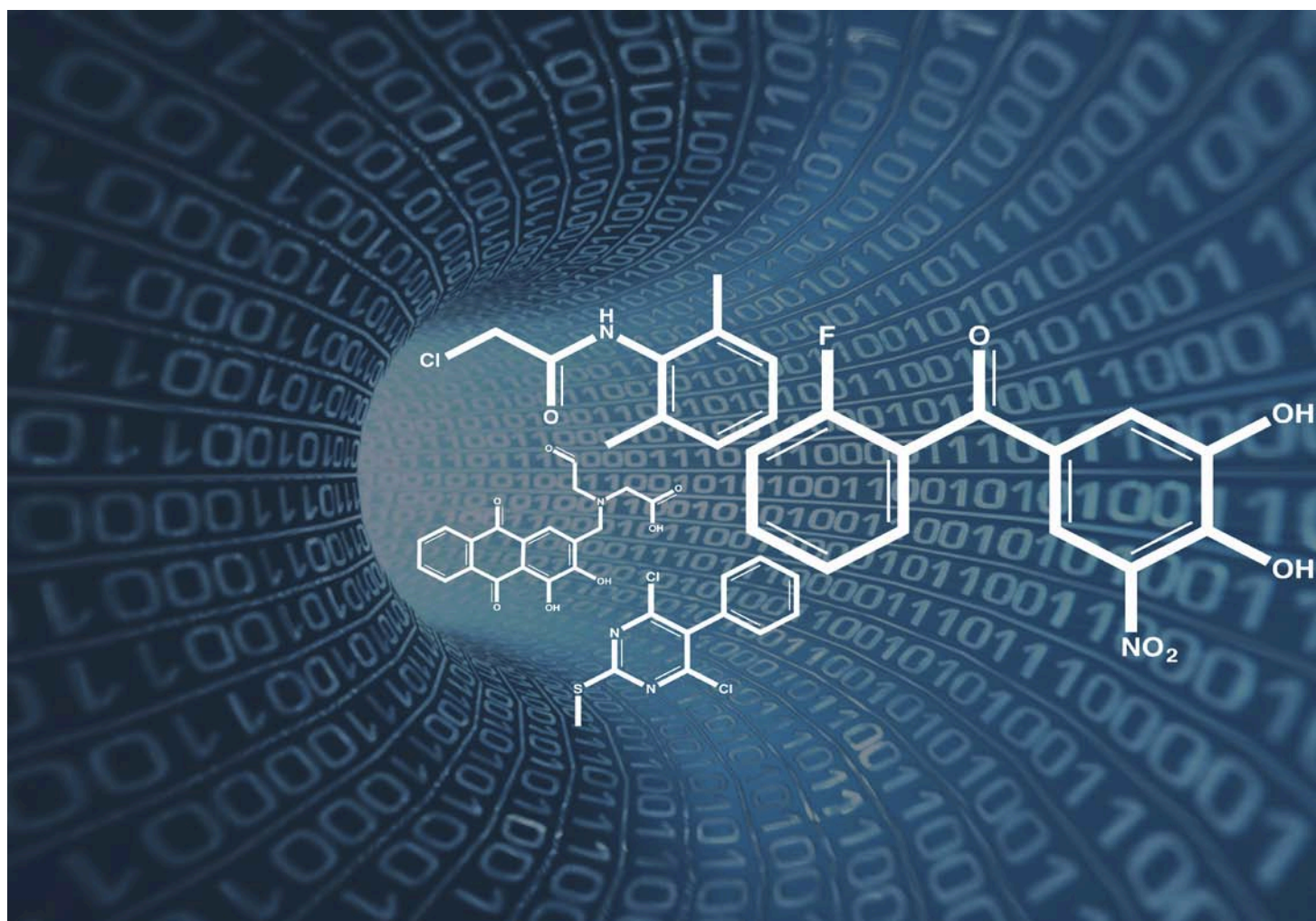




UNIVERSITAT DE VALENCIA

Doctorado en Biomedicina y Farmacia

**DISEÑO POR TOPOLOGÍA MOLECULAR Y ENSAYOS FARMACOLÓGICOS  
DE NUEVOS ANTIINFLAMATORIOS**



Directores:

Mari Carmen Recio Iglesias

Ramón García Domenech

Tesis Doctoral

María Gálvez Llompарт

Valencia, 2016





D. Ramón García Domenech Catedrático de Química Física y Dña. M<sup>a</sup> del Carmen Recio Iglesias Profesora titular de Farmacología, de la Universitat de València.

**CERTIFICAN:**

Que el trabajo presentado por la Lda. **María Gálvez Llompert**, titulado “**Diseño por Topología molecular y ensayos farmacológicos de nuevos antiinflamatorios**”, ha sido realizado en los departamentos de Química Física y Farmacología de la Universitat de València, bajo nuestra dirección y asesoramiento.

Concluido el trabajo experimental y bibliográfico, autorizamos la presentación de la Tesis, para que sea juzgada por el tribunal correspondiente.

Lo que firmamos en Valencia, Mayo de dos mil dieciséis

Prof.Dr. Ramón García Domenech

Prof.Dra. M<sup>a</sup> del Carmen Recio Iglesias



A mi hija  
A mi marido  
A mi familia



## AGRADECIMIENTOS

A mi directora, **Mari Carmen**, por su cercanía, su amabilidad y por haberme transmitido la pasión por la Farmacología.

A mi codirector, **Ramón**, por su paciencia, su afabilidad y por haber sabido guiarme en todas las etapas de mi carrera investigadora sin dejar de creer en mí.

A la **Universitat de Valencia**, por los medios puestos a mi disposición y permitirme disfrutar del programa de Atracció de Talents durante 4 maravillosos años.

A la **Università di Bologna**, por la gran acogida que siempre me ha brindado. Concretamente, a la Prof. Patrizia Romualdi, la coordinadora erasmus que cambió mi vida. Gracias a ti conocí al que es hoy mi marido y gracias a ti conocí lo que significa trabajar por primera vez en un laboratorio extranjero. Gracias a todas las personas con las que compartí mi estancia en Italia. Gracias al Dr. Alberto del Rio, por su disponibilidad, su comprensión y su paciencia, me enseñaste que se puede emprender y ser científico al mismo tiempo.

A todos los **compañeros/as** y **profesores visitantes** que he podido conocer a lo largo de estos años en el laboratorio de Farmacología.

A todos los del **departamento de Farmacología** y del de **Química Física** que he tenido el gusto de conocer: Rosa, José Luís, Salva, Mamen, Lionello, Olga, Facundo, Julián, Juan Luis, Nacho, Pepe...

A mi padre, **Jorge**, por haberme transmitido la pasión por la Topología molecular. He hecho mía tu lucha porque esta disciplina de la química matemática se reconozca algún día como el paradigma nuevo que representa. Dicen que tenemos el mismo carácter (yo espero que no sea así), pero algo de razón tienen que llevar porque somos igual de testarudos.

A mi madre, **María José**, por ser la mujer que más admiro la que desde pequeña me transmitió la cultura del esfuerzo y del trabajo diario. Gracias a ti he logrado llegar a este momento. Gracias por ser amiga, compañera y, sobre todo, madre. Tú me has enseñado que hay que ser humilde y buena persona por encima de todo. Ahora que soy madre me doy cuenta de que la ciencia es apasionante, pero formar parte de una familia es lo más importante para mí. En eso tú has sido mi “directora”.

A mi marido, **Riccardo**, el amor de mi vida y la persona que tiene la paciencia de soportarme día tras día. Nos conocimos allá por septiembre de 2006 y quien nos iba a decir a nosotros que un Erasmus cambiaría tanto nuestras vidas. Allá por mayo de 2007 comenzamos a escribir juntos nuestra historia. Hoy te has convertido no solo en mi mejor amigo, mi marido y el padre de la luz de mi vida, sino que haces que cada día aprenda de ti y me vaya haciendo una mejor persona. Me transmites tu pasión por los pequeños detalles, tu cercanía con las personas, tu alegría y tu sentimiento de la justicia. Ti amo.

A mi hija, **Claudia**, mi luz desde que llegaste a este mundo un 21 de febrero de 2014 y la lente con la cambié mi visión del mundo. Gracias porque cuando vuelvo del trabajo y he tenido un día duro y te veo de lejos se me dibuja una sonrisa que nadie puede borrar. Te amo.

Al resto de **mi familia**, mi hermano Jorge, mi tía Helen, por compartir mis vivencias durante estos años.

A mis **amigos**, vosotros sabéis quienes sois. Por estar siempre a mi lado y transmitirme vuestra pasión por vuestras profesiones, por apoyarme y escucharme tantas y tantas veces sobre lo duro que es la vida del investigador.

Alla mia **famiglia italiana**, perche sempre ci aiutate dove e come sia per sopraporci ad ogni imprevisto. Perche mi avete fatto amare l' Italia, la gente, il cibo e la cultura.

A todas las personas que veo a diario y que me han hecho tan agradables estos años en la Universidad como el servicio de limpieza y el personal del bar de Farmacia.

A todos a los que no he citado y habéis contribuido a que hoy esté escribiendo estas líneas.

Por último, a todos los que ya no están y me hubiera gustado compartir con ellos este momento. Concretamente a ti, **Marga**.



La presente Tesis Doctoral ha sido financiada por el proyecto SAF2009-13059-C03-02 concedido por el Ministerio Español de Ciencia e Innovación (MICINN) junto al fondo europeo de Desarrollo Regional y la ayuda concedida por la Universitat de València (UV-INV-AE14-265774).

La realización de la presente Tesis Doctoral ha sido posible gracias al contrato predoctoral VLC-Campus d' Atracció de talents, concedida por la Universitat de València.



La presente Tesis Doctoral ha dado lugar a las siguientes publicaciones:

### ARTÍCULOS

M. Gálvez-Llompарт; R. M. Giner; M. C. Recio; S. Candeletti; R. García-Domenech. Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity. *Letters in Drug Design & Discovery*, **2010**, 7, 438-445.

M. Gálvez-Llompарт; M. C. Recio; R. García-Domenech. Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B. *Molecular Diversity*, **2011**, 15, 917–926.

J. Galvez; M. Galvez-Llompарт; R. García-Domenech. Molecular topology as a novel approach for drug discovery. *Expert Opinion On Drug Discovery*, **2012**, 7(2), 133-153.

J. Gálvez; M. Gálvez-Llompарт; R. García-Domenech. Introduction to Molecular Topology: Basic Concepts and Application to Drug Design. *Current Computer-Aided Drug Design*, **2012**, 8(3), 196- 223.

M. Galvez-Llompарт; M. C. Recio Iglesias; J. Gálvez; R. García-Domenech. Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW264.7 cell lines. *Molecular Diversity*, **2013**, 17(3),573–593.

R. García-Domenech; M. Gálvez-Llompарт; R. Zanni; M. C. Recio; J. Gálvez. QSAR methods for the discovery of new inflammatory bowel disease drugs. *Expert Opinion On Drug Discovery*, **2013**, 8(8), 933- 949.

R. Zanni; M. Galvez-Llompарт; R. García-Domenech; J. Galvez. Latest advances in molecular topology applications for drug discovery. *Expert Opinion On Drug Discovery*, **2015**, 10(9), 945-957.

M. Gálvez-Llompарт; M. C. Recio; R. García-Domenech; J. Gálvez. Molecular topology: A strategy to identify novel compounds against ulcerative colitis. *Molecular Diversity* **2016**, XXX. *Enviado*.

## COMUNICACIONES A CONGRESOS

J. Gálvez Alvarez; M. Gálvez Llompart; R. García Domenech. Application of Molecular Topology to the Search of New Compounds Active in Ulcerative Colitis and Crohn's Disease. 11th Annual Drug Discovery, Montreux (Suiza), 2010. **COMUNICACIÓN ORAL.**

M. Galvez-Llompart; R. Zanni; R. Garcia-Domenech. New inhibitors of IL-6 production in Caco-2 cells through molecular topology methodology. XXXIII Congreso Nacional de la Sociedad Española de Farmacología, Malaga (España), 2011. **POSTER.**

M. Galvez-Llompart; M. C. Recio; R. Garcia-Domenech. New compounds against ulcerative colitis through molecular topology. European Crohn's and Colitis Organisation Congress, Barcelona (España), 2012. **POSTER.**

# ÍNDICE

<b>I. RESUMEN</b> .....	3
<b>II. INTRODUCCIÓN</b> .....	9
II.1. DISEÑO Y DESCUBRIMIENTO DE FÁRMACOS ASISTIDO POR ORDENADOR .	9
II.2. QSAR .....	13
II.3. TOPOLOGÍA MOLECULAR .....	17
<b>II.3.1. Teoría de Grafos</b> .....	17
II.3.1.1. Principios.....	17
II.3.1.2. Moléculas representadas como grafos.....	21
<b>II.3.2. Índices topológicos</b> .....	22
<b>II.3.3. Diseño de fármacos por Topología molecular</b> .....	28
II.4. INFLAMACIÓN .....	31
<b>II.4.1. Enfermedades inflamatorias</b> .....	32
<b>II.4.2. Principales mediadores inflamatorios</b> .....	34
II.5. ENFERMEDAD INFLAMATORIA INTESTINAL .....	36
<b>II.5.1. Colitis ulcerosa</b> .....	36
II.5.1.1. Factores implicados en la CU.....	37
II.5.1.2. Prevalencia e incidencia de la CU .....	39
II.5.1.3. Impacto socio-económico de la CU.....	39
II.5.1.4. Terapia actual de la CU .....	41
II.5.1.5. Mediadores implicados en la CU.....	44
<b>III. OBJETIVOS</b> .....	51
<b>IV. CAPÍTULO 1</b> .....	55
Artículo 1: “ <b>Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity</b> ” (En: <i>Letters in Drug Design &amp; Discovery</i> , <b>2010</b> , 7, 438-445). JCR [2010, CAT: Medicinal Chemistry; FI: 0, 668; Rank: 48 de 54 ( <b>Q4</b> )].....	55
<b>V. CAPÍTULO 2</b> .....	65
Artículo 2: “ <b>Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF-κB</b> ” (En: <i>Molecular Diversity</i> <b>2011</b> , 15:917–926). JCR [2011, CAT: Chemistry Applied; FI: 3,153; Rank: 9 de 71 ( <b>Q1</b> )].....	65
Artículo 3: “ <b>Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW264.7 cell lines</b> ” (En: <i>Molecular Diversity</i> <b>2013</b> , 17(3):573–593). JCR [2013, CAT: Chemistry Applied; FI: 2,544; Rank: 15 de 71 ( <b>Q1</b> )].....	75
Artículo 4: “ <b>Molecular topology: A strategy to identify novel compounds against ulcerative colitis.</b> ” (En: <i>Molecular Diversity</i> <b>2016</b> , XXX). Enviado.....	96

<b>VI. RESULTADOS Y DISCUSIÓN</b> .....	146
VI.1. CAPÍTULO 1 .....	146
VI.2. CAPÍTULO 2 .....	158
<b>VII. CONCLUSIONES</b> .....	205
<b>VIII. BIBLIOGRAFÍA</b> .....	209
<b>IX. ANEXOS</b> .....	219
<b>X. SUMMARY</b> .....	247
<b>XI. CONCLUSIONS</b> .....	251

## LISTADO DE FIGURAS

### I. INTRODUCCIÓN

**Figura 1.** Esquema típico de cribado virtual (*virtual screening*) por Topología molecular.

**Figura 2.** Tipos de estrategias empleadas en las técnicas computacionales aplicadas al descubrimiento de fármacos (CADD).

**Figura 3.** Esquema de los puentes de la ciudad de Königsberg (Prusia Oriental) en 1700.

**Figura 4.** Grafo de Euler basado en el esquema de los puentes de la ciudad de Königsberg (Prusia Oriental) en 1700.

**Figura 5.** Grafo simple con su correspondiente matriz de adyacencia (A).

**Figura 6.** Ejemplo de transformación del isopentano en un grafo y posteriormente, en una matriz de adyacencia.

**Figura 7.** Cálculo del índice de Wiener para la molécula del n-butano.

**Figura 8.** Cálculo del índice de Randić para la molécula del 3-metil-pentano.

**Figura 9.** Índices de conectividad de Kier y Hall.

**Figura 10.** Obtención de la matriz M, a partir de la cual se generan los ITC.

**Figura 11.** Enfermedades asociadas a procesos inflamatorios crónicos.

**Figura 12.** Principales mediadores inflamatorios.

**Figura 13.** Las características clínicas e histológicas de la CU y los factores causales que influyen en el riesgo de esta enfermedad.

**Figura 14.** Principales citocinas implicadas en la CU.

## V. RESULTADOS Y DISCUSIÓN GENERALES

**Figura 15.** Diagrama de distribución farmacológica para la actividad antiinflamatoria representando la expectativa de actividad o inactividad (E) frente al valor de la  $FD_1$  para las distintas moléculas que conforman el *training set*.

**Figura 16.** Compuestos ensayados que presentan un mejor perfil antiinflamatorio.

**Figura 17.** Diagrama de distribución farmacológica de inhibidores de NF- $\kappa$ B obtenida utilizando la  $FD_1$ .

**Figura 18.** Diagrama de distribución farmacológica de compuestos activos frente a CU obtenida utilizando la  $FD_2$ .

**Figura 19.** Ejemplo de compuestos del *training set* (grupo inactivo) con anillos de cinco miembros con heteroátomos.

**Figura 20.** Ejemplo de compuestos del *training set* (grupo activo) sin anillos de cinco miembros con heteroátomos.

**Figura 21.** Relación del grado de conjugación presente en las moléculas con actividad inhibitoria frente a IL-6.

**Figura 22.** Relación de la presencia de átomos electronegativos y la aromaticidad con la actividad inhibitoria frente a IL-6.

**Figura 23.** Gráfico de curva ROC para la  $FD_3$ .

**Figura 24.** Diagrama de distribución farmacológica de inhibidores de IL-6 obtenido mediante la  $FD_3$ .

**Figura 25.** Compuestos seleccionados como potenciales inhibidores de IL-6.

**Figura 26.** Diagrama de distribución farmacológica para la actividad inhibitoria frente a iNOS empleando el valor de la  $FD_4$ .

**Figura 27.** Representación gráfica de los valores de  $\log CI_{50}$  iNOS observados y calculados (arriba). Representación gráfica de los residuales calculados pertenecientes al *training set* versus los residuales obtenidos a partir de la validación cruzada (residual (cv) (abajo)).



**Figura 28.** Representación gráfica del coeficiente de predicción,  $Q^2$ , frente el coeficiente de correlación,  $r^2$ , obtenidos en el estudio de aleatoriedad del modelo de predicción del  $\log CI_{50}$  para iNOS. Punto negro: modelo seleccionado.

**Figura 29.** Representación gráfica del residual frente al valor observado para cada compuesto, empleando la Ecuación 5. Se considera *outlier* a todo compuesto con un valor residual predicho superior a  $\pm 2*EEE$ .

**Figura 30.** Compuesto del *training set* sin anillos hexagonales y sus valores de  $CI_{50}$  para iNOS.

**Figura 31.** Compuesto del *training set* con anillos hexagonales y sus valores de  $CI_{50}$  para iNOS.

**Figura 32.** Diagrama de distribución farmacológica de los inhibidores de TNF- $\alpha$  obtenidos usando la  $FD_5$ .

**Figura 33.** Diagrama de distribución farmacológica de los compuestos activos frente a CU inducida por DSS obtenido utilizando la  $FD_6$ .

**Figura 34.** Valores de puntuación del IAE para los compuestos en el modelo de CU inducida por DSS en ratones BALB/c.

**Figura 35.** Imagen representativa del efecto de los compuestos sobre el acortamiento del colon por CU.

**Figura 36.** Evaluación histológica del efecto protector de AMA, DLT y Ro 41-0960 frente a CU inducida por DSS en ratones.

**Figura 37.** Compuestos seleccionados por Topología molecular como potenciales agentes protectores frente a CU y sus valores de  $CI_{50}$  del inhibidor de iNOS predicho por la Ecuación 5.



## LISTADO DE TABLAS

### I. INTRODUCCIÓN

**Tabla 1.** Nuevos compuestos 'cabeza de serie' seleccionados por TM.

**Tabla 2.** Compuestos patentados por la “Unidad de Diseño de Fármacos y Topología Molecular”.

**Tabla 3.** Tratamientos farmacológicos de aplicación a la CU.

### V. RESULTADOS Y DISCUSIÓN GENERALES

**Tabla 4.** Matriz de clasificación obtenida a partir de la  $FD_1$  para el *training set*.

**Tabla 5.** Nuevos agentes antiinflamatorios seleccionados por Topología molecular.

**Tabla 6.** Efecto antiinflamatorio de indometacina y compuestos (5, 8, 9 y 14) en un modelo agudo de edema plantar inducido por carragenina en ratón.

**Tabla 7.** Efecto antiinflamatorio de indometacina y compuestos (5, 8, 9 y 14) en el modelo inflamatorio agudo del edema auricular inducido por TPA en ratón.

**Tabla 8.** Valores de  $FD_1$ - $FD_2$ , probabilidad de actividad, clasificación por parte de los modelos y actividad o categoría terapéutica encontrada en literatura para los compuestos seleccionados como potencialmente activos frente a CU.

**Tabla 9.** Matriz de clasificación obtenida a través de la  $FD_3$  para el *training* y *test set*.

**Tabla 10.** Valores de la  $FD_3$ , probabilidad de actividad, clasificación por el modelo y actividad biológica descrita en literatura para los compuestos seleccionados como potenciales inhibidores de IL-6 en el proceso de cribado virtual.

**Tabla 11.** Efecto de los compuestos sobre la producción de IL-6 en las líneas celulares Caco-2 y Raw 264.7 determinado por ELISA.

**Tabla 12.** Matriz de clasificación obtenida a través de la  $FD_4$ ,  $FD_5$  y  $FD_6$  para el *training* y *test set*.

**Tabla 13.** Valor experimental, predicho y test de validación cruzada para el modelo de regresión multilineal y valor de los descriptores para los compuestos que conforman el *training set*.

**Tabla 14.** Valores de  $FD_{4-6}$  y  $\log CI_{50}$  (iNOS): probabilidad de actividad y clasificación de los potenciales compuestos anti CU.

**Tabla 15.** Efecto de los compuestos seleccionados sobre la secreción de NO, IL-1 $\beta$  y TNF- $\alpha$  en el medio de cultivos de las células RAW 264.7 estimuladas con LPS (1  $\mu$ g/mL), 2 horas después de la administración de los productos.

**Tabla 16.** Pérdida de peso de los ratones BALB/c durante el modelo de CU inducida por DSS.

**Tabla 17.** Efecto de los compuestos sobre el acortamiento de la longitud del colon en la CU inducida por DSS en ratones BALB/c.

**Tabla 18.** Sinopsis de los ensayos *in vitro* e *in vivo* llevados a cabo para validar los modelos y las actividades farmacológicas predichas por estos, para los cuatro compuestos seleccionados.

## ABREVIATURAS

<b>5-ASA</b>	Ácido 5-aminosalicílico
<b>A</b>	Absorbancia
<b>ADME/Tox</b>	Absorción, distribución, metabolismo, excreción
<b>AINE</b>	Antiinflamatorios no esteroideos
<b>ALD</b>	Análisis linear discriminante
<b>AMA</b>	Ácido Alizarin-3-metilninodiacético
<b>ANOVA</b>	Análisis de varianza
<b>AOM</b>	Azoximetano
<b>ARML</b>	Análisis de regresión multilínea
<b>AUC</b>	Área bajo la curva ( <i>Area under the curve</i> )
<b>CADD</b>	Diseño de fármacos asistido por ordenador ( <i>Computer-Aided Drug Design</i> )
<b>CAS</b>	Chemical abstract service
<b>CCR</b>	Cancer colorrectal
<b>CI<sub>50</sub></b>	Concentración inhibitoria 50
<b>CoMFA</b>	Análisis comparativo de campo molecular ( <i>Comparative molecular field analysis</i> )
<b>CoMSIA</b>	Análisis comparativo de los índices de similitud molecular ( <i>Comparative molecular similarity index analysis</i> )
<b>COX</b>	Ciclooxigenasa
<b>CU</b>	Colitis ulcerosa
<b>CV</b>	Validación cruzada ( <i>cross-validation</i> )
<b>DDF</b>	Diagrama de distribución farmacológica
<b>DLT</b>	(+)-dibencil-L-tartrato
<b>DMSO</b>	Dimetilsulfóxido

<b>DSS</b>	Dextrano sulfato sódico
<b>Ea</b>	Expectativa de actividad
<b>EC</b>	Enfermedad de Crohn
<b>EEE</b>	Error de estimación estandard
<b>EEM</b>	Error estándar de la media
<b>Ei</b>	Expectativa de inactividad
<b>EII</b>	Enfermedad inflamatoria intestinal
<b>ELISA</b>	Ensayo por inmunoadsorción ligado a enzimas <i>(Enzyme-linked immunosorbent assay)</i>
<b>F</b>	Parámetro de Fisher-Snedecor
<b>FD</b>	Función discriminante
<b>FN</b>	Falsos negativos
<b>FP</b>	Falsos positivos
<b>FPF</b>	Fracción de falsos positivos
<b>HIF1<math>\alpha</math></b>	Subunidad alfa del factor 1 inducible por hipoxia
<b>IAE</b>	Índice de actividad de la enfermedad
<b>IFN</b>	Interferón
<b>IL</b>	Interleucina
<b>iNOS</b>	Óxido nítrico sintasa inducible
<b>IT</b>	Índices topológicos
<b>ITC</b>	Índices topológicos de carga
<b>JAK</b>	Janus quinasas
<b>LOO</b>	<i>Leave one out</i>
<b>LPS</b>	Lipopolisacárido
<b>MCC</b>	Coefficiente de correlación de Matthews
<b>MPO</b>	Mieloperoxidasa
<b>MTT</b>	3-(4,5-dimetiltiazol-2-il)-2,5-difenil-bromuro de tetrazolio

<b>NF-<math>\kappa</math>B</b>	Factor Nuclear kappa B
<b>NC</b>	No clasificado
<b>NO<sub>2</sub></b>	Dióxido de nitrógeno
<b>NT</b>	No testado
<b>PG</b>	Prostaglandina
<b>PPR</b>	Patógenos de la superficie celular
<b>QSAR</b>	Relaciones cuantitativas de estructura-actividad <i>(Quantitative structure–activity relationship)</i>
<b>QSPR</b>	Relaciones cuantitativas estructura-propiedad
<b>ARN</b>	Ácido ribonucleico
<b>RNA</b>	Redes neuronales artificiales
<b>ROC</b>	Característica operativa del receptor <i>(Receiver operating characteristics)</i>
<b>Se</b>	Sensibilidad
<b>Sp</b>	Especificidad
<b>STAT</b>	Transductor de señal y activador de la transcripción
<b>TGF-<math>\beta</math></b>	Factor de crecimiento transformante
<b>TM</b>	Topología molecular
<b>TNF</b>	Factor de necrosis tumoral
<b>TPA</b>	13-acetato-12-O-tetradecanoilforbol
<b>TPF</b>	Fracción de verdaderos positivos
<b>VN</b>	Verdaderos negativos
<b>VP</b>	Verdaderos positivos





# RESUMEN



## I. RESUMEN

El objetivo de la presente Tesis Doctoral es la selección de nuevos compuestos cabezas de serie con actividad antiinflamatoria empleando una innovativa estrategia basada en la aplicación de la Topología molecular (TM), un método que, esencialmente, permite describir estructuras químicas de un modo matemático.

Se han desarrollado numerosos modelos matemático-topológicos además de llevar a cabo diversos ensayos *in vitro* e *in vivo*, cuyos resultados han sido publicados en revistas especializadas.

El **Artículo 1** discute el desarrollo y elaboración de una ecuación predictiva dirigida a la identificación de nuevos compuestos antiinflamatorios no esteroideos (AINE) inhibidores de la ciclooxigenasa (COX). Este modelo se aplicó al cribado virtual de la base de datos comercial Sigma-Aldrich, seleccionándose un total de 14 compuestos con potencial actividad antiinflamatoria. Cuatro de ellos (compuestos **5**, **8**, **9** y **14**), mostraron además actividad *in vivo*, en un modelo de inflamación aguda en pata de ratón. Todos los compuestos fueron ensayados *in vitro*, a excepción de los insolubles en medios acuosos (compuestos 6, 10 y 12) y del compuesto 4 que ya estaba descrito como antiinflamatorio. A la concentración ensayada de 100  $\mu\text{M}$ , ninguno de los productos fue citotóxico. Los compuestos **8** y **9**, activos *in vivo*, también mostraron actividad *in vitro* al inhibir la producción de Prostaglandina E<sub>2</sub> (PGE<sub>2</sub>) y de nitritos.

A la vista de los resultados se decidió obtener nuevos modelos matemáticos que contemplaran la inhibición de otros mediadores proinflamatorios, tales como citocinas, factores de transcripción, enzimas, etc, implicados en general en procesos inflamatorios y en particular en la colitis ulcerosa. Así fue como se elaboró un segundo artículo (**Artículo 2**), centrado en la predicción de la actividad protectora frente a colitis ulcerosa (CU) mediante la inhibición del factor nuclear (NF)- $\kappa\text{B}$ . En este caso, se construyeron dos funciones discriminantes: la primera de ellas, diferencia compuestos inhibidores del NF- $\kappa\text{B}$  de aquellos que no lo eran, mientras que la segunda distingue compuestos activos e inactivos frente a la CU. Tras validar externamente los dos modelos, se llevó a cabo un cribado virtual en las bases de datos Microsource Pure Collection Natural Products (800 compuestos naturales) y Merck index

(aproximadamente 12000 compuestos), seleccionándose finalmente un grupo de 28 compuestos para su posterior ensayo *in vitro* e *in vivo*.

Por otra parte, la interleucina (IL)-6, es una citocina ampliamente presente en los tejidos y en el plasma de los enfermos que presentan inflamación crónica, en enfermedades tales como artritis, CU y también cáncer, lo que la convierte en una atractiva diana para el tratamiento de dichas enfermedades. Así pues, en el tercer artículo (**Artículo 3**), se construyó un modelo QSAR para predecir la inhibición de la IL-6. El modelo, que resultó estadísticamente significativo y demostró buena capacidad predictiva, se aplicó a los 28 potenciales agentes anticólicos seleccionados en el **Artículo 2**, para determinar la inhibición de la IL-6. La actividad anti IL-6 de los compuestos: ácido Alizarin-3-metiliminodiacético (**AMA**), (+)-dibencil-L-tartrato (**DLT**), **Calceína** y **Ro 41-0960**, fue corroborada experimentalmente *in vitro* frente a dos líneas celulares (Caco-2 y RAW 264.7). El compuesto **Ro 41-0960** resultó ser el más activo, al inhibir la producción de IL-6 en ambas líneas, obteniéndose valores de  $CI_{50}$  de 12  $\mu$ M en las Caco-2 y 0,4  $\mu$ M en los macrófagos RAW 264.7. Le sigue el compuesto **DLT** con una  $CI_{50}$  de 19  $\mu$ M en ambas líneas. Por último, la **Calceína** inhibió la producción de IL-6 solo en las Caco-2, mientras que **AMA** resultó inactivo.

El papel del factor de necrosis tumoral (TNF)- $\alpha$  y la óxido nítrico sintasa inducible (iNOS) en el desarrollo y cronicidad de la CU, nos animó a desarrollar 4 modelos topológicos capaces de predecir la actividad inhibitoria frente a dichos mediadores (**Artículo 4**). En este caso, para construir los modelos se emplearon las técnicas estadísticas del análisis lineal discriminante (LDA) y análisis de regresión multi-lineal (MLRA). El cribado virtual estableció la actividad de los 4 compuestos descritos como inhibidores de IL-6. En esta ocasión, **AMA** redujo la producción de NO y TNF- $\alpha$  *in vitro* y cuando se administró por vía oral (50 mg/kg, dosis única) protegió a los animales de la CU aguda inducida por DSS. Como en el caso anterior, el compuesto más activo y claro candidato a cabeza de serie fue **Ro 41-0960**, inhibiendo la producción de NO, IL-1 $\beta$  y TNF- $\alpha$  *in vitro* y protegiendo a los animales colíticos, reduciendo el índice de actividad de la enfermedad (IAE) más del 50%, lo que claramente demuestra el efecto protector del compuesto frente a la CU.

El conjunto de los resultados de los cuatro artículos, apoya la estrategia de cribado virtual basado en la TM que se ha seguido, centrándola en la identificación de moléculas inhibidoras de diferentes mediadores proinflamatorios, como COX, iNOS, IL-6, IL-1 $\beta$ , TNF- $\alpha$  y NF- $\kappa$ B, que pudiesen actuar como agentes terapéuticos frente a CU. Profundizar en las características farmacológicas de los compuestos seleccionados es un objetivo que queda fuera de la presente Tesis Doctoral, aunque los diferentes y preliminares ensayos *in vitro* e *in vivo* realizados, han corroborado la predicción hecha por TM, lo que demuestra su validez como una herramienta efectiva para el diseño y descubrimiento de nuevos compuestos cabezas de serie.



