



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

**EVALUACIÓN DEL POTENCIAL EFECTO ANALGÉSICO DE
SABORIZANTES UTILIZANDO MÉTODOS ESTADÍSTICOS Y DE
MODELADO MOLECULAR**

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

**PRESENTA
DIANA ESMERALDA TERRAZAS ALVAREZ**



MÉXICO, D.F.

2014



Universidad Nacional
Autónoma de México



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 Rubén Carreño Ortiz
VOCAL: Rosario Adelaida Muñoz Clares
SECRETARIO: Karina Martínez Mayorga
1er. SUPLENTE: Dulce María Gómez Andrade
2° SUPLENTE: Patricia Severiano Pérez

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

Asesor del tema:

Dra. Karina Martínez Mayorga

Supervisor técnico:

Dr. José Luis Medina Franco

Sustentante:

Diana Esmeralda Terrazas Alvarez

Índice

| | |
|---|-----------|
| Abreviaturas | 5 |
| Resumen | 6 |
| 1. Introducción | 7 |
| 2. Antecedentes | 9 |
| 2.1. Química computacional y quemoinformática | |
| 2.1.1. Representación de estructuras químicas | 9 |
| 2.1.2. Descriptores moleculares 1D, 2D, 3D | 11 |
| 2.1.3. Bases de datos | 13 |
| 2.1.4. Química computacional, quemoinformática y bioinformática | 16 |
| 2.1.5. Similitud molecular | 17 |
| 2.1.6. Análisis multivariable: Análisis de componentes principales, análisis de cluster o de agrupamiento. | 18 |
| 2.2. Modelado molecular | 19 |
| 2.2.1. Análisis conformacional | 19 |
| 2.2.2. Superposición molecular | 19 |
| 2.2.3. Acoplamiento molecular | 19 |
| 2.3. Base de datos GRAS | 20 |
| 3. Marco teórico | 22 |
| 4. Metodología | 25 |
| 4.1. Procedimiento | |
| 4.1.1. Comparación de bases de datos | 26 |
| 4.1.2. Espacio químico | 26 |
| 4.1.3. Búsqueda por similitud estructural | 27 |
| 4.1.4. Alineamiento o superposición molecular | 27 |
| 4.1.5. Comparación de secuencias de péptidos de alimentos | 27 |
| 5. Resultados y discusión | |
| 5.1. Caracterización de las bases de datos basada en propiedades fisicoquímicas | 28 |
| 5.2. Espacio químico por base de datos | 33 |
| 5.3. Búsqueda por similitud estructural | 36 |
| 5.4. Espacio estructural y químico de FEMA-GRAS y fármacos para el dolor | 50 |
| 5.5. Comparación de secuencia de péptidos de alimentos | 53 |

| | |
|---|-----------|
| 6. Conclusiones | 57 |
| 7. Bibliografía | 58 |
| 8. Glosario | 61 |
| 9. Anexos | |
| 9.1. Anexo 1. Definición de las propiedades fisicoquímicas evaluadas en los programas de ChemAxon y MOE para la caracterización de las bases de datos utilizadas en este trabajo | 62 |
| 9.2. Anexo 2. Métodos o tipos de búsqueda y descriptores utilizados en los programas de ChemAxon, PowerMV y MOE, para búsqueda por similitud | 64 |
| 9.3. Anexo 3. Aplicación de análisis de componentes principales y búsqueda por similitud en la industria alimentaria | 66 |

Abreviaturas

| | |
|---------------------|---|
| FEMA | <i>Flavor and Extract Manufacturers Association</i> |
| GRAS | <i>Generally Recognized as Safe</i> |
| PCA | Análisis de Componentes Principales (<i>Principal Component Analysis</i>) |
| CP | Componente Principal |
| MOE | <i>Molecular Operating Environment</i> |
| logP y SlogP | Coefficiente de partición octanol/agua |
| TPSA o PSA | Área de superficie polar topológica |
| a_don | Número de átomos donadores de enlaces de hidrógeno |
| a_acc | Número de átomos aceptores de enlaces de hidrógeno |
| b_rotN | Número de enlaces rotables |
| MW | Peso o masa molecular |
| RC | Número de anillos |
| a_nBr | Número de átomos de bromo |
| a_nCl | Número de átomos de cloro |
| a_nF | Número de átomos de flúor |
| a_nI | Número de átomos de yodo |
| COX | Ciclooxigenasa |
| UHT | <i>Ultra-High-Temperature</i> |

Resumen

El dolor es uno de los padecimientos que conducen con mayor frecuencia a la búsqueda de auxilio médico. El dolor puede ser de diferente intensidad. Diferentes estrategias han sido exploradas para su control incluyendo el uso de agentes analgésicos para dolor intenso y crónico, modificación de la dieta para dolor transitorio, reducción de estrés, etc. Se ha postulado la hipótesis de que ciertos componentes de los alimentos pueden tener efecto analgésico transitorio. Sin embargo, no se han realizado estudios sistemáticos de los posibles componentes de alimentos que puedan tener este efecto, así como sus posibles mecanismos de acción. En este trabajo se pretende contribuir a éste conocimiento; como primer paso se analizaron las propiedades estructurales y fisicoquímicas de las moléculas contenidas en seis bases de datos moleculares con el fin de identificar semejanzas y diferencias entre ellas. Posteriormente, se realizó un estudio de similitud molecular que consistió en la comparación estructural y de las propiedades fisicoquímicas entre analgésicos utilizados actualmente y saborizantes contenidos en la lista FEMA GRAS (GRAS 25). Cabe mencionar que se utilizaron los fármacos aprobados para uso clínico a la fecha en que se realizó este análisis.

De la búsqueda de similitud estructural se encontraron 20 moléculas FEMA GRAS relevantes. Entre ellas se encuentran la capsaicina y la piperina, que son sustancias que han sido utilizadas como analgésicos en medicina tradicional. Después de realizar la búsqueda de similitud molecular, tanto estructural como con propiedades fisicoquímicas, sólo dos moléculas (Teobromina y N-[(4-amino-2,2-dióxido-1H-2,1,3-benzotiadiazia-5-il)oxi]]-2,2-dimetil-N-propilpropanamida) de las 20 encontradas muestran similitud a las moléculas de los fármacos estudiados. Estas moléculas de aditivos saborizantes GRAS podrían ser exploradas como potenciales agentes analgésicos. La investigación de la bioactividad de componentes de alimentos puede tener implicaciones importantes en la industria alimentaria y farmacéutica.

1. Introducción

Los saborizantes son utilizados ampliamente en la industria de alimentos y bebidas con el fin de mejorar sus propiedades organolépticas o con fines de innovación. Los saborizantes pueden ser de origen natural o sintético y su uso está regulado según la ley de cada país, actúan sobre los receptores del gusto y del olfato. Dada la plasticidad y variedad de receptores en el sistema del olfato, la sensación de sabor se realiza, muy probablemente, a través de la interacción con varios receptores de manera concertada. 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.

El dolor es uno de los padecimientos que conducen con mayor frecuencia a la búsqueda de auxilio médico. La gran variedad de analgésicos actuales permite controlar el dolor en sus diferentes intensidades y causas. En el caso de dolor agudo, por ejemplo, postoperatorio o debido a cáncer, los analgésicos más utilizados actúan vía receptores de *opioides*. Sin embargo, estos analgésicos aun presentan importantes efectos secundarios no deseados. En el caso de dolor de baja intensidad se han buscado diferentes estrategias para su control incluyendo reducción de estrés, modificación de la dieta, etc. En este ámbito se ha postulado la hipótesis de que ciertos componentes de los alimentos son moduladores del dolor. En este trabajo se pretende contribuir a evaluar esta hipótesis para el caso particular de saborizantes autorizados para consumo humano.

El objetivo de este trabajo es identificar aquellos saborizantes con mayor similitud a fármacos utilizados para controlar el dolor. Este objetivo está sustentado en el “principio de similitud molecular” que indica que moléculas semejantes tienen propiedades semejantes.^{1,2} Existen excepciones a esta regla que son llamados “*property cliffs*” (propiedades acantiladas) éstos se refieren a pares de moléculas que tienen alta similitud estructural pero grandes variaciones en la propiedad de interés, que puede ser actividad biológica.³ En la primera parte de este trabajo se analizaron diversas bases de datos con fines comparativos y para evaluar la complementariedad entre ellas. Posteriormente, se exploró la base de datos FEMA (*Flavor and Extract Manufacturers Association*) GRAS para identificar moléculas con similitud a analgésicos actualmente en uso clínico. Se seleccionó la lista de FEMA GRAS de saborizantes (GRAS 25) por ser una base de datos adecuada para nuestro estudio y por su disponibilidad en nuestro grupo de trabajo. Para considerar que la base de datos es adecuada se consideran criterios como la veracidad y

exhaustividad de la información, remoción de duplicados, consideración de estructuras únicas (por ejemplo, sin sales o dímeros), consideración de la estereoquímica, etc. Este proceso es llamado “curación de datos” y fue realizado en estudios previos a éste trabajo.⁴ Además, se tiene conocimiento de que el perfil de las propiedades fisicoquímicas así como el espacio químico de ésta base de datos y los fármacos comercializados son parecidos.⁵ ⁶ Para realizar la comparación de las moléculas en forma sistemática y cuantitativa, en este trabajo se hace uso de herramientas de la *quemoinformática*; y se introducen los métodos, análisis, alcances y limitaciones de este campo.

2. Antecedentes

En este trabajo se utilizaron diferentes representaciones moleculares para evitar el sesgo que impondría el uso de una sola representación, haciendo así más robusta la búsqueda basada en similitud estructural. Para un mejor entendimiento de los objetivos, resultados y conclusiones, a continuación se detallan las metodologías y análisis de datos utilizados directa o indirectamente en este trabajo. En las primeras dos secciones (2.1.1 y 2.2.2) se describen representaciones moleculares, sus ventajas, desventajas y usos más comunes. Además de servir como introducción básica al tema tratado en esta tesis, algunas de estas representaciones estructurales son utilizadas en bases de datos descritas en la sección 2.1.3. Una vez generadas las bases de datos pueden realizarse estudios y análisis como los detallados en las secciones 2.1.4-2.1.6 y que engloban las metodologías más relevantes de la quemoinformática y del modelado molecular. En la segunda parte de los antecedentes se presentan metodologías computacionales que dependen directamente de la estructura tridimensional de las moléculas y que son utilizadas principalmente en estudios de reconocimiento molecular. En la tercera y última parte de los antecedentes se describirá la base de datos GRAS.

2.1. Química computacional y quemoinformática

2.1.1. Representación de estructuras químicas

Las estructuras químicas se representan tradicionalmente usando líneas que indican los enlaces químicos entre los átomos (fórmula estructural 2D). Esta representación de las estructuras químicas ha permitido la transmisión exitosa de conocimiento e información desde los inicios de la química como disciplina científica. Nuevas metodologías han demandado la implementación de formas diversas de representación de estructuras. Por ejemplo, los métodos de modelado molecular requieren el uso de representaciones basados en tres dimensiones y con optimización de los recursos de memoria. En este campo, las moléculas pequeñas (también llamadas ligandos en el área de diseño de fármacos) se representan usualmente mediante listas de átomos y su conectividad. Estas moléculas pueden agruparse en grandes bases de datos que permiten el almacenamiento, procesamiento y búsqueda de información sobre millones de moléculas ocupando terabytes de memoria física de la computadora. Las moléculas grandes como las proteínas pueden representarse de un modo más compacto usando secuencias de

aminoácidos. Ejemplos de representación estructural utilizados en modelado molecular incluyen:

- **Coordenadas cartesianas (XYZ):** Un sistema de coordenadas cartesianas se define por dos ejes ortogonales en un sistema bidimensional y tres ejes ortogonales en un sistema tridimensional que se cortan en el origen. Las coordenadas de un punto cualquiera están dadas por las proyecciones del vector de posición del punto sobre cada uno de los ejes.⁷
- **Coordenadas internas (Matriz Z):** Una matriz Z representa el conjunto de coordenadas internas que definen la posición de cada átomo dentro de una molécula. Se denominan coordenadas internas ya que la posición de cada punto queda definida por su posición relativa respecto de un número mínimo del resto de los puntos presentes. Su nombre proviene del hecho de que automáticamente se asigna como el eje cartesiano Z al primer vector de distancia encontrado respecto del átomo central. Esta representación consta de un identificador para cada átomo, de la distancia de este hacia otro ya definido con anterioridad, del ángulo que ambos forman respecto a un tercero y de ángulos de torsión respecto de un cuarto átomo. Los elementos empleados para la generación de una matriz Z pueden o no reflejar la conectividad de la molécula. Generalmente, el uso de una matriz Z permite realizar optimizaciones de geometría con mayor velocidad de cómputo respecto del uso de coordenadas cartesianas.⁷
- ***Simplified Molecular Input Line Entry Specification (SMILES) y Smiles arbitrary target specification (SMARTS)*:** Los sistemas de notación lineal de estructuras químicas tales como SMILES y SMARTS utilizan una combinación de caracteres alfanuméricos. La notación llamada SMILES fue desarrollada por David Weininger y colaboradores y es ampliamente utilizada por su simplicidad y fácil comprensión. El término SMILES se usa comúnmente para referirse tanto a una sola cadena como a varias cadenas de SMILES (*SMILES strings*).^{8, 9} La notación llamada SMARTS es una notación lineal para la especificación de esquemas subestructurales en moléculas. Los SMARTS hacen uso de muchos de los símbolos de SMILES y también permite la especificación de átomos y enlaces “comodines” que pueden ser usados para definir *consultas subestructurales* para búsqueda en una base de datos de compuestos químicos.¹⁰

Otra de las ramas de la química computacional, la quemoinformática, ha requerido el uso de formas más compactas y diversas de representación estructural que permitan procesar rápidamente bases de datos moleculares con miles o millones de moléculas en forma eficiente. Toda representación estructural captura ciertos elementos relevantes de las moléculas pero conlleva a pérdida de información. Por lo tanto, es necesario que la pérdida de información sea mínima, justificable y controlada. Las huellas digitales o *fingerprints* son representaciones estructurales comúnmente utilizadas en quemoinformática. Hay muchos tipos de *fingerprints* que pueden incluso combinarse, esto será discutido más adelante.

- **Huellas digitales (fingerprints):** Son elementos estructurales conocidos como claves, firmas, huellas digitales moleculares o simplemente, huellas dactilares. Representan un conjunto de características derivadas de la estructura de una molécula. Las características particulares calculadas a partir de la estructura pueden ser bastante arbitrarias y dependerán de la topología de la gráfica química o incluso la conformación, en el caso de fingerprints que involucren información de la molécula en tres dimensiones. Diferentes esquemas de huellas digitales capturan diferentes características moleculares, de acuerdo con la filosofía de diseño del sistema de huellas dactilares. El objetivo fundamental es codificar ciertas propiedades directamente o indirectamente en la huella dactilar y, a continuación, utilizar la huella digital como un sustituto de la estructura química. Las comparaciones entre las moléculas se reducen a la comparación de conjuntos de características. El grado en el que los conjuntos se superponen es a lo que se le conoce como similitud molecular.¹¹

2.1.2. Descriptores moleculares 1D, 2D, 3D

Los descriptores moleculares juegan un papel importante en la química, la industria farmacéutica, las leyes de protección ambiental y el control de calidad.¹² Los descriptores moleculares permiten la manipulación y el análisis de información química estructural, por ejemplo, correlaciones estructura-actividad, búsquedas basadas en similitud estructural o subestructural, etc. Los descriptores moleculares pueden representar propiedades fisicoquímicas o pueden ser derivados de la aplicación de algoritmos a las estructuras químicas. Los descriptores moleculares varían en complejidad y en tiempo de cómputo, algunos son representaciones matemáticas de una molécula obtenida por un algoritmo

específico aplicado a definir una representación molecular o bien a un procedimiento experimental específico. El uso de éstos nos permite interpretar las propiedades de las moléculas tomando parte importante en el desarrollo de modelos predictivos.

Si bien, varios parámetros de cuantificación molecular tienen su origen en la química cuántica y la *teoría de grafos*, el término “descriptor molecular” fue acuñado popularmente con el desarrollo de modelos donde se correlacionaban la estructura con una o varias propiedades. Los pioneros en el diseño de descriptores moleculares fueron Platt y Wiener.^{13, 14}

Los descriptores moleculares se dividen en dos grupos:

- **Medidas experimentales:** Son parámetros basados en medidas experimentales que representan a la molécula. Ejemplos: log P, dipolo, polarización, etc.
- **Descriptores moleculares teóricos:** Son derivados de representaciones simbólicas de una molécula obtenidos generalmente por métodos computacionales.

Los descriptores pueden a su vez dividirse de acuerdo al tipo de representación al que corresponden: 1D, 2D o 3D. El cálculo de descriptores 1D requiere únicamente de la representación estructural en una dimensión, por ejemplo, el cálculo del peso molecular, donde no es necesario conocer la conectividad entre los átomos para calcular apropiadamente el descriptor. Los descriptores 2D y 3D dependen de la representación estructural en 2 y 3 dimensiones, respectivamente.

En el caso de los descriptores 3D es necesaria la descripción geométrica espacial de las moléculas. Las representaciones 3D son las más completas, por ende, los descriptores 3D tienen un mayor contenido de información y son computacionalmente más demandantes. Ejemplos de descriptores tridimensionales son: descriptores geométricos, descriptores estéricos y descriptores de tamaño y forma.¹² La Tabla 1 resume ejemplos de los diferentes descriptores moleculares 1D, 2D y 3D.

Tabla 1. Descriptores moleculares y propiedades relacionadas.¹⁵

| Descriptor | | Propiedades estructurales y fisicoquímicas |
|------------|--|---|
| 1D | Filtro sencillo Filtro complejo | Peso molecular. Acidez, basicidad, etc. |
| 2D | Huellas dactilares Propiedades fisicoquímicas | Tipos de átomos y subestructuras en una representación binaria. Lipofilicidad, valores de pK _a , parámetros de toxicidad. |
| 3D | Descriptores químicos cuánticos. Patrones 3D de | Cambio de distribución en las moléculas. Interacciones en los sitios de unión (enlaces de |

| | |
|--|---|
| farmacóforos. Patrones SAR Coeficientes de auto- correlación Selección virtual | hidrógeno, electrostático, etc.) Actividad farmacológica. Similitud de forma. Actividad farmacológica. |
|--|---|

2.1.3. Bases de datos

Una base de datos es el conjunto de datos organizados en un mismo contexto para su uso y vinculación. Se le llama base de datos a los bancos de información que contienen datos relativos a diversas temáticas y categorizados de distinta manera, pero que comparten entre sí algún tipo de vínculo o relación que busca ordenarlos y clasificarlos en conjunto. Una base de datos puede ser de diverso tipo, desde una pequeña base de datos local o de grupo, hasta bases de datos complejas y de gran tamaño, de índole corporativo, gubernamental e incluso de organismos mundiales. Recientemente, el término base de datos comenzó a utilizarse casi exclusivamente en referencia a bases de datos construidas a partir de sistemas informáticos que permiten una organización fácil y rápida de los datos. Las bases de datos informáticas pueden crearse a partir de aplicaciones informáticas o incluso en línea usando Internet. Las bases de datos se clasifican como estáticas – exclusivamente para lectura y almacenamiento de datos – o dinámicas – la información se modifica y puede ser actualizada.¹⁶

Las bases de datos de compuestos químicos tienen la peculiaridad de almacenar información estructural de las moléculas y sus propiedades ya sean experimentales o teóricas. Diversos organismos y centros de investigación alrededor del mundo han colectado y puesto a disposición pública bases de datos de compuestos químicos. Ejemplos de estas iniciativas son: ChEMBL,¹⁷ NCI,¹⁸ PubChem,¹⁹ y UNIIQUIM.²⁰ Otras colecciones son comercialmente disponibles, tales como GRAS y CODEX ALIMENTARIUS, entre otras. Artículos de revisión de estas bases de datos se encuentran publicados en la literatura.²¹⁻²³

En el área de alimentos se han dedicado esfuerzos a la construcción de bases de datos conteniendo, por ejemplo, información nutricional. Estas bases de datos son una vasta fuente de información que puede ser utilizada para analizar y predecir patrones de la información que contienen. Estas bases de datos pueden también compararse con otras bases de datos de referencia, por ejemplo, aquellas que tienen relevancia en salud como las generadas en el campo de la farmacia. En la Tabla 2, se describe la información contenida en bases de datos relevantes en alimentos y farmacia así como organismos

reguladores de sanidad alimentaria y farmacéutica. En muchos de los casos, estos importantes organismos se guían por datos estadísticos poblacionales, por ejemplo, de brotes epidémicos. Esto recalca la importancia del adecuado registro, control y análisis de bases de datos de información.

Tabla 2. Bases de datos y fuentes de información relevantes en las áreas de alimentos y farmacia.

| Nombre | Descripción | URL |
|---|---|---|
| GRAS | Base de datos que contiene aditivos utilizados en la industria alimentaria, farmacéutica y cosmética. | http://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/ |
| INFOODS (International Network of Food Data Systems) | Se creó en 1984 y es una red mundial de expertos en composición de los alimentos cuyo fin es mejorar la calidad, la disponibilidad, la fiabilidad y la utilización de los datos de composición de alimentos. Esta base de datos contiene tablas de densidad, datos analíticos, composición de los alimentos en diferentes países. La información está disponible en archivos PDF, EXCEL y algunas tablas en HTML. Las personas que pueden utilizar esta base de datos son programadores, generadores de datos (por ejemplo, químicos), usuarios de datos (por ejemplo, expertos en educación nutricional, epidemiólogos nutricionales, biólogos, estudiantes, consumidores), los tomadores de decisiones para la localización de recursos a nivel nacional e internacional. | http://www.fao.org/infoods/en |
| BEDCA | Es una red de centros de investigación públicos de administración e instituciones privadas cuyo objetivo es el desarrollo y mantenimiento de la Base de Datos Española de Composición de Alimentos. Está disponible en línea para todas aquellas personas que quieran conocer la composición nutricional de los alimentos. | http://www.bedca.net/ |
| Comisión Federal para la Protección contra Riesgos Sanitarios (Cofepris) | Dependencia federal del gobierno de México que está vinculada al Departamento de Regulación y Fomento Sanitario de la Secretaría de Salud. Normas Oficiales Mexicanas elaboradas por el Comité Consultivo Nacional de Normalización de Regulación y Fomento Sanitario las cuales establecen, las reglas, especificaciones, atributos, directrices, características o prescripciones aplicables a un producto, proceso, instalación, sistema, actividad, servicio o método de producción u operación, así como aquellas relativas a terminología, simbología, embalaje, marcado o etiquetado y las que se refieran a su cumplimiento o aplicación en materia de control | http://www.cofepris.gob.mx |

| | | |
|--|--|---|
| | y fomento sanitario. Información disponible en PDF. | |
| CODEX ALIMENTARIUS | <p>Establecida por la FAO y la OMS en 1963, desarrolla normas internacionales armonizadas alimentarias, directrices y códigos de práctica para proteger la salud de los consumidores y asegurar prácticas equitativas en el comercio de alimentos; promueve también la coordinación de todas las normas alimentarias acordadas por las organizaciones gubernamentales y no gubernamentales.</p> <p>Se pueden buscar las disposiciones por aditivo alimentario (nombre, sinónimo, número del SIN), por clase funcional, así como por categoría de alimentos. Información disponible en PDF.</p> | http://www.codexalimentarius.org/ |
| EFSA (European Food Safety Authority) | La labor de la EFSA abarca el análisis de riesgos y seguridad de alimentos en todas las etapas de la producción y el suministro de alimentos, desde la producción primaria, pasando por la seguridad de los piensos, hasta llegar al suministro de los alimentos a los consumidores. Recoge información y analiza los nuevos avances científicos, a fin de poder identificar y determinar los riesgos potenciales que amenazan a la cadena alimentaria. | http://www.efsa.europa.eu/ |
| FDA | Agencia del gobierno de los Estados Unidos de América responsable de la regulación de alimentos (tanto para personas como para animales), suplementos alimenticios, medicamentos (humanos y veterinarios), cosméticos, aparatos médicos (humanos y animales), productos biológicos y derivados sanguíneos. La FDA está dividida en 6 centros mayores: CBER (Centro de Evaluación Biológica e Investigación), CDRH (Centro de Aparatos y Salud Radiológica), CDER (Centro de Investigación y Evaluación de Medicamentos) , CFSAN (Centro de Seguridad de Alimentos y Nutrición Aplicada) , CVM (Centro de Medicina Veterinaria), y CTP (Centro de Productos de Tabaco). | http://www.fda.gov/ |
| Farmacopea de los Estados Unidos Mexicanos (FEUM) | Documento expedido por la Secretaría de Salud que consigna los métodos generales de análisis y los requisitos sobre identidad, pureza y calidad de los fármacos, aditivos, medicamentos, productos biológicos y demás insumos para la salud. | http://www.farmacopea.org.mx/ |
| Portal Farma | Bases de datos de medicamentos, calendario de vacunaciones, consulta pública de medicamentos, farmacias de guardia, etc. De la organización farmacéutica Colegial de España. Tiene distintos niveles de acceso a: ciudadanos, | http://www.portalfarma.com |

| | | |
|-------------------------------------|---|---|
| | estudiantes, farmacéuticos y médicos. Información en línea. | |
| P. R. Vademécum | Es un directorio especializado en medicamentos, farmacología, farmacia, medicinas, enfermedades y enfermería. Información en línea. | http://mx.prvademecum.com/ |
| MOSBY'S DRUG CONSULT top 200 | Una lista de los 200 mejores medicamentos genéricos de 2003. Artículo disponible en línea, HTML o PDF. | http://www3.us.elsevierhealth.com/DrugConsult/Top_200/ |

2.1.4. Química computacional, quemoinformática y bioinformática

Química computacional. La química computacional es una rama de la química que involucra el uso de computadoras para el cálculo de estructuras y propiedades de las moléculas (energía, momento dipolar, frecuencias vibracionales, etc.) así como el entendimiento de reacciones químicas. Hay varias áreas en este campo, por ejemplo, la representación computacional de átomos y moléculas, almacenamiento, organización y búsqueda de datos sobre entidades químicas, identificación de patrones, tendencias y correlaciones cuantitativas entre estructuras químicas y sus propiedades (QSPR, por sus siglas en inglés de *Quantitative Structure-Property Relationships*), modelos de síntesis de compuestos, diseño de moléculas que interaccionen con otras de forma efectiva, especialmente en diseño de fármacos.²⁴

La química computacional puede dividirse en dos grandes categorías:

- La mecánica molecular: que aplica las leyes de la física clásica al núcleo molecular sin considerar explícitamente a los electrones.
- La mecánica cuántica: se basa en la ecuación de Schrödinger para describir una molécula con un tratamiento directo de la estructura electrónica.²⁵

Quemoinformática. Originalmente la quemoinformática, también conocida como quimioinformática se definió por Brown (1998) como “El uso de la tecnología y la gestión de la información se ha convertido en un componente crítico del proceso de descubrimiento de medicamentos”. La quemoinformática es el uso de los recursos informáticos para transformar datos en información y la información en conocimiento, con el propósito de tomar mejores y más rápidas decisiones en el ámbito de la identificación y optimización de medicamentos”.²⁶ Existen algunos conceptos que se relacionan a la quemoinformática se pueden ver con más detalle en la Tabla 3.

