UNIVERSIDAD NACIONAL DE LA PLATA

FACULTAD DE CIENCIAS EXACTAS

DEPARTAMENTO DE CIENCIAS BIOLÓGICAS

Tesis de Doctorado:

BÚSQUEDA DE NUEVOS ANTICONVULSIVOS A PARTIR DE MÉTODOS COMPUTACIONALES

Tesista: Pablo Hernán Palestro
Director: Luciana Gavernet
Co-director: Luis Enrique Bruno-Blanch

2016

A la memoria de mi padre
A mi familia
A Eve

Agradecimientos

En primer lugar quisiera agradecer a mi directora **Dra. Luciana Gavernet**, por la confianza que depositó en mi, por su apoyo y ayuda en este trabajo, pero mas que nada por ser una excelente persona.

A mi co-director **Dr. Luis Enrique Bruno-Blanch**, por su constante apoyo, tanto en el ámbito académico como personal.

A la **Facultad de Ciencias Exactas**, **UNLP** y **CONICET** por brindarme los medios para llevar a cabo esta tesis.

A los **Dres. Pedro Martin** y **Nicolás Enrique** y la **Dra. Verónica Milesi** por su colaboración en la realización de los ensayos in-vitro

A mis **compañeros del LIDeB**, por todos los momentos de diversión y esparcimiento dentro y fuera del laboratorio.

A mi **familia**, especialmente a mi **vieja** por todo su apoyo y esfuerzo para permitirme llegar hasta acá.

A mis **amigos** de la facultad y los de Bragado que, a pesar de la distancia y el tiempo, siempre están presentes.

A **Eve**, mi compañera de ruta.

Índice general

1. Epilepsia **5**

1.1. Generalidades. 5

1.2. Clasificación de la epilepsia. 6

1.3. Crisis epilépticas. 9

1.3.1. Crisis generalizadas. 10

1.3.1.1. Crisis tónico-clónicas. 10

1.3.1.2. Crisis de ausencia. 10

1.3.1.3. Crisis clónicas. 11

1.3.1.4. Crisis tónicas. 11

1.3.1.5. Crisis atónicas. 11

1.3.1.6. Crisis mioclónicas. 11

1.3.2. Crisis focales. 12

1.3.3. Crisis de origen desconocido. 12

1.4. Diagnóstico. 12

1.5. Tratamiento. 14

1.6. Fármacos antiepilépticos. 16

1.6.1. Fármacos de primera generación. 21

1.6.2. Fármacos de segunda generación. 23

1.6.3. Fármacos de tercera generación. 24

1.6.4. Mecanismos de acción de los fármacos antiepilépticos. 28

1.6.4.1. Modulación de canales iónicos dependientes de voltaje. . 29
1.6.4.2. Atenuación de la neurotransmisión excitatoria mediada por glutamato. 37
1.6.4.3. Activación de la neurotransmisión mediada por el acido γ-aminobutirico (GABA). 40
1.7. Epilepsia y farmacorresistencia. 42
1.7.1. Mecanismos de farmacorresistencia. 43
1.7.2. Hipótesis del transportador. 44
1.7.2.1. Glicoproteína P. 46

2. Métodos computacionales para el descubrimiento de fármacos 59
2.1. Introducción al descubrimiento de fármacos. 59
2.2. Métodos computacionales en el descubrimiento de fármacos. 62
2.2.1. Cribado virtual. 63
2.2.1.1. Cribado virtual basado en ligandos. 64
2.2.1.2. Cribado virtual basado en la estructura del receptor. . . 66
2.2.1.3. Validación del cribado virtual. 103

3. Modelo para la predicción de la afinidad de compuestos a P-gp 117
3.1. Modelo tridimensional de la P-gp humana. 117
3.2. Búsqueda bibliográfica y conjunto de evaluación. 123
3.3. Docking Molecular. 127

4. Modelo para la predicción de la afinidad de compuestos a Nav1.2 143
4.1. Modelado de la estructura del canal de sodio voltaje operado 1.2 143
4.1.1. Modelo de la estructura tridimensional de Nav1.2 en estado cerrado. 146
4.1.2. Modelo tridimensional del poro de Nav1.2 en estado abierto. . . . 152
4.2. Búsqueda bibliográfica y conjunto de evaluación. 156
4.3. Docking molecular. 157

5. Búsqueda, selección y evaluación de compuestos antiepilépticos. 165

5.1. Análisis de compuestos de síntesis. 166

5.2. Cribado virtual en bases de datos de fármacos. 167

5.3. Ensayos biológicos in-vitro. 173

5.3.1. Condiciones experimentales de los ensayos in-vitro. 175

5.3.2. Resultados de los ensayos in-vitro. 177

5.4. Ensayos biológicos in-vivo. 179

5.4.1. Condiciones experimentales de los ensayos in-vivo. 180

5.4.1.1. Resultados de los ensayos in-vivo. 183

6. Análisis de resultados y conclusiones. 189

Objetivos

El presente trabajo se encuentra enmarcado en la línea de investigación del Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB), dirigida hacia la búsqueda/diseño, síntesis y evaluación de compuestos con actividad antiepiléptica.

La epilepsia es uno de los desórdenes neurológicos crónicos de mayor prevalencia en el mundo. De acuerdo a la Organización Mundial de la Salud, afecta a alrededor de 50 millones de personas, presentando una mayor incidencia en países en vías de desarrollo. Hasta el momento la epilepsia no se cura, y el tratamiento apunta a controlar las manifestaciones sintomáticas del desorden epiléptico, que consisten comúnmente en convulsiones.

La farmacoterapia constituye la primera elección para el tratamiento de la epilepsia y resulta exitosa en el 70 % de los pacientes (pero no libre de los efectos secundarios presentes en este tipo de medicamentos). Por otra parte, el 30 % de los casos restantes no logra controlar los síntomas con fármacos y debe recurrirse a otras alternativas terapéuticas, como por ejemplo los tratamientos quirúrgicos, asumiendo el riesgo que implica una cirugía en el cerebro. Este tipo de epilepsia se denomina farmacorresistente o refractaria.

La problemática resumida anteriormente justifica la constante búsqueda de nuevos anticonvulsivos con mayor potencia, selectividad y menor toxicidad. Particularmente, la resistencia a los fármacos antiepilépticos actuales representa una de las mayores preocupaciones de los especialistas en el área; por lo que el descubrimiento de nuevas estrategias terapéuticas constituye uno de los principales objetivos a alcanzar en el campo de anticonvulsivos. En esta dirección, la identificación de los fenómenos involucrados en la

farmacorresistencia, así como un mayor entendimiento de los mecanismos de acción de los fármacos, resulta importante para el diseño racional de nuevos compuestos.

Se conoce que la resistencia a antiepilépticos es un fenómeno multifactorial y se han elaborado diferentes hipótesis que explican casos particulares de epilepsia farmacorresistente. Entre ellas se ha logrado identificar a la proteína de transporte glicoproteína P como responsable del eflujo de ciertos anticonvulsivos, impidiendo su llegada al sitio de acción. Adicionalmente, numerosos estudios han demostrado una importante asociación entre la epilepsia refractaria y la sobreexpresión de esta proteína en la barrera hematoencefálica, en tejidos epileptogénicos y en astrocitos. Por lo tanto, resulta interesante considerar a este "antitarget" de anticonvulsivos en las primeras etapas de diseño racional. Fármacos con alta afinidad a esta proteína serán excretados fácilmente, limitándose su concentración (y por lo tanto su acción) en el sistema nervioso central.

En relación a los mecanismos de acción de los fármacos antiepilépticos, en líneas generales, la acción de estas drogas genera un efecto conjunto que involucra más de un mecanismo, para finalmente modular la neurotransmisión excitatoria e inhibitoria. Entre los blancos moleculares más estudiados se encuentran los canales de sodio operados por voltaje, los cuales serán estudiados y utilizados en esta tesis.

Se plantea entonces como objetivo general de esta investigación, el descubrimiento de fármacos anticonvulsivos tendientes a satisfacer la demanda de los pacientes refractarios a la farmacoterapia actual. Para esto se hizo uso de estrategias de diseño racional, utilizando como herramienta el modelado molecular asistido por computadora. En particular se emplearon técnicas de modelado de macromoléculas y de cribado virtual secuencial mediante docking molecular.

Se llevaron a cabo los siguientes objetivos específicos:

1. *Desarrollo y validación de un modelo computacional de farmacorresistencia basado en las interacciones ligando–receptor frente a la glicoproteína P.* Como se enunció anteriormente, este antitarget es responsable de uno de los mecanismos de

farmacorresistencia asociados a los antiepilépticos. Al momento no se encuentra disponible la estructura de glicoproteína P humana, por lo que se construyeron modelos tridimensionales de esta macromolécula. Posteriormente se seleccionó el mejor programa de docking (y sus condiciones) capaz de discriminar sustratos y no sustratos conocidos de la glicoproteína.

2. *Desarrollo y validación de un modelo computacional capaz de identificar compuestos activos frente al canal de sodio dependiente de voltaje Nav1.2 humano mediante la metodología de docking molecular.* En forma similar a lo realizado para glicoproteína P, debieron construirse modelos tridimensionales del canal de sodio humano por no estar disponibles las estructuras experimentales. También se seleccionó el mejor programa de docking para discriminar entre bloqueantes y no bloqueantes del canal.

3. *Búsqueda y selección de posibles candidatos con actividad antiepiléptica en bases de datos de compuestos orgánicos y su posterior evaluación biológica.* Con la estructura de las macromoléculas construidas y determinadas las condiciones óptimas de la simulación por docking, se realizó un cribado secuencial sobre bases de datos. Inicialmente se buscaron compuestos que se predicen con alta afinidad al canal de sodio de acuerdo al docking molecular. Aquellas estructuras que superaron este filtro fueron sometidas a una segunda campaña de cribado para descartar compuestos que se predicen sustratos de la glicoproteína P. Finalmente los candidatos fueron evaluados mediante ensayos in-vitro y métodos in-vivo, orientados a determinar la capacidad de los modelos de seleccionar nuevos compuestos activos.

La presente tesis contiene seis capítulos. En el primero se tratan aspectos generales de la epilepsia. El capítulo siguiente aborda el desarrollo de fármacos por métodos computacionales y los aspectos más sobresalientes de este tipo de herramientas, empleadas en esta tesis. Los capítulos 3 y 4 muestran la obtención y los resultados de los modelos obtenidos para glicoproteína P y canal de sodio, respectivamente. En el capítulo 5 se reportan los resultados obtenidos del cribado virtual secuencial, la selección de candidatos

y la evaluación biológica. Finalmente en el último capítulo se enuncian las conclusiones finales y perspectivas del trabajo realizado.

Capítulo 1

Epilepsia

1.1. Generalidades.

La epilepsia es un desorden neuronal caracterizado por una predisposición crónica a sufrir crisis epilépticas y por las consecuencias neurológicas, cognitivas, fisiológicas y sociales de esta condición [1]. Una crisis epiléptica es una manifestación clínica que resulta de un episodio de actividad neuronal anormal en el cerebro [2]. Se caracteriza por síntomas transitorios como, perdida de la conciencia, alteraciones motoras y sensoriales, cambios de humor, entre otras alteraciones. Pueden presentarse como breves períodos de pérdida de la atención hasta convulsiones prolongadas e intensas [3].

Por definición, para enmarcar un caso clínico como epilepsia se requiere alguna de las siguientes condiciones: 1) Al menos dos crisis epilépticas no provocadas, separadas por un período de tiempo mayor a 24 horas; 2) Una crisis epiléptica no provocada y un riesgo de presentar nuevas crisis similar al producido tras dos crisis no provocadas (esto es, al menos, un 60 % de probabilidad de recurrencia en los próximos 10 años); 3) El diagnóstico de un síndrome epiléptico [4].

Por otra parte, se considera la epilepsia como resuelta en individuos que presentan un síndrome epiléptico con aparición dependiente de la edad y que han superado ya esa edad de susceptibilidad; o aquellos que han permanecido sin crisis durante al menos 10 años

y sin fármacos anticonvulsivos en los últimos 5 años. Sin embargo, esta consideración no es equivalente al concepto convencional de "curación" [4]. Esto implica que, si bien la persona no presenta las manifestaciones de la enfermedad durante un largo período de tiempo, no hay garantías que estas manifestaciones no vuelvan a aparecer.

De acuerdo a la Organización Mundial de la Salud, se estima que alrededor de 50 millones de personas en el mundo presentan epilepsia. Por lo cual, es considerada una de las enfermedades neurológicas crónicas más comunes [3]. En países desarrollados, se estima que la incidencia anual de epilepsia es de 50 casos por cada 100.000 personas, mientras que la prevalencia es de aproximadamente 700 pacientes por cada 100.000 habitantes [5]. En países en vías de desarrollo se estima que estos indicadores son generalmente más altos [3]. Esta patología se presenta a cualquier edad, pero con mayor frecuencia en pacientes jóvenes, en neonatos y también en adultos mayores (alrededor de los 75 años de edad) [2].

En nuestro país, el último estudio de la prevalencia de la enfermedad data de los años 1991-1993, realizado por Melcon y colaboradores [6]. En este estudio se comunicó que hay una prevalencia de 6,2 personas con epilepsia cada 1000 habitantes, de los cuales 3,8 de cada 1000 pacientes han sufrido al menos una crisis en los últimos dos años.

1.2. Clasificación de la epilepsia.

En el año 2010 la Liga Internacional contra la epilepsia (ILAE, por sus siglas en ingles) propuso cambios en la nomenclatura y clasificación de las crisis y tipos de epilepsia. Anteriormente, la clasificación se basaba en dos criterios fundamentales: el topográfico y el etiológico (o causal). En el primer caso, las epilepsias se dividían en tres tipos: parciales/focales, generalizadas y de localización desconocida. Según las causas, cada uno de estos grupos se subclasificaba como epilepsias idiopáticas, sintomáticas o criptogénicas. En la nueva clasificación se suprimió la división de base topográfica y los grupos vinculados a las causas fueron reemplazados por categorías denominadas genéticas, es-

tructurales/metabólicas y de causas desconocidas [7].

La clasificación genética es utilizada para epilepsias que son el resultado directo de un defecto genético conocido o presumible, en el cual la crisis es el síntoma central del desorden. El conocimiento de la contribución genética puede derivar de estudios moleculares genéticos específicos que se encuentran estandarizados, o puede provenir de estudios familiares correctamente diseñados [7]. Esta clasificación no excluye la posibilidad de que factores ambientales puedan contribuir a la enfermedad [7].

La epilepsia estructural o metabólica se desarrolla como resultado de una alteración estructural o metabólica determinada, o por una enfermedad que se ha demostrado que se asocia con un aumento del riesgo de desencadenar epilepsia. Las alteraciones estructurales incluyen accidentes cerebro-vasculares, lesiones físicas, infecciones, etc. Estas, pueden ser de origen genético tales como la esclerosis tuberular o malformaciones corticales en las que se presenta un desorden independiente de la epilepsia.

Por último, se clasifican como de causa desconocida aquellas epilepsias cuyo origen aún no ha podido establecerse. Podrían deberse a un trastorno genético o ser consecuencia de un desorden todavía no reconocido [7].

En forma independiente a las consideraciones etiológicas, la ILAE estableció una nueva clasificación para las epilepsias según el grado de especificidad sindrómica [7]. De esta manera se establece a los "síndromes electroclínicos" como el grupo de mayor especificidad, seguido de las "constelaciones", las de "causa estructural/metabólica" y finalmente las de "causa desconocida". En la tabla 1.1 se resume esta nueva clasificación y se enumeran los síndromes característicos de cada grupo.

Los síndromes electroclínicos son aquellos que presentan una serie de hallazgos clínicos, signos y síntomas que, en su conjunto, definen una alteración clínica específica y reconocible. Estos síndromes se encuentran definidos por los tipos de crisis, así como su severidad, la edad de presentación, la frecuencia, la respuesta a factores desencadenantes, la presencia o ausencia de otros síntomas acompañantes y la evolución de la enfermedad,

entre otros [7].

Por su parte las constelaciones epilépticas consisten en epilepsias que no se consideran como síndromes electroclínicos en un sentido estricto, pero que presentan características clínicas distintivas basándose en lesiones específicas u otras causas. Podrían tener implicaciones en el tratamiento, sobre todo quirúrgico. En esta clasificación se incluye la epilepsia del lóbulo temporal mesial (con esclerosis de hipocampo), el hamartoma hipotalámico con crisis gelásticas, la epilepsia con hemiconvulsión y hemiplejía y el síndrome de Rasmussen [7].

Finalmente, las epilepsias de causa estructural/metabólica carecen de un patrón electroclínico característico, y las epilepsias de causa desconocida pertenecen al grupo de menor especificidad sindrómica. Dentro de esta última clase se incluyen aquellas epilepsias que en el pasado se consideraban "criptogénicas".

Síndromes electroclínicos y otras epilepsias:
Síndromes electroclínicos (según la edad de comienzo): 1. Período neonatal: • Epilepsia neonatal familiar benigna • Encefalopatía mioclónica precoz • Síndrome de Ohtahara 2. Lactancia: • Síndrome de West • Síndrome de Dravet • Epilepsia mioclónica infantil/lactante • Epilepsia infantil/del lactante benigna • Encefalopatía mioclónica asociada a enfermedad no progresiva • Epilepsia del lactante con crisis focales migratorias 3. Infancia: • Epilepsia con crisis mioclónicas-atónicas/astáticas o síndrome de Doose • Epilepsia con crisis febriles plus • Síndrome de Panayiotopoulos o epilepsia occipital benigna de la infancia de inicio precoz • Epilepsia benigna de la infancia rolándica o con puntas centrotemporales • Epilepsia occipital benigna de la infancia de inicio tardío o tipo Gastaut • Epilepsia frontal nocturna autosómica dominante • Epilepsia con ausencias de la infancia • Epilepsia con ausencias mioclónicas • Síndrome de Lennox-Gastaut
Continua en la página siguiente

Síndromes electroclínicos y otras epilepsias:
• Síndrome Landau-Kleffner • Encefalopatía epiléptica con punta-onda continua durante el sueño 4. Adolescencia/adultez: • Epilepsia con ausencias juvenil • Epilepsia mioclónica juvenil o síndrome de Janz • Epilepsia con crisis generalizadas tónico-clónicas • Epilepsia temporal autosómica dominante con síntomas auditivos • Otras epilepsias temporales familiares • Epilepsias mioclónicas progresivas 5. Menos relacionadas con la edad: • Epilepsia focal familiar con focos variables • Epilepsias reflejas
Constelaciones: • Síndrome de Rasmussen • Epilepsia temporal mesial con esclerosis del hipocampo • Crisis gelásticas con hamartoma hipotalámico
Causas estructurales/metabólicas: • Malformaciones del desarrollo cortical • Síndromes neurocutaneos • Tumores • Infecciones • Traumatismos
Causas desconocidas (crisis epilépticas no diagnosticadas como epilepsias): • Convulsiones neonatales benignas • Convulsiones febriles

Tabla 1.1: Síndromes electroclínicos y otras epilepsias (clasificación según la ILAE) [7].

1.3. Crisis epilépticas.

Las crisis epilépticas son la manifestación clínica de la epilepsia y pueden clasificarse en dos grandes grupos según su origen: Las crisis generalizadas y las focales (anteriormente también denominadas parciales). Las primeras son originadas en algún punto dentro de la red neuronal distribuida bilateralmente y presentan una activación rápida. Esta red bilateral puede incluir estructuras corticales y subcorticales, pero no necesariamente incluye la corteza entera. En el caso de las segundas, estas se originan en una red limitada a uno de los hemisferios cerebrales. Las crisis focales también pueden originarse en

estructuras subcorticales [7]. A continuación se describirán brevemente estos dos tipos de crisis.

1.3.1. Crisis generalizadas.

Las crisis generalizadas pueden ser clasificadas como tónico-clónicas, crisis de ausencia, clónicas, tónicas, atónicas, o mioclónicas. A su vez las crisis de ausencia pueden ser divididas en típicas, atípicas y ausencias con características especiales (crisis de ausencia mioclónica o mioclonia palpebral). En el caso de las mioclónicas estas pueden dividirse en mioclónicas, mioclónicas-atonicas o mioclónicas-tonicas [8].

1.3.1.1. Crisis tónico-clónicas.

Consisten en una serie de respuestas motoras, incluyendo una extensión tónica de las extremidades, con una duración de entre 20–40 segundos, seguido de una fase clónica o movimientos espasmódicos rítmicos de las extremidades por 30 a 50 segundos. Durante la crisis los pacientes pierden la conciencia, pudiendo permanecer inconscientes durante un período de tiempo posterior a la finalización de los movimientos espasmódicos. Comúnmente durante la fase tónica puede presentarse un cese momentáneo en la respiración e incontinencia [9].

1.3.1.2. Crisis de ausencia.

Las crisis típicas de ausencia consisten en la pérdida momentánea de la consciencia con una duración generalmente menor a 15 segundos, a menudo indetectables. El grado de pérdida de la consciencia varía y es más pronunciado al inicio de la crisis [9, 10].

En el caso de las crisis de ausencia atípicas, el inicio no es tan marcado como en las típicas y suelen ser más prolongadas (mayor a 10 segundos), pero el grado de pérdida de consciencia suele ser menor. Pueden asociar componentes atónicos, tónicos y automatis-

mos [10].

En la crisis de ausencia mioclónica se produce una disminución de la consciencia asociada a mioclonías rítmicas de los hombros, brazos y piernas con contracción tónica predominante en hombros. Las mioclonías más frecuentes son las periorales, mientras que las palpebrales son más raras [11].

1.3.1.3. Crisis clónicas.

Es una contracción muscular repetitiva, regular y que afecta al mismo grupo muscular, prolongada en el tiempo. Puede afectar la consciencia [12].

1.3.1.4. Crisis tónicas.

Se produce por un aumento sostenido en la contracción muscular de segundos de duración, aunque ocasionalmente puede prolongarse unos minutos [12].

1.3.1.5. Crisis atónicas.

Se caracterizan por una reducción repentina del tono postural que ocasiona una pérdida de la postura. Se producen como consecuencia de una disminución brusca del tono muscular en músculos flexores del cuello, tronco y extremidades, sin aparente evento previo mioclónico o tónico [13].

1.3.1.6. Crisis mioclónicas.

Se debe a descargas corticales que producen contracciones involuntarias, únicas o múltiples, repentinas, breves, predominantemente en brazos, pero también en otras extremidades. Los movimientos generalmente son bilaterales pero también pueden observarse movimientos predominantes en un lado respecto del otro. Tienden a ocurrir en la

mañana, habitualmente al despertar [9].

1.3.2. Crisis focales.

No se recomienda una clasificación específica para este tipo de crisis. Estas se describen según las características de su semiología. Por ejemplo, si involucra o no perdida de la consciencia, si presenta o no componentes motores o autonómicos, etc. Esta descripción puede encontrarse en el glosario desarrollado por Blume y colaboradores [13].

1.3.3. Crisis de origen desconocido.

Los espasmos epilépticos son contracciones tónicas en flexión, extensión o mixtas, bilaterales breves y bruscas que afectan a la musculatura troncal y proximal que es usualmente mas sostenida que un movimiento mioclónico pero no tanto como una crisis tónica. La duración es de aproximadamente 1 segundo. Suele acompañarse después de una fase de falta de respuesta de 1-2 minutos que puede ser toda la sintomatología objetivable por el observador. Es muy habitual la presentación en clusters. Debido a que con los conocimientos actuales no se puede tomar una decisión acerca de si se trata de crisis generalizadas o focales, deben incluirse en el grupo de crisis de origen desconocido [7].

1.4. Diagnóstico.

El diagnóstico de la epilepsia es una tarea compleja, no solo por la importancia de encontrar las posibles causas de las crisis y así poder proveer el tratamiento apropiado, sino además porque otros desórdenes pueden causar síntomas parecidos a una crisis epiléptica generando confusión.

Debe realizarse una historia clínica detallada indicando antecedentes familiares y per-

sonales del paciente así como una examinación neurológica completa para determinar la etiología de la crisis, el tipo y la localización del foco epiléptico [14].

Durante el diagnóstico pueden incluirse estudios de laboratorio como análisis de sangre, orina y liquido cefalorraquídeo, estudios toxicológicos para descartar posibles causas metabólicas, toxicas, infecciosas o inflamatorias y estudios genéticos e inmunológicos [14].

Varias herramientas de diagnóstico son utilizadas para identificar y clasificar el tipo de crisis y la etiología, incluyendo el electroencefalograma (EEG), la resonancia magnética (RM), la tomografía computada de emisión de positrones (TEP), la tomografía computada por emisión de fotón único (SPECT), el magneto encefalograma (MEG) y estudios neuropsiquiátricos.

La prueba diagnóstica inicial más utilizada es el monitoreo por electroencefalograma (EEG), donde se registran las ondas cerebrales. Estas son detectadas por los electrodos colocados en el cuero cabelludo (EGG no invasivo) o a nivel cerebral profundo (EGG invasivo). La crisis puede ser confirmada solo cuando se produce durante el EEG, lo que dificulta el diagnóstico. Sin embargo, la presencia de descargas epileptiformes interictales (DEI) es un marcador electrofisiológico de epilepsia. Rara vez el EEG de un individuo sano muestra este tipo de descargas, mientras que los registros realizados a pacientes en el transcurso de las primeras 24 horas después de una crisis aumentan la probabilidad de detectar DEI [2].

Por otra parte, los estudios de neuroimagen proporcionan información de las posibles anomalías estructurales y funcionales relacionadas con el foco epileptogénico y, también, de la posible etiología. Entre ellos se destaca la tomografía computada (TC) y la resonancia magnética (RM). Estas herramientas muestran la estructura del cerebro y sirven para identificar lesiones en forma no invasiva. La tomografía por emisión de positrones (TEP) y la resonancia magnética funcional (RM funcional) monitorean la actividad cerebral y pueden detectar irregularidades en su funcionamiento. Particularmente, la tomografía computada por emisión de fotón único (SPECT) se utiliza para ubicar el foco epileptogénico. Esta prueba mide el flujo sanguíneo cerebral a través de agentes específicos que

indican cómo se distribuye la sangre en el cerebro. Se comparan diferencias regionales en el flujo de sangre durante la crisis (SPECT ictal) y entre las crisis (SPECT interictal) [2].

1.5. Tratamiento.

El tratamiento principal en la epilepsia es farmacológico. Los medicamentos existentes al momento no curan la epilepsia, pero evitan o reducen las crisis. Aproximadamente un 70-80 % de los pacientes logran entrar en remisión con los tratamientos actuales. Sin embargo, estos fallan en el control de las crisis del 20-30 % restante [15]. El tratamiento es empírico y a menudo basado en prueba y error, dependiendo de diversos factores como la frecuencia y severidad de las crisis, la edad de la persona y su historia clínica. Además, la dosis debe ajustarse en forma repetida teniendo en cuenta la respuesta del paciente y la aparición de efectos adversos. A lo largo del tiempo una gran variedad de fármacos han sido aprobados para el tratamiento de la epilepsia [15].

Cuando los medicamentos no pueden controlar adecuadamente las crisis se evalúa la posibilidad de realizar una cirugía. En este caso el paciente debe someterse a una evaluación prequirúrgica, donde se determinará si es beneficioso o no este tipo de opción terapéutica. La cirugía tiene como finalidad la eliminación o disminución de las crisis epilépticas mediante la resección de la zona epileptógena, sin producir déficits al paciente (cirugía resectiva). En el caso de que no sea posible una cirugía resectiva, puede recurrirse a procedimientos de desconexión, donde se realizan cortes en las terminales nerviosas para evitar la extensión de los impulsos convulsivos (transección subpial múltiple y callostomía). Aún cuando la cirugía detenga por completo las crisis epilépticas, es recomendable que se sigan administrando medicamentos durante un cierto período, para darle tiempo al cerebro de readaptarse [8].

La neuroestimulación es el tratamiento utilizado en pacientes fármaco resistentes que no son buenos candidatos para cirugía resectiva. Particularmente la estimulación vagal es

la metodología mas empleada. En ella se envía al cerebro impulsos pequeños y regulares de energía eléctrica a través de las fibras aferentes del nervio vago en la zona cervical, mediante un dispositivo implantado quirúrgicamente [16]. Una alternativa reciente es la estimulación vagal transdérmica, consistente en la estimulación transcutánea de la rama auricular de este nervio, usando un dispositivo externo ubicado en el pabellón del oído [16]. Otra opción es la estimulación cerebral profunda, que utiliza electrodos intracraneales para estimular el núcleo anterior del tálamo [17]. Se piensa que esta región puede actuar amplificando y sincronizando las descargas convulsivas provenientes del hipocampo y del tálamo central y posterior. Por lo tanto, la inhibición del tálamo mediante la estimulación eléctrica puede prevenir las crisis epilépticas [16]. Finalmente, la neuroestimulación responsiva consiste en la prevención de las crisis mediante un generador de pulsos, un software detector de convulsiones y electrodos intracraneales. Los electrodos se colocan en posiciones cercanas al foco epiléptico definido antes de la implantación. A diferencia de los otros métodos de estimulación, que envían estímulos pre programados para prevenir las crisis, la neuroestimulación responsiva funciona mediante la detección de la convulsión y su inhibición mediante un estimulo eléctrico [16].

A partir de estudios doble ciego, se ha demostrado que un 35 % de los pacientes tratados mediante estimulación vagal han logrado reducir al menos un 50 % la frecuencia de las crisis epilépticas. Este tratamiento es bien tolerado, presentando una baja tasa de efectos adversos graves (por ejemplo, parálisis de las cuerdas vocales o infección de la zona de implantación del dispositivo) [16]. En el caso de la estimulación cerebral profunda, el tratamiento demostró un 38,8 % de reducción de las crisis, mientras que el grupo control presentó un 22,8 %. Entre los efectos adversos más importantes se encuentran la depresión (14,8 %), alteraciones de la memoria (13 %), hemorragia (4,5 %) e infecciones (12,7 %) [16]. Finalmente, la neuroestimulación responsiva demostró una reducción de las crisis en un 37,9 % frente a un 17,3 % del grupo control [16].

A pesar de los avances en el tratamiento de la epilepsia mencionados anteriormente, todavía una gran cantidad de pacientes permanecen con crisis no controladas. Si el pa-

ciente no es candidato al tratamiento farmacológico, quirúrgico o de neuroestimulación, deben considerarse terapias alternativas. Entre ellas se encuentra la dieta cetogénica, que consiste en una alimentación rica en grasas y baja en carbohidratos y proteínas. Se ha demostrado que este tipo de régimen alimenticio controla o disminuye la frecuencia de las crisis epilépticas en niños [18]. Sin embargo, no se conoce al momento el mecanismo de inhibición de las crisis de esta terapia. Se ha postulado que favorece la síntesis de glutamato (precursor del neurotransmisor inhibitorio ácido γ-aminobutírico, GABA) y también de cuerpos cetónicos, similares al GABA, que podrían tener una acción directa anticonvulsiva [19, 20]. Adicionalmente, la inhibición de los canales iónicos o la estabilización de la membrana neuronal mediante una alteración del pH son otros de los posibles mecanismos implicados [19].

Por otra parte, la alimentación prolongada con tan alto contenido graso propicia efectos secundarios relacionados con el retardo en el crecimiento y la acumulación de ácido úrico en la sangre, lo que puede causar cálculos renales [21].

1.6. Fármacos antiepilépticos.

Históricamente, los fármacos antiepilépticos pueden ser clasificados en tres generaciones. La primera generación comprende aquellos comercializados entre 1857 y 1958 e incluye al bromuro de potasio, el fenobarbital y moléculas derivadas de los barbitúricos, como la fenitoína, la primidona, la trimetadiona y la etosuximida. La segunda generación incluye fármacos como la carbamazepina, el valproato y las benzodiacepinas, introducidos entre 1960 y 1975, químicamente diferentes a los barbitúricos. La tercera generación agrupa a fármacos comercializados a partir de 1980, entre los que se incluyen los primeros fármacos descubiertos de forma racional, como la progabida, la gabapentina, la vigabatrina y la tiagabina [22]. En la tabla 1.2 se enumeran los fármacos utilizados y sus características principales [15]. A continuación se enumeraran en mayor detalle los fármacos descubiertos en cada generación.

Fármaco	Año	Mecanismo de acción	Indicaciones aprobadas	Usos principales	Limitaciones
Fármacos de primera generación					
Bromuro de Potasio	1857	Potenciación inhibitoria GABA	Crisis convulsivas tónico-clónicas, mioclónicas	Amplio uso en convulsiones focales y generalizadas	Efecto sedante
Fenobarbital	1912	Potenciación inhibitoria GABA	Crisis convulsivas generalizadas y focales, sedante, trastornos de la ansiedad y del sueño	Amplio uso en convulsiones focales y generalizadas	Inductor enzimático, reacciones de hipersensibilidad dérmicas, sedación, hipnosis
Fenitoína	1938	Bloqueante de canales de Sodio	Crisis convulsivas generalizadas y focales	Uso de primera línea	Inductor enzimático, reacciones de hipersensibilidad dérmicas. Farmacocinética no lineal. No utilizado en mioclonias o crisis de ausencia
Trimetadiona	1946	Bloqueante de canales de Calcio tipo T	Crisis de ausencia	Uso poco habitual en crisis de ausencia	Teratogénico
Primidona	1954	Potenciación inhibitoria GABA	Crisis convulsivas generalizadas y focales	Amplio uso en convulsiones focales y generalizadas	Inductor enzimático, reacciones de hipersensibilidad dérmicas, sedación, hipnosis
Etosuximida	1958	Bloqueante de canales de Calcio tipo T	Crisis de ausencia	Uso de primera línea	Somnolencia, pérdida de apetito, nauseas, vómitos, hipo, depresión, psicosis, insomnio, anemia aplásica rara
Fármacos de segunda generación					

Continua en la página siguiente

Fármaco	Año	Mecanismo de acción	Indicaciones aprobadas	Usos principales	Limitaciones
Diazepam	1963	Potenciación inhibitoria GABA	Desórdenes convulsivos, estado epiléptico, ansiedad. Deshabituación alcohólica	Amplio uso en convulsiones focales y generalizada, uso intravenoso	Sedación, tolerancia. Sólo uso como adyuvante
Carbamazepina	1964	Bloqueante de canales de Sodio	Crisis convulsivas generalizadas y focales, desorden bipolar, dolor trigeminal	Uso de primera línea	Inductor enzimático, reacciones de hipersensibilidad dérmicas
Valproato	1967	Potenciación GABA, inhibición glutamatérgica NMDA, Bloqueo de canales de Sodio y de Calcio tipo T	Crisis convulsivas focales y generalizadas, crisis de ausencia, profilaxis de la migraña, desorden bipolar	Uso de primera línea. Amplio uso en convulsiones focales y generalizadas, uso intravenoso	Inductor enzimático, ganancia de peso, teratogenicidad considerable
Clonazepam	1968	Potenciación inhibitoria GABA	Síndrome de Lennox-Gastaut, convulsiones mioclónicas, pánico	Amplio uso en convulsiones focales y generalizada, uso intravenoso	Sedación, tolerancia, solo utilizado como adyuvante
Clobazam	1975	Potenciación inhibitoria GABA	Síndrome de Lennox-Gastaut, desorden de ansiedad	Amplio uso en convulsiones focales y generalizada, uso intravenoso	Sedación, tolerancia, solo utilizado como adyuvante
Fármacos de tercera generación					
Progabida	1985	Potenciación inhibitoria GABA	Crisis convulsivas focales y generalizadas y mioclónicas. Síndrome de Lennox-Gastaut, hipertonía muscular	Uso habitual en convulsiones focales	Hepatotoxicidad clínica, fuera de uso

Continua en la página siguiente

Fármaco	Año	Mecanismo de acción	Indicaciones aprobadas	Usos principales	Limitaciones
Vigabatrin	1989	Potenciación inhibitoria GABA	Espasmos epilépticos, convulsiones focales	Espasmos epilépticos	Perdida de la visión, ganancia de peso
Lamotrigina	1990	Bloqueante de canales de Sodio	Crisis convulsivas focales y generalizadas, Síndrome de Lennox-Gastaut, desorden bipolar	Uso de primera línea. Amplio uso en convulsiones focales y generalizadas	Inductor enzimático, reacciones de hipersensibilidad dérmicas
Oxcarbazepina	1990	Bloqueante de canales de Sodio	Convulsiones focales	Uso de primera línea	Inductor enzimático, reacciones de hipersensibilidad dérmicas. No utilizado en mioclonias o crisis de ausencia
Felbamato	1993	Potenciación GABA, inhibición glutamatérgica NMDA, Bloqueo de canales de Sodio y de Calcio tipo T	Crisis convulsivas focales y generalizadas, Síndrome de Lennox-Gastaut	Amplio uso en convulsiones focales y generalizadas	Sólo utilizado como adyuvante, anemia aplásica, hepatotoxicidad clínica, reacciones de hipersensibilidad dérmicas
Gabapentina	1993	Bloqueante de Canales de Calcio (subunidad $\alpha 2\delta$)	Crisis convulsivas focales y generalizadas, neuralgia diabética y post-herpética, síndrome de piernas inquietas	Síndrome de piernas inquietas, dolor neuropático	Sólo utilizado como adyuvante, ganancia de peso

Continua en la página siguiente

Fármaco	Año	Mecanismo de acción	Indicaciones aprobadas	Usos principales	Limitaciones
Topiramato	1995	Potenciación GABA, inhibición glutamatérgica NMDA, Bloqueo de canales de Sodio y de Calcio tipo T	Crisis convulsivas focales y generalizadas, Síndrome de Lennox-Gastaut, profilaxis de la migraña	Uso de primera línea. Amplio uso en convulsiones focales y generalizadas	Somnolencia, mareos, vértigo, deterioro cognitivo, trastornos del habla cálculos renales, pérdida de peso.
Tiagabina	1996	Potenciación inhibitoria GABA	Convulsiones focales	Convulsiones focales	Confusión, dificultad para hablar, tartamudeo, sedación leve, parestesia.
Levetiracetam	2000	Modulación SV2A	Crisis convulsivas focales, generalizadas, tónico-clónicas. Epilepsia Mioclónica Juvenil	Uso de primera línea	Somnolencia, astenia, cefalea, vértigo ataxia
Zonisamida	2000	Bloqueante de canales de Sodio	Convulsiones focales	Amplio uso en convulsiones focales y generalizadas	Solo utilizado como adyuvante, sedante
Stiripentol	2002	Potenciación inhibitoria GABA, bloqueo de canales de Sodio	Síndrome de Dravet	Síndrome de Dravet	Solo utilizado como adyuvante
Pregabalina	2004	Bloqueante de Canales de Calcio (subunidad $\alpha 2\delta$)	Convulsiones focales, dolor neuropático, trastornos de ansiedad, fibromialgia	Dolor neuropático	Solo utilizado como adyuvante, aumento de peso

Continua en la página siguiente

Fármaco	Año	Mecanismo de acción	Indicaciones aprobadas	Usos principales	Limitaciones
Rufinamida	2004	Bloqueante de canales de Sodio	Síndrome de Lennox-Gastaut	Síndrome de Lennox-Gastaut	Solo utilizado como adyuvante, somnolencia, vértigo, fatiga, cefalea, nauseas
Lacosamida	2008	Inactivación lenta de canales de Sodio	Convulsiones focales	Convulsiones focales	Solo utilizado como adyuvante
Eslicarbazepina	2009	Bloqueante de canales de Sodio	Convulsiones focales	Convulsiones focales	Inductor enzimático, somnolencia, mareos, hiponatremia, ataxia.
Retigabina	2011	Activador de canales de Potasio	Convulsiones focales	Adyuvante para tratar convulsiones focales	Coloración azul de labios y uñas, trastornos oculares, poco uso
Perampanel	2012	Agonista de receptor glutamatérgico AMPA	Convulsiones focales	Adyuvante para tratar convulsiones focales	Reacciones adversas conductuales y psiquiátricas graves, agresividad, irritabilidad, ira, pensamientos homicidas

Tabla 1.2: Fármacos antiepilépticos utilizados en clínica en la actualidad o en el pasado [15].

1.6.1. Fármacos de primera generación.

En 1857 se introdujo el bromuro de potasio como medicación capaz de controlar la epilepsia, luego de que el médico inglés Charles Locock observara su efecto sedante y su capacidad de reducir el número de crisis en algunos pacientes [23]. Este compuesto continuó como el único fármaco eficaz en el tratamiento de las crisis por 55 años, hasta la intro-

ducción del fenobarbital en 1912 por Alfred Hauptmann [24]. Hauptmann administraba fenobarbital a sus pacientes con epilepsia como tranquilizante, pero luego descubrió que este fármaco también causaba una reducción en las crisis [25]. Posteriormente, el fenobarbital se convirtió en la droga de elección para el tratamiento de la epilepsia, debido a su menor toxicidad y mayor efectividad respecto del bromuro de potasio, prevaleciendo su uso hasta la actualidad.

El siguiente compuesto introducido al mercado fue la fenitoína. Esta había sido sintetizada en 1908 por Heinrich Biltz, pero recién en 1938 se descubrió su actividad anticonvulsiva, debido a que carecía de efectos sedantes (propiedad considerada necesaria hasta ese momento en un anticonvulsivo) [26, 27]. La acción de la fenitoína fue descubierta por Merritt y Putnam, quienes en 1937 desarrollaron un modelo animal basado en la técnica de electroshock para la evaluación de las propiedades anticonvulsivas de las sustancias [28]. Este test resultó altamente exitoso para la búsqueda sistemática de fármacos más eficaces y mejor tolerados, utilizando bibliotecas de compuestos [22].

Por otra parte, luego del descubrimiento de la fenitoína, se iniciaron varios estudios sobre la fisiología de las crisis en animales y seres humanos, con los que se consiguieron una serie de pruebas para investigar la actividad antiepiléptica de nuevos compuestos. Entre ellas se destacan el ensayo de electroshock máximo y el de pentilentetrazol (PTZ) [22].

En el test de electroshock máximo, inicialmente se administra la sustancia a ensayar y luego se aplica un shock de corriente alterna de 60 Hz y 50 mA (en ratones) o 150 mA (en ratas) a través de electrodos localizados en la córnea por 0,2 segundos. El objetivo es verificar la eficacia del agente químico en abolir la extensión tónica de los miembros posteriores del animal [29]. En el modelo de PTZ, luego de la administración del compuesto se aplica una dosis convulsivante de pentilentetrazol (antagonista del receptor GABA-A) por vía subcutánea. Los animales son observados para verificar la presencia de espasmos clónicos que persistan por lo menos 5 segundos. La ausencia de esta actividad clónica indica que la sustancia ensayada elevó el umbral convulsivo [29].

La trimetadiona, un analgésico sintetizado por Richards y Everett en 1944 [30], fue reportada por su acción anticonvulsiva frente al ensayo PTZ, mientras que resulta inactivo frente al ensayo de electroshock [31]. Este fármaco se convirtió en el primer medicamento para las crisis de ausencia. Para tratar este tipo específico de crisis, se desarrollaron posteriormente otros fármacos como metosuximida, fensuximida y etosuximida [32]. El último de estos compuestos, menos tóxico, es considerado hasta el presente una opción de primera línea para el control de las crisis de ausencia [33].

Nuevos derivados encontrados en este período son la mefenitoína, la etotoína, el metharbital, la primidona y la parametadiona. Los dos primeros son análogos de la fenitoína, los dos siguientes del fenobarbital y el último de la trimetadiona. Todos presentan una actividad comparable a la del compuesto del que derivan y se comercializaron para un uso similar [34]. La figura 1.1 muestra la estructura química de los fármacos de primera generación. Otras drogas como fetenilato, benzocloropropamida y aminoglutetimida se introdujeron al mercado en esa época, pero fueron posteriormente retiradas debido a su toxicidad demostrada a largo plazo [34].

1.6.2. Fármacos de segunda generación.

A partir de los años 1950 surgieron nuevas drogas denominadas de segunda generación, entre las que se destacan la carbamazepina y el valproato de sodio. La figura 1.2 muestra la estructura química de los fármacos de segunda generación. El descubrimiento de la carbamazepina se produjo en 1953 durante la búsqueda de análogos del fármaco psicotrópico clorpromazina. Sus efectos antiepilépticos fueron reportados entre 1963 y 1964, y fue aprobada para su comercialización en Estados Unidos en el año 1974 [25].

El acido valproico fue sintetizado por primera vez a fines del siglo XIX, pero su acción anticonvulsiva fue detectada en Francia en 1963 durante una prueba de acción anticonvulsiva de compuestos heterocíclicos. Estos compuestos presentaban baja solubilidad en agua, por lo que debieron disolverse en diferentes solventes para los ensayos biológicos,

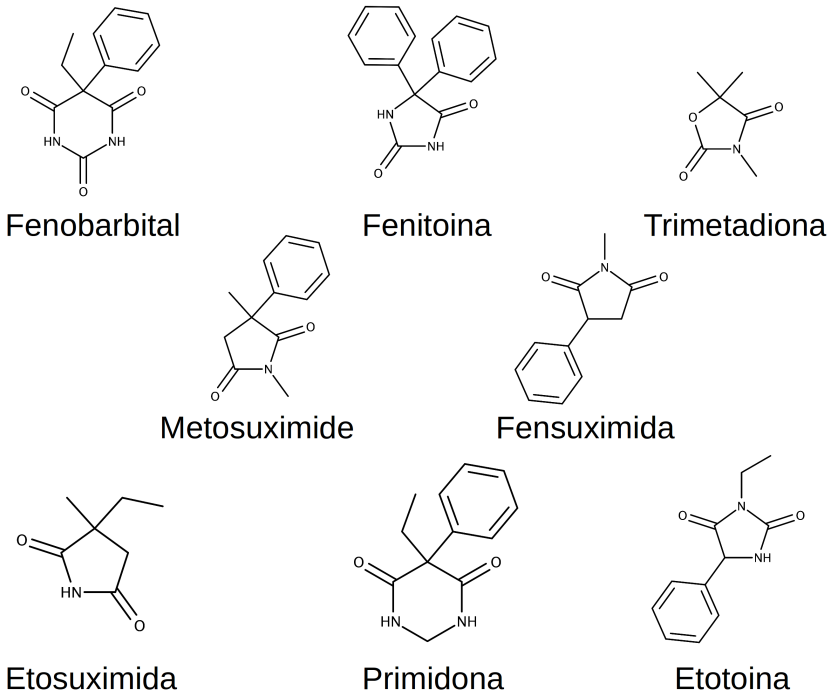


Figura 1.1: Fármacos antiepilépticos de primera generación.

entre los que se encontraba el acido valproico. Se encontró que varios de estos compuestos presentaban efectos anticonvulsivantes y que esta acción era, en realidad, dependiente del solvente [35]. Luego de estudios clínicos realizados en 1964, el valproato comenzó a comercializarse en Francia en 1967 [36]. Durante 1970, se inicio la comercialización en Europa, pero recién en 1978 se aprobó su uso en Estados Unidos.

En este período también fue descubierta la acción anticonvulsiva de las benzodiacepinas, como el diazepam, el clonazepam o el clobazam [37, 25].

1.6.3. Fármacos de tercera generación.

Luego de la introducción del ácido valproico y de las benzodiacepinas, tuvo lugar una pausa en el desarrollo de nuevos medicamentos que se prolongó hasta mediados de 1980. A partir de ese momento varios compuestos fueron aprobados, marcando el comienzo de una nueva generación de fármacos. Entre las moléculas pertenecientes a esta era

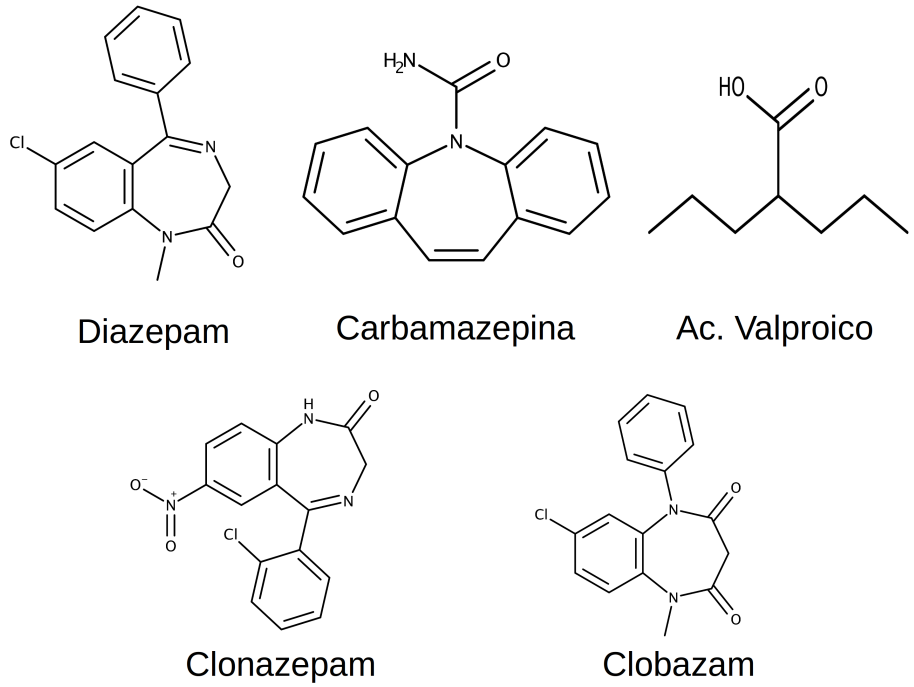


Figura 1.2: Fármacos antiepilépticos de segunda generación.

se encuentra la progabida (1985), vigabatrina (1989), lamotrigina (1990), oxcarbazepina (1990), gabapentina (1993), felbamato (1993), topiramato (1995), tiagabina (1996), zonisamida (2000), levetiracetam (2000), stiripentol (2002), pregabalina (2004), rufinamida (2004), lacosamida (2008), eslicarbazepina (2009), retigabina (2011) y perampanel (2012) [25]. La figura 1.3 muestra la estructura de los fármacos de tercera generación.

A diferencia de generaciones previas, donde el desarrollo de anticonvulsivos se ha caracterizado por la casualidad y el empirismo, dentro de la tercera generación de fármacos se encuentran los primeros ejemplos de diseño racional basado en el conocimiento de los posibles mecanismos fisiopatogénicos involucrados en las crisis epilépticas, así como también en el conocimiento del mecanismo de acción de fármacos conocidos. Como se detallara en la sección siguiente, se han desarrollado compuestos con el objeto de potenciar la vía GABAérgica (vigabatrina y tiagabina) y, más recientemente, la reducción de los mecanismos de hiperexcitabilidad mediados por neurotransmisores excitatorios que ejercen su acción sobre diferentes tipos de receptores (perampanel) [22].

Los nuevos antiepilépticos sin duda han logrado disminuir los efectos adversos de sus predecesores y han expandido las opciones terapéuticas, en particular para aquellos pacientes en los que se requiere un cambio de medicación. Sin embargo, la eficacia de muchos de estos sigue siendo similar a la de los compuestos clásicos, por lo que el porcentaje de pacientes resistentes a la medicación actual continúa sin disminuir. Es por eso que existe una creciente y genuina preocupación por parte de expertos en el área respecto de la falta de eficacia del tratamiento farmacológico, aun con la introducción de nuevos fármacos anticonvulsivos [38, 39]. En este contexto, un mayor estudio de los mecanismos de acción de los fármacos representa una estrategia fundamental para superar las limitaciones de los anticonvulsivos actuales.

