



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE QUÍMICA

**Desarrollo de modelos estadísticos y moleculares de
saborizantes y potencial efecto de saciedad**

**TESIS
QUE PARA OBTENER EL TÍTULO DE
QUÍMICA DE ALIMENTOS**

**PRESENTA
ARIADNA IRAIS RAMÍREZ HERNÁNDEZ**



MÉXICO, D.F.

2014



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

JURADO ASIGNADO:

PRESIDENTE: HUGO RUBEN CARREÑO ORTIZ

VOCAL: DULCE MARIA GOMEZ ANDRADE

SECRETARIO: KARINA MARTINEZ MAYORGA

1er. SUPLENTE: JOSE PEDRAZA CHAVERRI

2° SUPLENTE: PATRICIA SEVERIANO PEREZ

SITIO DONDE SE DESARROLLÓ EL TEMA: INSTITUTO DE QUÍMICA, UNAM

ASESOR DEL TEMA:

Karina Martínez Mayorga

SUPERVISOR TÉCNICO:

José Luis Medina Franco

SUSTENTANTE (S):

Ariadna Irais Ramírez Hernández

Índice

1. Introducción	3
2. Marco teórico	4
2.1 Química computacional	4
2.2 Quimioinformática	5
a) Descriptores moleculares en una, dos y tres dimensiones	6
b) Bases de datos moleculares	7
c) Similitud molecular	8
d) Análisis multivariable: Análisis de componentes principales y análisis de agrupamiento	9
2.3 Modelado molecular	11
a) Análisis conformacional	12
b) Superposición molecular	13
c) Acoplamiento molecular	14
2.4 Aplicaciones de química computacional y quimioinformática en alimentos	14
2.5 Control del apetito	15
2.6 FEMA GRAS	19
3. Metodología	21
3.1 Materiales	24
3.2 Procedimiento	24
4. Resultados	28
4.1 Estudio de bases de datos moleculares	28
4.2 Fármacos relacionados con el control del apetito	33
4.3 Similitud estructural	34
4.4 Propiedades fisicoquímicas	40
4.5 Similitud entre fármacos	44
5. Análisis de resultados	57
5.1 Estudio de bases de datos moleculares	57
5.2 Similitud estructural	62
5.3 Propiedades fisicoquímicas	63

5.4	Similitud entre fármacos	67
6.	Conclusiones	69
7.	Perspectivas	70
8.	Referencias	72
	Anexos	77
Anexo 1.	Descriptores estructurales	77
Anexo 2.	Descriptores fisicoquímicos	78
Anexo 3.	Coefficiente de Tanimoto	79
Anexo 4.	Distancia euclidiana	79
Anexo 5.	Referencias	79

1. Introducción

La obesidad es un problema de salud pública en nuestra sociedad. Los padecimientos asociados a sobrepeso y obesidad incluyen diabetes de tipo 2, presión arterial alta, síndrome metabólico, entre otras. Numerosos aspectos han sido evaluados y comercializados para el control de peso, incluyendo hábitos y componentes de la dieta. Sin embargo, los fármacos aprobados para el control de peso son escasos, incluso algunos de estos han sido descontinuados debido a efectos secundarios no deseados o falta de eficacia. Esto subraya la necesidad de un mejor entendimiento de los mecanismos biológicos involucrados en el control del apetito o estado de saciedad así como el papel de los alimentos en su regulación.

Los saborizantes son elemento central en la industria de alimentos y bebidas. Estos actúan en los receptores del gusto y del olfato. Los avances en métodos computacionales y en biología estructural están permitiendo un mejor entendimiento de los procesos involucrados en la percepción de aromas y sabores.

Las bases de datos de sustancias contenidas en alimentos pueden ser exploradas para identificar moléculas con posibles efectos en la salud. Estos efectos en salud han cobrado gran atención, tanto en el área académica como en la industria de alimentos.

El objetivo de este proyecto es identificar cuáles son los saborizantes de consumo humano similares a fármacos de uso actual relacionados con saciedad. La hipótesis de la tesis está centrada en el principio de similitud que indica que compuestos semejantes tienen propiedades semejantes.

En este proyecto se identificaron aquellos saborizantes con mayor similitud a fármacos aprobados para el control de peso. Esta información permitió postular la potencial bioactividad de dichos saborizantes.

2. Marco teórico

En este trabajo se utilizaron métodos de cómputo. Para una mejor comprensión del tema, en las primeras secciones de este apartado se describen conceptos básicos de la quimioinformática y del modelado molecular. Posteriormente, se describen estudios específicamente relacionados con el control del apetito y la saciedad que dan lugar al presente trabajo. Información complementaria se presenta en la sección de anexos.

2.1 Química computacional

La química computacional es una rama de la química que se centra en estudiar y resolver problemas de la química utilizando computadoras. La química computacional tiene aplicaciones en múltiples ramas de la ciencia, en particular en química orgánica e inorgánica, espectroscopia, bioquímica, biología, farmacia, química de alimentos, entre otras. Dentro de las distintas definiciones de química computacional destaca la propuesta por Boyd y colaboradores: “Aquellos aspectos de la química que son explicados o realizados mediante computadoras”.¹ La química computacional se extiende más allá de los límites tradicionales que separan la química, la física, la biología y la ciencia de la computación, permitiendo la investigación de átomos, moléculas y macromoléculas. La información derivada por métodos computacionales es particularmente útil cuando la investigación de laboratorio es impráctica (en condiciones de muy alta presión y/o temperatura), muy costosa o imposible (estado de transición inestables experimentalmente). Es importante mencionar que los métodos computacionales no pretenden sustituir a los métodos experimentales. Sin embargo, la información obtenida mediante estos métodos es valiosa, e incluso imprescindible, para complementar los modelos teóricos, así como optimizar o predecir ciertos estudios experimentales. Por lo tanto, los modelos computacionales pueden considerarse una guía y complemento de la experimentación; esto es particularmente relevante

en la búsqueda sistemática de compuestos con actividad biológica. La química computacional comprende un gran número de metodologías y aplicaciones, muchas de ellas están todavía en desarrollo. Las aplicaciones generales incluyen: modelado molecular, diseño molecular asistido por computadora, diseño de síntesis orgánica, construcción y búsquedas en bases de datos moleculares, por mencionar algunos ejemplos.

La química computacional puede dividirse en dos grandes categorías:

+ Mecánica molecular: Ésta aplica las leyes de la física clásica y no considera explícitamente a los electrones.

+ Mecánica cuántica: Se basa en la ecuación de Schrödinger para describir una molécula con un tratamiento directo de la estructura electrónica.

2.2 Quimioinformática

La quimioinformática es una combinación entre la química y las ciencias de la computación que poco a poco se está dando a conocer más allá de los círculos especializados. Hay varias definiciones de quimioinformática, una de ellas es la siguiente: “La mezcla de recursos de información para transformar los datos en información y la información en conocimiento, para los fines previstos de toma de mejores decisiones en el ámbito de identificación de fármacos y su optimización”.²

La aplicación primordial de la quimioinformática ha sido el descubrimiento de fármacos. Sin embargo, las metodologías son generales y recientemente se han comenzado explorar otras aplicaciones, como en el área de los alimentos.³ Una característica de la quimioinformática es que los métodos son generalmente aplicables a un gran número de moléculas,⁴ esto es debido a que la acumulación de una gran cantidad de datos requiere de métodos computacionales para hacer análisis y búsquedas en forma eficiente. La quimioinformática nace ante la necesidad de procesar gran cantidad de información química.

A continuación se detallan aspectos importantes a considerar en la quimioinformática.

a) Descriptores moleculares en una, dos y tres dimensiones

Los descriptores moleculares son el resultado de una lógica y un procedimiento matemático que transforma información química en representaciones simbólicas de una molécula o en números. Estos números o representaciones pueden dar una visión más clara de la interpretación de las propiedades moleculares y es capaz de formar parte en un modelo para la predicción de una propiedad interesante de otras moléculas. El desarrollo de descriptores moleculares es interdisciplinario por naturaleza e implica una mezcla de diferentes teorías. Para la implementación y el uso de descriptores moleculares, es necesario un conocimiento de estadística y quimiometría (disciplina metrológica que aplica conocimientos matemáticos y estadísticos a procesos químicos), además de los conocimientos específicos del problema.

Las moléculas pueden ser representadas utilizando descriptores moleculares en una, dos y tres dimensiones (1D, 2D, 3D). En términos generales, la dimensión es el número de coordenadas necesarias para especificar un punto en el espacio, de manera que un punto está en 1D, un cuadrado en 2D y un cubo en 3D. Las llamadas “huellas digitales moleculares” o “*molecular fingerprints*” son muy utilizadas en quimioinformática ya que permiten el análisis de gran cantidad de datos de manera rápida. Las huellas digitales moleculares son, generalmente, representaciones de estructuras químicas en una y dos dimensiones y constituyen un conjunto de características derivadas de la estructura de una molécula. Las características particulares calculadas a partir de la estructura pueden ser arbitrarias y dependerá de la topología del gráfico químico o incluso de la conformación 3D, en el caso de descriptores 3D. Diferentes representaciones de huellas digitales contienen distintas características moleculares. La idea fundamental es codificar ciertas propiedades directa o indirectamente en la huella

digital y utilizar dicha huella digital para representar a las moléculas. Las comparaciones entre las moléculas se reducen a la comparación de conjuntos de características, midiendo el grado en el que los conjuntos se superponen.

El nivel de complejidad de la representación estructural seleccionada se determina en relación de los objetivos planteados. Mientras más detallada sea la representación habrá menor pérdida de información. Sin embargo, los cálculos realizados con dichas representaciones serán más demandantes computacionalmente. Por lo tanto, es necesario establecer un balance entre rapidez de cálculo y nivel de detalle contenido en la representación estructural. Este balance está determinado por los objetivos del problema a estudiar.

Tanto en la química computacional como en la quimioinformática, es necesario tener representaciones adecuadas de las estructuras, para que estas sean posteriormente estudiadas de manera adecuada.

Las estructuras químicas suelen almacenarse en forma de gráficos moleculares. La teoría de grafos, utiliza no sólo la química, sino otras ciencias como las matemáticas e informática. Un gráfico es una estructura abstracta que contiene nodos conectados por los bordes.

Debido a que el uso de representaciones moleculares simplificadas conlleva a la pérdida de información, es aconsejable el uso de varias representaciones, de tal manera que la información perdida en una representación esté contenida en otra.

b) Bases de datos moleculares

Las bases de datos de compuestos químicos están diseñadas para contener información química. La información estructural, típicamente representada de manera gráfica mediante líneas, es generalmente almacenada con coordenadas atómicas y sus conectividades.

En el caso de bases de datos de muchas moléculas (en el orden de cientos, miles o millones de estructuras), es necesario representar a las moléculas de tal forma

que facilite su almacenamiento, análisis, selección y edición optimizando a su vez el uso de memoria.

Las bases de datos de compuestos químicos varían dependiendo de su origen o del tipo de moléculas que contienen. Respecto a su origen pueden ser bases de datos de compuestos químicos físicamente disponibles o de tipo virtual, es decir, compuestos que no han sido sintetizados pero que no son viables de obtener experimentalmente. En cualquiera de estos dos casos, las bases de datos pueden ser de diferentes tipos, por ejemplo, compuestos derivados de química combinatoria o compendios de moléculas estructuralmente diversas. Por otro lado, un tipo particularmente importante de bases de datos son las que contienen productos naturales. Los productos naturales han sido fuente de compuestos con actividad biológica.⁵ Esto ha llevado al diseño de análogos y de bibliotecas de compuestos mediante química combinatoria inspiradas en productos naturales.

Productos naturales y sintéticos han sido históricamente utilizados como aditivos alimentarios, la comparación de estas moléculas con bases de datos con importancia en la búsqueda de fármacos así como su potencial bioactividad está cobrando interés.^{6,7,8} Diversas iniciativas han contribuido a la creación de bases de datos públicas y comerciales, ejemplo de ellas son FEMA GRAS (*Generally Recognized as Safe*), ChEMBL (*European Molecular Biology Laboratory*), PubChem, PubMed, NCI. Excelentes revisiones pueden encontrarse en la literatura.^{3,7}

c) Similitud molecular

Similitud molecular es un concepto ampliamente utilizado en el descubrimiento y diseño de medicamentos. Su uso se basa en el principio de que las moléculas estructuralmente similares son más propensas a exhibir similares propiedades que las moléculas estructuralmente no similares. Por lo tanto, la búsqueda de moléculas funcionalmente similares se puede lograr mediante la búsqueda para las moléculas estructuralmente similares.⁹ Es necesario mencionar que este

principio de similitud tiene excepciones, como lo son los llamados “*property cliffs*”. Los *property cliffs* son pares de compuestos químicos con alta similitud estructural pero con propiedades y comportamiento (por ejemplo, actividad biológica) muy diferentes entre sí. Estos compuestos similares pueden tener algunas propiedades similares y otras completamente diferentes.¹⁰

La similitud es un concepto abstracto sin una definición única, aparente o categórica. Por lo tanto, es necesario establecer criterios que permitan cuantificar la similitud. Esos criterios se imponen al inicio del estudio y delimitan su alcance. El proceso de búsqueda o análisis de similitud estructural consiste en al menos tres etapas. En la primera etapa se define la representación o representaciones a utilizar, las cuales pueden ser estructurales o de propiedades fisicoquímicas. Se han desarrollado diversas maneras de representar moléculas y para realizar el cálculo de similitud.¹¹ Por ejemplo, las moléculas pueden ser representados por sus formas, mediante la proyección de sus propiedades sobre una superficie; o mediante el cálculo de propiedades moleculares.

La segunda etapa consiste en el cálculo de similitud, los más utilizados en aplicaciones químicas son el coeficiente de Tanimoto, y la distancia euclidiana (Anexo 2 y Anexo 3, respectivamente), el coeficiente de Tanimoto es el más utilizado.⁴

El tercer paso consiste en establecer el valor numérico límite utilizado para considerar que un par de moléculas son similares. Cuando se utiliza más de una representación se emplean métodos de fusión de datos (*data fusion*).¹²

d) Análisis multivariable: Análisis de componentes principales y análisis de agrupamiento

El análisis multivariable es de carácter multidimensional¹³ y es la parte de la estadística y del análisis de datos que estudia, analiza, representa e interpreta los datos que resultan de observar más de una variable estadística sobre una muestra

de datos. Estas variables deben ser homogéneas y correlacionadas, sin que alguna predomine sobre las demás. El análisis de componentes principales y el análisis de agrupamiento son ejemplos de análisis multivariantes.

El Análisis de Componentes Principales (PCA, por sus siglas en inglés) es un método estadístico de síntesis de la información o reducción de dimensiones (número de variables) entre algunos o todos los descriptores. Es decir, ante un conjunto de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible. Los nuevos componentes principales (CP) o factores serán una combinación lineal de las variables originales, y además serán ortogonales entre sí; los componentes principales generalmente se analizan mejor de manera gráfica.⁴

El número de componentes principales máximo que se pueden calcular es igual al número de variables (descriptores) en el conjunto de datos original. Para explicar el 100% de la varianza en los datos es necesario incluir todos los componentes principales. Sin embargo, para facilitar la visualización en un gráfico, se consideran dos o tres componentes principales que expliquen preferentemente 70% o más de la varianza.

El primer componente principal (el que explica la mayor parte de varianza), corresponde con el valor más grande; el segundo componente principal, al segundo valor más grande de la varianza, y así sucesivamente.⁴

Los pasos clave en la clasificación de compuestos utilizando el método de agrupamiento son los siguientes:

1. Generar descriptores para cada compuesto en el conjunto de datos.
2. Calcular la similitud o la distancia entre todos los compuestos en el conjunto de datos
3. Usar un algoritmo de agrupamiento para congrega los compuestos dentro del conjunto de datos.
4. Seleccionar un subconjunto representativo mediante la selección de uno (o más) de los compuestos.⁴

Los resultados de un análisis de agrupamiento pueden contribuir a:

1) la definición formal de un esquema de clasificación tal como una taxonomía para un conjunto de objetos.

2) sugerir modelos estadísticos para describir conjuntos de datos.

3) asignar nuevos datos a las clases para diagnóstico e identificación, etc.

Hay dos tipos fundamentales de métodos de clasificación: jerárquicos y no jerárquicos. En los primeros, la clasificación resultante tiene un número creciente de clases anidadas mientras que en el segundo las clases no son anidadas. Los métodos pueden dividirse en aglomerativos y divisivos. En los primeros se parte de tantas clases como objetos se tengan que clasificar y en pasos sucesivos se obtienen clases de objetos similares, mientras que en los segundos se parte de una única clase formada por todos los objetos que se va dividiendo en clases sucesivamente.¹⁴

2.3 Modelado molecular

El modelado molecular permite generar modelos moleculares tridimensionales, es decir arreglos geométricos de átomos y moléculas con sentido químico. En general se busca que las estructuras resultantes sean de mínima energía y/o con geometrías determinadas. Una vez calculadas, estas geometrías pueden ser utilizadas para calcular propiedades fisicoquímicas de interés.¹⁵

El modelado molecular es una descripción simplificada o idealizada de un sistema molecular o proceso entre moléculas, ideada para facilitar cálculos y predicciones; engloba métodos teóricos y técnicas computacionales para modelar, imitar y predecir el comportamiento de moléculas. Las técnicas y métodos utilizados se encuentran en un amplio rango de campos de la física (termodinámica, mecánica clásica, mecánica estadística, mecánica cuántica, física matemática y ciencia de materiales), la química computacional y la bioquímica. La característica principal

de este tipo de cálculos es la descripción a nivel atómico de los sistemas moleculares; el nivel de información más detallado corresponde al estudio de átomos individuales (o un pequeño grupo de átomos). Los métodos de modelado molecular son usados en la actualidad para investigar la estructura, dinámica y termodinámica de sistemas inorgánicos, biológicos y poliméricos. Por ejemplo, el modelado molecular se ha utilizado para investigar plegamiento proteico, catálisis enzimáticas, estabilidad de proteínas, cambios conformacionales asociados con la función biomolecular, y reconocimiento molecular de proteínas, ADN, y complejos de membranas.

En los métodos de modelado, se consideran las propiedades de las moléculas en 3D y son importantes, entre otros, el análisis conformacional, la mecánica cuántica, los campos de fuerzas y la termodinámica estadística.¹⁵

a) Análisis conformacional

Es la exploración de conformeros de una molécula dada que se pueden obtener al rotar enlaces sencillos (grados de libertad conformacionales), observando los cambios en la energía molecular asociados a esas torsiones.

Para llevar a cabo un análisis conformacional, es necesaria la generación y el estudio de posibles conformaciones que puede adquirir la molécula, así como el cálculo de las energías asociadas. Las conformaciones pueden estar representadas por mínimos locales en la superficie de energía potencial. En el análisis conformacional se toman en cuenta características geométricas y energéticas como son: enlaces, ángulos, torsiones impropias, diedros, interacciones electrostáticas. Los métodos de exploración del espacio conformacional incluyen: búsqueda sistemática, procedimientos de crecimientos de cadena, métodos estocásticos, distancias geométricas, algoritmos genéticos, búsqueda de modos normales de baja frecuencia, y simulaciones de dinámica molecular. Los pasos clave para estos cálculos son:

- 1) Construcción del modelo de la molécula.

- 2) Método elegido para generar las conformaciones.
- 3) Método de evaluación energética.
- 4) Criterio de finalización.

b) Superposición molecular

La superposición molecular se usa frecuentemente para comparar múltiples conformaciones de una misma proteína y para evaluar la calidad de los alineamientos producidos usando sólo información de las secuencias entre dos o más secuencias cuyas estructuras son conocidas. Los métodos de alineamiento operan mediante la búsqueda conformacional y rotación de las moléculas en el espacio en 3D. Por lo tanto, este método depende de la flexibilidad conformacional de las moléculas a ser considerados.¹⁶ El objetivo es determinar la alineación de las estructuras en el que la medida de similitud alcance un máximo, lo cual corresponde a la alineación óptima entre las dos estructuras. La flexibilidad conformacional puede ser tratada de forma directa mediante la generación de varias conformaciones para cada molécula y después explorar todas las posibles combinaciones de conformaciones. De manera alterna, las conformaciones de las moléculas a comparar se pueden variar al mismo tiempo así como las orientaciones relativas, en un intento de identificar las mejores superposiciones.⁴

c) Acoplamiento molecular

En el campo del modelado molecular, el acoplamiento es un método que propone posibles orientaciones y conformaciones preferentes de una molécula a una segunda cuando se unen entre sí para formar un complejo estable. El acoplamiento molecular es la técnica de elección para estudiar los posibles modos de unión entre moléculas pequeñas y macromoléculas o entre dos macromoléculas. El objetivo del acoplamiento molecular automatizado es simular el proceso de reconocimiento molecular. Esta información puede guiar la síntesis de nuevas moléculas con mejor afinidad predicha, así como permitir estudios de relaciones estructura actividad. Los dos pasos generales involucrados en el acoplamiento molecular automatizado son el algoritmo de búsqueda y la función de puntuación. El algoritmo de búsqueda sugiere posibles modos de unión. La función de puntuación parte de una posición del sistema “receptor-ligando” y da como resultado un valor numérico que evalúa dicha interacción. En algunos programas el valor de puntuación corresponde a la energía de unión de dicho sistema. Una función de puntuación debe ser capaz de discriminar las configuraciones correctas de las incorrectas. Existen alrededor de 30 implementaciones para la simulación del acoplamiento molecular, éstas se dividen en Acoplamiento Proteína-Ligando y Acoplamiento Proteína-Proteína.

