



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

FACULTAD DE QUÍMICA

MODULADORES DE DIANAS EPIGENÉTICAS PARA EL  
TRATAMIENTO Y/O PREVENCIÓN DE MÚLTIPLES  
ENFERMEDADES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN QUÍMICA FARMACEÚTICA BIOLÓGICA

PRESENTA:

HASSAN VILLEGAS QUINTERO

TUTORES:

DR. JOSÉ LUIS MEDINA FRANCO

MA. EN C. FERNANDA ISABEL SALDÍVAR GONZÁLEZ  
(ASESOR TÉCNICO)



Ciudad Universitaria, CDMX, 2023



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Agradecimientos

Al Programa de Apoyo a Proyectos de Investigación Tecnológica (UNAM-DGAPA-PAPIIT), proyecto IN201321 y al proyecto IV200121.

A los recursos UNAM-Huawei para el proyecto 7: "Desarrollo de modelos de inteligencia artificial para el tratamiento de diabetes mellitus y diabetes".

## **Jurado**

Presidente Dr. Gómez Martínez Atonatiu Edmundo

Vocal Dr. Hernández Luis Francisco

Secretario Dr. Medina Franco José Luis

Primer suplente Ma. en C. Medina Reyes Liz Jannet

Segundo suplente Dra. González Trujano María Eva

# Resumen

Los moduladores de dianas epigenéticas son compuestos que tienen el potencial de regular la expresión génica y, por lo tanto, pueden ser utilizados en el desarrollo de terapias y tratamientos para diversas enfermedades. En este sentido, comprender cómo los compuestos presentes en los alimentos pueden influir en la modulación de dianas epigenéticas, proporciona información relevante sobre los mecanismos de acción de los alimentos y podría ayudar a desarrollar intervenciones nutricionales personalizadas para mejorar la salud y prevenir enfermedades. Por esta razón, en el presente trabajo se analiza la importancia de compuestos químicos presentes en los alimentos mediante el uso de herramientas desarrolladas en *Google Colaboratory*, junto con un protocolo de acoplamiento molecular consenso que utiliza tres programas diferentes (AutoDock Vina, MOE y LeDock) para identificar posibles interacciones con dianas epigenéticas de importancia terapéutica (DNMT1, DNMT3A y DNMT3B). Para este propósito, se desarrollaron cuatro cuadernos de código (*notebooks*) de acceso libre en *Google Colaboratory*. El protocolo de acoplamiento molecular consenso se aplicó a 119 compuestos químicos alimentarios utilizando los tres programas mencionados anteriormente, evaluando su interacción con las tres dianas epigenéticas clasificadas como escritores (DNMT1, DNMT3A y DNMT3B). Se generaron huellas digitales de interacción ligando-proteína (PLIFs, por su abreviatura en inglés) para cada diana epigenética. Se identificaron los aminoácidos que presentaban una mayor interacción con los compuestos de la base de datos para cada diana, como Asn 1578 y Pro 1225 para DNMT1, Arg 891, Pro 709, Val 665 y Cys 710 para DNMT3A y Arg 832, Glu 605 y Gly 648 para DNMT3B). Utilizando un enfoque de consenso basado en los puntajes de acoplamiento (energía de afinidad), se estableció un orden relativo para los diferentes ligandos con los tres programas y las tres dianas epigenéticas. Los cinco compuestos con el mejor orden relativo fueron: amentoflavona, galato de epigallocatequina, hesperidina, ácido fólico e isoquercitrina. Se observó que estos cinco compuestos presentaban contactos clave predichos en los mapas de interacción ligando-proteína, los cuales han sido reportados en la literatura como capaces de modular la actividad de DNMT1, DNMT3A y DNMT3B.

# Abstract

Epigenetic target modulators are compounds that possess the potential to regulate gene expression and, consequently, can be employed in the development of therapies and treatments for various diseases. In this regard, comprehending how compounds found in food can influence the modulation of epigenetic targets provides relevant information regarding the mechanisms of action of food and could aid in the development of personalized nutritional interventions to enhance health and prevent diseases.

Therefore, this present study examines the importance of chemical compounds present in food using tools developed in Google Colaboratory, alongside a consensus molecular docking protocol that utilizes three different programs (AutoDock Vina, MOE, and LeDock) to identify potential interactions with therapeutically important epigenetic targets (DNMT1, DNMT3A, and DNMT3B). For this purpose, four freely accessible code notebooks were developed in Google Colaboratory. The consensus molecular docking protocol was applied to 119 food chemical compounds using the aforementioned three programs, assessing their interaction with the three epigenetic targets classified as writers (DNMT1, DNMT3A, and DNMT3B). Ligand-protein interaction fingerprints (PLIFs) were generated for each epigenetic target. The amino acids exhibiting the highest interaction with the compounds from the database were identified for each target, such as Asn 1578 and Pro 1225 for DNMT1, Arg 891, Pro 709, Val 665, and Cys 710 for DNMT3A, and Arg 832, Glu 605, and Gly 648 for DNMT3B.

Using a consensus approach based on docking scores (affinity energy), a relative ranking was established for the different ligands with the three programs and the three epigenetic targets. The top five compounds in terms of relative ranking were: amentoflavone, epigallocatechin gallate, hesperidin, folic acid, and isoquercitrin. It was observed that these five compounds exhibited predicted key contacts in the ligand-protein interaction maps, which have been reported in the literature as capable of modulating the activity of DNMT1, DNMT3A, and DNMT3B.

# Lista de Abreviaturas

ADN, Ácido Desoxirribonucleico;  
AML, Leucemia mieloide aguda;  
API, *application programming interfaces*;  
BRD4, Proteína 4 que contiene bromodominio;  
CTCL, Linfoma cutánea de células T;  
DIFAC, Diseño de fármacos asistido por computadora ;  
DNMT, ADN metiltransferasa;  
ID, Identificador;  
ECR, Clasificación Consenso Exponencial;  
EGCG, Galato de epigallocatequina;  
Ep300, Histona acetiltransferasa p300;  
EZH2, Histona-lisina N-metiltransferasa;  
FDA, Administración de Fármacos y Alimentos (de Estados Unidos);  
HDACs, Histona deacetilasas;  
KDM, Histona lisina desmetilasa;  
LBDD, *Ligand-Based Drug Design*;  
LSD1, Lisina específica desmetilasa 1;  
MDS, Síndrome Mielodisplásico;  
MOE, Molecular Operating Environment;  
MRCC, Carcinoma metastásico de células renales;  
PDB, *Protein Data Bank*;  
PLIFs, Huellas dactilares de interacciones proteína-ligando;  
PTCL, Linfoma Periférico de Células T;  
PNs, Productos naturales;  
ROC, Características operativas recibidas;  
RMSD, Desviación cuadrática media;  
SAH, S-Adenosil L-homocisteína;  
SBDD, *Structure-Based Drug Design*;  
SIRT, Proteína deacetilasa sirtuina dependiente de NAD;

SMILES, *Simplified Molecular Input Line Entry System*;  
TETS, *Ten Eleven Translocation enzymes*.

# Índice general

<b>Resumen</b>	<b>II</b>
<b>Abstract</b>	<b>III</b>
<b>Lista de Abreviaturas</b>	<b>IV</b>
<b>1 Marco teórico</b>	<b>1</b>
1.1 Introducción . . . . .	1
1.1.1 Epigenética . . . . .	1
1.1.2 Mecanismos de dianas epigenéticas . . . . .	2
1.1.3 Enfermedades relacionadas con dianas epigenéticas . . . . .	6
1.1.4 Tratamientos actuales: Inhibidores nucleosídicos vs inhibidores no nucleosídicos . . . . .	8
1.1.5 Diseño de fármacos asistido por computadora (DiFAC) . . . . .	11
1.1.6 Fundamentos del acoplamiento molecular automatizado . . . . .	12
1.1.7 Acoplamiento molecular consenso . . . . .	13
1.1.8 Informática de Química de Alimentos . . . . .	14
<b>2 Objetivo general</b>	<b>16</b>
<b>3 Objetivos particulares</b>	<b>17</b>
<b>4 Hipótesis</b>	<b>18</b>
<b>5 Metodología</b>	<b>19</b>
5.1 Creación de base de datos de compuestos alimenticios . . . . .	19
5.2 Búsqueda de estructuras . . . . .	20
5.3 Compuestos de referencia . . . . .	22
5.4 Preparación de ligandos . . . . .	22
5.5 Preparación de proteínas . . . . .	22



5.6	Acoplamiento molecular . . . . .	23
5.7	Validación . . . . .	23
5.7.1	Reproducción del ligando cristalográfico . . . . .	23
5.7.2	Curvas ROC . . . . .	24
5.7.3	Acoplamiento consenso . . . . .	25
<b>6</b>	<b>Resultados</b>	<b>27</b>
<b>7</b>	<b>Conclusiones</b>	<b>36</b>
<b>8</b>	<b>Perspectivas</b>	<b>37</b>
<b>9</b>	<b>Bibliografía</b>	<b>38</b>
<b>10</b>	<b>Anexo</b>	<b>48</b>

## Índice de figuras

1.1	Número de artículos disponibles en PubMed en los últimos veinte años relacionados con epigenética. . . . .	2
1.2	Metilación de novo vs. metilación. . . . .	3
1.3	Mecanismos epigenéticos de escritores, lectores y borradores. . . . .	5
1.4	Expresión de dianas epigenéticas en el cuerpo humano. . . . .	7
1.5	Estructuras químicas de ejemplos de inhibidores nucleosídicos y no nucleosídicos contra DNMT1, DNMT3A y DNMT3B. . . . .	9
5.1	Metodología implementada en esta tesis. . . . .	19
5.2	Número de estructuras acumuladas de DNMT1(azul), DNMT3A(morado) y DNMT3B (amarillo) en el PDB los últimos veinte años. En el PDB hasta el 2023 se encuentran depositadas un total de 21, 19 y 27 estructuras respectivamente. . . . .	21
5.3	Ecuación para el cálculo de eficiencia del ligando . . . . .	24
5.4	Ecuación de la clasificación consenso exponencial. . . . .	25
5.5	Ecuación del promedio de la clasificación por macromolécula generando una clasificación final. . . . .	26

6.1	En color amarillo se muestran los contactos polares amarillo con los aminoácidos, en color rojo se denota el ligando co-cristalizado (SAH), en naranja la primer pose de acoplamiento en Autodock Vina, en color verde la primer pose de acoplamiento para MOE y en azul la primer pose de acoplamiento de Ledock. (a) DNMT1, (b) DNMT3A (c) DNMT3B. . . . .	27
6.2	Curvas ROC para DNMT1 para los tres softwares. El color azul corresponde al programa de acoplamiento MOE, naranja LeDock y verde AutoDock Vina. En (a) se muestra la curva ROC contemplando la puntuación de cada programa (energía de afinidad). En (b) se muestra la curva ROC contemplando la eficiencia del ligando. . . . .	29
6.3	Clasificación consenso de 119 compuestos químicos alimentarios contra DNMT1, DNMT3A y DNMT3B. . . . .	29
6.4	PLIFs consenso de acoplamiento molecular de 119 ligandos con los tres programas contra (a) DNMT1 (b) DNMT3A (c) DNMT3B. . . . .	30
6.5	Diagramas 2D de (a) Amentoflavona (b) Galato de epigallocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT1. . . . .	31
6.6	Diagramas de interacciones 2D de (a) Amentoflavona (b) Galato de epigallocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT3A. . . . .	32
6.7	Diagramas de interacciones 2D de (a) Amentoflavona (b) Galato de epigallocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT3B. . . . .	33
6.8	Cinco compuestos principales con la mejor clasificación relativa. Los cinco compuestos acoplados para DNMT1, DNMT3A y DNMT3B, (1) Amentoflavona, (3) Hesperidina e (5) Isoquercitrina son compuestos clasificados estructuralmente como flavonoides, (2) galato de epigallocatequina clasificado estructuralmente como catequina y (4) ácido fólico, clasificado estructuralmente como folato. . . . .	34

## Índice de tablas

1.1	Fármacos dirigidos contra blancos epigenéticos. . . . .	10
1.2	Algoritmos de búsqueda de los programas de acoplamiento molecular utilizados. . . . .	13
5.1	Estructuras utilizadas para el acoplamiento molecular obtenidas del PDB. . . . .	21

5.2	Coordenadas para acoplamiento molecular de DNMT1, DNMT3A y DNMT3B para cada programa de acoplamiento molecular. . . . .	24
6.1	Valores de RMSD de la validación de ligandos co-cristalizados para DNMT1, DNMT3A y DNMT3B. . . . .	28
10.1	Clasificación final. . . . .	49

# Capítulo 1

## Marco teórico

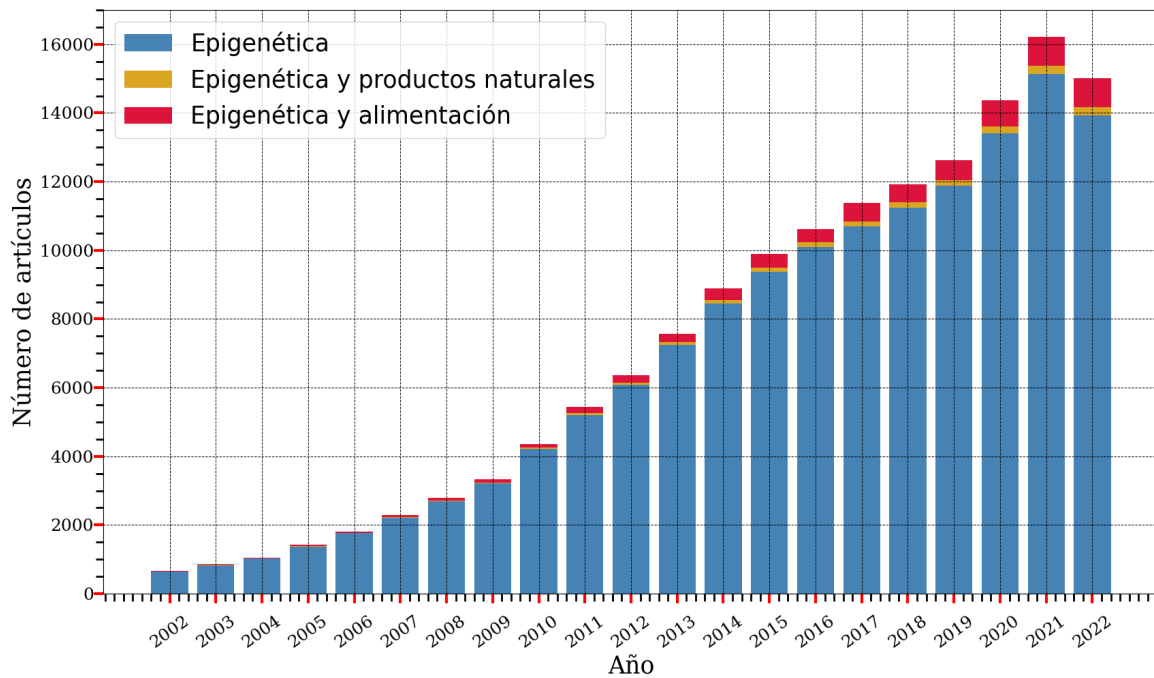
### 1.1. Introducción

#### 1.1.1. Epigenética

El término de epigenética fue introducido por primera vez por el biólogo Conrad Hal Waddington en el año 1942, en la cual resumía una nueva rama de la biología enfocada en analizar y comprender los vínculos entre la expresión génica y proteica. Esta definición ha cambiado a lo largo del tiempo como se elabora a continuación. Los genes que están incluidos en el ADN contienen la información y los planes moleculares que definen el conjunto de características bioquímicas de un organismo vivo, mientras que las proteínas y enzimas ejecutan dichos planes. Cuando investigadores notaron cambios en el fenotipo (carácter visible atribuido a uno o más genes) de algunas especies, los cuales no podían ser explicados exclusivamente a partir de la información genética, se abrió a la idea de la existencia de otro nivel de regulación y programación en el desarrollo de un organismo que va más allá de la información genética. Investigaciones posteriores demostraron que modificaciones epigenéticas (como la acetilación de las histonas o la metilación del ADN) conducen a la inhibición o sobreexpresión de un gen, lo que a su vez produce cambios fenotípicos y, en ocasiones, desajustes del metabolismo celular que afecta al funcionamiento de determinados órganos. En la actualidad la epigenética se puede entender como aquellas modificaciones por encima y más allá de la secuencia de ADN, las cuales, a diferencia de las mutaciones, no son atribuibles a los cambios ocurridos dentro de la secuencia de ADN (Peixoto et al., 2020).

En los últimos veinte años, se ha incrementado el número de artículos relacionados con la epigenética (**Figura 1.1**). De igual forma, ha aumentado el número de artículos científicos que hacen hincapié en la importancia de los productos naturales como estructuras químicas base para la optimización de nuevos compuestos con actividad epigenética. Debido a que los

cambios epigenéticos pueden ser influenciados por factores ambientales, incluyendo la dieta, nuevos campos de investigación como la nutrigenómica y la epigenética de los alimentos han emergido (Chen et al., 2011; Nasir et al., 2020). Sin embargo, se necesita más investigación para comprender los mecanismos moleculares que subyacen de dicha interacción (Stevens et al., 2018).



