



FACULTAD DE CIENCIAS
PROGRAMA DE LICENCIATURA EN CIENCIAS MENCIÓN
BIOLOGÍA O QUÍMICA

Diseño racional y síntesis mediante 3D-QSAR y Docking molecular de nuevos inhibidores
benzimidazólicos de *gap junction* de conexina 50.

Tesis para optar al grado de
LICENCIADO EN CIENCIAS MENCIÓN QUÍMICA

GIANFRANCO PAOLO SABADINI ALFARO

Director de tesis:

DR. JAIME MELLA RAIPAN

Comisión evaluadora:

DRA. MARISOL TEJOS R

DR. AGUSTÍN MARTÍNEZ C

Valparaíso, Chile.

2020

| | | |
|---------|---|----|
| 1 | ÍNDICE | |
| 2 | Nota aclaratoria contingencia pandemia COVID-19 | 4 |
| 3 | Resumen..... | 5 |
| 4 | Introducción | 6 |
| 4.1 | Conexinas 50: | 6 |
| 4.2 | QSAR:..... | 9 |
| 4.2.1 | DESCRIPTORES ¹⁴ :..... | 10 |
| 4.2.1.1 | Hidrofobicidad:..... | 10 |
| 4.2.1.2 | Parámetros Electroestáticos: | 10 |
| 4.2.1.3 | Componente estérico..... | 11 |
| 4.3 | QSAR 3D ¹⁴ | 12 |
| 4.3.1 | COMSIA: | 12 |
| 4.4 | Docking molecular..... | 13 |
| 4.4.1 | Sitemap..... | 15 |
| 4.4.2 | Preparación de proteína y ligando | 17 |
| 4.4.3 | Docking inducido..... | 17 |
| 4.4.3.1 | Muestreo del ligando | 18 |
| 4.4.3.2 | Muestreo de la proteína..... | 19 |
| 4.4.3.3 | Remuestreo del ligando | 19 |
| 4.4.4 | Puntuación | 19 |
| 4.5 | Síntesis de derivados de benzimidazol..... | 19 |
| 4.5.1 | Síntesis de 2-alkil-benzimidazol con ácidos carboxílicos sustituidos y 1,2-fenilendiamina..... | 20 |
| 4.5.2 | Síntesis de 2-alkil-benzimidazol mediante la formación de un haluro de acilo y posterior ciclación | 20 |
| 4.5.3 | Síntesis de 2-alkil-benzimidazol mediante la formación de un producto diacilado con DCC y posterior ciclación con ácido acético. | 21 |
| 4.5.4 | Síntesis de 1-alkil-benzimidazol con cloruros de alkilo y benzimidazol | 23 |
| 5 | Hipótesis:..... | 24 |
| 6 | Objetivos: | 24 |
| 7 | Materiales y métodos | 25 |
| 7.1 | Digitalización de las moléculas y minimización geométrica | 25 |

| | | |
|---------|---|----|
| 7.2 | 3D-QSAR COMSIA..... | 34 |
| 7.2.1 | Cálculos de CoMSIA..... | 34 |
| 7.3 | Validación estadística del modelo QSAR..... | 34 |
| 7.3.1 | Validación interna ⁵³ | 34 |
| 7.3.2 | Validación externa ⁵³ | 35 |
| 7.4 | Docking..... | 36 |
| 7.4.1 | Software..... | 36 |
| 7.4.2 | Sitemap..... | 36 |
| 7.4.3 | Preparación de la proteína..... | 37 |
| 7.4.4 | Docking inducido..... | 37 |
| 7.5 | Síntesis..... | 37 |
| 7.5.1 | Síntesis de 2-metil(difenil)benzimidazol..... | 37 |
| 7.5.1.1 | Síntesis mediante reacción de Phillips en medio acuoso..... | 37 |
| 7.5.1.2 | Síntesis de 2-alkil-benzimidazol mediante la formación de un haluro de acilo y posterior ciclación..... | 38 |
| 7.5.1.3 | Síntesis de 2-alkil-benzimidazol con DCC y posterior ciclación..... | 38 |
| 7.5.2 | Síntesis de 1-(di(3-fluorfenil))benzimidazol..... | 38 |
| 8 | Resultados y discusión..... | 39 |
| 8.1 | 3D QSAR-COMSIA:..... | 39 |
| 8.2 | Sitemap..... | 43 |
| 8.3 | Docking..... | 43 |
| 9 | Resultados esperados..... | 46 |
| 9.1 | Síntesis..... | 46 |
| 10 | Conclusión..... | 48 |
| 11 | Referencias:..... | 49 |

2 NOTA ACLARATORIA CONTINGENCIA PANDEMIA COVID-19

Tras el pedido de la rectoría de la Universidad De Valparaíso de prohibir el ingreso de personal no crítico a la Facultad de ciencias el día 16 de Marzo y luego la cuarentena en la región de Valparaíso el día 15 de Junio de 2020, yo no he podido acceder al laboratorio donde desarrollo mi investigación, lo que impidió terminar las síntesis de cual trata esta tesis, además de algunos resultados que solo se encuentran en los computadores del laboratorio. Por lo tanto, en esta tesis encontrarán secciones las cuales son resultados esperados o resultados aproximados, al no poder acceder a ellos.

3 RESUMEN

Las conexinas son proteínas de transmembrana, de las cuales seis de ellas pueden formar canales celulares en la membrana plasmática. Estos canales pueden unirse con otro canal de otra célula formando *gap junctions* (del inglés, “uniones estrechas”). Esta unión conecta ambos citoplasmas, permitiendo el intercambio de moléculas y iones. Existen diferentes conexinas, una de ellas es la conexina 50 (Cx50), la cual forma *gap junctions* en el desarrollo del cristalino ocular. Mutaciones en este gen generan cataratas congénitas, llevando a una inhibición de la conexina 50. El funcionamiento de la conexina 50 necesita ser mayormente estudiado, por ello en la presente tesis planteamos el diseño y síntesis de nuevos inhibidores de *gap junctions* de conexina 50 de estructura benzimidazol. Con este fin se realizarán estudios 3D-QSAR y Docking, y se llevará a cabo la síntesis de dos compuestos a partir de los estudios computacionales, utilizando el heterociclo benzimidazol como farmacóforo de partida.

Los Resultados de modelación computacional muestran prometedoras actividades biológicas para los compuestos diseñados. Por lo tanto, la síntesis y evaluación biológica de estos compuestos será llevada a cabo en una etapa ulterior.

4 INTRODUCCIÓN

4.1 CONEXINAS 50:

Las conexinas son proteínas de transmembrana, ordenadas en hexámeros formando hemicanales o conexones (Figura 1). Estos hemicanales pueden formar *gap junctions* (del inglés, “uniones estrechas”), las cuales son la unión de dos hemicanales de diferentes células. Los *gap junctions* poseen un poro, el cual permite el paso de pequeñas moléculas o iones como ATP, péptidos, calcio, etc.¹

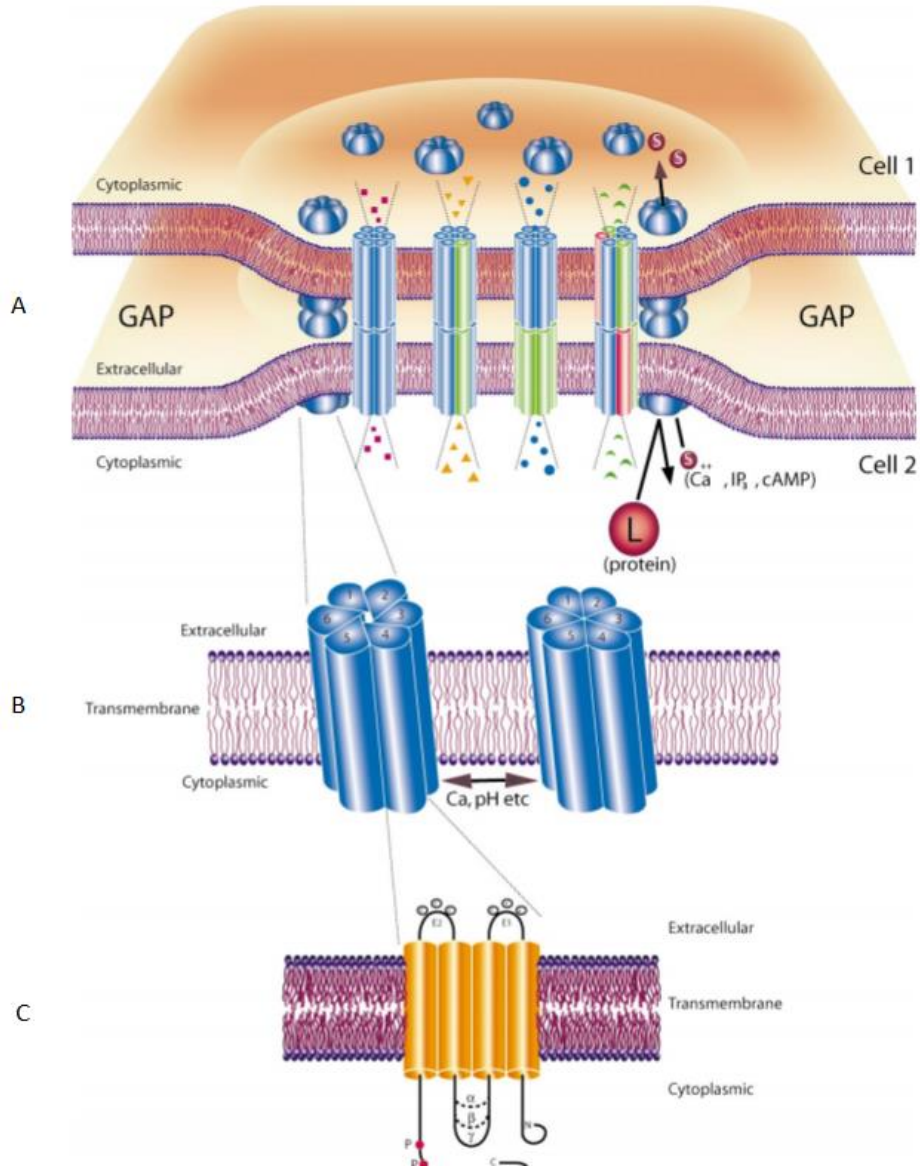


Figura 1 A. Esquema de un gap junction uniendo el citoplasma de ambas células. En el segundo y tercer esquemas B y C se muestran los hemicanales que conforman el gap junction².

Existen diferentes tipos de conexas, por lo cual se ha optado por una nomenclatura para diferenciarlas. Esta nomenclatura consiste en acompañar el nombre conexas (Cx) con su peso molecular en kilo-daltons, Ej. Conexas 26, 43, 50, etc.

Las conexas se expresan en diferentes tejidos humanos y están involucradas en diferentes patologías, como se muestra en la tabla 1:

Tabla 1. Conexinas con sus respectivos genes, lugares de expresión y patologías asociadas³.

| Conexina | Gen | Tipos de células u órgano en que se expresa | Patología |
|----------|-------------|---|---------------------------------------|
| Cx26 | <i>GJB2</i> | Microglía, hepatocito, célula de Kupffer, estómago, queratinocito, pulmón alveolar. | Problemas dérmicos. |
| Cx30.3 | <i>GJB4</i> | Queratinocito. | Síndrome de Clouston. |
| Cx31 | <i>GJB3</i> | Intestino delgado, colon, piel. | Sordera. |
| Cx32 | <i>GJB1</i> | Microglía, oligodendroglia. | Problemas dérmicos. |
| Cx40 | <i>GJA5</i> | Endotelio, corazón, células T, células b, Mioblastos. | Fibrilación arterial. |
| Cx43 | <i>GJA1</i> | Astroglia, endotelio, cardiacas, fibroblasto, célula T, célula B, monocito, macrófago, neutrófilo, célula dendrítica. | Displasia Oculodentodigital |
| Cx46 | <i>GJA3</i> | Epitelio pulmonar, lente ocular, osteoblasto. | Cataratas. |
| Cx47 | <i>GJC2</i> | Oligodendroglia. | Enfermedad tipo Pelizaeus-Merzbacher. |
| Cx50 | <i>GJA8</i> | Cristalino. | Cataratas. |

Como se puede apreciar diferentes tipos de conexinas pueden estar involucradas en diferentes enfermedades. Dentro de esta variedad se encuentra la Cx50, donde mutaciones en esta forma son causantes de cataratas congénitas⁴. Para este tipo de patología, la opacidad del cristalino se presenta una vez al nacer, la infancia o en el desarrollo fetal. Generalmente esta se presenta de forma bilateral. En fetos y embriones, a veces afectado el núcleo del lente y a veces presenta una difusión cortical⁵.

La Cx50 es responsable del correcto desarrollo del cristalino, sus gap junctions permiten la conexión intermolecular entre las células del cristalino. Las mutaciones disminuyen la probabilidad de encontrar los gap junction de conexina 50 funcionales o estas quedan atrapadas en compartimientos intercelulares⁴. El correcto funcionamiento de las gap junctions de Cx50, permite la correcta diferenciación del epitelio del cristalino, además actúa como una molécula adhesiva que es fundamental para mantener la integridad de la fibra de la lente⁶. Otros fenómenos observados a consecuencia de expresar Cx50 mutada (mutación L7Q) es el descenso de presión en ratas hipertensas, donde disminuyeron las proteína de colesterol de alta densidad, además de disminuir el transcriptoma de la Cx50 en corazón y hígado⁷.

La Cx50 no solo se puede encontrar en el cristalino, sino también está presente en: Túbulos seminíferos de rata⁸, Retina de rata⁹, riñón de rata¹⁰, células cromafín de rata¹¹, corazón de rata⁷, hígado de rata⁷ y aorta de rata⁷.

En estos casos, el desarrollo de un inhibidor selectivo de gap junction de Cx50, ayudará a una mejor comprensión del desarrollo y fisiología de cataratas. Además, debido a la presencia de la Cx50 en diferentes tejidos ayudará a una comprensión de la fisiología y fisiopatología de las Cx50.

Un estudio en base a ligandos puede ayudar a esclarecer un potencial sitio activo para la inhibición de gap junctions de Cx50 y las demás formas de conexinas. De esta forma se abre la posibilidad para el desarrollo de nuevas drogas más específicas para otras conexinas, las cuales están presentes en diferentes patologías producto de una abertura anómala de los hemicanales¹². Por lo que este estudio busca sentar las bases para el desarrollo racional de nuevas drogas para las diferentes conexinas de relevancia terapéutica.

A continuación, explicaremos las técnicas QSAR y Docking, empleadas en la presente tesis para proponer nuevos compuestos inhibidores de gap junctions.

4.2 QSAR:

Las relaciones cuantitativas de estructura- actividad, QSAR (de su nombre en inglés “quantitative structure activity relationships”), consisten en establecer una relación matemática entre la estructura de una serie de compuestos, con sus parámetros fisicoquímicos y la actividad biológica que poseen.

Dentro de las técnicas QSAR podemos encontrar los métodos 2D-QSAR, basados en descriptores calculados o experimentales y los métodos 3D-QSAR

como CoMSIA¹³ (“Comparative Molecular Similarity Indices Analysis”, esto es, “Análisis comparativo de índices de similitud molecular). CoMSIA relaciona la actividad y estructura de la molécula mediante una relación en 3D de diferentes campos. A diferencia de otros métodos CoMSIA utiliza funciones gaussianas. Dentro de los campos están: el electrostático, el estérico, el hidrofóbico, el dador y el receptor de puentes de hidrógeno.

Los principales tipos de descriptores empleados en los métodos QSAR son detallados a continuación.

4.2.1 DESCRIPTORES¹⁴:

Los descriptores corresponden a la variable independiente de los cálculos QSAR. Se derivan de los resultados de algunos hallazgos fisicoquímicos experimentales, o bien pueden ser calculados. Tienen una directa conexión con las propiedades fisicoquímicas de la molécula.

4.2.1.1 Hidrofobicidad:

La afinidad de un fármaco a un medio acuoso o lipídico es importante para determinar su actividad, transporte, absorción y excreción. Una forma de determinar la afinidad por un medio u otro es mediante el cálculo del coeficiente de partición (P):

$$P = \frac{[C]_{Octanol}}{[C]_{Acuoso}}$$

Donde, $[C]_{Octanol}$ es la concentración de la molécula en *n*-octanol y $[C]_{Acuoso}$ es la concentración de la molécula en la fase acuosa. Si $P > 1$ la molécula tiene un carácter lipofílico, si $P < 1$ tiene un carácter hidrofóbico.