2.4 Aplicaciones de química computacional y quimioinformática en alimentos

Los alimentos son sistemas complejos que en general están formados por una gran cantidad de componentes. Además, los procesos a los que son sometidos (secado, pasteurización, congelamiento, caramelización, etc.) determinan los componentes finales del alimento. Esto hace que la aplicación de métodos de cómputo esté limitada a componentes individuales, por ejemplo, aditivos alimentarios. Cabe mencionar que numerosos estudios QSAR han sido reportados

en el área de alimentos. Respecto a estudios de modelado molecular, los recientes avances en la cristalización de proteínas de membrana abren ampliamente las posibilidades, por ejemplo, de estudios de los receptores del olfato y gusto y con ello el mejor entendimiento de estos procesos. En analogía con las bases de datos moleculares con aplicaciones para el desarrollo de fármacos, es posible aplicar los métodos de la quimoinformática a bases de datos de alimentos. La base de datos FEMA GRAS es un compendio de saborizantes de consumo humano que pueden ser comparadas con otras bases de datos. Análisis de saborizantes contenidos en esta base de datos ha conducido a la identificación de patrones de descriptores sensoriales que pueden ser utilizados por saboristas. Esta base de datos también ha sido caracterizada y comparada con bases de datos relevantes en la química medicinal.¹⁷ Dicho estudio permitió conocer las diferencias y semejanzas entre las bases de datos analizadas.⁷ Esto muestra la viabilidad e interés actual de explorar la interacción de moléculas contenidas en la base de datos FEMA GRAS con dianas terapéuticas.

2.5 Control del apetito

Los saborizantes son elemento central en la industria de alimentos y bebidas. Estos actúan en los receptores del gusto y del olfato. Los avances en métodos computacionales están permitiendo un mejor entendimiento de los procesos involucrados en la percepción de aromas y sabores.¹⁸ El apetito es un término general que engloba tres conceptos: hambre, satisfacción y saciedad. El hambre es la sensación fisiológica o psicológica que induce a comer; satisfacción es el estado de plenitud que obliga a dejar de comer y saciedad, el período durante el cual la sensación de satisfacción se mantiene hasta que aparece nuevamente el hambre.¹⁹ El sistema fisiológico de regulación del apetito se encuentra en el hipotálamo. Se ha observado que lesiones en las diferentes áreas del hipotálamo afectan el comportamiento alimentario y por ende, la regulación del peso corporal. Así, lesiones bilaterales del hipotálamo ventromedial producen hiperfagia y obesidad. Sin embargo, también existen otras zonas del cerebro implicadas en la

regulación del apetito, sobre las cuales actúan neuropéptidos que se clasifican en neuropéptidos orexígenos y neuropéptidos anorexígenos. Adicionalmente a esta regulación central existen otras sustancias que se originan en la periferia, la mayoría constituidas por péptidos, glucosa y otras sustancias producidas por el metabolismo.²⁰

En el control de la ingesta intervienen una serie de señales que, a corto plazo, informan al cerebro del estado nutricional del organismo. Estas señales llegan al sistema nervioso central (SNC) a través de las innervaciones periféricas o a través de la circulación y acceden a sus receptores en el cerebro.²¹

Cuando el alimento llega al sistema digestivo, se producen y secretan una serie de péptidos saciantes, en cantidad proporcional a la comida que se está ingiriendo, y sinérgicamente con otras señales, como la distensión gástrica, actúan en el cerebro y contribuyen a la finalización de la ingesta.²²

El papel destacado de la colecitoquinina (CCK), como péptido regulador de la ingesta, ha sido muy estudiado; se sabe que la administración exógena de CCK en ratas reduce su consumo de alimento, al igual que en el hombre; se cree que actúa inhibiendo el vaciado gástrico, lo que produce la sensación de saciedad. Además, se han identificado una gran variedad de péptidos saciantes secretados por las células endocrinas gastrointestinales, como el péptido similar al glucagón tipo 1 (GLP-1), la insulina, el glucagón, la leptina o el neurotransmisor no peptídico y la serotonina, que se liberan en respuesta a la llegada de alimentos al estómago. La leptina, de origen gástrico, en sinergia con la CCK, podría estar involucrada en el control de la ingesta a corto plazo.²²

El GLP-1 es una hormona que facilita la liberación de insulina a partir del páncreas, por lo que se ha postulado como una hormona que puede ser útil en el tratamiento de la diabetes tipo 2. Adicionalmente, se ha observado que el GLP-1 tiene otras acciones como la de inhibir la secreción ácida gástrica y el apetito, pudiéndose utilizar para control de la obesidad.²²

La grelina es una hormona liberada principalmente en el estómago, ejerce un papel importante en la regulación de la ingesta. La concentración de grelina

aumenta al iniciar la ingesta y finalizan tras ella, lo que la convierte en una hormona clave en la señalización del comienzo de la ingestión de alimentos.

El péptido YY (PYY) ejerce un papel importante en la regulación de la ingesta, debido a su acción inhibitoria del apetito, disminuyendo la ingesta diaria.²¹

La serotonina es un neurotransmisor que se produce a partir del aminoácido esencial triptófano, el cual regula, entre otros, los efectos de las señales de saciedad a corto plazo.²³

Los factores antes mencionados, a excepción de la grelina (potente estimulante del apetito) inhiben el apetito. Estos factores actúan sobre los neuropéptidos del sistema nervioso central de manera directa sobre las neuronas cerebrales que los producen a través de una estimulación del nervio vago.

Es evidente que el consumo de alimentos es controlado por factores complejos que implican componentes metabólicos. Además, la dieta tiene un papel importante en la modulación de la liberación y la acción de hormonas intestinales.

De los tres macronutrientes, las proteínas se reconocen como el más fuerte inhibidor de la ingesta de alimentos. El efecto de saciedad de las proteínas parece estar mediada por péptidos intestinales, principalmente de CCK, GLP-1 y el PYY. Los aminoácidos tales como leucina, por su grupo amino de cadena ramificada, puede contribuir al control de la ingesta de alimentos a través de mecanismos centrales.²²

Diversos estudios han demostrado la capacidad de las dietas hiperprotéicas para limitar el consumo total de energía, el peso corporal y la deposición de lípidos. Se sabe que el contenido de proteína y aminoácidos de los alimentos es un factor determinante en el control de la ingesta de alimento.²²

Por otro lado, se han realizado estudios donde se demuestra que los olores en los alimentos tienen un efecto saciante en el consumidor. Por ejemplo, la sensación de saciedad aumenta al percibir olores más saturados.²⁴

Como se ha mencionado anteriormente, la ingesta de alimentos está sujeta a una regulación compleja por el hipotálamo y otros centros del cerebro

incluyendo el tronco cerebral y el hipocampo. Una intrincada red de mecanismos de retroalimentación involucran a los centros neurales ya mencionados, junto con el estómago, el intestino, el hígado, la tiroides, y tejido adiposo en la periferia, esto influye en el resultado final de la ingesta de alimentos y la saciedad

Los polifenoles son sustancias naturales, presentes en verduras, frutas, granos, té y vino. Algunos polifenoles actúan como potenciadores de insulina y efectos anti-inflamatorios. Se ha demostrado que algunos alimentos contienen polifenoles con efecto en los neuropéptidos asociados al control de la ingesta de alimentos (insulina, PYY, NPY, grelina, CCK, etc.), como son: canela, curcumina, extracto de semilla de uva (resveratrol), perejil, cebolla y apio (apigenina), aceite de oliva (oleuropeína aglicona), pulpa de algarrobo, té verde, bayas, chocolate y cocoa. Adicionalmente, se han hecho estudios donde se comprueba que al consumir canela y vinagre, el efecto de saciedad aumenta, además reduce la glucosa en sangre postprandial.^{25,26}

Es necesario estudiar con mayor precisión la interacción entre las diferentes áreas del cerebro y los circuitos neuronales tras la ingesta de dietas ricas en proteínas, al igual que al ingerir polifenoles así como las interacciones que hay entre los sistemas del olfato, el gusto y el sistema endócrino.

La regulación del apetito es un proceso complejo. Se puede actuar sobre los múltiples factores que lo afectan, y dichas acciones, como son el control de la ingesta calórica, el control del índice glucémico, la composición de los alimentos, los hábitos y el estilo de vida, deben estar interconectadas.

En la Tabla 1, se resumen algunos ligandos y receptores relacionados con el control del apetito.

Tabla 1. Ligandos relacionados con saciedad.

Ligandos	Familia	Información estructural	Mecanismo de señalización	Ref.
Neuropéptido Y	NPY	Polipéptido 36aa	NPY1R: G _{αi} inhibición AC; NPY2R: G _{αq} PLC activación; NPY4R: G _{αi} inhibición de AC y G _{αq} PLC activación; NPY5R: G _{αi} inhibición de AC	27 26 28 29 30

GLP	Glucagón	Polipéptido 29aa	G _{αS} AC activación; PKA activación	20 21 27 25 29
GLP-1	Glucagón	Polipéptido 30aa	G _{αS} AC activación	20 21 22 26
Insulina	Insulina	Polipéptido 51aa	PI3K activación; MAPcinasa activación	25 26 29 30
Leptina	Citoquina	Polipéptido 167aa	Jak-STAT activación	27 25 26 28 31 23 32
Grelina	Motilina	Polipéptido 28aa	NPY/AGRP activación	20 25 22 26 30 33
CCK	Gastrina	Polipéptido 33aa	CCKA: G _{αq} PLC activación and G _{αS} AC activación; CCKB(2): G _{αq} PLC activación	20 27 22 26 29 23
Serotonina	Monoaminas	Amina aromática, compuesta de un anillo indol hidroxilado en la posición 5, y una cadena lateral etilamínica,		22 29 30 33

2.6 FEMA GRAS

La categoría GRAS fue establecida en 1958 por la Ley Federal de Alimentos, Medicamentos y Cosméticos de Estados Unidos, consiste de productos químicos naturales o artificiales que pueden ser añadidos a los productos de alimentos y bebidas dentro de un nivel definido de uso.³⁴ Para que un compuesto o mezcla de compuestos pueda ser agregado a la lista de compuestos GRAS requiere de una extensa evaluación de seguridad, revisión y aprobación de la FDA.³⁵ Los aditivos inicialmente incluidos en la categoría GRAS fueron reconocidos como tal si habían estado en uso antes del 01 de enero 1958 y si su uso había sido lo suficientemente generalizado para garantizar que el compuesto no causaba alergias o intolerancias metabólicas.³⁶ Nuevos compuestos pueden obtener la aprobación GRAS mediante una revisión de material científico por un panel de expertos calificados mediante la formación y experiencia para determinar la seguridad alimenticia. Cualquier corporación o grupo puede convocar a un panel de expertos para determinar el estado de GRAS y notificar a la FDA. La notificación de la FDA no es requerida por la ley, pero es requerida para los nuevos aditivos GRAS.³⁵

Una fracción significativa de la lista GRAS son revisadas por la asociación de fabricantes de sabores y extractos (FEMA: *Flavor and Extract Manufacturers Association*), este panel de expertos que ha operado desde principios de 1960.³⁶ Los datos que el panel de expertos de FEMA requiere para evaluar un compuesto GRAS incluyen el destino metabólico, dosis repetida, estudio toxicológico y de efectos reproductivos para el nuevo compuesto. La dosificación en animales se lleva a cabo en un nivel elevado para asegurar un margen de seguridad.³⁵ FEMA ha desarrollado un programa innovador que utiliza el concepto GRAS para evaluar la seguridad de las sustancias aromatizantes. El panel de expertos de FEMA ha determinado el estado GRAS de sustancias aromatizantes que resulta en una lista de más de 2.500 ingredientes para su uso en humanos. Este panel revisa periódicamente el estado de las sustancias y da oportunidad para la introducción de nuevas sustancias aromatizantes. El programa FEMA GRAS es el programa de la industria GRAS de más larga duración y de mayor éxito.³⁷ FEMA responde a las actividades legislativas y reglamentarias sobre los niveles federales, estatales e internacionales que afectan a la fabricación, el etiquetado y el uso de aromas y sustancias aromatizantes. FEMA controla regulaciones estadounidenses y extranjeras relacionadas con la fabricación, el uso y el etiquetado de los sabores de todo el mundo. FEMA participa en el desarrollo y revisión de las especificaciones para las sustancias aromatizantes para su publicación en el Food Chemicals Codex y para su uso por las organizaciones internacionales.³⁷ En este trabajo se utilizó la base de datos FEMA GRAS (Flavor -Base Pro © 2010), comercialmente disponible por Leffingwell & Associates (LF).³⁸

3. Metodología

En esta tesis se realizó una búsqueda exhaustiva entre las moléculas de la base de datos FEMA GRAS y algunos fármacos relacionados con el control de peso, utilizando quimoinformática. El procedimiento se dividió en tres partes: estudio de bases de datos, similitud entre moléculas FEMA GRAS y fármacos, y similitud entre fármacos. A continuación se presentan los diagramas de flujo que representan las secciones mencionadas.

En la Figura 1, se muestra la metodología para el estudio de diversas bases de datos. La metodología general para comparación entre fármacos para control de peso y moléculas FEMA GRAS utilizando similitud molecular, se muestra en la Figura 2. La metodología general para la comparación entre fármacos para control de peso utilizando similitud molecular se muestra en la Figura 3.

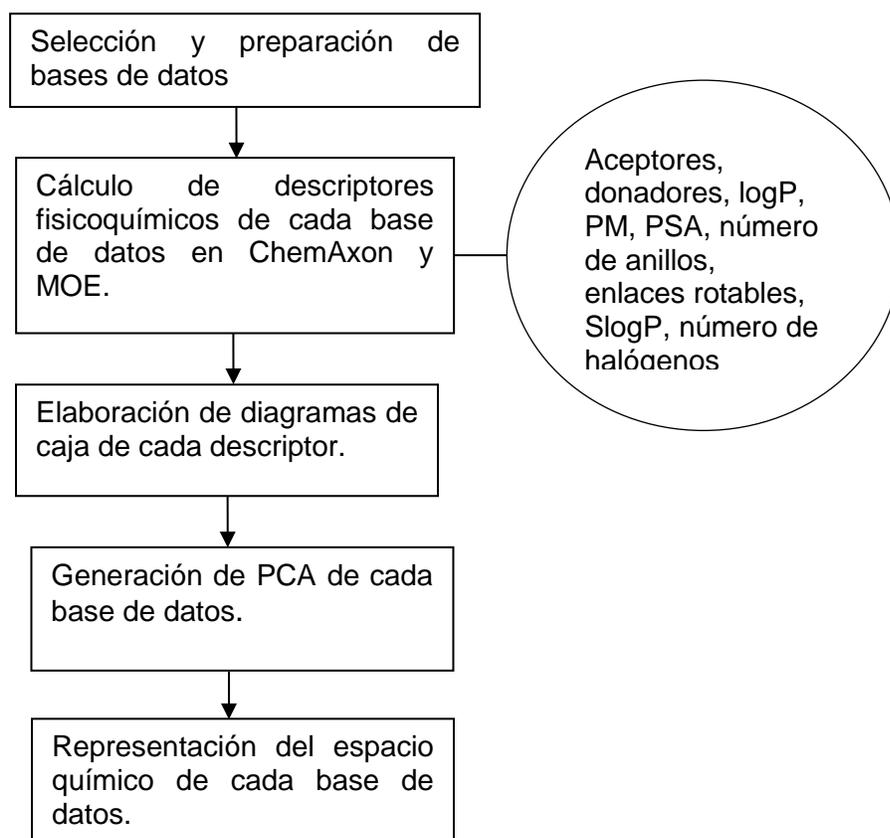


Figura 1. Diagrama de flujo para el estudio de bases de datos.

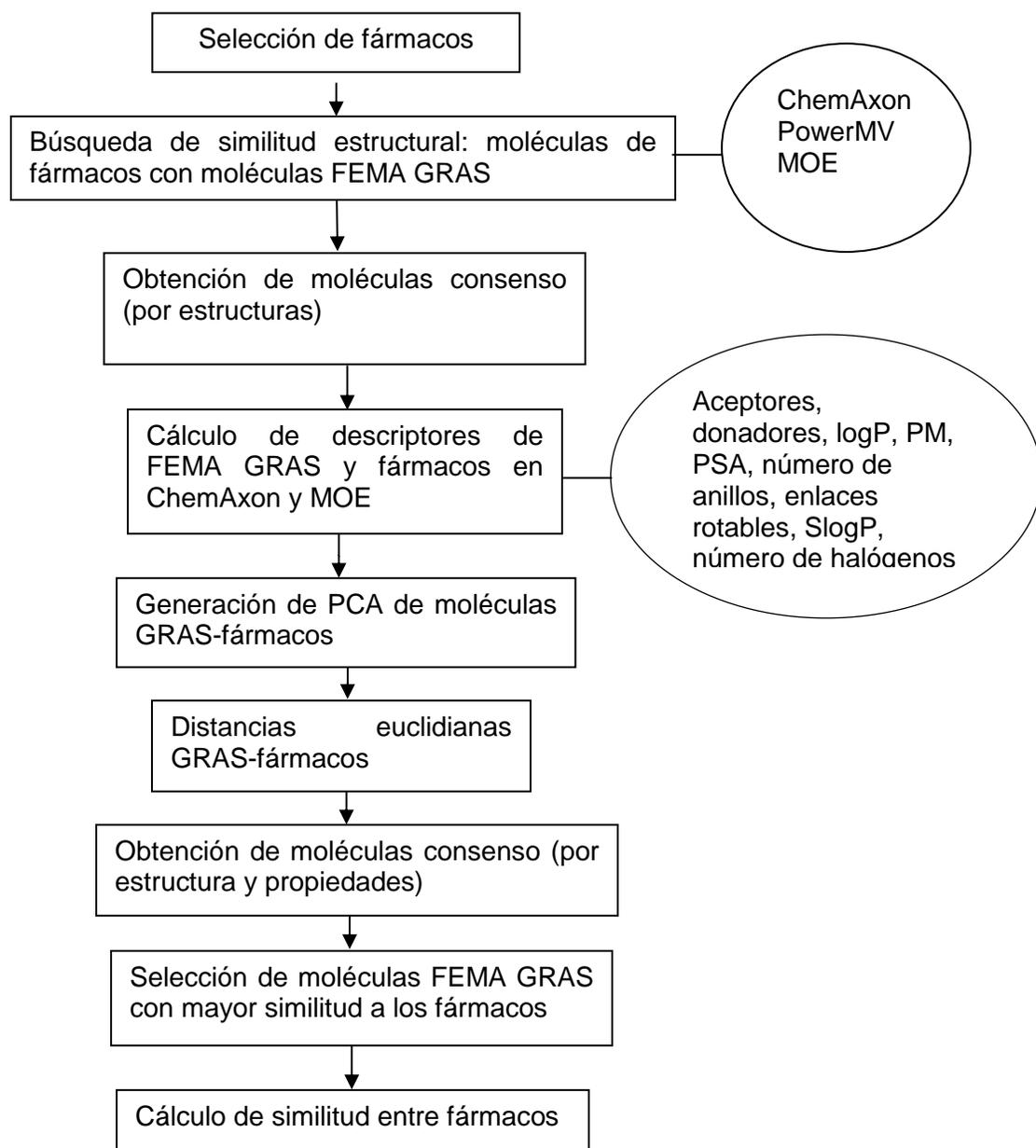


Figura 2. Diagrama de Flujo para la búsqueda de similitud molecular entre FEMA GRAS y fármacos.

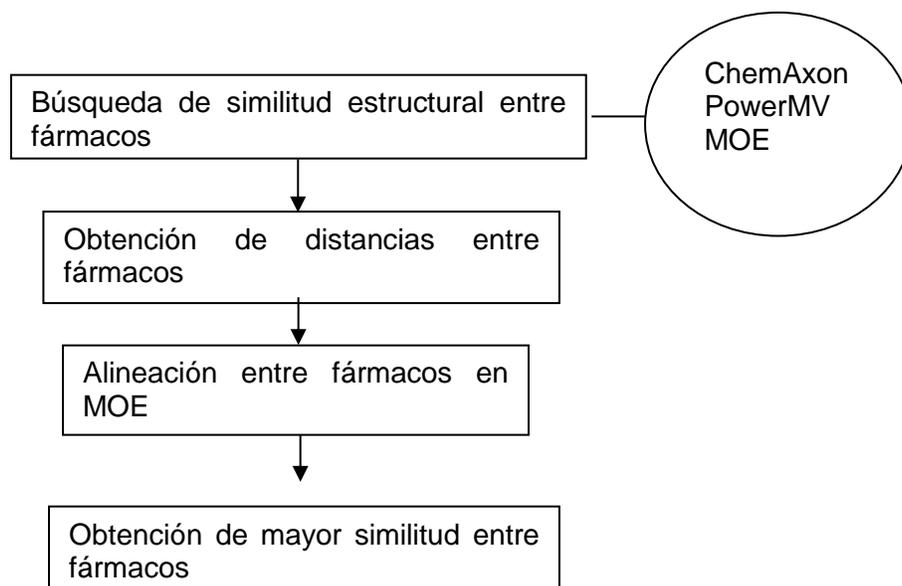


Figura 3. Diagrama de flujo para la búsqueda de similitud molecular entre fármacos para el control de peso.