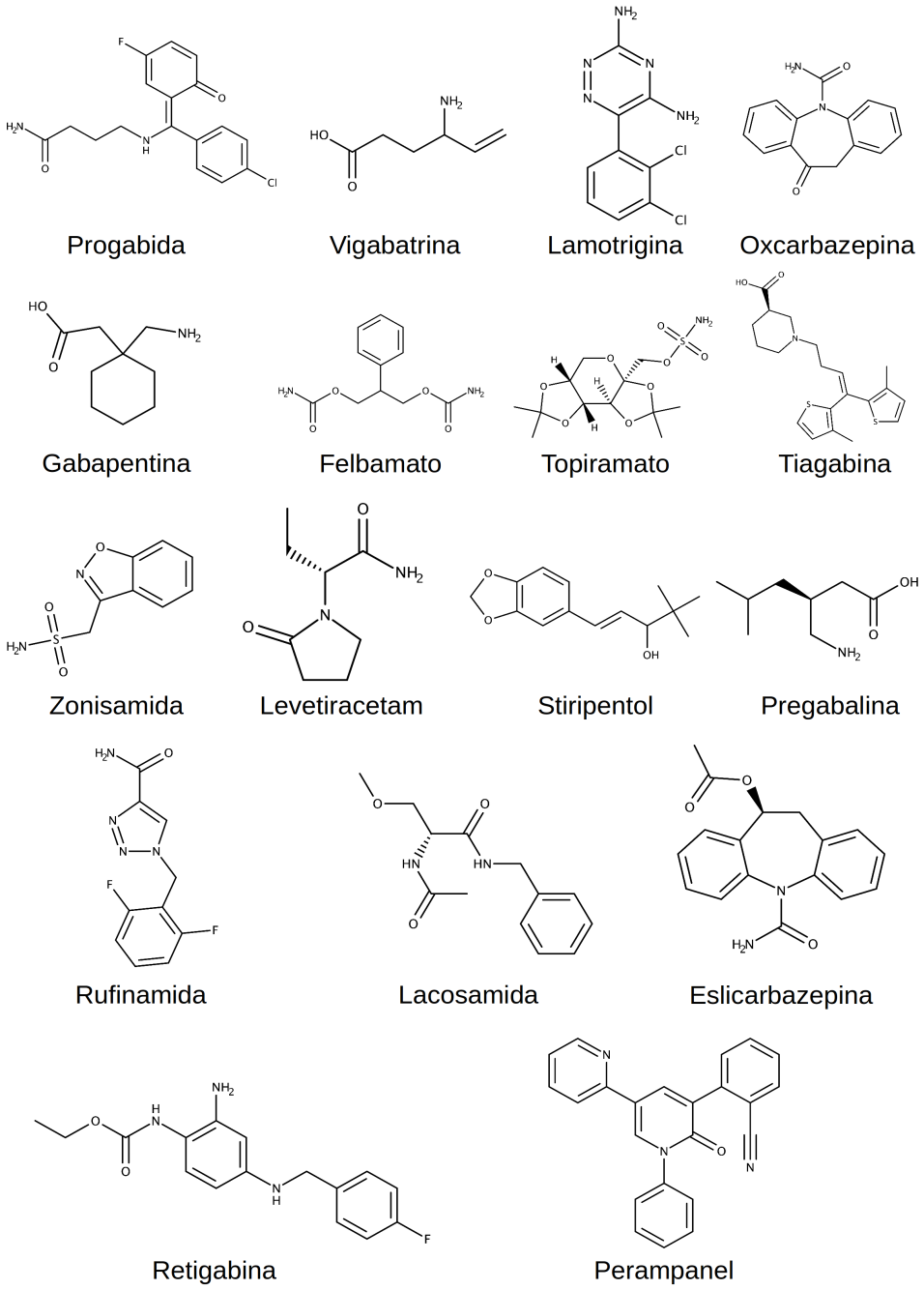


Figura 1.3: Fármacos antiepilépticos representativos de la tercera generación.

1.6.4. Mecanismos de acción de los fármacos antiepilépticos.

Los fármacos antiepilépticos disponibles son solo preventivos y no curativos. El objetivo de la fármacoterapia actual es controlar los síntomas, es decir, suprimir las crisis epilépticas. Si bien los mecanismos de acción de estas drogas no son completamente conocidos, se sabe que estos compuestos esencialmente actúan restableciendo el balance entre la excitación y la inhibición neuronal mediante su interacción con uno o más blancos moleculares en el cerebro [40]. Estos blancos comprenden diferentes tipos de macromoléculas: canales iónicos, receptores de neurotransmisores y enzimas encargadas de su metabolización y transporte [41]. El efecto final de estas interacciones es la alteración de la capacidad de descarga de las neuronas, la reducción de la sincronización de grupos de neuronas y la inhibición de la propagación de descargas anormales hacia sitios distales [40].

Se reconocen, al momento, tres mecanismos de acción principales para las drogas antiepilepticas: La regulación de canales iónicos voltaje operados (o dependientes de voltaje); el aumento de la inhibición de la neurotransmisión mediada por el acido γ-aminobutirico (GABA); y la atenuación de la neurotransmisión excitatoria mediada por glutamato [40]. La figura 1.4 resume los blancos moleculares conocidos de las drogas antiepilépticas. A continuación se describirán los mecanismos de acción clásicos, con un mayor detalle en el bloqueo de los canales de sodio operados por voltaje, dado que es el blanco molecular utilizado en esta tesis para el descubrimiento de nuevos anticonvulsivos.

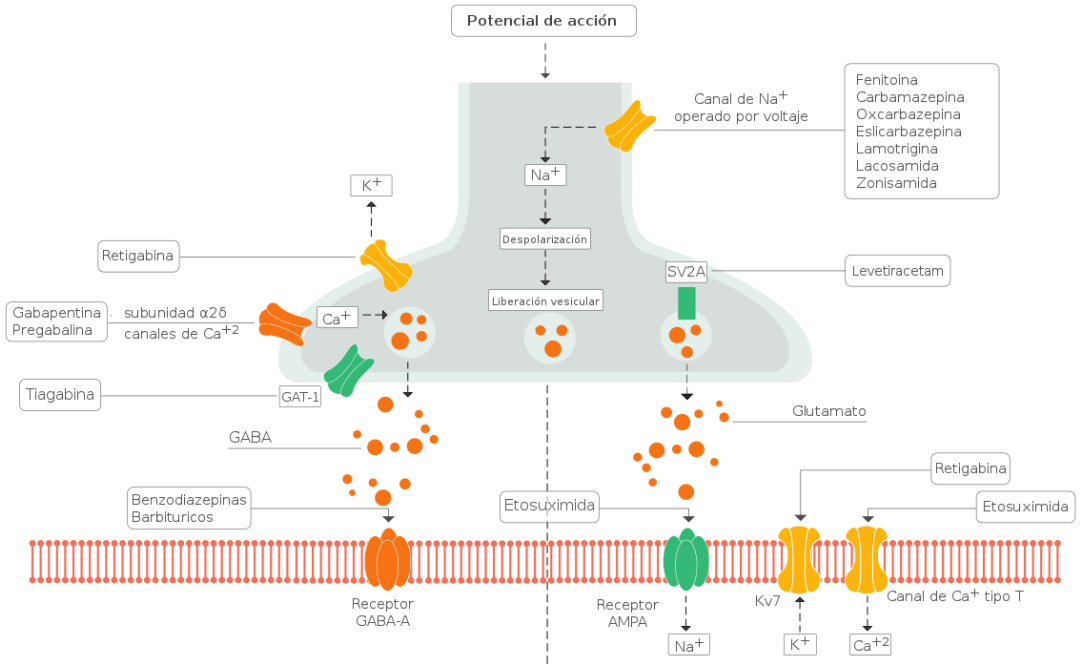


Figura 1.4: Mecanismos de acción de los fármacos antiepilépticos. Del lado izquierdo se muestran los blancos moleculares involucrados en la generación de efectos inhibitorios y del lado derecho los excitatorios. GAT-1: transportador GABA. SV2A: glicoproteína de la vesícula sináptica 2A. Adaptado con autorización de BMJ Publishing Group Ltd [42].

1.6.4.1. Modulación de canales iónicos dependientes de voltaje.

Los canales iónicos son proteínas que atraviesan la membrana plasmática y que permiten la difusión selectiva y pasiva de iones. La activación e inactivación de estas proteínas es inducida por estímulos externos, como el cambio del potencial eléctrico de la membrana celular (canales iónicos dependientes de voltaje) o la presencia de sustancias químicas (canales iónicos dependientes de ligandos). Durante la activación, el canal sufre un cambio conformacional que permite el pasaje de los iones a través de la membrana a favor de su gradiente de concentración.

Los canales iónicos dependientes de voltaje (entre los que se incluyen los canales de sodio, de calcio y los de potasio) modulan el comportamiento eléctrico neuronal,

permiten disparar los potenciales de acción y regulan la capacidad de estas células de responder a las señales sinápticas. Además, son elementos cruciales en la liberación de neurotransmisores, que se requieren para la transmisión sináptica. En consecuencia, son blancos moleculares clave para la inhibición del inicio, sincronización y propagación de las crisis epilépticas [40].

Desde el punto de vista estructural, los canales iónicos dependientes de voltaje comprenden una superfamilia de proteínas en la cual la subunidad que forma el poro para la transferencia de iones está compuesta por 4 dominios homólogos expresados en una misma estructura proteica (canales de sodio y de calcio) o por 4 subunidades idénticas (canales de potasio) [43]. Cada uno de estos dominios/subunidades está, a su vez, conformado por un motivo estructural de 6 hélices transmembrana (S1-S6). Los segmentos transmembrana S5-S6 constituyen el poro del canal (que permite el paso de iones) y, junto con el loop estructural presente entre ambos (loop P), forman la porción extracelular donde se encuentra la región selectora de iones. La hélice S4, con un patrón de 1 aminoácido con carga positiva cada 2 aminoácidos hidrofóbicos, funciona como sensor de voltaje. Es decir que el movimiento hacia arriba y de rotación de este segmento S4, debido al campo eléctrico de la membrana, induce el cambio conformacional que provoca la apertura y cierre del canal [44] (Figura 1.5).

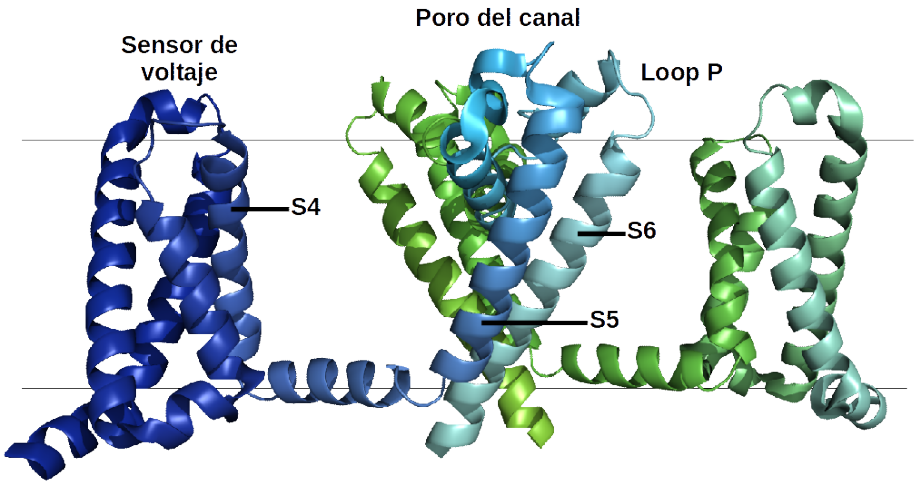


Figura 1.5: Organización estructural de los canales operados por voltaje. Para una mayor claridad solo se presentan los dominios I y III del poro del canal, perteneciente a la estructura experimental del canal de sodio de *Arcobacter butzleri* (código PDB: 3RVY) [45].

Canales de Sodio dependientes de voltaje.

Los canales de sodio dependientes de voltaje son responsables de la despolarización de la membrana de las células nerviosas y de la conducción del potencial de acción a través de la superficie neuronal [46]. Se encuentran expresados a lo largo de toda la membrana neuronal, en las dendritas, soma, axones y terminales nerviosas. Están compuestos por una subunidad proteica α (de un tamaño de aproximadamente 260 kDa), que constituye la región del poro y el sensor de voltaje; y de una o más subunidades proteicas β (30-40 kDa), que pueden modificar la función de la subunidad α pero no son esenciales para la actividad básica del canal (figura 1.6) [47].

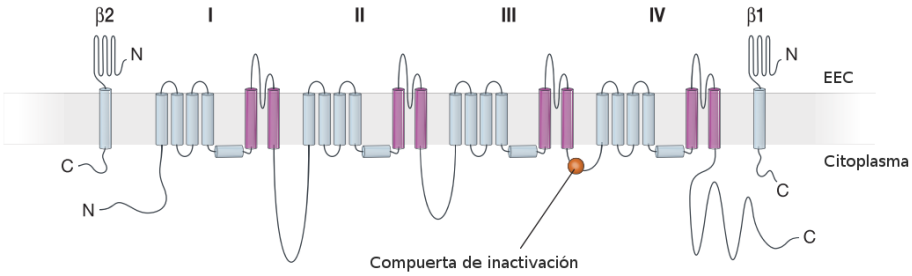


Figura 1.6: Topología de las subunidades α y β de los canales de sodio dependientes de voltaje. Se observa que consta de 4 dominios (I-IV), cada uno constituido por 6 segmentos transmembrana. Adaptada, con autorización, de Neurotherapeutics [43].

La familia de este tipo de canales comprende 9 isoformas resultantes de la variación genética de la subunidad α. Estas subunidades son codificadas por los genes SCN1A, SCN2A, SCN3A, SCN4A, SCN5A, SCN8A, SCN9A, SCN10A y SCN11A (figura 1.7). Las isoformas Nav1.1 (SCN1A, 2q), Nav1.2 (SCN2A, 2q), Nav1.3 (SCN3A, 2q) y Nav1.6 (SCN8A, 12q) predominan en el Sistema Nervioso Central de mamíferos. Particularmente, Nav1.2 y Nav1.6 se encuentran principalmente en el segmento inicial del axón de las principales neuronas excitatorias [48]. Las isoformas Nav1.7 (SCN9A, 2q), Nav1.8 (SCN10A, 3p) y Nav1.9 (SCN11A, 3q) predominan en el sistema nervioso periférico, Nav1.4 (SCN4A, 17q) está presente en el musculo esquelético y Nav1.5 (SCN5A, 2q) en musculo esquelético y cardíaco. La expresión de Nav1.3 esta mayormente restringida a las etapas tempranas del desarrollo, mientras que Nav1.1 es el principal canal de sodio en las interneuronas inhibitorias [49, 50].

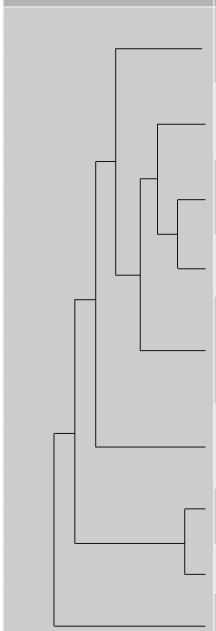
Relación filogenética	Subtipo	Gen	Cromosoma	Nro de aminoácidos	Sensibilidad a TTX	Tejido principal
	Nav1.6	SCN8A	12q13	1980	si	SNP, SNC
	Nav1.2	SCN2A	2q22-23	2005	si	SNC
	Nav1.1	SCN1A	2q24	2009	si	
	Nav1.3	SCN3A	2q23-24	1951	si	
	Nav1.7	SCN9A	2q24	1977	si	GRD, neuronas del trigémino
	Nav1.4	SCN4A	17q23-24	1836	si	Músculo esqueléti co
	Nav1.5	SCN5A	2q24	2016	no	Corazón
	Nav1.8	SCN10A	3p21-24	1957	no	GRD, neuronas del trigémino
	Nav1.9	SCN11A	3q21-24	1792	no	

Figura 1.7: Familia de canales de sodio operados por voltaje. Además de la relación filogenética, se incluye información genética y la principal ubicación de sus subunidades α (según Caterall y col. [51]). TTX: tetrodotoxina, SNC: sistema nervioso central, SNP: sistema nervioso periférico, GRD: Ganglios de la raíz dorsal.

Los canales de sodio dependientes de voltaje poseen al menos tres estados: desactivado, activado e inactivado (figura 1.8). Normalmente los canales se encuentran en un estado desactivado de conformación cerrada, pero cuando la despolarización de la membrana neuronal supera un valor umbral, se produce un cambio conformacional del canal que provoca el paso de este estado cerrado al estado activado (abierto), permitiendo el ingreso de iones sodio al interior celular. Luego de unos microsegundos el canal se inactiva (inactivación rápida), terminando el flujo de los iones sodio. Esta inactivación esta mediada por un loop intracelular entre los dominios III y IV de la subunidad α, donde el canal permanece en una conformación abierta pero impermeable a los iones [47]. Antes de poder ser reactivado por una nueva despolarización, es necesario que este vuelva al estado desactivado, el cual se alcanza luego de que la membrana se repolariza duran-

te un período de tiempo suficiente (tiempo de recuperación) [40]. Los canales de sodio pueden pasar rápidamente del estado desactivado al abierto y de este al inactivado, permitiendo a las neuronas disparar trenes de potenciales de acción de alta frecuencia, tanto en condiciones normales como en estados de actividad epiléptica. Cuando el proceso de despolarización de la membrana es más prolongado, puede iniciarse un mecanismo de inactivación adicional, conocido como inactivación lenta [40].

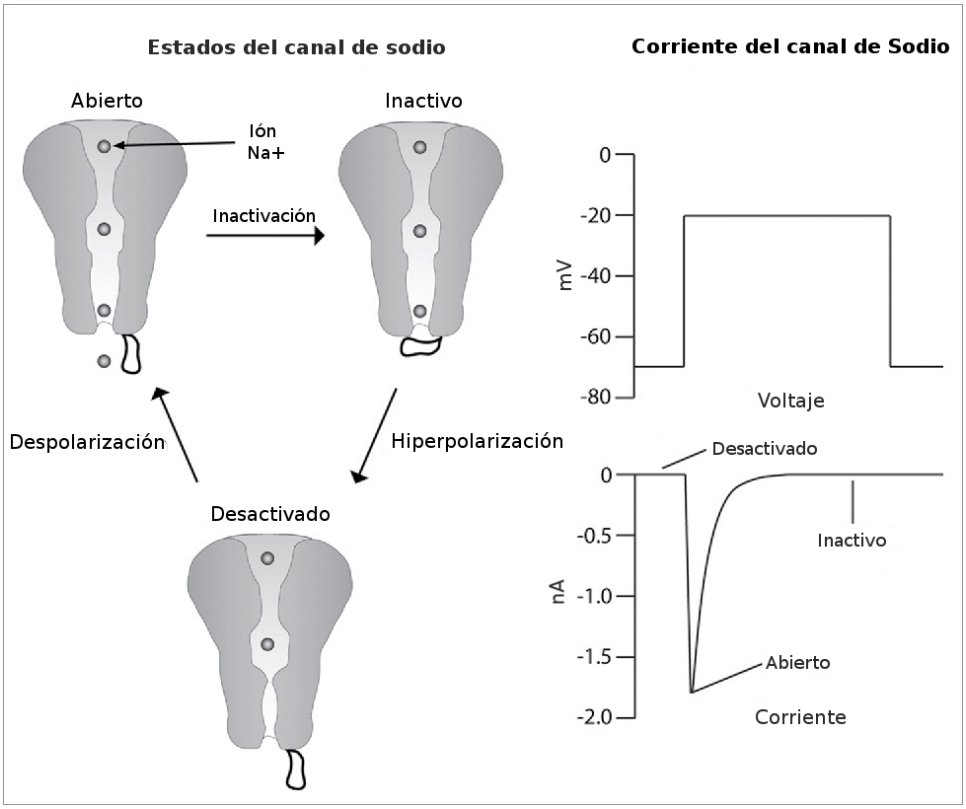


Figura 1.8: Estados del canal de sodio. Adaptado de Channels [46] bajo licencia Creative Commons Attribution.

En relación a su empleo como blancos moleculares, estas macromoléculas son bloqueadas por drogas de uso clínico como los anestésicos locales, anti arrítmicos y ciertos antiepilépticos [47]. El sitio de unión de estos fármacos está formado por residuos de aminoácidos del segmento S6 de los dominios I, III y IV. Por lo tanto, la acción farmacológica se basa en la unión a los canales en el interior del poro, impidiendo de esta forma

el pasaje de los iones [47].

La seguridad terapéutica de los bloqueantes de canales de sodio se produce debido a su selectividad funcional, ya que estos compuestos ejercen su máxima potencia solo en el caso de que el canal funcione bajo condiciones patológicas. En estas condiciones se producen potenciales de acción de alta frecuencia de disparo, que causan que los canales adopten el estado inactivado con mayor frecuencia. Por lo tanto, aquellos compuestos con alta afinidad por esta conformación presentan una mayor eficacia de unión [50]. En otras palabras, los diferentes estados del canal juegan un rol importante en la afinidad de unión de los fármacos, dado que la mayoría de los inhibidores ejercen su acción por estabilización del estado inactivado (lo que se conoce como inhibición dependiente de estado o dependiente de voltaje, ya que cada estado depende de este). Esta dependencia refleja un mecanismo alostérico por el cual el sitio de unión presenta una conformación de baja afinidad cuando se encuentra en estado cerrado y pasa a una conformación de alta afinidad cuando se encuentra abierto o en estado de inactivación [52].

La inhibición dependiente de estado puede resultar en inhibición dependiente de frecuencia (o de uso). Esta dependencia de la frecuencia se desarrolla durante pulsos de series rápidas de estimulación. Durante cada una de estas series rápidas, los canales experimentan aperturas e inactivaciones. Cuando el canal finalmente retorna al estado cerrado entre pulsos, estos bloqueos se disocian, pero este proceso es más lento que la recuperación normal desde el estado de inactivación. Así, cuando los pulsos se producen a una frecuencia suficiente, el bloqueo generado es acumulativo [50].

El bloqueo de los canales de sodio dependientes de voltaje es el mecanismo de acción más común de los fármacos antiepilépticos actualmente disponibles. Fenitoína, lamotrigina, carbamazepina, oxcarbazepina, eslicarbazepina, zonisamida y lacosamida son prototipos de bloqueantes de canales de sodio [53]. Se sospecha que compuestos asociados a otros mecanismos de acción anticonvulsiva como felbamato, topiramato y valproato podrían tener también efectos inhibitorios sobre estos canales [53]. Como se mencionó anteriormente para todos los inhibidores, los fármacos antiepilépticos con capacidad bloqueante

de canales de sodio tienen una alta afinidad por el estado inactivado y producen una reducción en la velocidad del cambio conformacional desde este estado al cerrado. Como consecuencia se produce una reducción en la conductancia del canal, limitando el disparo neuronal repetitivo, con un efecto muy bajo sobre la generación de potenciales de acción simples [50]. Debe mencionarse que, mientras que la mayor parte de los bloqueantes de los canales de sodio actúan bloqueando la ruta de inactivación rápida, la lacosamida parece incrementar el período de inactivación lenta [53].

Canales de calcio dependientes de voltaje.

Los canales de calcio dependientes de voltaje contribuyen a la excitabilidad eléctrica de las neuronas, están relacionados con la activación neuronal y son responsables del control de la liberación de los neurotransmisores de las terminales pre sinápticas [43, 54]. Están constituidos por una subunidad α-1 que forma el poro y el sensor de voltaje (codificada al menos por siete genes) y diversas subunidades accesorias, incluyendo las subunidades intracelulares β (β 1 – 4), las intramembrana γ (γ 1 – 8) y las subunidades intramembrana-extracelulares α-2-δ (tipos 1 – 4) que modulan la función y la expresión en la superficie celular [43].

Los canales de Calcio se distinguen en base al potencial de membrana al cual son activados [54]. Los canales de Calcio activados por alto voltaje (HVA, por sus siglas en ingles: High-voltage-activated) responden a despolarizaciones fuertes (del orden de los -40mV) y están involucrados tanto en la liberación pre-sináptica de los neurotransmisores (los tipos N, P/Q y R) como en el procesamiento de la sinapsis a nivel post-sináptico (tipo L). En contraste, los canales activados por bajo potencial (LVA, low-voltage-activated), también denominados tipo T, se abren en respuesta a despolarizaciones moderadas (-60mV) [40]. Estos últimos son altamente expresados en el soma y dendritas de neuronas del tálamo y en el sistema reticular, donde se ha postulado que descargas rítmicas de 3 Hz punta-onda son características de las crisis de ausencia [41].

Por lo tanto, los canales de calcio representan un blanco importante para varios agen-

tes antiepilépticos. Por ejemplo, la eficacia de etosuximida y zonisamida en el tratamiento de las crisis de ausencia generalizadas es atribuida al bloqueo de los canales LVA [40]. También se sospecha que lamotrigina, levetiracetam, topiramato, fenobarbital, felbamato, gabapentina, pregabalina y el valproato pueden actuar mediante el bloqueo de estos canales en mayor o menor medida [53].

Canales de potasio dependientes de voltaje.

Los canales de potasio dependientes de voltaje son los principales responsables de la repolarización de la membrana celular durante la finalización del potencial de acción, limitando así la excitabilidad neuronal [55]. Se han encontrado mutaciones en varios genes de canales de potasio en epilepsias genéticas humanas y en animales [55]. Están formados por un tetrámero de subunidades α que contienen el poro, el selector iónico y el sistema de activación; y por otras subunidades accesorias [43].

Los canales de potasio se clasifican en varias familias: los canales de potasio voltaje operados formados por 6 segmentos transmembrana (K_v), los canales de potasio tipo rectificador entrante (K_{IR}), los canales de potasio sensibles a la concentración de calcio intracelular (K_{Ca}) y los canales de potasio formados por 2 dominios (K_{2P}) [43]. En particular, se ha comprobado que el fármaco retigabina, aprobado para el tratamiento de la epilepsia focal refractaria en 2011, ejerce sus efectos antiepilépticos por la activación de la clase K_{v7} [56]. Esta activación hiperpolariza la membrana celular y reduce la excitabilidad de las neuronas.

1.6.4.2. Atenuación de la neurotransmisión excitatoria mediada por glutamato.

El glutamato es el principal neurotransmisor excitatorio en el sistema nervioso central, ejerciendo sus efectos en receptores ionotrópicos y metabotrópicos ubicados en la membrana celular. Los receptores ionotrópicos responden a la unión del glutamato incrementando la conductancia a un catión y se diferencian entre sí de acuerdo a su especificidad

a agonistas selectivos [43]. Por lo tanto, se subdividen en receptores del ácido α-amino-3–hidroxi-5-metil-4-isoxazolpropiónico (AMPA), del kainato y del N-metil-D-aspartato (NMDA).

Los receptores AMPA y kainato están involucrados en la transmisión sináptica excitatoria rápida, y son permeables tanto a iones sodio como potasio, pero no al calcio. La apertura del canal del receptor AMPA genera la despolarización de la membrana, predominantemente por el flujo intracelular de sodio. En contraste, los receptores NMDA son permeables al sodio, al potasio y al calcio, y son solo activados durante períodos largos de despolarización, como los que se producen durante las descargas epileptiformes [57]. Los receptores metabotrópicos glutamatérgicos están asociados a proteínas G y actúan predominantemente regulando canales iónicos. Algunos estudios parecen indicar que mutaciones en estos receptores podrían estar involucradas en el desarrollo de epilepsia [43].

Los receptores glutamatérgicos ionotrópicos forman una estructura tetramérica. Cada una de las familias de receptores tiene un conjunto de subunidades especificas: cuatro en el caso del receptor AMPA (GluR1-GluR4), cinco para el receptor Kainato (GluR5-7, KA1-2) y siete para el NMDA (NR1, NR2A-D, NR3A-B) [43]. Cada subunidad presenta una composición modular, incluyendo un dominio extracelular amino terminal (ATD, amino-terminal domain) que participa en el ensamblado y modulación del receptor; un dominio de unión de ligandos, constituido a su vez por dos lóbulos (L1 y L2); tres segmentos conectores (S1-M1, S3-M3 y S2-M4); un dominio transmembrana que está compuesto por tres segmentos transmembrana (M1 – 4, donde M3 forma el poro) y un dominio C-terminal citoplasmático [58]. Los agonistas se unen al lóbulo L1 y así el lóbulo L2 se mueve sobre este. Este movimiento se traslada, a través de los segmentos conectores, al dominio transmembrana generando un cambio conformacional que permite la apertura del poro y el flujo de iones (figura 1.9) [57]. Perampanel y la droga experimental talampanel, son los únicos fármacos antiepilépticos actuales con efecto selectivo sobre los receptores de glutamato [57]. Son antagonistas no competitivos del receptor AMPA, pues se unen a un sitio extracelular distinto al del glutamato, ubicado sobre los segmentos conectores S1-M1

y S2-M4 [57]. Esta unión induce un cambio conformacional en el receptor que limita su habilidad para trasladar la respuesta de la unión del agonista a la apertura del canal. De esta forma se reduce la excitabilidad neuronal, bloqueando la descarga epiléptica.

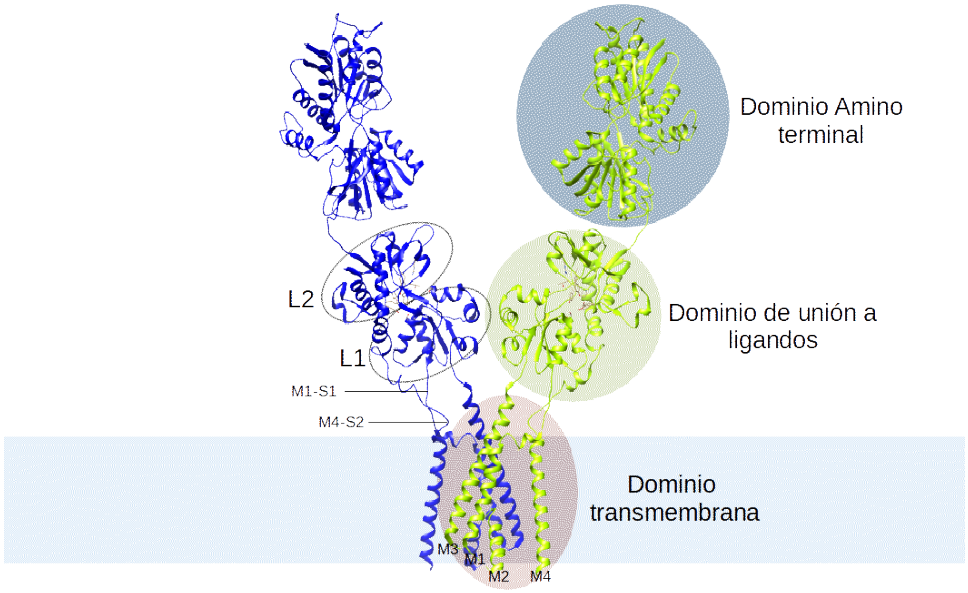


Figura 1.9: Estructura de los receptores ionotrópicos glutamatérgicos. Para una mayor simplicidad solo se representan dos subunidades.

Adicionalmente se ha encontrado que otros antiepilépticos ejercen su efecto, en parte, a través del bloqueo de la neurotransmisión glutamatérgica. Así, el bloqueo del receptor NMDA parece contribuir a la actividad del felbamato, mientras que el topiramato parece bloquear a receptores kainato y AMPA. Por su parte, el fenobarbital aparentemente inhibe al receptor AMPA [41]. Debe mencionarse que además de la inhibición directa de sus receptores, otros mecanismos de acción de los fármacos antiepilépticos pueden inhibir la neurotransmisión excitatoria del glutamato. Por ejemplo la lamotrigina parece reducir la liberación de glutamato mediante la inhibición de los canales de sodio y calcio presinapticos, y levetiracetam se une a la proteína 2A (SV2A) de la vesícula sináptica interfiriendo la exocitosis del neurotransmisor [43].

1.6.4.3. Activación de la neurotransmisión mediada por el acido γ-aminobutirico (GABA).

El ácido γ-aminobutírico (GABA) es el principal neurotransmisor inhibitorio en el sistema nervioso central. Se origina por decarboxilación de glutamato por acción de la enzima glutamato-descarboxilasa. Luego de su liberación, el GABA actúa en los receptores GABA-A y GABA-B produciendo la hiperpolarización de la membrana celular y, por lo tanto, un efecto inhibitorio [59].

El receptor GABA-A es un receptor ionotrópico dependiente de ligando. Está compuesto de 5 subunidades proteicas organizadas alrededor de un canal aniónico central que al activarse permite el pasaje de cloruro al interior celular. Se han identificado 19 subunidades posibles para la construcción del receptor (α 1-6, β 1-3, γ 1-3, δ, ϵ, θ, π, y ρ 1-3) codificadas por 19 genes diferentes. Cada subunidad contiene 4 segmentos transmembrana.

Los segmentos extracelulares forman los sitios de unión de GABA (dos por receptor) y en algunos casos se puede encontrar un sitio de unión para benzodiacepinas, las cuales actúan como moduladores alostéricos (figura 1.10) [43]. Los receptores GABA-A sensibles a las benzodiacepinas son responsables de la inhibición postsináptica de tipo "fásico", esto es, dependiente de las concentraciones de GABA liberado. Los receptores GABA-A que no son sensibles a las benzodiacepinas, se denominan "extrasinápticos", debido a que se encuentran alejados de la sinapsis y son responsables de una inhibición postsináptica de tipo "tónico" [43]. Es decir, serían responsables de la inhibición neuronal cuando exista, por ejemplo, un exceso de GABA, ya sea porque éste no es recaptado por los transportadores presinápticos o porque no es destruido por las enzimas pertinentes. La reducción de la eficacia de los receptores GABA-A por la presencia de agentes bloqueantes como por ejemplo, bicuculina o pentilenetetrazol, puede generar crisis convulsivas [40].

El receptor GABA-B, en cambio, es un receptor asociado a proteínas G que activa un canal de potasio. Se encuentran en la membrana presinaptica en las sinapsis GABA-

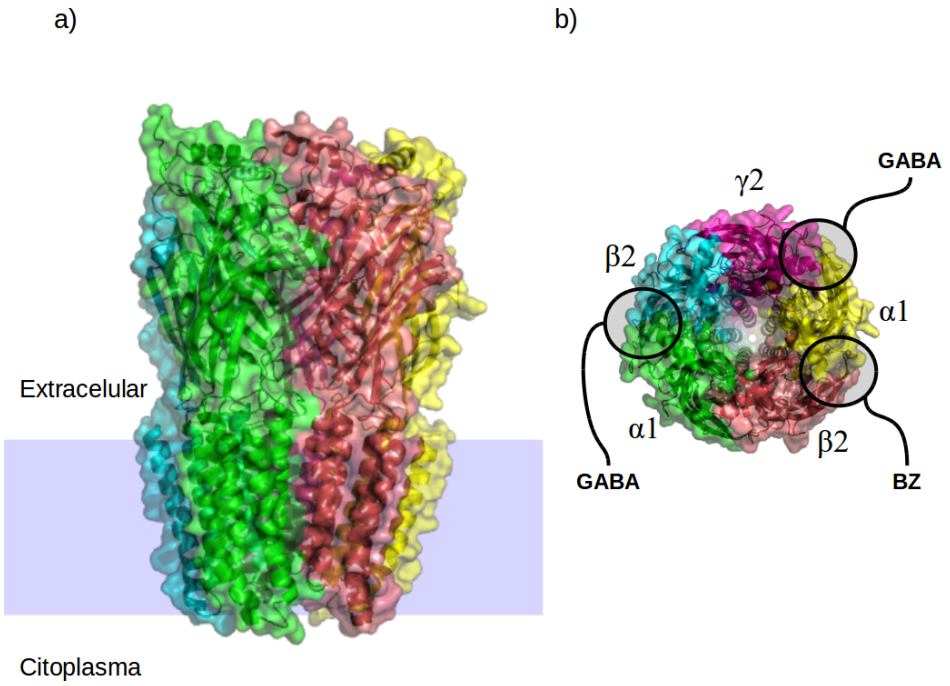


Figura 1.10: Estructura del receptor GABA-A humano. a) Vista lateral. b) Vista desde el espacio extracelular, donde se indica la composición más común, dada por dos subunidades α 1, dos β 2 y una γ 2 [60]. (BZ: benzodiacepinas). Código PDB: 4COF [61].

érgicas y gluatamatergicas, donde actúan reduciendo la liberación de neurotransmisor. También se expresan en la membrana postsináptica de la sinapsis GABA-érgica, donde producen una hiperpolarización tardía [43].

La concentración de GABA en el cerebro es controlada por la enzima glutamato descarboxilasa (GAD), que cataliza la conversión de glutamato a GABA; por la GABA-transaminasa (GABA-AT) que convierte el GABA a aldehído succínico y por la succínico aldehído deshidrogenasa que degrada este ultimo a succinato. Además, la acción del GABA es rápidamente terminada por la remoción desde el espacio sináptico hacia las terminales nerviosas y hacia células de la glía por los transportadores GAT-1, GAT-2 y GAT-3. De esta forma el GABA es reutilizado en la neurotransmisión o inactivado en las mitocondrias por la enzima GABA-AT [43].

La activación de los receptores GABA-A es uno de los mecanismos más importantes

de los fármacos antiepilépticos. Los barbitúricos (como fenobarbital y primidona) y las benzodiacepinas (diazepam, clonazepam, clobazam, etc.) comparten este efecto, aunque al parecer mediante la unión a sitios distintos del receptor [40]. Otros antiepilépticos, como el topiramato, el felbamato, el stiripentol y el levetiracetam pueden modular la actividad del receptor GABA-A [53]. El γ-vinil GABA (vigabatrin) actúa mediante la inhibición irreversible de la GABA transaminasa (GABA-AT), originando mayor cantidad de GABA disponibles para la descarga sináptica. La tiagabina, en cambio, inhibe a GAT-1, y reduce la captación neuronal y glial del mismo [43]. La gabapentina promueve la liberación de GABA desde el terminal presináptico [41].

1.7. Epilepsia y farmacorresistencia.

A pesar del exitoso descubrimiento de nuevos fármacos antiepilépticos con mejor perfil farmacocinético y, en algunos casos, con nuevos mecanismos de acción, la eficacia de los tratamientos farmacológicos para el control de la epilepsia no ha mejorado sustancialmente a través de los años. Se estima que un 30 % de los pacientes aún no logra controlar adecuadamente sus crisis mediante el uso de fármacos [15]. De acuerdo al consenso internacional, la epilepsia refractaria es descripta como un fracaso en lograr un paciente libre de crisis con dos o más fármacos antiepilépticos, siempre que estos sean los apropiados al tipo de epilepsia, administrados correctamente y no retirados por intolerancia [62].

Las crisis refractarias afectan la calidad de vida del paciente, ya que el desarrollo de crisis no controladas se asocia a un mayor riesgo de accidentes, lesiones físicas, trastornos psiquiátricos, deterioro cognitivo y de la memoria, así como consecuencias sociales [17]. Resulta evidente, entonces, que la resistencia a los fármacos antiepilépticos actuales representa una de las mayores preocupaciones de los especialistas en el área, y el descubrimiento de nuevas estrategias terapéuticas para el tratamiento de la epilepsia refractaria constituye uno de los principales objetivos a alcanzar por dichos investigadores. En esta dirección, la identificación de los mecanismos involucrados en la farmacorresistencia

resulta importante para el posterior diseño racional de nuevos fármacos con enfoques farmacológicos novedosos.

1.7.1. Mecanismos de farmacorresistencia.

Los mecanismos biológicos subyacentes a la epilepsia farmacorresistente no han sido completamente dilucidados aun, pero se proponen al momento cinco mecanismos principales para el desarrollo de la misma: La hipótesis del blanco molecular [63, 64], la hipótesis de la red neuronal [65], la hipótesis de la variación genética [66], la hipótesis de la gravedad intrínseca [67] y la hipótesis del transportador [68, 64]. Debe aclararse que la farmacorresistencia es un fenómeno multifactorial y que con estos cinco mecanismos ha sido posible explicar casos particulares de epilepsia farmacorresistente, pero ninguno es suficiente para dar una explicación general a este problema [69].

La *hipótesis del blanco molecular* propone que los fármacos anticonvulsivos fallan debido a alteraciones adquiridas sobre los canales iónicos y/o los receptores de neurotransmisores sobre los que actúan. Las crisis epilépticas, al parecer, ocasionan cambios en la transcripción de las subunidades de los receptores, así como también cambios postranscripcionales como glicosilaciones, fosforilaciones, etc.. Se han encontrado evidencias de que estos cambios probablemente afecten tanto la densidad sobre la membrana plasmática como la funcionalidad de los mismos [64, 70].

La *hipótesis de la red neuronal* sugiere que episodios recurrentes de actividad neuronal excesiva pueden inducir eventos que generan un remodelamiento de la red neuronal tales como degeneración neuronal, necrosis, gliosis, ramificación axonal y reorganización sináptica. Esta nueva red neuronal originada bajo condiciones patológicas podría estar relacionada con el fenómeno de farmacorresistencia [65].

La *hipótesis genética* sugiere que la resistencia a antiepilépticos es gobernada por variaciones genéticas de las proteínas involucradas en la farmacocinética y la farmacodinamia de estos. En este caso solo se consideran aquellas variaciones en la funcionalidad

y/o estructura producto del polimorfismo genético, dejándose de lado las alteraciones adquiridas [66].

La *hipótesis de la severidad intrínseca* sugiere que la farmacorresistencia es una propiedad inherente de la epilepsia relacionada a la gravedad de la enfermedad. Esta hipótesis pone de manifiesto observaciones epidemiológicas que indican que a mayor frecuencia de crisis epilépticas antes de iniciado el tratamiento mayor es el riesgo de sufrir farmacorresistencia [67]. El concepto general es que las crisis que se disparan fácilmente pueden repercutir en crisis frecuentes posteriores difíciles de controlar. Si bien los factores moleculares o neurobiológicos de esta hipótesis son aun desconocidos, un aspecto esencial de esta idea es que tanto la severidad de la epilepsia como la farmacorresistencia podrían encontrarse bajo el control de factores neurobiológicos comunes [71].

Por último, la *hipótesis del transportador* propone que la resistencia a los fármacos se produce debido a su incapacidad de acceder al blanco molecular en concentraciones suficientes como consecuencia de la hiperactividad de transportadores de eflujo de drogas presentes en la barrera hematoencefálica [15]. Este mecanismo de farmacorresistencia es uno de los más explorados y corroborados experimentalmente (junto con la hipótesis del blanco molecular). En la siguiente sección de este apartado se dará una revisión más detallada de esta hipótesis, dado que fue considerada para el diseño de anticonvulsivos en este trabajo de tesis.

1.7.2. Hipótesis del transportador.

Como se indicó anteriormente, la hipótesis del transportador propone que la farmacorresistencia se produce debido a la incapacidad del fármaco de acceder al receptor en concentraciones suficientes, y que este efecto se produce como consecuencia de la hiperactividad de transportadores de eflujo presentes en la barrera hematoencefálica.

La barrera hematoencefálica es una compleja interface entre la sangre y el sistema nervioso central que controla el intercambio entre ambos compartimientos, jugando un

rol esencial en la homeostasis y en la protección frente a agentes tóxicos y patógenos. Esta barrera se encuentra formada por las células endoteliales de los capilares nerviosos, las cuales se caracterizan por la formación de uniones estrechas (con escasa permeabilidad) entre ellas. Además de las células endoteliales que constituyen el elemento central, la barrera se halla compuesta por otras células como los astrocitos, neuronas y pericitos perivasculares (figura 1.11) [72].

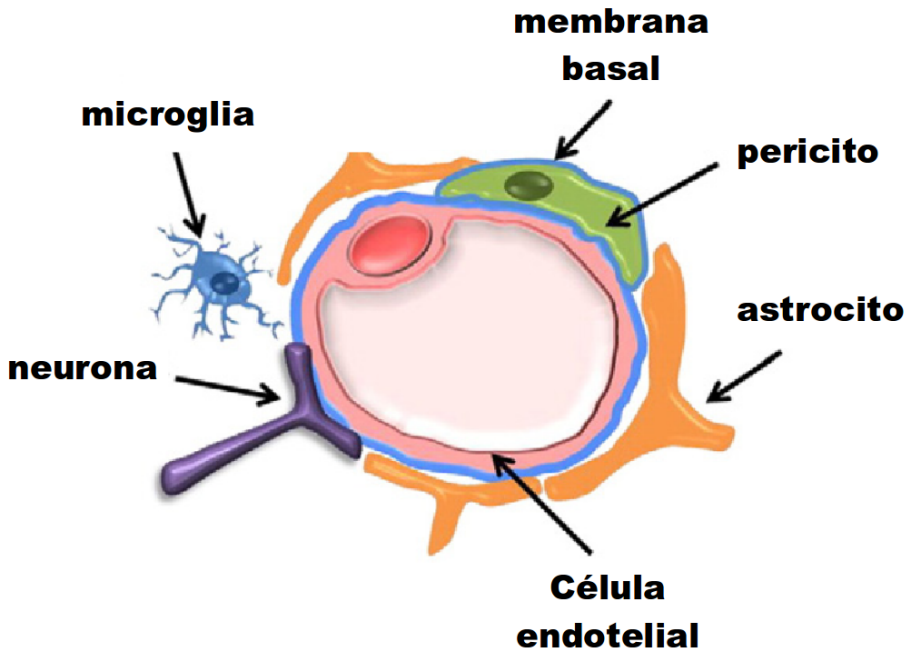


Figura 1.11: Esquema de la unidad neurovascular. Adaptado, con autorización, de Cardoso y col. [72].

Las células endoteliales expresan proteínas transportadoras, las cuales proveen rutas de entrada a nutrientes, iones y algunas macromoléculas; y también rutas de expulsión de metabólitos o xenobióticos potencialmente tóxicos [72]. Algunos de los transportadores encargados de la eliminación de sustancias están involucrados en la expulsión de fármacos. Este fenómeno, conocido como resistencia múltiple a drogas, juega un papel importante en la falla del tratamiento farmacológico de numerosos desórdenes, tales como enfermedades infecciosas o procesos tumorales [73].

A menudo la remoción de agentes terapéuticos es mediada por una gran superfamilia de proteínas denominadas proteínas de resistencia múltiple a fármacos, muchas de las cuales corresponden a la familia de proteínas transportadoras ABC, dependientes de adenosín trifosfato (ATP). Los genes encargados de su expresión se dividen en subfamilias según la organización de los dominios transmembrana y la homología de secuencia (ABCA, ABCB, ABCC, ABCD, ABCE, ABCF y ABCG) [73]. Entre los diferentes miembros que forman esta superfamilia, la más estudiada es la glicoproteína P (P-gp, codificada por el gen ABCB1), seguida de las proteínas relacionadas con la resistencia múltiple a fármacos (MRP 1, 2, 4 y 5; genes ABCC1-5) y la proteína relacionada con el cáncer de mama (BCRP; genes ABCG2) [71].

1.7.2.1. Glicoproteína P.

La glicoproteína P (P-gp) es uno de los transportadores de eflujo más intensamente estudiados debido a su inespecificidad de sustratos. Esta macromolécula es capaz de interaccionar con fármacos estructuralmente diferentes como inhibidores de la proteasa del HIV, antibióticos, inmunosupresores, entre otros [73]. Está compuesta por una única cadena polipeptídica de 1280 aminoácidos de longitud y aproximadamente 170 kDa, que se dispone de forma transversal a la membrana plasmática, formando dos mitades homólogas. Cada una de estas mitades, a su vez, está formada por un dominio transmembrana constituido por seis hélices transmembrana (TM1-TM6 y TM7-TM12) y un dominio citosólico (NBD1 y NBD2), denominado dominio de unión a nucleótidos (Figura 1.12 a). Los dominios transmembrana forman una cavidad hidrofóbica de alrededor de 6000 $\mathring{A}^3$, la cual alberga una gran región donde interactúan los sustratos [74].

El mecanismo por el cual esta proteína realiza la expulsión de las sustancias implica el cambio entre dos conformaciones diferentes, en cuyo pasaje de una a otra se encuentra involucrada la hidrólisis de ATP. La interacción de una sustancia con la P-gp, ocurre cuando ésta se encuentra en una conformación orientada hacia el interior celular [74]. Cuando un sustrato de la P-gp ingresa a la célula, es reconocido y retenido por el

transportador en el sitio de unión. Como consecuencia de esta interacción, la P-gp sufre un cambio conformacional para adoptar la forma orientada hacia el exterior celular, llevándose a cabo la expulsión del fármaco hacia el espacio extracelular [74] (Figura 1.12 b).

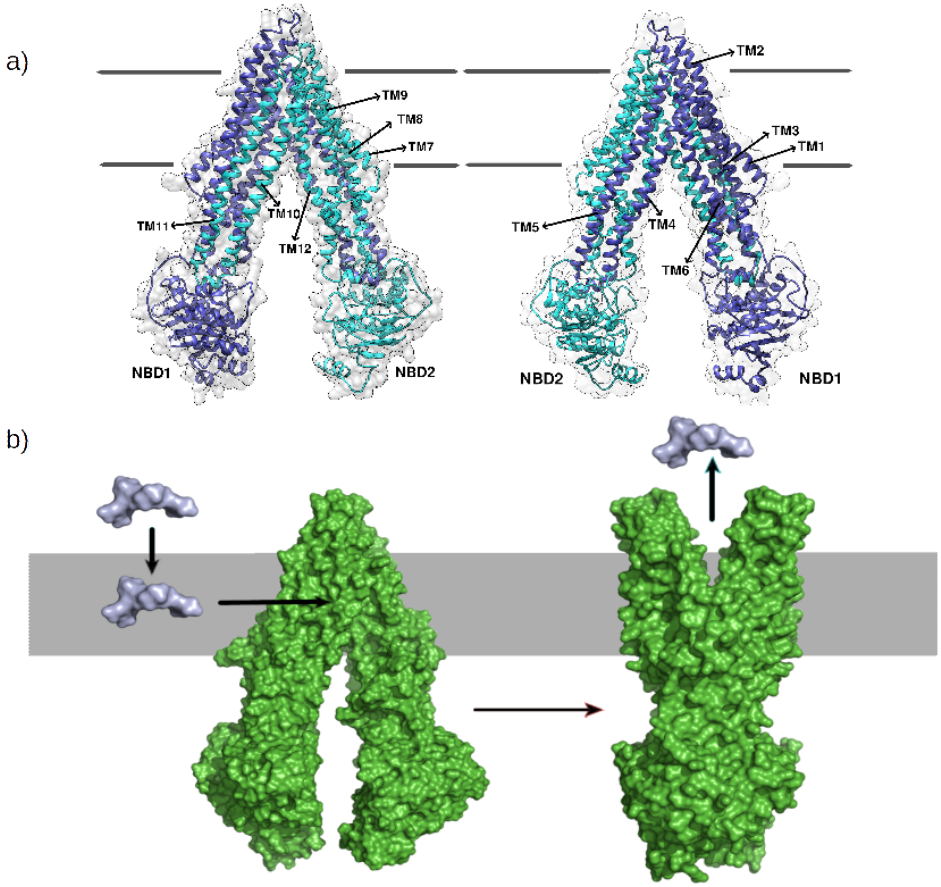


Figura 1.12: Glicoproteína P. a) Estructura y posición en la membrana plasmática de la P-gp. TM: hélices transmembrana, NBD: Dominios de unión a nucleótidos. b) Mecanismo de expulsión de sustancias de la P-gp.

