

Nutracéuticos derivados del aceite de oliva como posibles herramientas terapéuticas en angiogénesis

Estefanía Huergo Iglesias

Máster en Bioinformática y Bioestadística
Área de Bioinformática farmacéutica

Melchor Sánchez Martínez

Marc Maceira Duch

Enero de 2021



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada 3.0 España de Creative Commons

FICHA DEL TRABAJO FINAL

Título del trabajo:	Nutracéuticos derivados del aceite de oliva como posibles herramientas terapéuticas en angiogénesis
Nombre del autor:	Estefanía Huergo Iglesias
Nombre del consultor/a:	Melchor Sánchez Martínez
Nombre del PRA:	Marc Maceira Duch
Fecha de entrega (mm/aaaa):	01/2021
Titulación:	Máster en Bioinformática y Bioestadística
Área del Trabajo Final:	Bioinformática Farmacéutica
Idioma del trabajo:	Castellano
Palabras clave	Aceite de oliva virgen extra, fenoles bioactivos, angiogénesis
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i></p>	
<p>El cáncer es una de las principales causas de muerte a nivel mundial, lo que implica poner el foco en su prevención, detección temprana, diagnóstico y tratamiento. La inhibición de la angiogénesis se ha convertido en una estrategia clínica anticancerígena, pero los fármacos empleados no siempre son eficaces, muchos desarrollan resistencia y algunos son tóxicos. Por lo tanto, sería interesante identificar compuestos alternativos y, así, surge la “angioprevención” como el uso de alimentos ricos en moléculas inhibitoras de la angiogénesis o de estas moléculas en forma de fármacos para prevenir, retrasar o combatir el crecimiento de tumores. En este contexto, cada vez hay más evidencias de que los fenoles del aceite de oliva virgen extra (AOVE) tienen potentes actividades biológicas que podrían explicar algunos efectos beneficiosos para la salud de la dieta mediterránea, incluida la baja prevalencia de ciertos tipos de cáncer. En el presente trabajo se analizaron los fenoles del AOVE como posibles inhibidores de las enzimas ADN metiltransferasas (DNMTs) y catecol O-metiltransferasa (COMT), implicadas en procesos de angiogénesis, realizando estudios computacionales <i>in silico</i>. Varios fenoles podrían unirse eficazmente a los sitios de unión de estas proteínas, por lo que mostraron una potencial actividad antiangiogénica y podrían emplearse como nutraceuticos en terapia anticancerígena. Aunque la síntesis de productos naturales puede resultar complicada y costosa, se comprobó que varios de los fenoles seleccionados son comerciales. Además, se estudiaron compuestos comerciales similares a estos y se comprobó que algunos son potenciales inhibidores de estas proteínas, de forma que también podrían usarse como fármacos contra el cáncer.</p>	

Abstract (in English, 250 words or less):

Cancer is one of the leading causes of death worldwide, which implies focusing on its prevention, early detection, diagnosis and treatment. The inhibition of angiogenesis has become a clinical anticancer strategy, but the drugs used are not always effective, many develop resistance, and some are toxic. Therefore, it would be interesting to identify alternative compounds. Thus, “angioprevention” arises as the use of foods rich in angiogenesis inhibitor molecules or of these molecules in the form of drugs to prevent, delay or combat the growth of tumours. In this context, there is increasing evidence that extra virgin olive oil (EVOO) phenols have potent biological activities that could explain some beneficial health effects of the Mediterranean diet, including the low prevalence of certain types of cancer. In the present work, EVOO phenols were analysed as possible inhibitors of the enzymes DNA methyltransferases (DNMTs) and catechol O-methyltransferase (COMT), involved in angiogenesis processes, performing *in silico* computational studies. Various phenols could bind effectively to the binding sites of these proteins, thus showing potential antiangiogenic activity and could be used as nutraceuticals in anticancer therapy. Although the synthesis of natural products can be complicated and expensive, several of the selected phenols have been found to be commercial. In addition, commercial compounds similar to these phenols were studied and it was found that some are potential inhibitors of these proteins, so that they could also be used as anticancer drugs.

Índice

1. Introducción	1
1.1. Contexto y justificación del Trabajo	1
1.1.1. Descripción general	1
1.1.2. Justificación del TFM	1
1.2. Objetivos del Trabajo.....	4
1.2.1. Objetivos generales	4
1.2.2. Objetivos específicos.....	4
1.3. Enfoque y método seguido	5
1.4. Planificación del Trabajo	6
1.4.1. Tareas	7
1.4.2. Calendario	8
1.4.3. Hitos	9
1.5. Breve resumen de productos obtenidos	9
1.6. Breve descripción de los otros capítulos de la memoria	10
2. Antecedentes	11
2.1. Metilación del ADN.....	11
2.1.1. Enzimas DNMTs.....	12
2.1.1.1. <i>Reacción de metilación del ADN</i>	13
2.1.1.2. <i>Estructura y dominios</i>	13
2.1.1.3. <i>DNMT1</i>	14
2.1.1.4. <i>DNMT3</i>	15
2.1.2. DNMTs y la angiogénesis.....	16
2.1.3. Inhibición de las proteínas DNMTs	17
2.2. Enzima COMT	20
2.2.1. Formas de COMT	20
2.2.2. Estructura de COMT	21
2.2.3. Reacción de O-metilación.....	22
2.2.4. Polimorfismo de COMT	23
2.2.5. Inhibidores de la proteína COMT	24
2.3. Sustratos/inhibidores de las proteínas DNMTs y COMT	26
2.4. El AOVE y los fenoles	27
3. Materiales y métodos	29
3.1. Búsqueda bibliográfica	29
3.2. Compuestos fenólicos del AOVE.....	29
3.3. Softwares para la visualización y tratamiento de estructuras.....	32
3.3.1. PyMOL	32
3.3.2. UCSF Chimera	32
3.4. Estructuras 3D de las proteínas diana	32
3.5. Cálculos del cribado virtual.....	33

3.5.1. Preparación de las estructuras de las proteínas	33
3.5.2. Preparación de los ligandos.....	34
3.5.3. Definición del sitio de unión de la proteína.....	34
3.6. Interacciones ligando–proteína.....	35
3.7. Análisis de las propiedades ADMET.....	36
3.8. Búsqueda de moléculas comerciales similares a los fenoles.....	37
4. Discusión de resultados	39
4.1. Estructuras 3D de DNMT1 y COMT	39
4.2. Interacciones proteína–ligando en las estructuras de DNMT1 y COMT	40
4.3. Evaluación de inhibidores potenciales de DNMT1 y COMT en los fenoles del AOVE	41
4.3.1. Cribado virtual de las proteínas con sus inhibidores	41
4.3.2. Cribado virtual de los fenoles del AOVE con DNMT1 y COMT	43
4.3.3. Análisis de las propiedades ADMET de los <i>hits</i> de DNMT1 y COMT	46
4.3.4. Selección de los fenoles con más potencial para actuar como inhibidores	49
4.4. Estructuras 3D de DNMT3A y DNMT3B.....	50
4.5. Interacciones proteína–ligando en las estructuras de DNMT3.....	52
4.6. Conservación modos de unión de las proteínas DNMTs	53
4.7. Evaluación de los <i>hits</i> como inhibidores potenciales de DNMT3	55
4.7.1. Cribado virtual de las proteínas con sus inhibidores	55
4.7.2. Cribado virtual de los <i>hits</i> de los fenoles del AOVE con DNMT3.....	56
4.8. Análisis de moléculas comerciales con potencial antiangiogénico.....	57
4.8.1. Cribado virtual compuestos similares a los <i>hits</i> no comerciales.....	57
4.8.2. Cribado virtual compuestos similares a los <i>hits</i> comerciales.....	59
4.8.3. Análisis propiedades ADMET <i>hits</i> comerciales	62
4.8.4. Selección compuestos comerciales con más potencial para actuar como inhibidores.....	62
4.8.5. Evaluación de los <i>hits</i> comerciales como inhibidores de DNMT3.....	64
5. Conclusiones.....	65
6. Glosario.....	68
7. Bibliografía.....	69
8. Anexos	74
8.1. Anexo I.....	74
8.2. Anexo II.....	78
8.3. Anexo III.....	79

Lista de figuras

Figura 1. Estructura de la molécula de oleaceína	3
Figura 2. Planificación temporal realizada con <i>GanttProject</i>	8
Figura 3. Diagrama de Gantt realizado con <i>GanttProject</i>	9
Figura 5. Esquema de reacción de metilación de una citosina por las enzimas DNMTs	13
Figura 4. Arquitectura de los dominios de las proteínas DNMTs	14
Figura 6. Los diferentes roles de DNMTs en varios cánceres y la paradoja de DNMT	17
Figura 7. Mecanismo de acción de los inhibidores de DNMT azacitidina y decitabina	18
Figura 8. Inhibidores de DNMTs análogos no nucleósidos	19
Figura 9. Esquema de la estructura del gen COMT humano	21
Figura 10. Mecanismo propuesto para la reacción de O-metilación mediada por COMT	23
Figura 11. Ejemplos de los diferentes tipos de inhibidores de la enzima COMT	25
Figura 12. Esquema de un inhibidor bisustrato de la enzima COMT	26
Figura 13. Esquema de la metilación mediada por DNMTs y su modulación por la reacción de O-metilación de catecoles por COMT	27
Figura 14. Composición de la base de datos OliveNet™: 676 compuestos agrupados en 13 clases y 47 subclases	30
Figura 15. La base de datos OliveNet™ recoge 222 compuestos fenólicos divididos en 13 subclases	31
Figura 16. Script de Chimera con código Python para preparar los ficheros de los fenoles para el <i>docking</i>	34
Figura 17. Script de Python para calcular la caja de 6 Å para el <i>docking</i>	35
Figura 18. Script de Chimera Python para preparar los archivos ligando-proteína	36
Figura 19. Flujo de trabajo de la plataforma pkCSM	37
Figura 20. Interconversión del formato SMILES a las estructuras 2D y 3D del tirosol	38
Figura 21. Script de Chimera Python para convertir los SMILES en la estructura 3D	38
Figura 22. Estructura PDB 4WXX de la proteína DNMT1 con los cationes Zn ²⁺ y la molécula SAH	39
Figura 23. Estructura PDB 3BWM de la proteína COMT 108V con DNC, SAM, K ⁺ , Mg ²⁺ y la cadena lateral del residuo V108	40
Figura 24. Interacciones en la estructura 4WXX generadas por PLIP entre la proteína DNMT1 y la molécula de SAH	40
Figura 25. Interacciones en el cristal 3BWM por PLIP: en la imagen A entre la proteína COMT 108V y SAM, y en la B entre el ligando DNC, Mg ²⁺ y la proteína	41
Figura 26. Superposición con PyMOL de la molécula SAH obtenida en el <i>docking</i> con la molécula SAH del cristal 4WXX de la proteína DNMT1	42
Figura 27. Superposición con PyMOL de la molécula de DNC obtenida en el <i>docking</i> con el DNC del cristal de la COMT y en amarillo se representa el Mg ²⁺	43
Figura 28. Superposición de la mejor pose de la quercetina-7-O-glucósido obtenida en el <i>docking</i> con el ligando SAH en el cristal de DNMT1	44
Figura 29. Superposición de la mejor pose del isoacteósido obtenida en el <i>docking</i> con el ligando DNC en la estructura 3BWM de COMT 108V	46

Figura 30. Compuestos fenólicos del AOVE con potencial como inhibidores	50
Figura 31. Estructura PDB 6W8B. Se representa en verde el dímero de DNMT3A (<i>wild-type</i>), en naranja las subunidades DNMT3L y en azul la cadena de ADN.....	51
Figura 32. Estructura PDB 6KDA. Se representa en verde el dímero de la proteína DNMT3B, en naranja las subunidades DNMT3L y en azul la cadena de ADN.....	52
Figura 33. Interacciones entre la proteína y SAH por PLIP: en la imagen A en el cristal 6W8B de DNMT3A y en la B en el cristal 6KDA de DNMT3B	52
Figura 34. Proteínas DNMT1, DNMT3A y DNMT3B superpuestas con UCSF Chimera con la ampliación del sitio de unión del ligando SAH	53
Figura 35. Alineamiento de las proteínas DNMT1, DNMT3A y DNMT3B en una región de su dominio catalítico generado con UCSF Chimera	54
Figura 36. Superposición con PyMOL de la molécula de SAH obtenida en el <i>docking</i> con SAH del cristal en la estructura 6W8B de DNMT3A <i>wild-type</i>	55
Figura 37. Estructuras de los estereoisómeros de la catequina	62
Figura 38. Compuestos comerciales seleccionados como <i>hits</i> para DNMT1	63
Figura 39. Compuestos comerciales seleccionados como <i>hits</i> para COMT	64

Lista de tablas

Tabla 1. Compuestos fenólicos identificados como <i>hits</i> para la proteína DNMT1	43
Tabla 2. Compuestos fenólicos identificados como <i>hits</i> para la proteína COMT	45
Tabla 3. Mejores fenoles para las proteínas DNMT3A y DNMT3B.....	56
Tabla 4. Compuestos comerciales similares a quercetina-7-O-glucósido y <i>hellicoside</i> identificados como <i>hits</i> para la proteína DNMT1	58
Tabla 5. Compuestos comerciales similares a quercetina-7-O-glucósido y <i>hellicoside</i> como <i>hits</i> para la proteína COMT	59
Tabla 6. Compuestos comerciales similares a los fenoles del AOVE comerciales identificados como <i>hits</i> para la proteína DNMT1	60
Tabla 7. Compuestos comerciales similares a los fenoles del AOVE comerciales identificados como <i>hits</i> para la proteína COMT.....	61

1. Introducción

1.1. Contexto y justificación del Trabajo

1.1.1. Descripción general

La temática general en la que se enmarca este Trabajo Fin de Máster (TFM) es la búsqueda computacional de nuevas moléculas con actividad anticancerígena. En concreto, se busca analizar si los compuestos fenólicos presentes en el aceite de oliva virgen extra (AOVE) pueden actuar como inhibidores de las enzimas ADN metiltransferasas (DNMTs) y catecol O-metiltransferasa (COMT), implicadas en procesos de angiogénesis en humanos. Así, los fenoles que muestren una potencial actividad antiangiogénica, podrían ser empleados como nutracéuticos en terapia anticancerígena para la prevención y el tratamiento del cáncer.

En este estudio se analizan las distintas familias de los fenoles del AOVE, de manera que se pone de manifiesto diferencias entre algunas, al mostrar un mayor potencial como agentes antiangiogénicos que otras. En la literatura ya se han descrito precedentes de ácidos fenólicos, flavonoides y cumarinas con gran potencial para ser usados como agentes anticancerígenos.¹

Dado que en algunas ocasiones la síntesis de productos naturales puede resultar complicada y costosa, una alternativa es emplear compuestos disponibles comercialmente. Por tanto, en este trabajo se extiende el estudio a examinar, en primer lugar, si los compuestos fenólicos del AOVE seleccionados como potenciales inhibidores de las proteínas DNMTs y COMT son comerciales. En segundo lugar, determinar compuestos similares a estos fenoles que estén disponibles en el mercado para comprobar si también son capaces de inhibir a estas proteínas, de forma que puedan ser igualmente empleados como fármacos contra el cáncer.

1.1.2. Justificación del TFM

Según la plataforma web “Global Cancer Observatory” (GCO) desarrollada por el Centro Internacional de Investigaciones sobre el Cáncer que forma parte de la Organización Mundial de la Salud (OMS), el cáncer es la segunda causa principal de muerte a nivel mundial, siendo responsable en 2018 de aproximadamente 9.6 millones de muertes, lo que equivale a una de cada seis. También se estima que en ese mismo año se diagnosticaron casi 18.1 millones de nuevos casos.² Además, la GCO predice que en 2040 la incidencia aumentará hasta los 29.5 millones de nuevos casos y la

¹ Abotaleb, M.; Samuel, S.M.; Varghese, E.; Varghese, S.; Kubatka, P.; Liskova, A.; Büsselberg, D. Flavonoids in Cancer and Apoptosis. *Cancers* **2019**, *11*, 28–67.

² Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* **2019**, *144*, 1941–1953.

mortalidad será de alrededor de 16.4 millones.³ La OMS advierte que un alto índice de masa corporal, el bajo consumo de frutas y verduras, la falta de actividad física, y el consumo de tabaco y alcohol son los principales factores de riesgo, causantes de alrededor de un tercio de las muertes por cáncer.

Estos datos ponen de manifiesto la gran importancia de la prevención, la detección temprana, el diagnóstico y el tratamiento de esta enfermedad. Aunque hay distintos enfoques en el tratamiento del cáncer, a menudo tienen efectos secundarios adversos y en ocasiones son ineficaces por el aumento a la resistencia a los fármacos o la radioterapia.

Recientemente, la inhibición de la angiogénesis (crecimiento de nuevos vasos sanguíneos a partir de una vasculatura existente) se ha convertido en una estrategia clínica anticancerígena. Como tal, la angiogénesis tumoral es importante, ya que al igual que los tejidos sanos, los tumores necesitan nutrientes y oxígeno, así como eliminar los desechos metabólicos y el dióxido de carbono, pero también es clave para permitir la desregulación metabólica y la diseminación del tumor y la metástasis.⁴ Desafortunadamente, los fármacos antiangiogénicos que se emplean en la actualidad solo son eficaces en un subconjunto de pacientes, muchos acaban desarrollando resistencia y algunos son tóxicos, por lo que sería de gran importancia identificar compuestos alternativos que pudieran superar estos inconvenientes y limitaciones.

En este contexto, durante las últimas décadas, los fitoquímicos han generado una atención significativa por su potencial uso en prevención y tratamiento del cáncer, en particular, por su actividad antiangiogénica.⁵ Así, surge el concepto de “angioprevención”, que consiste en el uso de alimentos ricos en moléculas inhibitoras de la angiogénesis o de estas moléculas en forma de fármacos para prevenir, retrasar o combatir el crecimiento de posibles tumores.⁶ Es importante poner de manifiesto que estos compuestos naturales derivados de plantas destacan, en general, por su alta biodisponibilidad, seguridad, efectos secundarios mínimos y, lo que es más importante, rentabilidad.

En cuanto a los fitoquímicos, cabe destacar que muchos componentes naturales contenidos en la alimentación de la dieta mediterránea presentan un gran potencial antiangiogénico; en concreto, el AOVE, la principal fuente de grasas en esta dieta, se relaciona con la baja prevalencia de ciertos tipos de cáncer.⁷ Aunque, los efectos

³ *Global Cancer Observatory* [en línea] [fecha de consulta: 23 de Septiembre de 2020]. Disponible en: <https://gco.iarc.fr/>.

⁴ (a) Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell*. **2011**, *144*, 646–674. (b) Hanahan, D.; Weinberg, R.A. The hallmarks of cancer. *Cell*. **2000**, *100*, 57–70.

⁵ Wang, Z.; Dabrosin, C.; Yin, X.; Fuster, M.M.; Arreola, A.; Rathmell, W.K.; Generali, D.; Nagaraju, G.P.; El-Rayes, B.; Ribatti, D.; Chen, Y.C.; Honoki, K.; Fujii, H.; Georgakilas, A.G.; Nowsheen, S.; Amedei, A.; Niccolai, E.; Amin, A.; Ashraf, S.S.; Helferich, B.; Yang, X.; Guha, G.; Bhakta, D.; Ciriolo, M.R.; Aquilano, K.; Chen, S.; Halicka, D.; Mohammed, S.I.; Azmi, A.S.; Bilsland, A.; Keith, W.N.; Jensen, L.D. Broad targeting of angiogenesis for cancer prevention and therapy. *Seminars in Cancer Biology* **2015**, *35*, S224–S243.

⁶ Albin, A.; Tosetti, F.; Li, V.W.; Noonan, D.M.; Li, W.W. Cancer prevention by targeting angiogenesis. *Nature Reviews Clinical Oncology* **2012**, *9*, 498–509.

⁷ (a) Battino, M.; Forbes-Hernández, T.Y.; Gasparri, M.; Afrin, S.; Cianciosi, D.; Zhang, J.; Manna, P.P.; Reboledo-Rodríguez, P.; Lopez, A.V.; Quiles, J.L.; Mezzetti, B.; Bompadre, S.; Xiao, J.; Giampieri, F. Relevance of functional foods in the Mediterranean diet: the role of olive

beneficiosos para la salud de la ingesta de aceite de oliva se atribuían a la alta concentración de ácidos grasos monoinsaturados, en la actualidad, diversos estudios indican que sus compuestos fenólicos tienen efectos positivos sobre diferentes biomarcadores relacionados con la salud, así como para ser usados como nutraceuticos por sus efectos antiangiogénicos.⁸ De las investigaciones que se han llevado a cabo al respecto en los últimos años, un trabajo interesante es el realizado por Javier A. Menéndez y colaboradores, en el que se determinó que el fenol oleaceína (descarboximetil oleuropeína aglicona, representado en la Figura 1) puede ser útil como inhibidor de distintas proteínas involucradas en la angiogénesis.⁹

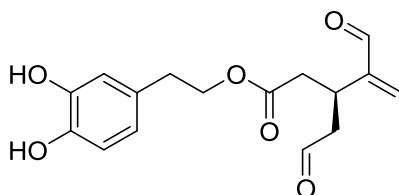


Figura 1. Estructura de la molécula de oleaceína

Entre las proteínas cuya actividad inhibe la molécula de oleaceína se encuentran las enzimas ADN metiltransferasas (DNMTs) y catecol O-metiltransferasa (COMT).^{9b,d} Por un lado, la función de las proteínas DNMTs es catalizar la metilación del ADN con S-adenosil-L-metionina (SAM) como donante de metilo y hay identificadas tres DNMTs activas en mamíferos, llamadas DNMT1, DNMT3A y DNMT3B. La enzima DNMT1 es una metiltransferasa de mantenimiento que mantiene los patrones de metilación del ADN en la replicación del ADN durante la división celular. En cambio, tanto DNMT3A como DNMT3B están involucradas en los procesos de metiltransferasa *de novo*,

oil, berries and honey in the prevention of cancer and cardiovascular diseases. *Critical Reviews in Food Science and Nutrition* **2019**, *59*, 893–920. (b) Colomer, R.; Menéndez, J.A. Mediterranean diet, olive oil and cancer. *Clin. Transl. Oncol.* **2006**, *8*, 15–21.

⁸ Reboredo-Rodríguez, P.; Varela-López, A.; Forbes-Hernández, T.Y.; Gasparrini, M.; Afrin, S.; Cianciosi, D.; Zhang, J.; Manna, P.P.; Bompadre, S.; Quiles, J.L.; Battino, M.; Giampieri, F. Phenolic Compounds Isolated from Olive Oil as Nutraceutical Tools for the Prevention and Management of Cancer and Cardiovascular Diseases. *Int. J. Mol. Sci.* **2018**, *19*, 2305–2326.

⁹ (a) Cuyàs, E.; Gumuzio, J.; Lozano-Sánchez, J.; Carreras, D.; Verdura, S.; Llorach-Parés, L.; Sanchez-Martinez, M.; Selga, E.; Pérez, G.J.; Scornik, F.S.; Brugada, R.; Bosch-Barrera, J.; Segura-Carretero, S.; Martin, A.G.; Encinar, J.A.; Menendez, J.A. Extra Virgin Olive Oil Contains a Phenolic Inhibitor of the Histone Demethylase LSD1/KDM1A. *Nutrients* **2019**, *11*, 1656–1672. (b) Cuyàs, E.; Verdura, S.; Lozano-Sánchez, J.; Viciano, I.; Llorach-Parés, L.; Nonell-Canals, A.; Bosch-Barrera, J.; Brunet, J.; Segura-Carretero, A.; Sanchez-Martinez, M.; Encinar, J.A.; Menendez, J.A. The extra virgin olive oil phenolic oleacein is a dual substrate-inhibitor of catechol-O-methyltransferase. *Food Chem. Toxicol.* **2019**, *128*, 35–45. (c) Verdura, S.; Cuyàs, E.; Lozano-Sánchez, J.; Bastidas-Velez, C.; Llorach-Parés, L.; Fernández-Arroyo, S.; Hernández-Aguilera, A.; Joven, J.; Nonell-Canals, A.; Bosch-Barrera, J.; Martin-Castillo, B.; Vellon, L.; Sanchez-Martinez, M.; Segura-Carretero, S.; Menendez, J.A. An olive oil phenolic is a new chemotype of mutant isocitrate dehydrogenase 1 (IDH1) inhibitors. *Carcinogenesis* **2019**, *40*, 27–40. (d) Corominas-Faja, B.; Cuyàs, E.; Lozano-Sánchez, J.; Cufí, S.; Verdura, S.; Fernández-Arroyo, S.; Borrás-Linares, I.; Martin-Castillo, B.; Martin, Á.G.; Lupu, R.; Nonell-Canals, A.; Sanchez-Martinez, M.; Micol, V.; Joven, J.; Segura-Carretero, A.; Menendez, J.A. Extra-virgin olive oil contains a metabolo-epigenetic inhibitor of cancer stem cells. *Carcinogenesis* **2018**, *39*, 601–613.

proporcionando una función importante durante el desarrollo (diferenciación). Por otra parte, la proteína COMT también es una enzima dependiente de SAM, que cataliza la transferencia de grupos metilo de esta molécula a uno de los grupos hidroxilo de un sustrato catecol. Cabe destacar que la metilación del ADN se ha convertido en un objetivo importante en la terapia anticancerígena.¹⁰

1.2. Objetivos del Trabajo

1.2.1. Objetivos generales

Teniendo en cuenta el contexto en el que se enmarca este trabajo y sus antecedentes, se proponen los siguientes objetivos generales:

1. Determinar los compuestos fenólicos presentes en el aceite de oliva virgen extra capaces de inhibir a las proteínas DNMTs y COMT, y con posible uso como nutracéuticos para la prevención y el tratamiento del cáncer.
2. Analizar moléculas comerciales similares a los fenoles elegidos que podrían ser otra alternativa como agentes antiangiogénicos.

1.2.2. Objetivos específicos

Una vez planteados los objetivos generales para este trabajo, se pueden desglosar en estos objetivos más concretos:

1. Determinar los compuestos fenólicos presentes en el aceite de oliva virgen extra capaces de inhibir a las proteínas DNMTs y COMT, y con posible uso como nutracéuticos para la prevención y el tratamiento del cáncer.
 - 1.1. Realizar una búsqueda bibliográfica en profundidad sobre la angiogénesis, en concreto, del papel de las proteínas DNMTs y COMT en este proceso y en el cáncer, así como para conocer los antecedentes de compuestos fenólicos con actividad antiangiogénica.
 - 1.2. Determinar las estructuras 3D disponibles de COMT y DNMT1.
 - 1.3. Identificar los inhibidores conocidos de estas proteínas y examinar su modo de unión.
 - 1.4. Seleccionar los fenoles del AOVE y realizar el cribado virtual de estos compuestos frente a DNMT1 y COMT. Analizar el modo de unión para compararlo contra los inhibidores conocidos. Esto sirve como control de que puede, computacionalmente, unirse igual y, por tanto, tener una actividad similar.

¹⁰ Henning, S.M.; Wang, P.; Carpenter, C.L.; Heber, D. Epigenetic effects of green tea polyphenols in cáncer. *Epigenomics* **2013**, *5*, 729–741.

- 1.5. Evaluar las propiedades farmacocinéticas ADMET (Absorción, Distribución, Metabolismo, Excreción y Toxicidad) de los *hits*.
 - 1.6. Determinar las estructuras 3D disponibles de DNMT3A y DNMT3B. Identificar los inhibidores conocidos y examinar su modo de unión.
 - 1.7. Comprobar si los modos de unión de DNMT3A y DNMT3B están conservados respecto a DNMT1.
 - 1.8. Chequear si los fenoles determinados como posibles inhibidores de DNMT1 y COMT (*hits*), también podrían inhibir a DNMT3A y DNMT3B.
2. Analizar moléculas comerciales similares a los fenoles elegidos que podrían ser otra alternativa como agentes antiangiogénicos.
 - 2.1. Buscar moléculas comerciales similares a los *hits* del anterior apartado.
 - 2.2. Realizar el cribado virtual de las moléculas frente a DNMTs y COMT para determinar si se unen con las interacciones necesarias.
 - 2.3. Evaluar las propiedades farmacocinéticas ADMET de los *hits* comerciales.

1.3. Enfoque y método seguido

El descubrimiento de un nuevo fármaco es un proceso complejo con riesgos, que requiere de una gran cantidad de tiempo y una elevada inversión económica. Este proceso implica la identificación, síntesis, caracterización, cribado y ensayos de candidatos potenciales, denominados *hits*, que tras un largo proceso de optimización se convierten en *leads*, además de estudios preclínicos y clínicos para conseguir la aprobación para comercializar el fármaco. El tiempo habitual de investigación hoy en día es de 10 a 15 años con costes alrededor de 800 millones de dólares para que un compuesto pase a ser un posible fármaco. Tras los ensayos preclínicos y la confirmación de su acción y toxicidad, el compuesto se aprueba para realizar los ensayos clínicos que implican de 1 a 2 años antes de su lanzamiento al mercado.

Sin embargo, se puede acelerar este proceso empleando métodos computacionales y, en concreto, el descubrimiento de fármacos asistido por ordenador (CADD, *Computer-Aided Drug Design*),¹¹ que implica: (1) procesos de descubrimiento y desarrollo de fármacos optimizados, (2) identificación y optimización de los nuevos fármacos, y (3) diseño *in silico* de filtros para la eliminación de compuestos indeseables con propiedades con actividad inadecuada y/o malas propiedades de absorción, distribución, metabolismo, excreción y toxicidad (ADMET). En general, estos métodos se pueden clasificar en tres grupos. Por un lado, el CADD basado en la estructura necesita conocer la estructura 3D del receptor para calcular las energías de

¹¹ Tomar, V.; Mazumder, M.; Chandra, R.; Yang, J.; Sakharkar, M.K. Small Molecule Drug Design. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* 2019, 3, 741–760.

interacción, mientras que el CADD basado en ligandos se basan en la información de las estructuras químicas de un conjunto de ligandos con actividad biológica conocida, como por ejemplo, el modelado predictivo de la relación cuantitativa entre la estructura y la actividad (QSAR, *Quantitative Structure Activity Relationships*). Por último, están los métodos híbridos, que son una combinación de los dos anteriores y se pueden emplear cuando se conoce la estructura del receptor y de las moléculas activas, un ejemplo de este tipo son ciertas técnicas de modelado farmacóforo.¹²

El objetivo principal del presente trabajo es identificar potenciales candidatos a fármacos capaces de unirse a varias proteínas a partir de sus estructuras 3D, por lo que el enfoque computacional seleccionado es el descubrimiento de fármacos basado en estructura usando el método de cribado virtual. Este proceso de cribado virtual identifica compuestos con una biblioteca de productos químicos como referencia, de manera que comienza con el *docking* para predecir la mejor orientación de un compuesto unido al receptor. A continuación, emplea una función de puntuación para reevaluar la afinidad de unión de las estructuras predichas. Este método de cribado virtual es más rápido, rentable y necesita menos recursos en comparación con los métodos experimentales, como el cribado de alto rendimiento. Por tanto, se utiliza ampliamente para el descubrimiento de fármacos en una etapa temprana por farmacéuticas y grupos de investigación.

Otra técnica muy empleada en el diseño y desarrollo de fármacos, y que se va a emplear también en este trabajo, es el modelado farmacocinético de base fisiológica para evaluar la toxicidad y predecir *in silico* las propiedades farmacocinéticas ADMET de fármacos en seres humanos.¹¹ Así, aunque no es sencillo, los métodos computacionales permiten anticipar cómo se va a comportar un fármaco dentro de un organismo, lo que puede ayudar en la selección de los candidatos más prometedores durante las primeras fases del fármaco reduciendo los costes y el tiempo del proceso antes de llegar a los ensayos preclínicos y clínicos. Esto es muy importante porque este análisis de ADMET es uno de los cuellos de botella en el desarrollo de fármacos. Además, la selección de estas propiedades y sus rangos permitidos se pueden ajustar para obtener filtros para aplicaciones específicas.¹³

1.4. Planificación del Trabajo

Tras plantear el trabajo a realizar y concretar los objetivos, el siguiente paso es organizar el trabajo en tareas y proponer la planificación temporal. Para ello, se han tenido en cuenta las entregas programadas de la asignatura con sus respectivas fechas de inicio y fin, las cuales se muestran a continuación:

¹² Prieto-Martínez, F.D.; López-López, E.; Juárez-Mercado, K.E.; Medina-Franco, J.L. Chapter 2 - Computational Drug Design Methods—Current and Future Perspectives. *In Silico Drug Design, Repurposing Techniques and Methodologies* **2019**, 19–44.