3.1 Materiales

Este proyecto fue computacional; la obtención, almacenamiento y análisis de datos se realizó en una computadora personal con el sistema operativo Windows 7 Professional. La paquetería de cómputo especializada que se utilizó fue la siguiente:

- ChemAxon (<http://www.chemaxon.com/products/online-tryouts/instant-ichem-via-webstart/>)
- PowerMV (<http://nisl05.niss.org/PowerMV/?q=PowerMV/>)
- Molecular Operating Environment (MOE) (http://www.chemcomp.com/MOE-Molecular_Operating_Environment.htm)
- ChemDraw

3.2 Procedimiento

La descripción detallada de la metodología se muestra a continuación:

❖ Estudio de base de datos moleculares

- Se seleccionaron distintas bases de datos para el estudio.
- Utilizando ChemAxon y MOE, se calcularon descriptores fisicoquímicos para cada base de datos: aceptores y donadores de puentes de hidrógeno, logP, PM, PSA, número de anillos, enlaces rotables. Además, el número de enlaces rotables y el número de halógenos se calcularon en MOE.
- Se elaboraron diagramas de caja de las bases de datos utilizadas para cada descriptor.
- Se realizó el Análisis de Componentes Principales de cada base de datos, utilizando las propiedades calculadas en ChemAxon y MOE.

- Se realizó la representación del espacio químico en dos dimensiones.

❖ Similitud fármacos – GRAS

- Se realizó la búsqueda bibliográfica de fármacos permitidos para su uso clínico o en vías de aprobación.
- Utilizando ChemDraw, se generaron las estructuras 3D de los fármacos seleccionados.
- En los programas ChemAxon, PowerMV y MOE, se buscó la similitud molecular entre los fármacos y las moléculas FEMA GRAS de la siguiente manera:
 - ChemAxon: Se buscó la similitud estructural de los fármacos con las moléculas FEMA GRAS, la similitud se calculó utilizando los descriptores: Substructure, Superstructure, similarity (Tanimoto 0.85 y distancia euclidiana 0.15), full, full fragment*.
 - PowerMV: Se buscó la similitud entre los fármacos y moléculas FEMA GRAS utilizando cinco descriptores: Atom pair, Atom pair Carhart, Fragment Pair, Pharmacophore Fingerprints y Weighted Burden Number*.
 - MOE: Se buscó la similitud entre fármacos y moléculas FEMA GRAS utilizando MACCS keys*, Esta búsqueda se realizó utilizando el coeficiente de Tanimoto, se buscó la similitud más grande posible para cada fármaco. La similitud utilizando el coeficiente de Tanimoto se explica en el Anexo.
- Con los resultados obtenidos se seleccionaron las moléculas FEMA GRAS *consenso*.
- En ChemAxon y MOE se calcularon los mismos descriptores calculados en el estudio de las bases de datos.

- Utilizando los descriptores generados, se realizó el Análisis de Componentes Principales, de las moléculas FEMA GRAS y los fármacos.
- Con los Componentes Principales anteriores, se calculó la distancia euclidiana entre los fármacos y cada una de las moléculas FEMA GRAS. La distancia se calcula de la siguiente manera: $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + \dots + (n_2 - n_1)^2}$
- Se realizó un gráfico para cada programa de los dos primeros componentes principales.
- Basándose en las distancias calculadas y normalizadas, se seleccionaron las moléculas FEMA GRAS con mayor similitud entre fármacos.
- Se seleccionaron las moléculas FEMA GRAS consenso, a partir de las distancias calculadas y las moléculas consenso resultantes.
- Se calculó la similitud entre los fármacos estudiados con ChemAxon, PowerMV y MOE.

❖ Similitud entre fármacos

- Al igual que la búsqueda de similitud entre fármacos y moléculas FEMA GRAS, utilizando los programas ChemAxon, PowerMV y MOE, se buscó la similitud entre fármacos.
 - ChemAxon: Se buscó la similitud estructural entre fármacos, utilizando los descriptores: Substructure, Superstructure, similarity, Tanimoto 0.85, distancia euclidiana 0.15, full, full fragment.
 - PowerMV: Se buscó la similitud entre fármacos utilizando los cinco descriptores: Atom pair, Atom pair Carhart, Fragment Pair, Pharmacophore Fingerprints y Weighted Burden Number.
 - MOE: Se buscó la similitud entre fármacos utilizando MACCS keys.

- Utilizando MOE, se realizaron alineaciones moleculares entre fármacos.
- Con base a las distancias obtenidas y a las alineaciones, se obtuvieron las mayores similitudes entre fármacos.

4. Resultados

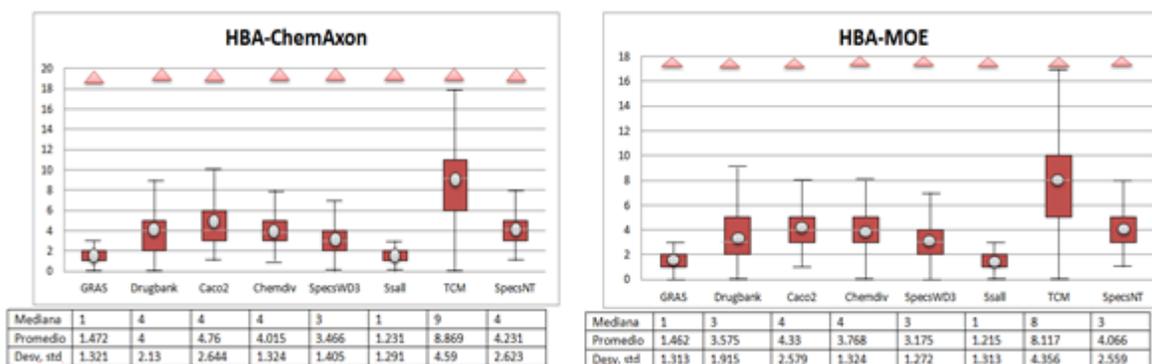
4.1 Estudio de bases de datos moleculares

Se estudiaron propiedades fisicoquímicas relevantes de ocho bases de datos, utilizando dos programas, ChemAxon y MOE. En la siguiente tabla se muestran las bases de datos estudiadas.

Tabla 2. Bases de datos utilizadas en este trabajo.

Base de datos	Contenido	Tamaño
FEMA GRAS	GRAS	2244
DrugBank	Fármacos aprobados	1713
Caco-2	Línea Celular	97
Chemdiv DNMT	Ligandos/inhibidores activos	26944
SpecsWD3	Fármacos aprobados	10000
SuperScent	Sabores y aromas	2116
TCM	Fármacos aprobados	32357
Specs NP	Productos naturales	467

Los resultados de la comparación de la distribución de propiedades fisicoquímicas de las bases de datos utilizando ChemAxon y MOE se muestran en la Figura 4 utilizando diagramas de caja. Se incluyen la mediana, el promedio y la desviación estándar, tanto numéricamente en forma de tabla en cada gráfica, como esquemáticamente en la figura.



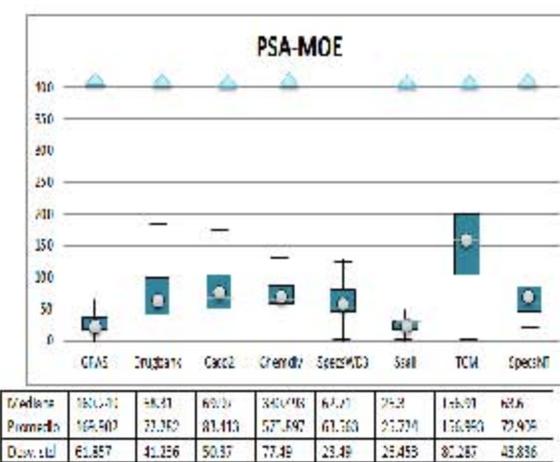
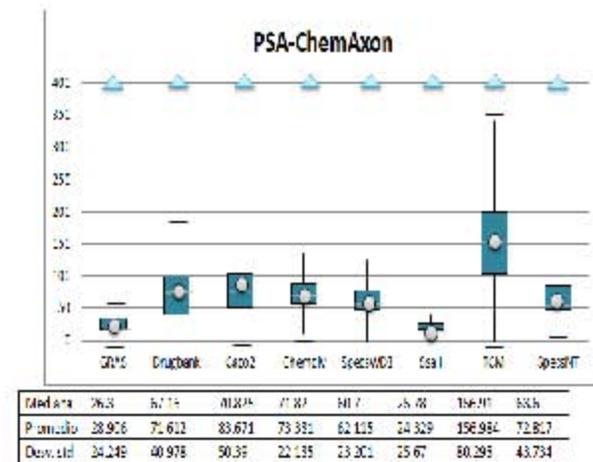
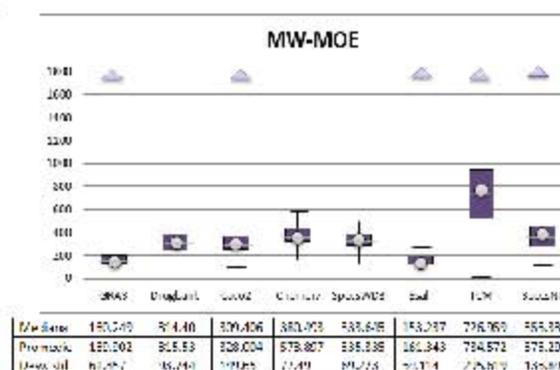
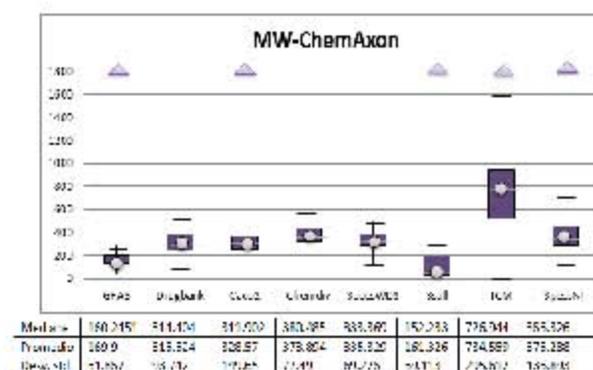
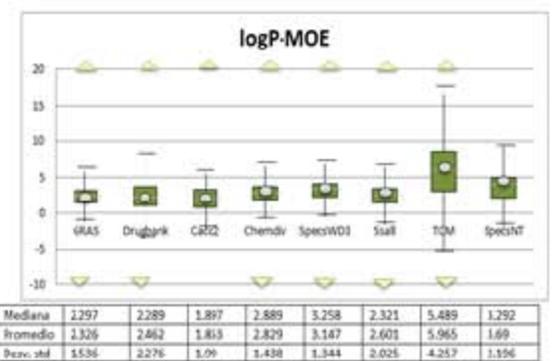
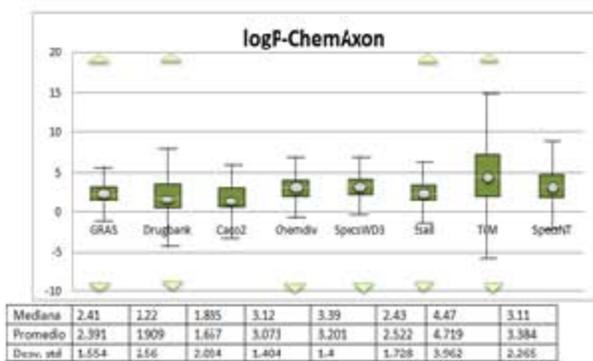
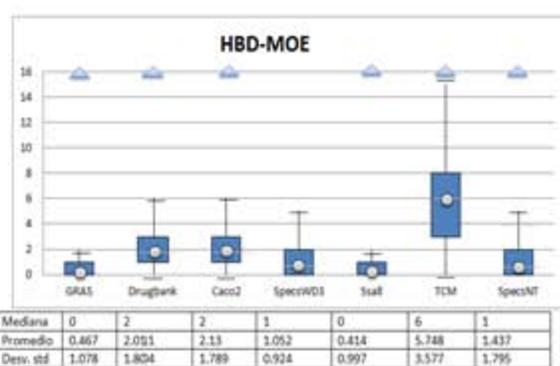
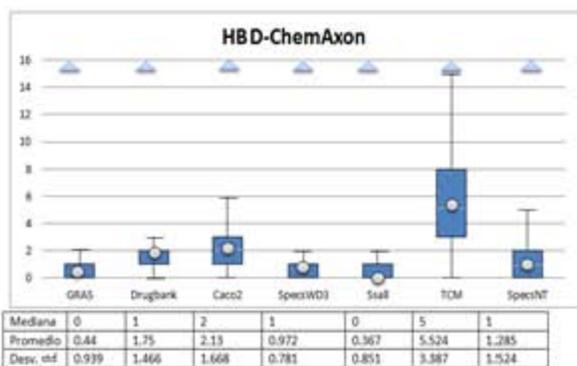
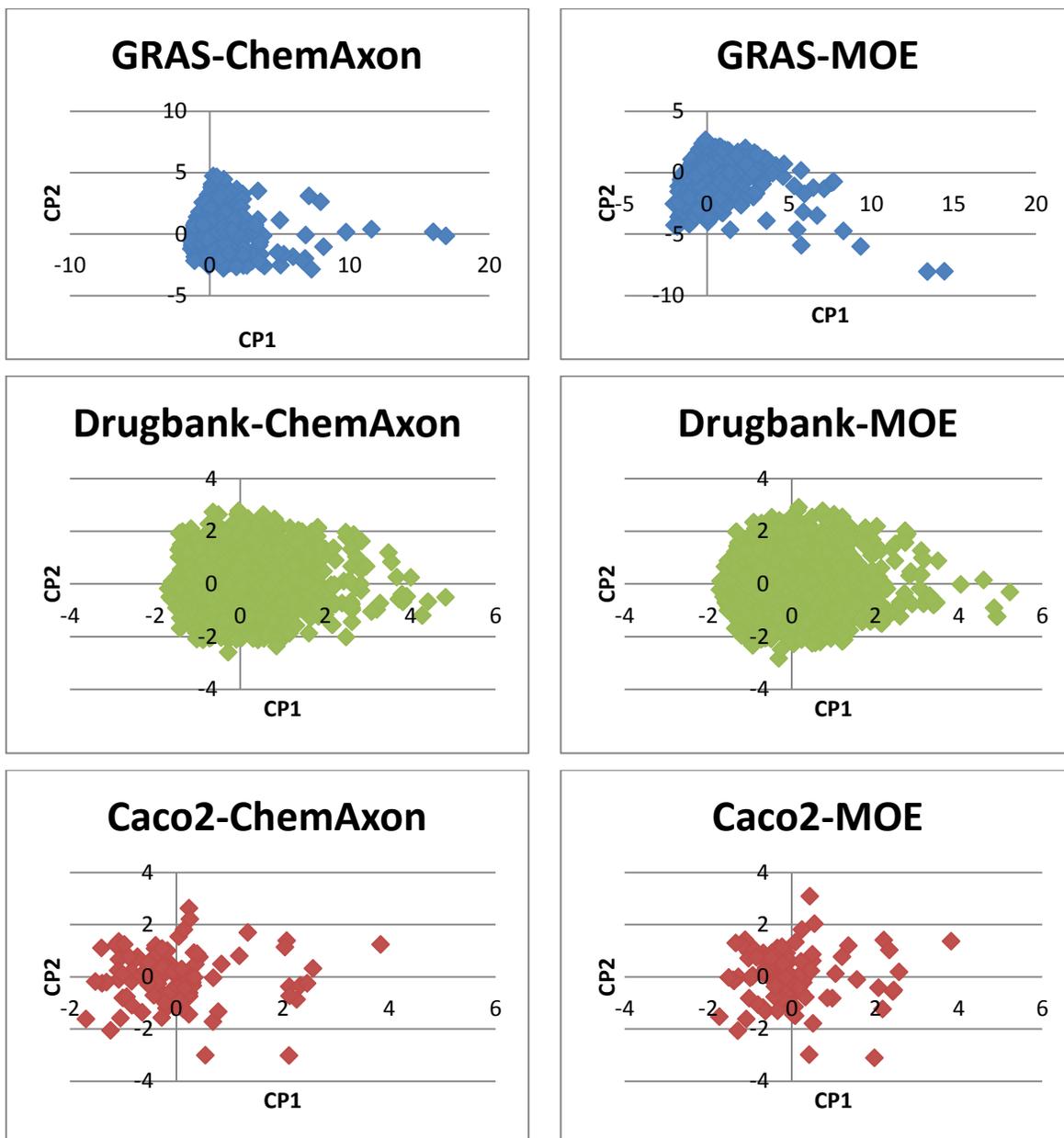


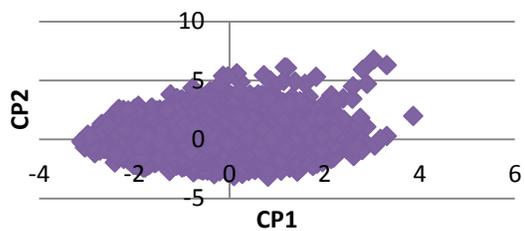


Figura 4. Diagramas de caja con la distribución de siete propiedades fisicoquímicas en las bases de datos. Los recuadros de colores representan el rango intercuartil (medida de extensión); los triángulos indican valores atípicos; las líneas debajo y sobre los recuadros, representan los valores adyacentes superior e inferior; el círculo gris indica la media; y la línea gris, representa la mediana.

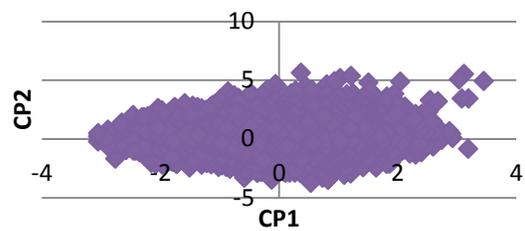
Con los datos obtenidos en el análisis de componentes principales, se realizó la representación del espacio químico de las bases de datos, para ello se utilizaron los dos primeros componentes principales en cada caso.