La P-gp se encuentra expresada en diferentes tejidos con capacidad excretora como el intestino delgado, hígado y riñón, donde actúa limitando la entrada de los fármacos en el cuerpo luego de la administración oral o promoviendo la eliminación hacia la bilis y la orina. También se encuentra en barreras presentes entre la sangre y algunos tejidos, como son la barrera hematoencefálica, la placenta y la barrera hematotesticular, limitando el ingreso a tejidos sensibles [75, 73].

En relación a su rol en la epilepsia refractaria, diversos estudios indican que varios fármacos antiepilépticos son sustrato o inhibidores de la P-gp [76, 71]. Por otra parte, estudios en pacientes farmacorresistentes han demostrado sobreexpresión de P-gp respecto a pacientes epilépticos sin fármaco resistencia [77, 78, 79, 68]. Además, a través de métodos inmunohistoquímicos fue detectada la presencia de P-gp en neuronas de pacientes resistentes, células en las que normalmente no se hallan presentes [68, 80]. Otro resultado interesante proviene del estudio realizado en ratones modificados genéticamente para no expresar la P-gp (knockout), donde se observa un incremento de los niveles de fármacos antiepilépticos en el cerebro respecto de los animales no modificados [81].

Del conjunto de los hallazgos enumerado previamente se desprende que la P-gp juega un rol importante en la epilepsia fármaco resistente. Sin embargo, aún no se conoce con certeza la causa de la sobreexpresión de esta y otras proteínas de la superfamilia ABC en el tejido cerebral de los pacientes con epilepsia refractaria. Se ha propuesto que puede ser adquirida principalmente como consecuencia de crisis convulsivas repetitivas, o de condiciones patológicas como inflamación o hipoxia [82].

Por otra parte, la liberación excesiva del neurotransmisor glutamato es un factor típico de las crisis epilépticas, lo que llevó a algunos investigadores a sospechar de su influencia en la expresión de P-gp. Así, utilizando capilares cerebrales de ratones incubados con glutamato, se observó que este neurotransmisor activa la transcripción genómica del gen que codifica la P-gp, con el consiguiente incremento de la expresión del transportador [83]. El mecanismo propuesto para esta acción parece estar mediado por los receptores NMDA. La activación de estos receptores genera un incremento del calcio intracelular que activa varias vías enzimáticas en neuronas. Una de estas enzimas, la fosfolipasa A2, provoca la liberación de acido araquidónico desde la membrana plasmática, activando su metabolismo mediante la acción de la ciclooxigenasa 2 (COX-2) que produce prostanoides (prostaglandinas, prostaciclinas y tromboxanos) [84]. La prostaglandina E2 es el mayor producto de la acción de la COX-2 en el cerebro y actúa sobre cuatro receptores nucleares acoplados a proteínas G (EP1-4), de los cuales se ha demostrado que el receptor EP1

juega un rol clave en los niveles de expresión de la P-gp [85] (Figura 1.13).

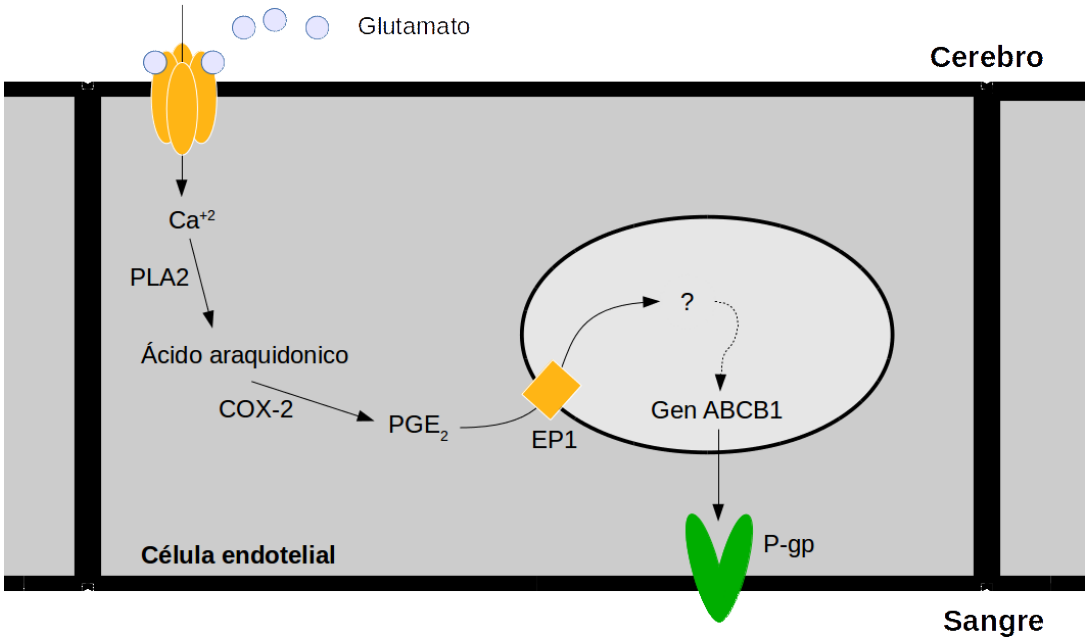


Figura 1.13: Mecanismo de sobreexpresión de P-gp en las crisis epilépticas. Adaptado de Potschka [84]

El reconocimiento temprano de compuestos que eviten interaccionar con la P-gp representa una estrategia útil para el diseño de nuevos fármacos anticonvulsivos tendientes a satisfacer la demanda de pacientes resistentes. Por otro lado, la identificación de estructuras con capacidad de interaccionar fuertemente con la P-gp puede ser implementada para el desarrollo de inhibidores para terapias adyuvantes. Esta última estrategia tiene serias limitaciones, dado que implicaría cancelar la función fisiológica de esta proteína (como primera línea de defensa en varias barreras) y requiere el diseño de compuestos altamente selectivos por la glicoproteína (para evitar la aparición de efectos adversos). Es por eso que en este trabajo de tesis se optará como criterio adicional de selección de candidatos a anticonvulsivos a aquellos compuestos que se predigan no sustratos de P-gp por cribado virtual.

Bibliografía

[1] R. S. Fisher, W. Van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, "Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[2] V. Jewells, H. W. Shin, H. Eldad, T. Fisher, and A. Hinn, "Review of Epilepsy - Etiology, Diagnostic Evaluation and Treatment," *International Journal of Neurorehabilitation*, vol. 01, no. 03, pp. 1–8, 2014.

[3] W.H.O, "Epilepsy," *Fact Sheet*, no. 999, 2015.

[4] R. S. Fisher, C. Acevedo, A. Arzimanoglou, A. Bogacz, J. H. Cross, C. E. Elger, J. Engel, L. Forsgren, J. a. French, M. Glynn, D. C. Hesdorffer, B. I. Lee, G. W. Mathern, S. L. Moshé, E. Perucca, I. E. Scheffer, T. Tomson, M. Watanabe, and S. Wiebe, "ILAE Official Report: A practical clinical definition of epilepsy.," *Epilepsia*, vol. 55, pp. 475–82, apr 2014.

[5] D. Hirtz, D. J. Thurman, K. Gwinn-Hardy, M. Mohamed, a. R. Chaudhuri, and R. Zalutsky, "How common are the çommon"neurologic disorders?," *Neurology*, vol. 68, no. 5, pp. 326–337, 2007.

[6] M. O. Melcon, S. Kochen, and R. H. Vergara, "Prevalence and clinical features of epilepsy in Argentina: A community-based study," *Neuroepidemiology*, vol. 28, no. 1, pp. 8–15, 2007.

[7] A. T. Berg, S. F. Berkovic, M. J. Brodie, J. Buchhalter, J. H. Cross, W. Van Emde Boas, J. Engel, J. French, T. a. Glauser, G. W. Mathern, S. L. Moshé, D. Nordli, P. Plouin, and I. E. Scheffer, "Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009," *Epilepsia*, vol. 51, no. 4, pp. 676–685, 2010.

[8] S. L. Moshé, E. Perucca, P. Ryvlin, and T. Tomson, "Epilepsy: new advances," *The Lancet*, vol. 385, no. 9971, pp. 884–898, 2015.

[9] R. H. Mattson, "Overview: idiopathic generalized epilepsies.," *Epilepsia*, vol. 44 Suppl 2, pp. 2–6, 2003.

[10] E. A. Felton, "The Generalized Epilepsies: Description, Pathophysiology, Treatment, and Prognosis," *THE NEUROLOGY REPORT*, pp. 11–18, 2015.

[11] P. Genton and M. Bureau, "Epilepsy with Myoclonic Absences," *CNS drugs*, vol. 20, no. 11, pp. 911–916, 2006.

[12] S. Noachtar and A. S. Peters, "Semiology of epileptic seizures: A critical review," *Epilepsy & Behavior*, vol. 15, no. 1, pp. 2–9, 2009.

[13] W. T. Blume, H. O. Lüders, E. Mizrahi, C. Tassinari, W. Van Emde Boas, and J. Engel J., "Glossary of descriptive terminology for ictal semiology: Report of the ILAE Task Force on classification and terminology," *Epilepsia*, vol. 42, no. 9, pp. 1212–1218, 2001.

[14] B. F. Shneker and N. B. Fountain, "Epilepsy," *Disease-a-Month*, vol. 49, pp. 426–478, jul 2003.

[15] W. Löscher, H. Klitgaard, R. E. Twyman, and D. Schmidt, "New avenues for anti-epileptic drug discovery and development.," *Nature reviews. Drug discovery*, vol. 12, pp. 757–76, oct 2013.

[16] C. M. DeGiorgio and S. E. Krahl, "Neurostimulation for drug-resistant epilepsy.," *Continuum (Minneapolis, Minn.)*, vol. 19, no. 3 Epilepsy, pp. 743–55, 2013.

[17] K. D. Laxer, E. Trinka, L. J. Hirsch, F. Cendes, J. Langfitt, N. Delanty, T. Resnick, and S. R. Benbadis, "The consequences of refractory epilepsy and its treatment," *Epilepsy & Behavior*, vol. 37, pp. 59–70, 2014.

[18] R. G. Levy, P. N. Cooper, P. Giri, and J. Pulman, "Ketogenic diet and other dietary treatments for epilepsy," *Cohocrane Epilepsi*, no. 3, p. 27, 2012.

[19] K. J. Bough and J. M. Rho, "Anticonvulsant Mechanisms of the Ketogenic Diet," *Epilepsia*, vol. 48, no. 1, pp. 43–58, 2007.

[20] M. Yudkoff, Y. Daikhin, T. M. Melø, I. Nissim, U. Sonnewald, and I. Nissim, "The ketogenic diet and brain metabolism of amino acids: relationship to the anticonvulsant effect.," *Annual review of nutrition*, vol. 27, no. 13, pp. 415–30, 2007.

[21] E. H. Kossoff, B. a. Zupec-Kania, P. E. Amark, K. R. Ballaban-Gil, a. G. Christina Bergqvist, R. Blackford, J. R. Buchhalter, R. H. Caraballo, J. Helen Cross, M. G. Dahlin, E. J. Donner, J. Klepper, R. S. Jehle, H. D. Kim, Y. M. Christiana Liu, J. Nation, D. R. Nordli, H. H. Pfeifer, J. M. Rho, C. E. Stafstrom, E. a. Thiele, Z. Turner, E. C. Wirrell, J. W. Wheless, P. Veggiotti, E. P. G. Vining, A. Anderson, B. Cohen, M. Currey, D. Donley, L. Dure, B. Garg, M. Goldstein, B. Grabert, D. Griesemer, E. Kovnar, R. Larson, A. Legido, L. A. Morrison, C. Parker, J. B. Renfroe, J. Schreck, S. Shinnar, R. Snyder, C. Tardo, G. D. Timmons, and G. Yim,

"Optimal clinical management of children receiving the ketogenic diet: Recommendations of the International Ketogenic Diet Study Group," *Epilepsia*, vol. 50, no. 2, pp. 304–317, 2009.

[22] W. Löscher and D. Schmidt, "Modern antiepileptic drug development has failed to deliver: ways out of the current dilemma.," *Epilepsia*, vol. 52, pp. 657–78, apr 2011.

[23] C. Locock, "ROYAL MEDICAL & CHIRURGICAL SOCIETY. TUESDAY, MAY 11TH, 1857.," 1857.

[24] A. Hauptmann, "Luminal bei epilepsie," *Munch Med Wochenschr*, vol. 59, pp. 1907–1909, 1912.

[25] E. Magiorkinis, A. Diamantis, K. Sidiropoulou, and C. Panteliadis, "Highights in the History of Epilepsy: The Last 200 Years," *Epilepsy Research and Treatment*, vol. 2014, pp. 1–13, 2014.

[26] H. H. Merritt and T. J. Putnam, "Sodium diphenyl hydantoinate in the treatment of convulsive disorders," *Journal of the American Medical Association*, vol. 111, no. 12, pp. 1068–1073, 1938.

[27] H. Merrit and T. Putnam, "Sodium diphenyl hydantoinate in the treatment of convulsive seizures. toxic symptoms and their prevention," *Arch Neurol Psychiatry*, vol. 42, no. 6, pp. 1053–1058, 1939.

[28] H. H. MERRITT and T. J. PUTNAM, "A new series of anticonvulsant drugs tested by experiments on animals," *Archives of Neurology & Psychiatry*, vol. 39, no. 5, pp. 1003–1015, 1938.

[29] J. P. Stables and H. J. Kupferberg, "The NIH Anticonvulsant Drug Development (ADD) Program: preclinical anticonvulsant screening project," in *Molecular and cellular targets for antiepileptic drugs* (G. Avanzani, G. Regesta, P. Tanganelli, and M. Avoli, eds.), ch. 16, pp. 191–198, 1997.

[30] R. K. Richards and G. M. Everett, "Tridione: a new anticonvulsant drug," *J Lab Clin Med*, vol. 31, no. 6, pp. 1330–1336, 1946.

[31] L. S. Goodman, W. C. Brown, E. A. Swinyard, and M. S. Grewal, "Comparison of maximal seizures evoked by pentylene-tetrazol (Metrazol) and electroshock in mice, and their modification by anticonvulsants," *Journal of Pharmacology and Experimental Therapeutics*, vol. 108, no. 2, pp. 168–176, 1953.

[32] S. Livingston and L. Pauli, "CELONTIN®(PM-396) IN THE TREATMENT OF EPILEPSY," *Pediatrics*, vol. 19, no. 4, pp. 614–618, 1957.

[33] S. Karceski, M. J. Morrell, and D. Carpenter, "Treatment of epilepsy in adults: Expert opinion, 2005," *Epilepsy and Behavior*, vol. 7, no. SUPPL. 1, pp. 1–64, 2005.

[34] R. J. Porter, J. J. Cereghino, G. D. Gladding, B. J. Hessie, H. J. Kupferberg, B. Scoville, and B. G. White, "Antiepileptic drug development program," *Cleve Clin Q*, vol. 51, no. 2, pp. 293–305, 1984.

[35] H. Meunier, G. Carraz, Y. Neunier, P. Eymard, and M. Aimard, "[pharmacodynamic properties of n-dipropylacetic acid.]," *Therapie*, vol. 18, pp. 435–438, 1962.

[36] P. J. Loiseau, "Clinical experience with new antiepileptic drugs: antiepileptic drugs in Europe.," *Epilepsia*, vol. 40 Suppl 6, pp. S3–S8; discussion S73–S74, 1999.

[37] G. H, "Propriétés antiépiléptiques exceptionelles et méconnues d'un anxiolytique de commerce: le clobazam.," *Concours Med.*, vol. 100, pp. 3697–701, 1978.

[38] P. Kwan and M. J. Brodie, "Early identification of refractory epilepsy," *New England Journal of Medicine*, vol. 342, no. 5, pp. 314–319, 2000.

[39] T. A. Glauser, A. Cnaan, S. Shinnar, D. G. Hirtz, D. Dlugos, D. Masur, P. O. Clark, E. V. Capparelli, and P. C. Adamson, "Ethosuximide, Valproic Acid, and Lamotrigine in Childhood Absence Epilepsy," *New England Journal of Medicine*, vol. 362, pp. 790–799, mar 2010.

[40] M. a. Rogawski and W. Löscher, "The neurobiology of antiepileptic drugs.," *Nature reviews. Neuroscience*, vol. 5, pp. 553–64, jul 2004.

[41] H. S. White, M. D. Smith, and K. S. Wilcox, "Mechanisms of action of antiepileptic drugs.," *International review of neurobiology*, vol. 81, pp. 85–110, jan 2007.

[42] D. Schmidt and S. Schachter, "Drug treatment of epilepsy in adults," *BMJ: British Medical Journal*, pp. 1–18, 2014.

[43] B. S. Meldrum and M. a. Rogawski, "Molecular targets for antiepileptic drug development," *Neurotherapeutics*, vol. 4, pp. 18–61, jan 2007.

[44] F. H. Yu, V. Yarov-Yarovoy, G. a. Gutman, and W. a. Catterall, "Overview of molecular relationships in the voltage-gated ion channel superfamily.," *Pharmacological reviews*, vol. 57, no. 4, pp. 387–395, 2005.

[45] J. Payandeh, T. Scheuer, N. Zheng, and W. a. Catterall, "The crystal structure of a voltage-gated sodium channel.," *Nature*, vol. 475, pp. 353–8, jul 2011.

[46] S. K. Bagal, B. E. Marron, R. M. Owen, R. I. Storer, and N. A. Swain, "Voltage gated sodium channels as drug discovery targets," *Channels*, vol. 9, pp. 360–366, nov 2015.

[47] W. a. Catterall and T. M. Swanson, "Structural Basis for Pharmacology of Voltage-Gated Sodium and Calcium Channels," *Molecular Pharmacology*, vol. 88, pp. 141–150, jun 2015.

[48] M. N. Rasband, "The axon initial segment and the maintenance of neuronal polarity," *Nature Reviews Neuroscience*, vol. 11, no. 8, pp. 552–562, 2010.

[49] W. a. Catterall, "Voltage-gated sodium channels at 60: structure, function and pathophysiology.," *The Journal of physiology*, vol. 590, pp. 2577–89, jun 2012.

[50] A. Nardi, N. Damann, T. Hertrampf, and A. Kless, "Advances in targeting voltage-gated sodium channels with small molecules.," *ChemMedChem*, vol. 7, pp. 1712–40, oct 2012.

[51] W. Catterall, A. Goldin, and S. Waxman, "International Union of Pharmacology. XLVII. Nomenclature and structure-function relationships of voltage-gated sodium channels," *Pharmacological reviews*, vol. 57, no. 4, pp. 397–409, 2005.

[52] M. Leuwer, G. Haeseler, H. Hecker, J. Bufler, R. Dengler, and J. K. Aronson, "An improved model for the binding of lidocaine and structurally related local anaesthetics to fast-inactivated voltage-operated sodium channels, showing evidence of cooperativity.," *British journal of pharmacology*, vol. 141, pp. 47–54, jan 2004.

[53] M. J. Brodie, "Antiepileptic drug therapy the story so far," *Seizure*, vol. 19, no. 10, pp. 650–655, 2010.

[54] J. Noebels, M. Avoli, M. Rogawski, R. Olsen, and A. Delgado-Escueta, eds., *Jasper's basic mechanisms of the epilepsies*, ch. Voltage-Gated Calcium Channels in Epilepsy. Oxford University Press, 2012.

[55] J. Noebels, M. Avoli, M. Rogawski, R. Olsen, and A. Delgado-Escueta, eds., *Jasper's basic mechanisms of the epilepsies*, ch. Potassium Channels (including KCNQ) and Epilepsy. Oxford University Press, 2012.

[56] L. J. Stephen and M. J. Brodie, "Pharmacotherapy of epilepsy: newly approved and developmental agents," *CNS drugs*, vol. 25, no. 2, pp. 89–107, 2011.

[57] M. a. Rogawski, "AMPA receptors as a molecular target in epilepsy therapy.," *Acta neurologica Scandinavica. Supplementum*, vol. 127, no. 197, pp. 9–18, 2013.

[58] A. I. Sobolevsky, M. P. Rosconi, and E. Gouaux, "X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor.," *Nature*, vol. 462, no. 7274, pp. 745–56, 2009.

[59] P. Kwan, G. J. Sills, and M. J. Brodie, "The mechanisms of action of commonly used antiepileptic drugs.," *Pharmacology & therapeutics*, vol. 90, no. 1, pp. 21–34, 2001.

[60] L. J. Greenfield, "Molecular mechanisms of antiseizure drug activity at GABAA receptors," oct 2013.

[61] P. S. Miller and A. R. Aricescu, "Crystal structure of a human GABAA receptor," *Nature*, vol. 512, no. 7514, pp. 270–275, 2015.

[62] P. Kwan, A. Arzimanoglou, A. T. Berg, M. J. Brodie, W. Allen Hauser, G. Mathern, S. L. Moshé, E. Perucca, S. Wiebe, and J. French, "Definition of drug resistant epilepsy: Consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies," *Epilepsia*, vol. 51, no. 6, pp. 1069–1077, 2009.

[63] D. Schmidt and W. Löscher, "Drug resistance in epilepsy: putative neurobiologic and clinical mechanisms.," *Epilepsia*, vol. 46, no. 6, pp. 858–77, 2005.

[64] S. Remy, "Molecular and cellular mechanisms of pharmacoresistance in epilepsy," *Brain*, vol. 129, no. 1, pp. 18–35, 2005.

[65] M. Fang, Z. Q. Xi, Y. Wu, and X. F. Wang, "A new hypothesis of drug refractory epilepsy: Neural network hypothesis," *Medical Hypotheses*, vol. 76, no. 6, pp. 871–876, 2011.

[66] W. Löscher, U. Klotz, F. Zimprich, and D. Schmidt, "The clinical impact of pharmacogenetics on the treatment of epilepsy," *Epilepsia*, vol. 50, no. 1, pp. 1–23, 2009.

[67] M. a. Rogawski, "The intrinsic severity hypothesis of pharmacoresistance to antiepileptic drugs.," *Epilepsia*, vol. 54 Suppl 2, pp. 33–40, 2013.

[68] P. Kwan and M. J. Brodie, "Potential role of drug transporters in the pathogenesis of medically intractable epilepsy," *Epilepsia*, vol. 46, no. 2, pp. 224–235, 2005.

[69] D. Schmidt and W. Löscher, "New developments in antiepileptic drug resistance: an integrative view," *Epilepsy Currents*, vol. 9, no. 2, pp. 47–52, 2009.

[70] M. A. Rogawski and M. R. Johnson, "Intrinsic severity as a determinant of antiepileptic drug refractoriness.," *Epilepsy currents / American Epilepsy Society*, vol. 8, no. 5, pp. 127–30, 2008.

[71] J. Noebels, M. Avoli, M. Rogawski, R. Olsen, and A. Delgado-Escueta, eds., *Jasper's basic mechanisms of the epilepsies*, ch. Drug Resistance. Oxford University Press, 2012.

[72] F. L. Cardoso, D. Brites, and M. A. Brito, "Looking at the blood-brain barrier: molecular anatomy and possible investigation approaches.," *Brain research reviews*, vol. 64, pp. 328–63, sep 2010.

[73] W. Löscher and H. Potschka, "Role of drug efflux transporters in the brain for drug disposition and treatment of brain diseases.," *Progress in neurobiology*, vol. 76, pp. 22–76, may 2005.

[74] S. G. Aller, J. Yu, A. Ward, Y. Weng, S. Chittaboina, R. Zhuo, P. M. Harrell, Y. T. Trinh, Q. Zhang, I. L. Urbatsch, and G. Chang, "Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding.," *Science (New York, N.Y.)*, vol. 323, pp. 1718–22, mar 2009.

[75] K. M. Giacomini, S. Huang, Tweedie DJ, and L. Z. Benet, "Membrane transporters in drug development," *Nature*, vol. 9, no. 3, pp. 215–236, 2012.

[76] C. Zhang, P. Kwan, Z. Zuo, and L. Baum, "The transport of antiepileptic drugs by P-glycoprotein," dec 2012.

[77] S. M. Dombrowski, S. Y. Desai, M. Marroni, L. Cucullo, K. Goodrich, W. Bingaman, M. R. Mayberg, L. Bengez, and D. Janigro, "Overexpression of multiple drug resistance genes in endothelial cells from patients with refractory epilepsy.," *Epilepsia*, vol. 42, no. 12, pp. 1501–1506, 2001.

[78] S. Sisodiya, W.-r. Lin, M. Squier, and M. Thom, "Multidrug-resistance protein 1 in focal cortical dysplasia," *The Lancet*, vol. 357, pp. 42–43, 2001.

[79] D. M. Tishler, K. I. Weinberg, D. R. Hinton, N. Barbaro, G. M. Annett, and C. Raffel, "MDR1 gene expression in brain of patients with medically intractable epilepsy.," *Epilepsia*, vol. 36, no. 1, pp. 1–6, 1995.

[80] A. Lazarowski, L. Czornyj, F. Lubienieki, E. Girardi, S. Vazquez, and C. D'Giano, "ABC transporters during epilepsy and mechanisms underlying multidrug resistance in refractory epilepsy.," *Epilepsia*, vol. 48 Suppl 5, pp. 140–9, jan 2007.

[81] M. Rizzi, S. Caccia, G. Guiso, C. Richichi, J. a. Gorter, E. Aronica, M. Aliprandi, R. Bagnati, R. Fanelli, M. D'Incalci, R. Samanin, and A. Vezzani, "Limbic seizures induce P-glycoprotein in rodent brain: functional implications for pharmacoresistance.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 22, no. 14, pp. 5833–5839, 2002.

[82] A. Lazarowski, L. Czornyj, and L. Rocha, "The complexity of roles of P-glycoprotein in refractory epilepsy: Pharmacoresistance, epileptogenesis, SUDEP and relapsing marker after surgical treatment," *Admet & Dmpk*, vol. 3, no. 2, pp. 110–121, 2015.

[83] B. Bauer, A. M. S. Hartz, A. Pekcec, K. Toellner, D. S. Miller, and H. Potschka, "Seizure-Induced Up-Regulation of P-Glycoprotein at the Blood-Brain Barrier through Glutamate and Cyclooxygenase-2 Signaling," *Molecular Pharmacology*, vol. 73, pp. 1444–1453, feb 2008.

[84] H. Potschka, "Targeting regulation of ABC efflux transporters in brain diseases: A novel therapeutic approach," *Pharmacology and Therapeutics*, vol. 125, no. 1, pp. 118–127, 2010.

[85] A. Pekcec, B. Unkru, J. Schlichtiger, J. Soerensen, A. M. S. Hartz, E. A. V. Vliet, J. A. Gorter, and H. Potschka, "Targeting Prostaglandin E 2 EP1 Receptors Prevents Seizure- Associated P-glycoprotein Up-Regulation," *Pharmacology*, vol. 330, no. 3, pp. 939–947, 2009.

Capítulo 2

Métodos computacionales para el descubrimiento de fármacos

2.1. Introducción al descubrimiento de fármacos.

Tradicionalmente, el descubrimiento de fármacos se ha realizado mediante una perspectiva con bases químicas o farmacológicas [1]. Gran parte del hallazgo de compuestos líderes se ha basado en la observación y el análisis de la medicina folklórica, de ligandos naturales, de segundos usos de drogas existentes o, simplemente, por serendipia [2]. También se han utilizado muy frecuentemente estudios de cribado de bibliotecas de compuestos de origen natural o sintético, los cuales implican un costo económico considerable [2]. Sin embargo, en las últimas décadas se ha producido una gran evolución en el descubrimiento de nuevos fármacos, dándole a este proceso un enfoque multidisciplinario que combina áreas tan diversas como la química, la biología, la biofísica, la informática y la matemática [3].

En líneas generales, el proceso actual de descubrimiento de nuevos fármacos se encuentra dividido en siete etapas básicas: 1 - Selección del blanco molecular a emplear; 2 - Validación de dicho blanco; 3 - Identificación de un compuesto líder; 4 - Optimización del líder; 5 - Ensayos a nivel pre clínico; 6 - Ensayos a nivel clínico; 7- seguimiento post-comercialización [4, 5] (figura 2.1).

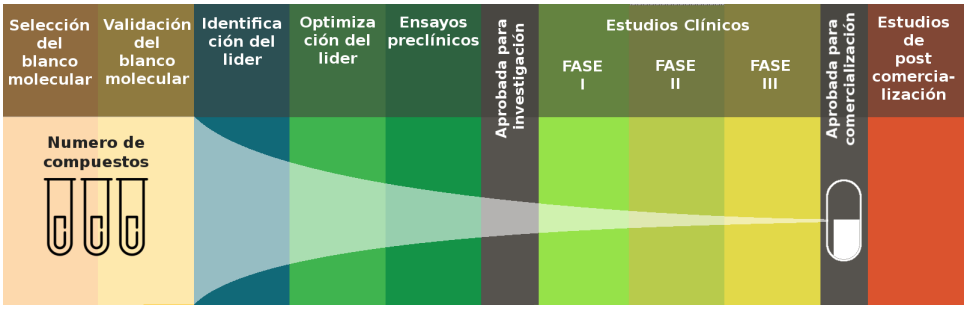


Figura 2.1: Esquema general del proceso de descubrimiento y desarrollo de fármacos.

La etapa inicial consiste en la identificación del sitio molecular especifico con el cual se pretende que interactúe la molécula candidato [6]. El término blanco molecular puede aplicarse a una amplia variedad de entidades biológicas, tales como los receptores de membrana, los canales iónicos, las enzimas, las proteínas de vías de señalización, los receptores nucleares y los genes o proteínas intervinientes en la regulación genética [6].

La elección de un blanco particular depende del nivel de conocimiento respecto a su influencia en la enfermedad. Además, debe ser eficaz, seguro y satisfacer las necesidades clínicas. Particularmente debe tenerse en cuenta en esta etapa que la macromolécula sea accesible para el fármaco, provocando una respuesta medible, ya sea in vivo como in vitro [7]. De acuerdo a literatura, el tiempo estimado para la esta etapa es de aproximadamente un año y presenta un costo de alrededor de 200 millones de dólares [8].

Una vez seleccionado el blanco molecular, la etapa siguiente corresponde a su validación. Esto implica demostrar la relevancia de la macromolécula en el proceso de la enfermedad estudiada, y que su modulación tiene efectos terapéuticos. La validación es abordada mediante el empleo de diversas técnicas, incluyendo estudios in vitro o modelos en animales. Entre las herramientas más utilizadas se encuentran los modelos genéticos en animales (knock in / ganancia de funciones; knock out / perdida de funciones), anticuerpos monoclonales, utilización de transgenes, etc [9, 7]. Por otra parte, debe mencionarse que algunos potenciales blancos más innovadores tienen un menor grado de validación, como por ejemplo una relación genética con la enfermedad o solo especulación como consecuencia de conocimiento aproximado [3]. Esta etapa suele tener un costo estimado de

250 millones de dólares y una duración aproximada de dos años [8].

La siguiente fase da inicio a la búsqueda de estructuras químicas adecuadas para interaccionar con el blanco terapéutico seleccionado. Esta etapa consiste en la identificación de compuestos que presentan la actividad biológica buscada (por interacción con el blanco molecular previamente determinado). Posteriormente, se seleccionan las estructuras activas que además tengan otras propiedades favorables, tales como baja toxicidad, alta permeabilidad de membrana, solubilidad etc. Estos compuestos se denominan líderes [5]. Se estima, en esta etapa, un costo de 40 millones de dólares y una duración menor al año [10].

Una vez identificada la estructura líder, comienza el proceso de optimización, que consiste en la realización de modificaciones estructurales sobre dicho compuesto con el objetivo de mejorar su eficacia terapéutica. De esta manera se busca un incremento de potencia frente a su blanco molecular, selectividad frente a otras estructuras relacionadas, mejoras en la farmacocinética, minimización de la toxicidad y efectos secundarios [11]. Como este proceso involucra la optimización simultanea de múltiples parámetros, es uno de los pasos que más tiempo y dinero consume (120 millones de dólares y 2,7 años de duración) [10]. Es por eso que se define como un paso limitante en el descubrimiento de fármacos [8].

Como resultado de las etapas mencionadas hasta aquí se dispone entonces de un conjunto de compuestos que presentan la actividad buscada. Estos son sometidos a una serie de evaluaciones preclínicas, entre las que se encuentran el desarrollo de métodos de síntesis a gran escala, estudios de seguridad en animales, carcinogénesis, estudios ADME (administración, distribución, metabolismo y excreción), entre otros [9].

Finalizados los ensayos preclínicos, el compuesto pasa a la fase de ensayos clínicos. Este periodo es considerado como el más costoso de todo el proceso (260 millones de dólares y 7 años de duración) y se subdivide en cuatro etapas: En la fase 1 se evalúa la seguridad y tolerancia del fármaco, sus propiedades farmacocinéticas y farmacodinámicas. Estos ensayos son realizados sobre voluntarios sanos. En la fase 2 se repiten estos estudios,

pero sobre pacientes que presentan la enfermedad a ser tratada. En la siguiente fase se compara el nuevo tratamiento con la terapia estándar sobre un gran número de pacientes. Estos estudios son multicéntricos, con ensayos doble ciego y al azar. Luego del éxito de esta última etapa el nuevo tratamiento es autorizado para su comercialización [5]. Durante la última fase, se evalúa la presencia de efectos adversos en etapa de comercialización.

2.2. Métodos computacionales en el descubrimiento de fármacos.

Desde la década del 90, la técnica de cribado de alto rendimiento (HTS, por sus siglas en ingles: high throughput screening) en conjunto con la química combinatoria ha permitido la evaluación biológica de un gran número de compuestos en forma automática. Sin embargo, la tasa de éxito del HTS para encontrar estructuras líderes es relativamente baja, restringiendo su uso a grandes librerías de moléculas [12]. Esta limitación, junto con el elevado costo de estas técnicas, propició la búsqueda de alternativas al HTS, potenciando otros métodos de diseño de fármacos como son aquellos asistidos por computadora [12]. Se conoce como métodos computacionales aplicados al descubrimiento de fármacos (Computer-aided drug design; CADD) al conjunto de herramientas y recursos computacionales para el almacenamiento, manejo, análisis y modelado de compuestos [13]. Esto incluye, por ejemplo, el desarrollo de programas para el diseño de estructuras químicas con determinadas características fisicoquímicas, el estudio virtual de interacciones químicas entre moléculas, o la generación de herramientas de análisis sistemático de potenciales candidatos líderes antes de su síntesis y ensayo [13].

En líneas generales los métodos computacionales son utilizados con tres propósitos principales en el proceso de descubrimiento de nuevos fármacos:

1. filtrar grandes librerías de compuestos en búsqueda de estructuras predichas como activas; que puedan ser luego evaluadas por métodos experimentales. Esta metodo-

logía se conoce como cribado virtual (en ingles virtual high throughput screening (vHTS) o simplemente virtual screening);

2. dirigir la optimización de compuestos lideres, ya sea mejorando su afinidad, potenciando su selectividad u optimizando su metabolismo y farmacocinética;

3. diseñar nuevos compuestos, ya sea mediante el "crecimiento" de una estructura de partida por adición o modificación de grupos funcionales o por la combinación de fragmentos pequeños [12].

En el desarrollo de la presente tesis se han utilizado métodos computacionales con la intención de filtrar compuestos activos mediante el cribado virtual, por lo que a continuación se describirá brevemente la teoría y los principios que sostienen estas herramientas.

2.2.1. Cribado virtual.

Uno de los usos más comunes de los métodos CADD es el cribado virtual de librerías de compuestos [12]. Esta metodología consiste en puntuar, ordenar, y luego filtrar un conjunto de estructuras mediante el uso de algoritmos computacionales [14]. Básicamente, el cribado virtual intenta mejorar la probabilidad de identificar moléculas activas, lo que permite enfocar los recursos en la evaluación de compuestos que probablemente presenten la actividad buscada.

Los métodos de cribado virtual se clasifican en dos categorías: los basados en el conocimiento de un conjunto de ligandos que presentan la misma acción biológica (también llamados métodos indirectos o ligand-based) y los que se fundamentan en la estructura tridimensional del receptor (métodos directos o structure-based) [14]. A continuación se explicaran ambas herramientas, haciendo mayor hincapié en los métodos directos ya que son los que se implementaron en esta tesis.

2.2.1.1. Cribado virtual basado en ligandos.

Los métodos de cribado virtual basados en el ligando requieren de un conjunto de compuestos que interaccionan con el blanco molecular de interés, para la realización de un análisis retrospectivo de sus estructuras. El objetivo final es representar estos compuestos de tal manera que sus características más importantes para la actividad biológica sean retenidas y codificadas, mientras que la información no relevante sea descartada [12]. Estas herramientas se clasifican en tres grandes grupos: las metodologías basadas en el uso de descriptores, de la similitud química, o en la superposición con farmacóforos.

Metodología basada en descriptores:

Los métodos QSAR (del inglés, quantitative structure activity relationships) consisten en obtener una correlación cuantitativa entre la actividad biológica y ciertas propiedades fisicoquímicas, estructurales y/o de carácter teórico de los ligandos [15]. Estas propiedades se denominan descriptores y codifican las características de las estructuras a diferentes niveles de complejidad [12].

El proceso general de los métodos QSAR inicia con la creación de un conjunto de compuestos activos e inactivos (conjunto de entrenamiento) y el cálculo de sus descriptores moleculares. Posteriormente se genera una ecuación matemática capaz de relacionar alguno de estos descriptores con la actividad experimental. Finalmente el modelo es validado y aplicado a una librería de compuestos (representados por los mismos descriptores), de modo de predecir cuál de ellos presentará la actividad buscada [12].

El éxito de los métodos QSAR, por lo tanto, depende no solo del conjunto de activos/inactivos seleccionados para generar el modelo sino también de la selección de descriptores y de la relación matemática apropiada. Una de las limitaciones más importantes de este método es que todos los modelos obtenidos van a depender del espacio de muestreo del conjunto de entrenamiento (también denominado dominio de aplicabilidad de la función QSAR) [12].

Metodología basada en la similitud química:

Las técnicas basadas en la similitud química consisten en medir el grado de similitud entre una o varias estructuras de referencia y los compuestos problema. El principio detrás de esta metodología radica en considerar que si una molécula es estructuralmente similar a un compuesto de actividad biológica conocida, entonces la primera tiene mayor probabilidad de ser activa (respecto a otra con menor grado de similitud) [16].

La medida de similaridad requiere de tres componentes: una representación para caracterizar las moléculas a ser comparadas; un sistema de ponderación para asignar diferentes grados de importancia a los distintos componentes de estas representaciones; y un coeficiente que es utilizado para determinar el grado de relación entre las dos representaciones estructurales [16]. Este valor por lo general oscila entre 0 y 1, siendo 1 el resultado para moléculas idénticas y 0 para entidades totalmente disimiles [17]. La comparación entre dos moléculas se realiza mediante el uso de secuencias de bits derivadas de las estructuras. En tales secuencias (denominadas fingerprints o huellas dactilares) cada bit indica la presencia o ausencia de una propiedad estructural determinada, normalmente una subestructura. Alternativamente, los bits pueden indicar la frecuencia con la que aparece la propiedad en las moléculas comparadas. Los bits compartidos por ambas moléculas son cuantificados y transformados en un valor que indica cuán similares son dichas estructuras (coeficiente de similitud) [17].

Esta metodología presenta un bajo costo computacional, lo que permite la búsqueda en grandes bases de datos en un tiempo relativamente corto. Además, la búsqueda puede llevarse a cabo incluso con una sola molécula de referencia [18]. Sin embargo, estas técnicas comparten el inconveniente encontrado en los métodos QSAR: las moléculas activas utilizadas inicialmente para la selección, limitan los resultados de la búsqueda [15].

Metodología basada en superposición con un farmacóforo:

Según la IUPAC, el farmacóforo se define como el "conjunto de propiedades estéricas y electrónicas necesarias para asegurar las interacciones supramoleculares óptimas con una

estructura biológica determinada y activar (o bloquear) su respuesta biológica" [19]. Más allá de esta definición, puede considerarse al farmacóforo como el arreglo espacial de los grupos funcionales (y/o otros fragmentos estructurales) que debe contener un compuesto para generar la respuesta biológica deseada. Por lo tanto, un farmacóforo efectivo debe brindar información respecto de qué porciones de la molécula activa interactúan con el blanco molecular, y debe ser capaz de proponer interacciones no covalentes entre los compuestos y la macromolécula [12].

Un modelo farmacofórico puede obtenerse mediante aproximaciones fundadas en el conocimiento de la estructura del receptor o basadas en los ligandos. Esta última consiste en extraer las propiedades químicas comunes de los compuestos activos que son esenciales para la actividad buscada.

En general, la obtención de un modelo farmacofórico a partir de múltiples ligandos activos (conjunto de entrenamiento) implica un análisis del espacio conformacional para cada uno de ellos, a fin de considerar su flexibilidad. Posteriormente debe realizarse un alineamiento de las conformaciones consideradas activas para establecer las características comunes relevantes que permitan generar el modelo farmacofórico [20].

Una vez definido el farmacóforo, puede realizarse el cribado a fin de encontrar nuevos candidatos que cumplan con los requerimientos propuestos por el modelo. Esta metodología presenta un alto costo computacional, por lo que su utilización en bases de datos muy grandes puede consumir mucho tiempo. Nuevamente el éxito del modelo depende del conjunto de entrenamiento seleccionado para generar el patrón farmacofórico [21].

2.2.1.2. Cribado virtual basado en la estructura del receptor.

Los métodos basados en la estructura del receptor requieren de modelos tridimensionales de la estructura de la macromolécula blanco (generalmente una proteína), para posteriormente evaluar su capacidad de interacción con pequeñas moléculas. La hipótesis central de esta aproximación es que la habilidad de un compuesto para ejercer un efecto

biológico determinado depende de su capacidad para interactuar favorablemente en un sitio particular de un blanco específico [12]. Por lo tanto, las moléculas que comparten estas interacciones deberían presentar efectos biológicos similares. En base a esta hipótesis es posible obtener nuevos compuestos con una acción buscada mediante el análisis del sitio de unión en la proteína. El docking molecular, capaz de simular el acoplamiento proteína-ligando desde el punto de vista estructural y energético, es una de las herramientas disponibles más utilizadas en esta metodología [22].

La información relacionada con la estructura tridimensional de los blancos moleculares generalmente se obtiene a partir de métodos experimentales, como la cristalografía por rayos X o la Resonancia Magnética Nuclear (RMN), entre otros [12]. El Protein Data Bank (PDB) [23], creado en 1977, alberga en la actualidad algo más de 114000 estructuras obtenidas por estos métodos experimentales, aunque muchas de ellas corresponden a diversas formas cristalinas de una misma macromolécula, (con lo que el número de macroestructuras diferentes es mucho menor). Muchas de estas estructuras juegan un rol vital en vías metabólicas críticas y pueden funcionar como blancos terapéuticos. Además, la base de datos PDB alberga complejos ligando-receptor, lo que proporciona una información estructural muy valiosa respecto del origen de las interacciones intermoleculares.

Si bien el número de estructuras depositadas en PDB para macromoléculas aumenta en forma rápida y continua con el paso del tiempo (figura 2.2), no resulta raro encontrar que las proteínas experimentalmente resueltas no pueden satisfacer por completo las necesidades en el campo del descubrimiento de fármacos. En ausencia de coordenadas espaciales de los átomos que definen la arquitectura del blanco molecular, una aproximación alternativa que puede utilizarse es la predicción de dicha estructura [12]. A partir de la secuencia de aminoácidos de la macromolécula de interés, pueden emplearse técnicas de modelado comparativo (o por homología), threading o también métodos Ab-Initio [24].

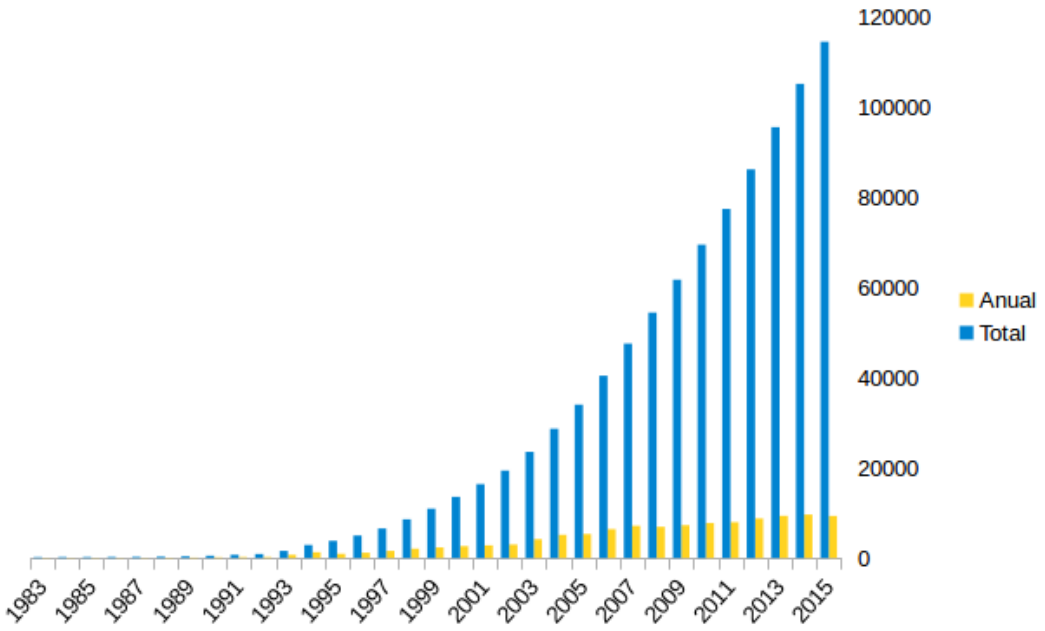


Figura 2.2: Evolución del contenido de la base de datos Protein Data-Base (PDB). En amarillo se indica el número de estructuras depositadas por año; en azul el número de estructuras totales (último acceso: 20/01/2016)

Modelado de proteínas.

Cuando no se dispone de la estructura del receptor, se pueden construir modelos tridimensionales teóricos, normalmente obtenidos por modelado comparativo, threading, o por métodos Ab-Initio [24]. Estas herramientas se inspiran en el trabajo realizado por Christian Anfinsen y colaboradores [25], quienes observaron que una cadena aminoácidica de ribonucleasa desnaturalizada en solución alcanza el estado nativo en un tiempo razonable. Estos experimentos sentaron las bases para demostrar que: (1) una proteína puede ser identificada únicamente por su secuencia de aminoácidos; (2) esta secuencia es la responsable de codificar la funcionalidad in-vivo de la proteína y (3) la secuencia de aminoácidos contiene la información necesaria para que la proteína adquiera su forma funcional, basándose exclusivamente en las interacciones intramoleculares y con el solvente que la rodea.

El modelado por homología es una metodología basada en la comparación de la secuencia de la proteína cuya estructura se quiere modelar (secuencia problema) con otra

(u otras) cuya estructura es conocida (secuencia molde). En esta técnica la proteína problema y la que actúa como molde se consideran homologás [26]. Debe tenerse en cuenta que la homología es una inferencia o una conclusión respecto de la existencia de una relación ancestral entre la secuencia problema y la secuencia molde; obtenida a partir de la comparación de las mismas [27]. Para llegar a esta conclusión inicialmente se debe realizar un alineamiento entre ambas, que permita establecer su parecido. Los parámetros a considerar durante la comparación son similitud e identidad, los cuales derivan de la observación del alineamiento. La similitud de secuencia hace referencia al porcentaje de residuos alineados que tienen propiedades fisicoquímicas similares, por lo que pueden ser más fácilmente sustituibles entre ellos. La identidad de secuencia se refiere al porcentaje de aminoácidos idénticos alineados entre las 2 secuencias. Cuando un alineamiento revela una similitud significativa entre un grupo de secuencias, estas pueden considerarse como homólogas. Si se conoce la estructura y función de una de estas secuencias "molde", esta información puede ser transferida a aquellas homólogas cuya arquitectura tridimensional no ha sido determinada aún.

El método de comparación de plegamientos, o threading, consiste en la asignación de las coordenadas del esqueleto carbonado de una estructura molde a la secuencia problema, y la medición de la plausibilidad del modelo tridimensional generado mediante el cálculo de una serie de parámetros empíricos [28]. Mediante esta herramienta se pueden identificar proteínas con estructura conservada sin que presenten un origen evolutivo común [28]. El fundamento de esta metodología es que, al momento, el número de plegamientos proteicos conocidos es muy reducido comparado con los millones de secuencias proteicas. En consecuencia, muchas proteínas pueden adoptar plegamientos similares incluso en ausencia de similitud en la secuencia [27]. Esto admite el desarrollo de métodos computacionales para predecir la estructura proteica más allá de relaciones de homología.

Como su nombre lo sugiere, los métodos de predicción Ab-Initio generan un modelo basado solo en la información obtenida a partir de la secuencia de aminoácidos de la proteína cuya estructura se quiere determinar. Los programas de predicción generalmente

son diseñados utilizando los principios de la minimización energética: Estos algoritmos buscan todas las posibles conformaciones que puede adoptar la secuencia, para encontrar aquella que presenta el mínimo de energía global [28]. La ventaja que presenta este tipo de metodologías es que la predicción no está restringida a las estructuras tridimensionales conocidas para macromoléculas. Sin embargo, dado que las leyes fisicoquímicas que gobiernan el plegamiento proteico aún no son bien entendidas, las funciones de energía utilizadas al presente son inexactas [26].

A continuación se detalla el proceso de modelado comparativo por homología, por tratarse de uno de los más empleados. Además, esta metodología se empleó en esta investigación para el modelado del blanco molecular (canal de sodio Nav1.2 en estado abierto) que fue finalmente utilizado en el cribado virtual de bibliotecas de compuestos.

Modelado por homología. Procedimientos generales.

El procedimiento general para el desarrollo del modelado por homología incluye las siguientes etapas: (1) Identificación de proteínas relacionadas a la problema que puedan ser utilizadas como molde; (2) Alineamiento de secuencias entre las proteínas molde y problema; (3) construcción del modelo tridimensional basado en las regiones alineadas y en la predicción de coordenadas de átomos no alineados; (4) refinamiento del modelo y evaluación del mismo (figura 2.3) [29].

