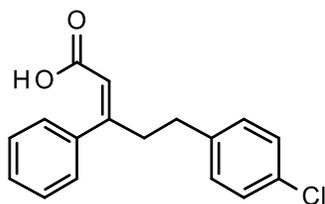


**Figura 11** Ácido 2 - ((3 - (4-clorofenil)-3-oxo-1-fenil-propil) tio) acético.

La conclusión de este estudio fue que un desarrollo racional de compuestos de bajo peso molecular como el compuesto 1 proporcionan los requerimientos mínimos necesarios para que haya una interacción con la región PIF y se logre una activación de la enzima PDK1.

#### *Segundo diseño de compuestos probablemente activadores de la PDK1*

Adriana Stroba y cols. (2009) propusieron una serie de compuestos similares a los de Matthias Engel y cols. Que mostraron actividad en la región PIF de la PDK1, de ellos sobresale el compuesto **2E-Z** presentado en la **Figura 12**.



**Figura 12** Estructura del compuesto **2E-Z** propuesto por Adriana Stroba y cols.

En base a lo que se investigó sobre compuestos que pudieran comportarse como activadores de la PDK1, se decidió tomar en cuenta los propuestos en la publicación de Matias, Engel y cols. De los que se les conoce la actividad  $AC_{50}$  frente a la PDK1, además algunas otras propiedades termodinámicas como constante de inhibición y energía de unión con la región PIF de la PDK1. Es muy importante destacar que para poder emplear las herramientas del DIFAC se puede abordar la investigación empleando métodos basados en el ligando, es decir emplear los compuestos anteriormente señalados para realizar estudios de estructura actividad (QSAR) además de emplear métodos basados en la estructura pues se cuenta con la información de estructuras cristalográficas de la PDK1 en complejo con algunos de los ligandos antes mencionados. Esto nos dio la pauta para poder plantear la metodología empleada durante toda la investigación.

Se realizó una búsqueda de algunos compuestos que ya hubieran sido probados como activadores o inhibidores de la PDK1 enfocándose mayoritariamente por aquellos que se comportan como activadores encontrando así el sitio alostérico PIF y la estructura básica de compuestos que pudieran tener interacciones en esa región.

## 7. PLANTEAMIENTO DEL PROBLEMA

La diabetes mellitus es una enfermedad no transmisible que está avanzando de forma rápida a pesar de las estrategias en su prevención, diagnóstico y manejo farmacológico. Se estima que para el año 2030 haya un incremento significativo en el número de personas que serán afectadas con esta enfermedad. La prevención debería de ser la mejor forma de evitar que aparezcan nuevos casos de la enfermedad, sin embargo, no ha sido suficiente debido en gran parte a las condiciones metabólicas que han evolucionado adaptándose a las condiciones hormonales y energéticas que enfrenta el cuerpo de un individuo con predisposición a la enfermedad. Es en parte por ese motivo que es necesario acelerar los procesos científicos que nos lleven a tener mejores estrategias en el campo de la farmacología con nuevos enfoques de dianas terapéuticas que no han sido exploradas esperando tener mejores resultados evitando los efectos adversos de los hipoglicemiantes ya conocidos e intentando atacar a la enfermedad en sus dos tipos.

## 8. JUSTIFICACIÓN

El papel de la PDK1 en la señalización de los efectos de la insulina tales como la translocación del transportador de glucosa GLUT4 abre un campo de investigación en el diseño y búsqueda de compuestos que puedan interactuar con la PDK1 en su región PIF, siendo así una probable nueva diana farmacológica de la cual necesitamos más estudios sobre el diseño de fármacos asistido por computadora (DIFAC), como se propone en el presente trabajo.

Por lo tanto, nos enfocaremos en el diseño y síntesis de nuevos compuestos que puedan interactuar con la PDK1 esperando que tengan un efecto activador provocando la translocación de GLUT4.

## 9. HIPÓTESIS

A partir del uso y combinación de estudios in silico aplicados al diseño de fármacos puede contribuir a la obtención de nuevos compuestos que puedan ser potencialmente activadores de la PDK1 para continuar con estudios que puedan demostrar la actividad biológica y su probable aplicación destinada a combatir la diabetes mellitus

## 10. OBJETIVOS

General:

Diseñar y encontrar compuestos que puedan interactuar la PDK1 a través de métodos in silico comunes del DIFAC

Específicos:

1. Predecir in silico la interacción que tendrían algunos compuestos propuestos derivados de la chalcona que pudieran presentar interacciones con la PDK1 a través de estudios de acoplamiento molecular.
2. Encontrar compuestos con estructura química diferente a la chalcona y a los ya descritos como activadores de la PDK1 a través de estudios de cribado virtual.
3. Predecir la actividad  $pAC_{50}$  que pueden tener de los mejores compuestos encontrados en el cribado virtual a través de estudios de relación estructura-actividad.

## 11. METODOLOGÍA

Para poder lograr los objetivos propuestos se emplearon como métodos del DIFAC:

Acoplamiento molecular (docking)

Modelado farmacofórico

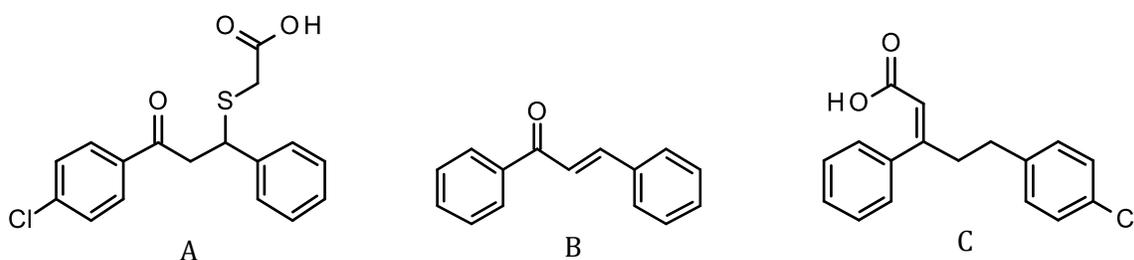
Cribado virtual

Filtrado virtual de compuestos utilizando PAINS

Estudios de relaciones estructura actividad (QSAR)

## 12. PRIMEROS ESTUDIOS EN BUSCA DE COMPUESTOS PROBABLEMENTE ACTIVADORES DE LA PDK1 CON DERIVADOS DE CHALCONA

El inicio sobre el estudio de compuestos que puedan ser probablemente activadores de la PDK1 comenzó en base a las estructuras químicas reportadas previamente por Mathias Engel y cols. que a simple vista tienen la base estructural similar a la chalcona (**Figura 13**).



**Figura 13** Comparación de las estructuras químicas entre los compuestos conocidos como activadores de la PDK1 (**A** y **C**) y la chalcona. (**B**)

Cabe señalar que para este tiempo en que estábamos desarrollando el trabajo, no se pensó en realizar alguna búsqueda de compuestos ya sintetizados que nos permitiera tener una gamma un poco más amplia de lo que probablemente estaríamos encontrando para la PDK1, así que se comenzó a desarrollar la búsqueda de algún compuesto derivado de chalcona por métodos computacionales accesibles y con un alto grado de eficiencia empleando análisis de acoplamiento molecular

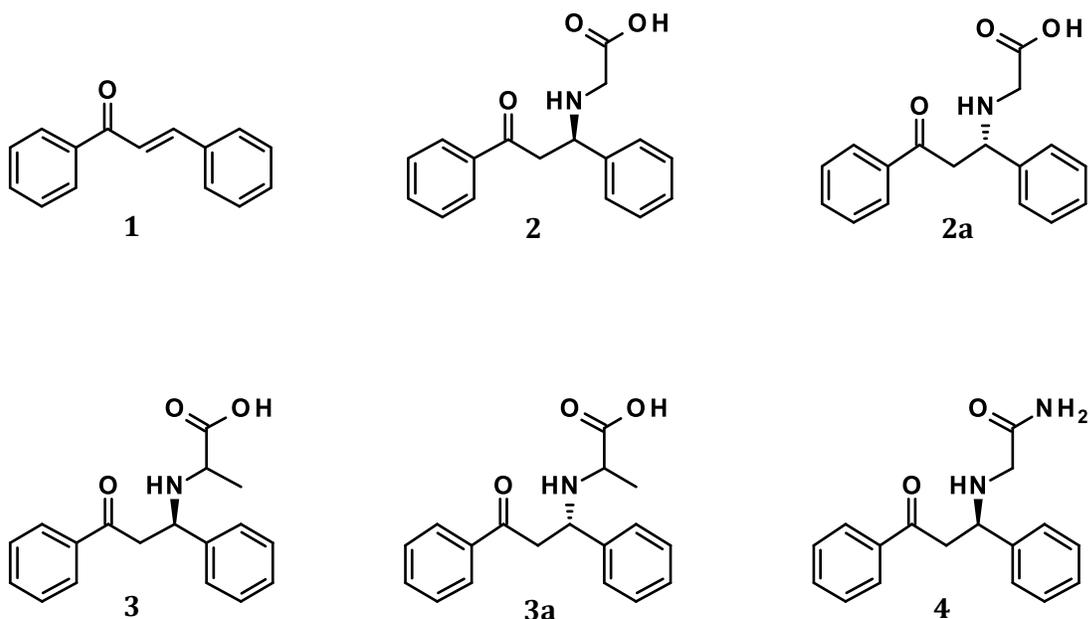
Este análisis condujo a la posibilidad de utilizar a la chalcona como probable estructura química de base para poder realizar modificaciones químicas aprovechando la cetona alfa-beta insaturada, además de que ya se había trabajado en el grupo de trabajo con chalconas, se conocía la reactividad de la misma.

*Propiedades farmacológicas de las chalconas.*

Las chalconas son consideradas precursores de los flavonoides<sup>40</sup>, las cuales están formadas por dos anillos bencénicos conectados por tres átomos de carbono formando una cetona aromática  $\alpha, \beta$ -insaturada. La presencia del grupo carbonilo y la unidad olefínica conjugados confieren a las chalconas gran reactividad, situándolas como una interesante clase de moléculas a estudiar tanto en las proyecciones en ciencia básica, como por las aplicaciones en áreas como la medicina, la agricultura y la industria. Investigaciones previas han reportado una importante actividad biológica para estos compuestos, donde cabe resaltar sus propiedades antimicrobianas anticancerígenas, antiinflamatorias y antineoplásicas, entre otras<sup>41 y 42</sup>.

**Material y métodos**

Se propuso una serie de nuevos compuestos derivados de la chalcona (**Figura 14**), a partir de las posibilidades de incluir nucleófilos a la cetona alfa beta insaturada.



**Figura 14** Compuestos propuestos de novo, derivados de la chalcona

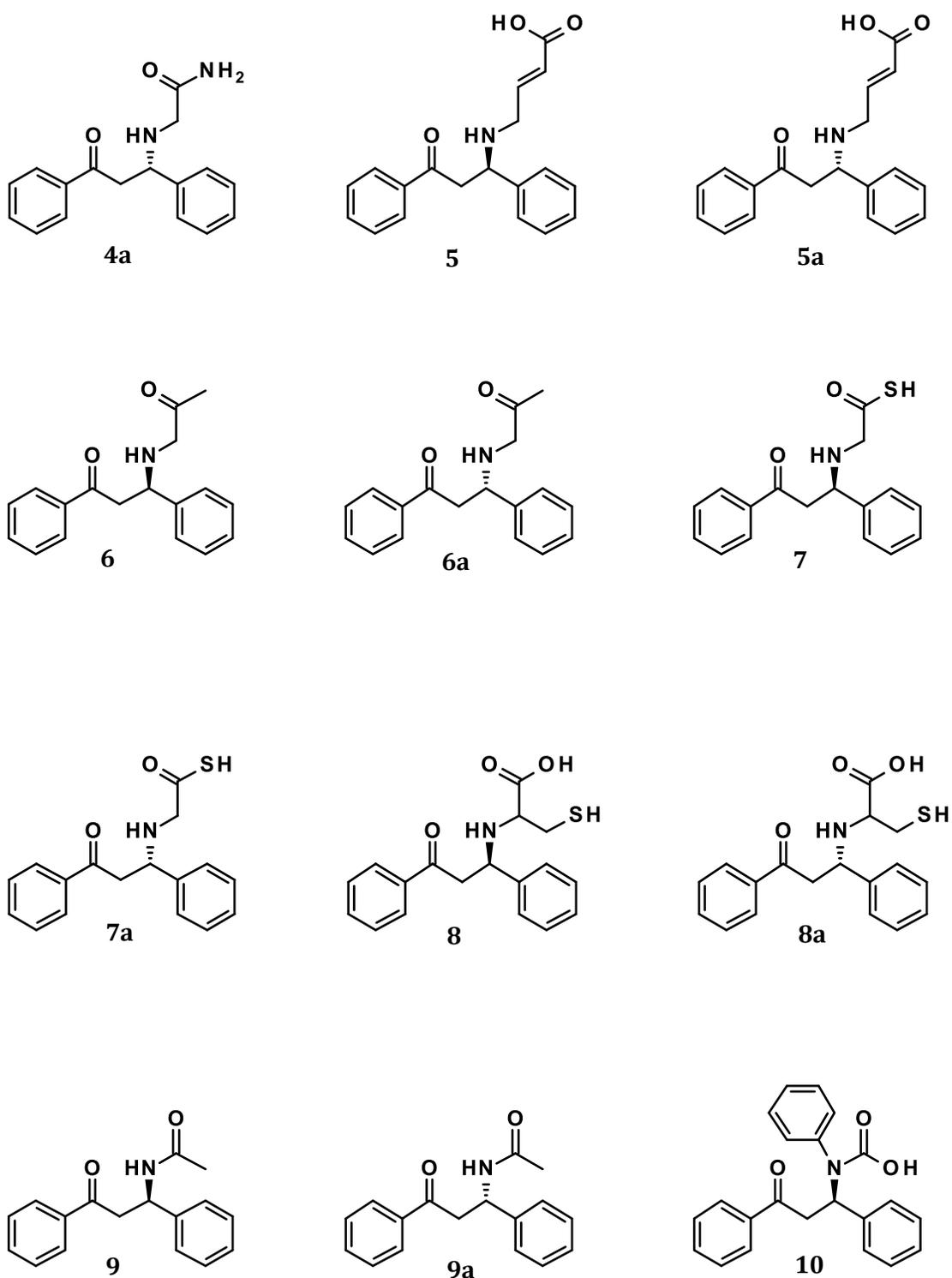
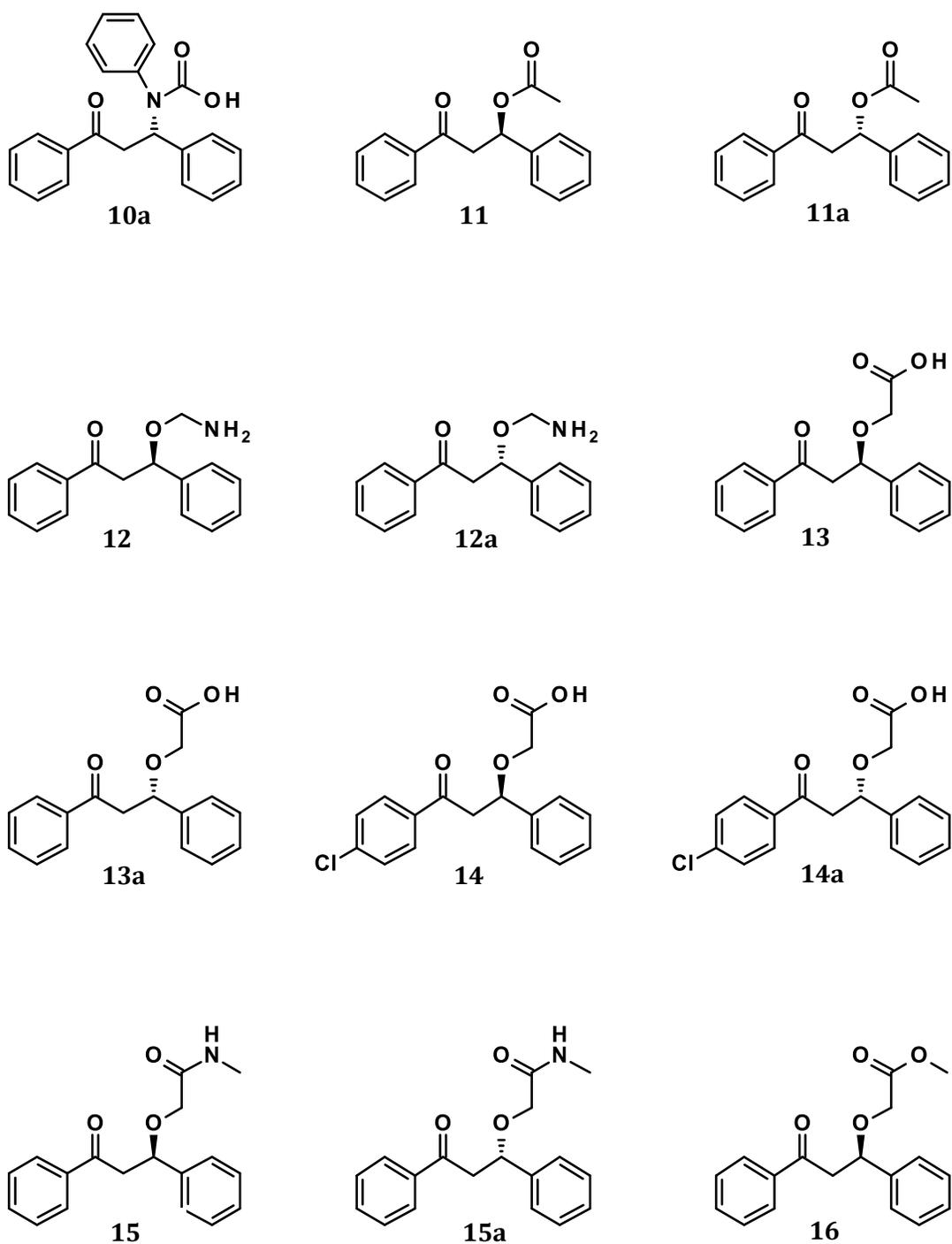


Figura 15 Compuestos propuestos de novo, derivados de la chalcona



**Figura 14.** Compuestos propuestos de novo, derivados de la chalcona (Continuación).

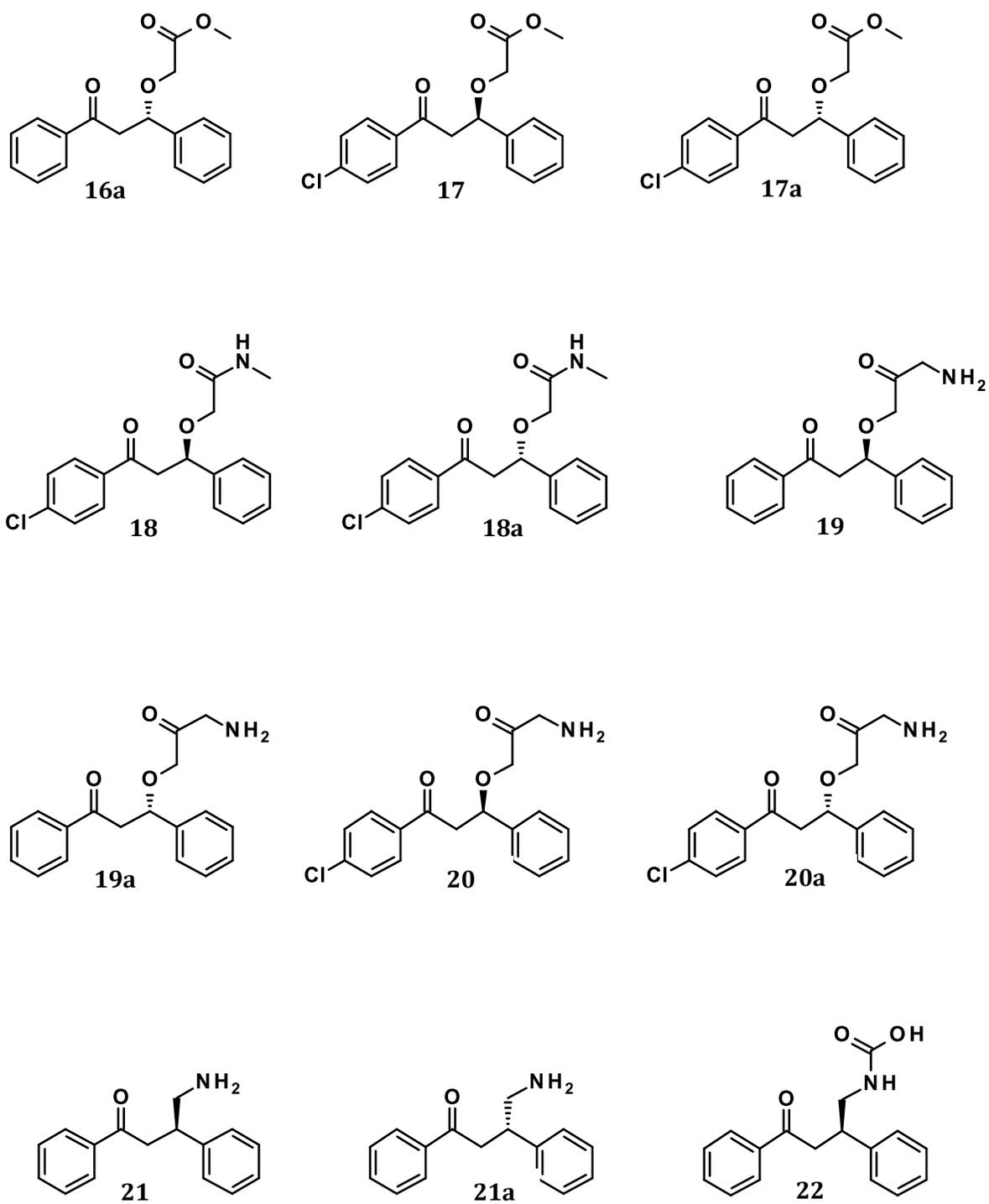


Figura 14. Compuestos propuestos de novo, derivados de la chalcona (Continuación).

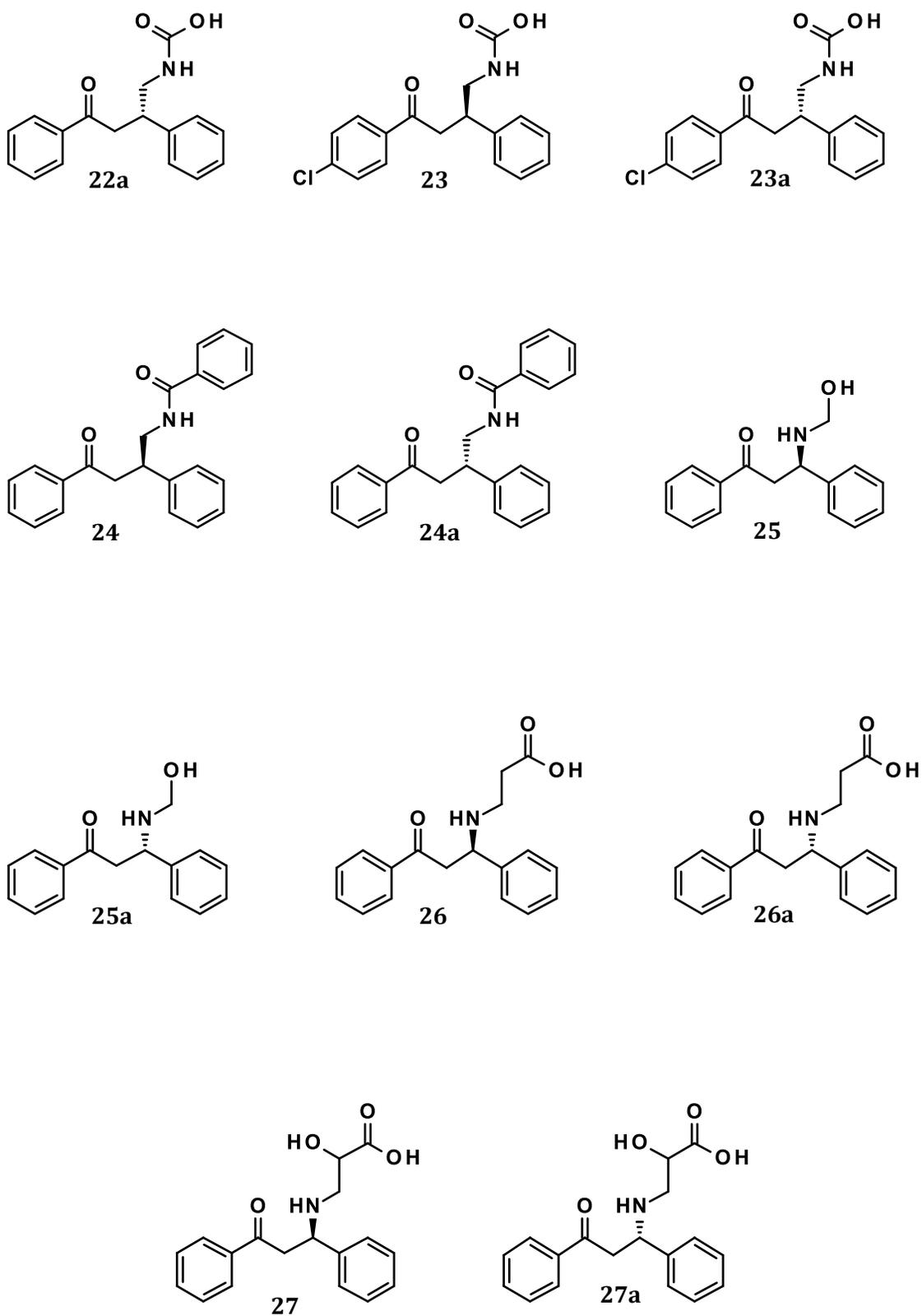


Figura 14. Compuestos propuestos de novo, derivados de la chalcona (Continuación).

Una vez propuestos los compuestos se modelaron utilizando métodos computacionales. Es bien conocido que en el DIFAC no es sólo apretar botones como en algunas ocasiones se afirma, cada proyecto de investigación que involucra estudios de este tipo, es muy diferente por tanto las condiciones en cómo modelar o cómo preparar un archivo que contiene información de una estructura química casi siempre está en función del objetivo que se pretende, atendiendo y cuidando desde luego los aspectos químicos importantes para la buena obtención de la mejor aproximación de los resultados.

#### *Modelado de los compuestos propuestos*

Los compuestos propuestos en la **figura 14** fueron dibujados en dos dimensiones empleando el programa ChemBioDraw Ultra 12.0, después en el programa Chems sketch se obtuvo un archivo con las coordenadas tridimensionales x, y, z en formato .mdl de donde se les determinó su estado de protonación a pH de 7.4 empleando el servidor chemicalize obteniendo así las estructuras de partida para poder realizar la optimización de su geometría con el programa Gaussian 03W empleando como base de cálculo el método semi empírico de tipo Dewar AM1. Es importante resaltar que en algunas ocasiones se cuestiona la forma de partida de las estructuras químicas en cuanto a sus coordenadas tridimensionales, puede ser lógico que deba partirse de una serie de confórmeros con la mínima energía intramolecular, sin embargo, dichos cálculos aplican para poder estudiar otras propiedades fisicoquímicas, de reactividad y de distribución electrónica, lo que no nos compete en los estudios y experimentos computacionales en el diseño de fármacos, pues no siempre, de hecho, es bien sabido que la conformación activa de un compuesto de bajo peso molecular cuando interactúa con su receptor, no siempre es la de la menor energía, pudiera estarse pensando en realizar cálculos ab initio tomando en cuenta el ambiente químico del sitio de interacción así como los solventes y/o iones que pudieran estar involucrados en dicha interacción siendo este tipo de estudios, los experimentos de dinámica molecular en el DIFAC.