# INTRODUCCIÓN





## II. INTRODUCCIÓN

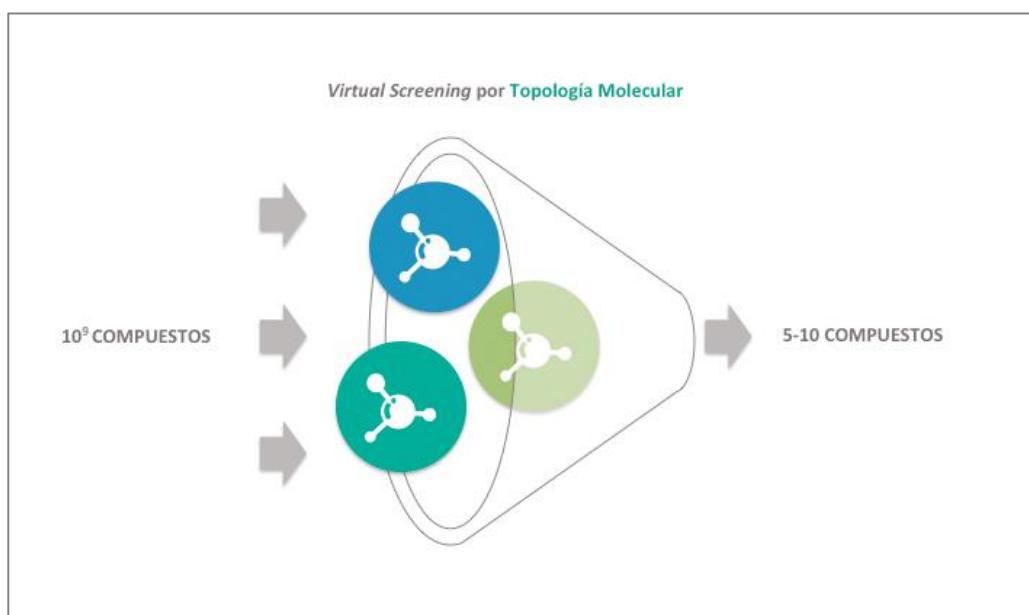
### II.1. DISEÑO Y DESCUBRIMIENTO DE FÁRMACOS ASISTIDO POR ORDENADOR

El proceso de diseño y descubrimiento de nuevos fármacos se puede dividir en distintas fases interconectadas entre sí. El primer paso consiste en seleccionar la diana o *target* terapéutico de interés, teniendo en cuenta el estado actual de la enfermedad, los posibles tratamientos, las tecnologías de las que se dispone y otros factores como el impacto social y económico (Song y col.s, 2009).

El proceso de descubrimiento de nuevos fármacos a lo largo de la historia, ha evolucionado sensiblemente en los últimos dos siglos. El siglo XX se caracterizó por el descubrimiento de nuevos principios activos por serendipia (Sneider 1985, 1990; Mann y Plummer, 1991).

Los avances en Bioquímica y Biología molecular dieron inicio a una nueva fase en el proceso de descubrimiento de fármacos, ya que se empezaron a identificar vías de señalización farmacológicas, enzimas y proteínas (receptores y ligandos) implicadas en las distintas enfermedades. El resultado fue la aparición de nuevos tratamientos que seguían vías de señalización enfocadas a actuar sobre unas determinadas proteínas diana. El desarrollo de estas disciplinas fue tan significativo y rápido que pronto se empezaron a refinar nuevas técnicas y aparecieron los primeros procesos computacionales, con la introducción de los primeros ordenadores (Rapaka y Hawks, 1993). De este modo, se entró en la fase actual de descubrimiento y diseño de fármacos, en la que las técnicas asistidas por ordenador o *Computer-Aided Drug Design (CADD)*, son el fulcro de todo proceso de búsqueda de nuevos medicamentos. El potencial de estas técnicas reside en su capacidad de hacer frente al cambio radical que el paradigma del proceso de diseño de fármacos ha sufrido en términos de competitividad y coste económico. La necesidad de desarrollar un diseño a medida de nuevas moléculas activas para las exigencias siempre crecientes de la industria farmacéutica ha acabado otorgando un papel relevante a las CADD en el ámbito del descubrimiento de nuevos *hits* y *leads* (Rapaka y Hawks, 1993).

El proceso global de desarrollo de un nuevo fármaco puede tardar alrededor de 10-15 años y suponer un coste de más de un millardo de dolares (Macalino y cols., 2015). El proceso suele constar de siete etapas principales: selección de la enfermedad, selección del *target*, identificación de nuevos *hits* y *leads* (*screening*), su optimización, ensayos pre-clínicos, ensayos clínicos y por último farmacogenómica, por lo que no es difícil entender que un retraso en una de estas fases, representaría un cuello de botella capaz de retrasar el proceso entero. Por esta razón, los procesos CADD, que permiten el rastreo de bases de datos de más de un millón de compuestos en tiempos muy breves, permitiendo la identificación y selección de unos pocos *hits* y *leads* para que se sigan estudiando en las fases siguientes, se han hecho indispensables (Figura 1) (Xu y Hagler, 2002; Prakash y Gareja, 2012).



**Figura 1.** Esquema típico de cribado virtual (*virtual screening*) por Topología molecular.

Las técnicas CADD se basan en la aplicación de sistemas computacionales altamente eficaces y complejos algoritmos matemáticos para resolver básicamente problemas de tipo biológico o fisicoquímico. Fundamentalmente se trata de una metodología *in silico* para la evaluación de un gran número de compuestos químicos y poder así identificar potenciales *hits* y *leads*.

Lógicamente, el éxito del diseño asistido por ordenador no es del 100%, y depende de la fiabilidad y robustez de los modelos predictivos que se emplean: tipo de técnica de relación cuantitativa estructura actividad o *Quantitative Structure-Activity Relationship* (QSAR), fiabilidad y reproductibilidad de los compuestos de entrenamiento empleados en la construcción de dichos modelos, tamaño molecular de los compuestos y complejidad intrínseca relacionada con la propiedad biológica bajo estudio.

Toda técnica CADD proporciona un enfoque basado en el conocimiento que puede llevar a obtener valiosa información acerca del patrón de interacciones proteína/receptor (complejo *target*/ligando), así como información sobre la afinidad del ligando por un determinado receptor.

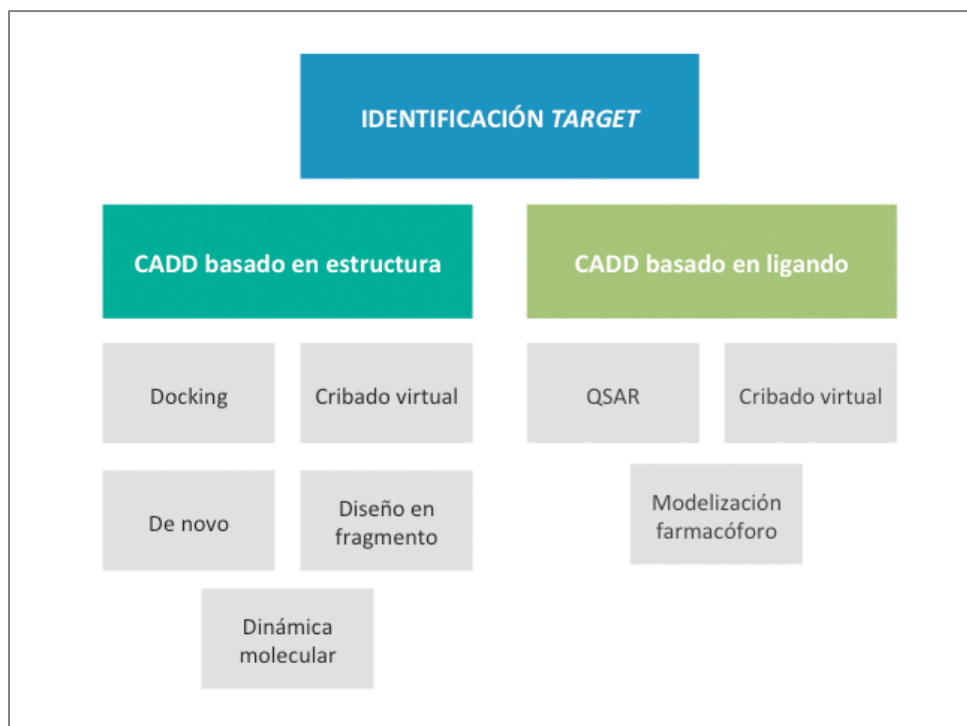
Hoy en día se pueden identificar distintas etapas específicas en el proceso de descubrimiento de un nuevo fármaco, entre las que se encuentran las siguientes (Macalino y cols., 2015):

- **Identificación del *target*** - descubrimiento y aislamiento de dianas farmacológicas o targets para investigar sus funciones y su implicación en una enfermedad determinada.
- **Validación del *target*** - etapa en la que la diana farmacológica se relaciona con la enfermedad de interés, así como con su capacidad de regular las funciones biológicas en el organismo después de la unión a una molécula (ligando).
- **Identificación de *leads*** - descubrimiento de un compuesto químico que muestre un grado de potencia y especificidad contra un target biológico y presente un perfil toxicológico y farmacocinético que le permita poder llegar a convertirse en un medicamento.
- **Optimización de *leads*** - mejora de la potencia y otras propiedades significativas a través de ciclos iterativos de evaluación del compuesto *lead* y sus análogos. Recopilación de los resultados de los ensayos *in vitro* e *in vivo* para priorizar y seleccionar a los candidatos con perfil óptimo y posibilidad de refinar los modelos matemáticos de predicción, aportando nueva información para el desarrollo de un potencial fármaco seguro y eficaz.

Dicho eso, es posible dividir las CADD en dos grandes grupos (Figura 2):

- **Diseño de fármacos basado en la estructura del receptor** - El principio del diseño, identificación y optimización de fármacos basado en la estructura del receptor consiste en identificar compuestos biológicamente activos mediante la información relacionada con la estructura 3D del receptor o *target* con el que interacciona, que normalmente es una proteína o una parte de ella (Anderson, 2003).
- **Diseño de fármacos basado en el ligando** - Se fundamenta en el conocimiento de moléculas que se unen a la diana biológica de interés. Estas técnicas pueden aplicarse tanto a casos en los que no se dispone de información relativa a la estructura 3D de la proteína diana como a casos de los que se dispone de esta información. El diseño de fármacos basado en el ligando puede aplicarse al rastreo de grandes bases de datos moleculares con el objetivo de identificar compuestos con patrones similares y que deben por tanto presentar actividades farmacológicas similares (Bacilieri y Moro, 2006).

Dentro de las técnicas CADD basadas en la estructura encontramos el *docking*, el *cribado virtual basado en la estructura del receptor*, el *diseño basado en fragmento moleculares*, el *diseño de novo* y la *dinámica molecular*. Entre las técnicas basadas en el ligando tenemos el *QSAR*, el *cribado virtual basado en ligandos* y la *modelización del farmacóforo*.



**Figura 2.** Tipos de estrategias empleadas en las técnicas computacionales aplicadas al descubrimiento de fármacos (CADD).

## II.2. QSAR

Las técnicas de modelización QSAR, del inglés *quantitative structure activity relationship*, representan uno de los principales instrumentos computacionales que se utiliza en todo proceso de diseño de fármacos. Los pioneros de esta disciplina, hace más de 50 años, fueron Corwin Hansch (Hansch y Steward, 1964; Hansch, 1971) y Toshio Fujita (Fujita y cols., 1967; Fujita, 2012). La evolución de esta metodología ha sido tan rápida y creciente, que desde un simple concepto de extensión de la fisicoquímica orgánica, ha pasado de ser una aplicación para el estudio de pequeñas series de compuestos químicos mediante métodos de regresión relativamente simples, a ser instrumento de análisis de grandes conjuntos de datos que comprenden miles de estructuras moleculares distintas, utilizando una amplia variedad de técnicas estadísticas y de métodos de *machine learning* (Cherkasov y cols, 2014).

El QSAR se suele usar para informatizar interacciones no específicas de los compuestos con propiedades físicoquímicas específicas como el punto de ebullición, el coeficiente pasivo de absorción intestinal, la permeabilidad de la barrera hematoencefálica así como otros parámetros farmacocinéticos y toxicológicos (ADME/Tox) (Young, 2009).

En la actualidad existen diferentes técnicas por las que se puede desarrollar un estudio QSAR dependiendo del tipo de tratamiento matemático que se le da a la muestra y el tipo de índices que se empleen. Entre ellas tenemos: el método tradicional, el método por redes neuronales y el tridimensional.

El QSAR tradicional incluye el tratamiento estadístico de los datos por distintas vías: análisis de regresión, análisis *cluster* y análisis de componentes principales. En ellos se valora la actividad biológica como variable dependiente del conjunto de descriptores moleculares que constituyen las variables independientes. Así se obtienen modelos que describen la actividad biológica como una determinada función matemática de los descriptores moleculares, bien sean estos estructurales o químico-físicos. Los métodos estadísticos principales y más utilizados en este ámbito son:

- Análisis lineal discriminante (ALD): permite clasificar cada compuesto de un conjunto en la categoría a la que pertenece, mediante un análisis de tipo cualitativo (Izenman, 2013).
- Análisis de regresión multilínea (ARML): identifica relaciones de tipo lineal entre un conjunto de descriptores y la propiedad de interés, permitiendo una predicción de tipo cuantitativo de dicha propiedad (Tranmer y Elliot, 2008).

En cuanto a las redes neurales artificiales (RNA), se definen dichas redes como un dispositivo de interconexión de neuronas que colaboran entre sí para producir estímulos de salida basados en distintos algoritmos matemáticos computacionales (Patel, 2013).

Finalmente, el QSAR 3D es una técnica moderna que combina los aspectos relacionados en el estudio QSAR clásico con los de los estudios SAR o de modelado molecular (Kubinyi, 1993; Garcia, 2013). En la representación SAR de los análogos activos, la elección se limita a si es activo o inactivo. Destacan entre las técnicas más comunes, el análisis comparativo de campo molecular (*Comparative Molecular Field*

*Analysis*, CoMFA) (Cramer y cols., 1988), el análisis comparativo de los índices de similitud molecular (*Comparative Molecular Similarity Index Analysis*, CoMSIA) (Klebe y Abraham, 1999) y el acoplamiento o Docking molecular (García, 2013).

En cualquier caso, el núcleo de todo modelo QSAR reside en los descriptores de la estructura química usados, que son una serie de parámetros que, a través de valores numéricos, codifican una información estructural, electrónica, estérica, así como una determinada propiedad fisicoquímica de la molécula bajo estudio. Hay distintos tipos de descriptores o índices, como por ejemplo constitucionales, topológicos, electrostáticos, geométricos o cuánticos. De acuerdo con Basak (Basak, 2011), se pueden clasificar en:

- Índices topoestructurales: codifican información sobre la conectividad de los átomos (vértices), más allá de su naturaleza química.
- Índices topoquímicos: consideran además sus características químicas (tipos de enlaces, heteroátomos, etc.).
- Índices tridimensionales: consideran parámetros geométricos.
- Índices o descriptores químico-cuánticos: consideran propiedades electrónicas.

Otra posible clasificación de los descriptores es aquella que tiene en cuenta sus niveles de complejidad dimensional (Todeschini y Consonni, 2009):

- Descriptores monodimensionales (1D): consideran tan solo la fórmula molecular.
- Descriptores bidimensionales (2D): consideran la fórmula estructural en dos dimensiones, son los más populares entre los químicos.
- Descriptores tridimensionales (3D): tienen en cuenta la conformación espacial.
- Descriptores tetradimensionales (4D): consideran la mutua orientación y las dinámicas tiempo-dependientes de las moléculas (Polanski y cols., 2006).

Una vez seleccionado el tipo de índices, la metodología QSAR ofrece un abanico muy amplio de técnicas quimiométricas y métodos estadísticos (Verma y cols., 2010) que se pueden aplicar para el diseño e identificación de nuevos compuestos *hits* y *leads*.

El proceso de construcción de un modelo QSAR se puede resumir en las siguientes etapas:

- Recogida de información sobre los ligandos y sus actividades biológicas o propiedades físicoquímicas objeto de estudio.
- Cálculo de los descriptores moleculares. Estos descriptores se relacionan con la actividad de los ligandos por medio de distintos métodos matemático-estadísticos de modelado para obtener los modelos de predicción QSAR.
- Finalmente, los modelos obtenidos son sujetos a procesos de validación interna y externa. El objeto principal de estas validaciones es asegurar, en lo posible, que los modelos sean de aplicación en la búsqueda de nuevas moléculas activas.



## II.3. TOPOLOGÍA MOLECULAR

### II.3.1. Teoría de Grafos

#### II.3.1.1. Principios

La topología surge como disciplina con entidad propia en el siglo XVII con Leibniz, el cual se dio cuenta de que determinados problemas geométricos dependían solo de la posición y no de métrica alguna, por lo que acuñó el término *geometría de la posición* o *análisis situs*, que equivale a lo que hoy llamamos topología (Mates, 1989; De Risi, 2007). Un ejemplo típico sería el camino para ir de un punto a otro de un conjunto de puntos que presentan ciertas conexiones entre sí.

Es importante señalar que cualquier objeto que se deforme, siempre que no se rompa, dará lugar a otro que es topológicamente equivalente. Por ejemplo un círculo y un cuadrado son dos figuras equivalentes ya que son superficies cerradas. Lo mismo podemos decir de un segmento y un arco de curva, que son figuras abiertas.

Cuando se aplica la topología a las moléculas, lo importante es como se conectan los átomos entre sí, si hay más de un camino para ir de un átomo a otro pasando por los enlaces que los interconectan, el número de enlaces entre átomos por el camino más corto, etc. Es decir, características geométricas fundamentales en química convencional, tales como distancias interatómicas, ángulos de enlace o diedros, estructura tridimensional o aspectos dinámicos, son aquí irrelevantes. En definitiva, se puede decir que la Topología molecular estudia las posiciones e interconexiones de los átomos en una molécula.

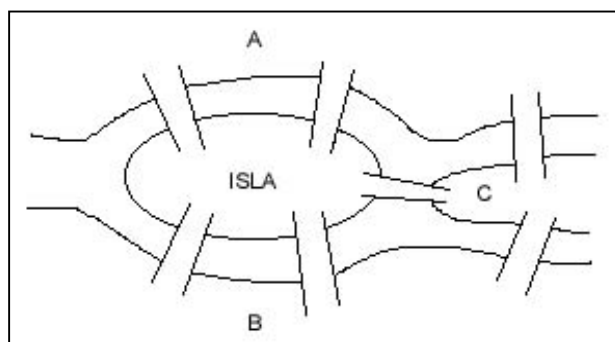
Aunque en el concepto más amplio, la topología corre pareja a otras áreas de la matemática como el álgebra, el cálculo y el análisis matemático, en la acepción que nos interesa aquí, incluye tres subdisciplinas:

- La teoría de grafos
- La teoría de nudos
- La teoría de superficies

La disciplina en la que se sustenta la presente Tesis Doctoral es la Teoría de grafos, basada fundamentalmente en la asimilación entre moléculas y grafos. Se puede definir como una parte de la Química matemática capaz de representar estructuras moleculares a través de índices grafo-teóricos (García-Doménech y cols., 2008; Gálvez y cols., 2012). Un grafo es una estructura abstracta que contiene nodos conectados con aristas (o arcos). En inglés: *nodes*, *edges*. Todas las moléculas pueden representarse como grafos. Para ello simplemente se representan a los átomos como nodos (puntos) y a los enlaces como aristas. Además, es posible asociar determinadas propiedades a cada nodo (como por ejemplo el número atómico), y a cada arista (como por ejemplo la multiplicidad y/o tipo de enlace). De este modo, las ventajas de la TM se pueden resumir de la siguiente manera (Zanni y cols., 2015):

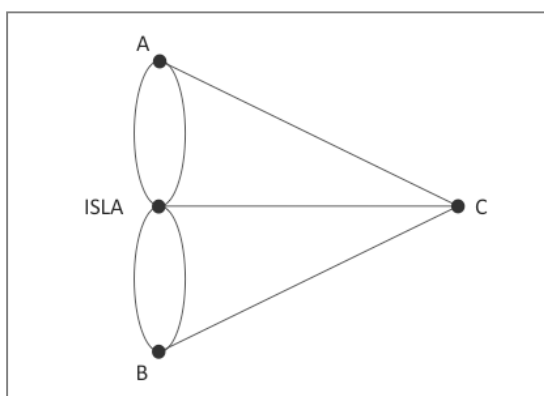
- La estructura molecular se representa en términos puramente matemáticos, es decir, como un conjunto de elementos (átomos) que interactúan por medio de segmentos (aristas).
- Permite el análisis rápido y preciso de un gran número de compuestos, así como el diseño de nuevos compuestos por proceso inverso (propiedad  $\rightarrow$  estructura).
- Es muy fácil de transformar en lenguaje informático (computación).

El primero en usar el concepto de grafo (aunque no el término, que es posterior), fue Euler (Pappas, 1989), a través del famoso problema de los puentes de Königsberg. Esa ciudad era la capital de la Prusia Oriental y hoy en día se llama Kaliningrado y es parte integrante de Rusia. La ciudad tenía a comienzos del siglo XVIII siete puentes que conectaban cuatro porciones distintas de tierra, incluyendo una isla sobre el río Pregel (ver Figura 3). Euler se planteó si era posible pasar por las cuatro porciones atravesando todo los puentes pero sin pasar dos veces por el mismo.



**Figura 3.** Esquema de los puentes de la ciudad de Königsberg (Prusia Oriental) en 1700.

Euler se percató de que para este problema era irrelevante la distancia o el tamaño de los puentes, sino que cada porción de tierra se podía representar por un punto y cada puente por un segmento que une dos de esos puntos. Había introducido así el concepto de grafo, como un conjunto de puntos (denominados vértices o nodos) unidos por segmentos (llamados aristas o arcos). También se percató de que era imposible hacer el recorrido completo sin pasar dos veces, al menos, por algún puente. Así, Euler constató que lo único importante era la secuencia de puentes (Alexanderson, 2006).



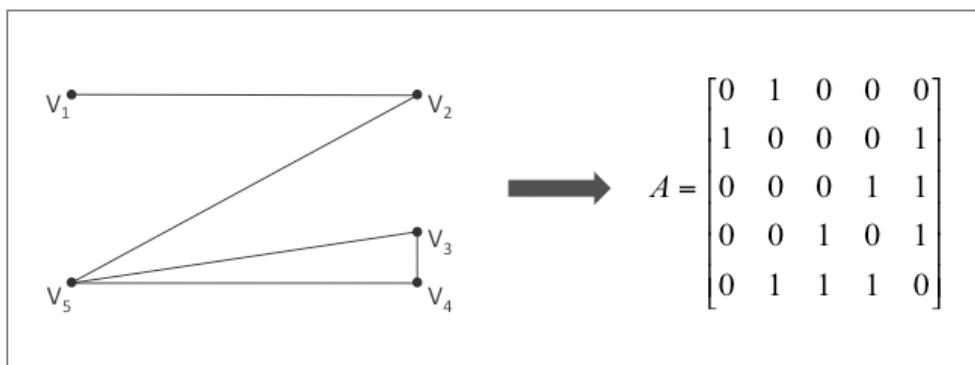
**Figura 4.** Grafo de Euler basado en el esquema de los puentes de la ciudad de Königsberg (Prusia Oriental) en 1700.

En la Figura 4 se aprecian los vértices A, B, C e ISLA y las aristas que los unen (AC, BC, etc.).

El posterior desarrollo de la teoría de grafos se debe a matemáticos como A. Cayley y J. J. Sylvester durante el siglo XIX (Cayley, 1878; Sylvester, 1878). Ya en el siglo XX, se ha usado en áreas diversas en las que la conectividad entre distintos elementos desempeña un papel esencial, como por ejemplo el diseño de redes de comunicación y de transporte (Cover, 1998), la construcción de circuitos eléctricos (por ejemplo, en los ordenadores) (Deo, 2004), la sincronización de osciladores con distintas topologías (Ren, 2008), o el análisis de redes sociales (Wasserman y Faust, 1994), por citar solo algunos ejemplos representativos.

Una de las ventajas principales de los grafos, es que pueden representarse siempre por una matriz. La más sencilla es la llamada matriz de adyacencia o matriz topológica (A).

En la Figura 5 se muestra un ejemplo de grafo, formado por cinco vértices ( $V_1$ ,  $V_2$ ,  $V_3$ ,  $V_4$ ,  $V_5$ ) y cinco aristas, a saber: ( $\{V_1, V_2\}$ ,  $\{V_2, V_5\}$ ,  $\{V_3, V_4\}$ ,  $\{V_3, V_5\}$ ,  $\{V_4, V_5\}$ ). Cada segmento o arista representa un vínculo directo entre dos nodos de la red. A su derecha vemos la matriz de adyacencia ( $A$ ) asociada al grafo.



**Figura 5.** Grafo simple con su correspondiente matriz de adyacencia ( $A$ )

Como se aprecia, los elementos  $a_{ij}$  de esta matriz adoptan valor 0 si no existe enlace (arista) entre el vértice  $i$  y el  $j$  y valor 1 si existe enlace entre ambos. Como cada elemento se considera no enlazado a sí mismo, los elementos de la diagonal principal, que es la que recorre la matriz de arriba a la izquierda hasta abajo a la derecha, tendrán valor 0.

Por otra parte, la matriz de adyacencia es siempre cuadrada, es decir, tiene el mismo número de filas que de columnas,  $n$ , que se llama *rango* de la matriz y que coincide con el número de vértices del grafo ( $n$ ). Además la matriz es simétrica, es decir, los elementos posicionados simétricamente respecto a la diagonal principal son iguales. Dicho de otro modo, la transpuesta (la resultante de cambiar filas por columnas) de la matriz de adyacencia es igual a sí misma. Otra propiedad importante de la matriz topológica es que pueden numerarse los vértices en cualquier orden, sin que cambien sus propiedades.

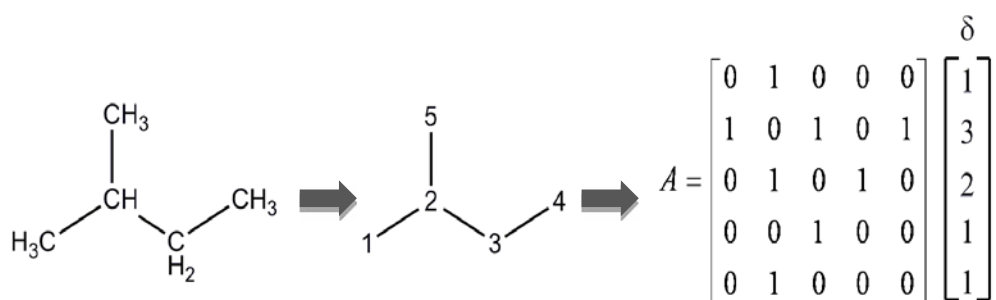
Las matrices, como objetos matemáticos, son muy versátiles ya que permiten una amplia variedad de operaciones a partir de las cuales generamos descriptores matemáticos específicos de la matriz y por tanto, del grafo asociado. Además de transponerse, las matrices pueden sumarse, restarse, multiplicarse, dividirse, generar sus autovalores o valores propios, etc. Además de la matriz de adyacencia, hay otras

muchas matrices que son de interés en la teoría de grafos, pero que quedan fuera del alcance de esta investigación.

### II.3.1.2. Moléculas representadas como grafos

El primero en proponer que las moléculas podían ser representadas por grafos fue el matemático inglés J.J. Sylvester (Sylvester, 1878). En la forma más simple de representación, los átomos de hidrógeno se suprimen, mientras que el resto de los átomos se representan como vértices y los enlaces como aristas. A continuación, se numeran los vértices en orden aleatorio y se ordena la matriz de la forma indicada anteriormente, es decir, asignando valor 1 al elemento  $ij$  si existe enlace y 0 en caso contrario.

Veamos, como ejemplo, la representación del isopentano en la Figura 6:



**Figura 6.** Ejemplo de transformación del isopentano en un grafo y posteriormente, en una matriz de adyacencia.

Se llama valencia de cada vértice al número de aristas que convergen en él. Es fácil apreciar que dicho número coincide con la suma de los elementos situados sobre la correspondiente fila (o columna) en la que se sitúa el vértice. Así, a la derecha de la matriz, vemos un vector columna formado por las valencias de cada vértice, de modo que las valencias de los vértices etiquetados como 1, 2, 3, 4 y 5 son 1, 3, 2, 1 y 1 respectivamente (Gálvez y cols., 2012).

### II.3.2. Índices topológicos

Una vez representada la molécula por un grafo, surge la pregunta de como asignar a ese grafo un número o descriptor que lo caracterice y por ende, a la propia molécula. Estos números reciben el nombre de índices topológicos. La idea es que se puedan establecer relaciones cuantitativas entre la estructura molecular caracterizada por ese número (o números) y cualquier propiedad experimental del compuesto químico que tiene esa estructura. Estas relaciones o ecuaciones, que se indicaron anteriormente, se conocen como QSAR o más genéricamente QSPR (relaciones cuantitativas estructura-propiedad), nos permiten predecir las propiedades o actividades de los compuestos químicos analizados.

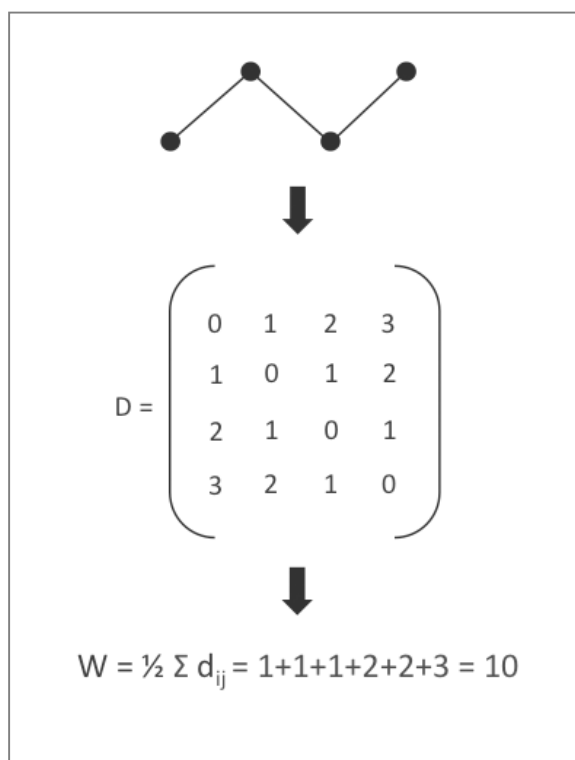
Los índices topológicos deben reunir una serie de requisitos, que han sido recogidos por Randić en un conocido decálogo (Randić, 1992). Entre esos requisitos se encuentran:

- Deben ser únicos para cada molécula/grafos y no deben depender del orden de numeración del grafo (invariantes de grafo).
- Deben ser de cálculo simple y fácilmente computarizable.
- Deben tener capacidad predictiva, particularmente para ciertas propiedades moleculares.

El primer intento de asignar un índice topológico a un grafo consistió en calcular los autovalores de la matriz de adyacencia. Los autovalores son las soluciones al polinomio característico de la matriz. No obstante, enseguida se vio que el criterio no era adecuado porque grafos diferentes presentaban los mismos autovalores. Así, Schwenk comprobó que aproximadamente uno cada tres grafos presentaban los mismos autovalores (Bogner y cols., 2006). Este hecho condujo a la búsqueda de otras vías, de las que se tratarán brevemente a continuación.

El primer índice topológico fue propuesto por Wiener en 1947 (Wiener, 1947). Se basaba en el concepto de distancia topológica, entendida como el número de aristas a recorrer entre dos vértices dados por el camino más corto. El índice de Wiener (W), se definió como la suma de todas las distancias topológicas entre cada par de vértice del grafo. Así, para el n-butano W vale 10, tal y como se ilustra en la Figura 7 (Bonchev, 2001). El índice de Wiener demostró una sorprendente eficacia para predecir las

temperaturas de ebullición de alcanos, una propiedad nada fácil de modelar incluso con métodos convencionales mucho más potentes.



**Figura 7.** Cálculo del índice de Wiener para la molécula del n-butano.

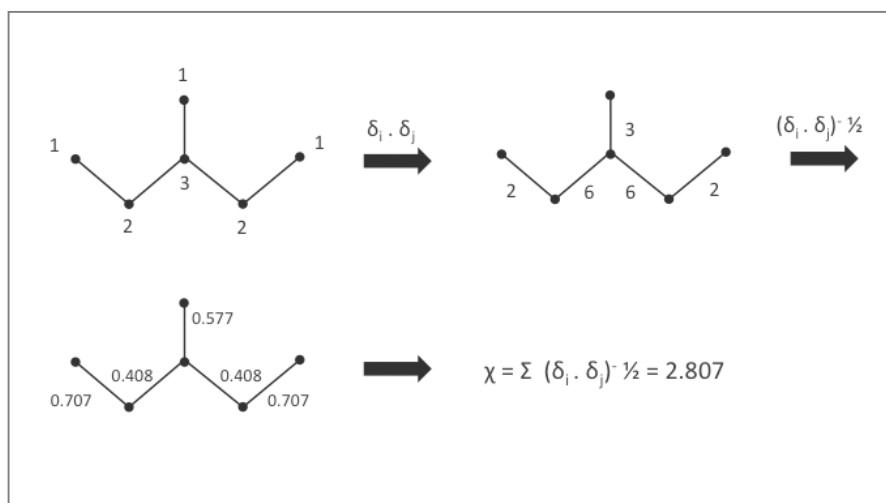
En las décadas de los 60 y 70 del pasado siglo se introdujeron otros índices como por ejemplo el de Hosoya (Hosoya, 1971), que representa la suma de los números de no-adyacencia de un grafo. Es decir una medida del número total de aristas que podemos introducir en un grafo de modo que ningún par de ellas sea adyacente.