También se puede contraer el coeficiente de partición en su forma logarítmica $\text{Log}(P)$. $\text{Log}(P)$ puede ser escrito como la contribución de cada sustituyente a la lipofilia total en una cadena o estructura base de una molécula:

$$\text{Log}P = \sum_i a_i f_i + \sum_j b_j f_j$$

Donde f_i es la constante del fragmento y f_j es el factor de corrección. La contribución es multiplicada el factor de ocurrencia del fragmento a_i . El factor de corrección f_j es multiplicado por el número de las instancias y corrección de la estructura b_j . La corrección del fragmento incluye flexibilidad, polaridad del fragmento o efectos estéricos.

La hidrofobicidad de un fragmento molecular puede ser cuantificada mediante la constante de hidrofobicidad de Hansch (π).

4.2.1.2 Parámetros Electroestáticos:

Definir la distribución eléctrica de una molécula es importante al momento de desarrollar un fármaco, debido a que las interacciones dipolo-dipolo o ion-dipolo, dependen netamente de esta propiedad. Cada átomo o grupos funcionales presentes puede perturbar esta

propiedad, aumentando o disminuyendo la densidad de carga eléctrica en la superficie de la molécula.

Para poder realizar un análisis cuantitativo del aporte electrostático de un sustituyente a una molécula, es posible recurrir al cálculo de la constante de disociación del ácido benzoico, como patrón para una estimación empírica de las propiedades electrónicas dadoras o aceptoras, de un determinado sustituyente:

$$K_a = \frac{[A^-][H^+]}{[HA]}$$

Donde H^+ es un protón y A^- la base conjugada de HA. K_a suele ser representado en su forma logarítmica $pK_a = -\text{Log}(K_a)$.

La pKa puede ser afectada si en alguna parte de la molécula se agrega algún sustituyente. Por ejemplo, El ácido benzoico tiene un constante de 4.20¹⁵, Si agregamos un bromo en posición orto su constante cambia a 2.85¹⁵, para la posición meta 3.81¹⁵ y para la posición para 3.97¹⁵.

Una forma de relacionar el cambio en la constante de acidez con un término relacionado con el componente electrostático del sustituyente es la ecuación propuesta por Hammett:

$$\text{Log} \frac{k}{k_0} = \rho\sigma$$

Donde ρ es una constante propia del ácido empleado ($\rho = 1$ para el ácido benzoico), mientras σ es una constante que depende de cada sustituyente. La constante de acidez del ácido benzoico es k_0 , por otra parte, k es la constante de acidez del compuesto derivado de ácido benzoico sustituido con el fragmento de interés.

Los sustituyentes que aporten carga al anillo tendrán un σ negativos, mientras que los sustituyentes que sustraen carga eléctrica al anillo aromático tendrán un σ positivo.

4.2.1.3 Componente estérico

El volumen de un sustituyente es importante para el correcto acoplamiento ligando-receptor. Sustituyentes muy voluminosos en el ligando, pueden provocar un impedimento estérico entre el ligando y el receptor.

Para cuantificar el factor estérico de un sustituyente, se aplica el mismo principio de la ecuación de Hammett. Para estimar la contribución estérica de un átomo o grupo, se compara la constante de hidrólisis de propanoato de metilo ($\text{CH}_3\text{CH}_2\text{COOMe}$) versus un éster con un sustituyente de interés, mediante la ecuación de Taft descrita a continuación:

$$E_s = \log \left(\frac{k_x}{k_{\text{CH}_3}} \right)_A$$

El núcleo de referencia del éster es: X-CH₂-COOCH₃, donde k_x es la constante de hidrólisis para el éster con un sustituyente "X"; y k_{CH_3} es la constante para el éster no sustituido (X=CH₃). "A" denota que es en hidrólisis ácida. Mientras más negativo es E_s , el efecto estérico es mucho mayor.

Charton¹⁶, por su parte introdujo un nuevo parámetro para describir el efecto estérico en función del radio de Van der Waals del sustituyente:

$$v_x = r_x - r_H$$

Donde r_x para sustituyentes de un solo átomo es el radio de Van Der Waals, mientras para sustituyentes con más de un átomo, se considera la distancia del eje central hasta el radio de Van Der Waals más cercano. Por ejemplo, para el sustituyente metilbromo, su radio de Van Der Waals se mide por el eje Central(Este eje tiene la misma dirección y posición, que el enlace entre el carbono del metilbromo y el átomo de la cadena principal adyacente) hasta la distancia mínima con la superficie de Van Der Waals ¹⁷ (Figura 2)

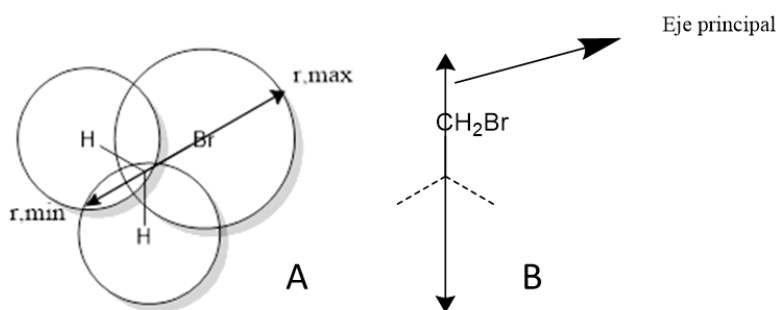


Figura 2 A. Donde r_{min} sería el radio de Van Der Waals más cercano y r_{max} el más lejano.
B. Visualización del eje y la cadena principales (línea entrecortada).

4.3 QSAR 3D¹⁴

En la metodología 3D-QSAR (QSAR tridimensional) se calculan los campos en tres dimensiones de un conjunto de moléculas orgánicas.

Las propiedades tridimensionales de la molécula son consideradas como un todo en lugar de tomar en cuenta los sustituyentes individuales. En esta técnica se asume que las características más importantes de una molécula son su tamaño, forma y sus propiedades electrónicas (campos electrostáticos).

4.3.1 COMSIA:

De sus siglas en inglés "Comparative Molecular Similarity Indices Analysis", es un método QSAR 3D, el cual se caracteriza por emplear descriptores que dependen de un campo gaussiano en tres dimensiones. Los campos pueden ser electrostáticos, estérico, hidrofóbico,

dador y aceptor puente de hidrogeno. Luego se ejecuta un análisis estadístico que correlacione dichos campos con la actividad biológica.

CoMSIA es similar a CoMFA (“Comparative Molecular Field Analysis”) en cuanto a que calcula todos los descriptores en un sistema de gradillas (generalmente espaciadas entre 1.1 y 2 Å). La función de campo utilizada por CoMSIA es calculada en cada punto q para la molécula j de acuerdo con la siguiente fórmula:

$$A_{F,k}^q(j) = \sum_{i=1} w_{probe,k} w_{ik} e^{-\alpha r_{iq}^2}$$

Donde A es el índice del descriptor, i el índice de suma de todos los átomos, w_{ik} es la propiedad fisicoquímica k sobre el átomo i , $w_{probe,k}$ es el valor del átomo sonda: +1 en carga, +1 en hidrofobicidad, radio 1 Å; α es el factor de atenuación, con un valor de 0.3; r_{iq}^2 es la distancia del átomo sonda en la gradilla q al átomo i ¹³.

CoMSIA tiene la ventaja de solucionar algunos problemas de alineamiento que presenta CoMFA. CoMFA emplea funciones de campo que tienden al infinito en las cercanías de la superficie molecular, como los potenciales de Lennard-Jonnes o Coulomb, esto implica que al momento de realizar los cálculos en cada gradilla, una leve desalineación puede afectar significativamente el análisis estadístico.

El análisis estadístico (descrito con detalle posteriormente), nos ayuda a determinar la capacidad predictiva del modelo planteado, si este modelo es validado, se puede empezar a predecir nuevas actividades para las diferentes moléculas con la ayuda de las ecuaciones encontradas.

El modelo arroja mapas en 3D donde se visualizan las regiones en que sería conveniente tener más ponderación de un descriptor y donde no, esto último es llamado holograma y nos ayuda a proponer nuevas moléculas activas.

Las técnicas QSAR se caracterizan por no requerir conocer la estructura del receptor, sino únicamente la de las moléculas afines al target y su actividad biológica o afinidad. Con el fin de complementar estos estudios, llevamos a cabo estudios analizando la conexina Cx50. Para ello llevamos a cabo estudios de acoplamiento molecular inducido o “Docking”, los que se explican a continuación.

4.4 DOCKING MOLECULAR

El Docking molecular o acoplamiento molecular es una técnica computacional usada para cumplir principalmente dos objetivos: Predecir la afinidad y la orientación del ligando en el receptor. La afinidad es calculada a través de un sistema de puntuación (Scoring), la

puntuación es calculada para cada orientación encontrada, de esta forma la pose con mejor puntuación es la mejor candidata para predecir las posibles interacciones Ligando-Receptor. Por lo dicho anteriormente, este procedimiento predictivo juega un papel importante en el desarrollo de nuevos fármacos. Estos acoplamientos se pueden clasificar en dos: Los rígidos y flexibles. Los rígidos son aquellos donde el receptor y el ligando conservan su forma, siguiendo el modelo llave-cerradura (Figura 3A). Los flexibles por su parte consisten en que el ligando y el receptor son flexibles y se basan en el principio que el ligando puede inducir una conformación distinta del sitio activo del receptor, y a su vez, el ligando puede rotar sus enlaces rotables para acomodarse mejor (Figura 3B).¹⁸

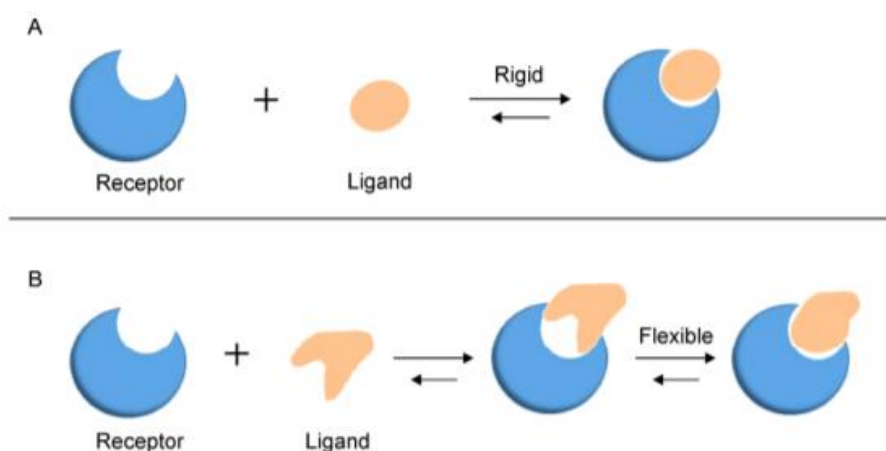


Figura 3 Esquema ilustrativo de los dos tipos de docking. A. Modelo llave-cerradura, o docking rígido y B. Modelo flexible o induced fit docking⁴¹⁸

El proceso de acoplamiento real es tan flexible que los receptores y ligandos cambian su conformación para encajar bien entre sí.¹⁸

Los algoritmos que debe seguir un software para poder realizar el proceso de docking se pueden clasificar en dos tipos : Los estocásticos y los sistemáticos¹⁹. Los algoritmos sistemáticos exploran todas las posibles conformaciones y todos los grados de libertad que puede llegar a tener un ligando. Ejemplo de estos algoritmos aplicados en docking son: Glide^{20 21}, FRED²² , FlexX²³ y DOCK²⁴. Por otra parte, los algoritmos estocásticos se basan en cambio aleatorios en los diferentes grados de libertad del ligando o receptor, este tipo de algoritmo se basan en el método de Monte Carlo, en algoritmos genéticos, en la búsqueda “Tabú” y en la optimización por enjambre de partículas. Ejemplo de estos algoritmos aplicados en docking son : DockVision²⁵, AutoDock²⁶ , PRO_LEADS²⁷ y SODOCK²⁸.

Por otra parte, cada algoritmo tiene que tener un sistema de puntuación que puede ser: Basado de campo de fuerza, empírico y “puntuación basada en el conocimiento”.

El sistema de puntuación por campos de fuerza se basa en la interacción intermolecular Ligando-Receptor, a través de fuerza de interacción de Van Der Walls, interacción Coulómbica, Puentes de hidrogeno, etc. Ejemplo el algoritmo de puntuación de DOCK²⁴ se basa solamente en la interacción de Van der Waalsy de Coulomb como se describe en la siguiente ecuación:

$$E = \sum_i \sum_j \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{4 r_{ij}} \right)$$

A_{ij} y B_{ij} son constantes propias de cada átomo. r_{ij} es la distancia que hay entre el átomo i al átomo j . q_i y q_j es la carga asignada para los átomos i y j . E es la puntuación recibida para esa pose. Los campos de fuerza más utilizados en docking son AMBER²⁹ y OPLS³⁰.

El sistema de puntuación empírico es similar al de campos de fuerzas, pero a diferencia de el sistema de campos de fuerza, cada interacción (fuerza de interacción de Van Der Walls, interacción Coulómbica, puentes de hidrogeno , etc) contribuye con un puntaje a la energía final (E), o sea cada interacción tiene un peso (W_i) sobre el cálculo final de la puntuación (E):

$$E = \sum_i W_i \Delta G_i$$

Por otra parte, la “puntuación basada en el conocimiento” se deriva directamente de valores experimentales previamente asignados en el software. La función que está detrás de este algoritmo es la ecuación de Boltzmann inversa^{31,32}.

Dado lo revisado anteriormente, existe múltiples tipos de algoritmos que van cambiando con el software que se realice el docking. Respecto el docking inducido o por su nombre en inglés *Induced Fit Docking* (IFD) , está basado en dos procedimientos de la compañía Schrodinger : Prime y Glide . El IFD será discutido más adelante.

En el caso de la conexina Cx50 no se sabe cuál o cuáles podrían ser los potenciales sitios relevantes dentro de la estructura donde realizar el docking de nuevos ligandos, por ello es necesario realizar un mapeo estructural de cavidades, lo que se describe a continuación.

4.4.1 Sitemap

En el campo de la química medicinal, hay veces en que el sitio activo se desconoce, esto limita mucho los estudios farmacológicos de la estructura a estudiar. Existen muchas estrategias para poder determinar estos sitios activos de forma computacional, las que se

pueden basar en la geometría, secuencia aminoacídica o interacción con una sonda exploratoria³³.

Los que se basan en el criterio de geometría buscan básicamente encontrar “bolsillos” o cavidades en el receptor. Generalmente los basados en la geometría usan un sistema de grilla y esferas.

Los basados en la interacción o energía, se basa en la interacción intermolecular entre ligando y receptor, de este modo determina la posible energía de interacción en cada posición del receptor. Ejemplo de este último caso es el algoritmo utilizado por Q-Sitefinder³⁴, el cual utiliza un grupo metil (CH₃) como una sonda para rastrear a través de todo el sistema las energías de interacción que puedan aportar a la unión ligando-receptor.

Sitemap es una herramienta adjunta al software Maestro de Schrödinger. El algoritmo que utiliza *Sitemap* tiene como objetivo determinar uno o más sitios de unión probables basándose en una mezcla de algoritmos geométricos y energéticos. Para cada sitio, hay un número de descriptores físicos que se utilizan para clasificarlo, estos incluyen: (i) el tamaño del sitio, (ii) el grado de “encerramiento” del posible sitio activo de la proteína, (iii) el grado de exposición al solvente, (iv) que tan juntos están los puntos con mejor puntuación en la proteína, (v) el carácter hidrofóbico del sitio, (vi) así como la posibilidad que un ligando tiene de donar o aceptar enlaces de hidrógeno³⁵.

El algoritmo utilizado por *Sitemap* consiste en tres pasos^{35,36}: (1) identificación de la cavidad, (2) caracterización de esta cavidad y (3) evaluación de la caracterización de la cavidad (“scoring”).

El primer paso consiste en fijar una gradilla de 1Å que incluya por completo al receptor, luego aquellos puntos que estén demasiado lejos del receptor son eliminados. Para poder seleccionar los puntos que estén dentro de una cavidad, se proyectan 110 vectores de 8Å perpendiculares desde cada punto de la gradilla distribuidos de forma esférica y homogénea. Aquellos puntos donde más del 50% de sus vectores penetre en la superficie del receptor (entiéndase superficie del receptor como la superficie de Van Der Waals) son seleccionados. Para eliminar aquellos puntos que estén muy cerca de los átomos de la superficie del receptor, se aplica el potencial de Lennard-Jones (Con un pozo potencial de 0.13 kcal/mol y una distancia a la cual el potencial es nulo de 1.5Å), aquellos átomos con un potencial de Lennard-Jones de -1.1 Kcal/mol o un potencial cercano a 0 son eliminados. La fase final del paso 1 es agrupar los puntos, aquellos puntos que este a una distancia menor a 1.76Å se agrupan, aquellos grupos que contengan menos de 3 puntos se descartan. De los posibles sitios activos que quedan, deben tener por lo menos de 3 puntos de gradillas de diferencia y una suma cuadrada de puntos de gradilla cuadrada mayor o igual a 5. Finalmente, si uno de los puntos está a 6.5 Å de otro punto que pertenezca a otro grupo, los dos grupos se unen, solo si esta unión no suma 100 puntos. La unión tiene que ser de tal forma que el recorrido sea el más cercano entre ambos grupos.