¹³ Jorgensen, W.L. The Many Roles of Computation in Drug Discovery. *Science* **2014**, *303*, 1813–1818.

PEC1 – Plan de trabajo: 29/09/2020 – 13/10/2020

PEC2 – Desarrollo del trabajo (Fase 1): 14/10/2020 – 16/11/2020

PEC3 – Desarrollo del trabajo (Fase 2): 17/11/2020 – 14/12/2020

PEC4 – Cierre de la memoria: 15/12/2020 – 05/01/2021

PEC5a – Elaboración de la presentación: 06/01/2021 – 10/01/2021

PEC5b – Defensa pública: 13/01/2021 – 20/01/2021

1.4.1. Tareas

En la siguiente tabla se recogen las tareas realizadas en el presente trabajo a partir de los objetivos específicos propuestos para conseguir el cumplimiento de cada uno de ellos:

Objetivo	Tarea
1.1.	1.1.1. Realizar una revisión bibliográfica 1.1.2. Seleccionar la bibliografía para incluir en la memoria
1.2	1.2.1. Buscar las estructuras 3D de las proteínas DNMT1 y COMT
1.3.	1.3.1. Identificar los ligandos a los que se unen estas proteínas 1.3.2. Analizar el modo de unión proteína-ligando
1.4.	1.4.1. Examinar la base de datos OliveNet™ 1.4.2. Elaborar el listado de los fenoles del AOVE 1.4.3. Realizar el cribado virtual de los fenoles con DNMT1 y COMT 1.4.4. Hacer la lista de los fenoles más relevantes para estas proteínas
1.5.	1.5.1. Evaluar las propiedades farmacocinéticas ADMET de los <i>hits</i>
1.6.	1.6.1. Conocer las estructuras 3D disponibles de DNMT3A y DNMT3B 1.6.2. Identificar sus inhibidores y examinar su modo de unión
1.7.	1.7.1. Investigar sobre las herramientas para estudiar la conservación de proteínas 1.7.2. Estudiar la conservación en los modos de unión de las DNMTs
1.8.	1.8.1. Realizar el cribado virtual de los <i>hits</i> con DNMT3A y DNMT3B 1.8.2. Analizar los resultados obtenidos en el cribado virtual
2.1.	2.1.1. Buscar si los fenoles seleccionados como <i>hits</i> son comerciales 2.1.2. Elaborar una lista con moléculas comerciales similares a los <i>hits</i>
2.2.	2.2.1. Hacer el cribado virtual de estas moléculas 2.2.2. Determinar los <i>hits</i> comerciales
2.3.	2.3.1. Evaluar sus propiedades farmacocinéticas ADMET

1.4.2. Calendario

Para realizar la planificación temporal del trabajo se ha elaborado un diagrama de Gantt, empleando la herramienta *GanttProject*, una aplicación de gestión de proyectos de código abierto.¹⁴ En la Figura 2 se recoge la lista de las tareas propuestas con las fechas de inicio y fin de cada una, así como la duración en días de cada una. Para ello, se ha teniendo en cuenta las fechas de entrega de cada una de las actividades para entregar de la asignatura (PECs), que también se han añadido al calendario.



Nombre	Fecha de inicio	Fecha de fin	Duración
PEC1 - Plan de trabajo	29/09/20	13/10/20	15
PEC2 - Desarrollo del trabajo (Fase 1)	14/10/20	16/11/20	34
• 1.1.1. Realizar una revisión bibliográfica	14/10/20	16/10/20	3
• 1.1.2. Seleccionar la bibliografía para incluir en la memoria	16/10/20	16/10/20	1
• 1.2.1. Buscar las estructuras 3D de las proteínas DNMT1 y COMT	17/10/20	18/10/20	2
• 1.3.1. Identificar los ligandos a los que se unen estas proteínas	18/10/20	19/10/20	2
• 1.3.2. Analizar el modo de unión proteína-ligando	20/10/20	22/10/20	3
• 1.4.1. Examinar la base de datos OliveNet	23/10/20	23/10/20	1
• 1.4.2. Elaborar el listado de los fenoles del AOVE	23/10/20	25/10/20	3
• 1.4.3. Realizar el cribado virtual de los fenoles con DNMT1 y COMT (I)	26/10/20	13/11/20	19
• Preparar informe entregable PEC2	14/11/20	16/11/20	3
PEC3 - Desarrollo del trabajo (Fase 2)	17/11/20	14/12/20	28
• 1.4.3. Realizar el cribado virtual de los fenoles con DNMT1 y COMT (II)	17/11/20	18/11/20	2
• 1.4.4. Hacer la lista de los fenoles más relevantes para estas proteínas	19/11/20	20/11/20	2
• 1.5.1. Evaluar las propiedades farmacocinéticas ADMET de los hits	21/11/20	22/11/20	2
• 1.6.1. Conocer las estructuras 3D disponibles de DNMT3A y DNMT3B	23/11/20	24/11/20	2
• 1.6.2. Identificar sus inhibidores y examinar su modo de unión	25/11/20	25/11/20	1
• 1.7.1. Investigar sobre las herramientas para estudiar la conservación	26/11/20	26/11/20	1
• 1.7.2. Estudiar la conservación en los modos de unión de las DNMTs	27/11/20	28/11/20	2
• 1.8.1. Realizar el cribado virtual de los hits con DNMT3A y B	29/11/20	01/12/20	3
• 1.8.2. Analizar los resultados obtenidos en el cribado virtual	02/12/20	02/12/20	1
• 2.1.1. Buscar si los fenoles seleccionados como hits son comerciales	03/12/20	03/12/20	1
• 2.1.2. Elaborar una lista con moléculas comerciales similares a los hits	04/12/20	06/12/20	3
• 2.2.1. Hacer el cribado virtual de estas moléculas con DNMT y COMT	07/12/20	10/12/20	4
• 2.2.2. Determinar los hits comerciales	11/12/20	12/12/20	2
• 2.3.1. Evaluar sus propiedades farmacocinéticas ADMET	13/12/20	13/12/20	1
• Preparación informe entregable PEC3	13/12/20	14/12/20	2
PEC4 - Cierre de la memoria	15/12/20	05/01/21	22
PEC5a - Elaboración de la presentación	06/01/21	10/01/21	5
PEC5b - Defensa pública del TFM	13/01/21	20/01/21	8

Figura 2. Planificación temporal realizada con *GanttProject*

El diagrama de Gantt se muestra en la Figura 3, que recoge el calendario con el número de cada tarea a realizar en la parte izquierda de los bloques. Para facilitar la visualización del calendario, se ha empleado un código de colores según la PEC a la que pertenece cada tarea.

¹⁴ *GanttProject* [en línea] [fecha de consulta y descarga: 07 de Octubre de 2020]. Disponible en: <https://www.ganttproject.biz/>.

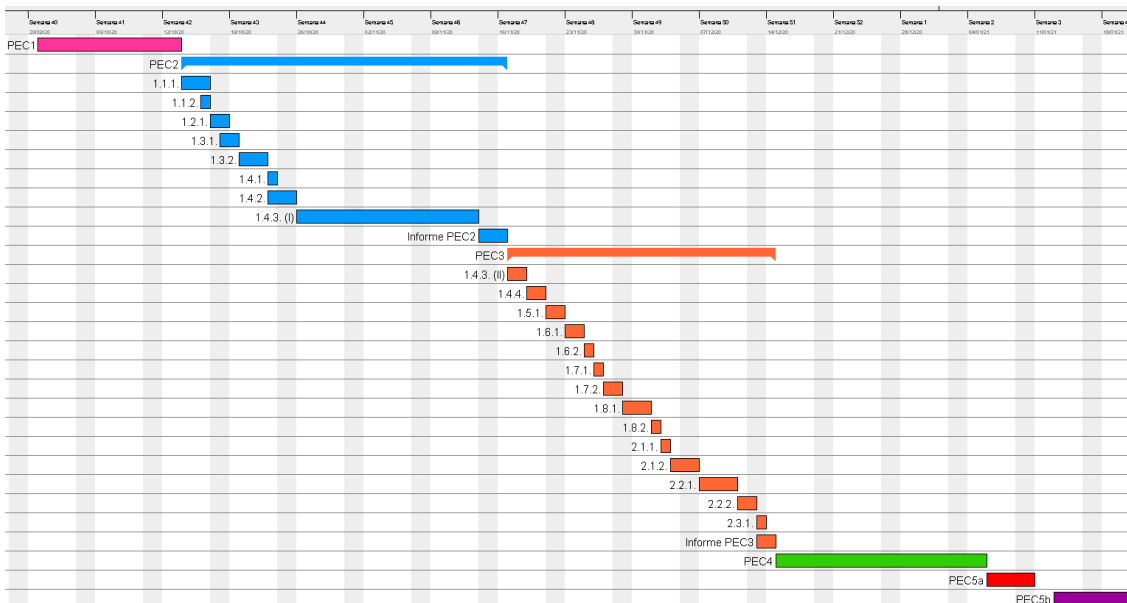


Figura 3. Diagrama de Gantt realizado con *GanttProject*

1.4.3. Hitos

Los hitos marcados para este trabajo se corresponden con las fechas de entrega de cada actividad entregable de la asignatura establecidas en el Plan Docente. Estas fechas se recogen en el diagrama de Gantt de la Figura 3, donde cada PEC se representa con un color para facilitar la visualización.

1.5. Breve resumen de productos obtenidos

A continuación, se mencionan los productos obtenidos durante el desarrollo del trabajo. Por un lado, estos son los archivos necesarios para superar la asignatura:

- **Memoria del trabajo final:** Se trata del presente documento y está elaborada siguiendo la estructura predeterminada en la asignatura.
- **Presentación:** Se realizará una exposición y defensa del trabajo realizado, para lo cual se elaborará una presentación empleando la herramienta Microsoft PowerPoint. Esta presentación recogerá un resumen del trabajo recogido en la memoria, incluyendo los resultados obtenidos y las conclusiones.

Por último, estos son los archivos de Microsoft Excel con algunos de los resultados obtenidos durante la realización del trabajo:

- Un documento que recoge los fenoles presentes en el AOVE con entrada en PubChem, denominado *fenoles_AOVE.xlsx*. También se especifica el nombre del fenol, la familia a la que pertenece, *InChi Key*, SMILES y el código CID de PubChem, así como el enlace del compuesto a esta página.

- Un documento con los resultados del cribado virtual de los fenoles del AOVE con las proteínas DNMT1 y COMT. Para cada una, se proporciona el nombre del fenol, su ID de modelo, la energía de enlace predicha en kcal/mol y el número de rotámeros (nRot). Además, para los 25 primeros resultados de cada estructura se indican las interacciones que se mantienen respecto de las del cristal y una última columna con la posición de ese fenol en la otra proteína. El nombre de este archivo es *output_table_fenoles.xlsx*.
- Un documento con el resultado las propiedades ADMET de los mejores fenoles del AOVE para DNMT1 y COMT, el cual se ha llamado *fenoles_pkcsm.xlsx*.
- Un documento con los compuestos comerciales similares a los fenoles del AOVE seleccionados, que han sido proporcionados por SwissSimilarity con una similitud igual o superior al 0.8, *swisssimilarity.xlsx*.
- Un documento con los resultados del cribado virtual de los compuestos comerciales similares a los fenoles del AOVE con las enzimas DNMT1 y COMT, *output_table_comerciales.xlsx*.
- Un documento con las propiedades ADMET calculadas para los mejores compuestos similares para DNMT1 y COMT, *comerciales_pkcsm.xlsx*.
- Un documento con los resultados del cribado virtual de los *hits* de los fenoles del AOVE y de los compuestos comerciales similares a estos fenoles con las proteínas DNMT3A y DNMT3B, *output_table_dnmt3.xlsx*.

1.6. Breve descripción de los otros capítulos de la memoria

Una vez realizada la introducción en la que se establecen los objetivos de la investigación, el contexto en el que se plantea y la planificación del trabajo, parece razonable realizar una breve revisión sobre el conocimiento actual de las proteínas DNMTs y COMT, así como en los procesos en las que se encuentran involucradas, con foco en el cáncer y los procesos de angiogénesis. Esta información se recoge en el segundo capítulo 2 de la Memoria (Antecedentes). En este capítulo también se realiza una ampliación de los compuestos con actividad antiangiogénica descritos en la literatura a los ya comentados en la Introducción, con especial atención en los compuestos naturales y los fenoles.

En el tercer capítulo (Materiales y métodos) se describen en detalle las herramientas y procedimientos seguidos para la realización de las tareas propuestas. En el cuarto capítulo se recoge la discusión de los resultados obtenidos en el trabajo realizado, donde se analizan en detalle cada uno de los pasos ejecutados y los resultados obtenidos en cada uno de ellos. En el quinto capítulo se muestran las principales conclusiones derivadas de la investigación desarrollada y en el sexto capítulo se recopila el glosario con las abreviaturas y los acrónimos empleados. Finalmente, en el séptimo capítulo se enumeran las referencias bibliográficas utilizadas durante la elaboración de la presente Memoria y en el octavo están los anexos del trabajo.

2. Antecedentes

2.1. Metilación del ADN

La epigenética estudia los mecanismos que regulan la actividad genética sin cambiar la secuencia del ADN y conducen a modificaciones que pueden transmitirse a las células hijas, aunque algunos cambios epigenéticos pueden revertirse. Estos procesos son esenciales para muchas funciones del organismo, pero si ocurren de manera incorrecta, pueden producirse importantes efectos adversos en la salud y el comportamiento. De hecho, alteraciones en las modificaciones epigenéticas se han relacionado con varias enfermedades, como trastornos neurodegenerativos, diabetes, cáncer y enfermedades inmunológicas.¹⁵

Uno de los principales procesos epigenéticos es la metilación del ADN cuyo papel es crucial para el desarrollo normal, ya que se ha demostrado que participa en muchos procesos, como la diferenciación de tejidos y la impronta genómica. En células normales asegura la regulación adecuada de la expresión génica y el silenciamiento génico estable, mientras que patrones de metilación del ADN aberrantes están asociados con muchas enfermedades, como el inicio y progreso de varios tipos de cáncer.¹⁶ En los mamíferos, la metilación del ADN se produce mediante un grupo de enzimas llamadas ADN metiltransferasas (DNMTs) con S-adenosil-L-metionina (SAM) como donante de metilo. Este grupo metilo se transfiere a la posición C5 del anillo pirimidina de una citosina, normalmente de un dinucleótido citosina-fosfato-guanina (CpG). Como resultado, se genera 5-metilcitosina (5mc) y S-adenosil-L-homocisteína (SAH), que se libera.

En el genoma humano, entre un 3% y un 5% de las citosinas se encuentran metiladas como 5mc, del cual el 70% está dentro de secuencias CpG, siendo estos dinucleótidos menos abundantes respecto a otros. Además, su distribución es desigual, ya que suelen localizarse en regiones promotoras de genes. En concreto, las regiones cortas ricas en CpG, denominadas “islas CpG”, se localizan en más del 50% de los promotores de genes humanos. Así, la hipermetilación de las islas CpG en estas regiones conduce al silenciamiento de genes, mientras que los promotores de genes transcripcionalmente activos suelen estar hipometilados.¹⁷

¹⁵ Weinhold; B. Epigenetics: The Science of Change. *Environ. Health. Perspect.* **2006**, *114*, A160–A167.

¹⁶ Jurkowska, R.Z.; Jurkowski, T.P.; Jeltsch, A. Structure and Function of Mammalian DNA Methyltransferases. *ChemBioChem* **2011**, *12*, 206–222.

¹⁷ Wang, Y.; Leung, F.C. An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics* **2004**, *20*, 1170–1177.

2.1.1. Enzimas DNMTs

En los seres humanos, se encuentran tres ADN metiltransferasas activas, llamadas DNMT1, DNMT3A y DNMT3B, y una proteína relacionada que carece de actividad catalítica, denominada DNMT3L, pero necesaria para potenciar la actividad de DNMT3A y DNMT3B actuando como cofactor. Aunque también existe la proteína DNMT2, en realidad, se trata de una enzima ARN metiltransferasa que actúa sobre el ARN de transferencia y ahora se la conoce como t-ARN metiltransferasa de ácido aspártico 1 (TRDMT1).¹⁸

El patrón de metilación inicial lo realizan las metiltransferasas *de novo* DNMT3 sobre ADN no metilado, proceso que continúa el resto de la vida del organismo con pequeños cambios específicos de los tejidos. Así, tras cada replicación del ADN, se genera ADN hemimetilado, ya que la hebra parental está metilada, pero la hebra hija recién sintetizada no lo está. Luego, la hebra hija es metilada por la metiltransferasa de mantenimiento DNMT1, que tiene una mayor preferencia por el ADN hemimetilado para evitar la metilación de sitios previamente no metilados y conservar así el patrón a través de las siguientes divisiones celulares.¹⁶ En cambio, las enzimas DNMT3A/B tienen afinidades similares por los sustratos de ADN tanto no metilados como hemimetilados.¹⁹

En un principio parecía que había una redundancia parcial entre las dos metiltransferasas *de novo* en el establecimiento de patrones de metilación en todo el genoma. Sin embargo, un *knockout* de DNMT3A o DNMT3B produce letalidad postnatal o embrionaria, respectivamente, lo que indica las distinciones funcionales de las DNMTs *de novo*. Por un lado, DNMT3A es fundamental para establecer la metilación de secuencias críticas para el desarrollo tardío o después del nacimiento, mientras que DNMT3B desempeña un papel importante en el desarrollo embrionario temprano. Además, las mutaciones de DNMT3A son frecuentes en cánceres hematológicos, como la leucemia mieloide aguda (LMA), y también ocurren en un síndrome de sobrecrecimiento del desarrollo. Por el contrario, mutaciones de DNMT3B conducen al síndrome de inmunodeficiencia, inestabilidad centromérica y anomalías faciales (ICF), lo que indica que DNMT3B está especializado en la metilación del ADN de regiones particulares del genoma.²⁰

¹⁸ Ghanbarian, H.; Wagner N.; Polo, B.; Baudouy, D.; Kiani, J.; Michiels, J.-F.; Cuzin, F.; Rassoulzadegan, M.; Wagner, K.-D. Dnmt2/Trdmt1 as Mediator of RNA Polymerase II Transcriptional Activity in Cardiac Growth. *PLoS One*. **2016**, 11, e0156953.

¹⁹ Okano, M.; Xie, S.; Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **1998**, 19, 219–220.

²⁰ Gao, L.; Emperle, M.; Guo, Y.; Grimm, S.A.; Ren, W.; Adam, S.; Uryu, H.; Zhang, Z.-M.; Chen, D.; Yin, J.; Dukatz, M.; Anteneh, H.; Jurkowska, R.Z.; Lu, J.; Wang, Y.; Bashtrykov, P.; Wade, P.A.; Wang, G.G.; Jeltsch, A.; Son, J. Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive *de novo* DNA methylation mechanisms. *Nat. Commun.* **2020**, 11, 3355.

2.1.1.1. Reacción de metilación del ADN

La reacción de metilación del ADN catalizada por estas enzimas, se inicia por el ataque nucleofílico de un residuo cisteína conservado en el motivo PCQ (motivo IV, Figura 5) de estas proteínas a la posición C6 del anillo de pirimidina de la citosina, tal como se muestra en el esquema de reacción de la Figura 4. Este paso es necesario porque la citosina no es lo suficientemente nucleófila y requiere que la enzima DNMT active su posición C5. Esto da lugar a un intermedio rico en electrones similar a una enamina (**A** en la Figura 4), que luego ataca a la molécula SAM para transferir el grupo metilo al C5 de la citosina y con la consecuente formación de SAH. El intermedio metilado tras una β eliminación regenera la enzima y 5mc.²¹

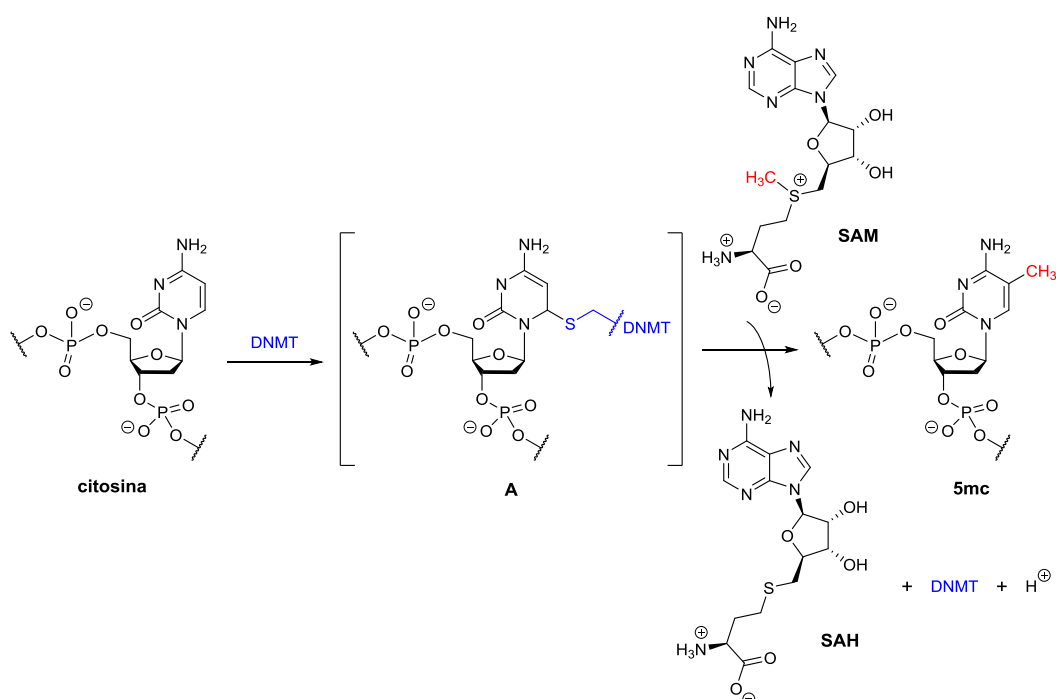


Figura 4. Esquema de reacción de metilación de una citosina por las enzimas DNMTs

2.1.1.2. Estructura y dominios

Todas las proteínas DNMTs constan de una región grande multidominio N-terminal con funciones reguladoras y un dominio metiltransferasa C-terminal más pequeño, tal como se muestra en la Figura 5. Tanto DNMT1 como DNMT3A y DNMT3B tienen un dominio catalítico (CD) conservado entre las enzimas eucariotas y procariontas, denominado pliegue de Rossmann, que consiste en una hoja β mixta de siete hebras con tres hélices α a cada lado y que está implicado en la unión del cofactor (motivos I y X, Figura 5) y en la catálisis (motivos IV, VI y VIII). En cambio, el dominio de reconocimiento de diana (TRD), localizado entre la hoja β central y la última hélice α

²¹ Ganesan, A. Epigenetic drug discovery: a success story for cofactor interference. *Phil. Trans. R. Soc. B.* **2018**, 373–388.

del núcleo catalítico (entre los motivos VIII y IX), no tiene similitud de secuencia entre DNMT1 y DNMT3, a pesar de que participa en el reconocimiento y especificidad del ADN. De hecho, excepto por el CD conservado, las regiones N-terminales reguladoras son totalmente diferentes en secuencia y estructura en DNMT1 y DNMT3.²²

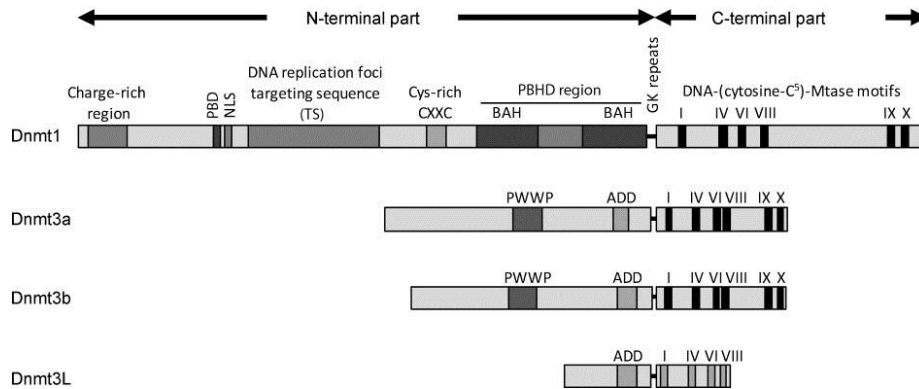


Figura 5. Arquitectura de los dominios de las proteínas DNMTs

2.1.1.3. DNMT1

La enzima DNMT1 humana es una proteína grande formada por 1616 aminoácidos y codificada por el gen DNMT1 localizado en el cromosoma 19. Las regiones N-terminal y C-terminal de DNMT1 están unidas por una serie de repeticiones lisina-glicina (KG), tal como se muestra en la Figura 5. Además, el dominio N-terminal está formado por varios subdominios que median interacciones moleculares:

- Dominio de interacción rico en carga (DMAP1) involucrado en la estabilidad de la proteína y en su interacción con el represor transcripcional DMAP1 (una proteína que une la DNMT1 con la acetilación de histonas).
- Dominio de unión (PBD) de antígeno nuclear celular en proliferación (PCNA), que media la interacción con PCNA y se une a otras proteínas.
- Dominio de focalización de focos de replicación (TS), necesario para dirigir DNMT1 a la cromatina centromérica y a los focos de replicación.
- Dominio CXXC, un dominio de dedo de zinc que contiene ocho residuos de cisteína conservados. Se une específicamente a CpG no metilado del ADN.
- Dos dominios de homología bromo-adyacente (BAH1 y BAH2) que pertenecen al dominio de homología polibromático (PBHD). Aunque aún se desconoce su papel funcional en la proteína.²³

²² Liang, Z.; Zhu, Y.; Liu, X.; Hu, G. Role of protein-protein interactions in allosteric drug design for DNA methyltransferases. *Adv. Protein Chem. Struct. Biol.* **2020**, *121*, 49–84.

²³ Lyko, F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* **2018**, *19*, 81–92.

2.1.1.4. DNMT3

La proteína DNMT3A está formada por 912 aminoácidos y está codificada por el gen DNMT3A humano ubicado en el cromosoma 2, mientras que la proteína DNMT3B codificada por el gen DNMT3B del cromosoma 20 está formada por 853 aminoácidos. Finalmente, la proteína DNMT3L tiene una longitud de 386 aminoácidos y el gen que la codifica es el DNMT3L del cromosoma 21.

Las proteínas DNMT3A y DNMT3B tienen estructuras similares con dominios catalíticos conservados para mantener la función de la enzima y catalizar la metilación del ADN, con aproximadamente un 85% de similitud de secuencia. Además, la región N-terminal de cada una está formada por dos subdominios (Figura 5):

- Dominio ADD (ATRX-DNMT3-DNMT3L) es la región de unión a los iones zinc enriquecida con cisteína, que también media la interacción con otras proteínas, así como con histonas no modificadas para mejorar la actividad de la enzima.
- Dominio prolina-triptófano-triptófano-prolina (PWWP), media la unión específica del ADN y el direccionamiento a la cromatina pericentromérica. Esta región de 100-150 aminoácidos está débilmente conservada entre DNMT3A y DNMT3B, a excepción del motivo prolina-triptófano.

En cuanto a la proteína DNMT3L, presenta una clara homología con las enzimas DNMT3A y DNMT3B, aunque no tiene el dominio PWWP, los motivos IX y X, ni todos los residuos catalíticos del dominio C-terminal. Por tanto, es catalíticamente inactiva e incapaz de unirse al cofactor SAM y se une débilmente al ADN. Sin embargo, es capaz de unirse a DNMT3A y DNMT3B interactuando con sus CD para estimular la actividad de estas enzimas, presumiblemente induciendo un cambio conformacional que facilita la unión y catálisis del ADN y SAM. Por tanto, actúa como regulador positivo de las DNMTs *de novo*.¹⁶

Las estructuras de DNMT3A y DNMT3B en complejo con DNMT3L son un heterotetrámero lineal formado por dos moléculas de DNMT3L (en los bordes del tetrámero) y otras dos de DNMT3A o DNMT3B (en el centro), como se ha demostrado con la resolución de sus estructuras cristalinas. El dominio C-terminal de DNMT3A/B proporciona dos interfaces para los contactos proteína-proteína: una interfaz FF hidrófoba y una interfaz polar RD. La interfaz FF también está presente en DNMT3L y crea el contacto entre ambas proteínas. En cambio, la interfaz RD no se encuentra en DNMT3L y es la que interviene en la interacción central del dímero DNMT3A o DNMT3B y forma el sitio de unión al ADN.

Finalmente mencionar que además de la metilación de CpG, las enzimas DNMT3A y DNMT3B producen metilación en otros dinucleótidos, principalmente CpA, en ovocitos, células madre embrionarias y células neuronales. Se ha relacionado la presencia de metilación sin CpG con la represión transcripcional y el estado epigenético asociado a la pluripotencia, lo que apoya la metilación sin CpG como una marca epigenética emergente en la definición de patrones de expresión génica específicos de tejido, particularmente en el cerebro.²⁰

2.1.2. DNMTs y la angiogénesis

Como se ha comentado, la metilación adecuada del ADN es esencial para el buen desarrollo y funcionamiento celular, por lo que una anomalía en este proceso puede provocar diversas enfermedades, incluido el cáncer. De hecho, las células tumorales muestran un metiloma diferente al de las células normales, que causa un aumento de la inestabilidad genómica y la activación de oncogenes silenciados, así como una disminución de la expresión de genes supresores de tumores.²⁴ En células normales, las islas CpG de los promotores de genes supresores de tumores no están metiladas, pero la hipometilación de todo el genoma y la hipermetilación de promotores ocurren en muchos cánceres y están asociadas con la tumorigénesis. La hipermetilación de las islas CpG en los promotores de genes puede silenciar los genes supresores de tumores y afectar procesos celulares clave como el ciclo celular, la apoptosis, la reparación del ADN, las interacciones celulares y la angiogénesis.²⁵

La desregulación de la angiogénesis, el crecimiento de nuevos vasos sanguíneos a partir de una vasculatura existente, es un proceso esencial en el cáncer, además de en otras muchas enfermedades humanas. Como tal, la angiogénesis tumoral es importante para suministrar oxígeno y nutrientes a los tumores, al mismo tiempo que participa en la desregulación metabólica y la diseminación/metástasis tumoral.⁵ El papel de las proteínas DNMTs en este proceso es muy importante, ya que están involucradas en la regulación de la transcripción durante la angiogénesis, como se ha mencionado anteriormente.²⁶

Estudios realizados sobre las características patológicas de DNMTs *in vivo* e *in vitro* en modelos animales muestran que la pérdida de estas proteínas tiene un impacto importante en los organismos. Como el realizado por Belinsky y col. para analizar su papel funcional en la formación de tumores usando un modelo de ratón con un alelo del gen DNMT1 alterado. Observaron que la reducción de DNMT1 conduce a una disminución en la incidencia de cáncer de pulmón inducido por carcinógenos del tabaco y a una reducción en la actividad de DNMT en los neumocitos que dan lugar a tumores.²⁷ Sin embargo, otros estudios ponen de manifiesto que estas enzimas DNMTs tienen diferentes roles dependiendo del tipo de cáncer, un resumen se muestra en la Figura 6. De hecho, para DNMT3A se observó que un aumento de su expresión en el cáncer de pulmón conduce a la proliferación y metástasis de células tumorales y, en contraste, la falta de DNMT3A facilita la progresión del tumor. Estos resultados opuestos sugieren que puede actuar como oncogén y como un gen supresor de tumores en el cáncer de pulmón, consecuencia de las diferentes características de esta enzima en diferentes procesos, lo que se conoce como la paradoja de DNMT.²⁵

²⁴ Kulis, M.; Esteller, M. DNA methylation and cancer. *Adv. Genet.* **2010**, *70*, 27–56.