Probablemente, el índice que ha generado las mejores predicciones en un mayor número de propiedades y por ello posiblemente el más representativo de todos, es el índice de ramificación de Randić ( $\chi$ ). Se define como la suma, para todas las aristas del grafo, de los inversos de las raíces cuadradas de los productos de las valencias de los pares de vértices que delimitan cada una de las aristas del grafo. O sea:

$$\chi = \sum_i \sum_j (\delta_i \delta_j)^{-1/2}$$

Dada la relevancia de este índice, a continuación se describirá su cálculo para el grafo del 3-metil-pentano (Figura 8). Las etapas para el cálculo son:

- Anotar en cada vértice el valor de su valencia topológica.
- Colocar sobre cada arista los productos de las valencias de los dos vértices que la limitan.
- Obtener las inversas de las raíces cuadradas de cada uno de esos productos.
- Sumar dichas inversas.



**Figura 8.** Cálculo del índice de Randić para la molécula del 3-metil-pentano.

El índice de Randić ha demostrado su eficacia tanto para propiedades fisicoquímicas (densidad, solubilidad, entalpía de vaporización, etc.), como biológicas (potencia anestésica, actividad antibacteriana, inhibición enzimática, etc.) (Kier y Hall, 1986).

No obstante, el índice de Randić presenta dos notables limitaciones:

- La degeneración, es decir, la existencia de valores iguales para grafos diferentes.
- La no consideración de enlaces múltiples y heteroátomos.



En realidad, esta última carencia es propia de todos los índices topológicos en la medida en que se calculan sobre grafos simples (es decir aquellos sin vértices ponderados y sin aristas múltiples), de modo que los índices que tienen en cuenta esas características químicas de las moléculas se conocen como *índices topoquímicos*. Cuando una molécula posee enlaces múltiples (cosa que como se sabe ocurre en la mayoría de interés biológico), el modo más sencillo de resolver el problema es asignar a los términos de la matriz topológica un número igual a la multiplicidad del enlace entre el vértice  $i$  y el  $j$ , en lugar de poner 0 y 1. Por ejemplo si entre el vértice  $i$  y el  $j$  hay un doble enlace, en la matriz pondremos un 2 y si hay un triple enlace pondremos un 3. Por lo que se refiere a los heteroátomos, se asigna un “peso” (un valor numérico dado para cada átomo) en la posición de la diagonal principal correspondiente al heteroátomo. Uno de los pesos habitualmente utilizados para caracterizar la valencia de heteroátomos es:

$$\delta^v = \frac{Z^v - h}{Z - Z^v - 1}$$

Donde  $Z$  es el número atómico del heteroátomo.  $Z^v$  es el número de electrones de valencia y  $h$  el número de hidrógenos a los que se une el heteroátomo. De esta manera se puede ya caracterizar a cada elemento del sistema periódico.

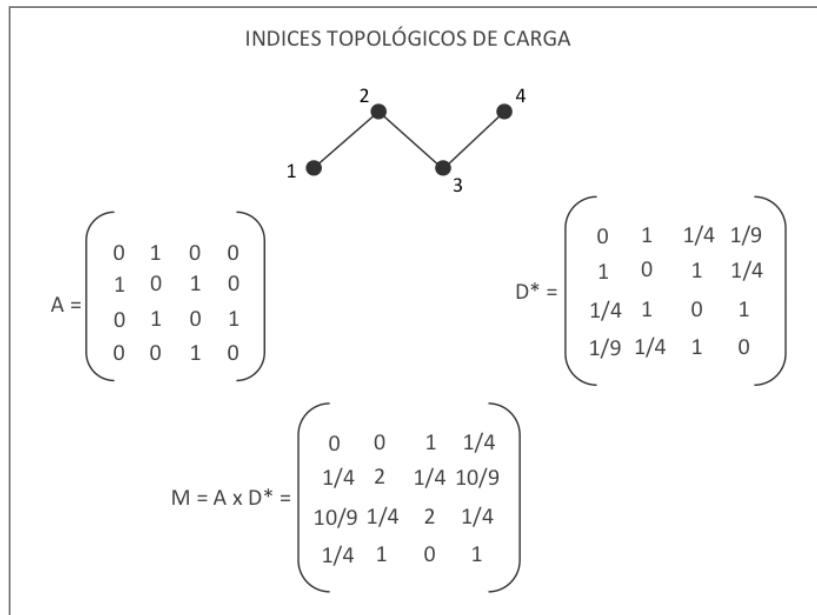
Posteriormente, Kier y Hall (Kier y Hall, 1977; Kier, 2012) generalizaron el índice de Randić convirtiéndolo en una familia de nuevos índices, llamados índices de conectividad molecular, dando así un paso fundamental en su capacidad predictiva. En esencia, los índices de conectividad se calculan igual que el índice de Randić, pero para cada uno de los subgrafos del grafo en cuestión. Por subgrafo se entiende cualquier parte de un grafo constituida por vértices aislados (orden 0) o aristas aisladas (orden 1) o  $n$  asociadas entre sí (orden  $n$ , donde  $n=2,3,\dots$ ). La Figura 9 muestra los valores de los índices de conectividad. Es fácil ver que el índice de conectividad de orden 1 coincide con el índice de Randić.

$$\begin{aligned}
m\chi_t = \sum_{j=1}^{n_m} m\delta_j &\rightarrow {}^0\chi = \sum_{i=1}^n (\delta_i)^{-1/2} \\
&\rightarrow {}^1\chi = \sum_{s=1}^{n_m} (\delta_i \cdot \delta_j)_s^{-1/2} \\
&\rightarrow {}^2\chi = \sum_{s=1}^{n_m} \left[ \delta_i \cdot \delta_j \cdot \delta_k \right]_s^{-1/2}
\end{aligned}$$

**Figura 9.** Índices de conectividad de Kier y Hall.

Por último, se mencionarán los índices topológicos de carga (ITC), introducidos por nuestro grupo (Gálvez y cols., 1994) que se definen como la suma en valor absoluto para subgrafos de un determinado orden, de los términos de carga. Entendiéndose por tales, las diferencias entre los elementos simétricos de la matriz producto entre la matriz topológica y la matriz de inversos de los cuadrados de las distancias topológicas, en la que los términos de la diagonal se dejan inalterados para evitar división por cero. En la matriz producto, denominada M ( $M = A \times D^*$ ), los términos de la diagonal principal coinciden con las valencias de los vértices. Se trata de una matriz asimétrica, M, cuyas diferencias entre los términos recíprocos ( $m_{ij}-m_{ji}$ ), reciben el nombre de *términos de carga*, porque representan una medida de la carga eléctrica neta transferida entre el átomo “i” y el “j”. La suma, en valor absoluto, de los términos para átomos ubicados a una cierta distancia topológica, constituye los llamados índices topológicos de carga,  $G_k$ . El valor del índice ponderado por enlace,  $J_k$  es  $G_k/N-1$ . Donde N es el número de vértices del grafo.

La Figura 10 ilustra el procedimiento para cálculo de los índices de carga  $G_k$  y el  $J_k$  (valor de  $G_k$  ponderado por arista o enlace).



**Figura 10.** Obtención de la matriz M, a partir de la cual se generan los ITC.

### II.3.3. Diseño de fármacos por Topología molecular

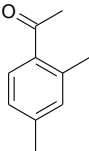
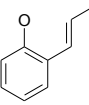
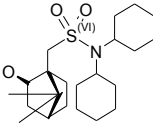
Aplicando la metodología QSAR denominada Topología molecular anteriormente descrita, nuestro grupo ha identificado numerosos y novedosos compuestos con actividad farmacológica *in vitro* (*hits*) e *in vivo* (*leads*). En la Tabla 1 se reflejan algunos de los campos en los que nuestro grupo ha descubierto nuevos compuestos activos usando la TM.

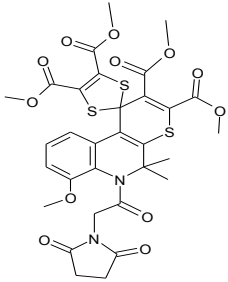
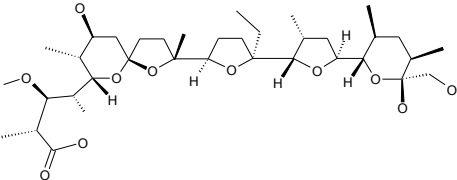
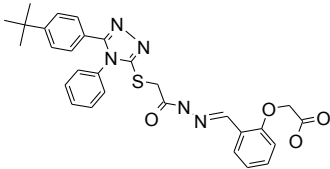
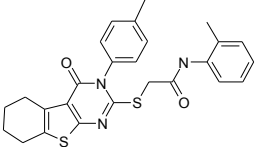
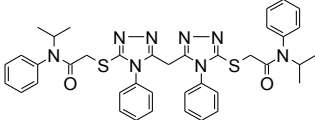
**Tabla 1.** Principales areas en las que nuestro grupo ha descubierto nuevos compuestos 'cabeza de serie' por TM.

Areas farmacológicas
Analgésicos (Gálvez y cols., 1994)
AntiAlzheimer (Gálvez y cols., 2010)
Antibacterianos (Mishra y cols., 2001)
Antidiabéticos (Anton-Fos y cols., 1994)
Antihistamínicos (Casaban-Ros y cols., 1999)
Antimaláricos (Mahmoudi y cols., 2006)
Antineoplásicos (Gálvez y cols., 1996)
Antivíricos (de Julian-Ortiz y cols., 1999)
Broncodilatadores (Rios-Santamarina y cols., 2004)
Hipolipemiantes (Cercos-Del Pozo y cols., 1996)

Una de las contribuciones más relevantes de nuestro grupo en la selección de nuevos compuestos cabezas de serie, ha consistido en el empleo de los índices topológicos en sentido inverso al convencional, es decir, en lugar de predecir propiedades de moléculas ya conocidas, tratar de generar nuevos compuestos a partir de propiedades predeterminadas (García-Domenech y cols., 1996). Esto se ha logrado desarrollando algoritmos propios y originales de semejanza topológica capaces de reproducir un patrón matemático asociado a cada molécula/s, a modo de “huella dactilar”, permitiendo la búsqueda de nuevos compuestos activos con mejores perfiles farmacológicos, toxicológicos y/o farmacocinéticos. Gracias a este novedoso enfoque, la “Unidad de Diseño de Fármacos y Topología Molecular” ha patentado distintas moléculas en campos como las enfermedades neurodegenerativas (Alzheimer), parasitarias (malaria), cáncer, inflamación y dolor (ver Tabla 2).

**Tabla 2.** Compuestos patentados por la “Unidad de Diseño de Fármacos y Topología Molecular”.

Compuesto	Actividad farmacológica
	Actividad analgésica y antiinflamatoria (Gálvez y cols., 1991) Patente Nacional (1991)
	Actividad analgésica y antiinflamatoria (Gálvez y cols., 1991) Patente Nacional (1991)
	Antineoplásico (Gálvez y cols., 2004) Patente Internacional (2004)

	<p>Antineoplásico (Llompart y cols., 2006) Patente Internacional (2006)</p>
	<p>Infecciones asociadas a Plasmodium (Mazier y cols., 2009) Patente Internacional (2009)</p>
	<p>Anti Alzheimer (Gálvez y cols., 2010) Patente Internacional (2010)</p>
	<p>Anti Alzheimer (Gálvez y cols., 2010) Patente Internacional (2010)</p>
	<p>Anti Alzheimer (Gálvez y cols., 2010) Patente Internacional (2010)</p>

En la presente Tesis Doctoral, se ha seguido la metodología QSAR basada en la TM como herramienta para la búsqueda *in silico* de nuevas moléculas con carácter antiinflamatorio (Capítulo 1) que, tras un posterior refinamiento, permitiera la selección de nuevas moléculas protectoras frente a colitis ulcerosa (CU) (Capítulo 2).

## II.4. INFLAMACIÓN

Normalmente, la inflamación es básicamente una actividad beneficiosa que rejuvenece el tejido lesionado y elimina a los agentes extraños que perturban la homeostasis, lo que conduce a la recuperación del estado de equilibrio. Esta tarea se logra a través de una compleja respuesta inflamatoria que implica un perfecto equilibrio entre un gran grupo de moléculas bioactivas, pro y antiinflamatorias, como la interleucina (IL)-6, el factor nuclear kappa B (NF-kB), el factor de necrosis tumoral (TNF)- $\alpha$  y el factor de crecimiento transformante (TGF)- $\beta$ , entre otros, producidas a partir de la infiltración de células inflamatorias constitucionales o de infiltración. Sin embargo, una persistente o inadecuada inflamación crónica no resuelta, inclina la balanza a favor de los agentes proinflamatorios, puede aumentar el riesgo de varias patologías (Jurjus y cols., 2016; Ellinghaus y cols., 2016). Los procesos inflamatorios se suelen dividir en agudos y crónicos. Las reacciones inflamatorias agudas suelen ser autolimitadas y se resuelven rápidamente, debido a la implicación de los mecanismos de retroalimentación negativa. Si bien las respuestas inflamatorias reguladas se pueden considerar esenciales para mantener la homeostasis, las respuestas inflamatorias que no logran autorregularse pueden llegar a cronicarse y contribuir a la perpetuación y progresión de la enfermedad. Las características típicas de las respuestas inflamatorias crónicas incluyen la pérdida de la función de la barrera, la capacidad de respuesta a un estímulo normalmente benigno, la infiltración de células inflamatorias en los compartimientos en los que no se encuentran normalmente en cantidad tan elevada y la sobreproducción de oxidantes, citocinas, quimiocinas, eicosanoides y metaloproteinasas de la matriz. Los niveles de estos mediadores amplifican la respuesta inflamatoria, son destructivos y contribuyen a los síntomas clínicos (Calder y cols., 2009).

### II.4.1. Enfermedades inflamatorias

Los procesos inflamatorios crónicos pueden derivar en condiciones patológicas graves, como la enfermedad inflamatoria intestinal (EII), la artritis reumatoide (AR), la psoriasis, la obesidad y la espondilitis anquilosante, además de algunas enfermedades neurodegenerativas como la enfermedad de Alzheimer (Figura 11).



**Figura 11.** Enfermedades asociadas a procesos inflamatorios crónicos.

La EII es una familia de enfermedades idiopáticas, crónicas, inflamatorias que afectan al colon y otras partes del tracto gastrointestinal. Su etiología no está perfectamente identificada, pero por lo general se asocia a un desequilibrio de la homeostasis luminal de la mucosa en personas genéticamente susceptibles. Lo que deriva en una respuesta inflamatoria inapropiada y prolongada, probablemente causada por una disfunción de la integridad microfloral y diversos factores ambientales. Las EII abarcan una vasta gama de patologías intestinales, aunque destacan dos enfermedades: colitis ulcerosa (CU) y enfermedad de Crohn (EC). Ambas muestran características etiopatológicas y un perfil sintomático similar que pueden incluir: diarrea, sangrado rectal, dolor abdominal y pérdida de peso. Diferenciándose fundamentalmente en la porción del aparato digestivo implicada (Jurjus y cols., 2016).

La psoriasis y la artritis psoriásica son enfermedades autoinmunes de la piel y otros tejidos, que presentan graves procesos inflamatorios asociados con el desarrollo de comorbilidades, que tienen un impacto significativo tanto en la salud como en la calidad de vida del paciente (Reich, 2012). Por otro lado, la artritis reumatoide se caracteriza por la inflamación crónica de las articulaciones y de los huesos, pudiendo causar la



destrucción articular permanente, junto con graves deformidades. Además, se ha demostrado que la formación de las placas ateroscleróticas puede ser exacerbada y acelerada por la inflamación sistémica, en enfermedades como la AR (Yang y cols., 2016). En el ámbito de las enfermedades neurodegenerativas, se ha visto como la inflamación puede jugar un papel fundamental en el desarrollo de algunas patologías. En pacientes con Alzheimer se han detectado en las placas seniles neuríticas, microglía activada, factores del complemento e interleucinas que pueden interactuar con el péptido  $\beta$ A4 y provocar la acumulación de  $\beta$ -amiloide. Además, se ha demostrado un calo de incidencia de dicha enfermedad en los pacientes que toman antiinflamatorios no esteroideos (AINE) (Imbimbo, 2009). En última instancia, cabe resaltar el papel de la inflamación en la obesidad traduciéndose en un desequilibrio energético crónico, asociado con la inflamación crónica de bajo grado dentro de los tejidos metabólicos, a menudo referido como inflamación metabólica o "meta-inflamación" (Lyons y cols., 2016).

## II.4.2. Principales mediadores inflamatorios

Todo proceso inflamatorio, tanto agudo como crónico, está caracterizado por la presencia de mediadores proinflamatorios y otros reguladores (ver Figura 12).



**Figura 12.** Principales mediadores inflamatorios.

El perfil de las citocinas que rigen la inflamación de la EII, difiere en función de la enfermedad de Crohn o de la CU. Predominando en la primera una respuesta de tipo Th17 y en la segunda una respuesta inmune adaptativa de tipo Th2. Por lo que en la EII podemos encontrar un gran pool de citocinas: IL-17, IFN- $\gamma$ , TNF- $\alpha$ , IL-6, IL-1, TGF, IL-22, IL-4, IL-5 e IL-13 entre otras; estas provocan una reacción inflamatoria persistente y mantenida (Jurjus y cols., 2016). Siempre en ámbito de la EII, encontramos una serie de mediadores proinflamatorios muy implicados, tanto a nivel local como sistémico, tales como: la ciclooxigenasa 2 (COX-2), la prostaglandina E2 (PGE<sub>2</sub>), la óxido nítrico sintasa inducible (iNOS) entre otros (Jurjus y cols., 2016).

Junto a las citocinas, se hallan diferentes quimiocinas como CCL y CXCL que pueden atraer monocitos y células dendríticas y que son fundamentales en los procesos inflamatorios de la EII (Fina y Pallone, 2008). Por último, la activación de factores de transcripción, principalmente, NF-kB (Borm y cols., 2005) y el transductor de señal y activador de la transcripción 3 (STAT3) (Sugimoto, 2008) y la subunidad alfa del factor 1 inducible por hipoxia (HIF1 $\alpha$ ) (Flück y cols., 2016).

En el caso de la inflamación asociada a la psoriasis y la AR, los mediadores principales son las interleucinas (IL-12 y IL-23) y las integrinas ( $\alpha$  y  $\beta$ ) (Mease, 2012; Meyer y Levine, 2014) así como la vía de señalización Janus quinasas (JAK)/STAT que regula la expresión de diferentes genes que a su vez regulan la expresión de citocinas y receptores como IL-23R, IL-10, IFN- $\gamma$  o IL-12B (Zundler y Neurath, 2016).

Una citocina que adquiere un papel clave en la resistencia a la insulina asociada a obesidad es el TNF (Sanz y Moya-Pérez, 2014; Khodabandehloo y cols., 2016). Además, se ha descubierto que en la obesidad algunas vías para la detección de patógenos comparten mecanismos comunes de señalización dentro de la célula con las vías de los nutrientes, como por ejemplo los receptores Toll-like 2 y 4 (TLR2/4) (Clarke y cols., 2012), que son los receptores de reconocimiento de los patógenos de la superficie celular (PPR) como por ejemplo el lipopolisacárido (LPS) que activan al NF- $\kappa$ B y la transcripción de este al núcleo, provocando la secreción de citocinas proinflamatorias (Hulsmans y cols., 2012).

## II.5. ENFERMEDAD INFLAMATORIA INTESTINAL

Hoy en día, la enfermedad inflamatoria intestinal juega un papel principal en el área de las enfermedades inflamatorias crónicas, solo superada por la artritis reumatoide. Esto se debe a que cada año el número de pacientes afectados por EII está aumentando, afectando a millones de personas en todo el mundo (Qin, 2012).

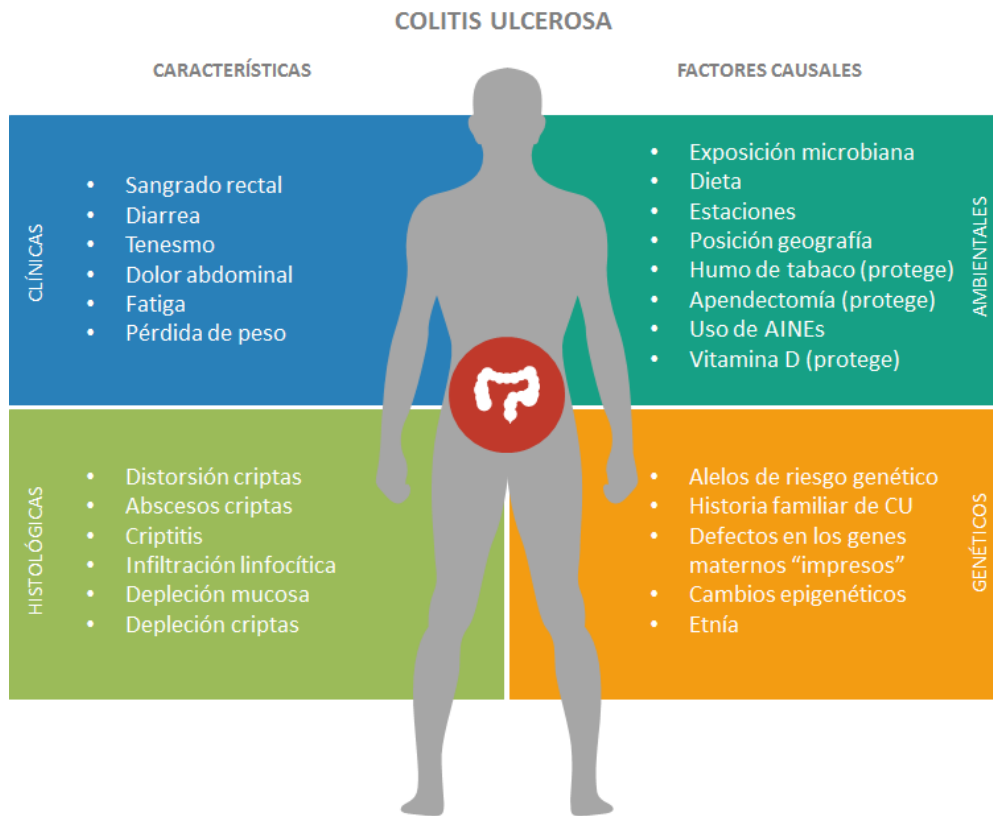
Dentro del término enfermedad inflamatoria intestinal se engloban dos afecciones fundamentalmente: la enfermedad de Crohn y la colitis ulcerosa. Estas enfermedades se diferencian fundamentalmente en la zona del aparato digestivo en la que se desarrolla el fenómeno inflamatorio; afectando en el caso de la CU únicamente al intestino grueso o colon mientras que en la EC puede verse involucrada cualquier zona del aparato digestivo entre la boca y el ano. El diagnóstico de la EII normalmente se determina mediante sigmoidoscopia flexible y colonoscopia (Dignass y cols., 2012).

### **II.5.1. Colitis ulcerosa**

La CU, objeto de esta tesis, se define como un proceso inflamatorio crónico que causa inflamación continua de la mucosa del colon sin granulomas, afectando al recto e intestino grueso y caracterizada por una sintomatología intermitente alternándose periodos de mayor actividad o brotes con periodos de inactividad o de remisión de la enfermedad (Dignass y cols., 2012).

Los síntomas asociados a esta enfermedad son sangrado rectal, diarrea, dolor abdominal, tenesmo, fiebre, fatiga y pérdida de peso (Bodger y cols., 2014). Además pueden aparecer síntomas extra-intestinales de carácter dermatológico, reumatológico, oftalmológico o hepático (<http://www.ua-cc.org/>).

A nivel histológico, la CU se caracteriza por presentar daño a nivel mucoso y submucoso del colon fundamentalmente en las criptas intestinales (Figura 13).



**Figura 13.** Las características clínicas e histológicas de la CU y los factores causales que influyen en el riesgo de esta enfermedad (Low y cols., 2013).

La colitis ulcerosa afecta fundamentalmente a adultos jóvenes (entre la adolescencia tardía y la edad adulta temprana), no obstante puede ser diagnosticada a cualquier edad. Sin embargo, no existen diferencias entre sexos en cuanto al riesgo de padecer CU (Dignass y cols., 2012).

#### II.5.1.1. Factores implicados en la CU

A pesar de que la CU es una enfermedad idiopática, la etiología de la enfermedad parece responder a factores genéticos y ambientales fundamentalmente, que alteran o rompen la barrera de la mucosa intestinal (Figura 13) (Roda y cols., 2011; Low y cols., 2013).

Los aspectos genéticos ligados a la enfermedad están ampliamente demostrados en la actualidad, ya que a pesar de no ser una enfermedad hereditaria los pacientes con CU parecen estar genéticamente más predispuestos a sufrirla. De hecho, factores como la étnia, la herencia familiar (especialmente ligada a la herencia materna) pueden

aumentar la incidencia de padecer esta enfermedad (Brophy y cols., 2003; Ellinghaus y cols., 2015).

Los factores ambientales y el estilo de vida también influyen en la incidencia de la enfermedad. Así se conoce que la baja exposición en edades tempranas a microbios puede favorecer el desarrollo de CU en etapas posteriores de la vida. Esto guarda una estrecha relación con la teoría de que las mejoras de las condiciones higiénicas en los países desarrollados pueden conducir a la inmadurez del sistema inmune de un individuo durante su niñez y, posteriormente, aumentar la susceptibilidad a la CU en la edad adulta. Además el abuso de determinados medicamentos como la *isotretinoína* (Crockett y cols., 2010), AINE (Klein y cols., 2010; Long y cols., 2016), anticonceptivos orales, terapias hormonales (Khalili, 2015) y antibióticos (Lucke y cols., 2006) también parecen guardar relación con el desencadenamiento de la CU. Otro aspecto relacionado con el estilo de vida y que podría ser un agente desencadenante de la enfermedad es adoptar dietas pobres en fibra y altas en ingestas masiva de grasas poli-saturadas y carne roja (Molodecky y Kaplan, 2010). Además infecciones provocadas por bacterias como la *Salmonella*, *Shigella* y *Campylobacter* podrían generar un desorden en la microbiota intestinal, iniciando un proceso inflamatorio autoinmune posterior (Jess y cols., 2011; Stecher, 2015).

En contraposición a estos factores que podrían predisponer a desarrollar la CU, cabe destacar los factores ambientales y de estilo de vida que parecen adoptar un efecto protector frente a la enfermedad. Fundamentalmente son cuatro: la apendicetomía (extracción del apéndice), la lactancia materna, vitamina D y el consumo de tabaco, ya que en personas exfumadoras la tasa de incidencia de CU aumenta considerablemente (Dignass y cols., 2012). De hecho, los fumadores activos de tabaco no sólo presentan un claro efecto protector frente al desarrollo de la enfermedad sino también frente a la gravedad de esta. Por el contrario, los ex fumadores presentan alrededor de un 70% mayor de riesgo de padecer CU, a menudo además la enfermedad se presenta con un carácter más extenso y refractaria que en aquellos que nunca han fumado. Las tasas de ingreso en el hospital y la colectomía también son más altas en los ex fumadores que en los no fumadores (Dignass y cols., 2012).

En resumen, esta compleja interacción de factores genéticos, ambientales y microbianos culmina en una activación sostenida de la respuesta inmune y no inmune de la mucosa, probablemente facilitado por defectos en la barrera del epitelio intestinal y el sistema inmune de la mucosa, lo que causa inflamación activa y destrucción del tejido.

#### II.5.1.2. Prevalencia e incidencia de la CU

La CU presenta unas elevadas tasas de prevalencia e incidencia en Europa Occidental y Norteamérica (Dignass y cols., 2012), situándose la prevalencia media en estas zonas entre los 150-550 casos por cada 100.000 habitantes al año mientras que la incidencia entre 6-8 casos nuevos/100.000 habitantes al año. A su vez dentro del continente europeo, encontramos diferencias entre los países nórdicos y Reino Unido respecto a países meridionales como España, Grecia e Italia, ya que los primeros presentan más casos que los segundos (Ernst & Young, 2016).

En España, las tasas de prevalencia e incidencia han ido aumentando paulatinamente año tras año (Gomollón y cols., 2012). En 2012 había 100 casos nuevos por cada 100.000 habitantes al año, situándose tras Reino Unido y Alemania en cuanto a prevalencia de la enfermedad (Ernst & Young, 2016).

#### II.5.1.3. Impacto socio-económico de la CU

Esta enfermedad supone un elevado impacto social, dado que los pacientes con CU padecen graves limitaciones en su vida cotidiana en los periodos activos de la enfermedad. De hecho, se estima que el 75% de los pacientes se han sentido gravemente deprimidos en algún momento a causa de la enfermedad (Bodger y cols., 2014).

Por lo que respecta al impacto económico habría que diferenciar entre los costes asociados a la enfermedad de tipo directo e indirecto. Los directos se relacionan fundamentalmente con los gastos derivados de la atención del paciente (visitas médicas, pruebas de diagnóstico, visitas a urgencias, hospitalizaciones, cirugía y medicamentos). Los costes indirectos implican los relacionados con ausencias laborales temporales, costes por incapacidad laboral, pérdida de productividad de padres por cuidado a menores y gastos asumidos directamente por el paciente (Fedorak y cols., 2010).

En España, se estima que el coste directo total anual de la EII se sitúa alrededor de los 580 millones de euros, distribuyéndose de manera muy similar para ambas patologías (294 millones de euros para la CU y 285 millones de euros para la EC). Cabe destacar que el 59% del importe corresponde al tratamiento farmacológico, fundamentalmente relacionado al elevado coste de las terapias biológicas (Ernst & Young, 2016).

Lo anteriormente expuesto justifica sobradamente los esfuerzos encaminados al descubrimiento de mejores fármacos para el tratamiento de la EII, especialmente en lo referente a la CU, que además del impacto sanitario y científico, tendrían una repercusión en el ámbito económico, colaborando en la reducción de los gastos derivados de la enfermedad.

En el territorio español, se estima que los costes indirectos de la EII ascienden a un poco más de 503 millones de euros. Tal y como se ha citado en párrafos anteriores, la EII tiene un gran impacto sobre la calidad de vida del paciente y puede llegar a resultar muy incapacitante. Por lo que cuando el paciente tiene un brote sintomático, esta enfermedad puede impedir al paciente acudir al trabajo, ocasionando ausencias laborales. Derivándose esta ausencia laboral temporal en un coste anual total por ausencias laborales temporales a 41 millones de euros (Ernst & Young, 2016).

Sin embargo, la mayor parte de los costes indirectos guardan relación con los costes anuales derivados de la incapacidad laboral. Ya que la EII puede llegar a imposibilitar en algunos casos el desarrollo de la actividad profesional de los pacientes. Representando un coste anual de 380 millones de euros al Estado (Ernst & Young, 2016).



#### II.5.1.4. Terapia actual de la CU

Debido a que la colitis ulcerosa es una enfermedad que en la actualidad no tiene cura, los tratamientos están dirigidos a remitir los síntomas, evitando complicaciones como la necesidad de cirugía y contribuyendo a aumentar la calidad de vida del paciente.

Dentro de los tratamientos farmacológicos podemos distinguir entre los convencionales y los biológicos (Tabla 3). Los tratamientos convencionales se basan en la administración de aminosalicilatos (terapia de primera elección), corticoides e inmunomoduladores. Las terapias biológicas se emplean en pacientes que no responden a los tratamientos convencionales y están asociadas a efectos adversos de mayor gravedad (Tabla 3) (Dignass y cols., 2012). Finalmente, cuando cada uno de estos enfoques fallan, la colectomía representa el último recurso (Dignass y cols., 2012).