El segundo paso consiste en muestrear a través de campos de fuerza electrostáticos, estéricos, hidrofílicos e hidrofóbicos.

Finalmente, el último paso consiste en ordenar los sitios activos según el siguiente sistema de puntuación (*SiteScore*):

$$\text{SiteScore} = 0.0733\sqrt{n} + 0.668e - 0.2p$$

Donde n es el número de puntos que componen a ese sitio activo, e corresponde a la fracción de 110 vectores de 10Å que son proyectados esféricamente y uniformemente de cada punto de ese sitio activo, y p corresponde a la razón entre el puntaje hidrofílico e hidrofóbico, que se calcula a través de los campos de fuerza descritos en el trabajo de Halgern³⁶.

Este algoritmo tiene una 85.9% de exactitud para determinar el posible sitio activo y un 98.5% de exactitud para ligando con una afinidad (IC₅₀) menor a 1nm³⁶.

En la presente tesis ejecutamos la búsqueda Sitemap, gracias a lo cual identificamos un potencial sitio de unión para ligandos afines a Cx50.

4.4.2 Preparación de proteína y ligando

Para la correcta predicción de un docking es necesario hacer tratamientos especiales a la proteína como al ligando. Esta preparación preliminar se debe evitar omitir todos los tautómeros de ligandos para el análisis, así como todos los estereoisómeros, de haberlos³⁷. Dentro de la preparación de ligando encontramos Epik³⁸ un algoritmo que permite predecir el pKa de las moléculas, el cual se basa en las ecuaciones Taft y Hammett. Por otra parte la preparación de la proteína incluye la eliminación de las moléculas de agua que pueden haber incrustado en la estructura cristalográfica de la proteína, la optimización de las interacciones de puentes de hidrógenos intramoleculares y una minimización estructural global con la ayuda del campo de fuerza OPLS_2005³⁷.

4.4.3 Docking inducido

El Docking inducido o *induced fit docking (IFD)*, es un protocolo computacional adjunto al software Maestro de Schrödinger. Este protocolo consiste en generar un conjunto diverso de poses del ligando, el procedimiento se basa en las interacciones de Van Der Waals y las Coulombicas entre ligando y proteína.

El procedimiento general (descrito en la Figura 5) tiene cuatro pasos³⁹:

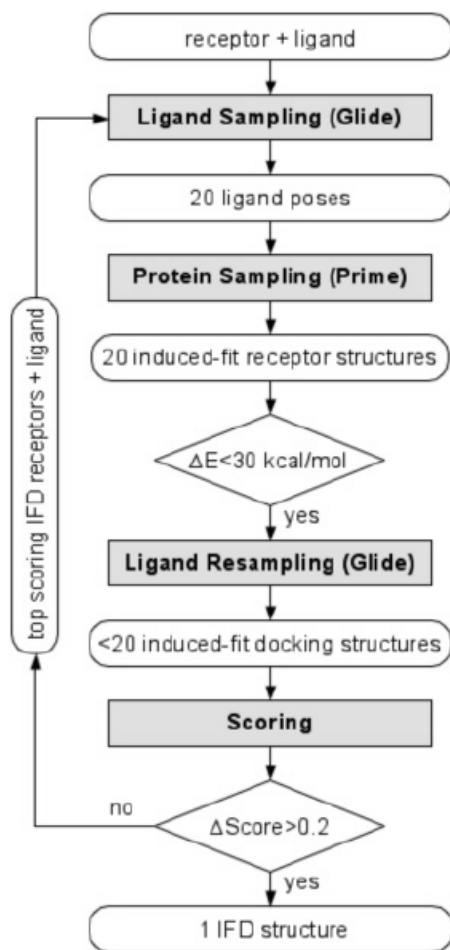


Figura 5 Diagrama de flujo del algoritmo de docking inducido³⁹

4.4.3.1 Muestreo del ligando

La primera etapa consiste en reducir el radio de van Der Waals de los átomos del ligando y de la proteína en un 50%, para ello los residuos de la proteína son cambiados por alaninas. Esto es con la finalidad que el sitio activo no sea bloqueado estéricamente. Las etapas secuenciales son las siguientes: (1) Acoplamiento con gradiente, para generar diferentes poses; (2) Se analizan las poses generadas en el paso (1); (3) Se realiza un docking rígido de

las diferentes poses y se selecciona las de menor energía; y (4) se efectúa el cálculo de la puntuación (GlideScore).

4.4.3.2 Muestreo de la proteína

El segundo paso del IFD consiste en utilizar el modelo de refinamiento y predicción de estructura Prime⁴⁰, que consiste en un minimizado seguido de un cálculo de la energía de solvatación del sistema mediante el Modelo de solvatación de Born generalizado de superficie aplicando a mecánica molecular con solvente implícito⁴¹. Al final de este paso las moléculas son ordenadas por su energía de solvatación.

4.4.3.3 Remuestreo del ligando

Los complejos que tienen una energética más baja que $30 \frac{kJ}{mol}$ son mantenidos para esta etapa. En esta etapa se realiza Glide Docking, el cual puede tener dos modos^{20,42,43}: modo SP (precisión estándar) el cual acopla de manera confiable decenas a cientos de miles de ligandos con alta precisión, al modo XP (Precisión Extra) donde la eliminación de falsos positivos se logra mediante un muestreo más extenso y una función de puntuación mucho más compleja, lo que resulta en una puntuación más precisa.

4.4.4 Puntuación

Finalmente, la puntuación final es calculada con la siguiente formula :

$$Puntos = GlideScore + 0.05 * PrimeScore$$

Como se puede apreciar, el puntaje final es una suma en el puntaje dado por Glide docking y Prime, nótese que el puntaje del Glide docking tiene un peso mucho mayor, esto debido a que la funciones de cálculo de Glide docking son más “suaves” y por tanto menos susceptibles al ruido estérico. Este 5% de contribución del Prime score a la puntuación final es suficiente para minimizar el ruido contribuido por Prime, de esta forma el puntaje que contribuye a la unión Ligando-Receptor pueda elevarse por encima de las fuentes de ruido.

El núcleo prometedor en base a los estudios de docking e Isosterismo molecular para nuestro estudio, es el benzimidazol. Por ello a continuación discutimos brevemente como efectuar su síntesis.

4.5 SÍNTESIS DE DERIVADOS DE BENZIMIDAZOL

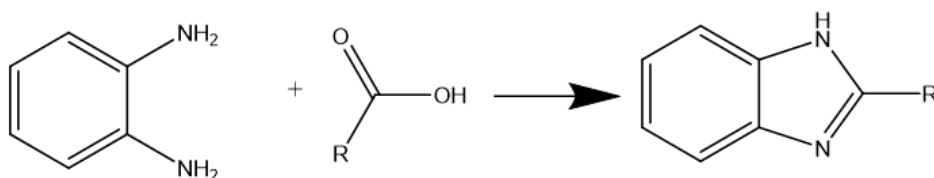
El benzimidazol o 1,3-benzodiazol y sus derivados son estructuras de gran importancia en el área de la química medicinal, pues presentan propiedades anticancerígenas, antiulcerosas, antiinflamatorias, antiprotozoarias, antimicrobianas, etc⁴⁴. Debido a estas propiedades tan diversas, es un sistema de gran interés para el diseño de nuevos compuestos bioactivos.

A continuación, explicamos la síntesis de moléculas de estructura benzimidazol que han sido propuestas en base a nuestros estudios computacionales como prometedores ligandos bloqueantes de conexinas. En términos generales nuestro objetivo es la obtención de benzimidazoles sustituidos en posición 1 ó 2.

4.5.1 Síntesis de 2-alkil-benzimidazol con ácidos carboxílicos sustituidos y 1,2-fenilendiamina

La 1,2-fenilendiamina reacciona con ácidos carboxílicos en medio ácido. Está reportada la ciclación con ácido fórmico y con ácido propanoico⁴⁵ con un rendimiento promedio de 68 a 85% (reacción 1).

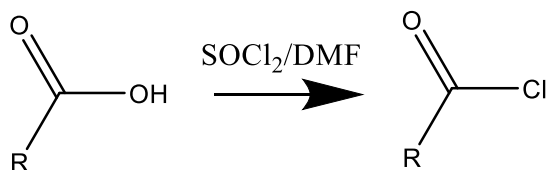
Este tipo reacción, conocida como “reacción de Philips” se puede llevar a cabo en un medio acuoso⁴⁶ o un medio orgánico⁴⁷ siempre y cuando ambos estén fuertemente acidificados.



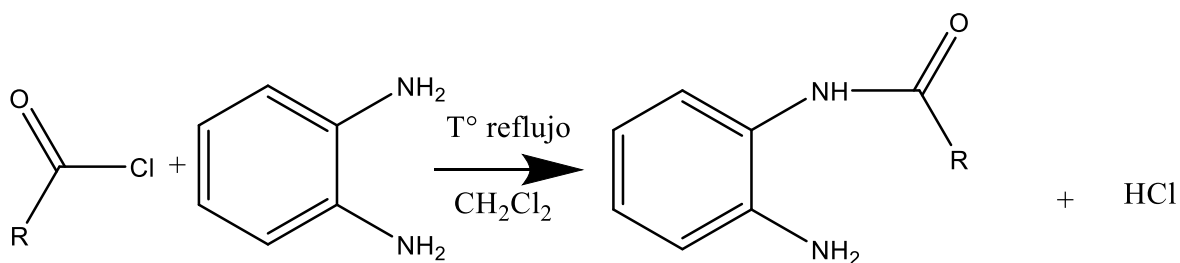
Reacción 1 Obtención de derivados de benzimidazoles a partir de ácido carboxílicos y 1,2-fenilendiamina.

4.5.2 Síntesis de 2-alkil-benzimidazol mediante la formación de un haluro de acilo y posterior ciclación

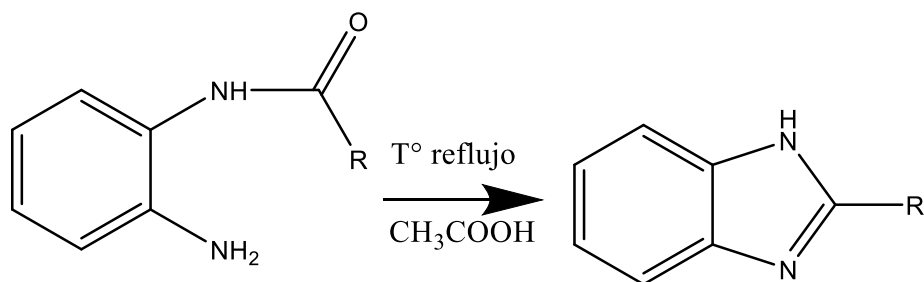
Esta técnica consiste en dos etapas. Primero se activa el grupo carbonilo haciéndolo más reactivo, mediante la formación de un cloruro de ácido por SOCl_2 y catalizado por dimetilformamida (DMF) (reacción 2). El cloruro de ácido es más propenso al posterior ataque nucleofílico por parte de 1,2-fenilendiamina. El producto del ataque de la amina es una *N*-(2-aminofenil)alquilamida intermedia (Reacción 3). Posteriormente, los productos acilados pueden ciclarse intramolecularmente formando el producto final 2-alkil-benzimidazol⁴⁸. Esta ciclación es realizada bajo reflujo en ácido acético glacial⁴⁷(reacción 4).



Reacción 2 Obtención de cloruro de acilo mediante SOCl_2 y DMF, a partir de una ácido carboxílico



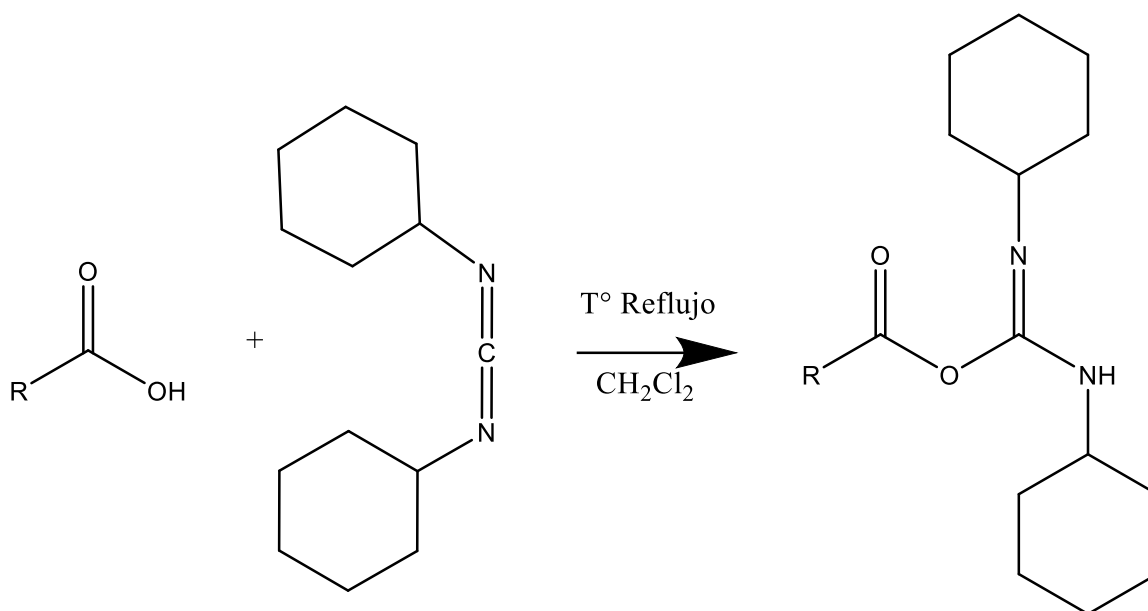
Reacción 3 Obtención de N-(2-aminofenil)alquilamida, a partir de un cloruro de ácido. Se puede emplear una base como trietilamina o similar en el medio de reacción para captar el HCl liberado.



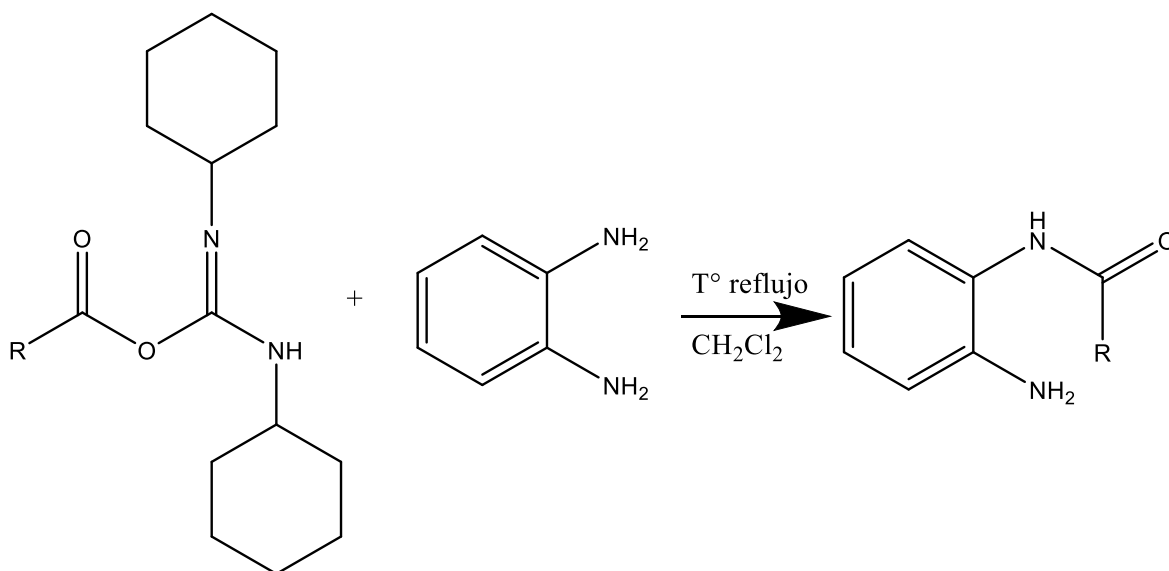
Reacción 4 Obtención del producto 2-alquilbenzimidazol, a partir de 2-acil fenilendiamina por ciclación en ácido acético glacial.