²⁵ Zhang, J.; Yang, C.; Wu, C.; Cui, W.; Wang, L. DNA Methyltransferases in Cancer: Biology, Paradox, Aberrations, and Targeted Therapy. *Cancers* **2020**, *12*, 2123.

²⁶ Ribatti, D.; Tamma, R. Epigenetic control of tumor angiogenesis. *Microcirculation* **2020**, *27*, e12602.

²⁷ Belinsky, S.A.; Klinge, D.M.; Stidley, C.A.; Issa, J.P.; Herman, J.G.; March, T.H.; Baylin, S.B. Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer Res.* **2003**, *63*, 7089–7093.

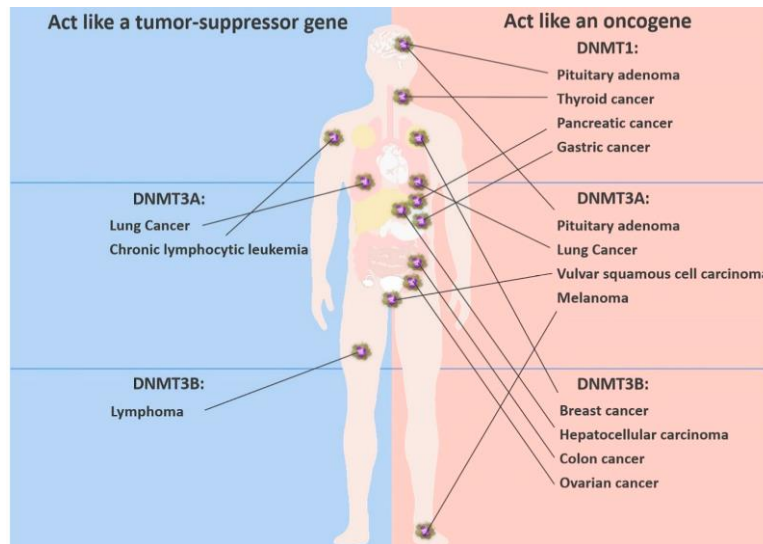


Figura 6. Los diferentes roles de DNMTs en varios cánceres y la paradoja de DNMT

2.1.3. Inhibición de las proteínas DNMTs

Como otros procesos epigenéticos, la metilación del ADN es reversible y se puede modular por agentes químicos, a diferencia de los cambios genéticos, lo que crea un objetivo para estrategias terapéuticas anticancerígenas. Se ha demostrado que los inhibidores de las enzimas DNMTs pueden reactivar genes supresores de tumores silenciados epigenéticamente y disminuir así el crecimiento de células tumorales, además de inhibir la angiogénesis *in vitro* e *in vivo*.²⁸ Cabe destacar que en los últimos años, la inhibición de la angiogénesis tumoral se ha convertido en una estrategia anticancerígena en consonancia con la quimioterapia, la radioterapia y la cirugía, dada la importancia del interruptor angiogénico durante el desarrollo temprano del tumor.

En este contexto, la interferencia de cofactores ha surgido como una estrategia con gran potencial para la búsqueda de fármacos que inhiban enzimas epigenéticas. Hay varios procedimientos para conseguir la interferencia con un cofactor de sitio activo que son: (1) desplazamiento competitivo del cofactor de su sitio de unión, (2) unión reversible con el cofactor que evita la ocupación del sustrato o la catálisis productiva, y (3) inactivación irreversible del cofactor. Aunque los tres enfoques han logrado un progreso significativo contra los objetivos epigenéticos, en el caso de las proteínas DNMTs el más empleado es el desplazamiento del cofactor SAM para conseguir la inhibición.²⁵ En primer lugar, mencionar que el subproducto de la transferencia del grupo metilo SAH es un potente inhibidor de las proteínas DNMTs muy bien conocido, ya que es capaz de unirse a ellas.²⁹

²⁸ Hellebrekers, D.M.; Jair, K.W.; Viré, E.; Eguchi, S.; Hoebbers, N.T.; Fraga, M.F.; Esteller, M.; Fuks, F.; Baylin, S.B.; van Engeland, M.; Griffioen, A.W. Angiostatic activity of DNA methyltransferase inhibitors. *Mol. Cancer Ther.* **2006**, *5*, 467–475.

²⁹ Isakovic, L.; Saavedra, O.M.; Llewellyn, D.B.; Claridge, S.; Zhan, L.; Bernstein, N.; Vaisburg, A.; Elowe, N.; Petschner, A.J.; Rahil, J.; Beaulieu, N.; Gauthier, F.; MacLeod, A.R.; Delorme, D.;

Los inhibidores de DNMTs azacitidina (Vidaza®) y 5-aza-2'-deoxicitidina (Decitabina) (Figura 7), análogos de la citidina, han sido aprobados para el tratamiento del síndrome mielodisplásico y la leucemia. Estos compuestos pueden atrapar DNMT y evitar que se produzca la metilación del ADN de manera efectiva. Esto tiene lugar ya que experimentan una activación metabólica del nucleótido de la molécula (**1** en Figura 7), que se incorpora al ADN e inhibe irreversiblemente a DNMTs con la formación de un aducto covalente (**3** en Figura 7).²¹ Sin embargo, estos fármacos son inestables en disolución acuosa, muestran baja especificidad, no son eficaces en tumores sólidos y tienen importantes efectos secundarios debido a su toxicidad. Aunque se pueden utilizar en dosis bajas, lo que reduce la toxicidad. Por otra parte, el desarrollo de zebularina, otro análogo de nucleósido que también inhibe DNMTs, es más estable en solución acuosa y sus efectos son menos tóxicos, pero se requiere una dosis mucho mayor para un tratamiento eficaz. Cabe destacar que con una combinación de zebularina y tratamiento con decitabina se consigue mejorar la actividad antitumoral.²⁴

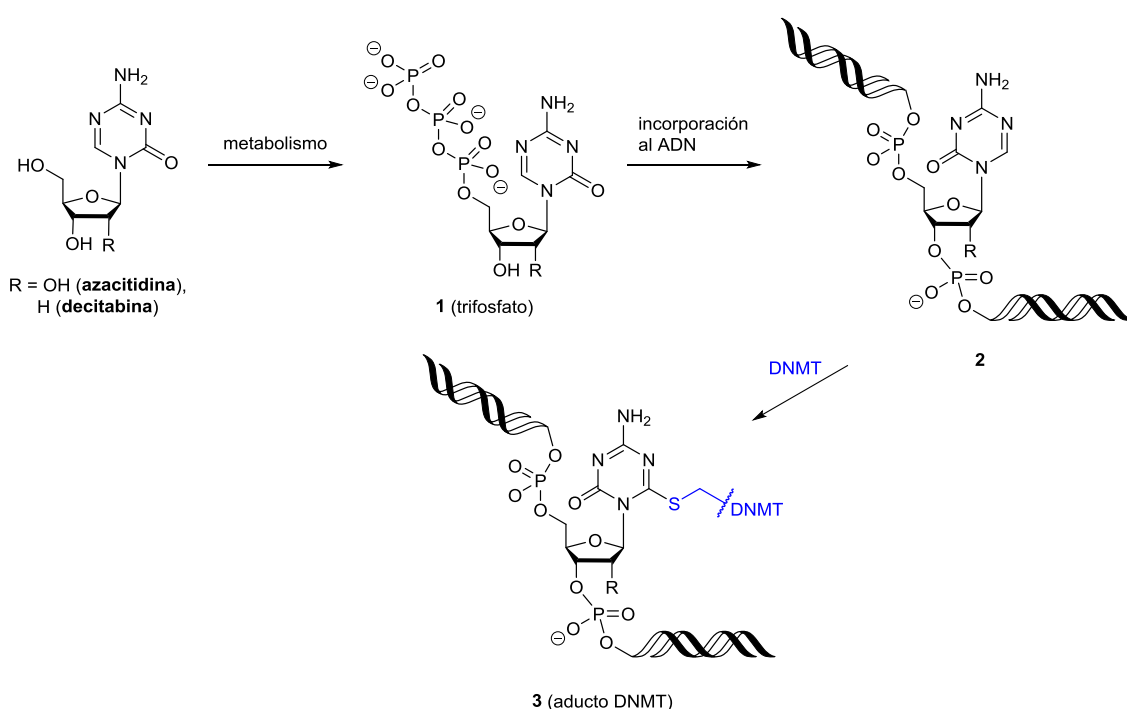


Figura 7. Mecanismo de acción de los inhibidores de DNMT azacitidina y decitabina

El riesgo de usar análogos de nucleósidos como inhibidores de DNMTs se debe a que se incorporan al ADN, lo que puede provocar toxicidad y efectos secundarios. Por esta razón, ha aumentado el interés de los análogos no nucleósidos, como los productos naturales genisteína, (-)-epigallocatequina-3-O-galato (EGCG) y ácido lacaico A; fármacos reutilizados como hidralazina (antagonista de la hipertensión), procainamida y procaína (agente antiarrítmico y anestésico, respectivamente); moléculas pequeñas sintéticas como el ftalimido-L-triptófano (RG108) y la quinolona

Besterman, J.M.; Wahhab, A. Constrained (L-)-S-adenosyl-L-homocysteine (SAH) analogues as DNA methyltransferase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2742–2746.

1027 (Figura 8). En comparación con los análogos de nucleósidos, es menos probable que estos se incorporen al ADN y, por lo tanto, inhiben la metilación del ADN de una forma más segura, aunque con una menor potencia. Además, ninguno de ellos se ha introducido hasta ahora con éxito en la práctica clínica.³⁰

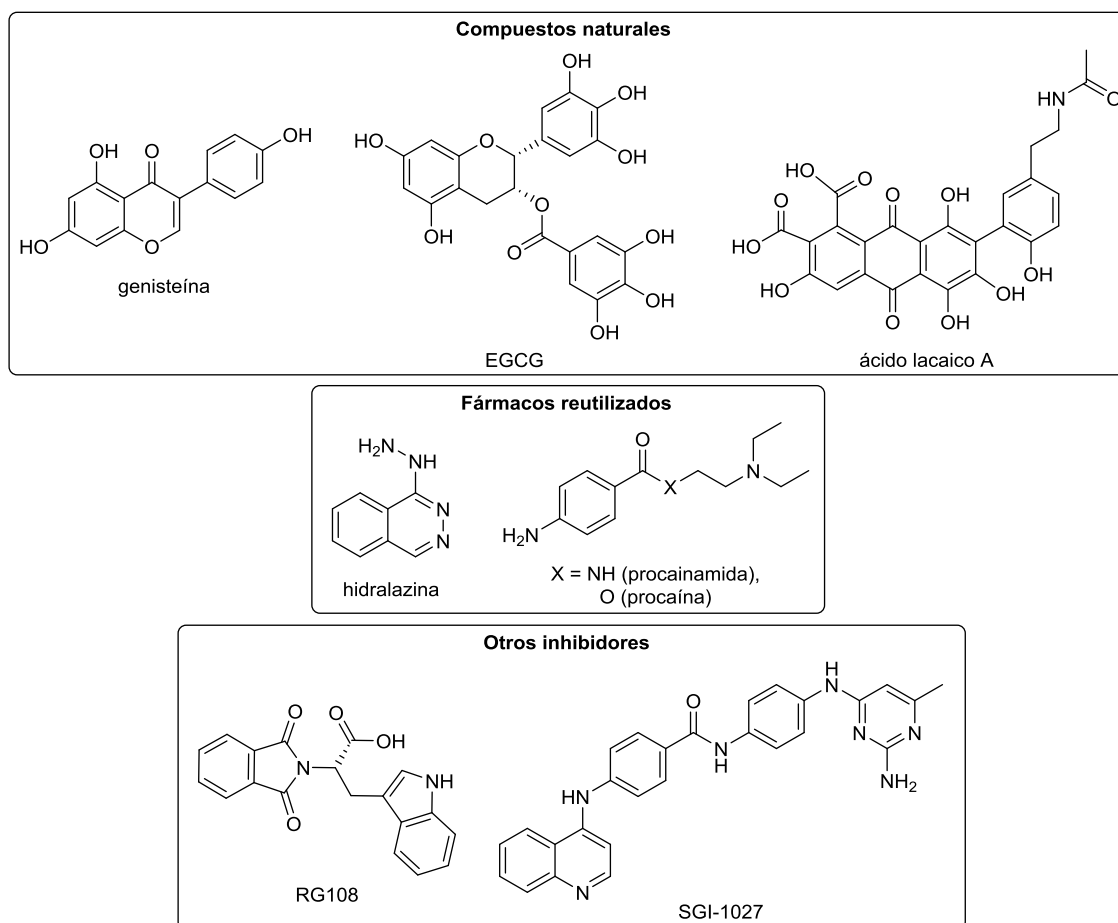


Figura 8. Inhibidores de DNMTs análogos no nucleósidos

Teniendo en cuenta todo esto, es importante identificar compuestos alternativos que superen estos inconvenientes y limitaciones de la terapia actualmente disponible. También sería interesante tener en cuenta que el cáncer también es un problema en aumento en los países en desarrollo, de manera que los tratamientos de bajo coste tienen un gran potencial. En este contexto, los fitoquímicos pueden ser una buena opción para el tratamiento anticancerígeno, ya que, en general, constituyen una clase de fármacos naturales baratos, con muy baja toxicidad y de acción amplia, especialmente si se administran combinados en dosis relativamente bajas.

³⁰ Chen, S.; Wang, Y.; Zhou, W.; Li, S.; Peng, J.; Shi, Z.; Hu, J.; Liu, Y.C.; Ding, H.; Lin, Y.; Li, L.; Cheng, S.; Liu, J.; Lu, T.; Jiang, H.; Liu, B.; Zheng, M.; Luo, C. Identifying novel selective non-nucleoside DNA methyltransferase 1 inhibitors through docking-based virtual screening. *J. Med. Chem.* **2014**, *57*, 9028–9041.

2.2. Enzima COMT

La proteína catecol-O-metiltransferasa (COMT) es una enzima que cataliza la reacción de O-metilación de una gran variedad de sustratos catecol endógenos y exógenos, incluidas las catecolaminas, en presencia de un catión magnesio y usando SAM como dador del grupo metilo, dando lugar a catecol O-metilado y SAH como productos de la reacción. Esta proteína se ha encontrado en la mayoría de los organismos estudiados, incluyendo levaduras, plantas, insectos, peces, anfibios, aves y todos los mamíferos, indicando su relevancia evolutiva y fisiológica.³¹

En humanos, además de participar en el metabolismo de las catecolaminas, la enzima COMT está implicada en la inactivación de catecolestrógenos, xenobióticos que contienen catecol, como catequinas y bioflavonoides, e intermedios indol del metabolismo de la melanina. Ejerce protección como barrera enzimática contra los catecoles exógenos al evitar su oxidación a o-quinonas y la formación de electrófilos. La O-metilación de catecol disminuye la tendencia al ciclo redox y aumenta la lipofilicidad facilitando el transporte transmembrana fuera de la célula. También tiene una gran importancia farmacológica, al estar involucrada en el metabolismo de fármacos con catecol, como L-DOPA, dobutamina, carbidopa, isoprenalina, rimiterol, benserazida, α -metildopa y apomorfina, para el tratamiento del asma, la hipertensión y la enfermedad de Parkinson. Sin embargo, en el sistema nervioso central, COMT desactiva metabólicamente los neurotransmisores de catecolaminas (dopamina y norepinefrina), lo que tiene un efecto potencialmente dañino que contribuye a un mayor riesgo de trastornos neurodegenerativos, como las enfermedades de Parkinson y Alzheimer. Además, contribuye al desarrollo de trastornos cardiovasculares y enfermedades con influencia hormonal, incluido el cáncer.

2.2.1. Formas de COMT

La COMT humana existe en dos formas diferentes, una soluble en el citoplasma (COMT soluble: S-COMT) y otra asociada a membranas (COMT unida a membrana: MB-COMT). Entre ambas existen algunas diferencias, como su ubicación celular, ya que en la mayoría de los tejidos, la gran parte de COMT se encuentra en su forma soluble y solo una pequeña fracción como MB-COMT. Sin embargo, en el cerebro alrededor del 70% se encuentra como MB-COMT y el resto como S-COMT.³² Además, S-COMT está constituida por 221 aminoácidos con un peso molecular de 24.4 kDa, mientras que la forma MB-COMT contiene 50 aminoácidos adicionales en la región N-terminal, de los cuales unos 20 sirven como un anclaje de membrana hidrofóbico, y tiene un peso molecular de 30 kDa. De manera que ambas formas tienen la misma

³¹ Bastos, P.; Gomes, T.; Ribeiro, L. Catechol-O-methyltransferase (COMT): an update on its role in cancer, neurological and cardiovascular diseases. *Rev. Physiol., Biochem. Pharmacol.* **2017**, *173*, 1–39.

³² Bai, H.-W.; Shim, J.-Y.; Yu, J.; Zhu, B.T. Biochemical and Molecular Modeling Studies of the O-Methylation of Various Endogenous and Exogenous Catechol Substrates Catalyzed by Recombinant Human Soluble and Membrane-Bound Catechol-O-Methyltransferases. *Chemical Research in Toxicology* **2007**, *20*, 1409–1425.

secuencia de aminoácidos a excepción de esos últimos 50 residuos, así que también comparten varias propiedades bioquímicas.

Las dos formas de la proteína COMT están codificadas por un solo gen ubicado en el cromosoma 22, denominado gen COMT, que se encuentra representado en la Figura 9. Este gen está regulado por dos promotores (P1 y P2) y tiene 6 exones, los dos primeros no codificantes. En el sexto exón se encuentra el codón de terminación (TGA) compartido por las dos formas de la proteína.³³

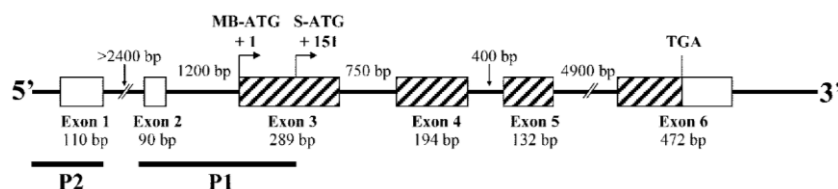


Figura 9. Esquema de la estructura del gen COMT humano

Los sitios activos de ambas formas tienen estructuras idénticas, afinidad similar por SAM y un pH óptimo cercano (7.4 para MB-COMT y 7.8 para S-COMT). Pero difieren en la cinética y, en comparación, MB-COMT tiene una afinidad entre 10 y 100 veces mayor hacia catecolaminas, mientras que S-COMT tiene una mayor afinidad por catecolestrógenos y xenobióticos que contienen catecol, y los metila dos o tres veces más rápido que a las catecolaminas. Esto sugiere que MB-COMT actúa a concentraciones fisiológicas bajas preferentemente, mientras que S-COMT lo hace en condiciones de saturación del sustrato. En consecuencia, MB-COMT parece tener un papel muy relevante en la inactivación de la neurotransmisión catecolaminérgica, mientras que S-COMT desempeña un papel más importante en la inactivación de catecolestrógenos y xenobióticos que contienen catecol.³¹

2.2.2. Estructura de COMT

La proteína COMT se compone de un único dominio de ocho hélices α formando dos subconjuntos que rodean siete hojas β con la sexta en sentido antiparalelo. Esta topología es característica de la familia de metiltransferasas dependientes de SAM (pliegue de Rossmann), al igual que las proteínas DNMTs, pero que difiere en la presencia de un bucle extendido entre las hojas β 6 y 7, que se corresponde con la región C-terminal de la estructura y que forma parte del sitio de unión del catecol. Por otra parte, las hélices 6 y 7 contienen una vuelta de una estructura helicoidal π y una estructura helicoidal 310, respectivamente, posiblemente debido a la presencia de residuos de prolina en las posiciones 124 y 149, que rompen las hélices.³⁴

³³ Tenhunen, J.; Salminen, M.; Lundström, K.; Kiviluoto, T.; Savolainen, R.; Ulmanen, I. Genomic organization of the human catechol O-methyltransferase gene and its expression from two distinct promoters. *Eur. J. Biochem.* **1994**, *223*, 1049–1059.

³⁴ Rutherford, K.; Le Trong, I.; Stenkamp, R.E.; Parson, W.W. Crystal Structures of Human 108V and 108M Catechol O-Methyltransferase. *J. Mol. Biol.* **2008**, *380*, 120–130.

El sitio activo de COMT está ubicado en la superficie de la enzima y consta de dos partes distintas, una región más profunda para la unión del cofactor SAM y el sitio catalítico para la unión del sustrato donde el ion Mg^{2+} ocupa la posición central.³⁴ El ligando SAM únicamente se une cuando la enzima está formando un complejo con el ion Mg^{2+} , aunque no hay interacción directa entre ambos cofactores. Además, el magnesio participa en la reacción enzimática y ayuda a retener en la orientación adecuada del sustrato catecol, al unirse a los dos grupos hidroxilo del sustrato, de manera que se propone que el reconocimiento del catecol depende de esta coordinación con el magnesio. Este catión completa su entorno octaédrico mediante la unión a las cadenas laterales de dos residuos de ácido aspártico y uno de asparagina y la sexta posición de coordinación está ocupada por una molécula de agua.³⁵

2.2.3. Reacción de O-metilación

La reacción de O-metilación mediada por la enzima COMT comienza con la unión del catión Mg^{2+} y de la molécula SAM a la proteína, y por último del sustrato catecol. Así, el magnesio se une a los dos oxígenos de los grupos hidroxilo del catecol (**1** en Figura 10), de manera que uno de los hidroxilos queda colocado en las proximidades tanto del grupo metilo de SAM como del grupo amina de la lisina 144 del sitio activo de la proteína. Esto facilita la desprotonación del catecol por este grupo amina para convertirlo en un nucleófilo más potente (**2** en Figura 10). El grupo hidroxilo restante está unido por enlaces de hidrógeno al oxígeno carboxílico del glutamato en posición 199, mientras que el anillo de catecol se mantiene en la posición correcta por los residuos triptófano 38 y 143 y por la prolina 174 que forman una barrera hidrofóbica y definen la selectividad enzimática hacia las cadenas laterales de catecol. Por lo tanto, otro papel del Mg^{2+} es reducir el pK_a de los hidroxilos del catecol para facilitar su ionización.³⁶ Tras este paso, se propone un ataque nucleófilo mediante un mecanismo S_N2 del oxígeno del hidroxilato al grupo metilsulfonio deficiente de electrones del cofactor SAM. El grupo metilo que se transfiere procede con una inversión neta de la configuración, adoptando una geometría similar a sp^2 en el estado de transición (**TS** en Figura 10). Después de la reacción, el producto desmetilado SAH es el último en disociarse.³⁷

³⁵ Vidgren, J.; Svensson, L.A.; Liljas, A. Crystal structure of catechol O-methyltransferase. *Nature* **1994**, *368*, 354–358.

³⁶ Moschovou, K.; Melagraki, G.; Mavromoustakos, T.; Zacharia, L.C.; Afantitis, A. Cheminformatics and virtual screening studies of COMT inhibitors as potential Parkinson's disease therapeutics. *Expert Opinion on Drug Discovery* **2020**, *15*, 53–62.

³⁷ Czarnota, S.; Johannissen, L.O.; Baxter, N.J.; Rummel, F.; Wilson, A.L.; Cliff, M.J.; Levy, C.W.; Scrutton, N.S.; Waltho, J.P.; Hay, S. Equatorial Active Site Compaction and Electrostatic Reorganization in Catechol-O-methyltransferase. *ACS Catalysis* **2019**, *9*, 4394–4401.

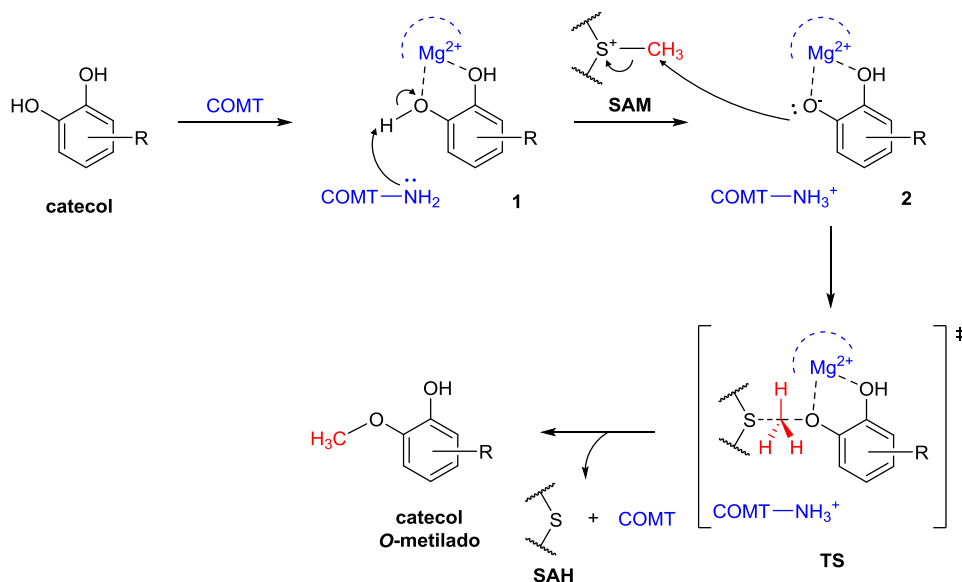


Figura 10. Mecanismo propuesto para la reacción de O-metilación mediada por COMT

A pesar de ser esencial para la catálisis, el ion magnesio puede ser intercambiado por otros cationes, como Mn^{2+} , Co^{2+} , Zn^{2+} , Cd^{2+} , Fe^{2+} , Sn^{2+} o Ni^{2+} , aunque con cambios en la cinética enzimática, como en el caso de Mn^{2+} y Co^{2+} que aumentan la actividad enzimática. En cambio, los metales calcio, hierro (III) y aluminio inhiben la actividad de la enzima.³¹

2.2.4. Polimorfismo de COMT

En 1995, Lotta y col. identificaron una transición de guanina a adenina en el gen COMT que conduce a un polimorfismo de un único nucleótido (SNP) en la proteína COMT humana, denominado Val^{108/158}Met y que da lugar al intercambio de un residuo valina por metionina en la posición 158 de la cadena polipeptídica de la forma MB-COMT y en la posición 108 de S-COMT.³⁸ Este polimorfismo ha sido ampliamente estudiado y hace que la enzima sea más propensa a la distorsión del sitio activo y la agregación de proteínas a temperatura fisiológica, lo que conduce a una reducción de los niveles de proteína *in vivo*, así como de la actividad enzimática en cuatro veces para individuos homocigóticos Met108/158, respecto a Val108/158. Aunque, se ha determinado que el cofactor SAM protege a la proteína COMT con Met108/158 contra la inactivación termal y la desnaturalización. Los genotipos Val/Val, Val/Met y Met/Met están asociados con fenotipos de actividad de la enzima alta, intermedia y baja, respectivamente.³⁹ La distribución de los alelos es distinta entre grupos étnicos, encontrándose el alelo de baja actividad con una frecuencia aproximada del 50% en

³⁸ Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J. Kinetics of human soluble and membrane-bound catechol O-methyltransferase: A revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* **1995**, *34*, 4202–4210.

³⁹ Shield, A.; Thomaes, B.; Eckloff, B.W.; Wieben, E.D.; Weinshilboum, R.M. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Molecular Psychiatry* **2004**, *9*, 151–160.

caucásicos, 49% en el suroeste de Asia, 18–30% en asiáticos orientales, 2.6% en afroamericanos y 2.6–3.6% en africanos.³¹

El polimorfismo Val^{108/158}Met de COMT se asocia con muchas enfermedades psiquiátricas y estudios realizados indican que participa en la obesidad y la diabetes *mellitus*. De hecho, el alelo de alta actividad se asocia a una mayor protección contra la diabetes de tipo II,⁴⁰ mientras que Met^{108/158} se relaciona con un mayor riesgo de trastorno obsesivo-compulsivo, algunas manifestaciones de esquizofrenia y una mayor sensibilidad al dolor, pero también se correlaciona con una mejor función cognitiva prefrontal.³⁴ Sin embargo, a pesar de que la proteína COMT tiene un papel central en la reparación del ADN y la carcinogénesis inducida por estrógenos, las observaciones realizadas hasta la fecha apuntan a dos direcciones contradictorias en cuanto al polimorfismo Val^{108/158}Met. Por un lado, el alelo de baja actividad se asocia con un mayor riesgo de cáncer de mama y progresión tumoral, mientras que otros estudios sugieren que este polimorfismo no está relacionado con este tipo de cáncer.⁴⁰ Además, hay otras evidencias que indican que este polimorfismo podría estar involucrado en el desarrollo de otros tipos de cáncer, como el de vejiga y esófago.⁴¹

2.2.5. Inhibidores de la proteína COMT

La O-metilación rápida mediada por la enzima COMT de grandes cantidades de catecoles no solo disminuye significativamente las reservas intracelulares de SAM, sino que también forma cantidades equimolares de SAH tras la transferencia del grupo metilo al sustrato. Sin embargo, SAH es un potente inhibidor no competitivo de COMT al interactuar con la proteína a través del mismo sitio de unión que el cofactor SAM y con una afinidad de unión similar. Cuando la enzima está unida a SAM es el estado activo y es capaz de catalizar la reacción de O-metilación, pero cuando está unida a SAH es el estado inactivo y no es capaz de completar la reacción de metilación al no tener el grupo metilo.⁴² Además de este inhibidor de retroalimentación, hay otros tipos de inhibidores que se pueden dividir en dos grandes grupos.

- Inhibidores de primera generación

Un tipo de inhibidores de primera generación de la proteína COMT son los competitivos que en su estructura contienen catecol o algunos segmentos bioisostéricos. Algunos son derivados del pirogalol, ácido cafeico, ácido gálico o flavonoides, como rutina o quercetina (Figura 11). Otros compuestos no catecólicos, como el ácido ascórbico y las tropolonas, también se han identificado como inhibidores de COMT (Figura 11). Sin embargo, muchos no han demostrado ser fármacos

⁴⁰ Srivastava, S.P.; Goodwin, J.E. Cancer Biology and Prevention in Diabetes. *Cells* **2020**, *9*, 1380.

⁴¹ Zhou, Q.; Wang, Y.; Chen, A.; Tao, Y.; Song, H.; Li, W.; Tao, J.; Zuo, M.; Mo, F.; Li, C. Association between the COMT Val158Met polymorphism and risk of cancer: evidence from 99 case-control studies. *Onco. Targets Ther.* **2015**, *8*, 2791–2803.