#### *Estudios de acoplamiento molecular*

Como los cálculos computacionales con estructuras químicas son aproximaciones que se realizan basados en las propiedades químicas que son estudiadas ya sea métodos de mecánica molecular o clásica o bien por métodos ab initio. Por lo tanto, las condiciones de partida en

cualquier cálculo computacional que involucre el uso de estos métodos, es decir, las coordenadas tridimensionales, la carga eléctrica, etc. son importantes para poder llevar a cabo una predicción lo más próxima a la resolución de las matemáticas que emplea cada método del DIFAC, por ejemplo, nuestro trabajo de acoplamiento molecular involucra solamente el uso de ecuaciones de la mecánica clásica tomando en cuenta las fuerzas de atracción y de repulsión entre otras como se presentan en la siguiente ecuación:

$$\begin{aligned} \Delta G = & \Delta G_{vdw} \sum_{i,j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \\ & + \Delta G_{hbond} \sum_{i,j} E(t) \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ & + \Delta G_{elec} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{tor} N_{tor} \\ & + \Delta G_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{\epsilon (r_{ij}^2 / 2 \sigma^2)} \end{aligned}$$

**Ecuación 1** Cálculo de energía libre total de un sistema derivado de un estudio de acoplamiento molecular (este representa a la función de puntaje utilizada en Autodock4)

Donde los cinco términos de  $\Delta G$  a la derecha después del signo de igualdad, son coeficientes empíricamente determinados utilizando análisis de regresiones lineales a partir de una serie de complejos proteína-ligando con las constantes de unión determinadas experimentalmente y que fueron tomadas para el desarrollo y desempeño de Autodock4 que fue utilizado durante los experimentos de acoplamiento molecular en esta parte de la investigación. La sumatoria de los diferentes términos de  $\Delta G$  se determinan para cada par de átomos del ligando ( $i$ ) además de los átomos del receptor ( $j$ ) además de todos los átomos que están separados por tres enlaces dentro del ligando.

También para la adecuada preparación de los ligandos y receptores es necesario tomar en cuenta las interacciones de atracción y repulsión entre el ligando y el receptor, por lo que como en la mayoría de los programas de acoplamiento molecular, se emplea el uso de los potenciales de Lennard-Jones 12-6 de dispersión/repulsión incluyendo la medida de los ángulos de enlace, además de los cálculos del potencial electrostático coulombico<sup>53</sup>.

### *Preparación de los ligandos*

Para los estudios de acoplamiento molecular, la preparación de los ligandos fue empleando la herramienta MGLTools de Autodock en donde se tomó en cuenta para cada ligando las coordenadas de salida de la optimización de su geometría, se adicionaron cargas Gasteiger así mismo se verificó la selección de los enlaces rotables (de libre giro) para los cálculos conformacionales iniciales al final se obtuvieron archivos .pdbqt necesarios para los estudios de acoplamiento molecular.

### *Elección de las estructuras cristalográficas como receptores*

Otra parte importante de los estudios de acoplamiento molecular es en relación a las características de partida del receptor en el que se pretende investigar las interacciones de los ligandos, así que se exploró la Base de Datos de Proteínas ampliamente conocida y de gran utilidad desde 1971 que contiene estructuras tridimensionales de proteínas y ahora de ácidos nucleicos en base a cristalografía de rayos X o bien resultados de Resonancia Magnética Nuclear, siendo la primera nuestra principal elección para tomar en cuenta un archivo con esta información para la PDK1, y para seleccionar de una forma adecuada la estructura cristalográfica se tomó en cuenta lo siguiente: que contenga un ligando en el sitio de interés que en este caso es la región PIF; que tenga una resolución no mayor a 2Å ; que esté completa la secuencia de aminoácidos en el sitio de interés; que las estructuras cristalográficas tengan una identidad mayor al 95%

En base a lo anterior la **tabla 2** muestra las características de cada estructura cristalográfica.

**Tabla 2.** Coordenadas de muestreo y dimensiones de la malla generada en Autodock Tools y utilizada en los estudios de acoplamiento molecular.

Receptor	Centro			Dimensiones (Å)			Distancia de búsqueda (Å)
	X	Y	Z	X	Y	Z	
<b>3HRF</b>	24.157	-22.870	-25.608	126	126	126	0.37
<b>4A06</b>	24.013	-23.030	-25.849	126	126	126	
<b>4AW0</b>	-19.997	20.977	2.856	126	126	126	

*Preparación de las estructuras cristalográficas de los receptores*

Se decidió incluir a los tres mejores receptores encontrados para la PDK siendo los códigos PDB: **3HRF**, **4A06** y **4AW0** los más aptos para nuestro estudio de acoplamiento molecular. De igual forma que los ligandos, se hizo la preparación de estas estructuras cristalográficas adicionando primero los átomos de hidrógeno que son de carácter polar y que pueden estar ausentes en las estructuras cristalográficas como es bien conocido, además se seleccionó realizar un estudio de acoplamiento molecular de tipo ciego para todos los ligandos con cada uno de los receptores teniendo como medida de una malla de cálculo con espacio de 126 X 126 X 126Å y con un espaciado de cálculo cada 0.37Å , los archivos de los receptores se obtuvieron en .dpbqt para los cálculos posteriores.

Para poder realizar los cálculos de acoplamiento molecular, es necesario conocer la naturaleza de los ligandos que serán explorados y estudiados en cuanto a las interacciones con el receptor, por lo que se conocen bien los enlaces rotables como se dijo anteriormente, y en sin dudas para los cálculos de la energía y para las interacciones electrostáticas colúmbicas, los parámetros atómicos dados para cada átomo componente tanto de los ligandos como del receptor, por lo que se incluyó en los cálculos, el archivo de parametrización AD4 que incluye toda la información necesaria para los ligandos y receptores, y una vez con esto, se procedió a preparar los archivos .gpf y .dpf que son los archivos que contienen la información de cada uno de los ligandos con cada receptor y que son necesarios para que se lleven los cálculos de forma separada para cada ligando con diferente receptor. Los cálculos de acoplamiento molecular fueron realizados en un equipo Mac con un sistema operativo OS X con un procesador de 2.6 GHz Core i7 y una memoria de 8GB. Por último, para la preparación de los archivos de entrada del estudio de acoplamiento molecular, se empleó como método de búsqueda un algoritmo genético Lamarckiano que incluyó los siguientes parámetros de búsqueda: número de algoritmos genéticos: 100; total de la población inicial de búsqueda; 100; número de eventos permitidos:  $1 \times 10^6$ ; tasa de mutación: 0.08; tasa de entrecruzamiento: 0.2; número máximo de mejores poses o resultados posibles; 100. Como función de puntaje el programa Autodock4 emplea el cálculo de la energía libre en base a las energías empíricas calculadas durante el desarrollo y la validación del programa, además de que realiza el cálculo de una constante de inhibición  $K_i$  a partir de la ecuación 2:

$$\Delta G = RT \ln K_i$$

**Ecuación 2.** Expresión termodinámica general que nos permite el cálculo de cantidades

Donde R es la constante de los gases  $1.987 \text{ cal K}^{-1} \text{ mol}^{-1}$  y T es la temperatura absoluta

#### *Validación de los estudios de acoplamiento molecular*

Para el tiempo en que fueron realizados estos primeros cálculos con los ligandos propuestos y los receptores mencionados arriba, se conocía dentro del grupo de investigación acerca de las validaciones de los estudios de acoplamiento molecular por lo que la validación de estos estudios fue realizando el cálculo del RMSD (Root Mean Square Deviation) por sus siglas en inglés, que es la raíz de la desviación cuadrática media entre los átomos pesados tanto del ligando en complejo con el receptor del cual se conocen las coordenadas por medio de cristalografía de rayos X y el mismo ligando pero sometido a el protocolo de acoplamiento molecular llevado a cabo en nuestro laboratorio para todos los ligandos propuestos por lo tanto, se sometió a los ligandos **P47**, **A06** y **MJF** al estudio de acoplamiento molecular con las características y condiciones antes descritas, y una vez obtenidos los resultados se calculó el RMSD que se presentan en la tabla 3 y la comparación tridimensional se presenta en la figura 10.

#### **Resultados**

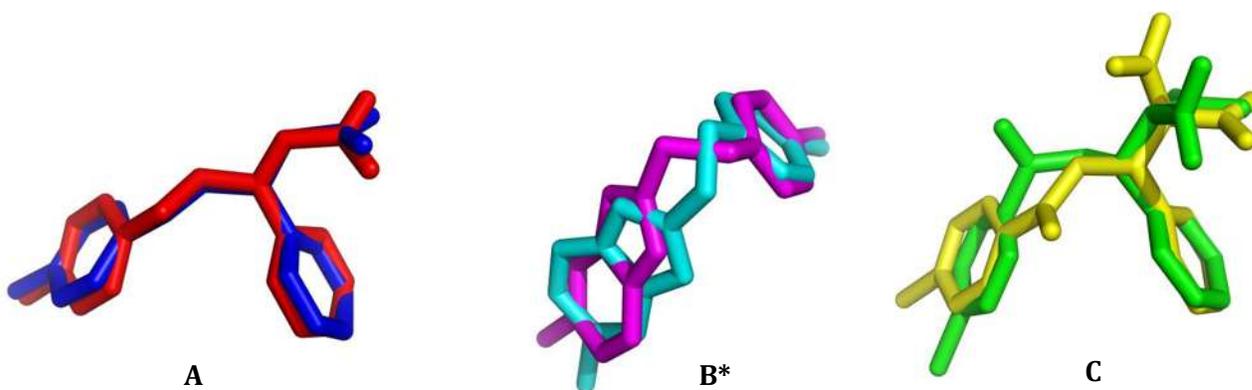
Validación de los estudios de acoplamiento molecular, resultados de RMSD

Una vez que se hicieron los cálculos de acoplamiento molecular, se obtuvieron las posiciones junto con las interacciones de cada ligando con su respectivo receptor, de los que se les calculó el valor de RMSD empleando el programa Pymol tomando como ligando base o de referencia aquel que se obtuvo de las coordenadas de la estructura de rayos x, y como analizado, aquellos ligandos que fueron preparados y sometidos al análisis de acoplamiento molecular con las condiciones descritas en el material y métodos de esta parte del trabajo. Los resultados obtenidos para la validación del protocolo a utilizar del acoplamiento molecular se resumen en la **tabla 3** y la conformación obtenida para cada ligando se presenta en la **figura 15**

**Tabla 3.** Resultados del cálculo de RMSD teniendo como referencia el ligando de su correspondiente estructura cristalográfica.

Ligandos analizados		Número de átomos alineados	RMSD (Å) <sup>a</sup>
Referencia	Analizado		
P47	3HRC_P47	20	1.673
A06	4A06_A06	17*	1.499
MJF	4AWO_MJF	23	1.852

<sup>a</sup> Los resultados fueron obtenidos en el software PyMOL. \*Ver nota en el pie de la figura 1



**Figura 15** Representación 3D de los ligandos utilizados en el cálculo del RMSD para la validación del estudio docking. **A:** P47; rojo = referencia; azul = analizado, **B:** A06; ciano = referencia; magenta = analizado, **C:** MJF; amarillo = referencia; verde = analizado.

El cálculo del RMSD que se realizó con cada ligando y su respectivo receptor, fueron satisfactorios siendo de 1.673, 1.499 y 1.852 para los ligandos **P47**, **A06** y **MJF** respectivamente, lo que nos dio paso a realizar los estudios de acoplamiento molecular con las mismas condiciones de cálculo para los demás ligandos derivados de chalcona propuestos.

Después de realizar la validación de los estudios de acoplamiento molecular, para cada ligando se hizo el cálculo de los potenciales electrostáticos empleando la función autogrid4, y posterior a esto se utilizó la función autodock4 para realizar el cálculo de la energía intramolecular e intermolecular de los complejos ligando-receptor. Después de terminados los cálculos, se obtuvieron los correspondientes archivos .dlg que contienen la información sobre la predicción de las poses, la energía libre y el Ki calculados. Los resultados fueron analizados de forma manual utilizando AutodockTools como herramienta de visualización gráfica para apreciar las mejores 100 poses de cada ligando, así como la energía libre de cada uno y su respectiva Ki

calculada siendo guardado el mejor resultado, es decir aquel que mostró menor energía libre y un valor menor de Ki.

En las **tablas 4, 5 y 6** se presentan los resultados obtenidos de los estudios de acoplamiento molecular con los compuestos propuestos derivados de la chalcona.

**Tabla 4.** Resultados del acoplamiento molecular de los 53 ligandos analizados con el receptor **3HRF** se incluye en el inicio los resultados del ligando de referencia.

Ligando	¿Hubo unión a la región PIF?	Rank_subrank	Energía libre (kcal/mol)	Ki (μM)
<b>P47</b>	SI	1_1	-8.04	1.29
<b>1</b>	SI	8_1	-5.76	59.99
<b>2</b>	SI	12_1	-5.24	145.06
<b>2a</b>	SI	38_1	-4.35	642.26
<b>3</b>	SI	15_1	-4.13	934.35
<b>3a</b>	SI	2_1	-6.4	20.23
<b>4</b>	NO	-	-	-
<b>4a</b>	NO	-	-	-
<b>5</b>	NO	-	-	-
<b>5a</b>	SI	22_1	-5.97	42.39
<b>6</b>	NO	-	-	-
<b>6a</b>	NO	-	-	-
<b>7</b>	NO	-	-	-
<b>7a</b>	SI	56_1	-4.5	493.5
<b>8</b>	SI	4_1	-5.85	51.75
<b>8a</b>	SI	2_1	-6.42	19.67
<b>9</b>	SI	19_1	-5.54	86.27
<b>9a</b>	SI	9_1	-5.86	50.31
<b>10</b>	SI	1_1	-7.43	3.59
<b>10a</b>	SI	13_1	-6.31	23.7
<b>11</b>	SI	3_1	-6.1	33.95
<b>11a</b>	SI	9_1	-5.71	65.15
<b>12</b>	NO	-	-	-
<b>12a</b>	NO	-	-	-
<b>13</b>	SI	1_1	-7.23	5.01
<b>13a</b>	SI	2_1	-7.17	5.53
<b>14</b>	SI	1_1	-7.4	3.75
<b>14a</b>	SI	1_1	-7.4	3.74
<b>15</b>	NO	-	-	-
<b>15a</b>	SI	8_1	-5.81	55.24
<b>16</b>	SI	8_1	-5.6	78.86
<b>16a</b>	SI	7_1	-5.62	75.64

**Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1**

17	SI	2_1	-6.39	20.59
Ligando	¿Hubo unión a la región PIF?	Rank_subrank	Energía libre (kcal/mol)	Ki (μM)
17a	SI	5_1	-6.19	29.25
18	SI	7_1	-6.31	23.55
18a	SI	11_1	-6.07	35.28
19	NO	-	-	-
19a	NO	-	-	-
20	NO	-	-	-
20a	SI	1_1	-7.13	5.96
21	NO	-	-	-
21a	NO	-	-	-
22	SI	1_1	-7.6	2.69
22a	SI	1_1	-7.8	1.92
23	SI	1_1	-7.6	2.68
23a	SI	1_1	-8.15	<b>1.06</b>
24	SI	9_1	-6.87	9.2
24a	SI	4_1	-6.53	16.42
25	NO	-	-	-
25a	NO	-	-	-
26	SI	3_1	-6.35	22.21
26a	SI	6_1	-6.24	26.71
27	SI	5_1	-5.98	41.13
27a	NO	-	-	-

**Tabla 5.** Resultados del estudio de acoplamiento molecular de los 53 ligandos analizados con el receptor 4A06 se incluye en el inicio los resultados del ligando de referencia.

Ligando	Unión a PIF	Rank_subrank	Energía libre (kcal/mol)	KI (μM)
A06	SI	1_1	-6.74	11.52
1	SI	10_1	-5.20	120.84
2	SI	22_1	-5.01	176.05
2a	SI	37_1	-4.23	656.26
3	SI	16_1	-4.00	956.87
3a	SI	4_1	-6.01	36.54
4	NO	-	-	-
4a	NO	-	-	-
5	NO	-	-	-
5a	SI	23_1	-5.20	52.59
6	NO	-	-	-
6a	NO	-	-	-
7	NO	-	-	-
7a	SI	57_1	-4.30	502.5
8	SI	6_1	-5.20	60.85
8a	SI	3_1	-6.20	23.34
9	SI	21_1	-5.03	92.62

**Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región  
PIF De La PDK1**

<b>9a</b>	SI	10_1	-5.46	60.41
<b>Ligando</b>	<b>Unión a PIF</b>	<b>Rank_subrank</b>	<b>Energía libre (kcal/mol)</b>	<b>Ki (µM)</b>
<b>10</b>	SI	1_1	-7.35	4.27
<b>10a</b>	SI	15_1	-6.01	45.8
<b>11</b>	SI	6_1	-5.63	69.82
<b>11a</b>	SI	10_1	-5.23	76.35
<b>12</b>	NO	-	-	-
<b>12a</b>	NO	-	-	-
<b>13</b>	SI	2_1	-6.96	7.06
<b>13a</b>	SI	3_1	-7.00	6.86
<b>14</b>	SI	3_1	-7.30	3.95
<b>14a</b>	SI	2_1	-7.45	3.86
<b>15</b>	NO	-	-	-
<b>15a</b>	SI	9_1	-5.62	65.24
<b>16</b>	SI	10_1	-4.90	98.59
<b>16a</b>	SI	8_1	-5.23	86.64
<b>17</b>	SI	3_1	-6.01	28.36
<b>17a</b>	SI	6_1	-5.63	41.32
<b>18</b>	SI	8_1	-6.01	29.65
<b>18a</b>	SI	12_1	-5.63	52.58
<b>19</b>	NO	-	-	-
<b>19a</b>	NO	-	-	-
<b>20</b>	NO	-	-	-
<b>20a</b>	SI	3_1	-7.01	6.23
<b>21</b>	NO	-	-	-
<b>21a</b>	NO	-	-	-
<b>22</b>	SI	3_1	-7.34	2.95
<b>22a</b>	SI	2_1	-7.60	2.03
<b>23</b>	SI	1_1	-7.89	1.65
<b>23a</b>	SI	2_1	-7.88	1.66
<b>24</b>	SI	10_1	-5.86	12.2
<b>24a</b>	SI	5_1	-6.01	19.23
<b>25</b>	NO	-	-	-
<b>25a</b>	NO	-	-	-
<b>26</b>	SI	3_1	-6.21	23.45
<b>26a</b>	SI	6_1	-6.03	27.23
<b>27</b>	SI	6_1	-5.53	44.63
<b>27a</b>	NO	-	-	-

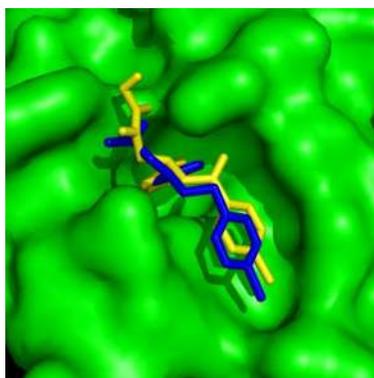
**Tabla 6.** Resultados de los estudios de acoplamiento molecular de los 53 ligandos analizados con el receptor **4AW0** se incluye en el inicio los resultados del ligando de referencia.

Ligando	Unión a PIF	Rank_subrank	Energía libre (kcal/mol)	KI ( $\mu$ M)
<b>MJF</b>	SI	1_1	-8.04	1.29
<b>1</b>	SI	8_1	-5.76	59.99
<b>2</b>	SI	12_1	-5.24	145.06
<b>2a</b>	SI	38_1	-4.35	642.26
<b>3</b>	SI	15_1	-4.13	934.35
<b>3a</b>	SI	2_1	-6.4	20.23
<b>4</b>	NO	-	-	-
<b>4a</b>	NO	-	-	-
<b>5</b>	NO	-	-	-
<b>5a</b>	SI	22_1	-5.97	42.39
<b>6</b>	NO	-	-	-
<b>6a</b>	NO	-	-	-
<b>7</b>	NO	-	-	-
<b>7a</b>	SI	56_1	-4.5	493.5
<b>8</b>	SI	4_1	-5.85	51.75
<b>8a</b>	SI	2_1	-6.42	19.67
<b>9</b>	SI	19_1	-5.54	86.27
<b>9a</b>	SI	9_1	-5.86	50.31
<b>10</b>	SI	1_1	-8.35	0.762
<b>10a</b>	SI	13_1	-6.31	23.7
<b>11</b>	SI	3_1	-6.1	33.95
<b>11a</b>	SI	9_1	-5.71	65.15
<b>12</b>	NO	-	-	-
<b>12a</b>	NO	-	-	-
<b>13</b>	SI	1_1	-7.27	4.7
<b>13a</b>	SI	2_1	-7.17	5.53
<b>14</b>	SI	1_1	-8.03	1.3
<b>14a</b>	SI	1_1	-7.98	1.42
<b>15</b>	NO	-	-	-
<b>15a</b>	SI	8_1	-5.81	55.24
<b>16</b>	SI	8_1	-5.6	78.86
<b>16a</b>	SI	7_1	-5.62	75.64
<b>17</b>	SI	2_1	-6.39	20.59
<b>17a</b>	SI	5_1	-6.19	29.25
<b>18</b>	SI	7_1	-6.31	23.55
<b>18a</b>	SI	11_1	-6.07	35.28
<b>19</b>	NO	-	-	-
<b>19a</b>	NO	-	-	-
<b>20</b>	NO	-	-	-
<b>20a</b>	SI	1_1	-7.23	4.98
<b>21</b>	NO	-	-	-
<b>21a</b>	NO	-	-	-
<b>22</b>	SI	1_1	-7.71	2.22
<b>22a</b>	SI	1_1	-8.07	1.22

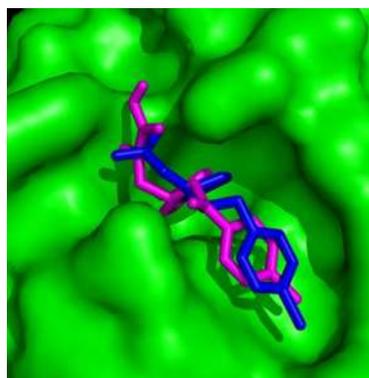
**Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1**

<b>23</b>	SI	1_1	-8.22	0.950
<b>23a</b>	SI	1_1	-8.46	0.628
<b>24</b>	SI	9_1	-6.87	9.2
<b>24a</b>	SI	4_1	-6.53	16.42
<b>25</b>	NO	-	-	-
<b>25a</b>	NO	-	-	-
<b>26</b>	SI	3_1	-6.35	22.21
<b>26<sup>a</sup></b>	SI	6_1	-6.24	26.71
<b>27</b>	SI	5_1	-5.98	41.13
<b>27<sup>a</sup></b>	NO	-	-	-

En la **tabla 7** se presenta la energía libre y el KI de los ligandos que obtuvieron mejores resultados siendo ellos los siguientes: **10, 13, 14, 14a, 20a, 22, 22a, 23, 23a**



**A**



**B**

*Figura 16. Representación de la forma de unión de los ligandos con menor energía libre y menor KI. A: Ligando **23a** = amarillo; B: Ligando **23** = magenta; ligando de referencia **P47** = azul*

**Tabla 7.** Resumen de los mejores resultados de energía libre y KI de los ligandos 10, 13, 14, 14a, 20a, 22, 22a, 23, 23a

Receptor	Ligando	Energía libre (kcal/mol)	KI ( $\mu\text{M}$ )
3HRF	P47 (Ref.)	-8.04	1.29
	10	-7.43	3.59
	13	-7.23	5.01
	14	-7.4	3.75
	14a	-7.4	3.74
	20a	-7.13	5.96
	22	-7.6	2.69
	22a	-7.8	1.92
	23	-7.6	2.68
4A06	23a	-8.15	1.06
	A06 (Ref.) *	-6.74	11.52
	10	-8.25	0.903
4AW0	23	-7.89	1.65
	MJF (Ref.)	-9.36	0.136
	10	-8.35	0.762
	13	-7.27	4.7
	14	-8.03	1.3
	14a	-7.98	1.42
	20a	-7.23	4.98
	22	-7.71	2.22
	22a	-8.07	1.22
	23	-8.22	0.9574
23a	-8.46	0.6287	

\*La energía del ligando de referencia A06 es muy alta en comparación de los ligandos de novo, lo cual pudiera explicarse por la conformación que se encontró en la estructura cristalográfica con código PDB 4A06.