4.5.3 Síntesis de 2-alquil-benzimidazol mediante la formación de un producto diacilado con DCC y posterior ciclación con ácido acético.

En este método se hace reaccionar ácido difenilacético con diciclohexilcarbodiimida (DCC) en diclorometano (reacción 5) con el fin de activar el carbonilo. Posteriormente es añadida la orto-fenilendiamina para formar el producto acilado⁴⁹ (reacción 6). Finalmente, la ciclación es hecha en reflujo con ácido acético⁴⁷ (reacción 4).



Reacción 5 Activación de ácido carboxílico con DCC en diclorometano. El DCC es un electrófilo que permite la formación de un intermediario similar a un anhídrido.

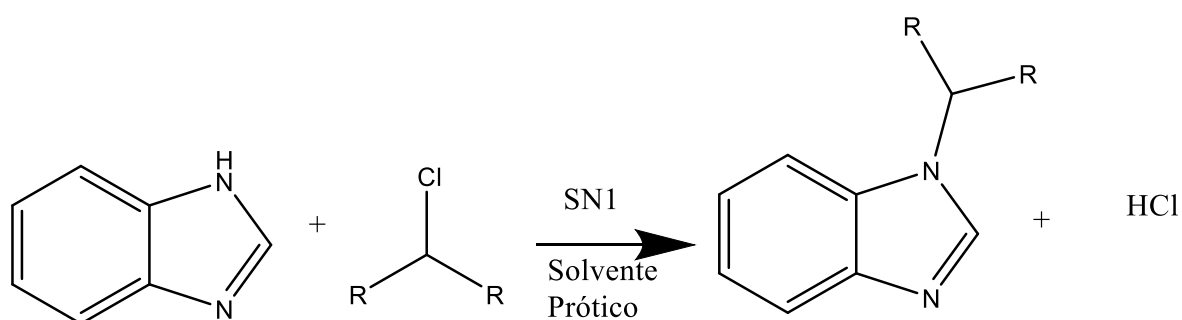


Reacción 6 Obtención del producto 2-acil-o-fenilendiamina, a partir de la reacción del derivado activado de ácido carboxílico y o-fenilendiamina. Este producto puede posteriormente ciclarse como se mostró en la reacción 4.

4.5.4 Síntesis de 1-alkil-benzimidazol con cloruros de alquilo y benzimidazol

Debido a que algunos de los compuestos propuestos por nuestros estudios de modelado computacional son benzimidazoles *N*-alquilados, a continuación, explicamos la ruta de obtención de estos compuestos.

El benzimidazol es un nucleófilo, sus átomos de nitrógeno electronegativos son tautoméricos y pueden generar un ataque nucleofílico tipo SN_1 , al carbono tetraédrico de un cloruro de alquilo, alquilándose en posición 1, en condiciones de un disolvente polar prótico⁵⁰(reacción 7).



Reacción 7 Obtención de derivados de benzimidazol a partir de un compuesto halogenado y benzimidazol. Se puede adicionar una base como trietilamina al medio de reacción para captar el HCl generado.

A continuación presentamos la hipótesis de trabajo y los resultados obtenidos así como las proyecciones futuras de nuestro estudio.

5 Hipótesis:

- Se postula que es posible diseñar, en base a estudios QSAR y de acoplamiento molecular inducido, y sintetizar inhibidores de conexinas de estructura benzimidazol.

6 OBJETIVOS:

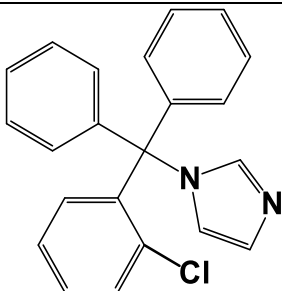
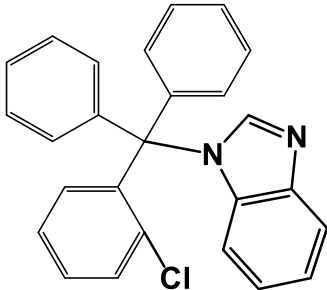
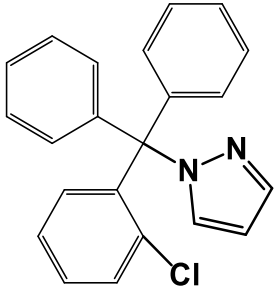
- Construir un modelo 3D-QSAR/COMSIA a base de inhibidores de *gap junction* de conexina 50.
- Predecir un sitio de unión en una *gap junction* de conexina 50.
- Generar una molécula con un elevado puntaje de unión.
- Sintetizar 2 compuestos derivados de benzimidazoles en base a las predicciones computacionales 3D-QSAR/COMSIA.

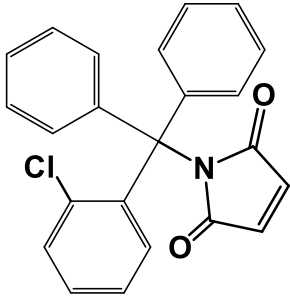
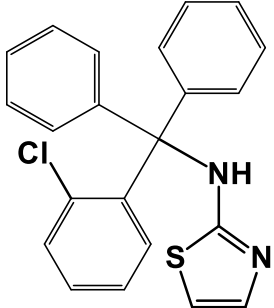
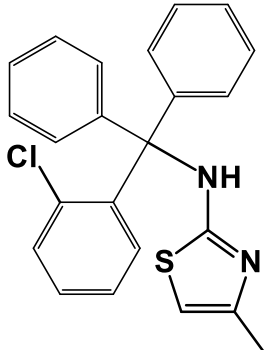
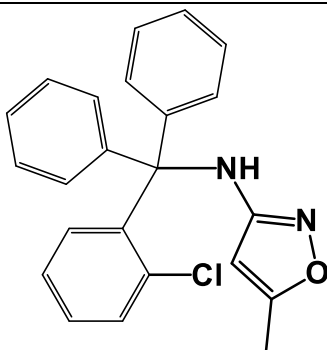
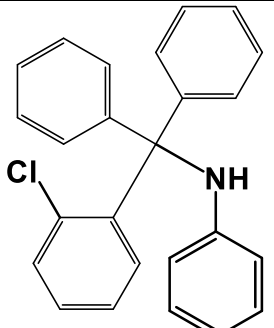
7.1 DIGITALIZACIÓN DE LAS MOLÉCULAS Y MINIMIZACIÓN GEOMÉTRICA

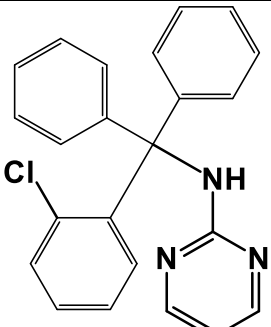
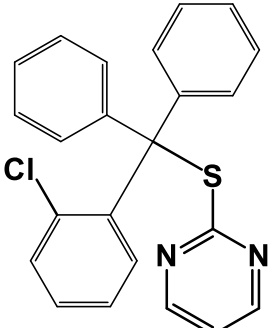
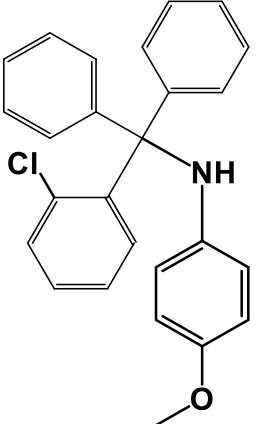
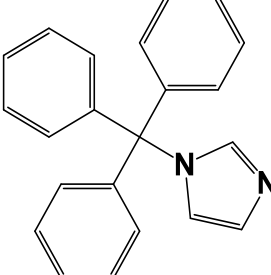
Con ayuda del software Chemdraw se digitalizaron las moléculas en 2D, siguiendo las estructuras descrita en el artículo titulado “*Triarylmethanes, a New Class of Cx50 Inhibitors*”⁵¹, posteriormente en el software Maestro de Schrödinger se las transformó a su conformación en 3D (Tabla 2).

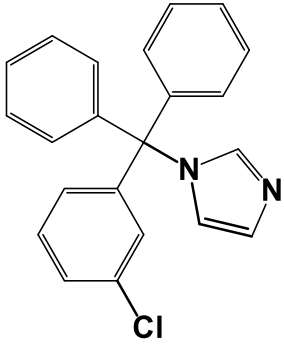
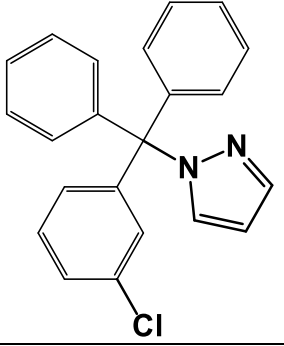
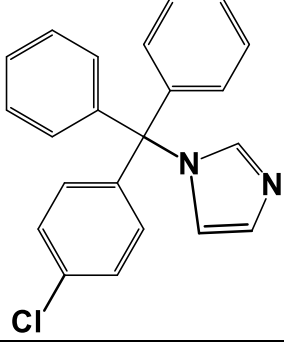
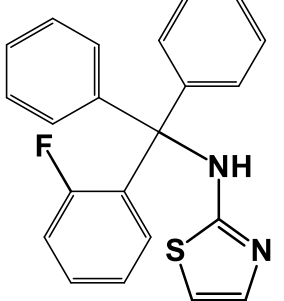
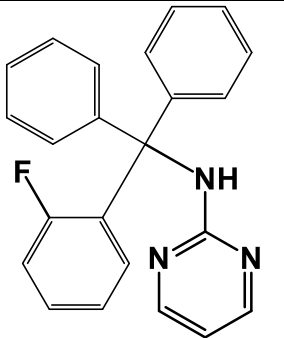
Luego de la digitalización se finalizó la fase preparatoria realizando un minimizado geométrico empleando el campo de fuerza OPLS₂₀₀₅.

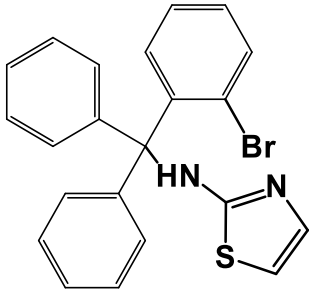
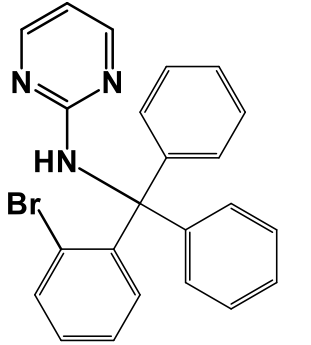
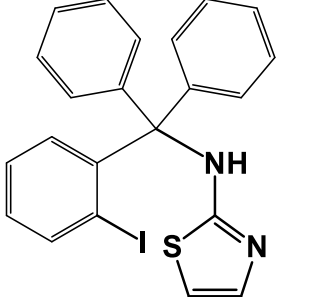
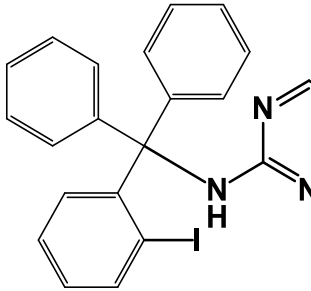
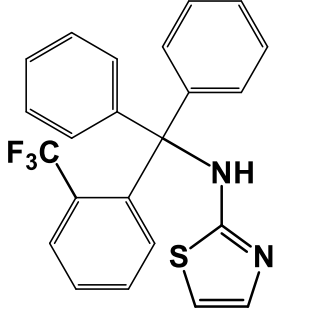
Tabla 2. Estructuras y actividades de las moléculas empleadas en la formulación del modelo QSAR.⁵¹

| Nombre del compuesto en el artículo ⁵¹ | Número para hechos de este estudio | Estructura | IC ₅₀ (μ M) | pIC ₅₀ |
|---|------------------------------------|--|-----------------------------|-------------------|
| Clotrimazol | 1 |  | 5 | 5.3010 |
| T85 | 2 |  | 4.2 | 5.3768 |
| T34 | 3 |  | 8.3 | 5.0809 |

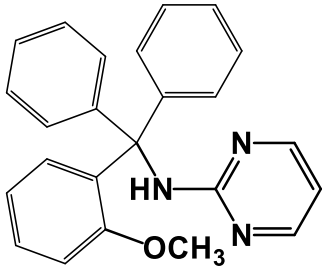
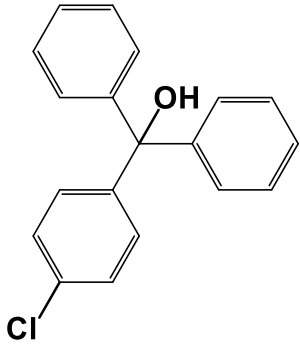
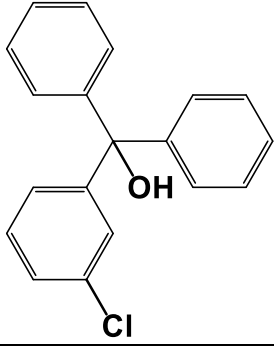
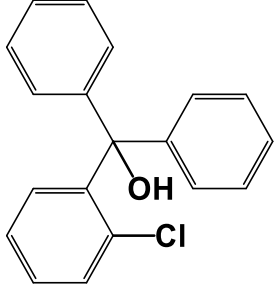
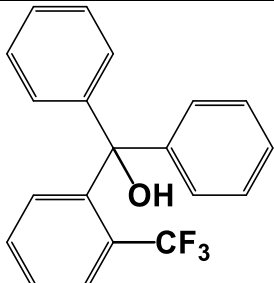
| | | | | |
|-----|---|--|-----|--------|
| T91 | 4 |  | 7.3 | 5.1367 |
| T66 | 5 |  | 3 | 5.5229 |
| T79 | 6 |  | 5.5 | 5.2596 |
| T80 | 7 |  | 11 | 4.9586 |
| T67 | 8 |  | 5 | 5.3010 |

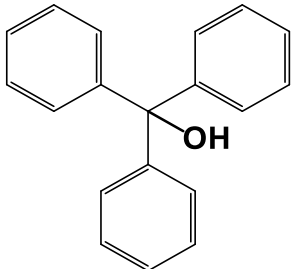
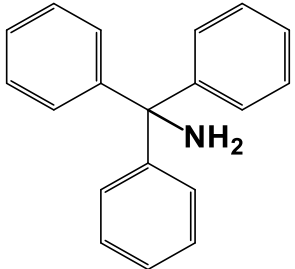
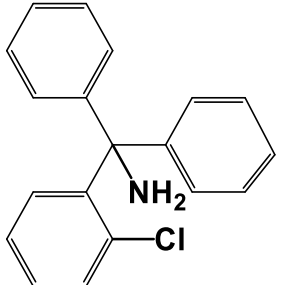
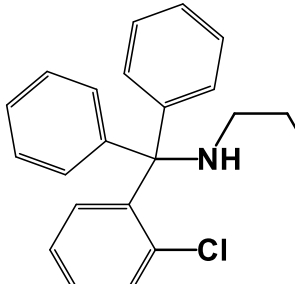
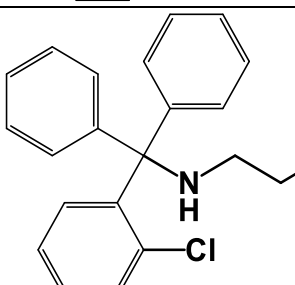
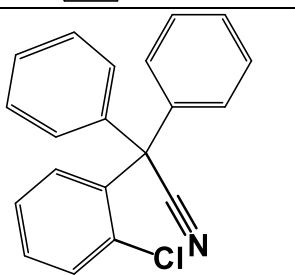
| | | | | |
|------|----|---|-----|--------|
| T68 | 9 |  <p>Chemical structure of 2-(2-chlorophenyl)phenylhydrazine: A benzene ring with a chlorine atom at the 2-position is connected at the 1-position to a phenyl ring, which is further connected to the 2-position of a pyridine ring.</p> | 8 | 5.0969 |
| T112 | 10 |  <p>Chemical structure of 2-(2-chlorophenyl)phenylsulfonamide: A benzene ring with a chlorine atom at the 2-position is connected at the 1-position to a phenyl ring, which is further connected to the 2-position of a pyridine ring.</p> | 8 | 5.0969 |
| T106 | 11 |  <p>Chemical structure of 2-(2-chlorophenyl)phenylmethoxyaniline: A benzene ring with a chlorine atom at the 2-position is connected at the 1-position to a phenyl ring, which is further connected to the 4-position of a benzene ring with a methoxy group at the 1-position.</p> | 10 | 5.0000 |
| T97 | 12 |  <p>Chemical structure of 2-(2-phenylphenyl)imidazole: A benzene ring is connected at the 1-position to a phenyl ring, which is further connected to the 2-position of an imidazole ring.</p> | 5.1 | 5.2924 |