Tabla 3. Conceptos asociados a la quemoinformática. ²⁶

| Término | Campo de acción |
|----------------------------|---|
| Quemoinformática | Recolección, análisis y gestión de datos químicos. Representación gráfica de análisis de datos y modelado molecular. Diseño y organización de bases de datos. Predicción de la estructura química y propiedades biológicas de análogos químicos. Análisis de la similitud y diversidad molecular. Optimización y diseño de compuestos y librerías químicas. Diseño de algoritmos y estrategias de búsqueda en bases de datos. Estudio cualitativo y cuantitativo de la estructura-actividad de las librerías químicas. Disertación teórica sobre la información aplicada a problemas químicos. Desarrollo de modelos estadísticos y nuevos descriptores en química orgánica. Predicción del comportamiento de moléculas orgánicas en seres vivos. |
| Química informática | Aplicación química de tecnologías de información. |
| Quimiométrica | Utilización de métodos estadísticos a datos químicos. |

Bioinformática. Según la definición del Centro Nacional para la Información Biotecnológica —*National Center for Biotechnology Information*—NCBI por sus siglas en inglés,¹⁹ "Bioinformática es un campo de la ciencia en el cual confluyen varias disciplinas tales como: biología, computación y tecnología de la información". El fin último de este campo es facilitar el descubrimiento de nuevas ideas biológicas así como crear perspectivas globales a partir de las cuales se puedan discernir principios unificadores en biología.²⁷

2.1.5. Similitud molecular

La similitud molecular proporciona las bases para clasificar, caracterizar y describir sistemáticamente a las moléculas. Esta organización del conocimiento ha originado teorías químicas que cada vez ganan más reputación. La similitud molecular puede entenderse como los patrones comunes sobre un grupo de moléculas representadas por alguna o algunas de sus propiedades.² Para cuantificar la similitud molecular se requiere al menos de dos elementos: una representación y una métrica.

La similitud molecular es uno de los conceptos más importantes en la quemoinformática. Ésta desempeña un papel importante en la predicción de las propiedades de los compuestos químicos y el diseño de productos químicos con un conjunto predefinido de propiedades. También es especialmente importante en la realización de estudios de

diseño y búsqueda de fármacos por medio de análisis de grandes bases de datos que contienen estructuras químicas disponibles o potencialmente disponibles.¹

2.1.6. Análisis multivariante: Análisis de componentes principales, análisis cluster o de agrupamiento.

El análisis multivariante es un conjunto de métodos estadísticos destinados a describir e interpretar los datos que provienen de la observación de varias variables estadísticas estudiadas conjuntamente.²⁸ Para que un análisis sea considerado multivariante, todas las variables deben de ser aleatorias y deben de estar interrelacionadas de tal manera que los diferentes efectos no puedan ser interpretados significativamente de manera independiente. Ejemplos de técnicas multivariantes son el Análisis de Componentes Principales y el Análisis de Agrupamiento (cluster). El análisis de componentes principales es una técnica estadística de síntesis de la información o reducción de dimensiones (número de variables).²⁹ La dimensionalidad de un conjunto de datos es el número de variables que se utilizan para describir cada objeto. El análisis de componentes principales proporciona un nuevo conjunto de variables además que se encuentra a menudo que gran parte de la variación en el conjunto de datos se puede explicar por un pequeño número de estos componentes principales. Los componentes principales son también mucho más convenientes para la visualización gráfica de los datos y el análisis.³⁰ El análisis de cluster (o análisis de agrupamiento) es una técnica de análisis exploratorio de datos para resolver problemas de clasificación. Su objetivo consiste en ordenar objetos (personas, cosas, animales, plantas, variables, etc.) en grupos, de forma que el grado de asociación/similitud entre miembros del mismo grupo sea más fuerte que el grado de asociación/similitud entre miembros de diferentes grupos. Cada agrupamiento o grupo se describe como la clase a la que sus miembros pertenecen. El análisis de agrupamiento es un método que permite descubrir asociaciones y estructuras en los datos que no son evidentes *a priori* pero que pueden ser útiles una vez que se han encontrado. Los resultados de un análisis de agrupamiento pueden contribuir a la definición formal de un esquema de clasificación, a sugerir modelos estadísticos para describir poblaciones, a asignar nuevos individuos a las clases para diagnóstico e identificación, etc.

2.2. Modelado molecular

Es un conjunto de métodos teóricos y técnicas computacionales con el fin de simular diferentes propiedades de sistemas moleculares y macromoleculares. Es una rama de la Química Computacional.³¹ En los métodos de modelado molecular se consideran las propiedades de las moléculas en tres dimensiones alguno de éstos métodos son el análisis conformacional, la mecánica cuántica, los *campos de fuerzas*, la termodinámica estadística y los gráficos moleculares interactivos. Estos métodos permiten la representación y manipulación de las moléculas en tres dimensiones, lo que proporciona una información espacial que es esencial para comparar moléculas y para estudiar la interacción entre ligandos y receptores macromoleculares.^{32, 33}

2.2.1. Análisis conformacional

El análisis conformacional es la exploración de los conformeros que se pueden obtener de una molécula dada al realizar torsiones alrededor de enlaces sencillos (grados de libertad conformacionales), observando los cambios en la energía molecular asociados a esas torsiones.³⁴

2.2.2. Superposición molecular

Consiste en el alineamiento o superposición de varias estructuras moleculares con el objetivo de observar cual o cuales moléculas tienen un alto grado de similitud o disimilitud con la molécula o moléculas de referencia. Para ello se superponen varias moléculas buscando el máximo grado de coincidencias posibles. Para series homólogas con similar actividad biológica la región molecular de máxima superposición podría contener al farmacóforo.³⁵

2.2.3. Acoplamiento molecular

El acoplamiento (*docking*) molecular es un método computacional que busca modos de unión entre ligandos y un blanco macromolecular (normalmente proteínas). Esta técnica requiere información precisa de la estructura tridimensional del blanco molecular, la cual se puede obtener utilizando las técnicas de cristalografía de rayos X, resonancia magnética nuclear o modelado por homología.³⁶⁻³⁸ El acoplamiento molecular es un método que predice la conformación y orientación preferida de una molécula cuando está en interacción con una biomacromolécula con el fin de formar un complejo estable. El

conocimiento de la orientación preferida a su vez puede ser usada para predecir la afinidad entre las dos moléculas usando las funciones de puntuación (o *scoring functions*). El objetivo de esto es simular computacionalmente el proceso de reconocimiento celular. Durante el estudio de acoplamiento molecular (*docking*) se busca alcanzar una conformación óptima tanto para la proteína como para su ligando.^{39, 40}

2.3. Base de datos GRAS

La base de datos GRAS contiene moléculas de aditivos utilizados en la industria alimentaria, farmacéutica y cosmética; un aditivo alimentario es aquella sustancia que se añade a los alimentos y bebidas con el objeto de proporcionar o intensificar aroma, color o sabor, prevenir cambios indeseables o modificar en general su aspecto físico quedando prohibido su uso para ocultar defectos de calidad.⁴¹

El término GRAS es la categoría de ingredientes de alimentos establecido por la Ley Federal de Alimentos, Medicamentos y Cosméticos de Estados Unidos. Esta ley se aplica a los productos químicos naturales o artificiales que puedan añadirse a los alimentos y bebidas dentro de un nivel de uso definido.⁴²⁻⁴⁵ Los aditivos alimentarios que fueron reconocidos como GRAS antes del 1 de Enero de 1958 fueron aquellos que cumplían con los siguientes requisitos: estar en uso antes de esa fecha, ser un ingrediente alimenticio, no un medicamento o remedio a base de hierbas, y el uso debe haber sido lo suficientemente amplio para dar garantías de que el compuesto no indujo ninguna alergia o intolerancia metabólica.⁴⁴ La categoría GRAS complementa la categoría “aditivos alimentarios” que requiere de una evaluación extensa de seguridad y revisión de la FDA y el proceso de aprobación. Una fracción significativa a la lista GRAS son revisadas por el Panel de Expertos FEMA que ha operado desde la década de 1960.⁴²⁻⁴⁶ Los datos que el panel de expertos de FEMA requiere para evaluar el reconocimiento GRAS están estandarizados y suelen incluir datos sobre el destino metabólico, un estudio de toxicología de dosis repetida y pruebas de efectos de enfermedades crónico degenerativas para un nuevo compuesto. La dosificación de los animales se lleva a cabo en un nivel elevado para demostrar un margen de seguridad. Sin embargo, el panel de FEMA evita el enfoque memorístico.⁴³ En el caso de que un nuevo compuesto tenga una fuerte similitud con compuestos seguros conocidos existentes, este proceso puede ser reducido.⁴⁶ FEMA se ha ganado una reputación distinguida para una acción eficaz y la excelencia científica, tanto en Estados Unidos como en el extranjero, a través de la prestación de sus

servicios a la industria del sabor durante más de 100 años. FEMA atribuye esto al alto grado de participación de sus miembros en las actividades sustantivas de la asociación.⁴⁷

FEMA ha desarrollado un programa innovador que utiliza el concepto generalmente reconocido como seguro (GRAS) para evaluar la seguridad de las sustancias aromatizantes. El Panel de Expertos de FEMA independiente determina el estado de GRAS de sustancias aromatizantes que resulta en una lista GRAS de más de 2.500 ingredientes para su uso por la industria. El panel de expertos revisa periódicamente el estado de las sustancias y proporciona una 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 y continúa ganando cada vez mayor aceptación en todo el mundo.⁴⁷

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. También participa en un Grupo Técnico Asesor EE.UU. que ayuda a desarrollar y revisar las normas ISO para los aceites esenciales.⁴⁷

FEMA apoya la armonización de las regulaciones estadounidenses y extranjeras relacionados con la fabricación, el uso y el etiquetado de los sabores de todo el mundo.⁴⁷

3. Marco teórico

El dolor es un mecanismo de alerta que todo ser humano ha experimentado en algún momento de su vida, éste permite identificar daño o peligro. Es también un padecimiento con diferentes grados de intensidad. A lo largo del tiempo, la humanidad ha identificado diversas sustancias particularmente para dolor intenso. Por ejemplo, extractos del opio se ha empleado desde hace alrededor de 5,000 años. Los extractos del opio se han utilizado por sus propiedades analgésicas, antidiarreicas y su *acción antitusígena*. El opio se obtiene de la planta solanácea *Papaver somniferum*, comúnmente conocida como amapola. La palabra “opio” por sí misma, deriva del nombre griego que significa “jugo” y el fármaco se obtiene del exudado lechoso que se extrae del tallo de la amapola. En 1806, el químico alemán Friederich Sertürner aisló por primera vez el principio activo del opio dándole el nombre de morfina en referencia a Morfeo dios de los sueños en la literatura griega.⁴⁸

En 1973, se descubrieron los receptores de opioides. Se encontraron en el sistema nervioso central y periférico, en el sistema inmune y en el sistema endócrino de mamíferos. Hay tres tipos de receptores de opioides: μ , δ y κ .⁴⁹ Los receptores de opioides interactúan con una gran variedad de ligandos, siendo los más estudiados los de origen alcaloide. Por otro lado, la mayoría de los ligandos endógenos son péptidos, los cuáles se comportan casi exclusivamente como agonistas. Los receptores y ligandos endógenos de los *sistemas opioidérgicos* son biosintetizados dentro de células, tejidos o compartimentos del sistema inmune, el sistema endocrino y el sistema nervioso central y periférico. La búsqueda de ligandos exógenos de los receptores opioides acompañó a la persecución de sus homólogos endógenos desde el inicio de la investigación sobre los sistemas opioidérgicos.

La detección de los productos alimenticios básicos con actividad opioide agonista o antagonista pronto tuvo éxito en su caracterización farmacológica.⁴⁹ La mayoría de los productos alimenticios básicos: leche, cereales, maíz, arroz, soya, huevos, y sangre como un componente de la carne, han sido estudiados como fuente de péptidos con actividad opioide agonista o antagonista. El primer componente alimentario identificado como un ligando del receptor opioide fue un fragmento de β -caseína bovina, más tarde seguido por la identificación de otros opioides agonistas o antagonistas derivados de proteínas de la leche. De hecho, una variedad de digeridos enzimáticos de proteínas de la leche, α -caseína, β -caseína, κ -caseína, α -lactoalbúmina, β -lactoglobulina, y lactotransferrina tienen

actividad agonista o antagonista.⁴⁹ La leche contiene numerosas secuencias de péptidos que afectan funciones fisiológicas cruciales y modulan muchos procesos de regulación. Por lo tanto, péptidos bioactivos de alimentos podrían ser utilizados como suplementos dietéticos, conservadores naturales seguros y eficaces, y nutracéuticos a base de leche. Este campo de investigación promete aportar nuevos péptidos biológicamente activos derivados de proteínas de la leche que afectan significativamente tanto las industrias de alimentos y salud pública.⁵⁰ Sin embargo, la aplicación generalizada de estos y otros ingredientes para fines alimenticios depende de la disponibilidad de datos científicos rigurosos que demuestren su eficacia y seguridad a las concentraciones y aplicaciones específicas utilizadas en alimentos de consumo humano.⁵¹ Dado que se ha comprobado que la leche contiene numerosas secuencias de péptidos con actividad opioide agonista o antagonista, se cree que varios derivados de alimentos podrían tener también esta actividad. En la Tabla 4 se muestra la caracterización de derivados proteicos de alimentos. Para dolor de menor intensidad se han desarrollado una amplia gama de analgésicos, así como el uso de plantas tal como *Eugenia caryophyllata* (*Syzygium aromaticum* L. Myrtaceae), de la cual se extrae el aceite esencial de clavo, que se usa en aplicaciones tópicas para aliviar el dolor y promover la curación; también se utiliza en la industria de saborizantes y fragancias.⁵² Otra fuente de origen natural es *Curcuma aromatica* Salisb, la cual ha sido utilizada como una planta medicinal y condimento. Esta planta tiene varias funciones fisiológicas incluyendo el alivio de dolor e inflamación, contribuye al sabor y prevención del cáncer.⁵³

Estudios recientes sugieren que la base de datos FEMA GRAS (GRAS 25) contiene moléculas que son similares a las moléculas de los fármacos aprobados en la base de datos DrugBank.⁶ Por tal razón, se puede postular que los aditivos de alimentos de consumo humano puedan tener algún efecto biológico de interés en salud o bienestar humano.

Tabla 4. Péptidos de alimentos con actividad opioide agonista o antagonista. ⁴⁹

| Péptido | Fragmento de la proteína | Secuencia | Actividad agonista / antagonista | Selectividad del receptor opioide (μ, δ, κ) |
|---|--|-------------------------|----------------------------------|---|
| α B-casein exorphin (1-7) | α B-Casein (90-96) | RYLGYLE | Agonista | $\mu / \delta \ll \kappa$ |
| α B-casein exorphin (2-7) | α B-Casein (91-96) | YLGYLE | Agonista | μ / δ |
| Casoxin D | α S1-(H)-Casein (158-164) | YVPFPPF | Antagonista | μ / δ |
| α S1(H) -Casomorphin (1-5) | α S1-(H)-Casein (158-162) | YVPFP | Agonista / Antagonista | $\mu / \delta \ll \kappa$ |
| α S1(H) -Casomorphin (1-5) - NH ₂ | α S1-(H)-Casein (158-162) | YVPFP-NH ₂ | Agonista / Antagonista | $\mu \ll \delta / \kappa$ |
| β B-Casomorphin (1-7) | β B- Casein A2 (60-66) | YPFPGPI | Agonista | $\mu > \delta \gg \kappa$ |
| Morphiceptin | β B- Casein A2 (60-63) | YFPF-NH ₂ | Agonista | $\mu \gg \delta \gg \kappa$ |
| Neocasomorphin | β B- Casein (114-119) | YPVEPF | Agonista | (μ) |
| β WB-Casomorphin (1-8) | β WB- Casein (60-67) | YPFPGPIP | Agonista | $\mu \gg \delta, \kappa$ |
| β H-Casomorphin (1-5) | β H- Casein (51-55) | YPFVE | Agonista | ($\mu / \delta \gg \kappa$) |
| β H-Casorphin | β H- Casein (41-44) | YPSF-NH ₂ | Agonista | (μ) |
| Valmuceptin | β H- Casein (51-54) | YPFV-NH ₂ | Agonista | $\mu > \delta$ |
| N.N. β H- Casein fragment | β H- Casein (59-63) | YGFLP | Agonista | (μ) |
| Casoxin A | κ B- Casein (35-41) | YPSYGLN | Antagonista | (μ) |
| Casoxin C | κ B- Casein (25-34) | YIPIQYVLSR | Antagonista | μ |
| Casoxin-5 | κ B- Casein (34-38) | RYPSYO-CH ₃ | Antagonista | $\mu > \kappa$ |
| Casoxin-6 | κ B- Casein (33-38) | SRYPSY-OCH ₃ | Antagonista | $\mu > \kappa$ |
| α B- Lactorphin | α B- Lactalbumin (50-53) | YGLF-NH ₂ | Agonista | (μ) |
| Des-NH ₂ - α B- Lactorphin | α B- Lactalbumin (50-53) | YGLF | Agonista | (μ) |
| α H- Lactorphin | α H- Lactalbumin (50-53) | YGLF-NH ₂ | Agonista | (μ) |
| Des-NH ₂ - α H- Lactorphin | α H- Lactalbumin (50-53) | YGLF | Agonista | (μ) |
| β B- Lactorphin | β B- Lactoglobulin (102-105) | YLLF-NH ₂ | Agonista | (μ) |
| Des-NH ₂ - β B- Lactorphin | β B- Lactoglobulin (102-105) | YLLF | (¿Agonista?) | (μ) |
| Lactoferrin A | (H) Lactoferrin (318-323) | YLGSGY-OCH ₃ | Antagonista | μ |
| Lactoferrin B | (H) Lactoferrin (536-540) | RYYGY-OCH ₃ | Antagonista | μ |
| Lactoferrin C | (H) Lactoferrin (673-679) | KYLGQY-OCH ₃ | (¿Antagonista?) | (¿ μ ?) |
| Gluten exorphin A5 | Gluteninas de alto peso molecular | GYPT | Agonista | $\mu \ll \delta$ |
| Gluten exorphin B5 | | YGGWL | Agonista | $\mu \ll \delta$ |
| Gluten exorphin C5 | Precursor de Glutenina HMW cadena 1By9 (624-628) | YPISL | Agonista | $\mu < \delta$ |
| Gliadorpin (1-7) | α -Gliadin (43-49) | YPQPQPF | Agonista | |
| Oryzatensin (1-9) | Precursor alergeno RA5B (47-55) | GYPMYPLPR | Agonista | μ |
| Oryzatensin (5-9) | Precursor alergeno RA5B (51-55) | YPLPR | Agonista | (¿ μ ?) |
| Serorphin | Albúmina de suero de bovino (399-404) | YGFQNA | Agonista | $\mu < \delta$ |
| Hemorphin (1-4) | Hemoglobina de bovina cadena β (34-37) | YPWT | Agonista | $\mu / \delta > \kappa$ |

4. Metodología

En este trabajo se analizaron moléculas de aditivos de la base de datos FEMA GRAS (Flavor-Base Pro © 2010)⁴ y péptidos de limentos. En la Figura 1, se muestra la metodología utilizada en este trabajo que se divide en tres partes, bases de datos, búsqueda y selección, y derivados proteicos alimenticios.

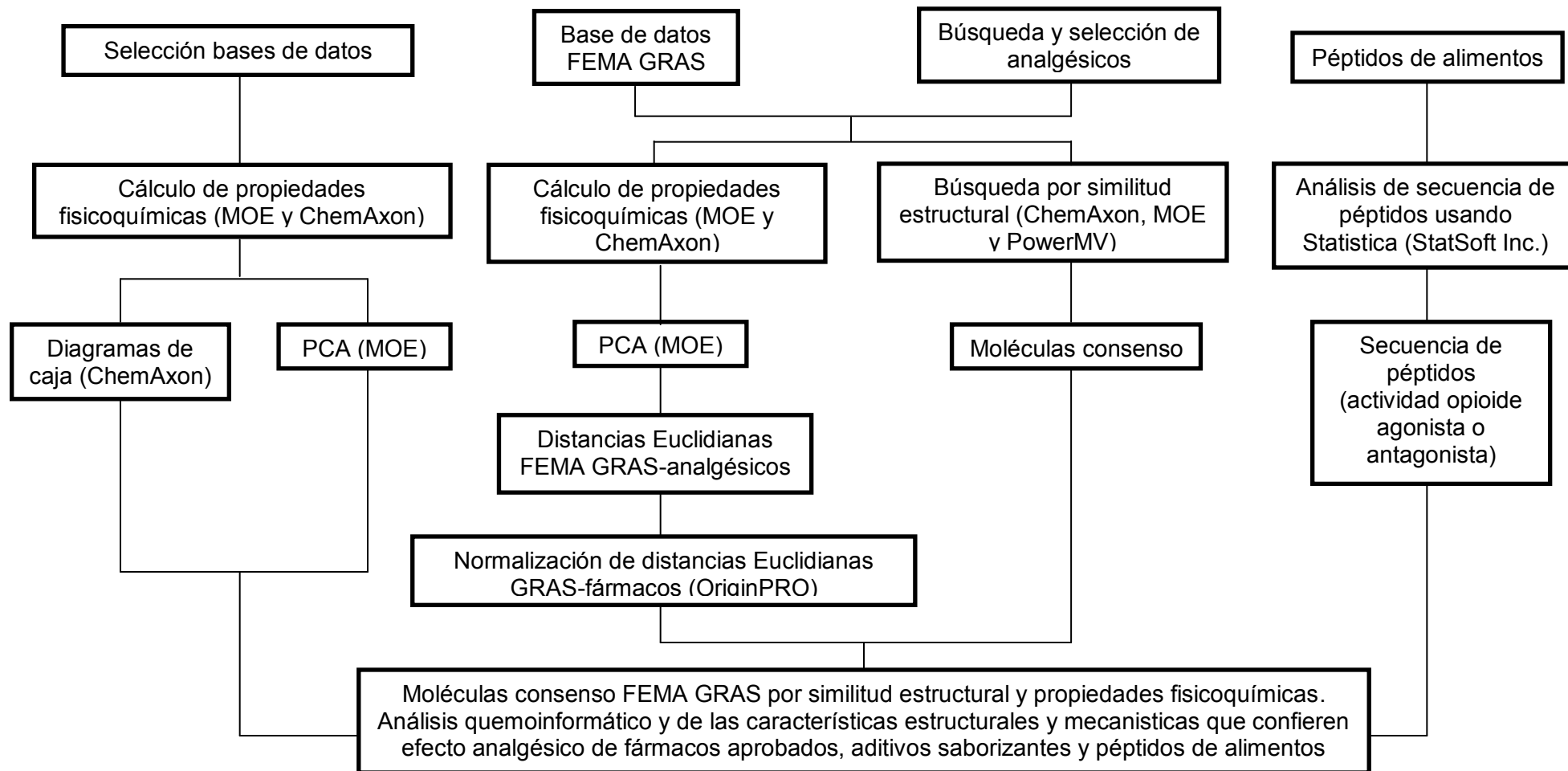


Figura 1. Etapas involucradas en este estudio.