⁴² Zhu, B.T. Catechol-O-Methyltransferase (COMT)-mediated methylation metabolism of endogenous bioactive catechols and modulation by endobiotics and xenobiotics: importance in pathophysiology and pathogenesis. *Curr. Drug. Metab.* **2002**, *3*, 321–349.

adecuados debido a su baja selectividad y, aunque no muestran alta toxicidad, presentan signos de hepatotoxicidad.³⁶

- Inhibidores de segunda generación

Por otro lado, aunque el catecol y sus derivados endógenos se metilan fácilmente, los derivados de catecol que poseen sustituyentes electronegativos son potentes inhibidores de esta enzima, ya que estos sustituyentes hacen que los hidroxilos ionizados sean menos nucleófilos que los correspondientes grupos hidroxilo del catecol no sustituido. Un ejemplo de este tipo de inhibidores es el compuesto 3,5-dinitrocatecol (DNC), que se ha conseguido cocrystalizar con COMT y determinar su estructura mediante difracción de rayos X.³⁴ El diseño de este tipo de inhibidores de COMT no catecolicos tiene cierto interés, porque las propiedades farmacocinéticas, como las reacciones metabólicas y la penetración de la barrera hematoencefálica, podrían modificarse.⁴³ En este contexto, los fármacos tolcapona (Tasmar[®]) y entacapona (Comtan[®]) recogidos en la Figura 11, que contienen un grupo 3-nitrocatecol como unidad estructural central, son inhibidores reversibles de la proteína COMT y se han aprobado para el tratamiento de la enfermedad del Parkinson. Son efectivos en bajas concentraciones, selectivos y activos por vía oral.⁴⁴

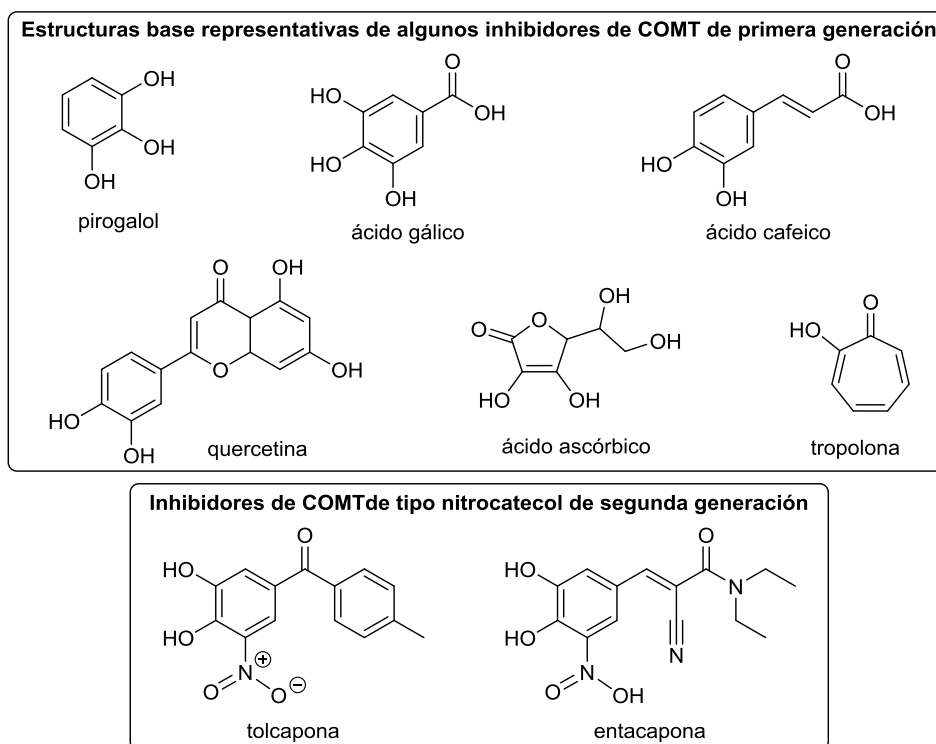


Figura 11. Ejemplos de los diferentes tipos de inhibidores de la enzima COMT

⁴³ Ovaska, M.; Yliniemelä, A. A semiempirical study on inhibition of catechol O-methyltransferase by substituted catechols. *J. Comput. Aided Mol. Des.* **1998**, *12*, 301–307.

⁴⁴ Masjost, B.; Ballmer, P.; Borroni, E.; Zürcher, G.; Winkler, F.K.; Jakob-Roetne, R.; Diederich, F. Structure-Based Design, Synthesis, and in vitro Evaluation of Bisubstrate Inhibitors for Catechol O-Methyltransferase (COMT). *Chem. Eur. J.* **2000**, *6*, 971–982.

Finalmente, mencionar los inhibidores bisustrato que tienen dos fragmentos conjugados, un cofactor y un análogo del sustrato unidos. En este caso concreto de los inhibidores bisustrato de COMT, estos están formados por un resto catecol (“bolsillo catecol”) unido covalentemente a la adenosina de SAM (“bolsillo adenosina”) a través de un enlazador adecuado, tal como se muestra en la Figura 12.³⁶

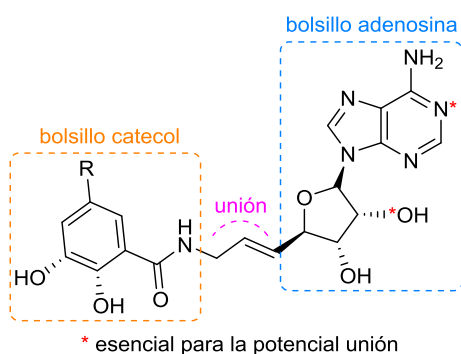


Figura 12. Esquema de un inhibidor bisustrato de la enzima COMT

2.3. Sustratos/inhibidores de las proteínas DNMTs y COMT

El consumo de polifenoles de origen vegetal se ha relacionado con la protección dietética contra enfermedades relacionadas con el envejecimiento, incluido el cáncer. En concreto, se han descrito varios polifenoles naturales que contienen catecol como sustratos excelentes de la *O*-metilación mediada por COMT. A su vez los polifenoles naturales que contienen catecol también pueden comportarse como inhibidores de COMT, impidiendo la *O*-metilación de una variedad de sustratos de catecol.^{10,45} Este mecanismo dual se ha explicado en términos de inhibición competitiva directa del centro catalítico de COMT por los propios polifenoles catecolícos o inhibición no competitiva debido a los niveles elevados de SAH, generado durante la reacción de *O*-metilación catalizada por COMT. Como se ha comentado previamente, SAH es un inhibidor de retroalimentación muy potente de varias metiltransferasas dependientes de SAM, incluida la propia COMT y las enzimas DNMTs. De hecho, dado que COMT y DNMTs pertenecen a esta misma familia con estructuras catalíticas centrales comunes, los polifenoles dietéticos que contienen catecol pueden afectar a la tasa de metilación del ADN no solo a través de mecanismos indirectos debidos a SAH, sino también a través de la ocupación inhibitoria directa de los centros catalíticos de las enzimas DNMTs (Figura 13).⁴⁶

⁴⁵ Zhu, B.T.; Wang, P.; Nagai, M.; Wen, Y.; Bai, H.-W. Inhibition of human catechol-*O*-methyltransferase (COMT)-mediated *O*-methylation of catechol estrogens by major polyphenolic components present in coffee. *J. Steroid. Biochem. Mol. Biol.* **2009**, *113*, 65–74.

⁴⁶ Lee, W.J.; Zhu, B.T. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. *Carcinogenesis* **2006**, *27*, 269–277.

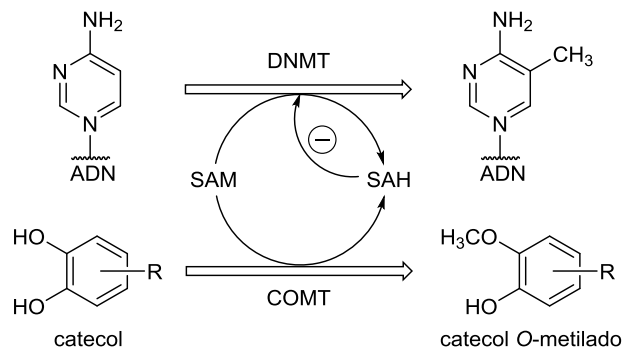


Figura 13. Esquema de la metilación mediada por DNMTs y su modulación por la reacción de O-metilación de catecoles por COMT

Este efecto se ha demostrado en diversos estudios, como el realizado con el ácido cafeico y el ácido clorogénico (polifenoles comunes del café que contienen catecol), que son dos inhibidores de la metilación del ADN. Se observó que la presencia de COMT mejora en gran medida la potencia de inhibición de ambos compuestos, lo que indica el mecanismo dual de inhibición. Los autores indicaban que aunque la ingesta diaria de ácido cafeico y ácido clorogénico no es suficiente, hay muchos polifenoles catecólicos en nuestra dieta, como las catequinas del té y los bioflavonoides, excelentes sustratos para la metilación por COMT. Por lo que la ingesta total de este tipo de polifenoles en la dieta podría ser un factor acumulativo significativo que contribuye al aumento de la cantidad intracelular de SAH y, como consecuencia, a la inhibición de la metilación del ADN de manera significativa.⁴⁶

2.4. El AOVE y los fenoles

La evidencia científica indica una gran influencia de la dieta en la salud, siendo la dieta mediterránea muy interesante. En concreto, los beneficios potenciales para la salud de la dieta mediterránea asociados en gran medida al consumo del aceite de oliva virgen extra (AOVE), principal fuente de grasa, incluyen la prevención del cáncer, así como una reducción del riesgo de padecer enfermedades cardiovasculares, neurodegenerativas y diabetes. En este contexto, la angioprevención es un concepto importante a tener en cuenta, ya que la prevención de la angiogénesis a través de compuestos bioactivos del AOVE podría explicar en parte el efecto quimiopreventivo de esta dieta en el cáncer.⁴⁷

El aceite de oliva contiene una mezcla compleja de diferentes tipos de compuestos, cuya proporción varía en función del tipo de aceituna, maduración, condiciones de cultivo, almacenamiento, método de extracción y grado de refinamiento. El AOVE es el aceite extraído de la aceituna (*Olea europaea*) solo por medios mecánicos y consumido sin más refinamiento, cuyos componentes principales,

⁴⁷ Martínez-Poveda, B.; Torres-Vargas, J.A.; Ocaña, M.C.; García-Caballero, M.; Medina, M.A.; Quesada, A.R. The Mediterranean Diet, a Rich Source of Angiopreventive Compounds in Cancer. *Nutrients* **2019**, *11*, 2036.

en general, son los ácidos grasos (98-99%), sobre todo ácido oleico (55-83%) y ácido linoleico (hasta 20%). En cambio, la fracción restante (1-2%) está formada por una mayor diversidad de compuestos, como alcoholes, fenoles, esteroides y pigmentos, responsables de estabilizar y proteger la integridad del aceite, conservando sus características organolépticas. Históricamente, las propiedades nutracéuticas del AOVE se atribuían a su alto contenido de ácidos grasos. Sin embargo, ahora se conoce que muchos de sus efectos beneficiosos son debidos a los compuestos fenólicos, los cuales poseen una amplia gama de propiedades bioactivas, como actividades antioxidantes, antiinflamatorias y quimiopreventivas.⁴⁸

Algunos de los compuestos fenólicos del AOVE más investigados han resultado ser agentes quimiopreventivos con capacidad para inhibir la angiogénesis, como la oleuropeína (un fenol de tipo secoiridoide). En el caso del hidroxitirosol y tirosol, que son los fenoles simples más abundantes del aceite de oliva y también los principales constituyentes fenólicos del vino, han mostrado actividades antitumorales y antiinflamatorias,⁸ pero solo el hidroxitirosol inhibe la angiogénesis *in vivo* e *in vitro*. Además, se han encontrado otros fenoles del AOVE menos estudiados con actividad antiangiogénica, como el flavonoide taxifolina. También se han determinado otro tipo de compuestos minoritarios en el aceite de oliva con este tipo de propiedades, que incluyen carotenoides, vitamina E, diterpenoides y esteroides.⁴⁷

Para el presente trabajo, uno de los antecedentes fenólicos del AOVE más destacables es el secoiridoide oleaceína (Figura 1). En varios estudios realizados por Javier A. Menéndez y col. se observó que la oleaceína puede ser un potencial nutracéutico para el tratamiento antiangiogénico, al ser un potente inhibidor de algunas proteínas involucradas en este proceso.⁹ En concreto, se ha demostrado que es un inhibidor competitivo de las enzimas DNMTs,^{9d} pero también muestra un comportamiento dual sustrato/inhibidor de la proteína COMT.^{9b} Por lo que la oleaceína puede actuar como un inhibidor dual DNMTs/COMT y, como consecuencia, es capaz de producir efectos acumulativos sobre la tasa de metilación del ADN con la inhibición no competitiva de DNMTs a través de la formación de SAH durante su metilación catalizada por COMT y a través de la inhibición directa de DNMTs.

Sin embargo, a pesar de los diversos estudios realizados con otros compuestos fenólicos naturales, como los contenidos en el café y el té, la mayoría de los compuestos fenólicos del AOVE aún no han sido analizados como posibles inhibidores de la angiogénesis para usarse como fármacos en terapia anticancerígena. Por lo que en este proyecto, se ha analizado el potencial de todos estos compuestos naturales para inhibir las proteínas DNMTs y COMT implicadas en procesos de angiogénesis mediante experimentos de unión *in silico* utilizando el cribado virtual.

Finalmente, señalar que estos compuestos fenólicos se encuentran en el AOVE en dosis bajas, por lo que la ingesta diaria promedio puede no ser una carga abrumadora para el sistema de metilación mediado por las proteínas COMT y DNMTs. Por lo que se cree que los beneficios preventivos del AOVE en la angiogénesis podrían residir en la combinación de varios compuestos bioactivos.

⁴⁸ Olive Wellness Institute™ [en línea] [fecha de consulta: 29 de Diciembre de 2020]. Disponible en: <https://olivewellnessinstitute.org/>

3. Materiales y métodos

3.1. Búsqueda bibliográfica

En este trabajo se ha realizado una búsqueda bibliográfica sobre la estructura y funciones en el organismo de las proteínas DNMTs y COMT, así como de su papel en el cáncer y en los procesos de angiogénesis. También se ha realizado una búsqueda sobre los antecedentes con actividad antiangiogénica descritos previamente en la literatura, con foco en los de compuestos fenólicos. Un resumen de la información recopilada durante este proceso se ha recogido en el segundo capítulo de la presente Memoria. Además, esta búsqueda ha permitido conocer el contexto y asentar las bases del proyecto, y ha ayudado a la realización del resto de los objetivos propuestos.

Esta búsqueda ha sido realizada principalmente mediante la base de datos Scopus de la empresa Elsevier. Aunque hay otras alternativas disponibles, esta herramienta es la mayor base de datos de citas y resúmenes de bibliografía revisada por pares (revistas científicas, libros y actas de conferencias). Además, es rápida y fácil de usar, y es un recurso disponible en la biblioteca de la Universitat Oberta de Catalunya.

3.2. Compuestos fenólicos del AOVE

La evidencia sobre los efectos beneficiosos para la salud del consumo de AOVE se debe, al menos en parte, a los potentes efectos antioxidantes y antiinflamatorios que han demostrado los compuestos derivados de las aceitunas (*Olea europaea*). Por lo que ha habido un aumento en la investigación en este campo, sin embargo, la diversidad y complejidad de los compuestos derivados de la aceituna, así como la matriz en la que se encuentran (fruto, hoja, aceite, orujo y aguas residuales), han dificultado su caracterización y aislamiento. Además, aunque en la literatura está disponible esta información para la mayoría de ellos, no es fácil buscarla, sobre todo para encontrar compuestos con ciertas propiedades químicas o biológicas. En este contexto y con el fin de recopilar toda esta información disponible, recientemente, se ha creado la base de datos gratuita OliveNetTM por los investigadores E. McCord, T. Karagiannis y col., que incorpora todos los compuestos derivados de las diversas matrices de la aceituna identificados hasta la fecha.⁴⁹

⁴⁹ (a) Bonvino, N.P.; Liang, J.; McCord, E.D.; Zafiris, E.; Benetti, N.; Ray, N.B.; Hung, A.; Boskou, D.; Karagiannis, T.C. OliveNetTM: a comprehensive library of compounds from *Olea europaea*. *Database* **2018**, 1–9. (b) OliveNetTM [en línea] [fecha de consulta: 23 de Octubre de 2020]. Disponible en: <https://mccordresearch.com.au/>.

A través de la revisión de 181 publicaciones científicas, OliveNet™ describe un total de 676 compuestos distintos asociados con la aceituna, cuya información se obtuvo de las bases de datos públicas PubChem, Human Metabolome Database (HMDB) y Chemical Entities of Biological Interest (ChEBI) con los respectivos enlaces a cada una. Estos compuestos se dividen en 13 clases principales (fenoles, ácidos grasos, alcoholes alifáticos y aromáticos, esteroides, fosfolípidos, ácidos triterpénicos, volátiles, hidrocarburos, azúcares, pigmentos, tocoferoles, aminoácidos y una pequeña cantidad de otros compuestos no clasificados) y 47 subclases, como se muestra en la Figura 14, donde cada clase se representa por un color.

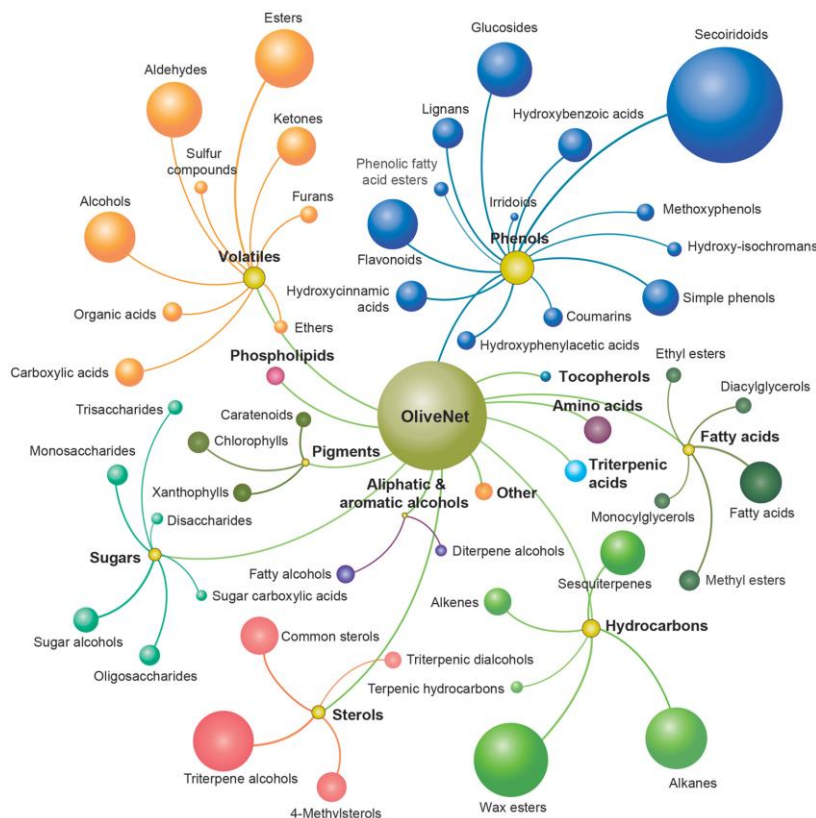


Figura 14. Composición de la base de datos OliveNet™: 676 compuestos agrupados en 13 clases (en negrita y por colores) y 47 subclases (el tamaño de la bola refleja el número de compuestos)

En concreto para la clase fenol, OliveNet™ recoge 222 compuestos divididos en 13 subclases, las cuales son cumarinas, flavonoides, glucósidos, hidroxiisocromanos, ácidos hidroxibenzoicos, ácidos hidroxicinámicos, ácidos hidroxifenilacéticos, iridoides, lignanos, metoxifenoles, ésteres de ácidos grasos fenólicos, secoiridoides y fenoles simples. En la Figura 15 se recoge la representación realizada para esta clase con una breve información sobre cada subclase.

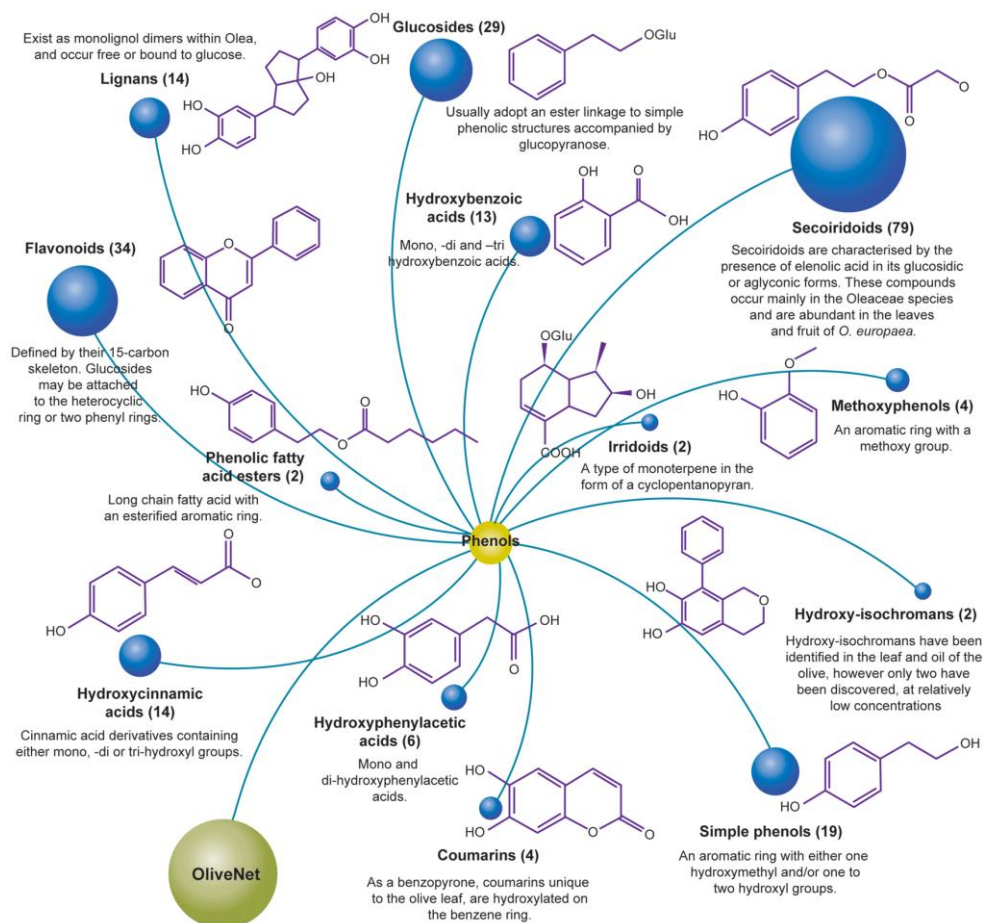


Figura 15. La base de datos OliveNet™ recoge 222 compuestos fenólicos divididos en 13 subclases. Se proporciona el número de compuestos de cada subclase, las estructuras químicas representativas y la información clave relevante

A partir de esta información, se ha empleado esta base de datos para obtener la lista de los compuestos fenólicos del AOVE. Sin embargo, aunque OliveNet™ emplea una interfaz web simple y fácil de usar, la base de datos no se puede descargar y hay que realizar la búsqueda manualmente de cada fenol. Por lo tanto, se ha decidido emplear únicamente los fenoles que tengan entrada a la base de datos abierta de química PubChem del National Institutes of Health (NIH),⁵⁰ un recurso clave que proporciona información sobre estructuras químicas, identificadores, propiedades químicas y físicas, entre otros muchos, principalmente de moléculas pequeñas, pero también recoge moléculas más grandes como nucleótidos, carbohidratos, lípidos, péptidos y macromoléculas modificadas químicamente. Siguiendo este procedimiento, se ha elaborado una lista con 129 fenoles del AOVE (Anexo I). Para cada uno se ha descargado su estructura 3D en formato SDF, además de recopilar en un archivo de Microsoft Excel su nombre, subclase a la que pertenece, *InChi Key*, SMILES, código CID de PubChem y el enlace a esta página.

⁵⁰ Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E.E. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **2019**, *8*, D1102–D1109.

3.3. Softwares para la visualización y tratamiento de estructuras

3.3.1. PyMOL

Las figuras de las proteínas y ligandos recogidas en la presente Memoria se han realizado empleando el software PyMOL un programa para la visualización de moléculas a partir de varios tipos de ficheros (The PyMOL Molecular Graphics System, Version 2.0.7 Schrödinger, LLC; en <https://pymol.org/2/>). PyMOL también incorpora distintas herramientas para trabajar con las moléculas, pero en este trabajo solo se ha utilizado para la visualización y la generación de imágenes.

3.3.2. UCSF Chimera

En este trabajo se ha empleado el programa de visualización UCSF Chimera (versión 1.14), gratuito para el uso académico, que además de la visualización interactiva, permite el análisis de estructuras moleculares. Ha sido desarrollado por UCSF Resource for Biocomputing, Visualization and Informatics y apoyado en parte por el NIH.⁵¹ En concreto, se ha utilizado para preparar las estructuras de las proteínas para el *docking*, así como de los ligandos a partir de los ficheros SDF de los fenoles y de los SMILES de los compuestos comerciales obtenidos con SwissSimilarity. Otro uso ha sido el análisis de los resultados del *docking* y la preparación de los complejos proteína–ligando tras el *docking*. Finalmente, UCSF Chimera también ha servido para ayudar a comprobar si los modos de unión de las proteínas DNMT3A y DNMT3B se encuentran conservados respecto a DNMT1, junto con una búsqueda bibliográfica.

3.4. Estructuras 3D de las proteínas diana

Para realizar el cribado virtual es necesario disponer de la estructura 3D de las proteínas diana. En este trabajo, la búsqueda de estas estructuras se ha realizado mediante la base de datos UniProtKB (Universal Protein Resource Knowledgebase),⁵² el cual proporciona información funcional sobre proteínas con anotaciones precisas y completas. Además de datos básicos para cada entrada, como la secuencia de aminoácidos, el nombre y descripción de la proteína, UniprotKB también contiene un resumen sobre las estructuras 3D disponibles, junto con los enlaces a la base de datos PDB (Protein Data Bank), a través de RSCB PDB (Research Collaboratory for Structural Bioinformatics PDB).⁵³ Este último recurso permite que los datos de estructuras 3D para grandes moléculas biológicas, como proteínas, ADN y ARN, de la base de datos PDB estén disponibles para toda la comunidad científica.

⁵¹ Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, *25*, 1605–1612.

⁵² The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515.

⁵³ Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

3.5. Cálculos del cribado virtual

Todos los cálculos de cribado virtual con las proteínas DNMTs y COMT se han realizado empleando el servidor web MTiOpenScreen (<https://bioserv.rpbs.univ-paris-diderot.fr/services/MTiOpenScreen/>).⁵⁴ El cribado virtual a través de MTiOpenScreen usa el programa AutoDock Vina para realizar el *docking*, el cual emplea una función empírica y un enfoque de búsqueda conformacional basado en gradientes a partir de la estructura 3D de una proteína para predecir la orientación de los ligandos en los sitios de unión de las cavidades de la proteína y su energía de enlace.⁵⁵ AutoDock Vina define el espacio de búsqueda mediante una caja (*grid box*) definida por las coordenadas del centro de la caja y sus dimensiones. La puntuación de las poses generadas en el *docking* y la clasificación de los ligandos está basada en la función de puntuación empírica de Vina que se aproxima a la afinidad de unión en kcal/mol.

Una vez finalizado el cribado virtual, se puede visualizar en 3D la proteína con la pose mejor puntuada para los 100 ligandos mejor clasificados seleccionados mediante el *docking* con AutoDock Vina clasificados por energía de enlace. Además, los resultados se pueden descargar con lo que se obtienen las estructuras y energías de las 3 poses mejor puntuadas para los 1500 compuestos mejor clasificados. En cuanto a la entrada de MTiOpenScreen, para realizar el cribado virtual es necesario proporcionar la estructura de la proteína y los ligandos, y la definición del sitio de unión. A continuación, se recoge el procedimiento seguido para obtener cada uno.

3.5.1. Preparación de las estructuras de las proteínas

La estructura de la proteína se puede proporcionar a MTiOpenScreen en formato PDB o MOL2. Si se utiliza en PDB, la estructura se preprocesa automáticamente. En cambio, si se quiere preparar la estructura manualmente, en formato MOL2 la estructura se usa sin realizar modificaciones, por lo que es la opción escogida en este caso. Para ello, se emplea USCF Chimera comenzando por eliminar los iones y ligandos de cada estructura no necesarios. Después, se preparan para el *docking* con la herramienta *Dock Prep*, que realiza las siguientes tareas:

- Eliminación de moléculas de disolvente, generalmente agua.
- Si hay ubicaciones alternativas, mantiene solo la ocupación más alta.
- Convierte los residuos modificados en residuos estándar para los que dispone de parámetros.
- Repara cadenas laterales truncadas.
- Añade hidrógenos generando estados de protonación a pH fisiológico.
- Asigna cargas parciales a los átomos.
- Escribe los archivos en formato MOL2.

⁵⁴ Labbé, C.M.; Rey, J.; Lagorce, D.; Vavruša, M.; Becot, J.; Sperandio, O.; Villoutreix, B.O.; Tufféry, P.; Miteva, M.A. MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic Acids Res.* **2015**, *43*, W448–W454.

⁵⁵ Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.

3.5.2. Preparación de los ligandos

Para realizar el *docking*, MTiOpenScreen ofrece varias librerías de compuestos químicos disponibles para seleccionar los ligandos. También da la opción de usar una librería dada por el usuario que contenga hasta un máximo de 5000 moléculas orgánicas pequeñas, en cuyo caso se debe proporcionar una única conformación en 3D con todos los hidrógenos en un archivo en formato MOL2 o SDF. Por lo que, en este caso, se ha creado un fichero con todos los ligandos en formato MOL2. Para ello, una vez descargadas las estructuras 3D de los fenoles del AOVE de PubChem, hay que protonarlos a pH fisiológico. Debido a la gran cantidad de compuestos, se ha utilizado el archivo de comandos (script) de UCSF Chimera con código Python de la Figura 16, que ha permitido automatizar el proceso de preparación (añadir hidrógenos, añadir cargas y guardar los archivos en MOL2). Después, se han juntado los archivos generados en un único fichero para emplearlo en el cribado virtual.

```
import os
from chimera import runCommand as rc # use 'rc' as shorthand for runCommand
from chimera import openModels
from AddCharge import estimateNetCharge

# gather the names of .pdb files in the folder
file_names = [fn for fn in os.listdir(".") if fn.endswith(".sdf")]

# loop through the files, opening, processing, and closing each in turn
for fn in file_names:
    rc("open " + fn)
    molecule = openModels.list()[0]
    net_charge = estimateNetCharge(molecule.atoms)
    rc("delete H")
    rc("addh")
    rc("addcharge nonstd #0 " + str(net_charge) + " method gas")
    # save
    new_name = "prot_chimera" + fn[: -3] + ".mol2"
    rc("write format mol2 0 " + new_name)
    rc("close all")
# uncommenting the line below will cause chimera to exit when the script is done
# rc("stop now")
```

Figura 16. Script de Chimera con código Python para preparar los ficheros de los fenoles para el *docking*

3.5.3. Definición del sitio de unión de la proteína

La definición del sitio de unión se puede proporcionar mediante los parámetros de la caja (centro y dimensiones en el espacio) o se pueden calcular automáticamente basándose en la lista de los residuos de la proteína del sitio de unión dada por el usuario. En el presente trabajo se ha seguido la primera opción, de manera que se selecciona todo lo que se encuentre a 6 Å del centro geométrico de la cavidad en la que se localiza el ligando. Estos valores se calculan automáticamente empleando el script de Python mostrado en la Figura 17, junto con las coordenadas del ligando en la estructura cristalina de la proteína en un fichero en formato PDB.

```

import sys, os
##Usage python get_center_size.py pdb_file.pdb

def get_coordinates(p):
    reader=open(p, 'r')
    coordinates=[[], [], []]
    for line in reader:
        if 'ATOM' in line or 'HETATM' in line:
            line=line.split()
            coordinates[0].append(float(line[5]))
            coordinates[1].append(float(line[6]))
            coordinates[2].append(float(line[7]))
    reader.close()
    return coordinates

def get_box_center_and_box_size(coords):
    ##Returns the cavity##
    xs = coords[0]
    ys = coords[1]
    zs = coords[2]
    bc = [(min(xs)+max(xs))/2, ((min(ys))+max(ys))/2, ((min(zs))+max(zs))/2]
    bs = [(max(xs)+6)-(min(xs)-6), ((max(ys)+6)-(min(ys)-6)), ((max(zs)+6)-(min(zs)-6))]
    return bs, bc

pdb=sys.argv[1]
coordinates=get_coordinates(pdb)
grid_size,grid_center=get_box_center_and_box_size(coordinates)
with open('./grid_center_and_size.dat', 'w') as fw:
    fw.write('Grid center: ' + str(grid_center)+'\n')
    fw.write('Grid size: ' + str(grid_size)+'\n')
fw.close()

```

Figura 17. Script de Python para calcular la caja de 6 Å para el *docking*

3.6. Interacciones ligando–proteína

El recurso PLIP (Protein-Ligand Interaction Profiler) es una herramienta para analizar y visualizar las interacciones no covalentes en complejos proteína–ligando.⁵⁶ Se puede emplear a través de una entrada PDB existente o con un archivo en formato PDB sin preparación previa de la estructura, a diferencia de otras herramientas. Tras analizar el complejo, se da como resultado una lista con todas las interacciones no covalentes detectadas (enlaces de hidrógeno, enlaces de halógeno, interacciones hidrófobas, π -stacking, π -catión, puentes de sal, puentes de agua, y complejos metálicos) entre proteínas y moléculas pequeñas, como son iones, polímeros y cadenas de ADN/ARN en detalle a nivel de átomo.