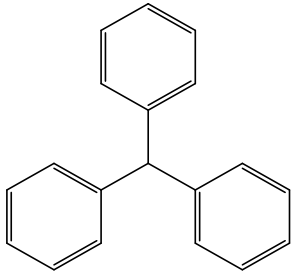
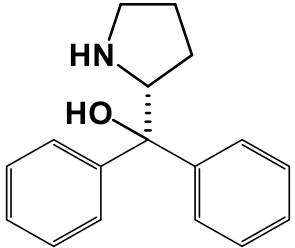
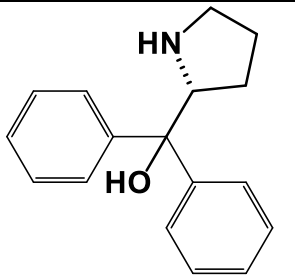
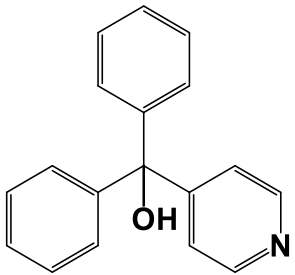
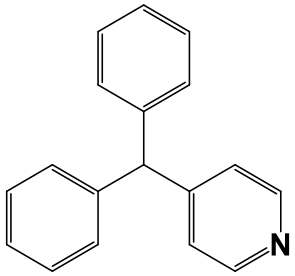
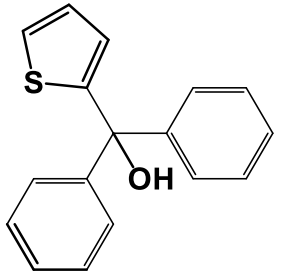
| | | | | |
|------|----|---|-----|--------|
| T143 | 13 |  | 4.3 | 5.3665 |
| T142 | 14 |  | 7.1 | 5.1487 |
| T144 | 15 |  | 6.1 | 5.2147 |
| T78 | 16 |  | 8.3 | 5.0809 |
| T128 | 17 |  | 9 | 5.0458 |

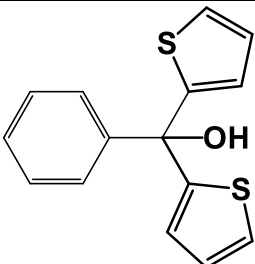
| | | | | |
|------|----|--|-----|--------|
| T131 | 18 |  | 3.2 | 5.4949 |
| T132 | 19 |  | 8 | 5.0969 |
| T136 | 20 |  | 2.4 | 5.6198 |
| T137 | 21 |  | 2.6 | 5.5850 |
| T72 | 22 |  | 5 | 5.3010 |

| | | | | |
|------|----|--|-----|--------|
| T73 | 23 | | 3 | 5.5229 |
| T126 | 24 | | 7.4 | 5.1308 |
| T127 | 25 | | 5.1 | 5.2924 |
| T124 | 26 | | 7.2 | 5.1427 |
| T122 | 27 | | 1.2 | 5.9208 |

| | | | | |
|------|----|---|-----|--------|
| T123 | 28 |  | 3 | 5.5229 |
| T1 | 29 |  | 2.8 | 5.5528 |
| T2 | 30 |  | 3.6 | 5.4437 |
| T3 | 31 |  | 2 | 5.6990 |
| T54 | 32 |  | 3 | 5.5229 |

| | | | | |
|-----------------|----|---|-----|--------|
| Trifenilmetanol | 33 |  | 1 | 6.0000 |
| T162 | 34 |  | 1.5 | 5.8239 |
| T41 | 35 |  | 4 | 5.3979 |
| T95 | 36 |  | 6 | 5.2218 |
| T94 | 37 |  | 7.3 | 5.1367 |
| T39 | 38 |  | 4 | 5.3979 |

| | | | | |
|----------------|----|---|-----|--------|
| Trifenilmetano | 39 |  | 0.9 | 6.0458 |
| T52 | 40 |  | 10 | 5.0000 |
| T53 | 41 |  | 10 | 5.0000 |
| T50 | 42 |  | 10 | 5.0000 |
| T160 | 43 |  | 4 | 5.3979 |
| T9 | 44 |  | 4 | 5.3979 |

| | | | | |
|-----|----|---|----|--------|
| T35 | 45 |  | 10 | 5.0000 |
|-----|----|---|----|--------|

7.2 3D-QSAR COMSIA

7.2.1 Cálculos de CoMSIA

Para la construcción del modelo CoMSIA se establece una cuadrícula de 3x3x3. La contribución electrostática, estérica, lipofílica, dadora y aceptora de puentes de hidrógeno es establecida por la ecuación gaussiana expuesta en la introducción, con carga electrostática +1, radio +1, lipofobicidad +1, y capacidad dadora y aceptora de puentes de hidrogeno +1.

La parte en común de cada molécula (farmacóforo) tiene que estar alineada para comparar los valores de los campos y construir la ecuación de regresión. Esto permite correlacionar adecuadamente los valores de IC₅₀ (variable dependiente) con los valores de los campos (variables independientes).

Estos campos se superponen con el de cada molécula, por ejemplo, el campo hidrofóbico de la molécula 1 con el hidrofóbico de la molécula 2, hasta la molécula N. La superposición tiene como fin en hacer un análisis parcial de mínimos cuadrados (PLS) con cada punto de la gradilla, para finalmente poder observar los hologramas de predicción.

Posteriormente, en base a los mapas de contornos obtenidos, se pueden proponer nuevas moléculas y predecir su actividad biológica mediante el análisis PLS previamente realizado⁵².

7.3 VALIDACIÓN ESTADÍSTICA DEL MODELO QSAR

La validación del modelo QSAR se basa en la validación cruzada, donde se dividen los datos en dos grupos: *Training set*, que se refiere al conjunto de datos que contribuyen a la generación del modelo, mientras que el *test set* contribuye a la evaluación de la predictibilidad del modelo. Para establecer el poder predictivo de este modelo, se realiza una validación interna y una validación externa.

7.3.1 Validación interna⁵³

La validación interna consiste en que la actividad de la molécula es predicha por el grupo *training set*. El valor del coeficiente de validación cruzada llamado Q² se utiliza para determinar la capacidad predictiva interna del modelo:

$$Q^2 = 1 - \frac{\sum(Y_{obs(train)} - Y_{pred(train)})^2}{\sum(Y_{obs(train)} - \bar{Y}_{train})^2}$$

Donde:

Y_{obs} e Y_{pred} son las actividades observadas y predichas, \bar{Y}_{train} por su parte es el promedio de las actividades del training set.

El mínimo aceptable de Q^2 es 0.5, donde los valores posibles son entre 1 y valores negativos.

Además, el error de desviación estándar de predicción (SDEP: "Standard deviation error of prediction") es calculada con la siguiente ecuación:

$$SDEP = \sqrt{\frac{\sum(Y_{obs(train)} - Y_{pred(train)})^2}{n}}$$

En la ecuación n corresponde al número de datos

Sin embargo, un valor de $Q^2 \geq 0.5$ es condición necesaria pero no suficiente para que un modelo QSAR sea considerado estadísticamente significativo y predictivo. Por ello se debe realizar la validación externa del modelo prediciéndole la actividad al test set, que son las moléculas que se dejaron fuera de la construcción de la ecuación QSAR, lo que se explica a continuación.

7.3.2 Validación externa⁵³

En la validación interna el conjunto de datos del *test set* no están involucrados en la validación del modelo QSAR, por lo que estos compuestos dejados afuera, se emplean en la validación externa del modelo. Los criterios más recomendados para la validación externa se explican a continuación.

El parámetro más utilizado para evaluar el rendimiento predictivo externo del modelo es el coeficiente de determinación para el test set:

$$R^2 = 1 - \frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2}$$

$Y_{obs(test)}$, $Y_{pred(test)}$ y \bar{Y}_{train} es la actividad observada, predicha y promedio del conjunto test y training, respectivamente.

Los valores de R^2 varían entre 0 y 1. Además, R^2 debe de ser mayor a 0.6. O sea, al menos el 60% de la actividad biológica del test set debe ser explicado por el modelo QSAR.

El valor R^2 no es suficiente para poder evaluar el modelo QSAR. Por ello Golbraikh and Tropsha⁵⁴ propusieron los siguientes criterios .

$$\frac{R^2 - R_0^2}{R^2} < 0.1 \text{ y } 0.9 \leq k \leq 1.1$$

$$\frac{R^2 - R_0'^2}{R^2} < 0.1 \text{ y } 0.9 \leq k' \leq 1.1$$

$$|R_0^2 - R_0'^2| < 0.3$$

Donde R^2_o es el coeficiente de determinación para el test set forzando a la recta a pasar por el punto (0,0). Mientras que R'^2_o es calculado invirtiendo ordenada y abscisa. Mientras que el error cuadrático medio de predicción (RMSEP: Root Mean Square Error of Prediction) se calcula por la siguiente ecuación:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{nEXT} (Y_{obs(test)} - Y_{pred(test)})^2}{nEXT}}$$

Donde $nEXT$ es el número de datos usados para la validación externa. Otros criterios matemáticos para la evaluación del poder predictivo de un modelo QSAR vienen dados por las siguientes ecuaciones:

$$\begin{aligned} \overline{r_m^2} &= \frac{(r_m^2 + r_m'^2)}{2} \\ \Delta r_m^2 &= |r_m^2 - r_m'^2| \\ r_m^2 &= r^2 \left(1 - \sqrt{r^2 - r_o^2} \right), r_m'^2 = r^2 \left(1 - \sqrt{r^2 - r_o'^2} \right) \\ r_o^2 &= 1 - \frac{\sum (Y_{obs} - k Y_{pred})^2}{\sum (Y_{obs} - \bar{Y}_{train})^2} \\ r_o'^2 &= 1 - \frac{\sum (Y_{pred} - k' Y_{obs})^2}{\sum (Y_{obs} - \bar{Y}_{pred})^2} \\ k &= \frac{\sum (Y_{obs} Y_{pred})}{\sum (Y_{obs})^2} \\ k' &= \frac{\sum (Y_{obs} Y_{pred})}{\sum (Y_{pred})^2} \end{aligned}$$

7.4 DOCKING

7.4.1 Software

El software utilizado es Glide, incorporado en la suite Maestro de Schrödinger, LLC, New York, NY, 2019.

7.4.2 Sitemap

Se fija un mínimo de 15 puntos para ser considerado un sitio activo. Se fija una hidrofobicidad estricta y una Gradilla de 1 Å.

7.4.3 Preparación de la proteína

Se eliminan las aguas residuales de la estructura y se minimiza utilizando el campo de fuerza OPLS_2005.

7.4.4 Docking inducido

Para el *Induced fit docking*, se plantea un sistema cúbico de 20 Å de lado centrado en el aminoácido Asp3 de la cadena peptídica. Los escalados del radio de Van der Waals de los átomos del receptor y ligando son fijados en 70% y 50%, respectivamente. Se fija un número máximo de poses de 10. Se retiene como distancia máxima Ligando-Receptor un valor de 5 Å. Se permite un optimizado de la cadena principal del receptor. Para el último docking GLIDE se fijan las 20 mejores estructuras y 30 kcal/mol de diferencia con respecto a la mejor estructura. Finalmente, se especifica el algoritmo de GLIDE en el modo XP (*Extra precision*).

7.5 SÍNTESIS.

La síntesis de los compuestos será realizada en el laboratorio de Química Medicinal de la Facultad de Ciencias de la Universidad de Valparaíso Chile.

Los reactivos con calidad para análisis serán adquiridos desde AK Scientific® y Matrix Scientific®.

El avance de la reacción será medido en cromatografía en capa fina en láminas de aluminio PF₂₅₄, siempre comparando su elución con los reactantes de partida.

La caracterización de los productos será analizada por RMN (¹H y ¹³C) en un equipo Bruker AM-400, los compuestos serán disueltos en dimetilsulfóxido deuterado. También se efectuará caracterización mediante espectrometría de masas, infrarrojo y medición del punto de fusión.

7.5.1 Síntesis de 2-metil(difenil)benzimidazol

7.5.1.1 Síntesis mediante reacción de Phillips en medio acuoso

Se agregan 150 mL de HCl 5.5M en un balón de 250mL, posteriormente se adicionan 10g de ácido difenilacético y 5.25g de 1,2-fenilendiamina. El balón es conectado a un sistema de reflujo, luego se agita y calienta hasta temperatura de reflujo por 72h. La disolución resultante se enfría y se neutraliza con NaOH. La solución final se deposita en un embudo de separación en conjunto con, luego se procede a agregar 50 mL de acetato de etilo y se extrae. Este último proceso se repite 3 veces, generando un volumen de 150mL, el cual es enjuagado con 10 mL de solución saturada de NaCl. La fase orgánica es evaporada al vacío hasta sequedad.

7.5.1.2 Síntesis de 2-alkil-benzimidazol mediante la formación de un haluro de acilo y posterior ciclación.

7.5.1.2.1 Generación de cloruro de acilo por cloruro de tionilo en exceso y sin tolueno
Se agrega 10g de ácido difenilacético, 20 mL de SOCl₂ y 0.5mL de dimetilformamida (DMF) como catalizador, en un matraz cerrado de 100mL. La mezcla es agitada y calentada a 50C° por 3h. Posteriormente la solución es llevada a rotavapor hasta sequedad.

7.5.1.2.2 Generación de cloruro de acilo por cloruro de tionilo con tolueno
Se agrega 10g de ácido difenilacético, 5.1 mL de SOCl₂ y 0.5mL de dimetilformamida (DMF), en un matraz de 100mL con 50 mL de tolueno como disolvente. La mezcla es agitada y calentada a 50C° por 5h, durante toda la reacción se mantuvo un tapón de vidrio. Posteriormente la solución es llevada a rotavapor hasta sequedad.

Los productos generados en 6.4.1.3.1 y 6.4.1.3.2, son juntados en un matraz de 250 mL. Posterior a esto se purifica el producto mediante una cromatografía en columna de 15cm de alto eluyendo con diclorometano. El solvente es descartado por rotavapor, el sólido es depositado en un matraz de 250mL y llevado al desecador por 24h.

7.5.1.3 Síntesis de 2-alkil-benzimidazol con DCC y posterior ciclación

Se agrega 1g de ácido difenilacético y 2.4g de DCC en 50ml de THF, se agita la solución por 30min. Luego se agrega 0.6g de *o*-difenilamina y se deja en agitación por 12h. Posteriormente el disolvente se evapora a presión reducida y se purifica el producto por cromatografía en columna.^{55,56}

7.5.2 Síntesis de 1-(di(3-fluorfenil))benzimidazol

Se agrega 1.25g de benzimidazol y 3.75g de Carbonato de potasio en 50ml de acetonitrilo. Luego se agrega 0.25g de Di(3-fluorfenil)metano y se agita por 4h a temperatura ambiente. Luego de esto se evapora el disolvente a presión reducida y se purifica el producto por cromatografía en columna.⁵⁷

8.1 3D QSAR-COMSA:

Ante de analizar los mapas de contornos generados por el modelo QSAR construido en base a lo explicado en la sección de métodos, calculamos los parámetros de validación externa para corroborar que la estadística obtenida era adecuada.

Los resultados de la validación del modelo QSAR se muestran en la siguiente Tabla 3:

Tabla 3: Validación estadística

| Condicion | Parametros | Valor limite | QSAR |
|-----------|----------------------|-------------------------|--------|
| 1 | q^2 | >0.5 | 0.8234 |
| 2 | r^2 | >0.6 | 0.8615 |
| 3a | r_0^2 | Close to value of r^2 | 0.7362 |
| 3b | $r_0'^2$ | Close to value of r^2 | 0.8393 |
| 4a | k | $0.85 < k < 1.15$ | 0.9973 |
| 4b | k' | $0.85 < k' < 1.15$ | 1.0023 |
| 5a | $(r^2 - r_0^2)/r^2$ | <0.1 | 0.1454 |
| 5b | $(r^2 - r_0'^2)/r^2$ | <0.1 | 0.0258 |
| 6 | $ r_0^2 - r_0'^2 $ | <0.3 | 0.1253 |

Como se aprecia en la tabla, el modelo cumple todas las condiciones. En el caso de la condición 5a no la cumple, pero la 5b sí. Una vez validado el estudio QSAR, se procede a proponer nuevas moléculas basados en los hologramas de campos (estérico, electrostático, hidrofóbico, dador y receptor de puente de hidrógeno) (Figura 6) los que se muestran a continuación.

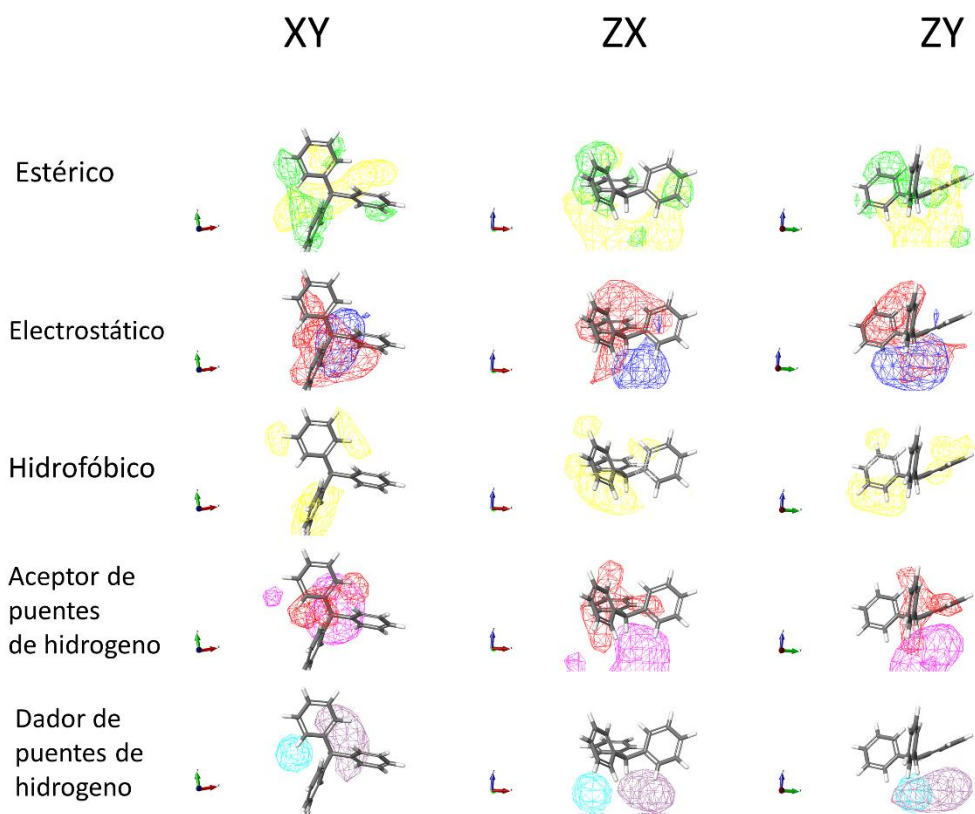


Figura 6 Hologramas obtenidos. Molécula de referencia Trifenilmetano.

Tabla 5. Significado de cada color de lo hologramas de la figura 6

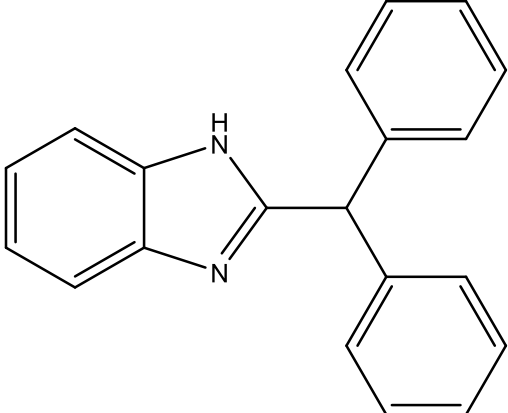
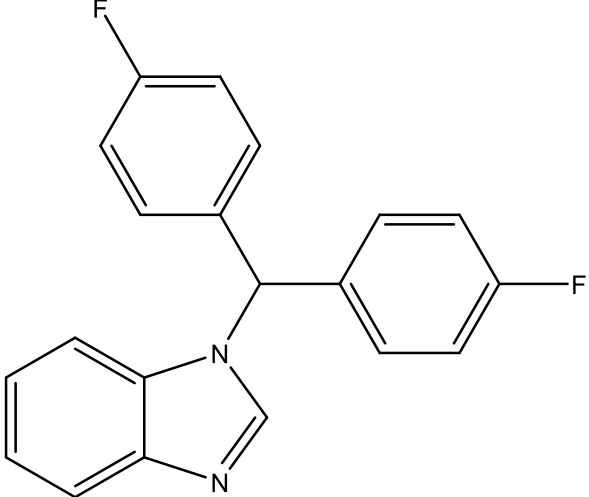
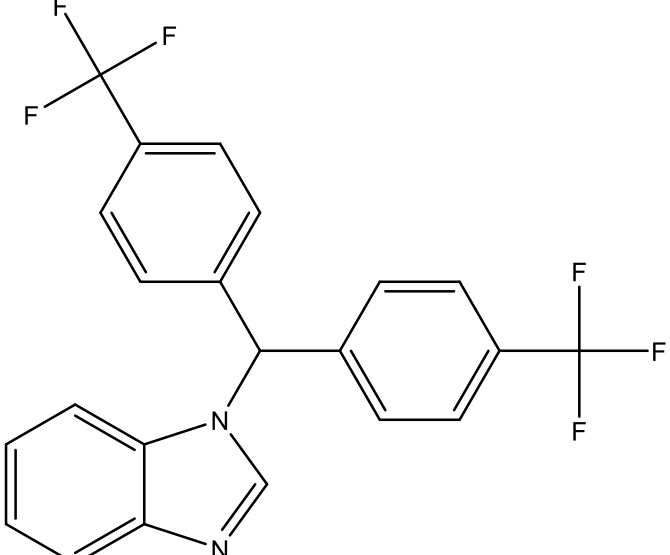
| | Favorable | Desfavorable |
|---------------------------------|----------------------------|----------------------------|
| Estérico | | |
| Electrostático | Densidad de carga negativa | Densidad de carga positiva |
| Hidrofóbico | | |
| Aceptor de puentes de hidrogeno | | |
| Dador de puentes de hidrogeno | | |

Como se aprecia hay claras zonas en torno a trifenilmetano en que se pueden hacer modificaciones como aumento o reducción de volumen, insertar o suprimir átomos dadores o aceptores de puente de hidrógeno, entre otros. Un heterociclo que se ajustó adecuadamente a estas condiciones fue el benzimidazol, motivo por el que lo proponemos en las nuevas estructuras diseñadas. A través de los hologramas expuestos en la figura 6 y su explicación declarada en la tabla 5, se proponen las siguientes moléculas reportadas en la Tabla 6.

Como se puede apreciar en la Tabla 6. Las moléculas propuestas son altamente lipofílicas, lo que tiene sentido dado que el ligando Trifenilmetano es la molécula más activa de la serie estudiada. Aunque su alto valor puede ser debido a una interacción indirecta con el *gap junction*, por ejemplo, alterando la membrana plasmática y de esta forma alterar la conformación de la proteína.

La molécula 3x de la tabla 6 es difícil de sintetizar, por lo que en este trabajo se priorizó la síntesis de las moléculas 1x y 2x de la Tabla 6. La molécula 3x supera en 0.1 unidades de pIC_{50} a la predichas por la publicación en que base el estudio QSAR 3D, mientras que las moléculas 1x y 2x, se encuentran dentro del rango, a pesar de esto, su estructura basal dista de los triarilmetanos, por lo que abre la posibilidad de nuevos inhibidores no descritos para conexina 50.

Tabla 6. Moléculas predichas con su respectivo pIC₅₀

| Numero de la molécula | Molécula | pIC ₅₀ predicho |
|-----------------------|--|----------------------------|
| 1x |  <p>Chemical structure of 1,2-diphenyl-1H-benzotriazole, consisting of a benzotriazole ring system substituted with two phenyl groups at the 1 and 2 positions.</p> | 5.5 |
| 2x |  <p>Chemical structure of 1-(4-fluorophenyl)-2-(3-fluorophenyl)-1H-benzotriazole, featuring a benzotriazole core with a 4-fluorophenyl group at position 1 and a 3-fluorophenyl group at position 2.</p> | 5.8 |
| 3x |  <p>Chemical structure of 1-(4-(trifluoromethyl)phenyl)-2-(3,3,4,4-tetrafluorophenyl)-1H-benzotriazole, featuring a benzotriazole core with a 4-(trifluoromethyl)phenyl group at position 1 and a 3,3,4,4-tetrafluorophenyl group at position 2.</p> | 6.1 |

8.2 SITEMAP

Con el fin de identificar un nuevo motivo estructural de inhibidor de conexina, llevamos a cabo un mapeo de la proteína y el diseño posterior de una molécula inhibidora afín a dicho sitio.

En este proceso se obtuvo 2 posibles sitios de unión. Uno está ubicado en la parte intracelular de la conexina, mientras que el segundo está ubicado en la zona extracelular (Figura 7). Por lo general, las conexinas tienen secuencias de aminoácidos bastante diferentes en su parte intracelular, esto podría ser favorable para este estudio, ya que la posible molécula obtenida por *docking* puede tener un carácter selectivo entre la familia de conexinas 50 (Comparado con la afinidad del ligando por otras de las estructuras del *Protein Data Bank*⁵⁸). En este sitio destaca el aminoácido ácido aspártico 3 (Asp3), ubicado en la cara intracelular del poro de un hemicanal de conexina (Figura 7), el que escogemos como centro de anclaje de los potenciales ligandos.

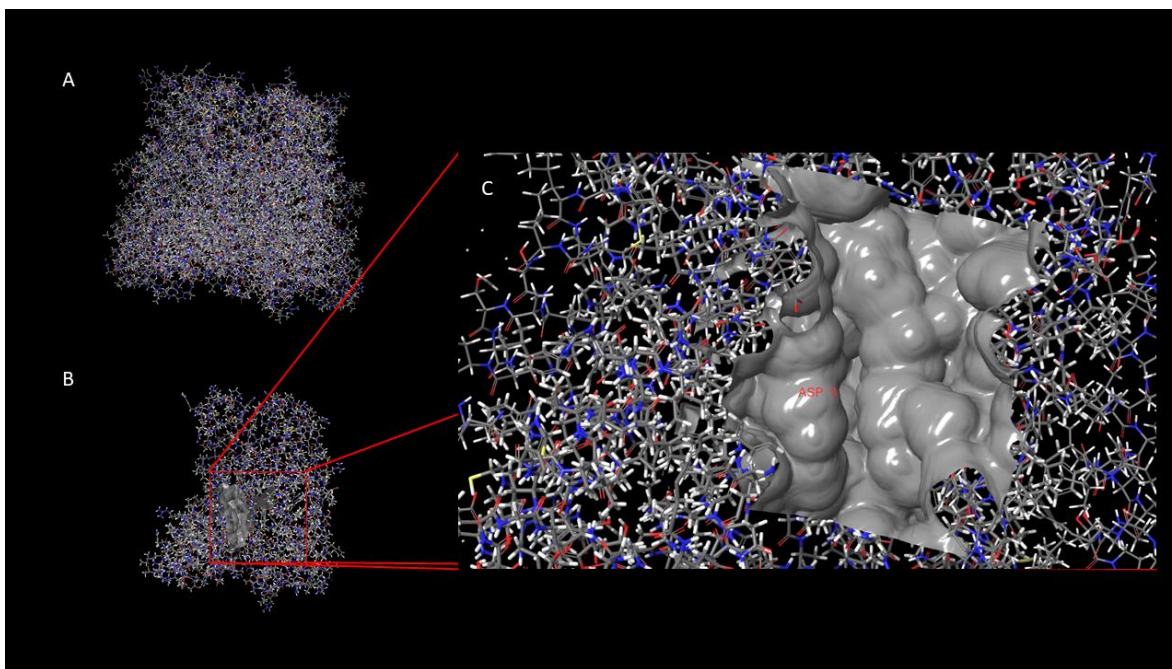


Figura 7. (A). Hemiacanal de conexina50, vista perpendicular al poro. (B). Parte interior expuesta del hemicanal y sitio activo. (C). Acercamiento de la imagen en B.

8.3 DOCKING

La ubicación dada por sitemap se centra en el aminoácido ácido aspártico 3 (Asp3). por lo que se procede a tomarlo como centro de la caja de análisis. En el centro acoplamos el anillo de benzimidazol como núcleo farmacofórico, y a partir del estudio del entorno aminocídico diseñamos las siguientes moléculas enlistadas en la Tabla 7. Todos los puntajes y poses de las moléculas mostradas en la Tabla 7, son para la mejor molécula según su función de ranqueo *IFD-SCORE*.

Tabla 7. Resumen de las moléculas diseñadas para el sitio descubierto en torno a Asp3 y sus interacciones con la conexina 50.

| Moléculas | Energía <i>Prime</i> | XP <i>Glidescore</i> | IFD- SCORE | Estructura diseñada |
|-----------|----------------------|----------------------|---------------|---------------------|
| 1 | -40090.39 | -14.566 | - 2019.09 | |
| 2 | -40098.75 | -10.810 | - 2015.75 | |
| 3 | -40103.62 | -9.810 | - 2014.99 | |
| 4 | -40886.32 | -10.429 | - 2054.74 | |

La interacción de cada una de las moléculas se visualiza en la Figura 8.

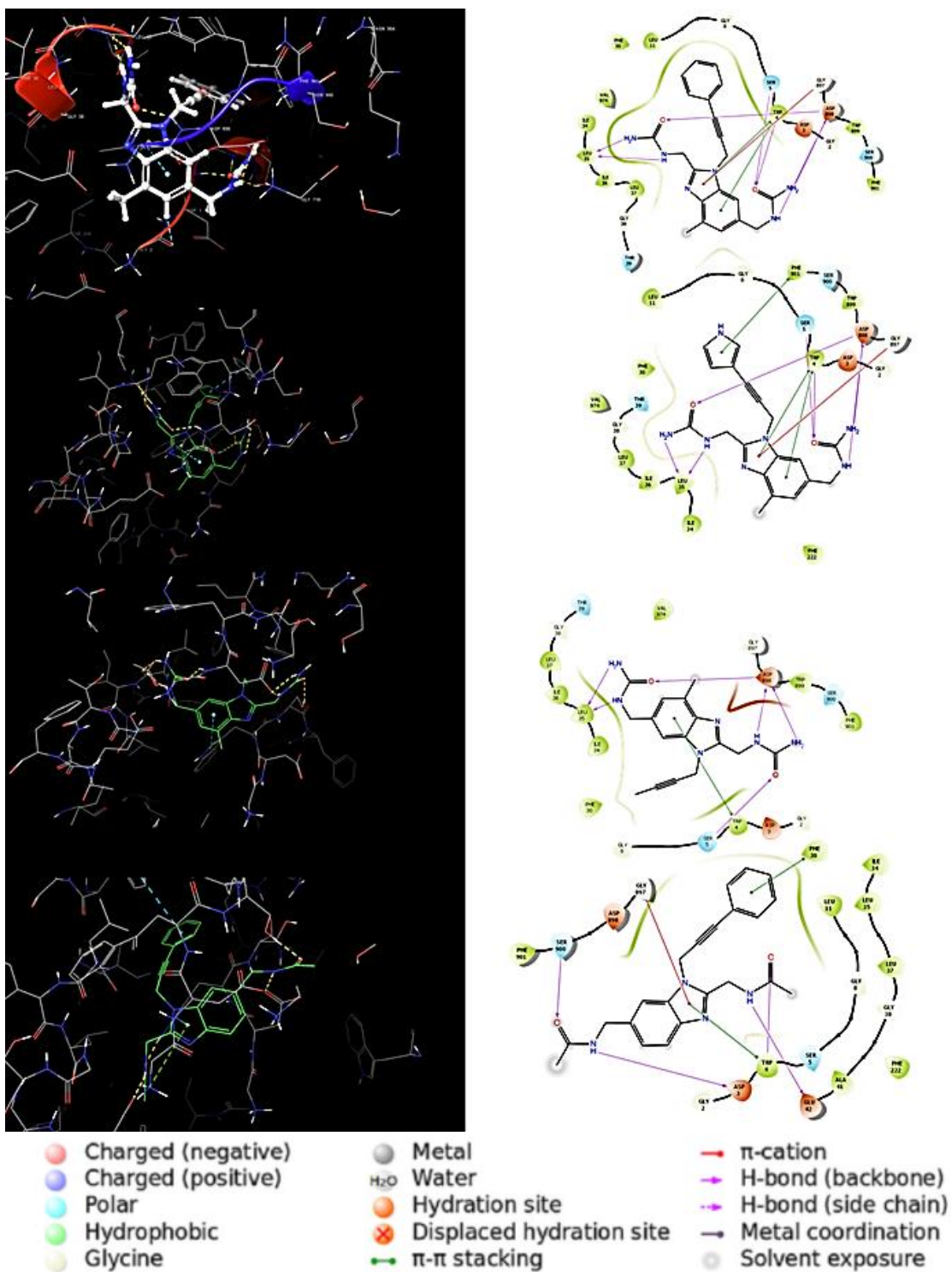


Figura 8 interacciones predichas para las moléculas diseñadas como ligandos para conexina 50. Primera columna: Imagen tridimensional de las interacciones en la cavidad. Segundo columna: representación 2D de las interacciones en 3D.