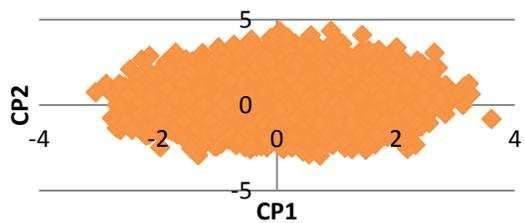
Chemdiv-ChemAxon



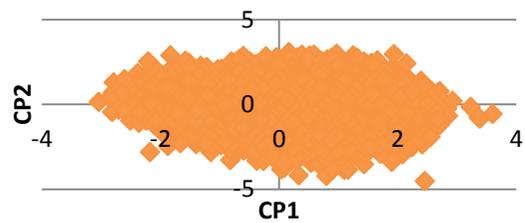
Chemdiv-MOE



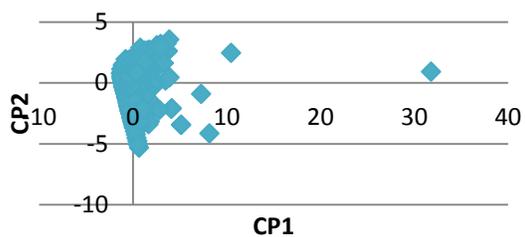
SpecsWD3-ChemAxon



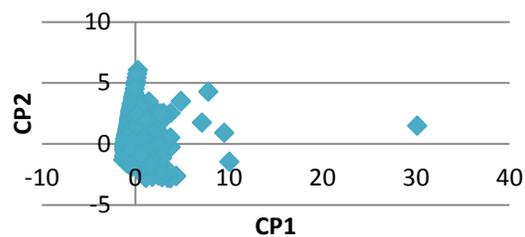
SpecsWD3



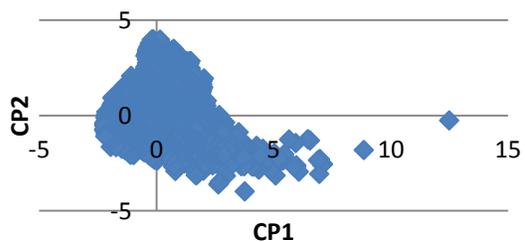
Ssall-ChemAxon



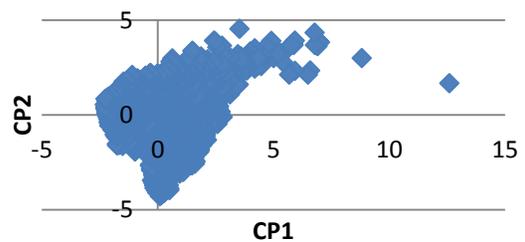
Ssall-MOE



TCM-ChemAxon



TCM-MOE



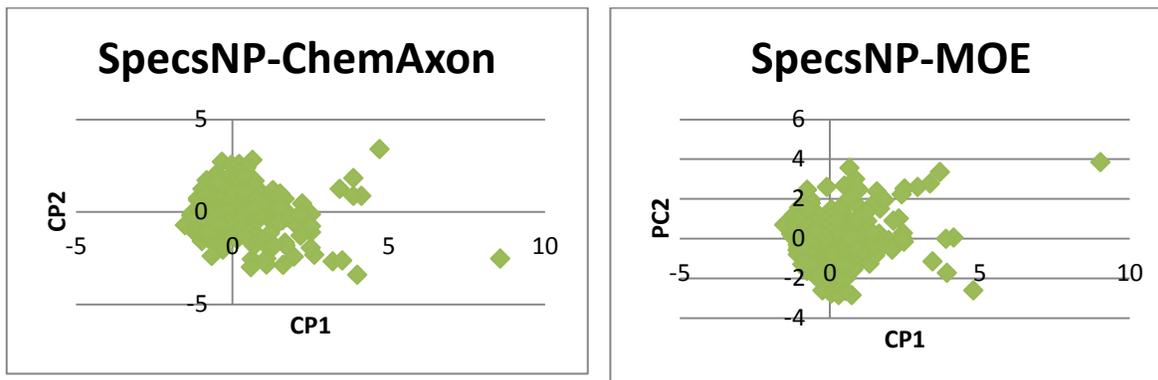
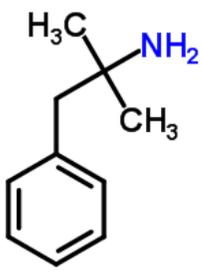
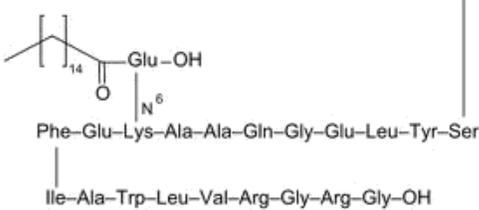


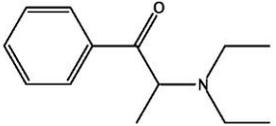
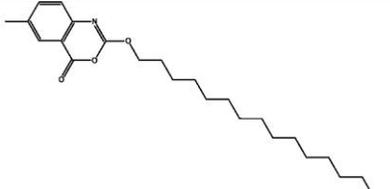
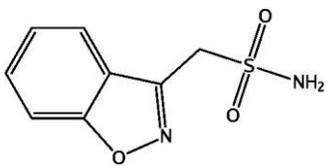
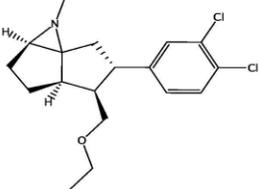
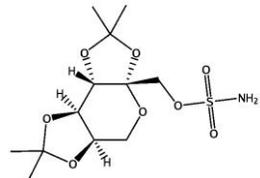
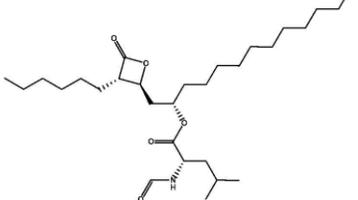
Figura 5. Representación del espacio químico de las distintas bases de datos, generado a partir del análisis de componentes principales calculado en ChemAxon y MOE. En la Tabla 11 se muestran las varianzas obtenidas en cada análisis de componentes principales calculado.

4.2 Fármacos relacionados con el control del apetito

En los últimos años un gran número de fármacos han sido aprobados para el tratamiento de la obesidad. Sin embargo, la mayoría de ellos han sido retirados del mercado debido a sus efectos adversos³⁹. En la actualidad se cuenta únicamente con cinco fármacos permitidos para su uso clínico y tres que están en estudio para su aprobación para el control de la obesidad. En la Tabla 3, se muestran los nombres y estructuras de los fármacos utilizados en este trabajo. La diversidad estructural de los fármacos es un indicativo de la diversidad en sus mecanismos de acción.

Tabla 3. Fármacos utilizados para el control de la obesidad y su situación.³⁹

Permitidos	En estudios clínicos (fase II y fase III)
Fentermina 	Liraglutida H-His-Ala-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser 

Dietilpropión 	Cetilistat 
Zonisamida 	Tesofensina 
Topiramamarato 	
Orlistat 	

4.3 Similitud estructural

Se realizó la búsqueda de estructuras contenidas en la base de datos FEMA GRAS similares a los fármacos de la Tabla 3. Para evitar la dependencia de los resultados con la representación,³ se utilizaron los diferentes descriptores calculados en los programas: ChemAxon, PowerMV y MOE.

En esta búsqueda, los resultados más relevantes fueron las moléculas con estructura privilegiada, aquellas que presentan similitud con más de un fármaco, y las moléculas consenso, aquellas que presentan similitud con un fármaco en más de un método. En la Tabla 4, se muestran los números CAS de las moléculas FEMA GRAS resultantes en esta búsqueda, como moléculas privilegiadas. Mientras que en la Tabla 5, se muestran los números CAS correspondientes a las

moléculas consenso. En ambas tablas se incluye también el programa y el *fingerprint* con el que fueron identificadas.

Tabla 4. Compuestos FEMA GRAS identificados por su número CAS resultantes como moléculas privilegiadas.

CAS	Fármaco	Programa / fingerprint
107-10-8	Fentermina Dietilpropión Orlistat Liraglutida Tesofensina	ChemAxon
19342-01-9	Fentermina Dietilpropión Liraglutida	PowerMV – Atom pair PowerMV – Atom pair Carhart ChemAxon
51-67-2	Fentermina Liraglutida	PowerMV – Atom pair ChemAxon
3777-69-3	Fentermina Dietilpropión	PowerMV – Atom pair PowerMV – Fragment pair
67-63-0	Fentermina Topiramarato Orlistat Liraglutida	PowerMV – Fragment pair ChemAxon
105-57-7	Fentermina Topiramarato	PowerMV – Pharmacophore fingerprints ChemAxon
75-07-0	Fentermina Dietilpropión Orlistat Liraglutida	PowerMV – Pharmacophore fingerprints ChemAxon
7493-57-4	Fentermina Tesofensina	PowerMV – Pharmacophore fingerprints PowerMV – Atom pair
102-76-1	Fentermina Topiramarato	PowerMV – Pharmacophore fingerprints
78-93-3	Dietilpropión Liraglutida	ChemAxon
123-38-6	Dietilpropión Orlistat Liraglutida	ChemAxon
75-50-3	Dietilpropión Tesofensina	ChemAxon
67-64-1	Dietilpropión Liraglutida	ChemAxon
7492-37-7	Dietilpropión Tesofensina	PowerMV – Atom pair Carhart PowerMV – Pharmacophore fingerprints
32974-92-8	Dietilpropión Zonisamida Topiramarato	PowerMV – Pharmacophore fingerprints PowerMV – Atom pair PowerMV – Weighted burden number
645-13-6	Dietilpropión Tesofensina	PowerMV – Weighted burden number
122-00-9	Dietilpropión Tesofensina	PowerMV – Weighted burden number
577-16-2	Dietilpropión Tesofensina	PowerMV – Weighted burden number
7446-09-5	Zonisamida Topiramarato	ChemAxon
7783-06-4	Zonisamida	ChemAxon

	Topiramarato	
23787-80-6	Zonisamida Topiramarato	PowerMV – Atom pair PowerMV – Weighted burden number
22047-25-2	Zonisamida Topiramarato	PowerMV – Atom pair PowerMV – Weighted burden number
847565-09-7	Zonisamida Liraglutida	PowerMV – Atom pair PowerMV – Atom pair Carhart
54300-08-2	Zonisamida Topiramarato	PowerMV – Atom pair PowerMV – Weighted burden number
102-04-5	Zonisamida Tesofensina	PowerMV – Atom pair Carhart PowerMV – Weighted burden number
104-53-0	Zonisamida Liraglutida	PowerMV – Atom pair Carhart ChemAxon
1217341-48-4	Zonisamida Liraglutida	PowerMV – Weighted burden number PowerMV – Atom pair
852379-28-3	Zonisamida Topiramarato Liraglutida	PowerMV – Weighted burden number PowerMV – Atom pair Carhart
1119831-25-2	Zonisamida Liraglutida	PowerMV – Weighted burden number PowerMV – Atom pair
121746-18-7	Zonisamida Liraglutida	PowerMV – Weighted burden number
65416-14-0	Zonisamida Cetilistat	PowerMV – Weighted burden number
71-41-0	Topiramarato Orlistat Liraglutida Cetilistat Tesofensina	ChemAxon
71-36-3	Topiramarato Orlistat Liraglutida Cetilistat Tesofensina	ChemAxon
64-17-5	Topiramarato Orlistat Liraglutida Cetilistat Tesofensina	ChemAxon
111-27-3	Topiramarato Orlistat Cetilistat Tesofensina	ChemAxon
71-23-8	Topiramarato Orlistat Liraglutida Cetilistat Tesofensina	ChemAxon
2679-87-0	Topiramarato Orlistat	ChemAxon
6032-29-7	Topiramarato Orlistat	ChemAxon
64-19-7	Fentermina Orlistat Liraglutida	PowerMV – Pharmocophore fingerprints ChemAxon
123-72-8	Orlistat Liraglutida	ChemAxon
107-92-6	Orlistat	ChemAxon

	Liraglutida	
112-31-2	Orlistat Liraglutida	ChemAxon
64-18-6	Orlistat Liraglutida	ChemAxon
111-71-7	Orlistat Liraglutida	ChemAxon
334-48-5	Orlistat Liraglutida	ChemAxon PowerMV – Fragment pair
111-70-6	Orlistat Cetilistat Tesofensina	ChemAxon
112-30-1	Orlistat Cetilistat Tesofensina	ChemAxon
36653-82-4	Orlistat Cetilistat	ChemAxon
66-25-1	Orlistat Liraglutida	ChemAxon
143-07-7	Orlistat Liraglutida	ChemAxon PowerMV – Fragment pair
97-96-1	Dietilpropión Orlistat	PowerMV – Fragment pair ChemAxon
112-54-9	Orlistat Liraglutida	ChemAxon
112-53-8	Orlistat Cetilistat	ChemAxon
124-25-4	Orlistat Cetilistat	ChemAxon
124-19-6	Orlistat Cetilistat	ChemAxon
79-09-4	Orlistat Cetilistat	ChemAxon
143-08-8	Orlistat Cetilistat Tesofensina	ChemAxon
124-13-0	Orlistat Liraglutida	ChemAxon
111-87-5	Orlistat Cetilistat	ChemAxon
78-83-1	Orlistat Tesofensina	ChemAxon
112-44-7	Orlistat Liraglutida	ChemAxon
104-76-7	Orlistat Tesofensina	ChemAxon
112-42-5	Orlistat Cetilistat	ChemAxon
110-62-3	Orlistat Liraglutida	ChemAxon
107-85-7	Orlistat Liraglutida Tesofensina	ChemAxon
109-52-4	Orlistat Liraglutida	ChemAxon
109-73-9	Orlistat Liraglutida	ChemAxon

	Tesofensina	
56-40-6	Orlistat Liraglutida	ChemAxon
4747-07-3	Orlistat Cetilistat Tesofensina	ChemAxon
10486-19-8	Orlistat Liraglutida	ChemAxon
137-32-6	Orlistat Tesofensina	ChemAxon
78-81-9	Liraglutida Tesofensina	ChemAxon
96-15-1	Liraglutida Tesofensina	ChemAxon
110-58-7	Liraglutida Tesofensina	ChemAxon
111-26-2	Liraglutida Tesofensina	ChemAxon
180964-47-0	Liraglutida Tesofensina	PowerMV – Pharmacophore fingerprints PowerMV – Atom pair Carhart
1323-75-7	Cetilistat Tesofensina	PowerMV – Atom pair PowerMV – Atom pair Carhart
5444-75-7	Cetilistat Tesofensina	PowerMV – Atom pair PowerMV – Atom pair Carhart
1119711-29-3	Dietilpropión Liraglutida	PowerMV – Atom pair PowerMV – Atom pair Carhart
1139-30-6	Topiramarato Tesofensina	PowerMV – Atom pair Carhart PowerMV – Atom pair
3782-00-1	Zonisamida Cetilistat	PowerMV – Pharmacophore fingerprints

Tabla 5. Compuestos FEMA GRAS con estructura consenso.

CAS	Fármaco	Programa
99-87-6	Fentermina	PowerMV – Atom pair PowerMV – Weighted burden number
100-86-7	Fentermina	PowerMV – Atom pair Carhart PowerMV – Fragment pair PowerMV – Weighted burden number
6304-24-1	Fentermina	PowerMV – Atom pair Carhart PowerMV – Fragment pair
14159-61-6	Fentermina	PowerMV – Atom pair Carhart PowerMV – Fragment pair
766-92-7	Fentermina	PowerMV – Atom pair Carhart PowerMV – Weighted burden number
103-05-9	Fentermina	PowerMV – Atom pair Carhart PowerMV – Weighted burden number
121-44-8	Dietilpropión	ChemAxon PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints MOE (Tanimoto 0.65)
26643-92-5	Dietilpropión	PowerMV – Atom pair PowerMV – Weighted burden number
6066-49-5	Dietilpropión	PowerMV – Atom pair Carhart PowerMV – Pharmacophore fingerprints
2315-68-6	Dietilpropión	PowerMV – Atom pair Carhart

		PowerMV – Pharmacophore fingerprints
3194-17-0	Dietilpropión	PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints
4265-16-1	Zonisamida	PowerMV – Atom pair Carhart PowerMV – Fragment pair
83-67-0	Zonisamida	PowerMV – Pharmacophore fingerprints PowerMV – Weighted burden number
67715-82-6	Topiramarato	PowerMV – Atom pair Carhart PowerMV – Fragment pair
38284-11-6	Topiramarato	PowerMV – Fragment pair PowerMV – Weighted burden number
67952-65-2	Topiramarato	PowerMV – Weighted burden number
68127-22-0	Orlistat	ChemAxon (Tanimoto 0.7) PowerMV – Atom pair PowerMV – Atom pair Carhart PowerMV – Weighted burden number MOE (Tanimoto 0.70)
77341-67-4	Orlistat	PowerMV – Atom pair Carhart PowerMV – Weighted burden number
18916-17-1	Liraglutida	PowerMV – Atom pair Carhart PowerMV – Pharmacophore fingerprints
39537-23-0	Liraglutida	PowerMV – Atom pair PowerMV – Atom pair Carhart
714229-20-6	Liraglutida	PowerMV – Atom pair Carhart PowerMV – Pharmacophore fingerprints
122-45-2	Cetilistat	PowerMV – Atom pair PowerMV – Atom pair Carhart PowerMV – Fragment pair
5421-17-0	Cetilistat	PowerMV – Atom pair PowerMV – Atom pair Carhart
7493-78-9	Cetilistat	PowerMV – Atom pair PowerMV – Atom pair Carhart
10024-57-4	Cetilistat	PowerMV – Atom pair Carhart PowerMV – Fragment pair
39251-88-2	Cetilistat	PowerMV – Atom pair Carhart PowerMV – Fragment pair
64577-91-9	Tesofensina	PowerMV – Atom pair PowerMV – Atom pair Carhart
13109-70-1	Tesofensina	PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints
94200-10-9	Tesofensina	PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints
2756-56-1	Tesofensina	PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints
19464-94-9	Tesofensina	PowerMV – Fragment pair PowerMV – Pharmacophore fingerprints

4.4 Propiedades fisicoquímicas

Se calcularon propiedades fisicoquímicas relevantes de la base de datos FEMA GRAS y de los fármacos, posteriormente se hizo el análisis de componentes principales y se generó la representación gráfica del espacio químico. En las Figuras 6 y 7, se muestran estas representaciones del espacio químico generadas en ChemAxon y MOE, respectivamente.

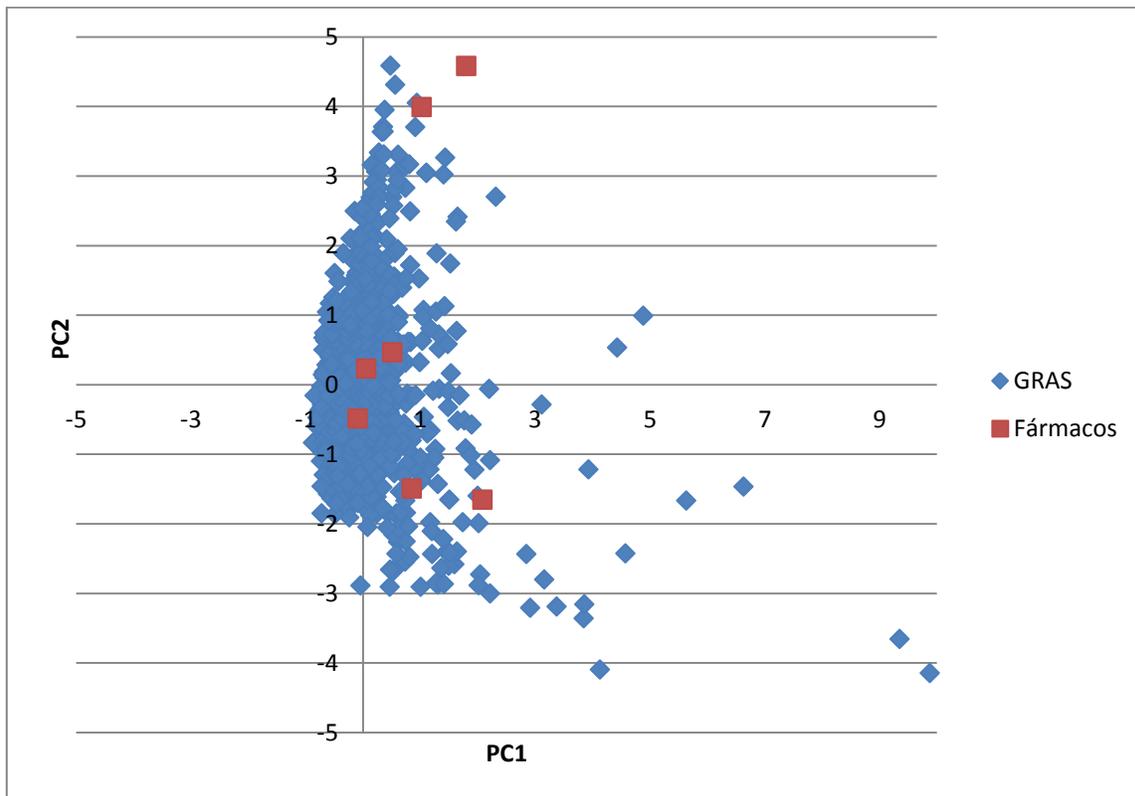


Figura 6. Representación de espacio químico FEMA GRAS – Fármacos, utilizando propiedades fisicoquímicas generadas en ChemAxon.

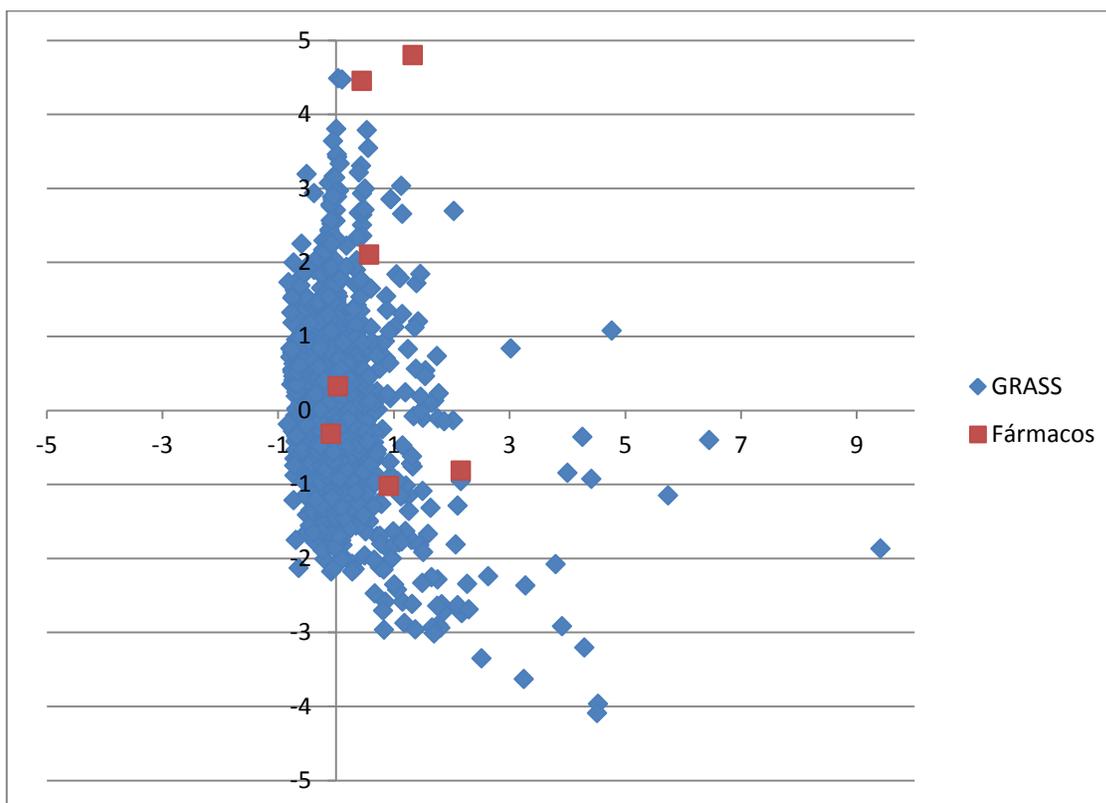


Figura 7. Representación del espacio químico GRAS – Fármacos, utilizando propiedades fisicoquímicas generadas en MOE.