PLIP es de libre acceso desde projects.biotec.tu-dresden.de/plip-web/plip/index y su código fuente Python está disponible como software de código abierto para realizar los cálculos de forma local. Esta última opción presenta algunas ventajas, ya que, además de las mismas características que el servidor web, ofrece configuraciones avanzadas para los archivos de salida y los parámetros de detección (*thresholds*). En este trabajo se ha empleado PLIP localmente en el modo de línea de comandos (versión 1.3.4) para realizar procesamientos por lotes y cambiar parámetros de detección cuando ha sido necesario. Así, esta herramienta ha permitido determinar los sitios de unión y las interacciones de las estructuras PDB de las proteínas, así como evaluar los resultados del *docking* para filtrar las poses de alta puntuación a partir del análisis de los patrones de interacción en el complejo de la estructura cristalina de la proteína y el complejo con las poses del *docking* identificando poses incorrectas de los ligandos.

⁵⁶ Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucl. Acids Res.* **2015**, *43*, W443–W447.

Por último, mencionar que para analizar las interacciones con PLIP de los ligandos generados en el cribado virtual con las proteínas, hay que preparar previamente los archivos en formato PDB de cada pose de los ligandos con la proteína. Para intentar automatizar el proceso, se ha creado el script de Chimera con código Python de la Figura 18 para crear estos archivos empleando la herramienta *ViewDock*, que facilita el análisis interactivo de los resultados del docking proteína-ligando. Una lista de ligandos tras el *docking* se abre en una nueva ventana que muestra los nombres de cada compuesto y la energía de enlace predicha, entre otra información.

```
import os
from chimera import runCommand as rc # use 'rc' as shorthand for runCommand

rc("viewdock output.pdbqt AutoDock")
rc("open 3bwy.pdb")
rc("delete H")
rc("addh")

for i in range(1, 91):
    n = str(i)
    rc("combine #1|#0." + n + " name a." + n)
    # save
    newname = "output_" + n + ".pdb"
    rc("write format pdb 2 " + newname)
    rc("close 2")
# uncommenting the line below will cause Chimera to exit when the script is done
# rc("stop now")
```

Figura 18. Script de Chimera Python para preparar los archivos ligando-proteína

3.7. Análisis de las propiedades ADMET

Para evaluar *in silico* las propiedades farmacocinéticas y de toxicidad ADMET de los compuestos seleccionados como *hits* se ha utilizado el servidor web de libre acceso pkCSM (<http://biosig.unimelb.edu.au/pkcsml/>). Se trata de una plataforma de *machine learning* para predecir y optimizar propiedades farmacocinéticas y de toxicidad de moléculas pequeñas, que utiliza el concepto de firmas estructurales basadas en gráficos, construyendo 30 predictores divididos en las cinco clases principales: absorción (siete predictores), distribución (cuatro predictores), metabolismo (siete predictores), excreción (dos predictores) y toxicidad (diez predictores).⁵⁷ En la Figura 19 se muestra el flujo de trabajo de pkCSM, de manera que a partir de las moléculas de partida, se calculan y combinan dos conjuntos principales de descriptores para utilizarlos en el siguiente paso de *machine learning*: propiedades generales de la molécula y una firma gráfica basada en la distancia.

⁵⁷ Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072.

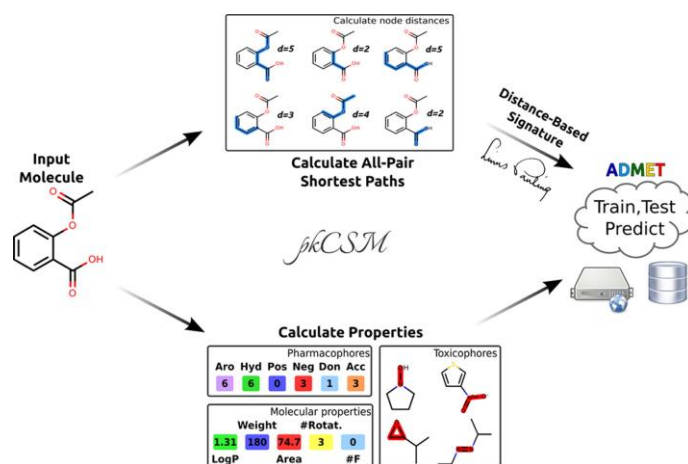


Figura 19. Flujo de trabajo de la plataforma pkCSM

3.8. Búsqueda de moléculas comerciales similares a los fenoles

A partir de los *hits* en los compuestos fenólicos, se ha realizado una búsqueda de compuestos comerciales similares empleando la herramienta web SwissSimilarity (<http://www.swiss similarity.ch/index.php>). Este recurso ofrece la posibilidad de llevar a cabo un cribado virtual basado en ligandos en varias librerías de productos químicos usando diferentes tecnologías.⁵⁸ Las librerías de moléculas pequeñas incluyen fármacos, compuestos bioactivos y comerciales, así como 205 millones de compuestos virtuales sintetizables a partir de reactivos disponibles comercialmente. En este caso se buscan compuestos comerciales, para lo que SwissSimilarity emplea ZINC, una base de datos gratuita de compuestos disponibles comercialmente para el cribado virtual, así como varios proveedores cuyos catálogos también están disponibles, que son Aldrich^{CPR}, Asinex, AsisChem, ChemBridge, ChemDiv, Enamine, InnovaPharm, Maybridge, Otava, Selleckchem, Sigma-Aldrich, SPECS, TimTec y Vitas, para un total de 9.7 millones de compuestos. En este trabajo, se han empleado todas estas bases de datos para realizar la búsqueda de compuestos comerciales.

La búsqueda en SwissSimilarity se ha realizado con el SMILES (Simplified Molecular Input Line Entry Specification) de la molécula a consulta. El SMILES es un lenguaje de notación química diseñado para un uso informático, que describe la conectividad y quiralidad de una molécula, de manera que las estructuras moleculares se especifican de forma única y precisa. Además, la conversión a la estructura 2D o 3D de una molécula se puede hacer con la mayoría de editores moleculares. En la Figura 20 se muestra como ejemplo el fenol simple tirosol [4-(2-hidroxi-etil)fenol].⁵⁹

⁵⁸ Zoete, V.; Daina, A.; Bovigny, C.; Michielin, O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *J. Chem. Inf. Model.* **2016**, *56*, 1399–1404.

⁵⁹ Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.

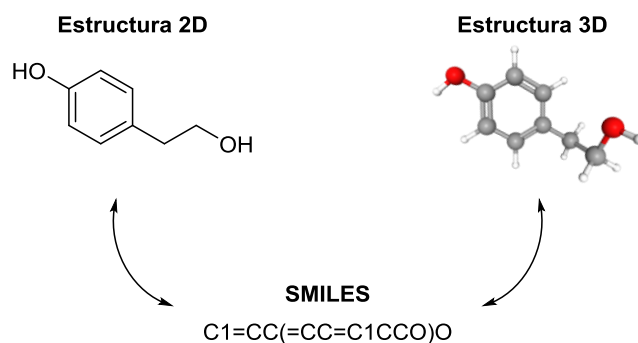


Figura 20. Interconversión del formato SMILES a las estructuras 2D y 3D del tirosol

Como resultado del análisis con SwissSimilarity se obtiene una lista con los compuestos más similares al de referencia, junto con la puntuación de similitud y un enlace a la entrada correspondiente a la base de datos de origen. La puntuación de similitud varía de 0 para moléculas totalmente diferentes a 1 para compuestos idénticos. Además, los resultados obtenidos se pueden descargar en distintos formatos, en concreto, el archivo CSV contiene los SMILES de todos los compuestos obtenidos, junto con el ID del compuesto y su puntuación de similitud.

Para convertir las cadenas SMILES en una estructura 3D, se ha empleado USCF Chimera, que realiza la conversión empleando el traductor SMILES proporcionado por el grupo del *National Cancer Institute CADD* o, si con esto no se produce una estructura, el servicio web *smi23d* del grupo de quimioinformática de la Universidad de Indiana. Este último implementa el mismo procedimiento que se describe para la base de datos Pub3D.⁶⁰ Para facilitar la conversión, se ha creado el script con código Python de la Figura 21, que también permite añadir los hidrógenos a pH fisiológico para cada molécula y el resultado se guarda en un fichero en formato MOL2.

```
import os
from chimera import runCommand as rc # use 'rc' as shorthand for runCommand

# the name of the file with the smiles
file_name = open("smiles_similares_fenoles_no_comerciales.txt")

# loop through the rows, processing and closing each in turn
for fn in file_name:
    line = fn.split()
    rc("open smiles:" + line[2])
    rc("setattr m name " + line[0])
    rc("delete H")
    rc("addh")
    # save
    new_name = line[0] + ".mol2"
    rc("write format mol2 0 " + new_name)
    rc("close all")
# uncommenting the line below will cause Chimera to exit when the script is done
# rc("stop now")
```

Figura 21. Script de Chimera Python para convertir los SMILES en la estructura 3D

⁶⁰ Willighagen, E.L.; O'Boyle, N.M.; Gopalakrishnan, H.; Jiao, D.; Guha, R.; Steinbeck, C.; Wild, D.J. Userscripts for the Life Sciences. *BMC Bioinformatics* **2007**, *8*, article number 487.

4. Discusión de resultados

4.1. Estructuras 3D de DNMT1 y COMT

Las estructuras 3D de las proteínas COMT y DNMT1 se encuentran disponibles en la base de datos PDB a través de UniProtKB. Para la proteína DNMT1 humana (código Uniprot P26358), se ha seleccionado la estructura con código PDB 4WXX, ya que se trata de la estructura cristalina de la enzima con todos sus dominios estructurales (hDNMT1, residuos 351-1600) en complejo con una molécula de SAH, un potente inhibidor de esta proteína, como se ha comentado. Esta estructura ha sido resuelta mediante difracción de rayos X a una resolución de 2.62 Å.⁶¹ Cabe destacar que hay dos moléculas de hDNMT1₃₅₁₋₁₆₀₀ en cada unidad asimétrica, en la Figura 22 se muestra una de estas moléculas y el sitio de unión del ligando SAH.

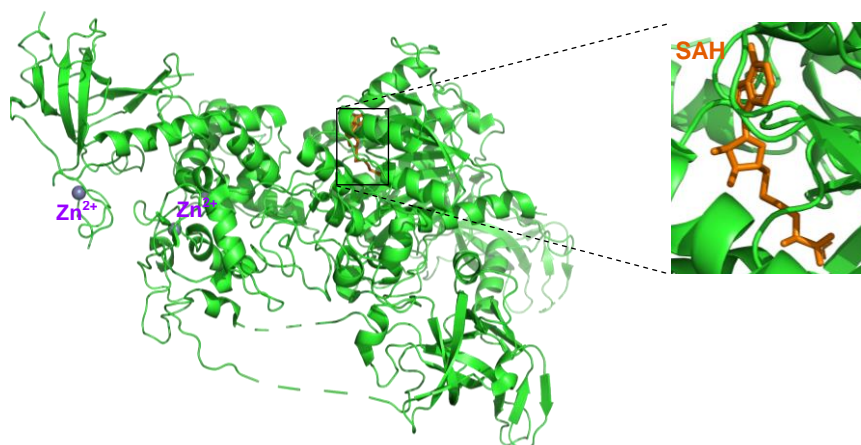


Figura 22. Estructura PDB 4WXX de la proteína DNMT1 con los cationes Zn^{2+} (morado) y la molécula SAH (naranja)

En cuanto a la proteína COMT humana (código Uniprot P21964), hay que tener en cuenta su polimorfismo Val^{108/158}Met, en el que se intercambia una valina por una metionina en la posición 158 de COMT unida a membrana (MB-COMT) y en la 108 de COMT soluble (S-COMT). Para este trabajo, se ha decidido emplear ambas isoformas de la proteína debido a los antecedentes reportados en literatura, para lo que se han escogido las estructuras con códigos PDB 3BWM y 3BWY de la forma S-COMT con el polimorfismo valina (108V) y metionina (108M), respectivamente. En ambas, la proteína está unida al cofactor SAM y al sustrato 3,5-dinitrocatecol (DNC), un potente inhibidor de la actividad enzimática de COMT. Están resueltas mediante difracción de rayos X con una resolución de 1.98 Å para 3BWM y 1.30 Å para 3BWY.³⁴

En general, ambas estructuras son muy similares, con un valor de RMSD (*Root Mean Square Deviation*) de 0.2 Å, y sus sitios activos son prácticamente idénticos, lo

⁶¹ Zhang, Z.-M.; Li, S.; Lin, K.; Luo, Y.; Perry, J.J.; Wang, Y.; Song, J. Crystal Structure of Human DNA Methyltransferase 1. *J. Mol. Biol.* **2015**, *427*, 2520–2531.

que está de acuerdo con la observación de que el cofactor SAM estabiliza a COMT 108M. Sin embargo, hay desviaciones de empaquetamiento alrededor del residuo polimórfico, localizado a unos 16 Å del sitio de unión a SAM, debido a que en 108M la cadena lateral de la metionina está más empaquetada que la valina. Como consecuencia, tiene una mayor interacción con los residuos Ala22 y Arg78, y esto genera un desplazamiento de 0.7 Å en la estructura principal de la proteína que se propaga hacia el sitio de unión de SAM. Además, la estructura de COMT 108V (Figura 23) contiene un catión K^+ unido en un pequeño bolsillo definido por la hélice $\alpha 8$, la hoja $\beta 6$ y su bucle de conexión, debido a la presencia de este catión en el medio de cristalización. En cambio, la estructura de COMT 108M tiene una molécula de agua coordinada en su lugar, ya que el potasio no estuvo presente durante la cristalización.

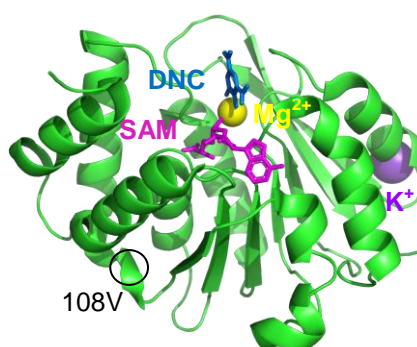


Figura 23. Estructura PDB 3BWM de la proteína COMT 108V con DNC (azul), SAM (rosa), K^+ (morado), Mg^{2+} (amarillo) y la cadena lateral del residuo V108 (circulo negro)

4.2. Interacciones proteína–ligando en las estructuras de DNMT1 y COMT

Se ha analizado el modo de unión de las proteínas en el cristal con sus respectivos inhibidores, SAH para DNMT1 y DNC para COMT. Para la estructura 4WXX de DNMT1, PLIP ha detectado entre la proteína y la molécula de SAH una interacción de tipo π -stacking con el residuo Phe1145 y varios enlaces de hidrógeno con los residuos Ser1146, Cys1148, Gly1150, Leu1151, Met1169, Trp1170, Asp1190, Cys1191 y Val1580. En la Figura 24 se muestran estas interacciones gráficamente.

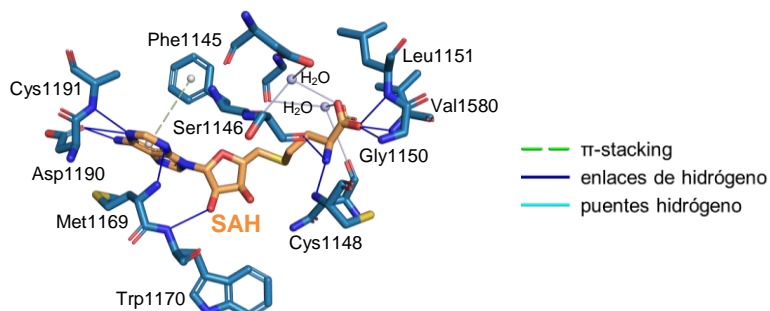


Figura 24. Interacciones en la estructura 4WXX generadas por PLIP entre la proteína DNMT1 (azul) y la molécula de SAH (naranja)

En las estructuras de la proteína COMT, la molécula de SAM se enlaza a la cadena polipeptídica mediante interacciones hidrofóbicas, enlaces de hidrógeno, interacciones π -stacking y puentes salinos, como se muestra en la imagen **A** de la Figura 25. Sin embargo, las interacciones más interesantes para el presente estudio son las que mantiene la molécula DNC. Para la estructura 3BWM de COMT 108V, este derivado del catecol se une a la proteína a través de interacciones hidrofóbicas con los residuos Trp38 y Pro174, y enlaces de hidrógeno con Lys46, Lys144 y Asn170. Además, se enlaza al catión Mg^{2+} mediante sus dos grupos hidroxilo. Este catión también se une a la cadena proteica, en concreto, a los residuos Asp141 y Asp169, y a una molécula de agua. Sin embargo, PLIP no detecta la interacción del magnesio con el residuo Asn170 que completa su entorno octaédrico, tal como se indica en el artículo bibliográfico en el que se describen estas estructuras de COMT.³⁴ Estas interacciones se recogen en la imagen **B** de la Figura 25. Finalmente, las interacciones en el cristal 3BWY para COMT 108V con DNC son análogas a las descritas para COMT 108M, a excepción de que no presenta el enlace de hidrógeno con el residuo Lys46.

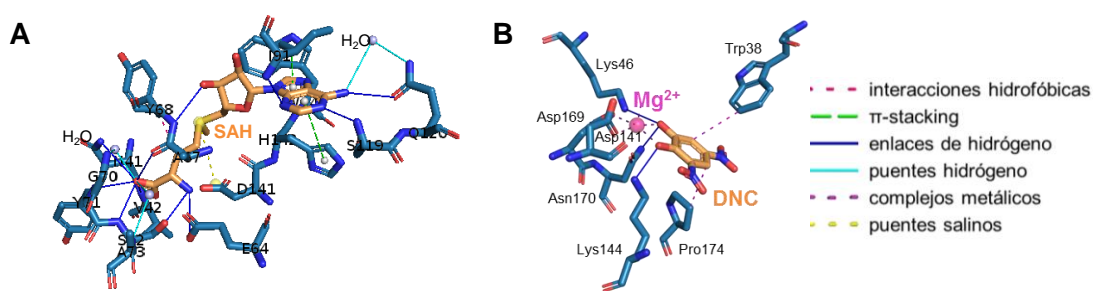


Figura 25. Interacciones en el cristal 3BWM por PLIP: en la imagen **A** entre la proteína COMT 108V y SAM, y en la **B** entre el ligando DNC, Mg^{2+} y la proteína

4.3. Evaluación de inhibidores potenciales de DNMT1 y COMT en los fenoles del AOVE

Para determinar la capacidad de los fenoles del AOVE de inhibir las proteínas DNMT1 y COMT se ha realizado el cribado virtual de estos compuestos. En concreto, se buscaba evaluar si la interacción de estos compuestos fenólicos tendrá lugar preferentemente en el sitio de unión catalítica, es decir, del cofactor SAM en DNMT1 y del sustrato catecol en COMT. Sin embargo, previamente es recomendable realizar el *docking* entre las proteínas y sus ligandos en el cristal, para analizar lo reproducibles que son los cálculos y esto también permite tener un control de los resultados.

4.3.1. Cribado virtual de las proteínas con sus inhibidores

Para realizar el cribado virtual con el servidor MTiOpenScreen, se ha seguido el procedimiento explicado en la sección 3.5.3 del capítulo Materiales y métodos. Por lo que se han preparado los ficheros con las proteínas y los ligandos, y se ha definido el

sitio de unión de los ligandos seleccionando todo lo que se encuentre a 6 Å del centro geométrico de la cavidad en la que está el ligando. Estos son los resultados del sitio de unión (centro y dimensiones en el espacio de la caja) para cada estructura:

- Para la estructura 4WXX de DNMT1, las coordenadas del centro del sitio de unión del ligando SAH son en el eje x -46.693, en el eje y 61.339 y en el eje z 6.563, y las dimensiones son 19 x 24 x 20 Å.
- Para la estructura 3BWM de COMT 108V, las coordenadas del sitio de unión del ligando DNC son en el eje x -4.016, en el eje y -12.232 y en el eje z -17.591, y las dimensiones son 14 x 19 x 17 Å.
- Para la estructura 3BWY de COMT 108M, las coordenadas del sitio de unión del ligando DNC son en el eje x -3.847, en el eje y -12.281 y en el eje z -17.657, y las dimensiones son 14 x 19 x 17 Å.

A continuación, se ha realizado el cribado virtual entre cada proteína y su ligando con el servidor MTiOpenScreen. Como resultado para la molécula SAH con la proteína DNMT1 se ha obtenido una energía de enlace predicha de -8.4 kcal/mol y se han analizado las interacciones con PLIP de la mejor pose del ligando con la proteína. Se ha observado que dichas interacciones se mantienen prácticamente igual respecto a las determinadas en el cristal, al unirse correctamente al sitio de unión de este ligando. Además, se puede comprobar visualmente que la molécula de SAH generada se solapa bastante bien con la del cristal, tal como se muestra en la Figura 26.

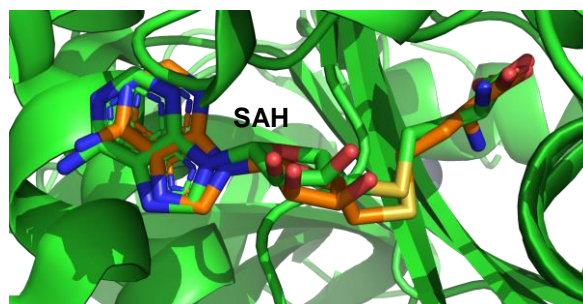


Figura 26. Superposición con PyMOL de la molécula SAH obtenida en el *docking* (naranja) con la molécula SAH del cristal 4WXX de la proteína DNMT1 (verde)

Por otra parte, tras el cribado virtual entre las dos estructuras de la proteína COMT y la molécula DNC se ha obtenido una energía de enlace predicha de -5.5 kcal/mol para COMT 108V y -5.6 kcal/mol para COMT 108M. En la Figura 27 se superponen la estructura del cristal 3BWM de COMT 108V y el ligando generado tras el *docking*. El resultado de esta superposición es similar para la estructura de COMT 108M, por lo que en ambos casos la molécula de DNC se asemeja bastante al ligando en el cristal. Sin embargo, no son tan similares como ocurría anteriormente con el ligando SAH en DNMT1, debido a la presencia del magnesio en la estructura de la proteína. Por otra parte, mientras que todas las interacciones proteína-DNC se mantienen respecto a las del cristal, las interacciones de este ligando con el catión magnesio no han sido detectadas por el programa PLIP.

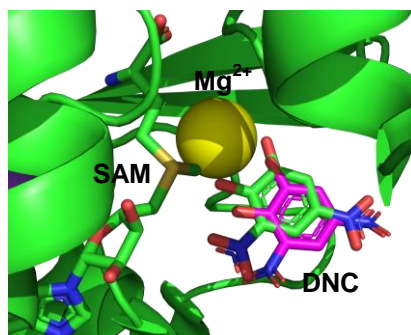


Figura 27. Superposición con PyMOL de la molécula de DNC obtenida en el *docking* (rosa) con el DNC del cristal de la COMT (verde) y en amarillo se representa el Mg^{2+}

4.3.2. Cribado virtual de los fenoles del AOVE con DNMT1 y COMT

El siguiente paso es realizar los estudios computacionales usando los fenoles del AOVE y las proteínas DNMT1 y COMT para analizar su posible unión *in silico* mediante el cribado virtual. Previamente, hay que preparar el fichero con los 129 fenoles del AOVE protonados a pH fisiológico, siguiendo el procedimiento indicado en la sección 3.5.2 a través del script de Chimera con código Python de la Figura 16, que ha permitido automatizar este proceso. La lista de fenoles se recoge en el Anexo I.

Como resultado del *docking*, para cada proteína se ha obtenido un listado de los fenoles según su energía de enlace para sus tres poses mejor puntuadas. Tras analizar los resultados de la estructura 4WXX para la proteína DNMT1, se han identificado un total de 12 *hits* entre los compuestos fenólicos del AOVE estudiados. Estos han mostrado capacidad para unirse a la proteína con poses ubicadas en las proximidades del sitio catalítico y energías de enlace entre -10.4 y -9.6 kcal/mol. Esta selección se recoge en la Tabla 1 y está basada tanto en la energía de enlace predicha al analizar los 25 primeros resultados, como en las interacciones dadas por PLIP que mantiene cada una de las tres poses de estos fenoles con la proteína respecto a las observadas en el cristal 4WXX.

Tabla 1. Compuestos fenólicos identificados como *hits* para la proteína DNMT1

Compuesto	PubChem CID	Mejor pose	Energía (kcal/mol)	nRot
Isorhoifolina	5377847	2	-10.4	14
Isoacteósido	6476333	1	-10.3	20
Luteolina 7-glucósido	5280637	1	-10.1	11
Apigenina-7-O-glucósido	12304093	1	-10.0	10
Hesperidina	10621	3	-10.0	15

Quercitrina	5280459	3	-9.9	12
Quercetina-7-O-glucósido	91820114	3	-9.8	11
<i>Hellicoside</i>	5281778	1	-9.8	23
Beta-hidroxi-acteósido	10009317	2	-9.7	21
Suspensaside	5281798	3	-9.7	21
Escolimosido	5282152	3	-9.6	15
Crisoeriol-7-O-glucósido	11294177	2	-9.6	11

De entre estos mejores fenoles se pueden destacar crisoeriol-7-O-glucósido, *hellicoside*, quercetina-7-O-glucósido e isoacteósido, ya que mantienen la mayoría de los enlaces de hidrógeno respecto a los determinados en el cristal, junto con la interacción π -stacking. Además, se solapan bastante bien con el ligando SAH del cristal de la proteína DNMT1, tal como se muestra en Figura 28 a modo de ejemplo para la molécula de quercetina-7-O-glucósido. En cuanto a los fenoles hesperidina y escolimosido, aunque no muestran la interacción π -stacking del cristal, mantienen 8 de los 9 enlaces de hidrógeno, por lo que se ha decidido incluirlos en esta primera selección. El resto de los fenoles de la lista mantienen la interacción π -stacking, pero el número de enlaces de hidrógeno respecto a los del cristal es menor a las de los anteriores fenoles, ya que solo mantienen entre 6 y 3 de estas interacciones.

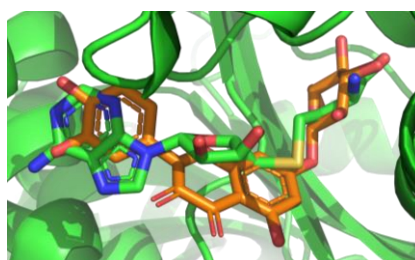


Figura 28. Superposición de la mejor pose de la quercetina-7-O-glucósido obtenida en el *docking* (naranja) con el ligando SAH en el cristal de DNMT1 (verde)

Cabe mencionar que ya se había descrito previamente en la literatura que el flavonoide luteolina 7-glucósido es capaz de inhibir la proliferación de células cancerosas y endoteliales en proliferación. Los autores de este estudio indican que estos resultados podrían indicar que se trata de un agente antiangiogénico, por comparación con la luteolina, otro flavonoide que muestra actividad antiangiogénica *in vitro* e *in vivo* al inhibir la señalización inducida por el factor de crecimiento endotelial vascular (VEGF) en las células endoteliales.⁶²

⁶² Goulas, V.; Exarchou, V.; Troganis, A.N.; Psomiadou, E.; Fotsis, T.; Briasoulis, E.; Gerothanassis, I.P. Phytochemicals in olive-leaf extracts and their antiproliferative activity

Para el cribado virtual de la proteína COMT se emplearon las estructuras 3BWM y 3BWY con el polimorfismo 108V y 108M, respectivamente. Se ha priorizado que visualmente los fenoles estén colocados con dos grupos hidroxilo hacia el catión magnesio para que la interacción entre ellos sea posible, igual que ocurre con el ligando DNC en el cristal, ya que PLIP no ha detectado estas interacciones. La selección se muestra en la Tabla 2 con un total de 9 *hits*, para lo que se ha tenido en cuenta los resultados de ambas estructuras, de manera que se revisaron los 25 mejores fenoles de cada una, analizando también las interacciones con la proteína. Esto ha llevado a descartar el fenol *orbanchoside*, que estaba entre los primeros de la estructura 3BWM y con interacciones muy similares a las del cristal, pero no se encontraba entre los mejores de 3BWY.

Tabla 2. Compuestos fenólicos identificados como *hits* para la proteína COMT. Los datos de mejor pose y energía de enlace predicha son de la estructura 3BWM de 108V

Compuesto	PubChem CID	Mejor pose	Energía (kcal/mol)	nRot
Delfinidina	128853	1	-7.9	0
Taxifolina	12309272	3	-7.4	0
Eriodictiol	90657147	3	-7.1	5
Isoacteósido	6476333	2	-7.5	20
Luteolina 7-glucósido	5280637	2	-7.3	11
Quercetina-7-O-glucósido	91820114	3	-7.1	11
<i>Suspensaside</i>	5281798	1	-7.2	21
Delfinidin-3-O-glucósido	443650	1	-7.2	13
Isoquercitrina	5280804	3	-6.7	11

De estos fenoles, aunque todos se colocan en las proximidades del sitio catalítico, los que muestran interacciones más similares a las de las estructuras 3BWM y 3BWY son taxifolina, eriodictiol, isoacteósido, quercetina-7-O-glucósido y *suspensaside*, ya que mantienen las dos interacciones hidrofóbicas y dos de los tres enlaces de hidrógeno del cristal 3BWM. En la Figura 29 se muestra la superposición del fenol isoacteósido, a modo de ejemplo, con la estructura 3BWM de COMT 108V. El resto de los fenoles seleccionados presentan una de las dos interacciones hidrofóbicas o solo uno de los enlaces de hidrógeno del cristal, aunque, como se ha comentado, todos podrían tener las dos interacciones de los grupos hidroxilo con el magnesio.

against cancer and endothelial cells. *Molecular Nutrition and Food Research* **2009**, 53, 600–608.

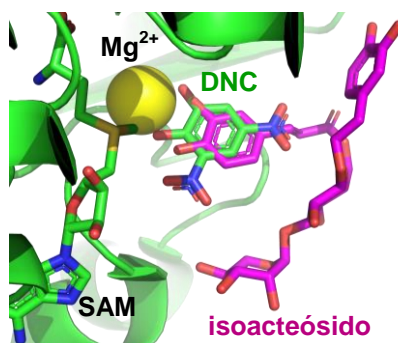


Figura 29. Superposición de la mejor pose del isoacteósido obtenida en el *docking* (rosa) con el ligando DNC (verde) en la estructura 3BWM de COMT 108V

Estudios previos indican que el flavonoide taxifolina muestra actividad antiangiogénica, ya que inhibe la formación *in vitro* de estructuras de células endoteliales, así como la formación *in vivo* de nuevos vasos sanguíneos en varios modelos.⁶³ En otro estudio realizado, se explica que la taxifolina suprime la angiogénesis, al menos en parte, al inhibir la autofosforilación de varios residuos de tirosina en VEGFR-2.⁶⁴ Según los resultados obtenidos en el presente trabajo, la taxifolina es un potencial inhibidor de la proteína COMT, por lo que esto también podría explicar su potente actividad antiangiogénica.

Con los resultados obtenidos para las proteínas DNMT1 y COMT se pueden determinar los fenoles del AOVE que podrían actuar como inhibidores duales. En total son 12 inhibidores de la enzima DNMT1 y 9 de COMT, de los cuales 4 tienen altas probabilidades de ser inhibidores duales, que son quercetina-7-O-glucósido, isoacteósido, luteolina 7-glucósido y *suspensaside*. Comentar de nuevo brevemente, que los inhibidores de la proteína COMT, aunque no se unan a DNMT1, también van a actuar de forma indirecta como inhibidores duales al aumentar la concentración de SAH en el medio, que es un inhibidor de DNMT1.

4.3.3. Análisis de las propiedades ADMET de los *hits* de DNMT1 y COMT

Una vez identificados los *hits* entre los compuestos fenólicos del AOVE para DNMT1 y COMT, se ha procedido a evaluar sus propiedades farmacocinéticas y de toxicidad ADMET usando el servidor web pkCSM, que proporciona el valor de 30 parámetros divididos en las cinco clases principales (absorción, distribución, metabolismo, excreción y toxicidad):

- **Absorción:**
 - Permeabilidad de Caco-2: La línea Caco-2 son células de adenocarcinoma colorrectal humano usada como modelo *in vitro* de la mucosa intestinal

⁶³ Haque, W.; Pattanayak, S.P.; Sinha, B.N. Evaluation of taxifolin and phloretin as antiangiogenic flavonoids: An *in vivo*, *in vitro* experimental analysis. *Int. J. Pharm. Pharm. Sci.* **2015**, *7*, 72–79.