**Tabla 3.** Tratamientos farmacológicos de aplicación a la CU.\*

	Tratamientos convencionales			Tratamientos biológicos
	Aminosalicilatos	Corticoides	Inmunomoduladores	
<b>Utilidad</b>	Brote agudo Mantenimiento	Brote agudo	Brote agudo Mantenimiento	Brote agudo Mantenimiento
<b>Gravedad</b>	L - M	L – M – G	M – G	M – G
<b>Vía de administración</b>	Oral o rectal	Oral, intravenosa y rectal	Oral, subcutánea o intravenosa	Intravenosa o subcutánea
<b>Función</b>	Controlan y disminuyen la inflamación más leve	Regulan la respuesta del s.i. y disminuyen la producción celular de las sustancias que provocan la EII	Reducen la actividad del s.i., alterada por la EII	Disminuyen la inflamación y mejoran la cicatrización de la mucosa intestinal
<b>Efectos secundarios</b>	De carácter leve, como problemas gastrointestinales	Estéticos, <sup>a</sup> trastornos nerviosos <sup>b</sup> y osteoporosis	Riesgo de contraer infecciones u otras inflamaciones del organismo	Depresión del s.i. <sup>c</sup>
<b>Otros aspectos</b>	Más efectivos en CU que en EC	No se emplean de forma prolongada por sus efectos secundarios a largo plazo.	Requieren de análisis de sangre periódicos para ajustar la dosis a las necesidades del paciente, con el fin de potenciar la eficiencia del fármaco y reducir sus efectos secundarios	Objeto de investigación de la industria farmacéutica, para mejorar su eficiencia y reducir sus efectos secundarios

s.i.: sistema inmunitario; L: leve; M: moderado; G: grave.

<sup>a</sup> acné, cara hinchada o posible incremento de peso.

<sup>b</sup> variabilidad en el humor o alteración del sueño.

<sup>c</sup> aumento del riesgo de infecciones, neoplasias, reacciones autoinmunes y alérgicas.

\*(Ernst & Young, 2016).

El tratamiento convencional de la CU consiste en: aminosalicilatos, tales como ácido 5-aminosalicílico (5-ASA) o sulfasalazina; corticosteroides como la prednisona o budesonida; agentes inmunosupresores como la azatioprina, 6-mercaptopurina (6-MP), metotrexato o ciclosporina y, finalmente, el metronidazol. La elección de uno u otro fármaco perteneciente al grupo de los tratamientos convencionales depende del objetivo clínico (inducción o mantenimiento de la remisión), la extensión y gravedad de la enfermedad, la respuesta a la medicación actual o anterior y la presencia de complicaciones (Ardizzone y Porro, 2005).

Sin embargo, entre un 20% - 40% de los pacientes no responden a la medicación convencional y reciben terapia biológica, basada en anticuerpos monoclonales tales como infliximab y adalimumab. Estos anticuerpos monoclonales tienen como diana terapéutica al factor de necrosis tumoral. Por supuesto, no solo se han introducido anticuerpos monoclonales frente al TNF- $\alpha$  sino también frente a IL-6 como por ejemplo, tocilizumab. El cual, no sólo reduce la inflamación de los pacientes con EII sino que también previene el desarrollo de CCR asociado a colitis (Williams y Goh, 2011).

No obstante, a pesar de que los anticuerpos monoclonales disminuyen la inflamación y mejoran la cicatrización de la mucosa intestinal están asociados a efectos secundarios relacionados con la depresión del sistema inmune (aumento de riesgo de infecciones, neoplasias, reacciones autoinmunes y alérgicas). Por lo que la industria farmacéutica está sumando esfuerzos para identificar nuevos anticuerpos monoclonales con mayor eficacia y menores efectos adversos. Focalizando la actividad de los nuevos fármacos en la modulación de la respuesta inmune adquirida intestinal y, en menor grado, la innata para regular el proceso inflamatorio (Williams y Goh, 2011), ya que hoy en día, la teoría de que la EII se debe a una desregulación del sistema inmune, junto con una predisposición genética y determinados factores ambientales está ampliamente aceptada.

### II.5.1.5. Mediadores implicados en la CU

El sistema inmune de la mucosa intestinal es el efector central de la inflamación intestinal y sus lesiones. En situaciones normales, la mucosa intestinal se encuentra en un estado de inflamación "controlada" regulado por un delicado equilibrio de citocinas proinflamatorias y antiinflamatorias (Ardizzone y Porro, 2005). Es por esto que las citocinas son un lógico *target* frente a la EII. Siendo ejemplo de esto, los anticuerpos monoclonales los cuales tiene como diana terapéutica al TNF- $\alpha$  y a la IL-6 entre otros (Ardizzone y Porro, 2005).

Las citocinas, proteínas que actúan como pequeñas moléculas de señalización celular secretadas por diversos tipos de células incluyendo las células inmunes y gliales, son los principales mediadores de la cicatrización de la mucosa en la EII (Roda y cols., 2011).

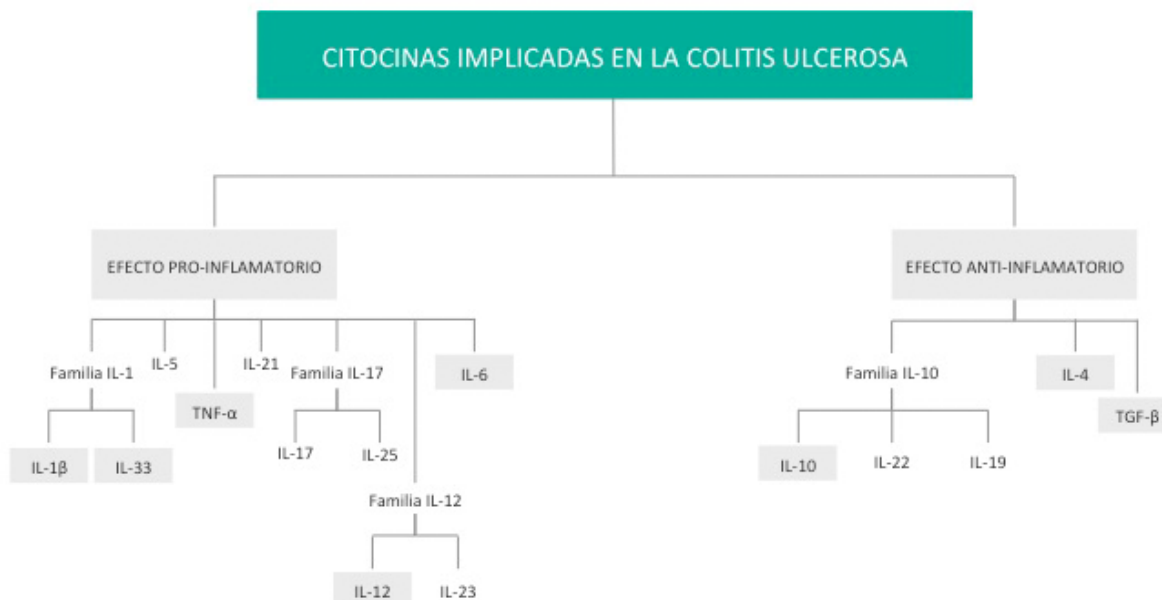
Cuando se produce una respuesta inmune aberrante mediada por diferentes citocinas, estas estimulan la proliferación de células efectoras específicas de antígeno con el fin de activar el sistema inmune adaptativo, con la consiguiente explosión de la inflamación local y sistémica (Roda y cols., 2011).

Las citocinas juegan muchos papeles diferentes en la EII, como la producción de mediadores inflamatorios, óxido nítrico, leucotrienos, factor activador de plaquetas, activación de la vía NF-kappa B e inhibición de la apoptosis; por consiguiente, las citocinas en sí son directamente responsables de la lesión de la mucosa y el consiguiente daño a los tejidos (Roda y cols., 2011).

Por lo que emplear a las citocinas como dianas terapéuticas frente a CU con el fin de interrumpir la cascada inmunológica y así prevenir los efectos proinflamatorios y por lo tanto el reclutamiento de leucocitos es una atractiva opción, debido a que son muchas las citocinas involucradas (Roda y cols., 2011).

Las citocinas implicadas en la CU se pueden clasificar según su efecto anti o pro inflamatorio (Figura 14). Dentro de las citocinas proinflamatorias destacan: el TNF- $\alpha$ ; la familia de las IL-1 (IL-1 $\beta$  y IL-33); la IL-6; la familia de la IL-12 ( IL-12 e IL-23) y la IL-13. Otras de menor importancia en la CU pertenecientes a esta clase son: IL-5, IL-21, IL-8, IL-17 y la IL-25. Mientras que dentro del grupo de las citocinas con actividad

antiinflamatoria destacan la familia de la IL-10 (IL-10, IL-19 e IL-22), el TGF- $\beta$  y la IL-4 (Roda y cols., 2011).



**Figura 14.** Principales citocinas implicadas en la CU.

Una de las citocinas proinflamatorias más relevantes en la CU es el TNF- $\alpha$ . Esta citocina, producida por las células presentadoras de antígeno APC y los macrófagos, activa las células T en respuesta a la lesión intestinal (Fiorino y cols., 2011). Su producción excesiva conduce al reclutamiento de células inmunes en el intestino, síntesis y liberación de citocinas proinflamatorias como la IL-1 $\beta$  e IL-6 y a la activación de fibroblastos y células endoteliales (Delhalle y cols., 2004; Atreya y cols., 2008; Arora y Shen, 2015).

La relación entre el TNF- $\alpha$  y la CU está ampliamente documentada por la elevada expresión de TNF- $\alpha$  en biopsias de la mucosa intestinal de las zonas afectadas en pacientes con CU respecto a pacientes sanos. Por último, los niveles séricos de TNF- $\alpha$  se correlacionan con los índices clínicos y de laboratorio de actividad de la enfermedad, tales como la velocidad de sedimentación globular, proteína C-reactiva (PCR), y el índice de actividad de la enfermedad (IAE) (Roda y cols., 2011).

Por lo tanto, el bloqueo o inactivación del TNF- $\alpha$  es una diana atractiva para la terapia de la CU. De hecho, nuevos anticuerpos monoclonales anti-TNF- $\alpha$  han sido aprobados en los últimos años para el tratamiento de la CU (Park y Jeon, 2015).

La familia de la IL-1 representa dos formas estructuralmente distintas de citocinas: IL-1 $\alpha$  e IL-1 $\beta$ . La principal fuente de IL-1 en la EII es el sistema de monocitos/macrófagos, ya que puede activar la enzima convertidora de IL-1, liberando IL-1 activa en la mucosa del colon (Múzes y cols., 2012). Dicha citocina a su vez, activa la expresión de COX-2, iNOS y fosfolipasa A2 (PLA2) por diversos tipos de células (Múzes y cols., 2012).

Niveles altos de IL-1 $\beta$  han sido detectados en la mucosa del colon de los pacientes con CU (Gionchetti P et al.), estos se correlacionan con el nivel de gravedad de la enfermedad. Además el aumento de la relación de IL-1/antagonista endógeno del receptor de IL-1 (IL-1Ra) guarda una relación directa con el grado de inflamación y por tanto, con la actividad de CU (Múzes y cols., 2012).

Los efectos proinflamatorios de TNF- $\alpha$  e IL-1 se deben principalmente a su capacidad para activar al NF- $\kappa$ B, a través de una serie de cascadas de señalización complejas que conducen a la degradación de la I $\kappa$ -B $\kappa$  (un inhibidor de la activación de NF- $\kappa$ B). Cuando se activa, el NF- $\kappa$ B se transloca al núcleo y comienza su transcripción (Atreya y cols., 2008). El NF- $\kappa$ B también está involucrado en la transcripción de genes que codifican enzimas tales como iNOS, COX-2, lipooxigenasa-5 y 12, quimiocinas y citocinas (IL-1, IL-6 y TNF- $\alpha$ ) (Delhalle y cols., 2004; Atreya y cols., 2008). Por otra parte, hay evidencias que indican que la activación del NF- $\kappa$ B de tipo constitutivo conduce a la aparición de cáncer asociado a CU (Delhalle y cols., 2004).

Una citocina ampliamente relacionada con la CU es la IL-6, secretada a nivel intestinal por los macrófagos y esencial en la evolución entre la inmunidad innata y adquirida. Niveles elevados de IL-6 se han asociado con pacientes con CU (Waldner y Neurath, 2009; Múzes y cols., 2012).

Esta citocina activa un complejo de señalización celular superficial compuesto por IL-6, el receptor soluble de IL-6 (sIL-6R) y el receptor de señalización compartida gp130 (Múzes y cols., 2012). IL-6 ejerce la mayoría de sus efectos proinflamatorios mediante la unión a su receptor soluble (sIL-6R), y esto estimula la unión sólo con

células que expresan gp130 pero no IL-6R; este proceso se llama trans-señalización, y se produce a través de la molécula de STAT3 (Roda y cols., 2011). Por todo ello, el bloqueo de la vía de señalización de IL-6/STAT3 ha sido sugerido como una opción terapéutica prometedora (Múzes y cols., 2012).

Otras citocinas proinflamatorias cuyos niveles se encuentran en cantidades elevadas en la mucosa de los pacientes con CU (Roda y cols., 2011) son la IL-12 y IL-13. Son típicas citocinas tipo Th2 producidas por las células NK (*natural killer*). La IL-13 es responsable del daño a la barrera epitelial, la velocidad de reparación de la mucosa y la alteración de las uniones estrechas, con el consiguiente deterioro de la permeabilidad de la mucosa (Múzes y cols., 2012).

El otro gran grupo de citocinas implicados en el desarrollo de la CU son las de tipo antiinflamatorio: IL-10, IL-4 y TGF- $\beta$ .

La IL-10 es la citocina antiinflamatoria más importante en humanos, producida por células T, B y monocitos. En presencia de un estímulo antigénico, IL-10 inhibe la presentación de antígenos y la liberación de citocinas proinflamatorias (IL-1 $\beta$ , IL-6 y TNF- $\alpha$ ) atenuando con todo ello el proceso inflamatorio en la mucosa (Roda y cols., 2011; Múzes y cols., 2012).

El análisis de los niveles de ARNm de esta citocina en pacientes con CU mostró un aumento significativo de IL-10 en la *lamina propria* del colon. Sin embargo, este aumento en la concentración de IL-10 no es suficiente para inhibir adecuadamente la actividad secretora de monocitos y macrófagos (Roda y cols., 2011; Múzes y cols., 2012).

La citocina antiinflamatoria IL-4, producida por linfocitos es un factor estimulador de células T y B que genera un efecto inmunosupresor en el colon. Su papel principal se encuentra en la inhibición de la formación de colonias de macrófagos, de la producción de H<sub>2</sub>O<sub>2</sub> por monocitos y de la liberación de mediadores tales como IL-1 $\beta$  y TNF- $\alpha$  (Roda y cols., 2011; Múzes y cols., 2012). Los niveles de dicha citocina aparecen disminuidos en los pacientes con CU respecto a individuos sanos, demostrando la pérdida de equilibrio entre citocinas anti y proinflamatorias en favor de estas últimas (Roda y cols., 2011).

Finalmente, el TGF- $\beta$  es una citocina inhibidora, que representa un papel clave en el mantenimiento de las respuestas inmunes e inflamatorias. La evidencia sugiere que el TGF- $\beta$  actúa junto con factores de crecimiento en la protección de los tejidos del huésped de los cambios luminales y puede ayudar a la curación de la mucosa en la EII (Roda y cols., 2011). De hecho, los pacientes con CU muestran un aumento de la síntesis de TGF- $\beta$ 1 por células mononucleares de la lámina propia en comparación con los controles sanos (Roda y cols., 2011).

Recientemente, se ha publicado el efecto protector frente a CU de fármacos inhibidores de las vías de señalización IL-6/STAT-3, como es el anticuerpo tocilizumab (Li y cols., 2010; Pandurangan y cols., 2016) y de la vía TLR4/NF- $\kappa$ B, como la atorvastatina (Rashidian y cols., 2016).



# OBJETIVOS



### III. OBJETIVOS

De acuerdo con lo expuesto anteriormente, los dos grandes objetivos perseguidos en la presente Tesis Doctoral son:

1. La identificación, a través de la Topología molecular, de nuevos agentes antiinflamatorios (Capítulo 1).
2. El descubrimiento, por la misma metodología, de nuevos agentes protectores frente a CU (Capítulo 2).

Para ello se llevan a cabo las siguientes tareas:

- Obtención y validación de modelos topológico-matemáticos.
- Identificación de compuestos con actividad antiinflamatoria y protectora frente a CU.
- Validación experimental de los compuestos seleccionados mediante ensayos *in vitro* e *in vivo*.



# CAPÍTULO 1

## **Selección de nuevos agentes antiinflamatorios por Topología molecular:**

Artículo 1: “**Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity**” (En: *Letters in Drug Design & Discovery*, **2010**, 7, 438-445). JCR [2010, CAT: Chemistry, Medicinal; FI: 0, 668; Rank: 48 de 54 (**Q4**)].



# Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity

María Galvez-Llompарт<sup>1</sup>, Rosa M. Giner<sup>2</sup>, María C. Recio<sup>2</sup>, Sanzio Candeletti<sup>3</sup> and Ramón García-Domenech\*<sup>1</sup>

<sup>1</sup>Molecular Connectivity and Drug Design Research Unit, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, Spain; <sup>2</sup>Department of Pharmacology, Faculty of Pharmacy, University of Valencia, Spain; <sup>3</sup>Department of Pharmacology, Faculty of Pharmacy, University of Bologna, Italy

Received October 27, 2009; Revised February 05, 2010; Accepted March 15, 2010

**Abstract:** A topological-mathematical model obtained by linear discriminant analysis has been used to the search of new nonsteroidal antiinflammatory drugs (NSAIDs). After carrying out an *in silico* screening based on such a model, on the Aldrich database, new structures potentially active were selected. Among these structures stand fourteen compounds, from which only one had been previously recorded as NSAID in the literature. The experimental tests performed on the remaining substances demonstrated that several compounds showed either *in vitro* or *in vivo* or both activity. Moreover, four compounds, namely 1,3-bis(benzyloxycarbonyl)-2-methyl-2-thiopseudourea, 4,6-dichloro-2-methylthio-5-phenylpyrimidine, 2-chloro-2',6'-acetoxylidide and trans-1,3-diphenyl-2-propen-1-ol, showed a significant *in vivo* antiinflammatory activity as compared to the reference drug (indomethacin). These results reinforce the role of Molecular Topology as a useful tool for drug discovery.

**Keywords:** COX-1 and COX-2 Inhibitors, Linear Discriminant Analysis, Molecular Topology, Nonsteroidal Antiinflammatory.

## INTRODUCTION

The non steroidal antiinflammatory drugs (NSAIDs) stand among the most largely used worldwide. Just in the USA, about 1% of population consumes them daily, what is due to its triple action: analgesic, antiinflammatory and antipyretic. The NSAIDs reduce or cure types of pain related to different diseases, including acute or chronic pain.

Most of pains likely to be suffered along the lifetime are sensible to these drugs, as for instance myalgia, artralgia, cephalalgia, neuralgia, dysmenorrheal and acute or chronic inflammatory processes such as otitis, periodontitis, adnexitis, arthritis and different types of arthrosis. Furthermore, they are often useful in the unrest linked to viral and bacterial processes. They constitute the first level treatment of pain in the World Health Organization (WHO) strategy.

Within the NSAIDs, the so called arylpropionic derivatives, namely ibuprofen, mainly COX-1 selective, is one of the drugs having experienced a dramatic increase in the marketplace demand.

To sum up, all what mentioned above accounts for the selection of this therapeutic group on our goal to achieve new *hits* and *lead* compounds.

Several mathematical approaches have been proposed to ease the search for new active compounds. Equation systems linking quantitative structure-activity -QSAR relationships- are particularly relevant and can be applied to large libraries of compounds for virtual computational screening [1-3].

However, these models require good structural descriptors that reliably represent the molecular features responsible for the relevant pharmacological activity. Molecular topology (MT) is a very useful approach to describe molecular structures so that the structure of each molecule is represented by specific subsets of topological indices (TIs) [4]. These indices are able to provide a unique framework to characterize the structures. Moreover, they correlate with many physical, chemical, and biological properties of structurally heterogeneous groups of compounds and can be used to find out new drugs [5-9].

The aim of this study is to develop and assess a QSAR model based on molecular topology capable of identifying new NSAIDs.

## MATERIAL AND METHODS

A training set comprised of 21 compounds with known antiinflammatory activity (active set) and 27 compounds not showing such activity (inactive set) (see Table I) were selected from a literature search in several medical databases (as for instance the 13th Edition of The Merck Index .Whitehouse Station, N.J., USA). The chemical structure of each drug was depicted with the aid of the Chemdraw software package, version 10. Each compound was characterized by a set of 434 TIs obtained with the Dragon software, version 5.4 [10].

The regression algorithm employed was the Linear Discriminant Analysis, LDA, which is a statistical tool providing a classification model based on the combination of variables that best predicts the category or group to which a given compound belongs [11]. In our case, the training set compounds were allocated to *active* or *inactive* groups according to their antiinflammatory activity or inactivity,

\*Address correspondence to this author at the Molecular Connectivity and Drug Design Research Unit, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia Avd. V.A. Estellés, s/n 46100-Burjassot, Valencia, Spain; Tel: +34 963 544 291; Fax: +34 963 544 892; E-mail: ramon.garcia@uv.es

Table 1. Results of Prediction of Antinflammatory Activity Obtained by Linear Discriminant Analysis

Compound	Topological Indices			DF	Clas.(calc)
	nCl	PCR	MATS8e		
<b>Active group training</b>					
Aceclofenac	2	1.452	0.232	4.63	A
Amias	0	1.519	-0.114	3.20	A
Amolmetin guacil	0	1.446	-0.192	2.66	A
Bipotrol	0	1.501	0	1.53	A
Celecoxib	0	1.391	-0.455	4.60	A
Chlortenoxicam	1	1.353	-0.283	5.24	A
Clonidine	2	1.321	0	4.71	A
Etoricoxib	1	1.468	-0.002	4.27	A
Felbinac	0	1.397	-0.510	5.35	A
Flurbiprofen	0	1.361	-0.264	1.81	A
Indomethacin	1	1.370	-0.521	8.30	A
Ketamine	1	1.270	0	0.35	A
Ketoprofen	0	1.404	-0.287	2.92	A
Lenalidomide	0	1.298	-0.440	2.60	A
Lorazepam	2	1.412	-0.635	13.8	A
Melatonin	0	1.292	-0.492	3.08	A
Mofezolac	0	1.384	-0.523	5.24	A
Naproxene	0	1.360	-0.757	7.46	A
Nepafenac	0	1.411	-0.319	3.43	A
Piroxicam	0	1.427	-0.199	2.37	A
Vedaprofen	0	1.311	-0.337	1.67	A
<b>Inactive group training</b>					
Adiphenine	0	1.261	0.007	-3.27	I
Azastreptonigrin	0	1.459	0.041	0.24	A
Busulfan	0	1.111	-0.210	-3.72	I
Carbutamide	0	1.272	0.172	-4.95	I
Cetiedil	0	1.075	-0.052	-6.25	I
Cyclopentamine	0	1.000	0	-8.32	I
Dicyclomine	0	1.010	0.040	-8.58	I
Diphenadione	0	1.342	-0.050	-1.02	I
Dopamine	0	1.261	0	-3.19	I
Fluorouracil	0	1.205	0	-4.29	I
Iobenguane	0	1.288	0	-2.66	I
Labetalol	0	1.348	-0.118	-0.12	I
Megalomicin	0	1.005	0.016	-8.41	I
Menthoval	0	1.019	0.133	-9.47	I
Nesapidil	0	1.363	0.123	-2.59	I
Nilutamide	0	1.209	-0.375	0.10	A
Phthalylsulfathiazole	0	1.434	0.200	-2.08	I
Pronetalol	0	1.393	0.024	-0.87	I



(Table 1). Contd.....

Compound	Topological Indices			DF	Clas.(calc)
	nCl	PCR	MATS8e		
Protheobromine	0	1.168	0	-5.02	I
Racephedrine	0	1.265	0	-3.11	I
Sisomicin	0	1.050	-0.003	-7.30	I
SNAP 6201	0	1.263	0.348	-7.15	I
Spiperone	0	1.266	-0.082	-2.15	I
Sulfadiazole	0	1.309	0.418	-7.04	I
Sulfaethidole	0	1.339	0.339	-5.55	I
Sulfonmethane	0	1.089	0	-6.57	I
Taurolidine	0	1.074	-0.460	-1.58	I

respectively. Hence, the discriminatory property was the anti-inflammatory activity and independent variables were the TIs. The software employed to perform the LDA was the BMDP New System 2 package, module 7M [12].

From the selected discriminant function, the corresponding distribution diagram of anti-inflammatory activity, PDD [13], was drawn. This diagram is arranged just to establish the intervals of the discriminant function in which the expectancy (E) of finding active compounds is maximum. PDDs are histogram-like plots of connectivity functions in which expectancies appear on the ordinate axis. For an arbitrary interval of values of a given function, we define the expectancy of activity as:  $E_a = a/(i+1)$ ; where "a" is the number of active compounds in the interval divided by the total number of active compounds, and "i" is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity is defined in a similar way, as  $E_i = i/(a+1)$ . This plot provides a good visualization of minimum overlap regions, and allows the selection of frames in which the probability of finding active compounds is maximum.

## In Vitro Assays

### Cell Culture

The murine macrophage cell line RAW 264.7 (*ECACC*, *Salisbury, GB*) were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 10% fetal bovine serum (FBS). All cells were grown in a humidified atmosphere containing 5% CO<sub>2</sub> at 37°C. Macrophage RAW 264.7 were plated at a density of 1x10<sup>6</sup> cells per ml. For stimulation, the medium was replaced with fresh DMEM 0.5 % FBS, and the cells were then stimulated with 1 µg/ml lipopolysaccharide (LPS).

### Determination of Cell Viability

The cytotoxicity of compounds was tested by the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl-tetrazolium bromide (MTT) assay [14]. Murine RAW 264.7 macrophages (1x10<sup>6</sup> cells per ml) were exposed to the compounds (100 µM) in a microplate and then incubated at 37°C for 24 h. The medium was then removed and 100 µl per well of a 0.5 mg/ml solution of MTT were added and then solution was incubated at

37°C until blue deposits were visible. The colored metabolite was dissolved in dimethyl sulfoxide (DMSO). Absorbance was measured at 490 nm using a Labsystems Multiskan EX plate reader (Helsinki, Finland). Results were expressed in absolute absorbance readings; a decrease indicated a reduction in cell viability.

### Nitrite (NO<sub>2</sub><sup>-</sup>) Production in RAW 264.7 Macrophages

NO production in culture supernatant was spectrophotometrically evaluated by measuring nitrite, an oxidative product of NO. Nitrite was determined with the Griess reaction [15] by mixing 100 µl of culture supernatant with 100 µl of Griess reagent containing equal volumes of 1% (w/v) sulphanilamide in 5% (w/v) phosphoric acid and 0.1% (w/v) of N-(1-naphthyl)ethylenediamine solution. Absorbance was measured at 540 nm against a calibration curve with sodium nitrite standards. Compounds were assayed at 10, 50 and 100 µM. In this experiment, nitrite production was assessed as the index of NOS activity.

### Prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) Production in RAW 264.7 Macrophages

Prostaglandin E<sub>2</sub> production was determined in culture supernatant by a specific enzyme immunoassay kit from Cayman Chemical Co, used according to the manufacturer's instructions. Compounds were assayed at 10, 50 and 100 µM. In this experiment, PGE<sub>2</sub> production was measured in order to determine COX-2 activity.

## In Vivo Assays

### Animals

Groups of six Swiss female mice weighting 25–30 g from Harlan Interfauna Ibérica (Barcelona, Spain) were used. Housing conditions and the *in vivo* experiments were approved by the Institutional Ethics Committee of the University of Valencia (Spain) in accordance with the guidelines established by the European Union on Animal Care (CEE Council 86/609).

### Paw Edema Induced by Carrageenan

Carrageenan-induced paw edema test as described by Sugishita *et al.* [16] was assayed. Compounds were dissolved in EtOH/Tween 80/H<sub>2</sub>O (1:1:10) and orally adminis-

tered. The control group received the vehicle only, the reference group received indomethacin (10 mg/kg) as a standard drug and the rest of groups, were treated with the compounds (10 mg/kg). After one hour, the edema was induced by subcutaneous injection of 25  $\mu$ l of a 3 % (w/v) suspension of carrageenan in saline solution in the subplantar region of the right hind paw in all the animals. The paw edema was measured plethysmographically (Ugo Basile 7150, Varese Italy plethysmograph) at intervals of 1, 3 and 5 h after injection of carrageenan. The edema volume is expressed in each animal as the difference found between the right hind paw volume compared against the left hind one. The edema inhibition is expressed as the percentage of volume reduction relative to the control.

### 12-O-Tetradecanoylphorbol Acetate (TPA)-Induced Ear Edema

Topical antiinflammatory activity of the compounds was studied using the method described by De Young *et al.* [17]. Edema was induced on the right ear by topical application of 2.5  $\mu$ g/ear of TPA in 20  $\mu$ l acetone. The left ear was left untreated and used as control. Compounds and indomethacin (0.5 mg/ear) were dissolved in 20  $\mu$ l acetone and applied to right ear simultaneously with TPA. Four hours after the inflammation induction the animals were sacrificed and a 6 mm diameter ear biopsy was performed with a leather punch. Edema was measured as the difference in weight between the punches from right and left ears due to the TPA agent application. The inflammation inhibition percentage was evaluated as the weight difference between treated and non-treated ears of each animal compared to the control group (vehicle). Details of the method have been described earlier by Giner *et al.* [18].

### Reagents

Biochemicals, chemicals, reagents and materials were purchased from Invitrogen, Cayman Chemical Company, Merck, Panreac, Fluka and Sigma-Aldrich.

## RESULTS AND DISCUSSION

In order to obtain the discriminant function, we applied the LDA to a training set comprised of 48 compounds distributed in two subsets: Active (21 compounds with activity) and an inactive (27 compounds not showing, or at least not described as such in the literature, antiinflammatory activity).

The discriminant function selected was:

$$DF = 3.36 nCl - 11.49 MATS8e + 19.67 PCR - 27.99$$

$$N = 48 \quad \lambda \text{ (Wilks' lambda)} = 0.3140 \quad F\text{-statistic} = 32.0$$

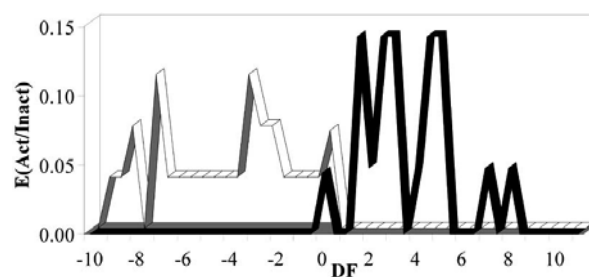
where nCl = number of atoms of chlorine; MATS8e = Moran autocorrelation indices weighted by the Saunders electro-negativity; and PCR = Ratio of multiple path count over path count.

These statistical parameters demonstrate the significance of the equation achieved. Moreover, the function was also checked as for its stability, randomness as well as its predictive capability over test sets of active and inactive compounds.

The classification matrix for DF is very significant for the training set: 100% of correct prediction for the active group, i.e. 21 out of 21 correctly classified, and 92.6% for the inactive group, i.e. 25 out of 27 were correct (see Table 1). The  $\lambda$  value was 0.314 in the cross validation analysis and the percentage of correct classification was 95.2% and 88.9% within the active and inactive, respectively (very close to the training set results). The internal validation yields 98.8, 93.6, 95 and 96 percent for active/inactive in the training and test sets, respectively.

Once confirmed the function's robustness, it was applied so that a given compound was selected as a potential antiinflammatory if  $DF > 0$ , otherwise it was classified as inactive.

Fig. (1) shows the PDD obtained. It is apparent that the regions with minimum overlap for the compounds with theoretical antiinflammatory activity occur between  $DF > 0$  and  $DF < 10$ , so the highest expectancy of activity takes place within that range of DF values.



**Fig. (1).** Pharmacological distribution diagram for antiinflammatory activity by plotting expectancy (E) versus DF function (the black and white bars represent the active and inactive compounds, respectively).

Based upon the topological model described above, a virtual screening in the Sigma Aldrich database was carried out in order to select new potential NSAID compounds. Table 2 shows the outcome.

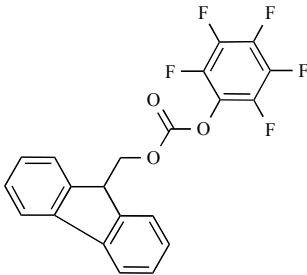
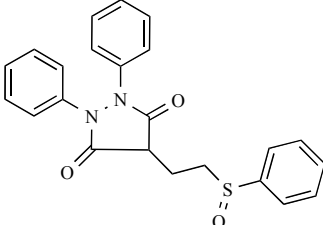
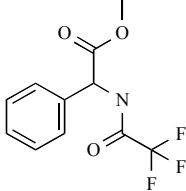
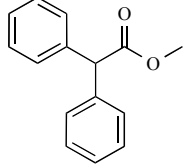
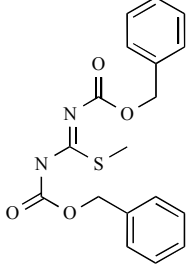
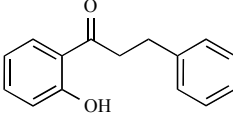
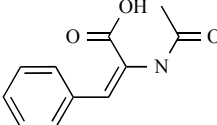
It is important to mention that several molecular candidates were discarded for testing due to different reasons: Lack of solubility, (compounds 6, 10 and 12) and previously reported NSAID activity (compound 4) [19]. The remaining compounds, namely # 1, 2, 3, 5, 7, 8, 9, 11, 13 and 14 were tested.