De acuerdo a los estudios de alineación de los receptores se puede observar que tienen una identidad del 99% teniendo como base la estructura cristalográfica **4AW0**, sin embargo, se puede notar que la energía libre tanto de los ligandos de referencia como de los ligandos propuestos, **10, 13, 14, 14a, 20a, 22, 22a, 23, 23a** el receptor 4A06 muestra menos interacciones favorables en la región PIF de la PDK1 lo que nos indica que los ligandos propuestos y que tienen interacción con dicho receptor pudieran también interactuar con otros receptores que sean de la familia de las AGC cinas como la PDK1, así en futuros estudios docking

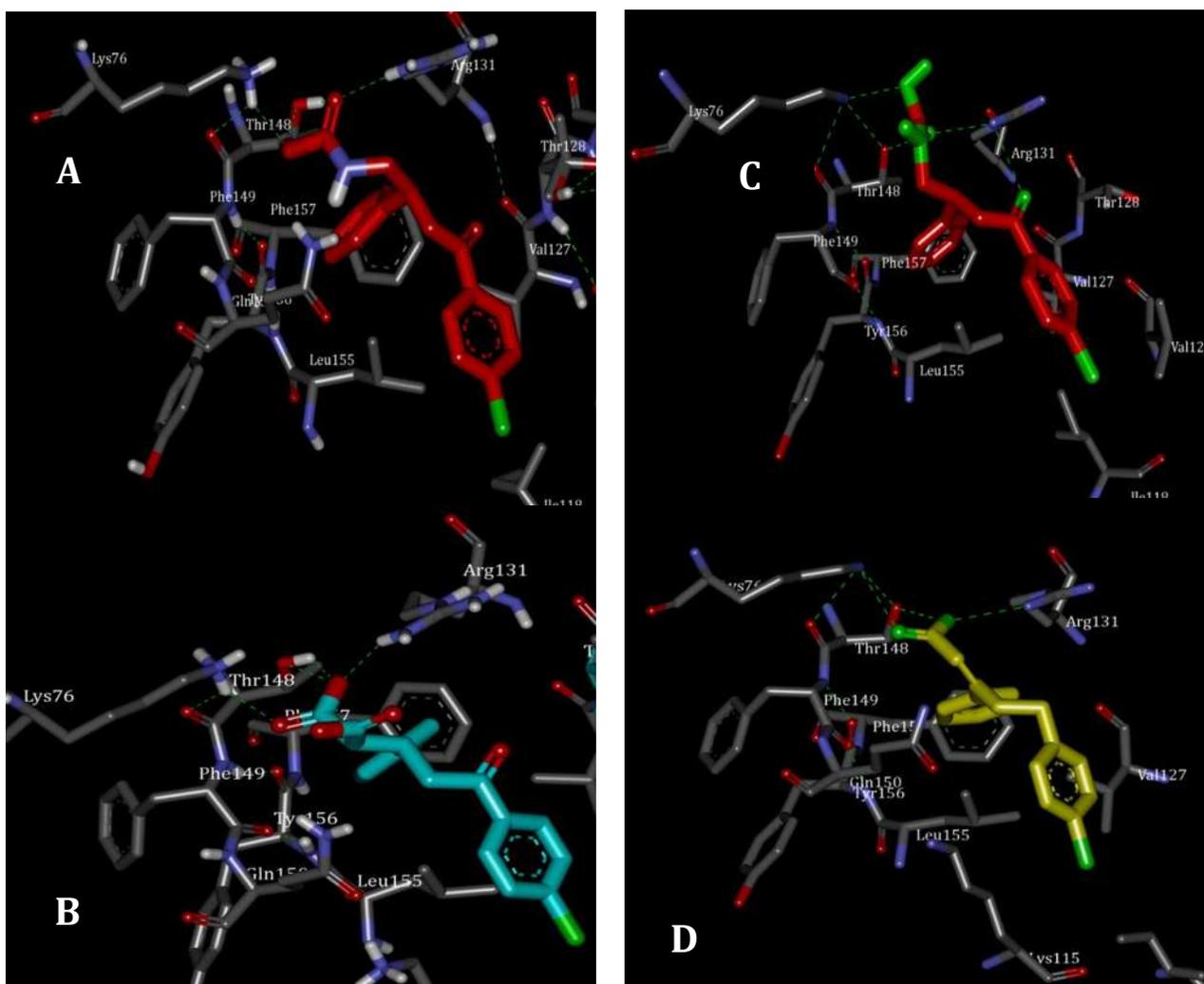
se puede tener en cuenta analizar diferentes ligandos con distintas cinasas que compartan la región PIF.

En lo que respecta a los receptores **3HRF** y **4AW0**, se puede observar claramente cómo disminuye la energía libre de la interacción de los ligandos con el segundo y que es proporcional a la energía que presentan los ligandos de referencia lo que nos puede sugerir que haya algunas diferencias mínimas en la región PIF y tal vez en otras regiones de la estructura cristalográfica, pensando en realizar un subsecuente modelado homólogo de la secuencia de aminoácidos de la PDK1.

Podemos observar que el ligando **23a** fue el que tuvo mejores resultados de energía libre y de KI en los 3 receptores por lo que se analizaron las interacciones y se compararon con los tres ligandos de referencia para poder establecer las diferencias en el modo de unión lo cual nos podría explicar mejor la energía libre que presenta. En el presente trabajo se muestran sólo las interacciones del mejor ligando pudiendo observar que para el caso de la unión con **4AW0** (**Figura 17 A**), **23a** tiene sólo tres interacciones tipo puente de hidrógeno, una con la Thr 148, otra con Lys 76 y una tipo puente de hidrógeno con Arg 131 en contraste con el ligando de referencia **MJF** (Fig. 16 B) que presentó 4 interacciones tipo puente de hidrógeno una con Lys 76, dos con Thr 148 y una con Arg 131, lo cual nos indica una mayor estabilidad de **MJF** presentando así una energía libre menor que **23a**.

Para el caso de las interacciones de **23a** con el receptor **3HRF** (**Figura 17 C**) presentó cuatro interacciones tipo puente de hidrógeno dos con Arg 131, una con Thr 148 y una con Lys 76 a diferencia del ligando **P47** (**Figura 17 D**) que sólo presentó 3 interacciones puentes de hidrógeno una con Arg 131, otra con Thr 148 y otra con Lys 76 lo cual le da mayor estabilidad a **23a** teniendo una energía libre menor de unión que **P47**.

Las interacciones de los ligandos **10**, **13**, **14**, **14a**, **20a**, **22**, **22a**, **23**, no se representan en este trabajo, sin embargo, pueden ser buenos candidatos a seguir su estudio ya que mostraron buenos resultados de energía libre comparados con sus ligandos de referencia, aunque no son menores que los que presentaron estos últimos.

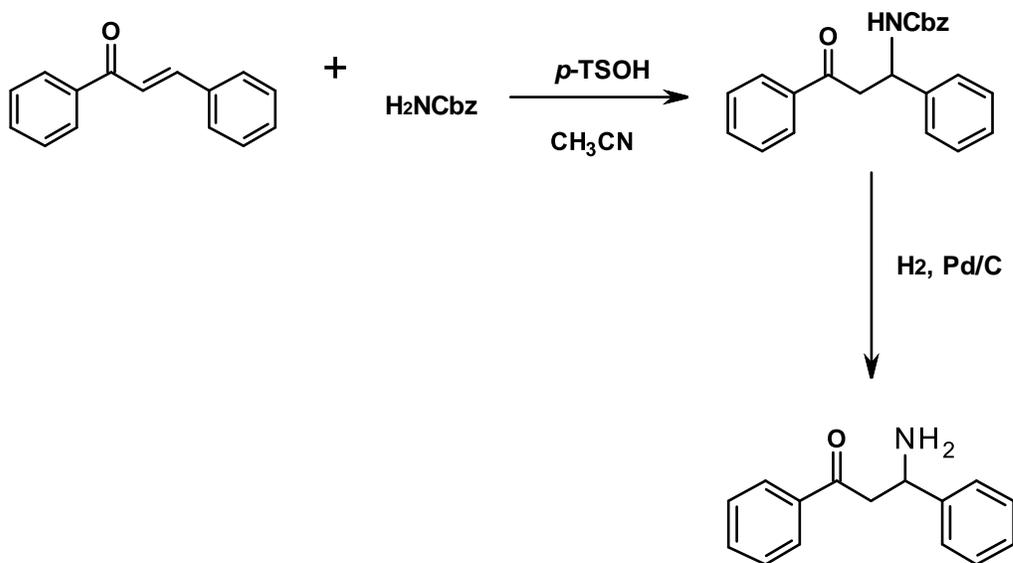


**Figura 17.** Interacciones entre el ligando 23a y los ligandos de referencia. A: 23a en la región PIF del receptor 4AW0. B: Ligando de referencia MJF en la región PIF del receptor 4AW0; C: 23a en la región PIF del receptor 3HRF; D: Ligando de referencia P47 en la región PIF del receptor 3HRF.

Primera propuesta se síntesis

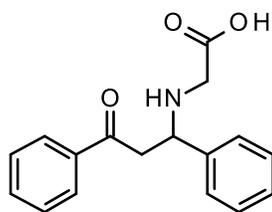
La obtención del compuesto a 3-amino-1,3-difenilpropan-1-ona se puede realizar a partir de la chalcona utilizando bencil carbamato ( $\text{NH}_2\text{Cbz}$ ) y utilizando ácido para toluensulfónico según reportes de Tobias C. Wabnitz<sup>43</sup>. La 3-amino-1,3-difenilpropan-1-ona será utilizada para la posterior adición de sustituyentes electroattractores en el grupo amino.

**Esquema 1** Primera propuesta de síntesis orgánica para generar un precursor que nos lleve a la obtención de los compuestos derivados de la chalcona que mostraron los mejores resultados de energía libre y de Ki en los estudios de acoplamiento molecular.



Segunda propuesta de síntesis orgánica de los mejores compuestos derivados del primer estudio de acoplamiento molecular

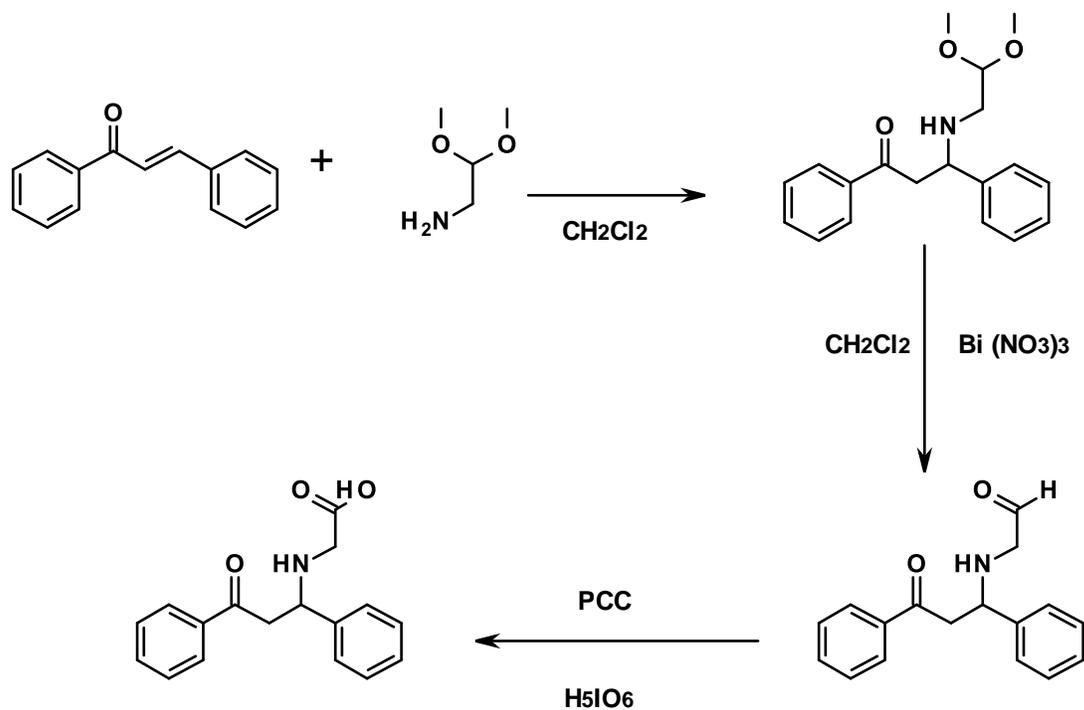
En base a un estudio previo docking ciego empleando el cristal de la PDK1 con el código 4AWO se obtuvo un  $\Delta G = -7.3$  Kcal/mol con el compuesto mostrado en la **figura 22**.



**Figura 16.** Estructura química del ácido 2-((3-oxo-1,3-difenilpropil)amino)acético.

Se pensó en llevar a cabo la síntesis de este derivado de chalcona empleando el aminoácido glicina, sin embargo, se pueden presentar problemas de solubilidad al momento de querer llevar a cabo la reacción, por lo que se plantea una ruta sintética representada en el **esquema 2**.

**Esquema 2.** Propuesta sintética del compuesto 22 y 22a que obtuvieron buenos resultados de energía libre y de  $K_i$  en los estudios de acoplamiento molecular.



### 13. CRIBADO VIRTUAL EN BUSCA DE COMPUESTOS COMO PROBABLES ACTIVADORES DE LA PDK1

Uno de los objetivos importantes dentro del DIFAC a través del cribado virtual es encontrar compuestos que hayan sido ya sintetizados, de los que se pueda conocer propiedades fisicoquímicas y en muchas ocasiones la actividad biológica frente a uno o varios receptores sean éstos últimos proteínas o bien ácidos nucleicos. Esto pudiera ahorrar el tiempo de desarrollo en la búsqueda de compuestos potencialmente activos e incluso pudiera contribuir al reposicionamiento de algunos compuestos que ya fueron aprobados para su uso en humanos pero que han quedado huérfanos. Esto se puede lograr con una búsqueda exhaustiva en grandes bases de datos que contienen almacenada la información fisicoquímica y de actividad biológica antes señalada. El cribado virtual no sustituye al cribado de alta eficiencia “in vitro” pues éste último, toma en cuenta la mayor cantidad de variables posibles que pueden interferir en cuanto a la forma de interactuar de un ligado con un receptor, sin embargo, el cribado virtual puede analizar grandes cantidades de compuestos en un tiempo considerable de gasto computacional siempre y cuando los métodos que se apliquen sean adecuados y que cumplan con las variables químicas mínimas requeridas. Así, el cribado virtual o in silico utiliza enfoques y métodos bien conocidos y que han sido aplicados desde hace más de 20 años, ente ellos encontramos aquellos que son basados en la estructura como los estudios de acoplamiento molecular y el modelado del farmacóforo que puede hacerse basado en la estructura o bien en el ligando. Y recientemente la minería de datos que aplica herramientas de cálculo estadístico, así como de propiedades de agrupamiento de compuestos en base a diferentes características fisicoquímicas y de actividad biológica.

En esta etapa de la investigación se contaba con la información sobre 23 estructuras químicas de compuestos de bajo peso molecular que mostraron actividad frente a la PDK1<sup>44</sup> y que fueron utilizados para realizar un cribado virtual empleando estudios de acoplamiento molecular y modelado del farmacóforo a través de un previo filtrado por PAINs que nos previene de buscar en aquellos compuestos que pudieran presentar interferencias en pruebas in vitro, o bien presentar efectos adversos conocidos<sup>45</sup>. La evaluación de los protocolos de acoplamiento molecular y del modelado del farmacóforo fue utilizando compuestos conocidos como *señuelos* que su nombre mejor conocido en inglés es *decoys* cuyas principales características son que

poseen una alta identidad química estructural y propiedades fisicoquímicas pero que no tuvieron interacción o no presentaron mejor actividad biológica in vitro a diferencia de compuestos que son conocidos como activos para cierto receptor<sup>46</sup>. Así estas evaluaciones nos permiten saber si un método computacional empleado puede distinguir bien entre aquellos compuestos que son activos de los que no lo son, a la vez que se emplean herramientas estadísticas como el teorema de Bayes donde se calcula la sensibilidad y especificidad que son ampliamente conocidos en el área de la química analítica aplicada al diagnóstico y apoyo en identificación de enfermedades, sólo que ahora en el DIFAC nos permite evaluar los protocolos antes mencionados y de donde se puede obtener información valiosa para construir curvas ROC que muestran el desempeño del trabajo que se pretende hacer durante el cribado virtual, evitando así el tiempo de gasto computacional además de la optimización del tiempo y la sensibilidad y especificidad que tenga nuestro método para que pueda hacer una buena aproximación o predicción de los resultados in *silico* que desde luego debe haber un seguimiento para llevar a cabo las pruebas in vitro.

Así, en esta parte de la investigación se presenta el cribado virtual realizado en el Departamento de Química de la Universidad de Cambridge en el Reino Unido bajo la asesoría del Dr. Andreas Bender con el apoyo de su grupo de investigación en el Unilever Centre.

El trabajo consistió que a partir de las 23 estructuras químicas con actividad conocida como activadoras de la PDK1 se sometieron a estudios de acoplamiento molecular, modelado de farmacóforo, y a partir de los mejores resultados obtenidos se desarrolló el cribado virtual obteniendo sólo nueve compuestos con alta probabilidad de tener interacción en la región PIF de la PDK1.

### ***Material y métodos***

#### *Fuentes de bases de datos*

Fueron considerados 23 compuestos con valores de actividad  $AC_{50}$  obtenidos de forma experimental los cuales son presentados en la figura 3. Para poder comenzar con los estudios de acoplamiento molecular y modelado de farmacóforo, se emplearon los complejos cristalográficos con los códigos PDB **3HRF**, **4AWO** y **4A06**, y que se encuentran en complejo con compuestos de bajo peso molecular con un comportamiento como activadores de la PDK1, aunque la estructura cristalográfica 4A06 fue descartada de los estudios iniciales por carecer de las coordenadas cristalográficas en algunos átomos del ligando. Para realizar el cribado

virtual se empleó la base de datos Zinc<sup>47</sup> analizando un total de 22,724,825 compuestos además de que de esta misma base de datos se obtuvieron los compuestos *señuelo* (*decoys*) empleando el programa DecoyFinder<sup>48</sup>. Además, dentro del modelado del farmacóforo también se empleó el servidor en línea ZINCPharmer<sup>49</sup> como método alternativo de modelado del farmacóforo.

Para comenzar con los estudios de acoplamiento molecular se realizó un alineamiento estructural de los códigos PDB: **4AWO** y **3HRF**, empleando el programa MOE con la finalidad de comparar la similitud espacial tridimensional con la finalidad de ubicar a los aminoácidos en la región PIF y así poder realizar un estudio basado en la estructura con una mayor aproximación, el resultado de RMSD fue obtenido y calculado para ambas estructuras. La estructura **3HRF** fue utilizada como base en cuanto a coordenadas cristalográficas para comenzar a realizar los estudios de acoplamiento molecular a los 23 compuestos conocidos puesto que no se cuenta con estructuras cristalográficas en complejo con cada uno de ellos.

#### *Estudios de acoplamiento molecular*

Como se ha comentado en la introducción al DIFAC, son varias las herramientas para poder llevar a cabo estudios de acoplamiento molecular, en esta parte de la investigación estos estudios se realizaron empleando el programa GOLD<sup>50</sup> distribuido por *The Cambridge Crystallographic Data Centre* (CCDC) que es una plataforma con interfaz gráfica de muy fácil manejo y para la preparación de los ligandos y el receptor, se empleó el programa MOE distribuido por Chemical Computing Group<sup>51</sup>.

#### *Preparación de los ligandos*

Los 23 ligandos con su actividad  $AC_{50}$  conocida fueron modelados en dos dimensiones (2D) utilizando el programa ChemBioDraw Ultra versión 12.0, después fueron transformados en estructuras tridimensionales (3D) empleando el programa ChemSketch obteniendo así archivos para cada compuesto en formato .mol, después en el programa MOE se hizo la adición de los átomos de hidrogeno faltantes y se calculó su estado protonado a pH de 7.4 seguido de la minimización de su energía y optimización geométrica utilizando el campo de fuerza MFF94 con un gradiente de RMSD de 0.1 kcal/mol.

### *Preparación del receptor*

La estructura cristalográfica de rayos X con el código PDB **3HRF** fue obtenida del PDB y fue preparada con la ayuda de la interfaz gráfica Hermes en el programa GOLD adicionando los átomos de hidrógeno faltantes junto con las cargas Kollman, las moléculas de agua estructural fueron removidas de todo el receptor, el ligando **P47** fue removido y guardado como ligando de referencia para posteriores validaciones de los estudios de acoplamiento molecular. El sitio de unión fue seleccionado a 10Å alrededor de las coordenadas del ligando **P47** como referencia para llevar a cabo los cálculos de acoplamiento molecular<sup>52</sup>.

Selección y preparación de los *decoys* para la evaluación de los estudios de acoplamiento molecular y de modelado de farmacóforo:

Dentro de la validación y evaluación del desempeño de los protocolos y métodos computacionales en el DIFAC existen varias formas de hacerlo, dentro de los estudios de acoplamiento molecular las más utilizadas son la comparación de la posición espacial obtenida por los cálculos computacionales con la reportada por cristalografía de rayos X entre otras, sin embargo cada vez más existe la tendencia de evaluar el desempeño de estos métodos a través del cálculo de su sensibilidad y especificidad con la ayuda del teorema bayesiano que es ampliamente aplicado al área médica en el diagnóstico de enfermedades y en el seguimiento de las mismas, pero que puede ser utilizado en otras áreas como el DIFAC<sup>54</sup>. La sensibilidad se define como la capacidad de un método de distinguir o elegir a los resultados positivos o reactivos como reactivos o positivos y la especificidad se define como la capacidad de un método de detectar ahora a aquellos resultados negativos o no reactivos como negativos o no reactivos, por tanto en la evaluación de los métodos computacionales debemos tener en cuenta este aspecto tan importante para poder llevar a cabo protocolos de estudios de acoplamiento molecular por ejemplo de una forma adecuada con las características de búsqueda o cálculo de puntaje con el objetivo de obtener resultados cada vez más próximos a la exactitud, o bien puede ser empleado en el modelado farmacofórico para poder encontrar compuestos activos. Por lo tanto, para poder decidir sobre las condiciones de un buen estudio de acoplamiento molecular y las características de un buen modelo de farmacóforo, se necesita tener información experimental sobre una serie de compuestos conocidos como activos o “positivos” frente a la PDK y con una serie de compuestos conocidos como inactivos o negativos, en este sentido, emplear compuestos inactivos para la PDK fue un reto pues sólo se cuenta con

información de los 23 compuestos activos frente a la PDK1, y que fueron tomados como positivos para el cálculo de la sensibilidad y especificidad, para los compuestos inactivos o negativos, se utilizó la herramienta DecoyFinder<sup>48</sup> donde es posible obtener compuestos *decoys* puesto que en la base de datos Directory of Useful Decoys (DUD) ([www.dud.docking.org](http://www.dud.docking.org))<sup>55</sup>. No se cuenta con información para la PDK1, el método que emplea esta herramienta es la similitud molecular calculando el coeficiente de Tanimoto que nos habla de la similitud que guardan dos compuestos entre sí, siendo 1 la máxima identidad por lo que se utilizó un valor de 0.9 como criterio más importante para poder hacer la búsqueda de compuestos que tuvieran casi las mismas características químicas además de que tuvieran un valor de diferencia entre las características químicas responsables de la interacción entre el ligando y el receptor como lo son: grupos donadores de átomos de hidrógeno=0; grupos aceptores de puentes de hidrógeno=0; peso molecular mayor o menor a los compuestos activos= 40 g/mol; y un LogP mayor o menor a una unidad según los compuestos activos, seleccionando como número total 36 compuestos decoys por cada compuesto activo, no fue posible obtener el número total de decoys por lo que sólo se emplearon 595 compuestos como inactivos o negativos frente a la PDK1 que fueron guardados en el formato .sdf que nos permite analizar la información de los compuestos a la vez. Después en el programa MOE los 595 compuestos decoys fueron sometidos bajo un análisis de probables errores, se les retiraron los átomos que pudieran tener de sobra y se les calculó el estado protonado a pH de 7.4 seguido de la minimización de su energía y optimización geométrica utilizando el campo de fuerza MFF94 con un gradiente de RMSD de 0.1 kcal/mol.

#### *Estudios de acoplamiento molecular*

Una vez que el receptor, los 23 compuestos activos y los 595 compuestos *decoys* estuvieron listos fueron incluidos dentro del programa GOLD para los estudios de acoplamiento molecular empleando los criterios de análisis mostrados en la **tabla 8** usando un 33% como capacidad de rendimiento de gasto computacional que es sugerido por de faul en el programa GOLD y guardando solo las tres poses más favorables por cada ligando, es decir, sólo se obtuvieron un total de 1854 resultados para ser analizados posteriormente.

**Tabla 8.** Condiciones y características de los nueve protocolos empleadas en los estudios de acoplamiento molecular en el programa GOLD

Protocolo	Función de puntaje	Base de referencia útil para el receptor	Refinamiento de la función de puntaje
1	CHEMPLP_1	Kinase	--
2	CHEMPLP_2	--	--
3	CHEMPLP_3	--	Chemscore.kinase
4	ChemScore_1	Kinase	--
5	ChemScore_2	--	--
6	ChemScore_3	--	Chemscore.kinase
7	GoldScore_1	Kinase	--
8	GoldScore_2	--	--
9	GoldScore_3	--	Chemscore.kinase

A continuación, se hace una descripción breve de cada una de las funciones de puntaje utilizadas en cada protocolo de acoplamiento molecular<sup>52</sup>.

Potencial lineal por fragmentos (Piecewise Linear Potential o CHEMPLP)

Es una función empírica optimizada para realizar la predicción de la mejor pose del ligando dentro del sitio activo del receptor, se utiliza para modelar el espacio de interacción estérica entre el ligando y el receptor. Estas interacciones están dadas en la **ecuación 3**.

$$fitness_{PLP} = -(W_{PLP} \cdot f_{PLP} + W_{lig-clash} \cdot f_{lig-clash} + W_{lig-tors} \cdot f_{lig-tors} + f_{chem-cov} + W_{prot} \cdot f_{chem-prot} + W_{cons} \cdot f_{cons})$$

$$fitness_{CHEMPLP} = fitness_{PLP} - (f_{chem-hb} + f_{chem-cho} + f_{chem-met})$$

**Ecuación 3.** Cálculo y predicción de la mejor pose e interacciones del ligando con el receptor que utiliza la función de puntaje CHEMPLP

### *GoldScore*

La función de puntaje Gold está compuesta por los siguientes cuatro componentes:

- a) La energía de enlace de los puentes de hidrógeno entre el receptor y el ligando (puente de hidrógeno externo)
- b) Energía externa del potencial de acción tipo Van der Waals (vdw) entre el receptor y el ligando (vdw externo)
- c) Energía interna del potencial de acción tipo Van der Waals del ligando (vdw interna)
- d) Energía de tensión torsional del ligando (torsión interna)

### *ChemScore*

Esta función de puntaje fue derivada de forma empírica a partir de un conjunto de 82 complejos proteína-ligando en las que se midieron las afinidades de unión experimentalmente. Esta función ChemScore fue valorada por métodos de regresión lineal con los datos medidos de afinidad calculada como se define en la **ecuación 4**.

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{hbond} + \Delta G_{metal} + \Delta G_{ligo} + \Delta G_{rot}$$

**Ecuación 4.** Cálculo de la energía total del sistema ligando-receptor empleada en la función de puntaje ChemScore

### *Funciones de puntaje refinadas*

El uso del refinamiento estadístico de los potenciales de acción en una función de puntaje permite la creación de funciones con aptitud dirigidos a ciertos receptores tipo cinasas como la PDK1 que se presenta en este trabajo. La función de puntaje refinada para las cinasas permite el reconocimiento de los grupos CH que pueden formar ciertas interacciones falsas evitando que sean tomadas en cuenta para los cálculos. grupos CH activos son aquellos en los anillos aromáticos próximos a nitrógeno.