**Figura 1.1:** Número de artículos disponibles en PubMed en los últimos veinte años relacionados con epigenética.

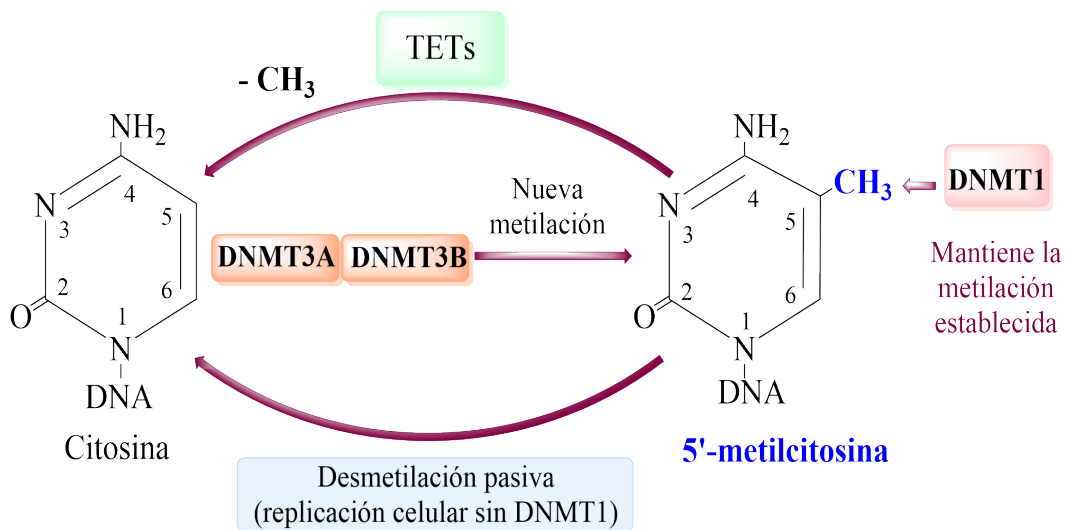
Las palabras claves de búsqueda corresponden a epigenética (azul), epigenética y productos naturales (naranja) y epigenética y alimentos (rojo). De acuerdo con la base de datos de PubMed, en el 2002 había un total de 635 artículos publicados relacionados con la epigenética, 3 con epigenética y productos naturales y 17 con epigenética y alimentos. En el 2022, estas cifras incrementaron a 13933, 228 y 849 artículos, respectivamente.

### 1.1.2. Mecanismos de dianas epigenéticas

Existen tres miembros de familias de dianas epigenéticas (escritores, lectores y borradores). Los escritores introducen distintas modificaciones químicas en el ADN e histonas. Los lectores se encargan de reconocer e interpretar las modificaciones realizadas por los escritores a través de dominios especializados y los borradores se encargan de eliminar las marcas epigenéticas agregadas por los escritores (Biswas Rao, 2018).

### ■ Escritores

DNMT1, DNMT3a y DNMT3b corresponden a dianas epigenéticas clasificadas como escritores. La DNMT1 actúa sobre sitios CpG hemimetilado debido a que la citosina es altamente propensa a sufrir metilación y se cree que en gran parte es responsable de mantener patrones de metilación preexistentes después de la replicación del ADN (**Figura 1.2**). Por otra parte, DNMT3a y DNMT3b tienen funciones importantes en la modificación de regiones genómicas no modificadas (metilación de novo) (Stankevičius et al., 2022; Ren et al., 2018).



**Figura 1.2:** Metilación de novo vs. metilación.

### ■ Lectores

Dentro de esta clasificación se incluyen a los bromodominios (BRDs), cuya función es el reconocimiento de grupos funcionales acetilo en las cadenas laterales de histona. Se destaca a BRD4, la cual puede identificar varias marcas de lisina acetilada en las histonas H3 y H4 y tiene un papel fisiológico diverso.

### ■ Borradores

La desmetilación del ADN en el genoma de los mamíferos se produce por dos procesos diferentes. Primero, se produce una eliminación pasiva de citosina metilada durante rondas sucesivas de replicación del ADN debido a la ausencia de la enzima DNMT1. Posteriormente, tiene lugar una eliminación activa o enzimática del grupo metilo de la 5-metilcitosina (5mC) del ADN. Dentro de esta clasificación, se destaca la importancia de varias enzimas como las histonas desacetilasas (HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, HDAC7, HDAC8, HDAC9, HDAC10, HDAC11), LSD1, SIRT1, SIRT3, KDM1B y KDM6A. Las HDACs son responsables de la desacetilación de las proteínas de histonas,

eliminando el grupo -aminoacetilo de los residuos de lisina en las histonas. Esto resulta en una organización compacta de la cromatina, lo cual tiene un efecto represivo sobre la transcripción génica.. HDAC1 desacetila residuos de lisina en las histonas centrales (H2A, H2B, H3 y H4), HDAC2 desacetila los residuos de lisina en las histonas centrales. Por otro lado, HDAC3 funciona como parte del complejo NCoRo SMRT y su capacidad para funcionar se ve afectada por la pérdida de unión con estos correpresores. HDAC8 es única entre las enzimas de clase I, ya que su actividad desacetilasa no se representa como parte del complejo represivo. Las HDAC de clase IIa (HDAC4, 5,7 y 9) contienen una región N-terminal larga además del dominio desacetilasa. Esta región N-terminal alberga un sitio de unión para el factor potenciador de miocitos 2 y las chaperonas 14-3-3, lo que hace que estas enzimas respondan a varias vías de señalización a través de la fosforilación por quinasas. HDA6 pertenece a la clase IIb, contiene dos dominios deacetilasa y un dominio de represión SE14. HDAC10 contiene un solo dominio de desacetilasa en la región N-terminal, mientras que el dominio C-terminal contiene una región rica en leucina. HDAC11 es el único miembro de HDAC de clase IV y tiene similitud de secuencia con HDAC de clase I y II. Las HDAC de clase III se conocen como proteínas Sir2 (regulador de información silenciosa 2) o sirtuinas (SIRT) que son desacetilasas dependientes de NAD, que constan de siete miembros de la familia (SIRT1-7). SIRT1 se localiza predominantemente en el núcleo y desacetila H1K16 y H4K26. Además de las proteínas histonas, SIRT1 desacetila muchas otras proteínas y se desplaza entre el núcleo y el citoplasma (“Drug Discovery for Epigenetics Targets,” 2022).

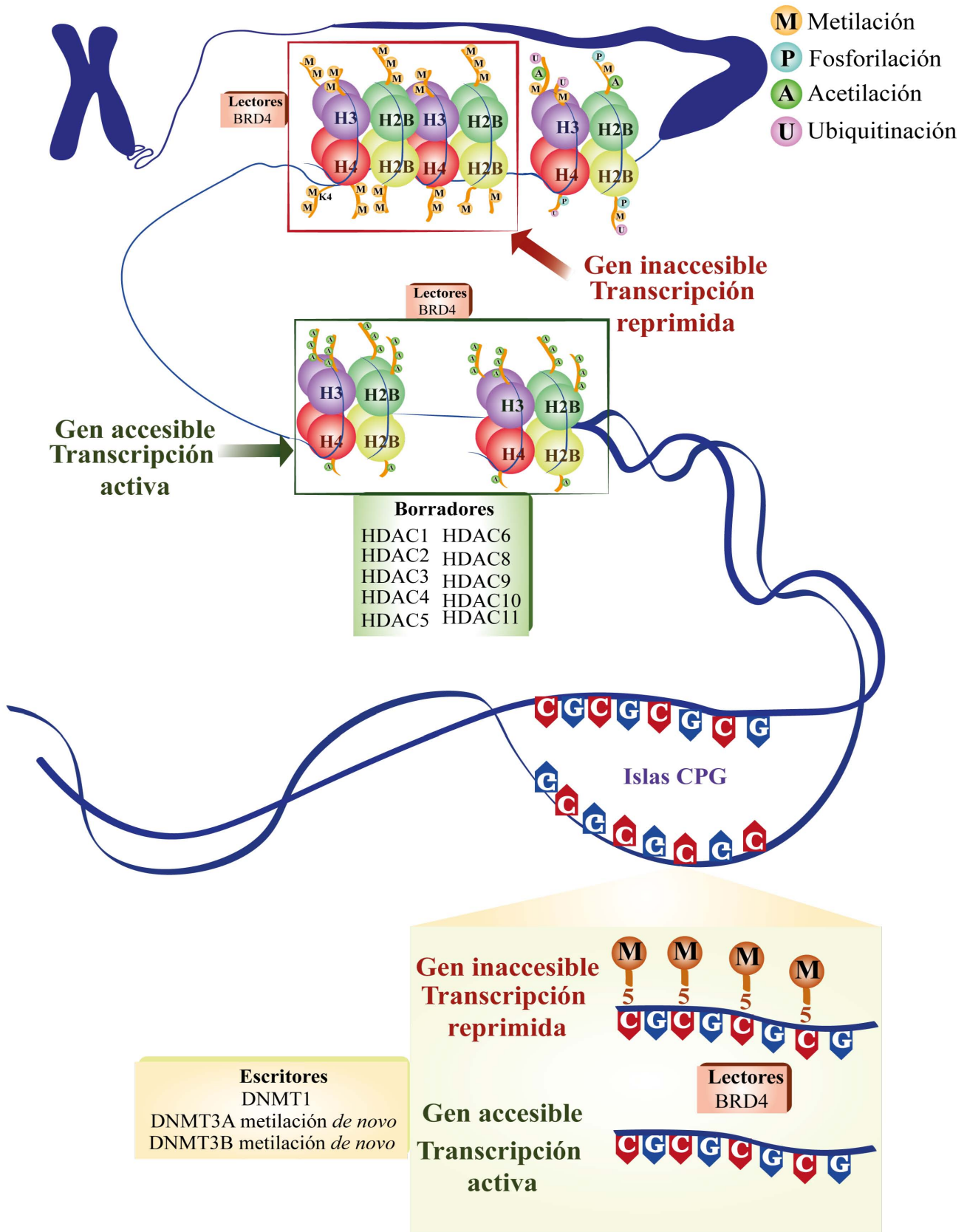


Figura 1.3: Mecanismos epigenéticos de escritores, lectores y borradores.



### 1.1.3. Enfermedades relacionadas con dianas epigenéticas

Las dianas epigenéticas se encuentran distribuidas ampliamente en distintos tejidos como se muestra en la **Figura 1.4**. Esta figura, hecha a partir de la base de datos de Pharos, proporciona una visión general de dónde se expresa probablemente la proteína, además, permite ver qué fuentes de evidencia apoyan la expresión en un tejido determinado (Santos et al., 2015). Como se observa, entre las dianas epigenéticas clasificadas como escritores, DNMT1 es la proteína que se expresa en un mayor número de tejidos (837), destacando el hígado, cerebro, testículos y médula ósea. Esto está asociado con enfermedades neurodegenerativas y el cáncer de testículos. Por otro lado, DNMT3A se expresa en un total de 747 tejidos, destacando a los testículos, pulmones y el cerebro. DNMT3B, por su parte, tiene un total de 754 tejidos de expresión, siendo el cerebro, testículos, el intestino delgado y la placenta los tejidos con un alto nivel de confianza de expresión. Entre las enfermedades que se han reportado con un enfoque epigenético, el cáncer fue una de las primeras enfermedades en ser investigadas. En pacientes con distintos tipos de cáncer se han encontrado patrones de metilación alterados, principalmente la hipermetilación del ADN. A través de la observación de patrones de hipermetilación se estableció la estrategia de identificar inhibidores competitivos para revertir los patrones de hipermetilación. Sin embargo, aún no está totalmente dilucidado el mecanismo de acción (Aldrich et al., 2022).

Otras enfermedades asociadas con modificaciones epigenéticas son: diabetes mellitus tipo I y II (Kaimala et al., 2022), Alzheimer (Coppedè, 2012; Kaimala et al., 2022), neurodegeneración (Chlamydas et al., 2021; Ghosh Saadat, 2021; Mahmoud, 2022), depresión (Begum et al., 2022), obesidad (Chlamydas et al., 2021; Mahmoud, 2022), Covid-19 (Chlamydas et al., 2021), entre otras. Debido a la distribución diversa en tejidos de distintas dianas epigenéticas, hay un campo de oportunidad significativo para la búsqueda de compuestos capaces de modular la actividad de las dianas epigenéticas y a través de esta estrategia prevenir y/o tratar distintas enfermedades.

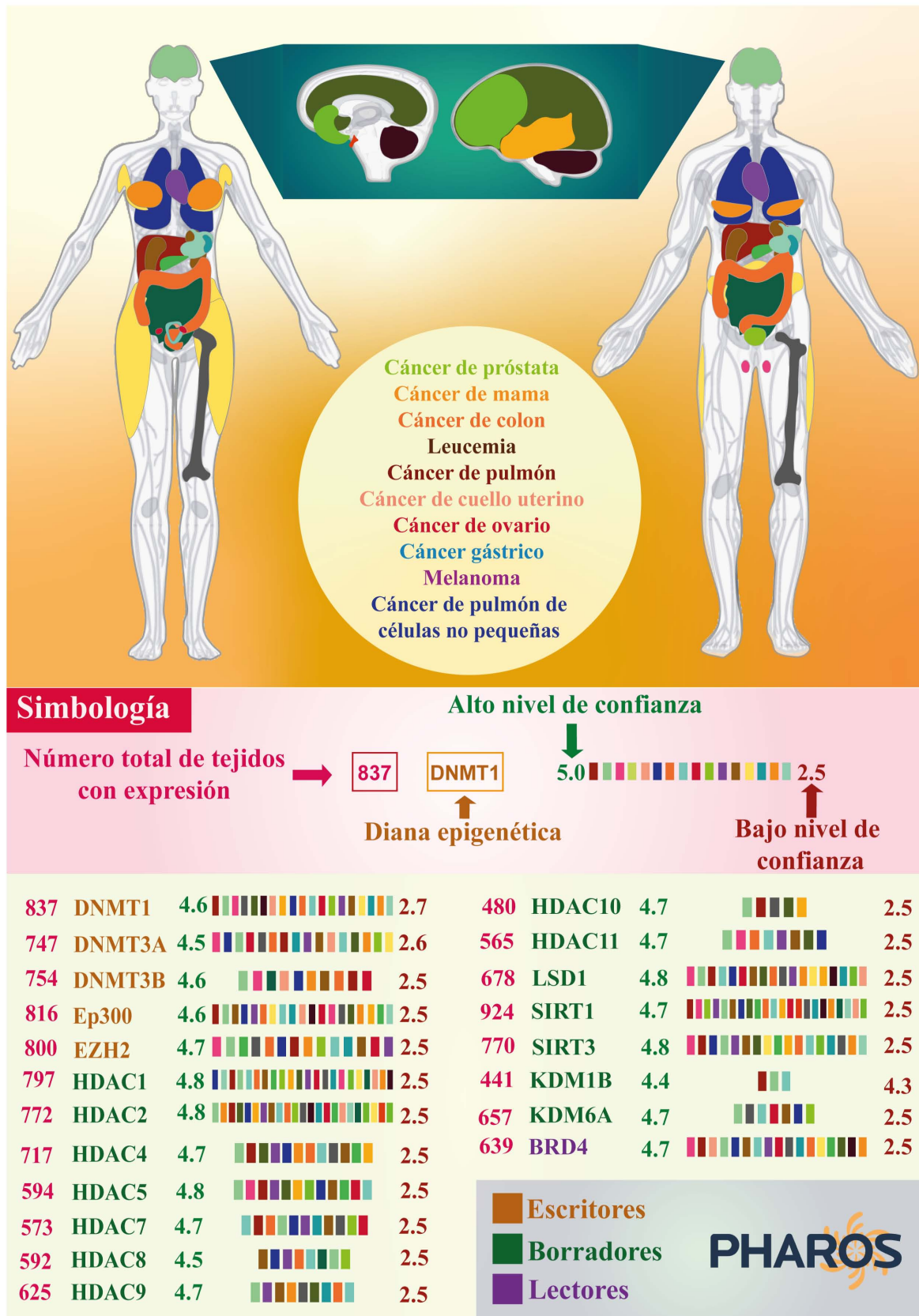


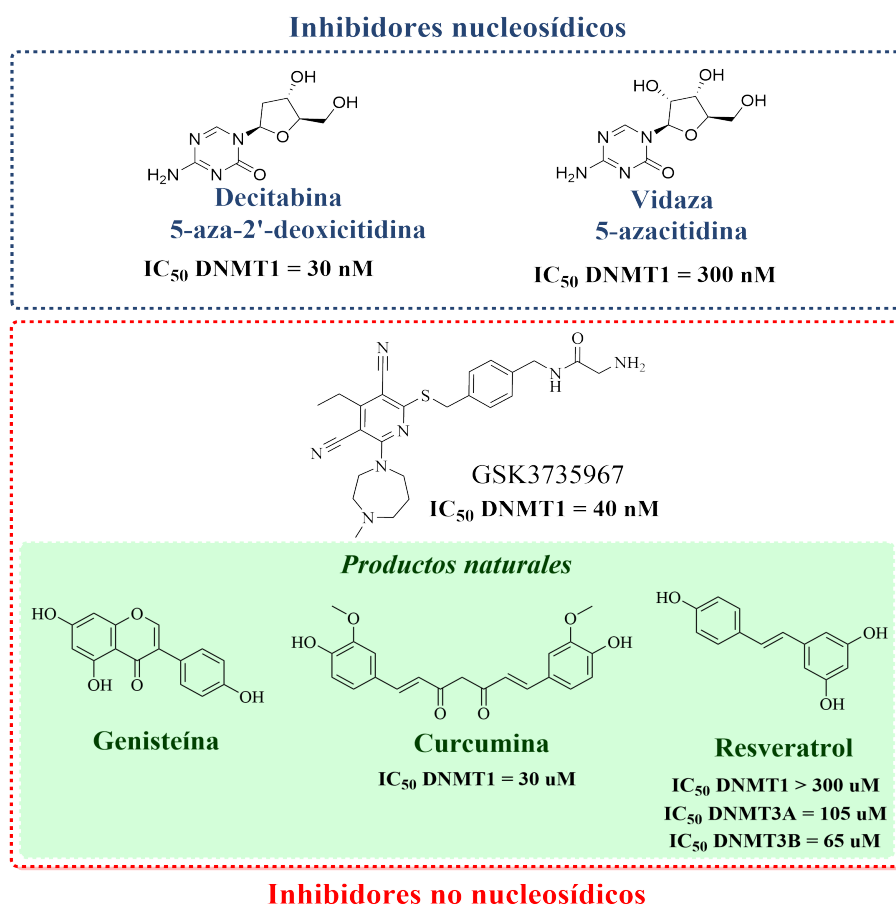
Figura 1.4: Expresión de dianas epigenéticas en el cuerpo humano.