1. **Identificación de proteínas molde:** Inicialmente la secuencia del blanco molecular es usada como patrón de búsqueda en bases de datos de secuencias con estructura tridimensional conocida (PDB). En esta etapa los moldes con alta similitud de secuencia son identificados por métodos de búsqueda del tipo heurísticos [30]. Estos algoritmos toman atajos para reducir el espacio de búsqueda, y de esta forma aumentar la velocidad del proceso. Por lo tanto, mediante estos métodos no se garantiza encontrar el mejor alineamiento entre la secuencia problema y las secuencias de la base de datos, ya que priorizan la rapidez del cálculo frente a otros algoritmos más exactos. Por otra parte los resultados deben presentar un

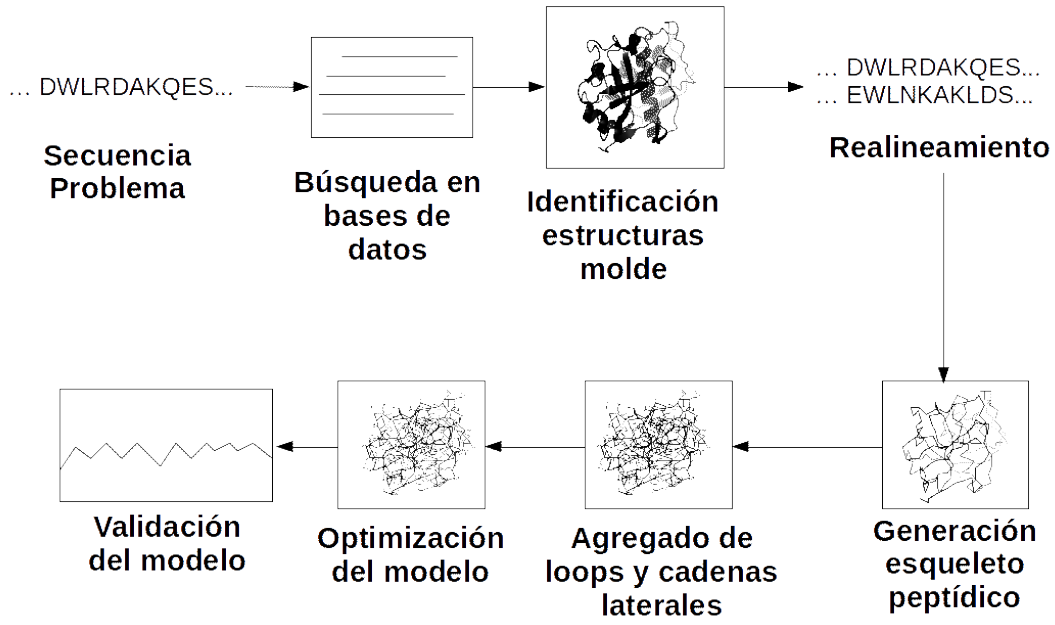


Figura 2.3: Diagrama de flujo del proceso de modelado comparativo.

balance entre sensibilidad (es decir, la capacidad de detectar el máximo número de verdaderos positivos), y especificidad (la capacidad de descartar los verdaderos negativos); de manera que se rechacen el máximo número posible de falsos positivos [26]. Los métodos heurísticos más utilizados son BLAST (Basic Local Alignment Search Tool) [31] y FASTA (Fast All) [30, 26].

Particularmente el programa BLAST mejora la velocidad de la búsqueda mientras mantiene una buena sensibilidad, a través del fraccionamiento de las secuencias. El primer paso entonces es crear una lista de posibles fragmentos de la secuencia problema y de las secuencias de la base de datos. Estos fragmentos se denominan "palabras" y típicamente se forman con 3 letras para proteínas (que codifican 3 aminoácidos), incluyendo todas las posibles combinaciones para cada secuencia.

El segundo paso es buscar en la base de datos secuencias que contengan la palabra buscada u otras similares. En esta etapa se cuantifica el grado de similitud entre las

palabras a comparar, mediante matrices que calculan la frecuencia de intercambio entre pares de aminoácidos (matrices de puntuación [32, 33]). En un tercer paso se seleccionan las palabras provenientes de secuencias de la base de datos que son lo suficientemente similares a las palabras de la secuencia problema. Esto se realiza mediante el análisis de la puntuación obtenida en la etapa anterior, eligiendo aquellas que superan un valor de corte. El cuarto paso consiste en la prolongación del alineamiento desde la palabra inicial en ambas direcciones de las secuencias. Esta extensión continúa hasta que la puntuación del alineamiento cae por debajo de un valor de corte (22 para proteínas) [34]. El segmento alineado resultante, originariamente sin gaps, es denominado HSP (segmento de alta puntuación o high scoring segment pair). En una nueva versión de BLAST el HSP se obtiene mediante la extensión del alineamiento en ambas direcciones usando métodos de alineamiento más exhaustivos, donde se permite la introducción de gaps [34]. La extensión continúa hasta que la puntuación se mantenga por encima de un valor de corte, de otra forma se termina. El puntaje, basado en matrices de puntuación, refleja el grado de similitud entre las secuencias [35, 36].

Luego de encontradas secuencias similares a la problema, surge una de las limitaciones más importantes del método, que es saber si las coincidencias entre aminoácidos son significativas, o si podrían deberse al azar (y por lo tanto no reflejar una relación de homología). En BLAST, se utiliza un indicador estadístico conocido como E-value (expectation value), que indica la probabilidad de que un alineamiento entre pares de secuencias resulte por efecto del azar. La expresión matemática de E-value está representada por la ecuación 2.1 [37]:

$$E = K \cdot m \cdot n \cdot e^{-\lambda \cdot S} \tag{2.1}$$

Donde m es el numero de residuos totales de la base de datos, n es el numero de residuos de la secuencia problema, K y λ son parámetros estadísticos determinados empíricamente y S es el valor de la puntuación del HSP. Mientras menor sea el

E-value, menor es la probabilidad de que el alineamiento sea por efecto del azar y por lo tanto más significativo es el apareamiento. La interpretación empírica para el E-value es [37]:

- $E < 1e^{-50}$ significa que hay una alta probabilidad que el apareamiento resulte en un homologo.
- Si $1e^{-50} < E < 0{,}01$ el apareamiento puede ser debido a homología.
- Si $0{,}01 < E < 10$ el apareamiento es no significativo, pero puede tratarse de un homologo remoto. Se necesitan más evidencias para inferir relación.
- Si $E > 10$ las secuencias no están relacionadas o su relación es muy distante y cae fuera de los limites de detección.

Debido a que E-value es proporcionalmente afectado por el tamaño de la base de datos, un problema observado es que cuando la base crece, el E-value para una secuencia dada también se incrementa.

El "bit score" es otro indicador estadístico usado en BLAST. Mide la similaridad de secuencia con independencia de la longitud de la secuencia problema y el tamaño de la base de datos. Es definido por la ecuación 2.2:

$$S' = \frac{\lambda \cdot S - \ln K}{\ln 2} \tag{2.2}$$

Por lo tanto, el bit score provee un indicador estadístico útil para realizar búsquedas en diferentes bases de datos [37]. Cuanto mayor es su valor, más significativo es el apareamiento entre secuencias.

Como regla general, una proteína de una base de datos tiene que tener al menos un 30 % de identidad de secuencia con la proteína problema para ser seleccionada. Ocasionalmente, a partir de un 20 % de identidad se considera aceptable y puede ser usada como molde. Debe tenerse en cuenta que cuanto mayor sea el número de aminoácidos que contienen las secuencias, mayor será la probabilidad de que

alcancen un determinado score por azar [38]. Cuando en la búsqueda, se encuentran múltiples estructuras con similitud significativa, es recomendable que se seleccionen aquellas con mayor porcentaje de identidad. Otro criterio a tener en cuenta es la resolución de la estructura tridimensional del molde.

2. **Alineamiento:** Una vez que la estructura con la mayor similaridad de secuencia es identificada como molde, toda la longitud de la secuencia patrón y la proteína problema necesitan ser realineadas usando algoritmos de alineamiento más exhaustivos. Este es el paso crítico en el modelado por homología, ya que afecta directamente al modelo final. Un alineamiento incorrecto puede llevar a una designación errónea de los residuos homólogos y, por lo tanto, a un modelo estructural inexacto. Los errores cometidos en esta etapa no pueden ser resueltos en las siguientes etapas.

   El objetivo final del alineamiento de secuencias es encontrar el mejor apareamiento entre las mismas, tal que haya la máxima correspondencia entre los aminoácidos. Maximizar el alineamiento significa encontrar las operaciones, es decir, sustituciones, deleciones e inserciones, que permitan transformar una secuencia en otra. A partir de una función *f*, se determina el peso de cada una de estas operaciones. Además, la probabilidad de que un residuo sea reemplazado por otro en dos secuencias a comparar se encuentra cuantificada en las matrices de sustitución. Las más comúnmente utilizadas son las matrices PAM [32] y BLOSUM [33], las cuales incluyen información de sustituciones que reflejan la posibilidad de una historia evolutiva común entre secuencias. A continuación se describen muy brevemente.

   Las matrices PAM (del inglés: point accepted mutation) fueron desarrolladas inicialmente por Margaret Dayhoff y se basan en el análisis de mutaciones permitidas en las secuencias de aminoácidos durante el proceso evolutivo [32]. En su desarrollo se utilizó una base de datos con 1572 mutaciones encontradas en los arboles filogenéticos de 71 familias de proteínas, donde cada familia consiste en un conjunto de macromoléculas muy relacionadas (con un porcentaje de identidad igual o mayor al 85 %) [32]. Estas matrices tienen un tamaño de 20 x 20, donde se coloca en

la primera fila y columna cada uno de los 20 aminoácidos esenciales. Las celdas en las que se interceptan cada par de aminoácidos corresponden a la probabilidad de intercambio de un aminoácido por otro. La matriz PAM1 correspondería a una divergencia de 1 % en la secuencia de una proteína respecto a otra (un aminoácido reemplazado por otro cada 100 residuos). Para obtener una matriz PAM que indique un proceso evolutivo más largo (un porcentaje de N mutaciones aceptadas en 100 aminoácidos), se multiplica la matriz PAM1 por sí misma N veces [32].

Las matrices BLOSUM (BLOcks SUbstitution Matriz) fueron creadas por Henikoff y Henikoff [33]. Estos investigadores derivaron las probabilidades de sustitución entre aminoácidos a partir de un conjunto de unos 2000 motivos conservados (blocks) encontrados en una base de datos de alrededor de 500 proteínas relacionadas. Para asignar la probabilidad de intercambio generaron n clusters de proteínas, de manera que todos aquellos motivos que presentan, por ejemplo, un 60 % de identidad, se agrupan para evaluar las probabilidades de mutación. De allí se deriva la correspondiente matriz BLOSUM60. A diferencia de las matrices PAM todas las matrices se calculan directamente y no se utilizan extrapolaciones [33]. Valores positivos corresponden a sustituciones que ocurren más frecuentemente, por el contrario, valores negativos indican sustituciones poco frecuentes [33].

Por otra parte, debe mencionarse que existen dos estrategias de alineamiento posibles: el alineamiento global y el local. En el alineamiento global, las dos secuencias a ser alineadas se asumen similares en cuanto a su tamaño. El proceso se lleva a cabo desde el principio hasta el final de las dos secuencias, para encontrar el mejor alineamiento posible. Este método es aplicado para alinear secuencias relacionadas y del mismo tamaño. Para secuencias más divergentes y de longitud variable, puede no generar resultados óptimos, ya que suele fallar en el reconocimiento de regiones locales con alta similitud.

Una estrategia general de alineamiento global es el algoritmo desarrollado por Needleman y Wunsch [39]. En una primera etapa se construye una matriz en cuyos ejes se colocan las secuencias a comparar. Posteriormente se van completando las celdas

de la matriz, comenzando por la primera fila de una de las secuencias, que es usada para escanear a través de toda la otra secuencia. Se coloca entonces en cada celda "(i,j)" un valor de puntuación resultante del análisis del apareamiento de los residuos ubicados en las posición "i" de una secuencia y "j" de la otra. La puntuación de cada celda, $F(i,j)$ está dada por la ecuación 2.3:

$$F(i,j) = max \begin{cases} F(i-1,j-1) + s(x,y) \\ F(i-1,j) - d \\ F(i,j-1) - d \end{cases} \tag{2.3}$$

Es decir, se elije el valor máximo entre tres números posibles relacionados. El valor *s(x, y)* corresponde al valor determinado por la matriz de sustitución utilizada para el par de aminoácidos *(x, y)* a analizar (BLOSUM62, por ejemplo) y d es la penalización que se decide adjudicar al alineamiento de un aminoácido de una de las secuencias con un hueco o gap. Como se muestra en la ecuación, la evaluación de la segunda fila tiene en cuenta los resultados obtenidos en la primera. Este proceso se repite hasta que se encuentran los valores de todas las celdas. El paso siguiente es encontrar el patrón que representa el alineamiento óptimo. Esto se realiza por el rastreo en orden inverso desde el borde inferior derecho de la matriz al borde superior izquierdo. El mejor alineamiento es el que tiene el máximo valor de puntuación [39].

El alineamiento local no asume que las secuencias sean similares en cuanto a su tamaño. Solo busca regiones locales con la mayor similitud y se alinean esas regiones sin considerar el resto de las secuencias. Este acercamiento resulta apropiado para alinear secuencias divergentes conteniendo solo módulos similares (motivos o dominios). Se basa en el algoritmo propuesto por Smith y Waterman [40] y utiliza una ecuación similar a la de Needleman y Wunsch [39], pero con algunas modificaciones. Al igual que en el modelo anterior, las dos secuencias se colocan en los ejes de una matriz rectangular, aunque se introduce una fila y una columna inicial cuyos elementos matriciales valen cero. Posteriormente las celdas *(i, j)* se completan

mediante una puntuación según la ecuación 2.4:

$$F(i,j) = max \begin{cases} 0 \\ F(i-1,j-1) + s(x,y) \\ F(i-1,j) - d \\ F(i,j-1) - d \end{cases} \tag{2.4}$$

En este caso, solo se asignan valores positivos o cero en los apareamientos. El recuento para obtener el mejor alineamiento se inicia en sentido inverso, pero a diferencia del alineamiento global, se puede iniciar y finalizar en el interior de la matriz, iniciando con la posición de mayor puntuación y continuando hasta dar con una celda con el cero [40].

3. **Construcción del modelo:** Una vez encontrado el alineamiento óptimo se inicia el proceso de generación de la estructura de la proteína problema. Los métodos para la construcción del modelo estructural se clasifican de acuerdo a cómo realizan la transferencia de la información de la estructura conocida a la secuencia problema. Entre los más utilizados se pueden mencionar: el ensamblado de cuerpo rígido (rigid body assembly) [41], los de evolución artificial (artificial evolution) [42], y los métodos de restricciones espaciales (spatial restraint) [43].

   En el método del ensamblado de cuerpo rígido, inicialmente se determina que regiones son conservadas y cuales varían en el alineamiento. Se considera que los residuos en las regiones conservadas de la proteína problema pueden asumir una estructura similar a los de la proteína molde. Por lo tanto, las coordenadas de los residuos correspondientes a la proteína molde pueden ser simplemente copiadas a los residuos de la secuencia problema. Si los dos residuos son idénticos, las coordenadas de las cadenas laterales son copiadas junto con las de los átomos de la cadena principal [41]. Si, en cambio, son diferentes, solo los átomos de la cadena principal son copiados y los átomos de la cadena lateral serán reconstruidos en un paso posterior. Las regiones no conservadas generalmente aparecen como regiones

de aberturas (gaps) producidas por inserciones o deleciones. Estos gaps no pueden ser directamente modelados, produciéndose "agujeros" en el modelo tridimensional. Cerrar estos agujeros requiere modelar las regiones faltantes, lo que genera la mayor fuente de error en el modelo. Posteriormente se abordará en mayor detalle esta etapa (apartados "modelado de loops" y "modelado de cadenas laterales") [41].

El método de evolución artificial, supone que la estructura molde es una especie de estructura "padre", a partir de la cual se fue generando la estructura del blanco debido a las modificaciones evolutivas [42]. En estas condiciones se supone improbable que este proceso se haya desarrollado en un solo paso. En cambio, es de esperar que la variación sea producto de múltiples pasos, con cambios mínimos en cada etapa. De esta manera, este método consiste en la generación de cambios (u operaciones) como mutaciones, inserciones y deleciones que modifiquen la estructura molde paso a paso hasta llegar a la estructura de la secuencia problema. Cada operación es evaluada en términos energéticos y es seguida de una pequeña minimización de energía para remover choques entre átomos. La estructura final es también minimizada. Las operaciones que producen un cambio favorable en la energía son aceptadas y el proceso continúa hasta que la estructura blanco es completamente modelada [42].

El método de satisfacción de restricciones espaciales es el que implementa el programa MODELLER [43, 44, 45] utilizado en esta tesis. Inicia generando restricciones en la estructura tridimensional de la proteína problema basadas en el alineamiento de está con la estructura de la proteína molde. Estas restricciones son obtenidas asumiendo que la distancia entre dos residuos del modelo de la estructura problema es similar a la distancia entre los dos residuos correspondientes al molde resultante del alineamiento. Además se adicionan restricciones estereoquímicas transfiriendo valores de ángulos de enlace, distancia de enlace, ángulo de torsión e interacciones no covalentes presentes en la estructura molde. Pueden agregarse también restricciones adicionales obtenidas por técnicas experimentales, incrementando la exactitud del modelo obtenido. Finalmente, la estructura es optimizada mediante la minimización

energética a partir del uso de campos de fuerza de mecánica molecular.

Modelado de loops:

La predicción de las estructuras tridimensionales de los loops es más difícil que para el caso de las α-hélices u otras regiones más regulares, debido a que exhiben una gran variabilidad estructural. Este modelado puede realizarse a través de métodos de búsqueda en bases de datos [46, 47], métodos Ab-Initio [48, 49, 50, 51, 52] o mediante la combinación de ambas aproximaciones [53].

Los métodos de búsqueda en bases de datos consisten en seleccionar partes de proteínas de estructura tridimensional conocida que puedan "encajar" entre las dos regiones troncales (ya modeladas) de la proteína problema. El procedimiento inicia midiendo la orientación y la distancia entre estas regiones y buscando en bases de datos segmentos del mismo tamaño que se ajusten a la zona a modelar. Usualmente, se encuentran muchos segmentos alternativos que pueden ubicarse en esta región. Por lo tanto, la mejor estructura de loop se puede seleccionar en base a la similitud de secuencia entre el fragmento encontrado y el fragmento "sin molde" de la secuencia problema. Otro criterio de selección es asignar aquella estructura capaz de minimizar choques estéricos con las regiones vecinas. Finalmente, la conformación del mejor fragmento es copiada en la región [45].

En el caso de los métodos ab-initio se generan las estructuras tridimensionales de los loops mediante diferentes estrategias de búsqueda conformacional como, por ejemplo, mediante métodos sistemáticos o por métodos estocásticos. Se buscan conformaciones mediante la comparación de funciones de puntuación o de energía, eligiéndose aquellas correspondientes a los valores más bajos [45].

Modelado de Cadenas laterales:

Una vez obtenido el esqueleto estructural de la proteína, las cadenas laterales pueden, en principio, ser construidas mediante un análisis conformacional de cada ángulo de torsión. Aquellas conformaciones más favorables en términos de energía de interacción serán las seleccionadas para incorporar al modelo. Sin embargo, esta

aproximación es computacionalmente prohibitiva en la mayoría de los casos. Muchos de los programas de modelado actuales utilizan librerías de rotámeros para resolver este inconveniente. En estas bases de datos se almacenan los ángulos diedros preferencialmente encontrados para las cadenas laterales de proteínas en estructuras cristalográficas. Cada uno de los confórmeros es puntuado según la frecuencia con que se encuentran experimentalmente. El empleo de estas librerías reduce significativamente los tiempos computacionales, debido a que solo se evalúa un pequeño número de ángulos de torsión favorables. Solo aquellas conformaciones que generen una energía favorable de interacción con los átomos vecinos son consideradas como estructuras posibles [54].

En muchas ocasiones aun con el uso de estas librerías de rotámeros el costo computacional necesario para resolver el modelo estructural es muy alto. Para reducirlo se puede tener en cuenta la conformación de la cadena principal. Se ha demostrado que hay una correlación entre ciertos rotámeros y la disposición espacial de la cadena peptídica. Por lo tanto, un gran número de rotámeros puede ser eliminado por el uso de estas correlaciones. Una vez agregado el rotámero más frecuente, la conformación debe ser evaluada y optimizada para minimizar los efectos estéricos con el resto de los átomos [54].

4. **Refinado y evaluación:** En los procesos de modelado, tanto de loops como de las cadenas laterales de aminoácidos, se aplican cálculos de la energía potencial para evaluar la calidad de los resultados. Sin embargo, esto no garantiza que el modelo final completo esté libre de irregularidades estructurales, tales como ángulos de enlace, distancias de enlace y hasta contactos entre átomos desfavorables. Esta clase de irregularidades estructurales se corrigen por la aplicación de un proceso de minimización energética sobre el modelo entero, el cual mueve los átomos de modo que toda la estructura adquiera una conformación de baja energía potencial. El objetivo es reducir las colisiones y la tensión estérica, sin alterar significativamente la estructura global. Este proceso se denomina refinamiento del modelo.

Otro procedimiento de refinamiento utilizado a menudo es la dinámica molecular. Esta práctica es derivada de la idea de que la minimización de energía solo mueve los átomos hacia un mínimo de energía local sin pasar por todas las posibles conformaciones, resultando muchas veces en una estructura subóptima. Mediante la dinámica molecular se hace un barrido conformacional más amplio, permitiendo a las macroestructuras acceder a mayor cantidad de estados diferentes y representativos de la superficie de energía potencial.

Finalmente, el modelo completo tiene que ser evaluado para asegurar la calidad de la estructura tridimensional obtenida. Por un lado es posible observar anomalías en los ángulos del esqueleto carbonado de la proteína (denominados psi (ψ) y phi (ϕ)) mediante gráficos de Ramachandran [55]. Otros programas evalúan la estereoquímica del modelo obtenido mediante la comparación de parámetros como tamaño de los enlaces, ángulos de enlaces, planariedad de enlaces peptídicos y de anillos aromáticos, etc.; con los valores encontrados experimentalmente. De esta manera se pueden detectar regiones de la secuencia que pueden considerarse bien modeladas y otras que no lo están. Ejemplos de estos softwares son PROCHECK [56] y WHATCHECK [57], entre otros. Adicionalmente existen programas como VERIFY3D [58] o PROSAII [59] que evalúan la capacidad de adaptación de la secuencia problema a la estructura adoptada. En estos casos se asigna una puntuación a cada aminoácido, lo que permite controlar la presencia de irregularidades residuo por residuo, e identificar zonas mal modeladas (o que requieran un nuevo refinamiento del modelo). Debe aclararse en este punto que el proceso de modelado debe repetirse si es necesario desde el comienzo hasta que los resultados obtenidos en la evaluación sean satisfactorios.

Como se mencionó anteriormente, la estructura tridimensional del blanco molecular (obtenida mediante métodos experimentales o por los modelos de predicción), resulta indispensable para el diseño racional de fármacos mediante cribado virtual basado en el receptor. Conseguida esta macromolécula, el paso siguiente es evaluar su capacidad

de interacción con una biblioteca de pequeños compuestos (posibles candidatos a fármacos) en forma virtual. Para ello generalmente se realizan simulaciones utilizando docking molecular.

Docking molecular.

El docking molecular, o simplemente docking, es una simulación computacional orientada a predecir la conformación del complejo ligando – receptor, donde el receptor es usualmente una proteína o un ácido nucleico y el ligando puede ser una molécula pequeña u otra proteína. El objetivo de esta herramienta es comprender y predecir el reconocimiento entre dos moléculas tanto a nivel estructural (encontrando el modo de unión entre ambos), como energético (prediciendo el valor de la afinidad de unión del complejo) [60].

Al momento, las propiedades moleculares responsables de este reconocimiento son difíciles de identificar, entender y sobre todo, de simular en una computadora [60]. Sin embargo, los programas de docking intentan encontrar la mejor forma de unión entre moléculas teniendo en cuenta diversos parámetros como la complementariedad geométrica, la flexibilidad de las estructuras y las interacciones intermoleculares. Como resultado estos softwares devuelven las orientaciones predichas del ligando en el sitio de unión del receptor (también denominadas "poses"). Usualmente el proceso de búsqueda del modo de unión devuelve varios resultados posibles. Para determinar cuál es el más probable se utiliza una función de evaluación. Este valor ordena las poses obtenidas asociándoles un valor numérico relacionado con la energía libre de unión del complejo, que permite seleccionar las mejores posiciones relativas ligando-receptor [61].

Por otra parte, existen diferentes aproximaciones respecto del grado de flexibilidad con que se considera el ligando y el receptor en la simulación. Se habla de docking rígido cuando se propone al ligando y al receptor como cuerpos rígidos, limitando el proceso de búsqueda de poses a encontrar la mejor complementariedad geométrica/fisicoquímica posible. Por su parte, el docking flexible involucra el acceso a diferentes conformaciones del

ligando, del receptor o de ambos al mismo tiempo [61]. Las capacidades computacionales actuales permiten el empleo de docking flexible durante el cribado virtual, pero con limitaciones que no son menores (como se detallará posteriormente). De hecho, considerar correctamente la flexibilidad del sistema es uno de los mayores desafíos en este tipo de herramientas de modelado [60].

De lo explicado anteriormente se desprende que en un proceso de simulación por docking conviven dos tipos de algoritmos: uno encargado de generar diferentes conformaciones relativas entre ligando–receptor, y otro capaz de evaluar numéricamente las interacciones moleculares en cada una de esas conformaciones; a fin de determinar cuáles son las más probables. A continuación se describen estos algoritmos.

Métodos de búsqueda:

El modo de unión del ligando con respecto al receptor es definido por la conformación adoptada (ángulos de torsión de cada enlace rotable), la posición (coordenadas *x, y, z*) y orientación (ángulos de Euler, cuaterniones, etc) en el sitio activo [62, 63]. Cada una de estas variables describe un grado de libertad en un espacio de búsqueda multidimensional. Un algoritmo de búsqueda riguroso debería muestrear todos los modos de unión entre las dos moléculas involucradas en el docking. Sin embargo, esto es impráctico en términos computacionales debido al gran tamaño del espacio conformacional del ligando [64]. Además, el problema se incrementa si se considera, por un lado, la flexibilidad del receptor y por otro el tamaño de las bases de datos de miles de ligandos que suelen utilizarse en el cribado virtual [22]. Consecuentemente, solo una pequeña parte del espacio conformacional suele ser muestreada, y esta exploración debe ser realizada con la suficiente exactitud como para identificar la conformación que mejor interacciona con la estructura del receptor. Además, debe ser adecuadamente rápida como para permitir la evaluación de miles de compuestos en una corrida de docking [60]. Los métodos de búsqueda más utilizados son los denominados sistemáticos y los estocásticos, los cuales se describirán brevemente a continuación.

1. Métodos sistemáticos: estas técnicas incorporan la flexibilidad del ligando mediante la exploración de todos los grados de libertad en la molécula. En algunos casos incluyen también movilidad en algunos residuos del receptor. Pueden dividirse en tres tipos principales: métodos de búsqueda de conformaciones, métodos de fragmentación y métodos de bases de datos [64].

   (a) Métodos de búsqueda de conformaciones: todos los enlaces capaces de rotar en el ligando son sistemáticamente modificados 360 grados usando un incremento fijo (por ejemplo cada 10 grados). Se generan, entonces, un gran número de conformaciones resultantes de todas las combinaciones posibles de ángulos de torsión, las cuales son evaluadas [12]. El mayor problema es el gran número de rotámeros creados, el cual se incrementa enormemente con el número de enlaces móviles. Este fenómeno es conocido como explosión combinatoria [65]. Por lo tanto, la aplicación de este tipo de métodos, en su forma pura, es muy limitado. Normalmente se aplican restricciones en los ángulos de torsión y limitaciones para reducir la dimensionalidad del problema [12].

   (b) Métodos de fragmentación: esta herramienta es una de las más utilizadas para introducir la flexibilidad del ligando en los programas de docking molecular. En este caso el ligando "se corta", mediante la ruptura de enlaces capaces de rotar, en fragmentos que son dockeados en forma separada en el sitio activo del receptor. Los fragmentos rígidos (o los de mayor tamaño) son añadidos inicialmente y funcionan como un "ancla" sobre el cual se van añadiendo los siguientes fragmentos. Estos últimos se van incorporando en todas las conformaciones posibles, y se van seleccionando aquellas que no presenten efectos estéricos negativos entre átomos del ligando o con el receptor. De esta forma el ligando se va construyendo gradual e incrementalmente sobre el sitio de unión [63]. Los programas que implementan este tipo de algoritmo son DOCK [66], FlexX [67], Hammerhead [68], entre otros. (figura 2.4)

   (c) Métodos de bases de datos: en este caso se utilizan librerías de conformaciones pre-generadas para tratar la flexibilidad del ligando. Estas conformaciones son

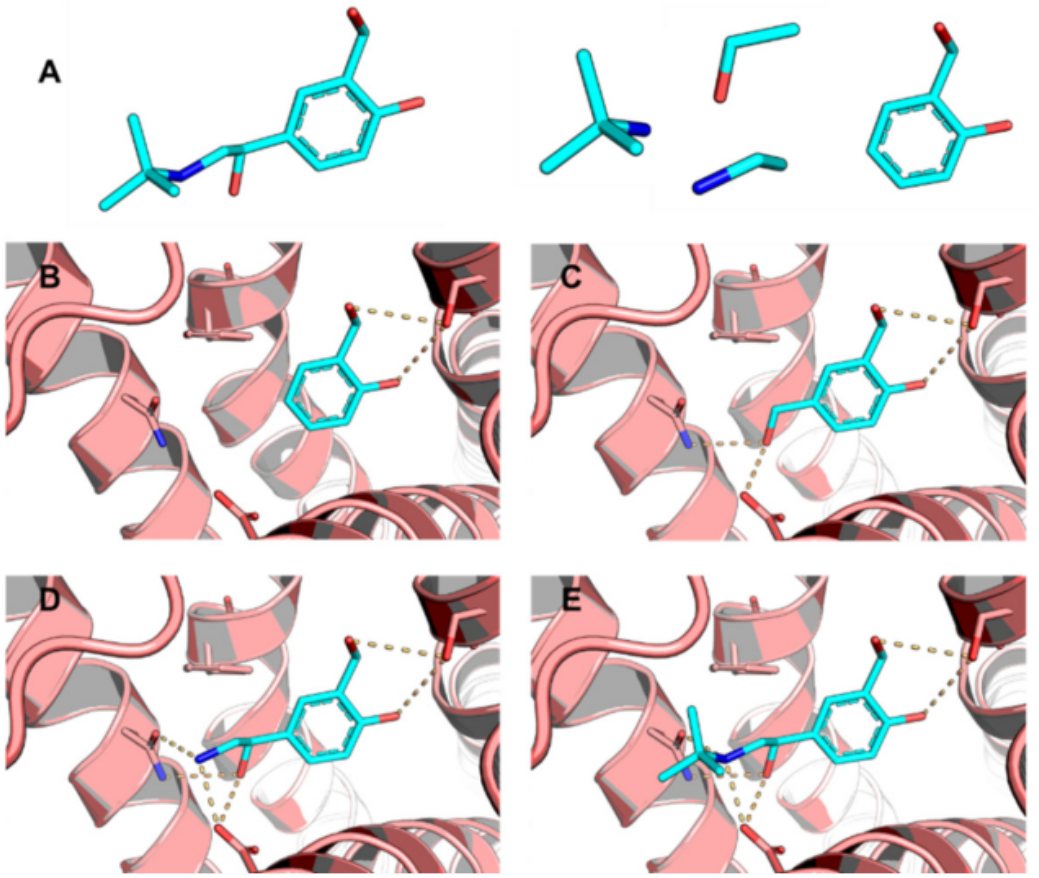


Figura 2.4: Método de fragmentación. A) El ligando es cortado en varios fragmentos; B) El fragmento "ancla" es posicionado sobre el sitio de unión del blanco; C) El siguiente fragmento se une luego del inicial; D y E) Los siguientes fragmentos se añaden secuencialmente. Adaptado de Ferreira y col. [69] bajo licencia Creative commons attribution.

subsecuentemente sometidas a un proceso de docking rígido [64]. Por ejemplo, el programa FRED [70] utiliza librerías de ángulos de torsión y de anillos para identificar y enumerar enlaces capaces de rotar y anillos flexibles. Los confórmeros que luego del docking muestran efectos estéricos negativos o una tensión elevada en sus ángulos de torsión son descartados. Por el contrario, aquellos con interacciones favorables son separados en clústers según su ubicación en el sitio activo, y se retienen los centroides de cada clúster [70].

2. Los métodos estocásticos operan generando cambios al azar en los grados de libertad del ligando en una conformación inicial y, en algunos casos, también residuos del

receptor. La nueva conformación obtenida es evaluada con la función de evaluación del docking, que analiza la energía de interacción en el complejo formado. Con este resultado se analiza si la conformación es aceptada o rechazada, mediante una función de probabilidad. Los métodos más utilizados son: el método de Metrópolis – Monte Carlo [71, 72], el método de simulated annealing [61] y los algoritmos genéticos [60].

(a) Método de Metrópolis – Monte Carlo [71, 72]: en esta técnica se genera una conformación inicial al azar del ligando en el sitio de unión *(e)*. Esta configuración inicial es puntuada mediante el empleo de la función de evaluación del docking. Luego se crea una nueva configuración al azar *(e')*, la cual también es puntuada. Según el criterio de Metrópolis [73], la probabilidad de aceptar una solución está dada por el valor de su puntuación: Si la nueva solución es mejor que la anterior (es decir que el valor de la función de evaluación de esta es menor a la anterior) la probabilidad de aceptación es igual a 1. En este caso la nueva solución es aceptada. Si por el contrario, la nueva solución es desfavorable (el valor de la función de evaluación es mayor), entonces el cambio conformacional se estima según la función de distribución de energías de Boltzmann (ecuación 2.5).

$$P(f(e), f(e')) = \begin{cases} 1 & \text{si } f(e') < f(e) \\ \exp\left(\dfrac{f(e) - f(e')}{kT}\right) & \text{si } f(e') > f(e) \end{cases} \qquad (2.5)$$

$$\text{si } P(f(e), f(e')) > n_A \quad \text{se acepta la nueva conformación}$$

Donde $f(e')$ es la función de puntuación aplicada a la conformación (e'), $f(e)$ es la correspondiente a la conformación *(e)*, k es la constante de Boltzmann y T es la temperatura de trabajo. En el caso de que $f(e')$ sea mayor que $f(e)$ se acepta la nueva conformación si la función de distribución de Boltzmann supera un numero aleatorio n_A entre 0 y 1 (ecuación 2.5).

La operación se repite y se da por finalizada una vez que se alcanza un número predeterminado de búsquedas o cuando se consigue un número suficiente de poses [12].

La principal ventaja del método es que los cambios generados en las conformaciones pueden ser muy grandes, permitiendo así atravesar barreras energéticas en la superficie de energía potencial, un aspecto difícil de alcanzar en otros métodos [63].

(b) Simulated Anneling o recocido simulado [74]: Al igual que en el método de Metrópolis-Monte Carlo, en esta simulación se produce una variación al azar de los grados de libertad del sistema generando, partir de una conformación inicial del ligando, una nueva conformación que se evalúa en su interacción con el receptor. Para esto se utiliza la función de evaluación proveniente del docking que luego es comparada con el valor obtenido para la conformación anterior mediante el criterio de Metrópolis [75]. Sin embargo, en este método el análisis contempla también la variación de la temperatura, la cual influye en el criterio de decisión de Metrópolis (ecuación 2.5). El proceso es dividido en un número predefinido de ciclos, los cuales permanecen a una temperatura constante. La temperatura se reduce al inicio de cada ciclo, que comienza con la conformación de menor energía obtenida durante el ciclo anterior. De esta manera las poses generadas son sucesivamente refinadas, ya que la simulación se encuentra controlada por la temperatura. Los métodos de simulated annealing realizan tanto una búsqueda global como a nivel local. A altas temperaturas, la búsqueda conformacional puede considerarse global ya que se permite la transición a través de las barreras energéticas que separan los valles de energía. Posteriormente, cuando se alcanzan temperaturas bajas, la búsqueda se restringe a zonas aledañas a un determinado valle (mínimo energético) [74].

(c) Los algoritmos genéticos adaptan los conceptos de la teoría evolutiva a la resolución de problemas computacionales [76]. En general el proceso de búsqueda se basa en la generación de una población de soluciones candidatas sobre las

que iteraciones sucesivas dan lugar a una selección competitiva, que elimina aquellas soluciones con un valor desfavorable de la función de evaluación (o fitness function) [76]. En la aplicación de los algoritmos genéticos al docking molecular, una posición particular de un ligando sobre la proteína puede ser definida por un conjunto de variables que describen la traslación, orientación y conformación de este sobre la macromolécula. Cada una de estas corresponde a un gen, que en su conjunto constituyen el genotipo o cromosoma; mientras que la posición del ligando relativa la proteína, corresponde al fenotipo (Figura 2.5) [77]. Dicho de otro modo, una determinada conformación de un ligando en un receptor puede codificarse en un cromosoma, el cual se divide en genes (representando cada una de las variables que describen la ubicación de los átomos del ligando). A su vez podemos agrupar a los genes en varios alelos, relacionados con los posibles valores que puede adoptar una variable concreta. Normalmente, los cromosomas se codifican mediante cadenas (strings) que pueden contener valores binarios, valores enteros o incluso valores reales en coma flotante. [77].

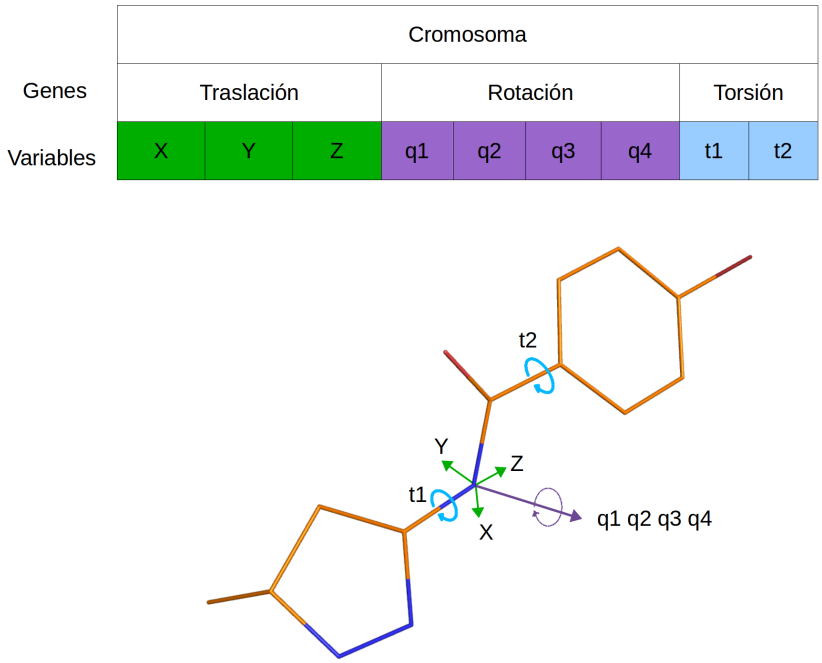


Figura 2.5: Representación de las variables de estado en cromosomas.

El algoritmo inicia con la generación de una población de conformaciones de partida (también llamadas cromosomas o individuos) generados al azar. Cada una de estas se evalúa según la función de evaluación del docking y se seleccionan los mejores candidatos (aquellos confórmeros que mejor interaccionan con el receptor). Sobre estos individuos se aplican operadores de recombinación y mutación, dando lugar a una nueva generación. En el proceso de recombinación se produce un entrecruzamiento de genes entre un par de individuos seleccionados al azar, lo que genera nuevas conformaciones. Las mutaciones consisten en el cambio al azar de un alelo por otro en un gen. Con el nuevo conjunto de candidatos obtenidos, la selección de los individuos para la nueva generación se realiza nuevamente por la función de evaluación del docking. El cálculo finaliza luego de un número pre-establecido de generaciones (Figura 2.6) [77].

Figura 2.6: Esquema de un algoritmo genético

Una variación a la metodología de búsqueda de conformaciones basada en algoritmos genéticos es la implementada en el programa de docking molecular Autodock desde la versión 3 en adelante [77, 78]. Esta modificación consiste en la aplicación de un método de minimización local de energía a cada individuo (Figura 2.7). Los candidatos construidos mediante mutaciones y entrecruzamientos en una generación son sometidos a una minimización local basada en la función de evaluación del docking. En esta implementación, denominada algoritmo genético "Lamarckiano", se produce un intercambio entre los espacios genotípico y fenotípico. El espacio genotípico es definido por los operadores genéticos (recombinación y mutaciones), mientras que el espacio fenotípico consiste en la función de energía a ser optimizada. La minimización energética (búsqueda local) es realizada en el espacio fenotípico, posteriormente a los cambios genotípicos sobre la población (búsqueda global) [77].

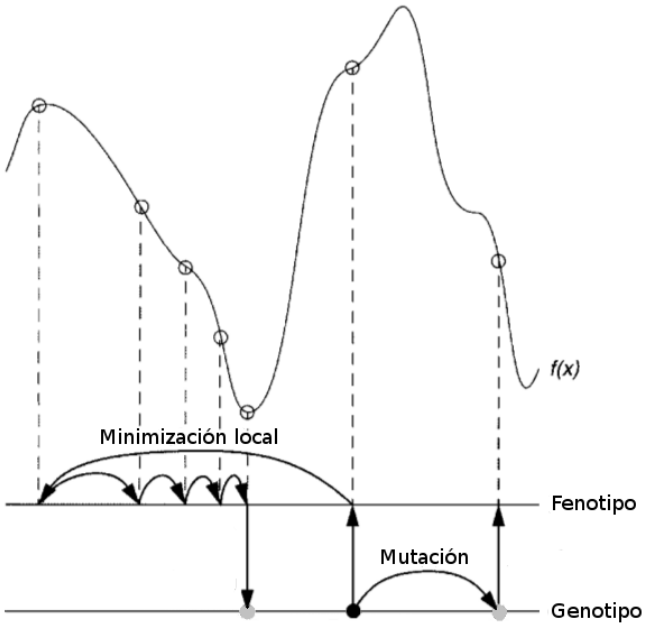


Figura 2.7: Esquema de búsqueda por algoritmos genéticos lamarckianos en relación a algoritmos genéticos puros

Como se detalló previamente, los métodos de búsqueda conformacional aplicados a la simulación por docking deben complementarse con una función de evaluación que clasifique los distintos modos de unión posibles entre el ligando y el receptor [79]. A continuación se enumeraran las funciones de puntuación más utilizadas.

Función de evaluación:

La asociación no covalente entre dos moléculas es gobernada por la termodinámica general. Al igual que otros procesos espontáneos, la interacción ocurre solo cuando se produce un cambio negativo de la energía libre (ΔG). La energía libre se obtiene a través de la ecuación de Gibbs-Helmholtz (ecuación 2.6) [80]:

$$\Delta G = \Delta H – T\Delta S \tag{2.6}$$

Donde ΔG es la energía libre de unión, ΔH es la entalpía, T es la temperatura en

grados Kelvin y ΔS es la entropía.

La entalpía en estos sistemas refleja la especificidad y la fuerza de las interacciones entre las moléculas. Estas interacciones pueden ser del tipo iónicas, puentes de hidrogeno, electrostáticas (coulómbicas), de van der Waals, entre otras. La entropía puede describirse, en forma simplificada, como una medida de la dinámica del sistema. La formación de un complejo estable refleja una pérdida de la entropía debido a la reducción del movimiento de sus moléculas componentes (y por lo tanto, en los grados de libertad de traslación, rotación y conformacionales de ambas partes). Sin embargo, los efectos de la desolvatación (como la reorganización del solvente o la liberación de las moléculas de agua unidas estrechamente a la superficie molecular) también pueden contribuir significativamente a la energía libre de unión [81, 82].

La constante de unión en el equilibrio (K_b) para la formación de un complejo (LR) entre un ligando (L) y un receptor (R) está relacionada con la energía libre a través de la ecuación 2.8:

$$L + R \rightleftharpoons LR \tag{2.7}$$

$$\Delta G = \Delta G^{\circ} + RT \ln K_b \tag{2.8}$$

Donde ΔG° es la energía libre estándar (298°K, 1 atm de presión y los componentes a 1 Molar). Bajo condiciones de equilibrio, ΔG es igual a cero, a partir de lo cual se obtiene 2.9:

$$\Delta G^{\circ} = -RT \ln K_b \tag{2.9}$$

La función de evaluación en un programa de docking se construye teniendo en cuenta los fenómenos fisicoquímicos involucrados en la energía libre de unión (como los descriptos anteriormente). Su capacidad para predecir la afinidad de la unión ligando-receptor es

una de las características más importantes de un programa de docking molecular. Incluso cuando el algoritmo de generación de conformaciones sea capaz de predecir correctamente aquellas poses que verdaderamente interaccionan con el blanco molecular, el método no tendrá éxito si la función de puntuación no es capaz de diferenciarlas de las incorrectas [60]. Mientras mayor sea el número de parámetros fisicoquímicos evaluados, mayor será la precisión de la función de evaluación. Sin embargo, el costo computacional se incrementa con el número de variables incluidas. Idealmente, un algoritmo eficiente debe ofrecer un balance entre la precisión y la velocidad, lo cual es crítico al trabajar con grandes conjuntos de ligandos [69].

Las funciones de evaluación aplicadas tradicionalmente se clasifican en tres grupos principales: basadas en campos de fuerza (force field – based), empíricas y basadas en el conocimiento (knowledge-based).

1. **Funciones de evaluación basadas en campos de fuerza (Force field – Based):** Estas funciones de puntuación aproximan la energía libre de unión en base a una ecuación que describe las interacciones atómicas (mediante la suma de las contribuciones electrostáticas, de van der Waals, puentes de hidrógeno, etc.) y las fuerzas torsionales, de estiramiento y flexión de los enlaces. En cuanto a su forma funcional, el termino correspondiente a las interacciones de van der Waals entre pares de átomos i y j es generalmente descripto mediante los potenciales de Lennard – Jones, dependientes de la distancia interatómica r_{ij} (ecuación 2.10). Estos potenciales, también conocidos como potenciales 6 – 12 (en alusión a los exponentes numéricos de ambos términos dentro de los paréntesis) comprenden la suma de dos términos, un término atractivo que prevalece a distancias largas y un término repulsivo que, por el contrario, predomina a cortas distancias. El término electrostático es descripto como una formulación Coulómbica entre las cargas q_i y q_j, con una función dieléctrica dependiente de la distancia ε_0, que reduce la contribución de las interacciones carga–carga (ecuación 2.11) [60]. Algunas funciones también incorporan un término correspondiente a las interacciones de puente de hidrógeno,

las cuales pueden representarse de distinta manera [60].

$$E_{vdW}(r) = \sum_{i=1}^{N} \sum_{j=1}^{N} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{2.10}$$

$$E_{coulomb}(r) = \sum_{i=1}^{N} {}_{A} \sum_{j=1}^{N} {}_{B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.11}$$

Por lo general, a cada componente energético se le agrega coeficientes obtenidos mediante cálculos ab-initio o de origen empírico (generados mediante el ajuste con datos experimentales), ya que los métodos de cálculo utilizados no tienen relación y, por lo tanto, presentan escalas diferentes [83, 84].

Una de las mayores limitaciones de los métodos basados en campos de fuerza se encuentra en la descripción de la contribución entrópica. Esta falla se debe a la ausencia de un modelo físico razonable que describa dicha contribución. Además, el solvente no es modelado explícitamente, lo que dificulta la estimación del fenómeno de desolvatación [69]. Para mejorar esta deficiencia algunas funciones incluyen términos empíricos, con el objetivo de representar al componente entrópico debido a la presencia del solvente y a los cambios conformacionales producto de la unión. Estas funciones de evaluación suelen denominarse también como semiempíricas [85].

Existen muchas funciones de evaluación basadas en campos de fuerza, entre las más populares se encuentran D-Score [86], G-Score [87] (basadas en el campo de fuerza Tripos), GoldScore [88] y Autodock [85] (basadas en el campo de fuerza AMBER) [89, 90].

2. **Funciones de evaluación empíricas:** Estas funciones de puntuación son desarrolladas para reproducir datos experimentales y se basan en el concepto de que la afinidad de unión puede aproximarse por la suma de varios términos individuales no relacionados [64]. Estiman la energía libre de unión de un complejo mediante la suma de términos parametrizados, como se muestra en la ecuación 2.12:

$$\Delta G = \sum_i W_i \cdot \Delta G_i \tag{2.12}$$

Donde ΔG_i representa diferentes términos energéticos, como las energías vinculadas a las interacciones de van der Waals, electrostáticas, puentes de hidrógeno, aromáticas, entre otras. También pueden incluir términos no entálpicos, como la entropía o, incluso, agregar un término relacionado únicamente con la desolvatación. Los coeficientes W_i son determinados mediante el ajuste con datos experimentales obtenidos de un conjunto de entrenamiento de complejos ligando–receptor con estructura tridimensional y constante de unión conocidas [83].

Las distintas funciones de evaluación empíricas se encuentran compuestas por términos similares. Sin embargo, la representación matemática de cada uno de ellas, los datos utilizados para la parametrización y los métodos de optimización de dichos parámetros suele variar. En comparación con las ecuaciones basadas en campos de fuerza, se observa que este tipo de formas funcionales empíricas es más simplificada, e intenta capturar las interacciones que dan origen al complejo sin tener en cuenta la física subyacente [84]. Debido a la mayor simplicidad de los términos energéticos, los cálculos de las funciones empíricas son más rápidos, lo que representa una ventaja sobre todo si se tienen que evaluar una gran cantidad de compuestos [69]. Una desventaja de estos métodos es su dependencia con los datos experimentales del conjunto de entrenamiento utilizado para generar la función, lo cual limita su aplicabilidad general [84]. Algunas de las funciones empíricas más utilizadas son LUDI [91], ChemScore [92], GlideScore [93, 94, 95] y X-Score [96].

3. **Funciones de evaluación basadas en el conocimiento (knowledge-based):** En este caso las funciones se basan en la observación de interacciones entre complejos ligando- receptor de estructuras resueltas experimentalmente. Son desarrolladas teniendo en cuenta la frecuencia con la que dos átomos son encontrados dentro de una distancia determinada en el conjunto de estructuras de entrenamiento [69]. Mediante el análisis estadístico, las frecuencias de una característica estructural

del complejo son convertidas en energía libre, en base a la suposición de que las características encontradas con mayor frecuencia corresponden a estados de menor energía [97]. Así, los diferentes tipos de interacciones observados son clasificados y ponderados de acuerdo con su frecuencia de aparición. La puntuación final de un complejo ligando–receptor es obtenida mediante la suma de estos potenciales individuales. La ecuación 2.13 define la energía libre de unión como la suma de todos los potenciales de interacción entre los átomos del ligando y del receptor.

$$\Delta G = \sum_{kl} A_{ij}(r) \tag{2.13}$$

Donde A_{ij} corresponde al potencial de interacción entre pares de átomos i-j en una distancia r [98].

Una desventaja de esta aproximación es que la función deriva de información correspondiente a un conjunto de estructuras ligando–receptor limitado. Una de las ventajas es su simplicidad computacional, lo cual permite su uso en la búsqueda en grandes librerías de compuestos [60].

DFIRE [99], DrugScore [100, 101], ITScore [102, 103], PMF-score [104], y SMoG [105] son algunas de estas funciones.

Como se enunció previamente, cada una de las funciones de evaluación presenta ventajas y limitaciones, por lo que, ninguna de ellas es perfecta en términos de exactitud, ni en cuanto a su capacidad de aplicación generalizada. Un intento de limitar las deficiencias individuales de las funciones plantea la combinación de las soluciones obtenidas a partir diferentes funciones individuales, de forma de aumentar la probabilidad de encontrar el resultado correcto. Esta estrategia se conoce como función de puntuación consenso o "Consensus Scoring" [106]. El punto crítico en este procedimiento es el diseño del consenso adecuado para reproducir la afinidad de unión [107]. Existen diferentes funciones de evaluación basadas en el consenso, algunas de ellas son: MultiScore [108], X-Cscore [96], GFscore [109], SCS [110] y SeleX-CS [111].