⁶⁴ Lamy, S.; Ouanouki, A.; Beliveau, R.; Desrosiers, R.R. Olive oil compounds inhibit vascular endothelial growth factor receptor-2 phosphorylation. *Exp. Cell Res.* **2014**, *322*, 89–98.

para predecir la absorción de fármacos (vía oral). Se predice el logaritmo del coeficiente de permeabilidad aparente en log cm/s. Si el valor es superior a 0.90, el compuesto tiene alta permeabilidad de Caco-2.

- Solubilidad en agua: Da la solubilidad de la molécula en agua a 25 °C como el logaritmo de la concentración molar (log mol/L).
 - Absorción intestinal: El intestino es el sitio principal de absorción de un fármaco y se predice su porcentaje absorbido a través del intestino. Si una molécula tiene una absorción inferior al 30%, se absorbe mal.
 - Sustrato de la glicoproteína-P (Pgp): La Pgp actúa de barrera biológica al expulsar toxinas y xenobióticos de las células, así que no se desea que ocurra con un fármaco. Se predice si es probable que sea sustrato de Pgp.
 - Inhibidores de la glicoproteína-P I y II: Modular el transporte mediado por Pgp puede dar lugar a ventajas o contraindicaciones para un fármaco. Se indica si es probable que un compuesto sea un inhibidor de la Pgp I/II.
 - Permeabilidad cutánea: Un factor importante para la eficacia de muchos fármacos. Se predice si es probable que un compuesto sea permeable a la piel, dado como la constante de permeabilidad de la piel logKp (cm/h).
- **Distribución:**
 - Volumen de distribución en estado estacionario (VDss): Volumen teórico que se necesita distribuir de la dosis total de un fármaco para dar la misma concentración que en el plasma sanguíneo. Cuanto mayor es, más fármaco se distribuye en los tejidos que en el plasma. El logaritmo predicho de VDss se expresa como log L/kg, y es bajo si su valor es inferior a 0.71 L/kg (log VDss < -0.15) y alto si está por encima de 2.81 (log VDss > 0.45).
 - Fracción libre: La mayoría de fármacos en el plasma están en equilibrio entre un estado libre o unido a proteínas séricas. Cuanto más se una a proteínas, menos eficientemente podrá atravesar las membranas celulares o difundirse. Se predice la fracción libre prevista en el plasma.
 - Permeabilidad BBB: El cerebro está protegido de compuestos exógenos por la barrera hematoencefálica (BBB). La capacidad de un fármaco de cruzarla es importante para reducir los efectos secundarios y toxicidad o para mejorar la eficacia de los fármacos con actividad farmacológica dentro del cerebro. Si logBBB (relación logarítmica de la concentración en cerebro y plasma) es menor de -1 se distribuye en el cerebro de manera deficiente.
 - Permeabilidad del sistema nervioso central (SNC): Medir la permeabilidad hematoencefálica puede ser difícil. Una medida más directa es el producto permeabilidad hematoencefálica-área superficial (logPS). Si logPS > -2 penetra en el SNC, pero si logPS < -3 se considera incapaz de penetrar.
 - **Metabolismo:**
 - Inhibidores del citocromo P450: Esta es una enzima de desintoxicación, localizada principalmente en el hígado, que desactiva fármacos y activa otros. Sus inhibidores pueden afectar al metabolismo del fármaco, así que es importante evaluar la capacidad para inhibirlo. Un compuesto la inhibe, si la concentración para dar una inhibición del 50% es inferior a 10 μM.

- Sustrato CYP2D6/CYP3A4: El citocromo P450 es responsable del metabolismo de muchos fármacos, sobre todo las isoformas 2D6 y 3A4, pero sus inhibidores pueden alterar la farmacocinética de estos fármacos, así que es importante saber si un compuesto es sustrato de esta enzima. Se evalúa si es probable que una molécula sea metabolizada por P450.
- **Excreción:**
 - Sustrato de OCT2 renal: El transportador catiónico orgánico 2 (OCT2) es un transportador de captación renal con un papel crucial en la eliminación renal de fármacos y compuestos endógenos. Sus sustratos pueden actuar en interacciones adversas con inhibidores de OCT2 coadministrados. Se evalúa si es probable que una molécula sea un sustrato de OCT2.
 - Eliminación total: Una combinación de la eliminación hepática y renal. Se relaciona con la biodisponibilidad y es importante para determinar las tasas de dosificación para lograr un estado estable de concentración. Se predice el logaritmo de la eliminación total prevista (CL_{tot}) en log (mL/min/kg).
- **Toxicidad:**
 - LD50 en ratas: Los valores de dosis letales (LD) son una medida estándar para evaluar la toxicidad relativa. LD50 es la cantidad de compuesto que causa la muerte del 50% de ratas. Se predice este valor en mol/kg.
 - Toxicidad AMES: La prueba de Ames evalúa el potencial mutagénico de un compuesto en bacterias. Si es positiva, el compuesto es mutagénico. Se predice si es probable que un compuesto sea positivo para Ames.
 - Toxicidad por *T. Pyriformis*: Esta es una bacteria protozoaria y su toxicidad se suele usar como criterio de valoración. Se predice el pIGC50 (logaritmo negativo de la concentración necesaria para inhibir un crecimiento del 50% en log µg/L). Si el valor es superior a -0.5 log µg/L se considera tóxico.
 - Toxicidad en *Minnow*: Se predice el logaritmo de CL50 en *F. Minnows*. Un valor inferior a 0.5 mM (log CL50 < -0.3) se considera alta toxicidad.
 - Dosis máxima tolerada recomendada (MRTD): Da una estimación del umbral de dosis tóxica de sustancias químicas en humanos. Se predice el logaritmo de MRTD (log mg/kg/día) que ayuda a saber la dosis inicial máxima recomendada para un fármaco en la fase I de los ensayos clínicos. Si es menor o igual a 0.477 se considera bajo y si es mayor de 0.477 alto.
 - Toxicidad crónica oral en ratas: Es importante tener en cuenta la exposición de químicos a dosis bajas o moderadas durante largos períodos de tiempo. Los estudios crónicos identifican la dosis más baja de un compuesto que provoca un efecto adverso (LOAEL) y la dosis más alta en la que no se observan estos efectos (NOAEL). Se predice el logaritmo de LOAEL, log(mg/kg_bw/día), pero los resultados se interpretan en función de la concentración bioactiva y la duración del tratamiento.
 - Hepatotoxicidad: La lesión hepática inducida por fármacos es un problema importante de seguridad. Se predice si es probable que un compuesto esté asociado con una función alterada del hígado.

- Sensibilización cutánea: Efecto adverso potencial de productos aplicados por vía dérmica. Se predice si es probable que un compuesto esté asociado con la sensibilización de la piel.
- Inhibidores de hERG I y II: La inhibición de los canales de potasio codificados por hERG es la principal causa de desarrollo del síndrome QT largo adquirido, que conduce a una arritmia ventricular mortal. Se determina si es probable que un compuesto sea un inhibidor de hERG I/II.

La administración de los fenoles sería por vía oral, por lo que no se tendrán en cuenta los parámetros referidos a la administración cutánea. Así, para este estudio, se tendrán en cuenta los siguientes requisitos de los parámetros analizados por pkCSM:

1. Permeabilidad de Caco-2: valor superior a 0.90.
2. Absorción intestinal: valor superior al 30%.
3. Sustrato de la Pgp: no sustrato de Pgp.
4. Permeabilidad BBB: no cruce la barrera, así que valores inferiores a -1.
5. Permeabilidad SNC: valores inferiores a -3.
6. Toxicidad AMES: no positivo para AMES (compuestos no mutagénicos).
7. Toxicidad por *T. Pyriformis*: valores inferiores a -0.5 log $\mu\text{g/L}$.
8. Toxicidad en *Minnow*: valores superiores a 0.5 mM (log CL50 > -0.3).
9. Hepatotoxicidad: no asociado a una función alterada del hígado.
10. Inhibidores de hERG I y II: compuestos no inhibidores de hERG I/II.

Los resultados de estos parámetros para los *hits* de los fenoles del AOVE se recogen en el Anexo II. De estos compuestos analizados, pkCSM ha predicho que todos son no mutagénicos (no positivos a la prueba AMES), son inhibidores de hERG I y la mayoría también lo son de hERG II, no son hepatotóxicos y no tienen toxicidad en *Minnow*. Sin embargo, a excepción de taxifolina, estos fenoles no muestran permeabilidad de Caco-2, un parámetro muy importante para la absorción de los compuestos. Además, todos son sustrato de la Pgp y tienen toxicidad por *T. Pyriformis*. En cuanto al resto de los parámetros examinados, hay fenoles que los cumplen y otros no, pero como mínimo cumplen todos y como máximo no cumplen dos, por lo que los resultados son bastante similares en todos estos compuestos estudiados.

4.3.4. Selección de los fenoles con más potencial para actuar como inhibidores

Debido a que en los resultados de ADMET para los fenoles seleccionados como posibles inhibidores de las proteínas DNMT1 y COMT no se han observado diferencias significativas, se ha decidido no descartar ningún fenol en función de sus propiedades farmacocinéticas y de toxicidad. Como parece que se han encontrado buenos *hits*, el siguiente paso para conseguir compuestos *leads* sería validarlos experimentalmente y optimizar las moléculas para resolver estos problemas de sus propiedades ADMET.

Entonces de estos fenoles se han seleccionado los mejores únicamente según sus interacciones con las proteínas, es decir, que los inhibidores de DNMT1 mantengan la interacción π -stacking y 7 o 8 enlaces de hidrógeno respecto al cristal, ya que aunque todos los fenoles tienen anillos aromáticos, no todos forman interacciones π -stacking con el residuo Phe1145; mientras que los de COMT, además de tener dos grupos hidroxilo próximos al magnesio, mantengan las dos interacciones hidrofóbicas y dos de los tres enlaces de hidrógeno del cristal 3BWM de COMT 108V. Con estos criterios, se obtiene una lista con dos inhibidores potenciales de la proteína DNMT1 (crisoeriol-7-O-glucósido y *hellicoside*), dos de COMT (taxifolina y eriodictiol) y tres duales (*suspensaside*, quercetina-7-O-glucósido e isoacteosido). Cabe destacar que todos estos compuestos pertenecen únicamente a dos familias de fenoles, flavonoides y glucósidos, tal como se muestra en la Figura 30. Los dos inhibidores de COMT son glucósidos, los tres duales son flavonoides y de los dos de DNMT1 cada uno pertenece a una de estas familias. Esto parece indicar que estas dos familias de fenoles tienen un mayor potencial como agentes antiangiogénicos respecto al resto.

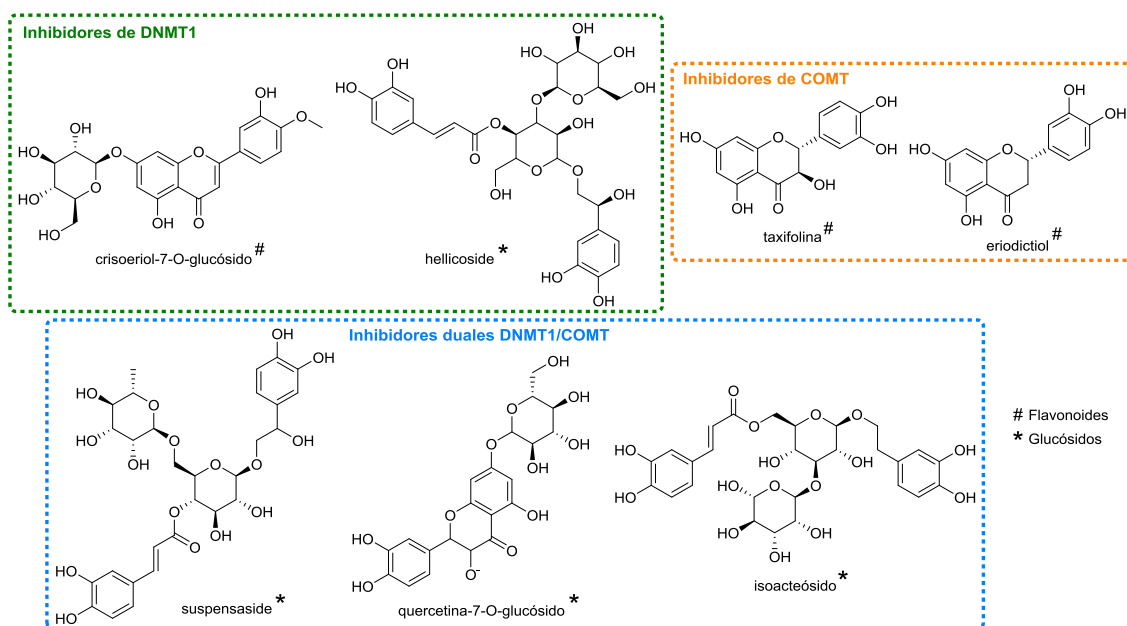


Figura 30. Compuestos fenólicos del AOVE con potencial como inhibidores

4.4. Estructuras 3D de DNMT3A y DNMT3B

Una vez seleccionados los fenoles que se consideran *hits* para las proteínas DNMT1 y COMT, se ha realizado el cribado virtual con las enzimas ADN metiltransferasas *de novo* DNMT3A y DNMT3B. Para ello, se ha seguido un procedimiento análogo al descrito previamente, así que, en primer lugar, hay que conocer las estructuras 3D disponibles de estas proteínas y seleccionar la más adecuada para cada una. Dado que estas proteínas se unen a través de su dominio catalítico a la proteína inactiva DNMT3L (código Uniprot Q9UJW3) para promover la

metilación, las estructuras 3D interesantes para el presente estudio son en las que están formando complejos con esta proteína.

Además, mencionar la mutación R882H de DNMT3A, ya que representa la mutación más frecuente de la leucemia mieloide aguda (LMA), detectada aproximadamente en el 25% de los pacientes con esta enfermedad. Esta mutación da lugar a una hipometilación del ADN y un aumento de la proliferación celular, ya que altera la actividad enzimática de DNMT3A y desestabilizada su forma tetramérica *in vitro* y en las células. Por lo que la mutación DNMT3A R882H puede contribuir a la patogénesis de la LMA.⁶⁵ Debido a esto, se ha decidido emplear también la estructura 3D de la proteína DNMT3A con esta mutación.

Para la proteína DNMT3A (código Uniprot Q9Y6K1) *wild-type* se ha seleccionado la estructura con código PDB 6W8B resuelta mediante difracción de rayos X a una resolución de 2.40 Å (Figura 31). Para la proteína DNMT3A mutada R882H se ha empleado la estructura PDB 6W89 resuelta también mediante difracción de rayos X con una resolución de 2.50 Å. Ambas estructuras análogas son la estructura cristalina del complejo DNMT3A–DNMT3L unido a ADN con el motivo CGA y a la molécula SAH. Como se ha comentado, la subunidad estructural de la proteína DNMT3A en complejo con DNMT3L es un homodímero de DNMT3A mediado por interacciones polares y unido a dos subunidades de DNMT3L a través de contactos hidrofóbicos (DNMT3L–DNMT3A–DNMT3A–DNMT3L). Además, la estructura 6W89 de DNMT3A R882H también contiene varias moléculas de ácido cítrico, debido a las condiciones de cristalización.⁶⁶

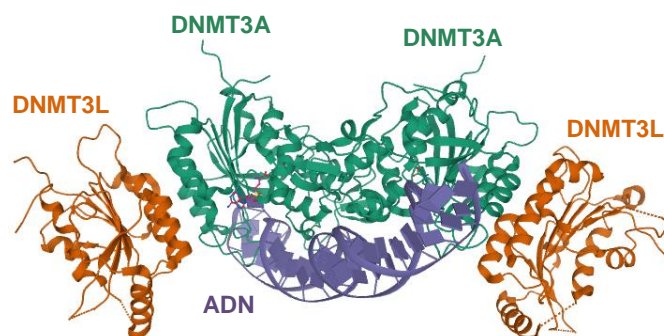


Figura 31. Estructura PDB 6W8B. Se representa en verde el dímero de DNMT3A (*wild-type*), en naranja las subunidades DNMT3L y en azul la cadena de ADN

Por último, para la proteína DNMT3B (código Uniprot Q9UBC3) se ha escogido la estructura con código PDB 6KDA resuelta mediante difracción de rayos X con una resolución de 2.91 Å. Se trata de la estructura cristalina del complejo heterotetramérico

⁶⁵ Ley, T.J.; Ding, L.; Walter, M.J.; McLellan, M.D.; Lamprecht, T.; Larson, D.E.; Kandoth, C.; Payton, J.E.; Baty, J.; Welch, J.; Harris, C.C.; Lichti, C.F. et al. DNMT3A Mutations in Acute Myeloid Leukemia. *N. Engl. J. Med.* **2010**, *363*, 2424–2433.

⁶⁶ Anteneh, H.; Fang, J.; Song, J. Structural basis for impairment of DNA methylation by the DNMT3A R882H mutation. *Nat. Commun.* **2020**, *11*, 2294.

DNMT3B–DNMT3L unido a una cadena de ADN y a la molécula SAH, similar a las estructuras cristalinas de DNMT3A, tal como se puede observar en la Figura 32.⁶⁷

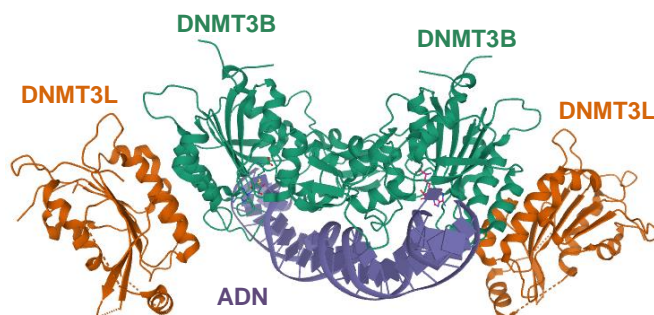


Figura 32. Estructura PDB 6KDA. Se representa en verde el dímero de la proteína DNMT3B, en naranja las subunidades DNMT3L y en azul la cadena de ADN

4.5. Interacciones proteína–ligando en las estructuras de DNMT3

A partir de las estructuras 3D seleccionadas para las enzimas DNMT3, se ha analizado el modo de unión de las proteínas con su inhibidor SAH en el cristal. Para la estructura 6W8B de DNMT3A *wild-type* con PLIP se han detectado una interacción de tipo π -stacking con el residuo Phe640 y enlaces de hidrógeno con otros ocho residuos (Ile643, Thr645, Val665, Asp686, Val687, Arg688, Arg891 y Trp893). La segunda estructura seleccionada para esta proteína con la mutación R882H, 6W89, muestra las mismas interacciones, a excepción del enlace de hidrógeno con el residuo Arg688. Finalmente, para la estructura 6KDA de DNMT3B se ha determinado también una interacción π -stacking de SAH con una fenilalanina de la proteína, pero que se encuentra en la posición 581. Además, de varios enlaces de hidrógeno con los residuos Gly583, Ile584, Thr586, Val606, Cys607, Asp627, Val628, Arg629 y 834Trp. En la Figura 33 se muestran gráficamente las interacciones para las estructuras cristalinas 6W8B (imagen A) y 6KDA (imagen B) generadas con PLIP.

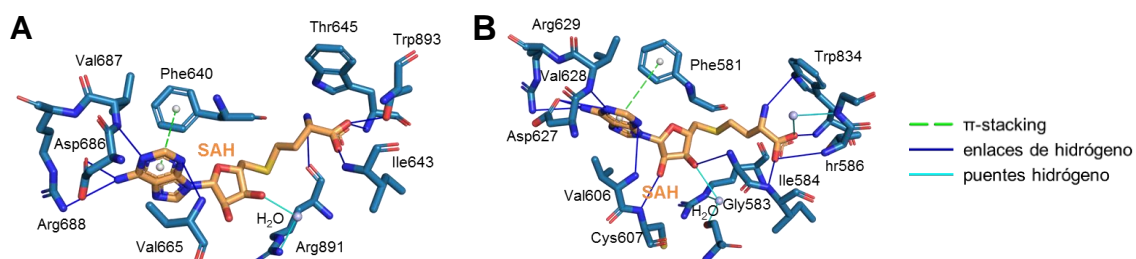


Figura 33. Interacciones entre la proteína (azul) y SAH (naranja) por PLIP: en la imagen **A** en el cristal 6W8B de DNMT3A y en la **B** en el cristal 6KDA de DNMT3B

⁶⁷ Lin, C.-C.; Chen, Y.-P.; Yang, W.-Z.; Shen, J.C.K.; Yuan, H.S. Structural insights into CpG-specific DNA methylation by human DNA methyltransferase 3B. *Nucleic Acids Res.* **2020**, *48*, 3949–3961.

4.6. Conservación modos de unión de las proteínas DNMTs

Se realizó un estudio sobre la conservación de los modos de unión de las proteínas DNMTs empleando distintas estrategias. Por un lado, se ha realizado una consulta bibliográfica para conocer que hay descrito en la literatura sobre este tema. Esta información se ha recogido en la sección 2.1.1.2 del capítulo Antecedentes, pero en resumen las ADN metiltransferasas activas, DNMT1, DNMT3A y DNMT3B, constan de una región multidominio N-terminal con funciones reguladoras y un dominio metiltransferasa C-terminal. El dominio catalítico está conservado en las tres enzimas, denominado pliegue de Rossmann, implicado en la unión del cofactor y en la catálisis. En cambio, el dominio de reconocimiento de diana no tiene similitud de secuencia entre DNMT1 y DNMT3, aunque participa en el reconocimiento y la especificidad del ADN. También las regiones N-terminales reguladoras son totalmente diferentes en secuencia y estructura en DNMT1 y DNMT3.²²

Además, para comprobar la conservación de los sitios de unión, se ha realizado la superposición de las tres proteínas y que se puedan comparar. Para ello, se ha empleado la opción *MatchMaker* de UCSF Chimera, que superpone estructuras por parejas alineando primero sus secuencias y luego ajustando los carbonos α de los residuos en las mismas columnas de la alineación de las secuencias. Por lo general, el ajuste se itera para que los pares de residuos alineados en secuencia pero muy separados en el espacio no se utilicen en la coincidencia 3D final. El resultado gráfico con las estructuras superpuestas se recoge en la Figura 34, lo que permite identificar que la superposición ha tenido lugar de forma adecuada y que la región más similar para las tres proteínas es alrededor del sitio de unión del ligando. Tras el alineamiento también se da el número de pares de C_α y el valor de RMSD en la iteración final de cada ajuste por pares. En el alineamiento de DNMT1 con DNMT3A se ha realizado en 264 pares de C_α y RMSD es igual a 16.71 Å, y con DNMT3B en 263 pares y RMSD es 17.22 Å. En cambio, el alineamiento de DNMT3A y DNMT3B ha tenido lugar en los 284 pares con un RMSD de 0.48 Å, un valor mucho más pequeño que para DNMT1 indicativo de un mejor alineamiento y de mayor similitud de secuencias.

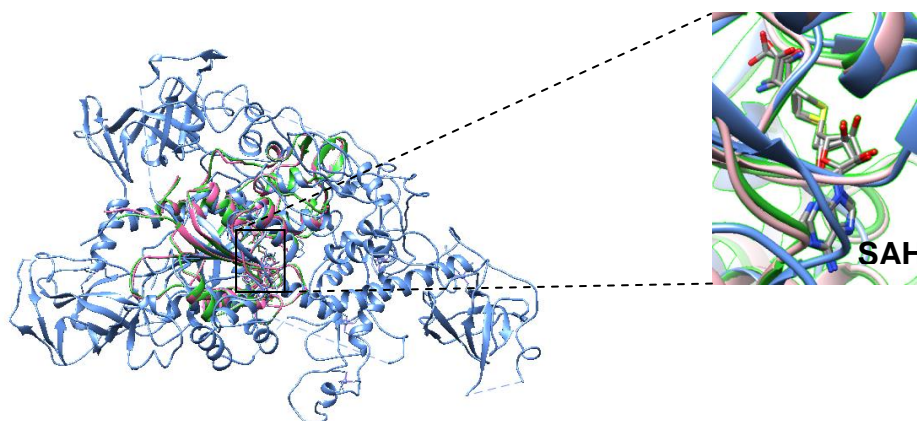


Figura 34. Proteínas DNMT1 (azul), DNMT3A (rosa) y DNMT3B (verde) superpuestas con UCSF Chimera con la ampliación del sitio de unión del ligando SAH

Una vez que las proteínas están superpuestas, se ha generado una alineación basada en la estructura de las tres secuencias usando *Match > Align*. Esta opción utiliza solo la distancia entre los C α para crear la alineación, sin tener en cuenta los tipos de residuos y cómo se superpusieron las estructuras. En la Figura 35 se muestra una parte de la alineación en la zona del dominio catalítico (regiones de los sitios de unión del ligando SAH), donde las columnas con residuos de las tres estructuras se resaltan como una región naranja y el encabezado “RMSD: ca” muestra el RMSD del C α como un histograma. Con esto, se corrobora que el dominio catalítico de DNMT1 está bastante conservado respecto a DNMT3A y DNMT3B, mientras que el resto de su estructura es bastante diferente.

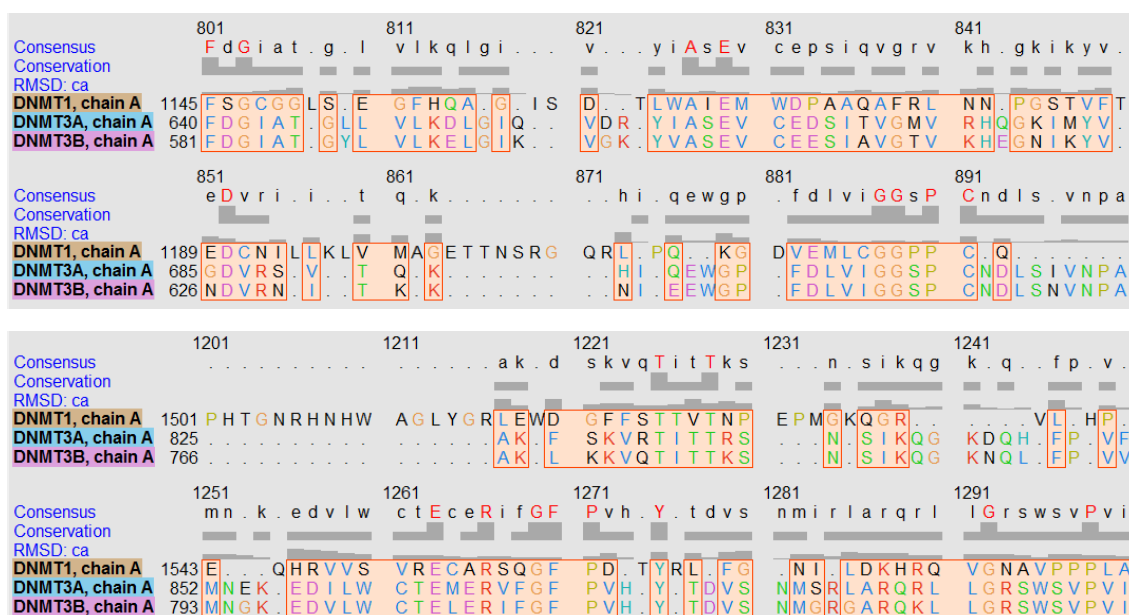


Figura 35. Alineamiento de las proteínas DNMT1, DNMT3A y DNMT3B en una región de su dominio catalítico generado con UCSF Chimera

Además, dadas las interacciones de cada proteína con el ligando SAH en el cristal generadas con PLIP, se puede analizar concretamente la conservación en sus sitios de unión. Esto pone de manifiesto similitudes, sobre todo entre las proteínas DNMT3A y DNMT3B, como era de esperar. Por un lado, el residuo que participa la interacción π -stacking se trata de un fenilo en las tres proteínas (posición 1146 para DNMT1, 640 para DNMT3A y 581 para DNMT3B, como se muestra en la Figura 35), por lo que está conservado. En cuanto a los residuos que realizan los enlaces de hidrógeno, únicamente se mantiene la asparagina (posición 1190 para DNMT1, 686 para DNMT3A y 627 para DNMT3B), mientras que el resto de residuos no están conservados entre DNMT1 y DNMT3A/B.

4.7. Evaluación de los *hits* como inhibidores potenciales de DNMT3

4.7.1. Cribado virtual de las proteínas con sus inhibidores

Para analizar si los fenoles del AOVE determinados como posibles inhibidores de las proteínas DNMT1 y COMT también podrían inhibir a DNMT3A y DNMT3B, se ha realizado el cribado virtual de estos fenoles. El procedimiento seguido es análogo al descrito para las proteínas DNMT1 y COMT, así que ha sido necesario definir el sitio de unión de los ligandos en estas proteínas, para lo que también se ha calculado una caja de 6 Å a partir del centro geométrico de la cavidad en la que se coloca el ligando y estos son los resultados obtenidos para cada estructura:

- Para la estructura 6W8B de DNMT3A *wild-type*, las coordenadas del centro del sitio de unión del SAH son en el eje x -4.820, en el eje y 44.700 y en el eje z -2.212, y las dimensiones son 19 x 24 x 19 Å.
- Para la estructura 6W89 de DNMT3A R882H, las coordenadas del sitio de unión del SAH son en el eje x 224.025, en el eje y -25.538 y en el eje z 7.368, y las dimensiones son 26 x 17 x 19 Å.
- Para la estructura 6KDA de DNMT3B, las coordenadas del sitio de unión del SAH son en el eje x 16.209, en el eje y -70.76 y en el eje z -10.674, y las dimensiones son 22 x 19 x 22 Å.

También se ha realizado el cribado virtual entre el ligando SAH y cada proteína para evaluar cómo funciona el método seguido y tener una referencia. La energía de enlace predicha para el ligando SAH en las dos estructuras de DNMT3A es de -8.6 kcal/mol y en la estructura de DNMT3B es de -9.1 kcal/mol. Además, en todos los casos, se mantienen prácticamente todas las interacciones proteína–ligando respecto a las del cristal, y la superposición entre la molécula SAH obtenida tras el *docking* y la del cristal es bastante buena, tal como se muestra en la Figura 36 para la estructura 6W8B de DNMT3A *wild-type* a modo de ejemplo.

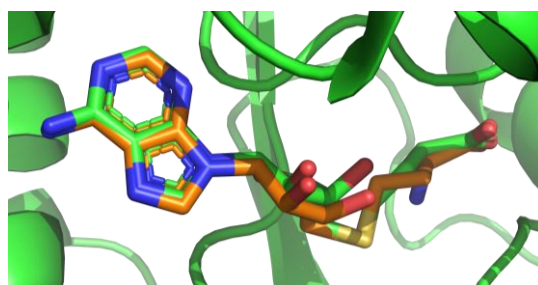


Figura 36. Superposición con PyMOL de la molécula de SAH obtenida en el *docking* (naranja) con SAH del cristal en la estructura 6W8B de DNMT3A *wild-type* (verde)

4.7.2. Cribado virtual de los *hits* de los fenoles del AOVE con DNMT3

A continuación, se ha llevado a cabo el cribado virtual de los siete fenoles del AOVE seleccionados como posibles candidatos para ser inhibidores de DNMT1 y COMT con las proteínas DNMTs de tipo 3. Los resultados obtenidos se han analizado según su energía de enlace predicha y a sus interacciones con las proteínas respecto a las del cristal. Los fenoles seleccionados como los mejores resultados para ambas se muestran en la Tabla 3 y se ha tenido en cuenta que se mantuviera la interacción π -stacking de la estructura cristalina.