El anatograma resume los datos de expresión recopilados de la base de datos de Pharos para distintas dianas epigenéticas, clasificadas como escritores, lectores y borradores (Sheils et al., 2021). En la base de datos se utilizaron los valores de la columna del laboratorio Jensen (Jensen Lab), donde los valores de 5.0 representan la probabilidad más alta (100%) de encontrar la proteína en ese tejido, y un punto de corte de 2.5 indica la probabilidad más baja (50%) de encontrar la proteína en el tejido. El color rosa (números del lado izquierdo) indica el número total de tejidos en los cuales se expresan las distintas dianas epigenéticas, lo cual se correlaciona con los 10 tipos de cáncer más estudiados desde un enfoque epigenético (cáncer de próstata, cáncer de mama, cáncer de colon, leucemia, cáncer de pulmón, cáncer de cuello uterino, cáncer de ovario, cáncer gástrico, melanoma y cáncer de pulmón de células no pequeñas) (Furtado et al., 2019). Además, se utilizan colores para diferenciar las dianas genéticas, siendo el color naranja se muestran las dianas genéticas clasificadas como escritores, en verde los lectores y en morado los borradores.

#### 1.1.4. Tratamientos actuales: Inhibidores nucleosídicos vs inhibidores no nucleosídicos

Los inhibidores nucleosídicos de las ADN metiltransferasas se derivan generalmente de la citosina y se unen de manera covalente e irreversible al ADN, lo cual puede llegar a generar efectos adversos. Los inhibidores nucleosídicos (5-azacitidina, Vidaza) y (decitabina, 5-aza-2'-deoxicitidina, Dacogen) son los dos fármacos que inhiben las tres clases de ADN metiltransferasa (DNMT1, DNMT3A y DNMT3B) y que han sido aprobados para el tratamiento del síndrome mielodisplásico (**Figura 1.5**) (Conery et al., 2022; Buocikova et al., 2022). Sin embargo, la unión covalente e irreversible de estos inhibidores nucleosídicos representa una de desventajas significativa, lo cual ha motivado la búsqueda de inhibidores no nucleosídicos (Horton et al., 2022). Los inhibidores no nucleosídicos ofrecen la ventaja de unirse de manera reversible al ADN (Cao et al., 2021), lo que permite la modulación de la actividad enzimática. El mecanismo catalítico de DNMT1 propone un ataque de la cisteína 1266 a la citosina (Cyt) C6 transfiriendo un grupo metilo de la S-adenosil-L-metionina (SAM) a la citosina C5, y la abstracción del protón del carbono 5 para dar lugar a la forma CpG metilada del ADN. El mecanismo de metilación del ADN involucra a tres factores principales a considerar: el ADN, la ADN metiltransferasa (DNMT) y el cofactor S-adenosil-L-metionina (SAM). Las DNMT escanean el ADN, reconocen sitios CpG y lanzan 2-desoxicitidina al bolsillo catalítico de la enzima. Un residuo de cisteína (C1226 en DNMT1, C711 en DNMT3A y C652 en DNMT3B en humano) se une a la posición 6 de la citosina, luego se produce la transferencia del grupo metilo del SAM en la posición C5 de la citosina, la cisteína se libera mediante eliminación y la enzima está lista para iniciar un nuevo ciclo catalítico. Las DNMT pueden inhibirse dirigiéndose a diferentes partes del bolsillo

catalítico (el sitio de unión de desoxicitidina, el sitio de unión de SAM o ambos), dirigiéndose a sitios alostéricos o interfiriendo con su unión al ADN. Cuando el ADN es el objetivo, por ejemplo, de los aglutinantes de ADN, es importante diseñar compuestos específicos que no inhiban otras enzimas que actúan sobre el ADN. Una posibilidad es utilizar ligandos de ADN específicos para sitios CpG. Los azanucleósidos, una vez fosforilados en las células e incorporados al ADN, reemplazan a la citosina en la bolsa catalítica de la enzima y actúan como sustratos suicidas de las DNMT. En el ADN, en los sitios CpG, 5azadC sale de la doble hélice hacia el bolsillo catalítico, la cisteína catalítica está unida covalentemente a la posición C6, pero la eliminación, esencial para liberar la enzima, no puede ocurrir después de la transferencia del grupo metilo debido a la presencia de un átomo de nitrógeno en la posición C5. De este modo, la enzima queda atrapada irreversiblemente en el ADN y el proteasoma la degrada aún más, lo que induce la desmetilación del ADN (Lopez et al., 2022). Por ejemplo, la curcumina, genisteína y resveratrol son productos naturales contenidos en los alimentos que pertenecen a la clase de inhibidores no nucleosídicos. En la **Tabla 1.1** se resumen algunos de los fármacos dirigidos hacia diversas dianas epigenéticas clasificadas como escritores, lectores y borradores.



**Figura 1.5:** Estructuras químicas de ejemplos de inhibidores nucleosídicos y no nucleosídicos contra DNMT1, DNMT3A y DNMT3B.

Tabla 1.1: Fármacos dirigidos contra blancos epigenéticos.

Fármacos dirigidos a escritores epigenéticos				
Inhibidores de ADN metiltransferasa (DNMT)				
Categoría	Compuestos	Etapas de desarrollo	Indicaciones o tipo de cáncer	Referencia
Análogo de nucleósido	Azacitidina	Aprobado	SMD	(Derissen et al., 2013)
	Decitabina	Aprobado	SMD	(Derissen et al., 2013)
	SGI-110	Clínico	SMD, LMA	NCT01261312
Análogo de no nucleósido	Nanaomicina	Pre clínico	Colon y pulmones	(Kuck et al., 2010)
	A	Preclínico	Colon	(Datta et al., 2009)
	SGI-1027	Preclínico	Colon	(Datta et al., 2009)
	MG98	Clínico	MRCC	(Amato et al., 2012)
Inhibidores de la histona lisina metiltransferasa (HKMTi)				
EZH2	EI1	Preclínico	Linfoma de células B	(Amato et al., 2012; Qi et al., 2012)
	CPI-1205	Clínico	Linfoma de células B	NCT02395601
	EPZ6438	Clínico	Linfoma de células B	(Kuntz et al., 2016)
Inhibidores de histona acetiltransferasa (HATi)				
Ep300	C646	Preclínico	Gástrico	(Y.-M. Wang et al., 2017)
Fármacos dirigidos a lectores epigenéticos				
BRD4	Fedratinib	Aprobado	Mielofibrosis asociada a neoplasias mieloproliferativas	(Talpaz Kiladjian, 2021)
Fármacos dirigidos a borradores epigenéticos				
Inhibidores de la histona lisina desmetilasa (HKDMi)				
Inhibidores de LSD1	Análogo de tranilcipromina	Clínico	Leucemia mieloide aguda	(Prusevich et al., 2014)
	Bizine	Preclínico	Pulmón, próstata	(Prusevich et al., 2014)
	PG11144	Preclínico	Seno	(Zhu et al., 2012)
	Namolina	Preclínico	Próstata	(Willmann et al., 2012)
Inhibidores de la histona desacetilasa (HDACi)				
Derivados del ácido hidroxámico	Vorinostat	Aprobado	CTCL	(Mann et al., 2007)
	Belinostat	Aprobado	PTCL	(Moskowitz, Horwitz, 2017)
	Panobinostat	Aprobado	Mieloma múltiple	(Laubach et al., 2015)
	Pracinostat	Aprobado	AML	(Van Veggel et al., 2018)
	Quisinostat	Clínico	Cáncer de ovario	NCT02948075
Tetrapeptido cíclico	Romidepsina	Aprobado	CTCL	NCT00106431
Benzamida	Entinostat	Clínico	NE tumor	NCT03211988
	Mocetinostat	Clínico	Linfoma	NCT02282358
Ácidos grasos de cadena corta	Ácido valproico	Clínico	Cáncer avanzado	NCT00496444
Inhibidor de la sirtuina	Sirtinol	Preclínico	Seno	(J. Wang et al., 2012)
	Nicotinamida	Preclínico	Seno	(T. Wang et al., 2013)

### 1.1.5. Diseño de fármacos asistido por computadora (DiFAC)

El diseño de fármacos asistido por computadora presenta ventajas significativas respecto a un diseño de fármacos tradicional en cuanto a la reducción de costos y tiempo de los ensayos experimentales. Esto se logra mediante la evaluación de grandes conjuntos de datos de moléculas a través de estudios de cribado virtual (Engel Gasteiger, 2018; Paul, 2017). Los enfoques basados en el ligando y la estructura son herramientas ampliamente utilizadas tanto en la industria farmacéutica como en la academia para el descubrimiento de fármacos. Además, los avances en los métodos espectroscópicos biomoleculares como la cristalografía de rayos X y la resonancia magnética nuclear han permitido el depósito de más de 10,000 estructuras de proteínas en la base de datos *Protein Data Bank* (PDB). El acceso a una base de datos tan amplia como el PDB proporciona información relevante de las distintas proteínas. Estos datos son de vital importancia para los métodos de diseño de fármacos basados en la estructura (SBDD, por sus siglas del inglés *Structure-Based Drug Design*).

#### Métodos basados en la estructura

Los métodos SBDD aprovechan la información estructural tridimensional obtenida de los métodos biológicos. La disponibilidad de estructuras macromoleculares tridimensionales permite examinar de manera minuciosa la topología de unión, la presencia de hendiduras, cavidades y bolsas secundarias (Maia et al., 2020). Los métodos SBDD permiten diseñar ligandos que cumplan con características necesarias para la modulación del receptor diana. Entre las estrategias SBDD destaca el cribado virtual o filtrado computacional basado en la estructura y la dinámica molecular (Ferreira et al., 2015). El cribado virtual se puede realizar con uno o varios programas (análisis consenso). Cabe destacar que no necesariamente el resultado obtenido como la pose de menor energía es la que biológicamente presenta las interacciones claves para que pueda tener actividad contra una o múltiples dianas moleculares. Por lo tanto, es necesario validar y analizar las estructuras y sus interacciones clave basándose en el conocimiento derivado de artículos científicos para predecir la actividad contra una o varias dianas biológicas (Gentile et al., 2023).

#### Métodos basados en ligando

Por otro lado, los métodos de diseño de fármacos basado en ligandos (LBDD, por sus siglas en inglés, *Ligand-Based Drug Design*) hacen uso de la diversidad química de bibliotecas bioactivas de moléculas pequeñas, estas bibliotecas representan el espacio químico ocupado por los ligandos que se sabe interactúan con una o distintas proteínas. Algunos ejemplos de métodos LBDD son el cribado virtual basado en ligandos, la similitud molecular, modelado de

relaciones estructura-actividad cuantitativas (QSAR, por sus siglas en inglés) y la generación de farmacóforos (Ye et al., 2022).

En la actualidad, la inteligencia artificial (IA) y los métodos computacionales han permitido hacer más eficiente el proceso de búsqueda de nuevos fármacos para distintas enfermedades como la diabetes mellitus tipo 2 (Huang et al., 2022), enfermedades infecciosas (Diakou et al., 2022), desórdenes mentales y el cáncer. Es por esta razón que resulta crucial desarrollar algoritmos de IA y herramientas computacionales para establecer protocolos validados y robustos para la búsqueda de fármacos (Krishnan et al., 2023).

### 1.1.6. Fundamentos del acoplamiento molecular automatizado

Desde el desarrollo de los primeros algoritmos computacionales en la década de 1980, el acoplamiento molecular se convirtió en una herramienta esencial para la investigación de compuestos bioactivos, incluyendo el modelado molecular de compuestos con actividad conocidos en contexto de su blanco terapéutico (Meng et al., 2011). El acoplamiento molecular también se ha empleado con mucha frecuencia para plantear hipótesis y proponer compuestos con potencial actividad que puedan avanzar a etapas de desarrollo preclínico (Prieto-Martínez et al., 2018). Los algoritmos de acoplamiento molecular ejecutan predicciones cuantitativas de la energía de unión, proporcionando clasificaciones de compuestos acoplados en función de la afinidad de unión de los complejos ligando-receptor (Sethi et al., 2019).

La identificación de las conformaciones de unión más probables requiere de dos pasos:

1. La exploración de un gran espacio conformacional que representa varios modos de unión potenciales.
2. La predicción precisa de la energía de interacción asociada con cada una de las conformaciones predichas.

Los programas de acoplamiento molecular realizan estos dos pasos a través de un proceso cíclico, en el que la conformación del ligando se evalúa mediante funciones de puntuación específicas. Este proceso, se realiza recursivamente hasta converger en una solución de mínima energía.

En cuanto a la búsqueda conformacional, los parámetros de los ligandos, como los grados de libertad torsional (diédrico), traslacional y rotacional, se modifican incrementalmente. Los algoritmos de búsqueda conformacional realizan esta tarea mediante la aplicación de métodos de búsqueda estocásticos y sistemáticos. En los métodos de búsqueda sistemáticos, el algoritmo sondea el panorama energético del espacio conformacional y después de miles de ciclos de búsqueda y evaluación converge a la solución de energía mínima correspondiente al modo de unión más probable.

Los métodos estocásticos llevan a cabo la búsqueda conformacional modificando aleatoriamente los parámetros estructurales de los ligandos. El algoritmo genera conjuntos de confor-

maciones moleculares y con esto se obtiene una amplia gama del panorama energético. Debido a que el algoritmo promueve una amplia gama del panorama energético, el costo computacional asociado con este procedimiento es una limitación que hay que tener en consideración. Para resolver está limitante en los gastos computacionales, se han desarrollado los algoritmos genéticos, estos algoritmos corresponden a métodos estocásticos mediante la aplicación de conceptos de la teoría de la evolución y la selección natural (Z. Wang et al., 2016).

En un primer paso, el algoritmo codifica todos los parámetros de la estructura inicial en un cromosoma que se representa mediante un vector. A partir de este cromosoma, el algoritmo de búsqueda aleatoria genera una población inicial de cromosomas que cubre una amplia área del panorama energético. Esta población es evaluada y los cromosomas más adaptados (es decir, aquellos con valores de energía más bajos) se seleccionan como plantillas para la generación de la siguiente población. Este procedimiento disminuye la energía promedio del conjunto cromosómico al transmitir características estructurales favorables de una población a otra, reduciendo el espacio conformacional a explorar. El algoritmo genético se ejecuta recursivamente y después de un número razonable de ciclos de búsqueda y evaluación conformacional, converge a una conformación (cromosoma) correspondiente al mínimo de energía global (Maia et al., 2020).

**Tabla 1.2:** Algoritmos de búsqueda de los programas de acoplamiento molecular utilizados.

Software	Algoritmo de búsqueda	Referencia
AutoDock Vina	Híbrido	(Trott Olson, 2010)
MOE	Híbrido	Molecular Operating Environment (MOE) (2014) Montreal, Quebec, Canada: Chemical Computing Group Inc. <a href="http://www.chemcomp.com">www.chemcomp.com</a>
LeDock	Alineamiento simulado	Xu et al., 2021

### 1.1.7. Acoplamiento molecular consenso

La clasificación consenso exponencial (ECR, por sus siglas del inglés, *exponential consensus ranking*) es una estrategia para combinar distintas funciones de puntuación de distintos programas de acoplamiento molecular contra múltiples dianas moleculares usando una distribución exponencial para cada rango individual (Palacio-Rodríguez et al., 2019; Triches et al., 2022). En distintos estudios, el ECR se ha utilizado para establecer posibles candidatos para el tratamiento de enfermedades, por ejemplo el COVID-19 (Cavasotto Di Filippo, 2021). En el trabajo de Triches (2022) se hace una propuesta de protocolo consenso, utilizando tres progra-



mas diferentes de acoplamiento molecular de múltiples ligandos contra diferentes proteínas. El protocolo consenso cuenta con ciertas ventajas y limitaciones, algunas de ellas son:

**Limitaciones:**

- Se asume que la pose de menor energía es la mejor.