Este proyecto es de carácter computacional. La obtención, almacenamiento y análisis de datos se realizó en una computadora personal con el sistema operativo Windows 7 Professional, Microsoft Office 2010. La paquetería de cómputo especializada que se utilizó fue ChemAxon,⁵⁴ PowerMV,⁵⁵ MOE (*Molecular Operating Environment*)¹¹ y ChemDraw.

4.1. Procedimiento

4.1.1. Comparación de bases de datos

La base de datos FEMA GRAS se comparó con bases de datos estándares, utilizadas de manera rutinaria en análisis quimoinformáticos. Éstas son: ChemDiv DNMT, DrugBank, GRAS, SpecsNP (*Specs Natural Products*), SpecsWD3 (*Specs World Diversity set 3*) y TCM (*Traditional Chinese Medicine*). Los programas utilizados fueron ChemAxon y MOE. El tipo y número de moléculas contenidas en cada base de datos utilizadas en este trabajo se resume en la Tabla 5. Para la comparación de estas bases de datos se calcularon las siguientes propiedades fisicoquímicas: coeficiente de partición octanol/agua (logP y SlogP), área de superficie polar topológica (TPSA o PSA), número de átomos donadores de enlaces de hidrógeno (a_don), número de átomos aceptores de enlaces de hidrógeno (a_acc), número de enlaces rotables (b_rotN), peso o masa molecular (MW) y número de anillos (RC); (Ver Anexo 1). Posteriormente, utilizando diagramas de caja elaborados en el programa de ChemAxon, se analizó la distribución de cada una de las propiedades fisicoquímicas evaluadas.

Tabla 5. Bases de datos analizadas en este trabajo.

| Base de datos | Compuestos | Número de moléculas |
|---------------------|---|---------------------|
| ChemDiv DNMT | Set de referencia de moléculas enfocadas a DNMT | 26944 |
| Drug Bank | Fármacos aprobados por la FDA | 1713 |
| FEMA GRAS (GRAS 25) | Aditivos saborizantes | 2244 |
| SpecsNP | Productos naturales | 467 |
| SpecsWD3 | Moléculas diversas | 10000 |
| TCM | Medicina tradicional china | 32357 |

4.1.2. Espacio químico

Una vez calculadas las propiedades fisicoquímicas de cada base de datos se llevó a cabo el análisis de componentes principales implementado en el programa MOE. La representación gráfica en dos dimensiones del espacio químico se realizó para cada una de las bases de datos utilizando las propiedades fisicoquímicas antes mencionadas, calculadas en ChemAxon y MOE.

También se realizó análisis de componentes principales para la base de datos FEMA GRAS incluyendo el conjunto de referencia de los fármacos para dolor. Las propiedades fisicoquímicas utilizadas fueron también las mencionadas anteriormente. Finalmente, se calculó la distancia euclidiana de cada molécula FEMA GRAS con cada uno de los fármacos utilizados considerando las 7 y 12 dimensiones generadas en el análisis de componentes principales para ChemAxon y MOE, respectivamente. Posteriormente estos datos fueron normalizados en una escala de 0 a 1.

4.1.3. Búsqueda por similitud estructural

Se llevó a cabo la comparación estructural de los fármacos para dolor respecto a la base de datos de FEMA GRAS utilizando tres programas: ChemAxon, PowerMV y MOE, el análisis se realizó mediante descriptores, éstos se describen en el Anexo 2. Una vez terminado el análisis se realizó un consenso para obtener aquellas moléculas que son similares a los fármacos que alivian el dolor. También se hizo una comparación por propiedades fisicoquímicas.

4.1.4. Alineamiento o superposición molecular

Se efectuó un análisis de alineamiento o superposición molecular entre los 10 fármacos para el dolor que se usaron en este proyecto con el programa MOE. El objetivo de este análisis fue identificar los pares de fármacos más similares dentro de los 10 identificados como más diversos. El alineamiento molecular implica la comparación en tres dimensiones, mientras que el análisis de diversidad se realizó utilizando descriptores en dos dimensiones.

4.1.5. Comparación de secuencias de péptidos de alimentos.

Usando la secuencia de péptidos descritos en la Tabla 4 se realizó un análisis estadístico en el programa StatSoft Inc. A través de este análisis se examinó la secuencia de péptidos y su relación con su actividad opioide agonista o antagonista así como la selectividad hacia receptores opioides μ y κ .

5. Resultados y discusión

5.1. Caracterización de las bases de datos basada en propiedades fisicoquímicas

La distribución de las propiedades fisicoquímicas de las bases de datos (ChemDiv DNMT, DrugBank, FEMA GRAS, SpecsNP, SpecsWD3 y TCM) se calcularon usando dos programas ChemAxon y MOE, para los cuales se utilizaron las siguientes propiedades fisicoquímicas: logP y SlogP, TPSA o PSA, a_don, a_acc, b_rotN, MW, RC. Además de estas propiedades, se calcularon el número de átomos de bromo (a_nBr), cloro (a_nCl), flúor (a_nF) y yodo (a_nI) en MOE. En las Figuras 2 a 5 se muestra la distribución de las 7 propiedades fisicoquímicas evaluadas en ambos programas para cada base de datos usando diagramas de caja.

La distribución de propiedades calculadas tanto con ChemAxon como con MOE muestra que la base de datos de TCM es la más diferente respecto a las otras bases de datos estudiadas (Figuras 2-5). En el caso de las bases de datos Chemdivdnmt, DrugBank, FEMA GRAS, SpecsNP y SpecsWD3 se observó que las distribuciones de las propiedades fisicoquímicas estudiadas fueron comparables, lo cual se ve reflejado en los valores de media y mediana de las distribuciones (Figuras 2-5). De manera interesante, en general, para la base de datos FEMA GRAS se obtuvieron distribuciones con media y mediana menores a las otras bases de datos. Esto muestra que la base de datos FEMA GRAS puede ser una fuente atractiva de moléculas para realizar modificaciones estructurales. Es decir, que dado su bajo peso molecular, número de enlaces rotables, etc (relativo por ejemplo, a DrugBank) es posible realizar sustituciones en estas estructuras y las propiedades fisicoquímicas de las moléculas resultantes estarán aún dentro de los valores admitidos para fármacos. En el Anexo 1 se describen cada una de las propiedades fisicoquímicas evaluadas para cada programa.

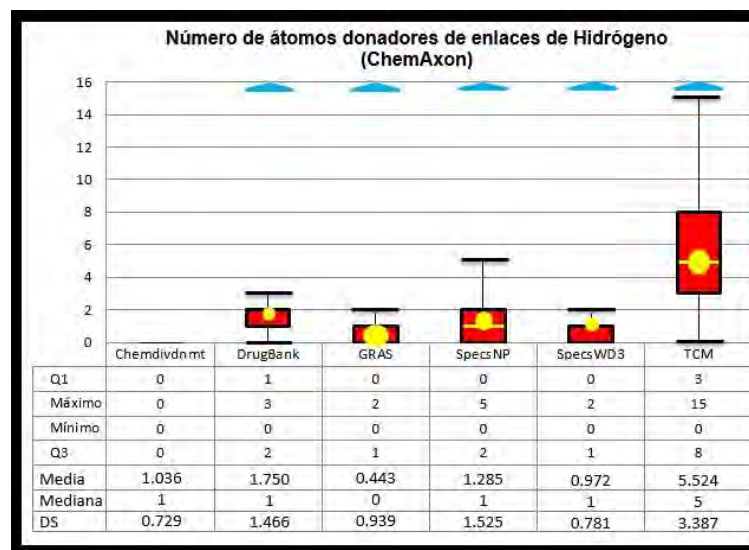
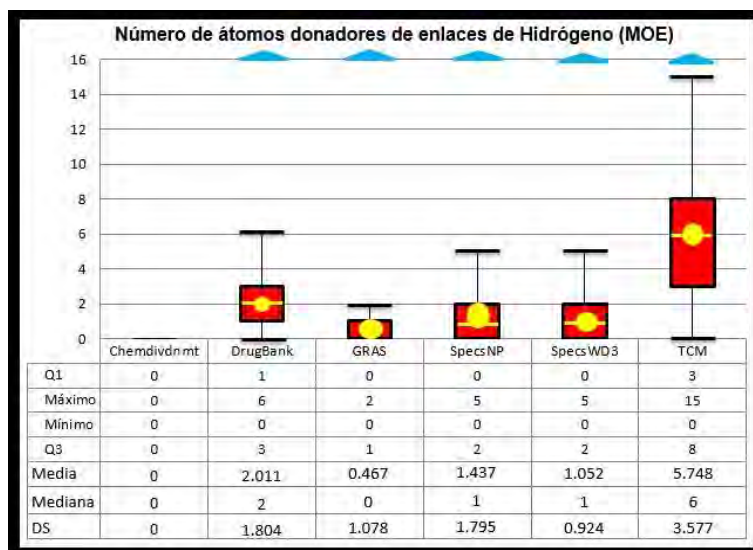
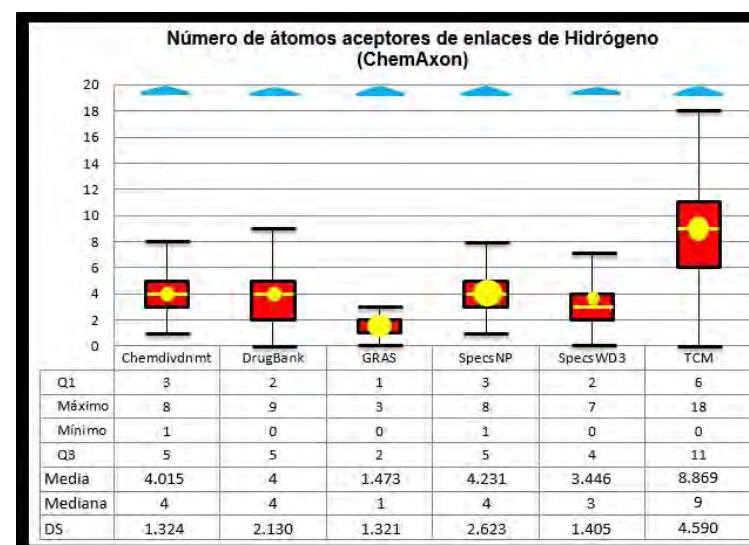
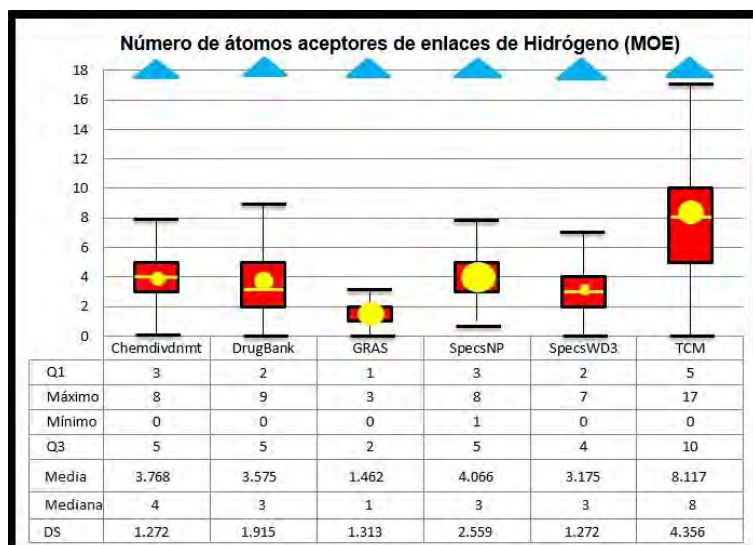


Figura 2. Diagramas de caja. Distribución de las propiedades fisicoquímicas, número de átomos aceptores y donadores de enlaces de hidrógeno de las bases de datos comparadas en este trabajo. Los cuadros de color rojo encierran los datos con valores dentro del primer y tercer cuartil; los bigotes (líneas de arriba y abajo) indican los valores máximo y mínimo de la muestra; las líneas amarillas muestran la mediana de todos los valores incluyendo los valores extremos; los círculos amarillos muestran la medida de todos los valores y los triángulos azules indica valores atípicos. También se muestran las estadísticas de cada distribución.

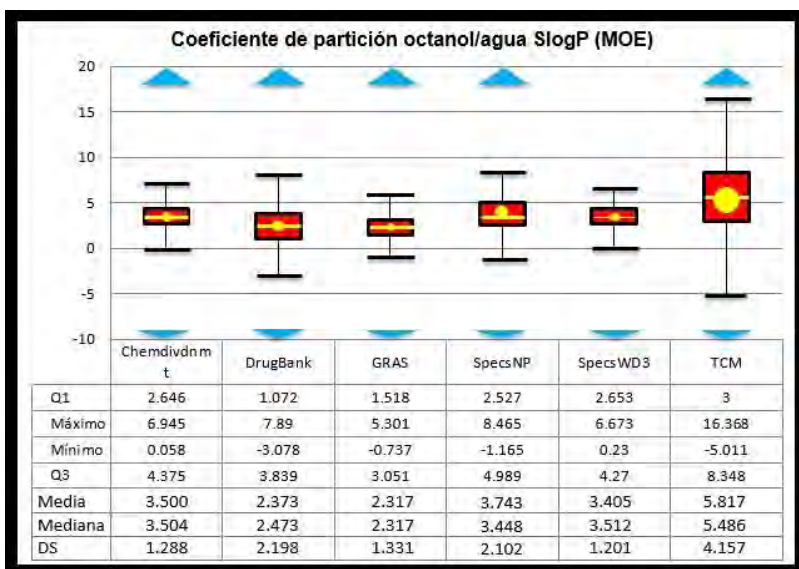
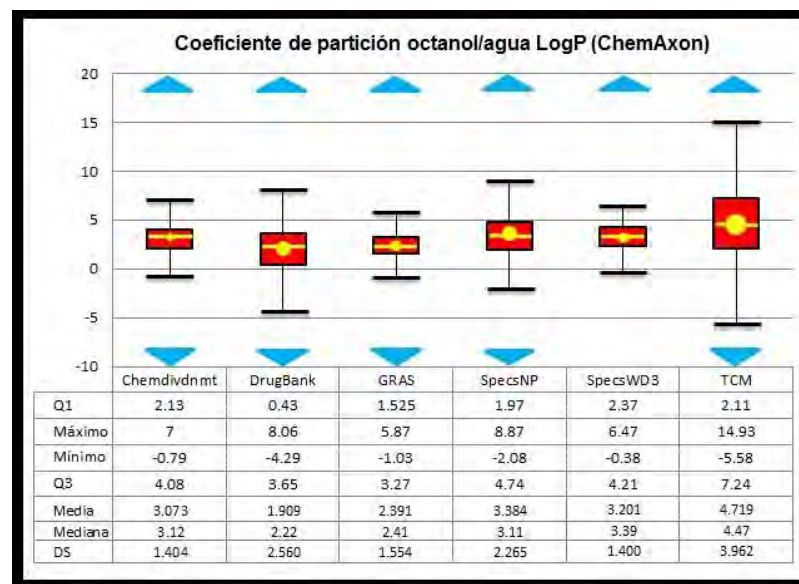
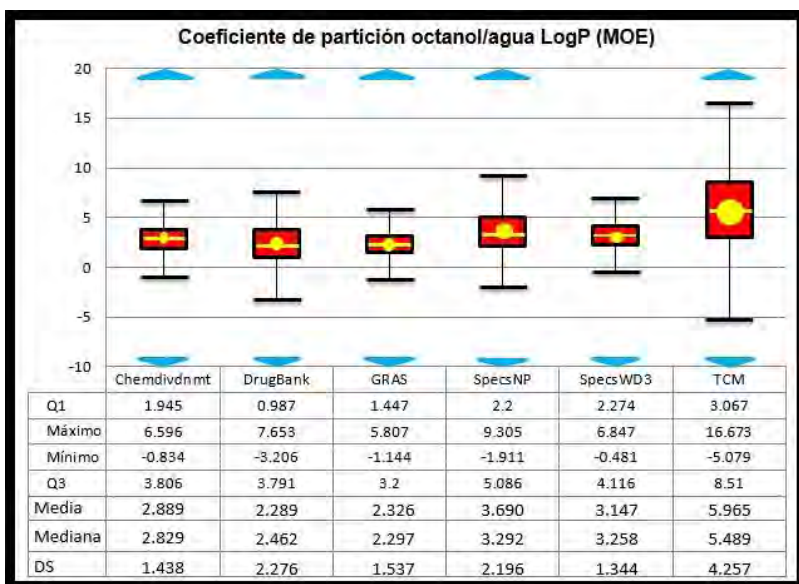


Figura 3. Diagramas de caja. Distribución de las propiedades fisicoquímicas, coeficiente de partición octanol/agua (logP y SlogP) de las bases de datos comparadas en este trabajo. Los símbolos y colores en los diagramas son los mismos que en la Figura 2.

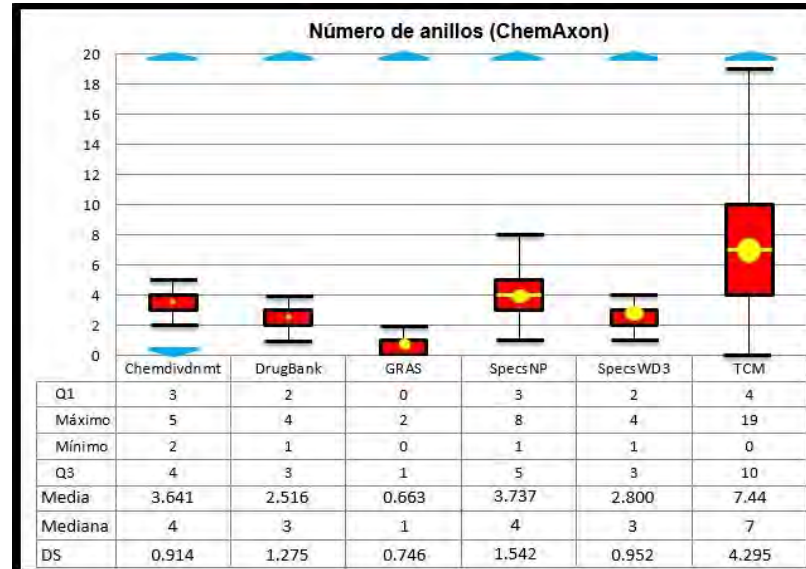
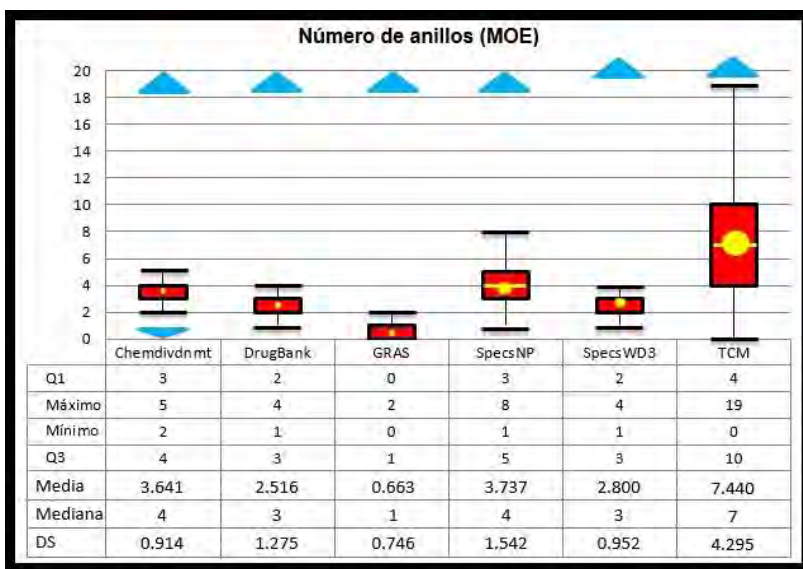
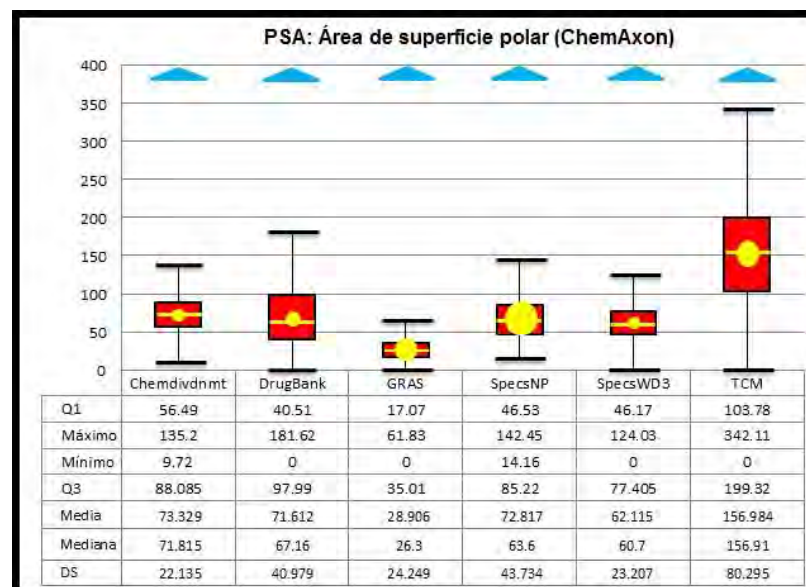
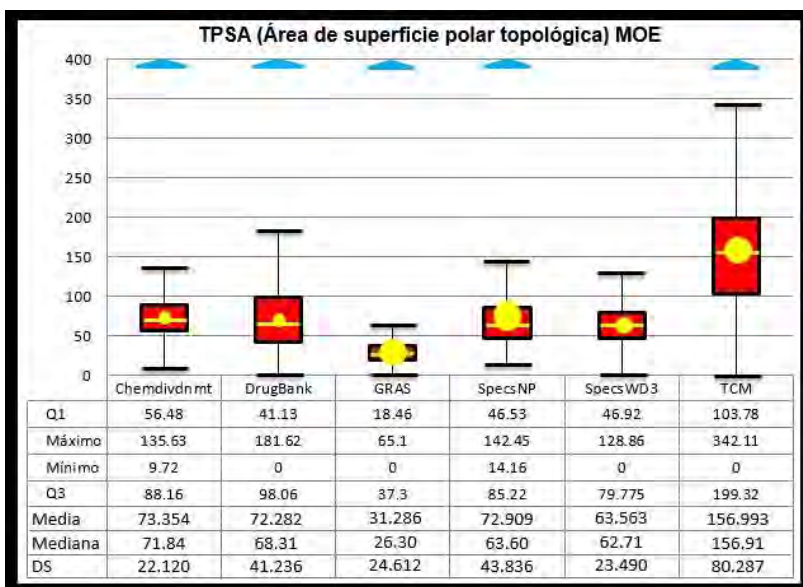


Figura 4. Diagramas de caja. Distribución de las propiedades fisicoquímicas, área de superficie polar topológica (TPSA y PSA); y número de anillos de las bases de datos comparadas en este trabajo. Los símbolos y colores en los diagramas son los mismos que en la Figura 2.