Table 2 illustrates all the selected compounds with their corresponding CAS numbers and DF values as well as the experimental results achieved.

Four compounds, namely 1,3-bis(benzyloxycarbonyl)-2-methyl-2-thiopseudourea (# 5), 4,6-dichloro-2-methylthio-5-phenylpyrimidine (# 8), 2-chloro-2',6'-acetoxylidide (# 9) and trans-1,3-diphenyl-2-propen-1-ol (# 14), showed a significant *in vivo* antiinflammatory activity as compared to the reference drug (indomethacin).

In particular, as regards the inhibition of carrageenan-induced paw edema (Table 3), the four active compounds displayed an activity between 33 (compound 8; 1h) and 64 %

Table 2. Selected Compounds and Experimental Results for NSAID Activity

N° Compound /CAS	Structure	<i>In silico</i>	<i>In vitro</i> (inhibition)		<i>In vivo</i> (inhibition)	
		DF1 values	NO <sub>2</sub> <sup>-a</sup>	PGE <sub>2</sub> <sup>b</sup>	Carrageenan <sup>c</sup>	TPA <sup>d</sup>
1/ 88744-04-1		2.29	- <sup>e</sup>	-	-	-
2 / 57-96-5		0.67	+ <sup>f</sup>	+	-	-
3 / 145513-97-9		10.53	+	+	-	-
4 / 3469-00-9		0.23	N.T. <sup>g</sup>	N.T.	N.T.	N.T.
5 / 25508-20-7		1.88	+	+	+	+
6 / 3516-95-8		1.61	N.T.	N.T.	N.T.	N.T.
7 / 5469-45-4		3.19	+	+	-	-

(Table 2). Contd.....

N ° Compound /CAS	Structure	<i>In silico</i>	<i>In vitro</i> (inhibition)		<i>In vivo</i> (inhibition)	
		DF1 values	NO <sub>2</sub> <sup>-a</sup>	PGE <sub>2</sub> <sup>b</sup>	Carrageenan <sup>c</sup>	TPA <sup>d</sup>
8 / 65515-11-8		4.02	+	+	+	+
9 / 1131-01-7		1.18	+	+	+	+
10 / 4124-41-8		0.05	N.T.	N.T.	N.T.	N.T.
11 / 128544-06-9		6.56	-	-	-	-
12 / 22287-56-5		10.68	N.T.	N.T.	N.T.	N.T.
13 / 84-62-8		5.54	N.T.	N.T.	-	-
14 / 62668-02-4		1.16	N.T.	N.T.	+	+
Indomethacin (reference drug)		8.3	+	+	+	+

<sup>a</sup>NO<sub>2</sub><sup>-</sup> production in RAW 264.7 macrophage at 50 μM, after stimulation with LPS.<sup>b</sup>PGE<sub>2</sub> production in RAW 264.7 macrophage at 50 μM, after stimulation with LPS.<sup>c</sup>Carrageenan induced mice paw edema test at 3 h. Compounds were tested at dosage of 10 mg/kg.<sup>d</sup>TPA induced ear edema test at 4 h. Compounds were tested at dosage of 0.5 mg/ear in 20 μl acetone.<sup>e</sup>Inhibition below 50 % of indomethacin.<sup>f</sup>+ Inhibition at least 50% of indomethacin.<sup>g</sup>N.T. Not tested.

**Table 3. Antinflammatory Effect of Indomethacin and Compounds (5, 8, 9 and 14) on Acute Carrageenan-Induced Mouse Paw Edema. Indomethacin and Compounds were Orally Administrated (10 mg/kg), 1 h Before Carrageenan Injection**

Product	1h		3h		5h	
	$\Delta V \pm SEM (\mu l)^a$	% I	$\Delta V \pm SEM (\mu l)$	% I	$\Delta V \pm SEM (\mu l)$	% I
Control	72 $\pm$ 5	-	77 $\pm$ 7	-	80 $\pm$ 8	-
Indomethacin	34 $\pm$ 5 <sup>**b</sup>	52	33 $\pm$ 4 <sup>**</sup>	58	42 $\pm$ 5 <sup>**</sup>	48
Compound 5	43 $\pm$ 6 <sup>**</sup>	40	53 $\pm$ 7 <sup>*</sup>	31	52 $\pm$ 3 <sup>**</sup>	35
Compound 8	48 $\pm$ 4 <sup>*</sup>	33	41 $\pm$ 5 <sup>**</sup>	47	36 $\pm$ 4 <sup>**</sup>	55
Compound 9	26 $\pm$ 2 <sup>**</sup>	64	49 $\pm$ 5 <sup>**</sup>	36	52 $\pm$ 8 <sup>**</sup>	35
Compound 14	45 $\pm$ 7 <sup>**</sup>	37	49 $\pm$ 6 <sup>**</sup>	36	36 $\pm$ 2 <sup>**</sup>	55

<sup>a</sup>Values represent the mean  $\pm$  standard error of mean (S.E.M.). ( $n=6$ ).

<sup>b</sup>\*\* $P < 0.01$ , \* $P < 0.05$  with respect to the control group (Dunnett's  $t$ -test).

(compound 9; 1h), at different times of determination. Whilst, Indomethacin displayed an activity between 48 (5h) and 58% (3h) at different times.

The different times showing the maximal inhibition for each compound (e.g. 1h for compound 9; 3h for compound 8) could probably reflect the different pharmacokinetic properties and characteristics of single molecules.

Concerning the TPA-induced ear edema the cited compounds showed about the 60 % of indomethacin antinflammatory activity (Table 4) expressed as percentage inhibition of ear weight increase.

**Table 4. Antinflammatory Effect of Indomethacin and Compounds (5, 8, 9 and 14) on Acute TPA-Induced Ear Mouse Edema. Indomethacin and Compounds were Topically Applied, at Dosage of 0.5 mg/ear in 20  $\mu$ l Acetone. The Weight of the Ear Punch (6 mm) was Measured After 4 h of Treatment with the Irritant**

Product	$\Delta P \pm SEM (mg)^a$	% I
Control	18 $\pm$ 1	-
Indomethacin	1 $\pm$ 0 <sup>**b</sup>	92
Compound 5	8 $\pm$ 2 <sup>**</sup>	52
Compound 8	8 $\pm$ 2 <sup>**</sup>	58
Compound 9	7 $\pm$ 2 <sup>**</sup>	61
Compound 14	7 $\pm$ 1 <sup>**</sup>	59

<sup>a</sup>Values represent the mean  $\pm$  standard error of mean (S.E.M.). ( $n=6$ ).

<sup>b</sup>\*\* $P < 0.01$  with respect to the control group (Dunnett's  $t$ -test).

As for the cytotoxicity and *in vitro* activity, the results (not shown here), demonstrated that none of the selected compounds was cytotoxic and 90% of the cells showed viability in the experimental culture conditions after the mixture. Furthermore, compounds # 8 and 9 exhibited also a significant *in vitro* activity (42 and 51% PGE<sub>2</sub> inhibition as well as 61 and 77% inhibition of nitrites, respectively).

## CONCLUSION

In this work a virtual screening by molecular topology, employing linear discriminant analysis, result in the selection of compounds with not yet known NSAIDs properties. Four

out of ten tested compounds showed a significant *in vivo* antinflammatory activity and two of them exhibited also *in vitro* activity, whilst none of the compounds was cytotoxic. The new compounds might serve as novel important *leads* for further pharmacological investigations. In summary, this paper shows how molecular topology is used as an outstanding and reliable tool for the discovery of novel NSAIDs.

## ACKNOWLEDGEMENTS

We thank the Ministerio de Educación e Innovación, Spain (projects SAF2009-13059-C03-01 and SAF2009-13059-C03-02) for support of this work. The authors want to express their acknowledgements to Profs. Patrizia Romualdi and Jose Luis Rios for their availability to experimental tests at the University of Bologna and Valencia, respectively, and to Prof. Jorge Galvez from the University of Valencia, for his help and useful comments about the new compounds.

## REFERENCES

- [1] Bruno-Blanch, L.; Gálvez, J.; García-Domenech, R. Topological virtual screening: a way to find new anticonvulsant drugs from chemical diversity. *Bioorg. Med. Chem. Lett.*, **2003**, *13*(16), 2749-2754.
- [2] Duart, M.J.; Anton-Fos, G.M.; Aleman, P.A.; Gay-Roig, J.B.; Gonzalez-Rosende, M.E.; Galvez, J.; Garcia-Domenech, R. Application of a mathematical topological pattern of antihistaminic activity for the selection of new drug candidates and pharmacology assays. *J. Med. Chem.*, **2006**, *49*, 3667-3673.
- [3] Garcia-Domenech, R.; Galvez, J.; de Julian-Ortiz, J.V.; Pogliani, L. Some new trends in chemical graph theory. *Chem. Rev.*, **2008**, *108*(3), 1127-1169.
- [4] Devillers, J.; Balaban, A.T.; Eds. *Topological Indices and Related Descriptors in QSAR and QSPR*; Gordon and Breach Science Publishers: Singapore, **1999**.
- [5] Gálvez, J.; Garcia-Domenech, R.; de Julián-Ortiz, V.; Soler, R. Topological approach to drug design. *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 272-284.
- [6] de Gregorio-Alapont, C.; Garcia-Domenech, R.; Gálvez, J.; Ros, M.J.; Wolski, S.; Garcia, M.D. Molecular topology: a useful tool for the search of new antibacterials. *Bioorg. Med. Chem. Lett.*, **2000**, *10*(17), 2033-2036.
- [7] Rios-Santamarina, I.; Garcia-Domenech, R.; Gálvez, J.; Cortijo, J.; Santamaria, P.; Morcillo, E. New bronchodilators selected by molecular topology. *Bioorg. Med. Chem. Lett.*, **1998**, *8*(5), 477-482.
- [8] Gálvez, J.; Gomez-Lechón, M.J.; Garcia-Domenech, R.; Castell, J.V. New cytostatic agents obtained by molecular topology. *Bioorg. Med. Chem. Lett.*, **1996**, *6*(19), 2301-2306.
- [9] Roy, K. Topological descriptors in drug design and modeling studies. *Mol. Divers.*, **2004**, *8*, 321-323.

- [10] Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim (GER), **2000**.
- [11] Wold, S.; Eriksson, L. In: *Statistical Validation of QSAR Results. Chemometric Methods in Molecular Design*; Van de Waterbeemd, H., Ed.; VCH: New York, **1995**, Vol. 2, pp. 309-318.
- [12] Dixon, W.J.; Brown, M.B.; Engelman, L.; Jennrich, R.I. *BMDP Statistical Software Manual*; University of California Press: San Francisco, **1990**.
- [13] Galvez, J.; Garcia-Domenech, R.; de Gregorio Alapont, C.; de Julian-Ortiz, J.V.; Popa, L. Pharmacological distribution diagrams: a tool for de novo drug design. *J. Mol. Graph.*, **1996**, *14*(5), 272-276.
- [14] Mosmann, T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods*, **1983**, *65*, 55-63.
- [15] Tarpey, M.; Wink, D.; Grishman, M. Methods for detection of reactive metabolites of oxygen and nitrogen: *in vitro* and *in vivo* considerations. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, **2004**, *286*, 431-444.
- [16] Sugishita, E.; Amagaya, S.; Ogihara, Y. Studies on the mechanism of antiinflammatory activities of papyriogenin A and papyriogenin C. *J. Pharmacobiodyn.*, **1983**, *6*, 287-294.
- [17] De Young, L.M.; Kheifets, J.B.; Ballaron, S.J.; Young, J. M. Edema and cell infiltration in the phorbol ester- treated mouse ear are temporally separate and can be differentially modulated by pharmacologic agents. *Agents Actions*, **1989**, *26*, 335-341.
- [18] Giner, R.M.; Villalba, M.L.; Recio, M.C.; Máñez, S.; Cerdá-Nicolás, M.; Ríos, J.L. Antiinflammatory glycoterpenoids from *Scrophularia auriculata*. *Eur. J. Pharmacol.*, **2000**, *389*, 243-252.
- [19] Benanti, G. Lysine salt of 4-diphenylacetic acid and its ophthalmic formulations. *U.S. Patent* 4,797,423, January 10, 1989.

## CAPÍTULO 2

### **Identificación de nuevos agentes protectores frente a colitis ulcerosa por Topología molecular:**

Artículo 2: “**Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B**” (En: Mol Divers **2011**, 15:917–926). JCR [2011, CAT: Chemistry Applied; FI: 3,153; Rank: 9 de 71 (**Q1**)].

Artículo 3: “**Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW264.7 cell lines**” (En: Mol Divers **2013**, 17(3):573–593). JCR [2013, CAT: Chemistry Applied; FI: 2,544; Rank: 15 de 71 (**Q1**)].

Artículo 4: “**Molecular topology: A strategy to identify novel compounds against ulcerative colitis.**” (En: Mol Divers **2016**, XXX). Enviado.





# Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B

María Gálvez-Llompert · María C. Recio ·  
Ramón García-Domenech

Received: 15 April 2011 / Accepted: 13 June 2011 / Published online: 30 June 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Ulcerative colitis and Crohn's disease are chronic, immune-mediated inflammatory diseases of the gastrointestinal tract. Nuclear Factor Kappa B (NF- $\kappa$ B) is a transcription factor that plays a key role in regulating expression of multiple inflammatory and immune genes. In this study, a Topological Virtual Screening study has been carried out to achieve a model capable of finding new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B. Different topological indices were used as structural descriptors, and their relation to biological activity was determined using linear discriminant analysis. A topological model consisting of two discriminant functions was built up. The first function focused in the discrimination between NF- $\kappa$ B active and inactive compounds, and the second one in distinguishing between compounds active and inactive on ulcerative colitis. The model was then applied sequentially to a large database of compounds with unknown activity. Twenty-eight of such compounds were predicted to be active and selected for *in vitro* and *in vivo* testing.

**Keywords** NF- $\kappa$ B · Inflammatory bowel disease · Ulcerative colitis · Linear discriminant analysis · Molecular topology

**Electronic supplementary material** The online version of this article (doi:10.1007/s11030-011-9323-4) contains supplementary material, which is available to authorized users.

M. Gálvez-Llompert · R. García-Domenech (✉)  
Molecular Connectivity and Drug Design Research Unit,  
Department of Physical Chemistry, Faculty of Pharmacy, University  
of Valencia, Avda. V.A. Estellés, s/n, 46100 Burjassot, Valencia, Spain  
e-mail: ramon.garcia@uv.es

M. Gálvez-Llompert · M. C. Recio  
Department of Pharmacology, Faculty of Pharmacy,  
University of Valencia, Burjassot, Valencia, Spain

## Abbreviations

NF- $\kappa$ B	Nuclear factor kappa B
IBD	Inflammatory bowel disease
UC	Ulcerative colitis
CD	Crohn's disease
TNF- $\alpha$	Tumor necrosis factor- $\alpha$
IL	Interleukin
LDA	Linear discriminant analysis
HTS	High throughput screening
QSAR	Quantitative structure–activity relationship
DF	Discriminant function
TI	Topological indices
PDD	Pharmacological distribution diagram

## Introduction

Inflammatory bowel disease, (IBD), is an immune-mediated chronic and relapsing inflammatory disease caused by the inflammation and sores in the lining of large intestine. The IBD is clinically characterized by recurrent episodes of diarrhea, rectal bleeding, cramping abdominal pain, and histologically characterized by mucosal inflammation and injury [1]. Ulcerative colitis (UC) and Crohn's disease (CD) are the two major types of IBD. The IBD may potentially cause lifelong disability and is associated with an increased risk of colorectal cancer [2].

The incidence and prevalence of UC and CD are beginning to stabilize in high-incidence areas such as northern Europe and North America, although they continue to rise in low-incidence areas such as southern Europe and Asia. Even so 1.4 million persons in the United States and 2.2 million persons in Europe suffer from these diseases [3].

Conventional therapy for UC includes sulfasalazine and other 5-aminosalicylic acid type of compounds, and in more

persistent and/or severe cases, corticosteroids and immunosuppressants are administered [1]. These medications alleviate inflammation and reduce symptoms, but do not provide a cure or prevent long-term complications. All of these have also significant side effects that limit their clinical application [4]. Hence, it is still a challenge to develop novel therapies for IBD.

Several agents used in the management of IBD, such as corticosteroids, sulfasalazine, and 5-aminosalicylic acid, have documented regulation of NF- $\kappa$ B function [1]. Extensive experimental studies have shown that a variety of the pro-inflammatory mediators associated with IBD [including tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), interleukin (IL)-1 $\beta$ , IL-6, and IL-8] [5], are regulated at the transcriptional level by the NF- $\kappa$ B [4].

Using immunohistochemistry methods in the inflamed intestinal mucosa of IBD patients provide evidence that NF- $\kappa$ B activation is significantly elevated in the intestinal mucosa of both CD and UC patients [6, 7].

The NF- $\kappa$ B is a major transcription regulator involved in a variety of immunological and/or inflammatory processes [8]. Conceivably, genes involved in the NF- $\kappa$ B pathway are to be interesting candidates for chronic inflammatory disorders, including UC [9].

Furthermore, Virtual Screening, or in silico screening, is an approach attracting increasing interest in the pharmaceutical industry as a productive and cost-effective technology for the search of novel *hit* or *lead* compounds [10, 11].

The principles involve the computational analysis of chemical databases, to identify those compounds that are most likely to show a given biological activity. Of course, these ideas are not new, but have been pursued for years by groups working in drug design and discovery. However, the availability of inexpensive high-performance computing platforms has transformed these processes in such a way that, at present, increasingly complex and more accurate analyses can be performed on very large data sets. The topological virtual screening is based on the analysis of a chemical diversity of molecules [12], which enables the selection of the best potential molecular choices. In principle, the molecules are not classified according to their biological activity, but depicted by their topological indices (TIs) [13] and after a computational study of their structures, only those ones complying with a desired topological model are chosen for further development.

The model above comes from a linear discriminant analysis (LDA) containing two sets of structures: One of them has a well-defined pharmacological activity. The other one is built from structures showing other biological activities different to the one for which the model is created.

The resulting model above, associated with the desired pharmacological activity, generates a set of topological

descriptors capable of differentiating potentially active compounds from those lacking activity.

The method above represents a rather detailed and relevant framework to search for leads, prioritizing the selection of compounds that are advisable to be tested in a biological assay. It offers a new option, a new method that shows itself powerful to face the hunting for new genes, new targets, new lead compounds that finally enable the securing of new drugs.

In the late 1980s, computational chemistry allowed the speed up of the drug discovery process. Afterward, combinatorial chemistry (including molecular evolution, multiple parallel synthesis, etc.) arrived combined with High Throughput Screening (HTS), in the mid-1990s.

The technologies above have partially delivered their promise but not in the manner initially expected by enthusiastic supporters [14].

This report deals with the search for new leads using a database of compounds with anti-UC activity and NF- $\kappa$ B inhibition activity. Furthermore, an aspect that makes us optimistic with the perspectives of this study is the research team's expertise in discovering new drugs applying Molecular Topology. In fact, literature sources support that assertion: Anticonvulsant [15], antimalarial [16], antimicrobial [17, 18], antifungal [19], antineoplastic [20], antihistaminic [21], bronchodilator agents [22], cytostatics [23], and anti-inflammatory compounds [24].

It should be pointed out that the topological method applied to NF- $\kappa$ B inhibition and anti-UC are used here for the first time to analyze the chemical diversity in this field. In addition, the method is significantly faster and less expensive to apply than other formalisms such as those based on molecular mechanics, quantum chemistry, similarity/dissimilarity approaches, or 3D-QSAR, just to mention some examples.

## Materials and methods

### Analyzed compounds

The model for searching new compounds with NF- $\kappa$ B inhibition activity was made up of 121 compounds (67 active as a NF- $\kappa$ B [8, 25] and 54 inactive). The inactive compounds were extracted from the Merck Index.

Moreover, the absence of reported activity for any compound belonging to the inactive group does not mean necessarily lack of activity but lack of experimental data on the activity for that compound. However, since the dataset of inactive compounds is significantly large, the risk of uncertain classification becomes negligible.

Furthermore, another model was developed for discovering new compounds potentially active in the treatment of UC. In this case, the database consisted of 25 compounds known to be active in UC. Among them stand those used in

the conventional therapy for UC, including 5-aminosalicylic acid compounds, corticosteroids, and immunosuppressants.

The compounds of the set are shown in Tables 1, 2, 3, and 4. All sets are characterized by a large structural diversity.

#### Molecular descriptors

The 2D structure of each compound was drawn using the ChemDraw Ultra package (version 10.0). Each compound was characterized by a set of TIs, standing among them the topological charge indices [26], quotients and differences between nonvalence and valence connectivity indices [27–29], topological descriptors [30], eigenvalue-based indices, and 2D autocorrelations descriptors. Other graph-theoretical descriptors were also used but are not depicted herein because of their lack of effectiveness. All descriptors were calculated with Desmol11 program [31] and Dragon software, version 5.4. [32]. In the supplementary material, the TIs are represented for all compounds used in this study.

#### Modeling techniques

LDA is a pattern recognition method which provides a classification model based on the combination of variables that best predict the category or group to which a given compound belongs. Database compounds were allocated to active and inactive groups according to their NF- $\kappa$ B inhibitory activity or their anti-UC activity. The LDA was then applied to these two groups to obtain two discriminant functions, DF1 and DF2. The compounds were classified either as active or inactive group as for their NF- $\kappa$ B inhibitor and anti-UC activities, respectively. The independent variables were the TIs, and the discriminatory property was the activity. The discriminant capability was assessed as the percentage of correct classifications in each set of compounds.

The classification criterion was the minimal Mahalanobis distance (distance of each case to the mean of all the cases in a category). The quality of the discriminant function was evaluated using the Wilks parameter,  $\lambda$ , which was obtained by multivariate analysis of variance that tests the equality of group means for the variable in the discriminant model. The method used to select the descriptors was based on the Fisher–Snedecor parameter ( $F$ ), which determines the relative importance of candidate variables. The software used for the LDA study was the BMDP New System 2 package, module 7M [33]. The variables used to compute the linear classification function are chosen in a stepwise manner: at each step, the variable that makes the largest contribution to the separation of the groups is entered into the discriminant equation (or the variable that makes the smallest contribution is removed).

The validation of the selected functions was done using an external test set.

#### Pharmacological-activity distribution diagrams

A pharmacological distribution diagram (PDD) is a graphical representation that provides a straightforward way of visualizing the regions of minimum overlap between active and inactive compounds, as well as the regions in which the probability of finding active compounds is at maximum [34]. Actually, a PDD is a frequency distribution diagram of dependent variables in which the ordinate represents the expectancy (probability of activity) and the abscissa represents the DF values in the range. For an arbitrary range of values of a given function, an “expectancy of activity” can be defined as  $E_a = a/(i+1)$ , where “ $a$ ” is the number of active compounds in the range divided by the total number of active compounds and “ $i$ ” is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity is defined in a symmetrical way, as  $E_i = i/(a+1)$ . Upon these diagrams, it is easy to visualize the intervals in which there is a maximum probability of finding new active compounds and a minimum probability of finding inactive compounds.

#### Topological virtual screening

The topological models resulting from DF1 and DF2 functions were used to find new NF- $\kappa$ B compounds active versus UC. Two databases, namely Merck Index and the Micro-Source Pure Natural Products Collection, were screened for the search of potential new drugs.

## Results and discussion

#### Mathematical modeling

The mathematical model achieved was developed from a training set of 95 compounds whose activities as NF- $\kappa$ B inhibitors had previously been determined. The training set was comprised of heterogeneous molecular structures with a similar number of active and inactive compounds.

This first equation, corresponding to DF1, distinguished compounds that were predicted to have NF- $\kappa$ B inhibitory activity. This equation comprised five independent variables:

$$DF1 = 0.633G_1 + 24.29^0C - 3.62^2D + 2.09^4Dp + 0.288V_3 - 30.94 \quad (1)$$

Statistical data and parameters accounting for the significance of this equation were as follows:

$$N = 95, F = 9.5, \text{ and } \lambda = 0.354.$$

**Table 1** Compounds used in the training set and results obtained applying the linear discriminant analysis (Eq. 1) to NF- $\kappa$ B inhibitors

Active group				Inactive group			
Compound	DF1	Prob (activ.) %	Class.	Compound	DF1	Prob (activ.) %	Class.
Andalusol	0.09	52	NC	Albocyclin	-1.89	13	I
Aspirin	-1.07	25	I	Allomethadione	-1.72	15	I
Biflavonoid-2',8''-biapigenin	4.12	98	A	Alonimid	-1.35	21	I
Bortezomib	0.58	64	A	Anisindione	-0.78	31	I
Cardamomin	1.64	84	A	Bensalan	-1.51	18	I
Celastrrol	2.73	94	A	Carbutamide	-1.81	14	I
Chelerythrine	-1.63	16	I	Cathine	-1.85	14	I
Curcumin	1.44	81	A	Clofexamide	-3.37	3	I
Cycloepoxydone	0.89	71	A	Cyclopentobarbital	-0.65	34	I
Dexamethasone	2.39	92	A	Delta-Tocopherol	-0.18	45	NC
DHMEQ	2.77	94	A	2',3'-Dideoxyadenosine	-0.44	39	I
Dihydroeponemycin	1.23	77	A	Diethylpropion	-3.23	4	I
1,3-diene prostaglandin E1	-1.85	14	I	Dihydrobenzofuranol	-0.07	48	NC
3,3'-Diindolylmethane	-0.93	28	I	Dihydroxyvitamin D3	-1.08	25	I
EGCG	3.00	95	A	Epiestriol	0.56	64	A
Ent-kaur-16-en-19 oic acid	0.73	68	A	Etiproston	-0.41	40	I
Eponemycin	1.27	78	A	Farrerol	0.79	69	A
Epoxomycin	1.98	88	A	Flumethasone	3.03	95	A
Epoxyquinomycin C	4.11	98	A	Hexenylsalicylate	-1.94	13	I
Flavopiridol	1.35	79	A	Ichthyone	-1.06	26	I
Foliol	1.71	85	A	Idramantone	-1.28	22	I
Halofuginone	-1.00	27	I	Isoxepac	-0.47	38	I
Helenaline	0.81	69	A	Lampterol	-0.03	49	NC
Hypoestoxide	1.18	77	A	L-Synephrine	-1.11	25	I
Indole-3-carbinol	-1.38	20	I	Lurtotecan	-1.21	23	I
Isodeoxyelephantopin	0.38	59	A	Mebutamate	-1.72	15	I
Kamebacetal A	2.65	93	A	Meobentine	-2.58	7	I
Kamebakaurin	3.44	97	A	(S)-Meptazinol	-2.25	10	I
Kamebarin	1.79	86	A	Metazocine	-1.51	18	I
Lactacystin	0.96	72	A	Nabilone	-0.32	42	I
Magnolol	0.02	50	NC	Naftopidil	-0.64	34	I
Mesalamine	1.80	86	A	Noreximide	-0.99	27	I
Morin	2.79	94	A	Norletimol	-0.86	30	I
Nitecapone	0.51	63	A	Omephin	0.09	52	NC
Panepoxydone	1.22	77	A	Oxepinac	-0.40	40	I
Parthenolide	-0.70	33	I	Pentizidone	-1.70	15	I
Piceatannol	1.99	88	A	Phanquone	-1.06	26	I
Prednisolone	1.60	83	A	Retinol	-0.31	42	I
Pristimerin	2.65	93	A	Subathizone	-1.90	13	I
Prostaglandin E1	-2.18	10	I	Tetrahydrobiopterin	-0.24	44	NC
Quercetin	2.52	93	A	Tiflorex	-3.60	3	I
Rengyolone	-1.36	20	I	Trepibutone	-4.03	2	NC
Resveratrol	1.27	78	A	Velaresol	-0.90	29	I
Sanguinarine	-1.87	13	I	Win 44441	-0.31	42	I

**Table 1** continued

Active group				Inactive group			
Compound	DF1	Prob (activ.) %	Class.	Compound	DF1	Prob (activ.) %	Class.
Siegeskaurolic acid	0.44	61	A				
Solasodine	1.53	82	A				
Sulfasalazine	1.14	76	A				
Sulindac	0.54	63	A				
Synalar	2.89	95	A				
Temozolomide	−0.50	38	I				
Tomatidine	1.43	81	A				

A active; I inactive; NC non classified

**Table 2** Compounds used in the training set and results obtained applying the linear discriminant analysis (Eq. 2) to Ulcerative colitis active drugs

Active group				Inactive group			
Compound	DF2	Prob (activ.) %	Class.	Compound	DF2	Prob (activ.) %	Class.
Azathioprine	1.62	84	A	Aliskiren	−3.34	4	I
Balsalazide	3.75	98	A	Alprazolam	−0.22	46	NC
Beclomethasone	0.93	73	A	Arformoterol	−5.31	1	I
Budesonide	3.14	96	A	Clevudine	−3.68	3	I
Ciprofloxacin	−1.36	22	I	Fluticasone	−0.71	33	I
Clarithromycin	6.50	100	A	Garenoxacin	−3.15	4	I
Cyclosporine	7.71	100	A	Imidafenacin	−9.55	0	I
Dexamethasone	4.00	98	A	Maraviroc	−6.44	0	I
Mercaptopurine	5.66	100	A	Medigoxin	−6.72	0	I
Mesalamine	3.43	97	A	Mifepristone	1.01	74	A
Methotrexate	4.18	99	A	Paliperidone	−0.50	39	I
Metronidazole	5.68	100	A	Raltegravir	−5.13	1	I
Prednisone	2.42	92	A	Retapamulin	−9.38	0	I
Sulfasalazine	4.12	99	A	Testosterone	−1.79	15	I
Tacrolimus	4.86	99	A	Tocopherols	−4.79	1	I
				Vildagliptin	−2.24	10	I

The TIs in Eq. 1 are:

$G_1$  = Topological charge index of order 1;  ${}^0C = {}^0\chi / {}^0\chi^v$  = quotient of connectivity indices of order 0;  ${}^2D = {}^2\chi - {}^2\chi^v$  = difference of connectivity indices of order 2;  ${}^4Dp = {}^4\chi_p - {}^4\chi_p^v$  = difference of connectivity indices of order 4 and type path;  $V_3$  = number of vertices with topological valence 3.

In the equation above, there are topological descriptors evaluating the molecular branching,  $V_3$ , the distribution of the intramolecular charge (charge index  $G_1$ ), molecular volume ( ${}^0C$ ), and topological docking of each compound ( ${}^2D$  and  ${}^4Dp$ ).

According to Eq. 1, a compound should be classified as active if  $DF1 > 0$ , otherwise it is labeled as inactive.

By applying this criterion to the DF1 training set (95 compounds), (see table 1), 40 out of 51 experimentally active compounds were correctly classified as such (78% accuracy),

and 40 out of 44 experimentally inactive compounds were also well classified (91% accuracy).

To calculate the percentage of correctly classified compounds within a particular category (active or inactive), the following formula was used:

$$\text{Classification accuracy(\%)} = (\text{CCC} * 100 / \text{TNC}),$$

where CCC is correctly-classified compounds and TNC is a total number of compounds.

Altogether, the average of correct classification for the entire set of compounds (active plus inactive) was 84%.

To establish the adequate range of activity, we analyzed the pharmacological distribution diagram obtained with DF1.

Looking at Fig. 1, we can appreciate that all the compounds studied show DF1 values in the range  $4.5 > DF1 > -3.8$ . In addition, we can realize that there is little overlap-

**Table 3** Compounds used in the test set and results obtained applying linear discriminant analysis (Eq. 1) to NF- $\kappa$ B inhibitors

Active group				Inactive group			
Compound	DF1	Prob (activ.) %	Class.	Compound	DF1	Prob (activ.)%	Class.
5,7-DHMC	0.84	70	A	Benzoic acid ion(1-)	-1.16	24	I
7,8-DHMC	0.28	57	NC	Eniluracil	-1.16	24	I
AIDS-006682	0.45	61	A	Glycidyl methacrylate	-2.17	10	I
Artemisinin	-0.16	46	NC	Guaiacol	-1.64	16	I
Aurine	1.38	80	A	Icofungipen	0.08	52	NC
Aurintricarboxylic	2.86	95	A	Isoleucine	-0.91	29	I
beta-D-Galactopyranoside	8.55	100	NC	<i>N</i> -Ethoxy- <i>N</i> -ethylthiocarbamic acid	-5.82	3	NC
BMCSB7i	0.29	57	NC	Octreotide	6.50	100	NC
Gallic acid	3.48	97	A	Scillaren A	3.60	97	A
LicochalconeA	0.89	71	A	Trimethadione	-2.12	11	I
Lucidone	1.79	86	A				
Oleandrin	1.68	84	A				
Poncirin	3.05	96	A				
Pyrrolidine dithiocarbamic acid	-5.94	3	NC				
Sodium Salicylate	0.79	69	A				
Theaflavin digallate	5.34	100	NC				

**Table 4** Compounds used in the test set and results obtained applying the linear discriminant analysis (Eq. 2) to drugs active on ulcerative colitis

Active group				Inactive group			
Compound	DF2	Prob (activ.) %	Class.	Compound	DF2	Prob (activ.) %	Class.
Apocynin	-10.89	0	I	Acetylcysteine	-2.89	6	I
Boldine	2.28	91	A	Almitrine	1.30	80	A
Curcumine	-7.18	0	I	Ambroxol	-2.87	6	I
Ellagic acid	5.52	100	A	Clomipramine	-1.74	16	I
Hydrocortisone	2.42	92	A	Dobutamine	-6.98	1	I
Ipsalazide	1.70	85	A	Dopamine	-15.8	0	NC
Methylprednisolone	2.87	95	A	Inamrinone	-13.3	0	NC
Naringine	8.11	100	A	Lisdexamfetamine	-16.3	0	NC
Olsalazine	7.77	100	A	Milrinone	-9.70	0	I
Sirolimus	3.25	97	A	Rufinamide	-6.02	0	I
Tixocortol	3.39	97	A	Tretinoin	-3.45	3	I

ping between compounds with probability of activity above 40% and below 60%. These percentages match the values in DF1 (Eq. 1) in the range  $0.4 > DF1 > -0.4$ . Therefore, a compound will be selected as NF- $\kappa$ B inhibitor if  $4.5 > DF1 > 0.4$  and as inactive if  $-0.4 > DF1 > -3.8$ . Outside these ranges the compound's classification is uncertain and it is labeled as "not-classified" (outliers), NC.