**Ventajas:**

- Ayuda a priorizar compuestos a través de su actividad contra múltiples dianas.

- Permite generar un ranking con distintos programas, un gran número de compuestos y múltiples dianas.

- Permite establecer métricas que tienen un sentido biológico más relevante en comparación con un acoplamiento consensus tradicional.

- Evita problemas de sesgo resultantes de escalas de las escalas de puntaje, compensación o unidades.

Requiere de menores recursos computacionales en comparación con otras estrategias consenso.

### 1.1.8. Informática de Química de Alimentos

La Informática de la Química de Alimentos es una subdisciplina que se puede definir como la aplicación de métodos informáticos para resolver problemas de Química de Alimentos. Siguiendo la discusión de las aplicaciones de la Quimioinformática, en general, la Informática de Alimentos tiene múltiples aplicaciones como la sistematización de la información en bases de datos de compuestos químicos alimentarios (un ejemplo notable es la base de datos FoodDB); el entendimiento y predicción de relaciones estructura-propiedad (por ejemplo su sabor, olor, propiedades biológicas.) Actualmente ya se habla del diseño racional de odorantes (e.g, contenidos en alimentos o que se pueden utilizar como aditivos alimentarios (Naveja et al., 2018). El campo de la Informática de los Alimentos se ha estado usando por varios años, pero alrededor de 2014 comenzó un esfuerzo por formalizar esta disciplina que poco a poco se está consolidando. En la actualidad, se han realizado distintos estudios en los que se evalúan los componentes presentes en los alimentos como estrategia para ayudar a prevenir o tratar algunas enfermedades, por ejemplo, el cáncer gástrico (Gigek et al., 2019). También se ha observado la relación que existe entre la epigenética, la alimentación y las enfermedades (Lévesque et al., 2019; Sapienza Issa, 2016). En otros casos, se ha estudiado la terapia combinatoria entre PNs para el tratamiento o prevención del cáncer (Arora et al., 2019). Por otro lado, algunos investigadores han estudiado el efecto de las dietas, por ejemplo, la dieta mediterránea se evaluó como probable estrategia para la prevención de enfermedades relacionadas con la epigenética; en este estudio se destaca la importancia del diseño de estrategias innovadoras dirigidas a la prevención, el manejo, el pronóstico y el tratamiento de enfermedades a través de enfoques personalizados (incluida la

nutrición personalizada) basados en la epigenética (Ramos-Lopez et al., 2021); esto sugiere que la informática de los alimentos es un área no explorada en forma exhaustiva, por lo tanto, el estudio de los componentes presentes en diversos alimentos ayudará a comprender su potencial influencia en la epigenética, así como también poner en evidencia que una alimentación balanceada puede ayudar a prevenir distintas enfermedades. Es por ello que la búsqueda e identificación de compuestos naturales presentes en los alimentos junto con métodos computacionales puede ayudar a identificar alimentos para prevenir ciertas enfermedades y/o generar moléculas similares, las cuales sean potenciales para el tratamiento de enfermedades.

## Capítulo 2

### Objetivo general

Identificar moduladores de las dianas epigenéticas DNMT1, DNMT3A y DNMT3B a partir de compuestos químicos alimentarios para el tratamiento y/o prevención de múltiples enfermedades. Esto se realizará mediante un protocolo de acoplamiento molecular consenso de tres programas diferentes (AutoDock Vina, MOE y LeDock), con los cuales se generará una puntuación de distintas sustancias alimentarias contra múltiples dianas epigenéticas.

# Capítulo 3

## Objetivos particulares

1. Crear una base de datos de compuestos alimentarios con actividad epigenética a partir de información disponible en la literatura científica.
2. Crear Notebooks en Google Colab para la automatización de la búsqueda de estructuras de proteínas en la base de datos Protein Data Bank y de protocolos de cribado virtual.
3. Validar los protocolos de acoplamiento molecular con los ligandos co-cristalizados y usando información de compuestos con datos de bioactividad disponibles en la base de datos de ChEMBL.
4. Generar modelos de acoplamiento molecular de los compuestos alimenticios contra distintas dianas epigenéticas usando los programas AutoDock Vina, MOE y LeDock.
5. Realizar un análisis consenso de los resultados usando la metodología propuesta por Triches 2022

# Capítulo 4

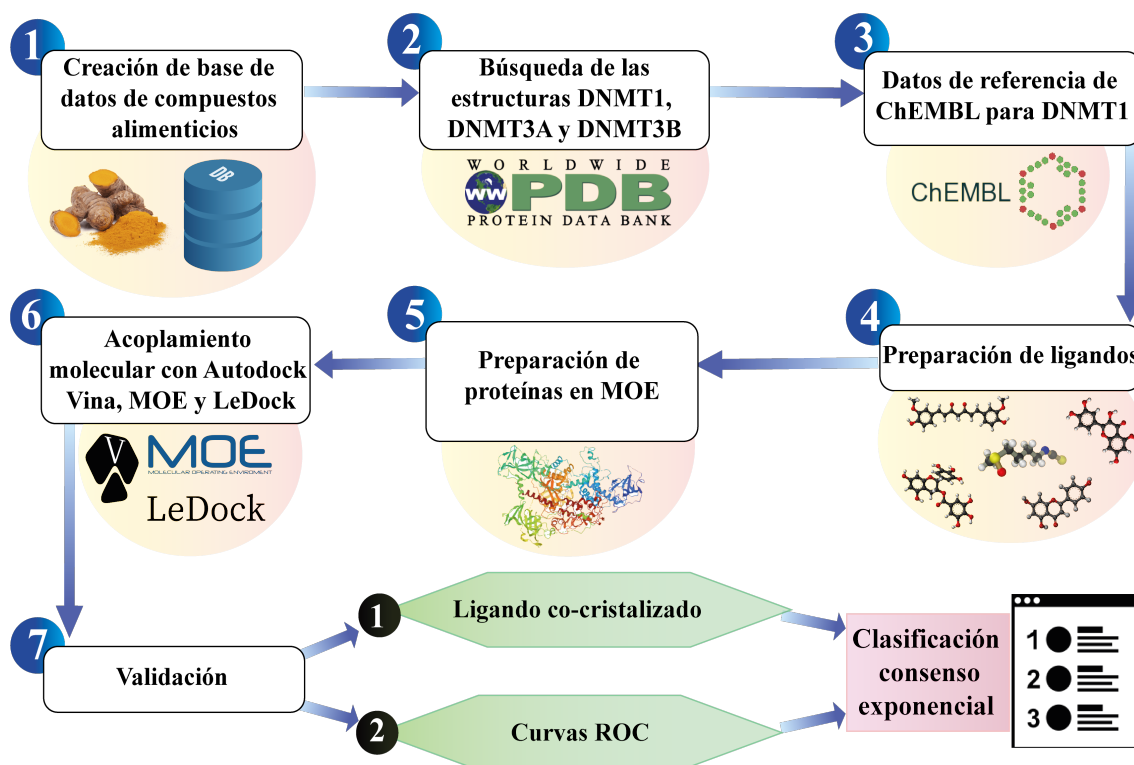
## Hipótesis

A través de un protocolo de acoplamiento molecular consenso se identificarán moduladores no nucleosídicos contenidos en diversos alimentos con actividad farmacológica reportada contra los blancos epigenéticos DNMT1, DNMT3a y DNMT3b. Esto permitirá generar y proponer nuevos líderes potenciales que modulen la actividad de dianas epigenéticas para el tratamiento y/o prevención de múltiples enfermedades.

# Capítulo 5

## Metodología

La metodología empleada en esta tesis se muestra en forma esquemática en la **Figura 5.1**. Cada uno de los pasos de la metodología se detallan a continuación.



*Figura 5.1:* Metodología implementada en esta tesis.

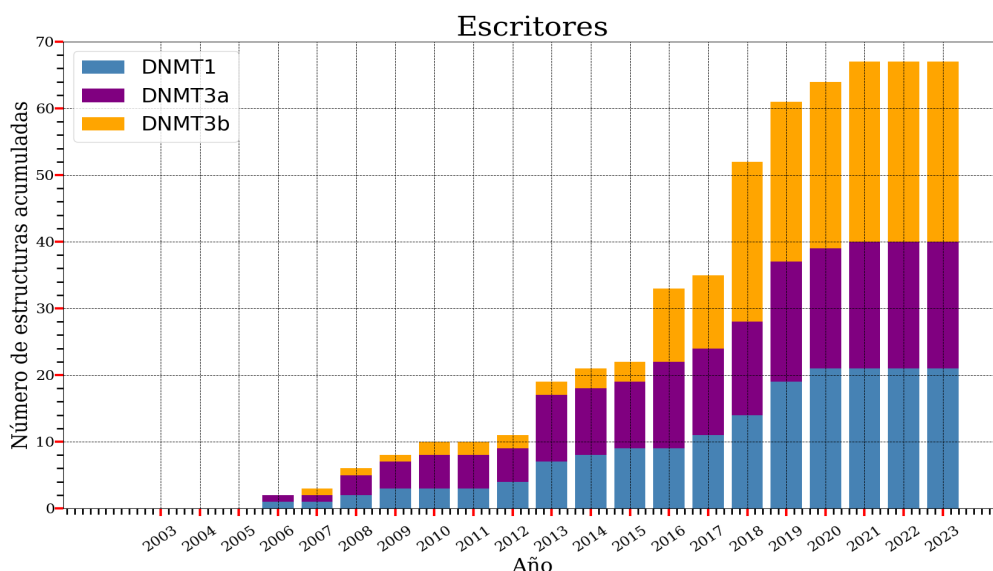
### 5.1. Creación de base de datos de compuestos alimenticios

Para la elaboración de la base de datos se realizó una búsqueda bibliográfica (meta-análisis) de las publicaciones de no más de cinco años (2017-2022) en PubMed y *Web of Science*, con las palabras claves: *epigenetics; food chemical(s); therapeutic application; disease; drug*

*discovery; epigenetic targets; epigenetic therapy; epigenetic mechanisms; epigenetic regulation; epigenetic modifiers; epi-drugs, nutritional epigenetics*. La quimioteca de sustancias químicas alimentarias fue elaborada con base en la información de 44 artículos. A partir de ellos se obtuvieron 119 sustancias alimentarias. La base de datos incluye el nombre del compuesto, notación SMILES, INChi e INChiKey (Heller et al., 2015; O'Boyle, 2012). La fuente principal y contenido fueron obtenidos de la base de datos FooDB (Harrington et al., 2019). La base de datos se puede consultar en el siguiente hipervínculo [base de datos de compuestos químicos alimentarios](#).

## 5.2. Búsqueda de estructuras

Se desarrolló un cuaderno de código en *Google Colaboratory* en el cual se obtuvo y se graficó el número total de proteínas, clasificadas de acuerdo con su herramienta epigenética (escritor, lector o borrador), depositadas en los últimos veinte años en PDB (véase el apartado de github en el anexo). En la **Figura 5.2** se muestran el número total de estructuras de DNMT1, DNMT3A y DNMT3B acumuladas por año desde el 2002 hasta el 2023. La **Tabla 5.1** resume las estructuras tridimensionales obtenidas del PDB de las proteínas utilizadas para realizar el acoplamiento molecular. Para los criterios de búsqueda para las proteínas DNMT1, DNMT3A y DNMT3B se consideró como método experimental la difracción de rayos X, una resolución no mayor a 3 Å y un peso molecular de 100 g/mol.



**Figura 5.2:** Número de estructuras acumuladas de DNMT1(azul), DNMT3A(morado) y DNMT3B (amarillo) en el PDB los últimos veinte años. En el PDB hasta el 2023 se encuentran depositadas un total de 21, 19 y 27 estructuras respectivamente.

**Tabla 5.1:** Estructuras utilizadas para el acoplamiento molecular obtenidas del PDB.

Diana epigenética	PDB ID	Resolución	Ligando co-cristalizado	Reporte de validación	Año
DNMT1	<a href="#">4WXX</a>	2.62 Å	<a href="#">SAH</a>	<a href="#">reporte 4WXX</a>	2014-11-14
DNMT3A	<a href="#">4U7T</a>	2.90 Å	<a href="#">SAH</a>	<a href="#">reporte 4U7T</a>	2022-10-11
DNMT3B	<a href="#">6U8W</a>	2.95 Å	<a href="#">SAH</a>	<a href="#">reporte 6U8W</a>	2019-09-06

### 5.3. Compuestos de referencia

Se diseñó un cuaderno de código en *Google Colaboratory* para obtener compuestos bioactivos para DNMT1 usando la interfaz de programación de aplicaciones (API por sus siglas en inglés, *application programming interfaces*) de la base de datos de ChEMBL. Estos compuestos se estandarizaron y curaron empleando otro cuaderno de código desarrollado dentro del grupo DIFACQUIM [curado de base de datos](#). Como se verá en la sección de validación (**5.7.2 Curvas ROC**), los compuestos obtenidos de ChEMBL sirven como punto de referencia para validar los protocolos de acoplamiento molecular.

### 5.4. Preparación de ligandos

Los 119 compuestos de productos químicos alimenticios con actividad reportada contra dianas epigenéticas que se obtuvieron de la búsqueda bibliográfica del punto **5.1 Creación de base de datos de compuestos alimenticios** se protonaron a un pH de 7.4, y se minimizaron utilizando el campo de fuerza MMFF94 (Halgren, 1999). El campo de fuerza MMFF94 fue desarrollado por Merck, el método proporciona una buena precisión de una gama de moléculas orgánicas y similares a medicamentos. El método incluye parámetros para una amplia gama de tipos de átomos, incluyendo los siguientes elementos orgánicos comunes: C, H, N, O, F, Si, P, S, Cl y Br. También soporta los siguientes iones:  $Fe^{2+}$ ,  $Fe^{3+}$ ,  $F^{-}$ ,  $Cl^{-}$ ,  $Br^{-}$ ,  $Li^{+}$ ,  $Na^{+}$ ,  $K^{+}$ ,  $Zn^{2+}$ ,  $Ca^{2+}$ ,  $Cu^{1+}$ ,  $Cu^{2+}$ ,  $Mg^{2+}$ .

La variante MMFF94 (estática) tiene parámetros de torsión diedro y flexión fuera del plano



ligeramente diferentes para parametrizar ciertos tipos de átomos de N trigonales deslocalizados, como la anilina aromática. Esto proporciona una mejor coincidencia con la geometría molecular promedio en el tiempo en soluciones o estructuras cristalográficas.

## 5.5. Preparación de proteínas

Debido a que las estructuras obtenidas de PDB pueden presentar redundancia, falta de átomos, residuos incompletos y ambigüedad, es importante prepararlas antes de utilizarlas en estudios de acoplamiento molecular. En este sentido, las estructuras DNMT1 (PDB ID: 4WXX), DNMT3A (PDB ID: 4U7T ) y DNMT3B (PDB ID: 6U8W) se prepararon usando la función de QuickPrep, implementada en el programa Molecular Operating Environment (MOE), versión 2022.02. QuickPrep es una interfaz para la optimización interactiva de ligandos en 3D en el sitio activo, proporciona comandos para la preparación de la estructura, el análisis del sitio activo, la propiedad/unión molecular, cálculos de afinidad y modificación y optimización de ligandos en el sitio activo (Vilar et al., 2008).

La conversión de proteínas previamente preparadas de archivos .pdb a .pdbqt se realizaron con el software MGLTools versión 1.5.6 (<https://ccsb.scripps.edu/mgltools/downloads/>).

Para la visualización de los complejos ligando-proteína y obtención de coordenadas necesarias para el acoplamiento en LeDock se utilizó PyMol 2.5.5 ([pymol.org](http://pymol.org)).

Para la obtención de coordenadas necesarias se implementó un plugin en PyMol 2.5.5, GetBox con una extensión de 5 Å ([GetBox-PyMOL-Plugin](#)).

## 5.6. Acoplamiento molecular

Para realizar el acoplamiento molecular se utilizaron tres programas diferentes:

1. AutoDock Vina versión 1.1.2 ([Manual – AutoDock Vina \(scripps.edu\)](#)).
2. MOE versión 2022.02 ([Chemical Computing Group \(CCG\) | Computer-Aided Molecular Design](#))
3. LeDock versión 1.0, la conversión de las estructuras previamente preparadas de .pdb se realizó con el módulo LePro (<http://www.lephar.com/download.htm>).

## 5.7. Validación

### 5.7.1. Reproducción del ligando cristalográfico

La reproducción del modo de unión del ligando co-cristalizado (SAH) se realizó utilizando las coordenadas obtenidas por el programa UCSF Chimera versión 1.16 con ayuda del tablero de comandos utilizando el comando *getcrd sel* centrado en el ligando co-cristalizado para las tres proteínas (DNMT1, DNMT3A y DNMT3B). El validar el modo de unión permitió determinar si el método utilizado fue capaz de reproducir los resultados experimentales previamente obtenidos del PDB.

**Tabla 5.2:** Coordenadas para acoplamiento molecular de DNMT1, DNMT3A y DNMT3B para cada programa de acoplamiento molecular.