### *Algoritmo genético utilizado en GOLD*

GOLD utiliza un algoritmo genético con el fin de obtener la mejor pose y la energía libre total más baja durante los cálculos de acoplamiento molecular. Al inicio una población de posibles soluciones de conformación del ligando es configurada de forma aleatoria. Cada miembro de la

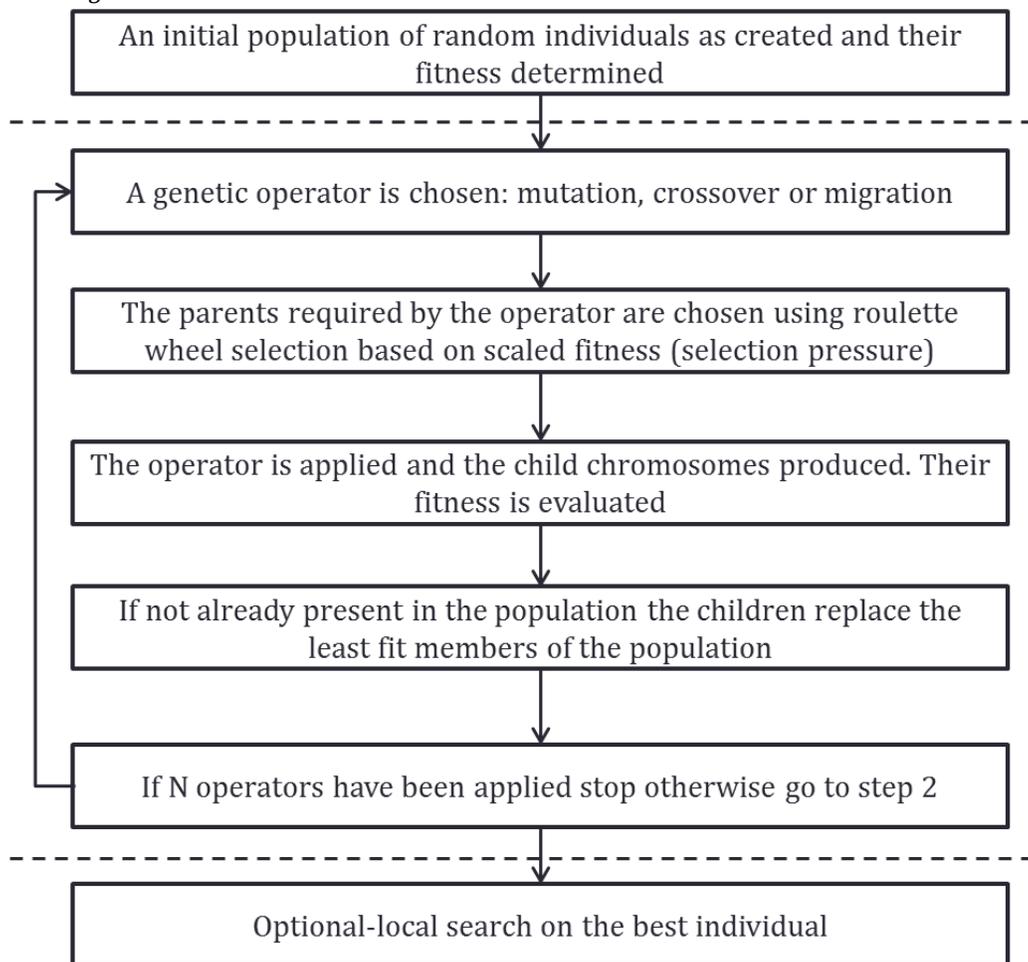
población se codifica como un cromosoma, que contiene la información sobre la asignación de los átomos en el ligando y en el receptor que son capaces de formar enlaces puentes de hidrógeno, también hace un mapeo de las regiones hidrofóbicas tanto en el ligando como en el receptor, y la exploración de los posibles conformeros alrededor de los enlaces de giro libre y de los grupos hidroxilo en los aminoácidos del sitio activo o de interés en el receptor. El proceso de algoritmo genético utilizado en el programa GOLD se muestra en el **esquema 3**.

Después de los cálculos de acoplamiento molecular, los resultados de cada protocolo se guardan en un solo archivo de tipo .sqlite para ser analizado con la herramienta GoldMine dentro del paquete de GOLD para el procesamiento, análisis y representación gráfica de los resultados.

Evaluación del desempeño de los protocolos empleados en los estudios de acoplamiento molecular

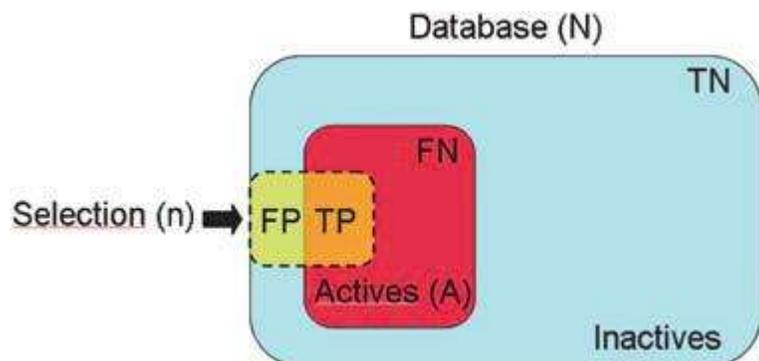
Utilizando la herramienta GoldMine, los resultados obtenidos se dividieron en dos grupos, a uno lo llamamos activos y al segundo decoys o compuestos de fondo, los compuestos activos correspondieron a los 23 compuestos activos frente a la PDK1, habiendo por lo tanto 69 resultados de las mejores poses con la menor energía libre, y de la misma forma para un total de 1785 resultados para los compuestos decoys, para evaluar la sensibilidad y especificidad de los protocolos, fue utilizado el método de curvas ROC y el factor de enriquecimiento que es un parámetro estadístico que nos muestra la capacidad predictiva de un método en una base de

**Esquema 3.** Algoritmo empleado dentro del programa GOLD para llevar a cabo la elección de la mejor pose de cada ligando en el sitio de unión tomando en cuenta los mejores resultados de energía libre



datos con múltiples variables, para dicho efecto, se consideró un valor de 10% de cohorte para realizar los cálculos de sensibilidad y especificidad, es decir, se tomó en cuenta un total de 185 resultados para dichos cálculos. También con la finalidad de revisar de forma gráfica el desempeño de los protocolos, se obtuvieron los gráficos correspondientes a los histogramas utilizando el resultado de energía libre como descriptor con la función de puntaje ChemScore para revisar los valores de energía libre mostrados por los 23 compuestos activos y por los decoys, por otro lado, para las funciones de puntaje CHEMPLP y GoldScore, se obtuvieron los gráficos correspondientes utilizando como descriptor el mejor resultado de pose que involucra a las interacciones sobresalientes en los resultados de acoplamiento molecular. Para conocer la sensibilidad y especificidad de los protocolos utilizados, se empleó el teorema bayesiano

aplicando el **conjunto de ecuaciones 5**, y el modelo de cálculo estadístico mostrado en la **figura 18**.



**Figura 18** Representación general del procedimiento para los cálculos utilizando el teorema bayesiano, ejemplifica la clasificación de una base de datos y las porciones sobre las que se hacen los cálculos,

$$Sp = \frac{N \text{ discarded inactives}}{N \text{ total inactives}} = \frac{TN}{TN + FP}$$

A

$$Se = \frac{N \text{ selected actives}}{N \text{ total actives}} = \frac{TP}{TP + FN}$$

B

$$EF = \frac{TP/n}{A/N}$$

C

$$Ya = \frac{TP}{n}$$

D

**Conjunto de ecuaciones 5.** Conjunto de ecuaciones empleadas durante los cálculos del teorema bayesiano. A: especificidad; B: Sensibilidad; C: Factor de enriquecimiento; D: Rendimiento del método para seleccionar compuestos activos.

Donde N es el número total de resultados analizados que en nuestro estudio corresponde a 1854 poses de ligandos obtenidas de los cálculos de acoplamiento molecular; la selección n corresponde al número de compuestos empleados para el cálculo del factor de enriquecimiento; FP corresponde a los resultados de los compuestos que resultaron falsos positivos; TP corresponde a los resultados de los compuestos que resultaron como verdaderos

positivos; FN corresponde a los resultados de los compuestos falsos negativos; A corresponde a los resultados de los compuestos que se consideró como activos que en este caso son los resultados de los 23 compuestos conocidos como activos para la PDK1; TN corresponde a los compuestos que se obtuvieron como verdaderos negativos o inactivos; EF es el factor de enriquecimiento que proviene de las curvas *Receiver Operating Characteristic* (ROC); Ya es el porcentaje de rendimiento de cada protocolo para obtener compuestos verdaderamente activos o positivos. El factor de enriquecimiento y el porcentaje de rendimiento para seleccionar a los compuestos activos como verdaderamente activos o positivos son los principales parámetros que nos permiten conocer el desempeño real de cada protocolo en el total de los compuestos utilizados en estos cálculos.

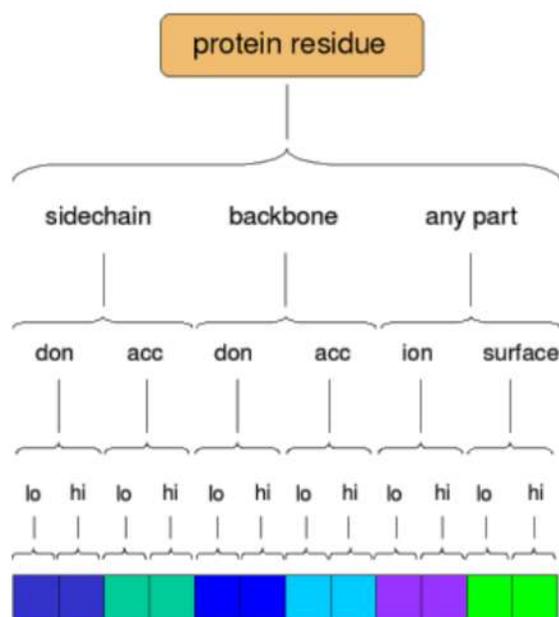
#### *Modelado farmacofórico*

Para efectuar el modelo de farmacóforo que se utilizó en el cribado virtual, principalmente se realizó basado en la estructura aprovechando que se conoce la información experimental estructural de la PDK1 en complejo con ligandos en la región PIF. La elección y construcción de las características del farmacóforo fueron realizados dentro del programa MOE en su sección de modelado farmacofórico que nos permite hacer el modelado de forma manual o bien es calculado por consenso de las características químicas y grupos funcionales implicados en la interacción ligando-proteína y que pudieran ser las responsables de la actividad biológica. Otra forma en que se llevó a cabo el modelado del farmacóforo fue basado en el ligando partiendo de un alineamiento tomando como base las coordenadas cristalográficas de interacción de uno de los ligandos con la PDK1.

El modelo farmacofórico basado en la estructura se realizó usando los códigos PDB: **3HRF** Y **4AW0**, a los cuales se les realizó el cálculo de las interacciones ligando-proteína basados en lo que se le conoce como *huellas digitales (PLIF)* dentro de la minería de datos empleada en la quimioinformática. Por otro lado, para el modelado farmacofórico basado en el ligando se tomó en cuenta a los mejores resultados de poses de los ligandos obtenidos por los estudios de acoplamiento molecular descritos en la primera sesión del material y métodos de esta parte del trabajo, se alinearon los mejores resultados de los 23 compuestos provenientes del acoplamiento molecular tomando como referencia al ligando **P47** para hacer un alineamiento flexible de tipo estocástico seguido por el método de consenso para generar el modelo farmacofórico en el programa MOE.

### *Interacciones ligando-proteína basados en su huella digital*

Este método genera un modelo farmacofórico utilizando las huellas digitales que se obtienen de las interacciones más importantes entre el ligando y el receptor, como enlaces puentes de hidrógeno, interacciones iónicas, y la superficie de contacto entre el ligando y el receptor. Cada aminoácido involucrado dentro del sitio de unión que es de interés en el receptor es clasificado dentro de una categoría de modo de unión según se muestra en la **figura 19**.



**Figura 19.** Clasificación de las interacciones de las huellas digitales en el programa MOE. Hay 6 tipos de interacciones y su respectivo orden jerárquico (bajo y alto) en el que un residuo podría participar como principales interacciones: la cadena lateral y los enlaces puentes de hidrógeno (aceptores o donadores) además de las interacciones iónicas y de superficie sobresalientes.

### *Método de consenso para el modelado farmacofórico empleado en el programa MOE*

La función de consenso utilizada es una medida de la proporción las conformaciones posibles del ligando y que es representado en forma de puntos de anotación comenzando con el cálculo de los puntos antes mencionados incluyendo a un solo ligando o bien al alineamiento calculado en base a las coordenadas cristalográficas de las interacciones del ligando con el sitio de unión de donde se establece un rango de tolerancia (T) que corresponde a la máxima distancia entre dos puntos calculados por el consenso entre cada característica de interacción o modo de interacción de las conformaciones posibles del ligando, después un mínimo puntaje de

consenso (C) es calculado debido a que se requiere para poder volver a rehacer un cálculo de nuevos puntos y características de interacción teniendo así un nuevo cálculo de T y C hasta que se logra un consenso promedio aceptable tomando en cuenta las principales características e interacciones importantes en el modo de unión del ligando por su puesto tomando en cuenta las conformaciones que puede adoptar. En seguida, se presenta una descripción gráfica de las características más importantes que se presentan en forma de farmacóforo llegando así a un mínimo valor de T y C que pueden ser despreciables y que es cuando aparece un *centroide* que es un punto de forma práctica para dar espacio y radio a cada característica farmacofórica. Las características sugeridas por medio del método de consenso muestran a los centroides con un radio de aproximadamente 0.5Å a partir del centro de cada característica

#### *Alineamiento flexible*

Este método emplea una forma de búsqueda de tipo estocástico que simultáneamente busca la conformación espacial de un conjunto de compuestos y las probables formas de alineamiento. El puntaje de los alineamientos está basado en una distribución gaussiana de densidades electrónicas representadas por características de los grupos funcionales contenidos en cada uno de los compuestos encaminadas a modelar las coordenadas cristalográficas del modo de unión en el sitio activo o de interés en el receptor. Las funciones de densidad electrónica calculadas se obtienen empleando la **ecuación 6**.

$$f(x) = s^3 (2\pi)^{-\frac{3}{2}} \exp\left(-\frac{1}{2} |x - x_0|^2 s^2\right)$$

**Ecuación 6.** Cálculo de las funciones de densidad electrónica durante el procedimiento de alineamiento de los ligandos

Para poder decidir la correcta asignación de un alineamiento propuesto por este método, la **ecuación 7** es utilizada para poder comprobar la medida de la similitud (F) entre los compuestos alineados, también generaliza a más de dos compuestos a través del promedio de su similitud topológica por fragmentos.

$$F = \sum_{i=1}^n \sum_{j=1}^{n'} \frac{C_u u_i u'_j + C_v v_i v'_j + C_w w_i w'_j}{nn'} \times \left(\frac{a^2}{2\pi(r_i^2 + r_j'^2)}\right)^{\frac{3}{2}} \exp\left(-\frac{a^2}{2} \frac{|x_i - x'_j|^2}{r_i^2 + r_j'^2}\right)$$

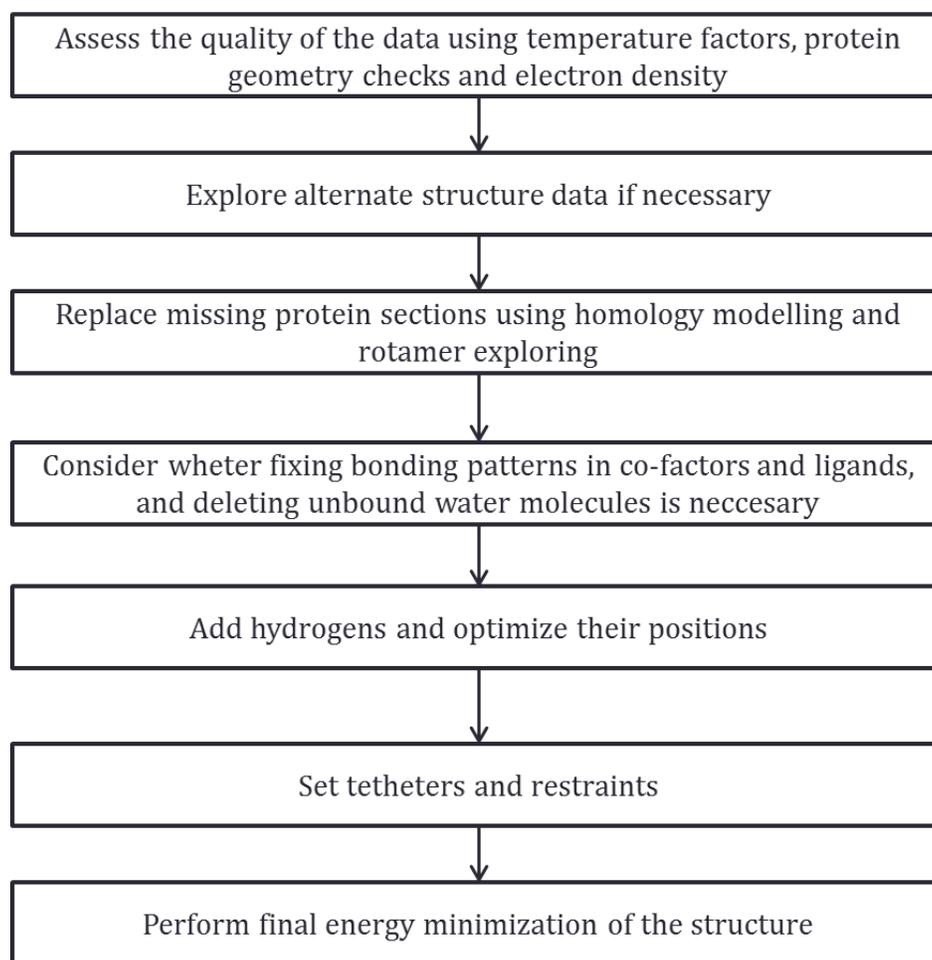
**Ecuación 7.** Ecuación empleada para la generación de los mejores modelos de alineamiento de los ligandos, tomando en cuenta la similitud molecular de cada uno de ellos comparándolos con el ligando base o de plantilla

Donde Cu, Cy y Cw son las características de mayor peso para poder enfatizar una propiedad en particular de los compuestos alineados. De esta forma, cualquier número de características pueden ser incluidas dentro de los cálculos de similitud molecular sin la necesidad de realizar cálculos computacionales adicionales.

Obtención de las plantillas farmacofóricas (*hipótesis farmacofórica*)

Antes de comenzar con la generación de los modelos farmacofóricos las estructuras de rayos x del receptor con el código PDB **3HRF**, el complejo ligando-proteína fue preparado dentro del programa MOE con el objetivo fue corregir las estructuras tanto del ligando como del receptor evitando los errores que pudieran venir de los experimentos de rayos x tal como se describe en el **esquema 4**.

*Esquema 4. Flujograma de trabajo para el modelado farmacofórico basado en la estructura*



Con la ayuda del editor para farmacóforo incluido en el programa MOE, dos diferentes requerimientos farmacofóricos fueron generados, el primero provino del análisis PLIF, y el segundo de la información estructural proveniente del alineamiento tomando como base las coordenadas cristalográficas de los receptores **3HRF** y **4AWO** que fueron superpuestas calculando el RMSD entre las dos estructuras cristalográficas para poder validar la construcción de un modelo farmacofórico adecuado. Después de esta validación, los mejores resultados del acoplamiento molecular provenientes del protocolo 4 realizados en este estudio fueron sometidos a un alineamiento flexible empleando como base estructural las coordenadas del ligando **P47**, dentro del sitio de unión en el receptor y tres diferentes requerimientos farmacofóricos fueron obtenidos.

#### *Evaluación del modelado farmacofórico*

Para poder decidir entre los requerimientos farmacofóricos obtenidos, al igual que con los estudios de acoplamiento molecular, se emplearon compuestos señuelo o decoys y se realizó el cálculo del desempeño a través de la sensibilidad y especificidad de cada farmacóforo. Los 23 compuestos activos frente a la PDK1 fueron seleccionados y guardados en un archivo tipo .sdf para facilitar su manejo dentro del programa MOE y convertidos a forma de base de datos para poder evaluar el modelado farmacofórico puesto que en teoría debería elegir a la mayoría como activos o al menos seleccionarlos puesto que los diferentes farmacóforos provienen de esos compuestos. De la misma forma, los compuestos decoys fueron empleados en un formato .sdf e incluidos a la base de datos junto con los 3 compuestos activos para llevar a cabo la selección de los compuestos como activos e inactivos dentro de un total de 617. Posteriormente, la sensibilidad y especificidad, así como las curvas ROC fueron calculadas empleando el teorema bayesiano.

#### *Cribado virtual*

Para poder identificar nuevos compuestos que tengan una alta probabilidad de presentar actividad frente a la PDK1, el protocolo 4 validado para los estudios de acoplamiento molecular junto con el farmacóforo llamado Query\_no\_plif\_1 fueron utilizados para hacer una búsqueda en la base de datos ZINC de donde se descargaron un total de 22, 724,852 compuestos en dos dimensiones (2D) que fueron guardados en un archivo tipo .sdf para convertirlos en un archivo tipo base de datos para facilitar su uso en el programa MOE, fueron generadas sus coordenadas tridimensionales para ser después sometidos bajo un análisis de probables errores donde se

les retiraron los átomos que pudieran tener de sobra y se les calculó el estado protonado a pH de 7.4 seguido de la minimización de su energía y optimización geométrica utilizando el campo de fuerza MFF94 con un gradiente de RMSD de 0.1 kcal/mol.

Es importante resaltar que antes de hacer un gasto computacional en estudios de acoplamiento molecular y con los farmacóforos generados, el total de los compuestos provenientes de la base de datos ZINC fue sometido a un filtrado previo en cuanto a los andamios o fragmentos de los que se tiene conocimiento que provocan interferencias en ensayos in vitro además de provocar efectos adversos conocidos y por lo tanto se recomienda que deben ser eliminados de los estudios computacionales descritos arriba. Para tal propósito, se utilizó la herramienta en línea PAIN filtre que nos ayuda a seleccionar compuestos que pueden presentar actividades cruzadas con múltiples receptores y por lo tanto generar falsos positivos, Después de este filtrado, solo 7,962 compuestos fueron obtenidos y guardados en un archivo tipo .sdf para su fácil manejo como base de datos dentro del programa MOE, después estos compuestos fueron sometidos a los estudios de acoplamiento molecular y de búsqueda farmacofórica según los procedimientos arriba señalados para cada método. Sin embargo, para garantizar que el modelado del farmacóforo pudiera hacer una adecuada búsqueda, se decidió llevar a cabo una búsqueda conformacional utilizando un método de simulación de dinámica molecular corta, (short molecular dynamics simulation) que emplea cálculos de energía cinética a bajas frecuencias de vibración en donde el tiempo de simulación es 1 ps a temperatura constante seguido de una minimización de energía atómica, los parámetros generales de la búsqueda conformacional fueron: un gradiente RMS de 0.005Å, número máximo de eventos de cálculo: 100, RMS límite máximo: 0.27Å y un rango de energía de 7 kcal/mol entre un conformero y otro.

Uso alternativo de la herramienta ZINCPharmmer:

Debido a que es importante el número de conformaciones posibles que puede presentar un compuesto en el sitio activo o de interés de un receptor, se decidió combinar otro método de generación de farmacóforo empleando el método implementado por David Ryan y cols<sup>49</sup> que es una herramienta abierta a la comunidad científica que lleva a cabo la generación de farmacóforos y refinamiento de los mismos, utiliza una base de datos de aproximadamente 17 millones de conformeros de 18.3 millones de compuestos. Los criterios generales que se utilizaron con este método de búsqueda fueron: número máximo de hits por estado conformacional: 1; número máximo de hits por molécula: 1; número máximo de hits totales:

100; máximo RMSD entre cada confórmero: 0.5Å; rango de peso molecular: 250 – 500 Da; número de enlaces rotables: 5 – 7. Después de obtener 100 compuestos como hits, estos fueron sometidos a un análisis de filtrado por PAINS de donde se obtuvieron sólo 92 compuestos como hits y ocho fueron descartados del estudio. Posteriormente dentro del programa MOE se realizó la predicción de su estado protonado a pH de 7.4 seguido de la minimización de su energía y optimización geométrica utilizando el campo de fuerza MFF94 con un gradiente de RMSD de 0.1 kcal/mol.

#### *Análisis de agrupamiento por el método k-means*

Es difícil elegir el compuesto líder de las predicciones realizadas por medio de los estudios de cribado virtual aquí presentados, puede esperarse que se obtengan sólo algunos compuestos potencialmente con las características de activación de la PDK1 que esperamos obtener, sin embargo es bien conocido que se obtienen docenas o incluso cientos compuestos clasificados como hits que comparten características de estructura química muy similar pero que no quiere decir que todos tengan la misma probabilidad de presentar actividad, por lo que se empleó una herramienta conocida como agrupamiento que es utilizada en la quimioinformática en cuanto a la clasificación de una base de datos de compuestos que comparten características estructurales muy similares. Para nuestro caso se sometieron los compuestos encontrados como hits por el cribado virtual realizado además de los 23 compuestos ya conocidos como activadores de la PDK1, empleamos el cálculo de descriptores topológicos, dichos descriptores son representados en un lenguaje binario para que sea leído en una computadora de forma inequívoca y única, por lo que se les conoce también con el nombre de *huellas digitales de conectividad extendida* o bien ECFP's (Extended Connectivity Fingerprints). Los ECFP's son una representación binaria como se comentó arriba que describe la topología extendida de un compuesto y que se extienden de forma radial comenzando con un átomo base dentro del núcleo de la estructura<sup>56</sup>. Para calcular los ECFP's, se utilizan varios algoritmos, en este caso el más utilizado y adecuado en cuanto a la similitud molecular es el algoritmo de Morgan<sup>57</sup> donde el proceso iterativo asigna identificadores numéricos a cada átomo desde el primero, utilizando una regla de codificación invariable, cada iteración utiliza la información de la precedente hasta que la información de identificación es única en toda la estructura e irrepetible. Como los compuestos son representados con un código binario, cada uno contiene un código único. El cálculo de los ECFP's fueron realizados en el programa Canvas desarrollado por Schrödinger Inc. Una vez que se calcularon ECFP's, fueron guardados en un archivo tipo .csv con el fin de

calcular el número de grupos de compuestos utilizando el programa R<sup>58</sup> utilizando el método K-means<sup>59</sup>. Posteriormente, se utilizaron 60 grupos en el programa Canvas para llevar a cabo el análisis de la agregación con el método K-means, utilizando 60 centros o núcleos de grupos de compuestos. Se realizó la construcción de un “gráfico de calor” (heat map) con el fin de conocer el número de grupos. Para conocer este dato tan importante, se utilizó el método de la mínima diferencia calculada de la suma del error cuadrado entre un núcleo de compuestos con estructura molecular muy similar y otro. Este fue uno de los retos más grandes dentro del cribado virtual puesto que a simple vista es muy difícil de observar y elegir a los compuestos estructuralmente muy similares, tal vez poseen una probable actividad frente a la PDK1, pero no sabemos ni conocemos su magnitud aproximada, por lo que dentro del mismo programa Canvas, empleamos una interfaz gráfica para elegir aquellos grupos de compuestos que estuvieran dentro de los compuestos activos para la PDK1 tuvo que la proximidad entre los 23 compuestos activos y los nuevos.

#### *Estudios cuantitativos de relaciones estructura actividad (QSAR)*

##### *QSAR en tres dimensiones (3DQSAR)*

Los estudios QSAR en tres dimensiones son útiles cuando se tiene información de actividad biológica y de la conformación de al menos un ligando en el sitio de interés en el receptor, puesto que la actividad biológica depende en gran medida de esta conformación y las interacciones ligando-receptor sobresalientes<sup>60</sup>. En nuestro caso, poseemos información de los dos tipos, por un lado, los 23 compuestos analizados como activadores frente a la PDK1 poseen una actividad AC<sub>50</sub> conocida, y por el otro, disponemos información cristalográfica de al menos 3 complejos ligando-receptor. Es bien conocido que las tres estructuras no poseen las mismas interacciones químicas en la región PIF, sin embargo, poseen ciertas características comunes que podemos observar en un análisis de farmacóforo. Dentro de los estudios de QSAR 3D, es importante hacer un alineamiento de las estructuras químicas de los ligandos para obtener un buen modelo de partida para poder hacer o bien el Análisis Comparativo de Campo Molecular (CoMFA) o el Análisis de los Índices de Similitud de Campo Molecular (CoMSIA) que nos hablan sobre las energías estéricas (van der Waals 6 – 12) y la constante dieléctrica entre los compuestos de interés y una *sonda atómica* generando varias intercesiones de un espacio tridimensional bien definido lo suficientemente largo para cubrir el área de densidad molecular de los ligandos con 2Å de separación entre cada cálculo<sup>61</sup>.