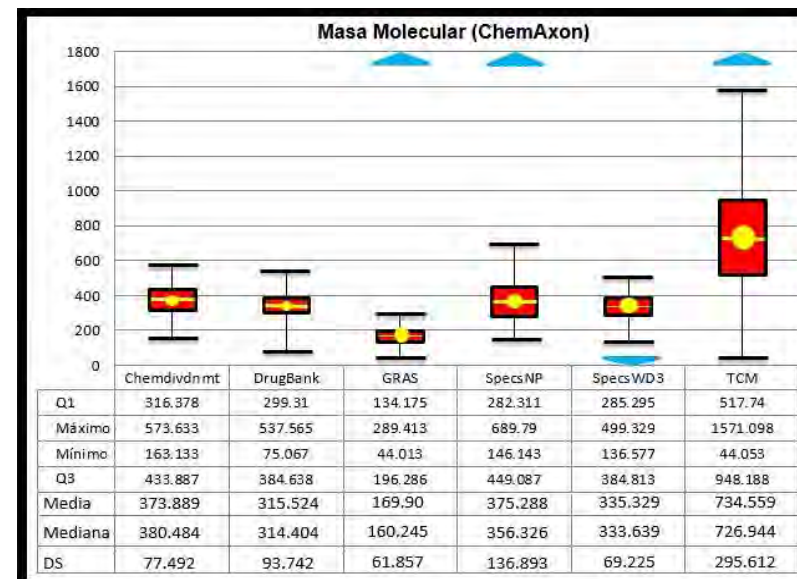
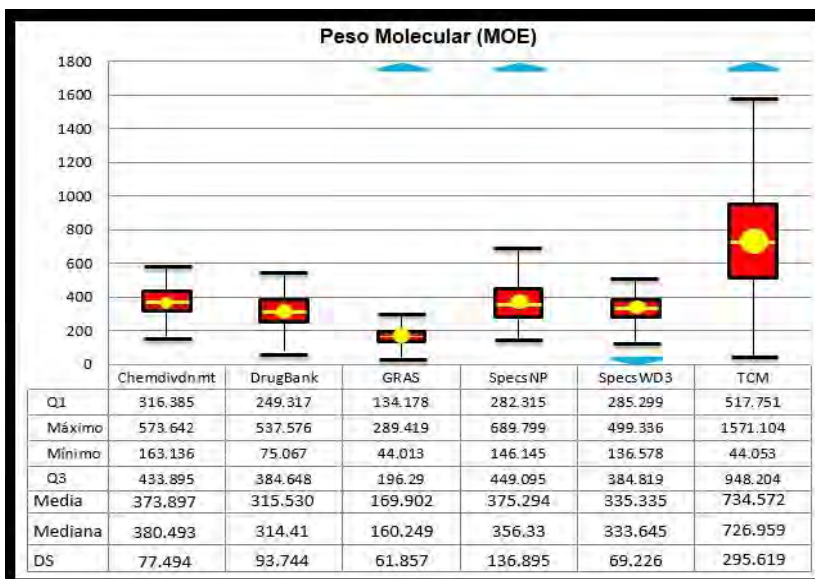
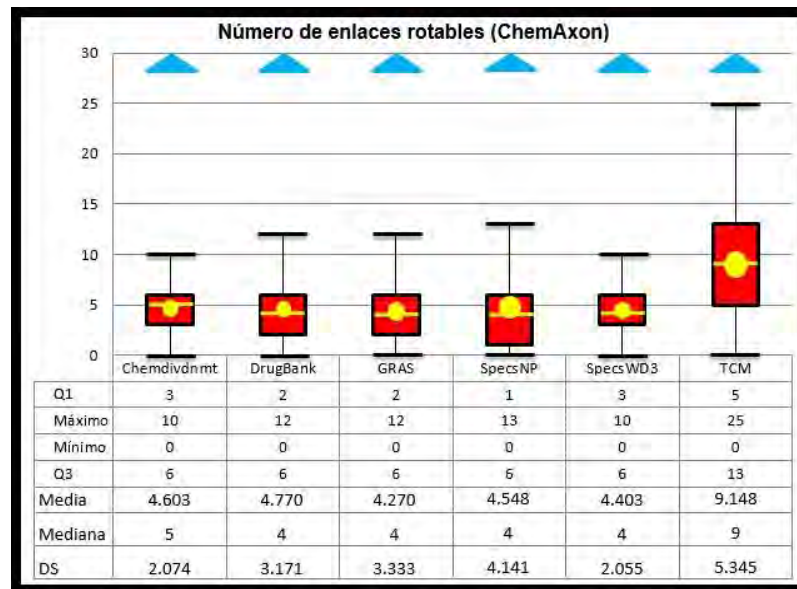
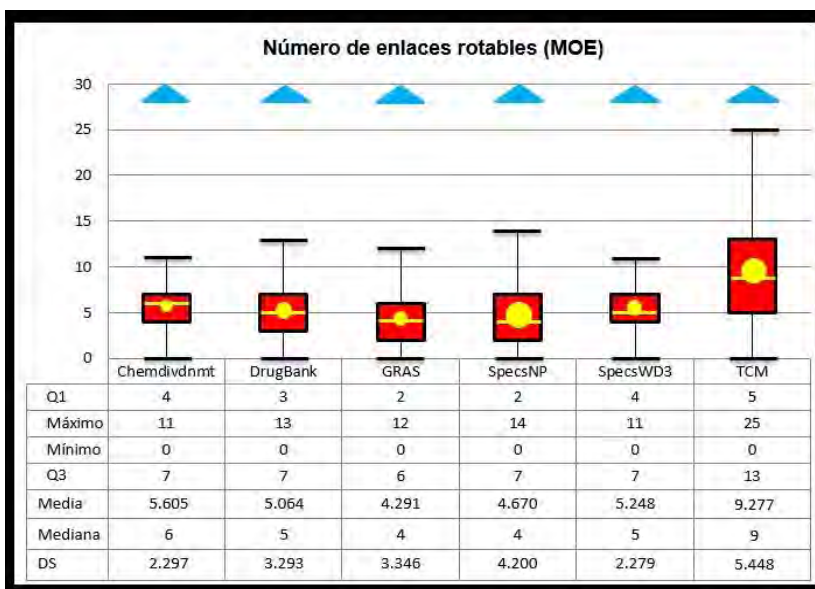


Figura 5. Diagramas de caja. Distribución de las propiedades fisicoquímicas, número de enlaces rotables y peso molecular de las bases de datos comparadas en este trabajo. Los símbolos y colores en los diagramas son los mismos que en la Figura 2.

Los resultados obtenidos de este análisis se compararon con dos artículos. El primer artículo utilizado fue “Chemoinformatic Analysis of GRAS (Generally Recognized as Safe) Flavor Chemicals and Natural Products”.⁶ En ese trabajo se describen cinco bases de datos: DrugBank, FEMA GRAS, AnalyticonNP (*Analyticon Natural Products*), SpecsNP y SpecsWD3, utilizando el programa Canvas.⁶ Las propiedades fisicoquímicas evaluadas en ese estudio fueron: logP, PSA, a_don, a_acc, b_rotN y MW. En ese trabajo se reportó que la base de datos GRAS tiene una diversidad estructural similar a las moléculas de los fármacos aprobados en la base de datos de DrugBank. El segundo artículo fue “A comparison of the chemical properties of drugs and FEMA/FDA notified GRAS chemical compounds used in the food industry”.⁵ En este estudio se analizaron tanto la base de datos FEMA GRAS como medicamentos comercializados. Para la descripción de la bases de datos utilizaron el programa MOE. Las propiedades fisicoquímicas evaluadas en ese estudio fueron: MW, enlaces rotables, logP, modelo de solubilidad Labute (logS), a_acc, a_don, número de átomos ácidos (a_acid) y básicos a_base, número de átomos de halógeno (F, Cl, Br o I) y número de átomos aromáticos. En ese trabajo se demostró que el espacio químico entre FEMA GRAS y los fármacos son diferentes y fácilmente reconocibles. Sin embargo, los compuestos más recientemente añadidos a la lista FEMA GRAS parecen más similares a los productos farmacéuticos con respecto a sus propiedades químicas. Por lo tanto, la distribución de las propiedades fisicoquímicas de los compuestos FEMA GRAS son similares a los reportados previamente en el trabajo anterior⁵ y en el realizado por Medina-Franco y colaboradores.

5.2. Espacio químico por base de datos

Se generó el espacio químico de las seis bases de datos utilizando los siguientes descriptores logP y SlogP, TPSA o PSA, a_don, a_acc, b_rotN, MW, RC, a_nBr, a_nCl, a_nF y a_nI. La visualización que se muestra en la Figura 6 se obtuvo de graficar los dos primeros componentes principales del análisis de componentes principales usando el programa de MOE. Para mayor claridad visual se representa cada base de datos individual. Las coordenadas corresponden a los dos primeros componentes principales, que reflejan a las propiedades fisicoquímicas que explican la mayor varianza. En la Tabla 6 se resume la varianza explicada por los dos primeros componentes principales y las propiedades fisicoquímicas que tienen mayor contribución para cada componente principal (CP). Un CP es una combinación lineal de las variables originales (en este caso propiedades fisicoquímicas), que tienen varianza máxima.

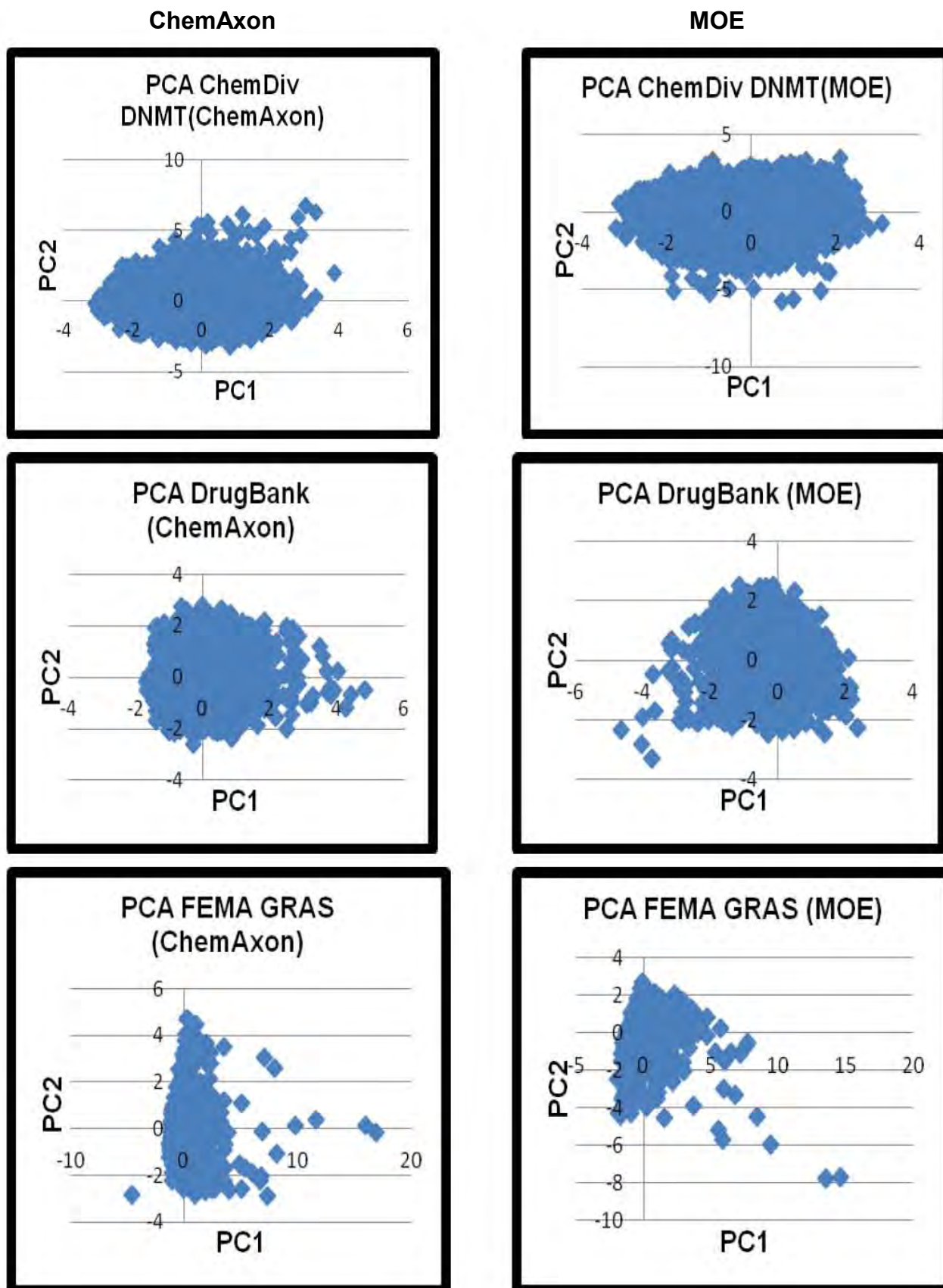


Figura 6. Representación de PCA en dos dimensiones. Espacio químico generado en base a 7 y 12 propiedades fisicoquímicas normalizadas. Ver detalles en la sección de Metodología.

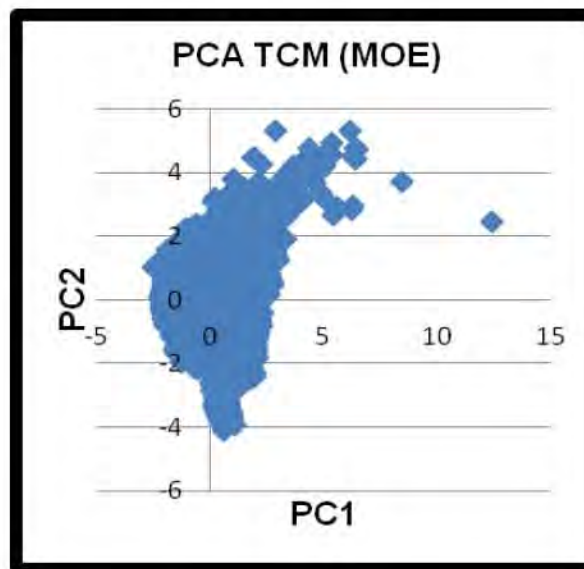
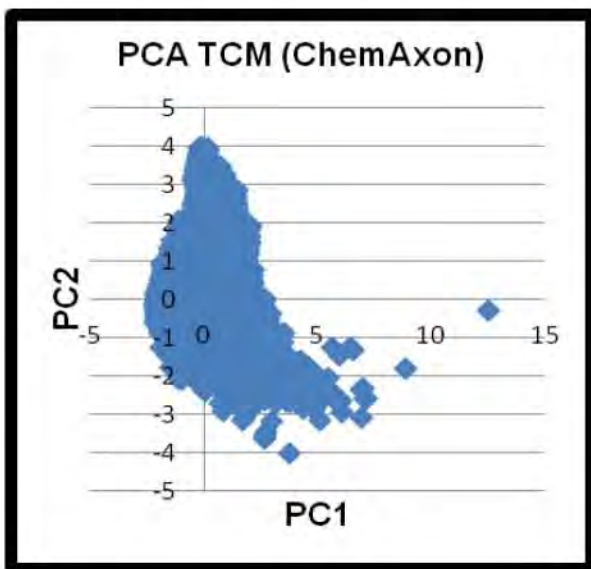
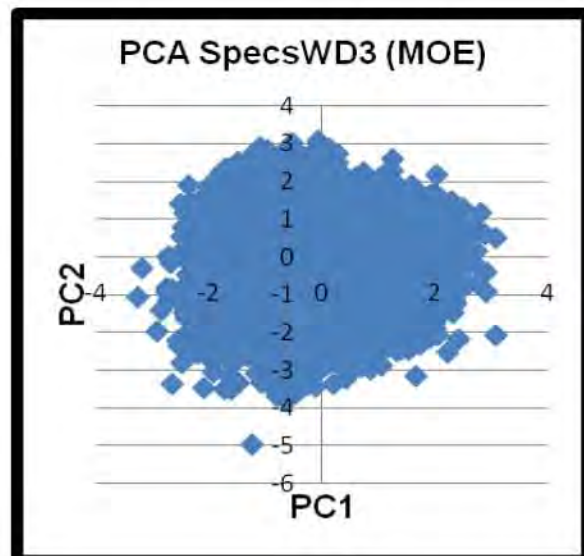
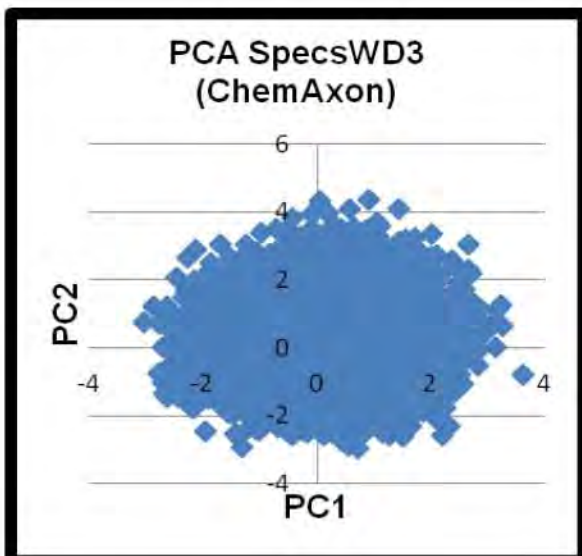
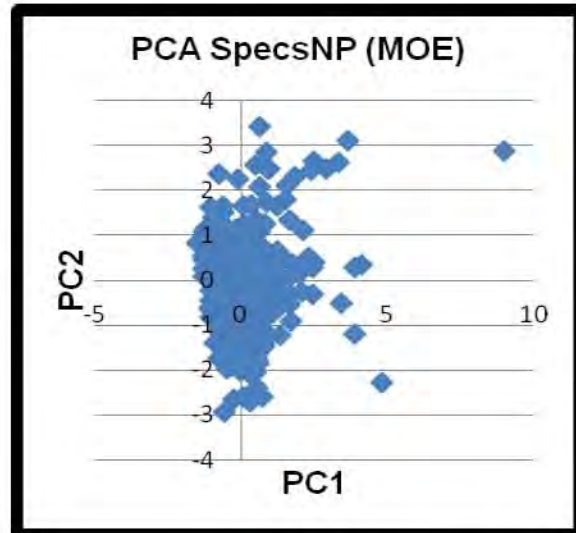
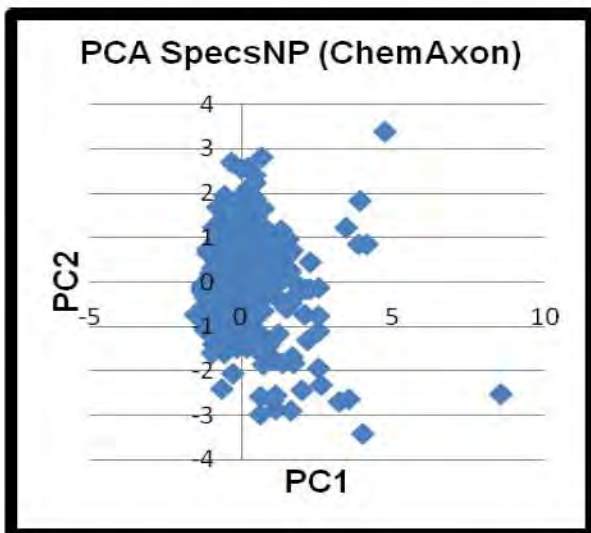


Figura 6. Continuación.

En la Figura 6 se observa que la representación 2D del espacio químico es dependiente del programa utilizado. Es importante entonces hacer notar que la utilización de diferentes programas y representaciones permite disminuir el riesgo de llegar a conclusiones sesgadas por el programa o representación empleados. Aquellas conclusiones que sean independientes del programa o representación serán consideradas como más confiables.

En la Figura 6 también se observa que en la distribución de los datos hay diferencias en cuanto a la ubicación en el plano de las moléculas contenidas en cada una de las bases de datos exploradas en este trabajo. Las diferencias observadas pueden ser debidas a que en el programa de ChemAxon se evaluaron siete propiedades mientras que para MOE fueron doce las propiedades evaluadas por lo tanto se esperaba que los datos no fueran idénticos. También se muestra que las propiedades a_don y a_acc son las que tienen mayor impacto en el porcentaje de varianza.

Tabla 6. Varianza explicada por los dos primeros PCs y propiedades fisicoquímicas que tienen mayor aportación para cada uno de ellos.

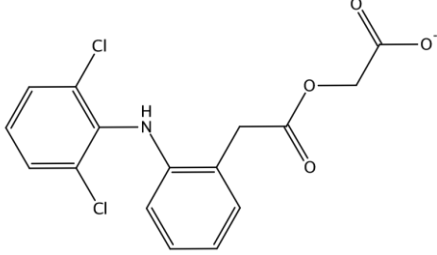
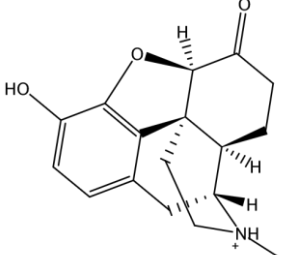
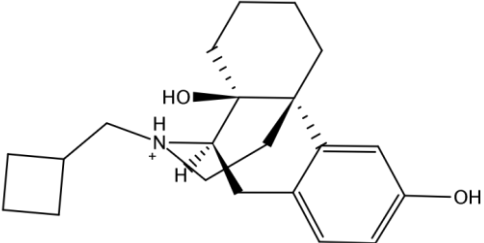
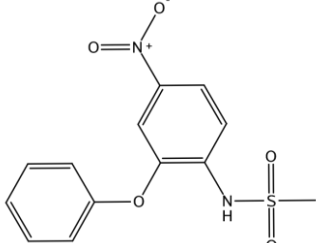
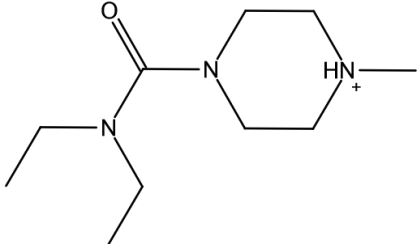
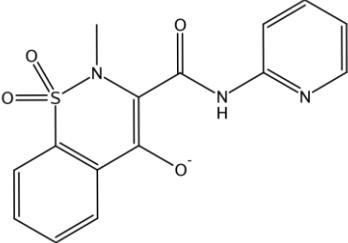
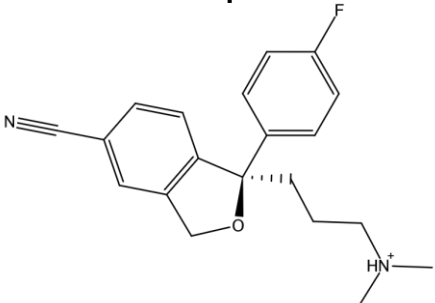
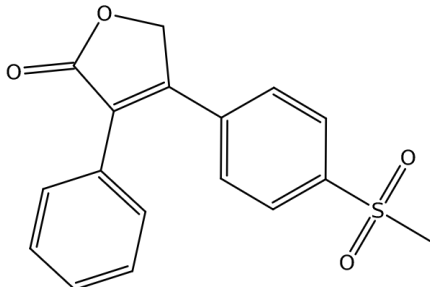
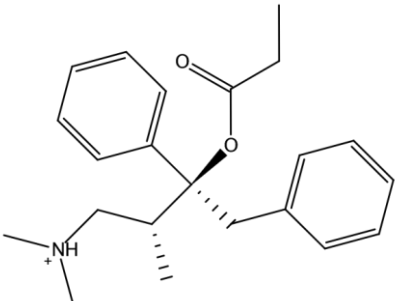
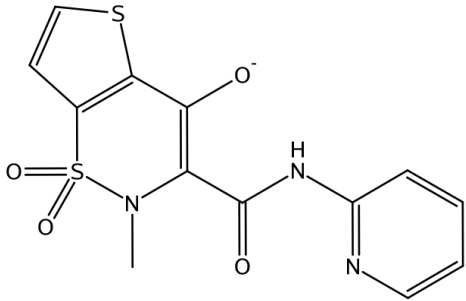
| Base de datos | ChemAxon | | | MOE | | |
|----------------------------|------------|---------------|---------------|------------|---------------|---------------|
| | % varianza | CP1 | CP2 | % varianza | CP1 | CP2 |
| Chemdiv DNMT | 65.999 | RC y a_acc | a_don y a_acc | 47.366 | RC y SlogP | a_nCl y a_nBr |
| DrugBank | 74.740 | a_don y a_acc | a_don y PSA | 52.367 | a_nI y a_nBr | a_nBr y TPSA |
| FEMA GRAS (GRAS 25) | 75.365 | a_don y a_acc | logP y b_rotN | 60.862 | a_nCl y a_don | TPSA y WM |
| SpecsNP | 72.795 | a_don y a_acc | logP y RC | 50.905 | a_acc y a_don | a_don y a_acc |
| SpecsWD3 | 63.129 | a_acc y a_don | a_don y a_acc | 45.229 | RC y SlogP | SlogP y logP |
| TCM | 82.280 | a_don y a_acc | logP y RC | 61.043 | a_don y a_acc | a_nCl y a_don |

5.3. Búsqueda por similitud estructural.

Desde hace mucho tiempo se han buscado fármacos que alivien el dolor. En la actualidad se tienen aprobados más de 100 para uso clínico. Como primer paso se realizó una búsqueda bibliográfica para identificar los analgésicos actualmente aprobados y se seleccionaron 10 en base a un análisis de diversidad estructural que se realizó utilizando el programa de MOE. Este análisis consiste en seleccionar una molécula al azar, después calcular la similitud de cada una de las moléculas de la base de datos de fármacos y seleccionar la molécula más disimilar a la molécula inicial, en la segunda iteración se selecciona a la molécula más disimilar a las dos anteriores y así sucesivamente hasta

seleccionar las 10 moléculas más diversas. Estos 10 fármacos de referencia se muestran en la Tabla 7.