Herramientas de docking molecular utilizadas:

En el presente trabajo se utilizaron los programas de docking molecular Autodock4.2 [77, 85], Autodock Vina [112] y Glide [93, 94, 95]. **Autodock4.2** utiliza una función de evaluación semiempírica en la que la energía libre de unión es estimada como la diferencia entre (1) la energía del ligando y de la proteína en un estado separado y (2) la energía del complejo ligando – proteína. Esta determinación es dividida en dos pasos: por un lado se estima el cambio de la energía intramolecular en la transición desde una conformación no enlazada a una enlazada para cada molécula por separado, y luego se evalúa la energía intermolecular de las dos moléculas unidas (figura 2.8) [85].

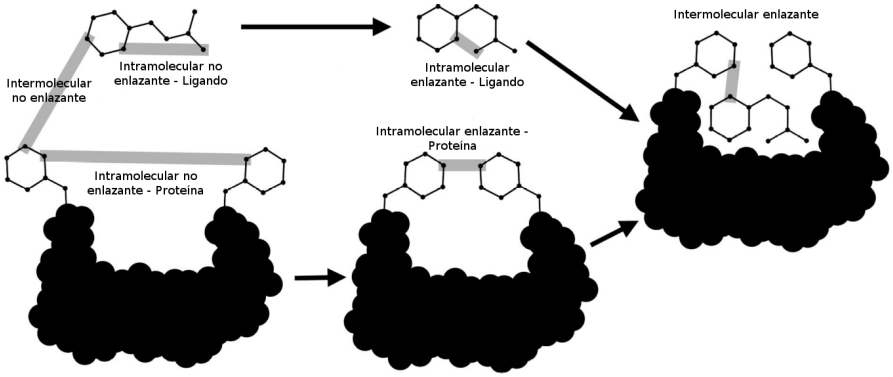


Figura 2.8: Método de cálculo de la interacción ligando-receptor de Autodock4.2.

La función de evaluación incluye seis términos que cuantifican la energía relacionada con interacciones de a pares, más un valor asociado a la pérdida de entropía (ΔS_{conf}) de acuerdo a la ecuación 2.14:

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}) \quad (2.14)$$

Donde L refiere al ligando y P a la proteína que forman el complejo. Los primeros cuatro términos corresponden a las energías intramoleculares para los estados enlazado (bound) y no enlazado (unbound) del ligando y de la proteína. El cambio de energía

intermolecular entre el estado enlazante y no enlazante se obtiene como diferencia entre el quinto y sexto término. Se asume que las dos moléculas en el estado no enlazante están a una distancia suficiente para que $V_{unbound}^{P-L}$ sea igual a cero. Si durante el docking la proteína se mantiene rígida, la diferencia entre los términos de cambio en la energía intramolecular, $V_{bound}^{P-P} - V_{unbound}^{P-P}$, es también igual a cero.

Cada uno de los términos V de la ecuación 2.14 incluyen la evaluación de interacciones no covalentes (van der Waals, puentes de hidrógeno, electrostáticas), junto con la contribución de la desolvatación, de acuerdo a la ecuación 2.15.

$$\begin{aligned} V = {} & W_{vdW} \sum_{i=1}^{N} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{PH} \sum_{i=1}^{N} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + \\ & W_{elec} \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{ij} \left(S_i V_j + S_j V_i \right) e^{\left(r_{ij}^2 / 2\sigma^2 \right)} \end{aligned} \tag{2.15}$$

En este algoritmo los parámetros W son coeficientes optimizados mediante regresión lineal, a partir de datos empíricos de energía libre de un conjunto de complejos ligando–proteína caracterizados experimentalmente. El primer término de esta ecuación corresponde al potencial 6 – 12 entre átomos denominados i y j, típico para la descripción de las interacciones de van der Waals a una distancia r_{ij}; y cuyos parámetros A y B corresponden al campo de fuerza AMBER [90]. El segundo término determina la contribución de las interacciones por puentes de hidrógeno, con un potencial similar al anterior pero con exponentes 10 – 12. El valor de *E(t)* depende del ángulo t formado por los átomos que participan de la interacción, ya que en este tipo de fuerzas influye la direccionalidad del enlace. El tercer término corresponde a las fuerzas electrostáticas, expresadas mediante un potencial del tipo coulómbico, que implica computar las cargas atómicas q_i y q_j. El término final describe la contribución energética de la desolvatación. Este cálculo se basa en el volumen de átomos *(V)* que rodean a un átomo dado y lo apantallan del solvente, ponderada por un término de solvatación *(S)* dependiente de la carga, y de una función exponencial que contiene una constante σ igual a 3,5$\mathring{A}$ y que está relacionada con la distancia entre los átomos [77, 85]. Finalmente, el valor de la pérdida de entropía conformacional en la ecuación 2.14 (ΔS_{conf}), es directamente

proporcional al número de ángulos de torsión del ligando (N_{tor}) de acuerdo al siguiente algoritmo (ecuación 2.16):

$$\Delta S_{conf} = W_{conf} \cdot N_{tor} \tag{2.16}$$

La función de evaluación de **Autodock Vina**, combina las ventajas de las ecuaciones basadas en el conocimiento y las empíricas. Este programa extrae información de la conformación de preferencia de complejos proteína–ligando conocidos y también de medidas experimentales de afinidad [112]. La forma de la función de puntuación en los términos dependientes de la conformación es una sumatoria de términos que tienen en cuenta las interacciones entre pares de átomos, cuyos movimientos se encuentran relacionados, excluyendo generalmente las parejas de átomos ubicados mas allá de 3 enlaces covalentes consecutivos (Ecuación 2.17).

$$c = \sum_{i<j} f_{t_i t_j}(r_{ij}) \tag{2.17}$$

A cada átomo i se le asigna un tipo t_i y se define un conjunto de funciones simétricas $f_{t_i t_j}$ que cuantifican las interacciones entre los átomos i y j a la distancia interatómica r_{ij}. La función de interacción $f_{t_i t_j}$ se define, a su vez, con respecto a la distancia superficial interatómica (d_{ij}) de la siguiente manera (ecuación 2.18):

$$f_{t_i t_j}(r_{ij}) = h_{t_i t_j}(d_{ij}) \tag{2.18}$$

Donde $h_{t_i t_j}$ es otra función definida como la suma parametrizada de las interacciones estéricas, hidrofóbicas y por puentes de hidrógeno. Por otra parte, d_{ij} puede definirse como (ecuación 2.19):

$$d_{ij} = r_{ij} - R_{ti} - R_{tj} \tag{2.19}$$

Donde R_{ti} corresponde al radio de van der Waals del tipo de átomo i y lo mismo sucede para R_{tj}. Los efectos estéricos en $h_{t_i t_j}$ están representados por tres términos, dos funciones gaussianas (ecuaciones 2.20 y 2.21) que representan las interacciones atractivas, mientras que el efecto repulsivo esta dado por el cuadrado de la distancia d_{ij}, si esta es menor que cero (ecuación 2.22):

$$gauss_1(d) = e^{-(d/0{,}5\mathring{A})^2} \tag{2.20}$$

$$gauss_1(d) = e^{-((d-3\mathring{A})/2\mathring{A})^2} \tag{2.21}$$

$$repulsion(d) = \begin{cases} d^2 & \text{si } d < 0 \\ 0, & \text{si } d \geq 0 \end{cases} \tag{2.22}$$

El término hidrofóbico en $h_{t_i t_j}$ es igual a 1 cuando d_{ij} es menor que $0,5\mathring{A}$; 0, cuando d_{ij} es mayor a $1,5\mathring{A}$ y entre ambos valores decrece linealmente. El término correspondiente a los enlaces de hidrógeno vale 1 para distancias menores a $-0,7\mathring{A}$; 0, cuando d_{ij} es mayor a $0\mathring{A}$ y también se comporta como lineal entre ambos valores. Los parámetros de las funciones que conforman los términos de $h_{t_i t_j}$ se presentan en la tabla 2.1.

Parámetro	Término
$-0,0356$	Gauss 1
$-0,00516$	Gauss 2
$0,840$	Repulsión
$-0,0351$	Hidrofóbico
$-0,587$	Enlaces de hidrógeno
$0,0585$	Ángulos de torsión

Tabla 2.1: Parámetros y términos de la función de evaluación de Autodock Vina. [112].

Todas las interacciones presentan un punto de corte a r_{ij} de 8 $\mathring{A}$. Finalmente, la contribución a la energía libre por parte de la pérdida conformacional del ligando es directamente proporcional al número de ángulos de torsión de este ($W_{tor} \cdot N_{tor}$).

Glide emplea una función de evaluación empírica, denominada GlideScore, la cual presenta dos formas funcionales: GlideScore SP (Standard Precision) y GlideScore XP

(Extra Precision). Estas dos funciones usan términos similares pero fueron desarrolladas con diferentes objetivos. GlideScore SP es una función "más permisiva", pensada para identificar ligandos propensos a unirse al receptor. Esto la hace apropiada para el cribado en grandes bases de datos [93]. En el caso de la versión XP, la forma funcional penaliza rigurosamente las poses que violan los principios fisicoquímicos de la interacción. Esta versión resultaría más adecuada para la etapa de optimización de líderes, debido a su mayor costo computacional [95].

La función de GlideScore SP [93], fue desarrollada a partir de la función ChemScore [113] y contiene 10 términos (ecuación 2.23):

$$\begin{aligned}
&\Delta G_{bind} = C_{lipo-lipo} \sum f(r_{ij}) + \\
&C_{hbond-neut-neut} \sum g(\Delta r_{DA}) h(\Delta \alpha_{DA}) + \\
&C_{hbond-neut-charged} \sum g(\Delta r_{DA}) h(\Delta \alpha_{DA}) + \\
&C_{hbond-charged-charged} \sum g(\Delta r_{DA}) h(\Delta \alpha_{DA}) + \\
&C_{max-metal-ion} \sum f(r_{ij}) + C_{rotb} H_{rotb} + \\
&C_{polar-phob} V_{polar-phob} + C_{coul} E_{coul} \\
&C_{vdW} E_{vdW} + \text{ terminos de solvatación}
\end{aligned} \tag{2.23}$$

El primer término corresponde a las interacciones lipofílicas de todos los pares de átomos lipófilos i y j entre proteína y ligando (generalmente carbono) y $f(r_{ij})$ es una función lineal de la distancia interatómica (r_{ij}). Las interacciones de puentes de hidrógeno están separadas en diferentes términos dependiendo de si el dador y el aceptor de protones son neutros, si uno es neutro y el otro está cargado o si ambos están cargados. Utilizan dos formas funcionales, una dependiente de la distancia de enlace (r) y otra del ángulo de enlace (α). Estos términos computan para cada combinación dador (D) – aceptor (A) la desviación de dichas variables respecto a un valor ideal y un valor máximo (Δr_{DA}, $\Delta \alpha_{DA}$) [93]. El siguiente término corresponde a las interacciones que pueden darse entre un metal presente en la proteína y los átomos del ligando que se encuentran como aniones. Además, cuando se encuentran dos o más interacciones con el metal, solo se tiene en cuenta la

de mejor energía. Por otra parte, el termino que contiene a H_{rot}, modela la pérdida de entropía conformacional por restricción de los enlaces rotables del ligando tras la unión. El séptimo término, penaliza las conformaciones donde un átomo polar, pero que no produce puentes de hidrógeno, es encontrado sobre una región hidrofóbica de la proteína. Los siguientes dos términos corresponden a las interacciones electrostáticas y de van der Waals, respectivamente. Finalmente, se considera los efectos del solvente, para esto Glide utiliza una puntuación empírica para determinar la exposición de los átomos al solvente [93].

La función de evaluación GlideScore XP [95] se muestra en la ecuación 2.24. Las contribuciones a la ecuación que favorecen la formación del complejo son del tipo electrostáticas (E_{coul}) y de van der Waals (E_{vdW}). Además se incluyen otras interacciones positivas (E_{bind}) enunciados en la ecuación 2.25. En esta última, el término $E_{hyd-enclosure}$ describe el comportamiento de un grupo hidrofóbico del ligando cuando está rodeado por grupos del mismo tipo pertenecientes a la proteína. Este término captura desviaciones del comportamiento promedio, el cual es representado por $E_{phobic-pair}$ (ecuación 2.25). Los términos $E_{hb-nn-motif}$ y $E_{hb-cc-motif}$ puntúan especialmente ciertas contribuciones energéticas por puentes de hidrógeno entre átomos neutros o cargados, respectivamente, más allá de las consideradas en GlideScore SP (denominadas $E_{hb-pair}$ en la ecuación 2.25). E_{PI} considera la formación de los enlaces π – catión y los apilamientos π – π [95].

Las contribuciones que desfavorecen la formación del complejo ($E_{penalty}$ en la ecuación 2.24) se relacionan con la pérdida de entropía del ligando y la proteína, efectos de desolvatación y de tensión estérica. Estos efectos se agrupan en dos términos definidos como E_{desolv} y $E_{ligand-strain}$ en la ecuación 2.26 [95].

$$GlideScoreXP = E_{coul} + E_{vdW} + E_{bind} + E_{penalty} \tag{2.24}$$

$$E_{bind} = E_{hyd-enclosure} + E_{hb-cc-motif} + E_{PI} + E_{hb-pair} + E_{phobic-pair} \tag{2.25}$$

$$E_{penalty} = E_{desolv} + E_{ligand-strain} \tag{2.26}$$

2.2.1.3. Validación del cribado virtual.

El desempeño y la exactitud de los algoritmos de cribado por docking son evaluados comúnmente utilizando un conjunto de testeo. Este incluye un subconjunto de compuestos que efectivamente interaccionan con el blanco molecular y cuya actividad es conocida, combinado con otro de estructuras que no presentan dicha actividad. Así, puede evaluarse la capacidad de los programas de docking de reconocer como activos (e inactivos) a aquellos que lo son dentro del conjunto de testeo. Una forma de cuantificar esta capacidad consiste en el uso de las curvas ROC (receiver operating characteristic). Estas curvas grafican la tasa de verdaderos positivos encontrados por la función de puntuación, en función de la tasa de falsos positivos, en el eje x, según varia el umbral de discriminación [114]. Como la relación de verdaderos positivos es equivalente a la sensibilidad del modelo, y la razón de falsos positivos es igual a 1 - especificidad, las curvas ROC también se conocen como la representación de la sensibilidad en función de (1-especificidad). Para el cálculo de la tasa de los compuestos activos clasificados correctamente (o verdaderos positivos, (VP)), debe hacerse el cociente entre este valor y los compuestos activos totales del conjunto de testeo (verdaderos positivos (VP) y falsos negativos (FN)), como se muestra en la ecuación 2.27. Para el cálculo de la proporción de falsos positivos clasificados por el modelo (FP) puede definirse la especificidad como la razón entre los verdaderos negativos clasificados correctamente por el modelo (VN) y los inactivos totales (verdaderos negativos (VN) y falsos positivos (FP)), de acuerdo a la ecuación 2.28.

$$\text{Sensibilidad } = \frac{VP}{(VP + FN)} \tag{2.27}$$

$$\text{Especificidad } = \frac{VN}{(VN + FP)} \tag{2.28}$$

Es decir que para la generación de cada punto de la curva se establece un valor de la función de evaluación del docking como umbral. En la figura 3.6 se muestra un

ejemplo de la función distribución de un conjunto de testeo, y se marca un valor de corte. Aquellos compuestos cuya función de evaluación sea menor que dicho valor se predicen activos, mientras que los que presentan una puntuación mayor se consideran inactivos. Luego, como las actividades son conocidas en el conjunto de testeo, pueden calcularse la sensibilidad y 1-especificidad en base a ese valor de corte, y se grafica dicho punto. Se repite el proceso para todos los valores umbral, de manera de construir la curva [114]. Una predicción al azar se representa como una recta desde el eje de coordenadas (0, 0) al valor máximo de ambos ejes (1, 1) y se asocia a un área bajo la curva de 0,5. A medida que el área bajo la curva del modelo se acerca al valor 1,0 (clasificación perfecta), mayor será su capacidad discriminante [114].

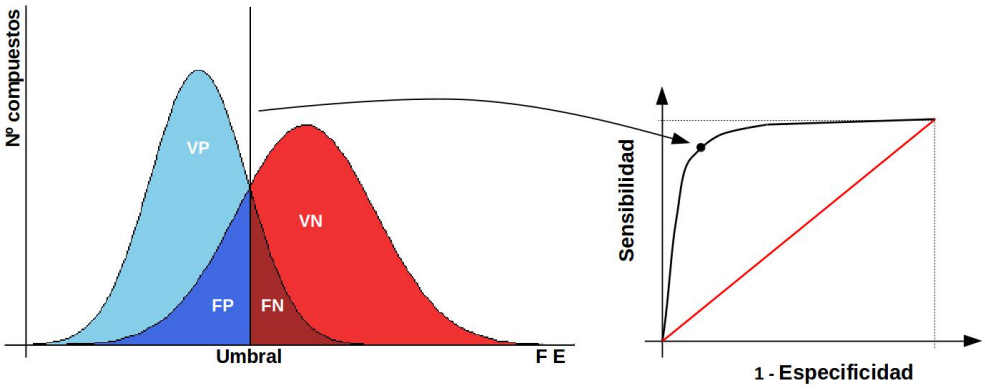


Figura 2.9: Esquema de construcción de una curva ROC.

Otra de las medidas esenciales para evaluar el funcionamiento de un modelo de docking molecular consiste en demostrar la capacidad del mismo para distinguir el modo de unión correcto de un ligando conocido. Así, si se cuenta con la estructura experimental del complejo ligando – receptor, una de las estrategias más comunes es analizar la capacidad de reproducir la conformación de dicho complejo con el modelo desarrollado [61]. Una predicción correcta del modo de unión es comúnmente determinada por el valor de RMSD (root-mean-square deviation). En este caso, se mide la distancia promedio

entre los átomos de la pose mejor puntuada y los correspondientes de la conformación experimental observada [61]. Si el docking puede reproducir exactamente la conformación experimental, el valor de RMSD es cero. Suele adoptarse como criterio general que si el RMSD es menor o igual a $1,5\mathring{A}$ la predicción es considerada como correcta. Una vez validado el modelo, este puede utilizarse sobre una base de datos, donde el número de moléculas activas (y sus modos de unión) es desconocido.

Bibliografía

[1] H. A. Lim, "Bioinformatics and cheminformatics in the drug discovery cycle," in *Bioinformatics*, pp. 30–43, Springer, 1997.

[2] G. L. Patrick, *An introduction to medicinal chemistry*. Oxford university press, 2013.

[3] P. L. Herrling, "The drug discovery process.," *Progress in drug research*, vol. 62, pp. 1–14, 2005.

[4] Z. Cai, G. Zhang, X. Zhang, Y. Liu, and X. Fu, "Current insights into computer-aided immunotherapeutic design strategies," *International Journal of Immunopathology and Pharmacology*, vol. 28, no. 3, pp. 278–285, 2015.

[5] N. Prakash and P. Devangi, "Drug Discovery," *Journal of Antivirals Antiretrovirals*, vol. 02, no. 04, pp. 475–560, 2010.

[6] J. Drews, "Drug Discovery: A Historical Perspective," *Science*, vol. 287, no. 5460, pp. 1960–1964, 2000.

[7] J. Hughes, S. Rees, S. Kalindjian, and K. Philpott, "Principles of early drug discovery," *British Journal of Pharmacology*, vol. 162, pp. 1239–1249, mar 2011.

[8] V. Umashankar and S. Gurunathan, "Drug discovery: an appraisal," *International Journal of Pharmacy and Pharmaceutical Sciences*, vol. 7, no. 4, pp. 59–66, 2015.

[9] A. Mohd and D. A. Jyoti, "CURRENT TRENDS IN DRUG DISCOVERY : TARGET IDENTIFICATION TO CLINICAL DEVELOPMENT OF THE DRUG," *International Research Journal of Pharmacy*, vol. 3, no. 4, pp. 23–27, 2012.

[10] P. Tollman, P. Guy, J. Altshuler, A. Flanagan, and M. Steiner, "A Revolution in RD. How Genomics and Genetics Are Tranforming the Biopharmaceutical Industry," tech. rep., 2001.

[11] S. Kalyaanamoorthy and Y. P. P. Chen, "Structure-based drug design to augment hit discovery," *Drug Discovery Today*, vol. 16, no. 17-18, pp. 831–839, 2011.

[12] G. Sliwoski, S. Kothiwale, J. Meiler, and E. Lowe, "Computational Methods in Drug Discovery," *Pharmacological reviews*, vol. 66, pp. 334–395, 2014.

[13] C. M. Song, S. J. Lim, and J. C. Tong, "Recent advances in computer-aided drug design," *Briefings in Bioinformatics*, vol. 10, no. 5, pp. 579–591, 2009.

[14] I. Muegge and S. Oloff, "Advances in virtual screening," *Drug Discovery Today: Technologies*, vol. 3, no. 4, pp. 405–411, 2006.

[15] M. N. Drwal and R. Griffith, "Combination of ligand- and structure-based methods in virtual screening," *Drug Discovery Today: Technologies*, vol. 10, no. 3, pp. e395–e401, 2013.

[16] P. Willett, "Similarity-based virtual screening using 2D fingerprints," *Drug Discovery Today*, vol. 11, no. 23-24, pp. 1046–1053, 2006.

[17] G. Maggiora, M. Vogt, D. Stumpfe, and J. Bajorath, "Molecular similarity in medicinal chemistry," *Journal of Medicinal Chemistry*, vol. 57, no. 8, pp. 3186–3204, 2014.

[18] J. Auer and J. Bajorath, "Molecular similarity concepts and search calculations," in *Bioinformatics* (J. Keith, ed.), vol. 453 of *Methods in Molecular Biology*^{TM}, pp. 327–347, Humana Press, 2008.

[19] C. Wermuth, C. Ganellin, P. Lindberg, and L. Mitscher, "Glossary of terms used in medicinal chemistry (iupac recommendations 1998)," *Pure and Applied Chemistry*, vol. 70, no. 5, pp. 1129–1143, 1998.

[20] G. Wolber, T. Seidel, F. Bendix, and T. Langer, "Molecule-pharmacophore superpositioning and pattern matching in computational drug design," *Drug Discovery Today*, vol. 13, no. 1-2, pp. 23–29, 2008.

[21] S.-Y. Yang, "Pharmacophore modeling and applications in drug discovery: challenges and recent advances.," *Drug discovery today*, vol. 15, no. 11-12, pp. 444–450, 2010.

[22] B. Waszkowycz, D. E. Clark, and E. Gancia, "Outstanding challenges in protein-ligand docking and structure-based virtual screening," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 1, no. 2, pp. 229–259, 2011.

[23] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The Protein Data Bank.," *Nucleic acids research*, vol. 28, no. 1, pp. 235–242, 2000.

[24] J. S. Fetrow, a. Giammona, a. Kolinski, and J. Skolnick, "The protein folding problem: a biophysical enigma.," *Current pharmaceutical biotechnology*, vol. 3, no. 4, pp. 329–47, 2002.

[25] C. B. Anfinsen, "Principles that govern the folding of protein chains.," *Science (New York, N.Y.)*, vol. 181, no. 96, pp. 223–230, 1973.

[26] J. Xiong, *Essential Bioinformatics*. Cambridge University Press, 2006.

[27] K. Ginalski, N. V. Grishin, A. Godzik, and L. Rychlewski, "Practical lessons from protein structure prediction," *Nucleic Acids Research*, vol. 33, no. 6, pp. 1874–1891, 2005.

[28] D. Baker and A. Sali, "Protein structure prediction and structural genomics," *Science*, vol. 294, no. October, pp. 93–96, 2001.

[29] C. N. Cavasotto and S. S. Phatak, "Homology modeling in drug discovery: current trends and applications," *Drug Discovery Today*, vol. 14, pp. 676–683, jul 2009.

[30] D. J. Lipman and W. R. Pearson, "Rapid and sensitive protein similarity searches.," *Science*, vol. 227, no. 4693, pp. 1435–1441, 1985.

[31] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool.," *Journal of molecular biology*, vol. 215, no. 3, pp. 403–410, 1990.

[32] M. O. Dayhoff and R. M. Schwartz, "Chapter 22: A model of evolutionary change in proteins," in *in Atlas of Protein Sequence and Structure*, 1978.

[33] S. Henikoff and J. G. Henikoff, "Amino acid substitution matrices from protein blocks.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 22, pp. 10915–10919, 1992.

[34] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, "Gapped BLAST and PS I-BLAST: a new generation of protein database search programs," *Nucleic Acids Res*, vol. 25, no. 17, pp. 3389–3402, 1997.

[35] E. Greenbaum, *Biological and medical physics, biomedical engineering*. Springer Science & Business Media, 2004.

[36] S. Batzoglou, "The many faces of sequence alignment," *BRIEFINGS IN BIOINFORMATICS*, vol. 6, no. 1, pp. 6–22, 2005.

[37] "The statistics of sequence similarity scores." http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html. Accedido: 19-01-2016.

[38] G. H. Lushington, "Comparative modeling of proteins," in *Molecular Modeling of Proteins*, pp. 199–212, Springer, 2008.

[39] S. B. Needleman and C. D. Wunsch, "a General Method Applicable To Search for Similarities in Amino Acid Sequence of 2 Proteins," *Journal of Molecular Biology*, vol. 48, no. 3, pp. 443–453, 1970.

[40] T. Smith and M. Waterman, "Identification of common molecular subsequences," *Journal of Molecular Biology*, vol. 147, no. 1, pp. 195–197, 1981.

[41] J. Greer, "Comparative model-building of the mammalian serine proteases," *Journal of Molecular Biology*, vol. 153, pp. 1027–1042, dec 1981.

[42] Z. Xiang, "Advances in homology protein structure modeling.," *Curr. Protein Pept. Sci.*, vol. 7, no. 3, pp. 217–27, 2006.

[43] A. Sali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints.," *Journal of molecular biology*, vol. 234, no. 3, pp. 779–815, 1993.

[44] A. Fiser and A. Sali, "ModLoop: Automated modeling of loops in protein structures," *Bioinformatics*, vol. 19, no. 18, pp. 2500–2501, 2003.

[45] M. Martí-Renom and A. Stuart, "Comparative protein structure modeling of genes and genomes," *Annual review of . . .*, pp. 291–325, 2000.

[46] R. Li, *Determining the molecular structure and function of sodium channel*. PhD thesis, 1999.

[47] J. Wojcik, J.-P. Mornon, and J. Chomilier, "New efficient statistical sequence-dependent structure prediction of short to medium-sized protein loops based on an exhaustive loop classification 1," *Journal of Molecular Biology*, vol. 289, no. 5, pp. 1469 – 1490, 1999.

[48] Q. Zheng and D. J. Kyle, "Accuracy and reliability of the scaling-relaxation method for loop closure: An evaluation based on extensive and multiple copy conformational samplings," *Proteins: Structure, Function, and Bioinformatics*, vol. 24, no. 2, pp. 209–217, 1996.

[49] C. S. Rapp and R. a. Friesner, "Prediction of loop geometries using a generalized born model of solvation effects.," *Proteins*, vol. 35, no. 2, pp. 173–83, 1999.

[50] A. Fiser, A. Sali, and S. Biology, "Comparative protein structure modeling," vol. 1, 2001.

[51] Z. Xiang, C. S. Soto, and B. Honig, "Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 7432–7, may 2002.

[52] M. P. Jacobson, D. L. Pincus, C. S. Rapp, T. J. Day, B. Honig, D. E. Shaw, and R. A. Friesner, "A hierarchical approach to all-atom protein loop prediction," *Proteins: Structure, Function, and Bioinformatics*, vol. 55, no. 2, pp. 351–367, 2004.

[53] K. Fidelis, P. S. Stern, D. Bacon, and J. Moult, "Comparison of systematic search and database methods for constructing segments of protein structure," *Protein Engineering*, vol. 7, no. 8, pp. 953–960, 1994.

[54] V. Vyas, R. Ukawala, M. Ghate, and C. Chintha, "Homology modeling a fast tool for drug discovery: current perspectives," *Indian journal of pharmaceutical sciences*, vol. 74, no. 1, p. 1, 2012.

[55] R. W. Hooft, C. Sander, and G. Vriend, "Objectively judging the quality of a protein structure from a Ramachandran plot.," *Computer applications in the biosciences : CABIOS*, vol. 13, pp. 425–30, aug 1997.

[56] R. A. Laskowski, "PROCHECK: a program to check the stereochemicai quality of protein structures.," *J. Appl. Cryst.*, vol. 26, pp. 283–291, 1993.

[57] R. W. W. Hooft, G. Vriend, C. Sander, and E. E. Abola, "Errors in protein structures," 1996.

[58] R. Lüthy, J. U. Bowie, and D. Eisenberg, "Assessment of protein models with three-dimensional profiles.," *Nature*, vol. 356, no. 6364, pp. 83–85, 1992.

[59] M. Sippl, "Recognition of errors in threedimensional structures of proteins," *Proteins: Structure, Function, and Bioinformatics*, 1993.

[60] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath, "Docking and scoring in virtual screening for drug discovery: methods and applications.," *Nature reviews. Drug discovery*, vol. 3, pp. 935–49, nov 2004.

[61] R. Dias and W. F. de Azevedo, "Molecular docking algorithms.," *Current drug targets*, vol. 9, no. 12, pp. 1040–1047, 2008.

[62] P. A. Ravindranath, S. Forli, D. S. Goodsell, A. J. Olson, and M. F. Sanner, "AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility," *PLOS Computational Biology*, vol. 11, no. 12, p. e1004586, 2015.

[63] X. Meng, H. Zhang, M. Mezei, and M. Cui, "Molecular docking: a powerful approach for structure-based drug discovery," *... computer-aided drug design*, vol. 7, no. 2, pp. 146–157, 2011.

[64] F. Sousa, P. A. Fernandes, and M. Joa, "Protein – Ligand Docking : Current Status and Future," *Bioinformatics*, vol. 26, no. February, pp. 15–26, 2006.

[65] A. Leach, "Molecular modelling: principles and applications," 1996.

[66] W. J. Allen, T. E. Balius, S. Mukherjee, S. R. Brozell, D. T. Moustakas, P. T. Lang, D. A. Case, I. D. Kuntz, and R. C. Rizzo, "DOCK 6: Impact of new features and current docking performance.," *Journal of computational chemistry*, vol. 36, no. 15, pp. 1132–56, 2015.

[67] M. Rarey, B. Kramer, T. Lengauer, and G. Klebe, "A fast flexible docking method using an incremental construction algorithm.," *Journal of molecular biology*, vol. 261, no. 3, pp. 470–89, 1996.

[68] W. Welch, J. Ruppert, and A. N. Jain, "Hammerhead : fast , fully automated protein binding sites docking of flexible ligands to," *Chemistry and Biology*, vol. 3, pp. 449–462, 1996.

[69] L. Ferreira, R. dos Santos, G. Oliva, and A. Andricopulo, *Molecular Docking and Structure-Based Drug Design Strategies*, vol. 20. 2015.

[70] M. McGann, "FRED pose prediction and virtual screening accuracy," *Journal of Chemical Information and Modeling*, vol. 51, no. 3, pp. 578–596, 2011.

[71] D. S. Goodsell, H. Lauble, C. D. Stout, and A. J. Olson, "Automated docking in crystallography: Analysis of the substrates of aconitase," *Proteins: Structure, Function, and Bioinformatics*, vol. 17, no. 1, pp. 1–10, 1993.

[72] T. N. Hart and R. J. Read, "A multiple-start monte carlo docking method," *Proteins: Structure, Function, and Bioinformatics*, vol. 13, no. 3, pp. 206–222, 1992.

[73] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[74] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, no. 4598, pp. pp. 671–680, 1983.

[75] D. S. Goodsell and a. J. Olson, "Automated docking of substrates to proteins by simulated annealing.," *Proteins*, vol. 8, pp. 195–202, jan 1990.

[76] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, no. 2, pp. 65–85, 1994.

[77] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function," *Journal of Computational Chemistry*, vol. 19, pp. 1639–1662, nov 1998.

[78] G. Morris, R. Huey, and W. Lindstrom, "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility," *Journal of*, vol. 30, no. 16, pp. 2785–2791, 2009.

[79] G. Morris and M. Lim-Wilby, "Molecular docking," in *Molecular Modeling of Proteins* (A. Kukol, ed.), vol. 443 of *Methods Molecular Biology^{TM}*, pp. 365–382, Humana Press, 2008.

[80] G. Ermondi and G. Caron, "Recognition forces in ligand-protein complexes: blending information from different sources.," *Biochemical pharmacology*, vol. 72, pp. 1633–45, dec 2006.

[81] C. a. MacRaild, a. H. Daranas, A. Bronowska, and S. W. Homans, "Global Changes in Local Protein Dynamics Reduce the Entropic Cost of Carbohydrate Binding in the Arabinose-binding Protein," *Journal of Molecular Biology*, vol. 368, no. 3, pp. 822–832, 2007.

[82] A. K. Bronowska, "Thermodynamics of ligand-protein interactions: Implications for molecular design,"

[83] S. Y. Huang, S. Z. Grinter, and X. Zou, "Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions.," *Phys. Chem. Chem. Phys.*, vol. 12, pp. 12899–12908, 2010.

[84] S. Z. Grinter and X. Zou, "Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design.," *Molecules (Basel, Switzerland)*, vol. 19, pp. 10150–76, jan 2014.

[85] R. Huey, G. Morris, A. Olson, and D. Goodsell, "A semiempirical free energy force field with charge-based desolvation," *Journal of computational chemistry*, vol. 28, no. 6, pp. 1145–1152, 2007.

[86] E. Meng, B. Shoichet, and I. Kuntz, "Automated docking with gridbased energy evaluation," *Journal of Computational ...*, vol. 13, no. 4, pp. 505–524, 1992.

[87] G. Jones, P. Willett, R. C. Glen, a. R. Leach, and R. Taylor, "Development and validation of a genetic algorithm for flexible docking.," *Journal of molecular biology*, vol. 267, no. 3, pp. 727–748, 1997.

[88] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, "Improved protein–ligand docking using gold," *Proteins: Structure, Function, and Bioinformatics*, vol. 52, no. 4, pp. 609–623, 2003.

[89] P. K. Weiner and P. A. Kollman, "Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions," *Journal of Computational Chemistry*, vol. 2, no. 3, pp. 287–303, 1981.

[90] S. J. Weiner, P. a. Kollman, D. a. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weinerl, "A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins," *Journal of the American Chemical Society*, vol. 106, no. 17, pp. 765–784, 1984.

[91] H.-J. Böhm, "Ludi: rule-based automatic design of new substituents for enzyme inhibitor leads," *Journal of Computer-Aided Molecular Design*, vol. 6, no. 6, pp. 593–606.

[92] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, "Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes," *Journal of Computer-Aided Molecular Design*, vol. 11, no. 5, pp. 425–445.

[93] R. a. Friesner, J. L. Banks, R. B. Murphy, T. a. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin, "Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.," *Journal of medicinal chemistry*, vol. 47, pp. 1739–49, mar 2004.

[94] T. a. Halgren, R. B. Murphy, R. a. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, and J. L. Banks, "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening," *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1750–1759, 2004.

[95] R. Friesner, R. Murphy, M. Repasky, L. Frye, J. Greenwood, T. Halgren, P. Sanschagrin, and D. Mainz, "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes," *J. Med. Chem.*, no. 49, pp. 6177–6196, 2006.

[96] R. Wang, L. Lai, and S. Wang, "Further development and validation of empirical scoring functions for structure-based binding affinity prediction," *Journal of Computer-Aided Molecular Design*, vol. 16, no. 1, pp. 11–26, 2002.

[97] X. Li and J. Liang, "Knowledge-based energy functions for computational studies of proteins," in *Computational Methods for Protein Structure Prediction and Modeling* (Y. Xu, D. Xu, and J. Liang, eds.), BIOLOGICAL AND MEDICAL PHYSICS BIOMEDICAL ENGINEERING, pp. 71–123, Springer New York, 2007.

[98] A. N. Jain, "Scoring Functions for Protein-Ligand Docking," *Current Protein Peptide Science*, vol. 7, no. 5, pp. 407–420, 2006.

[99] C. Zhang, S. Liu, Q. Zhu, and Y. Zhou, "A knowledge-based energy function for protein-ligand, protein-protein, and protein-DNA complexes," *Journal of Medicinal Chemistry*, vol. 48, no. 7, pp. 2325–2335, 2005.

[100] H. Gohlke, M. Hendlich, and G. Klebe, "Knowledge-based scoring function to predict protein-ligand interactions.," *Journal of Molecular Biology*, vol. 295, pp. 337–356, 2000.

[101] H. F. G. Velec, H. Gohlke, and G. Klebe, "DrugScoreCSD-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction," *Journal of Medicinal Chemistry*, vol. 48, no. 20, pp. 6296–6303, 2005.

[102] S.-Y. Huang and X. Zou, "An iterative knowledge-based scoring function to predict protein–ligand interactions: I. Derivation of interaction potentials," *Journal of Computational Chemistry*, vol. 27, pp. 1866–1875, nov 2006.

[103] S.-Y. Huang and X. Zou, "An iterative knowledge-based scoring function to predict protein–ligand interactions: II. Validation of the scoring function," *Journal of Computational Chemistry*, vol. 27, pp. 1876–1882, nov 2006.

[104] I. Muegge, Y. C. Martin, P. J. Hajduk, and S. W. Fesik, "Evaluation of PMF scoring in docking weak ligands to the FK506 binding protein.," *Journal of medicinal chemistry*, vol. 42, no. 14, pp. 2498–503, 1999.

[105] R. S. DeWitte and E. I. Shakhnovich, "SMoG: De novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence," *Journal of the American Chemical Society*, vol. 118, no. 1, pp. 11733–11744, 1996.

[106] P. S. Charifson, J. J. Corkery, M. a. Murcko, and W. P. Walters, "Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins," *Journal of Medicinal Chemistry*, vol. 42, no. 25, pp. 5100–5109, 1999.

[107] A. Oda, K. Tsuchida, T. Takakura, N. Yamaotsu, and S. Hirono, "Comparison of consensus scoring strategies for evaluating computational models of protein-ligand complexes," *Journal of Chemical Information and Modeling*, vol. 46, no. 1, pp. 380–391, 2006.

[108] G. E. Terp, B. N. Johansen, I. T. Christensen, and F. S. Jorgensen, "A new concept for multidimensional selection of ligand conformations (MultiSelect) and multidimensional scoring (MultiScore) of protein-ligand binding affinities," *J Med Chem*, vol. 44, no. 14, pp. 2333–2343, 2001.

[109] S. Betzi, K. Suhre, B. Chétrit, F. Guerlesquin, and X. Morelli, "GFscore: a general nonlinear consensus scoring function for high-throughput docking," *Journal of chemical information and modeling*, vol. 46, no. 4, pp. 1704–1712, 2005.

[110] R. Teramoto and H. Fukunishi, "Supervised consensus scoring for docking and virtual screening," *Journal of Chemical Information and Modeling*, vol. 47, no. 2, pp. 526–534, 2007.

[111] S. Bar-Haim, A. Aharon, T. Ben-Moshe, Y. Marantz, and H. Senderowitz, "SeleX-CS: a new consensus scoring algorithm for hit discovery and lead optimization.," *Journal of chemical information and modeling*, vol. 49, no. 3, pp. 623–33, 2009.

[112] O. Trott and A. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of computational chemistry*, vol. 31, no. 2, pp. 455–461, 2010.

[113] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, "Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes.," *Journal of computer-aided molecular design*, vol. 11, pp. 425–445, 1997.

[114] N. Triballeau, F. Acher, I. Brabet, J.-P. Pin, and H.-O. Bertrand, "Virtual screening workflow development guided by the receiver operating characteristicçurve

approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4.," *Journal of medicinal chemistry*, vol. 48, pp. 2534–47, apr 2005.

Capítulo 3

Modelo para la predicción de la afinidad de compuestos a P-gp

3.1. Modelo tridimensional de la P-gp humana.

Como se detalló anteriormente, uno de los objetivos de este trabajo es la predicción de la afinidad de los compuestos a la glicoproteína P (P-gp), mediante una aproximación basada en la estructura del receptor (utilizando métodos de docking molecular). Dado que no existen actualmente estructuras tridimensionales experimentales de la P-gp humana, se debe recurrir a modelos generados por métodos computacionales. Para ello se utilizó la secuencia de la P-gp (obtenida a partir de la base de datos UniProt [1], código P08183) como criterio de búsqueda en el servidor I-TASSER [2, 3, 4]. Este servidor predice la estructura y función de las secuencias proteicas ingresadas mediante la utilización de métodos de threading.

Al introducir una secuencia de aminoácidos, I-TASSER inicialmente busca en la base de datos Protein Data Bank (PDB) macromoléculas con secuencias similares mediante LOMETS (Local Meta-Threading-Server). Esta aplicación predice la estructura proteica mediante la clasificación y selección de modelos generados a partir de nueve programas de threading [5]. De esta manera se obtiene un conjunto de proteínas molde cuyo ple-

gamiento debe ser adquirido por la secuencia problema. Posteriormente los modelos son cortados y re-ensamblados, y las regiones no alineadas por threading son construidas por métodos Ab-Initio [4]. La etapa siguiente implica una optimización de geometría a través de simulaciones Replica Exchange Monte-Carlo guiadas por un campo de fuerza estadístico compuesto por tres componentes principales: 1) potenciales estadísticos genéricos; 2) optimización de la red de puentes de hidrógeno; 3) restricciones derivadas del proceso de threading [4]. Los confórmeros con menor energía son seleccionados y agrupados de acuerdo a su similitud estructural por el software SPICKER [6]. Cada uno de los grupos de modelos se caracteriza por una estructura de referencia (o centroide) que se utiliza para una segunda simulación Replica Exchange Monte-Carlo para remover contactos estéricos entre los átomos. La segunda simulación nuevamente clasifica los modelos y se seleccionan las estructuras de menor energía de los clusters más numerosos [3].

La confianza de cada modelo es medida en forma cuantitativa mediante el valor C-Score, calculado en base a los alineamientos de threading y a los parámetros de convergencia de la simulación. Este parámetro estima la calidad de los modelos predichos, basándose en la significancia de los alineamientos realizados sobre cada molde luego del threading [2]. Generalmente toma valores entre -5 y 2, donde un valor mayor indica un modelo más confiable [2]. Otro parámetro que también se utiliza es el valor de TM-score, el cual mide la similitud estructural entre dos macromoléculas. Se encuentra relacionado con el parámetro RMSD (root-mean-square deviation), pero presenta menor sensibilidad a errores locales (por ejemplo una cadena lateral ubicada en forma diferente que en la proteína molde), que podrían llevar a suponer que un modelo es incorrecto cuando en su topología global si lo es. Un TM-score menor a 0,17 indica que la similaridad global es muy pequeña y por lo tanto el modelo es incorrecto [7].

El servidor reporta los mejores cinco modelos que corresponden a las cinco agrupaciones más numerosas. Si la simulación converge, es posible que se obtengan menos de cinco clusters. Esto es usualmente un indicativo de buena calidad del modelo.

Particularmente para el caso del modelado de P-gp, en la etapa inicial de búsqueda

de estructuras moldes, I-TASSER devolvió 10 alineamientos correspondientes a macromoléculas de P-gp de ratón (códigos PDB: 3G61 y 3G5U) [8], del transportador humano mitocondrial ABCB10 (código PDB: 2YL4) y del transportador ABC MsbA de salmonella typhimurium (código PDB 3B60) (tabla 3.1).

Modelo	PDB	Iden 1	Iden 2	Cobertura	Z-score
1	3G61 A	0.88	0.82	0.92	7.81
2	3G5U A	0.89	0.82	0.92	2.36
3	3G5U A	0.89	0.82	0.92	11.38
4	3G5U A	0.87	0.82	0.78	8.14
5	3G61 A	0.88	0.82	0.92	11.05
6	3G5U A	0.89	0.82	0.92	2.06
7	3G5U A	0.89	0.82	0.92	3.61
8	3G61 A	0.89	0.82	0.44	8.84
9	2YL4 A	0.35	0.17	0.44	3.36
10	3B60 A	0.29	0.15	0.45	2.35

Tabla 3.1: Resultados del alineamiento de secuencias mediante threading.
Iden 1: porcentaje de identidad de secuencia del molde en las regiones alineadas con la secuencia problema. Iden 2: porcentaje de identidad de secuencia de toda la cadena molde con la secuencia problema. Cov. Representa la cobertura del alineamiento y es igual al número de residuos alineados dividido el tamaño de la secuencia problema. Z-score: Alineamientos con Z-score > 1 son considerados buenos alineamientos y viceversa

El análisis de los alineamientos de secuencias sugiere que existe un alto grado de conservación entre la P-gp humana y la de ratón, con un porcentaje de identidad del 82 % y un tamaño similar. En base a esta información podemos considerar a estas proteínas como homólogas. La tabla 3.2 muestra los cinco mejores modelos generados I-TASSER y ordenados según el valor C-score.

	Modelo 1	**Modelo 2**	**Modelo 3**	**Modelo 4**	**Modelo 5**
C-score	−0,20	−1,31	−1,43	−2,53	−3,07

Tabla 3.2: Valores de C-score de los modelos obtenidos en I-TASSER.

De acuerdo a los resultados obtenidos, se seleccionó el modelo 1 porque presenta un mayor C-score y un TM score de 0.79 (dentro de los valores aceptados). Posteriormente esta macromolécula fue refinada "relajando" su estructura. Los átomos de la proteína fueron rodeados por una caja periódica de moléculas de agua TIP3P32, ocupando hasta

$10\mathring{A}$ desde la estructura proteica, con el modulo TLEAP del software AMBER11 [9]. La geometría inicial fue minimizada utilizando el modulo SANDER mediante 1000 ciclos para la relajación de las moléculas de agua seguidos de 2500 para el sistema completo usando el campo de fuerza AMBER en su versión ff03.r1 [9]. La estructura final se muestra en la figura 3.1.

Figura 3.1: Estructura tridimensional del modelo de P-gp humana (vista lateral).

Posteriormente, el modelo de P-gp humana fue validado mediante el análisis de los

gráficos de Ramachandran utilizando el servidor Protein Structure Validation Software (PSVS) [10, 11, 12, 13, 14, 15]. En la figura 3.2 puede observarse que los ángulos diedros Ψ y Φ del esqueleto carbonado se encuentran distribuidos principalmente en las zonas favorables (88,5 %) o permitidas (97,9 %). El 2,1 % restante (ubicado en zonas desfavorables), no se encuentra ubicado en la región del sitio activo. Los valores porcentuales obtenidos son similares a los resultados recabados de la estructura experimental de ratón depositada en la Protein Data Bank (PDB: 3G60), donde el 64,9 % de sus residuos están en la zona favorable, mientras que el 91,3 % de los mismos están en regiones permitidas. Generalmente se acepta que un buen modelo tenga al menos un 90 % de sus aminoácidos en las regiones favorables.

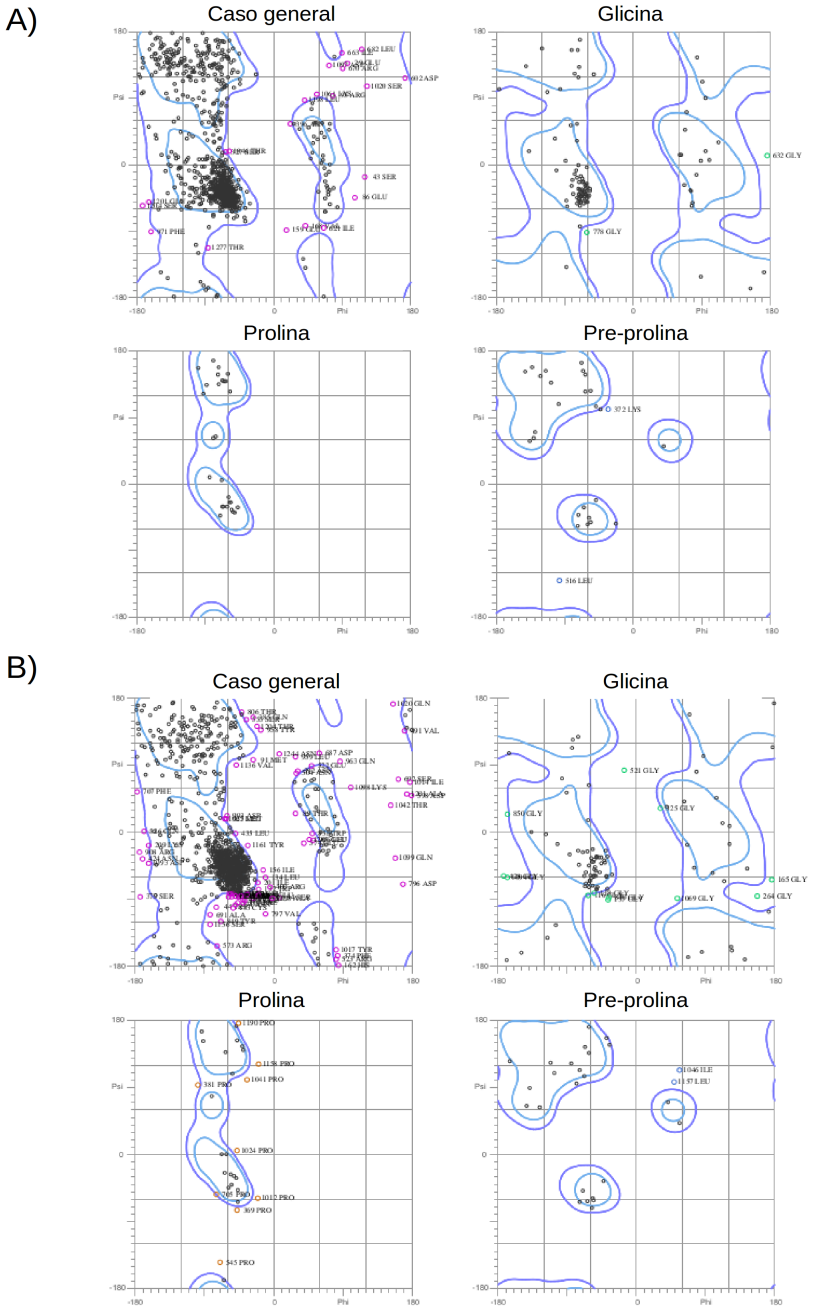


Figura 3.2: Gráficos de Ramachandran: A) Modelo de P-gp humana obtenido con I-TASSER; B) Estructura tridimensional de P-gp de ratón (PDB: 3G60)

Finalmente, para analizar la capacidad del servidor I-TASSER de reproducir los modelos experimentales se reconstruyó la estructura 3D de la P-gp de ratón y se superpuso

con la solución experimental (código PDB 3G60) [8]. Para ello se partió de su secuencia de aminoácidos tomada de la base de datos UniProt (P21447). El modelo con el mejor valor de C-score se superpuso con la estructura experimental, obteniéndose un valor de RMSD de 0,65 (figura 3.3). Estos resultados incrementan nuestra confianza en la capacidad de I-TASSER para predecir la estructura de la P-gp humana.

Figura 3.3: Superposición de la estructura tridimensional de la P-gp de ratón obtenida con I-TASSER (verde) frente a la estructura experimental (rojo).