Los compuestos presentan múltiples interacciones de tipo puente de hidrógeno y π -stacking con la cavidad, además de tener pocos enlaces rotables, por lo que se constituyen como promisorios compuestos para ser sintetizados en una etapa futura.

La molécula 4 tiene la mejor puntuación. Pero la molécula 1 tiene muchas más interacciones y de hecho tiene una mejor puntuación XP GLIDE-SCORE, pero el valor más bajo de puntuación puesto que es contrarrestado con la energía de solvatación.

Estas moléculas, a diferencia de las moléculas obtenidas por QSAR, no son necesariamente inhibidores, sino que también podrían excitar la actividad de los canales de conexina, por lo que, a diferencia de las moléculas inhibitoras, estas podrían llegar a tener actividad terapéutica, ya que las cataratas congénitas se caracterizan por falta de actividad de la conexina 50 en la membrana plasmática. La posible actividad terapéutica podría aplicarse si estas moléculas llegan a salir al medio extracelular.

9 RESULTADOS ESPERADOS

9.1 SÍNTESIS

Para la síntesis de 2-metil(Difenil)benzimidazol, se emplea la reacción de Phillips, que consiste en un aumento de la temperatura en medio fuertemente ácido para que ambas aminas ataquen nucleofílicamente al ácido carboxílico (Mecanismo 1).

Se espera que para esta reacción no se obtengan productos en medio acuoso, esto se debería a la pobre solubilidad de los reactivos en agua. Es por ello que se opta a un medio orgánico como tolueno, donde la solubilidad de ambos componentes es del 100% en el volumen escogido⁵⁹. Por tanto, el problema de la solubilidad debería ser superado. Ahora existe otro problema: El impedimento estérico, o sea el gran volumen de los dos bencenos conectados al centro electrofílico al momento de hacer ciclar la molécula, dado que pueden estorbar físicamente en este proceso, por lo que es de esperar un rendimiento de reacción moderado a bajo.

Por otra parte, la reacción con cloruro de tionilo catalizado con DMF para generar un cloruro de acilo (Mecanismo 2), que es mucho más reactivo que el ácido carboxílico, podría emplearse para superar el impedimento estérico de la molécula. Por lo que el rendimiento sería mayor que en la reacción de Phillips en medio orgánico. En este proceso se puede formar una especie monoacilada ó diacilada (Mecanismo 3), como intermediarios. Ambas especies al ser tratadas con ácido acético darían como producto 2-metil(difenil)benzimidazol (Mecanismo 4), pero la especie diacilada a parte generaría como subproducto ácido difenilacético, el cual nuevamente reaccionaría hasta generar finalmente 2-metil(Difenil)benzimidazol.

Al igual en el caso anterior, la reacción con DCC hace más reactivo el ácido carboxílico para que acile a la ortofenilendiamina. En este caso el DCC es atacado nucleofílicamente por el oxígeno hidroxílico del grupo carboxílico. El DCC unido al ácido lo convierte en un buen grupo saliente además de acentuar el carácter electrofílico del carbono carbonílico, permitiendo el ataque nucleofílico de la amina de la o-fenildiamina (Mecanismo 5), esto permitiría sobreponerse al impedimento estérico de los reactivos al momento de ciclar la molécula (Mecanismo 4), mejorando el rendimiento de reacción.

La síntesis de 1-(Di(3-Fluorfenil))benzimidazol, consiste en un ataque tipo *SN1* donde el cloro sale espontáneamente por ruptura térmica, que es un proceso lento y por lo tanto, es el que limita la velocidad de la reacción (Mecanismo 6). Para potenciar el carácter *SN1* se eligió como disolvente el 2 propanol, el cual es un disolvente polar prótico, que al ser un alcohol secundario posee un carácter de nucleófilo débil, que no competiría como nucleófilo con el benzimidazol. El benzimidazol puede atacar nucleofílicamente por sí solo, pero para potenciar aún más su carácter nucleofílico se puede adicionar carbonato de potasio, lo que permite que el nitrógeno se desprotone, quedando así con carga negativa, permitiendo que el nitrógeno ataque de forma efectiva el centro carbocatión electrofílico.

Por otra parte, el trazado de la ruta sintética para los compuestos de estructura más compleja diseñados por docking, se llevará a cabo una vez que los derivados previamente descritos sean sintetizados y evaluados en la enzima.

10 CONCLUSIÓN

Se puede concluir que

1. Fue posible construir un modelo 3D QSAR/COMSIA, que permitió el diseño de moléculas con actividad inhibitoria de conexina 50 y estructura benzimidazol, lo que corrobora nuestra hipótesis.
2. Se logró determinar un potencial sitio de unión para ligandos mediante *Sitemap* en el sitio intracelular del hemicanal de conexina centrado en el aminoácido Asp3.
3. Con el método de acoplamiento molecular inducido se diseñaron 3 moléculas potencialmente afines a la conexina 50, de estructura benzimidazol con buenos valores de función de scoring.
4. Proponemos que la mejor forma de llevar a cabo la síntesis de los inhibidores diseñados por QSAR3D, son los siguientes métodos: Para la molécula 1 mediante el empleo de DCC o cloruro de tionilo catalizado con DMF; y para la molécula 2 una reacción vía mecanismo *SN1* emplenado K_2CO_3 como base y 2 propanol como disolvente.

Finalmente, cabe destacar que a la fecha no hay reportes de nuevas familias de compuestos inhibidores de conexina en base a estructura de benzimidazol, por lo que el presente trabajo de tesis contribuye con la propuesta de un nuevo farmacóforo y el diseño predictivo racional de compuestos accesibles sintéticamente que serán obtenidos en una etapa posterior.

11 REFERENCIAS:

1. Laird DW. Life cycle of connexins in health and disease. *Biochem J.* 2006;394(3):527-543. doi:10.1042/BJ20051922
2. El-Sabban ME, Abi-Mosleh LF, Talhouk RS. Developmental regulation of gap junctions and their role in mammary epithelial cell differentiation. *J Mammary Gland Biol Neoplasia.* 2003;8(4):463-473. doi:10.1023/B:JOMG.0000017432.04930.76
3. Yamasaki R. Connexins in health and disease. *Clin Exp Neuroimmunol.* 2018. doi:10.1111/cen3.12433
4. Beyer EC, Ebihara L, Berthoud VM. Connexin mutants and cataracts. *Front Pharmacol.* 2013;4 APR. doi:10.3389/fphar.2013.00043
5. Srinivas M, Verselis VK, White TW. Human diseases associated with connexin mutations. *Biochim Biophys Acta - Biomembr.* 2018. doi:10.1016/j.bbamem.2017.04.024
6. Hu Z, Shi W, Riquelme MA, et al. Connexin 50 functions as an adhesive molecule and promotes lens cell differentiation. *Sci Rep.* 2017;7(1). doi:10.1038/s41598-017-05647-9
7. Šeda O, Liška F, Pravenec M, et al. Connexin 50 Mutation Lowers Blood Pressure in Spontaneously Hypertensive Rat. *Physiol Res.* 2017;66:15-28. www.biomed.cas.cz/physiolres.
8. Pelletier RM, Akpovi CD, Chen L, Kumar NM, Vitale ML. Complementary expression and phosphorylation of Cx46 and Cx50 during development and following gene deletion in mouse and in normal and orchitic mink testes. *Am J Physiol - Regul Integr Comp Physiol.* 2015;309(3):R255-R276. doi:10.1152/ajpregu.00152.2015
9. Dorgau B, Greb H, Herrling R, et al. Expression and distribution of connexin50 in the mouse retina. *undefined.* 2013.
10. Silverstein DM, Thornhill BA, Leung JC, et al. Expression of connexins in the normal and obstructed developing kidney. *Pediatr Nephrol.* 2003;18(3):216-224. doi:10.1007/s00467-002-1065-1
11. Willenberg HS, Schott M, Saeger W, Tries A, Scherbaum WA, Bornstein SR. Expression of connexins in chromaffin cells of normal human adrenals and in benign and malignant pheochromocytomas. *Ann N Y Acad Sci.* 2006;1073:578-583. doi:10.1196/annals.1353.060
12. Willebrords J, Maes M, Crespo Yanguas S, Vinken M. Inhibitors of connexin and pannexin channels as potential therapeutics. *Pharmacol Ther.* 2017;180:144-160. doi:10.1016/j.pharmthera.2017.07.001

13. Klebe G, Abraham U, Mietzner T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules To Correlate and Predict Their Biological Activity. *J Med Chem.* 1994;37(24):4130-4146. doi:10.1021/jm00050a010
14. Roy K, Kar S, Das RN. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment.* Elsevier Inc.; 2015. doi:10.1016/C2014-0-00286-9
15. Rappoport Z, Frankel M. *Handbook of Tables for Organic Compound Identification.* Chemical Rubber Co; 1967.
16. Charton M. Steric Effects . I . Esterification and Acid-Catalyzed Hydrolysis of Esters Ef. 1975;1552-1556. doi:10.1021/ja00839a047
17. Charton M. Organic and Biological Chemistry: The Nature of the ortho Effect. II. Composition of the Taft Steric Parameters. *J Am Chem Soc.* 1969;91(3):615-618. doi:10.1021/ja01031a016
18. Fan J, Fu A, Zhang L. REVIEW Progress in molecular docking. doi:10.1007/s40484-019-0172-y
19. Huang SY, Zou X. Advances and challenges in Protein-ligand docking. *Int J Mol Sci.* 2010;11(8):3016-3034. doi:10.3390/ijms11083016
20. Friesner RA, Banks JL, Murphy RB, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem.* 2004;47(7):1739-1749. doi:10.1021/jm0306430
21. Friesner RA, Banks JL, Murphy RB, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem.* 2004;47(7):1739-1749. doi:10.1021/jm0306430
22. McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK. Gaussian docking functions. *Biopolymers.* 2003;68(1):76-90. doi:10.1002/bip.10207
23. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol.* 1996;261(3):470-489. doi:10.1006/jmbi.1996.0477
24. Ewing TJA, Kuntz ID. Critical evaluation of search algorithms for automated molecular docking and database screening. *J Comput Chem.* 1997;18(9):1175-1189. doi:10.1002/(SICI)1096-987X(19970715)18:9<1175::AID-JCC6>3.0.CO;2-O
25. Hart TN, Read RJ. A multiple-start Monte Carlo docking method. *Proteins Struct Funct Bioinforma.* 1992;13(3):206-222. doi:10.1002/prot.340130304
26. Morris GM, Goodsell DS, Halliday RS, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem.* 1998;19(14):1639-1662. doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B

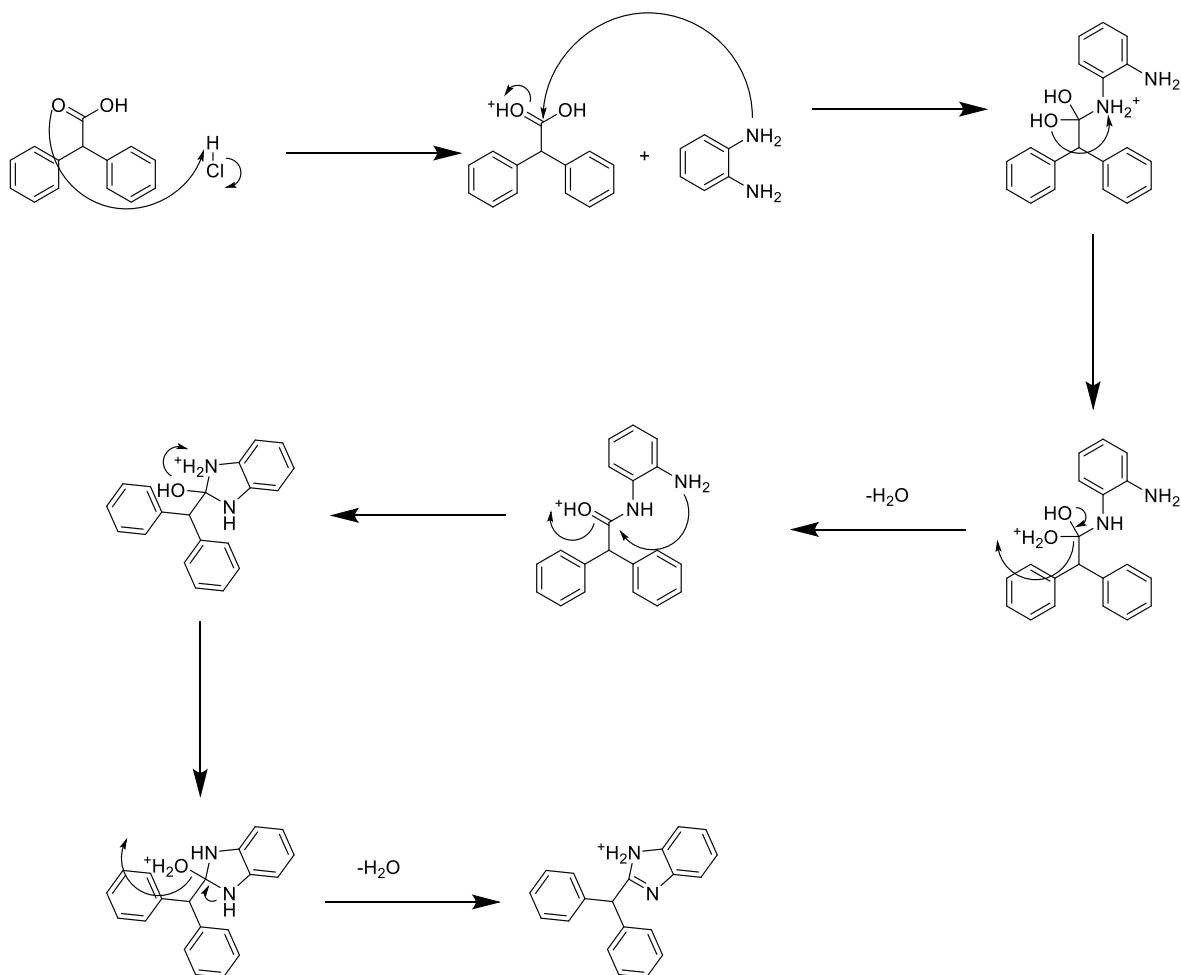
27. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD. Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins Struct Funct Genet.* 1998;33(3):367-382. doi:10.1002/(SICI)1097-0134(19981115)33:3<367::AID-PROT6>3.0.CO;2-W
28. Chen H-M, Liu B-F, Huang H-L, Hwang S-F, Ho S-Y. SODOCK: Swarm optimization for highly flexible protein–ligand docking. *J Comput Chem.* 2007;28(2):612-623. doi:10.1002/jcc.20542
29. Weiner PK, Kollman PA. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J Comput Chem.* 1981;2(3):287-303. doi:10.1002/jcc.540020311
30. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc.* 1996;118(45):11225-11236. doi:10.1021/ja9621760
31. Sippl MJ. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. *J Mol Biol.* 1990;213(4):859-883. doi:10.1016/S0022-2836(05)80269-4
32. Woetzel N, Karakaş M, Staritzbichler R, Müller R, Weiner BE, Meiler J. BCL::Score-Knowledge Based Energy Potentials for Ranking Protein Models Represented by Idealized Secondary Structure Elements. Deane CM, ed. *PLoS One.* 2012;7(11):e49242. doi:10.1371/journal.pone.0049242
33. Jiang M, Li Z, Bian Y, Wei Z. A novel protein descriptor for the prediction of drug binding sites. *BMC Bioinformatics.* 2019;20(1):478. doi:10.1186/s12859-019-3058-0
34. Laurie ATR, Jackson RM. Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics.* 2005;21(9):1908-1916. doi:10.1093/bioinformatics/bti315
35. Halgren T. New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des.* 2007;69(2):146-148. doi:10.1111/j.1747-0285.2007.00483.x
36. Halgren TA. Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model.* 2009;49(2):377-389. doi:10.1021/ci800324m
37. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des.* 2013;27(3):221-234. doi:10.1007/s10822-013-9644-8
38. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des.* 2007;21(12):681-691. doi:10.1007/s10822-007-9133-z

39. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J Med Chem*. 2006;49(2):534-553. doi:10.1021/jm050540c
40. Jacobson MP, Kaminski GA, Friesner RA, Rapp CS. Force field validation using protein side chain prediction. *J Phys Chem B*. 2002;106(44):11673-11680. doi:10.1021/jp021564n
41. Jacobson MP, Pincus DL, Rapp CS, et al. A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins Struct Funct Genet*. 2004;55(2):351-367. doi:10.1002/prot.10613
42. Halgren TA, Murphy RB, Friesner RA, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J Med Chem*. 2004;47(7):1750-1759. doi:10.1021/jm030644s
43. Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem*. 2006;49(21):6177-6196. doi:10.1021/jm051256o
44. Salahuddin, Shaharyar M, Mazumder A. Benzimidazoles: A biologically active compounds. *Arab J Chem*. 2017;10:S157-S173. doi:10.1016/j.arabjc.2012.07.017
45. wright JB. The chemistry of the benzimidazoles. *Chem Rev*. 1951;48(3):397-541. doi:10.1021/cr60151a002
46. Phillips MA. CCCXVII. - The formation of 2-substituted benzimidazoles. *J Chem Soc*. 1928:2393-2399. doi:10.1039/JR9280002393
47. Charton J, Girault-Mizzi S, Sergheraert C. Conversion of sterically hindered diacylated 1,2-phenylenediamines into 2-substituted benzimidazoles. *Chem Pharm Bull*. 2005;53(5):492-497. doi:10.1248/cpb.53.492
48. Preparation of Tetrahydroisoquinoline-3-ones via Cyclization of Phenyl Acetamides Using Eaton's Reagent. *Org Synth*. 2012;89:44. doi:10.15227/orgsyn.089.0044
49. Bremner JB, Samosorn S, Ambrus JI. N-acylation of 5-substituted indoles with carboxylic acids via DCC coupling. *Synthesis (Stuttg)*. 2004;2004(16):2653-2658. doi:10.1055/s-2004-831228
50. McMurry J. Organic Chemistry, 8th Edition. Belmont, CA Brooks Cole/Cengage Learn 2012. 2012:1376. doi:10.1016/B0-08-043751-6/08302-X
51. Bodendiek SB, Rubinos C, Trelles MP, et al. Triarylmethanes, a new class of Cx50 inhibitors. *Front Pharmacol*. 2012;3 JUN. doi:10.3389/fphar.2012.00106
52. Klebe G, Abraham U. Comparative Molecular Similarity Index Analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *J Comput Aided Mol Des*. 1999;13(1):1-10. doi:10.1023/A:1008047919606

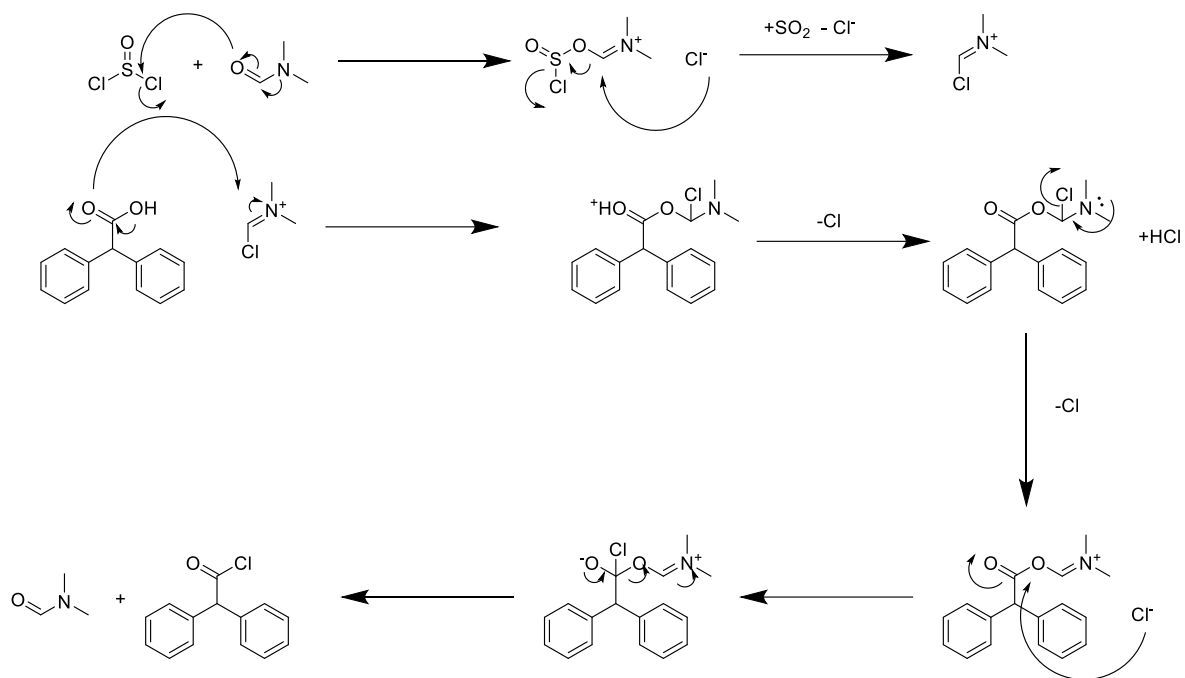
53. Dastmalchi S, Hamzeh-Mivehroud M, Sokouti B. *Quantitative Structure Activity Relationship : A Practical Approach*.
54. Golbraikh A, Tropsha A. Beware of q²! In: *Journal of Molecular Graphics and Modelling*. Vol 20. ; 2002:269-276. doi:10.1016/S1093-3263(01)00123-1
55. Gilles V, Vieira MA, Lacerda V, et al. A new, simple and efficient method of steglich esterification of juglone with long-chain fatty acids: Synthesis of a new class of non-polymeric wax deposition inhibitors for crude oil. *J Braz Chem Soc*. 2015;26(1):74-83. doi:10.5935/0103-5053.20140216
56. Bonsignore L, Cottiglia F, Maccioni AM, Secci D, Lavagna SM. Synthesis of coumarin-3-O-acylisoureas by dicyclohexylcarbodiimide. *J Heterocycl Chem*. 1995;32(2):573-577. doi:10.1002/jhet.5570320234
57. Barasa L, Yong A, Yoganathan S. An Efficient Chemo-Selective N-Alkylation Methodology for the Structure Diversification of Indolylbenzimidazoles. *ChemistrySelect*. 2020;5(11):3173-3178. doi:10.1002/slct.201904745
58. RCSB PDB: Homepage. <https://www.rcsb.org/>. Accessed September 2, 2020.
59. Bradley J-C, Friesen B, Mancinelli J, et al. Open Notebook Science Challenge: Solubilities of Organic Compounds in Organic Solvents. *Nat Preced*. February 2010. doi:10.1038/npre.2010.4243.1
60. Charville H, Jackson DA, Hodges G, Whiting A, Wilson MR. The Uncatalyzed Direct Amide Formation Reaction-Mechanism Studies and the Key Role of Carboxylic Acid H-Bonding. 2011. doi:10.1002/ejoc.201100714
61. Marson CM. *Synthesis Using Reagents.*; 1994.
62. Thompson QE. The Diacylation of Amides by Acyl Chloride-Pyridine Compounds. *J Am Chem Soc*. 1951;73(12):5841-5846. doi:10.1021/ja01156a115
63. Smith M, Moffatt JG, Khorana HG. Carbodiimides. VIII. Observations on the Reactions of Carbodiimides with Acids and Some New Applications in the Synthesis of Phosphoric Acid Esters. *J Am Chem Soc*. 1958;80(23):6204-6212. doi:10.1021/ja01556a015

Anexos.

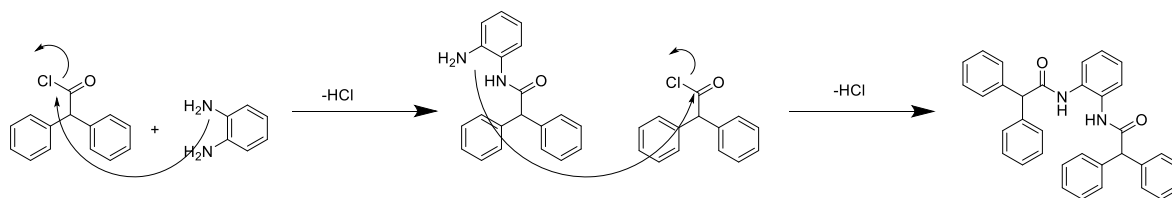
Mecanismos de las reacciones propuestas en la tesis.



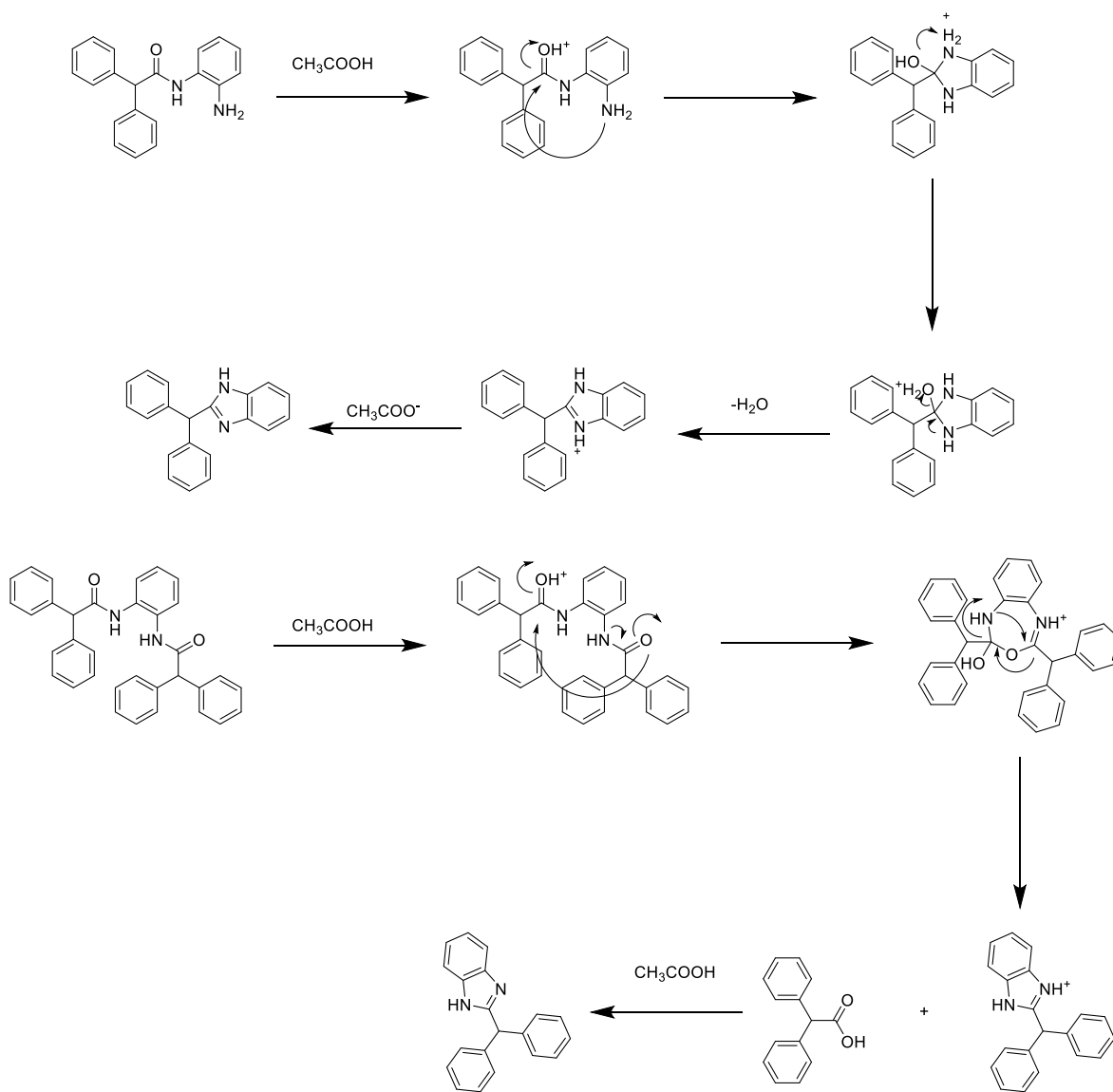
Mecanismo 1 Reacción de ácido difenilacético y o-fenilamina^{44,60}



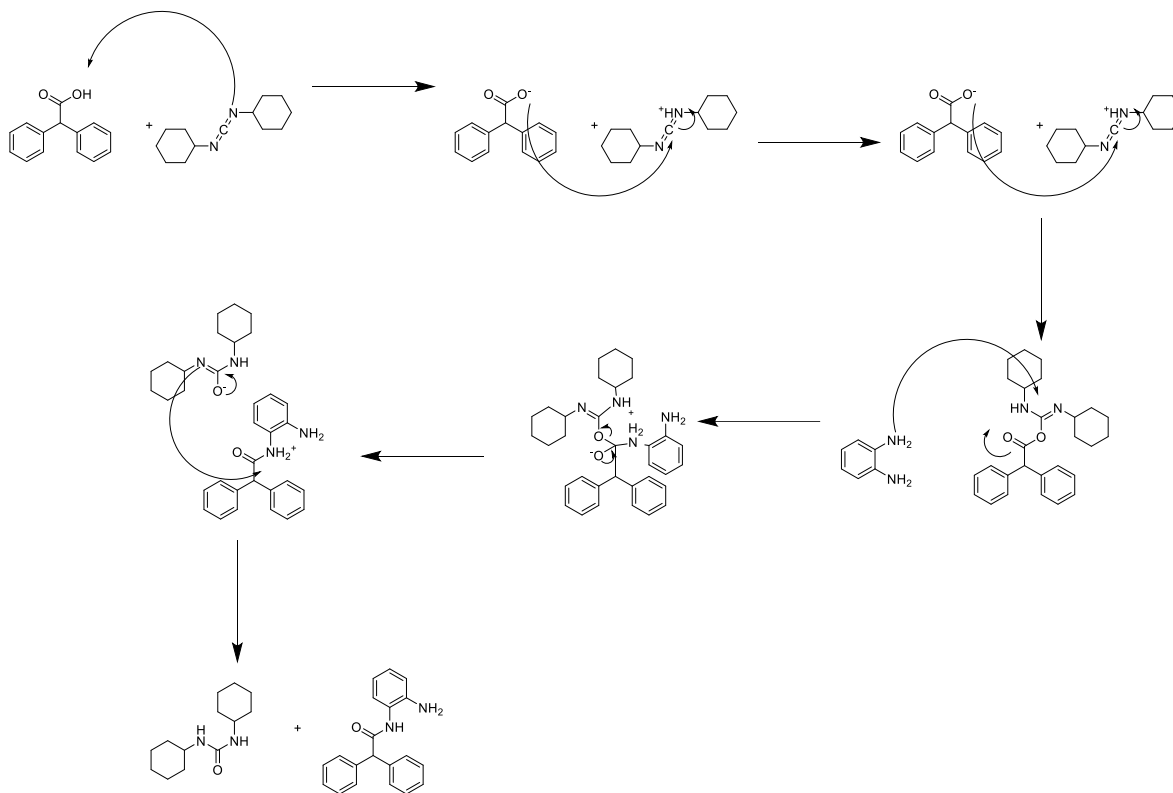
Mecanismo 2 Reacción de ácido difenilacético y cloruro de N-(clorometilamina)-N-metilmetanamina.⁶¹



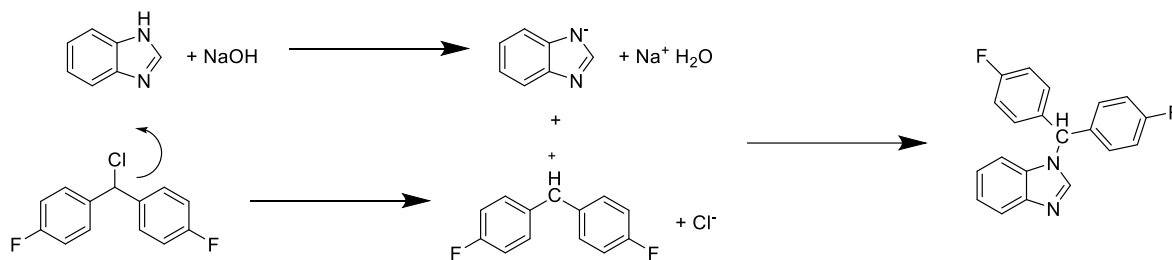
Mecanismo 3 Reacción cloruro de difenilacetoilo con o-difenilamina para dar un producto Mono o di acilado⁶²



Mecanismo 4 Ciclación con producto mono y di acilado con ácido acético⁴⁷



Mecanismo 5 Reacción de Ácido difenil acetico activado con DCC y o-fenilamina⁶³



Mecanismo 6 Reacción entre benzimidazol y Di(3-Fluorfenil)metano⁵⁰