Proteína	AutoDock Vina	MOE	LeDock
DNMT1 (PDB ID: 4WXX)	21.223 8.754 -56.784	Centrado en ligando co-cristalizado SAH	13.4 30.6 -3.8 19.4 -66.5 -47.2
DNMT3A (PDB ID: 4U7T )	20.014 61.261 -5.983	Centrado en ligando co-cristalizado SAH	11.8, 28.6 52.1, 72.8 -18.3, 3.0
DNMT3B (PDB ID: 6U8W)	15.986 69.265 8.903	Centrado en ligando co-cristalizado SAH	5.3 26.6 61.3 79.9 -0.1 20.8

### 5.7.2. Curvas ROC

La curva ROC se utiliza para evaluar el rendimiento de los algoritmos de clasificación binaria, es decir, entre dos clases o categorías (para este caso de estudio 1= activo, 0= inactivo). El área bajo la curva ROC (ROC-AUC, por sus siglas en inglés, receiver operating characteristic-area under the curve) nos permite observar la capacidad de predicción de un método; en este caso, explica la capacidad de cada programa de acoplamiento para diferenciar entre un compuesto activo y un compuesto inactivo. Un ROC-AUC superior a 0,5 indica que el método tiene una mejor capacidad predictiva que uno aleatorio. Cuanto mejor sea la capacidad predictiva del

método, más se acercará ROC-AUC a 1. Para minimizar el sesgo que suelen tener los resultados de acoplamiento molecular al no considerar el peso molecular de los compuestos, se graficaron las curvas ROC contemplando la eficiencia del ligando, cuya ecuación es la siguiente:

$$\text{Eficiencia del ligando} = \text{puntuación de acoplamiento} / \text{número de átomos pesados}$$

*Figura 5.3:* Ecuación para el cálculo de eficiencia del ligando

### 5.7.3. Acoplamiento consenso

Para la elaboración del acoplamiento consenso se siguió la metodología sugerida por Triches et al. 2022. Como primer paso se determinó la función de puntuación para cada uno de los programas (AutoDock Vina, MOE y LeDock) con los 119 ligandos contenidos en la base de datos de compuestos químicos presentes en alimentos. Como segundo paso se generó un listado asignando una numeración para cada una de las proteínas (DNMT1, DNMT3A y DNMT3B) y para los 119 ligandos acoplados, siendo el 1 el ligando con menor energía (puntaje con valor más negativo) y el 119 el ligando con mayor energía (puntaje con valor más positivo). En el tercer paso se hace uso de la ecuación de ECR para combinar los resultados de la puntuación por programa (**Figura 5.4**). Y como último paso se realizó un promedio por molécula y por proteína generando la clasificación final (**Figura 5.5**).

$$P(i) = \frac{1}{\sigma} \sum_{j=1}^p \exp\left(-\frac{R_{i,\lambda}^j}{\sigma}\right) = a_{i,\lambda}$$

- " $P(i)$ " corresponde a la posición del ligando del interesado
- Sigma ( $\sigma$ ) tiene un valor fijo de 10
- " $\exp$ " se refiere al número de Euler, igual a 2.718
- Los índices " $\lambda$ ", " $i$ ", " $j$ " indica cualquier macromolécula ( $M_\lambda$ ), cualquier ligando ( $L_i$ ), y cualquier programa ( $P_j$ ), respectivamente. Cada investigador puede definir el número de elementos en cada conjunto. Por ejemplo, en nuestro estudio hay 3 proteínas (DNMT1, DNMT3A y DNMT3B), 119 ligandos y tres programas de acoplamiento.
- El ranking viene dado por " $R_{i,\lambda}^j$ "
- El " $a_{i,\lambda}$ " corresponde a la suma de la puntuación ECR de todas las funciones de puntuación

**Figura 5.4:** Ecuación de la clasificación consenso exponencial.

$$L_i : \frac{M_1(a_{i,1}) + M_2(a_{i,2}) + \dots + M_\lambda(a_{i,\lambda}) + \dots + M_m(a_{i,m})}{m} = P_i$$

" $L_i$ " corresponden al ligando de los interesados;

" $\frac{1}{m}$ " es el número 1 dividido por el número total de macromoléculas ( $m$ );

" $M_\lambda$ " corresponden a cualquier macromolécula;

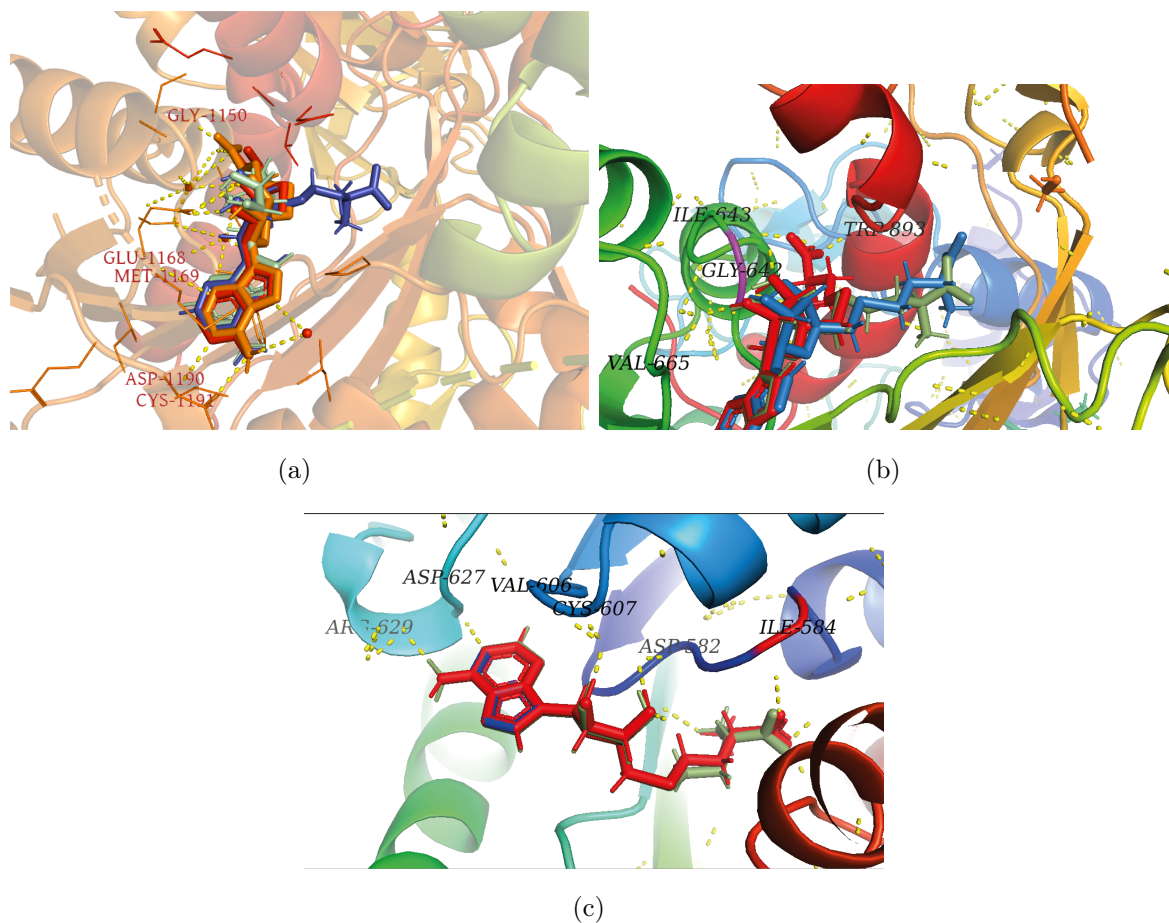
" $a_{i,\lambda}$ " es el resultado obtenido en la ecuación de la figura 5.4;

" $P_i$ " corresponden a la posición final del ligando de interés.

**Figura 5.5:** Ecuación del promedio de la clasificación por macromolécula generando una clasificación final.

# Capítulo 6

## Resultados

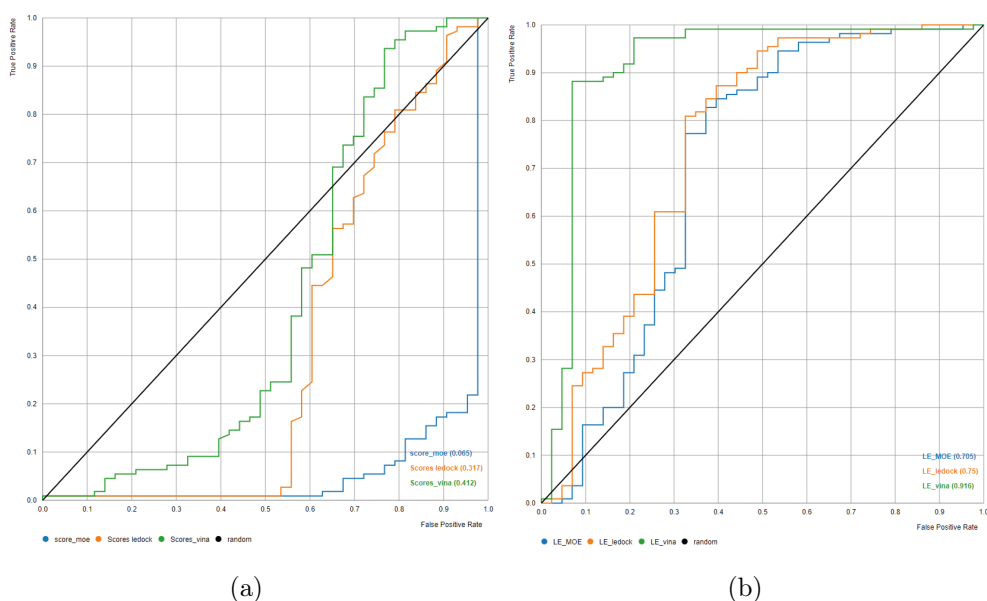


**Figura 6.1:** En color amarillo se muestran los contactos polares amarillo con los aminoácidos, en color rojo se denota el ligando co-cristalizado (SAH), en naranja la primer pose de acoplamiento en Autodock Vina, en color verde la primer pose de acoplamiento para MOE y en azul la primer pose de acoplamiento de Ledock. (a) DNMT1, (b) DNMT3A (c) DNMT3B.

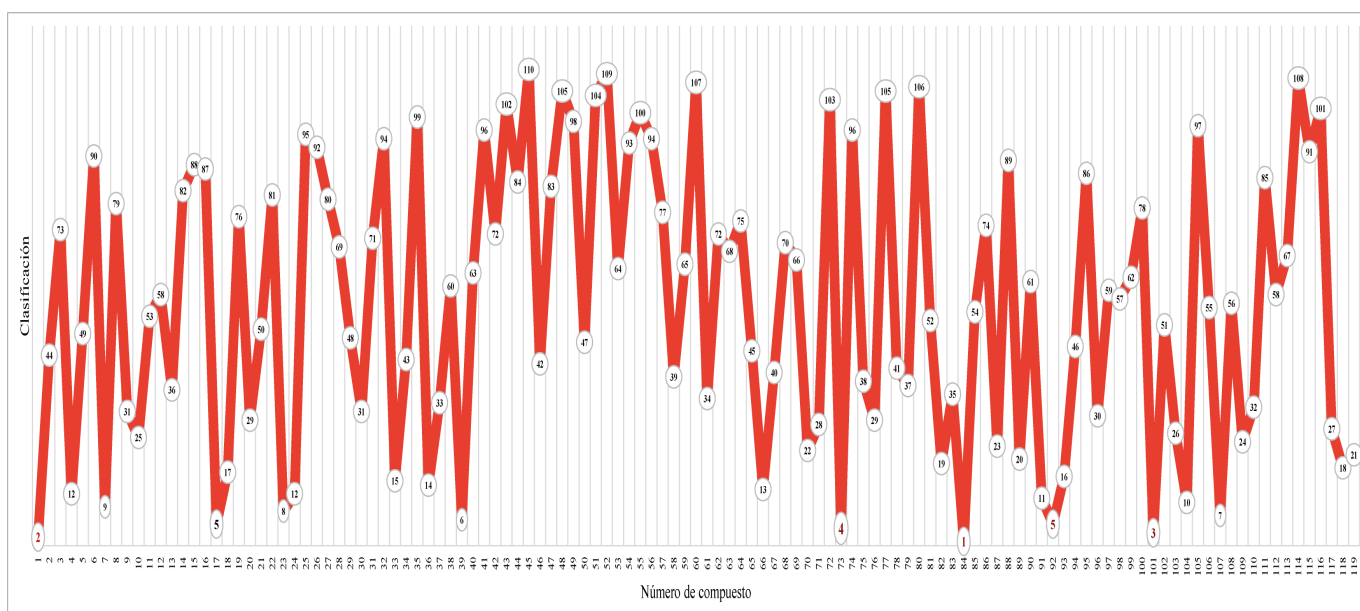
**Tabla 6.1:** Valores de RMSD de la validación de ligandos co-cristalizados para DNMT1, DNMT3A y DNMT3B.

<b>DNMT1 (PDB ID: 4WXX)</b>			
	Ligando co-cristalizado vs AutoDock Vina	Ligando co-cristalizado vs MOE	Ligando co-cristalizado vs LeDock
RMSD	0.97238 Å	1.65982 Å	3.46765 Å
<b>DNMT3A (PDB ID: 4U7T)</b>			
	Ligando co-cristalizado vs AutoDock Vina	Ligando co-cristalizado vs MOE	Ligando co-cristalizado vs LeDock
RMSD	0.7891 Å	1.25138 Å	3.35317Å
<b>DNMT3B (PDB ID: 6U8W)</b>			
	Ligando co-cristalizado vs AutoDock Vina	Ligando co-cristalizado vs MOE	Ligando co-cristalizado vs LeDock
RMSD	0.8576 Å	1.54941 Å	3.282895 Å

El validar el modo de unión permitió determinar si el método utilizado fue capaz de reproducir los resultados experimentales previamente obtenidos del PDB. Se puede observar que con los programas de AutoDock Vina y MOE se obtiene un valor de RMSD menor a 2.0 Å. Le Dock es el programa de acoplamiento que tiene un RMSD mayor de 2.0 para DNMT1, DNMT3A y DNMT3B. Estas diferencias se deben principalmente a los distintos métodos que cada programa utiliza para realizar el acoplamiento molecular. Por otro lado, el ligando co-cristalizado presenta enlaces rotables, esto ocasiona que haya diferencias significativas entre las poses de acoplamiento de los distintos programas.

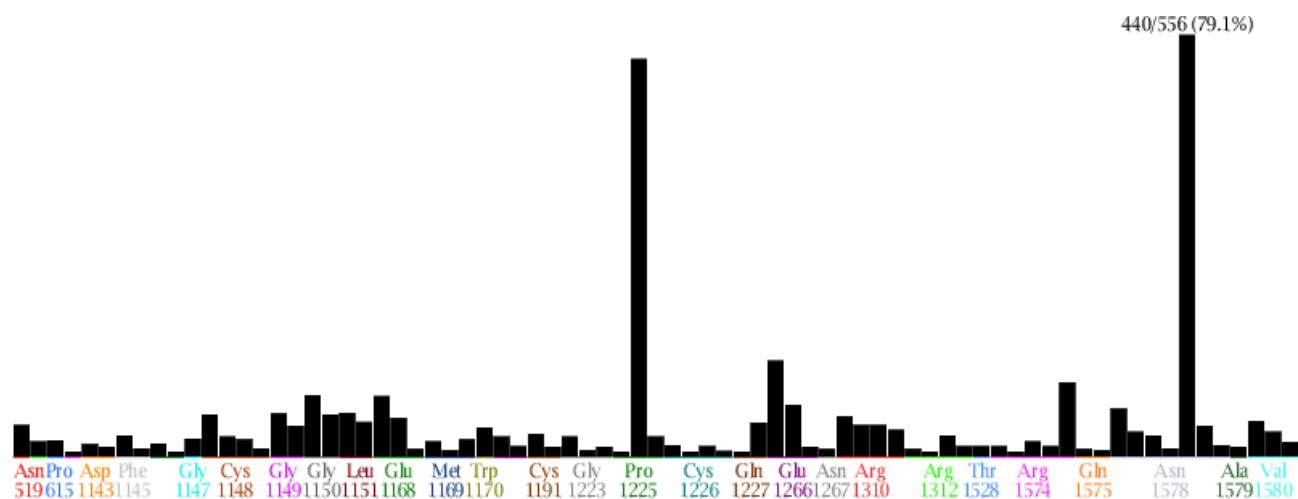


**Figura 6.2:** Curvas ROC para DNMT1 para los tres softwares. El color azul corresponde al programa de acoplamiento MOE, naranja LeDock y verde AutoDock Vina. En (a) se muestra la curva ROC contemplando la puntuación de cada programa (energía de afinidad). En (b) se muestra la curva ROC contemplando la eficiencia del ligando.