La forma de llevar a cabo el alineamiento de los ligandos que se emplean en los estudios 3DQSAR son diferentes y por consecuencia los resultados que se obtienen nos ofrecen diferente grado de exactitud.

Para el 3DQSAR en nuestro trabajo, se emplearon dos métodos de alineamiento diferentes, en el primero provino de las mejores poses derivadas de nuestros estudios de acoplamiento molecular utilizando el método de alineamiento flexible dentro del programa MOE, el segundo alineamiento provino del mejor resultado del farmacóforo obtenido en este trabajo y el tercer alineamiento fue con el método que utiliza la herramienta 3DALIGN<sup>62</sup>. La actividad AC<sub>50</sub> de los 23 compuestos fue convertida en pAC<sub>50</sub> utilizando la **ecuación 8**.

$$pAC_{50} = -\log_{10} (AC_{50})$$

**Ecuación 8.** Cálculo y conversión de la actividad AC<sub>50</sub> en pAC<sub>50</sub>

Las 23 actividades obtenidas se distribuyeron de la siguiente forma: 19 fueron para el conjunto de entrenamiento y los cuatro restantes para el conjunto de prueba.

para los cálculos de CoMFA, fueron creadas dos mallas tridimensionales, a 2.0 y 1.5 Å respectivamente. Como sonda se utilizó un carbono sp<sup>3</sup> y el límite de energía entre cada cálculo fue de 30 kcal/mol, el análisis estadístico se llevó a cabo a través del análisis de mínimos cuadrados parciales (PLS) usando como descriptores principales de las variables independientes al campo electrostático, el estérico, y la presencia de grupos funcionales donadores y aceptores de puentes de hidrógeno y los valores de pAC<sub>50</sub> como una variable dependiente, se permitió calcular cinco como número máximo de componentes principales (PC). La validación de los resultados se realizó mediante un análisis de validación cruzada conocido como *dejando uno afuera* (Leave One Out: LOO) con el fin de conocer el número óptimo de componentes principales. Se predijo la pAC<sub>50</sub> para el conjunto de compuestos de prueba y se obtuvieron los valores de R<sup>2</sup> y Q<sup>2</sup>.

#### *QSAR en dos dimensiones (2DQSAR)*

Los estudios 2DQSAR se realizaron dentro del programa MOE empleando la actividad pAC<sub>50</sub> dividiendo el conjunto de compuestos como se describió en la sección del estudio de 3DQSAR. Se llevó a cabo el cálculo de 151 descriptores en dos dimensiones para el conjunto de compuestos de entrenamiento. A continuación, se realizó el análisis estadístico mediante el

método PLS, como variable independiente se utilizó a los descriptores calculados y como variable dependiente a los valores de  $pAC_{50}$ . Se calcularon los valores de  $R^2$  y de RMSE. A continuación, los mismos resultados se sometieron a un proceso de selección de los mejores descriptores utilizando el análisis de contingencia Quasar implementado dentro del programa MOE para evitar el sobre entrenamiento de los cálculos utilizando el coeficiente de contingencia, V de Carmer, el coeficiente incierto T y  $R^2$  de los cuales se obtuvieron 128 descriptores principales para nuestros estudios. La validación del modelo final se llevó a cabo con un análisis de validación cruzada utilizando el método *dejar uno fuera*, de donde se obtuvieron valores de puntuación Z y los valores predichos para  $pAC_{50}$ . Después, la predicción de los valores de  $pAC_{50}$  fue evaluada con el modelo final y los valores predictivos de  $pAC_{50}$  fueron favorables.

## **Resultados**

### *Estudios de acoplamiento molecular*

#### *Evaluación de los protocolos empelados en los estudios de acoplamiento molecular*

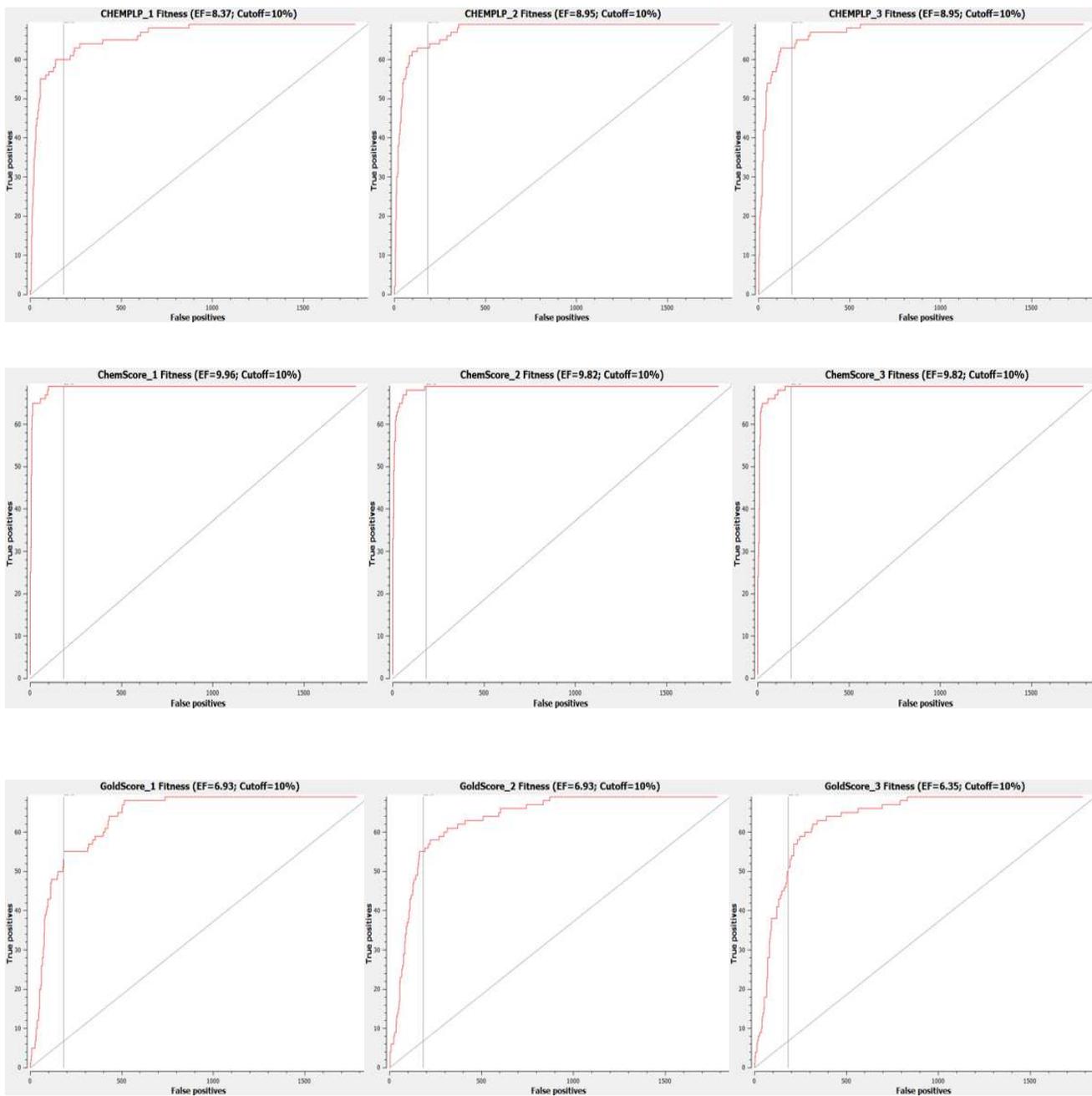
La evaluación de los diferentes protocolos de los estudios de acoplamiento molecular se realizó con la finalidad de encontrar el mejor de ellos siendo el protocolo número cuatro del que se obtuvieron los mejores resultados (**Tabla 9**), ya que los valores de sensibilidad y especificidad obtenidos por ese protocolo son mejores que los otros. Entendiendo que la especificidad (SP) se define como la capacidad para encontrar los verdaderos negativos o en el caso de este estudio los verdaderos inactivos, resultó que otros protocolos obtuvieron un valor mayor de SP que el protocolo número cuatro, sin embargo, el área bajo la curva (AUC) de esos protocolos refleja que la sensibilidad (SE) no es suficiente para encontrar los verdaderos positivos o los activos, por lo tanto el protocolo número cuatro fue la mejor opción para ser utilizado durante el proceso de cribado virtual para encontrar nuevos compuestos activos para PDK1 utilizando la función de puntuación ChemScore con un refinamiento de los resultados utilizando como plantilla de cálculo la función de cinasa implementada dentro del programa GOLD.

**Tabla 9.** Resultados de los estudios de acoplamiento molecular comparando el desempeño de nueve diferentes protocolos utilizando el teorema bayesiano para calcular sensibilidad y especificidad.

PROTOCOL	AUC	EF	N	n	A/N	TP	TN	A	FP	FN	Se	Sp	Ya
1	0.668	8.38	1854	185.4	0.037217	57.822	1796.178	69	57.822	11.178	0.838	0.968812	0.311877
2	0.715	8.95	1854	185.4	0.037217	61.755	1792.245	69	61.755	7.244997	0.895	0.966691	0.333091
3	0.689	8.95	1854	185.4	0.037217	61.755	1792.245	69	61.755	7.244997	0.895	0.966691	0.333091
4	0.925	9.96	1854	185.4	0.037217	68.724	1785.276	69	68.724	0.275997	0.996	0.962932	0.37068
5	0.921	9.82	1854	185.4	0.037217	67.758	1786.242	69	67.758	1.241997	0.982	0.963453	0.365469
6	0.903	9.82	1854	185.4	0.037217	67.758	1786.242	69	67.758	1.241997	0.982	0.963453	0.365469
7	0.391	6.93	1854	185.4	0.037217	47.817	1806.183	69	47.817	21.183	0.693	0.974209	0.257913
8	0.36	6.93	1854	185.4	0.037217	47.817	1806.183	69	47.817	21.183	0.693	0.974209	0.257913
9	0.351	6.35	1854	185.4	0.037217	43.815	1810.185	69	43.815	25.185	0.635	0.976367	0.236327

Las curvas ROC para este trabajo se obtuvieron con la herramienta GoldMine dentro del programa GOLD con un punto de corte de 10% del total de la población de los resultados obtenidos, los resultados gráficos se presentan en la figura 12. La función ChemScore muestra el mejor comportamiento con un valor de AUC de 0,903-0,925 que es similar a un comportamiento ideal dentro del análisis ROC (**Figura 20**). La función ChemPLP obtuvo una puntuación en segundo lugar entre los mejores resultados de sensibilidad, especialmente en el protocolo número tres que podría explicarse debido a que se utilizó un refinamiento de los cálculos de energía libre utilizando una plantilla para enzimas cinasas como función de puntuación específica.

**Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1**



*Figura 20. Gráficos de las curvas ROC que muestran el desempeño de los diferentes protocolos utilizados en los estudios de acoplamiento molecular*

PIF De La PDK1

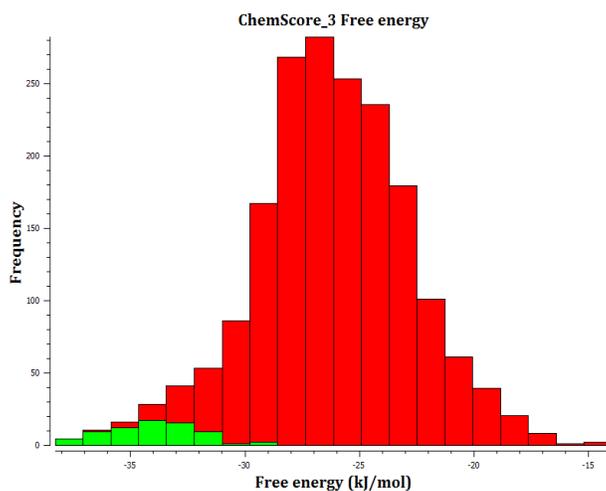
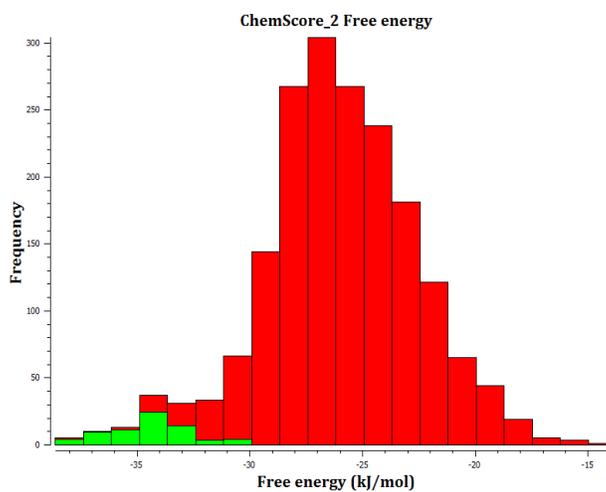
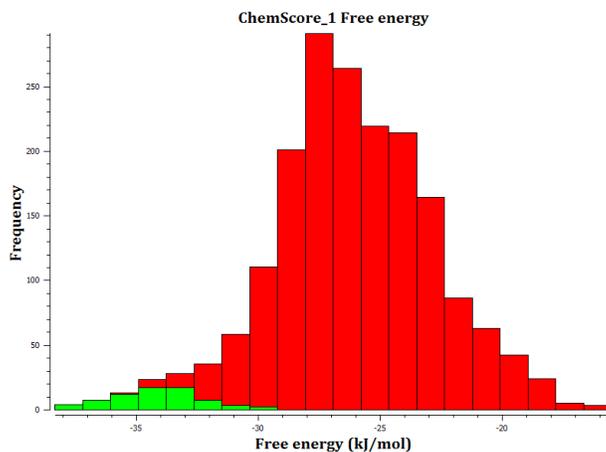


Figura 21. Gráficos correspondientes a la energía libre y la mejor pose de todos los ligandos empleados en los estudios de acoplamiento molecular. En verde se representan los 23 ligandos activos y en rojo todos aquellos que se les considera como fondo, o decoys

Los histogramas de la **figura 21** muestran la distribución de los resultados de energía libre en cal/mol y se pueden diferenciar a los activos (en verde) de los inactivos, (decoys o también llamado fondo en estos estudios mostrados en rojo). Es evidente la diferencia en la energía libre entre los activos y los decoys. El histograma A que se corresponde con el protocolo cuatro de los estudios de acoplamiento molecular, puede demostrar que es capaz de distinguir a los de los activos ya que la energía libre es inferior a un segundo plano, por lo tanto, por estas razones el protocolo 4 es el mejor procedimiento para llevar el cribado virtual usando acoplamiento y la función de puntuación ChemScore con una plantilla de quinasa.

#### *Modelado farmacofórico*

La generación de los modelos farmacóforos se hizo utilizando el método PLIF y método de consenso. Para ser capaz de generar los diferentes modelos farmacofóricos es necesario tener en cuenta todas las conformaciones probables que pueden adoptar los 23 ligandos conocidos como activadores de la PDK1, puesto que se conoce bien que una conformación de mínima energía de un ligando no siempre es su conformación activa. Algunas características importantes a considerar son las interacciones del ligando con el receptor, especialmente entre los grupos funcionales de los residuos involucrados en el sitio de interés y del ligando. Los 23 compuestos activos utilizados en este estudio, cuentan con grupos funcionales similares, por lo que este no es un gran problema, sin embargo, es importante llevar a cabo la generación de diferentes modelos farmacofóricos y seleccionar sólo al mejor, que en este caso vino del método PLIF.

Se formaron varios modelos farmacofóricos, pero aquí se presentan sólo los más importantes con su respectiva evaluación a través del cálculo de la sensibilidad y especificidad utilizando el teorema bayesiano.

Los resultados de sensibilidad y especificidad (**Tabla 10**), nos revelaron que el modelo tres fue el apropiado para llevar a cabo el cribado virtual ya que tiene un valor de especificidad de 0.8821 siendo este valor que para los modelos cuatro y cinco. Sin embargo, los modelos 1 y 2 tienen los valores más altos de especificidad, pero no de sensibilidad. Aunque éstos últimos modelos proceden del método PLIF, contienen características farmacofóricas un volumen de exclusión y que podría ser la razón de que no son capaces de encontrar a los compuestos activos debido tal vez a su impedimento estérico que se contrapone a las diferentes conformaciones de los ligandos durante el proceso de búsqueda. Por otro lado, como se comentó arriba, estos

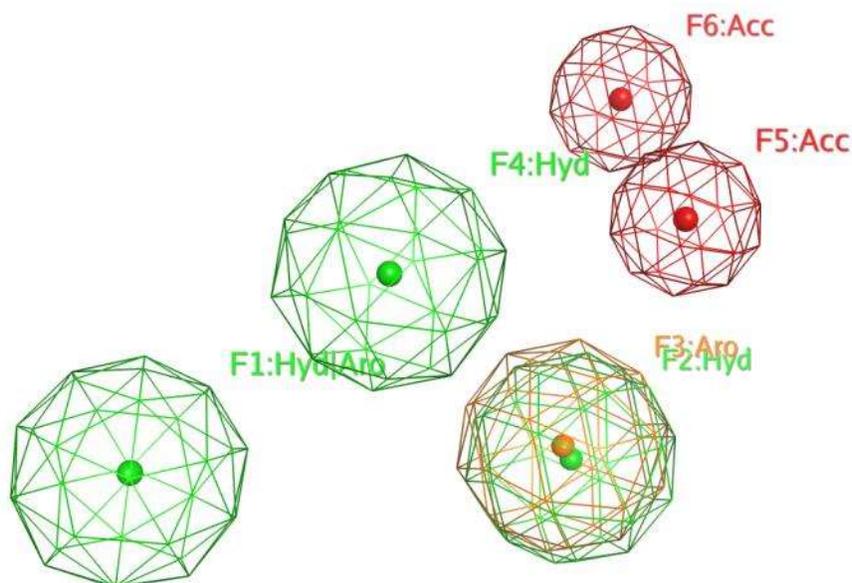
modelos son buenos para seleccionar a los inactivos, pero si no es capaz de seleccionar a los ligandos activos. Por otro lado, los modelos provenientes que no incluyen el método PLIF mostraron un mejor rendimiento durante el proceso de búsqueda con valores de especificidad y selectividad aceptables, las diferencias sustanciales entre estos modelos están en las características y el tamaño de cada uno, por ejemplo, la característica de grupo funcional aceptor de puentes de hidrógeno (Acc) es más corta en el modelo tres (**Figura 23**) que la de los modelos cuatro y cinco, el modelo tres tiene una característica aromática (Aro) de la cual carece el primer modelo, y que podría ser una probable explicación del por qué no es capaz de seleccionar a los compuestos activos.

**Tabla 10.** Modelos farmacofóricos con sus respectivas características y sus medidas, además de los resultados de sensibilidad y especificidad de cada una de ellas.

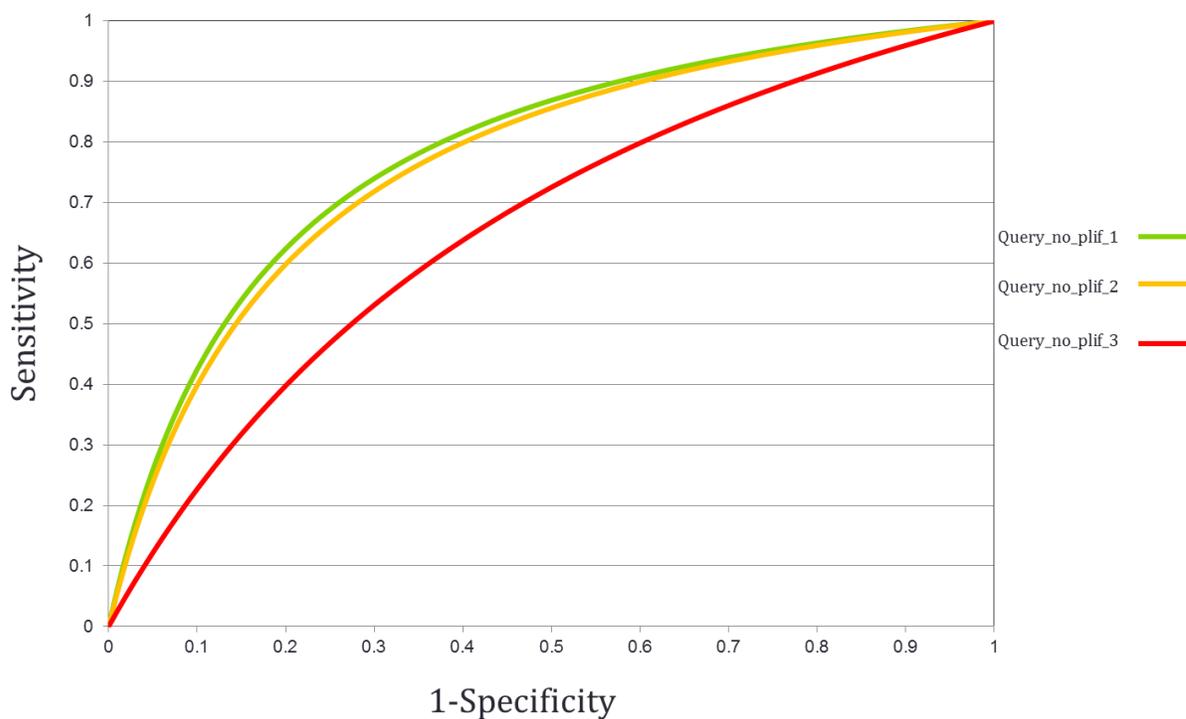
Query entry	Query name	Feature	size	N (total database)	Hits	Actives	Decoys	TP	TN	FN	FP	Se	Sp	n	EF	Ya
1	query_pdb_2	Ani&Acc&ML	1	617	19	23	594	0	575	23	19	0	0.9680	61.7	--	--
		HydA	1.5													
		Ani&Acc&ML	1.5													
		Hyd/Aro	1.5													
		Hyd	1.5													
13 Vol Excl	1.4 - 1.5															
2	query_form_docking_2	Ani&Acc&ML	1.5	617	15	23	594	0	579	23	15	0	0.9747	61.7	--	--
		Aro	1.5													
		Aro	1.5													
		Hyd	1.5													
14 Vol Excl	1.4 - 1.5															
3	query_no_plif_1	Hyd/Aro	1.5	617	88	23	594	18	524	5	70	0.7826	0.8821	61.7	7.8260	0.2917
		Hyd	1.5													
		Aro	1.5													
		Hyd	1.5													
		Acc	1													
Acc	1															
4	query_no_plif_2	Hyd/Aro	1.5	617	96	23	594	18	516	5	78	0.7826	0.8686	61.7	7.8260	0.2917
		Hyd	1.5													
		Aro	1.5													
		Acc	1.2													
		Acc	1.2													
		Hyd	1.5													
5	query_no_plif_3	Hyd/Aro	1.5	617	194	23	594	18	418	5	176	0.7826	0.7037	61.7	7.8260	0.2917
		Hyd	1.5													
		Acc	1													
		Acc	1													
		Hyd	1.5													

Entre los tres modelos que carecen del método PLIF, sus resultados son sobresalientes ya que tienen diferente capacidad para seleccionar los compuestos inactivos porque mientras que el número de consultas 3 encontraron 70 compuestos como "falsos positivos" o "falsas" las consultas activas 4 y 5 y encontraron 78 176 respectivamente. El análisis de las curvas ROC se representa en la **figura 23**.

Los resultados obtenidos en la evaluación del modelado farmacofórico nos sugieren que el mejor modelo es el número 3, debido a que posee una especificidad de 0.8821 que es mayor que el valor de los modelos cuatro y cinco.



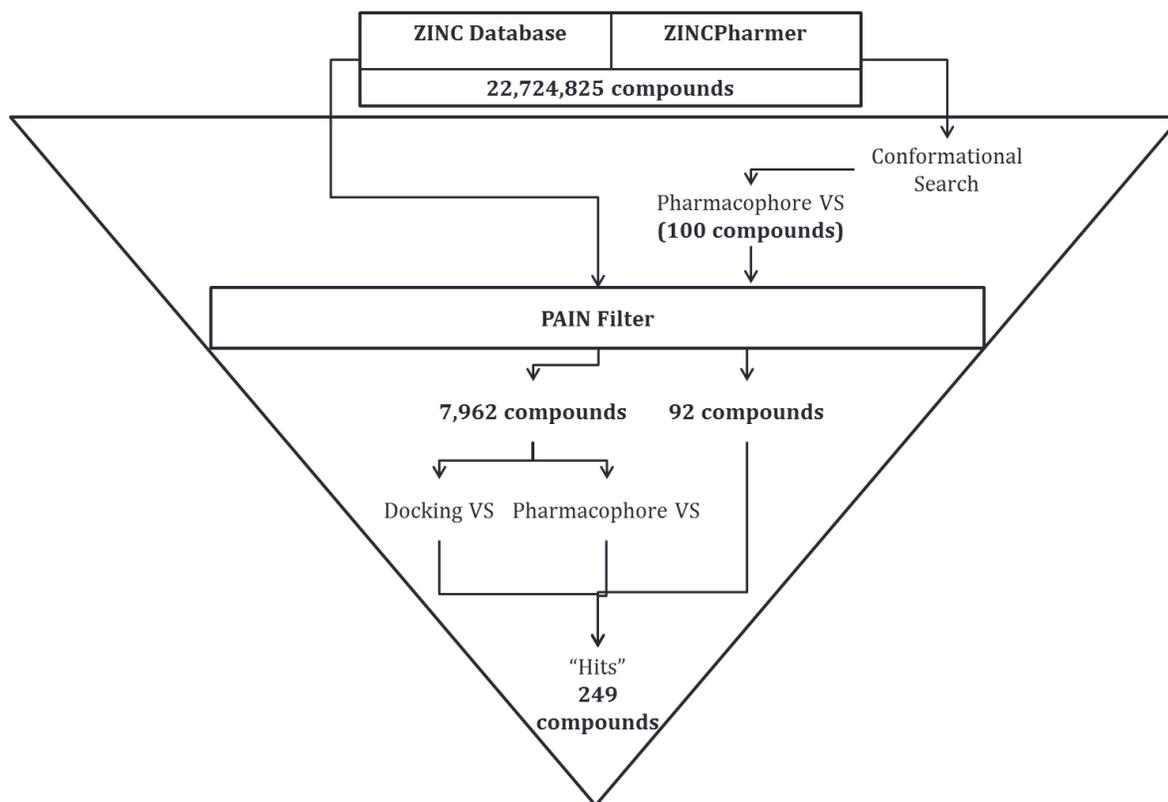
**Figura 22.** Representación del modelo farmacofórico número 3 derivado del alineamiento tomando como base el ligando contenido en el código PDB: 3HRF. La letra F corresponde al número de la característica farmacofórica; **Acc**: grupo acepto de puentes de hidrógeno; **Hyd**: región hidrofóbica; **Aro**: región aromática



**Figura 23.** Gráfico de la curva ROC obtenida a partir de los tres mejores modelados farmacofóricos que ejemplifica el desempeño de cada uno de ellos.