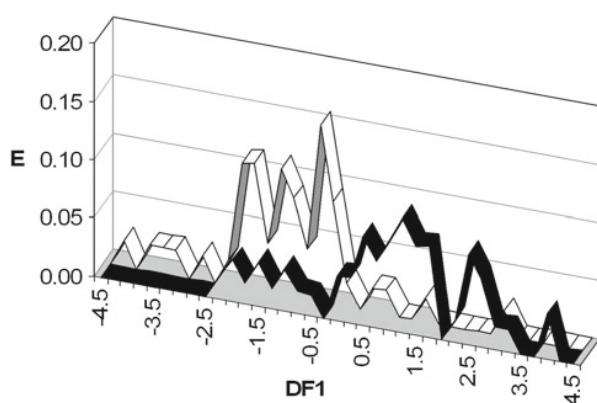
An easy way to evaluate the quality of the function above is to apply it into an external group. In our case, this group was made up of 26 compounds (16 actives and 10 inactives) which have not been included in the discriminant analysis, which is about 20% of the data. Table 3 shows the results of classification obtained for each test compound.

Although the success rate decreases in the active group, 63% (10 of 16 compounds analyzed are correctly classified), in the inactive group there is only one compound misclassified, indicating that DF1 function is highly sensitive in recognizing active compounds. Hence, we can ensure that the number of "false active" compounds is going to be minimum.

The second equation, DF2, was employed to discriminate compounds showing a profile of active in UC. Four independent variables composed DF2, as follows:

$$DF2 = 3.15SEige - 10.28GATS6p - 261.66 X3A - 0.06 D/Dr 05 + 52.73 \quad (2)$$





**Fig. 1** Pharmacological distribution diagram for NF- $\kappa$ B inhibitors obtained using the discriminant function DF1. (The black color represents the compounds with inhibitor activity and the gray color, the compounds without it)

The statistical data and parameters were:

$$N = 31, F = 13.3, \text{ and } \lambda = 0.328.$$

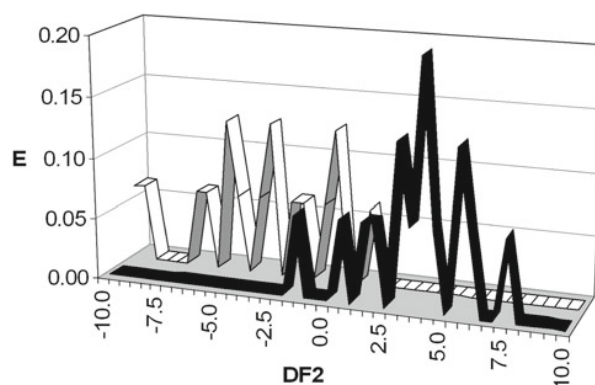
The meaning of the TIs are:

SEige = Eigenvalue sum from electronegativity weighted distance matrix. The term “SEige” is a measure of the strength of bonds in a molecule. GATS6p = Geary autocorrelation - lag 6/weighted by atomic polarizabilities. Geary indices are a measure of spatial autocorrelation, i.e., it measures how much similar are the molecules each other as for their atomic polarizabilities. X3A = Average connectivity index chi-3. It is a simple structural index evaluating the contributions of all atoms at topological distance 3, independently of their environment. D/Dr05 = Distance/detour ring index of order 5.

By employing this function, a compound was considered to be active if DF2 is  $> 0$ , whereas it is labeled as inactive, if DF2 is  $< 0$ . The interval of confidence is 95%. Therefore, if we see Table 2, 14 out of the 15 active compounds and 15 out of the 16 inactive compounds were correctly classified, which imply 93 and 94% correct classification, respectively.

The corresponding PDD for this equation (Fig. 2) shows that all studied compounds present DF2 values in the range  $8 > \text{DF2} > -10$ . There is little overlap between compounds with probability active above 40% and below 60%, which matches the values in DF2 (Eq. 2) within the range  $0.89 > \text{DF2} > -0.49$ . Hence, a compound is selected as active if it lies in the range:  $8 > \text{DF2} > 0.89$  and as inactive if the range is  $-0.49 > \text{DF2} > -10$ . Outside these ranges, the analyzed compound will be considered as uncertain (NC).

Although it is not easy to unfold the structural features explaining the discrimination above, some insight can be gained on the basis of the most relevant index, namely D/Dr05.



**Fig. 2** Pharmacological distribution diagram for ulcerative colitis active drugs obtained using the discriminant function DF2. (Black color represents the compounds with antiulcerative colitis activity and the gray color, the compounds without it)

The index above is related to the longest paths linking the atoms at lag 5. It encodes information about condensed rings and is related to molecule’s cyclicality [35]. In Eq. 2, this index’s regression coefficient shows a negative sign; hence, the larger its value is the lower the activity. Thus, aromatic compounds showing five-member rings with heteroatoms (especially oxygen) should be the least active (for example azathioprine or budesonide among the active or maraviroc and medigoxin among the inactive). On the contrary, those compounds not showing the five-member rings ( $D/\text{Dr}05 = 0$ ) are expected to be more potent, as for instance balsalazide or clarithromycin.

When Eq. 2 is applied to an external validation data set (22 compounds), as illustrated in Table 4, 9 out of 11 active compounds were correctly classified (82% accuracy) and 7 out of 8 experimentally inactive compounds were also well classified (88% accuracy). There are three outliers or uncertain compounds. It is interesting to realize that by means of DF2, there is only one compound misclassified in the inactive group, indicating that the function is highly specific.

In the case above, the test compounds were selected by applying a criterion of structural similarity. For example, if our data contains 7 molecules with structural similarity (glucocorticoids), 3 of them are chosen for the external data set and 4 for the training set, to make the training and the test sets comparable.

It must be emphasized the significant role played by connectivity and topological charge indices in the different achieved discriminant functions. Just watching DF1 and DF2, it may be noticed that structural and electronic properties of compounds determine their activity in both, UC and NF- $\kappa$ B inhibition. Given the high structural heterogeneity of the molecules used to build up the models, they can be applied into large databases including variegated compounds.

**Table 5** Values of DF1, DF2, probability of activity, and classification of the potential anti UC compounds, selected from the Merck Index database and the MicroSource Pure Natural Products Collection

Compound	DF1	Prob (activ.) %	Class.	DF2	Prob (activ.) %	Class.	Activity/therapeutics category
Aconitic acid	2.90	95	A	2.97	95	A	Anti-inflammatory [36]
Ajmaline diacetate	1.16	77	A	4.86	99	A	
Alizarin-3-methylinodiacetic acid	1.86	87	A	4.33	99	A	Red staining [37]
Apigenin	1.64	84	A	2.6	94	A	Anti-inflammatory [38]
Calcein	2.21	90	A	13.9	100	NC	Fluorescent dye [39]
Carapin	2.51	93	A	3.04	96	A	
Cromolyn sodium	0.33	58	NC	1.19	78	A	Anti-inflammatory [40]
(+)-Dibenzyl L-tartrate	0.62	65	A	-12.7	0	NC	
Emodic acid	2.69	94	A	4.44	99	A	
Evernic acid	1.42	81	A	7.09	99	A	
Fissinolide	2.38	92	A	4.39	99	A	
Folic acid	1.65	84	A	3.9	98	A	
Genkwanin	0.57	64	A	2.59	93	A	Anti-inflammatory [41]
Haematommic acid	2.22	90	A	5.62	100	A	Antioxidant [42]
Hesperidin	2.20	90	A	6.08	100	A	Anti-inflammatory [43]
10-Hydroxycamptothecin	0.37	59	NC	2.57	93	A	Anti-inflammatory [44]
Lonchocarpic acid	1.06	75	A	2.49	93	A	Antimicrobial [45]
3-Methylorsellinic acid	0.99	73	A	6.65	100	A	
Methyl orsellinate	0.14	54	NC	6.38	100	A	Anti-inflammatory [46]
2-Methoxyresorcinol	0.31	58	NC	1.94	88	A	
Physcion	1.17	76	A	2.91	95	A	
Purpurin	1.79	86	A	1.59	84	A	Anti-inflammatory [47]
Pyrocatechuic acid	2.21	90	A	3.24	97	A	
Ro 41-0960	1.75	85	A	-0.46	40	I	
Rosmarinic acid	1.99	88	A	-8.03	0	I	Anti-inflammatory [48]
Rotenone	0.26	57	NC	1.53	83	A	
Ursolic acid	2.66	94	A	3.33	97	A	Anti-ulcer [49]
Uvaol	2.98	95	A	3	96	A	Anti-inflammatory [50]

Compounds' therapeutic profile from the literature are also included

### Topological virtual screening

Based on the models described above, a virtual screening was carried out on a database of heterogeneous drug molecules. A library (MicroSource Pure Natural Products Collection) consisting of 800 natural products and the Merck index database (about 12,000 compounds) were screened. The composition of the library can be obtained from the MicroSource Discovery Systems website (<http://www.msdiscovery.com>).

The model above predicted that 28 compounds might be active against UC by NF- $\kappa$ B inhibition. Almost all of these are commercially available and are selected for future in vitro and in vivo testing, to assess the model's predictive capability. Table 5 illustrates the DF1 and DF2 values and the classification of each compound on the basis of the PDD diagrams. Five compounds (cromolyn sodium, rotenone, 10-hydroxy-

camptothecin, methylorsellinate, and 2-methoxyresorcinol) were classified as active by DF2 but not by DF1; two compounds (rosmarinic acid and Ro 41-0960) were classified as being active by DF1 but inactive by DF2 and two more compounds (calcein and (+)-dibenzyl L-tartrate) were classified as active by DF1 but as outliers by DF2. The compounds selected were those passing at least one of the filters, either the NF- $\kappa$ B inhibition or the UC drug treatment profile, because both represent a good choice in UC treatment.

As shown in the Table 5, most of the compounds selected had been described previously as anti-inflammatory in the literature (10-hydroxycamptothecin, purpurin, physcion, methyl orsellinate, aconitic acid, genkwanin, uvaol, cromolyn sodium, hesperidin, and rosmarinic acid) and one of them, namely ursolic acid, is described to show also anti-ulcerative properties.



The outcome above is highly encouraging to follow up applying the models described herein to the search of new/novel molecules active in UC.

Virtual screening is increasingly gaining acceptance in the pharmaceutical industry as a cost-effective and timely strategy for analyzing very large chemical data sets. This procedure is computationally intensive for analyzing large databases and it provides the most detailed basis for determining which compounds are likely to be potent hits or leads. The results outlined here demonstrate not only that the Topological Virtual Screening could accurately reproduce the well-known pharmacological activity, but that it also enables discriminating between reference drugs and random compounds, on the basis of a predicted DF value.

Our results herein represent a new step forward in the pathway to demonstrate the high efficiency of the *in silico* methods based in Molecular Topology.

## Conclusions

A combination of topological-structural description of compounds and statistical treatment by discriminant analysis is very efficient in the selection of new compounds effective against UC. The obtained mathematical models can readily be applied to the search of new drugs in large databases or even to the *de novo* drug design. These results confirm the usefulness of molecular topology as a powerful tool in the search for new drug candidates.

**Acknowledgements** We thank the Ministerio de Educación e Innovación, Spain (projects SAF2009-13059-C03-01 and SAF2009-13059-C03-02) for support of this study. The authors want to express their acknowledgements to Prof. Jorge Gálvez from the University of Valencia, for his help and useful comments.

## References

- Venkataranganna M, Rafiq M, Gopumadhavan S, Peer G, Babu U, Mitra S (2007) NCB-02 (standardized curcumin preparation) protects dinitrochlorobenzene-induced colitis through down-regulation of NF- $\kappa$ B and iNOS. *World J Gastroenterol* 13:1103–1107
- Munkholm P (2003) Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. *Aliment Pharmacol Ther* 18:1–5
- Loftus EV (2004) Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology* 126:1504–1517. doi:10.1053/j.gastro.2004.01.063
- Ukil A, Maity S, Das P (2006) Protection from experimental colitis by theaflavin-3, 3'-digallate correlates with inhibition of IKK and NF- $\kappa$ B activation. *Br J Pharmacol* 149:121–131. doi:10.1038/sj.bjp.0706847
- Lee JY, Kim JS, Kim JM, Kim N, Jung HC, Song IS (2007) Simvastatin inhibits NF- $\kappa$ B signaling in intestinal epithelial cells and ameliorates acute murine colitis. *Int Immunopharmacol* 7:241–248. doi:10.1016/j.intimp.2006.10.013
- Karban AS, Okazaki T, Panhuysen CIM et al (2004) Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. *Hum Mol Genet* 13:35–45. doi:10.1093/hmg/ddh008
- Jian YT, Mai GF, Wang J, Zhang Y, Luo R, Fang Y (2005) Preventive and therapeutic effects of NF- $\kappa$ B inhibitor curcumin in rats colitis induced by trinitrobenzene sulfonic acid. *World J Gastroenterol* 11:1747–1752
- Ríos J, Recio M, Escandell J, Andújar I (2009) Inhibition of transcription factors by plant-derived compounds and their implications in inflammation and cancer. *Curr Pharm Des* 15:1212–1237
- Born M, Bodegraven A, Mulder C, Kraal G, Bouma G (2005) A NFKB1 promoter polymorphism is involved in susceptibility to ulcerative colitis. *Int J Immunogenet* 32:401–405. doi:10.1111/j.1744-313X.2005.00546.x
- Xu J, Hagler A (2002) Chemoinformatics and drug discovery. *Molecules* 7:566–600. doi:10.3390/70800566
- Oprea TI (2002) Virtual screening in lead discovery: a viewpoint. *Molecules* 7:51–62. doi:10.3390/70100051
- Amigo JM, Galvez J, Villar VM (2009) A review on molecular topology: applying graph theory to drug discovery and design. *Naturwissenschaften* 96:749–761. doi:10.1007/s00114-009-0536-7
- Garcia-Domenech R, Galvez J, de Julian-Ortiz JV, Pogliani L (2008) Some new trends in chemical graph theory. *Chem Rev* 108:1127–1169. doi:10.1021/cr0780006
- Horrobin DF (2000) Innovation in the pharmaceutical industry. *J R Soc Med* 93:341–345
- Bruno-Blanch L, Galvez J, Garcia-Domenech R (2003) Topological virtual screening: a way to find new anticonvulsant drugs from chemical diversity. *Bioorg Med Chem Lett* 13:2749–2754. doi:10.1016/S0960-894X(03)00535-3
- Mahmoudi N, Garcia-Domenech R, Galvez J et al (2008) New active drugs against liver stages of plasmodium predicted by molecular topology. *Antimicrob Agents Chemother* 52:1215–1220. doi:10.1128/AAC.01043-07
- Mishra RK, Garcia-Domenech R, Galvez J (2001) Getting discriminant functions of antibacterial activity from physicochemical and topological parameters. *J Chem Inf Comput Sci* 41:387–393. doi:10.1021/ci000303c
- de Gregorio Alapont C, Garcia-Domenech R, Galvez J, Ros MJ, Wolski S, Garcia MD (2000) Molecular topology: a useful tool for the search of new antibacterial. *Bioorg Med Chem Lett* 10:2033–2036. doi:10.1016/S0960-894X(00)00406-6
- Garcia-Domenech R, Catala-Gregori A, Calabuig C, Anton-Fos G, del Castillo L, Galvez J (2002) Prediction of antifungal activity: a computational screening using topological descriptors. *Internet Electron J Mol Des* 1:339–350
- Jasinski P, Welsh B, Galvez J, Land D, Zwolak P, Ghandi L, Terai K, Dudek AZ (2008) A novel quinoline, MT477: suppresses cell signaling through ras molecular pathway, inhibits PKC activity, and demonstrates *in vivo* anti-tumor activity against human carcinoma cell lines. *Investig New Drug* 26:223–232. doi:10.1007/s10637-007-9096-x
- Duart MJ, Garcia-Domenech R, Galvez J, Aleman PA, Martin-Algarra RV, Anton-Fos GM (2006) Application of a Mathematical Topological Pattern of Antihistaminic Activity for the Selection of New Drug Candidates and Pharmacology Assays. *J Med Chem* 49:3667–3673. doi:10.1021/jm0580555
- Rios-Santamarina I, Garcia-Domenech R, Galvez J, Morcillo Esteban J, Santamaria P, Cortijo J (2004) Getting new bronchodilator compounds from molecular topology. *Eur J Pharm Sci* 22:271–277. doi:10.1016/j.ejps.2004.03.013
- Galvez J, Garcia-Domenech R, Gomez-Lechon MJ, Castell JV (2000) Use of molecular topology in the selection of new

- cytostatic drugs. *J Mol Struct* 504:241–248. doi:[10.1016/S0166-1280\(00\)00365-1](https://doi.org/10.1016/S0166-1280(00)00365-1)
24. Galvez-Llompарт M, Giner M, Recio C, Candeletti S, Garcia-Domenech R (2010) Application of molecular topology to the search of novel NSAIDs: experimental validation of activity. *Lett Drug Des Discov* 7:438–445. doi:[10.2174/157018010791306597](https://doi.org/10.2174/157018010791306597)
  25. Dolinnaya N, Kubareva EA, Kazanova E, Zigangirova N, Naroditsky B, Gintsburg A, Oretskaya TS (2008) Low-molecular-weight inhibitors of NF- $\kappa$ B signalling pathways. *Russ Chem Rev* 77:967–981. doi:[10.1070/RC2008v077n11ABEH003881](https://doi.org/10.1070/RC2008v077n11ABEH003881)
  26. Galvez J, Garcia R, Salabert M, Soler R (1994) Charge indexes. New topological descriptors. *J Chem Inf Comput Sci* 34:520–525. doi:[10.1021/ci00019a008](https://doi.org/10.1021/ci00019a008)
  27. Kier LB, Hall LH (1976) Molecular connectivity in chemistry by drug research. Academic Press, London, pp 46–79
  28. Kier LB, Hall LH (1977) The nature of structure–activity relationships and their relation to molecular connectivity. *Eur J Med Chem* 12:307–312
  29. Kier LB, Hall LH (1986) Molecular connectivity in structure-activity analysis. John Wiley & Sons Inc, New York
  30. Randić M (1977) On canonical numbering of atoms in a molecule and graph isomorphism. *J Chem Inf Comput Sci* 17:171–180. doi:[10.1021/ci60011a013](https://doi.org/10.1021/ci60011a013)
  31. DESMOL11 software, Molecular Connectivity and Drug Design Research Unit, Faculty of Pharmacy, University of Valencia, Spain.
  32. Todeschini R, Consonni V (2009) Molecular descriptors for Chemoinformatics. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
  33. Dixon WJ (1990) BMDP Statistical Software, Univeristy of California, Berkeley
  34. Galvez J, Garcia-Domenech R, de Gregorio Alapont C, de Julian-Ortiz JV, Popa L (1996) Pharmacological distribution diagrams: a tool for de novo drug design. *J Mol Graphics* 14:272–276. doi:[10.1016/S0263-7855\(96\)00081-1](https://doi.org/10.1016/S0263-7855(96)00081-1)
  35. Jain HK, Agrawal RK (2007) Topological descriptors in modeling tumor necrosis factor alpha inhibitory activity of xanthines, pteridinediones and related compounds. *Internet Electron J Mol Des* 6:218–228
  36. Huang L (1985) Pharmacological effect of the organic acid of *achillea alpina*. *Zhong Yao Tong Bao* 10:38–40
  37. McLeod MJ (1980) Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. *Teratology* 22:299–301. doi:[10.1002/tera.1420220306](https://doi.org/10.1002/tera.1420220306)
  38. El Alfy T, El Sawi S, Sleem A, Moawad D (2010) Investigation of flavonoidal content and biological activities of *Chorisia Insignis* Hbk. Leaves. *Aust J Basic Appl Sci* 4:1334–1348
  39. Parish CR (1999) Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunol Cell Biol* 77:499–508
  40. Zhang Y, Reenstra WW, Chidekel A (2001) Antibacterial activity of apical surface fluid from the human airway cell line calu-3: Pharmacologic alteration by corticosteroids and beta(2)-agonists. *Am J Respir Cell Mol* 25:196–202
  41. Pelzer LE, Guardia T, Juarez AO, Guerreiro E (1998) Acute and chronic antiinflammatory effects of plant flavonoids. *Farmacol* 53:421–424. doi:[10.1016/S0014-827X\(98\)00046-9](https://doi.org/10.1016/S0014-827X(98)00046-9)
  42. Marante FJT, Castellano AG, Rosas FE, Aguiar JQ, Barrera JB (2003) Identification and quantitation of allelochemicals from the lichen *lethariella canariensis*: phytotoxicity and antioxidative activity. *J Chem Ecol* 29:2049–2071. doi:[10.1023/A:1025682318001](https://doi.org/10.1023/A:1025682318001)
  43. Benavente-Garcia O, Castillo J (2008) Update on uses and properties of citrus flavonoids: new findings in anticancer, cardiovascular, and anti-inflammatory activity. *J Agric Food Chem* 56:6185–6205. doi:[10.1021/jf8006568](https://doi.org/10.1021/jf8006568)
  44. Cashman J, Burt H, Springate C, Gleave J, Jackson J (2004) Camptothecin-loaded films for the prevention of postsurgical adhesions. *Inflamm Res* 53:355–362. doi:[10.1007/s00011-004-1272-2](https://doi.org/10.1007/s00011-004-1272-2)
  45. Alencar N, Cavalcante C, Vasconcelos M, Leite K, Aragão K, Assrey A, Nogueira N, Cavada B, Vale M (2005) Anti-inflammatory and antimicrobial effect of lectin from *lonchocarpus sericeus* seeds in an experimental rat model of infectious peritonitis. *J Pharm Pharmacol* 57:919–922. doi:[10.1211/0022357056352](https://doi.org/10.1211/0022357056352)
  46. Du H, Matsushima T, Spyvee M, Goto M, Shirota H, Gusovsky F, Chiba K, Kotake M, Yoneda N, Eguchi Y (2009) Discovery of a potent, metabolically stabilized resorcylic lactone as an anti-inflammatory lead. *Bioorg Med Chem Lett* 19:6196–6199. doi:[10.1016/j.bmcl.2009.08.096](https://doi.org/10.1016/j.bmcl.2009.08.096)
  47. Mishchenko N, Fedoreev S, Bryukhanov V, Zverev YF, Lampatov V, Azarova O, Shkryl' YN, Chernoded G (2007) Chemical composition and pharmacological activity of anthraquinones from *rubia cordifolia* cell culture. *Pharm Chem J* 41:605–609. doi:[10.1007/s11094-008-0021-1](https://doi.org/10.1007/s11094-008-0021-1)
  48. Su P, Wang G, Wu D, Sheng X (2008) Progress in rosmarinic acid biological activities and its sources. *Shipin Yu Fajiao Gongye* 34:135–138
  49. Liu J (2005) Oleanolic acid and ursolic acid: research perspectives. *J Ethnopharmacol* 100:92–94. doi:[10.1016/j.jep.2005.05.024](https://doi.org/10.1016/j.jep.2005.05.024)
  50. Marquez-Martin A, Puerta RDL, Fernandez-Arche A, Ruiz-Gutierrez V, Yaqoob P (2006) Modulation of cytokine secretion by pentacyclic triterpenes from olive pomace oil in human mononuclear cells. *Cytokine* 36:211–217. doi:[10.1016/j.cyto.2006.12.007](https://doi.org/10.1016/j.cyto.2006.12.007)

# Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW 264.7 cell lines

María Galvez-Llompарт · María del Carmen Recio Iglesias ·  
Jorge Gálvez · Ramón García-Domenech

Received: 14 March 2013 / Accepted: 8 June 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** Ulcerative colitis (UC) is an immune-mediated chronic and relapsing intestinal inflammatory disease. Interleukin (IL)-6, a pro-inflammatory cytokine, plays a key role in the uncontrolled intestinal inflammatory process, which is a main characteristic of UC. In this work, a quantitative structure–activity relationship model based on molecular topology (MT) has been built up to predict the IL-6 mediated anti-UC activity. After an external validation of the model, a virtual screening of the MicroSource Pure Natural Products Collection and Sigma-Aldrich databases was carried out looking for potential new active compounds. From the entire set of compounds labeled as active by the model, four of them, namely alizarin-3-methylimino-*N,N*-diacetic acid (AMA), Calcein, (+)-dibenzyl-*L*-tartrate (DLT), and Ro 41-0960, were tested *in vitro* by determination of IL-6 production in two cell lines (RAW 264.7 and Caco-2). The results demonstrate that three of them were able to significantly reduce IL-6 levels in both cell lines and particularly one, namely Ro 41-0960. These results confirm MT's efficacy as a tool for the selection of compounds potentially active in UC.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11030-013-9458-6) contains supplementary material, which is available to authorized users.

M. Galvez-Llompарт · J. Gálvez · R. García-Domenech (✉)  
Molecular Connectivity & Drug Design Research Unit,  
Department of Physical Chemistry, Faculty of Pharmacy,  
University of Valencia, Avenida V.A. Estelles s/n, Burjasot,  
46100 Valencia, Spain  
e-mail: ramon.garcia@uv.es

M. del Carmen Recio Iglesias  
Department of Pharmacology, Faculty of Pharmacy,  
University of Valencia, Avenida V.A. Estelles s/n, Burjasot,  
46100 Valencia, Spain

**Keywords** Interleukin-6 (IL-6) · Quantitative structure–activity relationship (QSAR) · Molecular topology · Virtual screening · RAW 264.7 (mouse leukemic monocyte macrophage cell line) · Caco-2 (human epithelial colorectal adenocarcinoma cells)

## Introduction

Inflammatory bowel disease (IBD) is an immune-mediated chronic and relapsing inflammatory disorder caused by the inflammation and sores in the lining of large intestine. IBD is clinically marked by recurrent episodes of diarrhea, rectal bleeding, cramping, abdominal discomfort, and mucosal injury. Ulcerative colitis (UC) and Crohn's disease (CD) are two major types of IBD, they are mainly differentiated by their location, as while UC affects the colon, CD can strike the complete gastrointestinal tract.

IBD typically strikes two frames of population, one aged 15–25 and another around 60 years old. The incidence rates for ulcerative colitis in Europe range from 1.5 to 20.3 cases per one hundred thousand people-year, according to what reported by Loftus [1].

This disorder, whose etiology remains unknown, profoundly increases patient morbidity, decreases quality of life, increases patient susceptibility of developing colorectal cancers [2]. The incidence of colorectal cancer in UK is linked to patients suffering from UC at the following rate: 2% in people having UC for 10 years, which increases to 8% after 20 years and finally, rises to 18% after 30 years as UC's patient [3].

UC therapy includes sulfasalazine and other 5-aminosalicylic acid compounds, corticosteroids and finally if the patient does not respond to the above-cited drugs, then immunosuppressant drugs are administered [4]. However,

there is not still an efficient treatment to cure IBD, resulting at best in symptomatic improvements with severe side toxicity and opportunistic infections. Therefore, there is an urgent need for improving our therapeutic arsenal against IBD [3].

There are many cytokines taking part in the pathogenesis of IBD, but one of the major inflammatory cytokine is interleukin (IL)-6. As Matsumoto et al. [5] reported, the important role played by IL-6 in the development of Th1 cell-mediated chronic colitis is well known. This cytokine, secreted by T cells and macrophages, is essential in the evolution between innate and acquired immunity [6], in addition, elevated levels of IL-6 have been already associated with CD and UC patients [7].

A therapeutic strategy for targeting the IL-6 pathway has been already proposed by using a humanized monoclonal antibody facing the IL-6 receptor (anti-sIL-6R), Tocilizumab. This compound has been tested in patients with IBD with promising results [8].

Moreover, growing experimental evidence suggests a role for IL-6 not only in chronic inflammation, but also in inflammation-associated tumor development as seen in patients with IBD [7]. Accordingly, high levels of IL-6 were found in the peripheral blood of IBD patients and IL-6 was shown to promote T cell accumulation in the colon lamina propria by up-regulating anti-apoptotic factors such as Bcl-2 and Bcl-xl [6].

IL-6, released in the context of chronic inflammation such as that observed in human IBD, not only contributes to cancer initiation by releasing reactive oxygenated species (ROS) and inducing the continuous regeneration of the epithelium secondary to the inflammation-related tissue damage, but also directly affects the promotion and progression of dysplastic lesions by modulating proliferation of cancer cells [6].

Although during inflammation IL-6 is mainly expressed by dendritic cells and macrophages, recent studies have demonstrated that cancer cells may acquire the capability to express high levels of IL-6 being the higher expression observed in the least differentiated colorectal cancer cells. Hence, the capacity to express IL-6 emerges as an important progression step for colorectal cancer [6].

Therefore, targeting IL-6 in patients with IBD not only could reduce inflammation but also prevent the development of colitis-associated cancer [7].

Of course, apart from IL-6, there are other cytokines playing a key role in the acute phase of inflammation and diverse immunological processes, such as IL-1 $\beta$ , tumor necrosis factor (TNF)- $\alpha$ , and interferon (IFN)- $\gamma$ . They are released by various cell types including monocytes-macrophages, neutrophils, and endothelial cells and have been found in increased concentrations in the intestinal tissue of IBD patients [9].

In order to get a consistent in vitro model of the physiopathological action of enterocytes and macrophages concern-

ing the acute phase of IBDs, Caco-2 cells (human epithelial colorectal adenocarcinoma cells), a commonly used model of the human intestinal barrier, were exposed to IL-1 $\beta$ , TNF- $\alpha$ , and IFN- $\gamma$  and finally, murine RAW 264.7 macrophages were exposed to lipopolysaccharide (LPS). LPS, a cell wall component of gram-negative enteric bacteria, is a notorious IL-6 inducer in host immunocytes [5].

Several mathematical approaches have been proposed to find new *hit* or *lead* compounds. Equation systems for quantitative structure-activity relationships (QSARs) are frequently applied to large libraries of compounds for virtual computational screening [10].

Among the QSAR approaches, there is one based on molecular topology (MT). MT applies graph theory to molecular characterization. The MT discipline is still young as its origin dates from the 1970s, when Kier and Hall [11, 12] and other researchers started using 'indices' based on graph theory to study some physicochemical properties of organic compounds, such as formation heat and boiling temperature. Kier and Hall concluded that those properties can be formulated as linear combinations of a few such indices.

Later on, MT was also applied in pharmacological research [13, 14] yielding interesting results. Among the pharmacological topics in which MT has demonstrated high performance stand anticonvulsant [15], antifungal [16], antihistaminic [17], bronchodilator [18], cytostatic [19], hypoglycaemic [20], sedative/hypnotic [21], and anti-inflammatory drugs [22, 23].

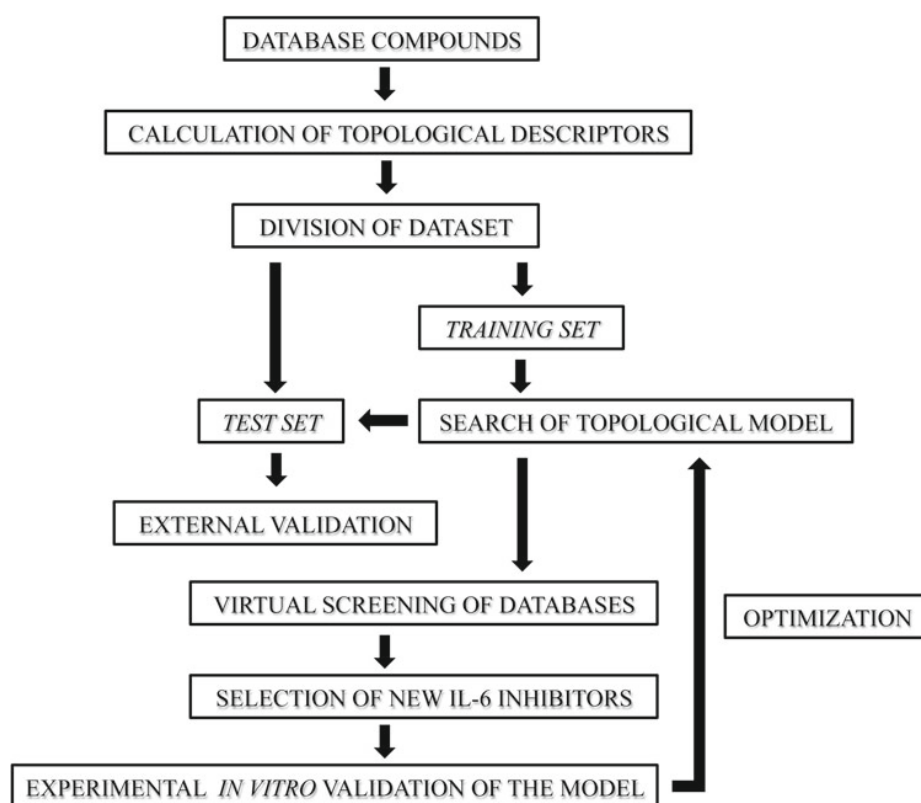
Our group has filed patent applications for the use of this methodology in drug discovery areas such as malaria [24], lung cancer [25, 26], and Alzheimer disease [27].