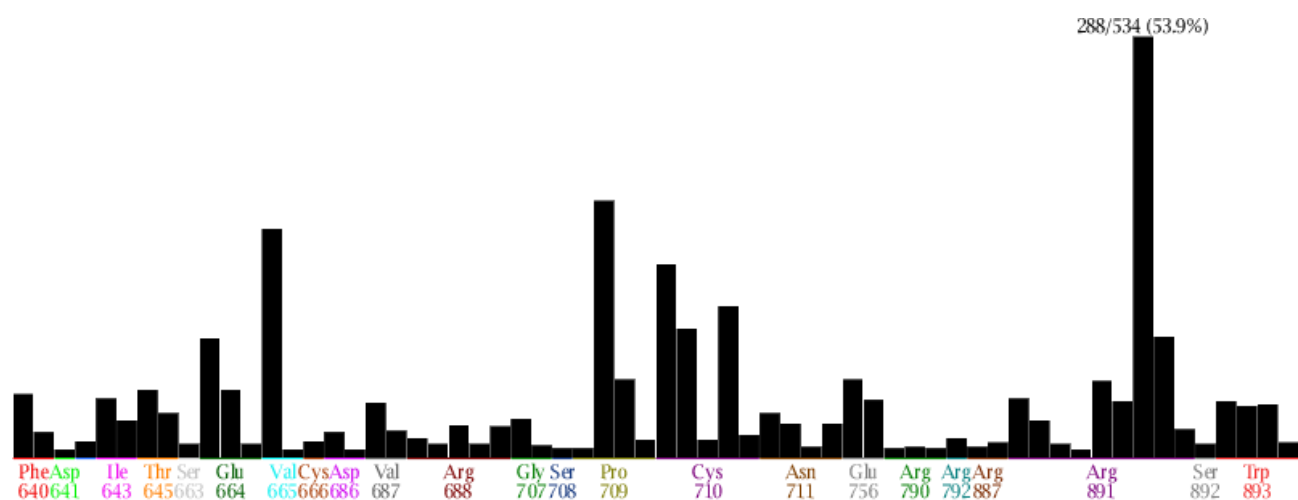


**Figura 6.3:** Clasificación consenso de 119 compuestos químicos alimentarios contra DNMT1, DNMT3A y DNMT3B.

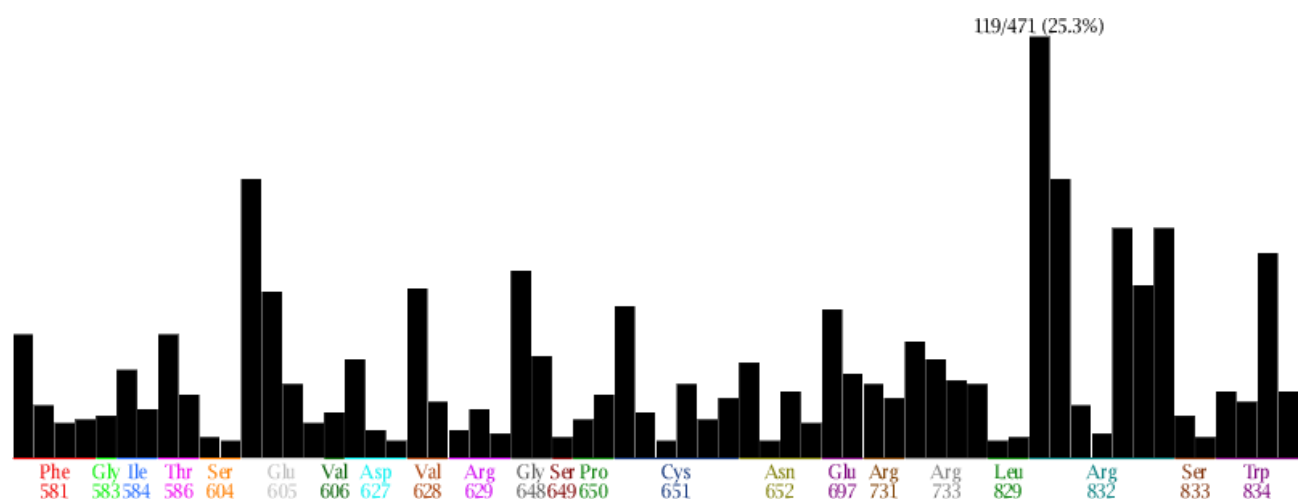
En la **Figura 6.3** se muestran la clasificación de los 119 compuestos químicos alimentarios con los tres programas de acoplamiento, la amentoflavona es el compuesto con la clasificación relativa con el número 1, galato de epigalocatequina con el número 2, hesperidina con el número 3, ácido fólico con el número 4 y la isoquercitrina con el número 5.



(a)



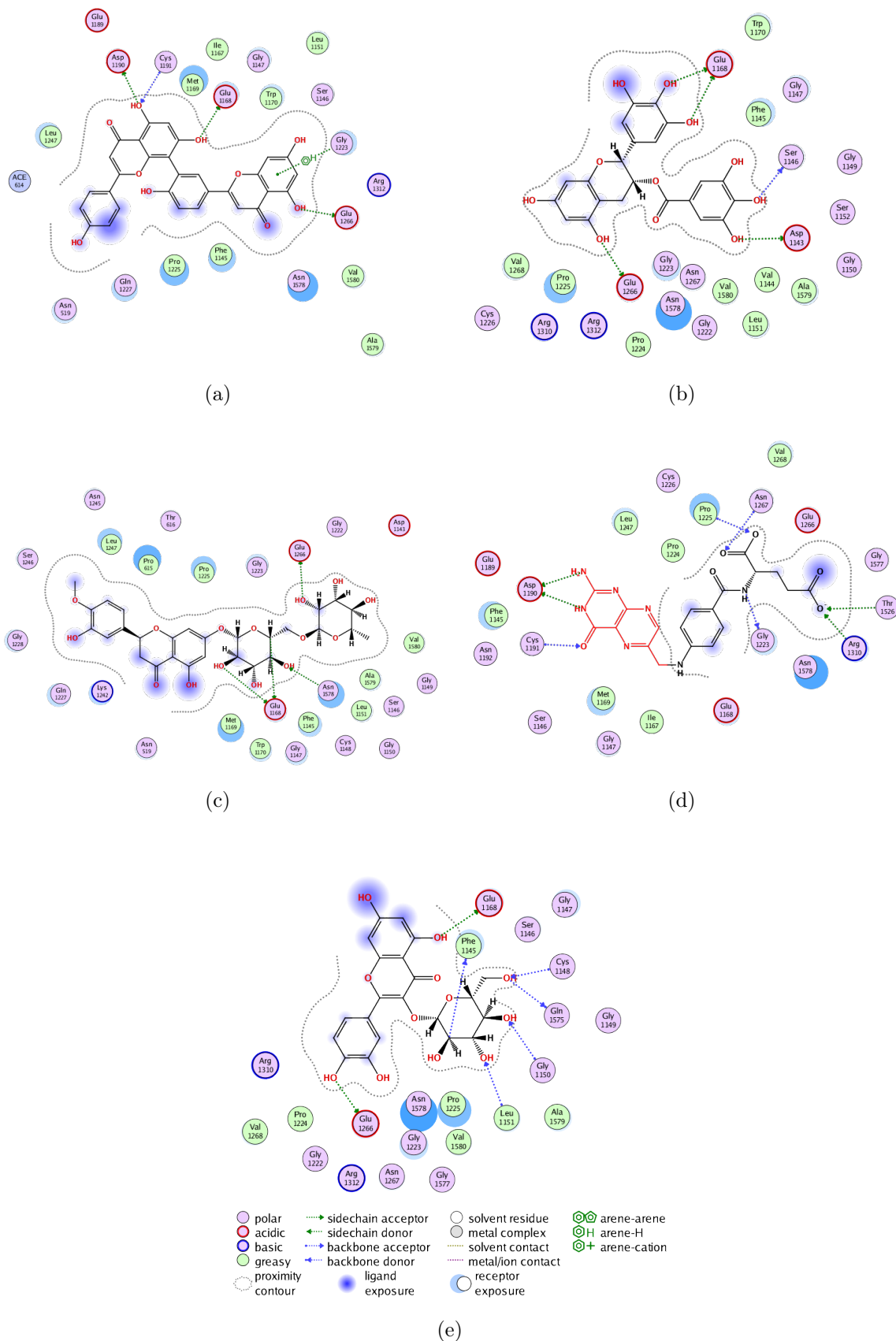
(b)



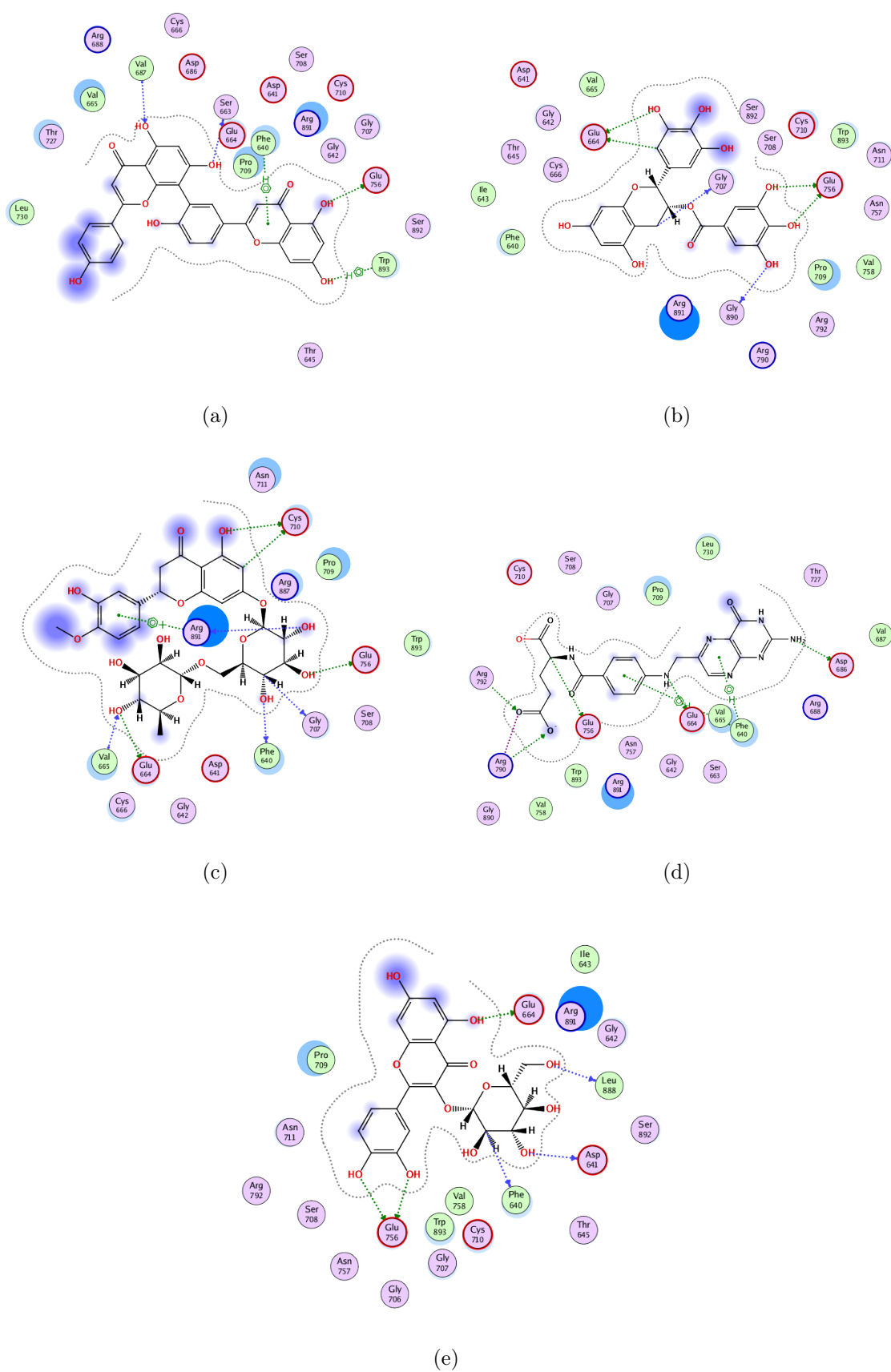
(c)

**Figura 6.4:** PLIFs consenso de acoplamiento molecular de 119 ligandos con los tres programas contra (a) DNMT1 (b) DNMT3A (c) DNMT3B.

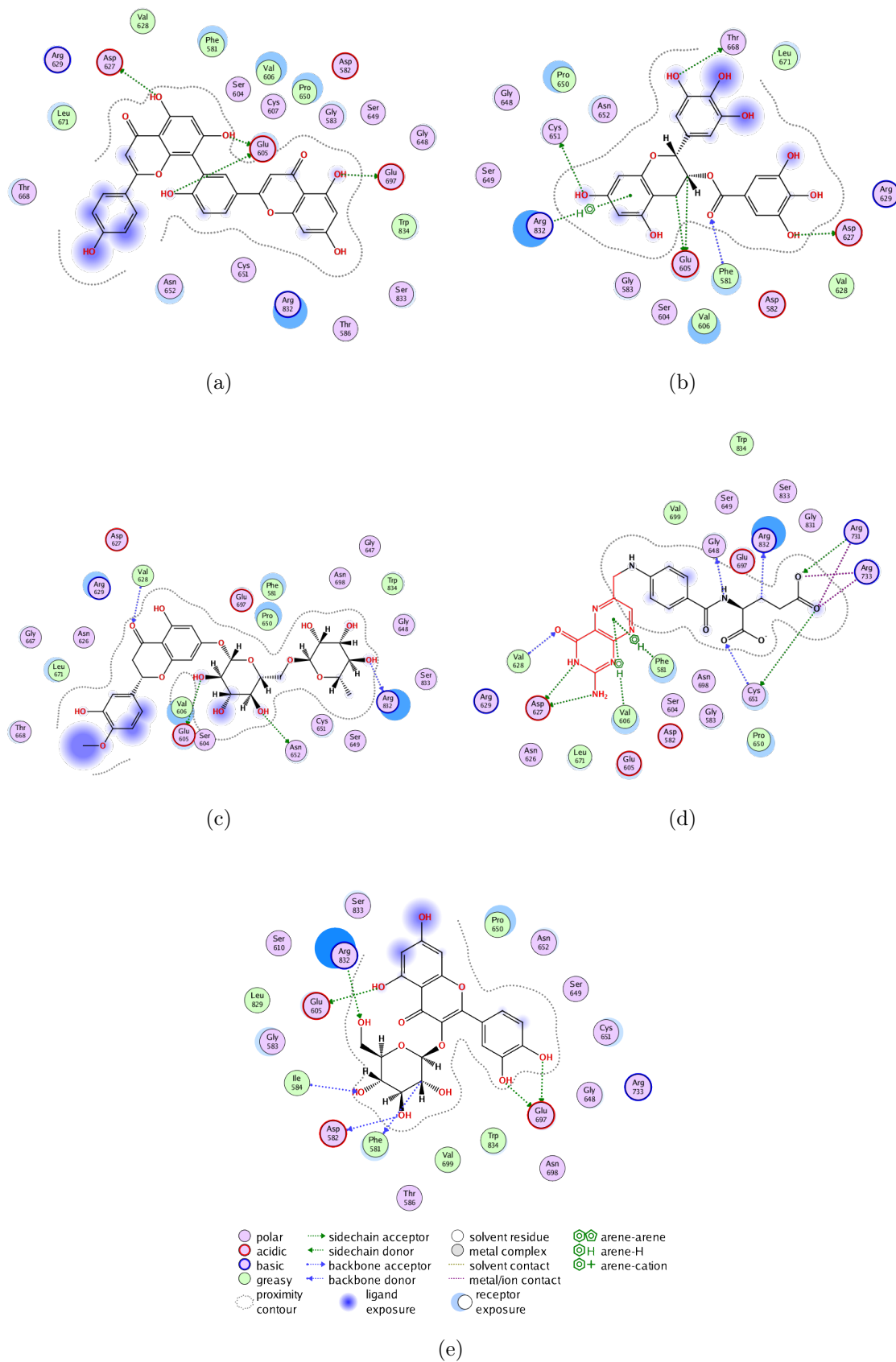




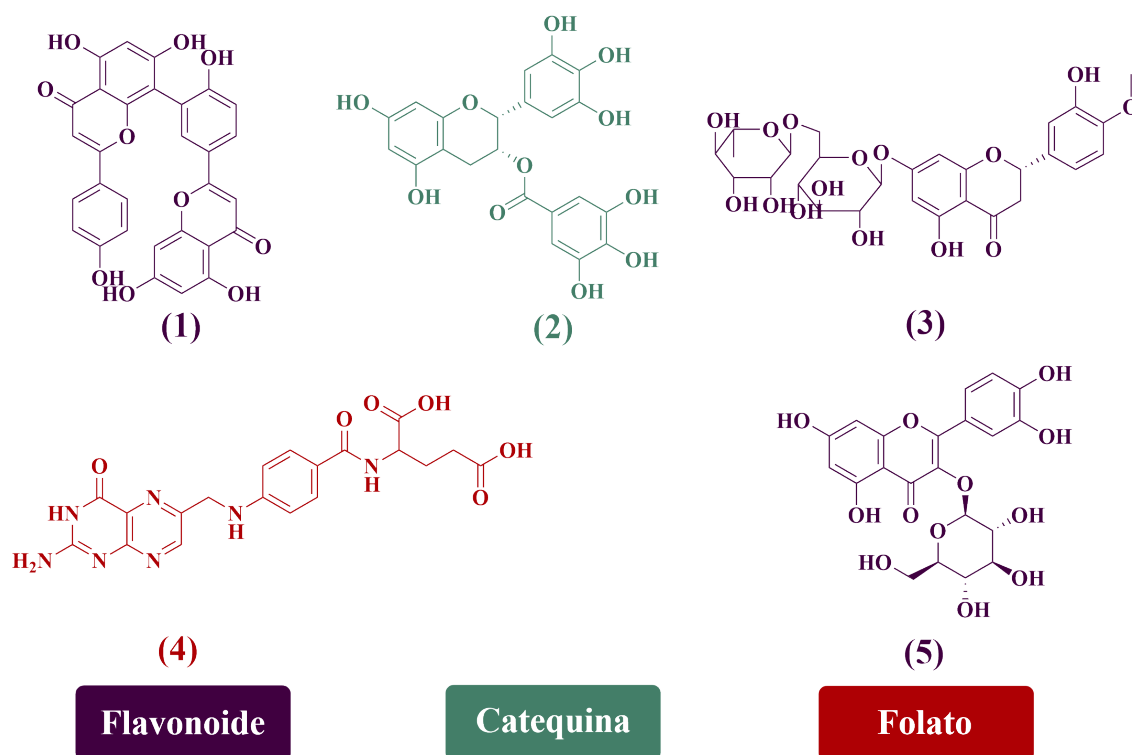
**Figura 6.5:** Diagramas 2D de (a) Amentoflavona (b) Galato de epigalocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT1.



**Figura 6.6:** Diagramas de interacciones 2D de (a) Amentoflavona (b) Galato de epigallocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT3A.



**Figura 6.7:** Diagramas de interacciones 2D de (a) Amentoflavona (b) Galato de epigallocatequina (c) Hesperidina (d) Ácido fólico (e) Isoquercitrina en DNMT3B.



**Figura 6.8:** Cinco compuestos principales con la mejor clasificación relativa. Los cinco compuestos acoplados para DNMT1, DNMT3A y DNMT3B, (1) Amentoflavona, (3) Hesperidina e (5) Isoquercitrina son compuestos clasificados estructuralmente como flavonoides, (2) galato de epigalocatequina clasificado estructuralmente como catequina y (4) ácido fólico, clasificado estructuralmente como folato.

La **Figura 6.8** muestra las estructuras químicas de los cinco compuestos con los mejores órdenes relativos globales obtenidos del consenso para DNMT1, DNMT3A y DNMT3B con los tres programas de acoplamiento molecular. Estos cinco compuestos se clasifican estructuralmente en tres categorías (flavonoides, catequinas y folatos). La figura muestra el total de aminoácidos en el cual hay interacción de los 119 compuestos alimentarios en el sitio activo de DNMT1. Pro 1225 y Asn 1578 son los aminoácidos de DNMT1 que tienen el mayor número de interacciones con los compuestos acoplados. No obstante, también se muestra interacción con otros aminoácidos que podrían ser claves para explicar el mecanismo de acción de compuestos alimentarios contra DNMT1 aunque en menor proporción, algunos de estos aminoácidos ya han sido estudiados por otro grupos de investigación, como Phe1145, Cys 119 y Glu1168 (X. Xu et al., 2023).

La amentoflavona presenta interacciones con los aminoácidos Asp 1190, Glu 1168, Glu 1266 y Gly 1147 en el sitio de acción de DNMT1(**Figura 6.5(a)**). En el sitio de DNMT3A presenta interacciones con Val 687, Ser 663, Phe 640, Glu 756 y Trp 893 (**Figura 6.6(a)**). En el sitio de acción de DNMT3B Asp 627, Glu 605 y Glu 697 (**Figura 6.7(a)**). Estos aminoácidos podrían ayudar a explicar el efecto que tiene la amentoflavona en tratamiento de enfermedades

estudiadas por otros investigadores como distintos tipos de cancer y el sars-cov2 (Xiong et al. 2021).

El galato de epigallocatequina presenta interacciones con los aminoácidos Glu 1168, Ser 1146, Asp 1143 y Glu 1266 en el sitio de acción de DNMT1 (**Figura 6.5(b)**). En el sitio de acción de DNMT3A presenta interacciones con Glu 664, Glu 756, Gly 890 y Gly 707 (**Figura 6.6(b)**). En el sitio de acción de DNMT3B presenta interacciones con Thre 668, Asp 627, Phe 581, Glu 605, Arg 832 y Cys 651 (**Figura 6.7(b)**).

La hesperidina presenta interacciones con los aminoácidos Glu 1266, Asn 1578 y Glu 1168 en el sitio de acción de DNMT1 (**Figura 6.5(c)**). En el sitio de acción de DNMT3A presenta interacciones con Cys 710, Glu 756, Gly 707, Phe 640, Glu 664 y Val 665 (**Figura 6.6(c)**). En el sitio de acción de DNMT3B presenta interacciones con Val 628, Glu 605, Asn 652 y Arg 832 (**Figura 6.7(c)**).

El ácido fólico presenta interacciones con los aminoácidos Pro 1225, Asn 1267, Thr 1526, Arg 1310, Gly 1223, Cys 1191 y Asp 1190 en el sitio de acción de DNMT1 (**Figura 6.5(d)**). En el sitio de acción de DNMT3A presenta interacciones con Asp 686, Phe 640, Glu 664, Glu 756, Arg 790 y Arg 792 (**Figura 6.6(d)**). En el sitio de acción de DNMT3B presenta interacciones con los aminoácidos Gly 648, Arg 832, Arg 731, Arg 733, Cys 651, Phe 581, Val 606, Asp 627 y Val 628(**Figura 6.7(d)**).

La isoquercitrina presenta interacciones con los aminoácidos Glu 1168, Cys 1148, Gln 1575, Gly 1150, Leu 1151 y Glu 1266 (**Figura 6.5(e)**). En el sitio de acción de DNMT3A presenta interacciones con Glu 664, Leu 888, Asp 641, Phe 640 y Glu 756 (**Figura 6.6(e)**). En el sitio de acción de DNMT3B Glu 697, Phe 581, Asp 582, Ile 584 , Glu 605 y Arg 832 (**Figura 6.7(e)**).

Estos resultados muestran cómo a través del acoplamiento molecular, junto con un protocolo consenso se puede predecir algunas interacciones de estos compuestos como una propuesta de mecanismo de acción que llevan a cabo a través de sus interacciones moleculares con el sitio de acción de DNMT1, DNMT3A y DNMT3B. Además, en el caso de los flavonoides (amentoflavona, hesperidina e isoquercitrina) se ha demostrado que presentan actividad contra la proteína 4 que contiene bromodominio (BRD4) (Xiong et al. 2021; Prieto-Martínez and Medina-Franco 2018) e histonas deacetilasas, estas proteínas corresponden a otras dianas epigenéticas que resultan de interés para comprender la influencia de las dianas epigenéticas en la prevención y desarrollo de múltiples enfermedades (Selvakumar et al. 2020). Cabe destacar que el folato, resulta tener una contribución importante durante el desarrollo de los bebés y se sugiere la necesidad urgente de asesoramiento personalizado previo a la concepción sobre el estado del folato, el estilo de vida y las condiciones médicas, en particular para las mujeres con obesidad (van der Windt et al. 2021).

# Capítulo 7

## Conclusiones

Se construyó una base de datos de 119 compuestos de origen alimenticio con actividad epigenética reportada. Esta base de datos es relevante en la investigación de fármacos y en el área de nutrición, ya que puede guiar a generar dietas personalizadas para el tratamiento y/o prevención de múltiples enfermedades y explicar el mecanismo de acción de compuestos presentes en alimentos. Se desarrollaron cuatro cuadernos de código en *Google Colaboratory* con la metodología empleada en este trabajo, la cual puede adaptarse a otros blancos terapéuticos y puede llegar a ser trasladado a otros lenguajes de programación. Con esto, se contribuye a fomentar la ciencia abierta y a la democratización de la ciencia. Se validaron los protocolos de acoplamiento molecular con las estructuras co-cristalizadas, obteniendo poses predichas que replican los resultados experimentales. Esto indica que los protocolos de acoplamiento molecular son robustos y tienen el potencial de predecir compuestos activos. El enfoque de acoplamiento consenso, junto con un filtrado basado en PLIFs para la validación externa, mostró ser un enfoque que proporciona mayor información biológica de relevancia que un acoplamiento tradicional. Por último, se propuso un mecanismo de acción con las probables interacciones con DNMT1, DNMT3A y DNMT3B de cinco compuestos presentes en los alimentos amentoflavona, galato de epigallocatequina, hesperidina, ácido fólico e isoquercitrina con mejor clasificación.

# Capítulo 8

## Perspectivas

Los compuestos alimenticios amentoflavona, galato de epigallocatequina, hesperidina, ácido fólico e isoquercitrina, pueden ser utilizados para la generación de dietas personalizadas para el tratamiento o prevención de múltiples enfermedades. Además, al realizar pruebas experimentales que demuestren que estos cinco compuestos pueden modular dianas epigenéticas (DNMT1, DNMT3A y DNMT3B) se puede evaluar su perfil ADME/TOX y de esta manera saber si requieren ser optimizados para su uso en la generación de nuevas moléculas. Los cuadernos de código desarrollados al ser un recurso de acceso libre pueden ser aplicados y adaptados para la necesidad u objetivos de otras dianas biológicas, no se limitan solo a las dianas epigenéticas. El estudio de los 119 compuestos químicos alimentarios se puede ampliar contra otras dianas epigenéticas (por ejemplo las histonas deacetilasas) para tener un panorama más extenso acerca del potencial de modulación de múltiples dianas epigenéticas.

# Capítulo 9

## Bibliografía

Aldrich, C. C., Calderón, F., Conway, S. J., He, C., Hooker, J. M., Huryn, D. M., Lindsley, C. W., Liotta, D. C., Müller, C. E. (2022). Virtual Special Issue: Epigenetics 2022. *ACS Chemical Biology*, 17(10), 2673–2678.

Amato, R. J., Stephenson, J., Hotte, S., Nemunaitis, J., Bélanger, K., Reid, G., Martell, R. E. (2012). MG98, a second-generation DNMT1 inhibitor, in the treatment of advanced renal cell carcinoma. *Cancer Investigation*, 30(5), 415–421.

Arora, I., Sharma, M., Tollefsbol, T. O. (2019). Combinatorial Epigenetics Impact of Polyphenols and Phytochemicals in Cancer Prevention and Therapy. *International Journal of Molecular Sciences*, 20(18). <https://doi.org/10.3390/ijms20184567>

Begum, N., Mandhare, A., Tryphena, K. P., Srivastava, S., Shaikh, M. F., Singh, S. B., Khatri, D. K. (2022). Epigenetics in depression and gut-brain axis: A molecular crosstalk. *Frontiers in Aging Neuroscience*, 14, 1048333.

Biswas, S., Rao, C. M. (2018). Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *European Journal of Pharmacology*, 837, 8–24.

Buocikova, V., Longhin, E. M., Pilalis, E., Mastrokalou, C., Miklikova, S., Cihova, M., Poturnayova, A., Mackova, K., Babelova, A., Trnkova, L., El Yamani, N., Zheng, C., Rios-Mondragon, I., Labudova, M., Csaderova, L., Kuracinova, K. M., Makovicky, P., Kucerova, L.,



Matuskova, M., ... Smolkova, B. (2022).

Decitabine potentiates efficacy of doxorubicin in a preclinical trastuzumab-resistant HER2-positive breast cancer models. In *Biomedicine Pharmacotherapy* (Vol. 147, p. 112662). <https://doi.org/10.1016/j.biopha.2022.112662>

Cabrera, A., Hernández, L. H., Chávez, D., Medina-Franco, J. L. (2018). Molecular Modeling of Potential Dual Inhibitors of HIV Reverse Transcriptase and Integrase. In *Computational Molecular Bioscience* (Vol. 08, Issue 01, pp. 1–41). <https://doi.org/10.4236/cmb.2018.81001>

Cao, M., Li, T., Chen, Y., Zhai, X. (2021). Nucleoside and Non-Nucleoside DOT1L Inhibitors: Dawn of MLLrearranged Leukemia. *Mini-Reviews in Medicinal Chemistry*, 21(11), 1337–1350.

Cavasotto, C. N., Di Filippo, J. I. (2021). In silico Drug Repurposing for COVID-19: Targeting SARS-CoV-2 Proteins through Docking and Consensus Ranking. In *Molecular Informatics* (Vol. 40, Issue 1, p. 2000115). <https://doi.org/10.1002/minf.202000115>

Chen, H., Hardy, T. M., Tollefsbol, T. O. (2011). Epigenomics of ovarian cancer and its chemoprevention. *Frontiers in Genetics*, 2, 67.

Chlamydas, S., Papavassiliou, A. G., Piperi, C. (2021). Epigenetic mechanisms regulating COVID-19 infection. In *Epigenetics* (Vol. 16, Issue 3, pp. 263–270). <https://doi.org/10.1080/15592294.2020.1796896>

Ciesielski, O., Biesiekierska, M., Balcerczyk, A. (2020). Epigallocatechin-3-gallate (EGCG) Alters Histone Acetylation and Methylation and Impacts Chromatin Architecture Profile in Human Endothelial Cells. *Molecules*, 25(10). <https://doi.org/10.3390/molecules25102326>

Conery, A. R., Rocnik, J. L., Trojer, P. (2022). Small molecule targeting of chromatin writers in cancer. *Nature Chemical Biology*, 18(2), 124–133. Coppedè, F. (2012).

The Epigenetics of Alzheimer's Disease. In *Epigenetics in Human Disease* (pp. 175–191). <https://doi.org/10.1016/b978-0-12-388415-2.00009-3>

Datta, J., Ghoshal, K., Denny, W. A., Gamage, S. A., Brooke, D. G., Phiasivongsa, P., Redkar, S., Jacob, S. T. (2009). A new class of quinoline-based DNA hypomethylating agents reactivates tumor suppressor genes by blocking DNA methyltransferase 1 activity and inducing its degradation. *Cancer Research*, 69(10), 4277–4285.

Derissen, E. J. B., Beijnen, J. H., Schellens, J. H. M. (2013). Concise drug review: azacitidine and decitabine. *The Oncologist*, 18(5), 619–624.

Diakou, I., Papakonstantinou, E., Papageorgiou, L., Pierouli, K., Dragoumani, K., Span-  
didos, D. A., Bacopoulou, F., Chrousos, G. P., Eliopoulos, E., Vlachakis, D. (2022).

Novel computational pipelines in antiviral structure-based drug design (Review). *Biome-  
dical Reports*, 17(6), 97. Drug discovery for epigenetics targets. (2022).

*Drug Discovery Today*, 27(4), 1088–1098. Engel, T., Gasteiger, J. (2018). *Applied Che-  
moinformatics: Achievements and Future Opportunities*. John Wiley Sons.

Ferreira, L., dos Santos, R., Oliva, G., Andricopulo, A. (2015). Molecular Docking and  
Structure-Based Drug Design Strategies. In *Molecules* (Vol. 20, Issue 7, pp. 13384–13421).  
<https://doi.org/10.3390/molecules200713384>

Furtado, C. L. M., Dos Santos Luciano, M. C., Da Silva Santos, R., Furtado, G. P.,  
Moraes, M. O., Pessoa, C. (2019). Epidrugs: targeting epigenetic marks in cancer treatment. In  
*Epigenetics* (Vol. 14, Issue 12, pp. 1164–1176). <https://doi.org/10.1080/15592294.2019.1640546>

Gentile, F., Oprea, T. I., Tropsha, A., Cherkasov, A. (2023). Surely you are joking, Mr  
Docking! *Chemical Society Reviews*, 52(3), 872–878. Ghosh, P., Saadat, A. (2021). Neurode-  
generation and epigenetics: A review. In *Neurología*. <https://doi.org/10.1016/j.nrl.2021.01.016>

Gigek, C. O., Chen, E. S., Smith, M. A. C. (2019). Epigenetic Alterations in Stomach  
Cancer: Implications for Diet and Nutrition. In *Handbook of Nutrition, Diet, and Epigenetics*  
(pp. 1005–1022). [https://doi.org/10.1007/978-3-319-55530-0\\_9](https://doi.org/10.1007/978-3-319-55530-0_9)

Halgren, T. A. (1999). MMFF VII. Characterization of MMFF94, MMFF94s, and other

widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *Journal of Computational Chemistry*, 20(7), 730–748.

Harrington, R. A., Adhikari, V., Rayner, M., Scarborough, P. (2019). Nutrient composition databases in the age of big data: foodDB, a comprehensive, real-time database infrastructure. In *BMJ Open* (Vol. 9, Issue 6, p. e026652). <https://doi.org/10.1136/bmjopen-2018-026652>

Heller, S. R., McNaught, A., Pletnev, I., Stein, S., Tchekhovskoi, D. (2015). InChI, the IUPAC International Chemical Identifier. *Journal of Cheminformatics*, 7, 23.

Hoo, Z. H., Candlish, J., Teare, D. (2017). What is an ROC curve? In *Emergency Medicine Journal* (Vol. 34, Issue 6, pp. 357–359). <https://doi.org/10.1136/emmermed-2017-206735>

Horton, J. R., Pathuri, S., Wong, K., Ren, R., Rueda, L., Fosbenner, D. T., Heerding, D. A., McCabe, M. T., Pappalardi, M. B., Zhang, X., King, B. W., Cheng, X. (2022). Structural characterization of dicyanopyridine containing DNMT1-selective, non-nucleoside inhibitors. *Structure*, 30(6), 793–802.e5.

Kaimala, S., Kumar, C. A., Allouh, M. Z., Ansari, S. A., Emerald, B. S. (2022). Epigenetic modifications in pancreas development, diabetes, and therapeutics. *Medicinal Research Reviews*, 42(3), 1343–1371.

Kuck, D., Caulfield, T., Lyko, F., Medina-Franco, J. L. (2010). Nanaomycin A selectively inhibits DNMT3B and reactivates silenced tumor suppressor genes in human cancer cells. *Molecular Cancer Therapeutics*, 9(11), 3015–3023.

Kuntz, K. W., Campbell, J. E., Keilhack, H., Pollock, R. M., Knutson, S. K., Porter-Scott, M., Richon, V. M., Sneeringer, C. J., Wigle, T. J., Allain, C. J., Majer, C. R., Moyer, M. P., Copeland, R. A., Chesworth, R. (2016). The Importance of Being Me: Magic Methyls, Methyltransferase Inhibitors, and the Discovery of Tazemetostat. *Journal of Medicinal Chemistry*, 59(4), 1556–1564.

Laubach, J. P., Moreau, P., San-Miguel, J. F., Richardson, P. G. (2015). Panobinostat

for the Treatment of Multiple Myeloma. In *Clinical Cancer Research* (Vol. 21, Issue 21, pp. 4767–4773). <https://doi.org/10.1158/1078-0432.ccr-15-0530>

Lévesque, N., Leclerc, D., Rozen, R. (2019). Folate and Epigenetics: Colorectal Cancer Risk and Detection. In *Handbook of Nutrition, Diet, and Epigenetics* (pp. 61–78). [https://doi.org/10.1007/978-3-319-55530-0\\_3](https://doi.org/10.1007/978-3-319-55530-0_3)

Lintas, C. (2019). Linking genetics to epigenetics: The role of folate and folate-related pathways in neurodevelopmental disorders. *Clinical Genetics*, 95(2), 241–252.

Liu, H.-Y., Liu, S.-M., Zhang, Y.-Z. (2020). Maternal Folic Acid Supplementation Mediates Offspring Health via DNA Methylation. *Reproductive Sciences*, 27(4), 963–976.

Lopez, M., Gilbert, J., Contreras, J., Halby, L., Arimondo, P. B. (2022). Inhibitors of DNA Methylation. *Advances in Experimental Medicine and Biology*, 1389, 471–513.

Mahmoud, A. M. (2022). An Overview of Epigenetics in Obesity: The Role of Lifestyle and Therapeutic Interventions. *International Journal of Molecular Sciences*, 23(3). <https://doi.org/10.3390/ijms23031341>

Maia, E. H. B., Assis, L. C., de Oliveira, T. A., da Silva, A. M., Taranto, A. G. (2020). Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Frontiers in Chemistry*, 8, 343.

Mann, B. S., Johnson, J. R., Cohen, M. H., Justice, R., Pazdur, R. (2007). FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *The Oncologist*, 12(10), 1247–1252.

Meng, X.-Y., Zhang, H.-X., Mezei, M., Cui, M. (2011). Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer-Aided Drug Design*, 7(2), 146–157.

Moskowitz, A. J., Horwitz, S. M. (2017). Targeting histone deacetylases in T-cell lymphoma. *Leukemia Lymphoma*, 58(6), 1306–1319.

Nasir, A., Bullo, M. M. H., Ahmed, Z., Imtiaz, A., Yaqoob, E., Jadoon, M., Ahmed, H., Afreen, A., Yaqoob, S. (2020). Nutrigenomics: Epigenetics and cancer prevention: A comprehensive review. *Critical Reviews in Food Science and Nutrition*, 60(8), 1375–1387.

Naveja, J. J., Rico-Hidalgo, M. P., Medina-Franco, J. L. (2018). Analysis of a large food chemical database: chemical space, diversity, and complexity. *F1000Research*, 7. <https://doi.org/10.12688/f1000research.15440.2>

O’Boyle, N. M. (2012). Towards a Universal SMILES representation - A standard method to generate canonical SMILES based on the InChI. *Journal of Cheminformatics*, 4(1), 22. Obuchowski, N. A. (2005). ROC analysis. *AJR. American Journal of Roentgenology*, 184(2), 364–372.

Palacio-Rodríguez, K., Lans, I., Cavasotto, C. N., Cossio, P. (2019). Exponential consensus ranking improves the outcome in docking and receptor ensemble docking. *Scientific Reports*, 9(1), 5142.

Paul, P. (2017). Food Informatics and Its Challenges and Opportunities - A Review. Peixoto, P., Cartron, P.-F., Serandour, A. A., Hervouet, E. (2020). From 1957 to Nowadays: A Brief History of Epigenetics. *International Journal of Molecular Sciences*, 21(20). <https://doi.org/10.3390/ijms21207571>

Prieto-Martínez, F. D., Arciniega, M., Medina-Franco, J. L. (2018). Acoplamiento Molecular: Avances Recientes y Retos. In *TIP Revista Especializada en Ciencias Químico-Biológicas* (Vol. 21). <https://doi.org/10.22201/fesz.23958723e.2018.0.143>

Prusevich, P., Kalin, J. H., Ming, S. A., Basso, M., Givens, J., Li, X., Hu, J., Taylor, M. S., Cieniewicz, A. M., Hsiao, P.-Y., Huang, R., Roberson, H., Adejola, N., Avery, L. B., Casero, R. A., Jr, Taverna, S. D., Qian, J., Tackett, A. J., Ratan, R. R., ... Cole, P. A. (2014). A selective phenelzine analogue inhibitor of histone demethylase LSD1. *ACS Chemical Biology*, 9(6), 1284–1293.

Qi, W., Chan, H., Teng, L., Li, L., Chuai, S., Zhang, R., Zeng, J., Li, M., Fan, H., Lin,

Y., Gu, J., Ardayfio, O., Zhang, J.-H., Yan, X., Fang, J., Mi, Y., Zhang, M., Zhou, T., Feng, G., ... Li, E. (2012). Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 109(52), 21360–21365.

Ramos-Lopez, O., Milagro, F. I., Riezu-Boj, J. I., Martinez, J. A. (2021). Epigenetic signatures underlying inflammation: an interplay of nutrition, physical activity, metabolic diseases, and environmental factors for personalized nutrition. *Inflammation Research: Official Journal of the European Histamine Research Society ... [et Al.]*, 70(1), 29–49.

Ren, W., Gao, L., Song, J. (2018). Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes*, 9(12), 620. Santos, A., Tsafou, K., Stolte, C., Pletscher-Frankild, S., O'Donoghue, S. I., Jensen, L. J. (2015). Comprehensive comparison of large-scale tissue expression datasets. *PeerJ*, 3, e1054.

Sapienza, C., Issa, J.-P. (2016). Diet, Nutrition, and Cancer Epigenetics. *Annual Review of Nutrition*, 36, 665–681.

Selvakumar, P., Badgeley, A., Murphy, P., Anwar, H., Sharma, U., Lawrence, K., Lakshmi-kuttyamma, A. (2020). Flavonoids and Other Polyphenols Act as Epigenetic Modifiers in Breast Cancer. *Nutrients*, 12(3). <https://doi.org/10.3390/nu12030761>

Sethi, A., Joshi, K., Sasikala, K., Alvala, M. (2019). Molecular Docking in Modern Drug Discovery: Principles and Recent Applications. In *Drug Discovery and Development - New Advances*. IntechOpen.

Sheils, T. K., Mathias, S. L., Kelleher, K. J., Siramshetty, V. B., Nguyen, D.-T., Bologna, C. G., Jensen, L. J., Vidović, D., Koleti, A., Schürer, S. C., Waller, A., Yang, J. J., Holmes, J., Bocci, G., Southall, N., Dharkar, P., Mathé, E., Simeonov, A., Oprea, T. I. (2021). TCRD and Pharos 2021: mining the human proteome for disease biology. *Nucleic Acids Research*, 49(D1), D1334–D1346.

Stankevičius, V., Gibas, P., Masiulionytė, B., Gasiulė, L., Masevičius, V., Klimašauskas, S., Vilkaitis, G. (2022). Selective chemical tracking of Dnmt1 catalytic activity in live cells.

Molecular Cell, 82(5), 1053–1065.e8.

Stevens, A. J., Rucklidge, J. J., Kennedy, M. A. (2018). Epigenetics, nutrition and mental health. Is there a relationship? *Nutritional Neuroscience*, 21(9), 602–613.

Talpaz, M., Kiladjian, J.-J. (2021). Fedratinib, a newly approved treatment for patients with myeloproliferative neoplasm-associated myelofibrosis. *Leukemia*, 35(1), 1–17.

Triches, F., Triches, F., Lino de Oliveira, C. (2022). Consensus combining outcomes of multiple ensemble dockings: examples using dDAT crystalized complexes. *MethodsX*, 9, 101788.

Trott, O., Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461.

van der Windt, M., Schoenmakers, S., van Rijn, B., Galjaard, S., Steegers-Theunissen, R., van Rossem, L. (2021). Epidemiology and (Patho)Physiology of Folic Acid Supplement Use in Obese Women before and during Pregnancy. *Nutrients*, 13(2). <https://doi.org/10.3390/nu13020331>

Van Veggel, M., Westerman, E., Hamberg, P. (2018). Clinical Pharmacokinetics and Pharmacodynamics of Panobinostat. *Clinical Pharmacokinetics*, 57(1), 21–29.

Vilar, S., Cozza, G., Moro, S. (2008). Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. In *Current Topics in Medicinal Chemistry* (Vol. 8, Issue 18, pp. 1555–1572).

<https://doi.org/10.2174/156802608786786624>

Wang, J., Kim, T. H., Ahn, M. Y., Lee, J., Jung, J. H., Choi, W. S., Lee, B. M., Yoon, K. S., Yoon, S., Kim, H. S. (2012). Sirtinol, a class III HDAC inhibitor, induces apoptotic and autophagic cell death in MCF-7 human breast cancer cells. *International Journal of Oncology*, 41(3), 1101–1109.

Wang, T., Cui, H., Ma, N., Jiang, Y. (2013). Nicotinamide-mediated inhibition of SIRT1 deacetylase is associated with the viability of cancer cells exposed to antitumor agents and

apoptosis. *Oncology Letters*, 6(2), 600–604.

Wang, Y.-M., Gu, M.-L., Meng, F.-S., Jiao, W.-R., Zhou, X.-X., Yao, H.-P., Ji, F. (2017). Histone acetyltransferase p300/CBP inhibitor C646 blocks the survival and invasion pathways of gastric cancer cell lines. *International Journal of Oncology*, 51(6), 1860–1868.

Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S., Hou, T. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. In *Physical Chemistry Chemical Physics* (Vol. 18, Issue 18, pp. 12964–12975). <https://doi.org/10.1039/c6cp01555g>

Willmann, D., Lim, S., Wetzel, S., Metzger, E., Jandausch, A., Wilk, W., Jung, M., Forne, I., Imhof, A., Janzer, A., Kirfel, J., Waldmann, H., Schüle, R., Buettner, R. (2012). Impairment of prostate cancer cell growth by a selective and reversible lysine-specific demethylase 1 inhibitor. *International Journal of Cancer. Journal International Du Cancer*, 131(11), 2704–2709.

Xiong, X., Tang, N., Lai, X., Zhang, J., Wen, W., Li, X., Li, A., Wu, Y., Liu, Z. (2021). Insights Into Amentoflavone: A Natural Multifunctional Biflavonoid. *Frontiers in Pharmacology*, 12, 768708.

Xu, S., Wang, L., Pan, X. (2021). An evaluation of combined strategies for improving the performance of molecular docking. *Journal of Bioinformatics and Computational Biology*, 19(2), 2150003.

Xu, X., Ji, S., Chen, Y., Xia, S., Li, Y., Chen, L., Li, Y., Zhang, F., Zhang, Z., Zheng, S. (2023). Induction of DNMT1-dependent demethylation of SHP-1 by the natural flavonoid compound Baicalein overcame Imatinib-resistance in CML CD34 cells. *Cell Communication and Signaling: CCS*, 21(1), 47.

Ye, J., Yang, X., Ma, C. (2022). Ligand-Based Drug Design of Novel Antimicrobials against *Staphylococcus aureus* by Targeting Bacterial Transcription. In *International Journal of Molecular Sciences* (Vol. 24, Issue 1, p. 339). <https://doi.org/10.3390/ijms24010339>



Zhu, Q., Huang, Y., Marton, L. J., Woster, P. M., Davidson, N. E., Casero, R. A., Jr. (2012). Polyamine analogs modulate gene expression by inhibiting lysine-specific demethylase 1 (LSD1) and altering chromatin structure in human breast cancer cells. *Amino Acids*, 42(2-3), 887-898

# Capítulo 10

## Anexo

Hipervínculo de cálculos de clasificación consenso: [Clasificación consenso](#)

Hipervínculo de GitHub de los cuadernos de códigos elaborados: [cuadernos de códigos](#)

Contenido:

1. Búsqueda y visualización de estructuras del PDB
2. Acoplamiento molecular Autodock VINA
3. Cálculo de RMSD

Bioactividad ChEMBL: [Bioactividad de ChEMBL](#)

Tabla 10.1: Clasificación final.

Ligando	Parcial final	Clasificación final
Galato de epigallocatequina	1,09	2
Quercetina	13,45	44
Kaempferol	20,00	73
Apigenina	4,45	12
Luteolina	14,27	49
Tangeretina	26,09	90
Silibinina	4,00	9
Genisteína	22,73	79
Delfinidina	10,45	31
Berberina	9,09	25
Evodiamina	15,27	53
3,3-diindolilmetano (DIM)	16,18	58
Mahanina	11,64	36
Harmalina	23,09	82
Harmina	25,36	88
Vincristina	25,27	87
Isofistularina-3	1,64	5
Betanina	6,18	17
Indicaxantina	20,73	76
Solamargina	9,82	29
Reserpina	14,36	50
Nicotinamida	23,00	81
Psammaplina A	3,36	8
Psammaplina F	4,45	12
Ácido boswélico	27,09	95
Ácido corosólico	26,36	92
Cucurbitacina B	22,82	80
Ácido ursólico	19,27	69
Eugenol	14,09	48
Grifolin	10,45	31
Ácido holo-transretinoico	19,73	71
Partenolida	26,82	94
Z-ligustilida	5,73	15
Beta-elemene	13,36	43
Hinokitiol	28,73	99

Antroquinonol D	5,64	14
Emodina	10,91	33
Ácido laccaico	16,55	60
Physcion	1,82	6
8-O-beta-glucopiranosido (PG)		
Shikonina	18,09	63
Naftazarina (Naph)	27,18	96
Nanaomicina A	19,82	72
Timoquinona	29,64	102
Isocianato de feniletil (PEICT)	23,36	84
Sulforafano (SFN)	32,45	110
Isoliquiritigenina	13,00	42
Resveratrol	23,27	83
Astaxantina	30,45	105
Fucoxantina	27,91	98
Licopeno	13,91	47
Beta-caroteno	30,36	104
Butirato	32,18	109
Ácido docosahexaenoico (DHA)	18,36	64
Ácido eicosapentaenoico (EPA)	26,45	93
Ácido linoleico	29,36	100
Hexafosfato de inositol (IP6)	26,82	94
Honokiol	20,82	77
Ácido nordihidroguaiarético	12,64	39
Peperomina E	18,55	65
Fucoidan	31,00	107
Poliflina I (IBP)	11,09	34
Oleaceína	19,82	72
Guggulsterona	18,82	68
Prostaglandina E2 (PGE2)	20,55	75
Witaferina A (WA)	13,73	45
Ácido elágico (EA)	5,09	13
Procianidina B2	12,82	40
Tanshinona IIA	19,64	70
Glucarato de calcio	18,64	66
Ácido clorogénico	8,45	22
Proantocianidinas	9,64	28
Curcumol	30,09	103
Ácido fólico	1,55	4
Vitamina d3	27,18	96
Vitamina E	12,36	38
Amarogentina	9,82	29

Ácido ascórbico	30,45	105
S-alilcisteína (SAC)	12,91	41
Gosipol	12,18	37
Indol-3-carbinol	30,82	106
Piceatannol	14,91	52
Curcumina	7,91	19
Triacetato de naringenina	11,27	35
Amentoflavona	0,27	1
Fisetina	15,36	54
Aristoyagonina	20,36	74
Alfa-viniferina	8,64	23
Ácido anacárdico	25,45	89
Garcinol	8,00	20
Epiberberina	17,00	61
Wogonósido	4,36	11
Isoquercitrina	1,64	5
Alfa-mangostina	5,82	16
Melatonina	13,82	46
Ácido geranilgeránico	24,73	86
Miricetina	9,91	30
Daidzeína	16,27	59
Cianidina	15,82	57
Malvidina	17,27	62
Pelargonidina	21,45	78
Hesperidina	1,27	3
Pterostilbeno	14,82	51
Ácido gálico	9,18	26
Xanthohumol	4,27	10
Ácido cafeico	27,27	97
Floretina	15,55	55
Florizina	2,36	7
Secoisolariciresinol	15,64	56
Matairesinol	8,73	24
Arctigenina	10,82	32
Alilo mercaptano	23,73	85
Ácido cumárico	16,18	58
Alicina	18,73	67
Aliina	31,64	108
Disulfuro de dialil	26,18	91
Zerumbone	29,45	101
Osajin	9,55	27
Pomiferina	6,82	18
Buteína	8,18	21