## Cribado virtual

Un total de 22,724,852 compuestos fueron descargados de la base de datos libre ZINC, para poder optimizar desde un inicio el cribado virtual, se llevó a cabo un filtrado por PAIN's de donde se obtuvieron sólo 7,962 compuestos que posteriormente se sometieron a los estudios de acoplamiento molecular utilizando el protocolo número cuatro y para el modelo número tres utilizando el método del farmacóforo. La **figura 24** muestra de forma resumida el número de compuestos obtenidos en cada paso de lo virtual cribado.



**Figura 24.** Diagrama de flujo del procedimiento general utilizado para el cribado virtual de este proyecto. Se presenta el número de compuestos obtenidos en cada fase, así como el orden de aparición de cada método del DIFAC.

Dentro de la búsqueda farmacofórica, debemos de tener en cuenta que se realizó una búsqueda conformacional a los 7,962 compuestos obteniendo una base de datos con 272,913 conformeros con el fin de encontrar todas las posibilidades en las que pudiera encajar alguno de ellos con la conformación que presenta la actividad de alguno de los 23 ligandos conocidos como activos.

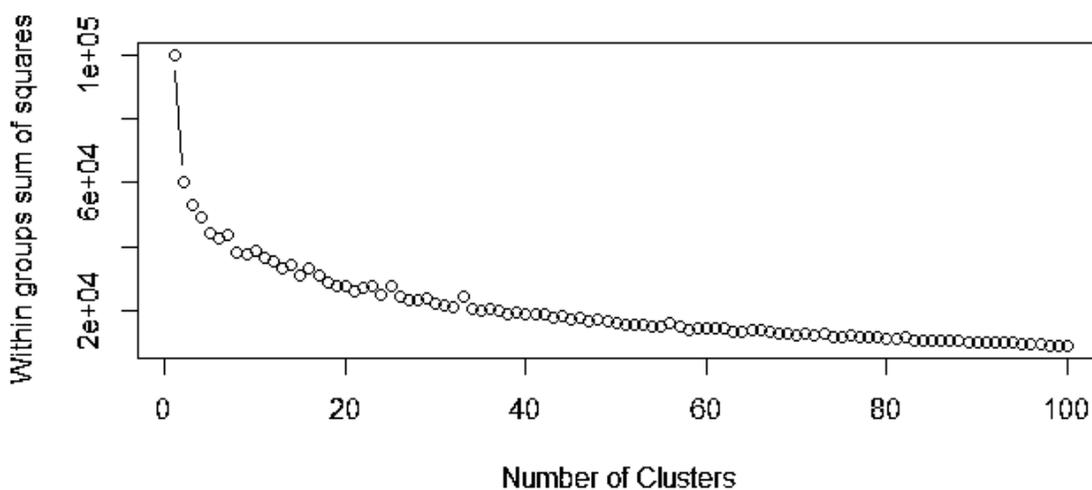
Los estudios de acoplamiento molecular fueron llevados a cabo conforme a las características del protocolo cuatro y los resultados fueron analizados comparando la energía libre obtenida para cada uno de ellos. Los mejores resultados fueron guardados junto con los que provienen del modelado farmacofórico para después ser analizados y filtrados para no tener duplicados.

Una propuesta de búsqueda farmacofórica diferente fue realizada para probar que el modelo farmacofórico número tres puede seleccionar desde un principio los compuestos obtenidos por la búsqueda farmacofórica con la base de datos preparada en este trabajo. ZINCPharmer es una base de datos libre en línea para llevar a cabo la búsqueda de farmacóforos utilizando diferentes criterios de filtrado como el peso molecular, la cantidad de grupos funcionales donadores de puentes de hidrógeno, grupos aceptores de puentes de hidrógeno, o el número máximo de *hits* por cada conformación y por cada compuesto analizado, siendo una buena opción para obtener una menor cantidad de falsos positivos. El modelo farmacofórico número tres fue utilizado como plantilla de búsqueda con los siguientes parámetros: número máximo de *hits* por búsqueda:100, número máximo de conformeros por compuesto:10; máximo valor de energía libre entre cada conformero: 1 kcal/mol. Después de hacer la búsqueda se obtuvieron un total de 100 compuestos como *hits*, sin embargo, estos compuestos todavía necesitaron llevar a cabo un filtrado por PAIN's, por lo que sólo 92 compuestos se obtuvieron como nuevos *hits* que se reunieron con los compuestos anteriormente obtenidos por el cribado virtual por acoplamiento molecular y modelado farmacofórico realizado en el programa MOE. Una vez que fueron llevados a cabo los estudios de acoplamiento molecular y de farmacóforo, se obtuvieron los mejores compuestos llamados como *hits* y que fueron agrupados y sometidos a un análisis de similitud molecular empleando el programa Canvas con un coeficiente de Tanimoto de 0.9, eliminando a los compuestos repetidos. Posterior a esto los mejores compuestos en común entre los resultados de los dos métodos fueron seleccionados en base a sus resultados de energía libre menor de -35 cal / mol, siendo 249 compuestos los que se encontraron con este criterio.

#### *Análisis de agrupamiento*

Para poder hacer una buena elección entre todos los compuestos hasta aquí clasificados como *hits* fue necesario hacer un análisis de agrupamiento de los mismos empleando herramientas de similitud molecular a través del método K-means calculando como descriptores las *huellas digitales de conectividad extendida* y clasificarlos por grupos según su similitud molecular.

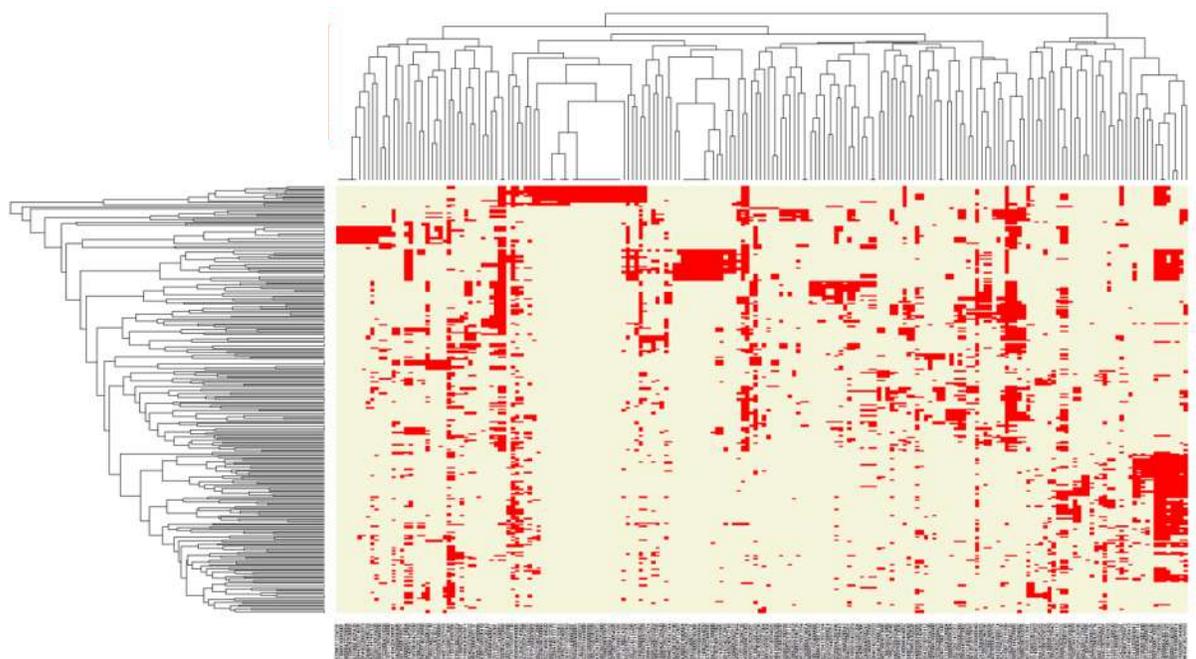
El análisis de agrupamiento se realizó utilizando los 249 compuestos obtenidos como *hits* dentro del programa Canvas. El cálculo del descriptor (ECFP's) se realizó como lo indica el procedimiento sugerido en el programa de donde se obtuvieron los 200 bits representativos, de cada compuesto, después todos los bits fueron arreglados en una matriz junto con los compuestos para poder hacer el cálculo del número de grupos necesarios para su clasificación por el método K-means dentro del programa R Studio, después de 60 grupos no hubo cambios significativos de valores de RMSD entre los centroides u origen de grupos en toda la base de datos de los bits como ECFP's (**Figura 25**). Se obtuvieron dos diagramas de clasificación jerárquica y con la ayuda de una interfaz gráfica se pudo seleccionar a aquellos compuestos que se parecían más a los 23 compuestos activos frente a la PDK1.



**Figura 25.** Gráfico obtenido en el programa R, empleando el método K-means y que representa el número de grupos tomando en cuenta la mínima diferencia de la raíz cuadrada de la distancia entre los centroides de cada grupo.

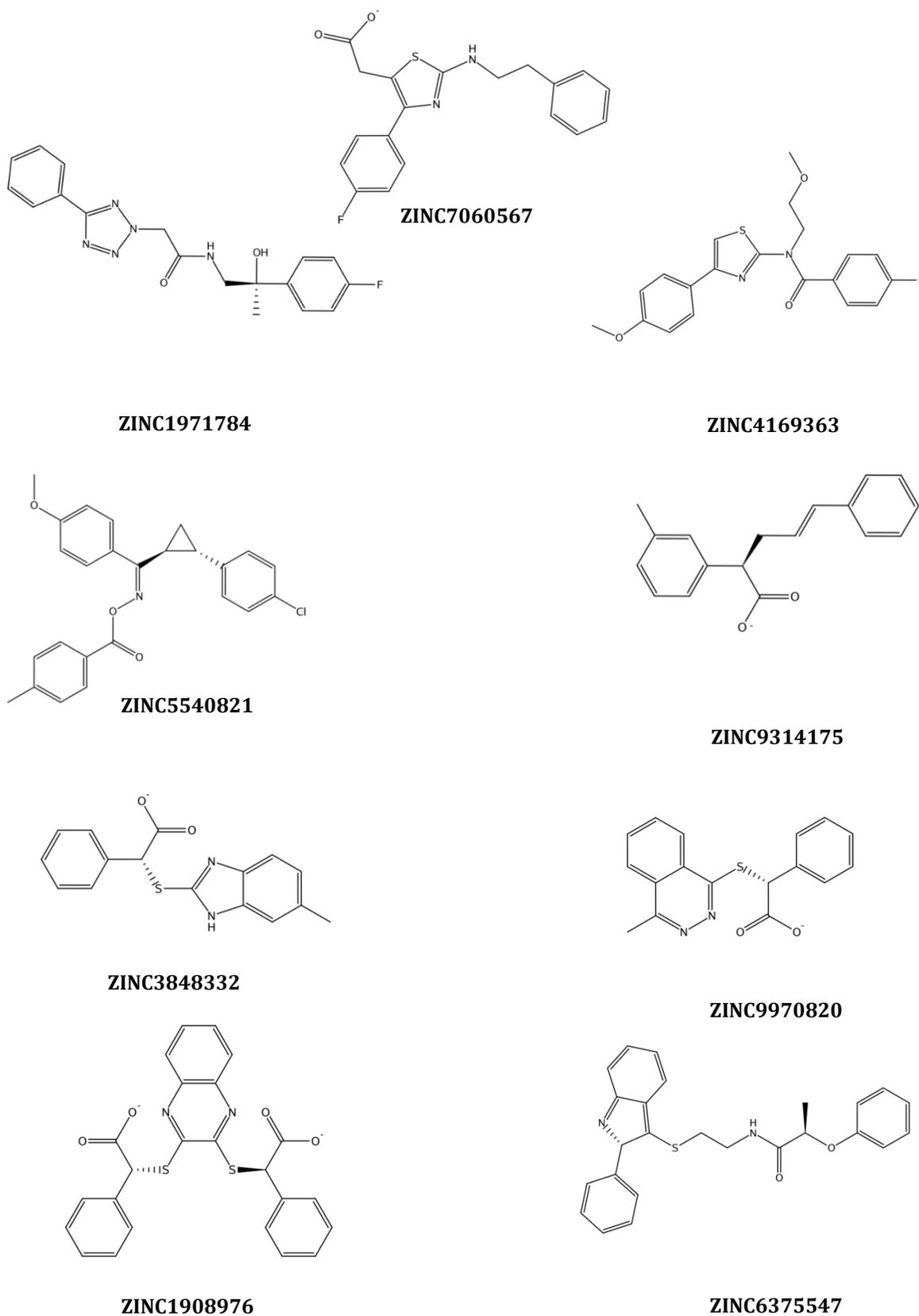
Además de realizó una gráfica tipo *mapas de calor*, donde se pueden visualizar de forma adecuada los grupos de compuestos relacionados entre sí de acuerdo a la similitud en su huella digital (figura 18). Después se decidió seleccionar de forma gráfica los que mejor relación tenían con los 23 activadores de la PDK1 siendo sumamente complicado por lo que esta tarea se realizó mediante una interfaz gráfica dentro del programa Canvas usando como 60 el número de grupos encontrando 9 nuevos compuestos como *hits* (**Figura 26**).

Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1



**Figura 26.** Gráfico que muestra la relación que tienen los compuestos encontrados en el cribado virtual como hits en base a su huella digital (ECFP's).

**Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1**



**Figura 27.** Estructura química de los nueve compuestos seleccionados después del análisis de agrupamiento.

### 3DQSAR

Los resultados de los estudios QSAR en tres dimensiones no fueron satisfactorios como se había pensado al inicio pues se contaba con información valiosa en cuanto a las interacciones que presentan todos los ligandos en la región PIF de la PDK1, los alineamientos fueron adecuados debido a la diferencia de RMSD entre cada ligando, sin embargo, en la **tabla 11** se presentan los resultados obtenidos, en donde podemos observar que se realizaron ocho diferentes protocolos de modelado 3DQSAR utilizando el programa 3DALIGN y 3DQSAR, en donde las variaciones importantes son el espaciado de la malla de cálculo siendo 1.0Å hasta 2.0Å. la diferencia de los cálculos fue en base a el método de alineamiento realizado, por ejemplo, el primero fue realizado a través de un alineamiento flexible dentro del programa MOE Y que proviene de la disposición espacial realizada en los estudios de acoplamiento molecular en este estudio, la segunda proviene de el algoritmo de alineamiento aleatorizado implementado dentro del programa 3DALIGN, y la tercera utilizando el método flexible dentro del programa MOE pero utilizando como plantilla las coordenadas de estructura cristalográfica con el código PDB: **3HRF**.

**Tabla 11.** Resumen de los resultados de los estudios 3DQSAR.

Entry	Box step (Å)	Pc's	R <sup>2</sup>	Q <sup>2</sup>
Alignment from docking poses				
1	2	5	-0.2756	0.1725
2	1.5	5	-0.4492	0.4347
3	1	5	-0.4703	0.5835
Alignment from 3DAlign algorithm				
4	2	5	-1.1687	0.5147
5	1	5	-0.8567	0.4826
Flexible alignment using a rigid template from X-ray				
6	2	5	-0.291	0.5227
7	1.5	5	-0.5799	0.7796
8	1	5	0.7116	-0.2112

Los resultados de los cálculos CoMFA muestran que del alineamiento proveniente de los resultados de acoplamiento molecular se obtienen valores de  $Q^2$  y  $R^2$  que son inaceptables lo que nos muestra que el estudio de acoplamiento molecular en nuestro caso no puede reproducir de la mejor manera resultados de actividad biológica. Además, casi de la misma forma los dos alineamientos que provienen del algoritmo utilizado en 3DALIGN muestran mejores resultados de  $Q^2$  pero no de  $R^2$  donde nos hace pensar que un alineamiento aleatorio no puede predecir de forma confiable los resultados de actividad biológica posiblemente porque algunos ligandos presentan pocos grados de libertad dentro de su estructura química lo cual hace que la posibilidad de encontrar un buen alineamiento se vean reducida. Por otro lado, los protocolos 6, 7 y 8 que provienen del alineamiento en el que se utilizó el ligando dentro del código PDB: 3HRF, mostró mejores resultados tanto de  $Q^2$  y  $R^2$ , sin embargo no correlacionan ambos en los valores obtenidos, mientras que el alineamiento 8 presenta un buen grado de correlación, carece de la capacidad predictiva por lo que no fue buena opción para continuar con los cálculos para predecir la actividad  $pAC_{50}$  de los nuevos compuestos encontrados mediante el cribado virtual. Podríamos pensar que el espacio de la malla de cálculo puede interferir en los resultados de  $Q^2$  y  $R^2$ , sin embargo, no tiene una correlación evidente. Al discutir los resultados anteriormente obtenidos, se tomó en cuenta el número de compuestos sometidos al análisis por CoMFA y se supuso que fueron pocos ligandos con dificultad de alineamiento por lo que se optó por no utilizar un modelo 3DQSAR específicamente con estos compuestos pues no son adecuados para llevar a cabo la predicción de la actividad de los *hits*.

## 2DQSAR

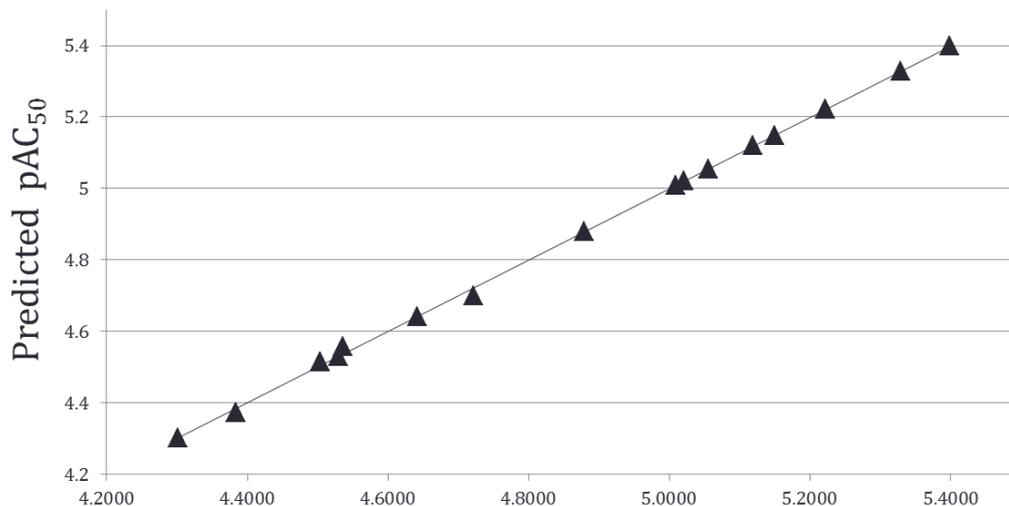
También fue propuesto un estudio 2DQSAR para poder llevar a cabo la predicción de los compuestos obtenidos como *hits* durante el cribado virtual, donde se calcularon 151 descriptores 2D, para obtener un buen modelo predictivo que fue sometido a un análisis de optimización utilizando el método QuaSAR-contingency dentro del módulo SAR del programa MOE que nos ayudó a mejorar el valor predictivo eliminando los descriptores que no aportan un resultado matemático significativo al modelo. La validación fue realizada a través del método *dejando uno fuera* (LOO), siendo aceptables los valores Z menores a 2.5 y con un valor de  $R^2$  de 0.993 (**Figura 28**) proveniente del análisis de regresiones lineales múltiples lo cual nos indica que es un buen modelo que predice adecuadamente la actividad  $pAC_{50}$  de los compuestos activos para la PDK1.

PIF De La PDK1

Training set plot

$$y = 0.9993x + 0.0032$$

$$R^2 = 0.9993$$



**Figura 28.** Gráfico de correlación lineal múltiple efectuado a los 23 compuestos activos para la PDK1 y que muestra el resultado de R2 que nos indica el poder predictivo del modelo de 2DQSAR. La ecuación con los componentes no es presentada debido a que dentro del programa MOE no es posible extraerla por el método empleado

**Tabla 12.** Resultados de la predicción de la actividad pAC50 de los 23 compuestos conocidos como activadores de la PDK1

Compuesto	pAC <sub>50</sub> Experimental	pAC <sub>50</sub> calculada
2Z	5.097	5.017
3Z	5.020	5.020
4Z	4.387	5.106
5Z	5.009	5.009
6Z	5.149	5.149
7Z	5.553	5.295
8Z	5.328	5.328
9Z	5.398	5.398
10Z	4.301	4.301
11Z	5.221	5.221
12Z	5.221	5.221
13E	5.055	5.055
13Z	5.119	5.119
14E	4.642	4.742
15E	4.642	4.642
16E	4.503	4.514
16Z	4.503	4.514
17E	4.384	4.373
17Z	4.384	4.373
18E	4.879	4.879

<b>19</b>	4.721	4.699
<b>20</b>	4.536	4.559
<b>21</b>	4.529	4.529

**Tabla 13.** Resultados de la actividad Pac50 calculada para los nueve compuestos obtenidos del cribado virtual

<b>Compuesto</b>	<b>pAC<sub>50</sub>-calculada</b>
ZINC7060567	3.6390
ZINC19717845	6.2412
ZINC4169363	0.6133
ZINC5540821	4.1221
ZINC93141755	6.1945
ZINC3848332	6.4299
ZINC9970820	6.3973
ZINC1908976	7.6644
ZINC63755473	4.2842

Después los nueve compuestos seleccionados con el método de análisis de agrupamiento (**Figura 27**) fueron sometidos a la predicción de la actividad biológica con el modelo obtenido en este estudio 2DQSAR, y los resultados se presentan en la **tabla 12**.

## 14. ESTUDIO SOBRE UNA PROPUESTA SINTÉTICA DE LOS MEJORES COMPUESTOS OBTENIDOS DEL CRIBADO VIRTUAL

El estudio del DIFAC de ninguna forma puede pensarse de forma aislada de otras disciplinas y campos de la química, es decir, siempre debe ir acompañado al menos de la síntesis de los compuestos encontrados como activos o inhibidores para una diana farmacológica, es por eso que los métodos computacionales contribuyen en gran medida al desarrollo de una investigación en el desarrollo de nuevos compuestos que sean sujetos de optimización incluso de obtener la candidatura a ser fármacos. Por lo anteriormente dicho, en esta parte se presenta el estudio de la propuesta sintética del compuesto **ZINC1908976** teniendo como base las predicciones realizadas en el cribado virtual.

Por lo tanto, se presenta la descripción del trabajo experimental de síntesis orgánica, en donde se muestran los retos en cuanto al manejo de compuestos azufrados como el reactivo de Lawesson en condiciones de reacción poco convencionales como el uso de las microondas que han cobrado gran importancia en el campo de la síntesis orgánica obteniendo buenos resultados.

### *Reactivo de Lawesson*

La transformación del grupo funcional carbonilo en un tiocarbonilo ha sido de gran interés dentro de la química medicinal gracias a la versatilidad que ha presentado el reactivo de Lawesson (LR)<sup>63</sup> (**Figura 29**) descrito en 1978 por Lawesson y colaboradores<sup>64</sup>, aunque algunos otros reactivos incluyen análogos del LR, no han mostrado mayor éxito que este último en cuanto a tiónación de grupos carbonilo se refiere.

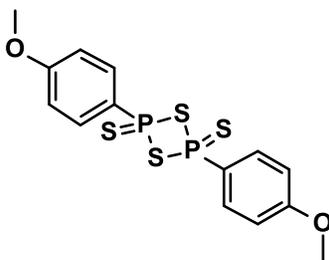
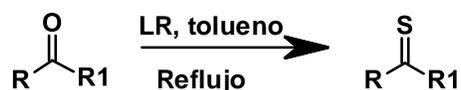


Figura 29. Estructura química del reactivo de Lawesson (LR).

El LR es comercial y ampliamente utilizado en síntesis orgánica aplicado a diferentes tipos de carbonilos como los de cetonas, amidas, lactamas y tiofenos entre otros.

En cuanto a la transformación del grupo carbonilo de las cetonas que es de lo que se encarga en este trabajo, la **figura 30** muestra las condiciones generales que han sido ampliamente exploradas y probadas por otros grupos de trabajo con el uso de solventes como tolueno y xileno en condiciones de reflujo<sup>65</sup>.



*Figura 30. Condiciones de la reacción general del LR con cetonas*

Aunque el tolueno es de los solventes más utilizados en este tipo de reacciones, existen ejemplos del uso de otros como benceno, piridina, THF, CH<sub>2</sub>Cl<sub>2</sub>, y en algunos casos la reacción se lleva a cabo a temperatura ambiente. Aunque también se han reportado productos no deseados en este tipo de reacciones. Existen casos en los que se ha reportado el uso de microondas en las reacciones de cetonas con el LR<sup>66</sup> libre de solventes con tiempos reducidos desde los tres hasta los diez minutos con buenos rendimientos.

Síntesis en microondas:

La energía de microondas además del uso que tiene en la vida cotidiana, ha sido aplicada en varias áreas de la investigación como la digestión, incineración y extracción en química analítica. El primer reporte del uso de microondas en síntesis orgánica fue publicado en 1986 por Gedye y colaboradores<sup>67</sup> y que se presenta en la **figura 31**. Estas reacciones fueron realizadas en hornos de microondas domésticos sin la medición de presión ni temperatura, pero utilizando solventes provocando en algunos casos que no hubiese control en las reacciones terminando en pequeñas explosiones.



Figura 31. Reacción reportada por Gedye y colaboradores con el uso de microondas en síntesis orgánica.

Hoy en día se ha vuelto familiar el uso de las microondas en el laboratorio de síntesis orgánica pues se ha puesto de manifiesto la optimización del rendimiento y reducción de los tiempos de reacción. Además del incremento de las publicaciones que involucran el uso de las mismas (Figura 32)<sup>68</sup>.