A partir del análisis de componentes principales es posible calcular la distancia euclidiana que hay entre cada fármaco y cada molécula FEMA GRAS, es decir, los “vecinos más cercanos” en el espacio químico FEMA GRAS – fármaco. Cabe mencionar, que para el cálculo de la distancia euclidiana se utilizó el total de dimensiones y no solamente los correspondientes a los dos primeros componentes principales. En la Tabla 6, se muestran las diez distancias menores entre cada fármaco y las moléculas FEMA GRAS.

Tabla 6. Distancias menores (mayor similitud) del análisis de componentes principales Fármacos – FEMA GRAS.

Fármaco	CAS (GRAS)	Distancia (MOE)	CAS (GRAS)	Distancia (ChemAxon)
Fentermina	589-18-4	0.6082	536-50-5	0.6480
	98-85-1	0.6281	536-60-7	0.6886
	17488-65-2	0.6530	598-18-4	0.6891
	100-86-7	0.6656	17488-65-2	0.7003
	1123-85-9	0.6807	98-85-1	0.7364
	103-05-9	0.7270	100-86-7	0.7507
	93-54-9	0.7602	1073-29-6	0.7783

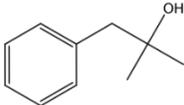
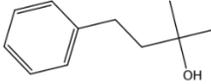
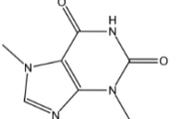
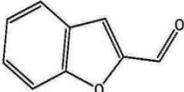
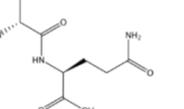
	1073-29-6	0.8001	93-54-9	0.7792
	104-51-1	0.8158	5502-75-0	0.8157
	536-50-5	0.8186	18479-68-0	0.8294
Dietilpropión	90-87-9	0.6651	90-87-9	0.7183
	93-15-2	0.7309	93-15-2	0.8860
	28804-53-7	0.8358	202188-43-0	0.9435
	94098-21-1	1.1162	1670-47-9	0.9864
	101-48-4	1.1179	101-48-4	1.0061
	6314-97-2	1.1270	6380-23-0	1.0206
	122-84-9	1.1277	28804-53-7	1.0379
	104-20-1	1.1928	7784-67-0	1.0392
	6380-23-0	1.2087	93804-64-9	1.1434
	103-60-6	1.2329	68258-95-7	1.1678
Zonisamida	35250-53-4	1.4391	611-13-2	2.0203
	59021-02-2	1.9134	65505-16-0	2.2761
	2044-73-7	1.9243	615-10-1	2.3548
	83-67-0	2.1064	53282-12-5	2.3992
	26486-13-5	2.2608	4208-49-5	2.4030
	623-17-6	2.2925	83-67-0	2.4881
	623-19-8	2.2992	623-19-8	2.5497
	13493-97-5	2.3444	623-17-6	2.5746
	4265-16-1	2.4546	50626-02-3	2.5838
	611-13-2	2.4591	623-22-3	2.5924
Topiramato	18114-49-3	2.7547	18114-49-3	2.7896
	2092-49-1	5.2913	2092-49-1	4.0385
	61683-99-6	5.5599	39156-54-2	4.1512
	180964-47-0	5.6436	61683-99-6	4.6070
	930587-76-1	5.7122	123-63-7	4.6487
	134-96-3	5.7562	180964-47-0	4.6671
	6290-17-1	5.7689	63253-24-7	4.6733
	123-63-7	5.7719	25773-40-4	4.7724
	57197-36-1	5.7979	68527-74-2	4.7961
	6413-10-1	5.9348	36871-78-0	4.8107
Orlistat	123-79-5	2.3996	123-79-5	1.9622
	123-95-5	2.9137	929222-96-8	2.3217
	142-75-6	3.0073	7384-98-0	2.4220
	111-61-5	3.0141	111-62-6	2.4593
	35274-05-6	3.0162	693-80-1	2.4785
	109-43-3	3.0838	111-61-5	2.4870
	502-21-1	3.1096	123-95-5	2.5366
	1227-51-6	3.1599	628-97-7	2.6505
	628-97-7	3.1780	109-43-3	2.6592
	7786-48-3	3.2062	502-26-6	2.7241
Liraglutida	58543-16-1	37.9971	58543-16-1	38.0441
	63550-99-2	38.4827	63550-99-2	38.3615
	791807-20-0	38.8295	791807-20-0	38.3757
	26446-38-8	39.1857	38837-70-6	38.7057
	686298-93-1	39.2008	714229-20-6	39.0640
	714229-20-6	39.5323	686298-93-1	39.1185
	38837-70-6	39.6829	39537-23-0	39.1712
	6138-23-4	39.8284	26446-38-8	39.2751
	5550-12-9	40.0636	14486-03-4	39.7530
	39537-23-0	40.0855	6138-23-4	39.7552
Cetilistat	1227-51-6	2.6323	929222-96-8	3.3814
	502-26-1	2.8139	10024-57-4	3.3831
	10024-57-4	2.8505	1227-51-6	3.5612

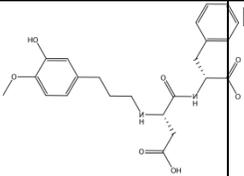
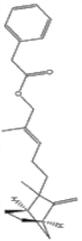
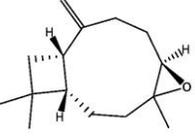
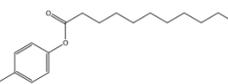
	123-95-5	3.1174	27113-22-0	3.5833
	7370-44-7	3.1386	57-11-4	3.6216
	111-61-5	3.2015	85392-05-8	3.6330
	628-97-7	3.3960	85392-06-9	3.6330
	123-79-5	3.4242	502-26-1	3.6480
	142-75-6	3.4385	112-80-1	3.6526
	7786-48-3	3.4897	123-95-5	3.7330
Tesofensina	67-03-8	19.9530	38325-25-6	2.9873
	1115-84-0	20.8938	94-62-2	3.4201
	2294-76-0	38.8046	11655-60-3	3.8354
	5421-17-0	38.8131	1323-75-7	4.0270
	10032-08-3	38.8138	3738-00-9	4.0527
	6789-88-4	38.8160	150-60-7	4.1667
	538-65-8	38.8165	101780-73-8	4.2253
	65505-25-1	38.8166	120-11-6	4.2543
	470-67-7	38.8169	103-50-4	4.2695
	28217-92-7	38.8170	1139-30-6	4.2915

Los números CAS en negritas en la tabla anterior, son *moléculas consenso* resultantes con las menores distancias en ChemAxon y en MOE.

Es posible conjuntar las moléculas FEMA GRAS que resultaron más similares en el análisis de similitud estructural y mediante la búsqueda de vecinos cercanos en el análisis de componentes principales basado en propiedades fisicoquímicas. El resultado de la unión de estos análisis nos lleva a las moléculas resumidas en la Tabla 7.

Tabla 7. Moléculas FEMA GRAS consenso, similitud estructural y del análisis de componentes principales

CAS	Nombre	Estructura	Fármaco	Método	Distancia
100-86-7	2-Metil-1-fenilpropan-2-ol		Fentermina	PMV – Carhart PMV – Fragment PMV– Weighted PCA ChemAxon PCA MOE	0.39 0.43 0.41 0.75 0.67
103-05-9	2-Metil-4-fenil-2-butanol		Fentermina	PMV –Carhart PMV – Weighted PCA MOE	0.59 0.41 0.73
83-67-0	Teobromina		Zonisamida	PMV – Pharmoco PMV – Weighted PCA ChemAxon PCA MOE	0.00 1.21 2.49 2.11
4265-16-1	1-benzofuran-2-carbaldehido		Zonisamida	PMV –Carhart PMV – Pharmoco PCA MOE	0.68 0.83 2.45
39537-23-0	L-Alanil-L-glutamina		Liraglutida	PMV – Atom pair PMV – Carhart PCA ChemAxon PCA MOE	0.37 0.63 39.19 40.08

714229-20-6	Advantame		Liraglutida	PMV –Carhart PMV – Pharmoco ChemAxon MOE PCA ChemAxon PCA MOE	0.48 0.71 0.5 0.57 39.06 39.53
1323-75-7	2-Eetil-5-(2-metil-3-metilo nebiciclo [2.2.1] hept-2-il)-2-fenilacetato enten-1-il		Tesofensina	PMV –Carhart PCA ChemAxon	0.66 4.027
1139-30-6	4,5-Epoxi-4,11,11-trimetil-8-metilenebiciclo(7.2.0)undecano		Tesofensina	PMV – Atom pair PCA ChemAxon	0.49 4.29
10024-57-4	4-Metilfenil laurato		Cetilistat	PMV –Carhart PMV – Fragment PCA ChemAxon PCA MOE	0.43 0.38 3.38 2.85

4.5 Similitud entre fármacos

Se calculó la similitud estructural entre fármacos con los métodos anteriormente utilizados, esto se hizo con el fin de hacer una comparación entre fármacos y de esa manera explorar si existe alguna relación entre ellos y las moléculas FEMA GRAS similares a éstos. En ChemAxon, no se encontró similitud entre fármacos para ninguno de los cinco descriptores.

En las Tablas 8 y 9, se muestran los resultados obtenidos en los programas PowerMV y MOE, respectivamente, en la búsqueda por similitud entre fármacos.

Tabla 8. Similitud estructural entre fármacos utilizando PowerMV y sus descriptores.

Fármaco	Atom Pair	Atom Pair carhart	Fragment Pair	Pharmocophere Fingerprints	Weighted Burden Number
Fentermina	Dietilpropión	Zonisamida	Dietilpropión	Cetilistat	Dietilpropión
	0.45	0.79	0.84	1.0	1.20
	Tesofensina	Dietilpropión	Zonisamida	Dietilpropión	Tesofensina
	0.62	0.85	0.92	1.0	1.31
	Zonisamida	Liraglutida	Liraglutida	Liraglutida	Orlistat
	0.68	0.89	0.93	1.0	2.61
	Cetilistat	Tesofensina	Topiramarato	Orlistat	Topiramarato
0.79	0.90	0.96	1.0	2.86	
Orlistat	Cetilistat	Cetilistat	Cetilistat	Tesofensina	Zonisamida
0.79	0.91	0.97	0.97	1.0	2.92

	Liraglutida 0.81 Topiramato 0.97	Orlistat 0.96 Topiramato 0.97	Orlistat 0.98 Tesofensina 1.00	Topiramato 1.0 Zonisamida 1.0	Cetilistat 3.50 Liraglutida 6.25
Dietilpropión	Fentermina 0.45 Orlistat 0.63 Zonisamida 0.64 Cetilistat 0.65 Liraglutida 0.67 Tesofensina 0.67 Topiramato 0.90	Cetilistat 0.70 Tesofensina 0.80 Liraglutida 0.82 Orlistat 0.83 Fentermina 0.85 Zonisamida 0.87 Topiramato 0.99	Fentermina 0.84 Cetilistat 0.88 Tesofensina 0.92 Liraglutida 0.94 Orlistat 0.95 Zonisamida 0.95 Topiramato 1.00	Liraglutida 0.90 Tesofensina 0.95 Cetilistat 1.00 Orlistat 1.00 Fentermina 1.00 Topiramato 1.00 Zonisamida 1.00	Tesofensina 0.55 Fentermina 1.20 Orlistat 1.76 Zonisamida 1.90 Topiramato 2.05 Cetilistat 2.42 Liraglutida 5.49
Zonisamida	Dietilpropión 0.64 Fentermina 0.68 Liraglutida 0.74 Cetilistat 0.74 Tesofensina 0.82 Orlistat 0.86 Topiramato 0.95	Fentermina 0.79 Cetilistat 0.80 Dietilpropión 0.87 Liraglutida 0.87 Tesofensina 0.91 Orlistat 0.94 Topiramato 0.97	Fentermina 0.92 Dietilpropión 0.95 Liraglutida 0.97 Cetilistat 0.97 Tesofensina 0.98 Orlistat 1.00 Topiramato 1.00	Cetilistat 1.00 Liraglutida 1.00 Dietilpropión 1.00 Orlistat 1.00 Fentermina 1.00 Tesofensina 1.00 Topiramato 1.00	Cetilistat 1.23 Topiramato 1.69 Dietilpropión 1.90 Tesofensina 2.08 Orlistat 2.44 Fentermina 2.92 Liraglutida 4.27
Topiramato	Orlistat 0.62 Liraglutida 0.76 Tesofensina 0.77 Cetilistat 0.82 Dietilpropión 0.90 Zonisamida 0.95 Fentermina 0.97	Orlistat 0.81 Tesofensina 0.83 Liraglutida 0.92 Cetilistat 0.93 Zonisamida 0.97 Fentermina 0.97 Dietilpropión 0.99	Tesofensina 0.85 Orlistat 0.92 Fentermina 0.96 Liraglutida 0.99 Cetilistat 1.00 Dietilpropión 1.00 Zonisamida 1.00	Orlistat 0.40 Liraglutida 0.95 Cetilistat 1.00 Dietilpropión 1.00 Fentermina 1.00 Tesofensina 1.00 Zonisamida 1.00	Zonisamida 1.69 Orlistat 1.87 Dietilpropión 2.05 Tesofensina 2.36 Cetilistat 2.38 Fentermina 2.86 Liraglutida 5.08
Orlistat	Liraglutida 0.50 Tesofensina 0.57 Cetilistat 0.61 Topiramato 0.62 Dietilpropión 0.63 Fentermina 0.79	Liraglutida 0.67 Cetilistat 0.75 Tesofensina 0.75 Topiramato 0.81 Dietilpropión 0.83 Fentermina 0.96	Cetilistat 0.66 Liraglutida 0.79 Tesofensina 0.85 Topiramato 0.92 Dietilpropión 0.95 Fentermina 0.98	Topiramato 0.40 Liraglutida 0.89 Cetilistat 1.00 Dietilpropión 1.00 Fentermina 1.00 Tesofensina 1.00 Zonisamida	Dietilpropión 1.76 Topiramato 1.87 Tesofensina 1.94 Zonisamida 2.44 Cetilistat 2.61 Fentermina 2.61

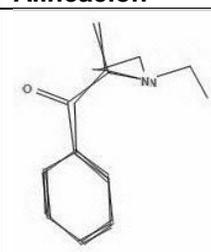
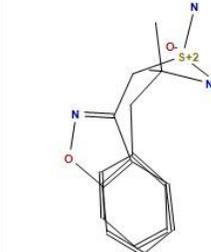
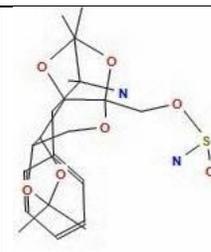
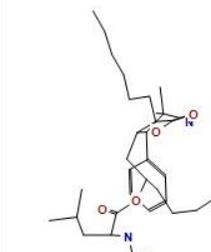
	Zonisamida 0.86	Zonisamida 0.96	Zonisamida 1.00	1.00	Liraglutida 5.70
Liraglutida	Orlistat 0.50 Cetilistat 0.53 Tesofensina 0.66 Dietilpropión 0.67 Zonisamida 0.74 Topiramarato 0.76 Fentermina 0.81	Orlistat 0.67 Tesofensina 0.78 Cetilistat 0.81 Dietilpropión 0.82 Zonisamida 0.87 Fentermina 0.89 Topiramarato 0.92	Orlistat 0.79 Cetilistat 0.82 Fentermina 0.93 Dietilpropión 0.94 Tesofensina 0.94 Zonisamida 0.97 Topiramarato 0.99	Orlistat 0.89 Dietilpropión 0.90 Cetilistat 0.94 Zonisamida 0.94 Topiramarato 0.95 Tesofensina 0.97 Fentermina 1.00	Cetilistat 3.69 Zonisamida 4.27 Topiramarato 5.08 Dietilpropión 5.49 Tesofensina 5.54 Orlistat 5.70 Fentermina 6.25
Cetilistat	Liraglutida 0.53 Orlistat 0.61 Tesofensina 0.62 Dietilpropión 0.65 Zonisamida 0.74 Fentermina 0.79 Topiramarato 0.82	Orlistat 0.75 Dietilpropión 0.77 Tesofensina 0.77 Zonisamida 0.80 Liraglutida 0.81 Fentermina 0.91 Topiramarato 0.93	Orlistat 0.66 Liraglutida 0.82 Tesofensina 0.87 Dietilpropión 0.88 Zonisamida 0.97 Fentermina 0.97 Topiramarato 1.00	Zonisamida 0.00 Liraglutida 0.94 Dietilpropión 1.00 Orlistat 1.00 Fentermina 1.00 Tesofensina 1.00 Topiramarato 1.00	Zonisamida 1.23 Topiramarato 2.38 Dietilpropión 2.42 Tesofensina 2.55 Orlistat 2.61 Fentermina 3.50 Liraglutida 3.69
Tesofensina	Orlistat 0.57 Cetilistat 0.62 Fentermina 0.62 Liraglutida 0.66 Dietilpropión 0.67 Topiramarato 0.77 Zonisamida 0.82	Orlistat 0.75 Cetilistat 0.77 Liraglutida 0.78 Dietilpropión 0.80 Topiramarato 0.83 Fentermina 0.90 Zonisamida 0.91	Orlistat 0.85 Topiramarato 0.85 Cetilistat 0.87 Dietilpropión 0.92 Liraglutida 0.94 Zonisamida 0.98 Fentermina 1.00	Dietilpropión 0.95 Liraglutida 0.97 Cetilistat 1.00 Orlistat 1.00 Fentermina 1.00 Topiramarato 1.00 Zonisamida 1.00	Dietilpropión 0.55 Fentermina 1.31 Orlistat 1.94 Zonisamida 2.08 Topiramarato 2.36 Cetilistat 2.55 Liraglutida 5.54

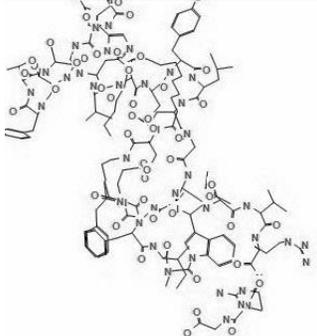
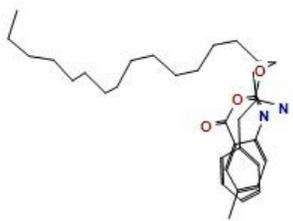
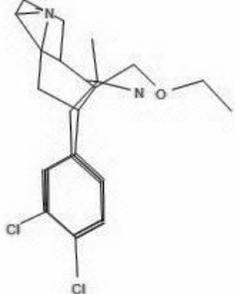
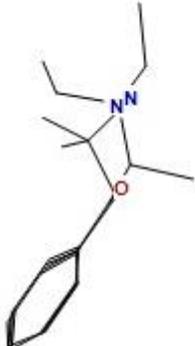
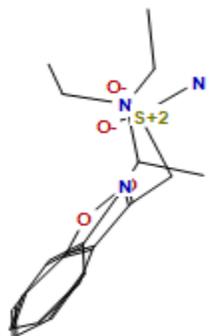
Tabla 9. Similitud estructural entre fármacos utilizando MOE.