3.2. Búsqueda bibliográfica y conjunto de evaluación.

Con el modelo de P-gp humano disponible se procedió a la selección del programa y condiciones de docking para la realización del cribado. Como criterio se utilizó la capacidad de dichos softwares en discriminar entre sustratos y no sustratos conocidos de la glicoproteína. En otras palabras, se creó por búsqueda bibliográfica un conjunto de evaluación que contiene compuestos con probada unión a P-gp y otros que no interac-

cionan con la proteína. Aquel programa que sea capaz de clasificarlos correctamente con su función de evaluación será el más confiable para la realización del cribado de bibliotecas de compuestos (cuya interacción con la macromolécula no se conoce). Existen varios tipos de ensayos biológicos que han sido desarrollados para analizar la habilidad de los compuestos de interaccionar con la P-gp [16, 17]. Entre estos, se encuentran las evaluaciones de transporte bidireccional a través de monocapas celulares [18, 19], los ensayos de medida de la actividad ATPasa de la P-gp [20, 21, 16] y los de determinación de la fluorescencia de calceina-AM [22] o de rodamina-123 [23].

Los ensayos de transporte bidireccional consisten en determinar la relación entre la permeabilidad basolateral-apical (B) y la apical-basolateral (A) de la sustancia a evaluar en monocapas celulares, y compararla con la tasa respectiva en presencia de un inhibidor del trasportador [18, 19]. Por otra parte, los test de actividad ATPasa de la P-gp implican un monitoreo de dicha enzima mediante la determinación de fosfato inorgánico liberado por una reacción de colorimetría [20]. En los ensayos de calceina-AM y rodamina-123 (ambos sustratos de la P-gp) se determina la acumulación de fluorescencia en el interior celular como consecuencia de la inhibición de la P-gp [22, 23]. La FDA (Food and Drug Administration) sugiere el ensayo de transporte bidireccional usando cultivos celulares (en ausencia y presencia de uno o más inhibidores) como prueba inicial para decidir si un compuesto es un potencial sustrato o inhibidor de P-gp [24]. Recientemente se ha encontrado que este tipo de ensayos en células Caco-2 se encuentran entre los métodos in vitro más utilizados, seguidos por las evaluaciones en células MDCKII-MDR1 (multi drug resistance protein 1 transfected-Madin Darby Canine Kidney cells) [17].

La preparación de un conjunto de evaluación de calidad y con una cantidad considerable de compuestos es uno de los pasos más importantes para la construcción de un modelo. En este sentido, se puso especial atención en la selección y categorización de las estructuras encontradas en literatura, dado que existen algunas controversias en cuanto a la clasificación en sustratos / inhibidores y no sustratos / no inhibidores. Por ejemplo, Polli y colaboradores clasifican al Verapamilo como no sustrato de P-gp de acuerdo a

experimentos de permeabilidad sobre monocapas de células MDCK [25], mientras que el grupo de Feng lo considera como sustrato en el mismo ensayo [26]. Por otra parte, Doan y coautores reportan la Fluoxetina como no sustrato en experimentos de eflujo en monocapas de células MDCKII-MDR1, pero este compuesto es clasificado como inhibidor en ensayos de inhibición de Calcein-AM (CAM) sobre la misma línea celular [27]. En el mismo sentido, Feng y colaboradores determinaron que Fluoxetina resulta inactiva en ensayos de permeabilidad y de ATPasa, pero que muestra actividad inhibitoria en ensayos CAM [26].

Con el objetivo de preparar un conjunto de evaluación representativo y diverso, se consideraron los resultados biológicos de múltiples publicaciones [25, 26, 28, 29, 30], dando prioridad a ensayos in vitro respecto a ensayos in vivo. Los ensayos in vivo son realizados sobre sistemas complejos (pacientes o animales), haciendo más difícil evaluar si una entidad química es sustrato o no de la P-gp. Adicionalmente, hay más información disponible de estudios in vitro, ya que estos permiten la evaluación de un mayor número de compuestos, en un determinado tiempo, que los ensayos in vivo. Adicionalmente se decidió definir como sustratos a aquellos compuestos que fueran detectados como tales en dos o más publicaciones en diferentes ensayos (si es posible más de tres). Por otra parte, se consideró como no sustratos a aquellos compuestos catalogados de esta manera en 2 o más ensayos, y que no se reportaron como sustratos (o solo se identificaron así en un solo ensayo). Las estructuras que presentaron alguna controversia en los resultados publicados (por ejemplo, un ensayo donde el compuesto fuera considerado sustrato o inhibidor y otro donde se lo clasifique como no sustrato) fueron descartados.

De una recopilación inicial de 53 compuestos, se seleccionaron 26 como sustratos y 13 como no sustratos, descartándose los 14 restantes. La tabla 3.3 muestra los compuestos elegidos y la clasificación obtenida.

Compuesto	monocapa	ATPasa	CAM	Rod123	In-vivo
Amantadina	N [28, 25]	N [25]	N [25]		

Continua en la página siguiente

Carbamacepina	N [30, 28, 26]	N [26]	N [28, 26]	N [30]	N [30]
Clorfeniramina	N [28, 25]	N [25]	N [28, 25]		
Etosuximida	N [30, 26]	N [26]	N [26]		
Felbamato				N[30]	
Fluvoxamina	N [28, 26]	N [26]	P [28] N [26]		
Lidocaina	N [28, 25]	N [25]	N [28, 25]		
Manitol	N [28, 25]	N[25]	N[28, 25]		
Propranolol	N [28, 25]	N [25]	N [28, 25]		
Ranitidina	N[25]	N[29, 25]	N[29, 25]	N[29]	
Sumatriptan	N [28, 25]	N[25]	N [28, 25]		
Topiramato	N[30]				N[30]
Vigabatrina	N[30]				
Amprenavir	P[25, 28]	P[25]	P[25, 27]		
Astemizol	P[28]	P[29]	P [29, 28]	P[29]	
Ciclosporina	P[25]	P[29]N[25]	P[29, 25]	P [29]	
Dexametasona	P[29, 25]	P [29] N [25]	N[29, 25]	N[29]	
Diltiazem	P [28, 25]	P[25]	P [28, 25]		
Indinavir	P [28, 25]	P [25]	N [28, 25]		
Ketoconazol	N[25]	P [29, 25]	P [29, 25]	P [29]	
Lamotrigina	N[26, 30] P [26]	N[26]	N[26]	P[30]	P [30]
Leviteracetam	N[30] P [30]			P[30]	
Loperamida	P[28, 25, 26]	P [25, 26]	P[28, 25, 26]		
Loratadina	P[25] N[26]	P[25, 26]	P[25, 26]		
Nelfinavir	P[25]	P [29, 25]	P[29, 25]	P [29]	
Neostigmina	P [27, 25]	N [25]	N[27, 25]		
Nicardipina	N[25]	P [29, 25]	P[29, 25]		
Oxcarbazepina	P[30]				P[30]N[30]
Fenobarbital	P[30] N[30]	P[30]		P[30]	P[30]
Fenitoina	P[30] N[30, 26]	N[26]	N[26]	N[30]	P [30]
Prazosina	P [25, 26]	P[25, 26]	P[25] N [26]		
Quinidina	P[25, 26]	P [25, 26]	P [25, 26]		
Risperidona	P[27, 26]	P[26]	P[27]N[26]		
Ritonavir	P[25, 26]	P[29, 25, 26]	P[29, 25]N[29, 26]	P[29]	
Saquinavir	P[29, 25]	P[29, 25]	P [29, 25]N[29]	P[29]	
Terfenadina	P[25]	P[25] N[29]	P[29, 25]	N[29]	
Trimetoprima	P[27, 25]	P[25]	N[27, 25]		
Verapamilo	P[29, 26]N[25, 26]	P[29, 25, 26]	P[29, 25, 26]	P[29]	
Vinblastina	P[25]	P[29, 25]	P[29, 25] N[29]	P[29]	

Tabla 3.3: Compuestos que integran el conjunto de evaluación y la clasificación en cada tipo de ensayo según la bibliografía consultada. "P" indica que la estructura se consideró como positiva en los ensayos, mientras que "N" indica los ensayos en los resultados fueron negativos.

3.3. Docking Molecular.

Los compuestos seleccionados como sustratos y no sustratos de P-gp (tabla 3.3) fueron utilizados para el análisis de una serie de modelos de docking molecular. Para esto, se dibujaron las estructuras y se realizó una optimización geométrica de las mismas utilizando el programa Avogadro [31]. En esta última etapa se empleo la técnica de gradientes conjugados y el campo de fuerza MMFF94 [32, 33, 34, 35, 35]. En todas las simulaciones el "sitio activo" utilizado fue definido por una grilla de 24 x 24 x 24 $\AA^3$, que comprende la cavidad presente en la región transmembrana e incluye los subsitios de unión reportados para sustratos e inhibidores [8] (figura 3.4). Para definir el centro de la grilla se utilizaron las coordenadas relativas de los ligandos presentes en las estructuras cristalográficas de P-gp de ratón (códigos PDB: 3G60 y 3G61).

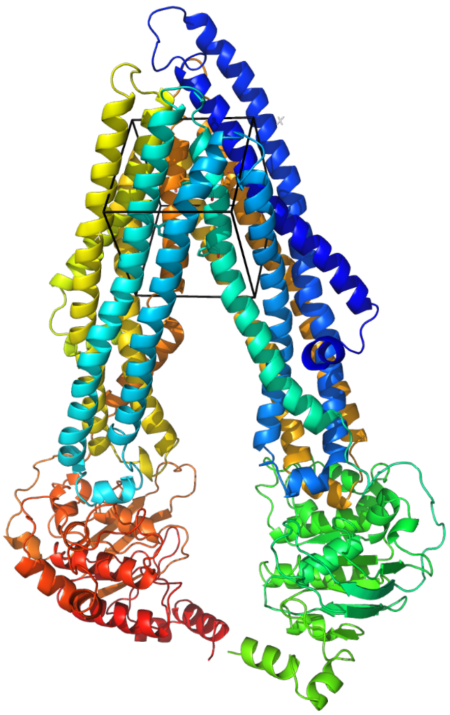


Figura 3.4: Posición del sitio activo definido para el docking

Las simulaciones se realizaron mediante los programas de docking molecular Au-

todock4.2 [36, 37], Autodock Vina [38] y Glide (versión 5.7, Schrodinger Suite 2011) [39, 40, 41].

En Autodock4.2 se utilizó el método de búsqueda conformacional por algoritmos genéticos "Lamarkianos", explicado en el capítulo 2. Los ligandos fueron tratados como flexibles, es decir que se consideraron activos todos los ángulos de torsión a excepción de los que forman parte de anillos y enlaces múltiples. El receptor se consideró rígido en su mayor parte, aunque se permitió la movilidad de un conjunto de residuos ubicados en el sitio de unión (como se detallara más adelante). Se empleó el espaciado de los puntos de la grilla establecido por default en el programa (0.375 $\mathring{A}$) y se realizaron 100 corridas de docking para cada compuesto. El resto de los parámetros del programa se fijaron por defecto, como por ejemplo las cargas parciales de los átomos, que fueron calculadas por el método Marsili-Gasteiger [42]. En el caso de AutoDock Vina, el sitio de unión y los ligandos fueron tratados de modo similar a AutoDock4.2. Se realizaron 20 corridas para cada compuesto utilizando los parámetros por defecto para el resto de las variables. Como se mencionó en el capítulo 2 de esta tesis, AutoDock Vina difiere de AutoDock4.2 en la función de evaluación y en líneas generales se encuentra una mejor performance tanto en términos de velocidad como en exactitud [38]. Sin embargo, en este trabajo se compararon los dos programas considerando que la exactitud de un software de este tipo es usualmente dependiente del sistema a simular (ligando-receptor).

Tanto para las simulaciones con Autodock como con Autodock Vina se permitió la movilidad de algunos residuos del sitio activo. Se definieron dos modelos, que incluyen distintos aminoácidos móviles. En uno de ellos se consideraron como flexibles los residuos Phe-336, Phe-343, Phe-728, Phe-732 y Phe-978 (modelo A), mientras que en el otro conjunto de simulaciones se consideraron los residuos Tyr-307, Tyr-953, Phe-343 y Phe-978 (modelo B). El criterio para la selección de los residuos móviles del modelo A sé baso en el análisis de los aminoácidos que interactúan con el ligando experimental presente en los complejos de ratón (códigos PDB: 3G60 y 3G61). En el caso del modelo B, los aminoácidos capaces de rotar surgen de un análisis retrospectivo del docking con el modelo A y

el conjunto de evaluación. Luego del docking, se encontró que Phe-343 y Phe-978 muestran diferentes conformaciones dependiendo del ligando unido, pero Phe-336, Phe-732 y Phe-728 adoptan prácticamente la misma conformación en su interacción con todos los compuestos activos testeados. Por lo tanto, se eligió otro conjunto de residuos flexibles que incluye Phe-343, Phe-978 pero no los otros aminoácidos móviles del modelo A. El conjunto se completó con dos residuos de Tyr, pues el docking con el modelo A predice una interacción entre dichos residuos y los sustratos del conjunto de evaluación (figura 3.5).

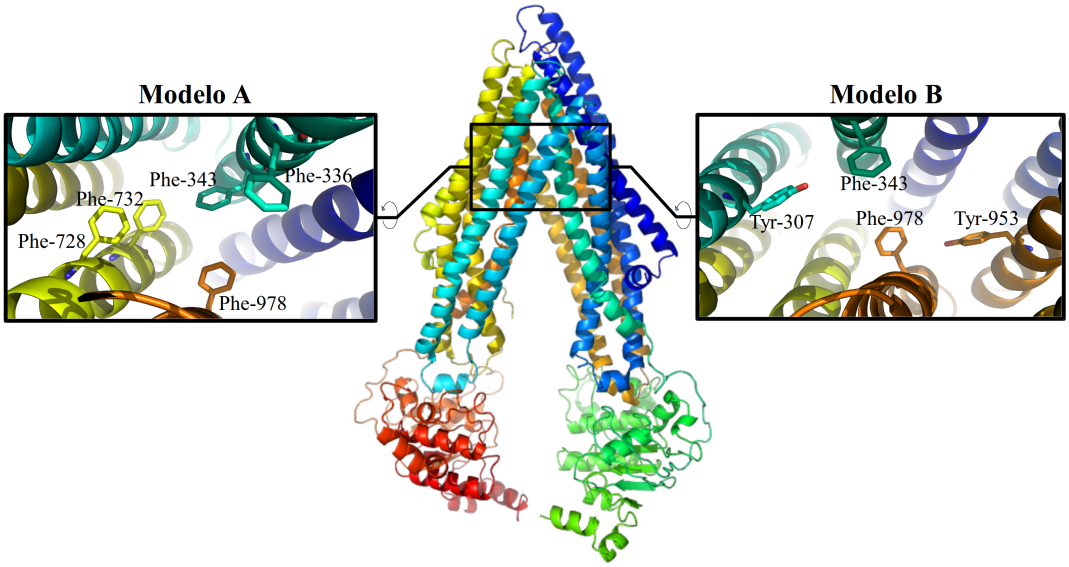


Figura 3.5: Conjuntos de residuos flexibles utilizados en las simulaciones con Autodock y Autodock Vina.

Como se detallará más adelante, con Autodock Vina también se realizó otra simulación con el modelo B, en la que los ligandos fueron protonados a pH fisiológico (pH 7,4). Además, se amplió el numero de compuestos mediante la inclusión de un grupo adicional de no sustratos al conjunto original.

Finalmente, en Glide se utilizó un procedimiento estándar. Los compuestos del conjunto de evaluación fueron preparados usando la herramienta Ligprep del paquete Schrodin-

ger11 para generar los posibles estados de protonación a diferentes pH y conformaciones de baja energía. Se utilizaron los protocolos "Standard precision (SP)" y "Extra precision (XP)" guardando las tres mejores conformaciones de cada compuesto en la corrida SP para construir la librería usada en la corrida XP.

Programa	**Residuos flexibles**	**Protonación**	**Nombre**	**AUC**
AutoDock4.2	Modelo A	Neutro	Simulación 1	0,737
AutoDock4.2	Modelo B	Neutro	Simulación 2	0,747
Autodock Vina	Modelo A	Neutro	Simulación 3	0,790
Autodock Vina	Modelo B	Neutro	Simulación 4	0,833
Autodock Vina	Modelo B	pH fisiológico	Simulación 5	0,833
Glide	Rígido	todos	Simulación 6	0,737

Tabla 3.4: Condiciones y funciones de evaluación en las simulaciones de docking molecular. AUC: área bajo la curva (Area Under the Curve)

Los resultados de cada simulación se compararon mediante el uso de las curvas ROC (Figura 3.6). En la columna 5 de la tabla 3.4 se muestran los valores de área bajo la curva (AUC, Area Under the Curve) para cada caso. Como era de esperarse, los mejores resultados se obtuvieron mediante modelos de docking flexible (residuos móviles: Tyr-307, Tyr-953, Phe-343 y Phe-978), en particular mediante el uso de la función de evaluación de Autodock Vina (simulaciones 4 y 5). Estas dos simulaciones solo difieren en el estado de protonación de los ligandos. La observación de que el estado de ionización de los sustratos no es importante en el proceso de docking es consistente con la naturaleza lipofílica del sitio activo, formado principalmente por residuos hidrofóbicos y aromáticos [8].

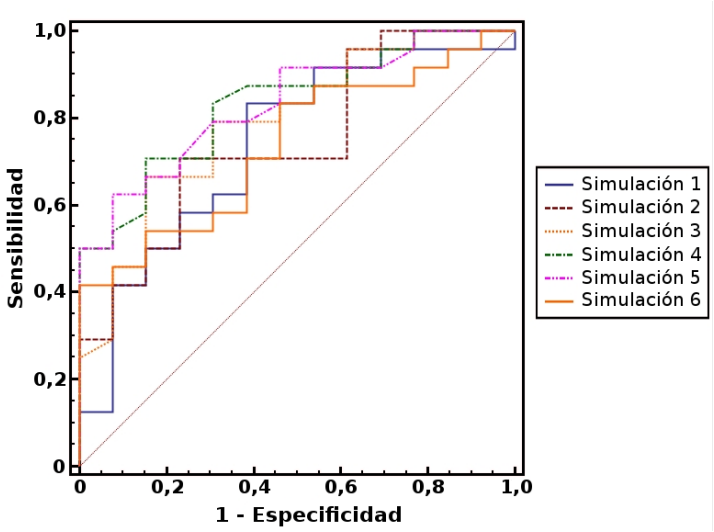


Figura 3.6: Curvas ROC obtenidas de las simulaciones de docking.

La figura 3.7 muestra el resultado obtenido de la simulación de Saquinavir en el sitio activo de la P-gp, como un ejemplo de las interacciones características encontradas en los sustratos. Este compuesto presentó uno de los mejores valores de función de evaluación en las simulaciones (energías de unión predicha más negativa), en relación a las estructuras del conjunto. Se observan interacciones de puentes de hidrógeno débiles con el enlace peptídico Val-982, pero el complejo se encuentra principalmente estabilizado mediante interacciones lipofílicas y apilamientos aromáticos π-π con residuos alifáticos y aromáticos que rodean la molécula. Esto explica el hecho de que los compuestos del conjunto no modifiquen sustancialmente la interacción predicha con el sitio de unión de acuerdo a su estado de protonación. Quizás la ionización afecte la habilidad de los sustratos para llegar al sitio activo, pero este fenómeno no es tenido en cuenta en la simulación de docking molecular.

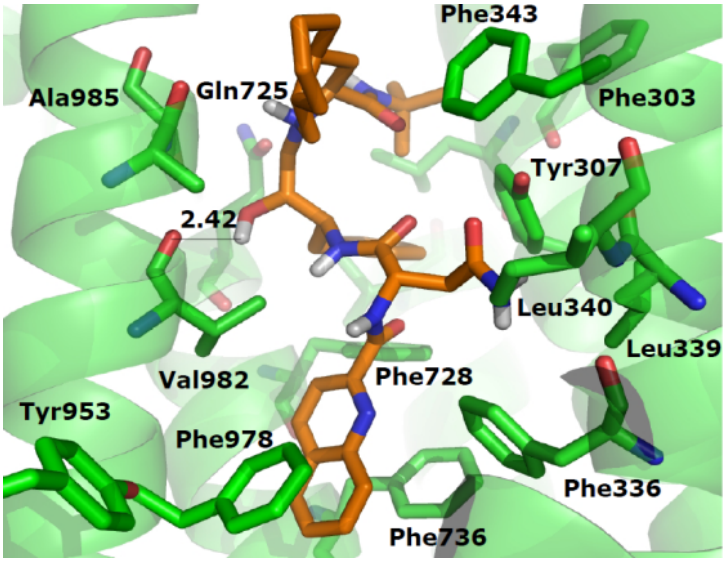


Figura 3.7: Geometría de unión de Saquinavir predicha mediante la simulación 5. Los residuos en el sitio de unión se representan en verde. Solo se muestran los átomos de hidrógeno unidos a nitrógeno y oxigeno. Los átomos de carbono de Saquinavir se representan en naranja. Los valores de las interacciones relevantes se indican en $\mathring{A}$.

También se evaluó la capacidad de interacción de las moléculas del conjunto de evaluación con la estructura experimental de la P-gp de ratón utilizando el programa Autodock Vina en las condiciones definidas en la simulación 5. El área bajo la curva obtenida fue de 0,821, similar pero menor que la encontrada para el modelo humano, lo que justifica el uso de un modelo humano de P-gp para la clasificación de compuestos.

Como método de validación adicional se dockeó el inhibidor co-cristalizado en la estructura experimental de ratón (QZ59, código PDB: 3G60) en el modelo por homología (simulación 5) para evaluar la capacidad del modelo de reproducir la conformación experimental. La figura 3.8 muestra la calidad del modelo mediante la superposición de la mejor conformación obtenida del docking y la estructura experimental (RMSD = 1,329 $\mathring{A}$).

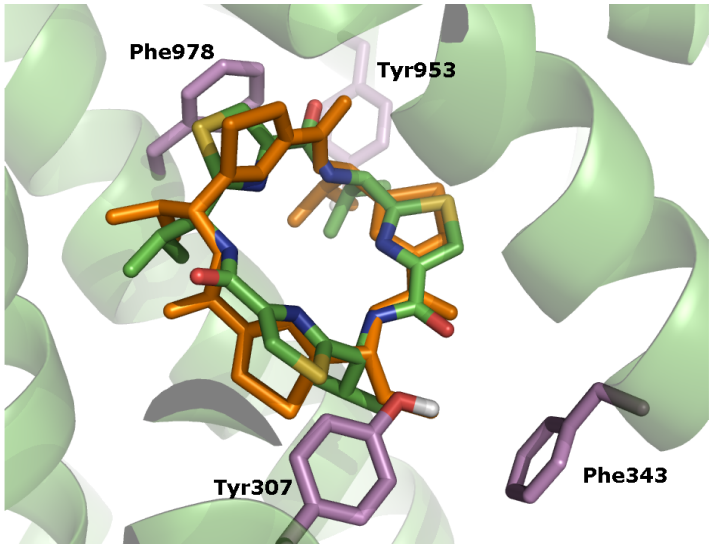


Figura 3.8: Superposición de la conformación del compuesto QZ59 resultante del docking (simulación 5) con la correspondiente a la estructura experimental de la P-gp de ratón (PDB: 3G60). Los átomos de carbono de los residuos flexibles son representados en violeta. La conformación experimental de QZ59 in 3G60 es representada en naranja.

Los resultados obtenidos hasta el momento fueron generados utilizando un conjunto de evaluación donde el número de sustratos es mayor al número de no sustratos (26 frente a 13). Con el fin de construir un conjunto de evaluación que contenga el mismo número de moléculas en ambos subsets usados para el análisis de las curvas ROC [43], se repitió el docking en las condiciones de la simulación 5 pero incluyendo en el conjunto de evaluación, moléculas endógenas de la base de datos KEGG [44]. Estas nuevas estructuras pueden ser consideradas como no sustratos debido a que su expulsión de la célula sería considerada ineficiente para su función. La base de datos KEGG ha sido usada con el mismo criterio por otros autores para incrementar el número de no sustratos [45]. Como solo se requerían 13 estructuras para equilibrar al subset de no sustratos con el de sustratos, se seleccionaron de la base de datos compuestos con peso molecular similar al conjunto inicial. Moléculas de bajo peso molecular tienden a ser mal puntuadas por las funciones de evaluación de docking, causando errores en el valor del área bajo la curva. La curva ROC obtenida de esta simulación se muestra en la figura 3.9, con un área bajo

la curva de 0,916, mejor que el reportado en la tabla 3.4. Esto significa que si se toma, por ejemplo, como punto de corte el valor de la función de evaluación de -9, este modelo es capaz de predecir el 65 % de los sustratos (sensibilidad: 0,65) y de excluir un 92 % de los no sustratos (especificidad: 0,92) con una exactitud del 81 %.

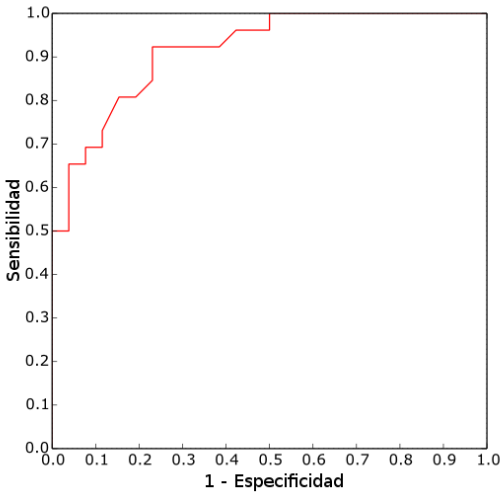


Figura 3.9: Curva ROC del conjunto de evaluación extendido con la base de datos KEEG en las condiciones de la simulación 5.

Finalmente, se procedió a evaluar el modelo obtenido en la simulación 5 frente a un conjunto de evaluación que incluye inhibidores y no inhibidores de P-gp. Esta librería incluye 1275 compuestos, 666 inhibidores y 609 no inhibidores de P-gp, y fue construida por Broccatelli y colaboradores [46]. Las estructuras de esta base de datos fueron analizadas y corregidas manualmente, tomándose como referencia el nombre indicado por los autores. Aquellos compuestos que no pudieron encontrarse en PUBCHEM [47], ya sea a partir del nombre o mediante su código SMILE, fueron eliminados. Del conjunto inicial de Broccatelli se obtuvieron 1206 estructuras correctamente identificadas. Con este conjunto se realizó el docking utilizando el modelo anteriormente seleccionado (simulación 5), lográndose un área bajo la curva de 0.836 (figura 3.10). En este caso, para un punto de corte de -9, se obtiene una sensibilidad de 0,80 y una especificidad de 0,73. Por lo tanto se concluye que la simulación por docking en condiciones de la simulación 5 posee

capacidad para identificar y clasificar razonablemente bien compuestos que interaccionan con P-gp, así como para reproducir interacciones experimentales en la estructura 3D de ratón disponible. Este modelo se utilizará para el cribado virtual de librerías de compuestos, como se detallara en el capítulo 5.

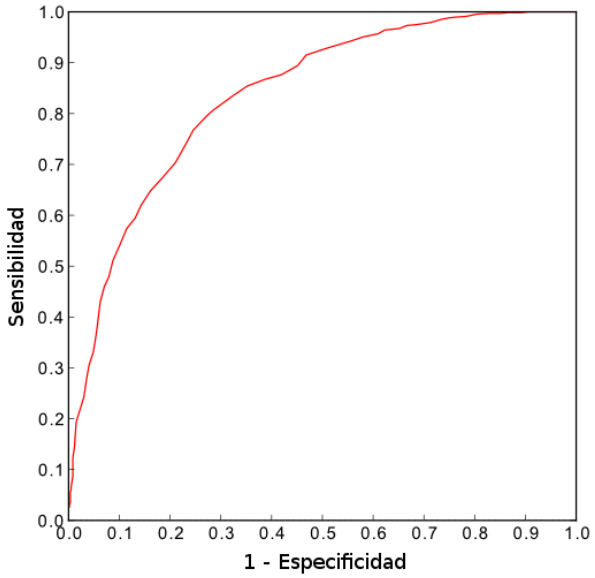


Figura 3.10: Curva ROC del modelo 5 utilizando el conjunto de evaluación de Broccatelli y colaboradores [46] en las condiciones de la simulación 5.

Bibliografía

[1] The UniProt Consortium, "UniProt: a hub for protein information," *Nucleic Acids Research*, vol. 43, no. Database issue, pp. D204–12, 2014.

[2] Y. Zhang, "I-TASSER server for protein 3D structure prediction.," *BMC bioinformatics*, vol. 9, p. 40, jan 2008.

[3] A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: a unified platform for automated protein structure and function prediction.," *Nature protocols*, vol. 5, pp. 725–38, apr 2010.

[4] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, "The I-TASSER Suite: protein structure and function prediction," *Nature Methods*, vol. 12, no. 1, pp. 7–8, 2014.

[5] S. Wu and Y. Zhang, "LOMETS: A local meta-threading-server for protein structure prediction," *Nucleic Acids Research*, vol. 35, no. 10, pp. 3375–3382, 2007.

[6] Y. Zhang and J. Skolnick, "SPICKER: A clustering approach to identify near-native protein folds," *Journal of Computational Chemistry*, vol. 25, pp. 865–871, 2004.

[7] Y. Zhang and J. Skolnick, "Scoring function for automated assessment of protein structure template quality," *Proteins: Structure, Function and Genetics*, vol. 57, no. 4, pp. 702–710, 2004.

[8] S. G. Aller, J. Yu, A. Ward, Y. Weng, S. Chittaboina, R. Zhuo, P. M. Harrell, Y. T. Trinh, Q. Zhang, I. L. Urbatsch, and G. Chang, "Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding.," *Science (New York, N.Y.)*, vol. 323, pp. 1718–22, mar 2009.

[9] D. a. Case, T. Darden, T. E. C. Iii, C. Simmerling, S. Brook, J. Wang, U. T. Southwestern, R. E. Duke, U.-c. Hill, R. Luo, U. C. Irvine, M. Crowley, R. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, U. C. S. Diego, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, N. C. State, V. Babin, P. a. Kollman, U. C. S. Francisco, D. a. Pearlman, R. V. Stanton, J. Pitera, I. Massova, A. Cheng, P. State, J. J. Vincent, P. Beroza,

V. Tsui, C. Schafmeister, W. S. Ross, R. Radmer, G. L. Seibel, J. W. Caldwell, C. Singh, P. Weiner, and P. Cieplak, "AMBER 11 Users' Manual," pp. 1–304.

[10] A. Bhattacharya, R. Tejero, and G. T. Montelione, "Evaluating protein structures determined by structural genomics consortia," *Proteins: Structure, Function, and Bioinformatics*, vol. 66, pp. 778–795, dec 2006.

[11] Y. J. Huang, R. Powers, and G. T. Montelione, "Protein NMR Recall, Precision, and F -measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics," *Journal of the American Chemical Society*, vol. 127, pp. 1665–1674, feb 2005.

[12] R. A. Laskowski, "PROCHECK: a program to check the stereochemicai quality of protein structures.," *J. Appl. Cryst.*, vol. 26, pp. 283–291, 1993.

[13] S. C. Lovell, I. W. Davis, W. B. Adrendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, "Structure validation by C alpha geometry: phi,psi and C beta deviation," *Proteins-Structure Function and Genetics*, vol. 50, no. 3, pp. 437–450, 2003.

[14] R. Lüthy, J. U. Bowie, and D. Eisenberg, "Assessment of protein models with three-dimensional profiles.," *Nature*, vol. 356, no. 6364, pp. 83–85, 1992.

[15] M. Sippl, "Recognition of errors in threedimensional structures of proteins," *Proteins: Structure, Function, and Bioinformatics*, 1993.

[16] S. V. Ambudkar, S. Dey, C. a. Hrycyna, M. Ramachandra, I. Pastan, and M. M. Gottesman, "Biochemical, cellular, and pharmacological aspects of the multidrug transporter.," *Annual review of pharmacology and toxicology*, vol. 39, pp. 361–398, 1999.

[17] S. Agarwal, V. Arya, and L. Zhang, "Review of P-gp inhibition data in recently approved new drug applications: utility of the proposed [I(1)]/IC(50) and [I(2)]/IC(50) criteria in the P-gp decision tree.," *Journal of clinical pharmacology*, vol. 53, pp. 228–33, feb 2013.

[18] R. B. Kim, M. F. Fromm, C. Wandel, B. Leake, A. J. Wood, D. M. Roden, and G. R. Wilkinson, "The drug transporter P-glycoprotein limits oral absorption and brain entry of HIV-1 protease inhibitors.," *Journal of Clinical Investigation*, vol. 101, pp. 289–294, jan 1998.

[19] J. W. Polli, J. L. Jarrett, S. D. Studenberg, J. E. Humphreys, S. W. Dennis, K. R. Brouwer, and J. L. Woolley, "Role of P-Glycoprotein on the CNS Disposition of Amprenavir (141W94), an HIV Protease Inhibitor," *Pharmaceutical Research*, vol. 16, no. 8, pp. 1206–1212, 1999.

[20] G. A. Scarborough, "Drug-stimulated ATPase activity of the human P-glycoprotein," *Journal of Bioenergetics and Biomembranes*, vol. 27, no. 1, pp. 37–41, 1995.

[21] T. Litman, T. Zeuthen, T. Skovsgaard, and W. D. Stein, "Structure-activity relationships of P-glycoprotein interacting drugs: Kinetic characterization of their effects on ATPase activity," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1361, no. 2, pp. 159–168, 1997.

[22] G. Liminga, P. Nygren, and R. Larsson, "Microfluorometric evaluation of calcein acetoxymethyl ester as a probe for P-glycoprotein-mediated resistance: effects of cyclosporin A and its nonimmunosuppressive analogue SDZ PSC 833.," 1994.

[23] G. D. Eytan, R. Regev, G. Oren, C. D. Hurwitz, and Y. G. Assaraf, "Efficiency of P-glycoprotein-mediated exclusion of rhodamine dyes from multidrug-resistant cells is determined by their passive transmembrane movement rate," *Eur. J. Biochem.*, vol. 248, no. 1, pp. 104–112, 1997.

[24] Human Food and Drug Administration (FDA) and U.S. Department of Health and Services., "Guidance for Industry. Study Design , Data Analysis , and Implications for Dosing and Labeling," Tech. Rep. September, 2006.

[25] J. W. Polli, S. a. Wring, J. E. Humphreys, L. Huang, J. B. Morgan, L. O. Webster, and C. S. Serabjit-Singh, "Rational use of in vitro P-glycoprotein assays in drug discovery.," *The Journal of pharmacology and experimental therapeutics*, vol. 299, pp. 620–8, nov 2001.

[26] B. Feng, J. Mills, R. Davidson, and R. Mireles, "In vitro P-glycoprotein assays to predict the in vivo interactions of P-glycoprotein with drugs in the central nervous system," *Drug Metabolism and*, vol. 36, no. 2, pp. 268–275, 2008.

[27] K. M. M. Doan, J. E. Humphreys, L. O. Webster, S. A. Wring, L. J. Shampine, C. J. Serabjit-singh, K. K. Adkison, and J. W. Polli, "Passive Permeability and P-Glycoprotein-Mediated Efflux Differentiate Central Nervous System (CNS) and Non-CNS Marketed Drugs," *Pharmacology*, vol. 303, no. 3, pp. 1029–1037, 2002.

[28] K. Doan, J. Humphreys, and L. Webster, "Passive permeability and P-glycoprotein-mediated efflux differentiate central nervous system (CNS) and non-CNS marketed drugs," *of Pharmacology and*, vol. 303, no. 3, p. 1029, 2002.

[29] D. Schwab, H. Fischer, A. Tabatabaei, S. Poli, and J. Huwyler, "Comparison of in vitro P-glycoprotein screening assays: recommendations for their use in drug discovery.," *Journal of medicinal chemistry*, vol. 46, pp. 1716–25, apr 2003.

[30] C. Zhang, P. Kwan, Z. Zuo, and L. Baum, "The transport of antiepileptic drugs by P-glycoprotein," dec 2012.

[31] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeerschd, E. Zurek, and G. R. Hutchison, "Avogadro: An advanced semantic chemical editor, visualization, and analysis platform," *Journal of Cheminformatics*, vol. 4, no. 8, pp. 1–17, 2012.

[32] T. A. Halgren, "Merck Molecular Force Field. I.," *J. Comput. Chem.*, vol. 17, pp. 490–519, 1996.

[33] T. A. Halgren, "Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions," *Journal of Computational Chemistry*, vol. 17, no. 5-6, pp. 520–552, 1996.

[34] T. A. Halgren, "Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94*," *J. Comput. Chem.*, vol. 17, no. 1 996, pp. 553–584, 1996.

[35] T. A. Halgren and R. B. Nachbar, "Merck molecular force field. IV. conformational energies and geometries for MMFF94," *Journal of Computational Chemistry*, vol. 17, no. 5-6, pp. 587–615, 1996.

[36] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function," *Journal of Computational Chemistry*, vol. 19, pp. 1639–1662, nov 1998.

[37] R. Huey, G. Morris, A. Olson, and D. Goodsell, "A semiempirical free energy force field with charge-based desolvation," *Journal of computational chemistry*, vol. 28, no. 6, pp. 1145–1152, 2007.

[38] O. Trott and A. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of computational chemistry*, vol. 31, no. 2, pp. 455–461, 2010.

[39] R. a. Friesner, J. L. Banks, R. B. Murphy, T. a. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin, "Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.," *Journal of medicinal chemistry*, vol. 47, pp. 1739–49, mar 2004.

[40] T. a. Halgren, R. B. Murphy, R. a. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, and J. L. Banks, "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening," *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1750–1759, 2004.

[41] R. Friesner, R. Murphy, M. Repasky, L. Frye, J. Greenwood, T. Halgren, P. Sanschagrin, and D. Mainz, "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes," *J. Med. Chem.*, no. 49, pp. 6177–6196, 2006.

[42] J. Gasteiger and M. Marsili, "Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges," *Tetrahedron*, vol. 36, no. 22, pp. 3219–3228, 1980.

[43] A. Rizzi and A. Fioni, "Virtual screening using PLS discriminant analysis and ROC curve approach: An application study on PDE4 inhibitors," *Journal of Chemical Information and Modeling*, vol. 48, no. 8, pp. 1686–1692, 2008.

[44] S. Goto and M. Kanehisa, "KEGG: Kyoto encyclopedia of genes and genomes," 1999.

[45] E. Dolghih, C. Bryant, A. R. Renslo, and M. P. Jacobson, "Predicting Binding to P-Glycoprotein by Flexible Receptor Docking.," *PLoS computational biology*, vol. 7, p. e1002083, jun 2011.

[46] F. Broccatelli, E. Carosati, A. Neri, M. Frosini, L. Goracci, T. I. Oprea, and G. Cruciani, "A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields.," *Journal of medicinal chemistry*, vol. 54, pp. 1740–51, mar 2011.

[47] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S. H. Bryant, "PubChem Substance and Compound databases.," *Nucleic acids research*, vol. 44, no. D1, pp. D1202–1213, 2015.

Capítulo 4

Modelo para la predicción de la afinidad de compuestos a Nav1.2

4.1. Modelado de la estructura del canal de sodio voltaje operado 1.2

Como se detalló en el capítulo 1 de la presente tesis, el mayor componente estructural de los canales de sodio de mamíferos es una proteína de 260 kDa de peso molecular, llamada subunidad α [1]. Esta subestructura forma el poro de conducción iónica del canal, así como las compuertas que regulan el flujo de sodio. En el cerebro, la subunidad α se asocia covalentemente con dos subunidades, $\beta 1$ (33 kDa) y $\beta 2$ (36 kDa) [1]. Estas subunidades β no son necesarias para que el canal sea funcional, pero modulan la expresión y la función del mismo [2]. El análisis de secuencia de la subunidad α muestra cuatro fragmentos repetidos (denominados I-IV) de un motivo estructural consistente en 6 segmentos α-hélice transmembrana S1-S6, conectados por loops intra y extracelulares. Se presume que estos cuatro dominios forman un arreglo en forma de cubo en la membrana plasmática, con el poro conductor iónico en el centro, en forma similar a la estructura del poro encontrada en bacterias (figura 4.1).

Al momento no se dispone de la estructura tridimensional experimental de canales de sodio humanos. Por esta razón, se procedió a modelar la estructura de la isoforma

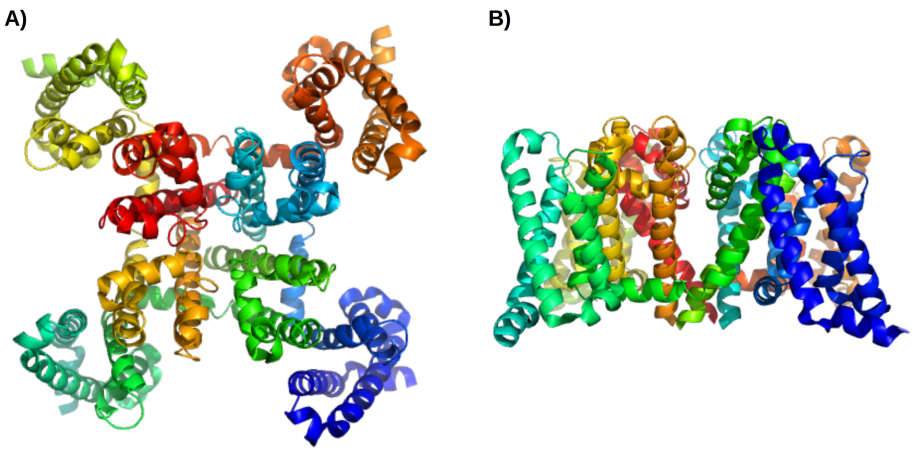


Figura 4.1: Estructura tridimensional de la subunidad α, ejemplificado mediante el canal de sodio voltaje operado de *Arcobacter butzleri* (PDB: 3RVY). A) Vista desde el espacio extracelular. B) Vista lateral. En tonos azules se muestra el dominio I, en verde el dominio II, en amarillo el dominio III y en rojo el IV.

Nav1.2, la cual se encuentra implicada en los fenómenos epilépticos [3]. Además este isotipo es blanco molecular de ciertos fármacos anticonvulsivos, los cuales interaccionan particularmente en la región del poro [4].

La secuencia de aminoácidos de la subunidad α humana de Nav1.2 se encuentra depositada en la base de datos UNIPROT [5] con el código Q99250. Con la intención de identificar posibles macromoléculas molde para la predicción estructural, esta secuencia se utilizo como patrón de búsqueda en la base de datos Protein Data Bank [6] mediante el servidor NCBI-BLAST [7]. Los mejores 15 resultados obtenidos, ordenados en forma decreciente respecto al valor de puntuación, se presentan en la tabla 4.1. Los primeros siete corresponden a alineamientos con moldes de fragmentos pequeños de canales iónicos humanos con un porcentaje de cobertura muy reducido (abarcan menos del 8 % del total de la secuencia). El siguiente (PDB: 1BYY), corresponde a una macromolécula patrón de la compuerta de inactivación de canales de sodio voltaje operados de rata. El resto de los moldes son canales iónicos de las bacterias *Arcobacter butzleri* (códigos PDB: 3RVY, 4MVZ, 4MS2, 4MW8, 4MVQ y 4MVU) y *Caldalkalibacillus thermarum* (PDB: 4BGN).

Estos canales bacterianos están compuestos por cuatro subunidades proteicas idénticas, de aproximadamente trescientos aminoácidos de longitud. Todas las secuencias de estas estructuras se alinearon con las mismas cuatro regiones de la secuencia problema, con un bajo valor E y un alto valor de puntuación.

PDB	P. Máxima	P. Total	% Cobertura	Valor E	% Identidad
4JPZ B	333	333	8	$8e^{-104}$	100
4DCK A	270	270	8	$5e^{-82}$	76
4JQO D	271	271	8	$7e^{-82}$	76
4OVN A	250	250	7	$6e^{-75}$	74
2KAV A	226	226	5	$3e^{-67}$	100
2KBI A	167	167	4	$9e^{-47}$	84
1BYY A	118	118	2	$3e^{-30}$	100
3RVY A	77	309	45	$1e^{-14}$	24
4MVZ A	73,6	261	37	$2e^{-13}$	24
4MVQ A	72,8	292	40	$2e^{-13}$	24
4MS2 A	72,8	258	40	$3e^{-13}$	24
4MW8 A	72,4	252	37	$3e^{-13}$	24
4BGN A	71,6	256	46	$5e^{-13}$	24
4MVU A	67,8	67,8	37	$6e^{-13}$	24

Tabla 4.1: Resultados de BLAST ordenados según el valor E (E-value). Los segmentos de alta puntuación o high scoring segment pair (HSP) de una misma secuencia de la base de datos son agrupados. Puntuación máxima corresponde al valor del HSP con el puntaje más elevado. Puntaje total corresponde a la suma de los puntajes individuales de cada HSP. Si ambos valores son iguales, esa secuencia contiene un solo HSP. % cobertura se refiere a la porción de la secuencia problema que se alinea con el fragmento de la base de datos. % de identidad indica el grado de identidad en el HSP de mayor puntaje.

El servidor NCBI-BLAST realiza también una búsqueda de dominios conservados en la base de datos CDD (Conserved Domain Database) [8]. Con esta exploración pueden generarse alineamientos de la secuencia con fragmentos representativos de familias, cuyas estructuras tridimensionales son conocidas. Esta búsqueda identificó también las cuatro regiones de la secuencia previamente mencionadas. En este caso se detectó que estas zonas están relacionadas con la familia pf00520, conformada por canales de sodio, potasio y calcio (figura 4.2). Por lo tanto, puede inferirse la presencia de un patrón repetitivo en la secuencia, que posiblemente se encuentra asociado a la función de la proteína. En base a este análisis se procedió a utilizar estas regiones alineadas para la obtención

de la estructura tridimensional de los dominios que forman el poro de la subunidad α. Se utilizaron los fragmentos ubicados en las posiciones 111-456, 741-1013, 1190-1504 y 1513-1611 que corresponderían a los dominios I, II, III y IV respectivamente.

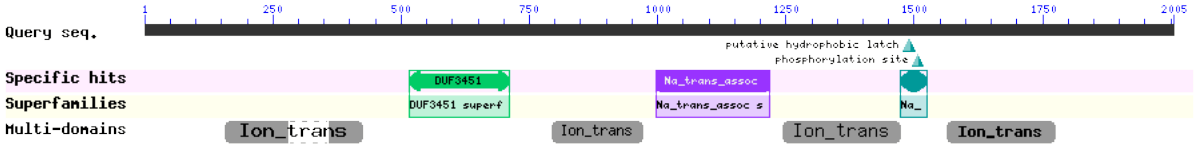


Figura 4.2: Resumen gráfico de la búsqueda de dominios conservados en CDD. Los segmentos identificados como Ion_trans corresponden a la familia pf00520. La familia Na_trans_assoc (pf06512) identifica a una región presente exclusivamente en canales de sodio eucariotas. DUF3451 corresponde a un dominio cuya función no ha sido caracterizada, este segmento se encuentra presente en proteínas eucariotas, asociado a las familias pf00520 y pf06512 [9]. Na_ (Na_channel_gate) es parte del segmento conector intracelular de los dominios III y IV de las subunidades α de los canales de sodio dependientes de voltaje, responsable de la inactivación rápida.

4.1.1. Modelo de la estructura tridimensional de Nav1.2 en estado cerrado.

El modelo estructural de la subunidad α del canal de sodio humano Nav1.2 se construyó en dos etapas. Primero, cada uno de los cuatro dominios se modeló por separado utilizando el servidor GPCR-I-TASSER, una extensión del programa I-TASSER, que utiliza un método de comparación de plegamientos (threading) [10, 11, 12]. En esta modificación se incorpora un campo de fuerza y un conjunto de restricciones específicas para proteínas transmembrana [13]. Al igual que en I-TASSER, la secuencia problema primero es utilizada para buscar posibles estructuras patrón en la base de datos PDB mediante nueve programas de threading diferentes. Si se encuentran estructuras homologas, se utiliza un proceso de ensamblado de fragmentos basado en la estructura de los patrones para la generación del modelo completo. En caso de no encontrarse patrones, se utiliza un procedimiento Ab-Initio.

Como se indicó previamente, para la generación de la estructura tridimensional de

los dominios se ingresaron al servidor los fragmentos ubicados en las posiciones 111-456, 741-1013, 1190-1504 y 1513-1611 de la secuencia de Nav1.2 humana por separado. Como resultado se obtuvieron cinco modelos estructurales de cada dominio, eligiéndose aquel con el mayor valor de C-Score (Tabla 4.2). Este parámetro estima la calidad de los modelos predichos, basándose en la significancia de los alineamientos realizados sobre cada molde luego del threading [10]. Generalmente toma valores entre -5 y 2, donde un valor mayor indica un modelo más confiable [10]. En la tabla 4.2 también se reporta el valor de TM-score, el cual mide la similitud estructural entre dos macromoléculas. Se encuentra relacionado con el parámetro RMSD (root-mean-square deviation), pero presenta menor sensibilidad a errores locales (por ejemplo una cadena lateral ubicada en forma diferente que en la proteína molde), que podrían llevar a suponer que un modelo es incorrecto cuando en su topología global si lo es. Un TM-score menor a 0,17 indica que la similaridad global es muy pequeña y por lo tanto el modelo es incorrecto [14].

	Dominio 1	**Dominio 2**	**Dominio 3**	**Dominio 4**
C-score	$-2,47$	$-0,59$	$-0,81$	$-0,82$
TM-score	$0,43$	$0,64$	$0,61$	$0,61$

Tabla 4.2: Valores de C-score y TM-score de los modelos obtenidos en I-TASSER.

Luego de la generación de cada dominio, el poro del canal fue ensamblado mediante la superposición estructural con la estructura completa del canal de sodio bacteriano 3RVY. Se acoplaron los 4 dominios en el orden del movimiento de las agujas de reloj (visto desde el espacio extracelular), siguiendo el criterio empleado en publicaciones anteriores para la construcción de este tipo de macroestructuras [15, 16]. Al realizar el ensamblado, se encontró que los dominios I y III presentaban un loop de aproximadamente 90 y 40 residuos, los cuales se superponían con la región a ocupar por el resto de los dominios al formar el poro. Estos loops son producto de gaps o segmentos de la secuencia no alineados con la secuencia patrón, generándose así una región sin estructura secundaria definida. Dado que el sitio activo para anticonvulsivos y anestésicos locales se encuentra en el vestíbulo interior del poro formado por los segmentos S6, se decidió eliminar estos loops ya que su ausencia no afectaría el modelado de la región del poro, necesario para

el cribado virtual.

Finalmente, el modelo estructural completo, compuesto por los cuatro dominios, fue refinado por métodos de minimización geométrica. Los átomos de la proteína fueron rodeados por una caja periódica de moléculas de agua del tipo TIP3P32 [17], ocupando hasta 10$\mathring{A}$ desde la macromolécula, con el módulo TLEAP del software AMBER11 [17]. La geometría inicial fue minimizada utilizando el módulo SANDER mediante 1000 ciclos para la relajación de las moléculas de agua seguidos de 2500 para el sistema completo usando el campo de fuerza AMBER en su versión ff03.r1 [17]. La estructura final obtenida se muestra en la figura 4.3.