Tabla 3. Mejores fenoles para las proteínas DNMT3A y DNMT3B. Los datos (mejor pose y energía de enlace predicha) de DNMT3A son de la estructura *wild-type* 6W8B

Proteína	Compuesto	PubChem CID	Mejor pose	Energía (kcal/mol)	nRot
DNMT3A	Taxifolina	12309272	3	-10.2	0
	Isoacteósido	6476333	2	-10.1	20
	<i>Hellicoside</i>	5281778	1	-10.0	23
	Eriodictiol	90657147	3	-9.5	5
DNMT3B	Quercetina-7-O-glucósido	91820114	3	-10.9	11
	Crisoeriol-7-O-glucósido	11294177	2	-10.7	11
	Taxifolina	12309272	3	-10.5	0
	Isoacteósido	6476333	2	-9.6	20

En general, los resultados de las proteínas DNMT3 son peores de los obtenidos para estos fenoles con DNMT1 porque se mantienen menos interacciones del cristal, en concreto, menos enlaces de hidrógeno, aunque las energías de enlace predichas son bastante similares. Aun así, estos fenoles siguen siendo posibles candidatos a unirse e inhibir estas proteínas DNMT3. Para la enzima DNMT3A, se pueden destacar los compuestos glucósidos isoacteósido y *hellicoside*, ya que son los que más enlaces de hidrógeno mantienen respecto del cristal, 5 y 6, respectivamente, además de la interacción π -stacking. En cambio, para DNMT3B, los fenoles quercetina-7-O-glucósido y crisoeriol-7-O-glucósido son los que mejor se solapan con el ligando SAH del cristal, manteniendo cada uno cuatro de los enlaces de hidrógeno y la interacción π -stacking.

4.8. Análisis de moléculas comerciales con potencial antiangiogénico

Una vez determinados los *hits* de los compuestos fenólicos del AOVE con potencial para actuar como inhibidores de las proteínas DNMTs y COMT, un paso lógico es determinar si estos compuestos están disponibles comercialmente, ya que a veces puede ser complicado preparar compuestos naturales. Una forma de saber si los fenoles seccionados son comerciales es empleando la base de datos PubChem, porque, entre toda la información que aporta sobre los compuestos, tiene un apartado sobre los vendedores de productos químicos. Tras realizar esta búsqueda se ha encontrado que de los siete *hits*, únicamente los glucósidos quercetina-7-O-glucósido y *hellicoside* no son comerciales, mientras que los otros cinco están a la venta:

- Crisoeriol-7-O-glucósido: 14 vendedores, como AvaChem Scientific (5 mg por 390 €).
- Taxifolina: 13 vendedores, como Matrix Scientific (500 mg por 119 €).
- Eriodictiol: 44 vendedores, como Sigma-Aldrich (10 mg por 193 €).
- Isoacteósido: 37 vendedores, como Selleck Chemicals (1 mg por 140 €).
- *Suspensaside*: solo se vende en BLDpharm bajo pedido.

Una vez determinados los compuestos fenólicos del AOVE con potencial para inhibir las proteínas DNMTs y COMT que son comerciales, en un siguiente paso del estudio deberían adquirirse y testarlos experimentalmente. Por otro lado, comentar que como solo son cinco compuestos comerciales, lo ideal sería tratar de tener una lista más amplia, ya que muchas veces no todos los *hits* computacionales se validan experimentalmente. Así que la siguiente parte del trabajo ha sido determinar compuestos comerciales similares a estos *hits*, para después analizar si también podrían actuar como inhibidores de las proteínas DNMTs y COMT, y ser usados como fármacos antiangiogénicos. Aunque no era necesario, se ha dividido la búsqueda de moléculas comerciales en dos grupos, según si son similares a los glucósidos quercetina-7-O-glucósido y *hellicoside* o a los cinco fenoles comerciales.

4.8.1. Cribado virtual compuestos similares a los *hits* no comerciales

Primero hay que realizar la búsqueda de los compuestos similares comerciales a través de SwissSimilarity usando la base de datos ZINC y el resto de proveedores disponibles (Asinex, AsisChem, ChemBridge, ChemDiv, Enamine, InnovaPharm, Maybridge, Otava, Selleckchem, Sigma-Aldrich, SPECS, TimTec y Vitas). Como se ha comentado en la sección 3.7 del capítulo Materiales y métodos, únicamente se seleccionan los compuestos con una puntuación de similitud superior a 0.8. Tras eliminar los duplicados, se han obtenido 51 compuestos comerciales similares a los glucósidos quercetina-7-O-glucósido y *hellicoside*. Después, se han convertido las cadenas SMILES en la estructura 3D con el script de Chimera de la Figura 21 para automatizar el proceso.

Tras analizar los resultados obtenidos para DNMT1, en la Tabla 4 se ha recogido la selección de los mejores compuestos en cuanto a su energía de enlace y a las interacciones que mantienen respecto a las del cristal entre los 25 primeros resultados. Como se puede observar, los cuatro *hits* determinados son similares al fenol quercetina-7-O-glucósido y la mejor pose para todos los compuestos es la primera, por lo que no se ha incluido esta información en la tabla.

Tabla 4. Compuestos comerciales similares a quercetina-7-O-glucósido y *hellicoside* identificados como *hits* para la proteína DNMT1

Compuesto	ID (nº CAS)	Energía (kcal/mol)	Fenol AOVE ^a	Similitud ^b
Apigenina 7-O-neohesperidosida	12676__SIGMA (CAS 17306-46-6)	-10.7	Quercetina ^c	0.871
Luteolina 7-glucósido	ZINC03947431 (CAS 5373-11-5)	-10.6	Quercetina ^c	0.937
Kaempferol 7-O-neohesperidosida	92996__SIGMA (CAS 17353-03-6)	-10.5	Quercetina ^c	0.985
-	ZINC12419107	-10.3	Quercetina ^c	0.922
Apiin	17784__SIGMA (CAS 26544-34-3)	-10.3	Quercetina ^c	0.863

^a Fenol del AOVE al que es similar el compuesto. ^b Puntuación de similitud dada por SwissSimilarity, de 1 para compuestos idénticos a 0 para los totalmente diferentes. ^c Se refiere a quercetina-7-O-glucósido.

Todos estos compuestos muestran energías de enlace comparables a las observadas con los fenoles del AOVE, además de interacciones con DNMT1 similares, ya que también mantienen la interacción π -stacking y varios enlaces de hidrógeno. En concreto, cabe mencionar el compuesto apigenina 7-O-neohesperidosida que mantiene 8 de los 9 enlaces de hidrógeno, mientras que kaempferol 7-O-neohesperidosida mantiene 7 y los otros dos compuestos 6 de estos enlaces. Además, luteolina 7-glucósido es un fenol del AOVE que ya había evaluado en el primer cribado virtual, pero con peores resultados. Por otra parte, el compuesto con código ZINC12419107, según la base de datos ZINC, ya no está disponible comercialmente y no se encuentra en la página del proveedor TimTec, desde cuyo catálogo se había encontrado esta molécula. Por lo tanto, aunque se ha incluido en esta lista no se tendrá en cuenta como posible candidato.

En cuanto a los mejores resultados con las dos estructuras de la proteína COMT, 3BWM para 108V y 3BWY para 108M, estos se recogen en la Tabla 5 siguiendo un procedimiento análogo a los anteriores casos. La mejor pose de todos los *hits* es la primera, por lo que en la tabla no se indica este dato. De nuevo, se han seleccionado únicamente los compuestos con dos grupos hidroxilo cerca del ion magnesio, de manera que puedan interactuar, ya que PLIP no proporciona estas interacciones. Comentar que los compuestos marcados con un asterisco son fenoles presentes en el

AOVE, por lo que ya se habían empleado para realizar el cribado virtual, pero ahora se han conseguido mejores resultados. Estos compuestos son los flavonoides verbascósido, cloruro de cianidina-3-O-rutinosido y cloruro de cianidina 3-O-glucósido, y el glucósido escolimosido.

Tabla 5. Compuestos comerciales similares a quercetina-7-O-glucósido y *hellicoside* como *hits* para la proteína COMT. Los datos de mejor pose y energía de enlace predicha son de la estructura 3BWM

Compuesto	ID (nº CAS)	Energía (kcal/mol)	Fenol AOVE ^a	Similitud ^b
Luteolin-7-O-β-D-glucurónido	04480585__FLUKA (CAS 29741-10-4)	-7.7	Quercetina ^c	0.853
Cloruro de cianidina-3-O-rutinosido*	36428__FLUKA (CAS 18719-76-1)	-7.6	Quercetina ^c	0.888
Cloruro de cianidina 3-O-galactósido	19050__SIGMA (CAS 27661-36-5)	-7.3	Quercetina ^c	0.830
Cloruro de cianidina 3-O-glucósido*	44689__FLUKA (CAS 7084-24-4)	-7.1	Quercetina ^c	0.824
Verbascósido*	00820580__FLUKA (CAS 61276-17-3)	-7.1	<i>Hellicoside</i>	0.985
Escolimosido*	SMB00200__SIGMA (CAS 20633-84-5)	-7.0	Quercetina ^c	0.916

^a Fenol del AOVE al que es similar el compuesto. ^b Puntuación de similitud dada por SwissSimilarity, de 1 para compuestos idénticos a 0 para los totalmente diferentes. ^c Se refiere a quercetina-7-O-glucósido.

En cuanto a las interacciones de los *hits* con la proteína COMT, todos mantienen las dos interacciones hidrofóbicas y dos de los tres enlaces de hidrógeno de la estructura 3BWM de COMT 108V. Sin embargo, los fenoles verbascósido y escolimosido no se encuentran entre los 25 mejores resultados de la proteína COMT 108M, pero dado que mantenían prácticamente todas las interacciones del cristal COMT 108V se ha decidido incluirlos en esta lista, aunque no se van a estudiar sus propiedades ADMET y no se tendrán en cuenta como posibles candidatos.

4.8.2. Cribado virtual compuestos similares a los *hits* comerciales

Por otra parte, se han buscado compuestos comerciales similares a los *hits* del AOVE comerciales con una similitud superior a 0.8. Se ha seguido un proceso análogo al anterior y, a continuación, se ha realizado el cribado virtual con las proteínas. Tras analizar los resultados obtenidos para DNMT1, en la Tabla 6 se recoge la selección de los mejores compuestos en cuanto a su energía de enlace predicha y a las

interacciones que mantienen respecto a las del cristal, de manera que todos se corresponden con la pose 1 del *docking*. A pesar de que las energías de enlace son similares a las obtenidas con los fenoles del AOVE, mantienen menos interacciones respecto a las del cristal, aunque siguen presentando la interacción π -stacking y varios enlaces de hidrógeno. Por lo tanto, se ha realizado la selección entre los 25 primeros resultados que mantienen dicha interacción π -stacking y seis de los enlaces de hidrógeno. Además, se han eliminado los compuestos con ID ZINC00668104, ZINC36046318, ZINC00668105 y ZINC12419107, que a pesar de que mostraban buenas interacciones con la proteína, no se han encontrado en la base de datos de ZINC, ni en la de sus respectivos vendedores.

Tabla 6. Compuestos comerciales similares a los fenoles del AOVE comerciales identificados como *hits* para la proteína DNMT1

Compuesto	ID (nº CAS)	Energía (kcal/mol)	Fenol AOVE ^a	Similitud ^b
-	ZINC04543717	-10.7	Crisoeriol ^c	0.996
-	ZINC04543718	-10.7	Crisoeriol ^c	0.996
Luteolina 7-glucósido	ZINC03947431 (CAS 5373-11-5)	-10.6	Eriodictiol	0.919
Kaempferol 7-O-neohesperidosida	92996__SIGMA (CAS 17353-03-6)	-10.5	Crisoeriol ^c	0.855
Apiin	17784__SIGMA (CAS 26544-34-3)	-10.3	Crisoeriol ^c	0.975
Apigenina 7-O-glucurónido	04490590__FLUKA (CAS 29741-09-1)	-10.3	Crisoeriol ^c	0.947
Apigenina 7-glucósido	00720585__FLUKA (CAS 578-74-5)	-10.3	Crisoeriol ^c	0.985
Luteolin-7-O- β -D-glucurónido	04480585__FLUKA (CAS 29741-10-4)	-10.3	Crisoeriol ^c	0.982
Procianidina B1	19542__FLUKA (CAS 20315-25-7)	-10.2	Taxifolina	0.841

^a Fenol del AOVE al que es similar el compuesto. Si lo es con más de un fenol se escoge al que se parece más. ^b Puntuación de similitud dada por SwissSimilarity, de 1 para compuestos idénticos a 0 para los totalmente diferentes. ^c Se refiere a crisoeriol-7-O-glucósido.

En cuanto a los resultados para la proteína COMT, se ha tenido en cuenta los resultados obtenidos para las dos estructuras estudiadas, 3BWM y 3BWY. Esto ha llevado a seleccionar nuevos compuestos comerciales como posibles inhibidores de esta proteína, los cuales se muestran en la Tabla 7. La selección se ha realizado entre los 25 primeros compuestos teniendo en cuenta la energía de enlace y las

interacciones que se mantienen respecto del cristal. Al igual que anteriormente, solo se han seleccionado compuestos en los que se podría mantener las interacciones entre dos grupos hidroxilo y el catión magnesio de la estructura. Después, se han tenido en cuenta el resto de interacciones, de manera que se mantengan el máximo número de ellas, lo que ha llevado a seleccionar aquellos en los que se mantienen las dos interacciones hidrofóbicas y dos de los tres enlaces de hidrógeno de COMT 108V.

Tabla 7. Compuestos comerciales similares a los fenoles del AOVE comerciales identificados como *hits* para la proteína COMT. Los datos de mejor pose y energía de enlace predicha son de la estructura 3BWM

Compuesto	ID (nº CAS)	Mejor pose	Energía (kcal/mol)	Fenol AOVE ^a	Similitud ^b
(+)-Catequina	ZINC00119978 (CAS 154-23-4)	1	-8.1	Taxifolina	0.958
(±)-Catequina-2,3,4- ¹³ C ₃	719579__ALDRICH (CAS 1261254-33-4)	3	-8.0	Taxifolina	0.951
(±)-Catequina hidrato	03910590__FLUKA (CAS 7295-85-4)	3	-8.0	Taxifolina	0.951
(-)-Catequina	ZINC00119985 (CAS 18829-70-4)	3	-8.0	Taxifolina	0.952
(-)-Leucofisetinidina	ZINC59727703 (CAS 34620-73-0)	2	-7.9	Taxifolina	0.946
Butin	ZINC04098238 (CAS 492-14-8)	1	-7.9	Eriodictiol	0.991
7-O-metileriodictiol	ZINC01081535 (CAS 51857-11-5)	1	-7.8	Eriodictiol	0.969
Dihidrotaxifolina	ZINC01668762 (CAS 6270-97-9)	3	-7.5	Taxifolina	0.928
<i>Cyrtometin</i>	ZINC00004749 (CAS 95272-99-4)	2	-7.4	Eriodictiol	0.980

^a Fenol del AOVE al que es similar ese compuesto comercial. ^b Puntuación de similitud dada por SwissSimilarity, que va de 1 para compuestos idénticos a 0 para compuestos totalmente diferentes.

Los primeros 4 compuestos de la Tabla 7 son esteroisómeros de la catequina, ya que esta molécula presenta dos centros quirales (Figura 37). De manera que se encuentra la (+)-catequina, que es el isómero más común, y la (-)-catequina, que es el otro isómero, pero también está el compuesto (±)-catequina, que es una mezcla de los dos enantiómeros, y la (±)-catequina-2,3,4-¹³C₃, que es igual pero está marcada con ¹³C, lo que carece de importancia para su potencia como inhibidor, además de aumentar el precio de venta, por lo que no se a tener en cuenta.

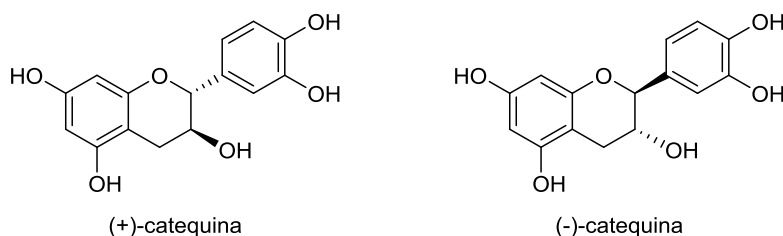


Figura 37. Estructuras de los estereoisómeros de la catequina

4.8.3. Análisis propiedades ADMET *hits* comerciales

También se ha realizado el análisis de las propiedades farmacocinéticas y de toxicidad ADMET de estos compuestos comerciales con el servidor web pkCSM y se han tenido en cuenta los mismo criterios que los comentados para los fenoles del AOVE. En el Anexo III se recoge una tabla con el resultado para los parámetros analizados en este estudio. Sin embargo, los resultados obtenidos no son mejores que para los fenoles del AOVE analizados anteriormente. Al igual que ocurría con estos fenoles del AOVE, todos estos compuestos son sustrato de la Pgp y tienen toxicidad por *T. Pyriformis*. Además, ninguno muestra permeabilidad a Caco-2, a excepción del butin, que a su vez es el único que no se considera incapaz de penetrar al SNC, por lo que tampoco muestra buenos resultados.

4.8.4. Selección compuestos comerciales con más potencial para actuar como inhibidores

Al igual que ocurría con los fenoles del AOVE, no hay grandes diferencias entre los resultados obtenidos en el análisis de sus propiedades ADMET, por lo que se ha decidido no descartar ninguno de los compuestos comerciales. Aunque estos fenoles tengan propiedades ADMET mejorables, el siguiente paso sería validar su capacidad inhibidora experimentalmente y, a continuación, se pasaría a modificar su estructura para aumentar su potencia si es posible y resolver estos problemas de sus propiedades ADMET (*lead to optimization*).

Por lo que de nuevo, la selección de los mejores *hits* se ha realizado teniendo en cuenta sus interacciones con las proteínas, de manera que los inhibidores de DNMT1 mantengan la interacción π -stacking, junto con entre 5 y 8 enlaces de hidrógeno respecto al cristal. Así, se han obtenido seis inhibidores potenciales de la proteína DNMT1, que son apigenina 7-O-neohesperidosida, luteolina 7-glucósido, kaempferol 7-O-neohesperidosida, apiin, apigenina 7-O-glucurónido y procianidina B1 (Figura 38). Cabe destacar que todos estos compuestos son fenoles de la familia de los flavonoides que se encuentran en la naturaleza. La apigenina 7-O-neohesperidosida está presente en varias plantas como *Paeonia suffruticosa* y *Buddleja albiflora*, entre otras. El kaempferol 7-O-neohesperidoside se encuentra en las semillas del fruto del lichi. La luteolina 7-glucósido aparte de en el AVOE, se encuentra en plantas como

Reseda luteola de la que se ha aislado. El apiin, también llamado apigenina-7-apiogluósido, es el componente flavonoide principal del apio y también está presente en el perejil. La apigenina 7-O-glucurónido se puede obtener de las plantas escutalaria barbada y escutalaria de Virginia. La procianidina B1 se puede encontrar en la canela de Ceilán, en la raíz de la uña de gato y en algunas frutas, como el melocotón.

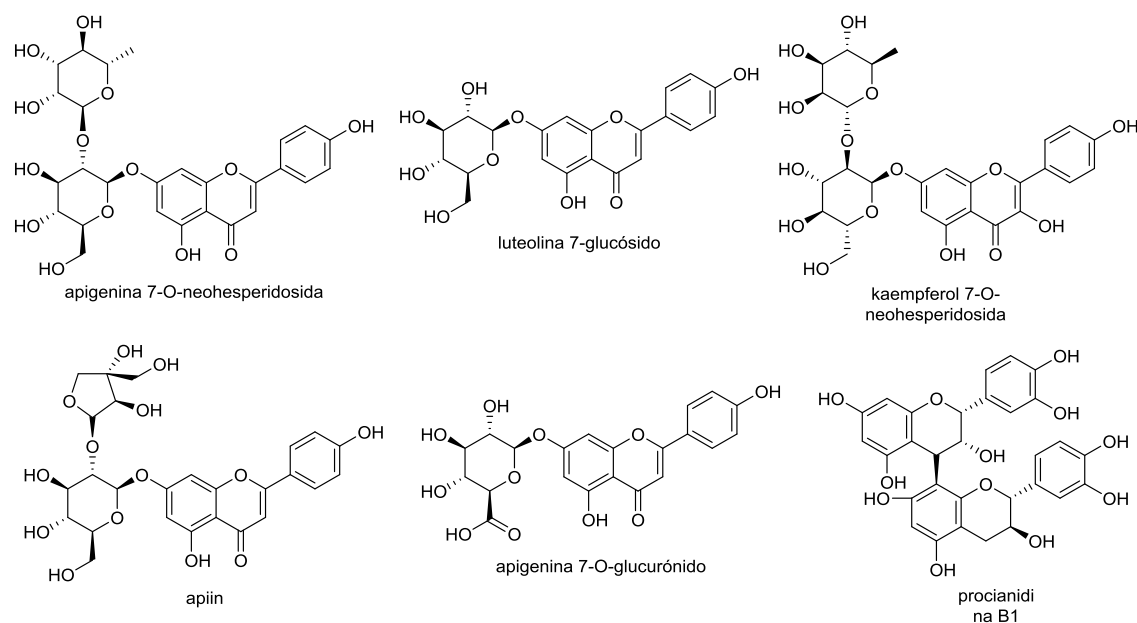


Figura 38. Compuestos comerciales seleccionados como *hits* para la proteína DNMT1

Por otra parte, para la proteína COMT se han seleccionado los compuestos con dos grupos hidroxilo cerca del magnesio y que mantengan las dos interacciones hidrofóbicas y dos enlaces de hidrógeno del cristal. Con esto, se han obtenido 10 inhibidores potenciales de esta enzima: luteolin-7-O- β -D-glucurónido, cloruro de cianidina-3-O-rutinosido, cloruro de cianidina 3-O-galactósido, cloruro de cianidina 3-O-glucósido, catequina, (-)-leucofisetinidina, butin, 7-O-metileriodictiol, dihidrotaxifolina y *cyrtominetin* (Figura 39). De nuevo, todos estos compuestos son fenoles naturales de la familia de los flavonoides. El luteolin-7-O- β -D-glucurónido se encuentra en la planta *Acanthus hirsutus* y en el centeno. El cloruro de cianidina 3-O-glucósido, además de en el AOVE, está en la planta *Antirrhinum majus* y en frutas, como el açai, la grosella negra, la frambuesa negra, el lichi y el higo. La catequina es producida por el árbol Cutch y se encuentra abundantemente en hojas de té, semillas de uva, vino y verduras. Además, las catequinas son los principales flavonoides presentes en el té verde, representando el 90% del total. La mezcla racémica, (\pm)-catequina se puede extraer, por ejemplo, de la raíz de *Centaurea maculosa*. La (-)-leucofisetinidina se encuentra, en la plantas *Acacia mearnsii* y *Schinopsis balansae*, entre otras. Finalmente, el 7-O-metileriodictiol, también llamado *sterubin*, se extrae de la planta Yerba Santa.

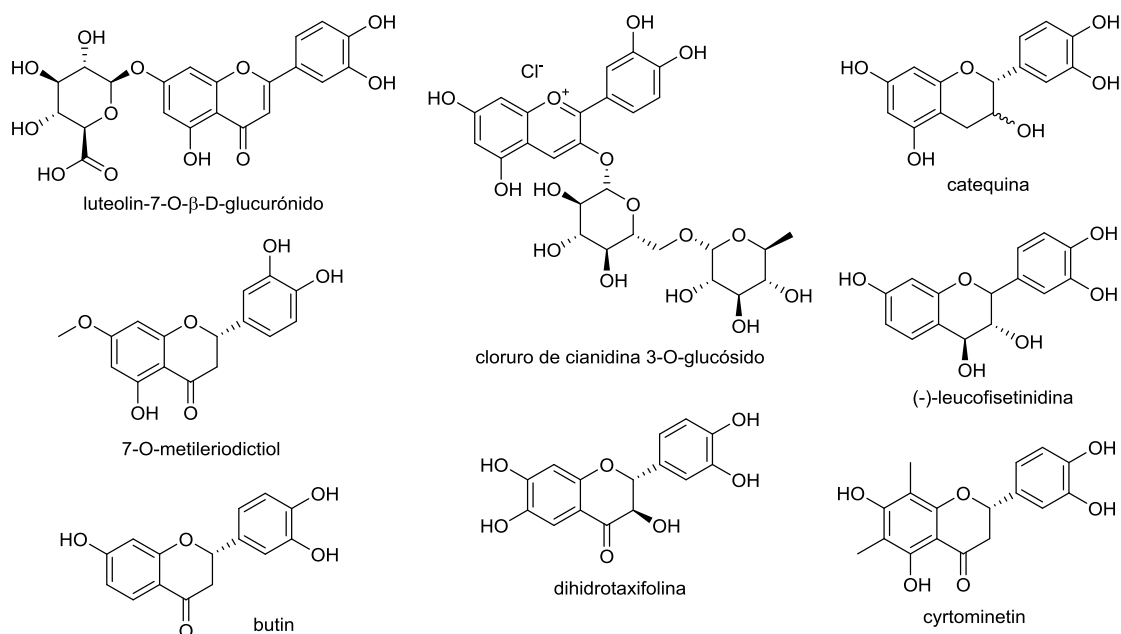


Figura 39. Compuestos comerciales seleccionados como *hits* para la proteína COMT

A modo de resumen, se han obtenido seis inhibidores potenciales de DNMT1 y diez de COMT, todos ellos flavonoides naturales que se encuentran en distintas plantas. Sin embargo, como se puede observar, no hay ningún compuesto con potencial para actuar como inhibidor dual DNMT1/COMT de manera directa.

4.8.5. Evaluación de los *hits* comerciales como inhibidores de DNMT3

Finalmente, se realizó de nuevo el cribado virtual empleando las proteínas DNMT3A y DNMT3B con los mejores compuestos comerciales seleccionados anteriormente. Sin embargo, al igual que ocurría con los fenoles del AOVE, las interacciones que se mantienen respecto al ligando SAH en las estructuras cristalinas son peores que para la proteína DNMT1. En la mayoría de los casos se mantiene la interacción π -stacking, pero menos enlaces de hidrógeno (tan solo entre dos y cinco de los 9 enlaces de hidrógeno del cristal). En cambio, se muestran algunas interacciones hidrofóbicas, que no existen en el cristal, en algunos casos con residuos que en el cristal están participando en los enlaces de hidrógeno.

5. Conclusiones

La angiopreención del cáncer basada en la dieta ha surgido como un enfoque interesante para prevenir y tratar esta enfermedad en sus fases iniciales, mediante el uso de inhibidores de la angiogénesis. El potencial antiangiogénico que presentan muchos compuestos naturales contenidos en la dieta mediterránea hace que este patrón dietético sea una fuente importante de agentes quimiopreventivos, con especial foco en el aceite de oliva. A continuación, se comentan las conclusiones derivadas del trabajo, así como otros aspectos relaciones con su realización.

- **Conclusiones del trabajo**

A partir de los hallazgos realizados en este trabajo, se puede concluir de forma general que varios compuestos fenólicos, tanto presentes en el AOVE como en otras fuentes naturales, tienen el potencial de ser utilizados como agentes antiangiogénicos que pueden ser efectivos en el tratamiento de varios tipos de cáncer.

En concreto, los enfoques computacionales *in silico* empleados sugieren que fenoles del AOVE podrían unirse al sitio de unión del cofactor SAM en las proteínas DNMTs y, como consecuencia, inhibir su actividad catalítica. También dihidroxifenoles del AOVE podrían actuar como sustrato/inhibidor dual de la O-metilación de catecoles mediada por la enzima COMT, al unirse a esta proteína en el sitio de unión del sustrato catecol. Así, se han seleccionado dos inhibidores potenciales de la proteína DNMT1 (crisoeriol-7-O-glucósido y *hellicoside*), dos de COMT (taxifolina y eriodictiol) y tres duales (*suspensaside*, quercetina-7-O-glucósido e isoacteosido). Los inhibidores de DNMT1 muestran potencial para inhibir a las metiltransferasas *de novo* DNMT3, por lo que estos fenoles no parecen ser selectivos. Además, los inhibidores de COMT, aunque no inhiban directamente a las enzimas DNMTs, pueden funcionar como moduladores indirectos del proceso de metilación del ADN con el aumento de la formación de SAH durante la reacción de O-metilación, un potente inhibidor de estas metiltransferasas dependientes de SAM.

Por otra parte, en este estudio se pone de manifiesto diferencias entre las distintas familias de los fenoles del AOVE, ya que los fenoles seleccionados como posibles inhibidores de las proteínas pertenecen únicamente a las familias de flavonoides y glucósidos. Esto parece indicar que estas dos familias de fenoles tienen un mayor potencial como agentes antiangiogénicos respecto a las otras. Además, todos los compuestos comerciales similares a los *hits* del AOVE son flavonoides naturales. En la literatura ya se han descrito varios precedentes de flavonoides que muestran un gran potencial como agentes anticancerígenos.

No se han observado grandes diferencias en los resultados del cribado virtual de la proteína COMT según su polimorfismo Val^{108/158}Met con energías de enlace predichas e interacciones ligando-proteína similares con los fenoles del AOVE y los compuestos comerciales similares a estos. Esto está de acuerdo con los estudios realizados con el fenol oleaceína, cuya potencia inhibidora de la O-metilación de sustratos de catecol no se vio alterada al sustituir la proteína COMT 108V por la

variante 108M. Esto es importante porque aunque este polimorfismo podría no ser un factor de riesgo para el cáncer, parece estar involucrado en el desarrollo de esta enfermedad en algunos grupos étnicos o tipos específicos de cáncer, por lo que es necesario verificar experimentalmente que la inhibición de este polimorfismo de COMT no altera la actividad catalítica de la proteína.

Aunque la ingesta diaria promedio de fenoles del AOVE no representa una carga abrumadora para el sistema de metilación mediado por las enzimas COMT y DNMTs dependiente de SAM, se pueden conseguir efectos más significativos si se administran como parte de mezclas que contienen inhibidores dirigidos a estas proteínas. Además, la combinación de este tipo de fármacos naturales y sintéticos podría ser más efectiva y tener menos efectos secundarios para prevenir o tratar el cáncer, ya que así se necesitarían dosis más bajas de estos agentes.

Finalmente, mencionar que el hecho de que los dihidroxifenoles de AOVE puedan actuar como inhibidores de la enzima COMT, también podría ser útil para desarrollar tratamientos destinados a prevenir o tratar la neurodegeneración en algunas patologías, como las enfermedades de Alzheimer y Parkinson.

- **Logro de los objetivos planteados**

En el Plan de Trabajo recogido en la PEC1 de la asignatura, se definieron los objetivos generales y específicos del proyecto. Durante el desarrollo de este trabajo se consiguió completar satisfactoriamente todos estos objetivos planteados, aunque se realizó algún cambio en el orden de varios objetivos específicos. En concreto, se decidió priorizar el estudio de las proteínas DNMT1 y COMT, y posponer el trabajo con las proteínas DNMT3A y DNMT3B, incluido el estudio de la conservación en los modos de unión de las proteínas DNMTs.

- **Seguimiento de la planificación y metodología**

En general, la planificación durante el desarrollo del trabajo se ha conseguido con un buen grado de cumplimiento, aunque ha habido retrasos en la realización de algunas tareas, que se han conseguido solucionar. Uno de los principales problemas ha sido el mal funcionamiento del servidor web MTiOpenScreen, empleado para realizar el cribado virtual entre las proteínas y los ligandos. Durante la primera fase del desarrollo, MTiOpenScreen funcionaba intermitentemente, provocando retrasos en la obtención de los resultados. A continuación, el servidor pasó a estar en mantenimiento durante varios días, lo que retrasó aún más el trabajo. Como consecuencia, se dedicó más tiempo del planeado inicialmente al uso de esta plataforma, pero también a analizar otras herramientas alternativas, como HADDOCK o PyRx, para realizar el *docking*. Finalmente, MTiOpenScreen volvió a funcionar, por lo que fue posible emplearla de nuevo para continuar con el estudio.

Debido a estos retrasos respecto a la planificación de las actividades, fue necesario reajustar el calendario propuesto y algunas de las tareas. Pero, a pesar de

este tiempo perdido, se consiguió completar todas las tareas propuestas, en parte, gracias a que otras tareas, como el análisis de las propiedades farmacocinéticas y de toxicidad ADMET, han requerido menos tiempo del propuesto para su realización. También el uso de los distintos scripts de Chimera, han reducido considerablemente el tiempo de preparación de los ligandos para el *docking* y del análisis de los resultados obtenidos en este proceso, lo que ha ayudado a disponer de más tiempo para completar el resto de tareas.