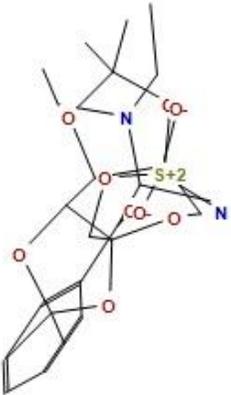
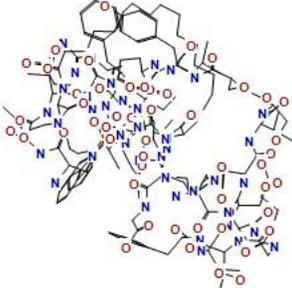
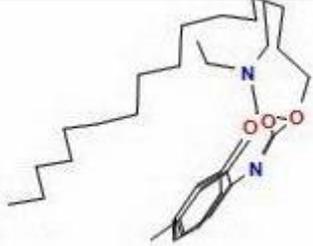
Fármaco	Similitud MOE	Fármaco	Similitud MOE
Fentermina	Dietilpropión (Tanimoto 0.29)	Orlistat	Cetilistat (Tanimoto 0.49)
Dietilpropión	Tesofensina (Tanimoto 0.35)	Liraglutida	Cetilistat (Tanimoto 0.49)
Zonisamida	Topiramarato (Tanimoto 0.54)	Cetilistat	Liraglutida Orlistat (Tanimoto 0.49)
Topiramarato	Zonisamida (Tanimoto 0.54)	Tesofensina	Cetilistat (Tanimoto 0.38)

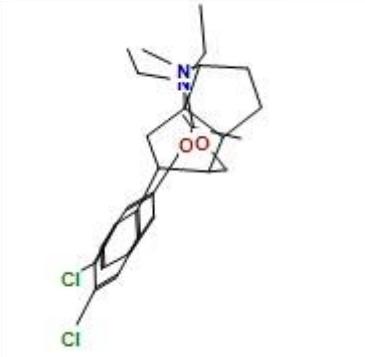
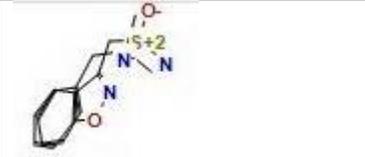
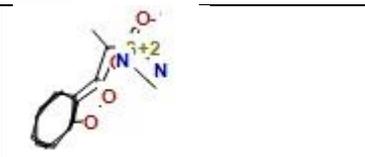
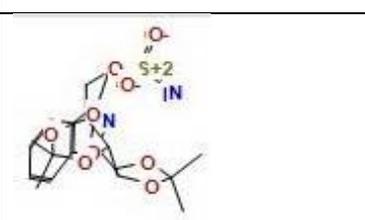
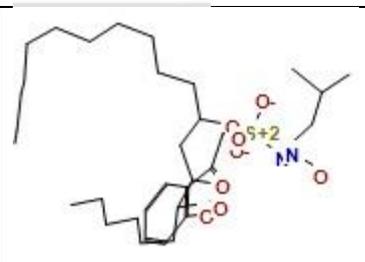
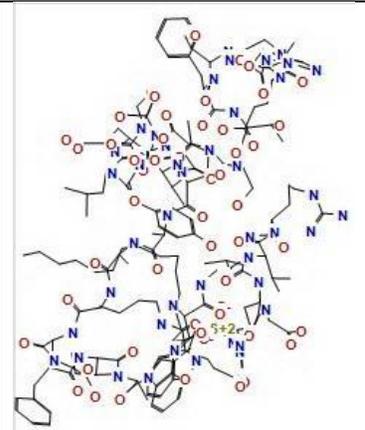
Posteriormente se realizó una superposición entre fármacos con el fin de identificar la mejor superposición entre cada fármaco. Al realizar la superposición molecular se utilizaron las moléculas en 3D. En la Tabla 10, se muestra la mejor alineación para cada fármaco; esto se sabe por los valores de "U" la cual es la energía media de la tensión de las moléculas en la alineación (kcal / mol), "F" que es a medida de similitud de la configuración, y "S" la cual es la suma de las columnas de U y F. Los valores más bajos se indican mejores alineaciones

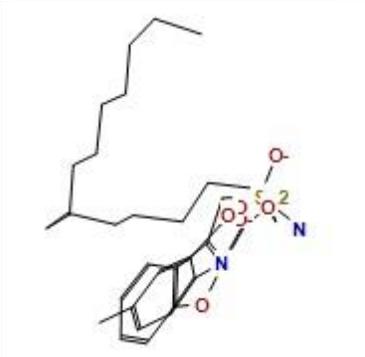
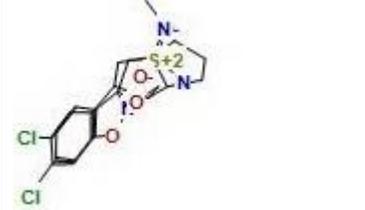
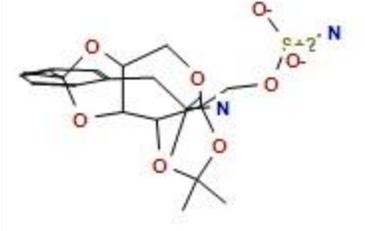
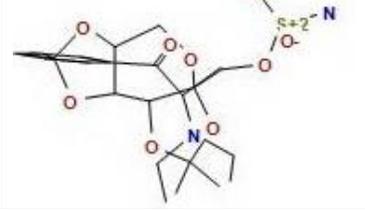
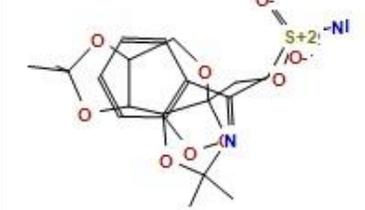
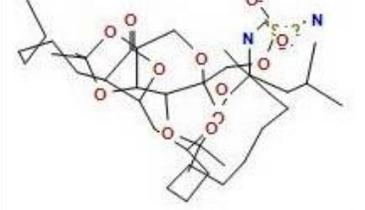
Tabla 10. Superposición entre fármacos.

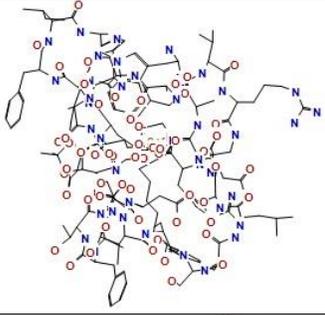
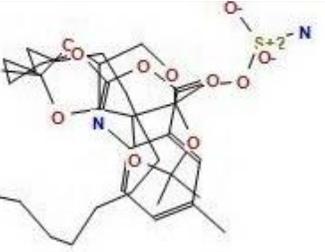
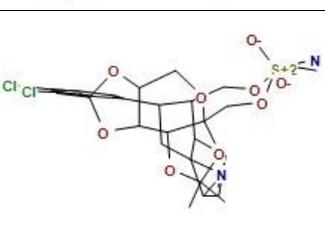
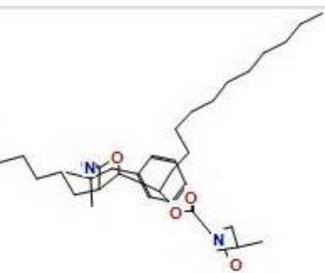
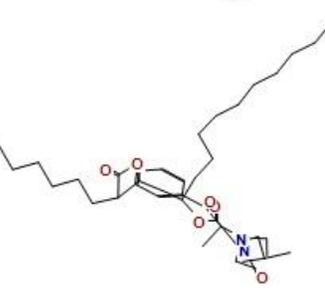
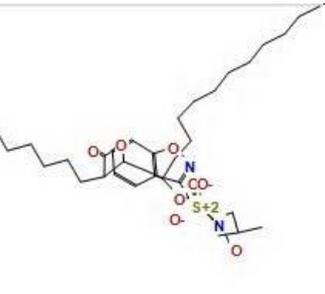
Fármacos	Alineación	U	F	S
Fentermina-Dietilpropión		24.1184	-74.8389	-50.7205
Fentermina-Zonisamida		-4.5913	-74.8151	-79.4064
Fentermina-Topiramato		9.5239	-59.6931	-50.1692
Fentermina-Orlistat		13.6312	-71.9678	-58.3366

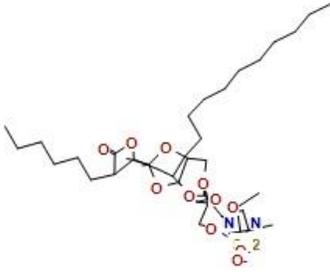
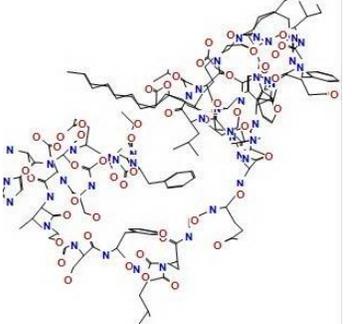
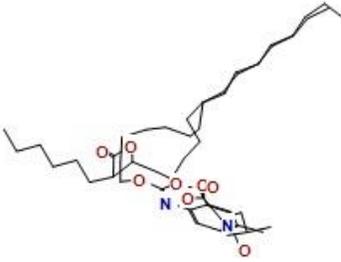
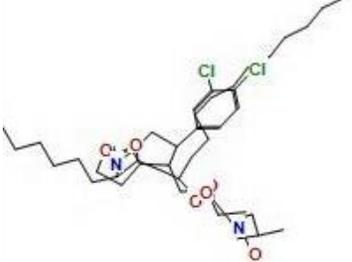
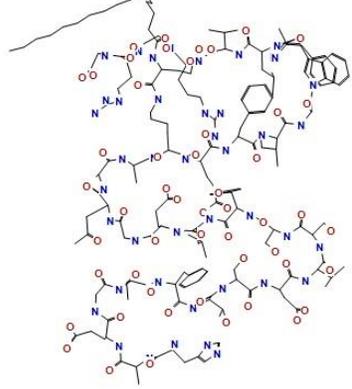
Fentermina-Liraglutida		144.3205	-73.1797	71.1408
Fentermina-Cetilistat		11.0911	-68.6604	-57.5694
Fentermina-Tesofensina		45.7022	-81.6971	-35.9949
Dietilpropión-Fentermina		12.8504	-71.3888	-58.5384
Dietilpropión-Zonisamida		-3.2434	-88.5216	-91.7649

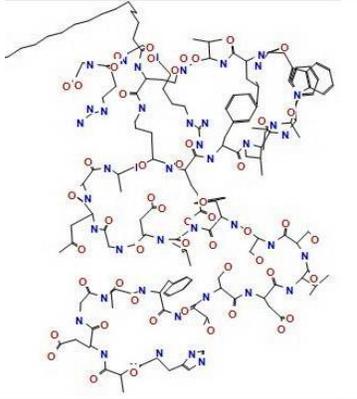
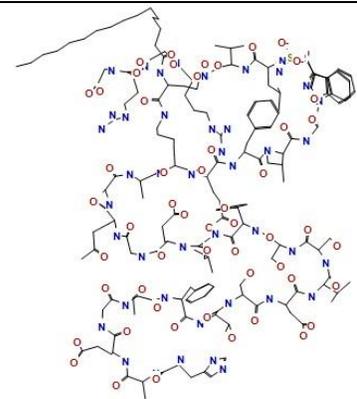
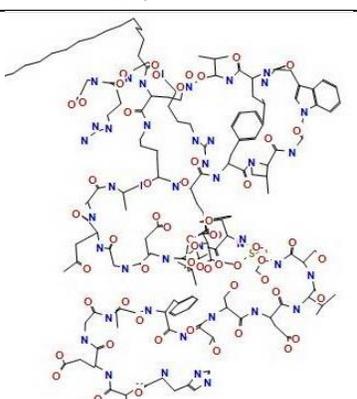
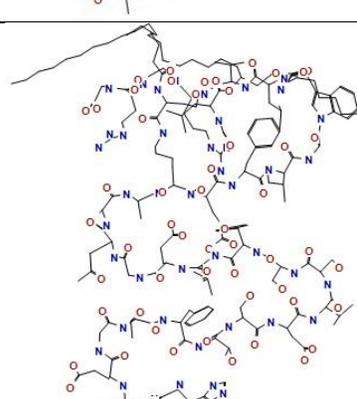
Dietilpropión- Topiramarato	 <p>The structure shows a complex bicyclic core with a sulfur atom labeled 'S+2' in green, and several nitrogen and oxygen atoms. It is a derivative of topiramate.</p>	11.0239	-78.6915	-67.6676
Dietilpropión- Orlistat	 <p>The structure features a bicyclic core with a long alkyl chain and a nitrogen atom bonded to an oxygen atom, characteristic of orlistat.</p>	17.0905	-87.5992	-70.5087
Dietilpropión- Liraglutida	 <p>The structure is a highly complex, multi-ring system with numerous nitrogen and oxygen atoms, representing the liraglutid derivative.</p>	160.6828	-99.8214	60.8614
Dietilpropión- Cetilistat	 <p>The structure shows a bicyclic core with a long alkyl chain and a nitrogen atom bonded to an oxygen atom, similar to orlistat but with a different side chain.</p>	11.8071	-92.0316	-80.2245

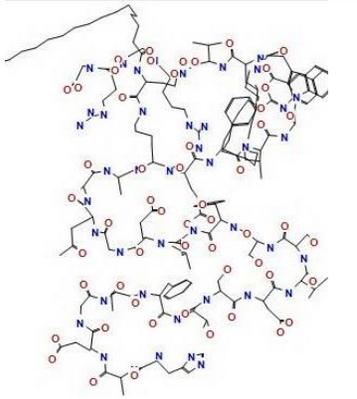
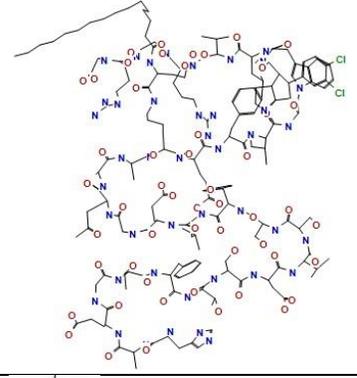
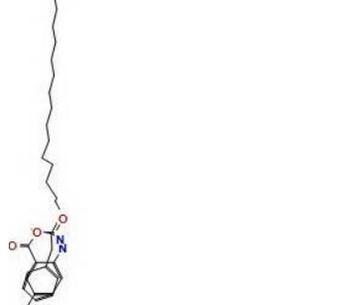
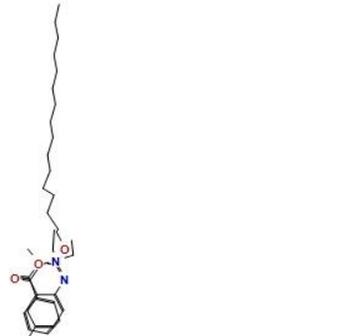
Dietilpropión-Tesofensina		49.2687	-95.7677	-46.4990
Zonisamida-Fentermina		13.4497	-73.9210	-60.4713
Zonisamida-Dietilpropión		24.3083	-82.8424	-58.5341
Zonisamida-Topiramato		14.8531	-95.9744	-81.1213
Zonisamida-Orlistat		14.7665	-87.0958	-72.3292
Zonisamida-Liraglutida		136.0674	-96.8693	39.1820

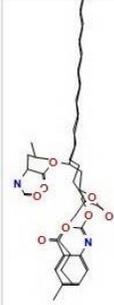
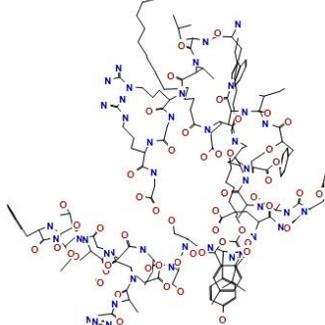
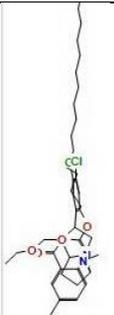
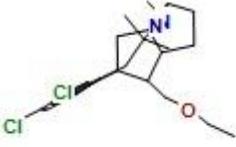
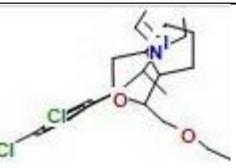
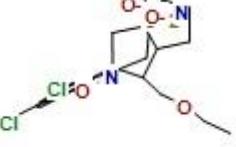
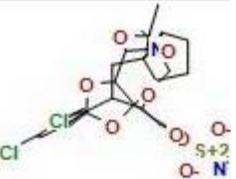
Zonisamida-Cetilistat		9.6411	-76.2515	-66.6104
Zonisamida-Tesofensina		46.4898	-83.5917	-37.1019
Topiramarato-Fentermina		14.1601	-54.8578	-40.6977
Topiramarato-Dietilpropión		25.4505	-73.3119	-47.8615
Topiramarato-Zonisamida		-3.8955	-97.0695	-94.6450
Topiramarato-Orlistat		19.7751	-93.4016	-73.6266

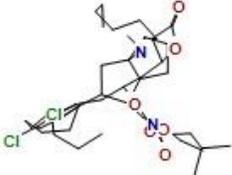
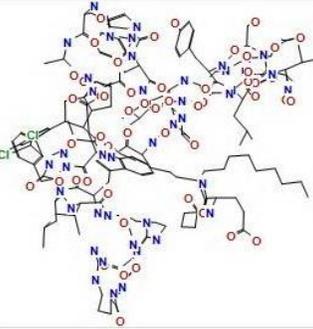
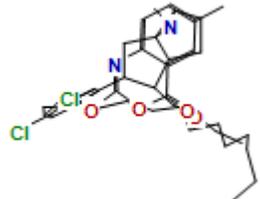
Topiramarato-Liraglutida		147.8911	-125.2973	22.5937
Topiramarato-Cetilistat		11.3486	-76.0230	-64.6743
Topiramarato-Tesofensina		48.0339	-83.1048	-35.0709
Orlistat-Fentermina		13.6410	-71.7417	-58.1007
Orlistat-Dietilpropión		26.8450	-82.0794	-55.2344
Orlistat-Zonisamida		-4.0567	-76.8192	-80.8759

Orlistat- Topiramato	 The image shows the chemical structure of Orlistat-Topiramato. It features a long, branched aliphatic chain on the left, connected to a complex bicyclic core. This core is further linked to a nitrogen-containing ring system that includes a quaternary ammonium group with a positive charge and a nitrogen atom with a negative charge, forming a zwitterionic salt.	8.6750	-84.5810	-75.9060
Orlistat- Liraglutida	 The image displays the chemical structure of Orlistat-Liraglutida. It is a highly complex molecule consisting of a long aliphatic chain from Orlistat, which is covalently bonded to the C-terminal region of the Liraglutida peptide chain. The peptide backbone is shown with various side chains, including hydroxyl, amide, and sulfur-containing groups.	151.6337	-149.2025	2.4312
Orlistat-Cetilistat	 The image shows the chemical structure of Orlistat-Cetilistat. It features a long aliphatic chain from Orlistat, which is linked to the Cetilistat moiety. The Cetilistat part consists of a bicyclic core with a nitrogen atom and a carbonyl group, and a long, branched aliphatic side chain.	14.7117	-147.4756	-132.7639
Orlistat- Tesofensina	 The image displays the chemical structure of Orlistat-Tesofensina. It shows a long aliphatic chain from Orlistat, which is connected to the Tesofensina moiety. The Tesofensina part includes a bicyclic core, a chlorine atom, and a carbonyl group.	48.1736	-112.7414	-64.5678
Liraglutida- Fentermina	 The image shows the chemical structure of Liraglutida-Fentermina. It is a complex molecule consisting of the Liraglutida peptide chain covalently bonded to the Fentermina moiety. The Fentermina part features a bicyclic core with a nitrogen atom and a carbonyl group, and a long, branched aliphatic side chain.	13.4748	-96.9906	-83.5158

Liraglutida-Dietilpropión		26.5373	-96.5154	-69.9780
Liraglutida-Zonisamida		-3.5325	-96.5325	-100.0650
Liraglutida-Topiramato		15.6905	-126.7458	-111.0553
Liraglutida-Orlistat		19.1497	-179.3141	-160.1644

Liraglutida-Cetilistat		13.4380	-161.3177	-147.8798
Liraglutida-Tesofensina		52.3732	-118.4146	-66.0413
Cetilistat-Fentermina		13.7322	-68.9139	-55.1817
Cetilistat-Dietilpropión		26.6532	-84.6365	-57.9833
Cetilistat-Zonisamida		-2.7810	-80.0690	-82.8500
Cetilistat-Topiramarato		7.2889	-68.4826	-61.1936

Cetilistat-Orlistat		14.1120	-152.7403	-138.6283
Cetilistat-Liraglutida		173.1006	-143.0182	30.0824
Cetilistat-Tesofensina		46.6716	-101.7197	-55.0481
Tesofensina-Fentermina		15.7653	-82.7521	-66.9868
Tesofensina - Dietilpropión		25.2194	-88.0424	-62.8230
Tesofensina - Zonisamida		-4.2539	-83.0275	-87.2813
Tesofensina - Topiramato		13.4424	-84.9835	-71.5411

Tesofensina - Orlistat		18.6473	-117.7607	-99.1134
Tesofensina- Liraglutida		129.7019	-92.0860	37.6159
Tesofensina- Cetilistat		13.7920	-108.1281	-94.3361

5. Análisis de resultados

5.1 Estudio de bases de datos moleculares

Las propiedades fisicoquímicas que se calcularon en ChemAxon y MOE fueron: enlaces aceptores de hidrógeno (HBA), enlaces donadores de hidrógeno (HBD), logaritmo del coeficiente de partición agua/ octanol (logP), peso molecular (MW), área de superficie polar (PSA), número de anillos (Num rings), número de enlaces rotables (RB). Adicionalmente, en MOE, se calculó el logaritmo del coeficiente de partición agua/octanol (SlogP) y el número de halógenos (Br, Cl, F, I). La descripción de estas propiedades se encuentra en el Anexo.