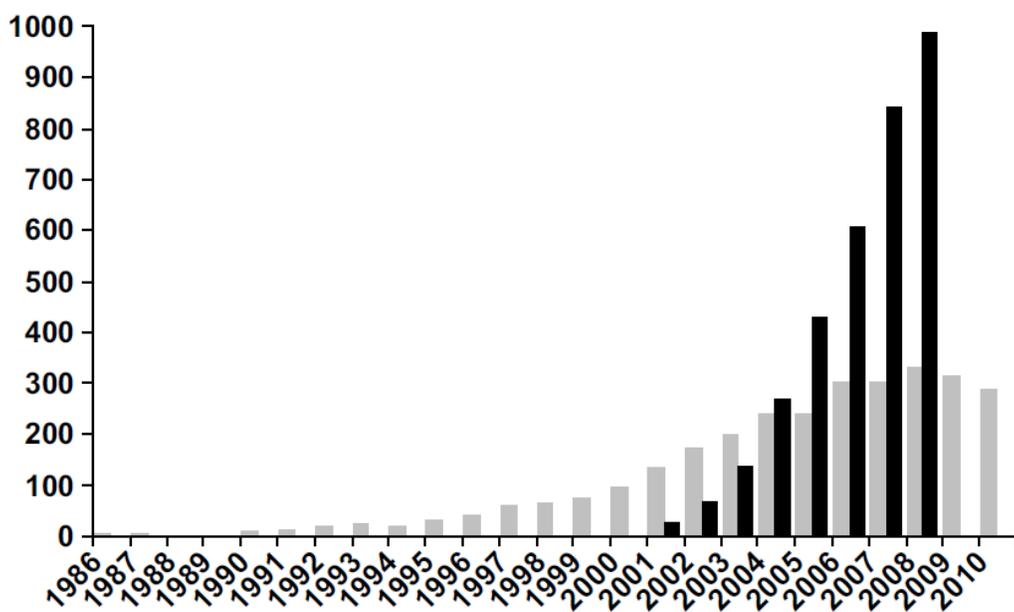
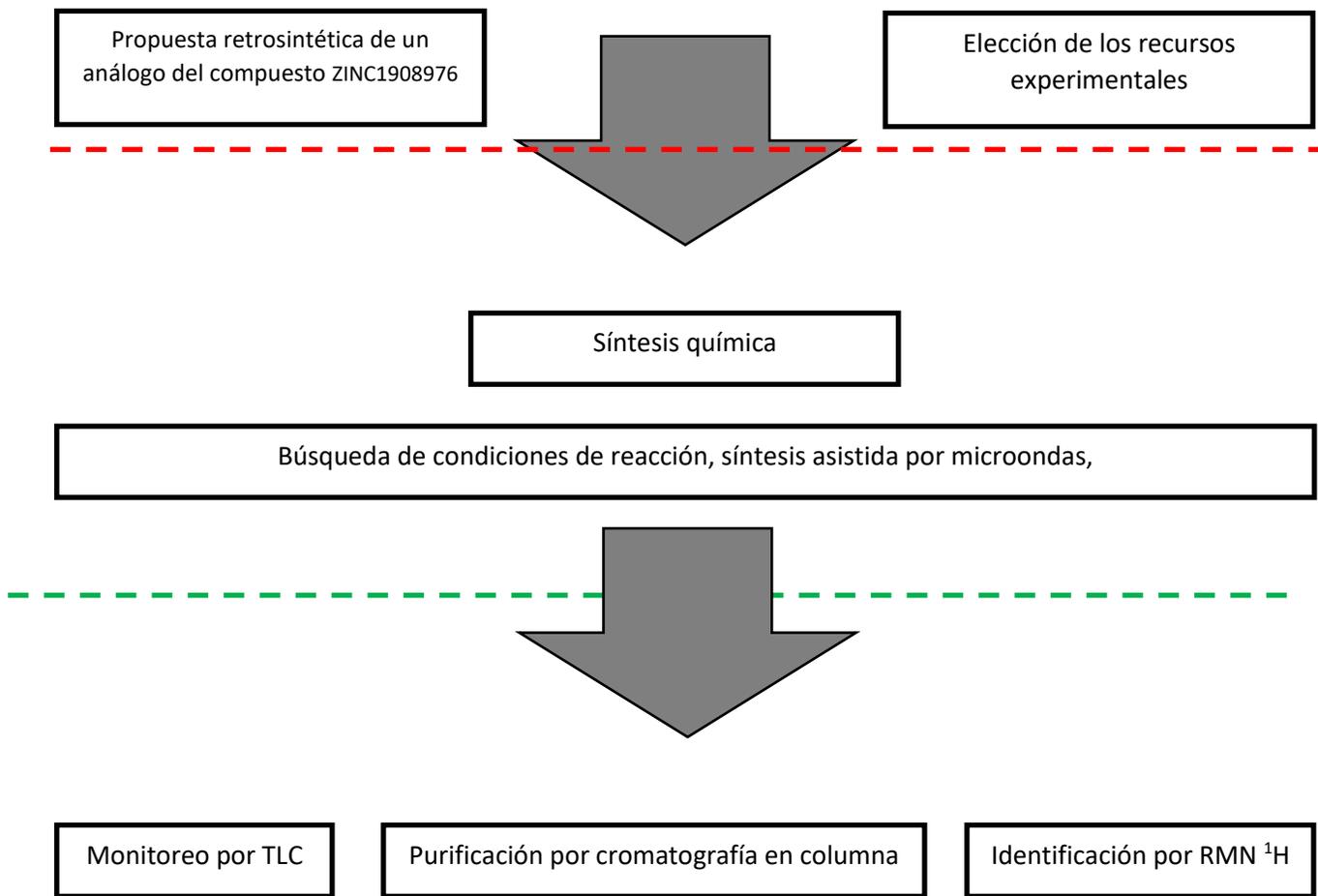


Figura 32. Representación gráfica del número de publicaciones por año desde la primera en 1986 hasta el año 2010. Las barras grises representan aquellos casos en los que no existe un control de la presión y temperatura en el reactor de microondas, y las barras negras representan los casos en donde hubo el control de las variables antes mencionadas

### Material y métodos

La metodología de síntesis general se presenta en el esquema 5.

*Esquema 5. Diagrama de flujo de la metodología empleada en los experimentos de síntesis orgánica*



Propuesta de síntesis orgánica del compuesto **ZINC1908976**

Para poder llevar a cabo una propuesta de síntesis, se realizó el análisis retrosintético del compuesto **ZINC1908976** (**1**) que se muestra en el **esquema 6**, que corresponde con la tiónación del carbonilo en posición 2 del ácido oxo-fenilacético (**7**) proponiendo intermediarios

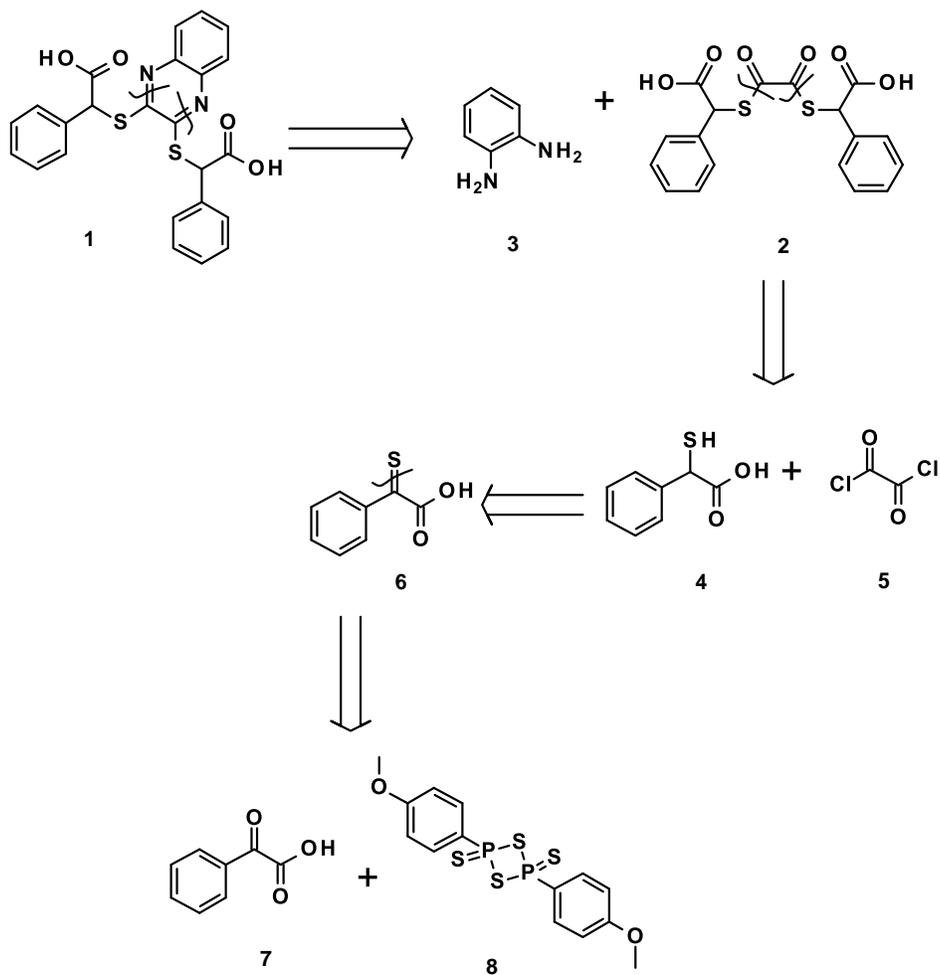
y materia prima convencionales y de manejo conocido en nuestro grupo de trabajo. Al mismo tiempo por la falta de materia prima de partida que corresponde al compuesto **7** se decidió que este ácido fuera cambiado por la acetona correspondiente siendo el compuesto **7'** presentado en el esquema 5 proponiendo así el mismo análisis retrosintético que nos llevaría a un compuesto análogo a **ZINC1908971'6** (**1'**). Las propuestas retrosintéticas de **1** y **1'** en los **esquemas 6 y 7** respectivamente.

#### *Procedimientos generales de síntesis*

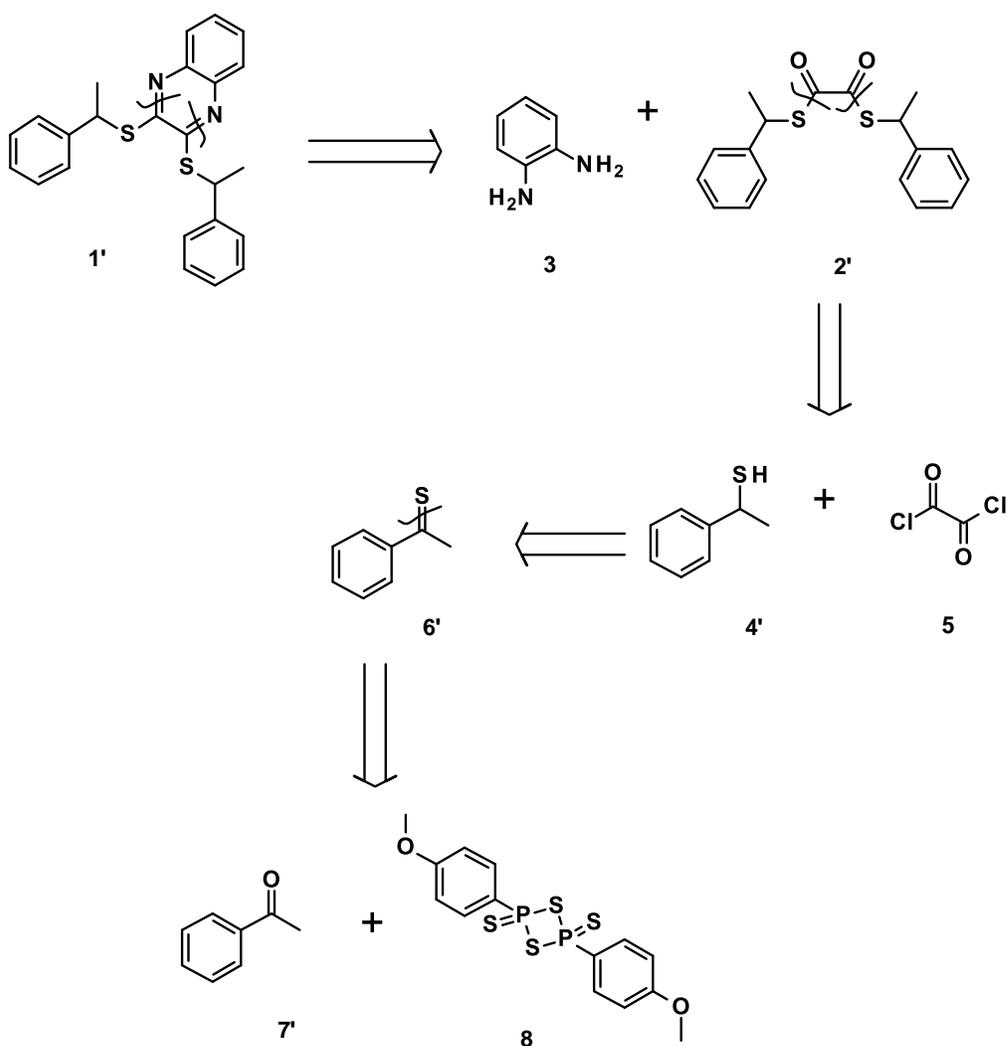
He de admitir que el trabajo de la síntesis orgánica puedo considerarlo un arte estando en la mesa de trabajo dentro de un laboratorio de química pura, pues en ocasiones podemos encontrar procedimientos o protocolos extensos en cuanto a una ruta de síntesis en donde resultados obtenidos no son los esperados llevándonos a explorar nuevas formas de lograr los objetivos, así, el uso de las microondas en el primer paso fue crucial del cual se obtuvo el producto de la primera reacción con buenos rendimientos dejando así la referencia de la tiónación de la acetofenona con el reactivo de Lawesson en ausencia de solventes orgánicos, por lo tanto, se llevaron a cabo condiciones en un microondas doméstico convencional con un voltaje de 1,000 watts. El material de cristalería de laboratorio empleado fue lavado, limpiado con acetona y secado en estufa antes de ser utilizado.

Los reactivos empleados en las etapas de síntesis química fueron adquiridos de forma comercial en la compañía Sigma-Aldrich y la mayoría de las reacciones fueron montadas dentro del ambiente de una campana de extracción. Para el monitoreo de las reacciones fue utilizado el método de cromatografía en capa fina (TLC) empleando cromatofolios con fase fija de gel F254 dentro de una cámara cromatográfica, como reveladores se emplearon vapores de yodo y una lámpara de luz UV a una longitud de onda de 254nm.

Esquema 6. Análisis retrosintético de 1 (compuesto ZINC1908976).



Esquema 7. Análisis retrosintético de 1' (análogo del compuesto ZINC1908976).

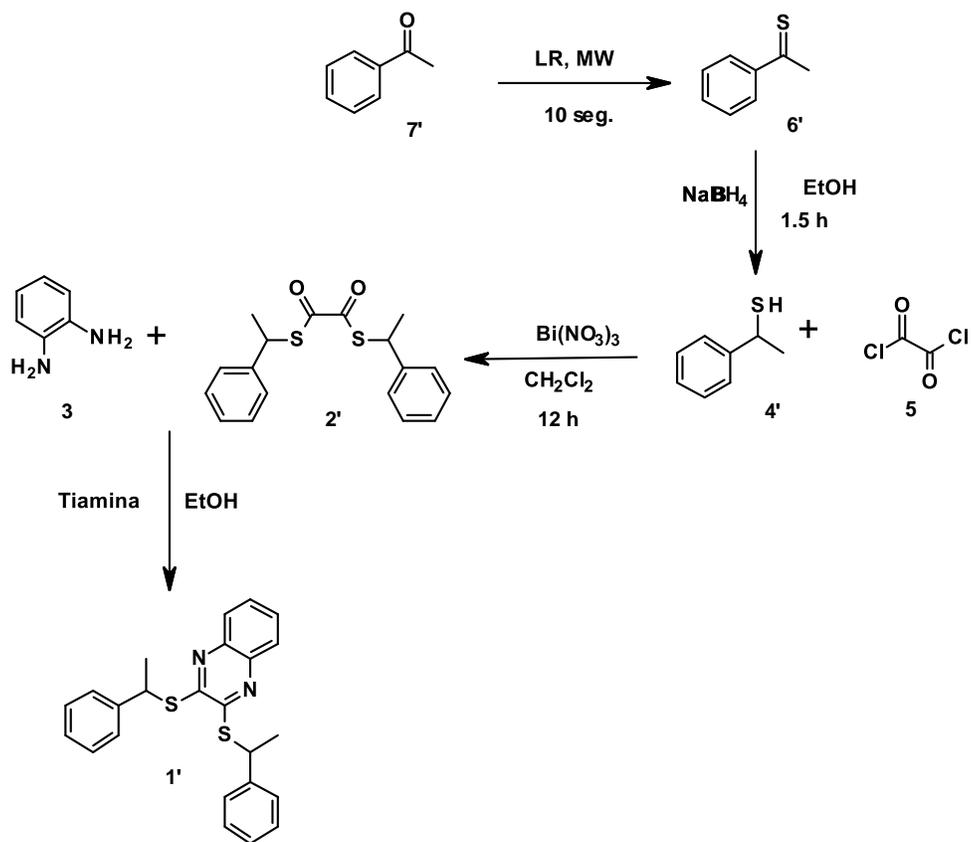


La purificación de los compuestos fue por cromatografía en columna empleando gel de sílice con número de malla: 70-320 de la marca Sigma-Aldrich

La espectroscopía de RMN  $^1\text{H}$  se obtuvo con el equipo Mercury Plus 400 a 100 MHz utilizando  $\text{CDCl}_3$  como solvente deuterado con tetrametilsilano (TMS) como referencia de medición. La propuesta de la síntesis del compuesto **1'** se presenta en el **esquema 8**.

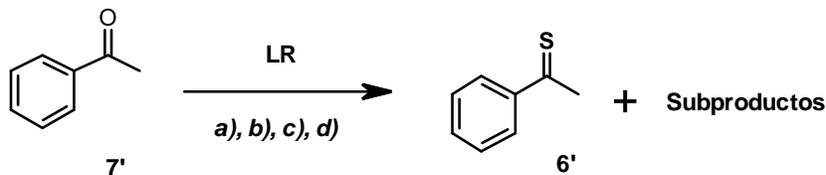
**Esquema 6. Ruta general de síntesis del compuesto 1'.**

**Esquema 8. Ruta general de síntesis del compuesto 1'.**



*Procedimiento de síntesis orgánica*

*Paso 1. Tionación del carbonilo de la acetofenona*



Para lograr el primer paso de la ruta sintética a seguir se consultó la bibliografía para tener referencias de tionación de cetonas con el reactivo de Lawesson (**8**), encontrando en la mayoría de los casos el uso de solventes como el tolueno o xileno en condiciones de reflujo, además del uso de microondas y condiciones de reacción suaves en la minoría 24. Al inicio se decidió emplear condiciones de reacción suaves empleando solventes como acetonitrilo a temperatura ambiente y 55°C y diclorometano a temperatura ambiente utilizando como catalizador al nitrato de bismuto monitoreando el curso de las reacciones por cromatografía en capa fina hasta observar cambios importantes e interesantes en diferentes tiempos, así mismo se optó por el uso de tolueno en condiciones de reflujo según se reporta en la mayoría de las referencias encontradas. La **tabla 14** proporciona un resumen de las condiciones de reacción probadas para el primer paso de la síntesis de 1'

**Tabla 14.** Resumen de las condiciones e intentos de la tionación del carbonilo de la acetofenona

Entrada	Cantidad de 8' (μL)	Cantidad de 9 (mg)	Solvente (a)	Temperatura (b)	Tiempo (c)	Catalizador (d)	Rendimiento (%)
1	14	50	CH <sub>3</sub> CN	Ambiente	12 h	Bi(NO <sub>3</sub> ) <sub>3</sub>	--
2	14	50	CH <sub>3</sub> CN	55 °C	12 h	Bi(NO <sub>3</sub> ) <sub>3</sub>	--
3	14	50	CH <sub>2</sub> Cl <sub>2</sub>	Ambiente	12 h	Bi(NO <sub>3</sub> ) <sub>3</sub>	--
4	14	50	Tolueno	79°C	12 h	Bi(NO <sub>3</sub> ) <sub>3</sub>	--
5	50	80	--	Microondas	3 min	--	--
6	50	89	--	Microondas	10 seg	--	66

El procedimiento de cada reacción se llevó a cabo de forma general como se describe a continuación variando solamente la temperatura y el disolvente.

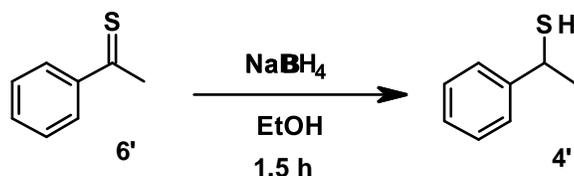
Se disolvió el reactivo de Lawesson (50 mg) en acetonitrilo (3 mL) en un vial de reacción, se colocó en agitación por espacio de 5 minutos y se adicionó la acetofenona (1 eq., 14 $\mu$ L) lentamente dejando reaccionar la mezcla a temperatura ambiente (reacciones 1 y 3) y a 55°C (reacción 2) y a 79°C (reacción 4 en un sistema de reflujo) haciendo un monitoreo por cromatografía en capa fina de la reacción cada 30 minutos durante 2 horas, al cabo de ese tiempo no se observó algún cambio que indicase que hubiera algún producto de interés por lo que se dejó en agitación durante 12 horas. Cumplido este tiempo, se evaporó el disolvente por presión reducida y se sometió el crudo de reacción a purificación por cromatografía en columna empleando como eluente una mezcla de hexano/acetato 8:2, se enviaron a RMN  $^1\text{H}$  las fracciones que probablemente pudieran tratarse del producto deseado. En el análisis espectroscópico por RMN  $^1\text{H}$  de las reacciones 1 - 6 mostró que no había señales de desplazamiento químico que pudiera mostrar la presencia del producto deseado. Fue hasta los productos de la reacción 6 que con el empleo de microondas y sin solventes como condiciones de reacción, se obtuvo información espectroscópica de señales con un desplazamiento químico que pudiera indicarnos la presencia del producto **6'**.

A continuación, se describe el procedimiento general de la reacción 6:

Se colocó 80 mg de reactivo de Lawesson (0.5 eq) en un tubo de ensayo al que posteriormente se adicionaron 50  $\mu$ L (1 eq.) de acetofenona procurando mezclar la totalidad de ambos reactivos obteniéndose una mezcla de color amarillo claro, posterior a esto, se colocó el tubo de ensayo con la preparación dentro de un vial que contenía alúmina que sirvió como soporte para evitar que la radiación de las microondas chocara directamente con el tubo de reacción. La fuente de microondas fue en un horno casero convencional de 300 W en donde se sometió la reacción a dos ciclos de 5 segundos cada uno, obteniéndose una mezcla de color morado el cual se analizó de forma completa, el sobrenadante y el precipitado por cromatografía en capa fina de los cuales tanto la mezcla como el sobrenadante mostraron una separación de un compuesto menos polar no así para el precipitado por lo que se filtró y el sobrenadante se sometió a purificación por columna empleando como eluente una mezcla de hexano/acetato 8:2 obteniendo 3 fracciones con un solo compuesto a las que se les retiró la mezcla de eluente con presión reducida obteniendo un producto semisólido que posteriormente se envió a RMN  $^1\text{H}$

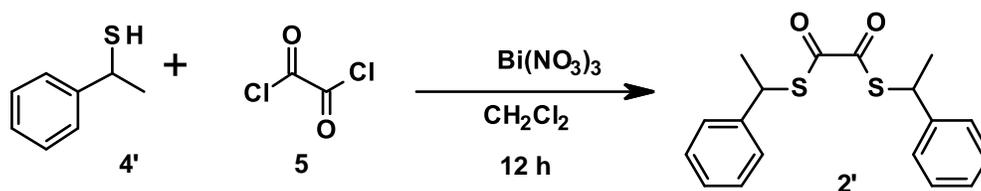
obteniendo un resultado de espectroscopía sugerente del producto deseado con un rendimiento de 66% aproximadamente (34 mg).

*Paso 2. Reducción del compuesto 6'*



Se disolvieron 34 mg del compuesto 6' en EtOH (2 mL) dentro de un vial de reacción, posteriormente se adicionó NaBH<sub>4</sub> (1 eq., 15 mg) dejándose en agitación a temperatura ambiente por espacio de 2 horas monitoreando la reacción cada 30 minutos por cromatografía en capa fina observando que no hubo algún cambio después de 1.5 horas de reacción, se retiró la reacción evaporando el solvente por presión reducida y se sometió a purificación en cromatografía en columna usando como eluente una mezcla de hexano:acetato 9:1 obteniendo 2 fracciones sugerentes del producto deseado con aspecto semisólido enviándose a RMN <sup>1</sup>H obteniendo la espectroscopía sugerente del compuesto 6' con un rendimiento de 11.5 % aproximadamente (3.9 mg)

*Paso 3. Formación de 2'*



Se disolvió el compuesto 4' (3.9 mg) en CH<sub>2</sub>Cl<sub>2</sub> (1 mL) en un vial de reacción en donde posteriormente se adicionó cloruro de oxalilo (0.03 eq., 10 µL) empleando Bi(NO<sub>3</sub>)<sub>3</sub> (0.1 eq., 5.7 mg) como catalizador y se sometió a agitación monitoreando la reacción cada 30 minutos por 2 horas, no habiendo cambios aparentes por lo que se dejó agitando 10 horas más. Transcurrido este tiempo, se detuvo la reacción agregando agua destilada (5 mL) y se mezcló con DCM (3x10 mL), de donde se recolectó la fase orgánica y se sometió a deshidratación con Na<sub>2</sub>SO<sub>4</sub> anhidro, se filtró la solución y se concentró evaporando el disolvente bajo presión reducida y se purificó en columna cromatográfica obteniendo tres fracciones que nos sugieren que puede ser el producto 2', sin embargo, no fueron suficientes las cantidades de estas fracciones por lo que no se obtuvieron resultados espectroscópicos.

### **Resultados**

Se presentan los resultados de la propuesta sintética para la obtención del compuesto 1' que se ha replanteado en este trabajo como se describe en el análisis retrosintético teniendo la formación de 6' y 4' como se presenta en la espectroscopía de RMN <sup>1</sup>H (**Figura. 33 y 34 respectivamente**)

Diseño Y Síntesis De Compuestos Potencialmente Translocadores De GLUT4 A Través Del Reconocimiento De La Región PIF De La PDK1

H\_ASC-Rx.06  
ASC-Rx.06  
Dr. Luis Chacon  
4-dic-15

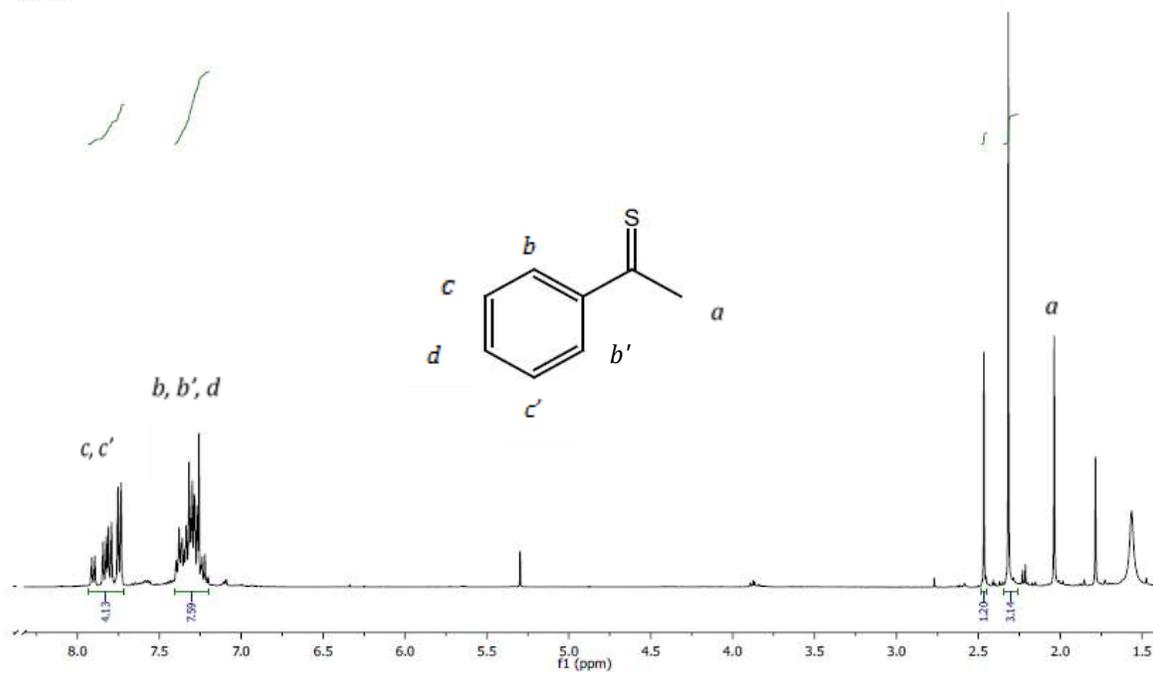


Figura 33. Espectro de <sup>1</sup>H RMN del compuesto 6' (400 MHz, CDCl<sub>3</sub>).