Finalmente, comentar que la metodología prevista en el Plan de Trabajo ha sido la adecuada para consecución de los distintos objetivos. Por lo tanto no ha sido necesario introducir ningún cambio.

- **Líneas de trabajo futuro**

En primer lugar, se podría extender el estudio a todos los fenoles del AOVE y no solo a los que tienen entrada a la base de datos PubChem. De esta forma, se podrían determinar otros compuestos fenólicos con potencial para ser empleados en terapia antiangiogénica, con especial foco en los flavonoides y glucósidos, que fueron los que mostraron un mayor potencial para inhibir las proteínas DNMTs y COMT.

Aunque no se ha entrado en detalle en la presente Memoria, las proteínas DNMTs pueden actuar como oncogén o como gen supresor de tumores dependiendo del tipo de cáncer, tal como se ha comentado en el capítulo de Antecedentes. En la práctica habría que tener en cuenta esta información en función de la aplicación que se esté buscando. Así, si en un determinado caso no se quiere inhibir todas las proteínas DNMTs, habría que buscar selectividad, lo que implicaría diseñar compuestos que exploten diferencias entre los sitios de unión de estas enzimas.

Por otra parte, también se pueden realizar simulaciones dinámicas moleculares, ya que, como se ha visto, los cálculos de *docking* se pueden emplear para explorar y predecir los modos de unión de un sustrato a una enzima, pero las simulaciones dinámicas pueden proporcionar información adicional sobre diferentes movimientos intra e intermoleculares a lo largo del tiempo. Por lo que pueden aportar distinta información sobre las moléculas a estudio.

Finalmente, como ya se ha comentado, esta línea de trabajo se podría continuar desarrollando. Para ello, habría que validar experimentalmente los resultados obtenidos, tanto *in vitro* como *in vivo*, algo que estaba fuera del alcance de este TFM.

6. Glosario

5mc	5-metilcitosina
ADMET	Absorción, Distribución, Metabolismo, Excreción y Toxicidad
ADN	Ácido Desoxirribonucleico
AOVE	Aceite de oliva virgen extra
CADD	Descubrimiento de fármacos asistido por ordenador (del inglés <i>Computer-Aided Drug Design</i>)
CD	Dominio catalítico
col.	Colaboradores
COMT	Enzima catecol O-metiltransferasa (MB-COMT forma asociada a membranas y S-COMT forma soluble en el citoplasma)
CpG	Citosina-fosfato-guanina
DNC	3,5-dinitrocatecol
DNMTs	Enzimas ADN metiltransferasas
GCO	Global Cancer Observatory
LMA	Leucemia mieloide aguda
NIH	National Institutes of Health
OMS	Organización Mundial de la Salud
PDB	Protein Data Bank
Pgp	Glicoproteína-P
PLIP	Protein-Ligand Interaction Profiler
RSCB PDB	Research Collaboratory for Structural Bioinformatics PDB
RMSD	<i>Root Mean Square Deviation</i>
SAH	S-adenosil-L-homocisteína
SAM	S-adenosil-L-metionina
SMILES	Simplified Molecular Input Line Entry Specification
SNC	Sistema nervioso central
TFM	Trabajo Fin de Máster
TRD	Dominio de reconocimiento de diana
UniProt	Universal Protein Resource
UniProtKB	Universal Protein Resource Knowledgebase
VEGF	Factor de crecimiento endotelial vascular

7. Bibliografía

1. Abotaleb, M.; Samuel, S.M.; Varghese, E.; Varghese, S.; Kubatka, P.; Liskova, A.; Büsselberg, D. Flavonoids in Cancer and Apoptosis. *Cancers* **2019**, *11*, 28–67.
2. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* **2019**, *144*, 1941–1953.
3. *Global Cancer Observatory* [en línea] [fecha de consulta: 23 de Septiembre de 2020]. Disponible en: <https://gco.iarc.fr/>.
4. (a) Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **2011**, *144*, 646–674. (b) Hanahan, D.; Weinberg, R.A. The hallmarks of cancer. *Cell* **2000**, *100*, 57–70.
5. Wang, Z.; Dabrosin, C.; Yin, X.; Fuster, M.M.; Arreola, A.; Rathmell, W.K.; Generali, D.; Nagaraju, G.P.; El-Rayes, B.; Ribatti, D.; Chen, Y.C.; Honoki, K.; Fujii, H.; Georgakilas, A.G.; Nowsheen, S.; Amedei, A.; Niccolai, E.; Amin, A.; Ashraf, S.S.; Helferich, B.; Yang, X.; Guha, G.; Bhakta, D.; Ciriolo, M.R.; Aquilano, K.; Chen, S.; Halicka, D.; Mohammed, S.I.; Azmi, A.S.; Bilsland, A.; Keith, W.N.; Jensen, L.D. Broad targeting of angiogenesis for cancer prevention and therapy. *Seminars in Cancer Biology* **2015**, *35*, S224–S243.
6. Albin, A.; Tosetti, F.; Li, V.W.; Noonan, D.M.; Li, W.W. Cancer prevention by targeting angiogenesis. *Nature Reviews Clinical Oncology* **2012**, *9*, 498–509.
7. (a) Battino, M.; Forbes-Hernández, T.Y.; Gasparri, M.; Afrin, S.; Cianciosi, D.; Zhang, J.; Manna, P.P.; Reboredo-Rodríguez, P.; Lopez, A.V.; Quiles, J.L.; Mezzetti, B.; Bompadre, S.; Xiao, J.; Giampieri, F. Relevance of functional foods in the Mediterranean diet: the role of olive oil, berries and honey in the prevention of cancer and cardiovascular diseases. *Critical Reviews in Food Science and Nutrition* **2019**, *59*, 893–920. (b) Colomer, R.; Menéndez, J.A. Mediterranean diet, olive oil and cancer. *Clin. Transl. Oncol.* **2006**, *8*, 15–21.
8. Reboredo-Rodríguez, P.; Varela-López, A.; Forbes-Hernández, T.Y.; Gasparri, M.; Afrin, S.; Cianciosi, D.; Zhang, J.; Manna, P.P.; Bompadre, S.; Quiles, J.L.; Battino, M.; Giampieri, F. Phenolic Compounds Isolated from Olive Oil as Nutraceutical Tools for the Prevention and Management of Cancer and Cardiovascular Diseases. *Int. J. Mol. Sci.* **2018**, *19*, 2305–2326.
9. (a) Cuyàs, E.; Gumuzio, J.; Lozano-Sánchez, J.; Carreras, D.; Verdura, S.; Llorach-Parés, L.; Sanchez-Martinez, M.; Selga, E.; Pérez, G.J.; Scornik, F.S.; Brugada, R.; Bosch-Barrera, J.; Segura-Carretero, S.; Martin, A.G.; Encinar, J.A.; Menendez, J.A. Extra Virgin Olive Oil Contains a Phenolic Inhibitor of the Histone Demethylase LSD1/KDM1A. *Nutrients* **2019**, *11*, 1656–1672. (b) Cuyàs, E.; Verdura, S.; Lozano-Sánchez, J.; Viciano, I.; Llorach-Parés, L.; Nonell-Canals, A.; Bosch-Barrera, J.; Brunet, J.; Segura-Carretero, A.; Sanchez-Martinez, M.; Encinar, J.A.; Menendez, J.A. The extra virgin olive oil phenolic oleacein is a dual substrate-inhibitor of catechol-O-methyltransferase. *Food Chem. Toxicol.* **2019**, *128*, 35–45. (c) Verdura, S.; Cuyàs, E.; Lozano-Sánchez, J.; Bastidas-Velez, C.; Llorach-Parés, L.; Fernández-Arroyo, S.; Hernández-Aguilera, A.; Joven, J.; Nonell-Canals, A.; Bosch-Barrera, J.; Martin-Castillo, B.; Vellon, L.; Sanchez-Martinez, M.; Segura-Carretero, S.; Menendez, J.A. An olive oil phenolic is a new chemotype of mutant isocitrate dehydrogenase 1 (IDH1) inhibitors. *Carcinogenesis* **2019**, *40*, 27–40. (d) Corominas-Faja, B.; Cuyàs, E.; Lozano-Sánchez, J.; Cufí, S.; Verdura, S.; Fernández-Arroyo, S.; Borrás-Linares, I.; Martin-Castillo, B.; Martin, Á.G.; Lupu, R.; Nonell-Canals, A.; Sanchez-Martinez, M.; Micol, V.; Joven, J.;

- Segura-Carretero, A.; Menendez, J.A. Extra-virgin olive oil contains a metabolo-epigenetic inhibitor of cancer stem cells. *Carcinogenesis* **2018**, *39*, 601–613.
10. Henning, S.M.; Wang, P.; Carpenter, C.L.; Heber, D. Epigenetic effects of green tea polyphenols in cancer. *Epigenomics* **2013**, *5*, 729–741.
 11. Tomar, V.; Mazumder, M.; Chandra, R.; Yang, J.; Sakharkar, M.K. Small Molecule Drug Design. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* **2019**, *3*, 741–760.
 12. Prieto-Martínez, F.D.; López-López, E.; Juárez-Mercado, K.E.; Medina-Franco, J.L. Chapter 2 - Computational Drug Design Methods—Current and Future Perspectives. *In Silico Drug Design, Repurposing Techniques and Methodologies* **2019**, 19–44.
 13. Jorgensen, W.L. The Many Roles of Computation in Drug Discovery. *Science* **2014**, *303*, 1813–1818.
 14. *GanttProject* [en línea] [fecha de consulta y descarga: 07 de Octubre de 2020]. Disponible en: <https://www.ganttproject.biz/>.
 15. Weinhold; B. Epigenetics: The Science of Change. *Environ. Health. Perspect.* **2006**, *114*, A160–A167.
 16. Jurkowska, R.Z.; Jurkowski, T.P.; Jeltsch, A. Structure and Function of Mammalian DNA Methyltransferases. *ChemBioChem* **2011**, *12*, 206–222.
 17. Wang, Y.; Leung, F.C. An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics* **2004**, *20*, 1170–1177.
 18. Ghanbarian, H.; Wagner N.; Polo, B.; Baudouy, D.; Kiani, J.; Michiels, J.-F.; Cuzin, F.; Rassoulzadegan, M.; Wagner, K.-D. Dnmt2/Trdmt1 as Mediator of RNA Polymerase II Transcriptional Activity in Cardiac Growth. *PLoS One.* **2016**, *11*, e0156953.
 19. Okano, M.; Xie, S.; Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **1998**, *19*, 219–220.
 20. Gao, L.; Emperle, M.; Guo, Y.; Grimm, S.A.; Ren, W.; Adam, S.; Uryu, H.; Zhang, Z.-M.; Chen, D.; Yin, J.; Dukatz, M.; Anteneh, H.; Jurkowska, R.Z.; Lu, J.; Wang, Y.; Bashtrykov, P.; Wade, P.A.; Wang, G.G.; Jeltsch, A.; Son, J. Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms. *Nat. Commun.* **2020**, *11*, 3355.
 21. Liang, Z.; Zhu, Y.; Liu, X.; Hu, G. Role of protein-protein interactions in allosteric drug design for DNA methyltransferases. *Adv. Protein Chem. Struct. Biol.* **2020**, *121*, 49–84.
 22. Ganesan, A. Epigenetic drug discovery: a success story for cofactor interference. *Phil. Trans. R. Soc. B.* **2018**, 373–388.
 23. Lyko, F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* **2018**, *19*, 81–92.
 24. Kulis, M.; Esteller, M. DNA methylation and cancer. *Adv. Genet.* **2010**, *70*, 27–56.
 25. Zhang, J.; Yang, C.; Wu, C.; Cui, W.; Wang, L. DNA Methyltransferases in Cancer: Biology, Paradox, Aberrations, and Targeted Therapy. *Cancers* **2020**, *12*, 2123.
 26. Ribatti, D.; Tamma, R. Epigenetic control of tumor angiogenesis. *Microcirculation* **2020**, *27*, e12602.
 27. Belinsky, S.A.; Klinge, D.M.; Stidley, C.A.; Issa, J.P.; Herman, J.G.; March, T.H.; Baylin, S.B. Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer Res.* **2003**, *63*, 7089–7093.
 28. Hellebrekers, D.M.; Jair, K.W.; Viré, E.; Eguchi, S.; Hoebbers, N.T.; Fraga, M.F.; Esteller, M.; Fuks, F.; Baylin, S.B.; van Engeland, M.; Griffioen, A.W. Angiostatic activity of DNA methyltransferase inhibitors. *Mol. Cancer Ther.* **2006**, *5*, 467–475.

29. Isakovic, L.; Saavedra, O.M.; Llewellyn, D.B.; Claridge, S.; Zhan, L.; Bernstein, N.; Vaisburg, A.; Elowe, N.; Petschner, A.J.; Rahil, J.; Beaulieu, N.; Gauthier, F.; MacLeod, A.R.; Delorme, D.; Besterman, J.M.; Wahhab, A. Constrained (L-)-S-adenosyl-L-homocysteine (SAH) analogues as DNA methyltransferase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2742–2746.
30. Chen, S.; Wang, Y.; Zhou, W.; Li, S.; Peng, J.; Shi, Z.; Hu, J.; Liu, Y.C.; Ding, H.; Lin, Y.; Li, L.; Cheng, S.; Liu, J.; Lu, T.; Jiang, H.; Liu, B.; Zheng, M.; Luo, C. Identifying novel selective non-nucleoside DNA methyltransferase 1 inhibitors through docking-based virtual screening. *J. Med. Chem.* **2014**, *57*, 9028–9041.
31. Bastos, P.; Gomes, T.; Ribeiro, L. Catechol-O-methyltransferase (COMT): an update on its role in cancer, neurological and cardiovascular diseases. *Rev. Physiol., Biochem. Pharmacol.* **2017**, *173*, 1–39.
32. Bai, H.-W.; Shim, J.-Y.; Yu, J.; Zhu, B.T. Biochemical and Molecular Modeling Studies of the O-Methylation of Various Endogenous and Exogenous Catechol Substrates Catalyzed by Recombinant Human Soluble and Membrane-Bound Catechol-O-Methyltransferases. *Chemical Research in Toxicology* **2007**, *20*, 1409–1425.
33. Tenhunen, J.; Salminen, M.; Lundström, K.; Kiviluoto, T.; Savolainen, R.; Ulmanen, I. Genomic organization of the human catechol O-methyltransferase gene and its expression from two distinct promoters. *Eur. J. Biochem.* **1994**, *223*, 1049–1059.
34. Rutherford, K.; Le Trong, I.; Stenkamp, R.E.; Parson, W.W. Crystal Structures of Human 108V and 108M Catechol O-Methyltransferase. *J. Mol. Biol.* **2008**, *380*, 120–130.
35. Vidgren, J.; Svensson, L.A.; Liljas, A. Crystal structure of catechol O-methyltransferase. *Nature* **1994**, *368*, 354–358.
36. Moschovou, K.; Melagraki, G.; Mavromoustakos, T.; Zacharia, L.C.; Afantitis, A. Cheminformatics and virtual screening studies of COMT inhibitors as potential Parkinson's disease therapeutics. *Expert Opinion on Drug Discovery* **2020**, *15*, 53–62.
37. Czarnota, S.; Johannissen, L.O.; Baxter, N.J.; Rummel, F.; Wilson, A.L.; Cliff, M.J.; Levy, C.W.; Scrutton, N.S.; Waltho, J.P.; Hay, S. Equatorial Active Site Compaction and Electrostatic Reorganization in Catechol-O-methyltransferase. *ACS Catalysis* **2019**, *9*, 4394–4401.
38. Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J. Kinetics of human soluble and membrane-bound catechol O-methyltransferase: A revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* **1995**, *34*, 4202–4210.
39. Shield, A.; Thomae, B.; Eckloff, B.W.; Wieben, E.D.; Weinshilboum, R.M. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Molecular Psychiatry* **2004**, *9*, 151–160.
40. Srivastava, S.P.; Goodwin, J.E. Cancer Biology and Prevention in Diabetes. *Cells* **2020**, *9*, 1380.
41. Zhou, Q.; Wang, Y.; Chen, A.; Tao, Y.; Song, H.; Li, W.; Tao, J.; Zuo, M.; Mo, F.; Li, C. Association between the COMT Val158Met polymorphism and risk of cancer: evidence from 99 case–control studies. *Onco. Targets Ther.* **2015**, *8*, 2791–2803.
42. Zhu, B.T. Catechol-O-Methyltransferase (COMT)-mediated methylation metabolism of endogenous bioactive catechols and modulation by endobiotics and xenobiotics: importance in pathophysiology and pathogenesis. *Curr. Drug. Metab.* **2002**, *3*, 321–349.
43. Ovaska, M.; Yliniemelä, A. A semiempirical study on inhibition of catechol O-methyltransferase by substituted catechols. *J. Comput. Aided Mol. Des.* **1998**, *12*, 301–307.

44. Masjost, B.; Ballmer, P.; Borroni, E.; Zürcher, G.; Winkler, F.K.; Jakob-Roetne, R.; Diederich, F. Structure-Based Design, Synthesis, and in vitro Evaluation of Bisubstrate Inhibitors for Catechol O-Methyltransferase (COMT). *Chem. Eur. J.* **2000**, *6*, 971–982.
45. Zhu, B.T.; Wang, P.; Nagai, M.; Wen, Y.; Bai, H.-W. Inhibition of human catechol-O-methyltransferase (COMT)-mediated O-methylation of catechol estrogens by major polyphenolic components present in coffee. *J. Steroid. Biochem. Mol. Biol.* **2009**, *113*, 65–74.
46. Lee, W.J.; Zhu, B.T. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. *Carcinogenesis* **2006**, *27*, 269–277.
47. Martínez-Poveda, B.; Torres-Vargas, J.A.; Ocaña, M.C.; García-Caballero, M.; Medina, M.A.; Quesada, A.R. The Mediterranean Diet, a Rich Source of Angiopreventive Compounds in Cancer. *Nutrients* **2019**, *11*, 2036.
48. Olive Wellness Institute™ [en línea] [fecha de consulta: 29 de Diciembre de 2020]. Disponible en: <https://olivewellnessinstitute.org/>
49. (a) Bonvino, N.P.; Liang, J.; McCord, E.D.; Zafiris, E.; Benetti, N.; Ray, N.B.; Hung, A.; Boskou, D.; Karagiannis, T.C. OliveNet™: a comprehensive library of compounds from *Olea europaea*. *Database* **2018**, 1–9. (b) OliveNet™ [en línea] [fecha de consulta: 23 de Octubre de 2020]. Disponible en: <https://mccordresearch.com.au/>.
50. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E.E. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **2019**, *8*, D1102–D1109.
51. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, *25*, 1605–1612.
52. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515.
53. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
54. Labbé, C.M.; Rey, J.; Lagorce, D.; Vavruša, M.; Becot, J.; Sperandio, O.; Villoutreix, B.O.; Tufféry, P.; Miteva, M.A. MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic Acids Res.* **2015**, *43*, W448–W454.
55. Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
56. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072.
57. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucl. Acids Res.* **2015**, *43*, W443–W447.
58. Zoete, V.; Daina, A.; Bovigny, C.; Michielin, O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *J. Chem. Inf. Model.* **2016**, *56*, 1399–1404.
59. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.

60. Willighagen, E.L.; O'Boyle, N.M.; Gopalakrishnan, H.; Jiao, D.; Guha, R.; Steinbeck, C.; Wild, D.J. Userscripts for the Life Sciences. *BMC Bioinformatics* **2007**, *8*, article number 487.
61. Zhang, Z.-M.; Li, S.; Lin, K.; Luo, Y.; Perry, J.J.; Wang, Y.; Song, J. Crystal Structure of Human DNA Methyltransferase 1. *J. Mol. Biol.* **2015**, *427*, 2520–2531.
62. Goulas, V.; Exarchou, V.; Troganis, A.N.; Psomiadou, E.; Fotsis, T.; Briasoulis, E.; Gerothanassis, I.P. Phytochemicals in olive-leaf extracts and their antiproliferative activity against cancer and endothelial cells. *Molecular Nutrition and Food Research* **2009**, *53*, 600–608.
63. Haque, W.; Pattanayak, S.P.; Sinha, B.N. Evaluation of taxifolin and phloretin as antiangiogenic flavonoids: An *in vivo*, *in vitro* experimental analysis. *Int. J. Pharm. Pharm. Sci.* **2015**, *7*, 72–79.
64. Lamy, S.; Ouanouki, A.; Beliveau, R.; Desrosiers, R.R. Olive oil compounds inhibit vascular endothelial growth factor receptor-2 phosphorylation. *Exp. Cell Res.* **2014**, *322*, 89–98.
65. Ley, T.J.; Ding, L.; Walter, M.J.; McLellan, M.D.; Lamprecht, T.; Larson, D.E.; Kandoth, C.; Payton, J.E.; Baty, J.; Welch, J.; Harris, C.C.; Lichti, C.F. et al. DNMT3A Mutations in Acute Myeloid Leukemia. *N. Engl. J. Med.* **2010**, *363*, 2424–2433.
66. Anteneh, H.; Fang, J.; Song, J. Structural basis for impairment of DNA methylation by the DNMT3A R882H mutation. *Nat. Commun.* **2020**, *11*, 2294.
67. Lin, C.-C.; Chen, Y.-P.; Yang, W.-Z.; Shen, J.C.K.; Yuan, H.S. Structural insights into CpG-specific DNA methylation by human DNA methyltransferase 3B. *Nucleic Acids Res.* **2020**, *48*, 3949–3961.

8. Anexos

8.1. Anexo I

En la siguiente tabla se recogen los compuestos fenólicos del AOVE recogidos en OliveNet™ y con entrada a la base de datos PubChem.

Familia	Fenol	PubChem CID
Cumarinas	Esculin	5281417
	Esculetin	5281416
	Scopoletin	5280460
	Scopolin	439514
Flavonoides	Cyanidin 3-O-glucoside	4481259
	Luteolin 7-glucoside	5280637
	Luteolin-4'-glucoside	5319116
	Luteolin-6-C-glucoside	49852298
	Luteolin-3',7-O-diglucoside	5490298
	Luteolin-8-C-glucoside	5382105
	Quercitrin	5280459
	Apigenin-7-O-glucoside	12304093
	Taxifolin	12309272
	Quercetin	5280343
	Chrysoeriol	5280666
	Diosmetin	5281612
	Methoxyluteolin	5317284
	Cyanidin (cation)	128861
	Eriodictyol	90657147
	Luteolin	5280445
	Quercetin 3-O-rutinoside	5280805
	Apigenin	5280443
	Delphinidin	128853
	Delphinidin-3-O-glucoside	443650
	Luteolin-7,4-O-diglucoside	44258093
	Vicenin-2	442664
	Cyanidin-3-O-rutinoside	441674
	Scolymoside	5282152
	Luteolin-4'-O-rutinoside	378305
	Isoquercitrin	5280804
	Chrysoeriol-7-O-glucoside	11294177
Hesperitin	72281	
Hesperidin	10621	

Glucósidos	Beta-hydroxy-acteoside	10009317
	Acteoside	5281800
	Isoacteoside	6476333
	Hellicoside	5281778
	Suspensaside	5281798
	Orbanchoside	6441894
	Wedelosin	156818
	Verucosin	10736226
	Quercetin-7-O-glucoside	91820114
	Oleoside	101042548
	Secologanoside	14136854
Salidroside	159278	
Hidroxi-isocromanos	1-(3-Methoxy-4-hydroxy)-phenyl-6,7-dihydroxy-isochroman	131750845
	1-Phenyl-6,7-dihydroxyisochroman	131750844
Ácidos hidroxibenzoicos	Syringic acid	10742
	Quinic acid	6508
	Shikimic acid	8742
	Gallic acid	370
	Vanillic acid	8468
	Phloretic acid	10394
	Protocatechuic acid	72
	Gentisic acid	3469
	4-hydroxybenzoic acid	135
	2,4-dihydroxybenzoic acid	1491
	2,6-Dihydroxybenzoic acid	9338
4-O-methyl- β -D-glucuronic acid	446874	
Ácidos hidroxicinámicos	Rosmarinic acid	5281792
	Chlorogenic Acid	1794427
	Sinapic acid	637775
	Ferulic acid	445858
	Caffeic acid	689043
	p-Coumaric acid	637542
	o-Coumaric Acid	637540
	m-Coumaric acid	637541
	Cinnamic acid	444539
	Caffeoylglucose	5281761
	Caftaric acid	6440397
Ácidos hidroxifenilacéticos	2-(4-hydroxyphenyl)acetic acid	127
	3,4-Dihydroxyphenylacetic acid	547
	Homoveratric acid	7139
	Homovanillic acid	7138
	2,5-Dihydroxyphenylacetic acid	780

Irridoides	Loganin	87691
	Loganic Acid	89640
Lignanós	Syringaresinol	100067
	1-Acetoxy-pinóresinol	442831
	(-)-Epipinóresinol	234817
	(+)-Epipinóresinol	637584
	(+)-Pinóresinol	73399
	Pinóresinol	17750970
	8-Hydroxypinóresinol	3010930
	(-)-Olivil	5273570
	Berchemol	14521044
Metoxifenóles	Guaiacol	460
	2-Methoxy-4-vinylphenol	332
	Isoeugenol	853433
	Homovanillin	151276
Secoiridoides	Neonuezhenide	6476723
	Nuezhenide	6440999
	10-Hydroxyoleuropein	6440747
	Oleuropein	5281544
	Oleuroside	102016333
	Demethyloleuropein	6450302
	Oleoside dimethylester	6476722
	Secologanic acid	439612
	Secologanol	14488076
	Secologanin	161276
	Oleuropein aglycone	56842347
	Oleacein	15553186
	Cornoside	3084796
	Oleocanthal	11652416
	p-HPEA-EDA	16681728
	Elenolic acid	169607
	Hydroxytyrosil Elenolate	101109285
	Ligstroside	14136859
	Ligstroside aglycone	71718370
Demethylligstroside	102461561	
Hydroxytyrosol acetate	155240	
Tyrosol acetate	11195402	
Fenóles simples	4-Vinylphenol	62453
	4-Methylcatechol	9958
	3,4,5-Trimethoxybenzoic acid	8357
	3,4-Dimethoxybenzoic acid	7121
	2,6-Dimethoxybenzoic acid	15109
	4-Hydroxybenzaldehyde	126

	Hydroxytyrosol	82755
	Homovanillyl Alcohol	16928
	Tyrosol	10393
	Syringaldehyde	8655
	3,4-Dihydroxyphenylglycol	91528
	Catechol	289
	Phenol	996
	p-cresol	2879
	m-cresol	342
	o-cresol	335
	4-Ethylguaiacol	62465
	4-ethylphenol	31242

8.2. Anexo II

A continuación, se recoge una tabla con los resultados obtenidos del análisis de las propiedades ADMET usando el servidor web pkCSM de los *hits* entre los fenoles del AOVE para las proteínas DNMT1 y COMT. Únicamente se muestran los parámetros que se han tenido en cuenta en este estudio, junto con “sustrato de Pgp” que ha sido positiva para todos los fenoles, por lo que no se ha incluido.

Compuesto	Caco-2*	Absorción intestinal	BBB*	SNC*	AMES tox.	hERG I inhibidor	hERG II inhibidor	Hepato-tox.	T.Pyrif. tox.	Minnow tox.
Isorhoifolina	0.285	32.39	-1.72	-4.73	No	No	Si	No	0.285	4.806
Quercitrina	0.048	52.71	-1.50	-4.16	No	No	No	No	0.285	4.954
Crisoeriol-7-O-glucósido	0.356	39.21	-1.61	-4.05	No	No	Si	No	0.285	4.631
Apigenina-7-O-glucósido	0.33	37.61	-1.39	-3.75	No	No	No	No	0.285	5.507
<i>Hellicoside</i>	-1.618	3.55	-2.25	-5.30	No	No	Si	No	0.285	10.15
Beta-hidroxacteósido	-1.619	29.38	-2.04	-4.90	No	No	Si	No	0.285	6.205
Delfinidina	-0.833	77.39	-1.62	-3.52	No	No	No	No	0.286	3.311
Taxifolina	0.924	64.71	-0.73	-3.20	No	No	No	No	0.286	4.688
Eriodictiol	-0.157	78.72	-0.80	-2.47	No	No	No	No	0.513	2.759
Delfinidin-3-O-glucósido	-1.121	32.50	-2.16	-4.45	No	No	Si	No	0.285	7.652
Isoquercitrina	0.242	48.00	-1.69	-4.09	No	No	Si	No	0.285	8.061
Quercetina-7-O-glucósido	0.311	35.85	-1.52	-4.09	No	No	No	No	0.285	6.324
Isoacteósido	-1.391	21.59	-1.94	-4.57	No	No	Si	No	0.285	5.196
Luteolina 7-glucósido	0.248	37.56	-1.56	-3.93	No	No	No	No	0.285	6.342
<i>Suspensaside</i>	-1.857	18.23	-2.40	-4.82	No	No	No	No	0.285	7.145
Criterio	> 0.90	> 30%	< -1.0	< -3.0	No	No	No	No	< -0.5	> -0.3

* Permeabilidad

8.3. Anexo III

En la siguiente tabla se dan los resultados del análisis de las propiedades ADMET usando el servidor web pkCSM de los *hits* entre los compuestos comerciales similares a los fenoles del AOVE para las proteínas DNMT1 y COMT. Solo se muestran los parámetros que se han tenido en cuenta en este estudio, junto con “sustrato de Pgp” que ha sido positiva para todos los fenoles, por lo que no se ha incluido.

Compuesto	Caco-2*	Absorción intestinal	BBB*	SNC*	AMES tox.	hERG I inhibidor	hERG II inhibidor	Hepato-tox.	T.Pyrif. tox.	Minnow tox.
Apigenina 7-O-neohesperidosida	0.372	34.03	-1.79	-4.92	No	No	Si	No	0.285	4.442
Luteolina 7-glucósido	0.245	37.67	-1.66	-4.51	No	No	Si	No	0.285	1.897
Kaempferol 7-O-neohesperidosida	-0.878	27.13	-1.99	-5.12	No	No	Si	No	0.285	5.741
Luteolin-7-O-β-D-glucurónido	0.245	9.02	-1.74	-4.64	No	No	No	No	0.285	1.030
Cloruro de cianidina-3-O-rutinosido	-1.239	32.73	-2.17	-5.03	No	No	Si	No	0.285	7.671
Cloruro de cianidina 3-O-galactósido	-0.858	35.63	-2.00	-4.50	No	No	Si	No	0.285	3.183
Cloruro de cianidina 3-O-glucósido	-0.858	35.63	-2.00	-4.50	No	No	Si	No	0.285	3.183
Apiin	0.376	33.13	-1.80	-5.03	No	No	Si	No	0.285	4.510
Apigenina 7-O-glucurónido	0.348	14.12	-1.47	-4.40	No	No	No	No	0.285	0.478
Apigenina 7-glucósido	0.348	42.77	-1.39	-4.27	No	No	Si	No	0.285	1.345
Procianidina B1	-1.080	59.14	-2.19	-4.08	No	No	Si	No	0.285	7.332
(+)-Catequina	-0.411	66.77	-1.01	-3.36	No	No	No	No	0.605	2.869
(±)-Catequina hidrato	-0.411	66.77	-1.01	-3.36	No	No	No	No	0.605	2.869
(-)-Catequina	-0.411	66.77	-1.01	-3.36	No	No	No	No	0.605	2.869
(-)-Leucofisetinidina	-0.356	72.30	-0.99	-3.37	No	No	No	No	0.440	2.744
Butin	0.955	89.85	-0.91	-2.30	No	No	No	No	0.834	2.285
7-O-metileriodictiol	0.536	78.83	-0.91	-3.21	Si	No	No	No	0.316	2.445
Dihidrotaxifolina	-0.235	63.43	-1.07	-3.60	No	No	No	No	0.287	2.955
Cyrtometinetin	0.318	75.68	-1.03	-3.06	No	No	No	No	0.472	1.911
Criterio	> 0.90	> 30%	< -1.0	< -3.0	No	No	No	No	< -0.5	> -0.3

* Permeabilidad