Las bases de datos que se analizaron se muestran en la Tabla 2. A continuación una breve descripción de estas bases de datos:

- * FEMA GRAS: contiene 2244 compuestos, saborizantes aprobados por la FDA (*Food and Drug Administration*) para consumo humano.³⁷

- * DrugBank: contiene 1713 compuestos, estos son medicamentos aprobados para uso humano. Esta base de datos es utilizada frecuentemente en bioinformática y quimioinformática, para obtener información sobre fármacos.⁴⁰
- * Caco-2: es la base de datos más pequeña que se estudia, con 97 compuestos. Está basada en una línea celular utilizada en la industria farmacéutica, como modelo para predecir absorción de fármacos en el intestino delgado.⁴¹
- * Chemdiv DNMT: contiene 26944 compuestos, es un conjunto de moléculas pequeñas seleccionadas de origen comercial como potenciales inhibidores de la enzima DNAmetiltransferasa.⁴²
- * Specs World Diversity Set 3 (Specs WD3): es una colección de 10000 compuestos utilizados en pruebas experimentales para encontrar compuestos bioactivos.⁷
- * SuperScent: contiene 2116 compuestos, es una base de datos que contiene compuestos volátiles utilizados en la industria farmacéutica, de alimentos y cosmética.⁴³
- * Traditional Chinese Medicine (TCM), es la base de datos más grande que se estudia en este trabajo, con 32357 compuestos; esta base de datos contiene los constituyentes de hierbas, productos animales y minerales utilizados en la medicina tradicional china.⁴⁴
- * Specs Natural Products (Specs NP), es una base de datos relativamente pequeña, de productos naturales.¹⁷

Estos descriptores se calcularon en ChemAxon y MOE, con el fin hacer una comparación entre los programas. A continuación el análisis de estas propiedades. En cuanto a enlaces aceptores de hidrógeno, la mayoría de las bases de datos tienen un valor promedio entre 3 y 4, a excepción de FEMA GRAS, SuperScent y TCM, esta última con un valor cercano a 8, mientras FEMA GRAS y SuperScent, entre 1 y 1.5; en ambos programas se observan tendencias similares, dando valores ligeramente más pequeños en el cálculo hecho en MOE. En los enlaces donadores de hidrógeno se presenta una tendencia similar a la propiedad anterior, donde FEMA GRAS y SuperScent, son las bases de datos con menor número de

enlace donadores de hidrógeno, con un promedio alrededor de 0.4, mientras que TCM tiene los valores mayores, alrededor de 6. Al contrario que en los aceptores de hidrógeno, en los donadores de hidrógeno, MOE arroja valores ligeramente mayores.

La base de datos que obtuvo mayor valor en el cálculo de logP, fue TCM. Mientras que FEMA GRAS y SuperScent tienen valores similares entre sí. LogP es el coeficiente de partición agua/octanol, indicando la polaridad de los compuestos.

Los compuestos FEMA GRAS, junto con los compuestos SuperScent, tienen menor peso molecular, TCM contiene los compuestos con mayor peso molecular. Al igual que peso molecular, FEMA GRAS y SuperScent tienen los compuestos con menor área de superficie polar, y Chemdiv DNMT contiene los compuestos con mayor PSA; de manera similar ocurre con en el número de anillos, FEMA GRAS y SuperScent contienen los compuestos con menor número de anillos y, en este caso, TCM contienen los compuestos con mayor número de anillos; en el caso de los enlaces rotables, SuperScen tiene el menor número de enlaces rotables, mientras que TCM el mayor número de estos, el resto de las bases de datos tienen un promedio entre 4 y 5.

Con lo anterior es posible observar que en la mayoría de las propiedades, las bases de datos FEMA GRAS y SuperScent, arrojan valores similares entre sí, en comparación al resto de las bases de datos. Esto es de esperarse debido a que estas dos bases de datos contienen compuestos que se utilizan como aditivos alimenticios, mientras que el resto están más estrechamente relacionadas con el área farmacéutica.

Adicionalmente en MOE, se calculó SlogP, y el número de halógenos (Cl, Br, F, I). En el caso de SlogP, se obtuvieron resultados similares al cálculo de logP. Para los halógenos no se obtuvieron resultados significativos para poder realizar un gráfico como en el caso de las demás propiedades; esto indica que en todas las bases de datos, la mayoría de los compuestos no tienen halógenos.

En general, se observa una diferencia mínima entre los resultados de los dos programas, ChemAxon y MOE. En los gráficos de la Figura 4, el tamaño de las

cajas representa el rango de los valores que arrojan los resultados del cálculo de las propiedades fisicoquímicas; en la mayoría de los casos, se observan triángulos, ya sea arriba o debajo de las cajas, lo que indica la presencia de valores atípicos muy grandes, esto debido a la variación de datos que se presentan.

Dada la naturaleza multivariable de los datos, se utilizaron métodos de reducción de dimensiones. La dimensionalidad de un conjunto de datos es el número de variables que se utilizan para describir cada objeto. El análisis de componentes principales es un método de uso común para reducir la dimensionalidad de un conjunto de datos. El análisis de componentes principales proporciona un nuevo conjunto de variables, así el conjunto de datos puede ser explicado por un pequeño número de estos componentes principales. Los componentes principales, son más convenientes para la visualización gráfica de los datos y su análisis.⁴

Con las propiedades calculadas anteriormente (HBA, HBD, logP, MW, PSA, Num rings, RB) se realizó un análisis de componentes principales y se calculó la distancia euclidiana utilizando todas las dimensiones, además se generó el gráfico con la representación del espacio químico. Para hacer comparables los dos programas utilizados, en el análisis de componentes principales no se incluyeron el número de halógenos y SlogP calculados en MOE.

En la Figura 5, se observan las representaciones del espacio químico, al igual que el caso anterior, el cálculo de componentes principales se realizó en los programas ChemAxon y MOE. El análisis de componentes principales se realizó utilizando los datos de las propiedades fisicoquímicas calculadas. Las diferencias que se observan entre los programas, se deben a que, en MOE, el análisis de componentes principales se realizó tomando en cuenta más propiedades; sin embargo, la tendencia es semejante para una misma base de datos en distintos programas.

Por razones de visualización, en general se grafican únicamente los dos o tres primeros componentes principales. En este caso, todas las representaciones del espacio químico se graficaron únicamente los dos primeros componentes principales. Cuando la varianza explicada por los dos primeros componentes

principales es mayor al 70%, la interpretación de la representación en dos dimensiones es válida.

En la Tabla 11, se presentan la varianza explicada por los dos primeros componentes principales, así como las propiedades fisicoquímicas con mayor contribución para cada base de datos. El cálculo de propiedades fisicoquímicas varía ligeramente dependiendo de cómo se calculan en los diferentes programas. Esto se ve reflejado tanto en las tendencias de la Figura 4 como en los gráficos del análisis de componentes principales. Esto también se observa en los diferentes valores obtenidos de varianza explicada y propiedades que contribuyen mayoritariamente a los componentes principales (Tabla 11).

Tabla 11. Varianzas explicadas en los dos primeros componentes principales (CP) del análisis de componentes principales y propiedades fisicoquímicas con mayor contribución en la bases de datos.

Base de datos	ChemAxon			MOE		
	Varianza	CP1	CP2	Varianza	CP1	CP2
GRAS	75.365%	HBD, HBA	logP, RB	73.988%	HBD, HBA	logP, RB
DrugBank	74.740%	HBD, HBA	HBD, PSA	75.552%	HBA, HBD	Num ring, logP
Caco-2	74.287%	HBA, HBD	HBD, logP	74.548%	HBD, HBA	logP, Num ring
Chemdiv DNMT	65.999%	Num ring, HBA	HBD, HBA	67.439%	HBA, Num ring	logP, Num ring
Specs WD3	63.129%	HBA, HBD	HBD, HBA	64.932%	HBA, HBD	Num ring, logP
SuperScent	72.920%	HBD, HBA	logP, RB	71.828%	HBD, HBA	logP, HBD
TCM	82.280%	HBD, HBA	logP, Num ring	82.444%	HBD, HBA	logP, Num ring
Specs NP	72.795%	HBD, HBA	logP, Num ring	74.093%	HBA, Num ring	LogP, HBD

En la tabla anterior se pueden observar las diferencias que hay en el cálculo del análisis de componentes principales en ChemAxon y MOE; las diferencias de varianza y contribuciones mayoritarias entre los programas se dan por diferencias entre estos. Se encontró que las bases de datos FEMA GRAS y SuperScent son las más semejantes entre sí, esto puede deberse a que el resto de las bases están relacionadas con fármacos, mientras que FEMA GRAS es una base de datos utilizada en alimentos y algunos compuestos SuperScent son utilizados en alimentos; SuperScent es la base de datos con mayor similitud a FEMA GRAS.

5.2 Similitud estructural

En la Tabla 4, se muestran las moléculas “privilegiadas”. En este trabajo se entiende como molécula privilegiada a aquella molécula FEMA GRAS, que presenta similitud con más de un fármaco en algún descriptor de cualquiera de los tres programas utilizados. En la Tabla 4, se muestran las moléculas privilegiadas FEMA GRAS, los fármacos a los que resultaron similares y el programa y descriptor asociado; en los casos donde el programa es ChemAxon, se da por entendido que el descriptor es Superstructure, debido a que, al realizar la búsqueda en base a la similitud estructural, entre fármacos y moléculas FEMA GRAS, se utilizaron cinco descriptores en ChemAxon, de los cuales en cuatro de estos (Substructure, similarity (Tanimoto 0.85 y distancia euclidiana 0.15), full, full fragment) no se obtuvo ningún resultado, es decir, las moléculas FEMA GRAS no son similares a los fármacos estudiados en base a las características estructurales que miden estos descriptores.

Se obtuvieron 81 compuestos FEMA GRAS con estructura privilegiada utilizando los descriptores de ChemAxon y PowerMV. En MOE, no se encontraron se obtuvieron resultados relevantes.

Las moléculas privilegiadas resultan importantes en este estudio, ya que si una molécula FEMA GRAS es privilegiada para varios fármacos, con el mismo o diferentes descriptores, esto podría significar dos cosas, una que los fármacos, para los que resultó privilegiada la molécula FEMA GRAS, podrían ser similares entre sí, dando así la pauta para posteriormente realizar el estudio entre los fármacos o que estas moléculas FEMA GRAS son prometedoras.

Las moléculas presentadas en la Tabla 5 son las más importantes en este trabajo, ya que éstas son las moléculas consenso, es decir, aquellas que resultaron similares a un determinado fármaco en dos o más programas y/o descriptores utilizados. Por lo que, se confirma la similitud entre esas moléculas con los fármacos. Es importante mencionar que en el caso de PowerMV, se seleccionan las diez moléculas FEMA GRAS más similares a cada fármaco para cada descriptor. Se obtuvieron 30 moléculas consenso. En la Tabla 5, se presenta el

número CAS de la molécula resultante como consenso, el fármaco con el que comparte similitud y los programas y descriptores correspondientes. Al igual que en caso anterior, cuando se menciona ChemAxon, se da por entendido que el descriptor utilizado fue Superstructure. En algunos casos se especifica que la similitud se encontró con ChemAxon y MOE, indicando con paréntesis el valor del coeficiente de Tanimoto con el cual se encontró la mayor similitud.

Es de esperarse que el análisis de similitud estructural proporcione moléculas semejantes en base a conectividades, tipos de átomos, etc. De manera complementaria el análisis de similitud puede realizarse un análisis en base a las propiedades fisicoquímicas de las moléculas. Este análisis puede o no arrojar resultados semejantes a los obtenidos en el análisis de similitud estructural. Dado que no son mutuamente excluyentes, es recomendable realizar ambos análisis.

5.3 Propiedades fisicoquímicas

Se generaron los descriptores mencionados en la metodología, en ChemAxon y MOE, con el fin de comparar los resultados obtenidos con ambos programas. Las características de los descriptores se especifican en el Anexo.

Se realizó el análisis de componentes principales de la base de datos FEMA GRAS junto con los fármacos que se estudian. Tomando los dos primeros componentes principales, se graficó el espacio químico de FEMA GRAS y los ocho fármacos. Al igual que en el estudio de las bases de datos, esto permitió visualizar la distribución de los datos en este espacio; estas representaciones se muestran en las Figuras 6 y 7.

En este caso, los dos primeros componentes principales explicaron el 80.40% de la varianza en el caso de ChemAxon, y 79.80% de la varianza para MOE. Para los dos primeros componentes principales, las contribuciones más representativas corresponden a donador de hidrógeno y aceptor de hidrógeno para el primer componente principal, y a logP y número de anillos para el segundo componente principal, en ChemAxon y en MOE. En ambos programas se obtuvieron resultados

semejantes, en cuanto a varianza y a la mayor contribución de las propiedades fisicoquímicas. La varianza explicada por los dos primeros componentes principales permitió explicar más del 70% de la variabilidad de los datos en los dos programas, por lo que la visualización de los dos primeros componentes principales es representativa de las propiedades calculadas en ambos programas.

La distancia euclidiana es la suma de las diferencias entre los valores de los atributos de cada entidad comparada, y no es más que una expresión simple en un espacio de varias dimensiones del Teorema de Pitágoras ⁴⁵. Se calcularon las distancias entre cada fármaco y cada molécula FEMA GRAS, de manera que, entre menor sea la distancia entre dos moléculas, mayor será la similitud entre estas.

Se seleccionaron las diez moléculas FEMA GRAS con menor distancia euclidiana entre cada fármaco, al igual que con los descriptores utilizados en PowerMV. Una vez que se tienen las moléculas más cercanas a los fármacos se realizó una comparación de estas, con las moléculas consenso de la Tabla 5, de esta manera, se obtuvieron las moléculas consenso resultantes en la comparación estructural que además, de acuerdo sus propiedades fisicoquímicas, muestran una cercanía con los fármacos.

En la última columna de la Tabla 7, se indica la distancia que hay entre cada fármaco y la molécula FEMA GRAS, como ya se había mencionado, entre menor sea la distancia, mayor será la similitud; el rango de valores de la distancia euclidiana se encuentra entre 0 y 1. Como se observa en la Tabla 7, el número de moléculas similares es mucho menor que en las Tabla 4 y 5, esto es porque, como ya se dijo, para obtener las moléculas privilegiadas y consenso (Tablas 4 y 5), sólo se buscó similitud estructural sin tomar en cuenta las propiedades fisicoquímicas, mientras que en la Tabla 7, se muestran las moléculas similares tanto en estructura como en propiedades fisicoquímicas.

La molécula FEMA GRAS 2-metil-1-fenilpropan-2-ol, es la molécula FEMA GRAS que tuvo mayor consenso con tres descriptores de PowerMV, es decir, tres descriptores utilizados en PowerMV fueron capaces de captar la similitud entre esta molécula FEMA GRAS y el fármaco Fentermina. La similitud entre estas dos

moléculas fue también observada en el espacio químico generado en base a propiedades fisicoquímicas utilizando los programas ChemAxon y MOE, es decir, bajo este protocolo la molécula 2-metil-1-fenilpropan-2-ol es similar a Fentermina, tanto estructuralmente como en sus propiedades fisicoquímicas.

Para la molécula FEMA GRAS 2-metil-4-fenil-2-butanol, dos representaciones estructurales la identificaron como similar al fármaco Fentermina. En el análisis de componentes principales se obtuvo una concordancia con este resultado, especialmente en el análisis de componentes principales generado utilizando las propiedades fisicoquímicas calculadas en MOE. Fentermina es un fármaco que actualmente se encuentra aprobado para su uso; se utiliza como supresor del apetito afectando SNC, es utilizado para personas con factores de riesgo como la presión arterial alta, colesterol alto o diabetes. Fentermina tiene varias funciones; sin embargo, la base principal de su uso es la reducción del apetito; a dosis clínicamente relevantes, libera serotonina.⁴⁶ La serotonina desempeña varias funciones fisiológicas, entre ellas la regulación del apetito.

Por otro lado, se encontró similitud entre la Teobromina y la Zonisamida; esta similitud fue encontrada por dos representaciones estructurales de PowerMV, en una de estas representaciones (Pharmacophore fingerprints), la distancia es 0.0, es decir que, en los aspectos medidos con este descriptor la Teobromina y la Zonisamida resultan idénticos. También se encontró similitud entre estas moléculas en el espacio químico generado en base a propiedades fisicoquímicas utilizando los programas ChemAxon y MOE, lo que indica que hay similitud significativa entre estas dos moléculas. Una de las funciones de la Zonisamida es el control del peso corporal. Sin embargo, el mecanismo para el control del apetito no es conocido, pero se sabe que aumenta niveles de serotonina y dopamina en el cerebro. También inhibe la actividad de la anhidrasa carbónica lo que podría alterar la percepción del gusto y puede regular lipogénesis *de novo*.⁴⁷ Este fármaco está permitido a dosis máximas de 400mg/día.³⁹ Aunque su uso es como anticonvulsivo, se está evaluando su reposicionamiento como fármaco supresor del apetito.

L-Alanil-L-glutamina tiene similitud con el fármaco Liraglutida, esta similitud fue encontrada con dos descriptores estructurales de PowerMV, y en las dos representaciones del espacio químico. Sin embargo, en comparación con las moléculas FEMA GRAS anteriores, los valores de las distancias calculadas a partir de los componentes principales, son mucho mayores; los valores obtenidos con los descriptores de PowerMV, son semejantes a los casos anteriores. Advantame, es otra molécula FEMA GRAS que resultó similar a Liraglutida; se encontró similitud con cuatro representaciones estructurales, en los tres programas utilizados, también en el análisis de componentes principales se obtuvieron resultados concordantes, debido a que se encontraron distancias pequeñas en ChemAxon y MOE. Liraglutida es utilizada para tratamientos de personas con diabetes tipo 2; se ha observado que induce a una reducción de peso corporal debido a una disminución de la ingesta calórica. La reducción en el peso corporal deriva sobre todo del tejido graso, particularmente de la grasa visceral, y no de la masa magra.⁴⁸ Actúa sobre los receptores GLP-1, favoreciendo la secreción de insulina y disminuyendo la de glucagón de forma dependiente de la glucosa además de retrasar el vaciamiento gástrico y producir una sensación de saciedad que reduce ligeramente el apetito. Este fármaco se encuentra aprobado para su uso en el tratamiento de la diabetes tipo 2, y en fase III para su aprobación como fármaco para control de peso.³⁹

Se encontraron seis compuestos FEMA GRAS similares a la Tesofensina. En los seis casos, se encontró similitud en un descriptor estructural, y en un análisis de componentes principales, ya sea de ChemAxon o MOE. La Tesofensina es un inhibidor de la noradrenalina, dopamina y la recaptación de serotonina.⁴⁹ La Tesofensina se encuentra en fase III para la aprobación para su uso.³⁹

La molécula FEMA GRAS 4-metilfenil laurato, resultó la más similar a Cetilistat para los descriptores calculados en PowerMV, así como en la representación del espacio químico, tanto en ChemAxon como en MOE. Por lo tanto, puede postularse que la molécula FEMA GRAS 4-metilfenil laurato puede posiblemente actuar como el fármaco Cetilistat. Éste, es un fármaco que se encuentra en la fase III para su aprobación, actúa inhibiendo la lipasa intestinal, de manera similar al

Orlistat, un fármaco ya aprobado, pero con menos efectos colaterales.⁵⁰ Sin embargo, el fármaco Orlistat, no presentó similitud relevante como Cetilistat.

5.4 Similitud entre fármacos

De acuerdo a los resultados obtenidos en la Tabla 8, las mayores similitudes estructurales entre fármacos se dio entre Fentermina y Dietilpropión, como se observa en la Tabla 3. Estos dos fármacos tiene el mismo anillo aromático seguido de cadenas alifáticas, por lo cual esta similitud se hace evidente. Otro par de compuestos con coincidencia estructural fueron Zonisamida y Fentermina, esto se debe también al anillo aromático. También se encontró cierta similitud entre Topiramato y Orlistat, en este caso, la coincidencia de átomos de oxígeno distribuidos en la misma región relativa es parcial, esto conduce a cierta similitud estructural. Estas similitudes se presentaron en fármacos ya aprobados para el control del peso.

Se encontró también cierta similitud entre Orlistat y Liraglutida, este último aun no aprobado para uso en el control del peso. Dado que Liraglutida es una molécula muy grande en comparación a los otros fármacos, la similitud medida con el coeficiente de Tanimoto no es apropiada, para tales casos, es más apropiado el coeficiente de Tversky. Cetilistat y Zonisamida, presentaron similitud, Zonisamida presenta similitud con Fentermina, Cetilistat y Dietilpropión debido al anillo aromático.⁴

En la Tabla 9, se muestran los resultados obtenidos con MOE, en los cuales se presentan algunas semejanzas con PowerMV. En este caso, se utilizó el coeficiente de Tanimoto, como se observa, los valores del coeficiente de Tanimoto son menores o cercanos a 0.50 lo que indica que la similitud es baja.⁴

Cuando se buscaron las menores distancias euclidianas entre fármacos y la base de datos FEMA GRAS, entre esos datos se encontró que Orlistat y Cetilistat presentan similitud tanto en propiedades fisicoquímicas como estructuralmente. Se sabe que Orlistat y Cetilistat actúan de manera similar por la inhibición de la

enzima lipasa pancreática. Sin embargo, Cetilistat aún se encuentra en proceso de aprobación, mientras que Orlistat si está aprobado para su uso a largo plazo para el control de peso.