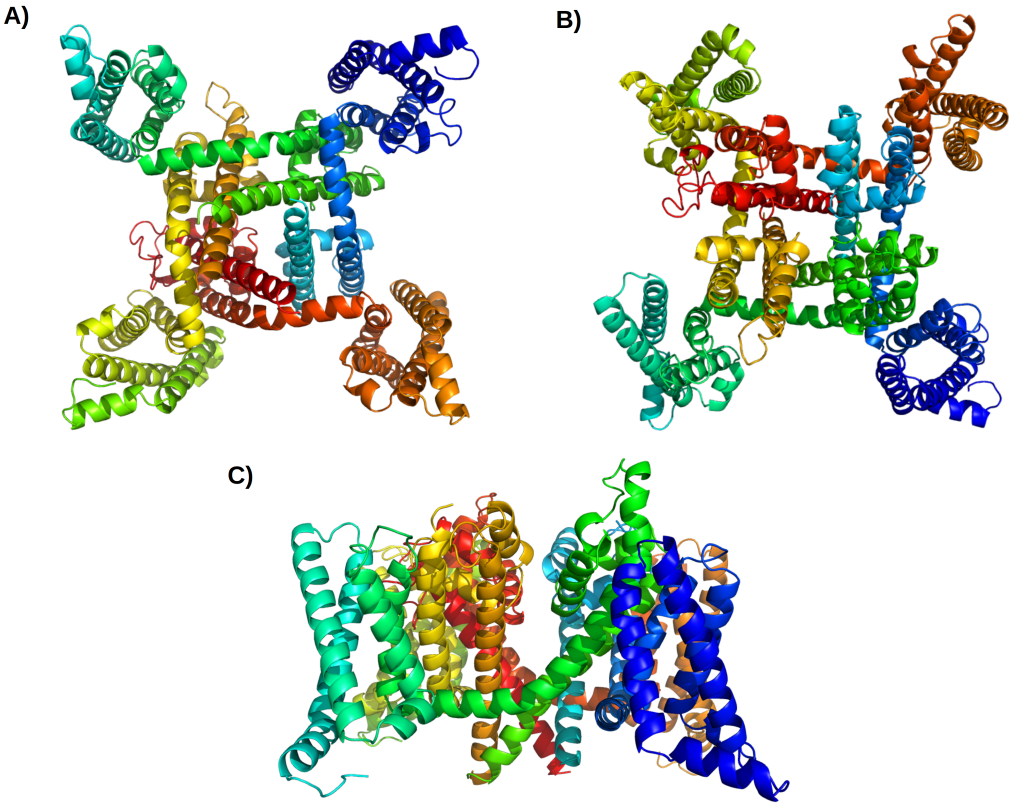


Figura 4.3: Estructura tridimensional del canal de sodio humano Nav1.2, en su conformación cerrada. A) Vista desde el interior celular. B) Vista desde el espacio extracelular. C) Vista lateral. En tonos azules se muestra el dominio I, en verde el dominio II, en amarillo el dominio III y en rojo el IV.

La calidad del modelo obtenido fue analizada mediante los gráficos de Ramachandran, utilizando el servidor Protein Structure Validation Software (PSVS) [18, 19, 20, 21, 22,

23]. Un 90,6 % de los residuos del modelo se encuentran en regiones favorables y un 97,6 % en regiones permitidas. Solo el 2,4 % del total está en zonas desfavorables (figura 4.4 A). Estos valores son similares a los de la estructura experimental de bacteria (código 3RVY), donde el 90,9 % de sus residuos están en la zona favorable, mientras que el 100 % de los mismos están en zonas permitidas (figura 4.4 B).

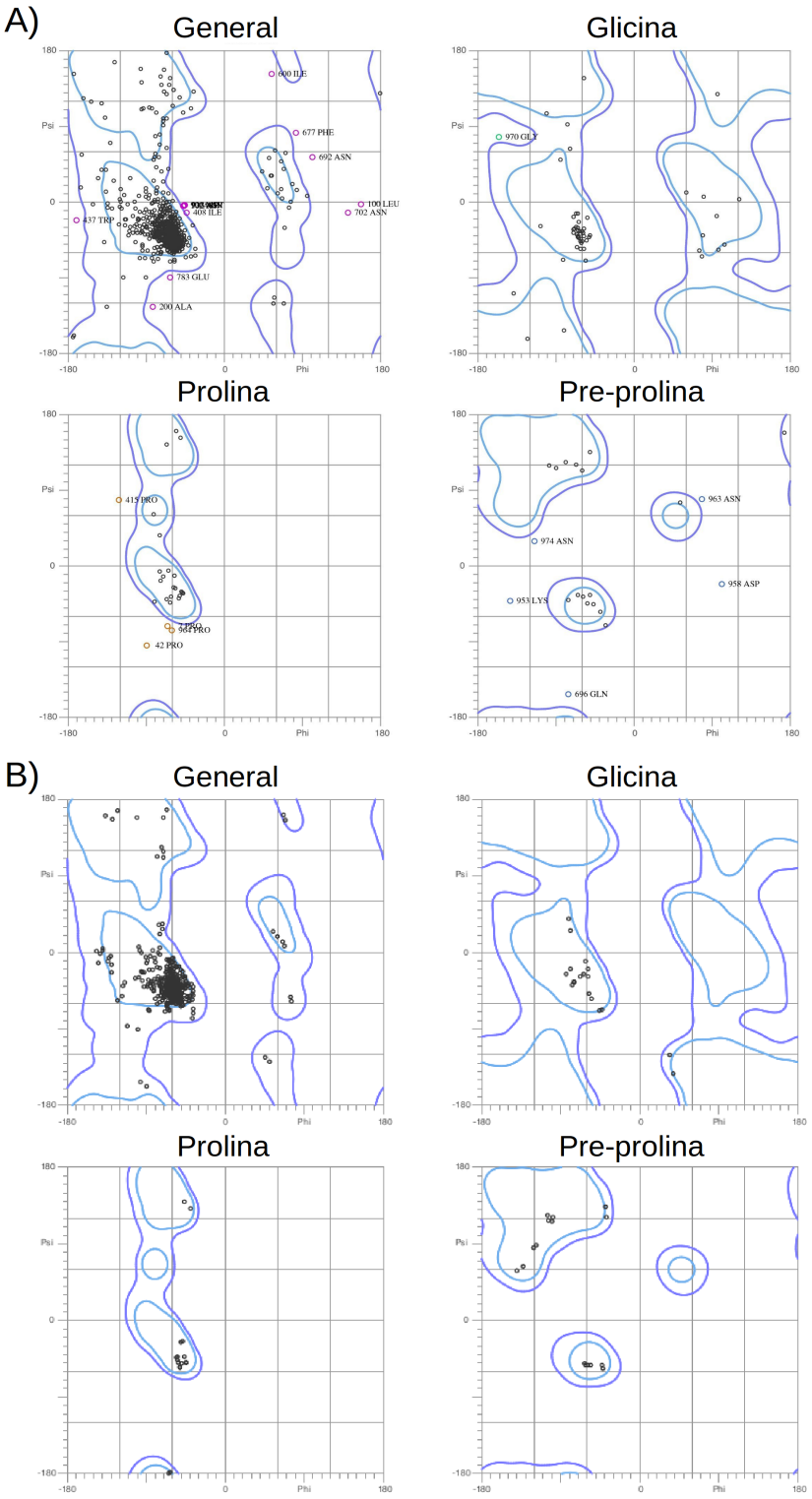


Figura 4.4: Gráficos de Ramachandran: A) Modelo de Nav1.2 humano obtenido con I-TASSER; B) Estructura tridimensional de NavAb Arcobacter butzleri (PDB: 3RVY).

Otro aspecto a considerar en la validación es la presencia en el modelo del "anillo DEKA". En los canales de sodio eucariotas, existen residuos altamente conservados de Asp (D), Glu (E), Lys (K) y Ala (A), en los dominios I a IV, respectivamente. Estos forman un anillo conocido como anillo DEKA, el cual sería uno de los responsables de la selectividad iónica del canal [24]. Como puede verse en la figura 4.5, el modelo reproduce la presencia del anillo DEKA, lo cual aumenta la confianza en la estructura tridimensional obtenida.

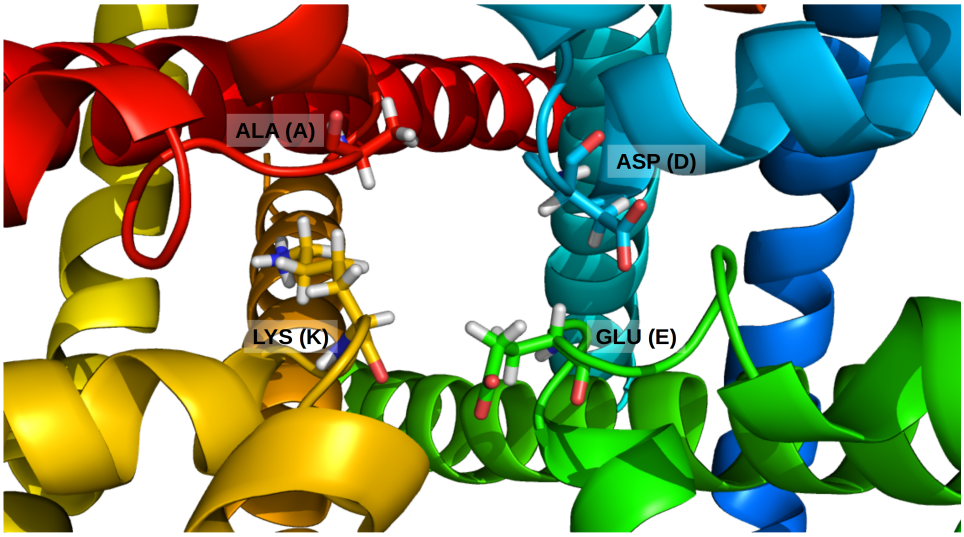


Figura 4.5: Anillo DEKA presente en el selector de iones del modelo del canal cerrado obtenido con I-TASSER. En tonos azules se muestra el dominio I, en verde el dominio II, en amarillo el dominio III y en rojo el IV.

Como se detalló anteriormente, los anticonvulsivos interaccionan en la región del poro del canal, y lo hacen más eficientemente cuando este se encuentra en una conformación abierta (inhibición dependiente de estado) [4]. La estructura molde empleada en el modelado presentado hasta aquí fue seleccionada en base a los parámetros encontrados en el alineamiento (Tabla 4.1), pero presenta una conformación cerrada. Es por eso que se decidió modificar la estructura del molde para lograr un modelo adicional del canal Nav1.2 humano de conformación abierta.

4.1.2. Modelo tridimensional del poro de Nav1.2 en estado abierto.

Para la construcción de un modelo de conformación abierta, inicialmente se analizó la estructura del canal iónico no especifico NaK de Bacilus cereus, del cual se encuentra disponible información estructural tanto de su conformación cerrada (código PDB: 2AHY) como abierta (código PDB: 3E86). Mediante un análisis por superposición de estas dos conformaciones, se observa que las mayores diferencias se dan en la zona interna del poro, debajo del filtro de selectividad DEKA. Por otra parte, las regiones en torno a este último se mantienen en las mismas posiciones (figura 4.6 A). Lo mismo sucede si se compara la estructura experimental del molde de NavAb empleado previamente (código PDB: 3RVY) y la conformación abierta de NaK (figura 4.6 B). Una comparación entre los estados conformacionales disponibles muestra diferencias en la zona del filtro de selectividad, y están relacionadas con la estructura secundaria. Teniendo en cuenta estas diferencias, se consideró que para la construcción de un modelo abierto debían reemplazarse las coordenadas el segmento S6 (que conforma el poro del canal) de la estructura NavAb (estado cerrado) por el correspondiente a NaK en estado abierto. Por lo tanto, la estructura resultante de la combinación entre 3RVY y 3E86 se utilizó como molde para la generación de los dominios del canal de sodio humano mediante modelado comparativo.

Debe aclararse que durante la apertura del canal se genera un cambio conformacional tal que influye también en el sensor de voltaje. En la estructura molde obtenida a través del reemplazo del segmento S6 del canal 3RVY por el correspondiente de 3E86, se produce una superposición de este fragmento con el segmento S4. Por lo tanto este último segmento fue eliminado de la estructura molde, quedando solo los segmentos S5 y S6, los cuales forman el poro del canal (y son los necesarios para el cribado virtual).

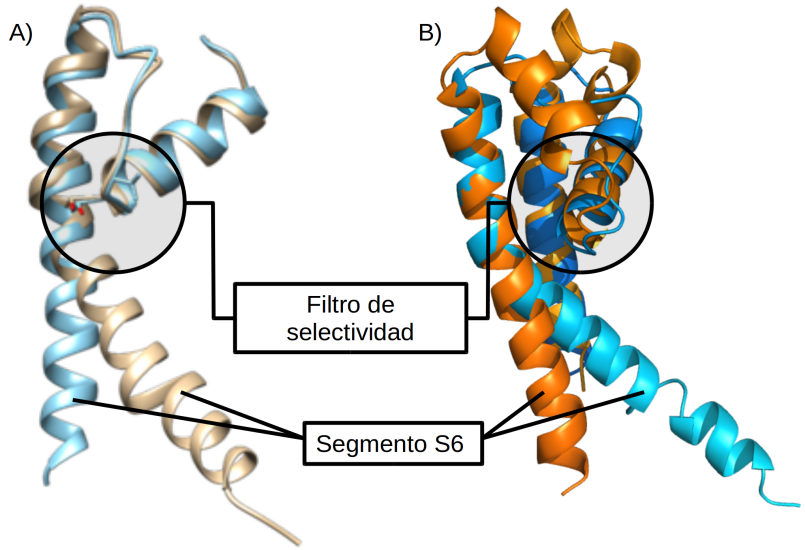


Figura 4.6: Superposición de estructuras: A) Canal NaK en su conformación cerrada (celeste) frente a la conformación abierta (naranja claro). B) Canal de sodio NavAb (celeste) frente a NaK en conformación abierta (naranja).

El modelo tridimensional del canal de sodio Nav1.2 en estado abierto se realizó mediante el software de modelado comparativo MODELLER 9.11 [25, 26, 27]. Para la obtención del modelo, la secuencia de la estructura molde se alineó a la correspondiente de cada uno de los dominios utilizando la matriz de puntuación BLOSUM62. A partir de este alineamiento se generaron 30 estructuras de cada uno, eligiéndose aquella con el menor valor DOPE (Discrete Optimized Protein Energy). DOPE consiste en un potencial estadístico desarrollado para evaluar modelos proteicos. Fue desarrollado mediante el análisis de 1472 estructuras cristalográficas [28]. A partir de la designación de potenciales estadísticos optimizados a cada aminoácido de un modelo, permite evaluar grupos de estos y comparar regiones problemáticas del modelo. Otra aplicación del DOPE es la evaluación de varios modelos de una misma proteína para discriminar entre ellos desde un punto de vista del potencial estadístico total de las mismas. En la tabla 4.3 se muestran los valores de DOPE para el modelo de cada dominio seleccionado.

	Dominio 1	**Dominio 2**	**Dominio 3**	**Dominio 4**
DOPE	1, 420	1, 391	0, 897	1, 678

Tabla 4.3: Valores de DOPE obtenidos para cada uno de los dominios del canal de sodio Nav1.2. Cuanto menor es el valor de DOPE, mejor es la calidad del modelo.

Una vez seleccionada la estructura a utilizar de cada dominio, se ensambló y minimizó el poro del canal de la misma forma que para el canal cerrado (ver sección anterior). El modelo final se muestra en la figura 4.7.

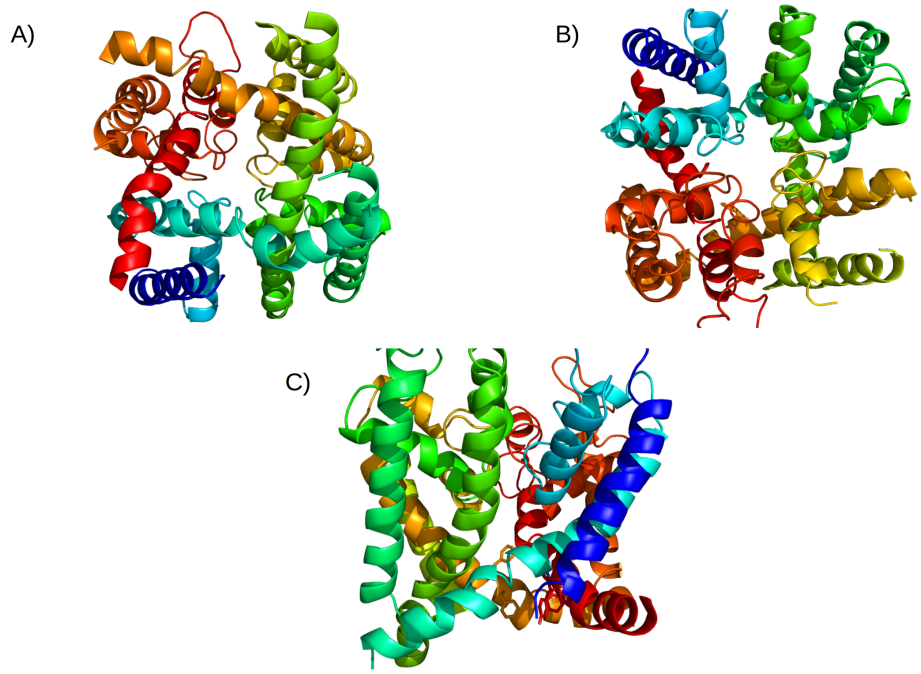


Figura 4.7: Estructura tridimensional del canal de sodio humano Nav1.2, en su conformación abierta. A) Vista desde el interior celular. B) Vista desde el espacio extracelular. C) Vista lateral.

Al igual que en el modelo abierto se comprobó la presencia del anillo DEKA (figura 4.8).

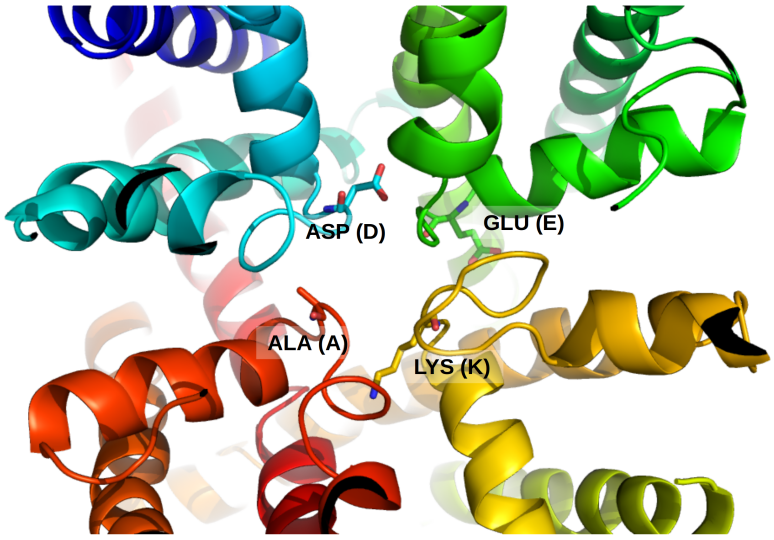


Figura 4.8: Anillo DEKA presente en el selector de iones del modelo del canal abierto. En azul se muestra el dominio I, en verde el dominio II, en amarillo el dominio III y en rojo el IV.

El gráfico de Ramachandran correspondiente al modelo de la estructura de Nav1.2 abierta muestra un 91.9 % de los residuos aminoácidicos dentro de zonas permitidas, mientras que un 76.8 % se encuentra en zonas favorables (figura 4.9).

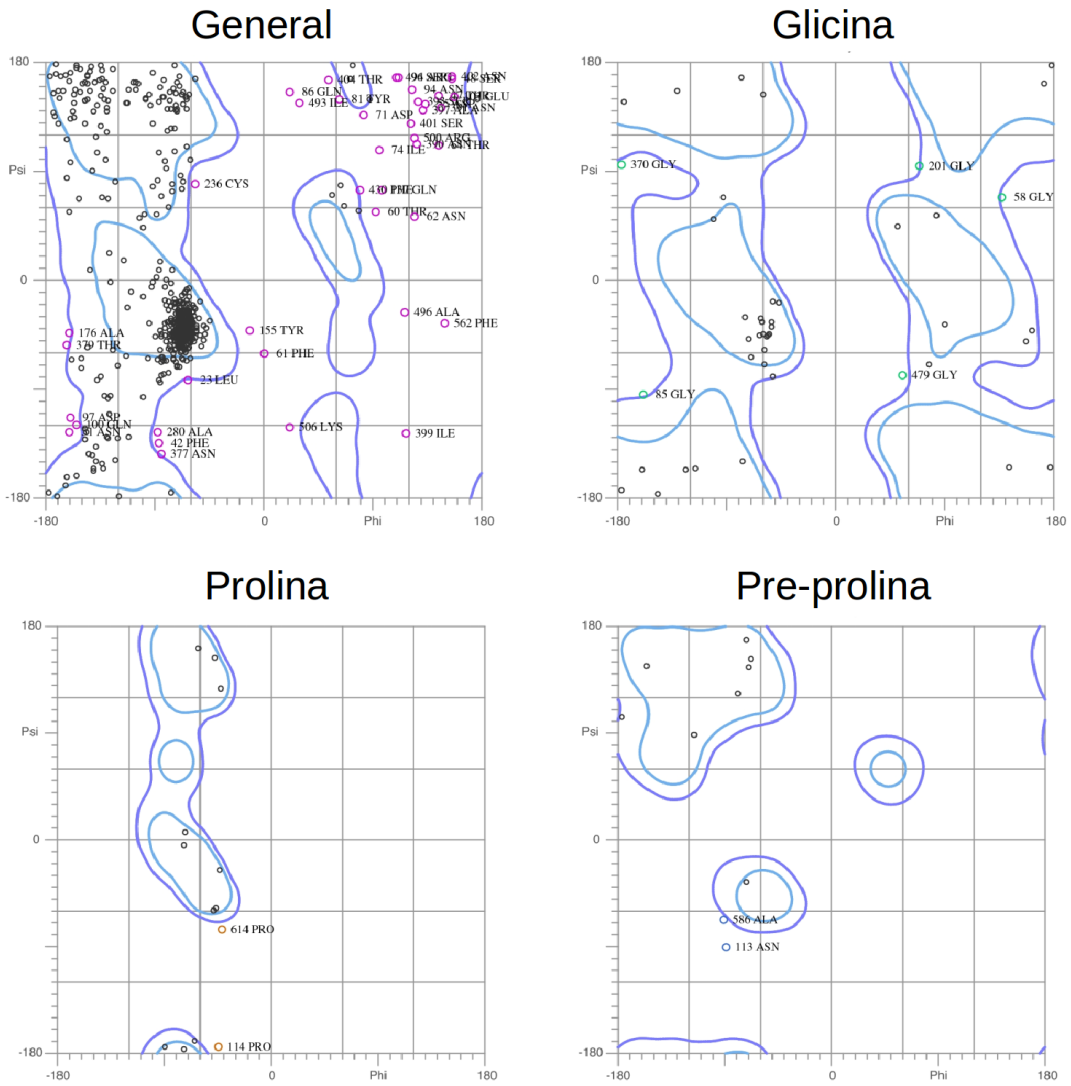


Figura 4.9: Gráficos de Ramachandran del modelo del canal de sodio en la conformación abierta.

4.2. Búsqueda bibliográfica y conjunto de evaluación.

Una vez obtenidas y validadas las estructuras tridimensionales del canal de sodio cerrado y abierto resta seleccionar el programa de docking (y sus condiciones) para la realización del cribado. Para ello inicialmente se analizó el desempeño del software en

discriminar entre compuestos activos e inactivos conocidos. Para la construcción del conjunto de evaluación se generaron dos subconjuntos de compuestos: activos e inactivos. Como activos se tomaron 180 estructuras con alta afinidad de unión al canal Nav1.2 proveniente de la base de datos CHEMBL [29] (valores de IC_{50} menores o iguales a 100 μM). Además se agregaron compuestos antiepilépticos con acción comprobada frente al canal de sodio (Fenitoina, Carbamazepina y Lamotrigina) [30, 31].

El conjunto de inactivos se construyó con compuestos con baja afinidad de unión al canal como: Deprenilo, gabapentina, mecamilamina, moclobemida, nialamida, procainamida, tiaprida y topiramato [31]. De manera de ampliar el número de compuestos en este subconjunto se realizó una búsqueda de estructuras similares a las antes mencionadas, asumiendo que estas también serían inactivas. De acuerdo al número de estructuras similares encontradas para cada compuesto, se fijo el límite de similaridad requerido. Por ejemplo, se utilizaron bajos valores de similaridad para nialamida (70 %) ya que a valores de 90 % no se encontraron compuestos similares. Por otro lado, en el caso de mecamilamina se utilizó un porcentaje de similaridad del 99 % ya que para 90 % se obtenían del orden de 1000 compuestos. De esta forma el conjunto de evaluación quedo formado por 506 compuestos, 183 activos y 323 inactivos.

4.3. Docking molecular.

Se procedió a desarrollar un modelo de docking molecular capaz de predecir la capacidad de interacción de los compuestos del conjunto de evaluación. Los modelos se generaron utilizando el software AutoDock Vina [32] definiendo el sitio activo en la zona interna del poro donde interaccionan los anestésicos locales y los antiepilépticos. Esta zona contiene a los residuos desde Phe-1754 hasta Tyr-1771 [33]. Para el canal cerrado se utilizó una grilla de 16 x 19 x 25 Å, mientras que en el modelo abierto el tamaño de la grilla fue de 20 x 20 x 20 Å. En cada caso se permitió la movilidad de diferentes residuos de aminoácidos. En el modelo cerrado los residuos flexibles fueron Gln-383, Ser-413, Phe-

1764 y Tyr-1771 mientras que en el modelo abierto fueron Phe-1462, Ser-1463, Phe-1756 y Tyr-1760. El sitio de unión seleccionado para cada modelo se indica en la figura 4.10.

Las moléculas a evaluar fueron protonadas a pH fisiológico con el software OpenBabel2.3.2 [34] y se utilizaron cargas parciales Marsili-gasteiger [35] calculadas mediante el software AutoDockTools 1.5.2 [36].

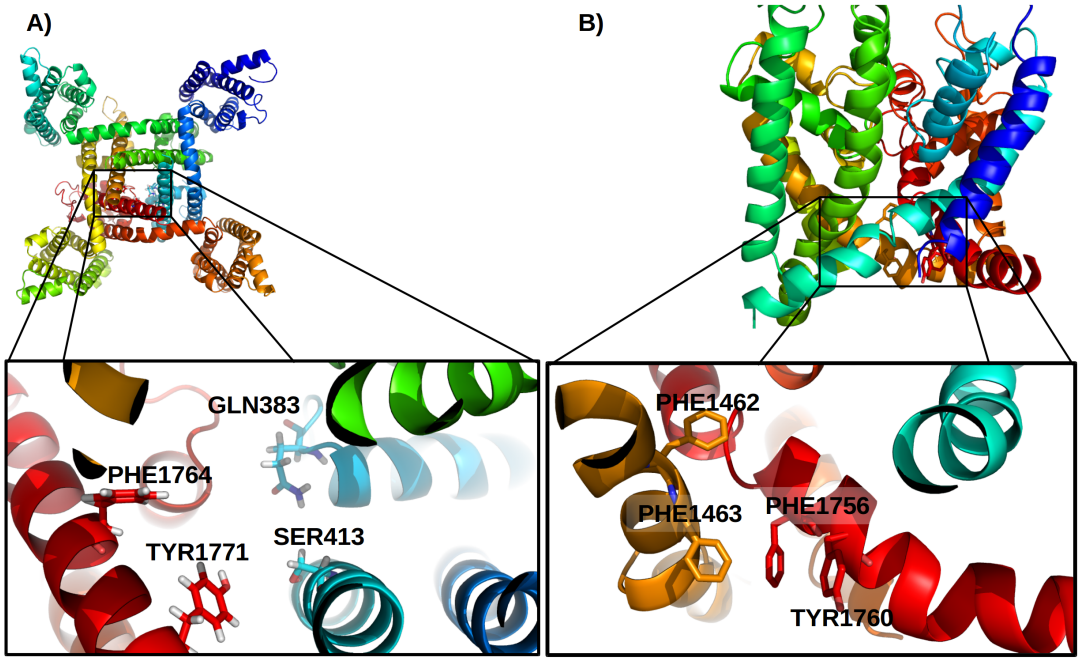


Figura 4.10: Sitio de unión propuesto y residuos flexibles para el docking molecular. A) Modelo cerrado. B) Modelo abierto.

La comparación de las dos simulaciones (con canal cerrado y abierto) se realizó mediante la evaluación de las curvas ROC generadas (figura 4.11). Se utilizó el valor de la función de evaluación de la conformación más estable de cada molécula del conjunto de evaluación como parámetro. En el caso del canal abierto, el área bajo la curva fue de 0,9125; mientras que para el canal cerrado este valor fue de 0,8779. Estos datos indican que, si bien ambos modelos dan resultados bastante similares, el que mejor se ajusta a la clasificación experimental es el que utiliza la estructura del canal abierto. Estos resultados se encuentran en línea con la conocida preferencia de los anticonvulsivos a

una conformación abierta del canal [4]. Por estas razones se decidió utilizar esta última estructura para la búsqueda de nuevos antiepilépticos en bases de datos.

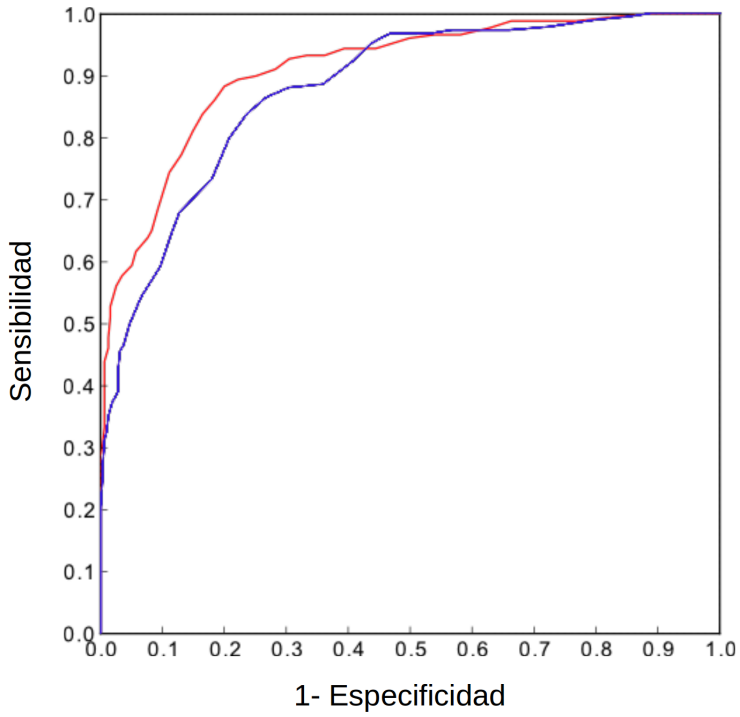


Figura 4.11: Curvas ROC obtenidas durante la simulación del conjunto de evaluación. Azul: modelo del canal cerrado; rojo: modelo del canal abierto

Bibliografía

[1] W. a. Catterall and T. M. Swanson, "Structural Basis for Pharmacology of Voltage-Gated Sodium and Calcium Channels," *Molecular Pharmacology*, vol. 88, pp. 141–150, jun 2015.

[2] D. S. Ragsdale and M. Avoli, "Sodium channels as molecular targets for antiepileptic drugs.," *Brain research. Brain research reviews*, vol. 26, pp. 16–28, mar 1998.

[3] X. Xu, F. Guo, X. Lv, R. Feng, D. Min, L. Ma, Y. Liu, J. Zhao, L. Wang, T. Chen, C. Shaw, L. Hao, and J. Cai, "Abnormal changes in voltage-gated sodium channels NaV1.1, NaV1.2, NaV1.3, NaV1.6 and in calmodulin/calmodulin-dependent protein kinase II, within the brains of spontaneously epileptic rats and tremor rats.," *Brain research bulletin*, vol. 96C, pp. 1–9, apr 2013.

[4] A. Nardi, N. Damann, T. Hertrampf, and A. Kless, "Advances in targeting voltage-gated sodium channels with small molecules.," *ChemMedChem*, vol. 7, pp. 1712–40, oct 2012.

[5] T. U. Consortium, "UniProt: a hub for protein information," *Nucleic Acids Research*, vol. 43, no. D1, pp. D204–D212, 2014.

[6] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The Protein Data Bank.," *Nucleic acids research*, vol. 28, no. 1, pp. 235–242, 2000.

[7] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool.," *Journal of molecular biology*, vol. 215, no. 3, pp. 403–410, 1990.

[8] a. Marchler-Bauer, M. K. Derbyshire, N. R. Gonzales, S. Lu, F. Chitsaz, L. Y. Geer, R. C. Geer, J. He, M. Gwadz, D. I. Hurwitz, C. J. Lanczycki, F. Lu, G. H. Marchler, J. S. Song, N. Thanki, Z. Wang, R. a. Yamashita, D. Zhang, C. Zheng, and S. H. Bryant, "CDD: NCBI's conserved domain database," *Nucleic Acids Research*, vol. 43, no. D1, pp. D222–D226, 2014.

[9] R. D. Finn, P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate,

and A. Bateman, "The Pfam protein families database: towards a more sustainable future," *Nucleic Acids Research*, vol. 44, pp. D279–D285, jan 2016.

[10] Y. Zhang, "I-TASSER server for protein 3D structure prediction.," *BMC bioinformatics*, vol. 9, p. 40, jan 2008.

[11] A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: a unified platform for automated protein structure and function prediction.," *Nature protocols*, vol. 5, pp. 725–38, apr 2010.

[12] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, "The I-TASSER Suite: protein structure and function prediction," *Nature Methods*, vol. 12, no. 1, pp. 7–8, 2014.

[13] J. Zhang and Y. Zhang, "GPCRRD: G protein-coupled receptor spatial restraint database for 3D structure modeling and function annotation," *Bioinformatics*, vol. 26, no. 23, pp. 3004–3005, 2010.

[14] Y. Zhang and J. Skolnick, "Scoring function for automated assessment of protein structure template quality," *Proteins: Structure, Function and Genetics*, vol. 57, no. 4, pp. 702–710, 2004.

[15] R. a. Li, I. L. Ennis, R. J. French, S. C. Dudley, G. F. Tomaselli, and E. Marbán, "Clockwise Domain Arrangement of the Sodium Channel Revealed by μ-Conotoxin (GIIIA) Docking Orientation," *Journal of Biological Chemistry*, vol. 276, no. 14, pp. 11072–11077, 2001.

[16] S. C. Dudley, N. Chang, J. Hall, G. Lipkind, H. a. Fozzard, and R. J. French, "mu-conotoxin GIIIA interactions with the voltage-gated Na(+) channel predict a clockwise arrangement of the domains.," *The Journal of general physiology*, vol. 116, no. 5, pp. 679–690, 2000.

[17] D. a. Case, T. Darden, T. E. C. Iii, C. Simmerling, S. Brook, J. Wang, U. T. Southwestern, R. E. Duke, U.-c. Hill, R. Luo, U. C. Irvine, M. Crowley, R. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, U. C. S. Diego, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, N. C. State, V. Babin, P. a. Kollman, U. C. S. Francisco, D. a. Pearlman, R. V. Stanton, J. Pitera, I. Massova, A. Cheng, P. State, J. J. Vincent, P. Beroza, V. Tsui, C. Schafmeister, W. S. Ross, R. Radmer, G. L. Seibel, J. W. Caldwell, C. Singh, P. Weiner, and P. Cieplak, "AMBER 11 Users' Manual," pp. 1–304.

[18] A. Bhattacharya, R. Tejero, and G. T. Montelione, "Evaluating protein structures determined by structural genomics consortia," *Proteins: Structure, Function, and Bioinformatics*, vol. 66, pp. 778–795, dec 2006.

[19] Y. J. Huang, R. Powers, and G. T. Montelione, "Protein NMR Recall, Precision, and F -measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics," *Journal of the American Chemical Society*, vol. 127, pp. 1665–1674, feb 2005.

[20] R. A. Laskowski, "PROCHECK: a program to check the stereochemicai quality of protein structures.," *J. Appl. Cryst.*, vol. 26, pp. 283–291, 1993.

[21] S. C. Lovell, I. W. Davis, W. B. Adrendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, "Structure validation by C alpha geometry: phi,psi and C beta deviation," *Proteins-Structure Function and Genetics*, vol. 50, no. 3, pp. 437–450, 2003.

[22] R. Lüthy, J. U. Bowie, and D. Eisenberg, "Assessment of protein models with three-dimensional profiles.," *Nature*, vol. 356, no. 6364, pp. 83–85, 1992.

[23] M. Sippl, "Recognition of errors in threedimensional structures of proteins," *Proteins: Structure, Function, and Bioinformatics*, 1993.

[24] D. B. Tikhonov and B. S. Zhorov, "Modeling P-loops domain of sodium channel: homology with potassium channels and interaction with ligands.," *Biophysical journal*, vol. 88, pp. 184–97, jan 2005.

[25] N. Eswar, B. Webb, M. a. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-Y. Shen, U. Pieper, and A. Sali, "Comparative protein structure modeling using MODELLER.," *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, vol. Chapter 2, p. Unit 2.9, 2007.

[26] M. Martí-Renom and A. Stuart, "Comparative protein structure modeling of genes and genomes," *Annual review of ...*, pp. 291–325, 2000.

[27] A. Sali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints.," *Journal of molecular biology*, vol. 234, no. 3, pp. 779–815, 1993.

[28] M. Y. Shen and A. Sali, "Statistical potential for assessment and prediction of protein structures," *Protein Science*, vol. 15, no. 11, pp. 2507–2524, 2006.

[29] a. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. a. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, and J. P. Overington, "The ChEMBL bioactivity database: An update," *Nucleic Acids Research*, vol. 42, no. D1, pp. 1083–1090, 2014.

[30] C. E. Stafstrom, "Mechanisms of action of antiepileptic drugs: the search for synergy.," *Current opinion in neurology*, vol. 23, pp. 157–63, apr 2010.

[31] N. Lenkey, R. Karoly, P. Lukacs, E. S. Vizi, M. Sunesen, L. Fodor, and A. Mike, "Classification of drugs based on properties of sodium channel inhibition: a comparative automated patch-clamp study.," *PloS one*, vol. 5, p. e15568, jan 2010.

[32] O. Trott and A. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of computational chemistry*, vol. 31, no. 2, pp. 455–461, 2010.

[33] A. Mike and P. Lukacs, "The enigmatic drug binding site for sodium channel inhibitors.," *Current molecular pharmacology*, vol. 3, pp. 129–44, nov 2010.

[34] N. M. O'Boyle, M. Banck, C. a. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An Open chemical toolbox," *Journal of Cheminformatics*, vol. 3, no. 10, p. 33, 2011.

[35] J. Gasteiger and M. Marsili, "Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges," *Tetrahedron*, vol. 36, no. 22, pp. 3219–3228, 1980.

[36] G. Morris, R. Huey, and W. Lindstrom, "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility," *Journal of*, vol. 30, no. 16, pp. 2785–2791, 2009.

Capítulo 5

Búsqueda, selección y evaluación de compuestos antiepilépticos.

Una vez construida la estructura de los blancos moleculares y seleccionados los modelos de docking a utilizar, se procedió a la búsqueda de compuestos que actúen por bloqueo de canales de sodio mediante cribado virtual. Posteriormente, aquellos que se predijeron activos se sometieron a un segundo tamizado, con el fin de analizar su interacción con P-gp.

Se utilizaron dos tipos de bibliotecas virtuales. Una de ellas se elaboró para esta investigación, a partir de un conjunto de compuestos sintetizados y evaluados en su acción anticonvulsiva mediante métodos in-vivo en el grupo de trabajo [1, 2]. Particularmente se seleccionaron las estructuras previamente ensayadas frente al test de electroshock máximo (maximal electroshock seizure test, o MES test) [3]. Como se detalló en el capítulo uno, este ensayo ha permitido detectar la acción anticonvulsiva de fármacos posteriormente introducidos al mercado, y forma parte del protocolo utilizado por el Instituto Nacional de Salud de Estados Unidos (ADD-NIH) para la identificación de compuestos activos [4]. Sin embargo, debe aclararse que no existe una correlación directa entre la actividad en MES test y el mecanismo de acción de los fármacos. Anticonvulsivos clásicos, cuya acción principal asociada es el bloqueo de los canales de sodio, son activos frente a este ensayo (fenitoína, carbamacepina y lamotrigina entre otros); pero muchos otros fármacos con

diferente mecanismo de acción también lo son.

Otra clase de bases de datos empleadas fueron aquellas formadas por compuestos aprobados por la FDA (U. S Food and Drugs Administration), nutracéuticos y drogas experimentales. En este caso el criterio utilizado fue analizar estructuras que en muchos casos están indicadas para el tratamiento de otras patologías, por lo que ya han sido evaluadas exitosamente en un gran número de test relacionados con su toxicidad y/o farmacocinética en humanos. El encontrar un segundo uso como antiepiléptico a drogas ya estudiadas permitirá reducir costos, tiempos y vidas de animales de experimentación implicadas en el proceso de descubrimiento de fármacos [5].

Finalmente los compuestos predichos como inhibidores del canal de sodio y no sustrato de P-gp fueron seleccionados según accesibilidad y ensayados mediante métodos in-vitro e in-vivo.

5.1. Análisis de compuestos de síntesis.

Los modelos desarrollados fueron utilizados para evaluar una serie de derivados de sulfamidas previamente obtenidos y ensayados en estudios in-vivo en nuestro laboratorio. El conjunto analizado está formado por 28 compuestos, 23 de los cuales fueron activos en MES test [1, 2].

Los cálculos de docking se realizaron de acuerdo a las condiciones seleccionadas previamente (descriptas en los capítulos tres y cuatro de esta tesis). Se definió un valor de corte de -8,1 Kcal/mol para la discriminación de bloqueantes y no bloqueantes del canal de sodio, lo que implica una sensibilidad del 88 % y una especificidad del 80 %. El modelo computacional clasificó como activos frente al canal de sodio Nav1.2 a tres de ellos (figura 5.1), indicando que la acción anticonvulsiva de los mismos podría estar asociada al bloqueo de dicho blanco. Sobre estos tres compuestos se aplicó el modelo de P-gp, utilizando un valor de corte de -9 Kcal/mol para la diferenciación de sustratos y no sustratos de

P-gp (especificidad 80 % y sensibilidad 73 %). En los tres casos los compuestos se predicen como no sustratos (figura 5.1). Por lo tanto, se procedió a determinar la capacidad inhibitoria sobre el canal de sodio mediante estudios in-vitro de aquella estructura con la función de evaluación más favorable en el canal de sodio (compuesto SIII).

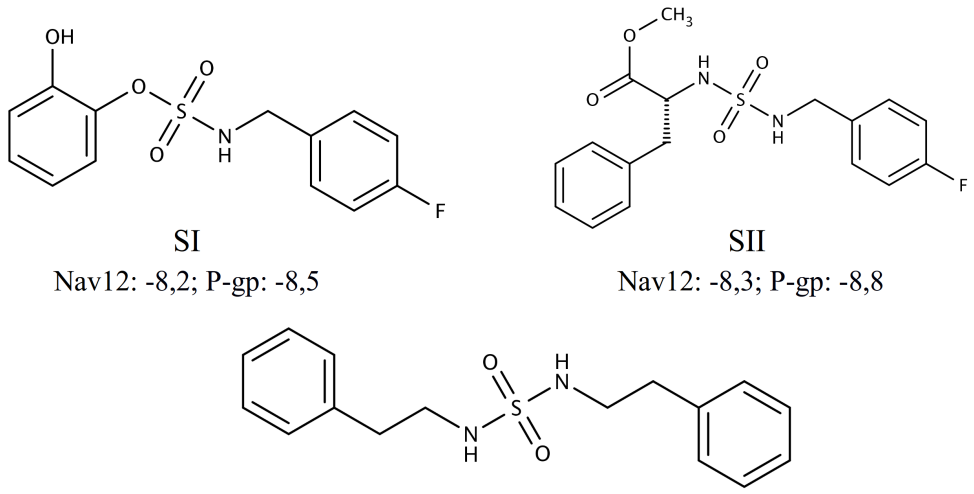


Figura 5.1: Compuestos sintetizados en el LIDeB clasificados como inhibidores del canal de sodio y no sustratos de P-gp. Se reportan los valores de la función de puntuación obtenidos en el cribado para ambos blancos. *SI: sulfamato de 2-hidroxifenil-N-p-fluorobencilamina; SII: N-(N'-p-fluorobencil)sulfamoil-L-fenilalaninato de metilo; SIII: N, N'-difenetilsulfamida.*

5.2. Cribado virtual en bases de datos de fármacos.

Se emplearon las bases de datos ZINC [6] y DrugBank4.0 [7]. En el caso de ZINC, se utilizó una sublibrería de compuestos aprobados para su comercialización por la FDA (ZINC-FDA) . Esta librería fue desarrollada por Stefano Forli, Alex Perryman y Arthur Olson del "The Scripps Research Institute" (TSRI, La Jolla, California), las estructuras se encuentran en coordenadas tridimensionales (3D) y bajo extensión "PDBQT", que es la requerida para utilizar Autodock o Autodock Vina. Esta base consta de 3358 estructuras,

incluyendo tautómeros e isómeros de los compuestos aprobados por la FDA.

En cuanto a DrugBank , se encuentra constituida por 7818 estructuras entre las que se encuentran compuestos aprobados por la FDA, nutracéuticos y drogas experimentales. En esta base de datos los compuestos se encuentran en formato "SDF" y con coordenadas en 2 dimensiones (2D). Las estructuras de la misma fueron inicialmente filtradas según peso molecular, descartando aquellas por debajo de 150 umas y también las que superaban las 600 umas. Como resultado de este filtrado, se obtuvieron 5775 sustancias. Estas se transformaron a coordenadas tridimensionales, donde se las protonó según pH 7,4 y se le asignaron cargas parciales a los átomos. Finalmente, se optimizaron mediante búsqueda conformacional implementada por 750 pasos de Monte-Carlo y la mejor conformación obtenida fue minimizada por 300 ciclos de gradiente conjugado utilizando el campo de fuerza de mecánica molecular MMFF94 [8, 9, 10, 11, 12]. Todas estas etapas previas al cribado fueron realizadas a través el paquete OpenBabel2.1.2 [13].

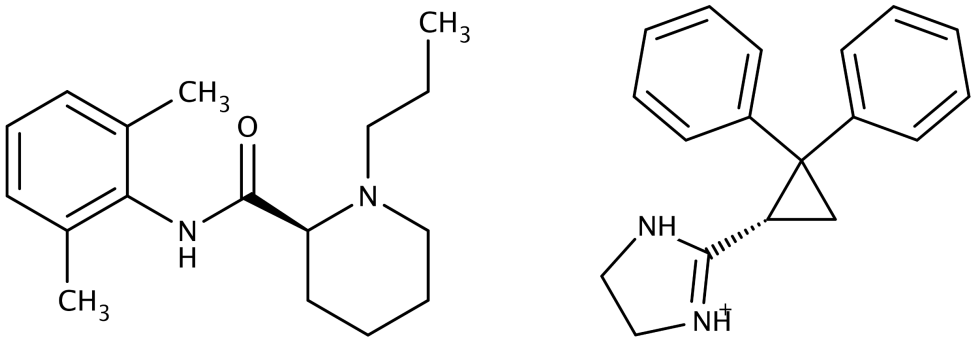
Ambas bases de compuestos se evaluaron con el modelo de canal de sodio y P-gp, a modo de tamizado secuencial. Inicialmente se aplicó el docking sobre el canal de sodio para detectar aquellos compuestos que tuvieran actividad frente al mismo.

En este caso se utilizó como valor de corte -9 Kcal/mol, a efectos de aumentar la especificidad del cribado y, de este modo, disminuir la tasa de falsos positivos. Con este umbral se obtiene una especificidad del 93 % y una sensibilidad del 63 %. Evidentemente, un incremento en la especificidad trae aparejada una pérdida de sensibilidad, esto es, un aumento en el número de falsos negativos. Sin embargo se utiliza este criterio debido a que pueden evaluarse experimentalmente un pequeño número de estructuras, por lo que resulta conveniente evitar realizar ensayos biológicos con compuestos que se predicen activos y no lo son, a expensas de perder compuestos activos en el tamizado. Bajo estas condiciones el modelo clasifica como activas 1181 estructuras para la base ZINC-FDA y 1596 compuestos provenientes de DrugBank.

Los compuestos seleccionados fueron evaluados mediante el modelo de docking frente a P-gp para filtrar y eliminar sustancias que tuvieran alta afinidad a la misma. Nuevamente,

se aplicó una restricción mayor para el valor de corte respecto de la utilizada para la biblioteca de derivados de sulfamidas (sección 5.1). En forma inversa al cribado del canal de sodio, en este caso se buscan estructuras que no interaccionen con el blanco, por lo que se eligió un valor umbral más positivo. Superaron el tamiz aquellas moléculas con un valor de función de evaluación por encima de -8.5 Kcal/mol (Sensibilidad de 89 % y especificidad de 55 %), ya que se consideró que las que tenían un valor menor presentarían una elevada afinidad predicha a P-gp. Esto permitió la selección de 190 estructuras en total.

Posteriormente los candidatos fueron sometidos a una búsqueda bibliográfica, con el fin de determinar si alguno de ellos ha sido previamente estudiado en relación a su interacción con los canales de sodio voltaje operados. Se encontró que el cribado virtual seleccionó la Ropivacaina, un anestésico de uso clínico con actividad inhibitoria sobre los canales de sodio [14, 15] y la Cibenzolina, que también ha sido informado como bloqueante de dicho blanco (Figura 5.2) [16]. La identificación de estas estructuras dentro de la base de datos da cuenta de la capacidad del modelo de encontrar estructuras activas.



Ropivacaina
Nav1.2: -9; P-gp: -7,8

Cibenzolina
Nav1.2: -9,7; P-gp: -8,5

Figura 5.2: Bloqueantes de canal de sodio detectados mediante cribado virtual. Se reportan los valores de la función de puntuación obtenidos en el cribado para este blanco y para P-gp.

Mediante un análisis de la estructura de los 188 compuestos restantes, se preseleccionaron 19 como los mejores candidatos. En este caso se utilizó como criterio la disponibilidad de los mismos en el mercado y/o la estructura química, ya que algunos de ellos presentan fragmentos estructurales comunes a fármacos bloqueantes del canal de sodio (por ejemplo a fenitoína, carbamacepina o lamotrigina). La figura 5.3 muestra los compuestos seleccionados.

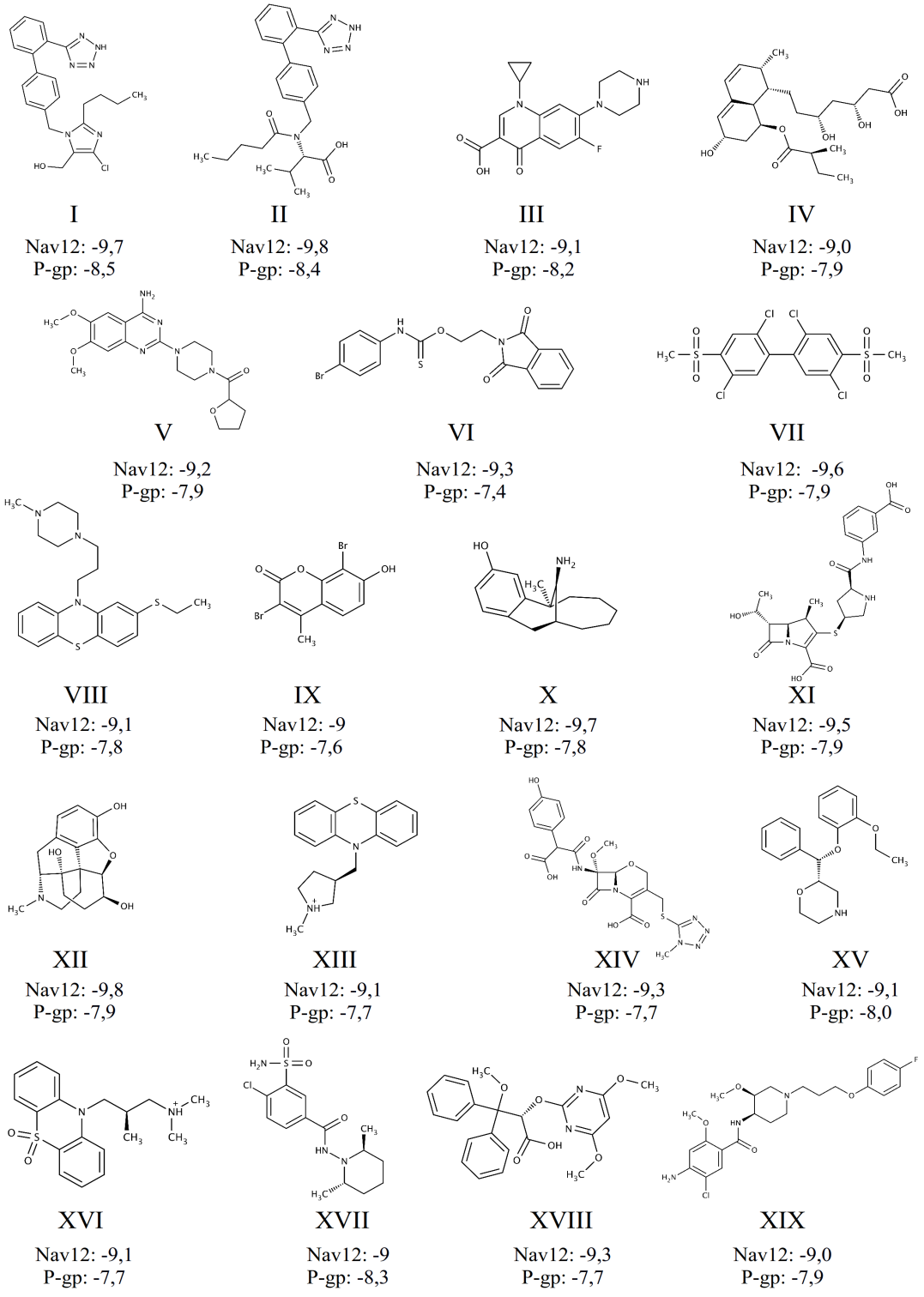


Figura 5.3: Compuestos preseleccionados para los ensayos biológicos. Se reportan los valores de la función de puntuación obtenidos para la interacción con Nav1.2 y para P-gp. I: Losartan ((2-butil-4-cloro-1-[2'-(1H-tetrazol-5-yl)bifenil-4-il]metil-1H-imidazol-5-il)metanol), antihipertensivo; II: Valsartan (ácido (S)-3-metil-2-(N-[2'-(2H-1,2,3,4-tetrazol-5-il)bifenil-4-yl]metilpentanamido)butanoico), antihipertensivo;

figura 5.3 (continuación): III: Ciprofloxacina (ácido 1-ciclopropil-6-fluoro-4-oxo-7-(piperazin-1-il)-quinolin-3-carboxilico), antibiótico; IV: Pravastatina (ácido (3R,5R)-7-[(1S,2S,6S,8S,8aR)-6-hidroxi-2-metil-8- [(2S)-2-metilbutanoil]oxi-1,2,6,7,8,8a-hexahidronaftalen-1-il]-3,5-dihidroxiheptanoico), antihipercolesterolemiante; V: Terazosina (1-(4-Amino-6,7-dimetoxi-2-quinazolinil)-4-((tetrahidro-2-furanil)carbonil)piperazina), antagonista del receptor α1; VI: O-[2-(1,3-dioxoisoindol-2-il)etil] N-(4-bromofenil) tiocarbamato, droga experimental; VII: 4,4'-Bis([H]metilsulfonil)-2,2',5,5'-tetraclorobifenil, droga experimental; VIII:Tietilpirazina, (2-(etiltio)-10-[3-(4-metilpiperazin-1-il)propil]-10H-fenotiazina), antiemético; IX:3,8-dibromo-7-hidroxi-4-metil-2H-cromen-2-ona, droga experimental; X: Dezocina ((1R,9S,15S)-15-amino-1-methyltriciclo[7.5.1.02,^{7}]pentadeca-2,4,6-trien-4-ol), analgésico opioide; XI: Ertapenem (ácido (4R,5S,6S)-3-[(3S,5S)-5-[(3-carboxifenil) carbamoil]pirrolidin-3-il]sulfanil-6-(1-hidroxietil)-4-metil-7-oxo-1-azabiciclo [3.2.0]hept-2-ene-2-carboxilico), antibióticocarbapenem; XII: Hidroximorfinol ((1S,5R,13R,14S,17S)-4-metil-12-oxa-4-azapentaciclo[9.6.1.0^{1},^{13}.0^{5},^{17}.0^{7},^{18}]octadeca-7(18),8,10-trieno-10,14,17-triol), experimental); XIII: Metodilazina (10-[(1-metilpirrolidin-3-il)metil]-10H-fenotiazina), antiinflamatorio; XIV: Moxalactam (ácido (6R,7R)-7-[2-carboxi-2-(4-hidroxifenil)acetamido]-7-metoxi-3-[(1-metil-1H-1,2,3,4-tetrazol-5-il)sulfanil]metil-8-oxo-5-oxa-1-azabiciclo[4.2.0]oct-2-en-2-carboxilico), antibiótico; XV: Reboxetina ((2S)-2-[(S)-2-etoxifenoxi(fenil)metil]morfolina), antidepresivo; XVI: Oxomemazina (3-(5,5-dioxofenotiazin-10-yl)-N,N,2-trimetilpropan-1-amina), antiinflamatorio; XVII: Clopamida (4-cloro-N-[(2S,6S)-2,6-dimetilpiperidin-1-il]-3-sulfamoilbenzamida), diurético; XVIII: Darusentan (ácido (2S)-2-[(4,6-dimetoxipirimidin-2-il)oxi]-3-metoxi-3,3-difenilpropanoico), antihipertensivo; XIX: Cisapride (4-amino-5-cloro-N-((3S,4S)-1-[3-(4-fluorofenoxi)propil]-3-metoxipiperidin-4-il)-2-metoxibenzamida), agente gastroprocinético.

En este conjunto final se encuentra la Terazosina (compuesto V, Figura 5.3) un antagonista del receptor α-1, y la Cisaprida (compuesto XIX), un agente gastroprocinético. Se encontró en literatura que estos fármacos han sido estudiados como inhibidores de canales de potasio operados por voltaje [17, 18, 19, 20]. Dada la similitud estructural existente entre los canales de sodio y de potasio, estos candidatos resultan interesantes para su estudio en el blanco molecular Nav1.2. Al momento estos compuestos no se encuentran disponibles en nuestro laboratorio y no pudieron ser adquiridos, por lo que se propone su estudio en una etapa posterior a esta investigación.

Finalmente pudieron adquirirse tres compuestos para la evaluación biológica, los cuales se encuentran disponibles en el mercado farmacéutico del país. Ellos son: Losartan (I) y Valsartan (II), antagonistas del receptor de angiotensina II utilizados como antihiper-

tensivos y Ciprofloxacina (III) antibiótico del grupo de las fluoroquinolonas. Previo a los ensayos biológicos se corroboró que estos candidatos presenten un coeficiente de partición (log P) en un rango entre 1-5 y una solubilidad en agua tal que permita la preparación de las soluciones necesarias.

5.3. Ensayos biológicos in-vitro.

Se utilizó la técnica de patch-clamp para analizar el efecto de los candidatos sobre la actividad del canal Nav1.2, expresado en un sistema heterólogo. La técnica de patch-clamp permite medir corrientes iónicas de transmembrana manteniendo el potencial de la misma en valores prefijados (en inglés, voltage-clamp), o la variación de potencial de membrana ante un estímulo fijo de corriente (en inglés, current-clamp) [21]. Esta técnica permite identificar los iones que transportan la corriente (y calcular la magnitud de la misma), así como evaluar la sensibilidad al voltaje y la modulación por sustancias endógenas o exógenas de los canales iónicos, lo que la hacen ideal para el estudio de canales iónicos.

La técnica consiste en la aproximación de una micropipeta de vidrio, por medio de un micromanipulador, a la superficie de la membrana de la célula elegida y el establecimiento (mediante una suave succión) de un sello de alta resistencia (en inglés, giga-seal), de aproximadamente 10 GΩ, entre el vidrio de la misma y la membrana celular [22]. Para el registro de la corriente de sodio se trabajó en la configuración denominada de célula completa (en inglés whole-cell, WC), donde se registran corrientes macroscópicas representativas de la totalidad de canales iónicos presentes en la membrana celular. Para la configuración WC es necesario inicialmente formar un giga-seal (figura 5.4 A) mediante una pequeña succión y luego se realiza una segunda presión negativa desde la pipeta (figura 5.4 B), que logra retirar la porción de membrana sellada (figura 5.4 C) [22]. De esta manera la solución de la pipeta dializa la célula y se establece una conexión eléctrica entre el electrodo (en contacto con la solución de la micropipeta) y el interior celular,

como se muestra en el esquema de la figura 5.5.

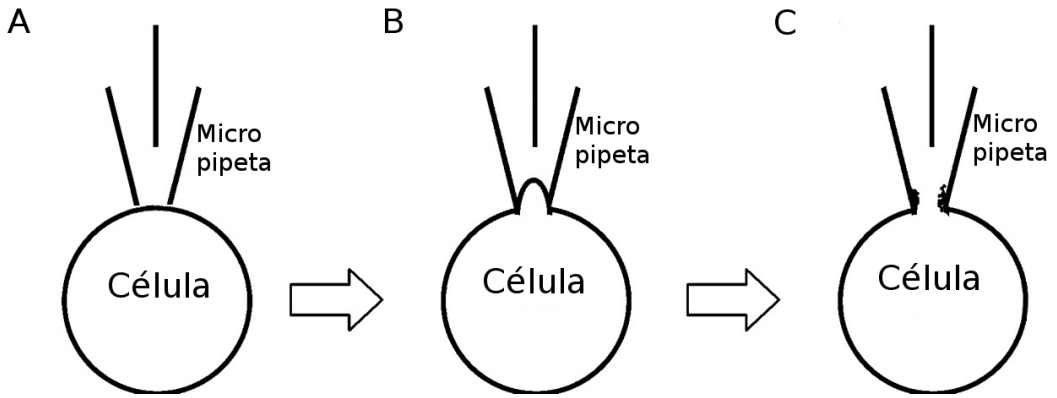


Figura 5.4: Obtención de la configuración whole-cell. A: Mediante la aplicación de una pequeña succión se genera el giga-seal (B). A partir de esta configuración (también conocida como cell-attached) una segunda succión logra retirar la porción de membrana en la pipeta. C: Configuración whole-cell [22].

La configuración WC representa un buen modelo para el estudio de canales iónicos en condiciones cercanas a las fisiológicas, dado que permite mantener gran parte de las estructuras intracelulares, respetar los gradientes de los distintos iones a ambos lados de la membrana y registrar las corrientes en una célula que mantiene su capacidad de establecer un potencial de membrana en el reposo. Estos ensayos fueron realizados por el Dr. Pedro Martin del Instituto de Estudio Inmunológicos y Fisiopatológicos (IIFP, CONICET-UNLP), bajo la dirección de la Dra. Verónica Milesi.

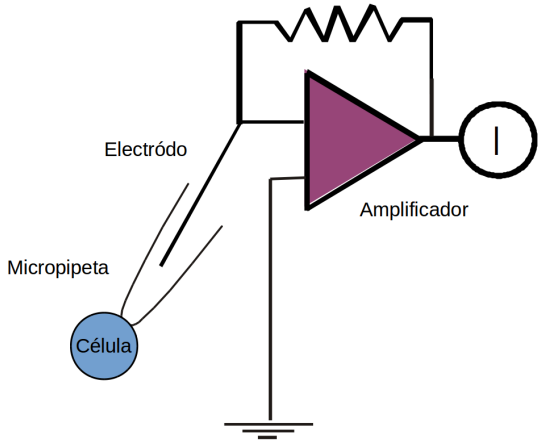


Figura 5.5: Esquema básico del circuito Patch-clamp.

5.3.1. Condiciones experimentales de los ensayos in-vitro.

Todos los experimentos fueron realizados en líneas celulares HEK293 con expresión estable de la subunidad α del canal Nav1.2 (gentilmente donadas por GlaxoSmithKline, Stevenage, UK). Esta expresión fue generada utilizando el vector pCIN5, el cual confiere la resistencia a la Geneticina como método de presión de selección de las células transfectadas [23, 24]. La línea celular fue cultivada en DMEM F12 (Emeve medios, Argentina) suplementado con Suero Fetal Bobino (SFB) al 10 % (Internegocios, Argentina) y Geneticina 50 μg/ml (Sigma-Aldrich). Las células crecieron en una estufa de cultivo REVCO a 37°C y en atmosfera controlada de 5 % de CO_2. El día de realización de los registros electrofisiológicos, las células fueron levantadas con Accutasa (Sigma-Aldrich), resuspendidas en medio sin SFB ni Geneticina, y colocadas en una cámara de fondo de vidrio de 3 ml apta para microscopio invertido. Se seleccionan por su aspecto para intentar el logro de sellos con los microelectrodos de patch-clamp, y se utilizaron durante las 5 a 6 horas posteriores a la preparación.

Las células son observadas a través de un microscopio invertido (Televal 3; Zeiss, Jena, Alemania) estabilizado mecánicamente sobre una mesa antivibratoria (Newport) recubierto con una caja de Faraday para filtrar los ruidos eléctricos del ambiente. La señal

de corriente iónica fue registrada en función del tiempo con un amplificador Axopatch 200A (Axon Instruments, Foster City, CA). Se filtró adecuadamente con un filtro de tipo Bessel de 8 polos a una frecuencia de 2 kHz y se digitalizó (Digidata 1400 Axon Instruments, Foster City, CA) con una frecuencia de muestreo de 200kHz.

Las soluciones ensayadas fueron perfundidas con una pipeta multivías colocada en la proximidad de la célula investigada, todas a temperatura ambiente. Luego de ensayar un protocolo en una célula, se reemplaza la cámara experimental por otra conteniendo una nueva muestra de células. Las distintas drogas ensayadas fueron agregadas sobre las soluciones estándares y, en el caso que correspondiera, los controles fueron realizados en presencia del solvente en el cual se disolvió la droga.

Como se mencionó anteriormente, la configuración de célula completa (WC) permite registrar las corrientes que atraviesan la membrana a través de todos los canales presentes, en respuesta a un pulso de voltaje (modo voltage-clamp). Para el ensayo de los compuestos candidatos se aplicaron protocolos de voltaje en forma de un pulso rectangular, que consisten en un cambio discreto instantáneo del valor de potencial desde un valor determinado (potencial de mantenimiento) a otro valor de potencial que se mantiene durante un tiempo y luego retorna en forma instantánea al valor inicial (figura 5.6).

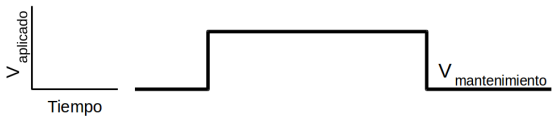


Figura 5.6: Esquema de protocolo de un pulso de voltaje.

El potencial de membrana de mantenimiento fue fijado en -70mV, potencial donde la mayor parte de los canales de sodio se encuentran en estado de reposo, pero que facilita la observación de bloqueos por fármacos Voltaje-dependientes [25]. Las corrientes macroscópicas fueron analizadas en condiciones control y luego del agregado de los distintos

candidatos probados a partir de la aplicación de un pulso de voltaje de 15 ms de duración a -20 mV (potencial de la corriente máxima de los canales Nav1.2) a partir del potencial de mantenimiento (de -70 mV). Se esperaron 10 segundos entre la aplicación de cada pulso para permitir la recuperación de los canales del estado inactivado. El valor de la corriente fue medido al pico negativo, momento en el cual la corriente alcanzó su máximo valor.

Se utilizaron las siguientes soluciones de registro:

- Solución intracelular de la pipeta (SI) en mM: CsCl 126; $MgCl_2$ 4; HEPES 10; EGTA 10; EDTA 1; Na_2ATP 4; llevada a pH 7,2 con CsOH
- Solución extracelular del baño (SE) en mM: NaCl 135; KCl 4,7; $MgCl_2$ 1,2; $CaCl_2$ 2,5; HEPES 10; Glucosa 10; llevada a pH 7,4 con NaOH.

5.3.2. Resultados de los ensayos in-vitro.

Los cuatro compuestos seleccionados mediante el cribado virtual fueron evaluados en el ensayo patch-clamp, siguiendo la técnica descripta con anterioridad. Los mismos se agregaron a la solución extracelular desde una solución madre concentrada (dilución 1:1000). La ciprofloxacina (III) fue disuelta en agua bi-destilada, mientras que el resto en dimetilsulfoxido (DMSO). En los casos donde fue necesario, el solvente se agregó en las condiciones control en la misma dilución de 1:1000.

Los resultados obtenidos se muestran en la figura 5.7. Se observa un esquema del protocolo de voltaje utilizado para generar la corriente típica de sodio y la influencia sobre ella de los compuestos ensayados (Figura 5.7 A). En la figura 5.7 B se observan los valores de corriente porcentual registrados en presencia de los candidatos con respecto al control. Los resultados fueron expresados como la media ± error estándar de la misma. Se realizaron test de t Student para comparar entre los 2 grupos (control y con droga). Las diferencias fueron consideradas estadísticamente significativas en los casos en que el

valor de p fue menor a 0,05.

Como puede observarse, los compuestos N,N'-difenetilsulfamida y ciprofloxacina presentan acción inhibitoria sobre el canal de sodio a las concentraciones ensayadas (100 μM y 300 μM respectivamente). En el caso de la N,N'-difenetilsulfamida se observa una disminución relativa de la corriente del 35,8 % $\pm$ 6,6 (n = 3, p $<$ 0,05), mientras que para la ciprofloxacina este valor es de 29,1 % $\pm$ 5,0 (n = 3, p $<$ 0,05). Por el contrario, losartan y valsartan producen un incremento en el registro de corriente respecto del control (136,0 % $\pm$ 2,3 y 115,9 % $\pm$ 4,1; n = 3, p $<$ 0,05). Si bien para el caso de los dos últimos compuestos el efecto no es el buscado, a partir de estos resultados puede inferirse que los cuatro candidatos parecen interaccionar con el canal Nav1.2, alterando su función. Estos resultados corroboran la capacidad de las estrategias propuestas para el descubrimiento de nuevos compuestos activos.

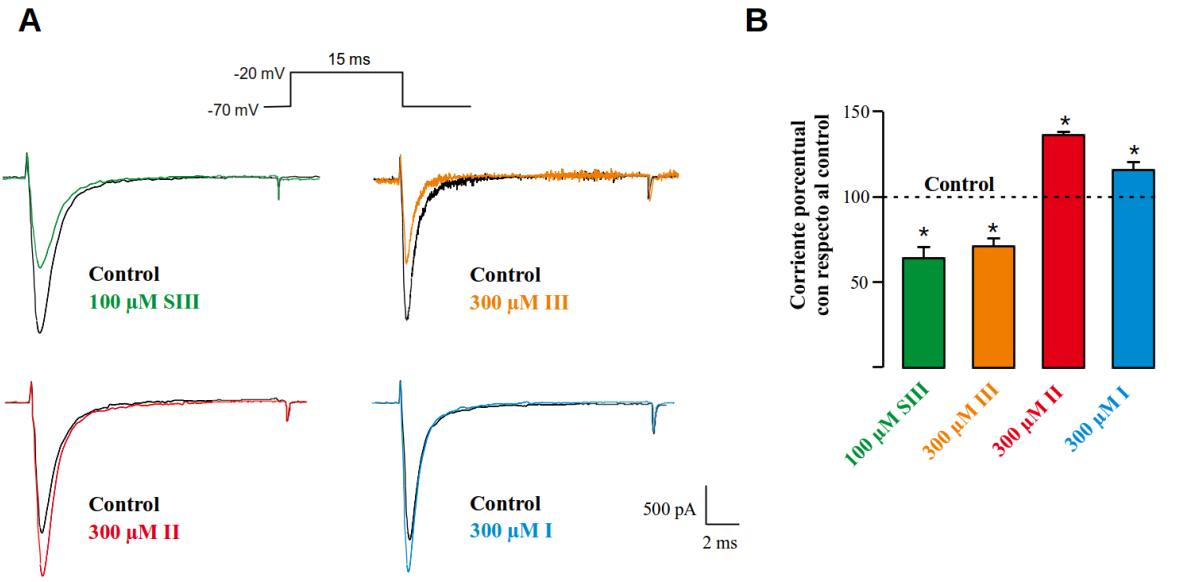


Figura 5.7: Efecto de los candidatos evaluados sobre el canal de sodio Nav1.2. Se grafica en color verde el comportamiento de la N,N'-difenetilsulfamida (compuesto SIII), en amarillo la ciprofloxacina (compuesto III), en rojo Valsartan (compuesto II) y en azul losartan (compuesto I). El control se muestra en negro. A: Corriente típica de sodio evocada por un pulso de voltaje desde un potencial de mantenimiento de -70mV hasta un potencial de -20mV en ausencia y en presencia del compuesto a evaluar. B: Valores medios del bloqueo de la corriente al pico de sodio para cada uno de los compuestos ensayados. Los resultados representan el valor medio de 3 ensayos independientes, las barras indican la desviación estándar y el asterisco indica diferencias significativas respecto al control ($p < 0{,}05$).

5.4. Ensayos biológicos in-vivo.

Para determinar la acción anticonvulsiva in-vivo de los compuestos seleccionados se realizaron estudios en animales, siguiendo los lineamientos propuestos por el programa de desarrollo de anticonvulsivos del Instituto Nacional de Salud de Estados Unidos (ADD-NIH) [4]. Este tipo de estudios son aceptados internacionalmente para la detección inicial de actividad anticonvulsiva en nuevas estructuras. Los ensayos estuvieron a cargo de especialistas en este campo que conforman el grupo de trabajo del LIDeB.

El proceso de desarrollo de nuevos fármacos anticonvulsivos requiere poder evaluar un gran número de compuestos en un tiempo relativamente corto. Para esto, se deben utilizar modelos experimentales sencillos, eficientes (en tiempo y costo) y con un gran poder predictivo [26]. La mayoría de los ensayos en la etapa de identificación de compuestos se lleva a cabo sobre animales sanos, a los que se inducen las crisis por medio de métodos eléctricos o químicos. Aunque existen otros ensayos que podrían ser potencialmente utilizados, los métodos MES (Maximal Electroshock Test) y PTZ (subcutaneous Pentilentetrazol Test) continúan, desde hace más de 60 años, siendo los procedimientos estándar en las etapas tempranas de búsqueda de nuevos fármacos anticonvulsivos [3].Como se explicó anteriormente en esta tesis, en el ensayo MES se induce el episodio convulsivo mediante la aplicación de una descarga eléctrica controlada. Por otra parte, en el test de PTZ, este episodio es provocado mediante la administración de un agente convulsivo: el pentilentetrazol (PTZ). En ambos casos se mide la capacidad de un fármaco, administrado previamente, de evitar la convulsión. Los compuestos con actividad en estos ensayos presentan una buena correlación respecto a su eficacia clínica [3]. Sin embargo, la falta de actividad en estos test no indica que el compuesto sea un mal candidato para el tratamiento de las crisis, ya que no son los únicos ensayos preclínicos para la detección de compuestos promisorios [3]. Dado que los anticonvulsivos actuales cuyo principal mecanismo de acción asociado es el bloqueo de los canales de sodio son activos frente a MES test [27], este ensayo fue el elegido para los estudios in-vivo de los compuestos seleccionados en esta tesis.

5.4.1. Condiciones experimentales de los ensayos in-vivo.

Para todos los ensayos biológicos se utilizaron ratones machos albinos Swiss adultos (18 a 23 gr de peso). Se eligieron animales de la misma edad y similar peso en cada ensayo, para minimizar la variabilidad biológica.

Los animales fueron adquiridos en la Facultad de Ciencias Veterinarias de la Universidad Nacional de La Plata y mantenidos en el bioterio de experimentación del LIDeB

siguiendo las recomendaciones éticas de la guía provista por el Instituto Nacional de Salud de Estados Unidos para el manejo y cuidado de animales de laboratorio [28].

Se permitió que los ratones se habituaran a dicho bioterio durante al menos 24-48 horas antes del ensayo biológico, para evitar la influencia en el mismo del estrés que sufren durante el traslado y cambio de habitat. Para esto se los colocó en una habitación exclusiva para el mantenimiento de ratones, diferente a la de ensayos biológicos. En esta etapa se mantuvieron en jaulas colectivas, con alimento y agua ad-libitum, bajo un régimen de ciclos luz/oscuridad de 12 hs.

Los ensayos se realizaron siempre a la misma hora del día. Los ratones seleccionados, que se utilizaron por única vez, se trasladaron a la sala de ensayos biológicos y se aislaron en compartimientos individuales. Luego del ensayo, los animales se colocaron nuevamente en jaulas colectivas y se mantuvieron en observación durante 72 horas para detectar posibles efectos tóxicos que ocasionen la muerte. Para una correcta evaluación de los compuestos, estos deben administrarse a los animales en una solución, suspensión o emulsión homogénea. Para ello se acepta la disolución de las drogas en solución fisiológica o polietilenglicol 400 (PEG 400), de acuerdo a su solubilidad. Las drogas a ensayar se administraron vía intraperitoneal (i.p.).

Ensayo MES.

Como se indicó anteriormente, en el ensayo MES debe inducirse la convulsión mediante una descarga eléctrica controlada. Según lo previsto por el ADD-NIH [4], se deben administrar tres dosis de la droga a evaluar: 30, 100 y 300 mg/Kg. Cada una de estas dosis debe ser evaluada a dos tiempos luego de la administración de la droga: 0,5 y 4 hs. Es por eso que inicialmente se administra el compuesto a ensayar al ratón y, al tiempo estipulado en el ensayo, se induce el estímulo eléctrico. Se mide entonces la capacidad de la droga de prevenir la crisis convulsiva.

La convulsión máxima por electrochoque es provocada en los ratones por medio del suministro de un estimulo eléctrico de 60 Hz/50 mA de 0,2 segundos de duración uti-

lizando electrodos de oreja y generada por una unidad ECT 57800 de UGO Basile®. Con estas magnitudes se obtiene un tren de ondas rectangulares correspondientes a un estímulo aplicado 4-5 veces mayor que el necesario para llegar al umbral de convulsión. En relación al comportamiento del animal, una convulsión máxima provocada por este equipo implica una flexión tónica de sus miembros posteriores por un corto tiempo, seguida de un período más largo de extensión tónica y finalmente un episodio clónico. Particularmente, si la extensión tónica no es mayor a 90 grados con respecto al eje de su cuerpo, indica que el compuesto administrado protege frente a la crisis [4].

Ensayo de Rotorod.

La toxicidad inducida por los compuestos a ensayar se determina inicialmente mediante el uso del ensayo de Rotorod [4]. Se busca detectar efectos neurotóxicos que se manifiestan como sedación o ataxia. Un ratón sano puede mantener el equilibrio durante un período prolongado de tiempo en un cilindro que gira a razón de 6 r.p.m.. El ensayo mide la capacidad del animal de mantener una caminata en el cilindro giratorio durante 1 minuto en 3 instancias sucesivas.

Se utilizó un aparato construido en la Facultad de Ingeniería de la Universidad Nacional de La Plata, en analogía con el equipo Rota-Rod trend mill formice Ugo Basile®. El equipo consiste en un cilindro estriado de 3 cm de diámetro. Los ratones se ubican en el cilindro en movimiento, separados por paneles circulares. Los animales que no mantienen el equilibrio caen en una bandeja situada debajo del cilindro. Esta evaluación se realiza previa a la administración de las drogas a evaluar, a fin de verificar el buen estado de los animales. Se repite el mismo ensayo en forma posterior a la administración del compuesto a evaluar y previo a la inducción convulsiva, con el objeto de verificar si se ha producido una depresión del sistema nervioso central por la administración de la droga (manifestada como ataxia o sedación).

De acuerdo a los resultados obtenidos en MES test y Rotorod, los compuestos pueden clasificarse en cuatro categorías [29]:

- Clase 1: Compuestos activos a dosis de hasta 100 mg/Kg.
- Clase 2: Compuestos activos a dosis mayores a 100 mg/Kg.
- Clase 3: Compuestos inactivos a dosis inferiores a 300 mg/Kg.
- Clase 4: Compuestos activos a dosis inferiores a 100 mg/Kg y tóxicos a esa misma dosis.

De este modo, aquellos compuestos que no manifiesten actividad apreciable o por el contrario, presenten una alta neurotoxicidad pueden descartarse rápidamente (clases 3 y 4), mientras que los más activos (clase 1) pasan a la siguiente fase del programa ADD-NIH [4]. Esta implica la cuantificación de la dosis anticonvulsiva en ratones mediante el cálculo de la dosis efectiva media (ED50) en MES y PTZ. Asimismo se calcula la dosis tóxica media (TD50) y con ella el índice terapéutico (PI) que presenta el compuesto. Las drogas con actividad intermedia (clase 2), pueden considerarse o no de acuerdo al criterio del investigador, generalmente se considera lo novedoso de su estructura para seguir siendo estudiadas.

5.4.1.1. Resultados de los ensayos in-vivo.

Se presentan a continuación los resultados obtenidos mediante la evaluación biológica frente a los ensayos MES y Rotorod de los candidatos seleccionados en esta tesis. Siguiendo los lineamientos del programa ADD-NIH se midió la respuesta frente a estos ensayos luego de 0,5 y 4 hs de administrado el compuesto; a concentraciones de 30, 100 mg/kg para cada tiempo. Debe aclararse que, como se indicó previamente, el programa incluye concentraciones mayores además de las aquí ensayadas. El criterio utilizado para prescindir de estas dosis altas fue que se buscan en este trabajo compuestos muy activos, capaces de superar la potencia de anticonvulsivos conocidos. Por lo tanto, encontrar compuestos activos a concentraciones mayores a los límites de clase 1 implicaría descartar el compuesto. De esta manera se evitó sacrificar animales, tiempo y dinero en resultados que no se utilizarán.

La Tabla 5.1 muestra los resultados obtenidos en el ensayo MES y Rotorod para los compuestos seleccionados por cribado virtual. Se observan propiedades anticonvulsivas in-vivo en tres de los cuatro compuestos considerados.

La N,N'-difenetilsulfamida (candidato perteneciente al cribado de la librería de compuestos de síntesis del laboratorio), ya había sido ensayada previamente por el grupo de trabajo [1]. Esta molécula y valsartan pueden definirse como candidatos de clase 1, es decir que pertenecen al grupo más activo de compuestos. En el caso de losartan, no se observa acción a las dosis ensayadas (clase 3). Para ciprofloxacina se encontró actividad anticonvulsiva a la menor concentración testeada, lo que da idea de una potencia interesante. Sin embargo uno de los ratones no superó el test de Rotorod a dosis de 30mg/Kg por lo que se proponen estudios posteriores a menores concentraciones para observar si la actividad anticonvulsiva prevalece sin la presencia de efectos neurotóxicos.

Compuesto	Clase MES	Dosis mg/kg	MES [a]		RotoRod[b]	
			0.5	4	0.5	4
SIII[c]	1	30	0/3	0/3	0/3	0/3
		100	3/3	0/3	0/3	0/3
I	3	30	0/2	0/2	0/2	0/2
		100	0/2	0/2	0/2	0/2
II	1	30	0/2	0/2	0/2	0/2
		100	1/2	1/2	0/2	0/2
III	4	30	0/2	2/4	1/2	0/4
		100	0/2	1/4	0/2	0/2

Tabla 5.1: Resultados de los ensayos realizados en ratones para determinar acción convulsiva de los compuestos seleccionados. a (Número de animales protegidos / número de animales ensayados). b (Número de animales con presencia de sedación o ataxia / número de animales ensayados). c Actividad anticonvulsiva determinada previamente [1].

Bibliografía

[1] L. Gavernet, I. a. Barrios, M. S. Cravero, and L. E. Bruno-Blanch, "Design, synthesis, and anticonvulsant activity of some sulfamides.," *Bioorganic medicinal chemistry*, vol. 15, pp. 5604–14, aug 2007.

[2] M. L. Villalba, A. V. Enrique, J. Higgs, R. A. Castaño, S. Goicoechea, F. D. Taborda, L. Gavernet, I. D. Lick, M. Marder, and L. E. Bruno Blanch, "Novel sulfamides and sulfamates derived from amino esters: Synthetic studies and anticonvulsant activity," *European Journal of Pharmacology*, feb 2016.

[3] M. A. Rogawski, "Diverse Mechanisms of Antiepileptic Drugs in the Development Pipeline," *Epilepsy Research*, vol. 69, no. 3, pp. 273–294, 2006.

[4] J. P. Stables and H. J. Kupferberg, "The NIH Anticonvulsant Drug Development (ADD) Program: preclinical anticonvulsant screening project," in *Molecular and cellular targets for antiepileptic drugs* (G. Avanzani, G. Regesta, P. Tanganelli, and M. Avoli, eds.), ch. 16, pp. 191–198, 1997.

[5] D.-L. Ma, D. S.-H. Chan, and C.-H. Leung, "Drug repositioning by structure-based virtual screening.," *Chemical Society reviews*, vol. 42, no. 5, pp. 2130–41, 2013.

[6] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, and R. G. Coleman, "ZINC: A free tool to discover chemistry for biology," *Journal of Chemical Information and Modeling*, vol. 52, no. 7, pp. 1757–1768, 2012.

[7] V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. MacIejewski, D. Arndt, M. Wilson, V. Neveu, A. Tang, G. Gabriel, C. Ly, S. Adamjee, Z. T. Dame, B. Han, Y. Zhou, and D. S. Wishart, "DrugBank 4.0: Shedding new light on drug metabolism," *Nucleic Acids Research*, vol. 42, no. D1, pp. 1091–1097, 2014.

[8] T. A. Halgren, "Merck Molecular Force Field. I.," *J. Comput. Chem.*, vol. 17, pp. 490–519, 1996.

[9] T. A. Halgren, "Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions," *Journal of Computational Chemistry*, vol. 17, no. 5-6, pp. 520–552, 1996.

[10] T. A. Halgren, "Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94*," *J. Comput. Chem.*, vol. 17, no. 1 996, pp. 553–584, 1996.

[11] T. A. Halgren and R. B. Nachbar, "Merck molecular force field. IV. conformational energies and geometries for MMFF94," *Journal of Computational Chemistry*, vol. 17, no. 5-6, pp. 587–615, 1996.

[12] T. a. Halgren, "Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules," *Journal of Computational Chemistry*, vol. 17, pp. 616–641, apr 1996.

[13] N. M. O'Boyle, M. Banck, C. a. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An Open chemical toolbox," *Journal of Cheminformatics*, vol. 3, no. 10, p. 33, 2011.

[14] B. G. Liu, X. L. Zhuang, S. T. Li, and G. H. Xu, "The effects of ropivacaine on sodium currents in dorsal horn neurons of neonatal rats.," *Anesthesia and analgesia*, vol. 90, no. 5, pp. 1034–1038, 2000.

[15] H. W. Cheng, H. T. Yang, J. J. Zhou, Y. H. Ji, and H. Y. Zhu, "Pharmacological modulation of brain Nav1.2 and cardiac Nav1.5 subtypes by the local anesthetic ropivacaine," *Neuroscience Bulletin*, vol. 26, no. 4, pp. 289–296, 2010.

[16] R. Niwa, H. Honjo, I. Kodama, K. Maruyama, and J. Toyama, "Na+ channel blocking effects of cibenzoline on guinea-pig ventricular cells.," *European journal of pharmacology*, vol. 352, no. 2-3, pp. 317–27, 1998.

[17] D. Thomas, A. B. Wimmer, K. Wu, B. C. Hammerling, E. K. Ficker, Y. A. Kuryshev, J. Kiehn, H. A. Katus, W. Schoels, and C. A. Karle, "Inhibition of human ether-a-go-go-related gene potassium channels by alpha 1-adrenoceptor antagonists prazosin, doxazosin, and terazosin," *Naunyn-Schmiedeberg's Archives of Pharmacology*, vol. 369, no. 5, pp. 462–472, 2004.

[18] B. D. Walker, C. B. Singleton, J. a. Bursill, K. R. Wyse, S. M. Valenzuela, M. R. Qiu, S. N. Breit, and T. J. Campbell, "Inhibition of the human ether-a-go-go-related gene (HERG) potassium channel by cisapride: affinity for open and inactivated states.," *British journal of pharmacology*, vol. 128, no. 2, pp. 444–450, 1999.

[19] J. Chen, G. Seebohm, and M. C. Sanguinetti, "Position of aromatic residues in the S6 domain, not inactivation, dictates cisapride sensitivity of HERG and eag potassium channels.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 19, pp. 12461–12466, 2002.

[20] J. Lin, J. Guo, H. Gang, P. Wojciechowski, J. T. Wigle, and S. Zhang, "Intracellular K+ is required for the inactivation-induced high-affinity binding of cisapride to HERG channels.," *Molecular pharmacology*, vol. 68, no. 3, pp. 855–65, 2005.

[21] O. P. Hamill, A. Marty, E. Neher, B. Sakmann, and F. J. Sigworth, "Improved Patch-Clamp Techniques for High-Resolution Current Recording from Cells and Cell-Free Membrane Patches," *Pflügers Archiv - European Journal of Physiology*, vol. 391, pp. 85–100, aug 1981.

[22] C. Li, "A reliable whole cell clamp technique.," *Advances in physiology education*, vol. 32, no. 3, pp. 209–211, 2008.

[23] Y. H. Chen, T. J. Dale, M. A. Romanos, W. R. Whitaker, X. M. Xie, and J. J. Clare, "Cloning, distribution and functional analysis of the type III sodium channel from human brain," *Eur J Neurosci*, vol. 12, pp. 4281–4289, 2000.

[24] S. A. Burbidge, T. J. Dale, A. J. Powell, W. R. Whitaker, X. M. Xie, M. A. Romanos, and J. J. Clare, "Molecular cloning, distribution and functional analysis of the NAV1.6. Voltage-gated sodium channel from human brain," *Molecular Brain Research*, vol. 103, pp. 80–90, jun 2002.

[25] X. Qiao, G. Sun, J. J. Clare, T. R. Werkman, and W. J. Wadman, "Properties of human brain sodium channel α-subunits expressed in HEK293 cells and their modulation by carbamazepine, phenytoin and lamotrigine.," *British journal of pharmacology*, vol. 171, pp. 1054–67, feb 2014.

[26] W. Löscher, "Critical review of current animal models of seizures and epilepsy used in the discovery and development of new antiepileptic drugs," *Seizure*, vol. 20, no. 5, pp. 359–368, 2011.

[27] a. M. Waszkielewicz, a. Gunia, N. Szkaradek, K. Słoczyńska, S. Krupińska, and H. Marona, "Ion channels as drug targets in central nervous system disorders.," *Current medicinal chemistry*, vol. 20, pp. 1241–85, jan 2013.

[28] National Academy of Sciences (US), *GUIDE FOR THE CARE AND USE OF LABORATORY ANIMALS*. Washington, D.C.: National Academies Press, dec 2011.

[29] B. Malawska, K. Kulig, A. Śpiewak, and J. P. Stables, "Investigation into new anticonvulsant derivatives of α-substituted N-benzylamides of γ-hydroxy- and γ-acetoxybutyric acid. Part 5: Search for new anticonvulsant compounds," *Bioorganic and Medicinal Chemistry*, vol. 12, no. 3, pp. 625–632, 2004.

Capítulo 6

Análisis de resultados y conclusiones.

En esta investigación se ha desarrollado y validado un protocolo de cribado virtual capaz de identificar sustancias con actividad anticonvulsiva. Para esto, se aplicaron estrategias de diseño racional de fármacos asistido por computadora para el descubrimiento de nuevas estructuras activas. Específicamente, se abordó la problemática de la falta de eficacia de los fármacos anticonvulsivos, los cuales no logran controlar las crisis epilépticas de alrededor del 30 % de los pacientes.

Se seleccionó como blanco molecular a los canales de sodio operados por voltaje. Se trabajó sobre el isotipo Nav1.2, dado que este se encuentra involucrado en fenómenos epilépticos y tiene probada interacción con fármacos anticonvulsivos. Como resultado de este trabajo se generaron dos modelos tridimensionales de la arquitectura del poro de dicho canal en humanos (estados cerrado y abierto). En ambos casos, y debido a la escasa información estructural disponible en el Protein Data Bank, debió realizarse un modelado independiente para cada uno de los segmentos del canal y luego ensamblar los resultados para la construcción del poro.

La estructura de conformación cerrada del poro fue predicha mediante un método de reconocimiento del plegado (threading). Este modelo fue luego validado mediante el uso de los gráficos de Ramachandran y de un análisis de la ubicación de residuos altamente

conservados en la zona del filtro de selectividad (anillo DEKA).

La construcción de un modelo abierto de Nav1.2 constituye uno de los mayores logros de este trabajo de tesis. Como se detalló anteriormente, los anticonvulsivos tienen una alta afinidad al poro del canal cuando este se encuentra en un estado abierto. Dada la ausencia de estructuras experimentales en esta conformación, el molde a utilizar para la predicción estructural debió "fabricarse". Para ello se realizó un análisis de las conformaciones de canales experimentales disponibles, seleccionándose el canal iónico no específico NaK de *Bacilus cereus*, que cuenta con estructuras en estado cerrado y abierto. Esta comparación permitió detectar que la zona interna del poro, debajo del filtro de selectividad DEKA es la más afectada por el cambio conformacional. Por lo tanto se construyó una estructura molde propia resultante del reemplazo de las coordenadas del segmento S6 (que conforma el poro del canal) de la estructura del canal de sodio de la bacteria *Arcobacter butzleri* (en estado cerrado) por el correspondiente a NaK en estado abierto. Finalmente el modelo fue construido mediante un método comparativo y validado de igual manera que para el estado cerrado.

Otra contribución de esta investigación fue la individualización de los programas y condiciones de docking capaces de identificar compuestos que se predicen bloqueantes del canal Nav1.2. Esto se logró mediante un análisis de los resultados de las simulaciones aplicadas a un conjunto de evaluación de compuestos que se conocen activos e inactivos en este blanco. Se concluyó que Autodock Vina es un programa capaz de clasificar en forma aceptable las estructuras de ambos subconjuntos, especialmente si se considera la conformación abierta del canal y se permite la movilidad de ciertos residuos de la zona interna del poro.

A pesar de que esta investigación está enmarcada en las primeras etapas del descubrimiento de estructuras activas para el tratamiento de las crisis convulsivas, se decidió considerar de alguna manera la problemática de la epilepsia refractaria. Esta estrategia está basada en la idea de contribuir al conocimiento no solo mediante el hallazgo de compuestos activos, sino también por encontrar estructuras que puedan ser el punto

de partida para satisfacer la demanda de pacientes resistentes a la medicación actual. Además, la consideración temprana de mecanismos de farmacorresistencia permite ahorrar tiempo, esfuerzo, dinero y vidas de animales de experimentación en candidatos que luego deben ser descartados por no superar los criterios de selección en las etapas subsiguientes del desarrollo de fármacos. Por lo tanto, en esta tesis se incorporó como anti-target a la glicoproteína P (P-gp), un transportador de eflujo parcialmente responsable de la incapacidad de ciertos fármacos de acceder al receptor en concentraciones suficientes para ejercer su acción. Para ello se construyó y validó un modelo tridimensional del transportador utilizando herramientas basadas en la reconocimiento del plegado.

Otro aporte importante de esta investigación fue la identificación de las condiciones de docking capaces de discriminar adecuadamente entre sustratos y no sustratos de P-gp. Debido a la controversia encontrada en literatura respecto a la clasificación de ciertos compuestos como sustratos o no sustratos, debió adoptarse un criterio de selección para la creación del conjunto de evaluación. Esto implicó un análisis de los resultados publicados en distintos tipos de ensayos para estructuras conocidas. Como resultado se obtuvo un conjunto pequeño de 53 compuestos que sirvió para seleccionar a Autodock Vina como el programa de docking capaz de discriminar en forma aceptable entre sustratos y no sustratos, considerando cierta flexibilidad en residuos presentes en la gran cavidad de la proteína que define el sitio activo. Posteriormente el conjunto de evaluación fue ampliado hasta contener unas 1200 estructuras, las cuales fueron clasificadas correctamente por el modelo de simulación previamente utilizado.

Los modelos tridimensionales y la determinación de las condiciones para la simulación por docking de las interacciones con P-gp y canal de sodio fueron las herramientas que debieron construirse para la realización del cribado virtual. Se seleccionó para ello una biblioteca propia que contiene compuestos preparados en el laboratorio y bases de datos que incluyen estructuras aprobadas por la FDA, nutracéuticos y drogas experimentales. Se pensó en analizar estructuras indicadas para el tratamiento de otras patologías porque estas ya han sido evaluadas exitosamente en un gran número de test relacionados con su

toxicidad y/o farmacocinética en humanos.

El cribado se realizó en forma secuencial. Inicialmente 28 compuestos propios y unos 10000 provenientes de bases de datos se sometieron a una simulación por docking sobre el canal de sodio, recuperándose unos 2700 como candidatos a interaccionar con Nav1.2 (3 de ellos compuestos preparados en el laboratorio). Estas fueron analizadas en su interacción con P-gp mediante docking, obteniéndose unas 200 estructuras que se predicen no sustratos de dicha proteína. Cuatro compuestos (uno de ellos proveniente de la biblioteca de compuestos disponible en el laboratorio) fueron evaluados en modelos in vitro e in vivo. Ellos son: N,N'-difenetilsulfamida, Losartan, Valsaltan y Ciprofloxacina.

Los estudios in vitro reflejan que en todos los casos existe una interacción de los candidatos con el canal de sodio. Esto valida la capacidad de la técnica de docking molecular implementada durante el cribado de detectar compuestos capaces de interaccionar con los mismos. Quedan pendientes ensayos adicionales capaces de describir el comportamiento de esta interacción con mayor profundidad. En particular, aquellos destinados a caracterizar el efecto de los compuestos respecto a la recuperación del estado de inactivación en función del tiempo (inhibición dependiente del uso) y del potencial de mantenimiento (inhibición dependiente de voltaje). Estos efectos se encuentran bien estudiados en los fármacos anticonvulsivos que actúan mediante el bloqueo de canales de sodio. Como ya se mencionó en el capítulo 1 de esta tesis, estos compuestos actúan estabilizando el estado inactivado, por lo que presentan una inhibición dependiente de la frecuencia de despolarización y del voltaje de mantenimiento durante el ensayo.

De acuerdo a las evaluaciones in vivo, la mayoría de los compuestos muestra actividad antiepiléptica frente a MES test, a excepción de losartan. En el caso de N,N'-difenetilsulfamida y ciprofloxacina estos resultados son consistentes con los obtenidos en el ensayo patch-clamp. Respecto a valsartan, si bien en principio los resultados podrían parecer contradictorios, cabe la posibilidad de que el efecto anticonvulsivo sea debido a otro mecanismo de acción o incluso estar asociado con un retardo en la recuperación de la inactivación de los canales de sodio.

Adicionalmente no se observa neurotoxicidad en forma de sedación o ataxia excepto para ciprofloxacina, donde un animal no superó el test de Rotorod. En este caso en particular, dado que el compuesto presentó actividad a bajas concentraciones (30 mg/kg) se debería ensayar a dosis más bajas para ver si continúa la actividad sin la presencia de neurotoxicidad.

A partir de los resultados obtenidos, es posible demostrar que los modelos tridimensionales construidos y la metodología basada en la estructura del receptor (bajo las condiciones ensayadas) han sido capaces de enriquecer una base de compuestos distinguiendo satisfactoriamente activos de los que no lo son. Este aspecto es de notable interés, ya que estos modelos quedan disponibles para futuras investigaciones que impliquen, por ejemplo, el cribado virtual de nuevas bases de datos.

Publicaciones realizadas durante el periodo de desarrollo de la presente tesis.

Durante este trabajo de tesis se han realizado 8 presentaciones a congresos (3 internacionales y 5 nacionales), se han publicado 7 artículos en revistas científicas con referato, 1 capítulo de libro y 1 revisión (ambas aun en redacción).

A continuación se detallan las publicaciones realizadas:

Artículos:

- C. Pizzo, C. Saiz, A. Talevi, L. Gavernet, P. Palestro, C. Bellera, L. B. Blanch, D. Benítez, J. J. Cazzulo, A. Chidichimo, P. Wipf, and S. G. Mahler, "Synthesis of 2-hydrazolyl-4-thiazolidinones based on multicomponent reactions and biological evaluation against Trypanosoma Cruzi.," *Chemical biology drug design*, vol. 77, pp. 166–72, mar 2011
- L. Gavernet, P. H. Palestro, and L. Bruno-Blanch, "Docking Applied to the Study of Inhibitors of c-Met Kinase," *ISRN Physical Chemistry*, vol. 2012, pp. 1–5, 2012
- M. E. Di Ianni, A. V. Enrique, P. H. Palestro, L. Gavernet, A. Talevi, and L. E. Bruno-Blanch, "Several New Diverse Anticonvulsant Agents Discovered in a Virtual Screening Campaign Aimed at Novel Antiepileptic Drugs to Treat Refractory Epilepsy," *Journal of Chemical Information and Modeling*, vol. 52, pp. 3325–3330, dec 2012
- L. Gavernet, J. L. Gonzalez Funes, P. H. Palestro, L. E. Bruno Blanch, G. L. Estiu, A. Maresca, I. Barrios, and C. T. Supuran, "Inhibition pattern of sulfamide-related compounds in binding to carbonic anhydrase isoforms I, II, VII, XII and XIV.," *Bioorganic medicinal chemistry*, vol. 21, pp. 1410–8, mar 2013

- P. H. Palestro, L. Gavernet, G. L. Estiu, and L. E. Bruno Blanch, "Docking Applied to the Prediction of the Affinity of Compounds to P-Glycoprotein," *BioMed Research International*, vol. 2014, pp. 1–10, 2014
- C. L. Bellera, D. E. Balcazar, M. C. Vanrell, a. F. Casassa, P. H. Palestro, L. Gavernet, C. a. Labriola, J. Gálvez, L. E. Bruno-Blanch, P. S. Romano, C. Carrillo, and A. Talevi, "Computer-guided drug repurposing: Identification of trypanocidal activity of clofazimine, benidipine and saquinavir," *European Journal of Medicinal Chemistry*, vol. 93, pp. 338–348, 2015
- M. L. Villalba, P. Palestro, M. Ceruso, J. L. Gonzalez Funes, A. Talevi, L. Bruno Blanch, C. T. Supuran, and L. Gavernet, "Sulfamide derivatives with selective carbonic anhydrase VII inhibitory action," *Bioorganic Medicinal Chemistry*, vol. 24, no. 4, pp. 894–901, 2016