The key step for developing a good QSAR model is finding the best structural descriptors that reliably represent the molecular features responsible for the pharmacological activity analyzed. In the MT realm, the structure of each molecule is characterized by topological indices (TIs) [28] capable to describe the structures by encoding the structural information pertinent to the property in question.

Virtual screening is a fast, cost-effective process for the discovery of novel *hit* or *lead* drugs [22]. This virtual screening of chemical databases permits to identify those compounds showing the best profiles (drug-like, cell permeability, etc.) for a given pharmacological activity. Today, the availability of high performance computing platforms with low economic cost, allows a much more complex and accurate analysis of very large data sets [22].

Topological virtual screening is focused on the analysis of the chemical diversity of given molecules, which enables the selection of the optimal physicochemical profile molecule among multiple candidates. The molecules are depicted by their TIs. After a computational study of their structures, only those ones complying with a desired topological pattern—

**Fig. 1** Scheme of IL-6 inhibitors research through molecular topology by *virtual screening* on databases



linked to a given biological or physicochemical profile—are chosen for further development [22].

The main strength of MT is just its capability of identifying structural similarities with known active compounds that are not evident from direct observation [13].

In addition, the method is significantly faster and less expensive than other approaches such as those based on molecular mechanics, quantum chemistry, similarity/dissimilarity, or 3D-QSAR, just to mention some examples [22].

The aim of this study is to develop and assess a mathematical model based on MT, capable of identifying new suppressors of IL-6 production with potential applicability on UC.

## Experimental

The strategy of a MT-QSAR based model involves the following steps (Fig. 1):

- Selection of the data set from the literature and databases. These data comprise all compounds (IL-6 inhibitors and non-inhibitors) used for building up the model.
- Calculation of topological descriptors using Dragon software [28].
- Splitting of the data set into two groups: *training set* and *test set*. The applied criteria used for such a division was just based on the structural heterogeneity of compounds.

- Application of linear discriminant analysis (LDA) to the *training set*.
- Validation of the LDA through an external *test set*.
- Application of the topological model to the selection of new IL-6 inhibitors.
- In vitro testing of the selected compounds and hence, an extra model's validation.

## Compound analysis

The QSAR model was developed from a *training set* composed by 53 compounds (25 actives and 28 inactives) with heterogeneous molecular structures. This model was then tested against a set of 108 compounds consisting of 37 IL-6 inhibitors and 71 inactive compounds.

It is necessary to clarify that whereas we are sure that the active compounds are actually IL-6 inhibitors, we cannot assure that none of the compounds in the inactive group is showing activity, simply because not all the compounds in the inactive set has been necessarily tested as IL-6 inhibitors.

Nevertheless, since the data set of inactive compounds is significantly large, the risk of uncertain classification becomes negligible.

All active compounds were obtained from literature [29–41], inactive compounds were selected from OTAVA Ltd collection [42] and exhibited various biological activities



(antituberculosic, antiobesity, antimycobacterial, fibrinogen receptor antagonist, and oxidoreductase inhibitor) and no mention was found in the literature of possible IL-6 activity.

Two databases named MicroSource Pure Natural Products Collection [43] and Sigma-Aldrich database, containing both natural and synthetic origin compounds, were employed to carry on a virtual screening to identify potential IL-6 inhibitors.

### Molecular descriptors

TIs, as other approaches based on fingerprints, codify information about molecular structure in a purely numerical way. Furthermore, their numerical format greatly simplifies the automatic search of other molecules with similar structural properties, hence strong candidates to share the physicochemical, biological, and/or pharmacological properties sought.

The 2D structure of each compound was drawn using the ChemDraw Ultra package (version 10.0) [44]. Each compound was characterized by a set of TIs, derived of the vertices adjacency matrix (e.g., the Randić index, also known as the branching index,  $\chi$ ), distance matrix (e.g., the Balaban distance connectivity index,  $J$ ), or topological molecular descriptors derived from the edge adjacency matrix (spectral moments,  $\mu$ ) [45]. Among the most significant indices, those derived from the spectral moments and the sum of eigenvalues of such graph-matrices are particularly relevant. Other graph-theoretical descriptors were also used but are not depicted here because of their lack of effectiveness. All indices were calculated with Dragon software version 5.4 [28]. All TIs values for every compound used in this study (*training*, *test*, and *virtual screening set*) are shown in the Supplementary Material—Annex I, II, and III.

### Modeling techniques

LDA is a statistic approach to find a linear combination of variables (TIs in our case) enabling a model classification discriminating the category or group to which a given compound belongs (in our case active or inactive as IL-6 inhibitors).

A database was built up so that all compounds were allocated into an active or inactive group according to their IL-6 inhibitory activity. LDA was then applied to these two groups to obtain a discriminant function (DF) using the statistical software Statistica 9.0 [46].

TIs were selected as the independent variables, and IL-6 inhibition as the discriminatory property. The discriminant capability was assessed as the percentage of correct classifications in each set of compounds. The classification criterion was the minimal Mahalanobis distance (distance of each case to the mean of all the cases in a category) [47] and the quality of the discrimination was evaluated using the Wilks

parameter,  $\lambda$  [48], which was obtained by multivariate analysis of variance that tests the equality of group means for the variable in the discriminant model. The shorter the Wilks' parameter value, the smaller the overlap of the active and inactive ( $\lambda = 0$  would mean a perfect separation between the groups).

Selection of descriptors was based on the Fisher–Snedecor parameter ( $F$ ) [49], which establishes the relevance of candidate variables. The descriptors employed to compute the linear classification function are chosen in a stepwise manner: at each step, the variable making the largest contribution to the division of the groups is introduced into the discriminant equation (or the variable that makes the smallest contribution is removed).

The validation of the selected function was done using an external *test set*. Compounds of the *test set* were structurally heterogeneous.

It is worth mentioning that another relevant criterion that usually offers a balanced evaluation of the model's predictive power: the Matthews correlation coefficient (MCC) [50].

This parameter is focused on the fact that in any prediction process there can be four different possibilities to account for:

TP (True positive): Active compounds correctly classified or predicted.

FP (False positive): Inactive compounds classified as IL-6 inhibitor.

TN (True negative): Inactive compounds correctly classified.

FN (False negative): IL-6 inhibitors classified as inactive.

Therefore, any single number that represents the predictive capability of the method must account for all the possibilities listed above. MCC fulfills these requests. Matthews' coefficient is defined as shown above in Eq. (1):

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TN+FN) \times (TN+FP) \times (TP+FP) \times (TP+FN)}} \quad (1)$$

The MCC ranges from  $-1 \leq MCC \leq 1$ . A value of  $MCC = 1$  indicates the best possible prediction, in which every compound in the model was correctly classified, whereas if  $MCC = -1$  then we are in the worst possible case (or anti-correlation), where not a single compound has been correctly labeled. Finally, a Matthews correlation coefficient of  $MCC = 0$  is what would be expected for a random prediction.

Finally, we used the receiver operating characteristic (ROC) curve to evaluate the accuracy of the DFs selected through the sensitivity (true positive fraction) and specificity (true negative fraction) for different thresholds of DF. A ROC curve is the representation of sensitivity versus  $1 - \text{specificity}$  (false positive fraction). The closer the curve follows the left-hand border and then the top border of the ROC space, the

more accurate the model is. The closer the curve comes to the 45° diagonal of the ROC space, the less accurate the test. Accuracy is measured by the area under the ROC curve, AUC [51,52]. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test.

#### Pharmacological activity distribution diagrams

A pharmacological distribution diagram (PDD) is a graphical representation that provides a direct way of visualizing the regions of minimum overlap between active and inactive compounds, as well as the regions in which the probability of finding active compounds is at a maximum [53].

From a different perspective, a PDD is a frequency distribution diagram of dependent variables in which the ordinate represents the expectancy (probability of activity) and the abscissa represents the DF values in the range. For an arbitrary range of values of a given function, an “expectancy of activity” can be defined as  $Ea = a/(i + 1)$ , where “ $a$ ” is the number of active compounds in the range divided by the total number of active compounds and “ $i$ ” is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity is defined in a symmetrical way, as  $Ei = i/(a + 1)$ . Presented with these diagrams, it is easy to visualize the intervals in which there is a maximum probability of finding new active compounds and a minimum probability of finding inactive compounds.

#### Topological virtual screening

A mathematical topological model consisting of one equation is used here to filter a structural database, composed by previously selected compounds with potential activity against UC [4], in order to find new IL-6 inhibitors. This topological virtual screening led to the identification of four compounds potentially active against IL-6 which were then tested in vitro for the assessment of their biological activity against IL-6.

#### Material

Unless otherwise specified, all chemicals including products and biochemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA). ELISA kit for IL-6 cytokine was purchased from eBioscience (San Diego, CA, USA). IL-1 $\beta$ , TNF- $\alpha$  and LPS were also purchased from Sigma-Aldrich while IFN- $\gamma$  was from Calbiochem (Darmstadt, DE).

#### In vitro assays

##### *Cell culture and treatment*

The murine macrophage cell line RAW 264.7 (ECACC, Salisbury, UK) between passages 20 and 40 was main-

tained in Dulbecco’s Modified Eagle’s Medium (DMEM) glucose concentration 1 g/L supplemented with 10 % fetal bovine serum, penicillin (100 U/mL), and streptomycin sulfate (100  $\mu$ g/mL) in an atmosphere of 5 % CO<sub>2</sub>/95 % air (v:v) at 37 °C.

Macrophage RAW 264.7 was plated at a density of  $5 \times 10^5$  cells per mL in 24-well plates. Then, 24 h after seeding the cells, the medium was replaced with fresh DMEM 0.5 % FBS. After this, cells were treated with the four selected compounds at different concentrations (10, 25, 50, and 100  $\mu$ M) and then, 2 h later cells were stimulated with 1  $\mu$ g/mL LPS. Finally, 24 h after stimulation, supernatants were collected for ELISA analysis.

The Caco-2 cell line, derived from a human colon adenocarcinoma (ATCC, Rockville, MD), was used between passages 30 and 45 and cultured in DMEM glucose concentration 4.5 g/L supplemented with 20 % fetal bovine serum, penicillin (100 U/mL) and streptomycin sulfate (100  $\mu$ g/mL) in a humidified 5 % CO<sub>2</sub> atmosphere at 37 °C.

For experiments, cells were seeded at a density of  $3 \times 10^5$  cells per mL on 24-well plates. Then, 24 h after seeding cells, the medium was replaced with fresh DMEM 0.5 % FBS. After this, cells were treated at different concentrations (range 10, 25, 50, and 100  $\mu$ M) with the four compounds and then, 2 h later cells were stimulated with IL-1 $\beta$  (25 ng/mL), TNF- $\alpha$  (50 ng/mL), and IFN- $\gamma$  (50 ng/mL) as Jacqueline Van De Walle published in 2010 [9]. Finally, 24 h after stimulation, supernatants were collected for ELISA analysis.

According to the solubility, the compounds assayed were dissolved in DMSO 5 % in DMEM + 0.5 % FBS or just DMEM + 0.5 % FBS.

##### *Cytotoxicity assay*

The effect of compounds on cell viability was evaluated with the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl-tetrazolium bromide (MTT) assay [54]. After 24 h seeding of murine RAW 264.7 macrophages and Caco-2 cells, the medium was replaced with fresh DMEM + 0.5 % FBS. Two hours later, cells were exposed to compounds at the assayed maximum concentration (25, 50, or 100  $\mu$ M) in a 96-well microplate at 37 °C for 24 h. The medium was then removed and 100  $\mu$ L per well of a 0.5 mg/mL solution of MTT was added. The resulting solution was incubated at 37 °C until blue deposits were visible and then, this colored metabolite was dissolved in dimethyl sulfoxide (DMSO). Absorbance was measured at 490 nm with a Labsystems Multiskan EX plate reader (Helsinki, Finland). The results were expressed in absolute absorbance readings; a decrease in absorbance indicated a reduction in cell viability. The percentage of cell viability was determined as follows:

Cell viability %

$$= \frac{\text{Optical density of treated cells}}{\text{Optical density of non-treated cells}} \times 100 \quad (2)$$

The experiments were performed a minimum of three times with three replicates per concentration.

*Interleukin-6 (IL-6) production in RAW 264.7 macrophages and Caco-2 cells*

The supernatants media were collected and IL-6 secretion was evaluated using a specific enzyme-linked immunosorbent assay kit (eBioscience, San Diego, CA, USA) in accordance with the manufacturer's instructions. IL-6 secretion was quantified in pg/mL using the standard provided with the kit with sensitivity limits of 2 pg/mL. Results were subsequently expressed in relative terms to the negative control (untreated cells) to facilitate comparison between groups. Optical density at 540 nm (OD<sub>540</sub>) was measured by plate reader. Then, we calculated inhibition (%) value, as:

$$I \% = \frac{\text{OD}_{540} \text{ of non-treated well} - \text{OD}_{540} \text{ of treated well}}{\text{OD}_{540} \text{ of non-treated well}} \times 100 \quad (3)$$

In this experiment, IL-6 production was measured as an ulcerative colitis mediator in the intestinal inflammatory process.

### Statistical analysis

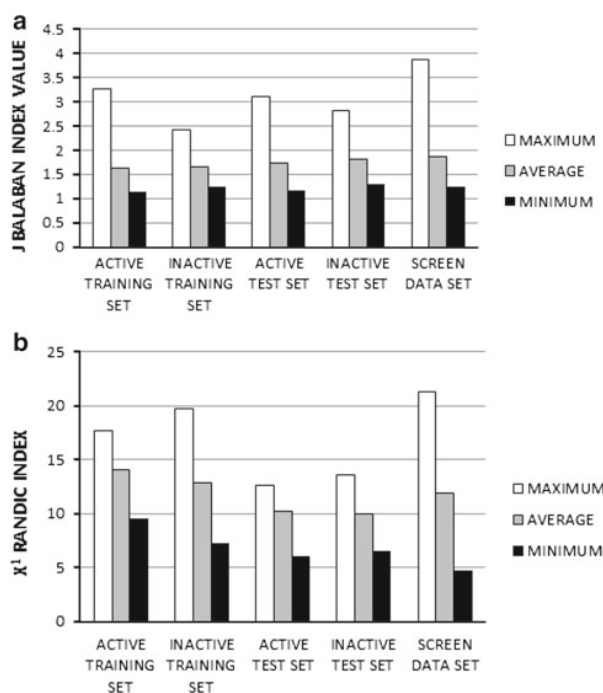
Statistical analysis was performed with a one-way analysis of variance (ANOVA) and Dunnett's *t* test. The results are presented as the mean ± SEM. GraphPad Prism 4.0 software (GraphPad Software Inc., San Diego, CA, USA) was used for all calculations.

## Results and discussion

### Topological similarity study

A study of molecular similarity was previously carried out in order to guarantee that no simple or evident structural features are discriminating between the molecules that make up the data set. We selected two TIs well known and widely used: the Randic branching index,  $\chi$ , and the Balaban distance connectivity index,  $J$ . These indices have been calculated for all compounds in the database with Dragon software [28].

For these parameters in the *training set*, we obtained an average value of 1.7 ( $J$ ) and 13.7 ( $\chi$ ) for the active compounds, whereas average values of 1.8 ( $J$ ) and 10.1 ( $\chi$ ) were obtained for the inactive ones. Hence, the set is well balanced and no obvious structural differences are expected to distort



**Fig. 2** Average, maximum and minimum values (gray, white, and black bars, respectively) obtained with Balaban index ( $J$ ) and Randic index ( $\chi$ ) for the *training*, *test* and *screened sets* compounds

the study. The results obtained with the *test set* are similar to those of the *training set* (see Fig. 2a, b).

If we compare these values to those obtained for the screened data set, i.e., 1.9 ( $J$ ), and 11.9 ( $\chi$ ) value, we can see that the screened compounds show similar values of the two parameters, and therefore the screened molecules are in the domain applicability of the model for identifying novel IL-6 inhibitors.

### Mathematical modeling

The best DF obtained, composed of three-variable, is shown below:

$$\text{DF} = -72.25 + 3.34 \times \text{ESpm15r} - 8.42 \times \text{SEigv} - 31.34 \times \text{SEige} \quad (4)$$

Statistical data and parameters accounting for the significance of this equation were as follows:

$$N = 53, \quad F = 54.04, \quad \lambda = 0.232, \quad p < 0.00001$$

where  $N$  is the number of data compounds;  $F$ , Fisher-Snedecor parameter;  $\lambda$ , Wilks' lambda;  $p$ , statistical significance.

According to Eq. 4, a compound is classified as active if  $\text{DF} > 0$ , otherwise it is labeled as inactive. By applying this criterion to the *training set* (53 compounds, see Supplementary Material—Annex I for details), 23 out of 25 experimentally



**Table 1** Classification matrix obtained through the selected DF for *training* and *test set*

	Percent-correct	Compounds		
		Classified as active	Classified as inactive	Non-classified
<i>Training set</i>				
Active group	92	23	2	–
Inactive group	100	0	28	–
<i>Test set</i>				
Active group	75	9	3	–
Inactive group	84	6	36	1

active compounds were correctly classified as such (92 % accuracy), and all of 28 experimentally inactive compounds were also well classified (100 % accuracy), as can be seen in Tables 1 and 2.

The following formula was used to calculate the percentage of correctly classified compounds within a particular category (active or inactive) as shown in Eq. 5:

$$\text{Classification accuracy \%} = \frac{\text{CCC}}{\text{TNC}} \times 100 \quad (5)$$

where CCC is correctly classified compounds and TNC is total number of compounds.

Altogether, the average of correct classification for the entire set of compounds (active and inactive) was 96 %.

Given the nature of the indices in Eq. 4 it is hard to have an easy interpretation of the suppressing activity on IL-6 production as a function of structural descriptors, which is not surprising, looking the complexity nature of the property studied. In fact, the three indices in Eq. 4 take into account the influence of the resonance integrals, that is a measure of the double bonds' strength related to its degree of conjugation, as well as the Van der Waals volume and the atomic electronegativity (indices ESpm15: Spectral moment 15 from edge adjacency matrix weighted by resonance integrals, SEigv: Eigenvalue sum from van der Waals weighted distance matrix and SEige: Eigenvalue sum from electronegativity weighted distance matrix, respectively) [45,55].

In general terms the active compounds show a low degree of conjugation so that if contain single aromatic rings they are isolated (for example, compounds # 6, 8, 18, and 20). On the contrary, the inactive compounds show generally a high conjugation (see for instance, compounds # 27, 33, and 35).

Moreover, most of compounds showing highly electronegative atoms (such as the halogens) are inactive (compounds # 34, 39, and 51, for example), whereas most of active compounds exhibit groups such as ester or hydroxyl (see compounds # 1, 5, 8, and 22).

Model's prediction was evaluated regarding the Matthews correlation coefficient, which returns a value between –1

and +1. Our model shows a value of 0.93, what ensures its reliability. Furthermore, the MCC was calculated in a slightly different way, i.e., by adding +1 to each scales value, so that the outcome can be expressed as % accuracy. In other words, 0 would mean no correlation at all, 1 represents 50 % and 2 stands for the maximum correlation (100 %). By doing so, our model's yield was 96.5 % (MCC modified = 1.93).

Another method to evaluate the accuracy of the DF selected, is building up the ROC curve, which gives us information about the sensitivity (true positive fraction) and specificity (true negative fraction) of the model, for different DF thresholds. ROC curve for the *training set* is shown in Fig. 3.

The area under the curve (AUC) is 0.9615 for Eq. 4 (see Fig. 3). This high value of AUC stands for the excellent accuracy of the selected DF.

To establish the adequate range of activity, we analyzed the PDD obtained with the DF. Looking at Fig. 4, we can appreciate that all compounds studied show DF values in the range  $14.5 > \text{DF} > -12$ . Outside that range, any compound classification is uncertain and it is labeled as “not-classified” (outliers), NC.

In addition, in Fig. 4 it is noteworthy that there is no overlapping between the active and inactive compounds for DF values above 0, so when applying to a *virtual screening*, the domain applicability to select a compound as active will be  $15 > \text{DF} > 0$ .

The best way to evaluate the quality of any DF is to apply it into an external group (*test set*). In our case, this group was made up of 55 compounds (12 active and 43 inactive) which were not included in the *training set*. *Test set* was randomly selected in a percentage of 30 and 60 from all active and inactive compounds, respectively, conforming *data set*. Table 3 outlines the results of the prediction obtained for every one of these compounds.

As appreciated in Tables 1 and 3, the success rate for active compounds in the *test set* was up to 75 % (9 out of 12 compounds were correctly classified) while, in the case of the inactive group, the rate of correctly classified compounds was 84 % (36 out of 43).

Furthermore, although the strict DF application leads to the loss of some of the active compounds (25 %), the important point is that from 43 test inactive compounds, 84 % were successfully recognized by our model, so that the risk of including “false active” compounds is minimized. That means that our model has a high specificity to predict IL-6 inhibitory activity.

#### Topological virtual screening

Based on the model described above (Eq. 4, Fig. 4), a *virtual screening* was carried out on a previously selected group of

**Table 2** Compounds used in the *training set* and corresponding values of the DF (Eq. 1) to IL-6 inhibitors

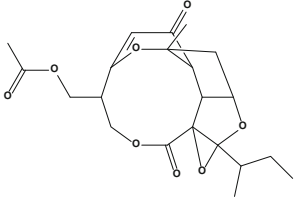
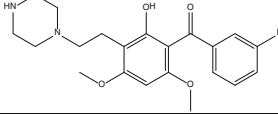
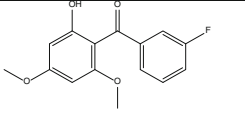
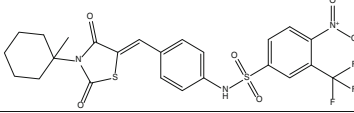
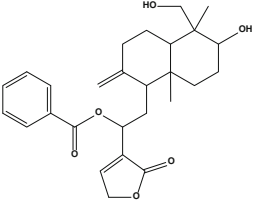
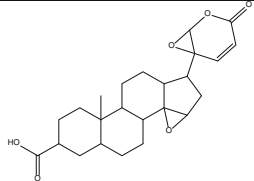
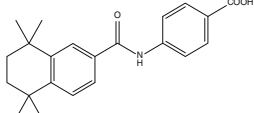
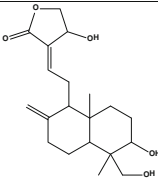
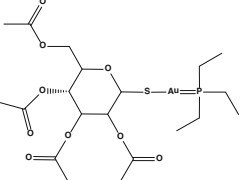
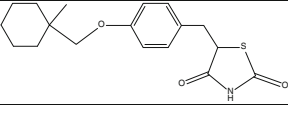
Compound	Structure	DF	Prob (activ.)	Class
<b>1</b>		11.65	1.00	A
<b>2</b>		3.39	0.97	A
<b>3</b>		3.46	0.97	A
<b>4</b>		6.69	1.00	A
<b>5</b>		8.67	1.00	A
<b>6</b>		11.28	1.00	A
<b>7</b>		5.30	1.00	A
<b>8</b>		8.28	1.00	A
<b>9</b>		7.98	1.00	A
<b>10</b>		-1.33	0.21	I

Table 2 continued

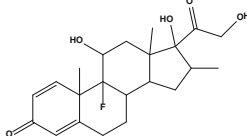
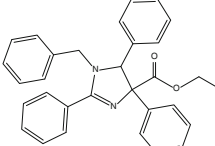
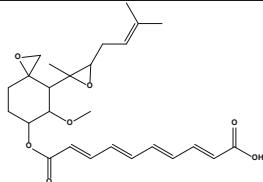
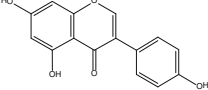
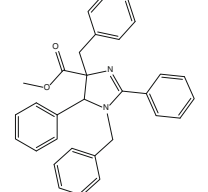
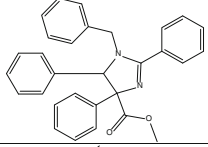
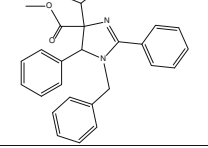
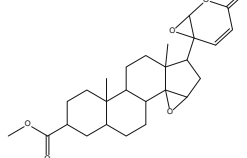
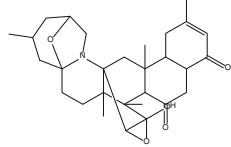
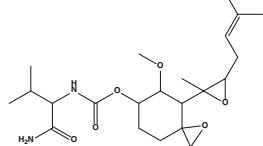
11		14.26	1.00	A
12		5.54	1.00	A
13		9.36	1.00	A
14		2.20	0.90	A
15		4.81	0.99	A
16		5.53	1.00	A
17		5.37	1.00	A
18		12.62	1.00	A
19		13.87	1.00	A
20		7.90	1.00	A

Table 2 continued

21		3.04	0.95	A
22		6.10	1.00	A
23		5.29	0.99	A
24		1.86	0.87	A
25		-3.41	0.03	I
<i>Inactive group</i>				
26		-9.69	0.00	I
27		-4.06	0.02	I
28		-7.01	0.00	I
29		-8.02	0.00	I
30		-8.22	0.00	I
31		-5.02	0.01	I
32		-6.90	0.00	I

Table 2 continued

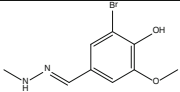
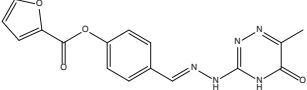
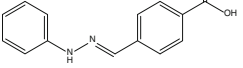
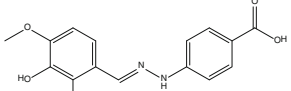
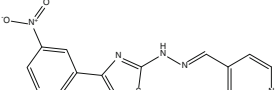
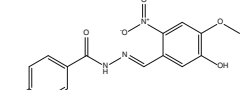
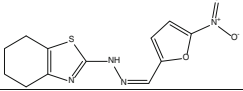
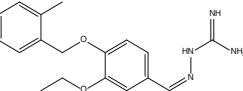
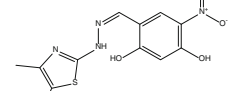
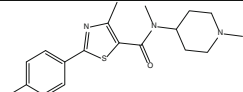
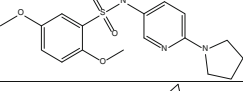
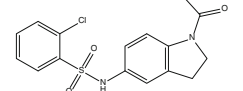
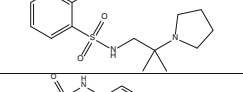
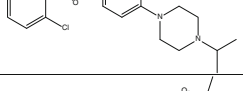
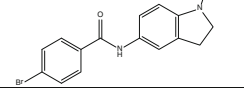
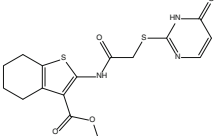
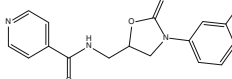
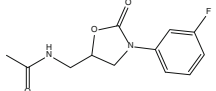
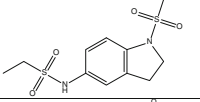
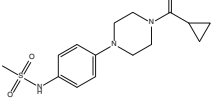
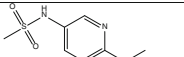
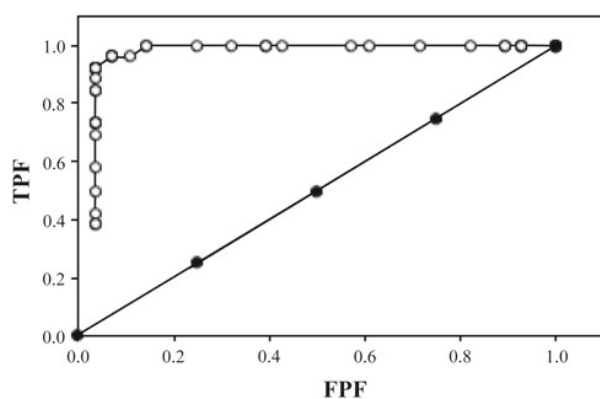
33		-11.00	0.00	I
34		-4.27	0.01	I
35		-4.97	0.01	I
36		-8.42	0.00	I
37		-7.79	0.00	I
38		-7.82	0.00	I
39		-6.93	0.00	I
40		-4.43	0.01	I
41		-5.35	0.00	I
42		-11.30	0.00	I
43		-1.47	0.19	I
44		-6.73	0.00	I
45		-6.66	0.00	I
46		-8.71	0.00	I
47		-7.88	0.00	I

Table 2 continued

48		-6.22	0.00	I
49		-8.82	0.00	I
50		-0.23	0.44	I
51		-2.89	0.05	I
52		-3.18	0.04	I
53		-4.84	0.01	I

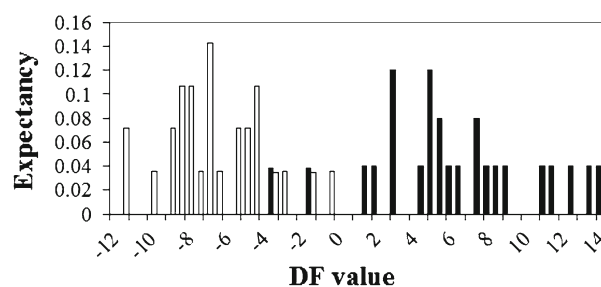


**Fig. 3** Receiver operating characteristic (ROC) curves for Eq. 4. In training set (white points) and random classifier (black points). TPF (true positive fraction) = sensitivity and FPF (false positive fraction) =  $1 - \text{specificity}$ , for different thresholds of classification function (between  $-1.0$  and  $+1.0$ )

28 molecules with hypothetical activity against UC [4]. As shown in Table 4, most of these 28 compounds were also classified as potentially IL-6 inhibitors based in our model's prediction (Eq. 4, Fig. 4).

From the overall 28 compounds selected, finally four compounds were chosen for in vitro testing according to the following criteria:

1. Be classified as active by our topological model.



**Fig. 4** Pharmacological distribution diagram for IL-6 inhibitors obtained using the DF (the black color represents IL-6 inhibitors and the white color, the compounds without IL-6 inhibitory activity)

2. Not already known anti-inflammatory, anti-oxidant or anti-ulcer activity. That is to guarantee a novel approach to UC.
3. Be commercially available.

The compounds selected were: alizarin-3-methyliminodiacetic acid (AMA), Calcein, (+)-dibenzyl-L-tartrate (DLT), and Ro 41-0960.