Se realizó la superposición entre fármacos. En la Tabla 10 se presentan los resultados del mejor alineamiento estructural para cada par de fármacos, en las columnas se muestran los valores de U, F y S; donde representan lo siguiente: “U” es la energía media de la tensión de las moléculas en la alineación (kcal / mol). Esto se calcula como la suma de las energías potenciales de campo de fuerza individuales, dividido por el número de moléculas, en este caso, dos. “F” es a medida de similitud de la configuración; es el valor negativo de la superposición. Los valores más bajos indican una mayor similitud. “S” es la suma de las columnas de U y F. Los valores más bajos se indican mejores alineamientos.

Las mejores alineamientos se dieron entre Liraglutida y Topiramato, Liraglutida y Cetilistat, Liraglutida y Orlistat, y Cetilistat y Orlistat. Los alineamientos entre algunos de los fármacos con Liraglutida reflejan que esta última es una molécula muy grande comparada con los demás fármacos, esto que hace que las posibilidades de un buen alineamiento con ella sean mayores. Los valores de la superposición entre Cetilistat y Orlistat son los más pequeños, indicando una similitud mayor entre estos fármacos. Esto es razonable, ya que en el análisis de componentes principales, se encontró una similitud relevante entre estos dos fármacos.

6. Conclusiones

Se encontraron nueve moléculas FEMA GRAS similares a fármacos para el control de peso, tanto estructuralmente como en propiedades fisicoquímicas. Dos de estas moléculas son similares a Fentermina, fármaco de uso actual. Dos moléculas FEMA GRAS, son similares a Zonisamida, fármaco aprobado para su uso y en estudio para su reposicionamiento para el control del peso. Otras dos moléculas FEMA GRAS son similares a Liraglutida, fármaco en vías de aprobación para su uso en control de peso, al igual que Tesofensina, fármaco con el cual se encontró similitud con dos moléculas FEMA GRAS, y una molécula FEMA GRAS similar a Cetilistat, fármaco que al igual que los anteriores se encuentra en estudio para su aprobación. Con lo anterior es posible postular la posible bioactividad de estas moléculas FEMA GRAS con posibles efectos en la salud, como lo es el control del apetito.

7. Perspectivas

En cuanto al tema de saciedad, es posible continuar con estudios para determinar si las moléculas GRAS propuestas en este trabajo, tienen una bioactividad semejante a los fármacos, realizando para esto estudios *in vitro* e *in vivo*. Esta información será un buen indicativo de la bioactividad de dichas moléculas. Sin embargo, para evaluar si desempeñan esa función en los alimentos, sería necesario tomar en cuenta las dosis permitidas en los mismos, el tipo de tratamiento al que se somete al alimento, la interacción con los demás componentes, etc. Por otro lado, de ser bioactivas, las moléculas FEMA GRAS encontradas en este trabajo pueden ser utilizadas como punto de partida para su uso como entidades individuales.

Los métodos utilizados en este proyecto, pueden ser de gran ayuda en otros tipos de trabajos; por ejemplo el análisis de componentes principales, método para reducir la dimensionalidad, puede ser aplicado en el área de evaluación sensorial, si se requiere un estudio de algún alimento donde se tienen variables como sabor, olor, color, textura, etc.; el uso de este análisis podría ser útil para determinar los factores que tienen mayor contribución en la condición final del alimento. Otro ejemplo puede ser en el área de calidad, en productos donde se tiene una referencia para comprar la calidad de otros productos, utilizando el análisis de componentes principales, se pueden, al igual que el caso anterior, determinar qué factores influyen más en el producto de referencia y en los demás, como humedad, temperatura, tiempos de cocción, etc. La búsqueda de similitud estructural puede ser útil en el área de aditivos, si se tienen la necesidad de cambiar el aditivo pero se desea algún otro con funciones similares.

Tanto los métodos computacionales utilizados como el tema de saciedad son temas con mucho potencial. El tema de saciedad cada vez cobra mayor interés debido a enfermedades como obesidad. Por su lado, estudios computacionales que involucran metodologías de la quimoinformática y el modelado molecular han demostrado su utilidad y tiene un gran potencial en diversas áreas de la química

de los alimentos. De la unión de estos campos de investigación surge la recientemente denominada área “Foodinformatics”.⁵¹

8. Referencias

1. Boyd, D., B.; Lipkowitz, K., B. Examination of the Employment Environment for Computational Chemistry. *Reviews in Computational Chemistry* **2002**, *18*, 293-319.
2. Brown, F., K., Chemoinformatics: What is it and How does it Impact Drug Discovery. In *Annual Reports in Medicinal Chemistry*, 1998; Vol. 33, pp 375-384.
3. Medina-Franco, J. L.; Martínez-Mayorga, K.; Bender, A., Marín, M. R.; Giulianotti, M. A.; Pinilla, C.; Houghten, R. A. Characterization of activity landscapes using 2D and 3D similarity methods: Consensus activity cliffs. *Journal of Chemical Information and Modeling* **2009**, *49*, 477-491.
4. Leach, A., R. ; Gillet, V. J., *An introduction to chemoinformatics*. The Netherlands: Springer.: 2007.
5. Blondeau, S. D., Q, T. Scior, T. Bernard, P. Morin-Allory, L. Reverse pharmacognosy: Another way to harness the geneosity of nature. *Current Pharmaceutical Design* **2010**, *16*, 1682-1696.
6. Martínez-Mayorga, K.; Medina-Franco, J. L., Chemoinformatics-Applications in food Chemistry. In research, A. i. f. a. n., Ed. Burlington: Academic Press: 2009; Vol. 58, pp 33-56.
7. Medina-Franco, J. L.; Martinez-Mayorga, K.; L., P. T.; Del Rio, A. Chemoinformatic Analysis of GRAS (Generally Recognized as Safe) Flavor Chemicals and Natural Products. *Public Library of Science ONE* **2012**, *7*, e50798.
8. Martínez-Mayorga, K.; Peppard, T., L.; Lopez-Vallejo, F.; Yongye, A., B.; Medina-Franco, J., L. Systematic mining of Generally Recognized as Safe (GRAS) flavor chemicals for bioactive compounds. *Agricultural and Food Chemistry* **2013**, *61*, 7507-7514.
9. Maza, J. R. Aproximación teórica de medidas de similitud molecular sobre laplacianos de la densidad electrónica. Universidad Nacional de Colombia, 2010.
10. Medina-Franco, J. L. Activity Cliffs: Facts or Artifacts? *Chemical Biology and Drug Desing* **2013**, *8*, 553-556.
11. Martínez-Bueno, M.; Valdivia, E.; Gálvez, A.; Coyette, J.; Maqueda, M., 1998. , , 27 (2), pp. 347-358. Analysis of the gene cluster involved in production

- and immunity of the peptide antibiotic AS-48 in *Enterococcus faecalis*. *Molecular Microbiology* **1998**, *27*, 347-358.
12. Ginn, C.; Willett, P.; Bradshaw, J. Combination of molecular similarity measures using data fusion. *Perspectives in Drugs Discovery and Desings* **2000**, *20*.
 13. Cuadras, C., M., *Nuevos métodos de análisis multivariante*. CMC Editions: 2012.
 14. Amaro, I., R.; Vicente-Villardón, J., L.; Galindo-Villardón, M., P. Manova biplot para arreglos de tratamiento de dos factores basado en modelos lineales generales multivariantes. *Interciencia* **2004**, *29*, 26-32.
 15. Darling, S., D., *Una guía para el uso y construcción de modelos moleculares*. Estados Unidos de Norteamérica, 2006.
 16. Lengauer, T.; Lemmen, C.; Rarey, M.; Zimmermann, M. Novel technologies for virtual screening. *Drug Discovery Today* **2004**, *9*, 27-34.
 17. J.L., M.-F.; K., M.-M.; Peppard, L. T.; Del Rio, A. Chemoinformatic Analysis of GRAS (Generally Recognized as Safe) Flavor Chemicals and Natural Products. *Public Library of Science ONE* **2012**.
 18. Palmer, R. K. The pharmacology and signaling. . *Molecular interventions* **2007**, *7*, 87-98.
 19. Mattes, R. D.; Hollis, J.; Hayes, D.; Stunkard, A. Appetite: Measurement and manipulation misgivings. *Journal of the American Dietetic Association* **2005**, *Supplement*, S87-S97.
 20. Duca, F. A.; Covasa, M. Current and emerging concepts on the role of peripheral signals in the control of food intake and development of obesity. *British Journal of Nutrition* **2012**, *108*, 778-793.
 21. Fromentin, G.; Darcel, N.; Chaumontet, C.; Marsset-Baglieri, A.; Nadkarni, N.; Tome, D. Peripheral and central mechanisms involved in the control of food intake by dietary amino acids and proteins. *Nutrition Research Reviews* **2012**, *25*, 29-39.
 22. Page, A. J.; Symonds, E.; Peiris, M.; Blackshaw, L. A.; Young, R. L. Peripheral neural targets in obesity. *British Journal of Pharmacology* **2012**, *166*, 1537-1558.

23. Palou, A.; Bonet, M. L.; Pico, C.; Rodriguez, A. M. Nutrigenomica y obsidad. *Revista Mdica de la Universidad de Navarra* **2004**, *48*, 36-48.
24. Rovner, S. L. Food that smells too god to eat. *Science & Technology* **2010**, *88*, 51-52.
25. Mettler, S.; Schwarz, I.; Colombani, P. C. Additive postprandial blood glucose-attenuating and satiety-enhancing effect of cinnamon and acetic acid. *Nutrition Research Journal* **2009**, *29*, 723-727.
26. Panickar, K. S. Effects of dietary polyphenols on neuroregulatory factors and pathways that mediate food and energy regulation in obesity. *Molecular Nutrition & Food Research* **2013**, *57*, 37-47.
27. Martin, B.; Maudsley, S.; White, C., M.; Egan, J., M. Hormones in the naso-oropharynx: endocrine modulation of taste and smell. *Trends in Endocrinology and Metabolism* **2009**, *4*, 163-170.
28. Norman, M., H.; Chen, N.; Chen, Z.; Fotsch, F.; Hale, C.; Han, N.; Hurt, R.; Jenkins, T.; Kincaid, J.; Liu, L.; Lu, Y.; Moreno, O.; Santora, V., J.; Sonnenberg, J., D.; Karbon, W. Structure-Activity Relationships of a Series of Pyrrolo[3,2-d]pyrimidine Derivatives and Related Compounds as Neuropeptide Y5 Receptor Antagonists. *Journal of Medicinal Chemistry* **2000**, *43*, 4288-4312.
29. Williams, G.; McKibbin, P., E.; McCarthy, H., D. Hypothalamic regulatory peptides and the regulation of food intake and energy balance: signals or noise? *Proceedings of the Nutrition Society* **1991**, *50*, 527-544.
30. Ulven, T.; Frimurer, T., M.; Receveur, J., M.; Little, P., B.; Rist, O.; Nørregaard, P., K.; Högberg, T. 6-Acylamino-2-aminoquinolines as Potent Melanin-Concentrating Hormone 1 Receptor Antagonists. Identification, Structure-Activity Relationship, and Investigation of Binding Mode. *Journal of Medicinal Chemistry* **2005**, *48*, 5684-5697.
31. Arvanitoyannis, I., S.; Vaitsi, O.; Mavromatis, A. Potato: A Comparative Study of the Effect of Cultivars and Cultivation Conditions and Genetic Modification on the Physico-Chemical Properties of Potato Tubers in Conjunction with Multivariate Analysis Towards Authenticity. *Food Science and Nutrition* **2008**, *48*, 799-823.
32. Bonaventure, P.; Letavic, M.; Dugovic, C.; Wilson, S.; Aluisio, L.; Pudiak, C.; Lord, B.; Mazur, C.; Kamme, F.; Nishino, S.; Carruthers, N.; Lovenberg, T.

- Histamine H3 receptor antagonists: From target identification to drug leads. *Biochemical Pharmacology* **2007**, *73*, 1084-1096.
33. Wang, X.; Richards, G., J. Computer-Based Strategy for Modeling the Interaction of AGRP and Related Peptide Ligands with the AGRP-Binding Site of Murine Melanocortin Receptors. *Current Pharmaceutical Design* **2005**, *11*, 345-356.
34. Smith, R., L.; Cohen, S., M.; Doull, J.; Feron, V., J.; Goodman, J., L.; Wagner, B., M.; Adams, T., B. Criteria for safety evaluation of flavoring substances. *Food and Chemical Toxicology* **2005**, *43*, 1141-1177.
35. Sprous, D., G.; Salemme, F., R. A comparison of the chemical properties of drugs and FEMA/FDA notified GRAS chemical compounds used in the food industry. *Food and Chemical Toxicology* **2007**, *45*, 1419-1427.
36. Burdock, G. The GRAS process. *Food Technology* **2003**, *57*, 17.
37. Flavor & Extract Manufacturers Association, 2013. *FEMA*. [En línea]. Disponible en: <http://www.femaflavor.org> [Último acceso el 15 de noviembre de 2013]
38. Martínez-Mayorga, K.; Peppard, T., L.; Yongye, A., B.; Santos, R.; Giulianotti, M.; Medina-Franco, J., L. Characterization of a comprehensive flavor database. *Journal of Chemometrics* **2011**, *25*, 550-560.
39. Kang, J. G.; Park, C.-Y. Anti-obesity drugs: A review about their effects and safety. *Diabetes & Metabolims Journal* **2012**, *36*, 13-25.
40. Genome Alberta, 2013. *Drugbank, open data drug & drug target database*. [En línea]. Disponible en: <http://www.drugbank.ca> [Último acceso el 22 de noviembre de 2013]
41. Hou, T., J.; Rhang, W.; Xia, K.; Quiao, X., B.; Xu, J. ADME Evaluation drug discover, correlation of Caco-2 permeation with simple molecular properties. *Journal of Chemical Information Computes Sciences* **2004**, *44*, 1585-1600.
42. Yoo, J. M.-F., J. L. Chemoinformatic approaches for inhibitors of DNAMethyltransferases: Comprehensive characterization of screening libraries. *Computational Molecular Biosciense* **2011**, *1*.
43. Charité, 2012, *Superscent*, [En línea]. Disponible en: <http://bioinf-applied.charite.de/superscent/> [Último acceso el 15 de noviembre de 2013]

44. TCM World Fundation, 2013. *TCM Traditional Medicine Chinesse*. Disponible en: <http://www.tcmworld.org/> [Último acceso el 24 de noviembre de 2013]
45. Herrera, A. La clasificación numérica y su aplicación en la ecología. *República Dominicana: Instituto Tecnológico de Santo Domingo*. **2000**.
46. Rothman, R., B.; Baumann, M., H.; Dersch, C., M.; Romero, D., V.; Rice, K., C.; Caroll, F., I.; Partilla, J., S. Amphetamine-type central nervous system stimulants release norepinephrine more potently than they release dopamine and serotonin. *Synapse* **2001**, 29, 32-41.
47. Gadde, K., M.; Yonish, G., M.; Foust, M., S.; Wagner, H., R. Combination therapy of Zonisamide and Bupropion for weight reduction in obese women: a preliminary, randomized, open-label study. *Journal Clinical Psychiatry* **2007**, 68, 1226-1229.
48. Ampudia-Blasco, F., J.; Calvo-Gómez, C.; Cos-Claramunt, X.; García-Alegría, J.; Jódar-Gimeno, E.; Mediavilla Bravo, J., J.; Mezquita-Raya, P., Navarro-Pérez, J.; Puig-Domingo, J. Liraglutida en el tratamiento de la diabetes tipo 2: recomendaciones para una mejor selección de los pacientes, desde una visión multidisciplinar. . *Avances de Diabetología* **2010**, 26, 226-234.
49. Astrup, A.; Madsbad, S.; Breum, L.; Jensen, T., J. ; Kroustrup, J., P.; Larsen, T., M. Effect of tesofensine on bodyweight loss, body composition, and quality of life in obese patients: a randomised, double-blind, placebo-controlled trial. *The Lancet* **2008**, 772, 1906-1913.
50. Alvarez, V. Tratamiento farmacológico de la obesidad. *Revista Medica Clinica Las Condes* **2012**, 23, 173-179.
51. Martínez-Mayorga, k.; Medina-Franco, J. L. Foodinformatics: Applications of Chemical Information to Food Chemistry. *Springer* **2014**, in preparation.

Anexos

Anexo 1. Descriptores estructurales

- * Descriptores utilizados en ChemAxon, PowerMV y MOE para la búsqueda de la similitud estructural entre moléculas GRAS y fármacos.
 - Descriptores ChemAxon
 - 8.1.1 Superstructure: encuentra todas las moléculas con superestructura similar a la consulta. En esta búsqueda las estructuras de la base de datos se utilizan sin hidrógenos
 - Descriptores PowerMV:
 - Atom pair: descriptores pares, útil para encontrar análogos cercanos.
 - Atom pair Carhart: se basa en 36 características; estas se basan en el número de hidrógenos, número de electrones, los halógenos unidos a moléculas orgánicas, etc.
 - Fragment Pair: se tienen 14 clases, basados en principios bioisostéricos.
 - Pharmacophore Fingerprints: basados en 6 clases; busca al vecino más cercano. Para los descriptores continuos, Tanimoto; se sustituye por la distancia euclidiana para medir similitud.
 - Weighted Burden Number: este descriptor está basado en la carga de la molécula.¹
 - MOE: Se buscó la similitud entre fármacos y moléculas GRAS utilizando MACCS keys, que encuentra moléculas similares a la estructura de consulta. Contiene un total de 166 claves estructurales, correspondientes a varios grupos y a compuestos orgánicos. La búsqueda se realizó utilizando el coeficiente de Tanimoto, se buscó la similitud más grande posible para cada fármaco

Anexo 2. Descriptores fisicoquímicos

- * Propiedades fisicoquímicas calculadas en ChemAxon y MOE para realizar el Análisis de Componentes Principales.

Descriptor	ChemAxon	MOE
Aceptores de hidrógeno	de Aceptor de enlace de hidrógeno en la molécula.	Número de átomos de aceptor de enlace de hidrógeno (sin contar los átomos de ácidos y contando átomos que son tanto de los donantes y aceptores de enlaces de hidrógeno, como -OH).
Donadores de hidrógeno	de Recuento átomo donador de hidrogeno en una molécula.	Número de átomos donantes de enlaces de hidrógeno (sin contar los átomos básicos y contando átomos que son los donantes y aceptores de enlaces de hidrógeno, tales como -OH).
logP (o/w)	Logaritmo del coeficiente de partición agua/octanol, para especies no cargadas.	Logaritmo del coeficiente de partición agua/octanol (incluyendo hidrógenos implícitos). Esta propiedad se calcula a partir de un modelo tipo de átomo lineal con $r^2=0,931$.
SlogP	No aplica	Logaritmo del coeficiente de partición agua/octanol. Esta propiedad es un modelo de contribución atómica que calcula logP de la estructura dada, es decir, el estado de protonación correcta.
Número de enlaces giratorios	Número de enlaces giratorios en la molécula. Esos enlaces simples se consideran giratorios, que no están conectados a hidrógenos, ni a átomos terminales.	Número de enlaces simples giratorios. Los enlaces simples conjugados no están incluidos.
Número de anillos	Número de anillos	Número de anillos
Peso molecular	Peso molecular en unidades de masa atómica con pesos atómicos.	Peso molecular en unidades de masa atómica con pesos atómicos.
Área de superficie polar	Mediante contribuciones de los grupos, calcula la aproximación de la superficie polar.	Mediante contribuciones de los grupos, calcula la aproximación de la superficie polar.
Número de halógenos	No aplica	Número de halógenos (Br, Cl, F, I) en la molécula.

Anexo 3. Coeficiente de Tanimoto

La similitud utilizando Tanimoto se calcula como:

$$S_{Tanimoto} = \frac{C}{A + B - C}$$

Donde: A: Número de fragmentos estructurales prese

B: Número de fragmentos estructurales presentes en la molécula B.

C: Número de fragmentos estructurales en común.

El coeficiente de Tanimoto oscila entre 0 y 1, donde 0 representa la mínima similitud y 1 representa máxima similitud.

Anexo 4. Distancia euclidiana

La distancia euclidiana es la suma de las diferencias entre los valores de los atributos de cada entidad comparada, y no es más que una expresión simple en un espacio de varias dimensiones del Teorema de Pitágoras.² La distancia se calcula de la siguiente manera:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + \dots + (n_2 - n_1)^2}$$

Anexo 5. Referencias

1. Liu, K.; Feng, J.; Young, S. PowerMV: software environment for molecular viewing, descriptor generation, data analysis and hit evaluation. *Journal of Chemical Information and Modeling* **2005**, *45*, 515-522.
2. Herrera, A. La clasificación numérica y su aplicación en la ecología. *República Dominicana: Instituto Tecnológico de Santo Domingo*. **2000**.