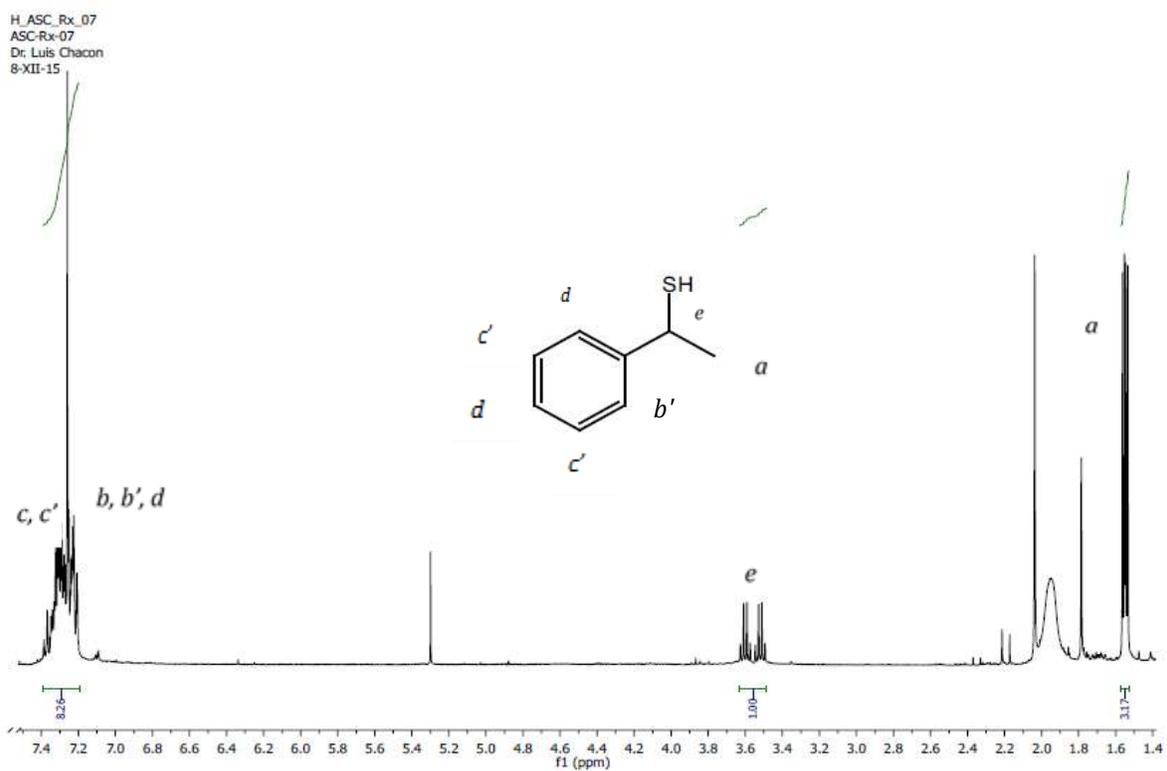


Figura 34. Espectro de  $^1\text{H}$  RMN del compuesto 4' (400 MHz,  $\text{CDCl}_3$ ).

Tabla 15. Resumen de los desplazamientos químicos de los compuestos 6' y 4'

Compuesto 6'		Compuesto 4'	
Señal	ppm	Señal	ppm
<b>a</b>	<b>2.0</b>	<b>a</b>	<b>1.6</b>
<b>b, b'</b>	<b>7.2</b>	<b>b, b'</b>	<b>7.2</b>
<b>c, c'</b>	<b>7.8</b>	<b>c, c'</b>	<b>7.8</b>
<b>d</b>	<b>7.2</b>	<b>d</b>	<b>7.2</b>
--	--	<b>e</b>	<b>3.6</b>

Para cumplir el objetivo de la síntesis del compuesto **1** se hizo la exploración de las condiciones de reacción de la tiónación de la acetofenona encontrando la aplicación de las microondas como una buena opción de síntesis del compuesto **6'** sin el uso de solventes y con un rendimiento del 66% aproximadamente del cual los resultados de espectroscopía de RMN  $^1\text{H}$  muestran un desplazamiento químico característico de un anillo de benceno monosustituído en 7.2 y 7.4 ppm y un desplazamiento químico del grupo metilo de la tiona a 2.1 ppm encontrándose a campo alto comparado con el desplazamiento químico del metilo de la acetofenona de 2.6 ppm.

Por otro lado, la reducción de **6'** para obtener **4'** se observa que tiene un rendimiento del 11.5 % aproximadamente mostrando una señal con un desplazamiento químico aproximadamente en 3.6 ppm correspondiente con un metino base de tiol, además se aprecia el desplazamiento a campo alto del grupo metilo con un desplazamiento químico a 1.6 ppm lo que equivale a 0.4 ppm aproximadamente con respecto del metilo del compuesto **6'**.

Durante la evaluación de los sinodales para esta parte del trabajo, se hicieron varias observaciones al respecto, como una posible exploración a las condiciones de reacción sobre todo en el segundo paso de la ruta de síntesis que es la reducción de la tiona, puesto que se obtiene un rendimiento del 11.5% aproximadamente, cabe señalar que la calidad del borhidruro de sodio tal vez no era la apropiada y por esa razón pudo haberse obtenido un rendimiento muy bajo. No obstante, no se decidió terminar el trabajo de síntesis hasta esta etapa, pensando también que alguien más puede retomar el trabajo de síntesis y además probar el comportamiento que tendría el compuesto frente a la PDK1 in vitro una vez que sea sintetizado.

## 15. CONCLUSIONES GENERALES Y PERSPECTIVAS

En este trabajo se llegó a varias conclusiones interesantes e importantes, en primer lugar, el haber explorado in silico una región alostérica de una cinasa como lo es su región PIF abre el camino para nuevas propuestas sobre estudios in silico que involucren a estos sistemas, pues se pudieron conocer las condiciones de un buen protocolo de estudios de acoplamiento molecular y la generación de una adecuada y eficiente hipótesis farmacofórica con el objetivo de realizar un cribado virtual. Entre las nuevas aportaciones podemos resaltar el inicio de la exploración in silico de la PDK1 enfocado al diseño de fármacos contra la diabetes mellitus. El trabajo aquí presentado se dedicó en su mayor parte al diseño y al estudio de nuevos

compuestos que pudieran activar a la PDK1 en cuanto a su similitud molecular a través del cribado virtual, a los grupos funcionales importantes en base a su farmacóforo y su probable interacción con la región PIF de la PDK1 por estudios de acoplamiento molecular. El inicio del proyecto fue realizado en base al ligando teniendo en cuenta los compuestos activadores que ya se conocían y de los que ya se cuenta con información experimental, tomando como base al núcleo de la chalcona, sobresaliendo algunos compuestos como 23a y que fueron propuestos de novo, sin embargo la propuesta sintética de algunos de ellos tal vez no se llevó a cabo porque la información predicha no es suficiente, de hecho, en cualquier proyecto de diseño y modelado de compuestos no basta con obtener información de las estructuras que pueden interactuar con el sitio de interés en el receptor sino que se necesita estar muy seguro el investigador que se tiene toda la información complementaria para proponer la síntesis de un nuevo compuesto.

Es así como se dio paso a la realización de un cribado virtual a partir de una base de datos muy grande como lo es el ZINC data base y aprovechando la estancia en la Universidad de Cambridge en el Reino Unido, donde se contó con buenas herramientas, sobre todo de gran capacidad computacional, haciendo el proceso y los cálculos de acoplamiento y de modelado farmacofórico ágiles y confiables, de donde se obtuvieron nueve compuestos con varios resultados en cuanto a la afinidad, modo de unión e interacciones importantes con la PDK1, además de que se realizaron algunos cálculos de similitud molecular como su huella digital, haciendo que la elección de los mejores compuestos fuera relativamente fácil entre 242 compuestos candidatos, que tal vez existan algunos con resultados prometedores y a los que no se les predijo la actividad por medio de los estudios QSAR. En la última parte de este proyecto, se realizó la propuesta de síntesis orgánica del compuesto ZINC1908976 en base a su resultado de actividad pAC50 predicha por medio de estudios QSAR, llegando hasta la segunda etapa de la ruta sintética después de un amplio análisis retrosintético y empleando los recursos con los que se cuenta en el laboratorio de nuestro grupo de trabajo. La síntesis de los compuestos propuestos requerirá de un trabajo de laboratorio y tiempo que están fuera del alcance del tiempo de conclusión de esta tesis. En principio se presenta la propuesta y los primeros pasos para llegar al compuesto objetivo. Mucho se ha dicho que la búsqueda de fármacos es un área interdisciplinaria y como he comentado durante el desarrollo de la última parte de esta tesis, no me considero un buen químico en síntesis orgánica por el manejo y habilidades a la hora de estar frente a la mesa de trabajo, no obstante considero que así como han compartido algunos expertos en el área de la química computacional, ahora se ve a ésta como apoyo para el químico

que se dedica exclusivamente a la síntesis orgánica, y por lo tanto considero que dentro del centro donde me he preparado y formado para obtener el grado de Maestría en Ciencias, se necesita tener un amplio criterio de incluir a todas las técnicas, herramientas y áreas de la química para que los proyectos de mis compañeros químicos puedan rendir más frutos que lleven en un futuro a contribuir con nuestra casa de estudios y después con la sociedad de la cual pertenecemos y que está cada vez más necesitada de soluciones a los problemas en donde se ve involucrada la química.

Para el caso de la investigación básica de tipo académico en el área de la química medicinal que es donde se fundamenta este trabajo, puedo concluir lo siguiente:

El DIFAC, ha tenido un gran avance durante los últimos 20 años, y en los que ha apoyado al área de la química medicinal y para aplicar los métodos a los que se tiene acceso en nuestro grupo de trabajo, se decidió estudiar otras probables dianas farmacológicas contra la diabetes mellitus de la cual podemos comentar que es una enfermedad preocupante de la cual se tienen predicciones de cómo estaremos como población afectada dentro de los próximos 10 años en cuanto a las complicaciones y el gasto económico que esta enfermedad representa. Tal vez con esta tesis no podemos afirmar con un grado elevado de certeza que la cinasa PDK1 que fue elegida como receptor durante todos nuestros estudios, es una nueva forma de combatir a la diabetes, porque las vías y formas de validación de nuevas dianas farmacológicas son diferentes a los estudios computacionales, sin embargo de varias fuentes pudimos constatar que efectivamente tiene una relación con dicha enfermedad, y que de ser encontrado algún mecanismo para su activación con compuestos de bajo peso molecular como los reportados por Engel y cols, además de los estudios de Adriana Stroba y cols, se puede avanzar en los nuevos diseños con métodos y tecnología mucho que ofrezca el conocimiento de los factores que se han olvidado tanto en los estudios experimentales como en los desarrollados en este trabajo, por lo que el valor más importante que tiene esta tesis es poder aportar una perspectiva nueva y diferente de atacar a la diabetes mellitus a través del DIFAC en donde otros investigadores se interesen y tomen como base nuestros resultados y logros. El haber empleado en un inicio a la chalcona como núcleo base fue en base a los requisitos mínimos para que un compuesto de bajo peso molecular pueda interactuar con la región PIF, y de los resultados obtenidos, aunque no sintetizados, nos dio la pauta para continuar con los estudios del DIFAC basados en el ligando, pero ahora a través de un cribado virtual empleando las técnicas de similitud y acoplamiento molecular además de una búsqueda farmacofórica. Después del cribado virtual puedo concluir

que el empleo y la combinación de los dos enfoques, algunos basados en la estructura y otros en el ligando, nos abrió el panorama de cómo con tan poca información experimental de sólo 23 compuestos conocidos como activadores de la PDK1, se pudieron conseguir buenos resultados, lo que al principio de ese trabajo, se veía muy complicado poder implementar los métodos del DIFAC, gracias a Dios y con la ayuda de algunos expertos del DIFAC puedo decir que fueron satisfactorios los resultados, así mismo discutidos con otros investigadores en congresos locales e internacionales, recibiendo críticas de mejoramiento de los distintos enfoques pero también comentarios alentadores en cuanto a la forma en cómo se combinaron las herramientas que ofrece el DIFAC.

Para finalizar, si tomamos en cuenta dentro de la investigación básica en el diseño de fármacos, que en total son años aproximadamente, el papel del DIFAC cuando es empleado, no le bastan sólo dos o tres años para poder lograr obtener resultados altamente confiables en cuanto a su predicción, por lo que considero que los resultados aquí obtenidos tienen un gran potencial que puede ser mejorado empleando otros métodos del DIFAC y así complementarlo logrando mejores resultados, con el fin de contribuir al desarrollo de la sociedad en la que vivimos. Así mismo, creo conveniente afirmar que entre más se obtienen los resultados, datos y pistas para resolver un problema dado, más dudas e incógnitas aparecen, por lo que por mi parte estoy abierto a que sea refutado y mejorado en cualquier momento en el futuro.

## 16. REFERENCIAS

- (1) José L. Medina-Franco, Eli Fernández-de Gortari y J. Jesús Naveja. Avances en el diseño de fármacos asistido por computadora. *Ed. Quim. UNAM*. **2015**, 26, 180 – 186
- (2) Peter W. Rose, Andreas Prli, Chunxiao Bi, Wolfgang F. Bluhm, Cole H. Christie Shuchismita Dutta, Rachel Kramer Green, David S. Goodsell, John D. Westbrook, Jesse Woo, Jasmine Young, Christine Zardecki, Helen M. Berman, Philip E. Bourne and Stephen K. Burley. The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *Nucleic Acids Research*, **2015**, 43, Database issue D345–D356.
- (3) José Luis Medina Franco. Aplicaciones exitosas del diseño de fármaco utilizando métodos computacionales *Ciencia* enero-marzo **2007**.

- (4) Gabriel Cuevas; Fernando Cortés. Introducción a la Química Computacional. Fondo de Cultura Económica "La ciencia en el Nuevo Milenio" Monografías de las Redes Latinoamericanas de Ciencias. México D.F. **2003**
- (5) Gregory Sliwoski, Sandeepkumar Kothiwale, Jens Meiler, and Edward W. Lowe, Jr. Computational Methods in Drug Discovery. *Pharmacological Reviews*, **2014**, 66, 334 – 395.
- (6) David C. Young. A Guide for Computational and Medicinal Chemists. Wiley & Sons, Inc. Hoboken New Jersey **2009**.
- (7) Bambini S and Rappuoli R. The use of genomics in microbial vaccine development. *Drug Discov Today* **2009**, 14:252–260.
- (8) Lundstrom K (2011) Genomics and drug discovery. *Future Med Chem.* **2011**, 3, 1855–1858.
- (9) Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei and Meng Cui, Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery, *Current Computer-Aided Drug Design*, 2011, 7, 146-157
- (10) Sam Z. Grinter and Xiaoqin Zou, Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design, *Molecules* **2014**, 19, 10150-10176
- (11) Elizabeth Yurieva, Jessica Holien and Paul A. Ramsland, Improvements, trends, and new ideas in molecular docking: 2012–2013 in review, *Molecular Recognition* **2015**, (wileyonlinelibrary.com) Doi: 10.1002/jmr.2471
- (12) Khaled M. Elokely and Robert J. Doerksen, Docking challenge: Protein sampling and molecular docking performance, *Chemical Informatics Modelling*, 2013 26, 1934–1945.
- (13) Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **1983**, 4, 187-217.
- (14) Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, **1990**, 112, 6127-6129.
- (15) Thierry Langer and Rémy D. Hoffmann, Pharmacophores and Pharmacophore Searches, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany **2009**.

- (16) Korolkovas A. Essentials of Medicinal Chemistry, Wiley Interscience New York, **1988**.
- (17) Marijn P. A. Sanders, Arménio J. M. Barbosa, Barbara Zarzycka, Gerry A.F. Nicolaes, Jan P.G. Klomp Jacob de Vlieg and Alberto Del Rio. Comparative Analysis of Pharmacophore Screening Tools, J. Chem. Inf. Model. **2012**, 52, 1607–1620
- (18) William R. Rowley et al. Mary Ann, Diabetes 2030: Insights from Yesterday, Today, and Future Trends. Liebert, Inc. **2016**
- (19) [www.who.int/diabetes/global-report](http://www.who.int/diabetes/global-report) World Health Organization 2016 WHO/NMH/NVI/16.3
- (20) World Health Organization, Global Status Report on noncommunicable diseases, **2014**
- (21) Lerman, I. Atención integral del paciente diabético; Mc Graw Hill México **2011**.
- (22) Encuesta Nacional de Salud y Nutrición **2012** (s.f.) recuperado en septiembre del 2013.
- (23) Organización Mundial de la Salud – Perfiles de los países para la diabetes, México **2016**
- (24) Standars of Medical Care in Diabetes, Diabetes Care, **2015** 38, 140–149
- (25) Goodman and Gilman Las bases farmacológicas de la TERAPÉUTICA; Mc Graw Hill, México **2012**.
- (26) Khalid Mahmood, Metformin: The hidden chronicles of a magic drug. European Journal of Internal Medicine **2013** 24, 20–26
- (27) American Medical Association, Update on SGLT2 inhibitor warning News & Analysis January 19, **2016** Volume 315, Number 3
- (28) Paromita King, Ian Peacock and Richard Donnelly, The UK Prospective Diabetes Study (UKPDS): clinical and therapeutic implications for type 2 diabetes, Journal of Clinical Pharmacology, **1999**, 48, 643–648
- (29) Carlos A. Aguilar Salinas, Sergio Hernández Jiménez, Mauricio Hernández Ávila, Juan Eugenio Hernández Ávila. Acciones para enfrentar a la Diabetes, Documento de postura. Academia Nacional de Medicina /México **2015**
- (30) Stephanie Aleskow Stein, Elizabeth Mary Lamos, and Stephen N Davis, review of the efficacy and safety of oral antidiabetic drugs Expert Opin. Drug Saf.; **2013**, 12, 153–175

- (31) K.L.R. Wick A New Molecular Target of Insulin Action: Regulating the Pivotal PDK1 Current Drug Targets - Immune, Endocrine & Metabolic Disorders, **2001**, 1, 209-22
- (32) Laura A. Lopez-García. Allosteric Regulation of Protein Kinase PKC $\zeta$  by the N-Terminal C1 Domain and Small Compounds to the PIF-Pocket Chemistry & Biology **2011**, 18, 1463-1473
- (33) Philip Cohen; PDK1 one of the missing links in Insulin signal transduction? FEBS Letters **1997** 410, 3-10.
- (34) Julie A Zorn & James A Wells Turning enzymes ON with small molecules nature Chemical Biology, 6; **2010**
- (35) Lawlor MA, Mora A, Ashby PR, Williams MR, Murray-Tait V, Malone L, Prescott AR, Lucocq JM, Alessi DR Essential role of PDK1 in regulating cell size and development in mice. EMBO J **2002**, 14, 3728-3738
- (36) Mora A, Davies AM, Bertrand L, Sharif I, Budas GR, Jovanovic S, Mouton V, Kahn CR, Lucocq JM, Gray GA, Jovanovic A, Alessi DR Deficiency of PDK1 in cardiac muscle results in heart failure and increased sensitivity to hypoxia. EMBO J **2003**, 18, 4666-4676
- (37) Bayascas JR, Wullschleger S, Sakamoto K, Garcia-Martinez JM, Clacher C, Komander D, Van Aalten DM, Boini KM, Lang F, Lipina C, Logie L, Sutherland C, Chudek JA, van Diepen JA, Voshol PJ, Lucocq JM, Alessi DR Mutation of the PDK1 PH domain inhibits protein kinase B/Akt, leading to small size and insulin resistance. Mol Cell Biol **2008** 28, 3258-3272
- (38) Adriana Stroba et al. Substrate-Selective Inhibition of Protein Kinase PDK1 by Small Compounds that Bind to the PIF-Pocket Allosteric Docking Site Chemistry & Biology, **2012**, 19,1152-1163
- (39) Matthias Engel; Allosteric activation of the protein kinase PDK1 with low molecular weight compounds, The EMBO Journal **2006**, 25, 5469-5480
- (40) João Henrique G. Lago Structure-Activity Association of Flavonoids in Lung Diseases Molecules **2014**, 19, 3570-3595.
- (41) S. Iwata, T. Nishino, H. Inoue, N. Nagata, Y. Satomi, H. Nishino, and S. Shibata, "Antitumorogenic activities of chalcones (II). Photoisomerization of chalcones and the correlation with their biological activities," Biological and Pharmaceutical Bulletin, **1997**, 18,1266-1270

- (42) Y. Kim, J. Kim, "Antiinflammatory activity of the synthetic chalcona derivatives: inhibition of inducible nitric oxide synthase-catalyzed nitric oxide production from lipopolysaccharide-treated RAW 264.7 cells," *Biological and Pharmaceutical Bulletin*, **2007**, 30, 1450-1453
- (43) Tobias C. Wabnitz and Jonathan B. Spencer\* A General, Brønsted Acid-Catalyzed Hetero-Michael Addition of Nitrogen, Oxygen, and Sulfur Nucleophiles *Org. Lett.*, **2003**, 5, 12, 2141-2144
- (44) Adriana Stroba, Francis Schaeffer, Valerie Hindie, Laura Lopez-Garcia, Iris Adrian, Wolfgang Frohner, Rolf W. Hartmann, Ricardo M. Biondi and Matthias Engel, 3,5-Diphenylpent-2-enoic Acids as Allosteric Activators of the Protein Kinase PDK1: Structure-Activity Relationships and Thermodynamic Characterization of Binding as Paradigms for PIF-Binding Pocket-Targeting Compounds, *Journal of Medicinal Chemistry* **2009**, 52, 4683 – 4693
- (45) Jonathan B. Baell and Georgina A. Holloway. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, 2719–2740
- (46) Marcel L. Verdonk, Valerio Berdini, Michael J. Hartshorn, Wijnand T. M. Mooij, Christopher W. Murray, Richard D. Taylor, and Paul Watson Virtual Screening Using Protein-Ligand Docking: Avoiding Artificial Enrichment *J. Chem. Inf. Comput. Sci.* **2004**, 44, 793-806
- (47) John J. Irwin and Brian K. Shoichet ZINC – A Free Database Commercially Available Compounds for Virtual Screening *J Chem Inf Model*, **2005** 45,177–182
- (48) Cereto-Massague, A., Guasch, L., Valls, C., Mulero, M., Pujadas, G., & Garcia-Vallve, S. DecoyFinder: an easy-to-use python GUI application for building target-specific decoy sets. *Bioinformatics*, **2012**, 2-3.
- (49) David Ryan Koes and Carlos J. Camacho ZINCPharmer: pharmacophore search of the ZINC database *Nucleic Acids Research*, **2012**, Doi:10.1093/nar/gks378
- (50) Marcel L. Verdonk, Jason C. Cole<sup>2</sup>, Michael J. Hartshorn, Christopher W. Murray and Richard D. Taylor Improved protein–ligand docking using GOLD *Proteins: Structure, Function, and Bioinformatics*, **2003**, 52, 609–623;
- (51) Molecular Operating Environment (MOE), 2012.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2012**

- (52) GOLD User Guide A Component of the GOLD Suite Cambridge Crystallographic Data Centre Registered Charity No 800579 **2014**.
- (53) Garrett M. Morris, David S. Goodsell, Robert S. Halliday, Ruth Huey, William E. Hart, Richard K. Belew, Arthur J. Olson, Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function, *Journal of Computational Chemistry*, **1998**,19, 1639 -1662
- (54) Lavecchia, A. Machine-learning approaches in drug discovery: methods and applications, *Drug Discov Today* **2014**, <http://dx.doi.org/10.1016/j.drudis.2014.10.012>
- (55) Huang, Shoichet and Irwin, *Journal of Medicinal Chemistry*, **2006**, 49, 6789-6801
- (56) David Rogers and Mathew Hahn Extended-Connectivity Fingerprints *J. Chem. Inf. Model* **2010**, 50, 742–754
- (57) Duan, J.; Dixon, S.L.; Lowrie, J.F.; Sherman, W., "Analysis and Comparison of 2D Fingerprints: Insights into Database Screening Performance Using Eight Fingerprint Methods," *J. Molec. Graph. Model.* **2010**, 29; 157-170
- (58) R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0; **2008**.
- (59) MacQueen, J. B. "Some Methods for classification and Analysis of Multivariate Observations". *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* University of California Press. **1967**,281–297
- (60) Hugo Kubinyi, *3D QSAR in Drug Design Volume 3 Recent Advances*, Kluwer academic publishers, New York **2002**.
- (61) Hugo Kubinyi, *Comparative Molecular Field Analysis (CoMFA)*,
- (62) Paolo Tosco, Thomas Balle, Fereshteh Shiri. Open3DALIGN: an open-source software aimed at unsupervised ligand alignment *J Comput Aided Mol Des.* **2011**, 25, 777–783
- (63) Michael P. Cava, Matthew I. Levinson Thionation reactions of lawesson's reagents. *Tetrahedron* **1985**, 22, 5061–5087
- (64) S. Scheibye, J. Kristensen, S.-O. Lawesson. Studies on organophosphorus compounds—XXVII: Synthesis of thiono-, thiolo- and dithiolactones. *Tetrahedron* **1979**, 35, 1339–1343

- (65) Turan Ozturk, Erdal Ertas and Olcay Mert Use of Lawesson's Reagent in Organic Syntheses Chem. Rev. **2007**, 107, 5210-5278
- (66) Varma RS, Kumar D. Microwave-accelerated solvent-free synthesis of thioketones, thiolactones, thioamides, thionoesters, and thioflavonoids. Org Lett. **1999**, 5, 697-700.
- (67) Richard Gedye, Frank Smith, Kenneth Westaway, Humera Ali, Lorraine Baldisera, Lena Laberge, John Rousell The use of microwave ovens for rapid organic synthesis. Tetrahedron Letters, **1986**, 27, 279-282.
- (68) C. Oliver Kappe, Alexander. Stadler, and Doris Dallinger Microwaves in Organic and Medicinal Chemistry 2012. Second, Completely Revised and Enlarged Edition. Wiley-VCH Verlag & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany.