



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS**

ESTUDIO *IN SILICO* DE METILTRANSFERASA DE ÁCIDO DESOXIRRIBONUCLEICO:  
MODELADO POR HOMOLOGÍA E IDENTIFICACIÓN DE POTENCIALES INHIBIDORES

**TESIS**

PARA OPTAR POR EL GRADO DE

**MAESTRO EN CIENCIAS**

PRESENTA

**QFB. ALEJANDRO GÓMEZ GARCÍA**

DR. JOSÉ LUIS MEDINA FRANCO  
DEPARTAMENTO DE FARMACIA, FACULTAD DE QUÍMICA, UNAM

Ciudad de México, Junio 2020



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.



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS**

**ESTUDIO *IN SILICO* DE METILTRANSFERASA DE ÁCIDO  
DESOXIRRIBONUCLEICO: MODELADO POR HOMOLOGÍA E  
IDENTIFICACIÓN DE POTENCIALES INHIBIDORES**

**TESIS  
PARA OPTAR POR EL GRADO DE**

**MAESTRO EN CIENCIAS**

**PRESENTA**

**QFB. ALEJANDRO GÓMEZ GARCÍA**



Ciudad de México, Junio 2020

El presente trabajo se realizó en el grupo DIFACQUIM ubicado en el cubículo 117 del edificio F de la Facultad de Química de la UNAM, durante el periodo agosto de 2018 - abril de 2020, bajo la supervisión del Dr. José Luis Medina Franco. El estudio *in silico* de la enzima NLRP3 se llevó a cabo en la Universidad de Turín, Italia, bajo la tutoría del Dr. Massimo Bertinaria durante el periodo septiembre - diciembre de 2019.

---

Vo.Bo. Dr. José Luis Medina Franco

---

QFB. Alejandro Gómez García

El jurado asignado a este trabajo está integrado por:

Dr. Miguel Antonio Costas Basín	Facultad de Química, UNAM
Dra. Adela Rodríguez Romero	Instituto de Química, UNAM
Dr. Alfonso Sebastián Lira Rocha	Facultad de Química, UNAM
Dra. Laura Domínguez Dueñas	Facultad de Química, UNAM
Dr. Marcelino Arciniega Castro	Instituto de Fisiología Celular, UNAM

Parte de los resultados de este proyecto se difundieron en el siguiente artículo y congresos.

### **Artículo:**

- Saldívar-González, F. I.; Gómez-García, A.; Chávez-Ponce de León, D. E.; Sánchez-Cruz, N.; Ruiz-Rios, J.; Pilon-Jiménez, B. A.; Medina-Franco, J. L. Inhibitors of DNA Methyltransferases from Natural Sources: A Computational Perspective. *Front. Pharmacol.* **2018**, 9, 1144.

### **Congresos:**

#### *Nacionales:*

- IV Simposio “Tendencias Actuales en la Búsqueda y Desarrollo de Fármacos”. 13-15 de junio de 2018. Ciudad de México, UNAM, Facultad de Química (Cartel).

#### *Internacionales:*

- American Chemical Society Fall 2019 National Meeting & Expo. Medicinal Chemistry Division. 25-29 de agosto de 2019. San Diego, California, Estados Unidos (cartel).
- Merck Young Chemists' Symposium 2019. 25-27 de noviembre de 2019. Rimini, Italia (cartel).

# Agradecimientos

A mi familia, ya que sin su apoyo incondicional nunca hubiera logrado llegar hasta aquí.

A mi tutor y compañeros de DIFACQUIM por su su asesoría, consejo y guía durante este trabajo.

A mis amigos, ya que ellos le dan ese toque especial a mi estancia en este espacio-tiempo.

A la UNAM y la Facultad de Química que me proporcionaron la mejor educación que puede encontrarse en este país.

A mis sinodales por su tiempo y consejo.

Beca CONACyT (No. 912137) y Beca de Movilidad Extranjera CONACyT 2019.

También se agradece el apoyo del Consejo Nacional de Ciencia y Tecnología (CONACyT), proyecto Ciencia Básica 282785 y del Programa de Nuevas Alternativas para el Tratamiento de Enfermedades Infecciosas (NUATEI), del Instituto de Investigaciones Biomédicas, UNAM.

# Resumen

La metilación del ADN juega un papel fundamental en la regulación de la expresión de genes. La familia de las ADN metiltransferasas (DNMTs) se encarga de establecer y mantener los patrones de metilación en el ADN. La pérdida de los patrones de metilación normales del ADN es observada en distintas patologías, entre ellas ciertos tipos de cáncer. La inhibición de las enzimas DNMT es una estrategia utilizada actualmente como terapia para el tratamiento de distintos tipos de tumores. A la fecha, existen sólo dos fármacos que inhiben DNMTs, azacitadina y decitabina. Sin embargo, su uso es restringido debido a su toxicidad y pobre estabilidad química. Actualmente, únicamente se conoce la estructura cristalográfica de DNMT1, DNMT2, DNMT3A1 y DNMT3L. Por lo tanto, es muy complicado hacer un diseño sistemático de inhibidores hacia el resto de las isoformas cuya estructura actualmente no se encuentra reportada. Cuando no se tiene acceso a la información experimental de las coordenadas tri-dimensionales de las estructuras, el modelado por homología permite predecir de manera *in silico* una estructura tridimensional aproximada de una proteína. Aún no se reportan las estructuras cristalográficas y/o por RMN de las isoformas catalíticamente activas DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Sin embargo, conocer la estructura del dominio catalítico de estas isoformas es conveniente si se pretende hacer un diseño basado en la estructura de inhibidores de estas enzimas. El modelado por homología es una opción para poder predecir el dominio catalítico de estas isoformas de DNMT. El objetivo de este trabajo es identificar potenciales inhibidores de DNMT1, DNMT2, DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Para las isoformas cuya estructura no ha sido reportada, se recurrió al modelado por homología.

## Índice general

<b>Abreviaturas y acrónimos</b>	<b>9</b>
<b>1. Introducción</b>	<b>10</b>
<b>2. Objetivos</b>	<b>12</b>
<b>3. Hipótesis</b>	<b>13</b>
<b>4. Antecedentes</b>	<b>13</b>
4.1 La familia de las DNMT	13
4.1.1 Función biológica y estructura	13
4.1.2 Mecanismo de acción	16
4.1.3 Isoformas	16
4.2 La búsqueda de inhibidores de DNMT	17
4.2.1 Patologías que involucran a las DNMT	17
4.2.2 Inhibidores aprobados actualmente para uso clínico	18
4.2.3 Inhibidores en desarrollo, compuestos cabezas de serie (hits)	19
4.3 Inhibidores de DNMT en los alimentos	20
<b>5. Procedimiento experimental</b>	<b>21</b>
5.1 Modelado por homología	21
5.1.1 Identificación del dominio catalítico	21
5.1.2 Alineamiento de las secuencias y construcción de la matriz de identidad	21
5.1.3 Construcción del modelo	21
5.1.4 Validación del modelo	23
5.2 Búsqueda de sitios de unión alternos al dominio catalítico	24
5.3 Cribado virtual	25
5.3.1 Curado	25
5.3.2 Filtrado	25
5.3.3 Acoplamiento molecular	26
5.4 Análisis de componentes principales	27
<b>6. Discusión de resultados</b>	<b>28</b>
6.1 Modelado por homología	28
6.1.1 Alineamiento de las secuencias y construcción de la matriz de identidad	28
6.1.2 Construcción del modelo y comparación con DNMT3A1	29
6.1.3 Validación del modelo	31
6.2 Búsqueda de sitios de unión alternos al dominio catalítico	36
6.3 Cribado virtual	41
6.4 Análisis de componentes principales	49
<b>7. Conclusiones</b>	<b>53</b>



<b>8.</b>	<b>Perspectivas</b>	<b>54</b>
<b>9.</b>	<b>Referencias</b>	<b>55</b>
<b>10.</b>	<b>Apéndice</b>	<b>61</b>
10.1	Alineamiento de la secuencia primaria de aminoácidos del dominio catalítico	61
10.2	Aminoácidos que conforman el sitio donde se lleva a cabo la metilación del ADN y otros potenciales sitios de unión para moléculas “tipo-fármaco”	65
10.3	Estudio in silico de NLRP3: Modelado por homología e identificación de potenciales inhibidores	68
10.4	Artículo publicado con parte de los resultados del estudio in silico de DNMT	70

## Abreviaturas y acrónimos

ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
C	Carbono
IC <sub>50</sub>	Concentración inhibitoria media máxima
DNMT	ADN metiltransferasa
ECFP6	<i>Extended Connectivity Fingerprint 6</i>
Fig.	Figura
MOE	<i>Molecular Operating Environment</i>
PC	<i>Principal Component</i>
PCA	<i>Principal Component Analysis</i>
PLB	<i>Propensity for Ligand Binding</i>
PME	<i>Particle Mesh Ewald</i>
TPSA	<i>Topological Polar Surface Area</i>

## 1. Introducción

La familia de las ADN metiltransferasas (DNMT, por las siglas en inglés) comprende a una serie de enzimas que tienen un papel central en la regulación epigenética del ADN (Lyko 2018). Principalmente se encargan de regular la expresión de genes a través de la metilación de ciertas regiones del ADN (Lyko 2018). La familia de las DNMT comprende principalmente a DNMT1, DNMT2, DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMT $\Delta$ 3B1, DNMT $\Delta$ 3B2, DNMT $\Delta$ 3B3, DNMT $\Delta$ 3B4, y DNMT3L (Wang et al. 2006). Se observa la pérdida de los patrones de metilación normales del ADN en distintas patologías, entre ellas ciertos tipos de cáncer (Choi et al. 2011).

La inhibición de las enzimas DNMT es una estrategia utilizada como terapia para el tratamiento de ciertos tipos de tumores. Actualmente, existen solo dos fármacos inhibidores de DNMTs, azacitadina y decitabina. Sin embargo, su uso es restringido debido a su toxicidad y estabilidad química deficiente (Gnyszka et al. 2013).

Es conveniente conocer la estructura tridimensional del dominio catalítico de estas enzimas para ayudar al diseño racional de inhibidores hacia las distintas isoformas. Actualmente, únicamente se conoce la estructura cristalográfica de DNMT1, DNMT2, DNMT3A1 y DNMT3L cuya información se resume en la **Tabla 1**. Por lo tanto, es complicado hacer un diseño sistemático de inhibidores selectivos hacia el resto de las isoformas cuya estructura actualmente no se encuentra reportada.

**Tabla 1. Isoformas de DNMT cuya estructura cristalográfica es conocida.\***

Enzima	Código PDB	Resolución (Å)	Longitud (aminoácidos)	Referencia
DNMT1	4WXX	2.622	1616	(Zhang et al. 2015)
DNMT2	1G55	1.8	391	(Dong et al. 2001)
DNMT3A1	5YX2	2.653	912	(Zhang et al. 2018)
DNMT3L	5YX2	2.653	386	(Zhang et al. 2018)

\*Todas las estructuras corresponden a DNMT de humano y se encuentran co-cristalizadas con S-Adenosil-L-homocisteína (SAH), a excepción de DNMT3L, la cual no se encuentra co-cristalizada con SAH, ya que no posee un dominio catalítico, más bien, se encuentra formando un trímero con DNMT3A. Para cada isoforma se muestra únicamente una estructura representativa.

La difracción de Rayos x, la Resonancia Magnética Nuclear (RMN) y la Criomicroscopía Electrónica son técnicas mediante las cuales se puede obtener la estructura de una macromolécula (Anderson 2003). Sin embargo, cuando no se tiene acceso a la información experimental de las coordenadas tridimensionales de las estructuras, el modelado por homología permite predecir de

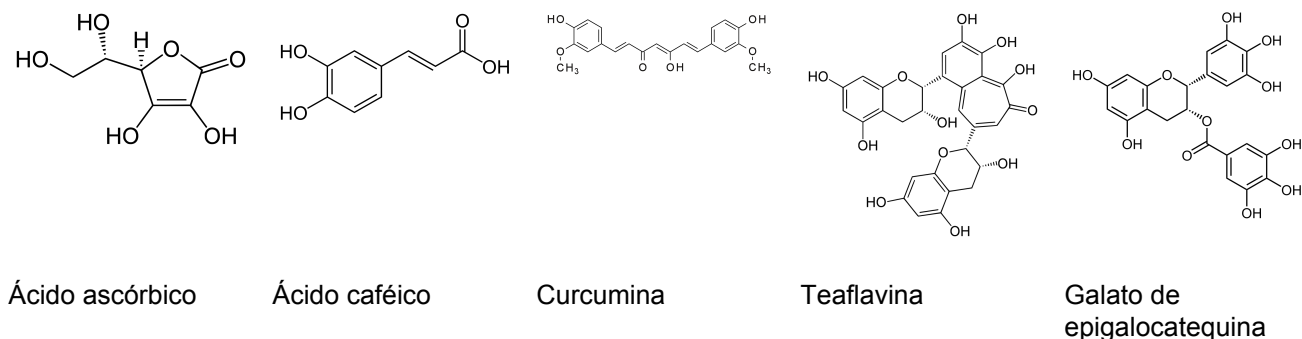
manera *in silico* la estructura tridimensional aproximada de una proteína (Cavasotto and Phatak 2009). Los algoritmos que utiliza el modelado por homología pueden dividirse en dos tipos: los que predicen la estructura utilizando como plantilla otra proteína con un alto grado de identidad y los algoritmos *de novo*, es decir, los que predicen la estructura de la proteína sin utilizar a otra proteína como plantilla (Kim et al. 2004). También existen algoritmos duales que combinan a los dos algoritmos anteriores, siendo estos los que tienen mayor exactitud en sus predicciones (Kim et al. 2004). Actualmente, los algoritmos que utilizan como plantilla otra proteína y los algoritmos duales, han logrado hasta un 78 % de exactitud en la predicción de la estructura de una proteína (di Luccio and Koehl 2011). Por lo tanto el modelado por homología es considerado como una opción para predecir la estructura de una proteína cuya estructura no ha sido publicada.

Se ha reportado el uso de modelado por homología para la búsqueda de inhibidores de DNMT1 (Yoo and Medina-Franco 2011) (Siedlecki et al. 2006) (Aldawsari et al. 2016) y DNMT3B (Kuck, Singh, et al. 2010). Además, con estos modelos por medio de estudios de cribado virtual, se ha intentado explicar las interacciones ligando-receptor de inhibidores de DNMT. Sin embargo, actualmente la estructura cristalográfica de DNMT1 ya se encuentra reportada (Zhang et al. 2015), así como la estructura de DNMT3A de humano en su conformación catalíticamente activa (Zhang et al. 2018), la cual es una plantilla mejor para la construcción de un modelo por homología de las isoformas de DNMT3B. Ya que el anterior modelo por homología de DNMT3B (Kuck, Caulfield, et al. 2010) fue hecho a partir de una estructura de DNMT3A de menor resolución (código PDB: 2QRV) (Jia et al. 2007).

Aún no se reportan las estructuras tridimensionales de las isoformas DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4. Sin embargo, por lo discutido anteriormente, conocer la estructura del dominio catalítico de estas isoformas es conveniente si se pretende hacer un diseño basado en la estructura de inhibidores de estas enzimas. El modelado por homología es una opción para poder predecir su dominio catalítico.

Aparte de azacitadina y decitabina que ya fueron aprobadas para su uso clínico, existen otras moléculas que han sido reportadas como inhibidores de DNMT, entre las cuales se encuentran compuestos provenientes de productos naturales, síntesis orgánica, cribado de alto rendimiento (del inglés *high-throughput screening*) y fármacos aprobados para otras indicaciones (Guianvarc'h and Arimondo 2014; Erdmann et al. 2015; Erdmann et al. 2016). Además, existe evidencia de que algunos de los nutrientes de nuestra dieta pueden inhibir la actividad de DNMT (Braicu et al. 2017). Cabe resaltar que la nutrigenómica es un área de investigación emergente, la cual se encarga de estudiar el efecto de los nutrientes en la salud humana a través de modificaciones epigenéticas (Braicu et al. 2017). En este contexto, ácido ascórbico (Young et al. 2015), ácido caféico (Lee and

Zhu 2006), curcumina (Maugeri et al. 2018), teaflavina (Rajavelu et al. 2011) y galato de epigallocatequina (EGCG) (Lee et al. 2005), son compuestos inhibidores de DNMT que se encuentran en la dieta **Fig. 1**.



**Fig. 1.** Inhibidores de DNMT que se encuentran en los alimentos.

## 2. Objetivos

### General

Identificar potenciales inhibidores de DNMT1, DNMT2, DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4.

### Específicos

1. Obtener el modelo tridimensional del dominio catalítico de las isoformas de DNMT catalíticamente activas cuya estructura no está reportada (DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4).
2. Determinar si existen sitios con afinidad por moléculas “tipo-fármaco” (*drug-like*), alternativos a la cavidad donde se lleva a cabo la metilación del ADN, los cuales puedan ser utilizados en un futuro para el diseño de inhibidores de DNMT.
3. De una base de datos pública de compuestos encontrados en los alimentos, determinar cuáles son los compuestos con propiedades “tipo-fármaco” con mayor afinidad por el sitio donde se lleva a cabo la metilación del ADN en las isoformas catalíticamente activas de DNMT.
4. De la base de datos pública de compuestos encontrados en los alimentos, determinar cuál es el espacio químico de los compuestos “tipo-fármaco” con mayor afinidad al sitio donde se lleva a cabo la metilación del ADN hacia cada isoforma de DNMT.

### 3. Hipótesis

1. Conociendo la secuencia primaria de aminoácidos del dominio catalítico de cada isoforma de DNMT es posible construir su respectivo modelo por homología.
2. En las enzimas DNMT existen sitios donde pueden ocurrir modificaciones postraduccionales, interacciones con otras proteínas o cambios conformacionales, cuya interacción con un ligando impedirá de manera indirecta que se lleve a cabo la metilación del ADN.
3. La base con compuestos encontrados en los alimentos, contiene compuestos con propiedades “tipo-fármaco” afines al sitio donde se lleva a cabo la metilación en las enzimas DNMT, que pueden ser utilizadas como prototipo para el desarrollo de inhibidores de DNMT.
4. El espacio químico de los compuestos más afines a cada isoforma de DNMT es definido y delimitado.

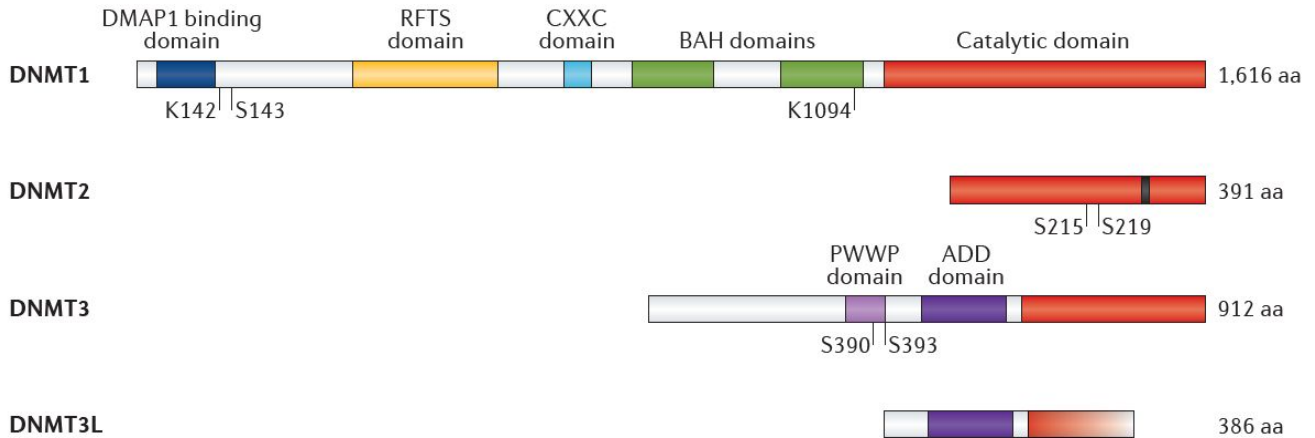
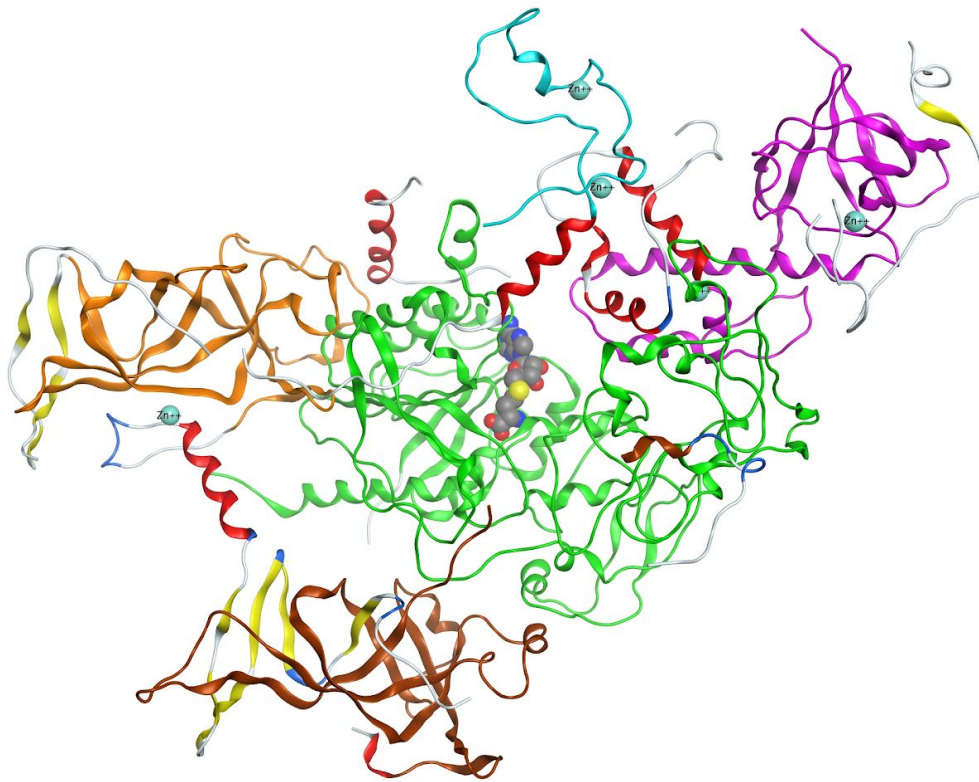
### 4. Antecedentes

#### 4.1 La familia de las DNMT

##### 4.1.1 Función biológica y estructura

La metilación del ADN juega un papel fundamental en la regulación de la expresión de genes (Lyko 2018). La familia de las DNMT se encarga de establecer y mantener los patrones de metilación en el ADN (Lyko 2018). DNMT1 se encarga de mantener el patrón de metilación durante la replicación del ADN (Jurkowska et al. 2011). DNMT2 se encarga de metilar el ARN de transferencia, el cual necesita estar metilado para llevar a cabo su función normal (Jeltsch et al. 2017). DNMT3A y DNMT3B se encargan de llevar a cabo la metilación *de novo* en el ADN (Jurkowska et al. 2011) y establecen los patrones de metilación en las células germinales durante el temprano desarrollo de los mamíferos (Jurkowska et al. 2011). DNMT3L es catalíticamente inactiva, sin embargo, funciona como factor regulador en DNMT3A y DNMT3B, siendo fundamental para que ambas enzimas puedan llevar a cabo su actividad de metiltransferasa (Jurkowska et al. 2011).

La **Fig. 2.A** ilustra esquemáticamente a los dominios estructurales que posee cada isoforma de DNMT. Todas las isoformas de DNMT poseen un dominio catalítico, sin embargo, en algunas isoformas de DNMT3B y en DNMT3L éste se encuentra trunco (Lyko 2018). En DNMT1 y DNMT3 los dominios adicionales al catalítico tienen la función de reconocimiento y unión al ADN (Lyko 2018).

**A****B**

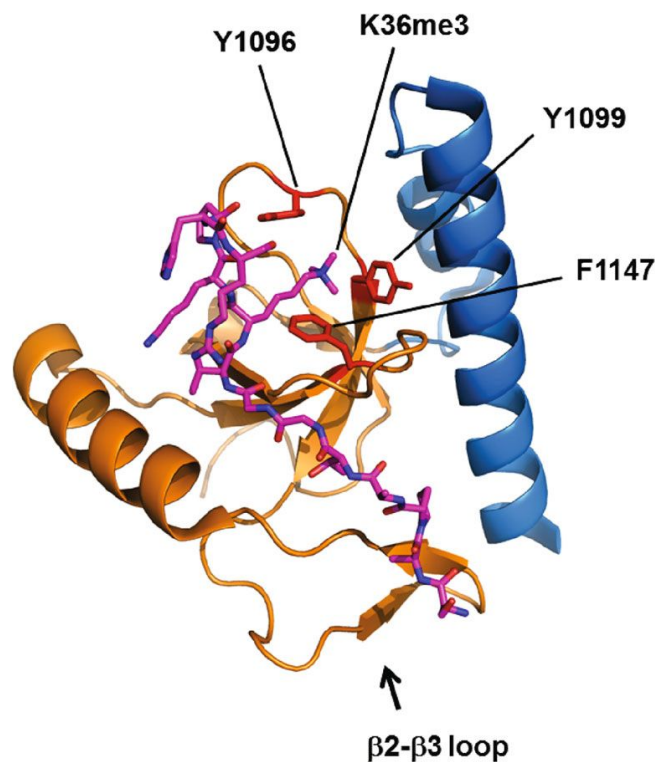
**Fig. 2. A.** Dominios estructurales de las isoformas de DNMT (Lyko 2018). **B.** Estructura de DNMT1: dominios BAH1 (naranja), BAH2 (café), CXX (azul claro), RFTS (magenta) y cofactor SAH representado con esferas (código PDB: 4WXX) (Zhang et al. 2015).

DNMT1 posee el dominio de unión a DMAP1, a través del cual se une a la proteína DMAP1, que es un represor transcripcional (Lyko 2018). El dominio RFTS (*Replication foci targeting sequence*) reconoce el foco de replicación del ADN y de esta manera DNMT1 es guiada hacia este sitio (Lyko 2018). El dominio CXX es un dominio de dedos de zinc el cual reconoce a los dinucleótidos no

metilados citosina y guanina en las islas CpG (Lyko 2018). En cuanto a los dominios BAH se ha propuesto que interactúan como moduladores de interacciones proteína-proteína (Lyko 2018).

En DNMT3 se encuentra el dominio PWWP (Pro-Trp-Trp-Pro), mediante el cual la enzima se une a moléculas de histona H3 trimetiladas en su residuo de lisina 36 (Rona et al. 2016). En este dominio se encuentra una pequeña cavidad conocida como la “caja aromática” (**Fig. 3**), la cual es fundamental para la unión de DNMT3 con la cromatina, ya que en ella se lleva a cabo el reconocimiento de moléculas de histona H3 trimetiladas en su residuo de lisina 36 (Rona et al. 2016). La caja aromática podría ser un sitio en el cual, mediante la inserción de un pequeño ligando, podría evitarse la unión ADN-DNMT3. Es por esto que esta región puede ser un sitio de unión alternativo al catalítico para un inhibidor de DNMT3. Sin embargo, hay que considerar que el dominio PWWP junto con la “caja aromática”, se encuentran en otras proteínas que interactúan con el ADN (Rona et al. 2016), por tanto, se podría anticipar que un inhibidor de DNMT3 dirigido hacia la “caja aromática”, interferirá con el funcionamiento normal de otras proteínas.

DNMT3 posee otro dominio de reconocimiento a moléculas de histona H3: el dominio ADD (ATRX-DNMT3-DNMT3L), ilustrado en la **Fig. 2**. Este es un dominio de dedos de zinc que reconoce a residuos de lisina que no se encuentran metilados.

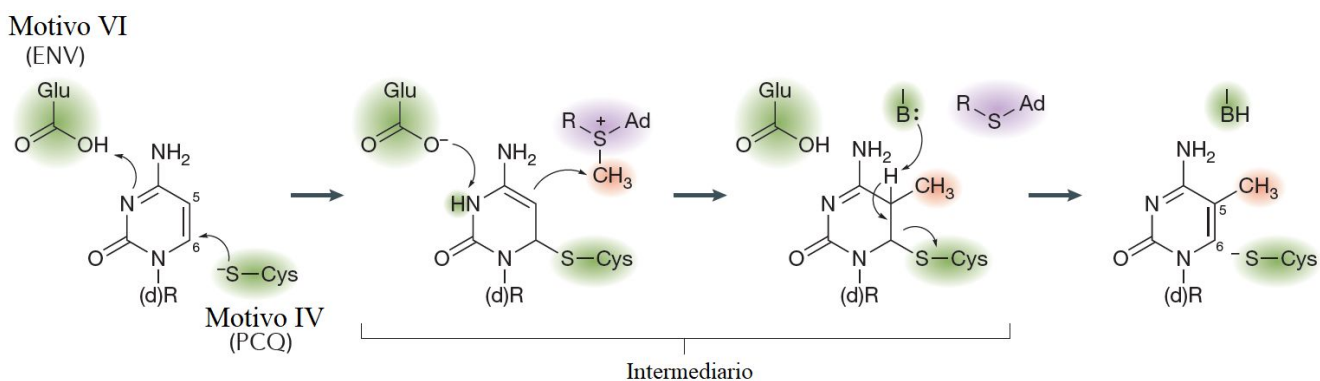


**Fig. 3.** Caja aromática del dominio PWWP de DNMT3A. En color rojo se observan los aminoácidos esenciales para el reconocimiento de la cromatina. En color rosa se observan los residuos que forman parte la histona trimetilada H3K36. Código PDB ID: 2X4W (Rona et al. 2016).



### 4.1.2 Mecanismo de acción

En DNMT1, DNMT2 y las isoformas catalíticamente activas de DNMT3, la reacción de metilación del ADN se lleva a cabo en el dominio catalítico, en una cavidad que contiene al cofactor S-Adenosil-L-Homocisteína (SAH), cuyo azufre se encuentra metilado (Klimasauskas et al. 1994). Se propone que el primer paso en la reacción de metilación del ADN es el ataque nucleofílico por parte de un residuo de cisteína que se encuentra conservado en el motivo IV, también conocido como motivo PCQ, hacia el C6 del anillo de citosina, el cual es facilitado por un residuo de ácido glutámico que se encuentra en el motivo VI, también conocido como motivo ENV (Klimasauskas et al. 1994). Posteriormente, el grupo metilo de SAH es transferido al C5 del anillo de citosina y subsecuentemente sufre una desprotonación, para dar como resultado la 5-metilcitosina (Klimasauskas et al. 1994). La **Fig. 4** ilustra al mecanismo.

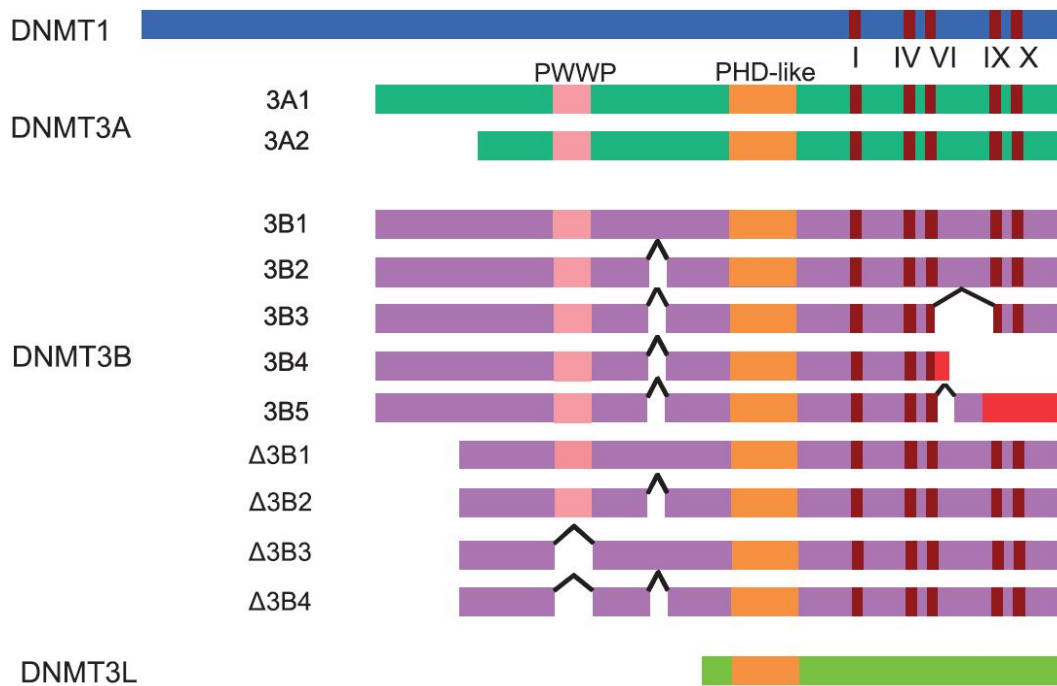


**Fig. 4.** Mecanismo de metilación de las DNMT. “B:” representa a un grupo funcional de carácter básico, cuya función es abstraer el protón del átomo de carbono 5, lo cual desencadena el movimiento de electrones que culmina con la escisión del enlace covalente entre el residuo de cisteína y el anillo de citosina. Se cree que “B:” es proveído por alguno de los aminoácidos de la cavidad catalítica.

### 4.1.3 Isoformas

Se han identificado a dos isoformas de DNMT3A: DNMT3A1 y DNMT3A2. En su dominio N-terminal ambas poseen el dominio PWWP y el dominio ADD (Jurkowska et al. 2011). El dominio C-terminal es idéntico en ambas isoformas (Choi et al. 2011). Existe un gran número de isoformas de DNMT3B, sin embargo, únicamente DNMT3B1 y DNMT3B2 son catalíticamente activas (Ostler et al. 2007). Similar a DNMT3A, DNMT3B1 y DNMT3B2 tienen un dominio de reconocimiento de histonas H3, PWWP y un dominio ADD en la región N-terminal. DNMT $\Delta$ 3B se diferencia de DNMT3B en que su región N-terminal se encuentra trunca debido a la ausencia de 200 aminoácidos (Lyko 2018). Las isoformas DNMT $\Delta$ 3B1-4 son catalíticamente activas, mientras que DNMT $\Delta$ 3B5-7 no poseen el dominio catalítico (Wang et al. 2006). DNMT $\Delta$ 3B se expresa principalmente durante el cáncer de pulmón de

células no pequeñas (Wang et al. 2006; Ostler et al. 2007). En la **Fig. 5** se esquematizan las diferencias entre las distintas isoformas.



**Fig. 5.** Comparación de la secuencia de las isoformas de DNMT (Choi et al. 2011). Se observan las isoformas de DNMT3B que no tienen el sitio activo trunco: DNMT3B1 y DNMT3B2.

## 4.2 La búsqueda de inhibidores de DNMT

### 4.2.1 Patologías que involucran a las DNMT

El síndrome mielodisplásico engloba a un grupo de enfermedades que afectan a las células inmaduras de la médula ósea, las cuales pierden su capacidad para madurar (Mohammad 2018). Como consecuencia se presenta citopenia (nivel inferior al normal de células sanguíneas maduras) (Zeidan et al. 2019). El mayor riesgo es que la enfermedad evolucione a una leucemia (Fozza 2019). El síndrome mielodisplásico tiende a evolucionar a leucemia mielomonocítica crónica o leucemia mieloide aguda (Mohammad 2018).

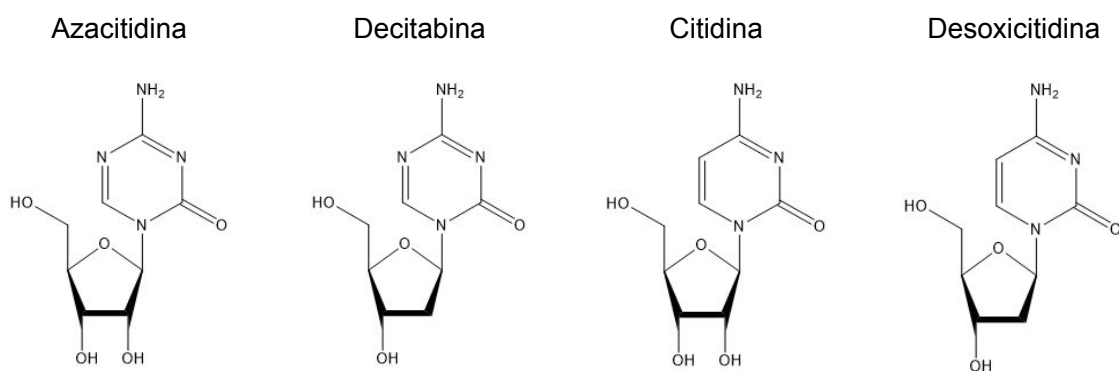
Los dos fármacos inhibidores de DNMT que actualmente se encuentran aprobados para uso clínico tienen como indicación el tratamiento del síndrome mielodisplásico. Sin embargo, existen otros tipos de cáncer en los cuales existe hipermetilación de genes supresores de tumores: cáncer de pulmón, nasofaríngeo, colorrectal y de mama (Zhang et al. 2014).

La sobreexpresión de DNMT se observa en distintos tipos de cáncer. Como consecuencia se tiene una hipermetilación del ADN (Gordon et al. 2013). Esta se lleva a cabo a nivel de la región

promotora, específicamente en las islas CpG, lo cual tiene como consecuencia el silenciamiento de genes, en especial los que son supresores de tumores, como es el caso de p53 y DAB2 (Melo et al. 2014; Wei et al. 2016; Zhang et al. 2014).

#### 4.2.2 Inhibidores aprobados actualmente para uso clínico

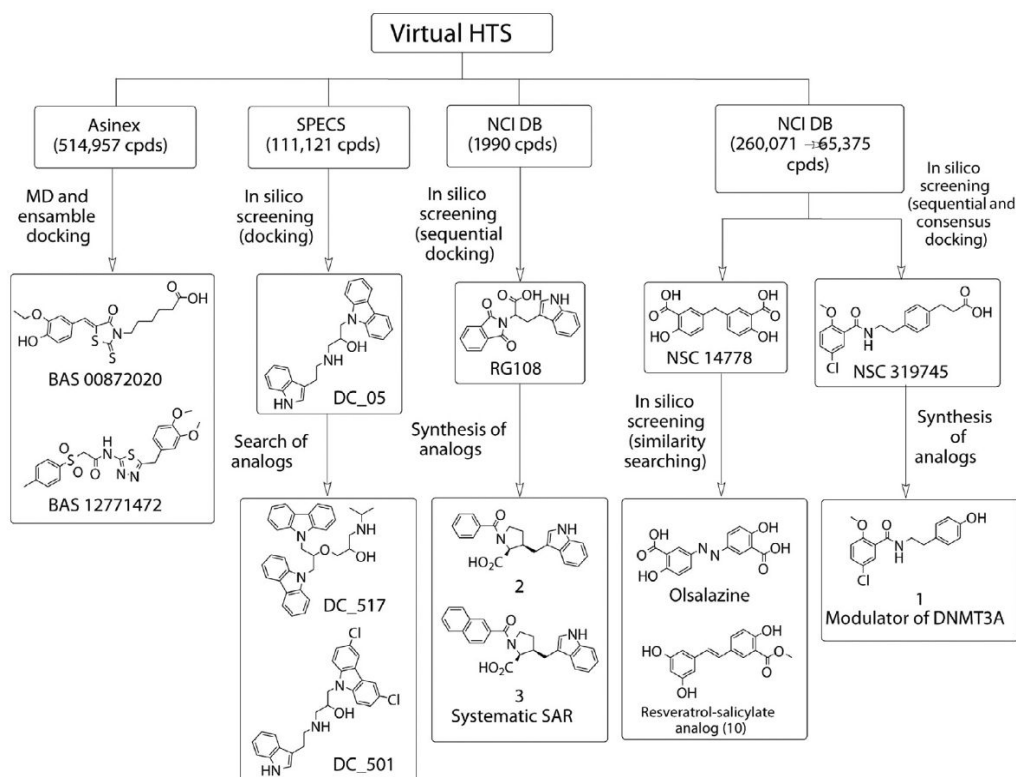
Actualmente, están aprobados dos inhibidores de DNMT para su uso clínico, azacitidina y decitabina (Estey 2013). La indicación terapéutica de ambos es el tratamiento de cualquiera de las enfermedades que engloba el síndrome mielodisplásico: anemia refractaria, anemia refractaria con sideroblastos en anillo, anemia refractaria con exceso de blastos y leucemia mieloide aguda con displasia de multilineaje. Está reportado que ambos fármacos se unen a DNMT1 (Estey 2013). La azacitidina y la decitabina son análogas de su respectivo nucleósido: azacitidina es un análogo de la citidina y decitabina es un análogo de la desoxicitidina. La **Fig. 6** muestra que la única diferencia entre ambos fármacos y su respectivo nucleósido es un nitrógeno que se encuentra dentro del anillo de pirimidina. La azacitidina tiene la capacidad de incorporarse tanto al ARN como al ADN, pero a este último en menor medida (Borodovsky et al. 2013). La decitabina tiene únicamente la capacidad de incorporarse al ADN (Kantarjian et al. 2006). La capacidad de poder incorporarse al ADN y/o al ARN es el motivo por el cual ambas moléculas son relativamente tóxicas, ya que esto tiene como consecuencia la inhibición de síntesis de proteínas y la síntesis de ADN lo cual termina causando muerte celular no sólo en células malignas (Kaminskas et al. 2005). En principio, el desarrollo de inhibidores de DNMT no nucleosídicos tendría como ventaja sobre estos dos fármacos, menor toxicidad debido a que se evitaría la incorporación al ADN y/o ARN.



**Fig. 6.** Azacitidina y decitabina (los dos inhibidores de DNMT aprobados para su uso clínico), citidina y desoxicitidina, nucleósidos que son sintetizados y posteriormente incorporados al ADN o al ARN.

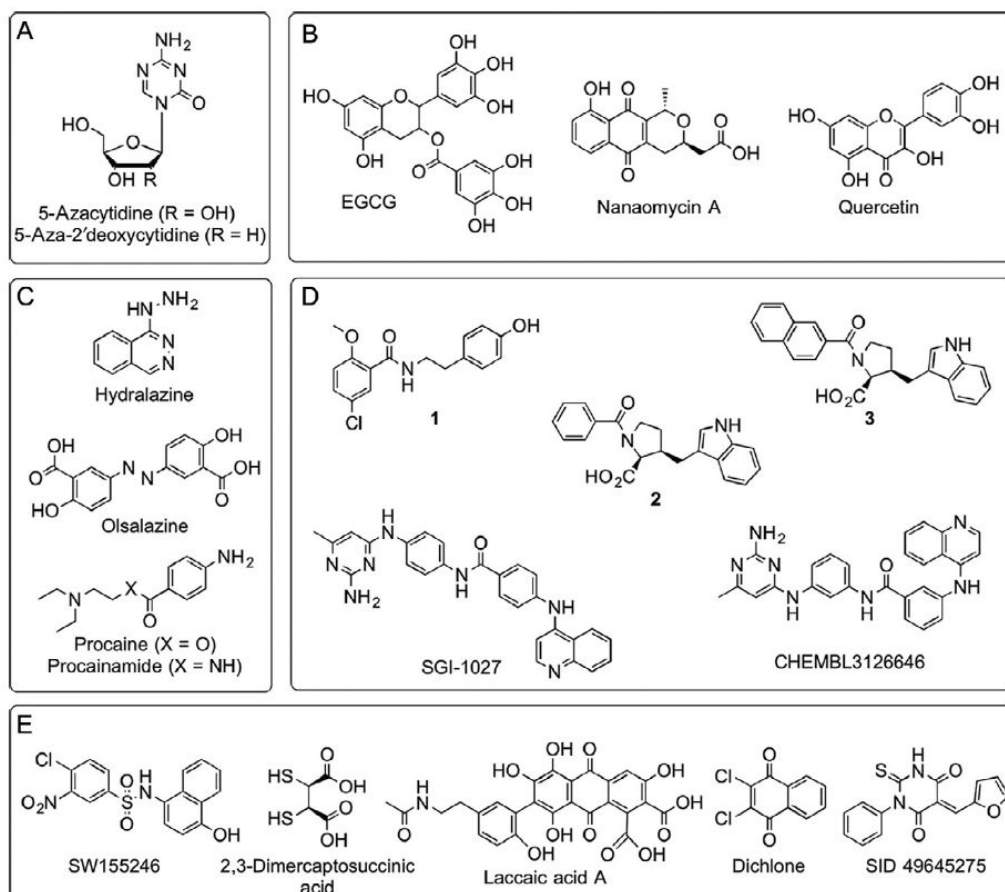
### 4.2.3 Inhibidores en desarrollo, compuestos cabezas de serie (*hits*)

En los últimos años se han identificado compuestos inhibidores de DNMT, algunos de ellos han provenido inicialmente de estudios de cribado virtual **Fig. 7** (Medina-Franco et al. 2015). En uno de estos estudios se identificaron a dos inhibidores de DNMT1: BAS 12771472 y BAS 00872020 con  $CI_{50}$  de 15  $\mu$ M y 18  $\mu$ M, respectivamente (Joshi et al. 2016). Por su parte, DC\_517 y DC\_501 fueron identificados como inhibidores de DNMT1, con valores de  $CI_{50}$  de 2.5  $\mu$ M y 1.7  $\mu$ M, respectivamente (Chen et al. 2014). También se han reportado compuestos con actividad inhibitoria de DNMT3A y DNMT3B. Un ejemplo es el análogo del resveratrol-salicilato, que fue etiquetado en el estudio original como compuesto **10** cuya  $CI_{50}$  contra DNMT3A y DNMT3B fue 25  $\mu$ M y 62  $\mu$ M, respectivamente (Aldawsari et al. 2016). Cabe resaltar que hoy en día hay más información cristalográfica de DNMT para hacer estudios de cribado virtual. En el cribado virtual que llevó a la obtención de DC\_517 y DC\_501 (Chen et al. 2014), se utilizó la estructura cristalográfica del dominio catalítico de DNMT1 de ratón. No obstante, la estructura cristalográfica de DNMT1 de humano ya se encuentra disponible. En el cribado virtual que llevó a la identificación del análogo de resveratrol-salicilato (Aldawsari et al. 2016), se utilizó un modelo por homología de DNMT3A. Sin embargo, actualmente se conoce la estructura cristalina del dominio catalítico de DNMT3A.



**Fig. 7.** Inhibidores de DNMT provenientes inicialmente de estudios de cribado virtual (Medina-Franco et al. 2015).

La **Fig. 8** ilustra a otros inhibidores de DNMT, los cuales no provienen de estudios de cribado virtual.



**Fig. 8.** Inhibidores de DNMT representativos (A) aprobados para uso clínico, (B) provenientes de productos naturales, (C) fármacos aprobados para otras indicaciones, (D) compuestos sintéticos provenientes de programas de optimización y (E) provenientes de estudios de cribado virtual de alto rendimiento (Erdmann et al. 2016; Erdmann et al. 2015; Guianvarc'h and Arimondo 2014). Actualmente, **SGI-1027** es el inhibidor que suele ser utilizado como referencia en los estudios preclínicos, debido a su mayor potencia sobre el resto de los inhibidores (Datta et al. 2009).

### 4.3 Inhibidores de DNMT en los alimentos

Se han identificado compuestos que se encuentran en nuestra dieta como inhibidores de DNMT, tal es el caso del ácido ascórbico (Young et al. 2015), el ácido caféico (Lee and Zhu 2006), la curcumina (Maugeri et al. 2018), la teaflavina (Rajavelu et al. 2011) y el galato de epigallocatequina -EGCG (Lee et al. 2005). Por lo tanto, los compuestos que forman parte de la dieta son una fuente de inhibidores de DNMT, los cuales pueden ser utilizados como punto de partida para programas de optimización de Química Farmacéutica. Actualmente, es posible acceder a bases de datos que contienen compuestos que se encuentran en los alimentos (Naveja et al. 2018) y por tanto llevar a cabo estudios de cribado virtual que lleven a la identificación de nuevos inhibidores de DNMT provenientes de los alimentos.

Existe evidencia de que los factores ambientales y la nutrición juegan un papel muy importante en el establecimiento de patrones anormales de metilación en el ADN (Prieto-Martínez et al. 2016). Por lo tanto, se considera que el consumo regular de inhibidores de DNMT presentes en los alimentos juega un papel quimiopreventivo (Prieto-Martínez et al. 2016) contra las patologías mencionadas en el apartado “4.2.1 Patologías que involucran a las DNMT”.

## 5. Procedimiento experimental

### 5.1 Modelado por homología

#### 5.1.1 Identificación del dominio catalítico

A partir de la secuencia primaria de aminoácidos de las isoformas de DNMT1, DNMT2, DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, se identificaron a los aminoácidos que conforman el dominio catalítico. Esta información se encuentra en la base de datos de “UniProt” (UniProt Consortium 2019).

#### 5.1.2 Alineamiento de las secuencias y construcción de la matriz de identidad

Se alineó la secuencia primaria de aminoácidos del dominio catalítico de las isoformas en el servidor de “Clustal Omega” versión 1.2.4 (Madeira et al. 2019). El alineamiento se encuentra en la sección correspondiente al Apéndice: “**10.1 Alineamiento de la secuencia primaria de aminoácidos del dominio catalítico**” Posteriormente, se construyó la matriz de identidad con el programa MOE *Molecular Operating Environment*, versión 2018.01 (Chemical Computing Group ULC 2018) (cf. **Fig. 10**).

#### 5.1.3 Construcción del modelo

Las isoformas cuya estructura cristalográfica no ha sido reportada son: DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4. En la matriz de identidad (**Fig. 10**), la cuál se construyó únicamente con las secuencias correspondientes al dominio catalítico, se observó que DNMT3A1 y DNMT3A2 poseen un 100 % de identidad, lo cual significa que poseen la misma secuencia primaria de aminoácidos. Por lo tanto, se esperaría que DNMT3A2 posea la misma estructura terciaria de DNMT3A1 a nivel del dominio catalítico, la cual está reportada. Por lo anterior, se determinó que no es necesario llevar a cabo la construcción del modelo por homología del dominio catalítico de DNMT3A2.

En los alineamientos de las secuencias se observó que DNMT3B3, DNMT3B4 y DNMT3B5 se encuentran truncas a nivel del dominio catalítico. Después de una revisión bibliográfica se confirmó que son isoformas catalíticamente inactivas (Choi et al. 2011). Debido a que se pretende utilizar los modelos por homología para el diseño de inhibidores de DNMT, se determinó que no tendría relevancia hacer el modelado por homología del dominio catalítico de estas isoformas. Por tanto, se llevaría a cabo únicamente el modelado por homología de las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4, las cuales son catalíticamente activas. A partir de la matriz de identidad (**Fig. 10**) se observó que a nivel del dominio catalítico poseen un 100 % de identidad, por lo cual el modelo construido representa a estas seis isoformas de DNMT.

Para seleccionar la plantilla a partir de la cual se construiría el modelo por homología, se buscaron en la base de datos de Protein Data Bank (Berman et al. 2000) las estructuras reportadas de DNMT3A1 de humano. Debido a que el dominio catalítico de DNMT3A1 tiene el mayor porcentaje de identidad con respecto al dominio catalítico de las seis isoformas catalíticamente activas de DNMT3B y DNMTΔ3B (80.6 % de identidad), se seleccionó como plantilla la estructura de DNMT3A1 con mejor resolución (código PDB ID: 5YX2).

La construcción del modelo por homología se hizo con el programa YASARA, versión 18.4.24, el cual utiliza el algoritmo de “Discriminación de clase de estructura secundaria” (Krieger et al. 2009). El modelo por homología fue afinado con el “*script*” que posee YASARA para afinar modelos por homología (Krieger et al. 2004), el cual consiste en hacer una minimización energética del sistema y posteriormente una simulación de dinámica molecular con disolvente explícito. La minimización energética es llevada a cabo con dos distintos algoritmos, gradiente descendiente y recocido simulado en presencia de disolvente explícito. Este “*script*” utiliza el campo de fuerza YASARA2 (Krieger et al. 2009), el cual está parametrizado para la optimización estructural de modelos proteicos. A continuación se describe cómo el “*script*” lleva a cabo la simulación del sistema.

El modelo se prepara en una caja de 15 Å alrededor de la proteína. Posteriormente, se añaden moléculas de agua, usando el modelo de potencial interatómico transferible de tres sitios (TIP3P). Se hace la predicción del  $pK_a$  de los aminoácidos para ajustar su estado de protonación al pH del medio que es de 7.4. Se añaden iones NaCl a la concentración fisiológica de 0.9%, con un exceso de iones  $Na^{+1}$  o  $Cl^{-1}$  para neutralizar la celda. La densidad es de 0.997 g/mL, la temperatura 298 K y la presión de 1 bar (ensamble NPT). La temperatura es ajustada mediante el termostato de Berendsen. La presión controlada mediante un densostato (Krieger and Vriend 2015). El tiempo de simulación es de 500 ps con un intervalo de integración múltiple de 1.0 fs para interacciones de enlace y 2.0 fs para las interacciones débiles. Las interacciones de Van der Waals son truncadas a los 8 Å, las interacciones electrostáticas no son truncadas y se calculan mediante el algoritmo PME (Essmann et al. 1995).

#### 5.1.4 Validación del modelo

La validación se llevó a cabo con el programa MOE versión 2018.01. Se generaron cuatro diagramas de Ramachandran distintos, el general, para residuos de glicina, prolina y pre-prolina (residuo que se encuentra antes de la prolina) **Fig. 14**. Posteriormente, se determinó la calidad estructural de la cadena principal de aminoácidos. Para ello se utilizaron sus longitudes de enlace y ángulos y a partir de estos se calculó el puntaje Z (**Fig. 15**, **Fig. 16** y **Fig. 17**). El puntaje Z puede ser interpretado como el número de desviaciones estándar respecto del valor esperado y se determina con la ecuación (1):

$$Z = \frac{x - \mu}{\sigma}$$

$x$ =Valor medido  
 $\mu$ =Valor medio  
 $\sigma$ =Desviación estándar

(1)

Los valores de  $\mu$  y  $\sigma$  se encuentran en la **Fig. 9**. Estos forman parte de la base de datos del programa MOE. Fueron determinados a partir de un grupo seleccionado de proteínas, las cuales poseen una resolución igual o menor a 1.3 Å,  $R_{\text{free}}$  igual o menor a 0.25, factor B de los átomos de la cadena principal igual o menor a 30 y un factor de ocupación del 100% para todos los átomos de la cadena principal. Los valores de Z mayores a 4 son considerados puntos atípicos (“*outliers*”), ya que son muy poco frecuentes en las estructuras de Protein Data Bank.

La calidad estructural de la cadena lateral de aminoácidos se evaluó mediante la energía asociada a cada rotámero. Las distintas combinaciones de ángulos phi-psi de cada aminoácido en la cadena principal tienen asociadas un determinado rotámero en la cadena lateral y cada rotámero tiene asociada una determinada energía. El programa MOE posee una base de datos de rotámeros, la cual fue hecha a partir de estructuras de Protein Data Bank, cuya resolución es mayor a 2 Å. Esta base de datos, para cada aminoácido contiene los rotámeros que se encuentran asociados a las distintas combinaciones phi-psi, así como su respectiva energía. De este modo, para evaluar la calidad estructural de la cadena lateral se compara la energía que posee cada rotámero del modelo por homología, contra la energía esperada. Para facilitar la visualización de la información, los valores de energía son normalizados, de tal manera que la energía esperada siempre será cero. Se graficó el valor de la energía asociada a cada rotámero del modelo por homología **Fig. 18**. Los valores de energía mayores a 5 (kcal/mol) son considerados como puntos atípicos debido a que son poco frecuentes en las estructuras de Protein Data Bank.

Por último, se determinó si había repulsión estérica entre los átomos del modelo **Fig. 19**. Para ello cada átomo fue considerado como una esfera de Van der Waals: cada esfera con su respectivo radio



de Van der Waals, dependiendo del tipo de átomo. La función para evaluar la energía entre dos esferas de Van der Waals es la siguiente:

$$E = w800(1 - r/R)^3$$

$E$ =Energía entre 2 esferas de Van der Waals

$w$ = Constante que depende de los 2 tipos de átomos involucrados

$r$ =Distancia interatómica entre las dos esferas

$R$ =Suma de los radios de Van der Waals de ambas esferas (2)

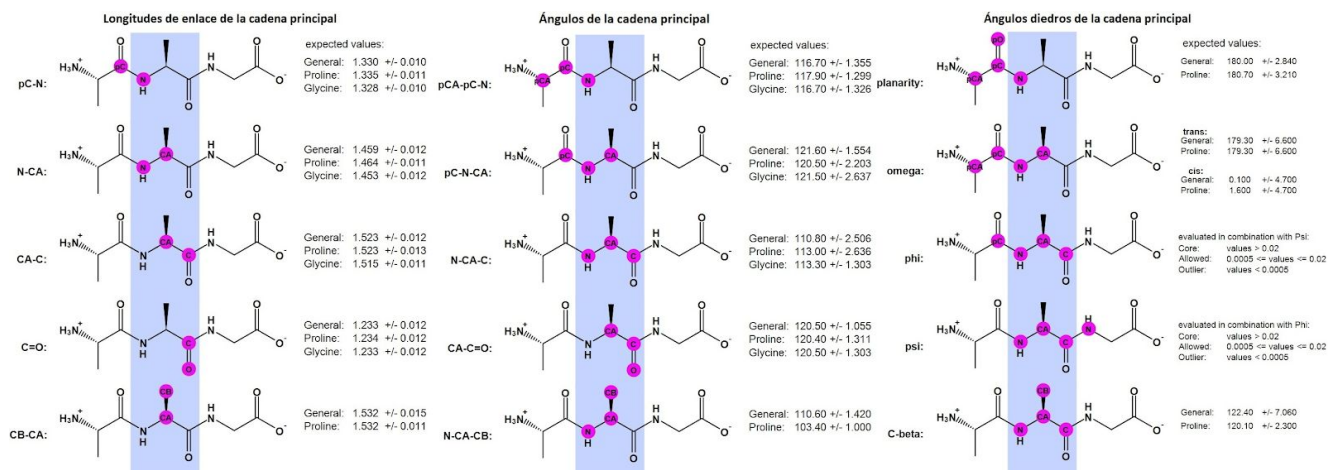


Fig. 9. Definiciones de las longitudes de enlace y ángulos de la cadena principal de aminoácidos.

## 5.2 Búsqueda de sitios de unión alternos al dominio catalítico

Se utilizó la aplicación “Site Finder” (Le Guilloux et al. 2009) del programa MOE versión 2018.01, para identificar potenciales sitios alostéricos. “Site Finder” identifica potenciales sitios de unión para moléculas con propiedades “tipo-fármaco”. Esta aplicación permea el espacio alrededor de la proteína con esferas alfa (esfera que no posee átomos en su interior y se encuentra en contacto con cuatro átomos en su perímetro, los cuales se encuentran equidistantes al centro de la esfera) mediante la triangulación de Delaunay. En la triangulación de Delaunay se genera una red de triángulos interconectados por sus vértices, en cada triángulo es circunscrita una circunferencia (esfera alfa) la cual no debe contener el vértice de otro triángulo. Posteriormente son eliminadas las esferas alfa que se encuentran en regiones inaccesibles para moléculas “tipo-fármaco” y en áreas muy expuestas al disolvente. Las esferas alfa que se encuentran en un sitio en el cual son susceptibles de ser aceptoras o donadoras de puentes de hidrógeno serán clasificadas como hidrofílicas, las demás esferas son clasificadas como hidrofóbicas. Las esferas hidrofílicas que no se encuentran adyacentes a una esfera hidrofóbica son eliminadas, ya que estos sitios suelen

corresponder únicamente a sitios de unión a moléculas de agua. Por último, en los sitios identificados se calcula el valor numérico de *Propensity for Ligand Binding* (PLB, por las siglas en inglés) (Soga et al. 2007), el cual es una medida cuantitativa de la probabilidad de que una molécula “tipo-fármaco” se una en el sitio.

Se llevó a cabo la búsqueda de sitios alostéricos en DNMT1 (PDB ID: 4WXX), DNMT2 (PDB ID: 1G55), DNMT3A1 (PDB ID: 5YX2) y el modelo por homología. Las proteínas utilizadas fueron previamente preparadas como se indica en la sección **5.3.3 Acoplamiento molecular**. Por último, se registraron los aminoácidos que conforman cada sitio alostérico identificado, se encuentran en el Apéndice: **10.2 Aminoácidos que conforman el sitio donde se lleva a cabo la metilación del ADN y otros potenciales sitios de unión para moléculas “tipo-fármaco”, Tabla S1**. Utilizando la base de datos “Uniprot” se determinó en qué dominio funcional se encuentran los sitios identificados y si existen modificaciones postraduccionales en estos sitios.

## **5.3 Cribado virtual**

### **5.3.1 Curado**

Se utilizó la base de datos FooDB (The Metabolomics Innovation Centre 2017), la cual está conformada por 23 883 compuestos que se encuentran en los alimentos. La curación de la base de datos se llevó a cabo con el programa MOE versión 2018.01. El proceso de curado consistió en la eliminación de sales, conservando únicamente el fragmento más grande, el cálculo del número de átomos de hidrógeno implícitos para cada átomo y posteriormente la adición de los átomos de hidrógeno faltantes. El estado de protonación de cada compuesto se ajustó al pH fisiológico de 7.4.

### **5.3.2 Filtrado**

En el programa KNIME versión 3.7.1, con los nodos de RDKit, fueron calculadas seis propiedades fisicoquímicas de interés farmacéutico: SlogP, área topológica polar superficial (TPSA, por sus siglas en inglés), masa molecular, número de átomos donadores de puente de hidrógeno, número de átomos aceptores de puente de hidrógeno y número de enlaces rotables. Se seleccionaron únicamente los compuestos cuyas propiedades fisicoquímicas son de “tipo-fármaco” (**Tabla 2**). Es decir, los compuestos que cumplían la regla de cinco de Lipinski (Lipinski et al. 2001), la extensión de la regla de cinco o reglas de Veber (Veber et al. 2002), la regla 3/75 de Pfizer (Hughes et al. 2008) y la regla 4/400 de GlaxoSmithKline (Gleeson 2008).

**Tabla 2. Propiedades fisicoquímicas asociadas a los compuestos “tipo-fármaco.”**

	Masa molecular	LogP	Número de átomos donadores de puente de hidrógeno	Número de átomos aceptores de puente de hidrógeno	Suma de número de átomos donadores y aceptores de puente de hidrógeno	Área topológica polar superficial	Número de enlaces rotables
Regla de cinco de Lipinski	≤500	≤5	≤5	≤10			
Reglas de Veber					≤12	≤140	≤10
Regla 3/75 de Pfizer		>3				<75	
Regla 4/400 de GlaxoSmithKline	<400	<4					

### 5.3.3 Acoplamiento molecular

En el Protein Data Bank se identificaron las estructuras de DNMT1, DNMT2 y DNMT3A1 que tuvieran el mayor grado de resolución y valores de R y  $R_{free}$  menores o iguales a 0.2 y 0.25 respectivamente. Para las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4, se utilizó el modelo por homología construido.

La preparación de las proteínas se hizo con el módulo de preparación de proteínas del programa MOE versión 2018.01. La preparación de las proteínas consistió en la adición de todos los átomos faltantes, incluyendo los átomos de hidrógeno, el estado de protonación se ajustó al pH fisiológico de 7.4 y se construyeron los bucles faltantes. El ligando co-cristalizado SAH, se preservó en todas las estructuras. Se llevó a cabo la minimización energética con el campo de fuerza Amber 14:EHT, el cual utiliza los parámetros del campo de fuerza Amber 14 (Maier et al. 2015) para las proteínas y los parámetros de la Teoría Extendida de Hückel “EHT por sus siglas en inglés” (Gerber and Müller 1995) para los ligandos. Los ligandos que fueron seleccionados como “tipo-fármaco” y sus estructuras fueron minimizadas energéticamente con el campo de fuerza Amber 14:EHT.

El estudio de acoplamiento molecular se hizo con dos programas distintos: Ledock (Zhang and Zhao 2016) y MOE, versión 2018.01 (Chemical Computing Group ULC 2018). Ledock es un programa de acceso libre el cual ha mostrado recientemente un desempeño superior al de programas gratuitos como rDock, AutoDock Vina, y AutoDock (Pagadala et al. 2017). Se utilizaron dos programas diferentes con dos distintos algoritmos de búsqueda con el propósito de aumentar la variedad de conformeros generados de cada ligando. Ledock utiliza el algoritmo de recocido simulado, el cual genera inicialmente un modo de unión (*pose*) aleatoria del ligando, a partir de la

cual se generará la vecindad de confórmeros en busca del que posea el  $\Delta G$  de unión más favorable. Sin embargo, durante las primeras iteraciones la generación de confórmeros no siempre se moverá en busca del  $\Delta G$  de unión más favorable, sino que puede moverse hacia la generación de confórmeros con un  $\Delta G$  de unión menos favorable. Esto con el objetivo de expandir la región de búsqueda en el espacio conformacional. Por otro lado, con MOE se utilizó el algoritmo de emparejamiento por triángulos, el cual inicialmente permea el espacio donde se llevará a cabo el acoplamiento molecular con esferas alfa. La posición inicial del ligando se generará mediante la superposición de tres tripletes de átomos aleatorios en el centro de tres esferas alfa aleatorias. A partir de la posición generada se producirá la vecindad de confórmeros en busca del confórmero con el  $\Delta G$  de unión más favorable. El estudio de acoplamiento molecular se realizó en la cavidad donde los nucleótidos se sitúan durante el proceso de metilación. La presencia de un ligando en la cavidad potencialmente podría impedir que los nucleótidos entren en la cavidad para ser metilados por SAH.

En Ledock, cada compuesto fue sujeto a 20 pasos del algoritmo de búsqueda y con MOE fueron sujetos a 30 pasos, los cuales corresponden a las condiciones estándar en cada programa. Las proteínas se mantuvieron rígidas durante el estudio de acoplamiento molecular.

Cabe destacar, que el porcentaje de acierto de Ledock (57.4%), en la predicción de la pose nativa de un ligando dentro de una proteína, es mayor al de MOE (45.6%) (Pagadala et al. 2017). Lo anterior, podría atribuirse a que Ledock abarca un mayor espacio conformacional que MOE, ya que la generación de confórmeros no se mueve en busca del  $\Delta G$  de unión más favorable durante las primeras iteraciones, de este modo se evita la convergencia prematura en un determinado confórmero. En cambio, en MOE, la generación de confórmeros se dirige a generar los que posean el  $\Delta G$  de unión más favorable, lo que favorece la convergencia prematura, limitando el espacio conformacional explorado.

#### **5.4 Análisis de componentes principales**

A partir de los seis descriptores: SlogP, TPSA, masa molecular, número de átomos donadores de puente de hidrógeno, número de átomos aceptores de puente de hidrógeno y número de enlaces rotables, se llevó a cabo el Análisis de Componentes Principales. Los valores de los descriptores fueron normalizados previamente.

## 6. Discusión de resultados

### 6.1 Modelado por homología

#### 6.1.1 Alineamiento de las secuencias y construcción de la matriz de identidad

El alineamiento de la secuencia primaria de aminoácidos se encuentra en el Apéndice, sección: **10.1 Alineamiento de la secuencia primaria de aminoácidos del dominio catalítico**. A partir de la matriz de identidad (**Fig. 10**) se observó que la secuencia primaria de aminoácidos del dominio catalítico de DNMT3A1 y DNMT3A2 es idéntica. Asimismo, la secuencia primaria de aminoácidos del dominio catalítico es idéntica en las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Consecuentemente, se espera que DNMT3A1 y DNMT3A2 tengan la misma estructura terciaria, al igual que DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Las isoformas de DNMT3A poseen un alto grado de identidad (80.6 %) con respecto a las isoformas catalíticamente activas de DNMT3B y DNMTΔ3B. DNMT1 y DNMT2 poseen un grado de identidad muy bajo (máximo 22.2 %) con respecto al resto de las isoformas.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1:DNMT1		20.5	12.8	18.3	18.3	17.6	18.1	16.1	16.1	16.1	16.1	16.1	16.1	17.1
2:DNMT2:	17.4		14.3	19.0	19.0	20.9	22.2	19.4	19.4	19.4	19.4	19.4	19.4	21.8
3:DNMT3L	5.4	7.2		23.3	23.3	27.2	29.6	23.3	23.3	23.3	23.3	23.3	23.3	33.5
4:DNMT3A1	11.1	13.6	33.2		100.0	81.6	83.3	80.6	80.6	80.6	80.6	80.6	80.6	81.8
5:DNMT3A2	11.1	13.6	33.2	100.0		81.6	83.3	80.6	80.6	80.6	80.6	80.6	80.6	81.8
6:DNMT3B5	9.1	12.8	33.2	69.9	69.9		100.0	85.7	85.7	85.7	85.7	85.7	85.7	99.4
7:DNMT3B3	8.5	12.3	32.7	64.5	64.5	90.4		77.4	77.4	77.4	77.4	77.4	77.4	99.4
8:DNMT3B1	9.8	13.8	33.2	80.6	80.6	100.0	100.0		100.0	100.0	100.0	100.0	100.0	99.4
9:DNMT3B2	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0		100.0	100.0	100.0	100.0	99.4
10:DNMTΔ3B1	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0		100.0	100.0	100.0	99.4
11:DNMTΔ3B2	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0		100.0	100.0	99.4
12:DNMTΔ3B3	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0	100.0		100.0	99.4
13:DNMTΔ3B4	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0		99.4
14:DNMT3B4	6.3	9.5	29.1	49.8	49.8	70.7	78.2	60.6	60.6	60.6	60.6	60.6	60.6	

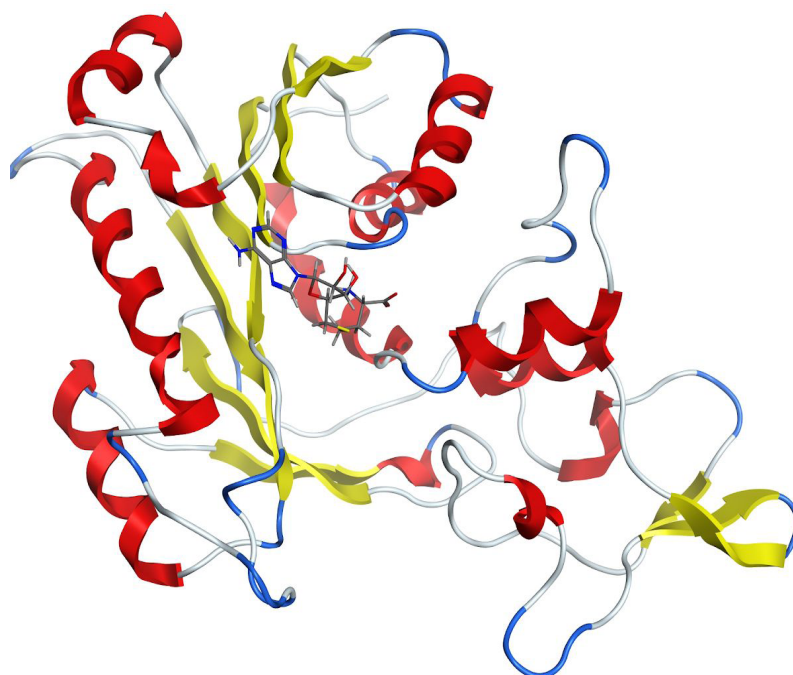
**Fig. 10.** Matriz de identidad del dominio catalítico de 14 isoformas de DNMT (Saldívar-González et al. 2018).

### 6.1.2 Construcción del modelo y comparación con DNMT3A1

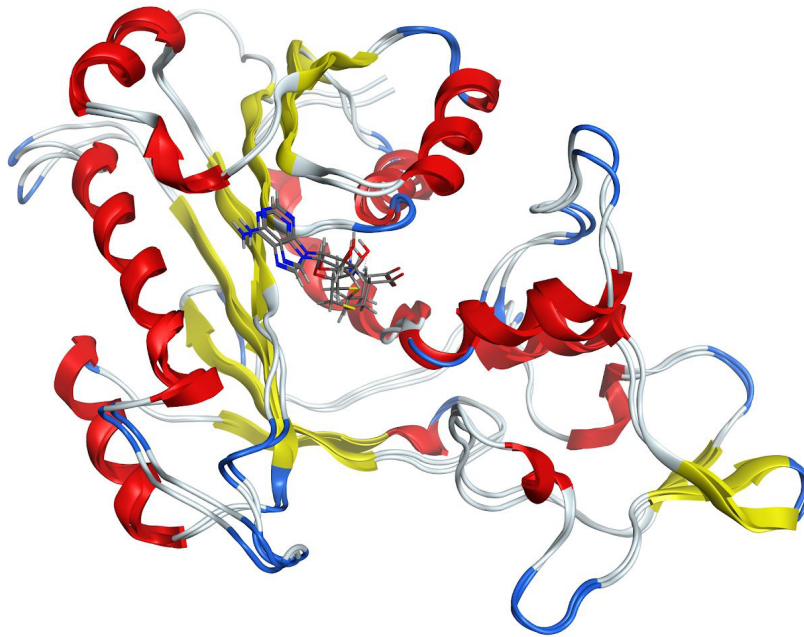
La **Fig. 11** muestra el modelo por homología construido. Al superponer la estructura del modelo por homología con la estructura de DNMT3A1 (PDB ID: 5YX2) **Fig. 12**, se observó que ambas isoformas poseen prácticamente la misma estructura terciaria. La similitud en la estructura terciaria era algo que se esperaba debido al alto grado de identidad de la secuencia primaria de aminoácidos (80.6 %).

En el modelo por homología, con la aplicación “Site Finder” de MOE, se identificó como potencial sitio de unión de una molécula “tipo-fármaco” la misma cavidad que se encuentra descrita como el sitio de metilación del ADN en DNMT3A **Fig. 13**.

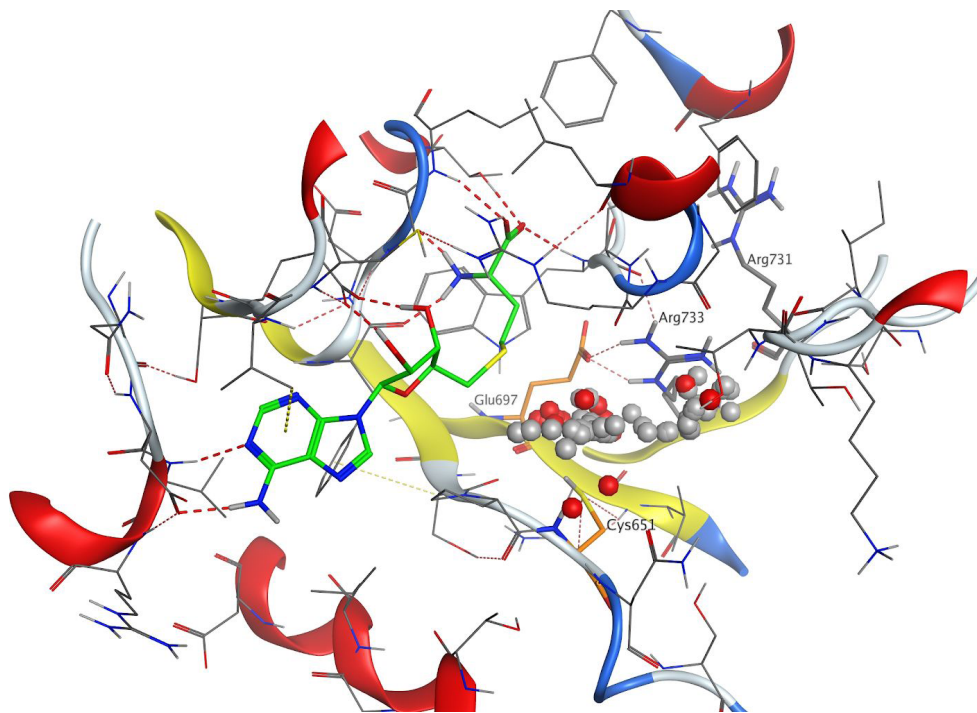
El cofactor SAH se encuentra en el mismo sitio en ambas estructuras. Además, la **Fig. 13** muestra que en el modelo por homología, en la cavidad en la cual se lleva a cabo la metilación del ADN, se encuentran los dos aminoácidos clave participantes en el mecanismo de metilación: cisteína y glutamato.



**Fig. 11.** Modelo por homología que representa el dominio catalítico de las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Se observa al cofactor SAH en su sitio de unión.



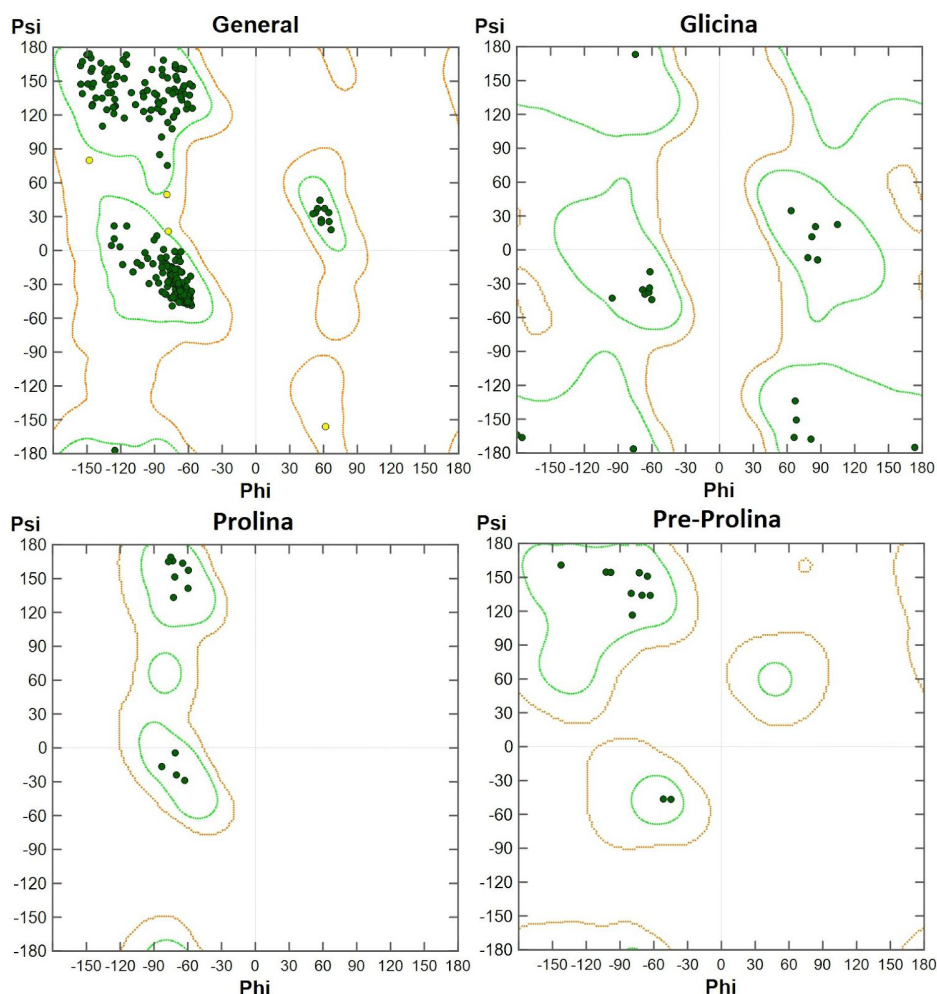
**Fig. 12.** Modelo por homología superpuesto con la estructura cristalográfica de DNMT3A1 (PDB ID: 5YX2). En ambas estructuras se muestra la estructura del cofactor SAH.



**Fig. 13.** Modelo por homología. Cavidad donde se lleva a cabo la metilación del ADN. Cofactor SAH (verde). La región identificada con “Site Finder” como probable sitio de unión de un ligando, que corresponde al sitio donde se lleva a cabo la metilación, se representa con esferas alfa: las regiones de contacto hidrofílicas son representadas con esferas rojas y las hidrofóbicas con esferas grises. Se destacan los dos aminoácidos clave que participan en el mecanismo de metilación: Cys651 y Glu697 en DNMT3B1.

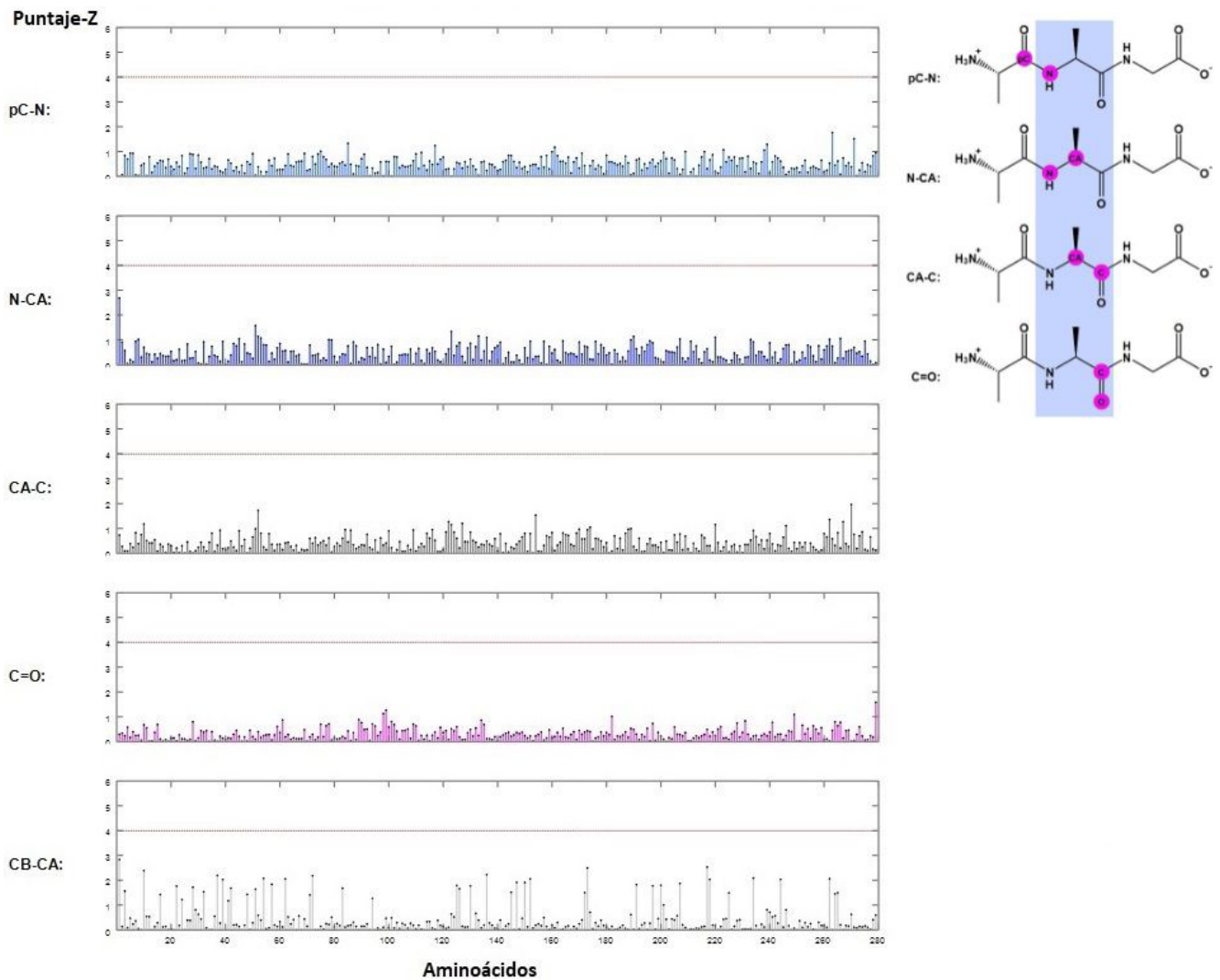
### 6.1.3 Validación del modelo

En cuanto a la evaluación de la calidad estructural de la cadena primaria de aminoácidos, en los cuatro diagramas de Ramachandran (general, glicina, prolina y pre-prolina) **Fig. 14**, no se observó ningún aminoácido en las regiones prohibidas. Considerando que cualquier valor de Z mayor a 4 es un punto atípico, en la cadena principal de aminoácidos, todas las longitudes de enlace (**Fig. 15**) y ángulos que involucran 3 átomos (**Fig. 16**) poseen valores de Z menores a 3. Ser259 tiene un valor de Z mayor a 4 en su ángulo diedro definido como “omega”, el cual es considerado como un punto atípico (**Fig. 17**). Los aminoácidos que se encuentran adyacentes a Ser259 (Arg258 y Trp260) interactúan mediante enlace de puente de hidrógeno con el cofactor SAH. La interacción de Arg258 y Trp260 con SAH son la causa de la deformación observada en el ángulo diedro “omega”. Es común observar deformaciones de los ángulos formados entre los átomos de la cadena principal en los sitios donde hay interacción con un ligando. El resto de los ángulos diedros tuvieron valores de Z menores a 4.

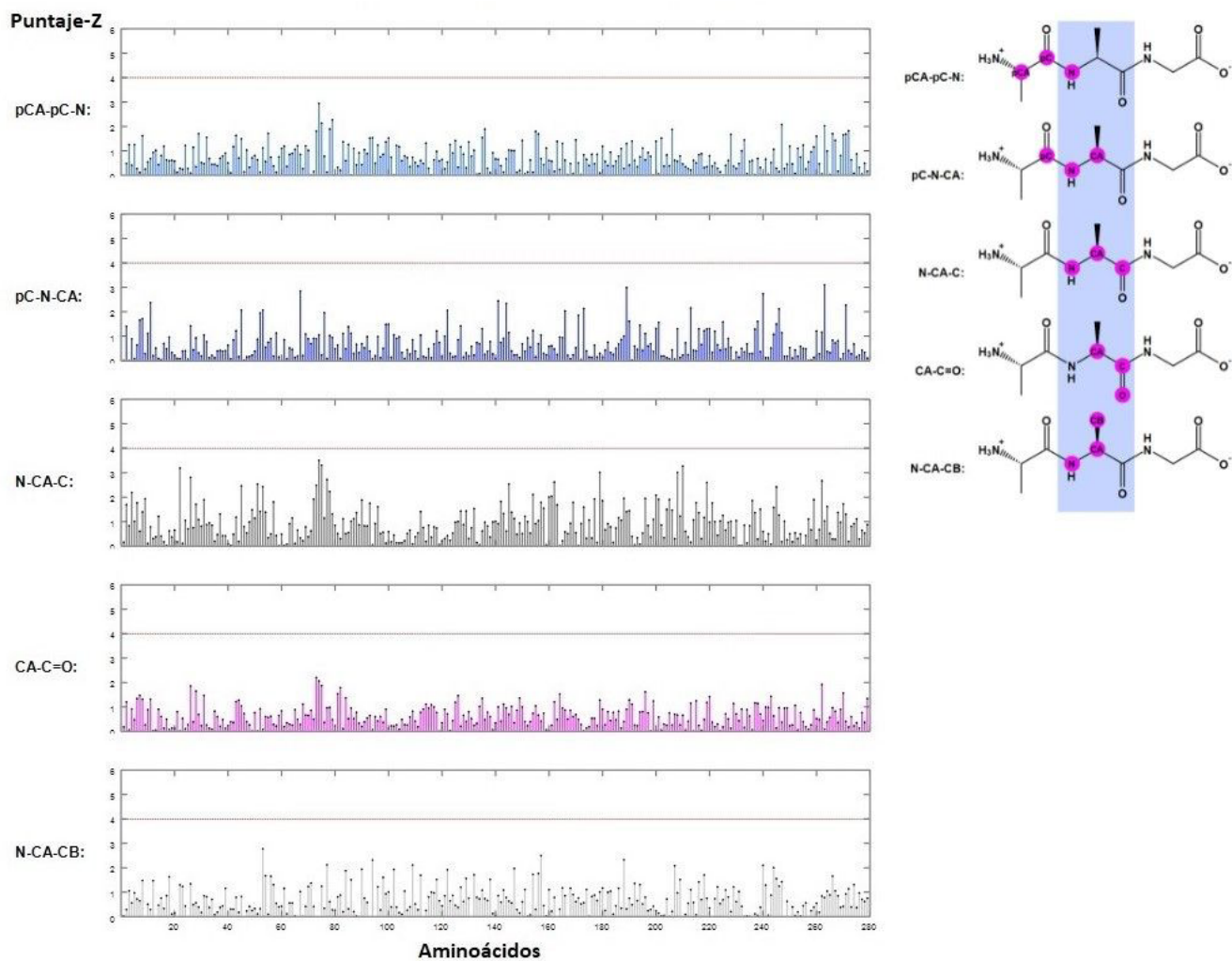


**Fig. 14.** Diagramas de Ramachandran general, de residuos de glicina, prolina y pre-prolina.

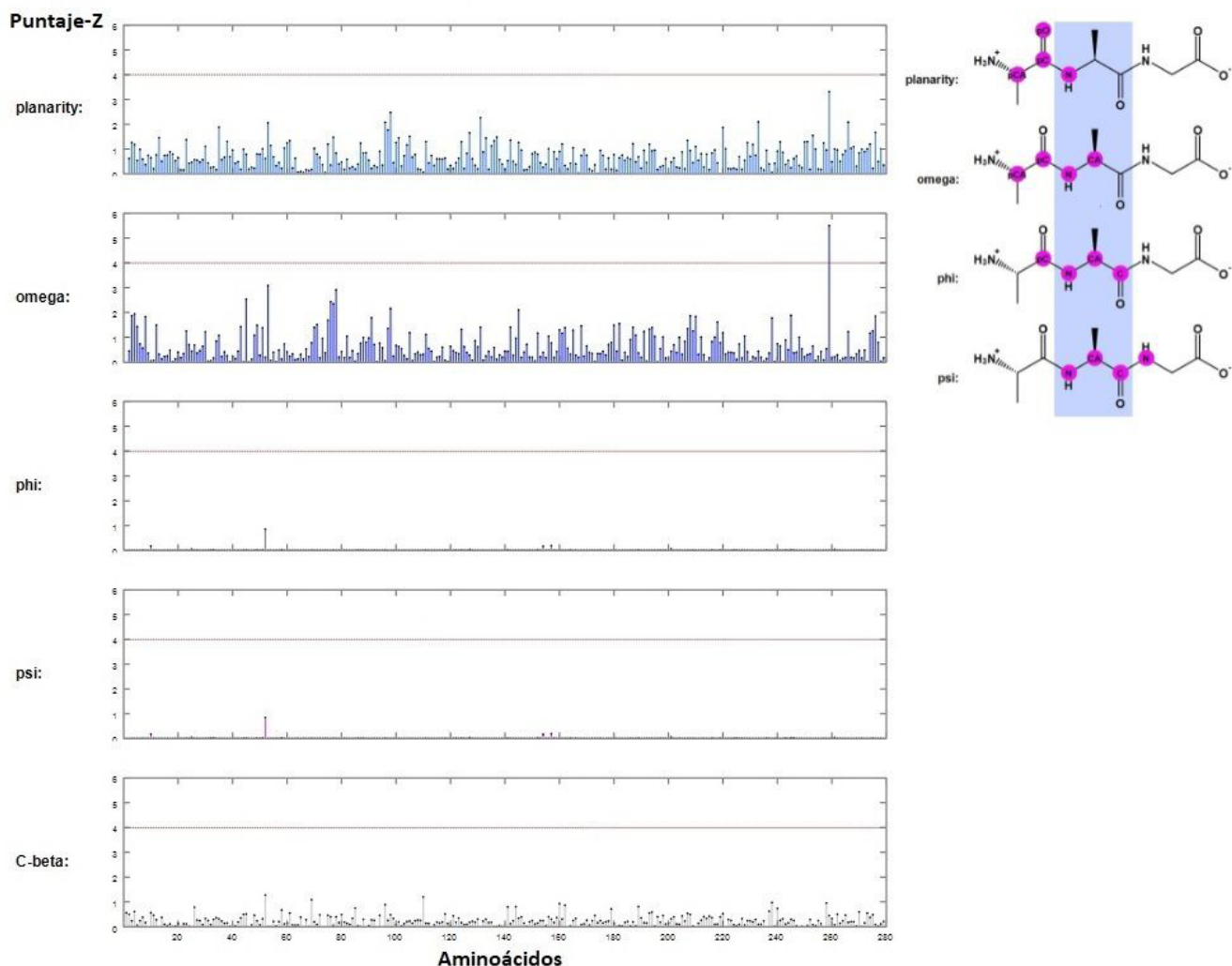




**Fig. 15.** Perfil del Puntaje-Z de las longitudes de enlace de los aminoácidos de la cadena principal. Los valores del Puntaje-Z mayores a 4 son considerados puntos atípicos.



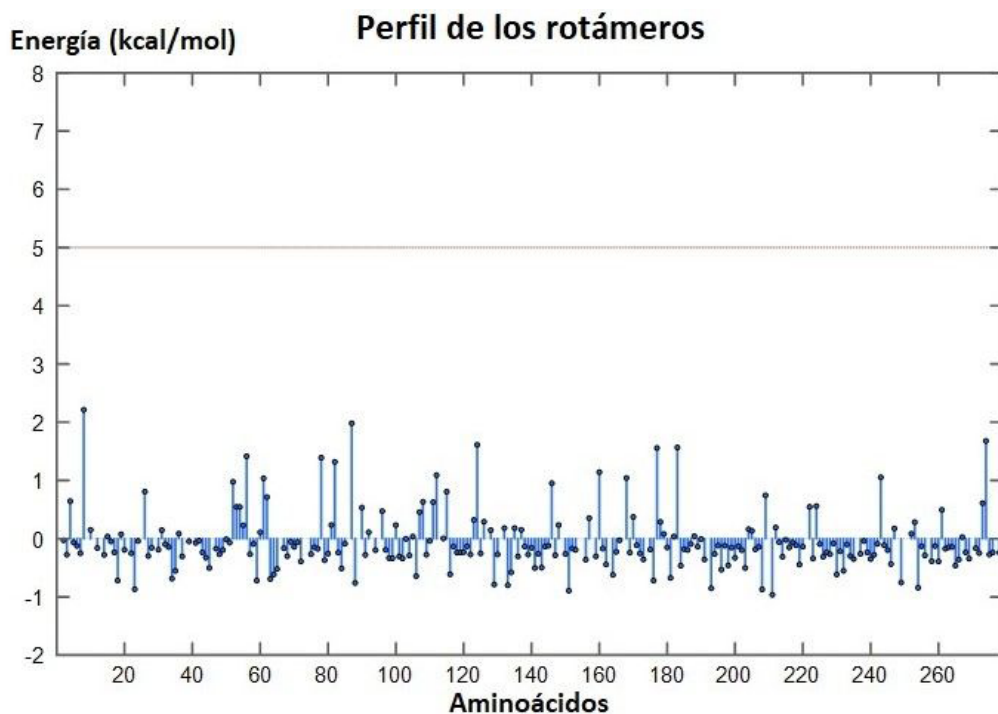
**Fig. 16.** Perfil del Puntaje-Z de los ángulos entre los aminoácidos de la cadena principal. Los valores del Puntaje-Z mayores a 4 son considerados puntos atípicos.



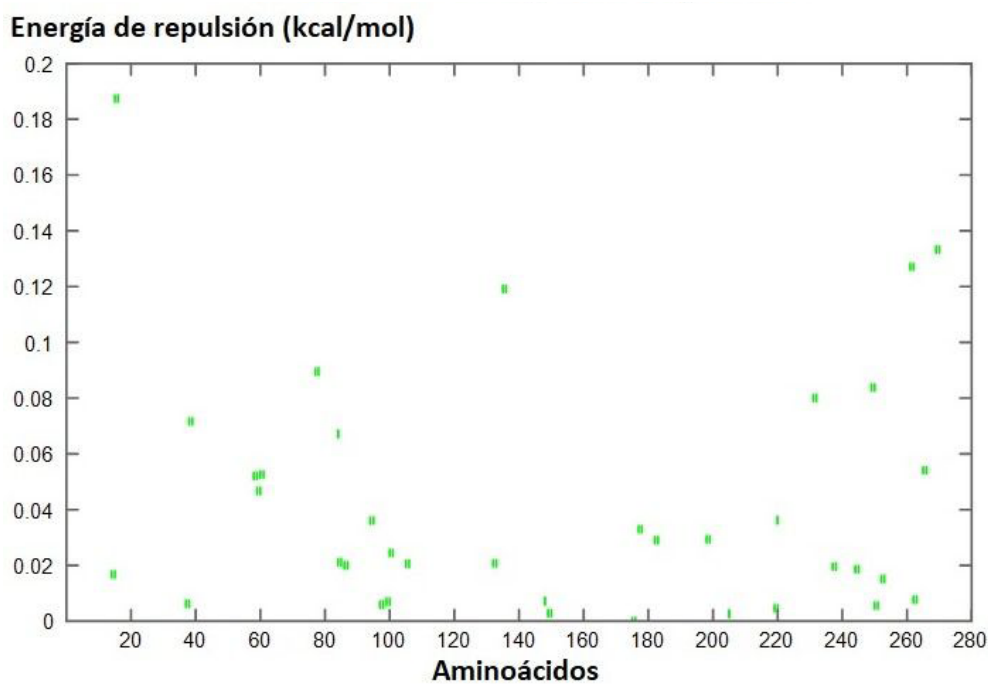
**Fig. 17.** Perfil del Puntaje-Z de los ángulos diedros de la cadena principal. Los valores del Puntaje-Z mayores a 4 son considerados puntos atípicos.

En cuanto a la evaluación de la calidad estructural de la cadena lateral, la energía asociada a los rotámeros oscila entre los -0.964 y los 2.214 (kcal/mol) (**Fig. 18**). Los valores considerados como puntos atípicos tienen una desviación de más de 5 (kcal/mol), por lo cual se considera que no hubo puntos atípicos en los valores de energía asociados a los rotámeros.

De acuerdo a la evaluación de la repulsión estérica entre los átomos del modelo (**Fig. 19**), todos los valores de repulsión fueron menores a 0.2 (kcal/mol), por lo cual se considera que no hay repulsión estérica entre los átomos del modelo.



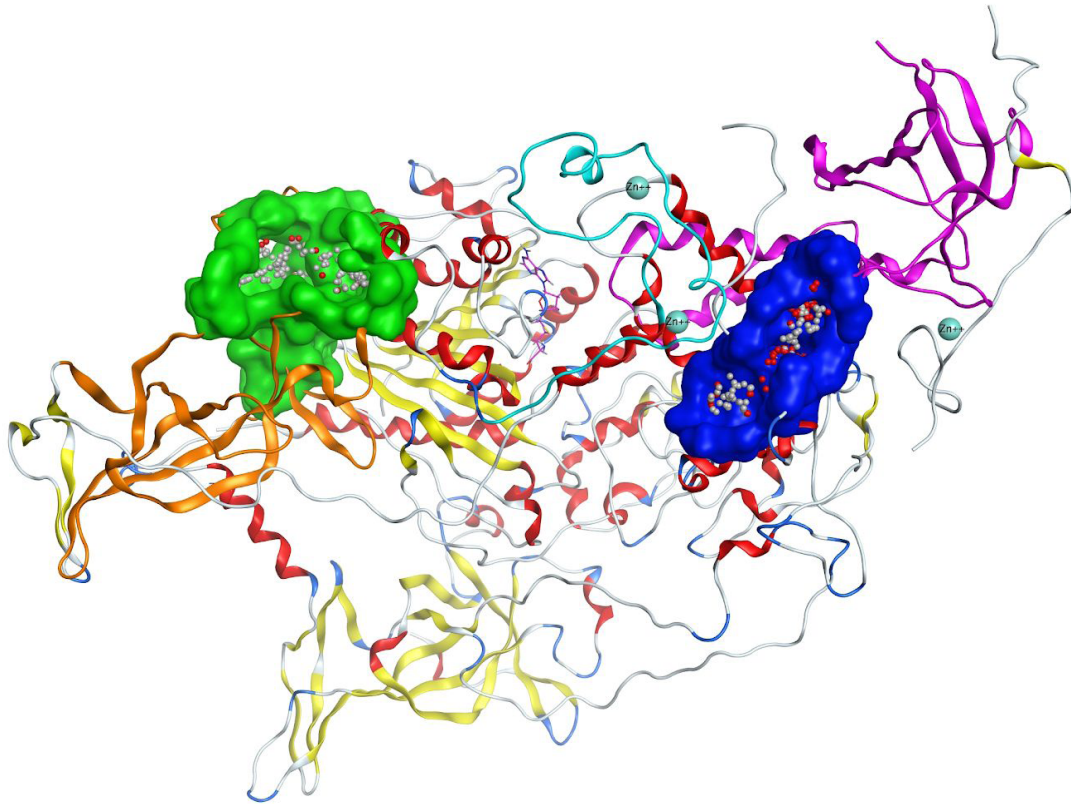
**Fig. 18.** Perfil de la energía de los rotámeros. Los valores del Puntaje-Z mayores a 5 son considerados puntos atípicos (*outliers*).



**Fig. 19.** Perfil de la energía de repulsión. Cada línea verde representa a un aminoácido. Dos líneas verdes adyacentes representan que hay repulsión entre 2 átomos de dos aminoácidos distintos. El valor de la energía de repulsión asociada se encuentra en el eje y de la gráfica.

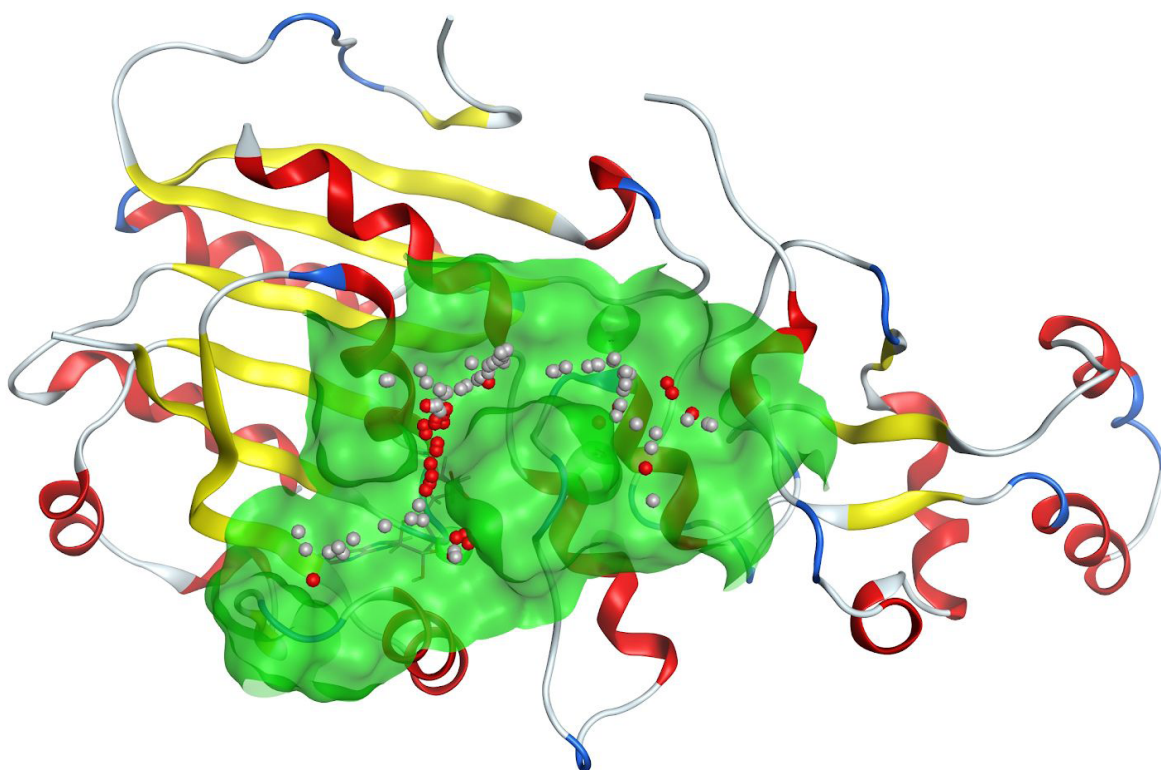
## 6.2 Búsqueda de sitios de unión alternos al dominio catalítico

Es necesario considerar que podrían existir otros sitios alostéricos en las isoformas de DNMT asociados a otros estados conformacionales, adicionales a los considerados en este apartado. En DNMT1 se identificaron dos sitios alostéricos potenciales (**Fig. 20**). El primero de ellos (sitio 1, cavidad verde, PLB = 2.28, **Fig. 20**) forma parte del dominio BAH1 (*Bromo-adjacent homology 1*). La función de BAH1 en DNMT1 es desconocida, sin embargo, se ha propuesto que actúa como modulador de interacciones proteína-proteína. El segundo sitio identificado (sitio 2, cavidad azul, PLB = 2.51, **Fig. 20**) se encuentra entre los dominios CXX y RFTS (*Replication foci targeting sequence*). El dominio CXX es un dominio de dedos de zinc el cual reconoce a los dinucleótidos no metilados citosina y guanina en las islas CpG. El dominio RFTS tiene como función reconocer el foco de replicación del ADN y de esta manera DNMT1 es guiada hacia este sitio. Como se mencionó en la sección de Métodos, el valor numérico de “Propensión de Unión de Ligando” (PLB, por sus siglas en inglés, *vide supra*), es una medida cuantitativa de la probabilidad de que una molécula “tipo-fármaco” se una al sitio identificado. Entre más grande sea el valor numérico de PLB, mayor es la probabilidad de que la molécula se una al sitio. De acuerdo a los resultados obtenidos, una molécula “tipo-fármaco” tendrá mayor afinidad por los dos potenciales sitios de unión identificados (sitio 1, PLB = 2.28; sitio 2, PLB = 2.51), que por el sitio de metilación del ADN (PBL = 1.26). A pesar de lo anterior, el efecto de la unión de un ligando a cualquiera de los dos sitios identificados es incierto, no obstante, hay que considerar que el primero de los sitios forma parte de un dominio que se ha propuesto como regulador de interacciones proteína-proteína y el segundo sitio se encuentra entre dos dominios que son esenciales para el reconocimiento del ADN. Además, se observó que en ninguno de los dos sitios propuestos se llevan a cabo modificaciones postraduccionales a las que es sometida DNMT1: sumoilación (Lee and Muller 2009), fosforilación (Lavoie and St-Pierre 2011), ubiquitinación, acetilación, desacetilación y metilación de residuos de lisina (Rathert et al. 2008) (Peng et al. 2011).



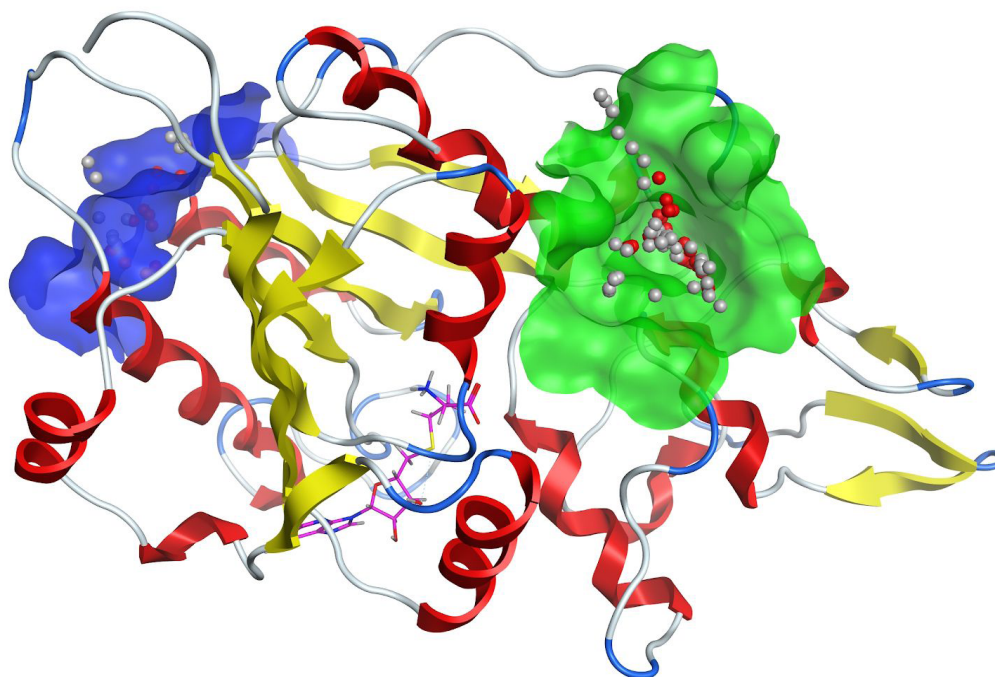
**Fig. 20.** Potenciales sitios alostéricos identificados en DNMT1 (PDB ID: 4WXX). Sitio 1, cavidad verde; Sitio 2, cavidad azul. El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas). Dominio BAH1 (naranja), dominio CXX (azul claro), dominio RFTS (magenta) y cofactor SAH (magenta).

En DNMT2 se identificó un sitio alostérico potencial (PLB=3.06, **Fig. 21**). DNMT2 cuenta únicamente con un solo dominio funcional, el catalítico, que es donde se encuentra el sitio de unión predicho. En UniProt se encontró que DNMT2 (UniProt ID: O14717) interacciona con 16 proteínas diferentes, además, no se encontraron modificaciones postraduccionales reportadas. De acuerdo a los valores de PLB calculados, es mayor la probabilidad de que una molécula “tipo-fármaco” se una al potencial sitio de unión identificado (PLB = 3.06), que al sitio catalítico (PLB = 2.59) (Soga et al. 2007). Queda por confirmar experimentalmente el efecto biológico funcional de la unión de un ligando a este sitio de unión.

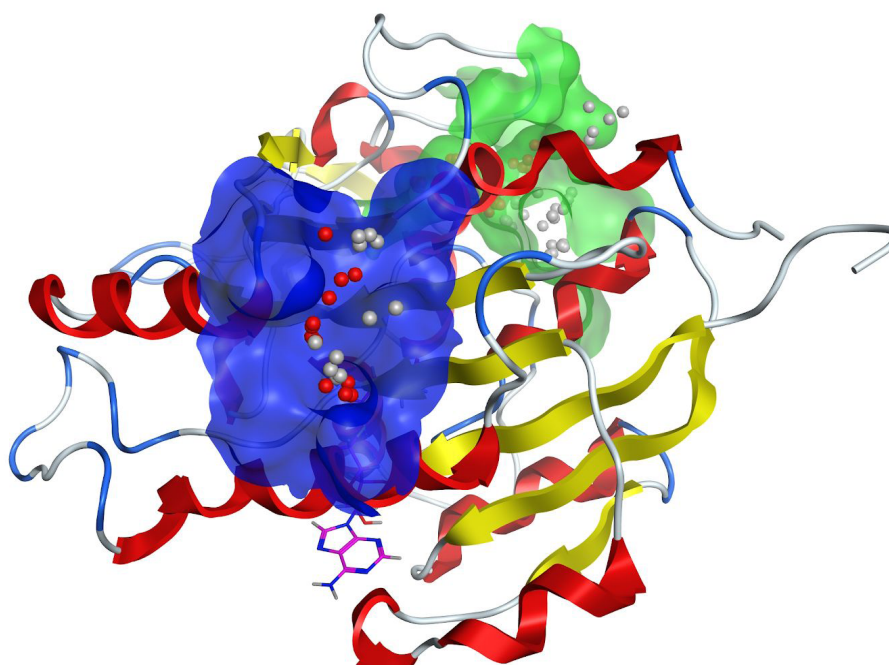


**Fig. 21.** Sitio alostérico potencial identificado en DNMT2 (PDB ID: 1G55). El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas).

En DNMT3A1 se identificaron dos sitios alostéricos potenciales en el dominio catalítico (**Fig. 22** y **Fig.23** sitio 1, cavidad verde, PLB = 1.98, y sitio 2, cavidad azul, PLB = 2.13). DNMT3A1 es activa catalíticamente cuando forma el complejo DNMT3L-DNMT3A-DNMT3A-DNMT3L. El primero de los sitios identificados se encuentra cerca de la interfaz de contacto DNMT3A-DNMT3A. El segundo sitio se encuentra cerca de la interfaz de contacto DNMT3A-DNMT3L. En UniProt se encontró que DNMT3A1 (UniProt ID: Q9Y6K1) sufre modificaciones postraduccionales, sin embargo, éstas se llevan a cabo en regiones cuya estructura no se ha elucidado. De acuerdo al valor de PLB es mayor la probabilidad de que una molécula “tipo-fármaco” se una al sitio catalítico (PLB = 2.38) que a los potenciales sitios de unión alternos (sitio 1, PLB = 1.98; sitio 2, PLB = 2.13). A pesar de que los dos sitios de unión potenciales se encuentran cerca de la interfaz de interacción proteína-proteína, queda por confirmar el efecto biológico funcional de la unión de un ligando al sitio de unión propuesto.



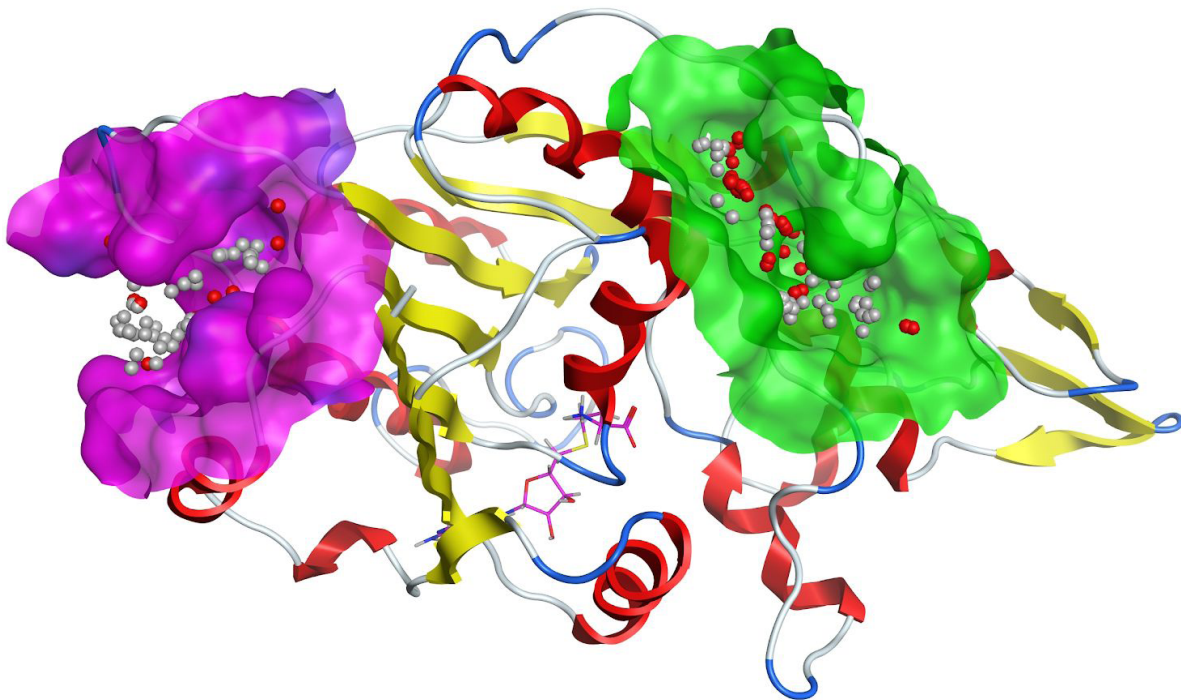
**Fig. 22.** Sitios alostéricos potenciales en DNMT3A1 (PDB ID: 5YX2). Sitio 1, cavidad verde; Sitio 2, cavidad azul. El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas). Cofactor SAH (magenta).



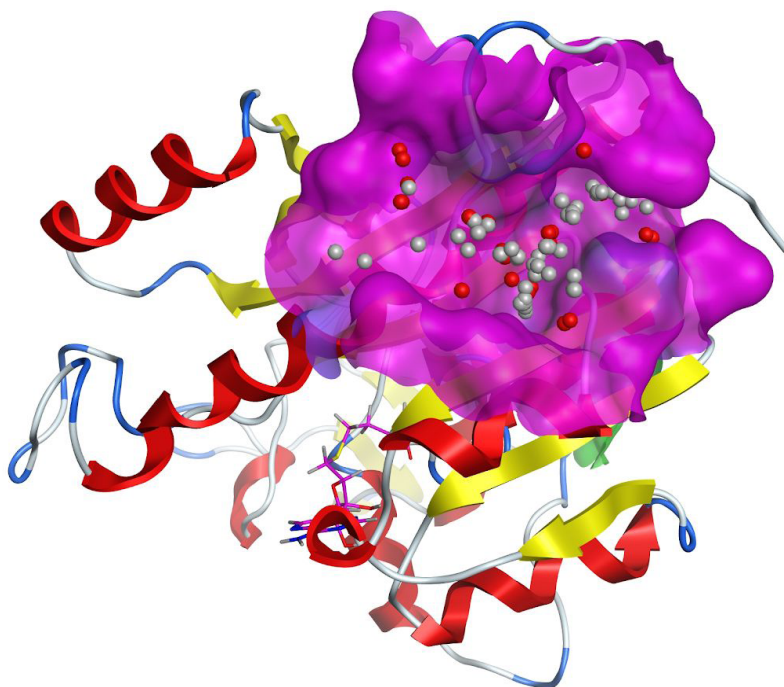
**Fig. 23.** Sitios alostéricos potenciales identificados en DNMT3A1 (PDB ID: 5YX2). Sitio 1, cavidad verde; Sitio 2, cavidad azul. El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas). Cofactor SAH (magenta).



En el modelo por homología que representa el dominio catalítico de las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4, se identificaron dos potenciales sitios alostéricos (**Fig. 24** y **Fig. 25** sitio 1, cavidad verde, PLB = 2.67, y sitio 2, cavidad magenta, PLB = 2.63). El primero de los sitios identificados corresponde con una de las cavidades identificadas en DNMT3A1 (sitio 1, cavidad verde **Fig. 23**). El segundo de los sitios identificados se encuentra adyacente a uno de los sitios identificados en DNMT3A1 (sitio 2, cavidad azul, **Fig. 23**). En UniProt se encontró que DNMT3B (UniProt ID: Q9UBC3) sufre modificaciones postraduccionales, sin embargo, estas se llevan a cabo en regiones que no corresponden al dominio catalítico. Queda por confirmar experimentalmente el efecto funcional de la unión de un ligando al potencial sitio de unión.



**Fig. 24.** Sitios alostéricos potenciales identificados en el modelo por homología que representa las isoformas: DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. Sitio 1, cavidad verde; Sitio 2, cavidad magenta. El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas). Cofactor SAH (magenta).



**Fig. 25.** Sitio alostérico potencial identificado en el modelo por homología (sitio 2), que representa las isoformas: DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. El centro de las esferas alfa es representado con círculos blancos (esferas alfa hidrofóbicas) y rojos (esferas alfa hidrofílicas). Cofactor SAH (magenta).

### 6.3 Cribado virtual

A partir de la base de datos pública FooDB (The Metabolomics Innovation Centre 2017), que contiene 23 883 compuestos, fueron identificados y seleccionados 1 455 compuestos con propiedades “tipo-fármaco”. Posteriormente, como se describe en la Sección de Procedimiento Experimental, se hizo con dos programas el acoplamiento molecular de los 1 455 compuestos seleccionados con DNMT1, DNMT2, DNMT3A1 y DNMT3B. El acoplamiento se centró en la cavidad donde los nucleótidos se sitúan para ser metilados por SAH. Las **Tablas 3 y 4** resumen los resultados para los diez compuestos con mejores puntajes para cada isoforma. Para cada estructura de DNMT, en el estudio de acoplamiento molecular, se utilizó como referencia **SIGI-1027**, el cual es un inhibidor conocido de las distintas isoformas de DNMT en un rango de  $IC_{50}$  de 6-13  $\mu$ M (Datta et al. 2009). Se prescindió del uso como referencia, para el estudio de acoplamiento molecular, de los dos fármacos aprobados como inhibidores de DNMT (azacitidina y decitabina) ya que ambos solo inhiben a DNMT1. Además el principal mecanismo de acción de ambos fármacos no es la unión con la enzima, sino su incorporación al ADN.

Para la isoforma DNMT1, se identificaron cuatro estructuras con un  $\Delta G$  de unión ligeramente mejor al inhibidor **SIGI-1027** (MOE: FDB006198 “charantina”, FDB018786 “lignano” y Ledock: FDB023598 “Rosiglitazona” y FDB023581 “Lansoprazol”). Las estructuras identificadas con el programa Ledock

(Fig. 26 y Fig. 27) presentan más interacciones por puente de hidrógeno e interacciones  $\pi$ , a comparación de las obtenidas con MOE. Los dos compuestos encontrados con Ledock presentan interacción mediante puente de hidrógeno con Glu226, el cual es esencial en el mecanismo de la metilación del ADN.

**Tabla 3.  $\Delta G$  de unión de los compuestos con mayor afinidad al sitio catalítico de DNMT1 y DNMT2.**

DNMT1				DNMT2			
MOE		Ledock		MOE		Ledock	
ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)
<b>FDB006198</b> Charanina	-7.66	<b>FDB023598</b> Rosiglitazona	-7.32	<b>FDB018786</b> Lignano	-8.13	<b>FDB021494</b> Nb-Ferulotriptamina	-7.43
<b>FDB018786</b> Lignano	-7.58	<b>FDB023581</b> Lansoprazol	-7.12	FDB019377	-7.95	FDB021493	7.18
FDB012453	-7.37	FDB021493	-6.59	<b>FDB006198</b> Charanina	-7.85	<b>FDB023598</b> Rosiglitazona	-7.13
FDB029185	-7.33	<b>FDB021494</b> Nb-Ferulotriptamina	-6.38	FDB020611	-7.83	FDB024123	-7.1
FDB002322	-7.30	FDB010985	-6.25	FDB021103	-7.81	<b>FDB023581</b> Lansoprazol	-6.94
FDB003148	-7.29	FDB020193	-6.21	<b>FDB020656</b> O-Etilcubebina	-7.79	FDB009779	-6.76
FDB019236	-7.23	FDB020725	-6.09	FDB011384	-7.78	FDB020725	-6.74
FDB020802	-7.22	FDB009770	-6.05	FDB012176	-7.75	FDB013257	-6.63
FDB019403	-7.19	FDB024123	-6.05	FDB029067	-7.72	FDB000900	-6.60
FDB003032	-7.19	FDB012099	-6.02	FDB019189	-7.72	FDB019132	-6.50
<b>SGI-1027</b>	-7.56	<b>SGI-1027</b>	-7.08	<b>SGI-1027</b>	-8.16	<b>SGI-1027</b>	-6.29

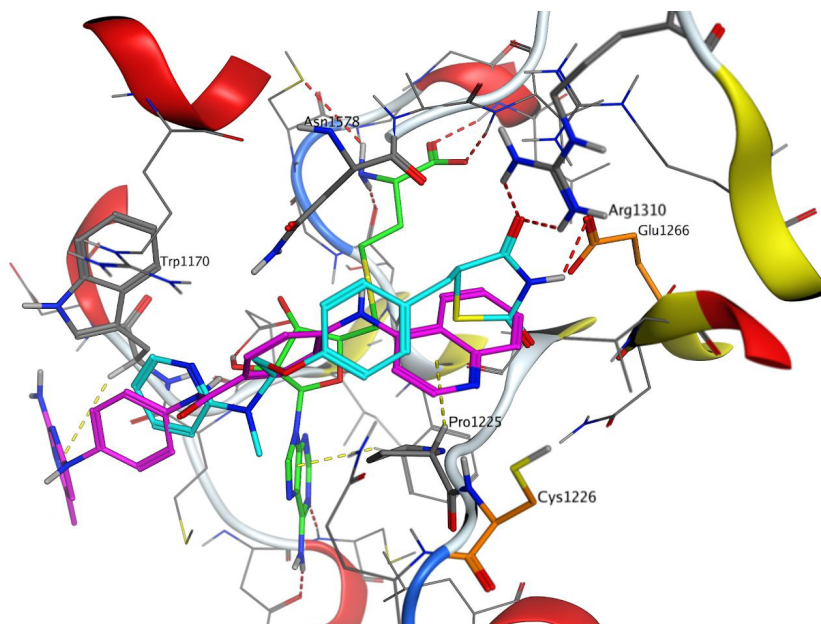
**Tabla 4.  $\Delta G$  de unión de los compuestos con mayor afinidad al sitio catalítico de DNMT3A1 y DNMT3B (modelo por homología).**

DNMT3A1				DNMT3B (Modelo por homología)			
MOE		Ledock		MOE		Ledock	
ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)	ID	$\Delta G$ (kcal/mol)
FDB011079	-7.91	<b>FDB023598</b> Rosiglitazona	-5.78	<b>FDB020656</b> O-Etilcubebina	-7.27	FDB024125	-6.47
FDB012453	-7.65	<b>FDB021494</b> Nb-Ferulotriptamina	-5.74	FDB030988	-7.25	FDB000900	-6.00
FDB011760	-7.57	FDB024123	-5.73	FDB012453	-7.16	FDB024123	-5.94
FDB001502	-7.50	FDB024125	-5.60	<b>FDB006198</b> Charanina	-7.11	FDB022574	-5.92
FDB019403	-7.49	FDB021493	-5.25	FDB011826	-7.10	FDB017961	-5.89
FDB019377	-7.41	FDB004118	-5.22	FDB015622	-7.06	FDB020725	-5.80
FDB010941	-7.41	FDB017961	-5.22	FDB010941	-7.01	FDB021493	-5.78
FDB014660	-7.40	FDB004087	-5.16	FDB013875	-7.00	<b>FDB021494</b> Nb-Ferulotriptamina	-5.72
FDB000398	-7.38	FDB012616	-5.15	FDB002303	-7.00	FDB020967	-5.67
FDB013705	-7.37	FDB000900	-5.14	FDB003032	-6.95	FDB013257	-5.63
<b>SGI-1027</b>	-8.14	<b>SGI-1027</b>	-5.07	<b>SGI-1027</b>	-7.40	<b>SGI-1027</b>	-6.61

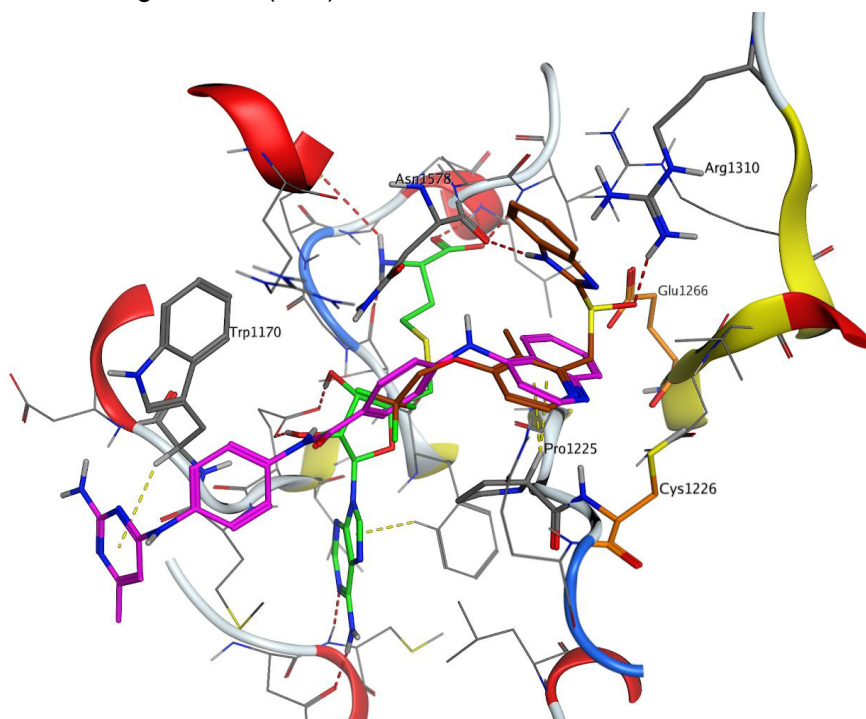
Para DNMT2, las estructuras identificadas en Ledock con los 10 mejores puntajes, tienen un  $\Delta G$  de unión mejor que **SGI-1027**. De entre ellas FDB021494 “Nb-Ferulotriptamina” posee el mejor  $\Delta G$  de unión. Se aprecian interacciones mediante puente de hidrógeno de FDB021494 con Cys79 y Glu119, los cuales son esenciales en el mecanismo de metilación (**Fig. 28**). Con MOE no se identificaron compuestos con un puntaje mejor que **SGI-1027**.

Los 10 compuestos identificados en Ledock con los mejores  $\Delta G$  de unión en DNMT3A1, tienen un mejor puntaje que **SGI-1027**. Se destaca FDB023598 “Rosiglitazona” que tuvo el mejor  $\Delta G$  de unión. A pesar de no interactuar con los aminoácidos Cys710 y Glu756 que son esenciales durante la metilación, FDB023598 tiene otras interacciones de tipo  $\pi$  y por puente de hidrógeno (**Fig. 29**). Con MOE no se identificaron compuestos con un  $\Delta G$  de unión mejor que **SGI-1027**.

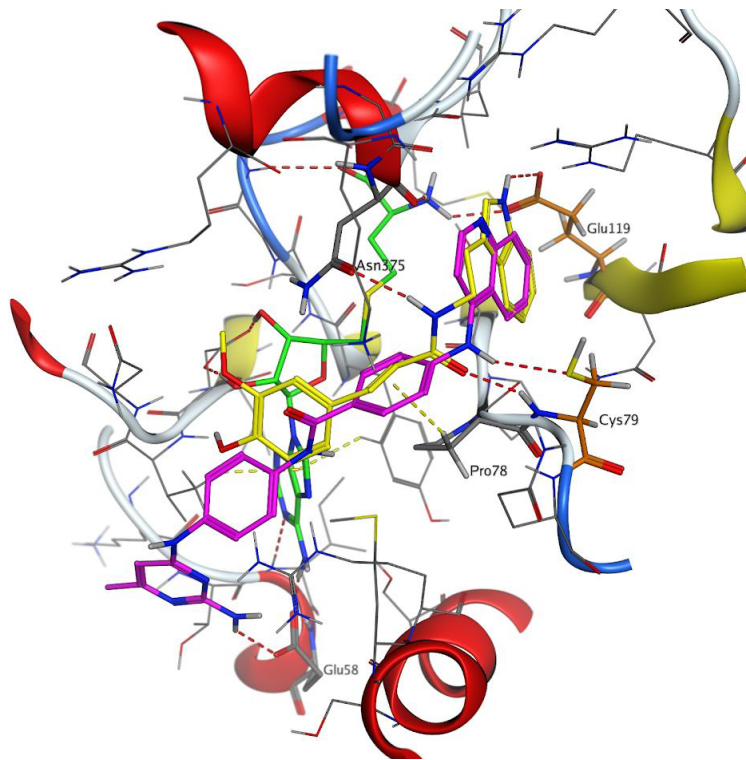
Para el modelo por homología de DNMT3B no se identificaron con ninguno de los dos programas compuestos con un  $\Delta G$  de unión mejor a **SGI-1027**. Sin embargo, con MOE se identificó al compuesto FDB020656 “O-Etilcubebina” con un  $\Delta G$  de unión similar a **SGI-1027**. FDB020656 presenta únicamente interacciones de tipo  $\pi$  con DNMT3B (**Fig. 30**).



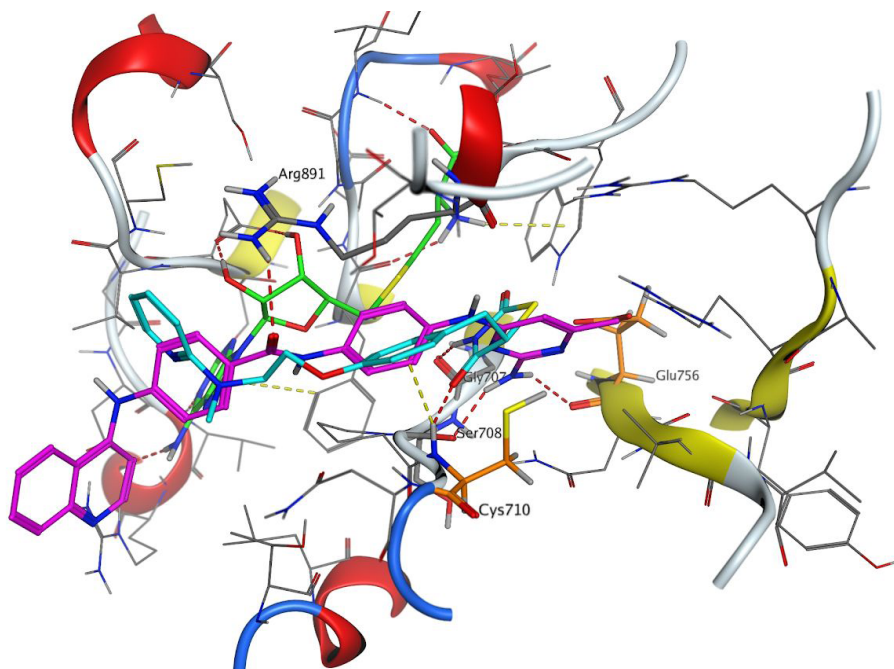
**Fig. 26.** Cavity donde se lleva a cabo la metilación del ADN en DNMT1. Cofactor SAH (verde), SGI-1027 (magenta), FDB023598 “rosiglitazona” (azul).



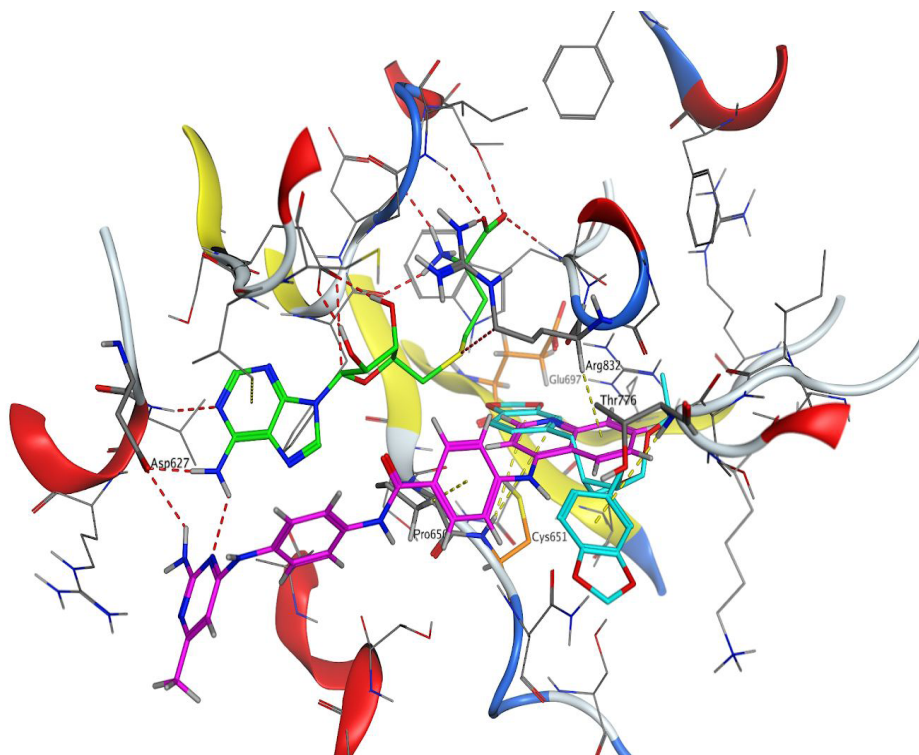
**Fig. 27.** Cavity donde se lleva a cabo la metilación del ADN en DNMT1. Cofactor SAH (verde), SGI-1027 (magenta), FDB023581 “lansoprazol” (azul).



**Fig. 28.** Cavidad donde se lleva a cabo la metilación del ADN en DNMT2. Cofactor SAH (verde), SGI-1027 (magenta), FDB021494 “Nb-Ferulotriptamina” (amarillo).



**Fig. 29.** Cavidad donde se lleva a cabo la metilación del ADN en DNMT3A1. Cofactor SAH (verde), SGI-1027 (magenta), FDB023598 “rosiglitazona” (azul).



**Fig. 30.** Cavity where the DNA methylation is carried out in DNMT3B. Cofactor SAH (green), SGI-1027 (magenta), FDB020656 “O-Etilcubebina” (blue). The numbering of the amino acids is with respect to the sequence of the catalytic domain of DNMT3B1.

La cavidad donde se lleva a cabo el proceso de la metilación está conservada en todas las isoformas, especialmente la disposición espacial de los aminoácidos (Lyko 2018). Por lo tanto, en la **Tabla 5** se observa que los compuestos seleccionados son afines a dos o tres isoformas de DNMT.

En DNMT1, los 10 compuestos con el mejor  $\Delta G$  de unión identificados con MOE son diferentes a los 10 compuestos con el mejor  $\Delta G$  de unión identificados con Ledock. Lo mismo se observó en DNMT2, DNMT3A1 y el modelo por homología de DNMT3B **Tablas 3 y 4**. Lo cual se puede atribuir, a que ambos programas de acoplamiento molecular utilizan dos distintos algoritmos para la generación de conformeros. De esta manera al utilizar los dos programas se aumentó la cantidad de conformeros generados de cada compuesto, permitiendo identificar compuestos distintos con cada programa. En otros estudios de acoplamiento molecular, se ha observado que se obtiene un mejor resultado utilizando dos programas, en lugar de uno solo (Houston and Walkinshaw 2013; Preto and Gentile 2019). Es posible obtener mejores resultados, utilizando tres programas, en lugar de dos, sin embargo, el tiempo que se requiere para llevar a cabo el estudio de acoplamiento molecular es mayor (Miteva et al. 2005).

Para explorar si los compuestos de la **Tabla 5** tienen actividad biológica reportada, se hizo una búsqueda bibliográfica en SciFinder. En el caso de de FDB006198 “Charantina”, el cual se encuentra en la calabaza amarga (The Metabolomics Innovation Centre 2017), no hay actividad biológica

reportada. Este compuesto posee en su estructura un sustituyente de naftalina el cuál podría ser tóxico, ya que este sustituyente potencialmente puede intercalarse en el ADN (Banks et al. 2016). Sin embargo, es necesario precisar que el fármaco propanolol, utilizado actualmente en el tratamiento terapéutico de la hipertensión (Bigucci et al. 2015), posee en su estructura un sustituyente de naftalina. Como alternativa podría llevarse a cabo un reemplazo bioisostérico. Dos reemplazos bioisostéricos que se encuentran reportados para la naftalina corresponden a la sustitución por benzaboronina (Rombouts et al. 2015) y por un indol (Abdellaoui et al. 1995).

FDB018786 “lignano” es un metabolito secundario que se encuentra en una gran variedad de plantas (The Metabolomics Innovation Centre 2017), el cual tiene reportada actividad antifúngica (Zacchino et al. 1997) y antioxidante (Konya 2001). En su estructura posee un alqueno, el cual resulta ser una alerta estructural debido a que potencialmente puede convertirse a través del metabolismo en un epóxido (Zhang and Elfarra 2017). Un reemplazo bioisostérico que podría llevarse a cabo es la sustitución del alqueno por una amida (Khalaf et al. 2011).

FDB020656 “O-Etilcubebina” se encuentra en la pimienta cubeba (The Metabolomics Innovation Centre 2017), la cual suele usarse como condimento y tiene reportada actividad antibacteriana en contra de *Mycobacterium tuberculosis*.

Inesperadamente, en la base de datos de FooDB se encontraron dos fármacos, rosiglitazona el cual es utilizado para tratar la diabetes mellitus tipo 2 (Alawi et al. 2020) y lansoprazol, indicado para tratar patologías relacionadas al exceso de secreción de ácido en el estómago (Iwahi et al. 1991).

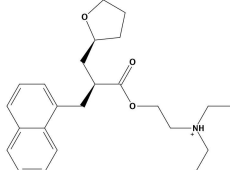
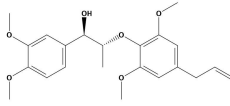
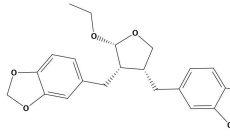
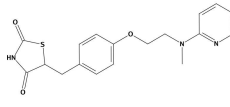
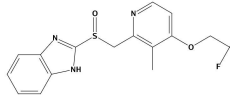
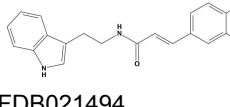
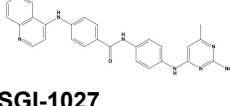
FDB021494 “Nb-Ferulotriptamina” se encuentra en los granos de maíz dulce (The Metabolomics Innovation Centre 2017). Esta molécula posee dentro de su estructura una porción carbonílica  $\alpha,\beta$ -insaturada, por lo cual resulta ser un aceptor de Michael. Los aceptores de Michael tienden a ser estructuras promiscuas debido a su carácter electrofílico. Además, esta molécula es potencialmente tóxica al presentar actividad inhibitoria sobre la acetilcolinesterasa (Young Sim and Ae Kim 2014).

Por lo tanto, FDB006198 “charantina”, FDB018786 “lignano” y FDB021494 “Nb-Ferulotriptamina”, previo a ser utilizados como prototipo para el diseño de inhibidores de DNMT, podrían requerir modificaciones estructurales para evitar un potencial efecto tóxico.

FDB006198 “charantina”, FDB018786 “lignano”, FDB020656 “O-Etilcubebina”, FDB021494 “Nb-Ferulotriptamina”, FDB023598 “Rosiglitazona” y FDB023581 “Lansoprazol” son estructuras que, de acuerdo al estudio de acoplamiento molecular, son afines al sitio de metilación del ADN y por lo tanto potenciales inhibidores de DNMT, ya que impedirían que el ADN se sitúe en la cavidad donde es metilado. Queda por confirmar experimentalmente el efecto inhibitorio sobre las respectivas isoformas de DNMT de estos cuatro compuestos encontrados en los alimentos y plantas y de los fármacos rosiglitazona y lansoprazol.



**Tabla 5. Potenciales inhibidores de distintas isoformas de DNMT.**

Compuesto <sup>a</sup>	Fuente natural	Programa <sup>b</sup>	Actividades biológicas reportadas	DNMT1	DNMT2	DNMT3A	DNMT3B
 FDB006198 Charantina	Calabaza amarga	MOE	Ninguna	✓ -7.66	✓ -7.85		✓ -7.11
 FDB018786 Lignano	Gran variedad de plantas	MOE	Antifúngica (Zacchino et al. 1997) Antioxidante (Konya 2001)	✓ -7.58	✓ -8.13		
 FDB020656 O-Etilcubebina	Pimienta cubeba	MOE	Antimicrobiana <i>Mycobacterium tuberculosis</i> (Baquero et al. 2011)		✓ -7.79		✓ -7.27 <b>Fig. 30</b>
 FDB023598 Rosiglitazona <sup>c</sup>	No procede de una fuente natural	Ledock	Hipoglucemiante (Alawi et al. 2020)	✓ -7.32 <b>Fig. 26</b>	✓ -7.13	✓ -5.78 <b>Fig. 29</b>	
 FDB023581 Lansoprazol <sup>d</sup>	No procede de una fuente natural	Ledock	Inhibidor de la bomba de protones (Iwahi et al. 1991)	✓ -7.12 <b>Fig. 27</b>	✓ -6.94		
 FDB021494 Nb-Ferulotriptamina	Granos de maíz dulce	Ledock	Inhibidor de la acetilcolinesterasa (Young Sim and Ae Kim 2014) Antioxidante (Georgiev et al. 2013)	✓ -6.38	✓ -7.43 <b>Fig. 28</b>	✓ -5.74	
 <b>SGI-1027</b>		MOE		-7.56	-8.16	-8.14	-7.40 <b>Fig. 30</b>
		Ledock		-7.08 <b>Fig. 26</b> <b>Fig. 27</b>	-6.29 <b>Fig. 28</b>	-5.07 <b>Fig. 29</b>	-6.61

<sup>a</sup> ID de FooFB, nombre común y estructura.

<sup>b</sup> Programa mediante el cual fue identificado el compuesto.

<sup>c</sup> Fármaco indicado para tratar la diabetes mellitus tipo 2.

<sup>d</sup> Fármaco indicado para tratar patologías relacionadas al exceso de secreción de ácido en el estómago.

✓ El compuesto presenta un  $\Delta G$  de unión mejor que **SGI-1027**.

#### 6.4 Análisis de componentes principales

El Componente Principal 1, por sus siglas en inglés “PC1”, describe el 64.95 % de la varianza, mientras que el Componente Principal 2, por sus siglas en inglés “PC2”, describe el 30.45% de la varianza, por lo tanto el 95.41 % de la varianza de los datos es descrita por PC1 y PC2. En la **Tabla 6**, se observa que TPSA (*Topological Polar Surface Area*), MW (*Molecular Weight*), HBA (*Hydrogen Bond Acceptor*) y HBD (*Hydrogen Bond Donor*) son los descriptores con mayor contribución al “PC1”. Los enlaces rotables y SLogP son los descriptores con mayor contribución al PC2. En la **Fig. 31** se observa la distribución de los compuestos de la base de datos FooDB en el espacio químico. Sin embargo en la **Fig. 31** no se aprecia que la zona que corresponde a compuestos “tipo-fármaco” se encuentra traslapada con compuestos que no son “tipo-fármaco”. En la **Fig. 31** en zona que se encuentra aproximadamente entre los valores de PC1 (-15,-5) y PC2 (-2,2), se observan compuestos que se encuentran dispuestos en una línea recta vertical, con una ligera pendiente positiva, a lo largo del PC2. Cada línea recta se encuentra conformada por compuestos cuyos valores de TPSA, MW, HBA y HBD (descriptores con mayor contribución al PC1) son muy similares. En contraste, sus valores de enlaces rotables y SLogP (descriptores con mayor contribución al PC2), aumentan de manera directamente proporcional, conforme disminuye el valor de PC2. Asimismo, en la zona que se encuentra aproximadamente entre los valores de PC1 (-18,-4) y PC2 (-4,-1) se observan compuestos que se encuentran dispuestos en una línea recta horizontal de pendiente negativa, a lo largo del PC1. La línea recta se encuentra conformada por compuestos cuyos valores de enlaces rotables y SLogP (descriptores con mayor contribución al PC2) son muy similares. En contraste, sus valores de TPSA, MW, HBA y HBD (descriptores con mayor contribución al PC1), aumentan de manera directamente proporcional, conforme disminuye el valor de PC1.

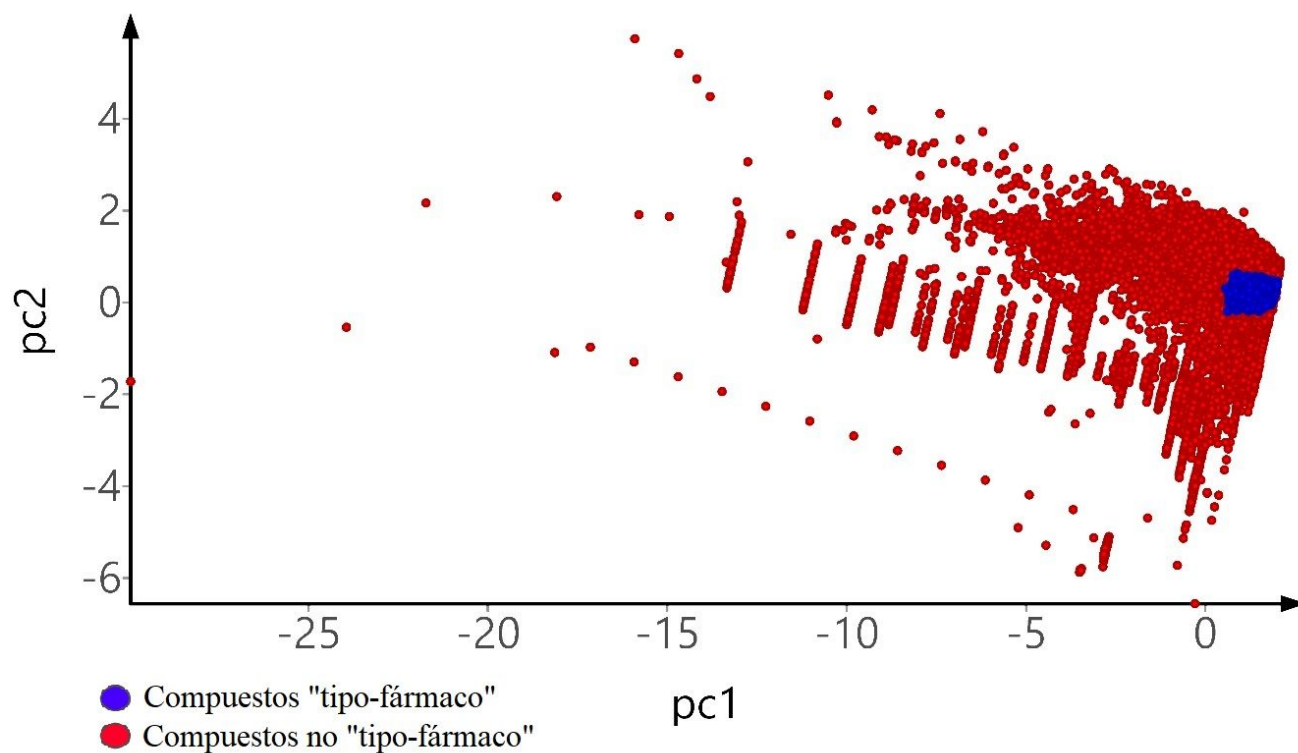
La zona que corresponde al espacio químico de los compuestos con mejores  $\Delta G$  de unión para DNMT1, DNMT2, DNMT3A1 y el modelo por homología de DNMT3B (MOE: **Fig. 32** y Ledock: **Fig. 33**), se encuentra bien delimitada en el caso de MOE **Fig. 32**. Esta zona es prácticamente la misma en el caso de las cuatro isoformas, lo cual puede ser atribuido a que la cavidad, en donde se llevó a cabo el estudio de acoplamiento molecular se encuentra bien conservada en las cuatro isoformas. La zona que corresponde al espacio químico de los compuestos con mejores  $\Delta G$  de unión de Ledock **Fig. 33** no se encuentra tan bien delimitada como en el caso de MOE **Fig. 32**, no obstante, se encuentra traslapada con la zona que cubren los compuestos de MOE.

**Tabla 6. Eigenvectores de los componentes principales.**

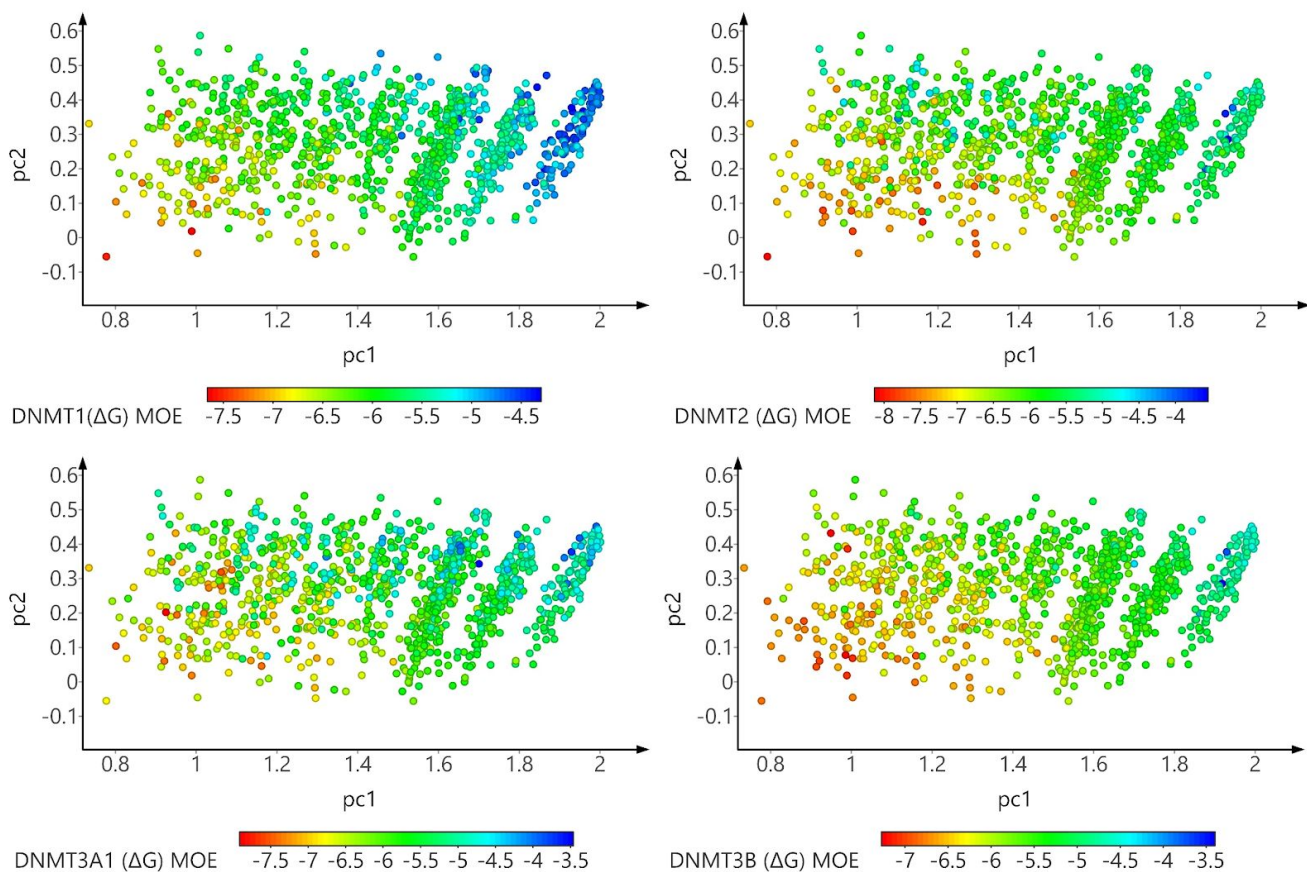
Descriptor	Eigenvector 1*	Eigenvector 2**
SLogP	0.113	-0.699
TPSA	-0.496	0.124
MW	-0.457	-0.293
HBA	-0.499	0.082
HBD	-0.453	0.252
Enlaces rotables	-0.278	0.582

\*Se observan las 6 componentes del eigenvector asociado a PC1.

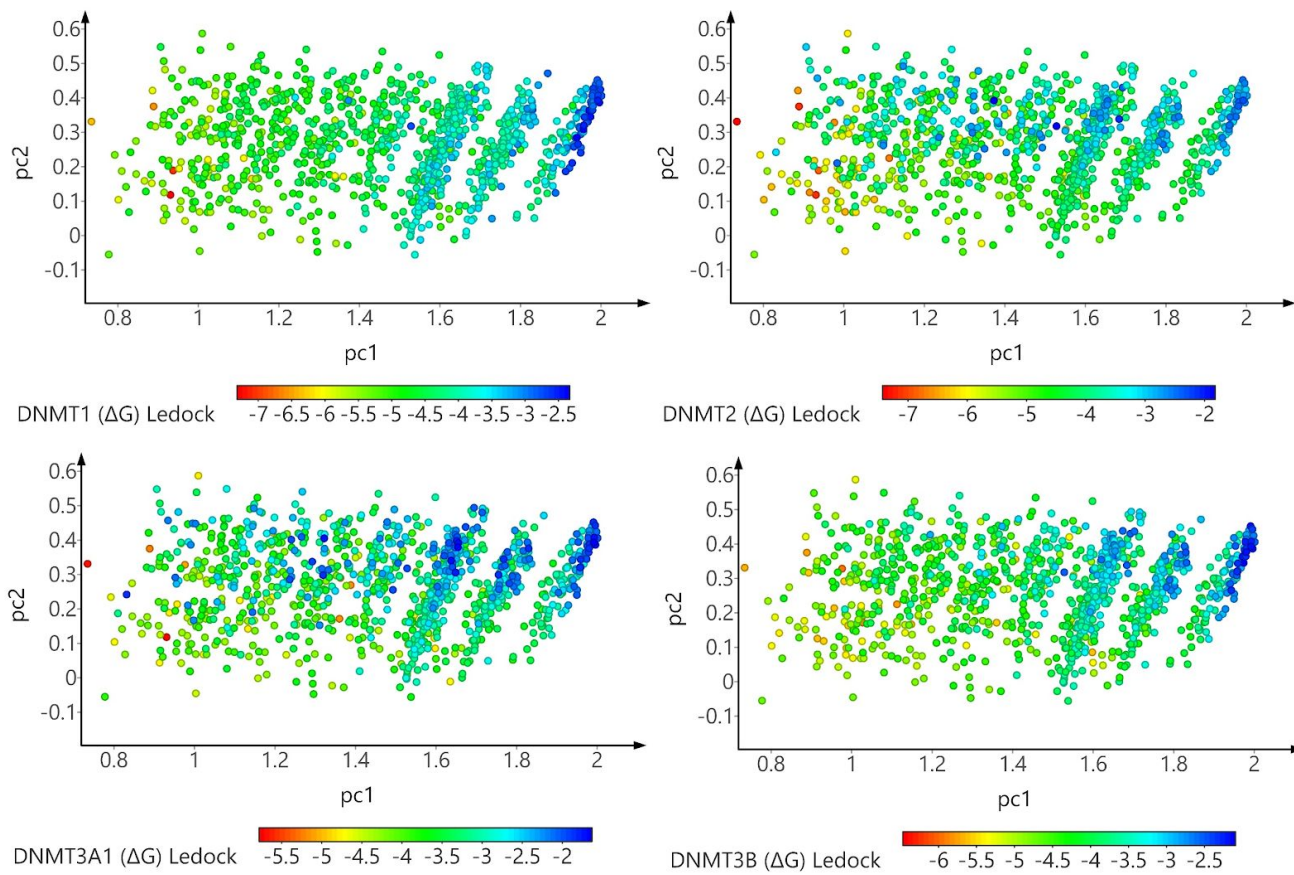
\*\*Se observan las 6 componentes del eigenvector asociado a PC2.



**Fig. 31.** Representación gráfica del espacio químico de los compuestos que conforman la base de datos FooDB.



**Fig. 32.** Representación gráfica del espacio químico de los 1455 compuestos “tipo-fármaco” de FooDB. Se representa con una escala de color, el valor del  $\Delta G$  de unión obtenido para cada compuesto en cada una de las cuatro isoformas con las que se llevó a cabo el estudio de acoplamiento molecular con MOE.



**Fig. 33.** Representación gráfica del espacio químico de los 1455 compuestos “tipo-fármaco” de FooDB. Se representa con una escala de color, el valor del  $\Delta G$  de unión obtenido para cada compuesto en cada una de las cuatro isoformas con las que se llevó a cabo el estudio de acoplamiento molecular con Ledock.

## 7. Conclusiones

- 1.a. Se hizo la construcción y validación de un modelo por homología que representa el dominio catalítico de las isoformas catalíticamente activas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4.
- 1.b. No se hizo la construcción del modelo por homología del dominio catalítico de DNMT3A2 debido a que posee un 100 % de identidad con respecto al dominio catalítico de DNMT3A1.
- 1.c. No se realizó la construcción del modelo por homología del dominio catalítico de DNMT3B3, DNMT3B4 y DNMT3B5 debido a que son isoformas catalíticamente inactivas por encontrarse truncas a nivel del dominio catalítico.
2. Fueron identificados potenciales sitios de unión para moléculas “tipo-fármaco” en DNMT1, DNMT2, DNMT3A1, DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4.
3. Mediante el uso de acoplamiento molecular con los programas MOE y Ledock, se identificaron seis compuestos con propiedades “tipo-fármaco” en la base de datos pública de compuestos de alimentos, FooDB, como potenciales inhibidores de DNMT1, DNMT2, DNMT3A1, DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4.
4. Mediante análisis de componentes principales, utilizando seis propiedades fisicoquímicas de interés farmacéutico, se hizo una representación del espacio químico de los compuestos de la base de datos FooDB “tipo-fármaco” y con mayor afinidad por la cavidad donde se lleva a cabo la metilación del ADN de las isoformas DNMT1, DNMT2, DNMT3A1, DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4.

## 8. Perspectivas

1. Construir el modelo por homología de otros dominios de DNMT, diferentes al catalítico, cuya estructura aún no ha sido determinada experimentalmente.
2. Seleccionar alguno o varios de los potenciales sitios de unión para moléculas “tipo-fármaco” identificados y a través del diseño, síntesis y evaluación biológica de moléculas afines a estos sitios, determinar si pueden ser utilizados para el diseño de inhibidores de DNMT.
3. Sintetizar o adquirir los seis compuestos encontrados y el ensayo biológico de inhibición con las respectivas isoformas de DNMT:
  - Síntesis (Colaboración: Dr. Massimo Bertinaria, Universidad de Turín, Italia)
  - Ensayo biológico (Colaboración: Dra. Hyang-Min Byun, Universidad de Newcastle, UK).
4. Identificar en otras bases de datos moleculares, a compuestos que compartan el mismo espacio químico que los compuestos de la base de datos FooDB “tipo-fármaco” y con mayor afinidad por la cavidad donde se lleva a cabo la metilación del ADN, como potenciales inhibidores de DNMT.

## 9. Referencias

- Abdellaoui, H., Depreux, P., Lesieur, D., Pfeiffer, B. and Bontempelli, P. 1995. Synthesis of the Naphthalenic Bioisostere of the Anti-Migraine Drug Sumatriptan. *Synthetic communications* 25(9), pp. 1303–1311.
- Alawi, L.F., Emberesh, S.E., Owuor, B.A., et al. 2020. Effect of hyperglycemia and rosiglitazone on renal and urinary neprilysin in db/db diabetic mice. *Physiological reports* 8(3), p. e14364.
- Aldawsari, F.S., Aguayo-Ortiz, R., Kapilashrami, K., et al. 2016. Resveratrol-salicylate derivatives as selective DNMT3 inhibitors and anticancer agents. *Journal of Enzyme Inhibition and Medicinal Chemistry* 31(5), pp. 695–703.
- Anderson, A.C. 2003. The process of structure-based drug design. *Chemistry & Biology* 10(9), pp. 787–797.
- Banks, T.M., Clay, S.F., Glover, S.A. and Schumacher, R.R. 2016. Mutagenicity of N-acyloxy-N-alkoxyamides as an indicator of DNA intercalation part 1: evidence for naphthalene as a DNA intercalator. *Organic & Biomolecular Chemistry* 14(15), pp. 3699–3714.
- Baquero, E., Quiñones, W., Ribon, W., Caldas, M.L., Sarmiento, L. and Echeverri, F. 2011. Effect of an Oxadiazoline and a Lignan on Mycolic Acid Biosynthesis and Ultrastructural Changes of Mycobacterium tuberculosis. *Tuberculosis research and treatment* 2011, p. 986409.
- Berman, H.M., Westbrook, J., Feng, Z., et al. 2000. The protein data bank. *Nucleic Acids Research* 28(1), pp. 235–242.
- Bigucci, F., Abruzzo, A., Cerchiara, T., Gallucci, M.C. and Luppi, B. 2015. Formulation of cellulose film containing permeation enhancers for prolonged delivery of propranolol hydrochloride. *Drug Development and Industrial Pharmacy* 41(6), pp. 1017–1025.
- Borodovsky, A., Salmasi, V., Turcan, S., et al. 2013. 5-azacytidine reduces methylation, promotes differentiation and induces tumor regression in a patient-derived IDH1 mutant glioma xenograft. *Oncotarget* 4(10), pp. 1737–1747.
- Braicu, C., Mehterov, N., Vladimirov, B., et al. 2017. Nutrigenomics in cancer: Revisiting the effects of natural compounds. *Seminars in Cancer Biology* 46, pp. 84–106.
- Cavasotto, C.N. and Phatak, S.S. 2009. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* 14(13–14), pp. 676–683.
- Chemical Computing Group ULC 2018. *Molecular Operating Environment*. Montreal, QC, Canada.
- Chen, S., Wang, Y., Zhou, W., et al. 2014. Identifying novel selective non-nucleoside DNA methyltransferase 1 inhibitors through docking-based virtual screening. *Journal of Medicinal Chemistry* 57(21), pp. 9028–9041.
- Choi, S.H., Heo, K., Byun, H.-M., An, W., Lu, W. and Yang, A.S. 2011. Identification of preferential target sites for human DNA methyltransferases. *Nucleic Acids Research* 39(1), pp. 104–118.
- Datta, J., Ghoshal, K., Denny, W.A., et al. 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), pp. 4277–4285.



- Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H. and Cheng, X. 2001. Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. *Nucleic Acids Research* 29(2), pp. 439–448.
- Erdmann, A., Arimondo, P.B. and Guianvarc'h, D. 2016. Structure-Guided Optimization of DNA Methyltransferase Inhibitors. In: *Epi-Informatics*. Elsevier, pp. 53–73.
- Erdmann, A., Halby, L., Fahy, J. and Arimondo, P.B. 2015. Targeting DNA methylation with small molecules: what's next? *Journal of Medicinal Chemistry* 58(6), pp. 2569–2583.
- Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103(19), p. 8577.
- Estey, E.H. 2013. Epigenetics in clinical practice: the examples of azacitidine and decitabine in myelodysplasia and acute myeloid leukemia. *Leukemia* 27(9), pp. 1803–1812.
- Fozza, C. 2019. Deciphering the prognostic significance of autoimmune disorders in myelodysplastic syndromes. *Annals of Hematology* 98(4), pp. 1025–1026.
- Georgiev, L., Chochkova, M., Totseva, I., et al. 2013. Anti-tyrosinase, antioxidant and antimicrobial activities of hydroxycinnamoylamides. *Medicinal chemistry research : an international journal for rapid communications on design and mechanisms of action of biologically active agents* 22(9), pp. 4173–4182.
- Gerber, P.R. and Müller, K. 1995. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *Journal of Computer-Aided Molecular Design* 9(3), pp. 251–268.
- Gleeson, M.P. 2008. Generation of a set of simple, interpretable ADMET rules of thumb. *Journal of Medicinal Chemistry* 51(4), pp. 817–834.
- Gnyszka, A., Jastrzebski, Z. and Flis, S. 2013. DNA methyltransferase inhibitors and their emerging role in epigenetic therapy of cancer. *Anticancer Research* 33(8), pp. 2989–2996.
- Gordon, C.A., Hartono, S.R. and Chédin, F. 2013. Inactive DNMT3B splice variants modulate de novo DNA methylation. *Plos One* 8(7), p. e69486.
- Guianvarc'h, D. and Arimondo, P.B. 2014. Challenges in developing novel DNA methyltransferases inhibitors for cancer therapy. *Future medicinal chemistry* 6(11), pp. 1237–1240.
- Houston, D.R. and Walkinshaw, M.D. 2013. Consensus docking: improving the reliability of docking in a virtual screening context. *Journal of Chemical Information and Modeling* 53(2), pp. 384–390.
- Hughes, J.D., Blagg, J., Price, D.A., et al. 2008. Physicochemical drug properties associated with in vivo toxicological outcomes. *Bioorganic & Medicinal Chemistry Letters* 18(17), pp. 4872–4875.
- Iwahi, T., Satoh, H., Nakao, M., et al. 1991. Lansoprazole, a novel benzimidazole proton pump inhibitor, and its related compounds have selective activity against *Helicobacter pylori*. *Antimicrobial Agents and Chemotherapy* 35(3), pp. 490–496.
- Jeltsch, A., Ehrenhofer-Murray, A., Jurkowski, T.P., et al. 2017. Mechanism and biological role of Dnmt2 in Nucleic Acid Methylation. *RNA Biology* 14(9), pp. 1108–1123.
- Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. and Cheng, X. 2007. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449(7159), pp. 248–251.

- Joshi, M., Rajpathak, S.N., Narwade, S.C. and Deobagkar, D. 2016. Ensemble-Based Virtual Screening and Experimental Validation of Inhibitors Targeting a Novel Site of Human DNMT1. *Chemical Biology & Drug Design* 88(1), pp. 5–16.
- Jurkowska, R.Z., Jurkowski, T.P. and Jeltsch, A. 2011. Structure and function of mammalian DNA methyltransferases. *Chembiochem* 12(2), pp. 206–222.
- Kaminskas, E., Farrell, A.T., Wang, Y.-C., Sridhara, R. and Pazdur, R. 2005. FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *The Oncologist* 10(3), pp. 176–182.
- Kantarjian, H., Issa, J.-P.J., Rosenfeld, C.S., et al. 2006. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer* 106(8), pp. 1794–1803.
- Khalaf, A.I., Anthony, N., Breen, D., et al. 2011. Amide isosteres in structure-activity studies of antibacterial minor groove binders. *European Journal of Medicinal Chemistry* 46(11), pp. 5343–5355.
- Kim, D.E., Chivian, D. and Baker, D. 2004. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research* 32(Web Server issue), pp. W526-31.
- Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. 1994. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76(2), pp. 357–369.
- Konya, K. 2001. Antioxidant properties of 8-O-4'-neolignans. *Phytomedicine* 8(6), pp. 454–459.
- Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A. and Vriend, G. 2004. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* 57(4), pp. 678–683.
- Krieger, E., Joo, K., Lee, Jinwoo, et al. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* 77 Suppl 9, pp. 114–122.
- Krieger, E. and Vriend, G. 2015. New ways to boost molecular dynamics simulations. *Journal of Computational Chemistry* 36(13), pp. 996–1007.
- Kuck, D., Caulfield, T., Lyko, F. and 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), pp. 3015–3023.
- Kuck, D., Singh, N., Lyko, F. and Medina-Franco, J.L. 2010. Novel and selective DNA methyltransferase inhibitors: Docking-based virtual screening and experimental evaluation. *Bioorganic & Medicinal Chemistry* 18(2), pp. 822–829.
- Lavoie, G. and St-Pierre, Y. 2011. Phosphorylation of human DNMT1: implication of cyclin-dependent kinases. *Biochemical and Biophysical Research Communications* 409(2), pp. 187–192.
- Lee, B. and Muller, M.T. 2009. SUMOylation enhances DNA methyltransferase 1 activity. *The Biochemical Journal* 421(3), pp. 449–461.
- Lee, W.J., Shim, J.-Y. and Zhu, B.T. 2005. Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Molecular Pharmacology* 68(4), pp. 1018–1030.
- Lee, W.J. and Zhu, B.T. 2006. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. *Carcinogenesis* 27(2), pp. 269–277.

- Le Guilloux, V., Schmidtke, P. and Tuffery, P. 2009. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* 10, p. 168.
- Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Journal of Pharmaceutical Sciences* 90(12), pp. 1646–1660. The article was originally published in *Advanced Drug Delivery Reviews* 23 (1997) 3–25.1. *Advanced Drug Delivery Reviews* 46(1–3), pp. 3–26.
- di Luccio, E. and Koehl, P. 2011. A quality metric for homology modeling: the H-factor. *BMC Bioinformatics* 12, p. 48.
- Lyko, F. 2018. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nature Reviews. Genetics* 19(2), pp. 81–92.
- Madeira, F., Park, Y.M., Lee, J., et al. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* 47(W1), pp. W636–W641.
- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. and Simmerling, C. 2015. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* 11(8), pp. 3696–3713.
- Maugeri, A., Mazzone, M.G., Giuliano, F., et al. 2018. Curcumin modulates DNA methyltransferase functions in a cellular model of diabetic retinopathy. *Oxidative medicine and cellular longevity* 2018, p. 5407482.
- Medina-Franco, J.L., Méndez-Lucio, O., Dueñas-González, A. and Yoo, J. 2015. Discovery and development of DNA methyltransferase inhibitors using in silico approaches. *Drug Discovery Today* 20(5), pp. 569–577.
- Melo, A., Artigas, C.G., Fritz, C., et al. 2014. Hipermetilación del gen supresor de tumores p53 en pacientes pediátricos con leucemia linfoblástica aguda. *International Journal of Morphology* 32(4), pp. 1243–1247.
- Miteva, M.A., Lee, W.H., Montes, M.O. and Villoutreix, B.O. 2005. Fast structure-based virtual ligand screening combining FRED, DOCK, and Surflex. *Journal of Medicinal Chemistry* 48(19), pp. 6012–6022.
- Mohammad, A.A. 2018. Myelodysplastic syndrome from theoretical review to clinical application view. *Oncology reviews* 12(2), p. 397.
- Naveja, J.J., Rico-Hidalgo, M.P. and Medina-Franco, J.L. 2018. Analysis of a large food chemical database: chemical space, diversity, and complexity [version 1; peer review: 2 approved, 1 approved with reservations]. *F1000Research* 7, p. 993.
- Ostler, K.R., Davis, E.M., Payne, S.L., et al. 2007. Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. *Oncogene* 26(38), pp. 5553–5563.
- Pagadala, N.S., Syed, K. and Tuszynski, J. 2017. Software for molecular docking: a review. *Biophysical reviews* 9(2), pp. 91–102.
- Peng, L., Yuan, Z., Ling, H., et al. 2011. SIRT1 deacetylates the DNA methyltransferase 1 (DNMT1) protein and alters its activities. *Molecular and Cellular Biology* 31(23), pp. 4720–4734.
- Preto, J. and Gentile, F. 2019. Assessing and improving the performance of consensus docking

strategies using the DockBox package. *Journal of Computer-Aided Molecular Design* 33(9), pp. 817–829.

Prieto-Martínez, F.D., Peña-Castillo, A., Méndez-Lucio, O., Fernández-de Gortari, E. and Medina-Franco, J.L. 2016. Molecular modeling and chemoinformatics to advance the development of modulators of epigenetic targets: A focus on DNA methyltransferases. *Advances in protein chemistry and structural biology* 105, pp. 1–26.

Rajavelu, A., Tulyasheva, Z., Jaiswal, R., Jeltsch, A. and Kuhnert, N. 2011. The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochemistry* 12, p. 16.

Rathert, P., Dhayalan, A., Murakami, M., et al. 2008. Protein lysine methyltransferase G9a acts on non-histone targets. *Nature Chemical Biology* 4(6), pp. 344–346.

Rombouts, F.J.R., Tovar, F., Austin, N., Tresadern, G. and Trabanco, A.A. 2015. Benzazaborinines as novel bioisosteric replacements of naphthalene: propranolol as an example. *Journal of Medicinal Chemistry* 58(23), pp. 9287–9295.

Rona, G.B., Eleutherio, E.C.A. and Pinheiro, A.S. 2016. PWWP domains and their modes of sensing DNA and histone methylated lysines. *Biophysical reviews* 8(1), pp. 63–74.

Saldívar-González, F.I., Gómez-García, A., Chávez-Ponce de León, D.E., et al. 2018. Inhibitors of DNA methyltransferases from natural sources: A computational perspective. *Frontiers in pharmacology* 9, p. 1144.

Sharif, H., Wang, L., Wang, W.L., et al. 2019. Structural mechanism for NEK7-licensed activation of NLRP3 inflammasome. *Nature* 570(7761), pp. 338–343.

Siedlecki, P., Garcia Boy, R., Musch, T., et al. 2006. Discovery of two novel, small-molecule inhibitors of DNA methylation. *Journal of Medicinal Chemistry* 49(2), pp. 678–683.

Soga, S., Shirai, H., Kobori, M. and Hirayama, N. 2007. Use of amino acid composition to predict ligand-binding sites. *Journal of Chemical Information and Modeling* 47(2), pp. 400–406.

Song, N., Liu, Z.-S., Xue, W., et al. 2017. NLRP3 phosphorylation is an essential priming event for inflammasome activation. *Molecular Cell* 68(1), pp. 185–197.e6.

The Metabolomics Innovation Centre 2017. The Metabolomics Innovation Centre: FooDB (Version 1) [Online]. Available at: <https://foodb.ca/> [Accessed: 12 March 2020].

UniProt Consortium 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* 47(D1), pp. D506–D515.

Veber, D.F., Johnson, S.R., Cheng, H.-Y., Smith, B.R., Ward, K.W. and Kopple, K.D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry* 45(12), pp. 2615–2623.

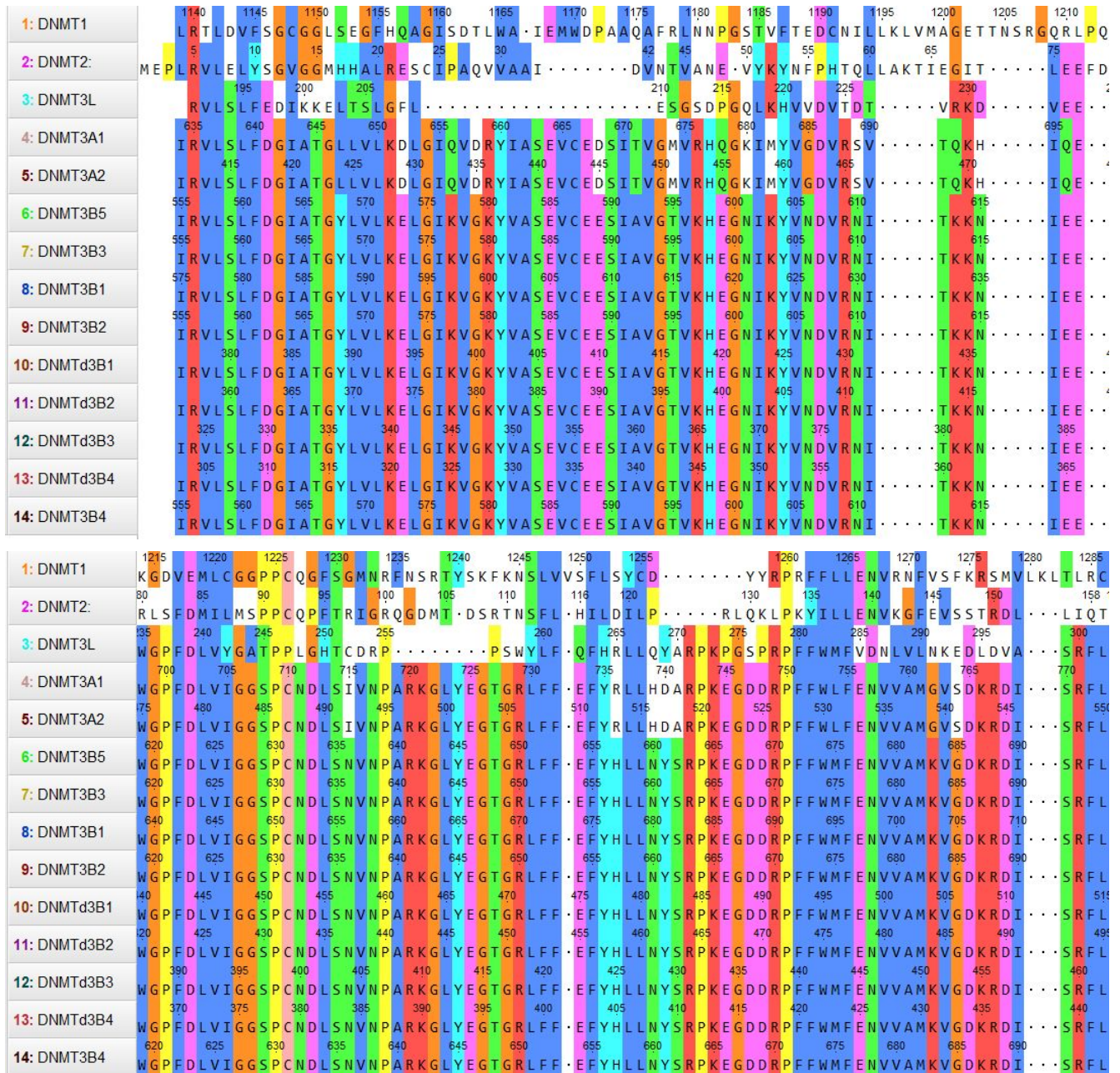
Wang, L., Wang, J., Sun, S., et al. 2006. A novel DNMT3B subfamily, DeltaDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer. *International Journal of Oncology* 29(1), pp. 201–207.

Wei, J., Li, G., Dang, S., Zhou, Y., Zeng, K. and Liu, M. 2016. Discovery and validation of

- hypermethylated markers for colorectal cancer. *Disease markers* 2016, p. 2192853.
- Yoo, J. and Medina-Franco, J.L. 2011. Homology modeling, docking and structure-based pharmacophore of inhibitors of DNA methyltransferase. *Journal of Computer-Aided Molecular Design* 25(6), pp. 555–567.
- Young, J.I., Züchner, S. and Wang, G. 2015. Regulation of the epigenome by vitamin C. *Annual review of nutrition* 35, pp. 545–564.
- Young Sim, J. and Ae Kim, M. 2014. Acetylcholinesterase Inhibitors from the Stem of *Zea mays*. *Natural Product Sciences* 20(1), pp. 13–16.
- Zacchino, S., Rodríguez, G., Pezzenati, G., Orellana, G., Enriz, R. and Gonzalez Sierra, M. 1997. In vitro evaluation of antifungal properties of 8.O.4'-neolignans. *Journal of Natural Products* 60(7), pp. 659–662.
- Zeidan, A.M., Shallis, R.M., Wang, R., Davidoff, A. and Ma, X. 2019. Epidemiology of myelodysplastic syndromes: Why characterizing the beast is a prerequisite to taming it. *Blood Reviews* 34, pp. 1–15.
- Zhang, N. and Zhao, H. 2016. Enriching screening libraries with bioactive fragment space. *Bioorganic & Medicinal Chemistry Letters* 26(15), pp. 3594–3597.
- Zhang, X.-Y. and Elfarra, A.A. 2017. Potential roles of myeloperoxidase and hypochlorous acid in metabolism and toxicity of alkene hydrocarbons and drug molecules containing olefinic moieties. *Expert Opinion on Drug Metabolism & Toxicology* 13(5), pp. 513–524.
- Zhang, Z.-M., Liu, S., Lin, K., et al. 2015. Crystal structure of human DNA methyltransferase 1. *Journal of Molecular Biology* 427(15), pp. 2520–2531.
- Zhang, Z.-M., Lu, R., Wang, P., et al. 2018. Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* 554(7692), pp. 387–391.
- Zhang, Z., Chen, Y., Tang, J. and Xie, X. 2014. Frequent loss expression of *dab2* and promoter hypermethylation in human cancers: a meta-analysis and systematic review. *Pakistan journal of medical sciences quarterly* 30(2), pp. 432–437.

# 10. Apéndice

## 10.1 Alineamiento de la secuencia primaria de aminoácidos del dominio catalítico

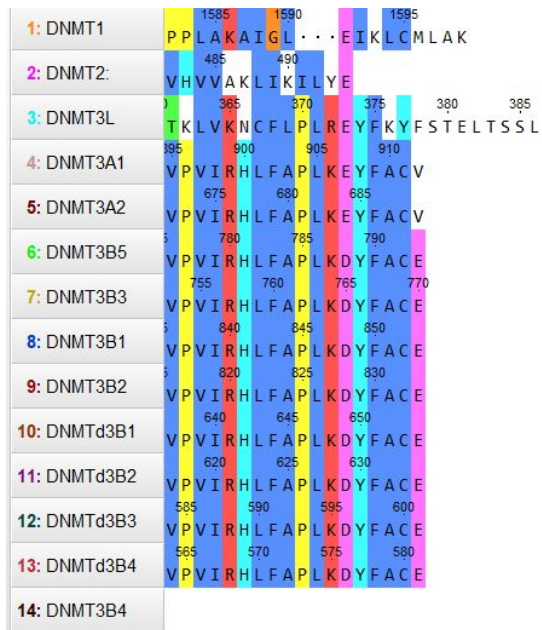




1: DNMT1	1425	1430	1435	1440	1445	1450	1455	1460	1465	1470	1475	1480	1485	1490	1495	1500
	RMRH	PLAP	GSDWR	LDLPI	NEVRL	SDGTM	ARKLRY	THHDK	RKNG	RSSSG	ALRG	VCS	CVEAG	KACD	PAARQ	FNTLIPW
2: DNMT2:	325	330	335	340	345	350						375	380	385		400
	DVNQ	YLLP	PKSLL	RYALL	L·DIVQ	PTCR	·	·	·	·	·	·	·	·	·	·
3: DNMT3L						E										
4: DNMT3A1						KLE					815	820	825			
						590					LQ	ECLE	HGR	IAK	FS	
5: DNMT3A2						KLE						595	600	605		
						735					LQ	ECLE	HGR	IAK	FS	
6: DNMT3B5						KLE						740	745			
											LQ	DCLE	YNRI	AKL		
7: DNMT3B3																
8: DNMT3B1						755						760	765			770
						KLE					LQ	DCLE	YNRI	AKL		K
9: DNMT3B2						735						740	745			750
						KLE					LQ	DCLE	YNRI	AKL		K
10: DNMTd3B1						555						560	565	570		
						KLE					LQ	DCLE	YNRI	AKL		K
11: DNMTd3B2						535						540	545	550		
						KLE					LQ	DCLE	YNRI	AKL		K
12: DNMTd3B3											505	510	515			
						KLE					LQ	DCLE	YNRI	AKL		K
13: DNMTd3B4											485	490	495			
						KLE					LQ	DCLE	YNRI	AKL		K
14: DNMT3B4																

1: DNMT1	1505	1510	1515	1520	1525	1530	1535	1540	1545	1550	1555	1560	1565	1570	1575	1580
	GNRH	NHWAG	LYGR	LEWD	GFFS	TTVT	NP	EP	MGKQ	GRVL	HE	QHRV	VS	VR	ECAR	SQGF
2: DNMT2:	405	415	420	425	432	435	440	445	450	455	460	465	470	475	480	
	GSVL	QT	·	·	AEDV	QVEN	IYKSL	TNLS	·	QEEQ	ITKL	LILK	LR	YFT	PK	IEAN
3: DNMT3L																ELSL
																LAQNK
4: DNMT3A1				830	835	840	845	850	855	860	865	870	875	880	885	890
				VRT	IT	TR	SNS	IK	GK	DQ	HFP	VFM	NK	ED	IL	WCT
5: DNMT3A2				810	815	820	825	830	835	840	845	850	855	860	865	870
				VRT	IT	TR	SNS	IK	GK	DQ	HFP	VFM	NK	ED	IL	WCT
6: DNMT3B5												750	755	760	765	770
												RIF	GFP	VHY	TD	VSNM
7: DNMT3B3												725	730	735	740	745
												RIF	GFP	VHY	TD	VSNM
8: DNMT3B1				775	780	785	790	795	800	805		810	815	820	825	830
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL
9: DNMT3B2				755	760	765	770	775	780	785		790	795	800	805	810
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL
10: DNMTd3B1				575	580	585	590	595	600	605		610	615	620	625	630
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL
11: DNMTd3B2				555	560	565	570	575	580	585		590	595	600	605	610
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL
12: DNMTd3B3				520	525	530	535	540	545	550		555	560	565	570	575
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL
13: DNMTd3B4				500	505	510	515	520	525	530		535	540	545	550	555
				VQ	IT	TK	SNS	IK	GK	NQ	LF	PV	VM	NG	KED	VL





**Fig S1.** Alineamiento de la secuencia primaria de aminoácidos del dominio catalítico de 14 isoformas de DNMT

**10.2 Aminoácidos que conforman el sitio donde se lleva a cabo la metilación del ADN y otros potenciales sitios de unión para moléculas “tipo-fármaco”**

**Tabla S1. Aminoácidos que conforman los potenciales sitios de unión identificados**

Isoforma	Número de esferas $\alpha$	PLB	No. átomos hidrofóbicos	No. átomos de la cadena lateral	Aminoácidos
DNMT1 (Sitio de metilación)  Código PDB ID: 4WXX	191	1.26	36	111	Asn519, Ser520, Asp521, Thr523, Glu525, Asp256, Thr616, Met1169, Trp1170, Asp1190, Gly1123, Pro1224, Pro1225, Cys1226, Gln1227, Gly1228, Phe1229, Ser1241, Lys1242, Asn1245, Ser1246, Leu1247, Val1248, Glu1266, Asn1267, Val1268, Arg1310, Arg1312, Thr1525, Thr1526, Thr1528, Lys1535, Gln1536, Arg1574, Gly1577, Asn1578, Ala1579, Val1580
DNMT1  Código PDB ID: 4WXX  (Fig. 20, cavidad azul)	168	2.51	27	80	Asp423, Thr424, Gly425, Leu426, Glu428, Lys429, Lys505, Leu551, Arg552, His553, Ala554, Gln555, Phe556, Glu559, Ile585, Lys586, Ala558, Gly589, Val590, Thr591, Gln594, Arg595, Cys1476, Cys1478, Lys1483, Ala1484, Cys1485, Asp1486, Pro1487, Arg1490, Trp1498, His1502, Thr1503, Arg1506, His1507
DNMT1  Código PDB ID: 4WXX	220	2.28	60	120	Thr621, Leu623, Gln626, Ile627, Thr630, Phe631, Glu634, Gln635, Ala807, Thr808, Ser809, Asp810, Pro811, Glu813, Ile835,

(Fig. 20, cavidad verde)					Tyr836, Lys837, Ala838, Pro839, Ser840, Trp843, Glu846, Gly847, Gly848, Met849, Pro851, Leu854, Leu855, Glu856, Asp859, Gly860, Lys861, Thr862, Tyr863, Phe864, Gln866, Leu867, Ser878, Pro879, Lys881, Asp1256, Arg1259, Leu1282, Arg1285, Val1288, Arg1289, Met1290, Gly1291, Gln1293, Pro1320, Gly1321
DNMT2 (Sitio de metilación)  Código PDB ID: 1G55	64	2.59	11	40	Ser76, Pro77, Pro78, Glu119, Asn120, Val121, Arg160, Arg162, Cys292, Phe293, Thr294, Lys295, Gly374, Asn375, Ser376
DNMT2 (Dominio catalítico)  Código PDB ID: 1G55  (Fig. 21, cavidad verde)	85	3.06	24	55	His18, Arg21, Glu22, Asn46, Phe47, Pro48, His49, Thr50, Asp255, Phe256, Leu257, Glu258, Lys346, Asn350, Gly353, Phe354, Pro355, Pro356, Glu357, Phe358, Val379, His380
DNMT3A1 (Sitio de metilación)  Código PDB ID: 5YX2	107	2.38	15	49	Gly707, Ser708, Pro709, Cys710, Asn711, Ser714, Val716, Glu756, Asn757, Val758, Ala760, Arg790, Arg792, Thr832, Ile833, Thr834, Thr835, Arg836, Gly890, Arg891
DNMT3A1 (Dominio catalítico)  PDB ID: 5YX2	22	2.13	13	29	Tyr735, Leu738, His739, Arg742, Trp753, Phe772, Leu773, Glu774, Ser775, Asn797

(Fig. 22 y Fig. 23 cavidad azul)					
DNMT3A1 (Dominio catalítico)  PDB ID: 5YX2  (Fig. 22 y Fig. 23 cavidad verde)	72	1.98	18	36	Leu648, Asp652, Val785, Ser786, Ala787, Thr808, Asn810, Asp811, Lys812, Cys818, Arg866, Val867, Gly869, Phe870, Pro871, Val895, Pro896, Arg899
Modelo por homología <sup>a</sup> (Sitio de metilación)	42	3.52	18	39	Gly648, Ser649, Pro650, Cys651, Asn652, Glu697, Asn698, Val699, Arg733, Ile774, Thr775, Thr776, Gly83, Arg832, Trp834
Modelo por homología (Dominio catalítico)  (Fig. 24, cavidad verde)	94	2.67	23	50	Glu593, Leu594, Val726, Ser726, Ala729, Arg731, Ile747, Ala748, Ser749, Asn751, Asp752, Lys753, Arg807, Ile808, Phe809, Gly810, Phe811, Pro812, Val813, Pro837, Val838, Arg840, His841
Modelo por homología (Dominio catalítico)  (Fig. 24 y Fig. 25 cavidad magenta)	69	2.63	25	55	Arg576, Lys633, Ile636, Glu637, Gly640, Pro641, Phe642, Asp643, Asn680, Tyr681, Ser682, Arg683, Pro684, Lys685, Glu686, Asp688, Arg690, Pro692

<sup>a</sup> El modelo por homología representa el dominio catalítico de las isoformas DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3 y DNMTΔ3B4. La numeración de la secuencia primaria de aminoácidos es respecto al dominio catalítico de DNMT3B1.

### 10.3 Estudio *in silico* de NLRP3: Modelado por homología e identificación de potenciales inhibidores

Este trabajo se llevó a cabo en la Universidad de Turín, Italia, bajo la tutoría del Dr. Massimo Bertinaria durante el periodo de septiembre del 2019 a diciembre del 2019.

Los inflamasomas son complejos supramoleculares citoplasmáticos involucrados en la respuesta a las invasiones microbianas exógenas y las señales de daño endógeno. NLRP3 (por sus siglas en inglés: NACHT, LRR and PYD domains-containing protein 3) es una proteína clave que desempeña un papel fundamental en la activación del inflamasoma. La sobreactivación de NLRP3 está directamente relacionada con muchas enfermedades inflamatorias. Se ha demostrado que la inhibición de NLRP3 tiene potencial terapéutico para el tratamiento de enfermedades como Alzheimer, aterosclerosis, asma, gota y enfermedades inflamatorias del intestino (Song et al. 2017). Al día de hoy no ha sido publicada la estructura 3D completa de NLRP3, por lo tanto, es difícil diseñar inhibidores de NLRP3 mediante enfoques computacionales basados en la estructura de la proteína. Basándose en la reciente estructura publicada en *Nature*, de algunos dominios estructurales de NLRP3 (Sharif et al. 2019), se modeló la estructura completa del complejo NLRP3 y NEK7 (proteína implicada en la activación de NLRP3), para ello se utilizó el servidor en línea Robetta. En base a este modelo, se identificó un sitio de unión potencial para el compuesto de referencia MCC950, el cual es un potente inhibidor de NLRP3. Con el objetivo de encontrar moléculas afines hacia el sitio de unión descubierto, de un grupo de 120 817 compuestos “tipo-fármaco” de la base de datos ChEMBL, se seleccionaron 28 676 compuestos para llevar a cabo un estudio de cribado virtual. Adicionalmente, a partir de los 120 817 compuestos “tipo-fármaco”, empleando el coeficiente de Tanimoto, se identificó un grupo de moléculas estructuralmente similares a MCC950 mediante el uso de 2 diferentes huellas dactilares: “MACS keys” y “ECFP6” (por sus siglas en inglés: Extended Connectivity Fingerprint 6).

Los resultados de este trabajo se presentaron en el siguiente simposio: “Merck Young Chemists’ Symposium 2019”, que tuvo lugar en la ciudad de Rimini en Italia del 25 al 27 de noviembre. Adjunto se encuentra una copia del póster presentado. En el póster se encuentra un diagrama de flujo con la metodología empleada para cada actividad realizada. Se puede apreciar la imagen del modelo por homología construido **Fig. 1.A**, los parámetros de validación, **Fig. 1.B-D** y la representación en el espacio químico del conjunto de compuestos similares a MCC950 **Fig. 2**. Se encuentran resaltados los compuestos que fueron identificados con un mayor grado de similitud estructural a MCC950 y los compuestos con afinidad predicha igual o mayor hacia el sitio de unión identificado, de acuerdo al estudio de acoplamiento molecular (magenta) **Fig.2**.

#### Logros

1. Se generó el modelo por homología del complejo NLRP3-NEK7, sin puntos atípicos (*outliers*) en el diagrama de Ramachandran. Longitudes de enlace y ángulos del complejo proteico dentro de los valores esperados.
2. A partir del estudio de cribado virtual, se identificaron más de 60 compuestos con un  $\Delta G$  de unión igual o mejor al de MCC950.
3. Se identificó un área bien definida en el espacio químico en donde se encuentran los compuestos más similares a MCC950.



# In silico study of NLRP3: towards selective inhibition with small molecules

Alejandro Gómez García,<sup>a,b</sup> Simone Gastaldi,<sup>b</sup> Giulia Fumero,<sup>b</sup>

Massimo Bertinaria,<sup>b</sup> and José Luis Medina Franco<sup>a</sup>

<sup>a</sup> Department of Pharmacy, National Autonomous University of Mexico, Avenue University 3000, Mexico City, México

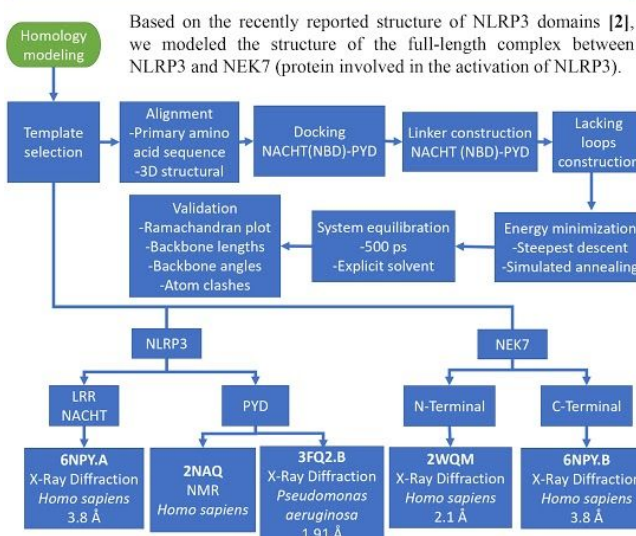
<sup>b</sup> Department of Drug Science and Technology, University of Turin, Via Pietro Giuria 9, Turin, Italy



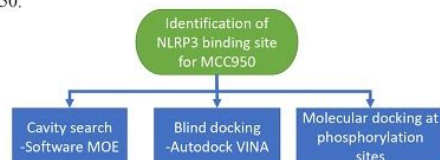
## Introduction

The Inflammasomes are cytoplasmic supramolecular complexes involved in the response to exogenous microbial invasions and endogenous damage signals. NLRP3 (NACHT, LRR and PYD domains-containing protein 3) is the key protein playing a pivotal role in the inflammasome activation. NLRP3 over-activation is directly connected with many inflammatory diseases. It has been shown that NLRP3 inhibition has therapeutic potential for the treatment of diseases such as Alzheimer's disease, atherosclerosis, asthma, gout and inflammatory bowel diseases [1]. Up to date, no experimental 3D structure of the full-length NLRP3 is available, therefore, it is difficult to design NLRP3 inhibitors by computational approaches.

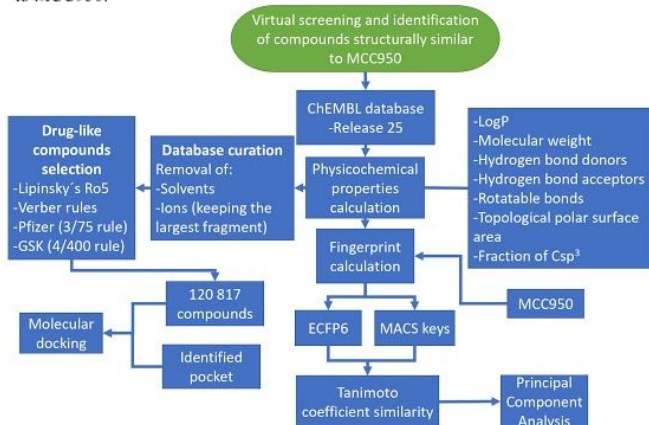
## Aim of the work



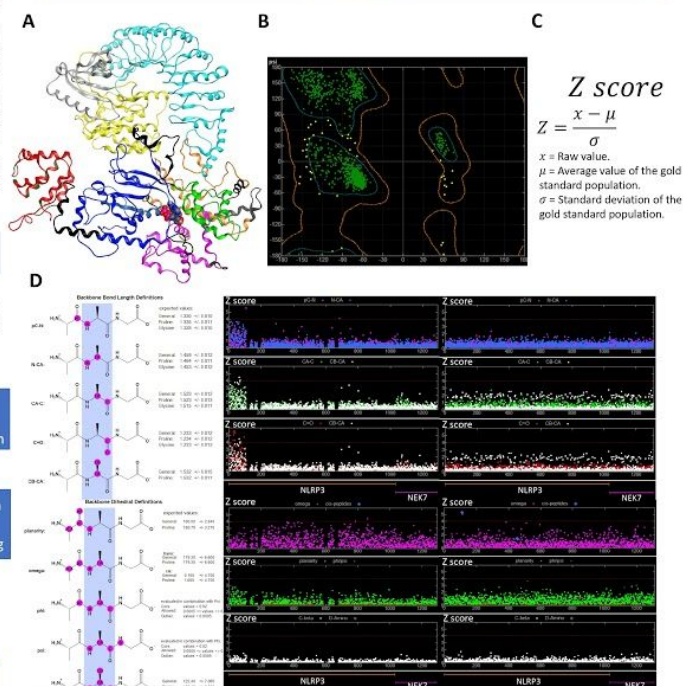
Based on this model, we identified a potential binding site for the reference compound MCC950.



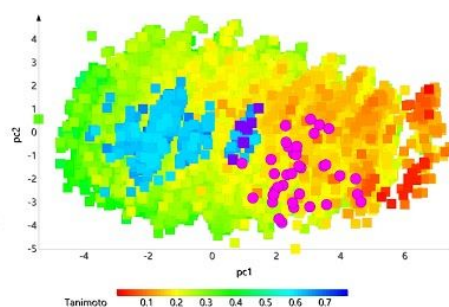
In order to find new binders for the discovered pocket we screened a selected set of drug-like substances and we identified a set of small molecules structurally similar to MCC950.



## Results



**Fig. 1.** A. Homology model of the full-length NLRP3 with the full-length NEK7. Leucine-rich-repeat "LRR" (light blue), nucleotide-binding domain (strong blue), helical domain 1 (green), winged helix domain (magenta), helical domain 2 (orange), pyrin domain "PYD" (red). NEK7 C-terminal (yellow) and N-terminal (gray). Depicted in green the contact surface of PYD with ASC. [3]. B. Ramachandran plot of the homology model. C. Definition of the Z score. D. Geometry inspection of the homology model, according to the Z score, angles and bonds definitions (first column), template profile (second column), homology model profile (third column).



**Fig. 2.** Representation of the chemical space of the set of 120,817 drug-like compounds. Color scale (red to blue) depicts the Tanimoto coefficient value of similarity compared to MCC950, with MACS keys fingerprint. Shown in purple, the compounds with the higher value of the Tanimoto coefficient with ECFP6. Depicted in magenta circles the compounds with predicted equal or higher affinity to the identified MCC950 binding pocket.

## Conclusions

- A full-length homology model of NLRP3-NEK7 was generated, with no outliers in the Ramachandran plot, backbone lengths and angles within expected values.
- More than 60 compounds with equal or better docking scores compared to MCC950 were identified from the virtual screening.
- It exists a well-defined area in the chemical space representation of the most similar compounds to MCC950.

## Bibliography

- [1] N. Song, Z.-S. Liu, W. Xue, Z.-F. Bai, Q.-Y. Wang, J. Dai, X. Liu, Y.-J. Huang, H. Cai, X.-Y. Zhan, Q.-Y. Han, H. Wang, Y. Chen, H.-Y. A.-J. Li, X.-M. Zhang, T. Zhou, T. Li. *Mol. Cell.* 68 (2017) 185-197.66
- [2] H. Sharif, L. Wang, W.L. Wang, V.G. Magupalli, L. Andreeva, Q. Qiao, A.V. Hansenstein, Z. Wu, G. Núñez, Y. Mao, and H. Wu. *Nature.* 570 (2019) 338-343.
- [3] J. Oroz, S. Barrera-Vilmaru, C. Alfonso, G. Rivas, E. de Alba. *J. Biol. Chem.* 291 (2016), 19487-19501.

## Contact information

Author: G-G. A. alex.go.gu21@hotmail.com  
Project leaders:  
UniTo: B. M. massimo.bertinaria@unito.it  
UNAM: M-F. J.L. medina@unam.mx  
Research group website: www.difacquim.com



## 10.4 Artículo publicado con parte de los resultados del estudio *in silico* de DNMT

De acceso libre en: <http://dx.doi.org/10.3389/fphar.2018.01144>

ResearchGate

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/328190081>

### Inhibitors of DNA Methyltransferases From Natural Sources: A Computational Perspective

Article in *Frontiers in Pharmacology* - October 2018

DOI: 10.3389/fphar.2018.01144

CITATION

1

READS

141

6 authors, including:



Fernanda I. Saldivar-Gonzalez  
Universidad Nacional Autónoma de México

9 PUBLICATIONS 16 CITATIONS

[SEE PROFILE](#)



Alejandro Gómez-García  
Universidad Nacional Autónoma de México

2 PUBLICATIONS 2 CITATIONS

[SEE PROFILE](#)



David Erick Chávez-Ponce de León  
Universidad Nacional Autónoma de México

1 PUBLICATION 1 CITATION

[SEE PROFILE](#)



B. Angélica Pílon-Jiménez

2 PUBLICATIONS 1 CITATION

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Activity landscape modeling and activity cliffs [View project](#)



Anti-infective compounds: design and development [View project](#)

All content following this page was uploaded by Jose L. Medina-Franco on 23 October 2018.

The user has requested enhancement of the downloaded file.



# Inhibitors of DNA Methyltransferases From Natural Sources: A Computational Perspective

Fernanda I. Saldivar-González, Alejandro Gómez-García, David E. Chávez-Ponce de León, Norberto Sánchez-Cruz, Javier Ruiz-Ríos, B. Angélica Pílon-Jiménez and José L. Medina-Franco\*

Department of Pharmacy, School of Chemistry, National Autonomous University of Mexico, Mexico City, Mexico

## OPEN ACCESS

### Edited by:

Shibashish Giri,  
University of Leipzig, Germany

### Reviewed by:

Jaigopal Sharma,  
Delhi Technological University, India  
Sujata Mohanty,  
All India Institute of Medical Sciences,  
India  
Rup Lal,  
University of Delhi, India

### \*Correspondence:

José L. Medina-Franco  
medinajl@unam.mx;  
jose.medina.franco@gmail.com

### Specialty section:

This article was submitted to  
Pharmacogenetics  
and Pharmacogenomics,  
a section of the journal  
Frontiers in Pharmacology

Received: 08 August 2018

Accepted: 21 September 2018

Published: 10 October 2018

### Citation:

Saldivar-González FI,  
Gómez-García A,  
Chávez-Ponce de León DE,  
Sánchez-Cruz N, Ruiz-Ríos J,  
Pílon-Jiménez BA and  
Medina-Franco JL (2018) Inhibitors  
of DNA Methyltransferases From  
Natural Sources: A Computational  
Perspective.  
Front. Pharmacol. 9:1144.  
doi: 10.3389/fphar.2018.01144

Naturally occurring small molecules include a large variety of natural products from different sources that have confirmed activity against epigenetic targets. In this work we review chemoinformatic, molecular modeling, and other computational approaches that have been used to uncover natural products as inhibitors of DNA methyltransferases, a major family of epigenetic targets with therapeutic interest. Examples of computational approaches surveyed in this work are docking, similarity-based virtual screening, and pharmacophore modeling. It is also discussed the chemoinformatic-guided exploration of the chemical space of naturally occurring compounds as epigenetic modulators which may have significant implications in epigenetic drug discovery and nutriepigenetics.

**Keywords:** chemical space, chemoinformatics, databases, DNMT inhibitors, drug discovery, molecular modeling, similarity searching, virtual screening

## SECTION 1: INTRODUCTION

Epigenetics has been defined as a change in phenotype without an underlying change in genotype (Berger et al., 2009). In the 1940s Waddington suggested the term “epigenetics” trying to describe “the interactions of genes with their environment, which brings the phenotype into being” (Waddington, 2012). Alterations in epigenetic modifications have been related to several diseases including cancer, diabetes, neurodegenerative disorders, and immune-mediated diseases (Dueñas-González et al., 2016; Tough et al., 2016; Hwang et al., 2017; Lu et al., 2018). Moreover, epigenetic targets are also attractive for the treatment of antiparasitic infections (Sacconnay et al., 2014).

In epigenetic drug discovery, epigenetic targets have been classified into three main groups (Ganesan, 2018). “Writers” are enzymes that catalyze the addition of a functional group to a protein or nucleic acid; “readers” are macromolecules that function as recognition units that can distinguish a native macromolecule vs. the modified one; and “erasers” that are enzymes that aid in the removal of chemical modifications introduced by the writers. Thus far, several targets from these three major families have reached different stages of drug discovery, ranging from lead discovery, preclinical development, clinical trials and approval. Currently, there are seven compounds approved for clinical use (Ganesan, 2018).

DNA methyltransferases (DNMTs) are a family of “writer” enzymes responsible for DNA methylation that is the addition of a methyl group to the carbon atom number five (C5) of cytosine. As surveyed in this work, since DNA methylation has an essential role for cell differentiation and



development, alterations in the function of DNMTs have been associated with cancer (Castillo-Aguilera et al., 2017) and other diseases (Lyko, 2017).

Several natural products have been identified as inhibitors of epigenetic targets including DNMTs. Most of these compounds have been uncovered fortuitously. However, there are recent efforts to screen systematically natural products as DNMT inhibitors. The vastness of the chemical space of natural products led to the hypothesis that many more active compounds could potentially be identified. Indeed, it has been estimated that more than 95% of the biodiversity in nature remains to be explored to identify potential bioactive molecules (Ho et al., 2018).

The aim of this work is to discuss a broad range of computational methods to identify novel inhibitors of DNMTs from natural products. The manuscript also discusses the chemical space of natural products as inhibitors of DNMTs. The manuscript is organized into nine sections. After this introduction, Section 2 reviews briefly the structure of DNMTs including different isoforms. The next section covers major aspects of the function of DNMTs including the mechanism of methylation. Section 4 reviews currently known inhibitors of DNMTs from natural sources including food chemicals. Section 5 discusses the epigenetic relevant chemical space of natural products comparing the chemical space of DNMT inhibitors from natural sources vs. other compounds. The next section reviews computational strategies that are used to identify natural compounds as potential epi-hits or epi-leads targeting DNMTs. Sections 7 and 8 presents Summary conclusions and Perspectives, respectively.

## SECTION 2: STRUCTURE OF DNMTs

The human genome encodes DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L. While DNMT1, DNMT3A, and DNMT3B have catalytic activity, DNMT2 and DNMT3L do not (Lyko, 2017). DNMT1 is a maintenance methyltransferase, responsible for duplicating the pattern of DNA methylation during replication. DNMT1 is essential for proper mammalian development and it has been proposed as the most interesting target for experimental cancer therapies (Dueñas-González et al., 2016). DNMT3A and DNMT3B are *de novo* methyltransferases. Human DNMT1 has 1616 amino acids whose structure can be divided into an N-terminal regulatory domain and a C-terminal catalytic domain (Jeltsch, 2002; Jurkowska et al., 2011). The N-terminal domain contains a replication foci-targeting domain, a DNA-binding CXXC domain, and a pair of bromo-adjacent homology domains. The C-terminal catalytic domain has 10 amino acid motifs. The cofactor and substrate binding sites in the C-terminal catalytic domain are comprised of motif I and X and motif IV, VI, and VIII, respectively (Lan et al., 2010). The target recognition domain which is maintained by motif IX and involved in DNA recognition, is not conserved between the DNMT family. **Figure 1A** shows a three-dimensional (3D) model of

a DNMT1 (PDB ID: 4WXX) (Zhang et al., 2015). **Figure 1B** shows a schematic diagram of human of DNMT1, 2, 3A, 3B, and L.

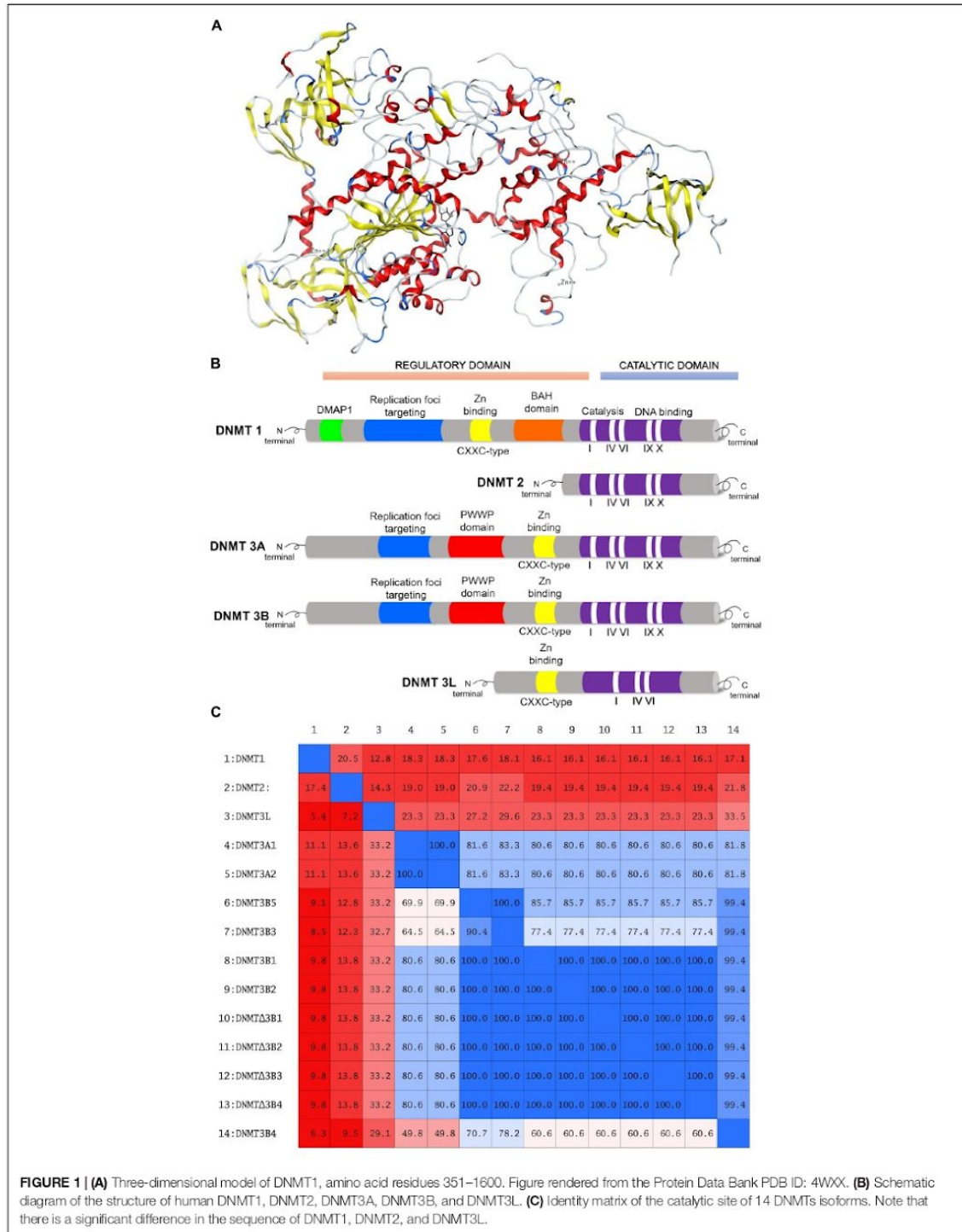
### Section 2.1: Isoforms

Two isoforms of DNMT3A have been identified, DNMT3A1 and DNMT3A2. At the N-terminal domain both isoforms have a PWWP (Pro-Trp-Trp-Pro) and an ADD (ATR-X-DNMT3-DNMT3L) domains (Jurkowska et al., 2011). The C-terminal domain is identical in the two isoforms (Choi et al., 2011).

There are more than 30 isoforms of DNMT3B, however, only DNMT3B1 and DNMT3B2 are catalytically active (Ostler et al., 2007). Similar to DNMT3A, DNMT3B1, and DNMT3B2 have a PWWP and ADD domains at the N-terminal region (Lyko, 2017). The rest of the isoforms are not catalytically active. Some of these such as DNMT3B3, DNMT3B4, and DNMT3B7 are overexpressed in many tumor cell lines (Gordon et al., 2013).  $\Delta$ DNMT3B has seven isoforms and lacks 200 amino acids from the N-terminal region of DNMT3B (Wang et al., 2006).  $\Delta$ DNMT3B1–4 possess catalytic activity whereas  $\Delta$ DNMT3B5–7 lacks the catalytic domain (Wang et al., 2006).  $\Delta$ DNMT3B is mainly expressed in non-small cell lung cancer (Wang et al., 2006; Ostler et al., 2007). **Figure 1C** shows the identity matrix of 14 DNMTs isoforms. The identity matrix indicates that the amino acid sequence at the catalytic site of DNMT3A1 and DNMT3A2 isoforms is identical. In the same manner, the amino acid sequence at the C-terminal domain of the catalytically active isoforms DNMT3B1, DNMT3B2, and  $\Delta$ DNMT3B1–4 are identical. DNMT1, DNMT2, and DNMT3L show a significant difference in the sequence of the catalytic site with respect to the rest of the isoforms. Therefore, it can be anticipated that is possible to identify or design selective inhibitors for these isoforms.

## SECTION 3: FUNCTION AND MECHANISM OF DNMTs

As outlined in Section 2, cytosine-5 DNMTs catalyze the addition of methylation marks to genomic DNA. All DNMTs have a related catalytic mechanism that is featured by the formation of a covalent adduct intermediate between the enzyme and the substrate base. All DNMTs use *S*-adenosyl-*L*-methionine (SAM) as the methyl group donor (Vilkaitis et al., 2001; Du et al., 2016). DNMT forms a complex with DNA and the cytosine which will be methylated flips out from the DNA (Klimasauskas et al., 1994). A conserved cysteine performs a nucleophilic attack to the six-position of the target cytosine yielding a covalent intermediate. The five-position of the cytosine is activated and conducts a nucleophilic attack on the cofactor SAM to form the 5-methyl covalent adduct and *S*-adenosyl-*L*-homocysteine (SAH). The attack on the six-position is aided by a transient protonation of the cytosine ring at the endocyclic nitrogen atom N3, which can be stabilized by a glutamate and arginine residues. The covalent complex between the methylated base and the DNA is resolved by deprotonation at the five-position to generate the methylated cytosine and the free enzyme.



## SECTION 4: KNOWN INHIBITORS OF DNMTs FROM NATURAL SOURCES

Thus far more than 500 compounds have been tested as inhibitors of DNMTs. The structural diversity and coverage in chemical space has been analyzed using chemoinformatic methods (Fernandez-de Gortari and Medina-Franco, 2015). The chemical space of DNMT inhibitors has been compared with inhibitors of other epigenetic targets (Naveja and Medina-Franco, 2018). Furthermore, the structure-activity relationships (SAR) of DNMT inhibitors using the concept of activity landscape has been documented (Naveja and Medina-Franco, 2015).

DNA methyltransferase inhibitors have been obtained from a broad number of different strategies including organic synthesis, virtual, and high-throughput screening (Medina-Franco et al., 2015). Organic synthesis has been employed in several instances for lead optimization (Castellano et al., 2008; Kabro et al., 2013; Davide et al., 2016). Natural products and food chemicals have also been a major source of active compounds. Natural products that are known to act as DNMT inhibitors or demethylating agents have been extensively reviewed by Zwergel et al. (2016). These natural products are of the type polyphenols, flavonoids, anthraquinones, and other classes. Some of the first natural products described were curcumin, (-)-epigallocatechin-3-gallate (EGCG), mahanine, genistein, and quercetin. Other natural products that have been described as inhibitors of DNMT or demethylating agents are silibinin, luteolin, kazinol Q, laccic acid, hypericin, boswellic acid, and lycopene. **Figure 2** shows the chemical structure of representative DNMT inhibitors with emphasis on compounds from natural origin.

The bioactivity profile and potency in enzymatic and/or cell-based assays of these natural products have been discussed in detail by Zwergel et al. (2016). Of note, it will be valuable if all natural products could have been screened under the same conditions. For few natural products the selectivity has been characterized being nanaomycin A an exception (*vide infra*). Indeed, for about eight natural products the  $IC_{50}$  has been measured in enzymatic based assays. Despite the fact that the potency of the natural products with DNMTs is not very high in enzymatic-based assays, e.g.,  $IC_{50}$  between 0.5 and 10  $\mu$ M, several natural products have shown promising activity in cell-based assays. Notably, natural products have distinct chemical scaffolds that could be used as a starting point in lead optimization efforts. Moreover, quercetin in combination with green tea extract has advanced into phase I clinical trials for the treatment of prostate cancer.

Most of the natural products with demethylating activity or ability to inhibit DNA methyltransferases in enzymatic assays have been identified fortuitously. However, as discussed in this work, there are efforts toward the identification of bioactive demethylating agents using systematic approaches such as virtual screening. Indeed, the natural product nanaomycin A (**Figure 2**) was identified from a virtual screening campaign initially focused on the identification of inhibitors of DNMT1. The quinone-based antibiotic isolated from *Streptomyces* showed antiproliferative effects in three human tumor cell lines, HCT116, A549, and HL60 after 72 h of treatment. Moreover, nanaomycin A showed

reduced global methylation levels in all three cell lines when tested at concentrations ranging from 0.5 to 5  $\mu$ M. Nanaomycin A reactivated the transcription of the RASSF1A tumor suppressor gene inducing its expression up to 18-fold at 5  $\mu$ M, higher than the reference drug 5-azacytidine (sixfold at 25  $\mu$ M). In an enzymatic inhibitory assay, nanaomycin A was selective toward DNMT3B with an  $IC_{50} = 0.50$   $\mu$ M.

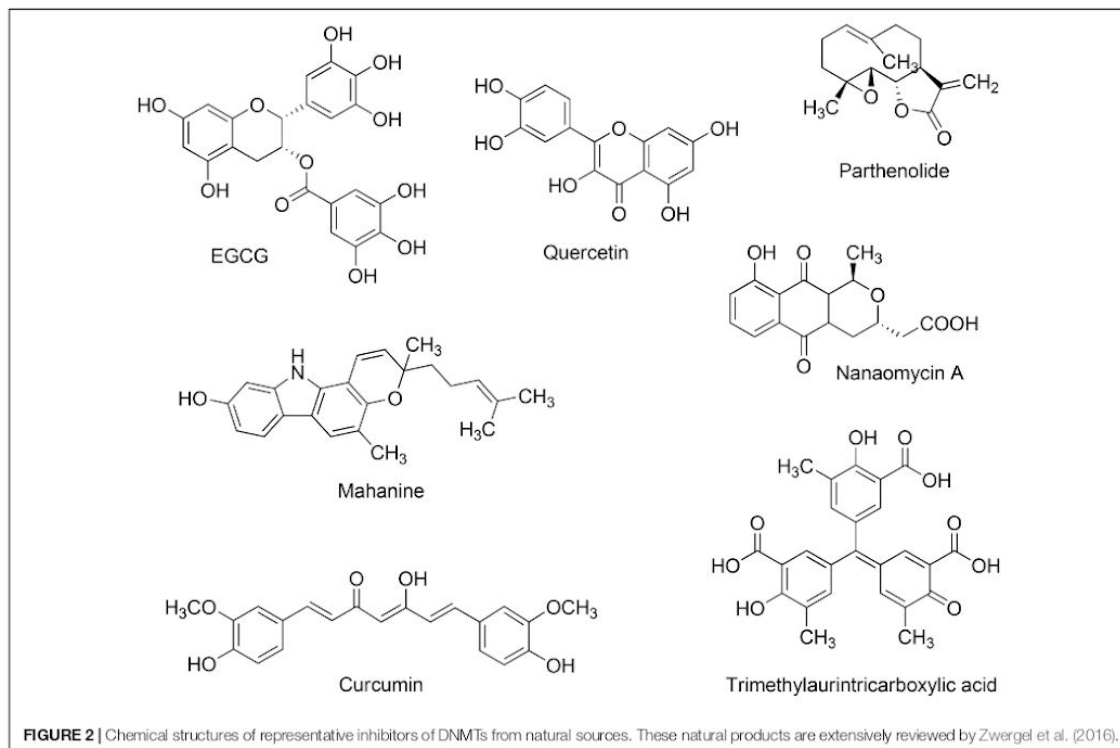
### Section 4.1: Natural Products and Food Chemicals

It is remarkable that several natural products are used as dietary sources such as curcumin, caffeic acid and chlorogenic acid found in *Coffea arabica*, genistein found in soybean, quercetin found in fruits, vegetables, and beverages. Of course, there is a large overlap between the chemical space of food chemicals and natural products (Naveja et al., 2018). This has given rise to systematically screen food chemical databases for potential regulators of epigenetic targets.

## SECTION 5: EPIGENETIC RELEVANT CHEMICAL SPACE OF NATURAL PRODUCTS: FOCUS ON DNMT INHIBITORS

In drug discovery it is generally accepted that a major benefit of natural products vs. purely synthetic organic molecules is, overall, the feasibility of the former to exert a biological activity and increased chemical diversity (Ho et al., 2018). The chemical space of natural products is vast and its molecular diversity has been quantified over the past few years (López-Vallejo et al., 2012; Olmedo et al., 2017; Shang et al., 2018). A major contribution to these studies has been the increasing availability of natural products collections in the public domain (Medina-Franco, 2015). Examples of major compound collections are the Traditional Chinese Medicine (Chen, 2011), natural products from Brazil – NuBBE (Pilon et al., 2017), AfroDb (Ntie-Kang et al., 2013) or collections available for screening in a medium to high-throughput screening mode. The large importance of natural products in drug discovery has boosted the development of open access applications to mine these rich repositories. Few examples are ChemGPS-NP, TCMAnalyzer, and other resources described elsewhere (Rosen et al., 2009; Chen et al., 2017; Gonzalez-Medina et al., 2017; Liu et al., 2018).

The chemical space of natural products from different sources has been compared to several other collections including the chemical space of drugs approved for clinical use and synthetic compounds (Olmedo et al., 2017; Shang et al., 2018). These studies demonstrate that the chemical space of natural products is vast, that there is a notable overlap with the chemical space of drugs, and that natural products cover novel regions of the chemical space. The overlap with the chemical space of approved drugs is not that surprising since there are a large percentage of drugs from natural origin. **Figure 3** shows a visual representation of the chemical space of 15 representative DNMT inhibitors from natural sources vs. 4103 compounds from a commercial



vendor library of natural products, 206 fungi metabolites, and 6253 marine natural products (Krishna et al., 2017). The visual representation was generated with principal component analysis of six physicochemical properties of pharmaceutical relevance, namely molecular weight (MW), topological surface area (TPSA), number of hydrogen bond donors and acceptors (HBD/HBA), number of rotatable bonds (RB), and octanol/water partition coefficient (logP). The first two principal components capture about 90% of the total variance. The visual representation of the chemical space in this figure indicates that marine natural products (data points in blue) cover a broader area of the chemical space followed by natural products in the vendor collection (orange) and by fungi metabolites (green). DNMT inhibitors from natural origin (purple) are, in general, inside the subspace of the DNMT1 inhibitors (red). This visualization of the chemical space indicates that there would be expected to identify more DNMT1 inhibitors in the marine and vendor collections, as well as in the data set of fungi metabolites.

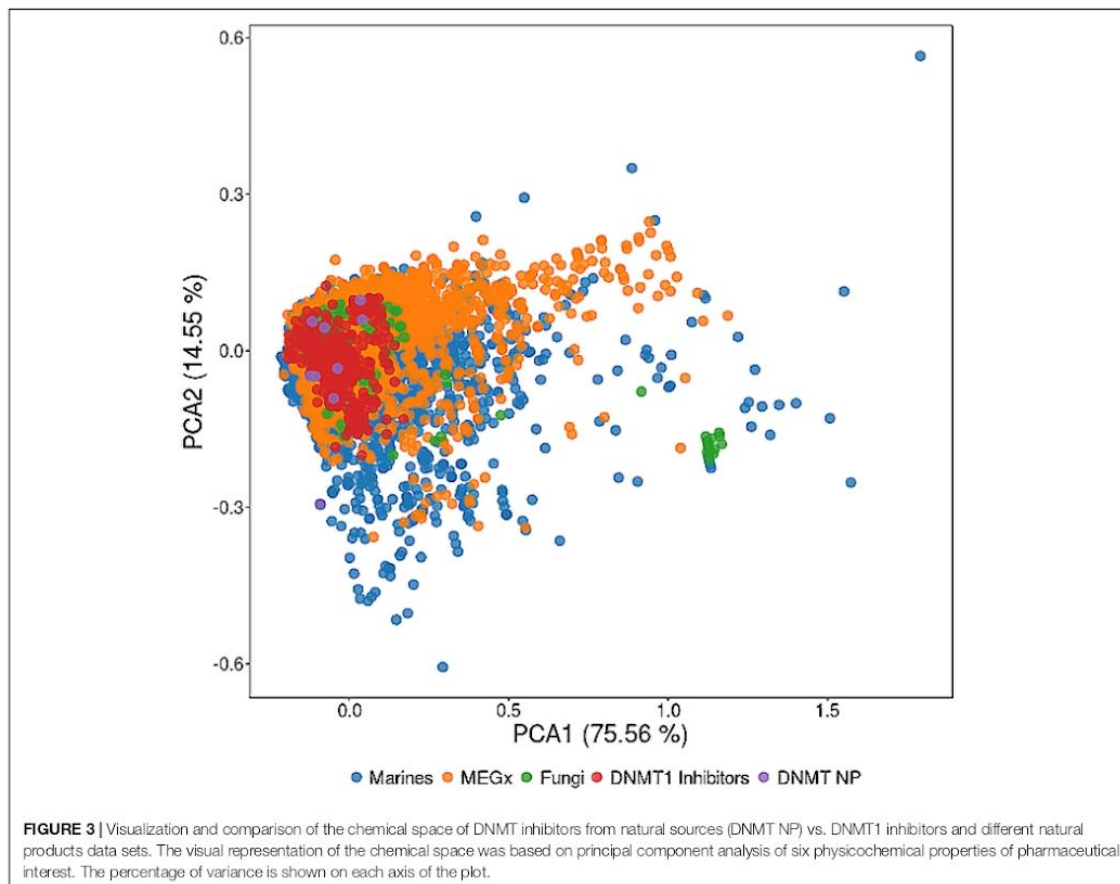
## SECTION 6: OPPORTUNITIES FOR SEARCHING FOR NATURAL PRODUCTS AS DNMT INHIBITORS

Most of the DNMT inhibitors from natural sources have been identified by serendipity. As discussed in Section 5, the chemical

space of natural products and food chemicals can be explored in a systematic manner using computational approaches. A classical and general approach is using virtual screening. The main aim of virtual screening is filtering compound data sets to select a reduced number of compounds with increased probability to show biological activity. Virtual screening has proven to be useful to identify hit compounds (Clark, 2008; Lavecchia and Di Giovanni, 2013). **Table 1** summarizes representative case studies where virtual screening has led to the identification of active compounds with novel scaffolds. In other studies, virtual screening has uncovered potential active compounds but experimental validation still needs to be conducted. Examples of these studies are further discussed in the following sections.

There are several published studies of virtual screening of natural products to identify DNMT inhibitors and/or demethylating agents. In an early work, Medina-Franco et al. (2011) reported the screening of a lead-like subset of natural products available in ZINC. Authors of that work implemented a multistep virtual screening approach selecting consensus hits identified from three different docking programs. One computational hit showed DNMT1 activity in a previous study. Other candidate compounds were identified for later experimental validation (Medina-Franco et al., 2011).

In a separate work, Maldonado-Rojas et al. (2015) developed a QSAR model based on linear discriminant analysis to screen 800 natural products. Hits selected were further

**TABLE 1** | Summary of virtual screening hits as inhibitors of DNMTs.

Study	<i>In silico</i> approach	Major outcome	Reference
Structure-based screening of a lead-like subset of NP from ZINC	Cascade docking followed by a consensus approach	One computational had reported activity. Additional natural products were identified for screening.	Medina-Franco et al., 2011
Ligand- and structure-based screening of 800 NP	QSAR model based on linear discriminant analysis and consensus docking.	Six consensus hits were identified as potential inhibitors.	Maldonado-Rojas et al., 2015
Structure-based screening of 111,121 molecules.	Docking-based screening of synthetic screening compounds.	Identification of a low micromolar hit with a novel scaffold. Further similarity searching led to the identification of two more potent hits.	Chen et al., 2014
Ligand-based screening of 500 compounds.	Pharmacophore-based virtual screening.	Identification of one inhibitor of DNMT1 with activity in the low micromolar range. The hit showed some selectivity vs. DNMT3B.	Hassanzadeh et al., 2017
Structure- and ligand-based screening of 53,000 synthetic compounds.	Pharmacophore model, a Naïve Bayesian classification model, and ensemble docking.	Two compounds showed DNMT1 inhibitory activity at single but low concentration of 1 $\mu$ M.	Krishna et al., 2017

NP: natural products.

docked with two crystallographic structures of human DNMT employing two docking programs. Six consensus hits were identified as potential inhibitors (Maldonado-Rojas et al., 2015).

Virtual screening of synthetic libraries has also been reported to identify active compounds with novel scaffolds and suitable for lead optimization. For instance, Chen et al. (2014) reported a docking-based virtual screening of the commercial screening compound library SPECS with 111,121 molecules (after filtering compounds with undesirable physicochemical properties). Results of that work led to the identification of a compound with a novel scaffold with low micromolar  $IC_{50}$  (10.3  $\mu$ M). Starting from the computational hit, similarity searching led to the identification of two more potent compounds.

Hassanzadeh et al. (2017) recently reported a pharmacophore-based virtual screening of a compound database with 500 compounds. The pharmacophore was generated using a ligand-based approach by superimposing a group of active nucleoside analogs. Selected hits, which are structurally related to the barbituric acid, were docked into the substrate binding site of DNMT1. One compound was identified with a novel chemical scaffold that inhibits DNMT1 in the low micromolar range ( $IC_{50} = 4.1 \mu$ M). The compound also showed some selectivity on DNMT1 over DNMT3 enzymes (Hassanzadeh et al., 2017).

Krishna et al. (2017) implemented a virtual screening protocol using several structure- and ligand-based approaches. Methods included a pharmacophore model, a Naïve Bayesian classification model, and ensemble docking. Three out of ten selected compounds from a commercial library of synthetic molecules (e.g., Maybridge with 53,000 small drug-like compounds), showed DNMT1 inhibitory activity at compound concentration of 20  $\mu$ M. Two of these molecules showed activity at 1  $\mu$ M (Krishna et al., 2017).

In addition to the studies discussed above and summarized in Table 1, the next subsections discuss other approaches that can be explored. Case studies for each strategy are outlined briefly.

### Section 6.1: Similarity-Based Virtual Screening of Natural Products

Similarity searching is a commonly used approach for identifying new hit compounds. Major goals are identifying starting points for later optimization or expand the SAR of analog series. Since similarity searching is fast it can be used to filter large chemical databases and it can be used in combination with other computational approaches such as molecular docking.

Similarity searching involves two major components: a molecular representation and a similarity coefficient. In practice, one of the most common molecular representations are two-dimensional (2D) fingerprints. A fingerprint is generally a string of zeros and ones that indicate the presence or absence of molecular features, respectively. In turn, one of the most common similarity coefficients is Tanimoto's (Bajusz et al., 2015). Full discussion of molecular representations and similarity coefficients are published elsewhere (Willett et al., 1998; Maggiora et al., 2014).

A novel approach to encode the chemical structures of data sets is the database fingerprint (DBFP) (Fernández-de Gortari et al., 2017). The rationale of DBFP is account for the most structural features encoded in bit positions of an entire data set. In principle, virtually any data set can be represented. For instance, it can be a small or large chemical database of screening compounds or a group of active compounds. DBFP can be used in visual representation of the chemical space (Naveja and Medina-Franco, 2018) and similarity searching (Fernández-de Gortari et al., 2017). More recently, this approach was further refined into the so-called statistical based database fingerprint (SB-DFP). This approach has the same underlying idea and application of DBFP. A key improvement is the approach to account for the most relevant structural features that are derived from a statistical comparison between the structural features of a data set of interest vs. a database of reference.

### Section 6.2: Pharmacophore-Based

Thus far, several pharmacophore modeling studies have been conducted for inhibitors of DNMT1. Different approaches and input molecules have been used to develop these models. Most of the pharmacophore models have been employed to virtually screen chemical databases and identify novel hit compounds.

Yoo and Medina-Franco (2011) reported one of the first pharmacophore models for inhibitors of DNMT1. The model was generated based on the docking poses of 14 known inhibitors available at that time. The docking was conducted with a homology model of the catalytic domain of DNMT1. Of note, at the time of that study the crystallographic structure of human DNMT1 was not available. Known DNMT inhibitors used to develop the pharmacophore model included the natural products curcumin, parthenolide, EGCG and mahanine (Yoo and Medina-Franco, 2011). A year later was reported that trimethylaurintricarboxylic acid (Figure 2) showed a good agreement with this structure-based pharmacophore model. This compound is structurally related to 5,5'-methylenedisalicylic acid that has an inhibition of DNMT1 in a low micromolar range ( $IC_{50} = 4.79 \mu$ M) (Yoo and Medina-Franco, 2012; Yoo et al., 2012).

More recently, as described in the first part of Section 6, Hassanzadeh et al. (2017) developed a pharmacophore model based on a ligand-based approach by 3D superimposition of active nucleoside analogs. That model was used to do virtual screening (*vide supra*). In the same year, with the aid of the Hypogen module of the software DS4.1, Krishna et al. (2017) developed a ligand-based pharmacophore model using the structures of 20 compounds obtained from the literature. The model was validated with the classification of an external set with known active and inactive compounds. The validated pharmacophore models were employed as part of a combined strategy to identify novel active molecules (Krishna et al., 2017).

### Section 6.3: De novo Design

*De novo* design is a technique currently explored for DNMT inhibitors on a limited basis. Here we briefly outline two promising perspectives related to natural product research.

The first one is a strategy that provides a structural diversity classification of natural products scaffolds through generative topographic map algorithm implementation often so-called chemographies. Chemographies allow the visualization of the landscape distribution of the chemical space of natural products and their synthetic mimetic compounds (Miyao et al., 2015). Since chemographies could be generated from pharmacophoric features and molecular descriptors, it would be feasible to do scaffold hopping based on the structures of natural products (Rodrigues et al., 2016). The second approach is based on scaffold simplification that could be adapted to generate fragment-like natural products focused on DNMT inhibitors. This strategy reduces the molecular framework of natural products through the implementation of a scaffold tree algorithm based on rule-based decomposition of ring systems (Bajorath, 2018).

## SECTION 7: CONCLUSION

Epigenetic targets are attractive to develop therapeutic strategies. DNA methyltransferases are the major enzyme family being one of the first epigenetic targets studied, in particular for the treatment of cancer. However, over the past few years, more therapeutic opportunities related to the modulation of DNMTs are emerging. Therefore, there is a growing interest in the scientific community to identify and develop small molecules that can be used as epi-drugs or epi-probes targeting DNMTs. Virtual screening has become more used in recent years to uncover natural products as inhibitors of DNMTs and/or demethylating agents. To this end, well established structure- and ligand-based virtual screening approaches are being used such as automated docking, QSAR and similarity searching. Also, novel chemoinformatic approaches are being developed. Of course, the computational methods should be validated with rigorous experiments *in vitro* and *in vivo* experiments to support their application.

Natural products have a well established history as inhibitors of DNMTs and demethylating molecules. However, most of the active natural products have been identified by serendipity. The knowledge of the three-dimensional structures of DNMTs in combination with increased *in silico* approaches and better computational resources are boosting the systematic search of bioactive molecules from natural origin. In addition, the increasing availability of natural product databases facilitates the discovery of epi-drugs and epi-probes targeting DNMTs.

## SECTION 8: PERSPECTIVES

Natural products inside or outside of the traditional drug-like chemical space represent a large promise to develop novel compounds with DNMT inhibitory activity or demethylating properties. This is because the traditional chemical space is highly represented by small molecules

that over the past few years have not been very successful. A notable example in this direction is the reemergence of peptide-based drug discovery. Indeed, linear, cyclic peptides and peptidomimetics are regaining interest in drug discovery (Fosgerau and Hoffmann, 2015; Henninot et al., 2018).

Other promising an emerging avenue are the modulators of protein-protein interactions (PPIs) (Díaz-Eufracio et al., 2018). DNMTs are known to be involved in several PPIs (Díaz-Eufracio et al., 2018). Modulation of such interactions can be conveniently achieved with natural products. This is because PPIs are “difficult targets” not easily addressed by small molecules from the traditional chemical space (Villoutreix et al., 2014). In other words, since PPIs have unique features these can be approached with novel chemical libraries. Natural products collections represent excellent candidates for this purpose.

We foresee an augmented hit and lead identification efforts based on natural products combining approaches such as high-throughput screening, structure-, ligand-based *in silico* screening, structure-based optimization, similarity searching, and scaffold hopping (Schneider et al., 1999). As part of the search for novel and more potent compounds is crucial to consider potential toxicity since toxicity issues play a major part in the lack of success of drug discovery projects.

## DISCLAIMER

A similar version of this manuscript was deposited in a Pre-Print server on July 6, 2018. The reference is: Saldívar-González, F. I.; Gómez-García, A.; Sánchez-Cruz, N.; Ruiz-Rios, J.; Pilón-Jiménez, B. A.; Medina-Franco, J. L. Computational Approaches to Identify Natural Products as Inhibitors of DNA Methyltransferases. *Preprints* 2018, 2018070116 (doi: 10.20944/preprints201807.0116.v1).

## AUTHOR CONTRIBUTIONS

All authors contributed to methodology and formal analysis. FS-G, JR-R, and BP-J contributed to data curation. AG-G, FS-G, DC-PdL, and JM-F contributed to writing-original draft preparation. AG-G, FS-G, NS-G, and JM-F contributed to writing-review and editing. AG-G, FS-G, and BP-J contributed to visualization. JM-F contributed to project administration.

## FUNDING

This research was funded by *Consejo Nacional de Ciencia y Tecnología* (CONACYT, Mexico) grant number 282785, the *Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica* (PAPIIT) grant IA203718, and by *Programa de Apoyo a Proyectos para*

la Innovación y Mejoramiento de la Enseñanza (PAPIME) grant PE200118, UNAM.

## ACKNOWLEDGMENTS

FS-G, AG-G, and NS-C acknowledge *Consejo Nacional de Ciencia y Tecnología* (CONACyT, Mexico) for the graduate

scholarships. DC-PdL and JR-R thanks the *Programa de Apoyo a Proyectos para la Innovación y Mejoramiento de la Enseñanza* (PAPIME) for the undergraduate scholarship. The authors also thank Chanachai Sae-Lee for providing the sequences used in **Figure 1**. They also acknowledge all current and past members of the DIFACQUIM research group for their comments and discussions that enriched this manuscript.

## REFERENCES

- Bajorath, J. (2018). Improving the utility of molecular scaffolds for medicinal and computational chemistry. *Future Med. Chem.* 10, 1645–1648. doi: 10.4155/fmc-2018-0106
- Bajusz, D., Rácz, A., and Héberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* 7:20. doi: 10.1186/s13321-015-0069-3
- Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes Dev.* 23, 781–783. doi: 10.1101/gad.1787609
- Castellano, S., Kuck, D., Sala, M., Novellino, E., Lyko, F., and Sbardella, G. (2008). Constrained analogues of procaine as novel small molecule inhibitors of DNA methyltransferase-1. *J. Med. Chem.* 51, 2321–2325. doi: 10.1021/jm7015705
- Castillo-Aguilera, O., Depreux, P., Halby, L., Arimondo, P., and Goossens, L. (2017). DNA methylation targeting: the DNMT/HMT crosstalk challenge. *Biomolecules* 7:3. doi: 10.3390/biom7010003
- Chen, C. Y.-C. (2011). TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening *in silico*. *PLoS One* 6:e15939. doi: 10.1371/journal.pone.0015939
- Chen, S. J., Wang, Y. L., Zhou, W., Li, S. S., Peng, J. L., Shi, Z., et al. (2014). Identifying novel selective non-nucleoside DNA methyltransferase 1 inhibitors through docking-based virtual screening. *J. Med. Chem.* 57, 9028–9041. doi: 10.1021/jm501134e
- Chen, Y., de Bruyn Kops, C., and Kirchmair, J. (2017). Data resources for the computer-guided discovery of bioactive natural products. *J. Chem. Inf. Model.* 57, 2099–2111. doi: 10.1021/acs.jcim.7b00341
- Choi, S. H., Heo, K., Byun, H.-M., An, W., Lu, W., and Yang, A. S. (2011). Identification of preferential target sites for human DNA methyltransferases. *Nucleic Acids Res.* 39, 104–118. doi: 10.1093/nar/gkq774
- Clark, D. E. (2008). What has virtual screening ever done for drug discovery? *Expert Opin. Drug Discov.* 3, 841–851. doi: 10.1517/17460440802281978
- Davide, G., Sandra, A., Emily, B., Mattia, C., Marta, G., Annalisa, C., et al. (2016). Design and synthesis of N-benzoyl amino acid derivatives as DNA methylation inhibitors. *Chem. Biol. Drug Des.* 88, 664–676. doi: 10.1111/cbdd.12794
- Diaz-Eufracio, B. I., Naveja, J. J., and Medina-Franco, J. L. (2018). Chapter three - protein-protein interaction modulators for epigenetic therapies. *Adv. Protein Chem. Struct. Biol.* 110, 65–84. doi: 10.1016/bs.apcsb.2017.06.002
- Du, Q., Wang, Z., and Schramm, V. L. (2016). Human DNMT1 transition state structure. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2916–2921. doi: 10.1073/pnas.1522491113
- Dueñas-González, A., Jesús Naveja, J., and Medina-Franco, J. L. (2016). Chapter 1 - Introduction of Epigenetic Targets in Drug Discovery and Current Status of Epi-Drugs and Epi-Probes, in *Epi-Informatics*. Boston, MA: Academic Press, 1–20. doi: 10.1016/B978-0-12-802808-7.00001-0
- Fernández-de Gortari, E., García-Jacas, C. R., Martínez-Mayorga, K., and Medina-Franco, J. L. (2017). Database fingerprint (DFP): an approach to represent molecular databases. *J. Cheminform.* 9:9. doi: 10.1186/s13321-017-0195-1
- Fernandez-de Gortari, E., and Medina-Franco, J. L. (2015). Epigenetic relevant chemical space: a chemoinformatic characterization of inhibitors of DNA methyltransferases. *RSC Adv.* 5, 87465–87476. doi: 10.1039/C5RA19611F
- Fosgerau, K., and Hoffmann, T. (2015). Peptide therapeutics: current status and future directions. *Drug Discov. Today* 20, 122–128. doi: 10.1016/j.drudis.2014.10.003
- Ganesan, A. (2018). Epigenetic drug discovery: a success story for cofactor interference. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170069. doi: 10.1098/rstb.2017.0069
- Gonzalez-Medina, M., Naveja, J. J., Sanchez-Cruz, N., and Medina-Franco, J. L. (2017). Open chemoinformatic resources to explore the structure, properties and chemical space of molecules. *RSC Adv.* 7, 54153–54163. doi: 10.1039/C7RA11831G
- Gordon, C. A., Hartono, S. R., and Chédin, F. (2013). Inactive DNMT3B splice variants modulate de novo DNA methylation. *PLoS One* 8:e69486. doi: 10.1371/journal.pone.0069486
- Henninot, A., Collins, J. C., and Nuss, J. M. (2018). The current state of peptide drug discovery: back to the future? *J. Med. Chem.* 61, 1382–1414. doi: 10.1021/acs.jmedchem.7b00318
- Ho, T. T., Tran, Q. T. N., and Chai, C. L. L. (2018). The polypharmacology of natural products. *Future Med. Chem.* 10, 1361–1368. doi: 10.4155/fmc-2017-0294
- Hwang, J.-Y., Aromolaran, K. A., and Zukin, R. S. (2017). The emerging field of epigenetics in neurodegeneration and neuroprotection. *Nat. Rev. Neurosci.* 18, 347–361. doi: 10.1038/nrn.2017.46
- Jeltsch, A. (2002). Beyond Watson and crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem* 3, 274–293. doi: 10.1002/1439-7633(20020402)3:4<274::AID-CBIC274>3.0.CO;2-S
- Jurkowska, R. Z., Jurkowski, T. P., and Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. *ChemBioChem* 12, 206–222. doi: 10.1002/cbic.201000195
- Kabro, A., Lachance, H., Marcoux-Archambault, I., Perrier, V., Dore, V., Gros, C., et al. (2013). Preparation of phenylethylbenzamide derivatives as modulators of DNMT3 activity. *MedChemComm* 4, 1562–1570. doi: 10.1039/c3md00214d
- Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. D. (1994). HHAL methyltransferase flips its target base out of the DNA helix. *Cell* 76, 357–369. doi: 10.1016/0092-8674(94)90342-5
- Krishna, S., Shukla, S., Lakra, A. D., Meeran, S. M., and Siddiqi, M. I. (2017). Identification of potent inhibitors of DNA methyltransferase 1 (DNMT1) through a pharmacophore-based virtual screening approach. *J. Mol. Graph. Model.* 75, 174–188. doi: 10.1016/j.jmgm.2017.05.014
- Lan, J., Hua, S., He, X. N., and Zhang, Y. (2010). DNA methyltransferases and methyl-binding proteins of mammals. *Acta Biochim. Biophys. Sin.* 42, 243–252. doi: 10.1093/abbs/gmq015
- Lavecchia, A., and Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: a critical review. *Curr. Med. Chem.* 20, 2839–2860. doi: 10.2174/09298673113209990001
- Liu, Z., Du, J., Yan, X., Zhong, J., Cui, L., Lin, J., et al. (2018). TCMAnalyzer: a chemo- and bioinformatics web service for analyzing traditional Chinese medicine. *J. Chem. Inf. Model.* 58, 550–555. doi: 10.1021/acs.jcim.7b00549
- López-Vallejo, F., Giulianotti, M. A., Houghten, R. A., and Medina-Franco, J. L. (2012). Expanding the medicinally relevant chemical space with compound libraries. *Drug Discov. Today* 17, 718–726. doi: 10.1016/j.drudis.2012.04.001
- Lu, W., Zhang, R., Jiang, H., Zhang, H., and Luo, C. (2018). Computer-aided drug design in epigenetics. *Front. Chem.* 6:57. doi: 10.3389/fchem.2018.00057
- Lyko, F. (2017). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* 19, 81–92. doi: 10.1038/nrg.2017.80
- Maggiora, G., Vogt, M., Stumpfe, D., and Bajorath, J. (2014). Molecular similarity in medicinal chemistry. *J. Med. Chem.* 57, 3186–3204. doi: 10.1021/jm401411z
- Maldonado-Rojas, W., Olivero-Verbel, J., and Marrero-Ponce, Y. (2015). Computational fishing of new DNA methyltransferase inhibitors from natural products. *J. Mol. Graph. Model.* 60, 43–54. doi: 10.1016/j.jmgm.2015.04.010



- Hassanzadeh, M., Kasymov, R., Mahernia, S., Adib, M., Emperle Michael Dukatz, M., Bashtrykov, P., et al. (2017). Discovery of novel and selective DNA methyltransferase 1 inhibitors by pharmacophore and docking-based virtual screening. *ChemistrySelect* 2, 8383–8392. doi: 10.1002/slct.201701734
- Medina-Franco, J. L. (2015). "Discovery and development of lead compounds from natural sources using computational approaches," in *Evidence-Based Validation of Herbal Medicine*, ed. P. Mukherjee (New York, NY: Elsevier), 455–475. doi: 10.1016/B978-0-12-800874-4.00021-0
- Medina-Franco, J. L., López-Vallejo, F., Kuck, D., and Lyko, F. (2011). Natural products as DNA methyltransferase inhibitors: a computer-aided discovery approach. *Mol. Divers.* 15, 293–304. doi: 10.1007/s11030-010-9262-5
- Medina-Franco, J. L., Méndez-Lucio, O., Yoo, J., and Dueñas, A. (2015). Discovery and development of DNA methyltransferase inhibitors using in silico approaches. *Drug Discov. Today* 20, 569–577. doi: 10.1016/j.drudis.2014.12.007
- Miyao, T., Reker, D., Schneider, P., Funatsu, K., and Schneider, G. (2015). Chemography of natural product space. *Planta Med.* 81, 429–435. doi: 10.1055/s-0034-1396322
- Naveja, J. J., and Medina-Franco, J. L. (2015). Activity landscape sweeping: insights into the mechanism of inhibition and optimization of DNMT1 inhibitors. *RSC Adv.* 5, 63882–63895. doi: 10.1039/C5RA12339A
- Naveja, J. J., and Medina-Franco, J. L. (2018). Insights from pharmacological similarity of epigenetic targets in epipolypharmacology. *Drug Discov. Today* 23, 141–150. doi: 10.1016/j.drudis.2017.10.006
- Naveja, J. J., Rico-Hidalgo, M. P., and Medina-Franco, J. L. (2018). Analysis of a large food chemical database: chemical space, diversity, and complexity. *FI000Res.* 7:993. doi: 10.12688/fi000research.15440.2
- Ntie-Kang, F., Zofou, D., Babiaka, S. B., Meudom, R., Scharfe, M., Lifongo, L. L., et al. (2013). AfroDb: a select highly potent and diverse natural product library from African medicinal plants. *PLoS One* 8:e78085. doi: 10.1371/journal.pone.0078085
- Olmedo, D. A., González-Medina, M., Gupta, M. P., and Medina-Franco, J. L. (2017). Cheminformatic characterization of natural products from panama. *Mol. Divers.* 21, 779–789. doi: 10.1007/s11030-017-9781-4
- Ostler, K. R., Davis, E. M., Payne, S. L., Gosalia, B. B., Expósito-Céspedes, J., Beau, M. M. L., et al. (2007). Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. *Oncogene* 26, 5553–5563. doi: 10.1038/sj.onc.1210351
- Pilon, A. C., Valli, M., Dametto, A. C., Pinto, M. E. F., Freire, R. T., Castro-Gamboa, I., et al. (2017). NuBBEDB: an updated database to uncover chemical and biological information from Brazilian biodiversity. *Sci. Rep.* 7:7215. doi: 10.1038/s41598-017-07451-x
- Rodrigues, T., Reker, D., Schneider, P., and Schneider, G. (2016). Counting on natural products for drug design. *Nat. Chem.* 8, 531–541. doi: 10.1038/nchem.2479
- Rosen, J., Lovgren, A., Kogej, T., Muresan, S., Gottfries, J., and Backlund, A. (2009). ChemGPS-NPWeb: chemical space navigation online. *J. Comput. Aided Mol. Des.* 23, 253–259. doi: 10.1007/s10822-008-9255-y
- Sacconny, L., Angleviel, M., Randazzo, G. M., Queiroz, M. M., Queiroz, E. F., Wolfender, J. L., et al. (2014). Computational studies on sirtuins from *Trypanosoma cruzi*: structures, conformations and interactions with phytochemicals. *PLoS Negl. Trop. Dis.* 8:e2689. doi: 10.1371/journal.pntd.0002689
- Schneider, G., Neidhart, W., Giller, T., and Schmid, G. (1999). Scaffold-hopping by topological pharmacophore search: a contribution to virtual screening. *Angew. Chem. Int. Ed.* 38, 2894–2896. doi: 10.1002/(SICI)1521-3773(19991004)38:19<2894::AID-ANIE2894>3.0.CO;2-F
- Shang, J., Hu, B., Wang, J., Zhu, F., Kang, Y., Li, D., et al. (2018). Cheminformatic insight into the differences between terrestrial and marine originated natural products. *J. Chem. Inf. Model.* 58, 1182–1193. doi: 10.1021/acs.jcim.8b00125
- Tough, D. F., Tak, P. P., Tarakhovskiy, A., and Prinjha, R. K. (2016). Epigenetic drug discovery: breaking through the immune barrier. *Nat. Rev. Drug Discov.* 15, 835–853. doi: 10.1038/nrd.2016.185
- Vilkaitis, G., Merkiene, E., Serva, S., Weinhold, E., and Klimauskas, S. (2001). The mechanism of DNA cytosine-5 methylation - kinetic and mutational dissection of Hhai methyltransferase. *J. Biol. Chem.* 276, 20924–20934. doi: 10.1074/jbc.M101429200
- Villoutreix, B. O., Kuenemann, M. A., Poyet, J. L., et al. (2014). Drug-like protein-protein interaction modulators: challenges and opportunities for drug discovery and chemical biology. *Mol. Inf.* 33, 414–437. doi: 10.1002/minf.20140040
- Waddington, C. H. (2012). The epigenotype. *Int. J. Epidemiol.* 41, 10–13. doi: 10.1093/ije/dyr184
- Wang, L., Wang, J., Sun, S., Rodriguez, M., Yue, P., Jang, S. J., et al. (2006). A novel DNMT3B subfamily,  $\Delta$ DNMT3B, is the predominant form of DNMT3B in Non-small cell lung cancer. *Int. J. Oncol.* 29, 201–207. doi: 10.3892/ijo.29.1.201
- Willett, P., Barnard, J., and Downs, G. (1998). Chemical similarity searching. *J. Chem. Inf. Comput. Sci.* 38, 983–996. doi: 10.1021/ci9800211
- Yoo, J., Kim, J. H., Robertson, K. D., and Medina-Franco, J. L. (2012). Molecular modeling of inhibitors of human DNA methyltransferase with a crystal structure: discovery of a novel DNMT1 inhibitor. *Adv. Protein Chem. Struct. Biol.* 87, 219–247. doi: 10.1016/B978-0-12-398312-1.00008-1
- Yoo, J., and Medina-Franco, J. L. (2011). Homology modeling, docking, and structure-based pharmacophore of inhibitors of DNA methyltransferase. *J. Comp. Aided Mol. Des.* 25, 555–567. doi: 10.1007/s10822-011-9441-1
- Yoo, J., and Medina-Franco, J. L. (2012). Trimethylaurintricarboxylic acid inhibits human DNA methyltransferase 1: insights from enzymatic and molecular modeling studies. *J. Mol. Model.* 18, 1583–1589. doi: 10.1007/s00894-011-1191-4
- Zhang, Z.-M., Liu, S., Lin, K., Luo, Y., Perry, J. J., Wang, Y., et al. (2015). Crystal structure of human DNA methyltransferase 1. *J. Mol. Biol.* 427, 2520–2531. doi: 10.1016/j.jmb.2015.06.001
- Zwergel, C., Valente, S., and Mai, A. (2016). DNA methyltransferases inhibitors from natural sources. *Curr. Top. Med. Chem.* 16, 680–696. doi: 10.2174/1568026615666150825141505

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Saldívar-González, Gómez-García, Chávez-Ponce de León, Sánchez-Cruz, Ruiz-Rios, Pilon-Jiménez and Medina-Franco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.