Tabla 7. Fármacos para dolor utilizados en este trabajo.

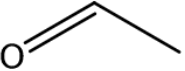
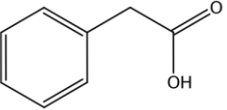

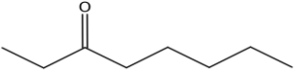
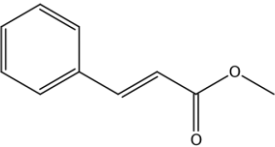
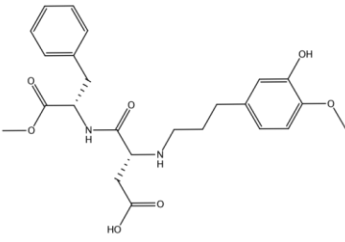
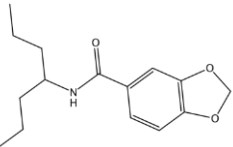
| | |
|---|---|
| <p style="text-align: center;">Aceclofenaco</p>  | <p style="text-align: center;">Hidromorфона</p>  |
| <p style="text-align: center;">Butorfanol</p>  | <p style="text-align: center;">Nimesulida</p>  |
| <p style="text-align: center;">Carbamazina</p>  | <p style="text-align: center;">Piroxicam</p>  |
| <p style="text-align: center;">Citalopram</p>  | <p style="text-align: center;">Rofecoxib</p>  |
| <p style="text-align: center;">Dextropropoxifeno</p>  | <p style="text-align: center;">Tenoxicam</p>  |

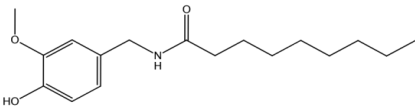
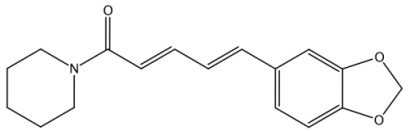
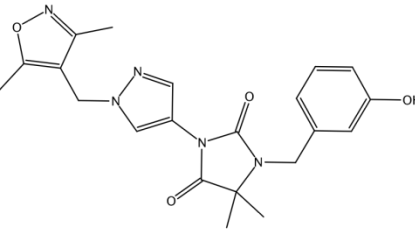
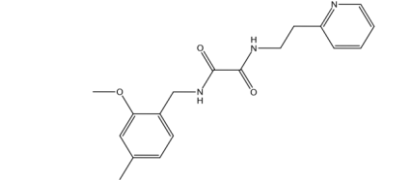
La búsqueda por similitud estructural se realizó entre los fármacos para dolor mostrados en la Tabla 7 y la base de datos FEMA GRAS, se utilizaron tres programas ChemAxon, PowerMV y MOE.

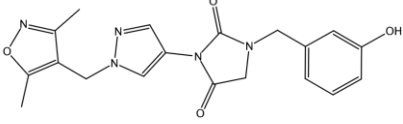
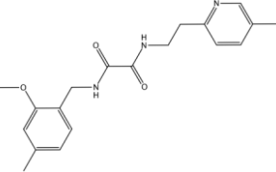
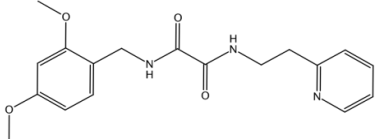
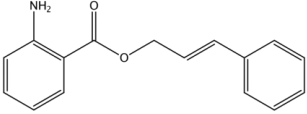
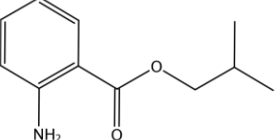
Con el programa ChemAxon, se calcularon cinco métodos o tipos de búsqueda que son: *substructure*, *superstructure*, *similarity* (Tanimoto 0.85 y Euclidian 0.15), *full* y *full fragment*. Sólo se obtuvieron datos para el método de *Superstructure*. En PowerMV, se calcularon cinco descriptores que son: Atom Pair, Atom Pair Carhat, Fragment Pair, Pharmacophore Fingerprints y Weighted Burden Number. De cada descriptor se obtuvieron 10 resultados. Los primeros cuatro descriptores fueron medidos con un coeficiente de Tanimoto de 0.85 y el quinto descriptor se midió en distancia Euclidiana. Los resultados de este programa están dados en distancias. En MOE, se obtuvieron datos con valores de similitud con un coeficiente de Tanimoto de 0.50 (ver Anexo 2). Con estos datos se hizo un consenso, es decir, aquellas moléculas de FEMA GRAS que resultaron similares al mismo fármaco de referencia por más de un método o por más de una medida de similitud. De esta forma se identificaron los compuestos de la bases de datos FEMA GRAS más similares a los diez fármacos analizados. Los resultados se muestran en la Tabla 8.

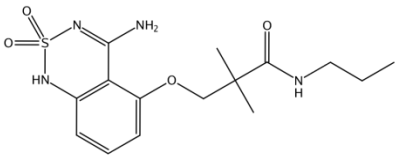
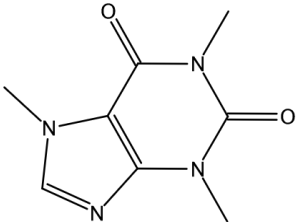
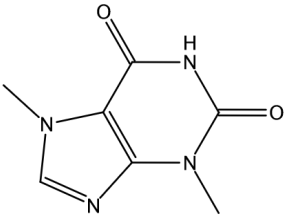
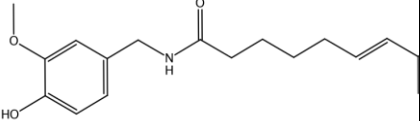
En la Tabla 8 se muestran las moléculas FEMA GRAS que resultaron estructuralmente más similares a los fármacos de referencia utilizados en este trabajo, cada molécula de aditivo saborizante GRAS es similar a uno o dos fármacos dando resultados en uno o dos programas/descriptores. Estas moléculas son las más atractivas para su posterior exploración experimental como posibles moduladoras de dolor.

Tabla 8. Moléculas consenso de aditivos saborizantes GRAS similares a las moléculas de los fármacos para el dolor.

| Número CAS | Nombre sistemático (IUPAC) | Estructura | Fármaco | Programa/Descriptor* |
|-------------|---|--|--|---|
| 75-07-0 | Etanal |  | Aceclofenaco Dextropropoxifeno Hidromorfona Piroxicam Rofecoxib Tenoxicam | ChemAxon ChemAxon ChemAxon ChemAxon ChemAxon ChemAxon |
| 103-82-2 | Ácido 2-fenilacético |  | Aceclofenaco Butorfanol Rofecoxib | ChemAxon PowerMV- burden number. ChemAxon Weighted |
| 110-89-4 | Piperidina |  | Butorfanol Carbamazina Hidromorfona | ChemAxon y MOE MOE ChemAxon |
| 106-68-3 | 3-Octanona |  | Butorfanol Carbamazina Hidromorfona | MOE MOE ChemAxon |
| 103-26-4 | Metil 3-fenilacrilato |  | Piroxicam Rofecoxib Tenoxicam | PowerMV- Fragment Pair ChemAxon PowerMV- Atom Pair, Fragment Pair PowerMV- Fragment Pair |
| 714229-20-6 | Advantame |  | Aceclofenaco Butorfanol Hidromorfona Piroxicam Tenoxicam | PowerMV- Atom Pair, Weighted burden number. MOE PowerMV- Atom Pair, Weighted burden number. MOE. PowerMV- Atom Pair PowerMV- Atom Pair |
| 745047-51-2 | N-(4-heptanilo)-1,3-benzodioxol-5-carboxamida |  | Aceclofenaco Citalopram Hidromorfona | PowerMV- burden number. MOE MOE Weighted |

| | | | | |
|----------------------------|---|--|--|--|
| <p>2444-46-4</p> | <p>Capsaicina sintética</p> |  | <p>Aceclofenaco Butorfanol Carbamazina Hidromorfona</p> | <p>PowerMV- Atom Pair Carhat MOE MOE PowerMV- Pharmacophore Fingerprints MOE</p> |
| <p>94-62-2</p> | <p>Piperina</p> |  | <p>Aceclofenaco Butorfanol Carbamazina Citalopram Hidromorfona Piroxicam Tenoxicam</p> | <p>PowerMV- Weighted burden number. MOE MOE MOE MOE PowerMV- Atom Pair PowerMV- Atom Pair</p> |
| <p>1217341-48-4</p> | <p>3-[1-[(3,5-dimetilisoxazol-4)-metil]pirazol-4]-1-[(3-hidroxifenil)metil]-5,5-dimetil-imidazolidona-2,4-diona</p> |  | <p>Citalopram Piroxicam Rofecoxib Tenoxicam</p> | <p>PowerMV- Weighted burden number. PowerMV- Atom Pair, Atom Pair Carhat, Pharmacophore Fingerprints, Weighted burden number. PowerMV- Weighted burden number. PowerMV- Atom Pair, Atom Pair Carhat, Pharmacophore Fingerprints, Weighted burden number. MOE</p> |
| <p>745047-97-6</p> | <p>N-(2-metoxi-4-metilbencil)-N'-[2-(2-piridinil)etil]etanodiamida</p> |  | <p>Butorfanol Carbamazina Hidromorfona Piroxicam Tenoxicam</p> | <p>MOE MOE PowerMV- Weighted burden number. PowerMV- Atom Pair PowerMV- Atom Pair</p> |

| | | | | |
|----------------------------|--|--|--|--|
| <p>1119831-25-2</p> | <p>3-[1-[(3,5-dimetilisoazol-4)metil]pirazol-4]-1-[(3-hidroksifenil)metil]imidazol-2,4-diona</p> |  | <p>Carbamazina Citalopram Piroxicam Rofecoxib Tenoxicam</p> | <p>MOE PowerMV- Weighted burden number. PowerMV- Atom Pair, Atom Pair Carhat, Pharmacophore Fingerprints, Weighted burden number. PowerMV- Weighted burden number. PowerMV- Atom Pair, Atom Pair Carhat, Pharmacophore Fingerprints, Weighted burden number. MOE</p> |
| <p>745047-94-3</p> | <p>N-(2-metoksi-4-metilbencil)-N'-[2-(5-metil-2-piridinil)etil]etanodiamida</p> |  | <p>Butorfanol Carbamazina Hidromorfona Piroxicam Tenoxicam</p> | <p>MOE MOE PowerMV- Weighted burden number. PowerMV- Atom Pair PowerMV- Atom Pair</p> |
| <p>745047-53-4</p> | <p>N-(2,4-dimetoksi-bencil)-N'-[2-(2-piridinil)etil]etanodiamida</p> |  | <p>Butorfanol Hidromorfona Piroxicam Tenoxicam</p> | <p>MOE PowerMV- Weighted burden number. PowerMV- Atom Pair PowerMV- Atom Pair</p> |
| <p>87-29-6</p> | <p>Cinamilo 2-aminobenzoato</p> |  | <p>Aceclofenaco Piroxicam Tenoxicam</p> | <p>MOE PowerMV- Atom Pair PowerMV- Atom Pair</p> |
| <p>7779-77-3</p> | <p>2-aminobenzoato de isobutilo</p> |  | <p>Aceclofenaco Piroxicam Tenoxicam</p> | <p>MOE PowerMV- Fragment Pair PowerMV- Fragment Pair</p> |

| | | | | |
|----------------------------|--|---|--|--|
| <p>1093200-92-0</p> | <p>N-[(4-amino-2,2-dióxido-1H-2,1,3-benzotiadiazia-5-il)oxi]-2,2-dimetil-N-propilpropanamida</p> |  | <p>Piroxicam Tenoxicam</p> | <p>PowerMV- Atom Pair, Fragment Pair. MOE PowerMV- Atom Pair, Fragment Pair. MOE</p> |
| <p>58-08-2</p> | <p>Cafeína</p> |  | <p>Carbamazina Piroxicam Rofecoxib Tenoxicam</p> | <p>MOE PowerMV-burden number. Weighted MOE PowerMV-burden number. Weighted PowerMV-burden number. Weighted MOE</p> |
| <p>83-67-0</p> | <p>Teobromina</p> |  | <p>Carbamazina Piroxicam Rofecoxib Tenoxicam</p> | <p>MOE PowerMV-burden number. Weighted MOE PowerMV-burden number. Weighted PowerMV-burden number. Weighted MOE</p> |
| <p>404-86-4</p> | <p>Capsaicina</p> |  | <p>Butorfanol Hidromorfona</p> | <p>MOE PowerMV- Pharmacophore Fingerprints</p> |

*Para ChemAxon el descriptor es *superstructure*.

De las 20 moléculas mostradas en la Tabla 8 se encontraron tres aditivos saborizantes con efecto analgésico reportado en la literatura. El primer aditivo saborizante es la piperina que se extrae de la pimienta negra. Esta ayuda a reducir el dolor activando el receptor transitor potencial vaniloide de tipo 1 y TPRV1, que son receptores que responden para contrarrestar el dolor.⁵⁶ El segundo compuesto FEMA GRAS con efecto analgésico reportado es la cafeína que tiene una eficacia analgésica en el tratamiento de la cefalea tensional y la migraña sin aura, además de aumentar y prolongar el efecto analgésico del paracetamol por la interacción farmacodinámica.⁵⁷ La cafeína tiene un efecto analgésico dosis-dependiente potenciado por los inhibidores de la serotonina y un efecto adyuvante en la analgesia.⁵⁸ La cafeína reduce los dolores de cabeza pues éste se debe a la tensión de los vasos sanguíneos del cerebro y la cafeína puede dilatarlos, reduciendo la intensidad del dolor, ayudando así a que los analgésicos proporcionen efecto más rápidamente.⁵⁹ El tercer y último aditivo saborizante encontrado en la literatura con efecto analgésico es la capsaicina que es una oleorresina, componente activo de los pimientos picantes (*Capsicum*). La capsaicina es un analgésico de uso tópico indicado para el alivio sintomático de dolores musculares o articulares localizados.⁶⁰ La capsaicina es eficaz en el tratamiento del dolor neuropático y del dolor asociado a condiciones como la osteoartritis, artritis reumatoide y soriasis.

Para identificar los pares de fármacos más similares dentro de los 10 identificados como más diversos se calculó la similitud estructural entre fármacos basada en los mismos programas y descriptores (ver Metodología y Anexo 2) que se usaron para el análisis de la base de datos FEMA GRAS con los fármacos de referencia.

En el programa de ChemAxon no se obtuvieron resultados de similitud entre fármacos para los cinco descriptores utilizados.

Los resultados obtenidos en los programas de PowerMV y MOE en la búsqueda por similitud estructural entre fármacos se muestran en las Tablas 9 y 10, respectivamente.

Tabla 9. Similitud estructural entre fármacos en PowerMV.

| Fármaco | Atom Pair | Atom Pair Carhat | Fragment Pair | Pharmacophore Fingerprints | Weighted Burden Number |
|---------------------|---------------------------|---------------------------|---------------------------|----------------------------|---------------------------|
| Aceclofenaco | Dextropropoxifeno 0.49 | Dextropropoxifeno 0.49 | Rofecoxib 0.92 | Rofecoxib 0.80 | Hidromorfona 0.54 |
| | Hidromorfona 0.53 | Hidromorfona 0.53 | Dextropropoxifeno 0.93 | Nimesulida 0.86 | Dextropropoxifeno 0.74 |
| | Rofecoxib 0.57 | Rofecoxib 0.57 | Nimesulida 0.94 | Dextropropoxifeno 0.89 | Butorfanol 1.21 |
| | Nimesulida 0.64 | Nimesulida 0.64 | Citalopram 0.95 | Citalopram 0.90 | Rofecoxib 1.55 |
| | | | | | |

| | | | | | |
|--------------------|---|---|---|---|---|
| | Piroxicam 0.67 Butorfanol 0.69 Tenoxicam 0.70 Citalopram 0.77 Carbamazina 0.86 | Piroxicam 0.67 Butorfanol 0.69 Tenoxicam 0.70 Citalopram 0.77 Carbamazina 0.86 | Piroxicam 0.98 Tenoxicam 0.98 Hidromorfona 0.98 Butorfanol 1.00 Carbamazina 1.00 | Piroxicam 0.92 Tenoxicam 0.92 Butorfanol 1.00 Carbamazina 1.00 Hidromorfona 1.00 | Nimesulida 1.75 Citalopram 1.76 Tenoxicam 1.85 Piroxicam 1.86 Carbamazina 2.29 |
| Butorfanol | Dextropropoxifeno 0.38 Hidromorfona 0.40 Citalopram 0.59 Rofecoxib 0.68 Aceclofenaco 0.69 Carbamazina 0.77 Nimesulida 0.79 Piroxicam 0.81 Tenoxicam 0.84 | Hidromorfona 0.60 Dextropropoxifeno 0.72 Citalopram 0.73 Aceclofenaco 0.83 Rofecoxib 0.86 Piroxicam 0.90 Tenoxicam 0.92 Carbamazina 0.93 Nimesulida 0.94 | Hidromorfona 0.77 Carbamazina 0.91 Citalopram 0.94 Rofecoxib 0.98 Aceclofenaco 1.00 Dextropropoxifeno 1.00 Nimesulida 1.00 Piroxicam 1.00 Tenoxicam 1.00 | Hidromorfona 0.69 Dextropropoxifeno 0.96 Aceclofenaco 1.00 Carbamazina 1.00 Citalopram 1.00 Nimesulida 1.00 Piroxicam 1.00 Rofecoxib 1.00 Tenoxicam 1.00 | Aceclofenaco 1.21 Dextropropoxifeno 1.23 Hidromorfona 1.37 Citalopram 1.86 Tenoxicam 1.87 Piroxicam 1.88 Carbamazina 1.96 Rofecoxib 1.99 Nimesulida 2.13 |
| Carbamazina | Dextropropoxifeno 0.74 Hidromorfona 0.75 Butorfanol 0.77 Aceclofenaco 0.86 Piroxicam 0.87 Citalopram 0.88 Tenoxicam 0.89 Rofecoxib 0.91 Nimesulida 0.91 | Dextropropoxifeno 0.86 Hidromorfona 0.89 Citalopram 0.91 Butorfanol 0.93 Rofecoxib 0.94 Piroxicam 0.95 Tenoxicam 0.95 Aceclofenaco 0.96 Nimesulida 0.98 | Dextropropoxifeno 0.81 Butorfanol 0.91 Hidromorfona 0.93 Citalopram 0.95 Rofecoxib 0.98 Piroxicam 0.98 Tenoxicam 0.98 Aceclofenaco 1.00 Nimesulida 1.00 | Citalopram 0.86 Dextropropoxifeno 0.94 Aceclofenaco 1.00 Butorfanol 1.00 Hidromorfona 1.00 Nimesulida 1.00 Piroxicam 1.00 Rofecoxib 1.00 Tenoxicam 1.00 | Butorfanol 1.96 Dextropropoxifeno 2.12 Aceclofenaco 2.29 Hidromorfona 2.37 Citalopram 2.88 Tenoxicam 2.89 Nimesulida 3.02 Rofecoxib 3.05 Piroxicam 3.07 |
| Citalopram | Butorfanol 0.59 Hidromorfona 0.62 Dextropropoxifeno 0.63 Rofecoxib 0.76 Aceclofenaco 0.77 Nimesulida 0.84 Piroxicam | Dextropropoxifeno 0.67 Hidromorfona 0.73 Butorfanol 0.73 Rofecoxib 0.74 Aceclofenaco 0.75 Nimesulida 0.85 Piroxicam | Hidromorfona 0.89 Dextropropoxifeno 0.90 Nimesulida 0.91 Butorfanol 0.94 Aceclofenaco 0.95 Carbamazina 0.95 Piroxicam | Nimesulida 0.75 Carbamazina 0.86 Dextropropoxifeno 0.88 Aceclofenaco 0.90 Butorfanol 1.00 Hidromorfona 1.00 Piroxicam | Rofecoxib 1.15 Piroxicam 1.21 Tenoxicam 1.25 Nimesulida 1.44 Aceclofenaco 1.76 Hidromorfona 1.77 Dextropropoxifeno |

| | | | | | |
|--------------------------|---|---|---|---|---|
| | 0.84 Tenoxicam 0.86 Carbamazina 0.88 | 0.91 Carbamazina 0.91 Tenoxicam 0.91 | 1.00 Rofecoxib 1.00 Tenoxicam 1.00 | 1.00 Rofecoxib 1.00 Tenoxicam 1.00 | 1.86 Butorfanol 1.86 Carbamazina 2.88 |
| Dextropropoxifeno | Hidromorfona 0.23 Butorfanol 0.38 Rofecoxib 0.47 Aceclofenaco 0.49 Citalopram 0.63 Nimesulida 0.68 Piroxicam 0.72 Carbamazina 0.74 Tenoxicam 0.75 | Hidromorfona 0.53 Citalopram 0.67 Rofecoxib 0.68 Butorfanol 0.72 Aceclofenaco 0.73 Nimesulida 0.80 Piroxicam 0.85 Tenoxicam 0.86 Carbamazina 0.86 | Carbamazina 0.81 Rofecoxib 0.88 Citalopram 0.90 Piroxicam 0.92 Tenoxicam 0.92 Nimesulida 0.92 Aceclofenaco 0.93 Hidromorfona 0.98 Butorfanol 1.00 | Rofecoxib 0.73 Citalopram 0.88 Piroxicam 0.89 Tenoxicam 0.89 Aceclofenaco 0.89 Hidromorfona 0.90 Carbamazina 0.94 Butorfanol 0.96 Nimesulida 1.00 | Hidromorfona 0.64 Aceclofenaco 0.74 Butorfanol 1.23 Rofecoxib 1.70 Nimesulida 1.78 Citalopram 1.89 Carbamazina 2.12 Tenoxicam 2.13 Piroxicam 2.14 |
| Hidromorfona | Dextropropoxifeno 0.23 Butorfanol 0.40 Aceclofenaco 0.53 Rofecoxib 0.57 Citalopram 0.62 Nimesulida 0.71 Piroxicam 0.74 Carbamazina 0.75 Tenoxicam 0.77 | Dextropropoxifeno 0.53 Butorfanol 0.60 Citalopram 0.73 Aceclofenaco 0.77 Rofecoxib 0.79 Piroxicam 0.87 Nimesulida 0.88 Tenoxicam 0.88 Carbamazina 0.89 | Butorfanol 0.77 Citalopram 0.89 Carbamazina 0.93 Nimesulida 0.98 Aceclofenaco 0.98 Rofecoxib 0.98 Dextropropoxifeno 0.98 Piroxicam 1.00 Tenoxicam 1.00 | Butorfanol 0.69 Dextropropoxifeno 0.90 Rofecoxib 0.92 Piroxicam 0.93 Tenoxicam 0.93 Aceclofenaco 1.00 Carbamazina 1.00 Citalopram 1.00 Nimesulida 1.00 | Aceclofenaco 0.54 Dextropropoxifeno 0.64 Butorfanol 1.37 Rofecoxib 1.39 Nimesulida 1.55 Citalopram 1.77 Tenoxicam 1.86 Piroxicam 1.87 Carbamazina 2.37 |
| Nimesulida | Rofecoxib 0.46 Piroxicam 0.49 Tenoxicam 0.58 Aceclofenaco 0.64 Dextropropoxifeno 0.68 Hidromorfona 0.71 Butorfanol 0.79 Citalopram 0.84 Carbamazina 0.91 | Tenoxicam 0.59 Rofecoxib 0.64 Tenoxicam 0.65 Aceclofenaco 0.76 Dextropropoxifeno 0.80 Citalopram 0.85 Hidromorfona 0.88 Butorfanol 0.94 Carbamazina 0.98 | Piroxicam 0.90 Tenoxicam 0.90 Citalopram 0.91 Dextropropoxifeno 0.92 Aceclofenaco 0.94 Hidromorfona 0.98 Butorfanol 1.00 Carbamazina 1.00 Rofecoxib 1.00 | Citalopram 0.75 Aceclofenaco 0.86 Butorfanol 1.00 Carbamazina 1.00 Dextropropoxifeno 1.00 Hidromorfona 1.00 Piroxicam 1.00 Rofecoxib 1.00 Tenoxicam 1.00 | Rofecoxib 0.72 Piroxicam 1.23 Tenoxicam 1.25 Citalopram 1.44 Hidromorfona 1.55 Aceclofenaco 1.75 Dextropropoxifeno 1.78 Butorfanol 2.13 Carbamazina 3.02 |

| | | | | | |
|------------------|--|--|--|--|--|
| Piroxicam | Tenoxicam 0.16 Nimesulida 0.49 Rofecoxib 0.57 Aceclofenaco 0.67 Dextropropoxifeno 0.72 Hidromorfona 0.74 Butorfanol 0.81 Citalopram 0.84 Carbamazina 0.87 | Tenoxicam 0.18 Nimesulida 0.59 Rofecoxib 0.70 Aceclofenaco 0.77 Dextropropoxifeno 0.85 Hidromorfona 0.87 Butorfanol 0.90 Citalopram 0.91 Carbamazina 0.95 | Tenoxicam 0.00 Nimesulida 0.90 Dextropropoxifeno 0.92 Rofecoxib 0.94 Aceclofenaco 0.98 Carbamazina 0.98 Butorfanol 1.00 Citalopram 1.00 Hidromorfona 1.00 | Tenoxicam 0.00 Rofecoxib 0.78 Dextropropoxifeno 0.89 Aceclofenaco 0.92 Hidromorfona 0.93 Butorfanol 1.00 Carbamazina 1.00 Citalopram 1.00 Nimesulida 1.00 | Tenoxicam 0.29 Rofecoxib 0.97 Citalopram 1.21 Nimesulida 1.23 Aceclofenaco 1.86 Hidromorfona 1.87 Butorfanol 1.88 Dextropropoxifeno 2.14 Carbamazina 3.07 |
| Rofecoxib | Nimesulida 0.46 Dextropropoxifeno 0.47 Piroxicam 0.57 Aceclofenaco 0.57 Hidromorfona 0.57 Tenoxicam 0.64 Butorfanol 0.68 Citalopram 0.76 Carbamazina 0.91 | Aceclofenaco 0.59 Nimesulida 0.64 Dextropropoxifeno 0.68 Piroxicam 0.70 Tenoxicam 0.73 Citalopram 0.74 Hidromorfona 0.79 Butorfanol 0.86 Carbamazina 0.94 | Dextropropoxifeno 0.88 Aceclofenaco 0.92 Piroxicam 0.94 Tenoxicam 0.94 Carbamazina 0.98 Hidromorfona 0.98 Butorfanol 0.98 Citalopram 1.00 Nimesulida 1.00 | Dextropropoxifeno 0.73 Piroxicam 0.78 Tenoxicam 0.78 Aceclofenaco 0.80 Hidromorfona 0.92 Butorfanol 1.00 Carbamazina 1.00 Citalopram 1.00 Nimesulida 1.00 | Nimesulida 0.72 Piroxicam 0.97 Tenoxicam 1.03 Citalopram 1.15 Hidromorfona 1.39 Aceclofenaco 1.55 Dextropropoxifeno 1.70 Butorfanol 1.99 Carbamazina 3.05 |
| Tenoxicam | Piroxicam 0.16 Nimesulida 0.58 Rofecoxib 0.64 Aceclofenaco 0.70 Dextropropoxifeno 0.75 Hidromorfona 0.77 Butorfanol 0.84 Citalopram 0.86 Carbamazina 0.89 | Piroxicam 0.18 Nimesulida 0.65 Rofecoxib 0.73 Aceclofenaco 0.78 Dextropropoxifeno 0.86 Hidromorfona 0.88 Citalopram 0.91 Butorfanol 0.92 Carbamazina 0.95 | Piroxicam 0.00 Nimesulida 0.90 Dextropropoxifeno 0.92 Rofecoxib 0.94 Aceclofenaco 0.98 Carbamazina 0.98 Butorfanol 1.00 Citalopram 1.00 Hidromorfona 1.00 | Piroxicam 0.00 Rofecoxib 0.78 Dextropropoxifeno 0.89 Aceclofenaco 0.92 Hidromorfona 0.93 Butorfanol 1.00 Carbamazina 1.00 Citalopram 1.00 Nimesulida 1.00 | Piroxicam 0.29 Rofecoxib 1.03 Nimesulida 1.25 Citalopram 1.25 Aceclofenaco 1.85 Hidromorfona 1.86 Butorfanol 1.87 Dextropropoxifeno 2.13 Carbamazina 2.89 |

Tabla 10. Similitud estructural entre fármacos en MOE

| Fármaco | Similitud |
|--------------------------|--|
| Aceclofenaco | Citalopram (Tanimoto 0.30) Dextropropoxifeno (Tanimoto 0.30) Hidromorfona (Tanimoto 0.30) |
| Butorfanol | Hidromorfona (Tanimoto 0.70) |
| Carbamazina | Hidromorfona (Tanimoto 0.46) |
| Citalopram | Dextropropoxifeno (Tanimoto 0.50) Hidromorfona (Tanimoto 0.50) |
| Dextropropoxifeno | Citalopram (Tanimoto 0.50) |
| Hidromorfona | Butorfanol (Tanimoto 0.70) |
| Nimesulida | Nimesulida (Tanimoto 0.60) |
| Piroxicam | Tenoxicam (Tanimoto 0.85) |
| Rofecoxib | Tenoxicam (Tanimoto 0.47) |
| Tenoxicam | Piroxicam (Tanimoto 0.85) |

Los pares de fármacos más similares son entre butorfanol e hidromorfona (Tabla 9) como se observa en la tabla 7 estos dos fármacos tienen en común un grupo fenol. Otro par de fármacos con similitud estructural fue piroxicam y tenoxicam (Tabla 9) esto se debe a su estructura ya que sólo cambia un anillo aromático (piroxicam) por un heterociclo insaturado (tenoxicam) por lo tanto, su similitud fue evidente. En MOE (Tabla 10) los pares de fármacos más similares fueron los mismos que en PowerMV (Tabla 9), la similitud entre butorfanol e hidromorfona es de 0.70, para piroxicam y tenoxicam la similitud fue de 0.85. En la literatura está reportado que butorfanol e hidromorfona son analgésicos opioides agonistas que actúan en el sistema nervioso central, el butorfanol es un agonista en los receptores opioides kappa y una mezcla de agonista-antagonista en los receptores opioides mu, actúa en la materia gris periventricular y periacueductal y en algunos sitios de la médula espinal. La hidromorfona es un analgésico narcótico que inhibe selectivamente

la liberación de neurotransmisores de las terminaciones nerviosas aferentes que conducen estímulos dolorosos, ejerce actividad agonista principalmente sobre los receptores opioides μ y en menor medida sobre los receptores opioides κ .

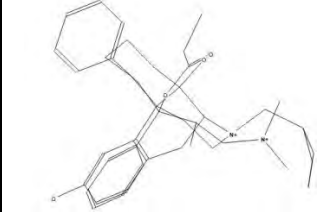
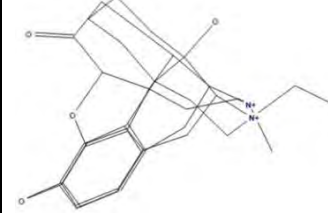
Piroxicam y tenoxicam pertenecen a la familia de los oxicam, son antiinflamatorios no esteroideos (AINES) que inhibe a la enzima ciclooxigenasa (COX), la cual regula la producción de las prostaglandinas.

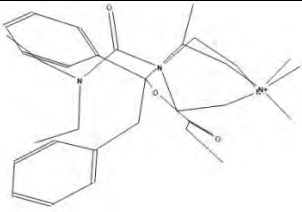
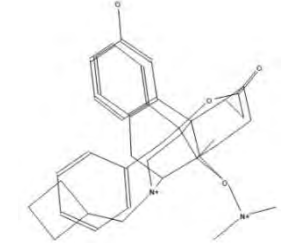
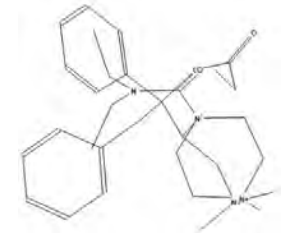
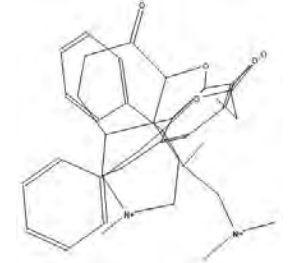
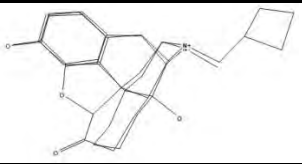
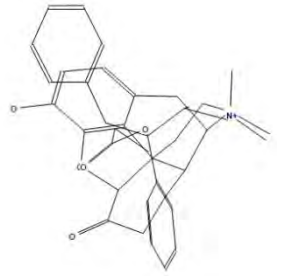
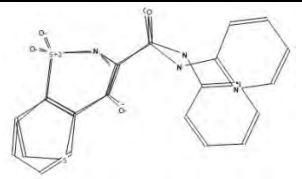
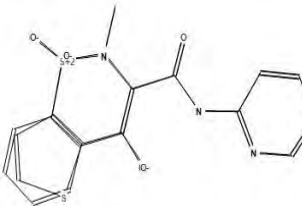
A pesar de que la mayoría de los fármacos de referencia obtenidos en este análisis actúan inhibiendo a COX-1 (aceclofenaco, piroxicam y tenoxicam), COX-2 (rofecoxib y nimesulida) y tres son opioides agonistas (butorfanol, dextropropoxifeno e hidromorfona), los análisis de similitud por estructura química muestran que el piroxicam y el tenoxicam son parecidos y esto se debe a que pertenecen a la misma familia que es oxicam.

Ésta comparación estructural se expandió utilizando la representación 3D de las moléculas. Es decir, se realizó un alineamiento molecular pareado entre los fármacos para buscar las conformaciones que se superponen en cada uno de los fármacos, este análisis se realizó utilizando el programa MOE.

Éste análisis indica que conformaciones y/o alineamientos son más similares, en la Tabla 11 se muestran las mejores alineaciones o superposiciones con valores bajos que indican que esa alineación o superposición es la indicada para deducir que tienen la misma actividad biológica y/o que su sitio de acción es el mismo.

Tabla 11. Mejores alineaciones o superposiciones moleculares de los fármacos por similitud estructural.

| Fármaco fijo | Fármaco superpuesto | Alineación molecular | F (similitud) | S (alineación) |
|-------------------|--------------------------|--|---------------|----------------|
| Butorfanol | Dextropropoxifeno |  | -121.31812 | -84.473526 |
| Butorfanol | Hidromorfona |  | -132.71721 | -85.83477 |

| | | | | |
|--------------------------|--------------------------|--|------------|------------|
| Carbamazina | Dextropropoxifeno |  | -83.598976 | -46.108971 |
| Dextropropoxifeno | Butorfanol |  | -116.02858 | -71.870346 |
| Dextropropoxifeno | Carbamazina |  | -81.341248 | -64.334297 |
| Dextropropoxifeno | Hidromorfona |  | -98.580147 | -58.717751 |
| Hidromorfona | Butorfanol |  | -137.7738 | -94.424492 |
| Hidromorfona | Dextropropoxifeno |  | -109.00567 | -71.497543 |
| Piroxicam | Tenoxicam |  | -156.03479 | -129.94365 |
| Tenoxicam | Piroxicam |  | -183.99379 | -159.55797 |

En la Tabla 11 se muestra la mejor alineación para cada par de los fármacos que resultaron ser muy similares estructuralmente, F es un valor negativo de P -overlap que es la función de densidad de la superposición, es decir, es una medida de similitud de la configuración estructural. Los valores más bajos indican una mayor similitud estructural, S es un valor de puntuación de la alineación, los valores más bajos indican mejores alineaciones. Para conocer los resultados del análisis de alineamiento molecular se basó principalmente en el valor de S .

5.4. Espacio estructural y químico de FEMA GRAS y fármacos para el dolor

Para comparar las propiedades fisicoquímicas de las moléculas contenidas en FEMA GRAS con cada uno de los fármacos primero se integraron estas dos bases de datos utilizando MOE y se calcularon las propiedades fisicoquímicas (Anexo 1), finalmente se realizó el análisis de componentes principales. Las representaciones visuales del espacio químico se muestran en las Figuras 7 y 8, realizadas en los programas de MOE y ChemAxon, respectivamente.

La varianza explicada de los dos primeros componentes principales para el espacio químico obtenido con MOE es de 60.79%, para el CP1 la propiedad fisicoquímica que contribuyó más a explicar la varianza es el número de átomos aceptores de hidrógeno seguido del número de átomos donadores de enlaces de hidrógeno, mientras que para el PC2 las propiedades fisicoquímicas con mayor contribución a la varianza son el número de anillos y el coeficiente de partición octanol/agua ($\log P$), (Figura 7).

El porcentaje de varianza de los dos primeros componentes principales es de 75.34%. Para el primer CP las propiedades fisicoquímicas con mayores contribuciones a la varianza son el número de átomos donadores de enlaces de hidrógeno y el número de aceptores de enlaces de hidrógeno mientras que para el segundo CP las propiedades con mayor contribución es el coeficiente de partición octanol/agua ($\log P$) seguido del número de enlaces rotables (Figura 8).

Puede observarse que la distribución de datos en las Figuras 7 y 8 son diferentes, particularmente en el CP2. Esto se debe a las pequeñas diferencias en que cada programa calcula los descriptores y a que se evaluaron diferentes propiedades fisicoquímicas entre los programas (ver Anexo 1).

Como se observa en las Figuras 7 y 8 existe una gran variedad de moléculas de aditivos saborizantes GRAS que están muy cerca de un fármaco. A partir del análisis de

componentes principales es posible calcular las distancias que hay entre cada fármaco y cada molécula FEMA GRAS. Dado que las propiedades fisicoquímicas pueden tomar valores en diferentes rangos, como primer paso se normalizaron los datos a valores entre 0 y 1. La normalización se realizó con el programa OriginPro. Posteriormente se calculó la distancia euclidiana considerando todas las dimensiones, no únicamente las visualizadas en las Figuras 7 y 8.

Una vez que se normalizaron las distancias, se seleccionaron las diez moléculas de aditivos saborizantes GRAS con menor distancia entre cada fármaco, puesto que entre menor sea la distancia mayor será la similitud entre cada par de moléculas. Como se hizo anteriormente, se evaluó el consenso de estas moléculas con las moléculas de la Tabla 8 y se obtuvieron las moléculas que se muestran en la Tabla 12.

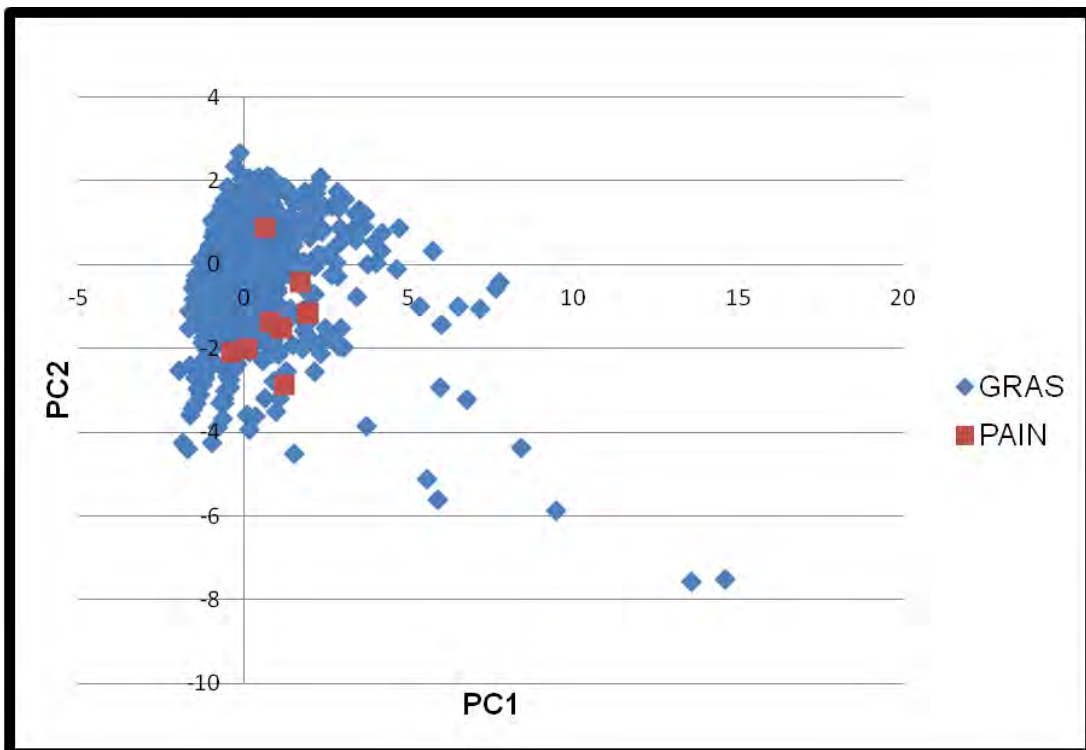


Figura 7. Espacio químico de la base de datos FEMA GRAS (GRAS 25) y Fármacos para dolor representativos. Las propiedades fisicoquímicas calculadas fueron logP, SlogP, TPSA, a_don, a_acc, b_rotN, MW, RC, (a_nBr), (a_nCl), (a_nF) y (a_nI) en MOE. La varianza explicada por los dos primeros componentes principales es de 60.79%.

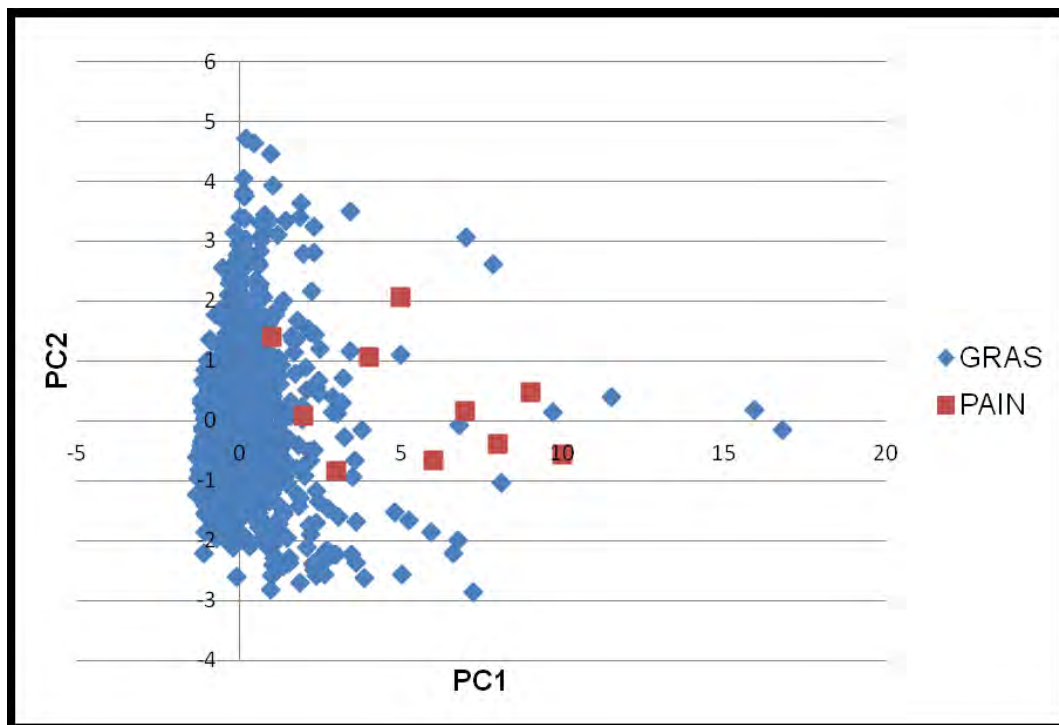
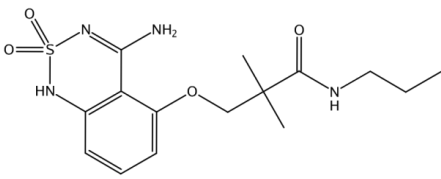
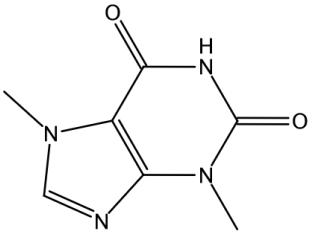


Figura 8. Espacio químico de la base de datos FEMA GRAS (GRAS 25) y Fármacos para dolor. Las propiedades fisicoquímicas calculadas fueron logP, PSA, a_don, a_acc, b_rotN, MW y RC en ChemAxon. La varianza explicada por los dos primeros componentes principales es de 75.34%

Tabla 12. Moléculas consenso FEMA GRAS, PCA y similitud estructural.

| CAS | Nombre sistemático (IUPAC) | Estructura | Fármaco Programa/Descriptor | Distancia |
|--------------|---|--|--|-----------|
| 1093200-92-0 | N-[(4-amino-2,2-dióxido-1H-2,1,3-benzotiadiazia-5-il)oxi]-2,2-dimetil-N-propilpropanamida |  | Piroxicam: PowerMV- Atom Pair 0.52 PowerMV- Fragment Pair 0.89 MOE 0.50 PCA MOE 0.08 PCA ChemAxon 0.19 Tenoxicam: PowerMV- Atom Pair 0.58 PowerMV- Fragment Pair 0.89 MOE 0.50 PCA MOE 0.09 PCA ChemAxon 0.20 | |
| 83-67-0 | Teobromina |  | Piroxicam: PowerMV- Weighted burden number. 0.77 MOE 0.50 PCA MOE 0.09 PCA ChemAxon 0.21 Tenoxicam: PowerMV- Weighted burden number. 0.86 MOE 0.50 PCA MOE 0.09 PCA ChemAxon 0.23 | |

En la Tabla 12 se observan las dos moléculas FEMA GRAS con mayor similitud a dos fármacos que alivian el dolor tanto en estructura como en propiedades fisicoquímicas. En la última columna de la Tabla 12, se indica la distancia euclidiana calculada mediante el análisis de componentes principales que hay entre el fármaco y la molécula FEMA GRAS, como ya se había mencionado, entre menor sea la distancia mayor será la similitud. Se observa que la distancia euclidiana calculada mediante el análisis de componentes principales en el programa ChemAxon es menor a 0.10 mientras que para MOE la distancia euclidiana es mayor a 0.10. Como ya se discutió anteriormente, esto es por el número de propiedades fisicoquímicas evaluadas en cada programa así como la manera de evaluarlas varía.

5.5. Comparación de secuencia de péptidos de alimentos

La última etapa de este proyecto consistió en analizar las secuencias de péptidos de alimentos con una longitud entre 4 y 10 aminoácidos y cuya afinidad a los receptores mu o kappa ha sido reportado. Dada la naturaleza química de los péptidos, este análisis se realizó de manera independiente a las secciones anteriores. Se realizó un análisis de las secuencias de péptidos usando el programa Statistica de StatSoft Inc.; se utilizó la información de la Tabla 4 la cual contiene 34 péptidos de alimentos. Para el análisis se excluyeron seis péptidos cuya afinidad a los receptores opioides mu o kappa no son confiables así como los péptidos con selectividad a receptores opioides kappa o delta los cuales eran cinco. Por lo tanto, se analizaron 23 péptidos con selectividad a receptores opioides mu y delta, de los cuales 16 tienen actividad agonista y 7 actividad antagonista. En las Tablas 13 a 16 se muestran los porcentajes de los aminoácidos según su posición en la secuencia. Para una visualización más clara de las secuencias de los péptidos con actividad opioide agonista y/o antagonista se realizaron diagramas circulares variando el eje de las ordenadas los cuales se muestran en las Figuras 9 a 11.

Tabla 13. Frecuencia de los aminoácidos en la secuencia de péptidos que se encuentran en la posición 1. Y (Tirosina), G (Glicina), S (Serina) y R (Arginina).

| Péptido | Frecuencia | Frecuencia acumulativa | Porcentaje | Porcentaje acumulativo |
|----------------|-------------------|-------------------------------|-------------------|-------------------------------|
| Y | 19 | 19 | 83 | 83 |
| G | 1 | 20 | 4 | 87 |
| S | 1 | 21 | 4 | 91 |
| R | 2 | 23 | 9 | 100 |

Tabla 14. Frecuencia de los aminoácidos en la secuencia de péptidos que se encuentran en la posición 2. Y, L (Leucina), G, V (Valina), P (Prolina), I (Isoleucina) y R.

| Péptido | Frecuencia | Frecuencia acumulativa | Porcentaje | Porcentaje acumulativo |
|---------|------------|------------------------|------------|------------------------|
| Y | 3 | 3 | 13 | 13 |
| L | 3 | 6 | 13 | 26 |
| G | 5 | 11 | 22 | 48 |
| V | 1 | 12 | 4 | 52 |
| P | 9 | 21 | 39 | 91 |
| I | 1 | 22 | 4 | 95 |
| R | 1 | 23 | 5 | 100 |

Tabla 15. Frecuencia de los aminoácidos en la secuencia de péptidos que se encuentran en la posición 3. Y, L, G, V, P, F (Fenilalanina), S y W (Triptófano).

| Péptido | Frecuencia | Frecuencia acumulativa | Porcentaje | Porcentaje acumulativo |
|---------|------------|------------------------|------------|------------------------|
| Y | 2 | 2 | 9 | 9 |
| L | 5 | 7 | 22 | 31 |
| G | 2 | 9 | 9 | 40 |
| V | 1 | 10 | 4 | 44 |
| P | 4 | 14 | 17 | 61 |
| F | 6 | 20 | 26 | 87 |
| S | 2 | 22 | 9 | 96 |
| W | 1 | 23 | 4 | 100 |

Tabla 16. Frecuencia de los aminoácidos en la secuencia de péptidos que se encuentran en la posición 4. Y, L, G, E (Ácido Glutámico), V, P, F, I, S, M (Metionina) y T (Treonina).

| Péptido | Frecuencia | Frecuencia acumulativa | Porcentaje | Porcentaje acumulativo |
|---------|------------|------------------------|------------|------------------------|
| Y | 2 | 2 | 9 | 9 |
| L | 1 | 3 | 4 | 13 |
| G | 1 | 4 | 4 | 17 |
| E | 1 | 5 | 4 | 21 |
| V | 2 | 7 | 9 | 30 |
| P | 4 | 11 | 17 | 47 |
| F | 7 | 18 | 30 | 77 |
| I | 1 | 19 | 4 | 81 |
| S | 2 | 21 | 9 | 90 |
| M | 1 | 22 | 5 | 95 |
| T | 1 | 23 | 5 | 100 |

En los estudios de la secuencia de péptidos se encontró que las moléculas tienen un grupo fenol en la parte inicial o terminal de la estructura, el fenol tiene buena actividad como grupo de unión en el sitio activo de los opioides, particularmente tiene afinidad por receptores opioides mu también necesitan de un grupo aromático como parte hidrofóbica y de un fragmento que confiera rigidez.

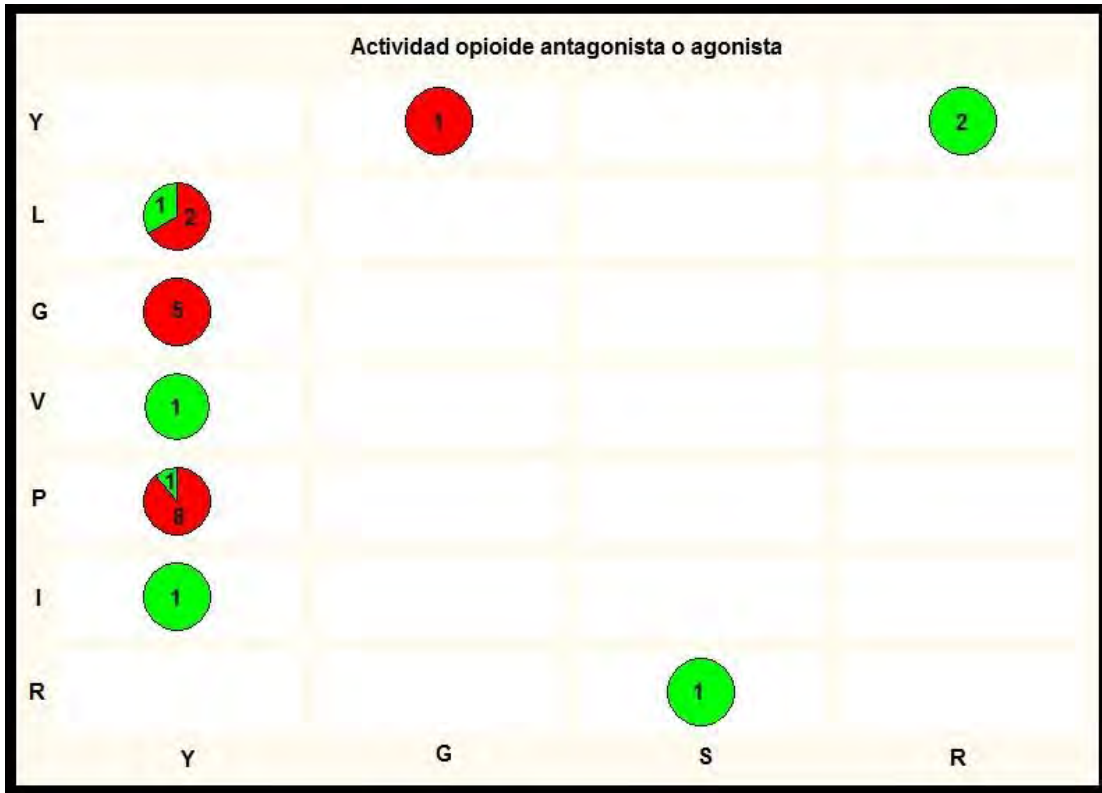


Figura 9. Gráfico circular. Actividad opioide agonista o antagonista. El círculo rojo indica que es un opioide con actividad agonista y el círculo verde indica que es un opioide con actividad antagonista; los números dentro de los círculos son el número de casos que tienen en la posición uno, alguno de los siguientes aminoácidos Y (Tirosina), G (Glicina), S (Serina) y R (Arginina), y en la posición dos, alguno de los siguientes aminoácidos R (Arginina), I (Isoleucina), P (Prolina), V (Valina), G (Glicina), L (Leucina) y Y (Arginina).

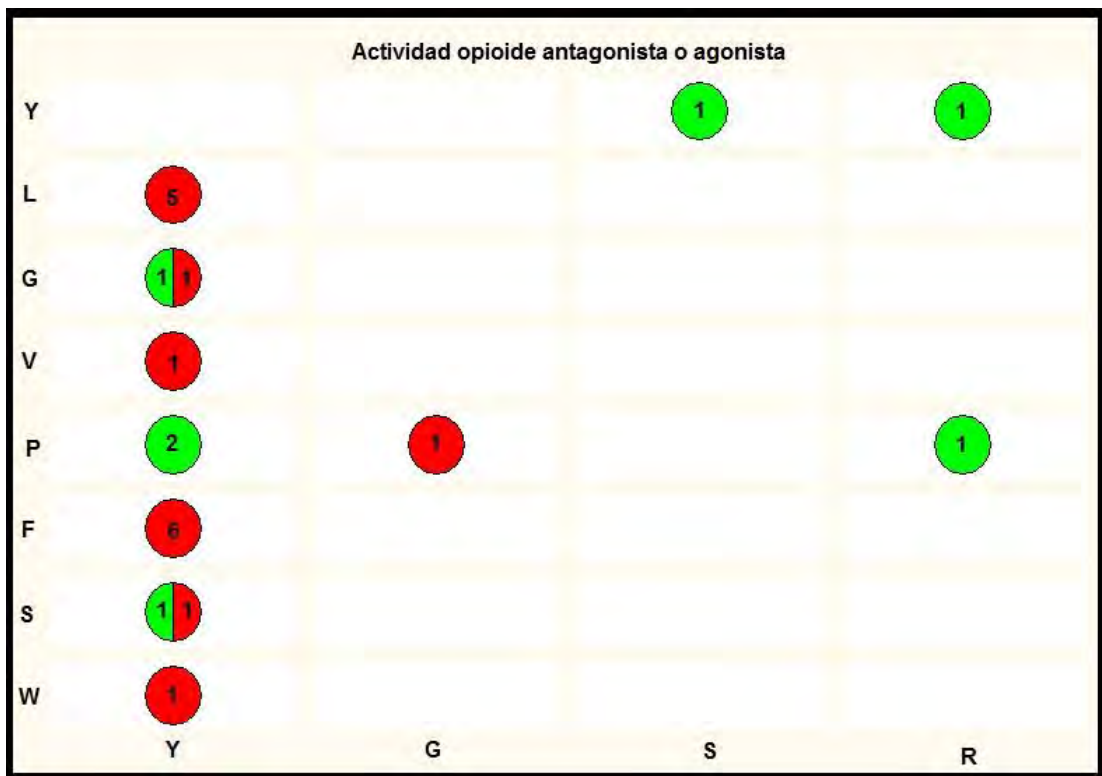


Figura 10. Gráfico circular. Actividad opioide agonista o antagonista. El círculo rojo indica que es un opioide con actividad agonista y el círculo verde indica que es un opioide con actividad antagonista; los números dentro de los círculos son el número de casos que tienen en la posición uno, alguno de los siguientes aminoácidos Y (Tirosina), G (Glicina), S (Serina) y R (Arginina), y en la posición tres, alguno de los siguientes aminoácidos W (Triptófano), S (Serina), F (Fenilalanina), P (Prolina), V (Valina), G (Glicina), L (Leucina) y Y (Arginina).

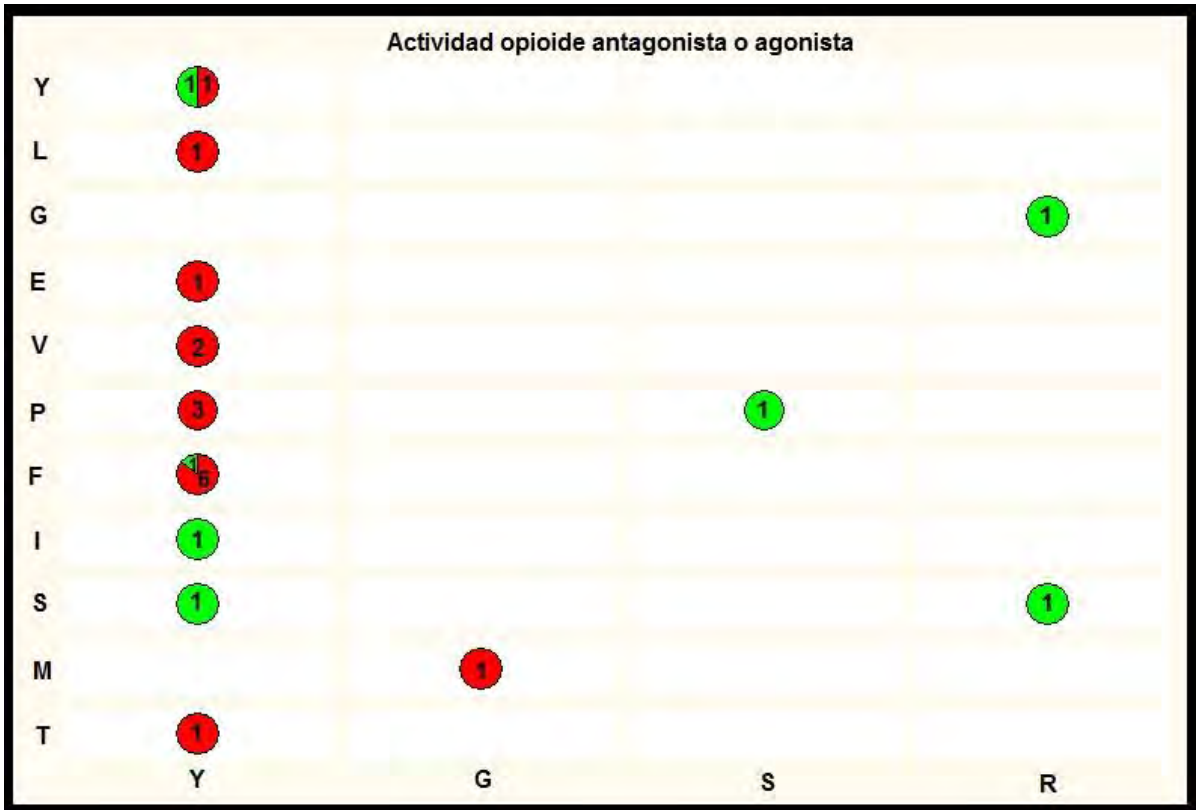


Figura 11. Gráfico circular. Actividad opioide agonista o antagonista. El círculo rojo indica que es un opioide con actividad agonista y el círculo verde indica que es un opioide con actividad antagonista; los números dentro de los círculos son el número de casos que tienen en la posición uno, alguno de los siguientes aminoácidos Y (Tirosina), G (Glicina), S (Serina) y R (Arginina), y en la posición tres, alguno de los siguientes aminoácidos T (Treonina), M (Metiona), S (Serina), I (Isoleucina), F (Fenilalanina), P (Prolina), V (Valina), E (Ácido Glutámico), G (Glicina), L (Leucina) y Y (Arginina).

En este análisis se encontró que la secuencia de péptidos de alimentos (ver Tabla 4) contienen al menos un aminoácido de tirosina, fenilalanina y prolina (ver Tablas 11 a 14), esto es debido a que la tirosina contiene un grupo fenol el cual se une al receptor opioide mu, la fenilalanina tiene un grupo aromático, es la parte hidrofóbica; y la prolina le confiere rigidez.

En las Figuras 9 a 11 se observa la actividad opioide de cada uno de los péptidos analizados. Los gráficos circulares muestran las posibles combinaciones, dejando en el eje de las abscisas los aminoácidos que se encuentran en la posición uno de la secuencia peptídica; únicamente cambia el eje de las ordenadas.

También se observa que el aminoácido Y en la posición 1 es el que mayoritariamente se encuentra al inicio de la secuencia peptídica y los aminoácidos G, S y R se encuentran en la posición 1 una sola vez (ver Figuras 9-11).

Con la información proporcionada en la Tabla 4 y con los gráficos circulares se sabe el número de casos que hay en cada combinación y la actividad opioide agonista o antagonista que tiene cada derivado.

6. Conclusiones

Se presenta una comparación de las entidades químicas dentro de la lista FEMA GRAS (GRAS 25) con productos naturales, fármacos aprobados, medicina tradicional china y una biblioteca de compuestos diversos utilizando criterios diferentes y complementarios. Las moléculas de aditivos saborizantes FEMA GRAS tienen diferentes propiedades moleculares y un perfil muy similar a los fármacos aprobados.

Se encontraron 20 moléculas de la base de datos FEMA GRAS (GRAS 25) que son estructuralmente similares a los fármacos que alivian el dolor. Sin embargo, sólo dos moléculas de la base de datos FEMA GRAS (GRAS 25), Teobromina (CAS 1093200-92-0) y la N-[(4-amino-2,2-dióxido-1H-2,1,3-benzotiadiazia-5-il)oxi)]-2,2-dimetil-N-propilpropanamida (CAS 83-67-0), son similares a los fármacos piroxicam y tenoxicam, tanto estructuralmente como en propiedades fisicoquímicas. Esto indica que dichas moléculas son atractivas para explorar su potencial efecto analgésico.

El análisis de péptidos de alimentos con actividad opioide muestra que los residuos asociados con actividad opioide son: tirosina, fenilalina y prolina. Estas características estructurales coinciden con el farmacóforo conocido para moléculas afines a receptores opioides.

El análisis quemoinformático de las bases de datos muestra que las moléculas contenidas en la lista FEMA GRAS tienen propiedades atractivas y comparables a fármacos o fármacos en desarrollo. El estudio de similitud estructural y de propiedades fisicoquímicas proporcionó aquellas moléculas con mayores posibilidades de tener actividad analgésica. Finalmente, el análisis de los péptidos de alimentos proporcionó las características estructurales asociadas a actividad opioide, la cuál es la vía de analgesia más potente conocida. En conjunto, este trabajo provee información relevante de la posible acción analgésica de aditivos saborizantes y las características estructurales asociadas a ella.

Los métodos computacionales aquí descritos permiten encontrar de forma sistemática y rápida compuestos empleados en la industria alimentaria con potencial actividad biológica.

7. Bibliografía

1. Johnson, M. A. y. M. G. M. Concepts and Applications of Molecular Similarity. *Journal of Molecular Structure* **1990**, *269*, 376-377.
2. Maza, J. R. Aproximación teórica de medidas de similitud molecular sobre laplacianos de la densidad electrónica. Universidad Nacional de Colombia, 2010.
3. Medina-Franco, J. L. Activity cliffs: Facts or Artifacts? *Chemical Biology & Drug Design* **2013**, *81*, 553-556.
4. Martínez-Mayorga, K., Peppard, T.L., Yongye, A.B., Santos, R., Giulanotti, M., Medina-Franco, J.L. Characterization of a comprehensive flavor database. *Journal of Chemometrics* **2011**, *25*, 550-560.
5. Sprous, D. G. a. S., 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.
6. Medina-Franco, J. L., Martínez-Mayorga, K., Peppard, T.L., Del Río, A. Chemoinformatic Analysis of GRAS (Generally Recognized as Safe) Flavor Chemicals and Natural Products. *Plos one* **2012**, *7*, 1-11.
7. Thomas, B. G., Weir, M.D., Hass, J., Giordano, F.R., *Cálculo: una variable*. Undecima ed.; Pearson Educación: Estado de Mexico, 2005.
8. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **1988**, *28*, 31-36.
9. Weininger, D., Weininger, A., Weininger, J.L. SMILES. 2. Algorithm for generation of unique SMILES notation. *Journal of Chemical Information and Computer Sciences* **1989**, *29*, 97-101.
10. Daylight Chemical Information Systems, I. SMARTS- A language for describing molecular patterns. <http://www.daylight.com> (6 de Agosto),
11. Group, C. C. MOE 2010.10. http://www.chemcomp.com/MOE-Molecular_Operating_Environment.htm (12 de Agosto),
12. Spegazzini, N. Diseño de un modelo de interacción de benzodiazepinas y flavonoides con acción ansiolítica derivado de su análisis QSAR. Tesina, Universidad de Belgrano, 2007.
13. Platt, J. R. Influence of Neighbor Bonds on Additive Bond Properties in Paraffins. *Journal of Chemical Physics* **1947**, *15*, 419-420.
14. Wiener, H. Structural Determination of Paraffin Boiling Points. *Journal of the American Chemical Society* **1947**, *69*, 17-20.
15. Martínez-Mayorga, K., Medina-Franco, J.L. Chemoinformatics-Applications in Food Chemistry. *Advances in Food and Nutrition Research* **2009**, *58*, 33-55.
16. Rodríguez Yunta, L., *Bases de datos documentales: estructura y uso*. CINDOC: Madrid, 2001.
17. ChEMBL EMBL-EBI. <https://www.ebi.ac.uk/chembl> (11 de Septiembre),
18. (NCI), N. C. I. Division of Cancer Treatment and Diagnosis. http://dctd.cancer.gov/About/major_initiatives_NExt.htm (26 de Octubre),
19. PubChem PubChem. <http://pubchem.ncbi.nlm.nih.gov/#> (22 de Octubre),
20. UNIIQUIM Unidad de Informática del Instituto de Química (UNIIQUIM). <http://uniiquim.iquimica.unam.mx> (18 de Septiembre),
21. Bolton, E. E., Wang, Y., Thiessen P.A., Bryant, S.H., PubChem: Integrated Platform of Small Molecules and Biological Activities. In *Annual Reports in Computational Chemistry*, Society, A. C., Ed. Washington, DC, 2008; Vol. 4.
22. Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., Overington, J.P. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research* **2012**, *40*, 1100-1107.
23. Wang, Y., Xiao, J., Suzek, T.O., Zhang, J., Wang, J., Zhou, Z., Han, L., Karapetyan, K., Dracheva, S., Shoemaker, B.A., Bolton. E., Gindulyte, A., Bryant, S.H. PubChem's BioAssay Database. *Nucleic Acids Research* **2012**, *40*, 400-412.
24. Sedici Química computacional-Estudio vibracional teórico. <http://sedici.unlp.edu.ar> (7 de Agosto),

25. UNAM Conceptos básicos de la química computacional. <http://www.fis.unam.mx> (6 de Agosto),
26. Peniche Trujillo, A. G. La quimioinformática, una herramienta eficiente para desarrollar los medicamentos del futuro. *Teoría y Praxis Investigativa* **2011**, 6, 77-86.
27. Martínez-Hormazábal, J. Función e importancia de la Bioinformática en el desarrollo de las ciencias, especialmente en Biotecnología y Medicina Molecular. *Ciencia & Trabajo* **2006**, 22, 159-163.
28. Cuadras, C. M., *Nuevos métodos de Análisis Multivariante*. Barcelona, 2012; p 297.
29. Cayuela, L., *Análisis multivariante*. Granada, 2010; p 26.
30. Leach, A. R. y. G., V.J., Molecular Descriptors. In *An introduction to Chemoinformatics*, Springer: Dordrecht, 2007.
31. Lucas, *Modelado Molecular en Química Medicinal*. Universidad de Buenos Aires: Buenos Aires, 2011.
32. Langridge, R., Ferrin, T.E., Kuntz, I.D., Connolly, M.L. Real-time color graphics in studies of molecular interactions. *Science* **1981**, 211, 661-666.
33. Gago Badenas, F., *Métodos computacionales de modelado molecular y diseño de fármacos*. Universidad de Alcalá de Henares: Madrid, 2003; p 75.
34. Nobelprize.org Nobel Lecture. The Principles of Conformational Analysis. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1969/barton-lecture.html (15 de Enero),
35. González-Díaz, H. Simulación Molecular. <http://sites.google.com/site/gonzalezdiaz> (26 de Junio),
36. Villoutreix, B. O., Renault, N., Lagorce, D., Sperandio, O., Montes, M., Miteva, M.A. Free resources to assist structure-based virtual ligand screening experiments. *Current Protein and Peptide Science* **2007**, 8, 381-411.
37. Krasky, A., Rohwer, A., Schroeder, J., Selzer, P.M. A combined bioinformatics and Chemoinformatics approach for the development of new antiparasitic drugs *Genomics* **2007**, 89, 36-43.
38. Hammami, R., Fliss I. Current trends in antimicrobial agent research: chemo- and bioinformatics approaches. *Drug Discovery Today* **2010**, 15, 540-546.
39. Lengauer, T., Rarey, M. Computational methods for biomolecular docking. *Curren Opinion in Structural Biology* **1996**, 6, 402-406.
40. Kitchen, D. B., Decornez, H., Furr, J.R. Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery* **2004**, 3, 935-949.
41. De la Madrid, M. REGLAMENTO DE LA LEY GENERAL DE SALUD EN MATERIA DE CONTROL SANITARIO DE ACTIVIDADES, ESTABLECIMIENTOS, PRODUCTOS Y SERVICIOS. <http://www.salud.gob.mx/unidades/cdi/nom/compi/rlgsmcsaeps.html> (27 de Octubre),
42. Woods, L. A. y. D., J. GRAS evaluation of flavoring substances by the expert panel of FEMA. *Regulatory Toxicology and Pharmacology* **1991**, 14, 48-58.
43. Hallagan, J. B. y. H., R.L. FEMA GRAS-A GRAS assessment program for flavor ingredients. *Regulatory Toxicology and Pharmacology* **1995**, 21, 422-430.
44. Burdock, G. A. The GRAS process regulatory impact. *Food Techonology* **2003**, 57, 17-17.
45. Smith, R. L., Cohen, S.M., Doull, J., Feron, V.J., Goodman, J.I., Marnett, L.J., Munro, I.C., Portoghese, P.S., Waddell, W.J., Wagner, B.M., Adams, T.B. Criteria for the safety evaluation of flavoring subdtances: the expert panel of the Flavor and Extract Manufacturers Association. *Food and Chemical Toxicology* **2005**, 43, 1141-1177.
46. Munro, I. C., Shubik, P., Hall, R. Principles for the safety evaluation of Flavoring Substances. *Food ans chemical Toxicology* **1998**, 36, 529-540.
47. Association), F. F. a. E. M. Activities. <http://www.femaflavor.org/> (10 de Enero),
48. Villarejo-Díaz, M., Dr. Murillo-Zaragoza, J.R., Dr. Alvarado-Hernández, H. Farmacología de los agonistas y antagonistas de los receptores opioides. *Educción e Investigación Clínica* **2000**, 1, 106-137.
49. Teschemacher, H. Opioid Receptor Ligands Derived from Food Proteins. *Current Pharmaceutical Design* **2003**, 9, 1331-1344.
50. Clare, D. A. S., H.E. Bioactive Milk Peptides: A prospectus. *Journal of Dairy Science* **2000**, 83, 1187-1195.

51. Meisel, H. Biochemical Properties of Peptides Encrypted in bovine Milk Proteins. *Current Medicinal Chemistry* **2005**, *12*, 1905-1919.
52. Chaieb, K., Hajlaoui, H., Zmantar, T., Kahla-Nakbi., A.B., Rouabhia, M., Mahdouani, K., Bakhrouf A. The Chemical Composition and Biological Activity of Clove Essential Oil, *Eugenia caryophyllata* (*Syzygium aromaticum* L. Myrtaceae): A Short Review. *Phytotherapy Resesearch* **2007**, *21*, 501-506.
53. Lee, Y.-L., Weng, Chun-Chun., Mau, Jeng-Leun. Antioxidant properties of ethanolic and hot water extracts from the rhizome of *Curcuma aromatic*. *Journal of Food Biochemistry* **2006**, *31*, 757-771.
54. ChemAxon Instant JChem via WebStart. <http://www.chemaxon.com/products/online-tryouts/instant-jchem-via-webstart> (9 de Mayo),
55. Sciences, N. I. o. S. PowerMV: A software environment for statistical analysis, molecular viewing, descriptor generation, and similarity search. <http://nisl05.niss.org/PowerMV/?q=PowerMV> (14 de Mayo),
56. Palazzo, E., Luongo, L., de Novellis, V., Berrino, L., Rossi, F., Maione, S. Moving towards supraspinal TRVP1 receptors for chronic pain relief. *Molecular pain* **2010**, *6*, 2-11.
57. Tavares, C. y. S., R.K. Cafeína para el tratamiento del dolor. *Revista Brasileira de Anestesiologia* **2012**, *62*, 387-401.
58. Laska, E. M., Sunshine, A., Mueller, F., Elvers, W.B., Siegel, C., Rubin, A. Caffeine as an analgesic adjuvant. *The Journal of the American Medical Association* **1984**, *251*, 1711-1718.
59. Calle Aznar, S. Determinación analítica de la cafeína en diferentes productos comerciales. Politecnica de Catalunya, Barcelona, 2011.
60. Gomez Alvis, A., Melucci, J.C., Quiroga, P., Mandrile, E. Capsaicina. *Acta Farmacéutica Bonaerense* **1995**, *14*, 63-70.
61. Liu, K., Feng, J., Young, S.S. PowerMV: A software environment for molecular viewing, descriptor generation, data analysis and hit evaluation. *Journal of Chemical Information and Modeling* **2005**, *45*, 515-522.

8. Glosario

- **Acción antitusígena:** fármaco empleado para tratar la tos seca irritativa, no productiva.
- **Campo de fuerzas:** suma de términos asociados a las tensiones de los enlaces, a sus ángulos, torsiones, interacciones de Van der Waals, electrostáticas y las interacciones por formación de puentes de hidrógeno.
- **Consulta subestructural:** es un método para acceder a los datos contenidos en una base de datos, con los cuáles se puede modificar, borrar y agregar datos; la consulta subestructural hace referencia a los cambios realizados en fragmentos, enlaces comodines y especificación de átomos en las moléculas.
- **CAS (*Chemical Abstracts Service*):** Identificación numérica única para compuestos químicos, polímeros, secuencias biológicas, preparados y aleaciones.
- **Endorfinas, encefalinas y dinorfinas:** son neurotransmisores opioides producidos en el Sistema Nervioso Central como moduladores de dolor, reproducción, temperatura corporal, hambre y funciones reproductivas.
- **Opioides:** sustancias endógenas con capacidad para unirse a receptores opéceos, por ejemplo, endorfinas, encefalinas o dinorfinas.
- **Quemoinformática:** método de computo que hace uso de la información química para diseñar, crear, organizar, manejar, recuperar, analizar, distribuir y visualizar de forma más precisa el contenido de una estructura molecular.
- ***Simplified Molecular Input Line Entry Specification (SMILES o especificación de introducción lineal molecular simplificado) y Smiles arbitrary target specification (SMARTS)*:** son notaciones lineales (métodos tipográficos con caracteres imprimibles) para introducir y representar moléculas y reacciones.
- **Sistema opioidérgico hipotalámico:** es una red neuronal bien definida en el interior del núcleo arqueado del hipotálamo mediobasal y se encuentra en íntimo contacto con el sistema de neuronas de GnRH.
- **Teoría de grafos:** (también llamada teoría de las gráficas) es un campo de estudio de las matemáticas y las ciencias de la computación, que estudia las propiedades de los grafos (también llamadas *gráficas*), estructuras que constan de dos partes, el conjunto de vértices, nodos o puntos; y el conjunto de aristas, líneas o lados (*edges* en inglés) que pueden ser orientados o no.

9. ANEXOS

9.1. Anexo 1

Definición de las propiedades fisicoquímicas evaluadas en los programas de ChemAxon y MOE para la caracterización de las bases de datos utilizadas en este trabajo.^{11, 54, 55}

| Propiedad fisicoquímica | Definición | |
|---|---|---|
| | ChemAxon | MOE |
| Peso molecular | Cálculo de la masa molecular. Opciones: -p , - precisión <floating precisión de punto como el número de fracciones digits: 0-8 o inf> (por defecto: la precisión de la masa atómica menos precisa) -s , - single [true false] en el caso de las moléculas de varios fragmentos: Toma fragmento más grande si es cierto, tiene toda la molécula si es falso (por defecto: false). | Peso molecular (incluyendo hidrógenos implícitos) en unidades de masa atómica con pesos atómicos tomadas de [CRC 1994]. |
| Log P (o/w) (coeficiente de partición octanol/agua). | Cálculo de logP: para el tipo logPTrue: logP de especies no cargadas, o, en el caso de iones híbridos, logd en pl; para el tipo logPMicro: logP de las especies de entrada. | Logaritmo del coeficiente de partición octanol / agua (incluyendo hidrógenos implícitos). Esta propiedad se calcula a partir de un modelo tipo de átomo lineal [LOGP 1998], con $r^2 = 0,931$, RMSE = 0,393 en el 1827 moléculas. |
| SlogP (coeficiente de partición octanol/agua). | No aplica | Logaritmo del coeficiente de partición octanol / agua (incluyendo hidrógenos implícitos). Esta propiedad es un modelo de contribución atómica [Crippen 1999] que calcula log P de la estructura dada, es decir, el estado de protonación correcta (estructuras lavados). Los resultados pueden variar de la (o / w) logP descriptor. El conjunto de entrenamiento para SlogP era ~ 7.000 estructuras. |
| b_rotN (número de enlaces rotables). | Contenido de enlaces rotables. Opciones: -s , - single [true false] en el caso de las moléculas de varios fragmentos: toma fragmento más grande si es cierto, tiene toda la molécula si es falso (por defecto: false). | Número de enlaces rotables. Un enlace es rotable si tiene orden 1, no está en un anillo, y tiene al menos dos vecinos pesados. |
| a_nF | No aplica | Número de átomos de flúor: # {Z i Z i = 9}. |
| a_nCl | No aplica | Número de átomos de cloro: # {Z i Z i = 17}. |

| | | |
|--|---|--|
| a_nBr | No aplica | Número de átomos de bromo: # {Z i Z i = 35}. |
| a_nI | No aplica | Número de átomos de yodo: #{Z i Z i = 53}. |
| TPSA o PSA (área de superficie polar topológica). | Cálculo de superficie topológica Polar (2D). Opciones: -p , - precisión <floating precisión de punto como el número de fracciones digits: 0-8 o inf> (por defecto: 2). -H , - pH <ph valor> toma grandes microespecies en este pH (por defecto: no pH, toma la molécula de entrada). -S , - excludesulfur [true false] excluir átomo de azufre de cálculo (por defecto: true). -P , - excludephosphorus [true false] excluir átomo de fósforo del cálculo (por defecto: true). | Superficie polar (\AA^2) calcula mediante contribuciones de los grupos para aproximar la superficie polar sólo de información de la tabla de conexión. La parametrización es la de Ertl et al. [Ertl 2000]. |
| a_acc | Aceptor de enlace de hidrógeno en el átomo de recuento molécula. Opciones: -H , - pH <ph valor> toma grandes microespecies en este pH (por defecto: no pH, toma la molécula de entrada). -e , - excludesulfur [true false] excluir átomo de azufre de los aceptantes (por defecto: true). -x , - excludehalogens [true false] excluir halógenos de aceptantes (por defecto: true). | Número de átomos aceptores de enlaces de hidrógeno (sin contar los átomos de ácidos pero contando átomos que son tanto de los donadores y aceptores de enlaces de hidrógeno, tales como –OH). |
| a_don | Enlace de hidrógeno recuento átomo donador en molécula. Opciones: -H , - pH <ph valor> toma grandes microespecies en este pH (por defecto: no pH, toma la molécula de entrada). | Número de átomos donadores de enlaces de hidrógeno (sin contar los átomos de básicos pero contando átomos que son los donadores y aceptores de enlaces de hidrógeno, tales como –OH). |
| Anillos | Número de anillos. Opciones: -s , - single [true false] en el caso de las moléculas de varios fragmentos: toma fragmento más grande si es cierto, tiene toda la molécula si es falso (por defecto: false). | Número de anillos. |

9.2. Anexo 2

Métodos o tipos de búsqueda y descriptores utilizados en los programas de ChemAxon, PowerMV y MOE, para búsqueda por similitud.

Para el programa de ChemAxon los métodos o tipos de búsqueda que se evaluaron fueron:

- **Substructure:** si una estructura molecular contiene a otra como una subestructura o no. Por definición, la molécula examinada se llama un objetivo, la estructura que estamos buscando se llama una consulta y una molécula diana correspondiente a la estructura de consulta se llama un éxito. Si las características moleculares especiales están presentes en la consulta (por ejemplo, estereoquímica, carga, etc), sólo aquellos objetivos que coinciden también contiene la característica. Sin embargo, si una característica falta en la consulta, no está activada de forma predeterminada.
- **Superstructure:** busca aquellas moléculas objetivo que podrían encontrarse en la molécula de consulta dada. (En este caso los papeles de las moléculas de consulta y objetivo simplemente cambiaron, así que las propiedades para la molécula de consulta deben ser especificadas en el objetivo. En el caso de la búsqueda de superestructura se nota que excepto las tablas de consulta de la normalización por defecto elimina hidrógenos explícitos. Esto tiene el efecto de que las estructuras de la base de datos se utilizan sin sus hidrógenos explícitos como las consultas en estos casos.
- **Similarity** (Tanimoto 0.85 y Eucliden 0.15): encuentra moléculas que son similares a la estructura de la consulta. Por defecto, la búsqueda utiliza coeficiente de Tanimoto. El coeficiente de Tanimoto tiene dos argumentos:
 - la huella digital de la estructura de la consulta.
 - la huella digital de la molécula en la base de datos.

La fórmula de disimilitud contiene el coeficiente de Tanimoto y mide que diferentes son las dos moléculas la una de la otra. La fórmula se define a continuación:

$$1 - \frac{N_{A \& B}}{N_A + N_B - N_{A \& B}}$$

donde N_A y N_B son el número de bits establecidos en la huella digital de la molécula de A y B, respectivamente, $N_{A \& B}$ es el número de bits que están activos en ambas huellas dactilares.

- **Full:** se encuentran moléculas que son iguales (en tamaño) a la estructura de la consulta. (No se permiten fragmentos adicionales o átomos pesados.) Características moleculares (por defecto) se evalúan de la misma manera como se ha descrito anteriormente para la búsqueda de subestructuras.
- **Full Fragment:** la consulta debe coincidir plenamente a todo un fragmento del objetivo. Otros fragmentos que pueden estar presentes en el objetivo, se omiten. Este tipo de búsqueda es útil para llevar a cabo una "búsqueda de la estructura completa" que hace caso omiso de las sales o solventes al lado de la estructura principal en el objetivo.⁵⁴

Los descriptores a evaluar en el programa de PowerMV fueron:

- **Atom pair / Tanimoto:** descriptores pares, útil para encontrar análogos cercanos.
- **Atom pair Carhat / Tanimoto:** 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 / Tanimoto:** se tienen 14 clases, basados en principios bioisostéricos.
- **Pharmacophore Fingerprints / Tanimoto:** 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 / Euclidian:** este descriptor está basado en la carga de la molécula.⁶¹

En el programa de MOE se utilizaron huellas digitales MACCS que consisten en un conjunto de indicadores que muestran si cada una de las 166 claves estructurales se encuentra presente en una molécula dada.¹¹ Se utilizó un coeficiente de Tanimoto de 0.85.

9.3. Anexo 3

Aplicación de análisis de componentes principales y búsqueda por similitud en la industria alimentaria.

Las técnicas empleadas en este proyecto como son el análisis de componentes principales y búsqueda de similitud, se pueden utilizar en la industria alimentaria a continuación se presenta de manera teórica un ejemplo en donde se pueden aplicar estas técnicas.

El siguiente ejemplo esta basado en una práctica realizada en la asignatura de Laboratorio de Tecnología de Alimentos (LABTEC). La práctica lleva por nombre "Control de Calidad en Leche".

En esta práctica se analizaron 4 diferentes tipos de leche:

- Leche Bronca: establo de Nezahualcoyotl, Edo. De México.
- Alpura entera pasteurizada.
- Lala entera pasteurizada.
- UHT (Alpura).

Se evaluaron 13 propiedades fisicoquímicas que son:

- 1) Temperatura de llegada.
- 2) Características sensoriales (evaluación sensorial).
- 3) Prueba lactométrica (peso específico).
- 4) Acidez titulable.
- 5) pH.
- 6) Prueba del alcohol.
- 7) Prueba de neutralizantes.
- 8) Prueba de reducción de azul de metileno.
- 9) Prueba de lactofermentación.
- 10) Detección de inhibidores en la leche (peróxido de hidrógeno).
- 11) Grasa.
- 12) Porcentaje de sólidos totales (%ST).
- 13) Porcentaje de sólidos no grasos (%SNG).

En la Tabla 17 se muestran los resultados de las pruebas realizadas a las 4 leches; la Tabla 18 muestra las especificaciones fisicoquímicas de la leche de acuerdo a la NOM-155SCFI la cual ayudó a establecer la calidad de las leches y en las Tablas 19 y 20 se muestra la Evaluación Sensorial.

Tabla 17. Resultados de las pruebas realizadas a las 4 diferentes tipos de leche.

| Prueba | Leche bronca | Alpura | LALA | UHT |
|----------------------------------|---------------------|-------------------|-------------------|-------------------|
| Temperatura de llegada °C | 17.5 | 15 | 5.5 | 21 |
| Temperatura de prueba °C | 20 | 18 | 13 | 20 |
| °Q (°Quevenne) | 25 | 30 | 31 | 29 |
| °Q corregidos | 24.6 | 29.47 | 31.58 | 28.08 |
| Densidad (mg/mL) | 1.024 | 1.029 | 1.032 | 1.028 |
| pH | 6.59 | 6.68 | 6.7 | 6.61 |
| mL de NaOH | 1.75 | 1.6 | 1.75 | 1.8 |
| % acidez | 0.168 | 0.163 | 0.178 | 0.184 |
| °D (°Dornic) | 16.66 | 16.1636 | 17.67 | 18.8 |
| °SH (°Soxhlet-Henkel) | 7.33 | 7.112 | 7.77 | 8 |
| Alcohol | Positiva | Positiva | Negativa | Negativa |
| Neutralizantes | Negativa | Negativa | Negativa | Negativa |
| Azul de metileno | Aceptable | Buena a excelente | Buena a excelente | Buena a excelente |
| Lactofermentación | Aceptable | Aceptable | Aceptable | Aceptable |
| Peróxido | Negativa | Negativa | Positiva | Negativa |
| % ST lacto. Formula | 9.39 | 11.3275 | 11.735 | 10.26 |
| % ST lacto. Tabla | NR | 11.22 | 11.6 | 10.24 |
| % SNG | 6.69 | 8.0275 | 8.535 | 7.56 |
| % Grasa | 2.7 | 3.3 | 3.2 | 2.7 |
| % Agua añadada | Positiva | Negativa | Negativa | Negativa |

Tabla 18. Especificaciones fisicoquímicas de la leche de acuerdo a la NOM-155-SCFI.

| Especificaciones | Límite |
|--|--|
| | Leche entera pasteurizada y UHT |
| Densidad a 15°C, g/mL | 1.029 mín. |
| Grasa butírica g/L | 30 mín. |
| Acidez (expresada como ac. Láctico) g/L | 1.3-1.7 |
| pH | 6.5-6.7 |
| Sólidos totales % m/m | 83 mín. |

Tabla 19. Evaluación Sensorial.

| Leche | Juez | Color | Olor | Sabor | Dulzor | Consistencia |
|---------------|-------------|--------------|-------------|--------------|---------------|---------------------|
| Bronca | 1 | 2 | 1 | 2 | 2 | 3 |
| | 2 | 1 | 2 | 2 | 2 | 3 |
| | 3 | 3 | 1 | 2 | 1 | 3 |
| | 4 | 5 | 1 | 1 | 1 | 4 |
| Alpura | 1 | 2 | 2 | 1 | 3 | 1 |
| | 2 | 1 | 1 | 1 | 1 | 1 |
| | 3 | 2 | 2 | 3 | 3 | 3 |
| | 4 | 3 | 3 | 3 | 3 | 1 |
| UHT | 1 | 1 | 4 | 3 | 5 | 1 |
| | 2 | 1 | 4 | 5 | 5 | 1 |
| | 3 | 5 | 5 | 5 | 5 | 1 |
| | 4 | 2 | 5 | 5 | 5 | 1 |
| Lala | 1 | 1 | 2 | 3 | 2 | 1 |
| | 2 | 1 | 1 | 2 | 2 | 1 |
| | 3 | 2 | 2 | 4 | 2 | 2 |
| | 4 | 1 | 2 | 4 | 4 | 2 |

Atributos sensoriales evaluados:

- Olor: fresco(1)-----cocido(5)
- Color: blanco(1)-----amarillo(5)
- Sabor: fresco(1)-----cocido(5)
- Dulzor: natural(1)-----dulce(5)
- Consistencia: líquida(1)-----Espesa(5)

Tabla 20. Promedio de los puntajes para cada atributo sensorial evaluado.

| Leche/Atributo | Color | Olor | Sabor | Dulzor | Consistencia |
|-----------------------|--------------|-------------|--------------|---------------|---------------------|
| Bronca | 2.75 | 1.25 | 1.75 | 1.5 | 3.25 |
| Alpura | 2 | 2 | 2 | 2.5 | 1.5 |
| UHT | 2.25 | 4.5 | 4.5 | 5 | 1 |
| Lala | 1.25 | 1.75 | 3.25 | 2.5 | 1.5 |

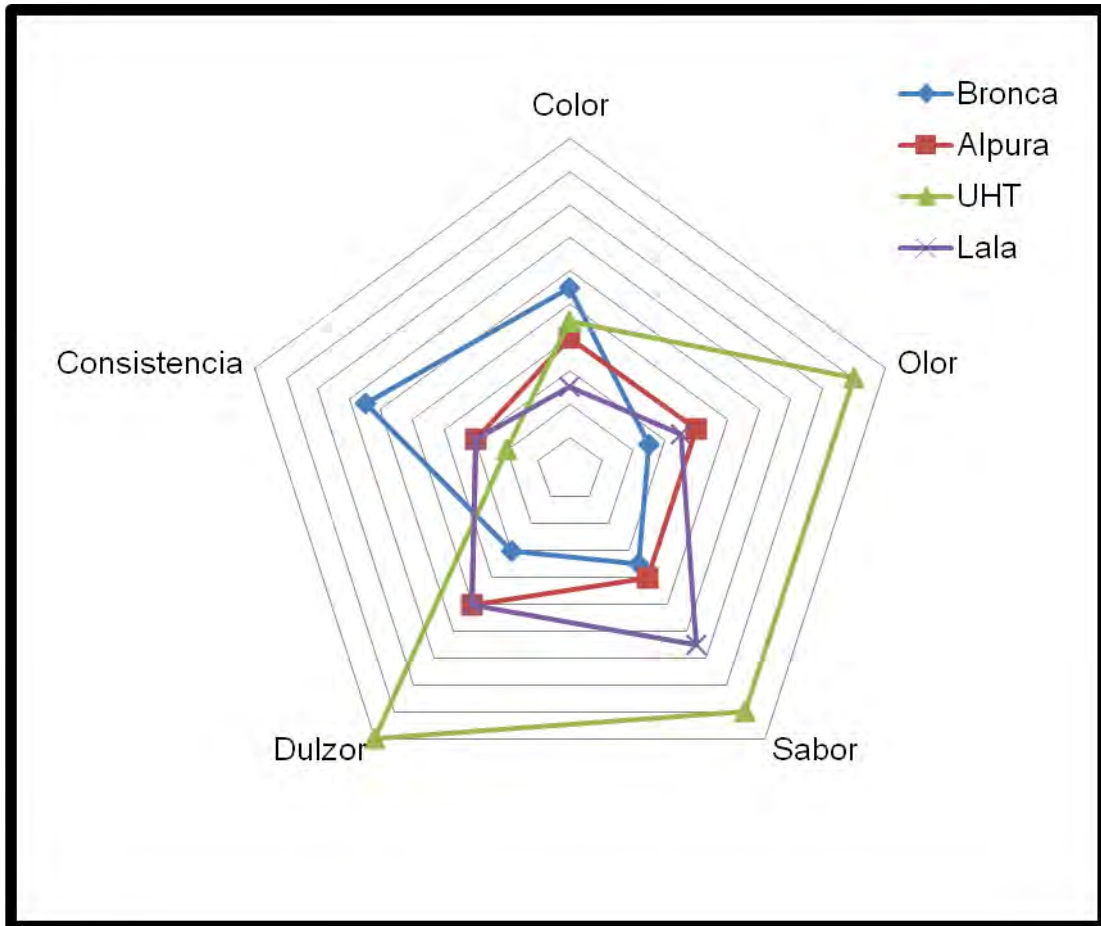


Figura 12. Gráfico radial. Características sensoriales evaluadas en las cuatro leches analizadas.

En la Tabla 17 se observa que se realizó una búsqueda de similitud por propiedades fisicoquímicas con las cuales se encontró que en general las leches comerciales presentaron una calidad buena lo cual habla de que tuvieron un buen tratamiento térmico y que la materia prima utilizada pasó las pruebas de calidad de campo y plataforma exigidas para la elaboración de leche. Se llegó a esta conclusión tomando en cuenta los parámetros que especifica la NOM-155SCFI mostrados en la Tabla 18.

Por último la Tabla 19 nos muestra el puntaje que cada juez (no entrenado) dio a cada atributo sensorial. Cabe mencionar que el equipo estaba conformado por 4 personas, la Tabla 20 indica el promedio para cada característica sensorial, en la Figura 12 se muestra un gráfico radial de cada atributo sensorial evaluado, el cual nos indica que dependiendo del tipo de leche la característica sensorial varía, para la leche bronca las características que afectan a que se acepte como una leche de buena calidad son la consistencia seguida por el color; en la leche pasteurizada Alpura se observa que siguió una tendencia y que ninguna característica sensorial afecta su calidad; en la leche UHT las características

sensoriales que la afectan son dulzor, sabor y olor; y para la leche Lala las características sensoriales que la afectan son sabor y dulzor.

En la Figura 12 también se observa que a pesar de que las leches Alpura y Lala tengan el mismo tratamiento térmico (pasteurización) hay diferencia aunque poco significativa (0.75 a 1) en cuanto a color y sabor, a su vez existe diferencia significativa (2.5 a 2.75, 0.25 a 2.5 y 2.5) para olor, sabor y dulzor, respectivamente. Con la leche con tratamiento térmico UHT; también existe diferencia significativa (1.75) en cuanto a consistencia con la leche bronca. Así como se observan diferencias también se observan semejanzas, entre las leches con diferente tratamiento térmico. Entre las leches Alpura, UHT y bronca; se observa semejanzas para color; en cuanto a la leche bronca se observa que tiene el mismo sabor que la leche Alpura.

En conclusión una leche de excelente calidad es aquella que cumple con lo establecido en la NOM-155SCFI así como aquella que tenga un olor y sabor fresco, dulzor natural, color blanco y consistencia líquida, esto está de acuerdo a lo evaluado en la práctica.