#### Biological assays

##### Cytotoxicity assay

Caco-2 and RAW 264.7 cells were exposed to AMA, Calcein, DLT, and Ro 41-0960 to determine compounds' cytotoxic

**Table 3** Compounds conforming the *test set* and results obtained applying the LDA (Eq. 1) to IL-6 inhibitors

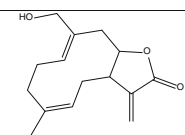
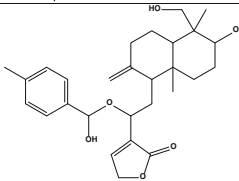
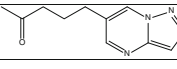
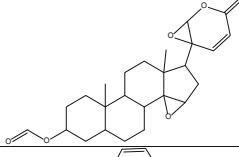
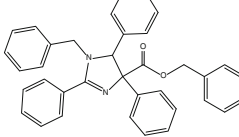
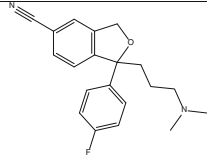
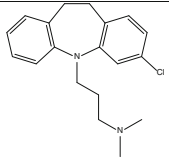
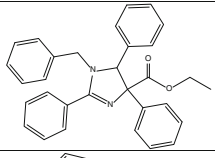
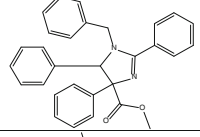
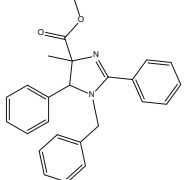
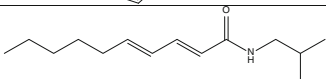
Compound	Structure	DF	Prob (activ.)	Class
<i>Active group</i>				
54		0.01	0.51	A
55		8.69	1.00	A
56		-5.38	0.00	I
57		12.60	1.00	A
58		5.57	1.00	A
59		5.96	1.00	A
60		-8.17	0.00	I
61		5.54	1.00	A
62		5.53	1.00	A
63		4.32	0.99	A
64		-10.00	0.00	I

Table 3 continued

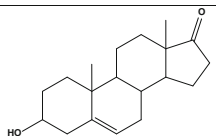
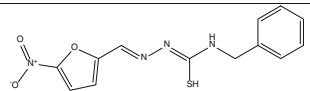
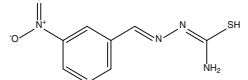
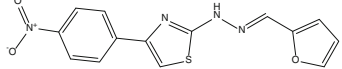
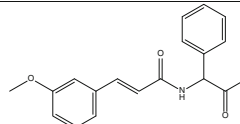
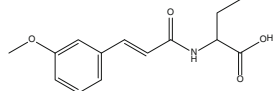
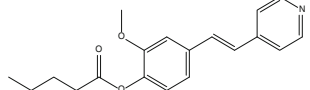
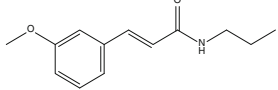
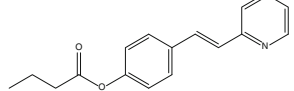
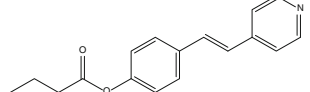
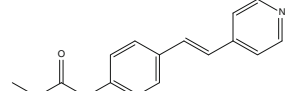
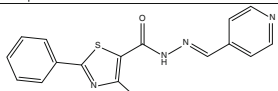
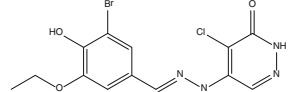
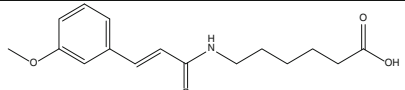
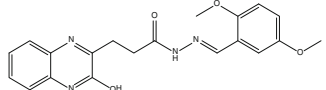
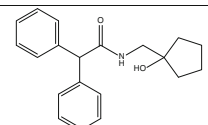
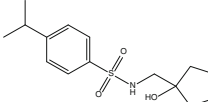
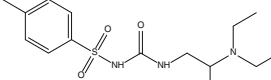
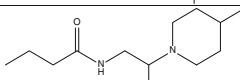
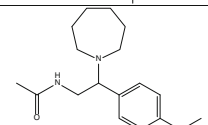
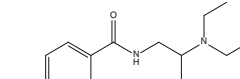
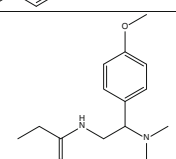
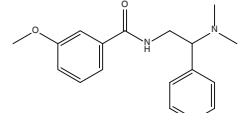
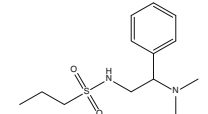
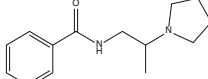
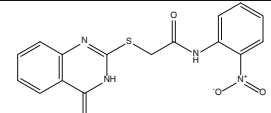
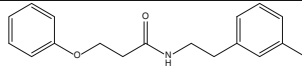
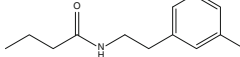
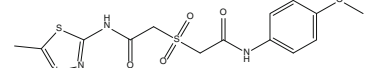
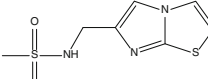
65		7.33	1.00	A
<i>Inactive group</i>				
66		-8.83	0.00	I
67		-9.83	0.00	I
68		-7.03	0.00	I
69		-1.60	0.17	I
70		-3.45	0.03	I
71		-3.03	0.05	I
72		-6.18	0.00	I
73		-5.23	0.01	I
74		-5.13	0.01	I
75		-4.15	0.02	I
76		-6.67	0.00	I
77		-16.56	0.00	NC
78		-5.19	0.01	I
79		-2.32	0.09	I



Table 3 continued

80		-8.51	0.00	I
81		-6.20	0.00	I
82		-4.43	0.01	I
83		-1.51	0.18	I
84		-6.60	0.00	I
85		-8.29	0.00	I
86		2.72	0.94	A
87		-3.67	0.02	I
88		0.73	0.68	A
89		-9.86	0.00	I
90		-2.10	0.11	I
91		1.19	0.77	A
92		0.87	0.70	A
93		-3.60	0.03	I

Table 3 continued

94		1.76	0.85	A
95		0.28	0.57	A
96		-2.60	0.07	I
97		-6.00	0.00	I
98		-3.25	0.04	I
99		-4.00	0.02	I
100		-3.72	0.02	I
101		-2.49	0.08	I
102		-3.70	0.02	I
103		-4.50	0.01	I
104		-4.67	0.01	I
105		-5.41	0.00	I
106		-6.71	0.00	I
107		-7.40	0.00	I
108		-7.98	0.00	I

**Table 4** Values of DF, probability of activity as IL-6 inhibitors, classification, and biological activity of the *virtual screening* data set

Compound	DF	Prob (activ.)	Class	Biological activity
3-Methylorsellinic acid	0.82	0.69	A	
(+)-Dibenzyl-L-tartrate (DLT)	0.71	0.67	A	
10-Hydroxycampothecin	5.12	0.99	A	Anti-inflammatory [56]
2-Methoxyresorcinol	-3.23	0.04	I	
Aconitic acid	-3.45	0.03	I	Anti-inflammatory [61]
Ajmaline diacetate	7.80	1.00	A	
Alizarin-3-methyliminodiacetic acid (AMA)	4.47	0.99	A	Red staining [57]
Apigenin	1.56	0.83	A	Anti-inflammatory [58]
Calcein	9.28	1.00	A	Fluorescent dye [59]
Carapin	10.85	1.00	A	
Cromolyn sodium	5.17	0.99	A	Anti-inflammatory [60]
Emodic acid	4.84	0.99	A	
Evernic acid	2.95	0.95	A	
Fissinolide	11.13	1.00	A	
Folic acid	-2.75	0.06	I	
Genkwanin	1.61	0.83	A	Anti-inflammatory [61]
Haematommic acid	1.32	0.79	A	Anti-oxidant [62]
Hesperidin	7.17	1.00	A	Anti-inflammatory [63]
Lonchocarpic acid	5.69	1.00	A	Antimicrobial [64]
Methyl orsellinate	-0.32	0.42	I	Anti-inflammatory [62]
Phycion	4.07	0.98	A	Anti-inflammatory [65]
Purpurin	4.09	0.98	A	Anti-inflammatory [66]
Pyrocatechuic acid	-1.47	0.19	I	
Ro 41-0960	3.43	0.97	A	Catechol-O-methyltransferase inhibitor [67]
Rosmarinic acid	0.69	0.67	A	Anti-inflammatory [68]
Rotenone	4.73	0.99	A	Cytotoxic [69]
Ursolic acid	11.22	1.00	A	Anti-ulcer [70]
Uvaol	10.77	1.00	A	Anti-inflammatory [71]

activity. All compounds were assayed at 100  $\mu$ M, except Ro 41-0960 which was tested at 50  $\mu$ M in Caco-2 cells and at 25  $\mu$ M in RAW 264.7 cells. As shown in Table 5, none of the compounds were cytotoxic to Caco-2 cells; however, Ro 41-0960 was slightly cytotoxic to RAW 264.7 cells.

#### *Interleukin-6 (IL-6) production in RAW 264.7 macrophages and Caco-2 cells*

RAW 264.7 and Caco-2 cell lines were exposed to four compounds (AMA, Calcein, DLT, and Ro 41-0960) to investigate the inhibitory effect on IL-6 production by enzyme-linked immunosorbent assay (ELISA). As shown in Table 6, three of them inhibited IL-6 production in Caco-2 cells and two of them in RAW 264.7.

**Table 5** Compounds cytotoxicity tested by MTT assay in Caco-2 and RAW 264.7 cell lines

Group	Caco-2 cells		RAW 264.7 cells	
	A $\pm$ SEM	Cell viability %	A $\pm$ SEM	Cell viability %
Control	0.46 $\pm$ 0.03	100	0.52 $\pm$ 0.07	100
AMA	0.40 $\pm$ 0.03	88	0.60 $\pm$ 0.06	116
Calcein	0.42 $\pm$ 0.03	92	0.63 $\pm$ 0.04	120
DLT	0.41 $\pm$ 0.03	89	0.61 $\pm$ 0.05	117
Ro 41-0960	0.38 $\pm$ 0.05	84	0.41 $\pm$ 0.07	78

Values were expressed in function of the control (untreated cells) and represent means absorbance  $\pm$  SEM, and are representative of at least three independent experiments per group. Differences with the control group (untreated cells) were determined by means of a one-way analysis of variance (ANOVA) followed by Dunnett's *t* test. No significant difference with the blank group ( $P > 0.05$ ,  $n = 3$ ; Dunnett's *t* test)

**Table 6** IL-6 secretion in cultured medium was evaluated by ELISA, in RAW 264.7 and Caco-2 cell lines exposed to inflammatory stimuli

Group	Concentration ( $\mu\text{M}$ )	IL-6 production			
		Caco-2 cells <sup>a</sup>		RAW 264.7 cells <sup>b</sup>	
		A $\pm$ SEM	IL-6 inhibition (%)	A $\pm$ SEM	IL-6 inhibition (%)
Control		20.56 $\pm$ 0.98	–	360.0 $\pm$ 10.24	–
AMA	10	33.06 $\pm$ 2.85*	–	319.0 $\pm$ 22.39	11
	50	22.11 $\pm$ 4.60	–	681.1 $\pm$ 26.03***	–
	100	16.49 $\pm$ 1.12	20	692.8 $\pm$ 11.13***	–
Calcein	10	22.07 $\pm$ 4.19	–	248.2 $\pm$ 26.07*	31
	50	9.82 $\pm$ 1.48*	52	385.0 $\pm$ 15.64	–
	100	3.55 $\pm$ 1.95***	83	564.10 $\pm$ 14.75***	–
DLT	10	13.43 $\pm$ 0.58	35	200.6 $\pm$ 49.05	44
	50	6.24 $\pm$ 2.76***	70	145.2 $\pm$ 33.34***	60
	100	0.47 $\pm$ 0.19***	98	126.5 $\pm$ 23.91***	65
Ro 41-0960	10	11.17 $\pm$ 1.29	46	71.31 $\pm$ 17.79***	80
	25	3.55 $\pm$ 0.81***	83	18.39 $\pm$ 7.83***	95
	50	1.72 $\pm$ 0.93***	92		

Values were expressed in function of the negative control (untreated cells) and represent means  $\pm$  SEM, and are representative of at least three independent experiments per group. Differences with the control group (untreated cells) were determined by means of a one-way analysis of variance (ANOVA) followed by Dunnett's *t* test

<sup>a</sup> Caco-2 cells were exposed to inflammatory stimuli [(IL-1 $\beta$  (25 ng/mL), TNF- $\alpha$  (50 ng/mL), and IFN- $\gamma$  (50 ng/mL)], 2 h after product administration;

<sup>b</sup> RAW 264.7 cells were exposed to the inflammatory stimuli LPS (1  $\mu\text{g}/\text{mL}$ ), 2 h after product administration

Significantly different from the control group (\**P* < 0.05; \*\*\**P* < 0.001; Dunnett's *t* test)

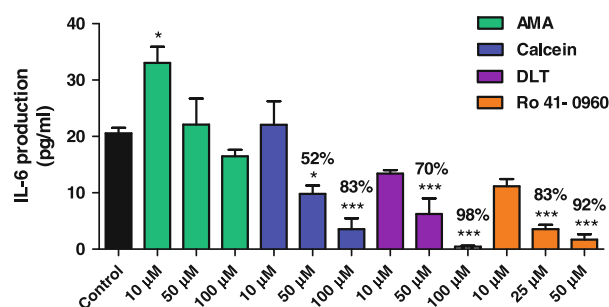
**Table 7** IC<sub>50</sub> of IL-6 inhibition for screened compounds assayed by ELISA, in RAW 264.7 and Caco-2 cell lines

Compound	IC <sub>50</sub> ( $\mu\text{M}$ ) RAW 264.7 cells	IC <sub>50</sub> ( $\mu\text{M}$ ) Caco-2 cells
Alizarin-3-methyliminodiacetic acid	–	248.3
Calcein	–	44.6
(+)-Dibenzyl-L-tartrate	18.6	18.8
Ro 41-0960	0.4	11.6

Ro 41-0960 was clearly the best IL-6 inhibitor on both cell lines, showing an IC<sub>50</sub> value of 0.4  $\mu\text{M}$  in RAW 264.7 cells and 11.6  $\mu\text{M}$  in Caco-2 cells, as shown in Table 7.

To check whether the suppressive effects of the four compounds on IL-6 production was related to cell viability, MTT assay was adopted. The most potent one in Caco-2 cells and RAW 264.7, namely Ro 41-0960, showed no cytotoxicity to Caco-2 cells but showed a little bit of cytotoxic effect in RAW 264.7 macrophages (78 % of cell viability at 25  $\mu\text{M}$ ).

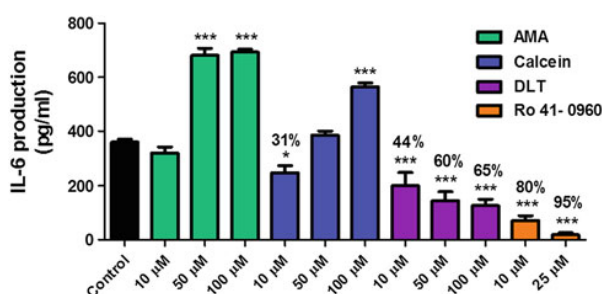
The second most active compound as IL-6 inhibitor was DLT, showing an IC<sub>50</sub> value of 18.6 and 18.8, in RAW 264.7 and Caco-2 cell lines, respectively. Finally, Calcein showed inhibitory activity in Caco-2 cells (IC<sub>50</sub> value of 44.6  $\mu\text{M}$ ) (Figs. 5, 6).



**Fig. 5** IL-6 secretion in culture medium (pg/mL) was evaluated by ELISA, in Caco-2 cells exposed to inflammatory stimuli [(IL-1 $\beta$  (25 ng/mL), TNF- $\alpha$  (50 ng/mL), and IFN- $\gamma$  (50 ng/mL)]. Values were expressed in function of the negative control and represent means  $\pm$  SEM, and are representative of at least three independent experiments per group. Differences with the control group (untreated cells) were determined by means of a one-way analysis of variance (ANOVA) followed by Dunnett's *t* test. \*Significantly different from the control group (\**P* < 0.05; \*\*\**P* < 0.001; Dunnett's *t* test)

DLT and Calcein, at the maximum concentration tested, showed no cytotoxic effect to cell lines used in this study.

The selected compounds (see Table 7) were comparable in IL-6 inhibitory activity with the compounds currently used in the treatment of UC, such as beclomethasone dipropionate [72], and budesonide [73] which have an IL-6 IC<sub>50</sub> value of >100 and 2.2  $\mu\text{M}$ , respectively [74].



**Fig. 6** IL-6 secretion in RAW 264.7 exposed to inflammatory stimuli LPS (1 µg/mL). IL-6 secretion in culture medium (pg/mL) was evaluated by ELISA. Values were expressed in function of the negative control and represent mean  $\pm$  SEM, and are representative of at least three independent experiments per group. Differences with the control group (untreated cells) were determined by means of a one-way analysis of variance (ANOVA) followed by Dunnett's *t* test. \*Significantly different from the control group (\* $P$  < 0.05; \*\*\* $P$  < 0.001; Dunnett's *t* test)

## Conclusions

The results outlined here not only demonstrate that the Topological Virtual Screening can accurately reproduce a well-known pharmacological activity, but also that it is able to discriminate between reference drugs and random compounds, through a predicted DF value.

Through the use of MT the identification of three *in vitro* active compounds has been possible. One of them, a natural dye (Calcein), with potent inhibitory activity against IL-6 and therefore potentially effective in UC, and two others with significant IL-6 inhibitory activity, namely DLT and Ro 41-0960.

These results represent a new step forward in the pathway to demonstrate the high efficiency of the *in silico* methods based on MT. The next step should be testing *in vivo* the active compounds to confirm their usefulness as potential new drugs in the scarce pool of anti-UC drugs.

**Acknowledgments** We thank the Ministerio de Economía y Competitividad, Spain (projects SAF2009-13059-C03-01 and SAF2009-13059-C03-02) for support of this study. M Galvez-Llompарт acknowledges the Atraccio de talents Fellowship provided by the University of Valencia to carry out this study.

## References

- Loftus EV Jr (2004) Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterol* 126:1504–1517. doi:10.1053/j.gastro.2004.01.063
- Fantini MC, Pallone F (2008) Cytokines: from gut inflammation to colorectal cancer. *Curr Drug Targets* 9:375–380. doi:10.2174/138945008784221206
- Eaden J (2004) Review article: colorectal carcinoma and inflammation bowel disease. *Alim Pharmacol Ther Suppl* 4:24–30
- Gálvez-Llompарт M, Recio MC, García-Domenech R (2011) Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B. *Mol Divers* 15:917–926. doi:10.1007/s11030-011-9323-4
- Matsumoto S, Hara T, Hori T, Mitsuyama M, Nagaoka M, Tomiyasu N, Suzuki N, Sata M (2005) Probiotic *Lactobacillus*-induced improvement in murine chronic inflammatory bowel disease is associated with the down-regulation of pro-inflammatory cytokines in lamina propria mononuclear cells. *Clin Exp Immunol* 140:417–426. doi:10.1111/j.1365-2249.2005.02790.x
- Grivennikov S, Karin E, Terzic J, Mucida D, Yu GY, Vallabhapurapu S, Scheller J, Rose-John S, Cheroutre H, Eckmann L (2009) IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer cell* 15:103–113. doi:10.1016/j.ccr.2009.01.001
- Waldner MJ, Neurath MF (2009) Novel cytokine-targeted therapies and intestinal inflammation. *Curr Opin Pharmacol* 9:702–707. doi:10.1016/j.coph.2009.07.005
- Iansante V, Capece D, Murgio S, Mancarelli M, Zazzaroni F, Alesse E (2009) Biotechnological approaches for the treatment of inflammatory diseases. *Anti-Inflamm Anti-Allergy Agents Med Chem* 8:51–71
- Van De Walle J, Hendrickx A, Romier B, Larondelle Y, Schneider YJ (2010) Inflammatory parameters in Caco-2 cells: effect of stimuli nature, concentration, combination and cell differentiation. *Toxicol In Vitro* 24:1441–1449. doi:10.1016/j.tiv.2010.04.002
- García-Domenech R, Gálvez J, de Julián-Ortiz JV, Pogliani L (2008) Some new trends in chemical graph theory. *Chem Rev* 108:1127–1169. doi:10.1021/cr0780006
- Kier LB, Hall LH (1986) Molecular connectivity in structure-activity analysis. Research Studies Press, Letchworth, Hertfordshire
- Kier LB, Hall LH (1976) Molecular connectivity in chemistry and drug research. Academic Press, New York
- Amigó JM, Galvez J, Villar VM (2009) A review on molecular topology: applying graph theory to drug discovery and design. *Naturwissenschaften* 96:749–761. doi:10.1007/s00114-009-0536-7
- Gálvez J, Gálvez-Llompарт M, García-Domenech R (2012) Molecular topology as a novel approach for drug discovery. *Expert Opin Drug Discov* 7:133–153. doi:10.1517/17460441.2012.652083
- Bruno-Blanch L, Galvez J, Garcia-Domenech R (2003) Topological virtual screening: a way to find new anticonvulsant drugs from chemical diversity. *Bioorg Med Chem Lett* 13:2749–2754. doi:10.1016/S0960-894X(03)00535-3
- García-Domenech R, Catala-Gregori A, Calabuig C, Anton-Fos G, Del Castillo L, Galvez J (2002) Predicting antifungal activity: a computational screening using topological descriptors. *Internet Electron J Mol Des* 1:339–350
- Duart MJ, García-Domenech R, Gálvez J, Alemán PA, Martín-Algarra RV, Antón-Fos GM (2006) Application of a mathematical topological pattern of antihistaminic activity for the selection of new drug candidates and pharmacology assays. *J Med Chem* 49:3667–3673. doi:10.1021/jm0580555
- Rios-Santamarina I, Garcia-Domenech R, Gálvez J, Morcillo Esteban J, Santamaria P, Cortijo J (2004) Getting new bronchodilator compounds from molecular topology. *Eur J Pharm Sci* 22:271–277. doi:10.1016/j.ejps.2004.03.013
- Gálvez J, Garcia-Domenech R, Gomez-Lechon M, Castell J (2000) Use of molecular topology in the selection of new cytostatic drugs. *J Mol Struct-THEOCHEM* 504:241–248. doi:10.1016/S0166-1280(00)00365-1
- Antón-Fos G, Garcia-Domenech R, Perez-Gimenez F, Peris-Ribera J, Garcia-March F, Salabert-Salvador M (1994) Pharmacological studies of the two new hypoglycaemic compounds 4-(3-methyl-5-

- oxo-2-pyrazolin-1-yl) benzoic acid and 1-(mesitylen-2-sulfonyl)-1H-1, 2, 4-triazole. *Arzneim Forsch* 44:821–826
21. Estrada E, Peña A, García-Domenech R (1998) Designing sedative/hypnotic compounds from a novel substructural graph-theoretical approach. *J Comput Aided Mol Des* 12: 583–595
  22. Galvez-Llompарт M, Zanni R, García-Domenech R (2011) Modeling natural anti-inflammatory compounds by molecular topology. *Int J Mol Sci* 12:9481–9503. doi:10.3390/ijms12129481
  23. Galvez-Llompарт M, Giner RM, Recio MC, Candeletti S, Garcia-Domenech R (2010) Application of molecular topology to the search of novel NSAIDs: experimental validation of activity. *Lett Drug Des Discov* 7:438–445. doi:10.2174/157018010791306597
  24. Mazier D, Mahmoudi N, Farhati K, Garcia-Domenech R, Galvez J, Derouin F, Danis M (2009) Monensin derivatives for preventing and treating plasmodium infections. *PCT Int Appl WO* 2009074649 A1 20090618
  25. Llompарт J, Galvez J, Pal K (2006) Treatment of cancer with MT477 derivatives (2006) US Patent Appl Publ US 20060014770 A1 20060119
  26. Llompарт J, Galvez J (2004) *N,N*-dicyclohexyl-(1S)-isborneol-10-sulfonamide MT103 family members as antitumor and other therapeutic agents and corresponding treatments and compositions. US Patent Appl Publ US 20040059000 A1 20040325
  27. Galvez J, Llompарт J, Land D, Pasinetti G (2009) Mount Sinai School of Medicine of New York University, USA, Inc Medisyn Technologies, editors. Patent Application Country: Application: WO; Priority Application Country: US patent WO2010114636. 2010 1007; Patent Application Date: 20100405; Priority Application Date: 20090403
  28. Galvez J, Villar VM, Galvez-Llompарт M, Amigo JM (2011) Chemistry explained by topology: an alternative approach. *Comb Chem High Throughput Screen* 14:279–283. doi:10.2174/138620711795222464
  29. Yamaguchi K, Yada M, Tsuji T, Hatanaka Y, Goda K, Kobori T (1999) 4-Phenylthiazole derivatives inhibit IL-6 secretion in osteoblastic cells and suppress bone weight loss in ovariectomized mice. *Bioorg Med Chem Lett* 9:957–960. doi:10.1016/S0960-894X(99)00122-5
  30. Lee SW, Kim MS, Park MH, Park SJ, Lee WS, Chang JS, Rho MC (2010) Alkamides from *Piper longum* and *Piper nigrum* as inhibitors of IL-6 action. *Bull Korean Chem Soc* 31:921–924. doi:10.5012/bkcs.2010.31.04.921
  31. Liu X, Wang J (2011) Anti-inflammatory effects of iridoid glycosides fraction of *Folium syringae* leaves on TNBS-induced colitis in rats. *J Ethnopharmacol* 133:780–787. doi:10.1016/j.jep.2010.11.010
  32. Tang JS, Zhao F, Gao H, Dai Y, Yao ZH, Hong K, Li J, Ye WC, Yao XS (2010) Characterization and online detection of surfactin isomers based on HPLC-MSn analyses and their inhibitory effects on the overproduction of nitric oxide and the release of TNF- $\alpha$  and IL-6 in LPS-induced macrophages. *Marine drugs* 8:2605–2618. doi:10.3390/md8102605
  33. Guh JH, Chang WL, Yang J, Lee SL, Wei S, Wang D, Kulp SK, Chen CS (2010) Development of novel adenosine monophosphate-activated protein kinase activators. *J Med Chem* 53:2552–2561. doi:10.1021/jm901773d
  34. Kuramoto M, Hayashi K, Yamaguchi K, Yada M, Tsuji T, Uemura D (1998) Structure–activity relationship of norzoanthamine exhibiting significant inhibition of osteoporosis. *Bull Chem Soc Jpn* 71:771–779
  35. Kagechika H, Kawachi E, Fukasawa H, Saito G, Iwanami N, Umemiya H, Hashimoto Y, Shudo K (1997) Inhibition of IL-1-induced IL-6 production by synthetic retinoids. *Biochem Biophys Res Commun* 231:243–248. doi:10.1006/bbrc.1997.6087
  36. Koch E, Klaas CA, Rüngeler P, Castro V, Mora G, Vichnewski W, Merfort I (2001) Inhibition of inflammatory cytokine production and lymphocyte proliferation by structurally different sesquiterpene lactones correlates with their effect on activation of NF-[kappa] B. *Biochem Pharmacol* 62:795–801. doi:10.1016/S0006-2952(01)00714-6
  37. Enomoto A, Rho MC, Komiyama K, Hayashi M (2004) Inhibitory effects of bufadienolides on interleukin-6 in MH-60 cells. *J Nat Prod* 67:2070–2072. doi:10.1021/np049950e
  38. Kahlon DK, Lansdell TA, Fisk JS, Tepe JJ (2009) Structural–activity relationship study of highly-functionalized imidazolines as potent inhibitors of nuclear transcription factor-[kappa] B mediated IL-6 production. *Bioorg Med Chem* 17:3093–3103. doi:10.1016/j.bmc.2009.03.002
  39. Enomoto A, Rho MC, Fukami A, Hiraku O, Komiyama K, Hayashi M (2004) Suppression of cancer cachexia by 20S, 21-epoxy-resibufogenin-3-acetate-a novel nonpeptide IL-6 receptor antagonist. *Biochem Biophys Res Commun* 323:1096–1102. doi:10.1016/j.bbrc.2004.08.196
  40. Bandgar BP, Patil SA, Totre JV, Korbad VL, Gacche RN, Hote BS, Jalde SS, Chavan HV (2010) Synthesis and biological evaluation of nitrogen-containing benzophenone analogues as TNF-[alpha] and IL-6 inhibitors with antioxidant activity. *Bioorg Med Chem Lett* 20:2292–2296. doi:10.1016/j.bmcl.2010.02.001
  41. Bharate SB, Mahajan TR, Gole YR, Nambiar M, Matan T, Kulkarni-Almeida A, Balachandran S, Junjappa H, Balakrishnan A, Vishwakarma RA (2008) Synthesis and evaluation of pyrazolo [3,4-*b*] pyridines and its structural analogues as TNF- $\alpha$  and IL-6 inhibitors. *Bioorg Med Chem* 16:7167–7176. doi:10.1016/j.bmc.2008.06.042
  42. OTAVA Ltd. <http://www.otavachemicals.com/>. Accessed 6 April 2012
  43. MicroSource Pure Natural Products Collection. <http://www.msdiscovery.com/home.html>. Accessed 14 Dec 2012
  44. ChemDraw Ultra package (version 10.0). CambridgeSoft (2009)
  45. Todeschini R, Consonni V (2008) Handbook of molecular descriptors. Wiley-VCH Verlag GmbH, Weinheim. doi:10.1002/9783527613106.oth2
  46. StatSoft I (2009) STATISTICA (data analysis software system), version 9. Tulsa, USA
  47. De Maesschalck R, Jouan-Rimbaud D, Massart D (2000) The mahalanobis distance. *Chemom Intell Lab Syst* 50:1–18. doi:10.1016/S0169-7439(99)00047-7
  48. Klecka WR (1980) Discriminant analysis. Sage Publications, Inc., Thousand Oaks
  49. Furnival GM (1971) All possible regressions with less computation. *Technometrics* 13:403–408
  50. Matthews BW (1975) Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim Biophys Acta* 405:442–451
  51. McClish DK (1989) Analyzing a portion of the ROC curve. *Med Decis Mak* 9:190–195. doi:10.1177/0272989X8900900307
  52. Pencina MJ, D’Agostino RB Sr, D’Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 27:157–172. doi:10.1002/sim.2929
  53. Gálvez J, García-Domenech R, de Gregorio Alapont C, de Julián-Ortiz J, Popa L (1996) Pharmacological distribution diagrams: a tool for de novo drug design. *J Mol Graph* 14:272–276. doi:10.1016/S0263-7855(96)00081-1
  54. Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 65:55–63. doi:10.1016/0022-1759(83)90303-4
  55. Estrada E (1996) Spectral moments of the edge adjacency matrix in molecular graphs. 1. Definition and applications to the prediction



- of physical properties of alkanes. *J Chem Inf Comput Sci* 36:844–849. doi:[10.1021/ci950187r](https://doi.org/10.1021/ci950187r)
56. Cashman J, Burt H, Springate C, Gleave J, Jackson J (2004) Camptothecin-loaded films for the prevention of postsurgical adhesions. *Inflammation Res* 53:355–362. doi:[10.1007/s00011-004-1272-2](https://doi.org/10.1007/s00011-004-1272-2)
57. McLeod MJ (1980) Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. *Teratology* 22:299–301
58. El Alfy T, El Sawi S, Sleem A, Moawad D (2010) Investigation of flavonoidal content and biological activities of *Chorisia insignis* Hbk. leaves. *Aust J Basic Appl Sci* 4:1334–1348
59. Parish CR (1999) Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunol Cell Biol* 77:499–508. doi:[10.1046/j.1440-1711.1999.00877.x](https://doi.org/10.1046/j.1440-1711.1999.00877.x)
60. Zhang Y, Reenstra WW, Chidekel A (2001) Antibacterial activity of apical surface fluid from the human airway cell line Calu-3: pharmacologic alteration by corticosteroids and beta(2)-agonists. *Am J Respir Cell Mol Biol* 25:196–202
61. Pelzer LE, Guardia T, Juarez AO, Guerreiro E (1998) Acute and chronic antiinflammatory effects of plant flavonoids. *Il Farmaco* 53:421–424. doi:[10.1016/S0014-827X\(98\)00046-9](https://doi.org/10.1016/S0014-827X(98)00046-9)
62. Marante FJT, Castellano AG, Rosas FE, Aguiar JQ, Barrera JB (2003) Identification and quantitation of allelochemicals from the lichen *Lethariella canariensis*: phytotoxicity and antioxidative activity. *J Chem Ecol* 29:2049–2071
63. Benavente-Garcia O, Castillo J (2008) Update on uses and properties of citrus flavonoids: new findings in anticancer, cardiovascular, and anti-inflammatory activity. *J Agric Food Chem* 56:6185–6205. doi:[10.1021/jf8006568](https://doi.org/10.1021/jf8006568)
64. Alencar N, Cavalcante C, Vasconcelos M, Leite K, Aragão K, Assreuy A, Nogueira N, Cavada B, Vale M (2005) Anti-inflammatory and antimicrobial effect of lectin from *Lonchocarpus sericeus* seeds in an experimental rat model of infectious peritonitis. *J Pharm Pharmacol* 57:919–922. doi:[10.1211/0022357056352](https://doi.org/10.1211/0022357056352)
65. Shen MY, Liu YJ, Don MJ, Liu HY, Chen ZW, Mettling C, Corbeau P, Chiang CK, Jang YS, Li TH (2011) Combined phytochemistry and chemotaxis assays for identification and mechanistic analysis of anti-inflammatory phytochemicals in *Fallopia japonica*. *PLoS ONE* 6:e27480. doi:[10.1371/journal.pone.0027480](https://doi.org/10.1371/journal.pone.0027480)
66. Mishchenko N, Fedoreev S, Bryukhanov V, Zverev YF, Lampatov V, Azarova O, Shkryl' YN, Chernoded G (2007) Chemical composition and pharmacological activity of anthraquinones from *Rubia cordifolia* cell culture. *Pharm Chem J* 41:605–609. doi:[0091-150X/07/4111-0605](https://doi.org/0091-150X/07/4111-0605)
67. Miller JW, Shukitt-Hale B, Villalobos-Molina R, Nadeau MR, Selhub J, Joseph JA (1997) Effect of L-Dopa and the catechol-O-methyltransferase inhibitor Ro 41-0960 on sulfur amino acid metabolites in rats. *Clin Neuropharmacol* 20:55–66
68. Su P, Wang G, Wu D, Sheng X (2008) Progress in rosmarinic acid biological activities and its sources. *Shipin Yu Fajiao Gongye* 34:135–138
69. Yoon IS, Au Q, Barber JR, Ng SC, Zhang B (2010) Development of a high-throughput screening assay for cytoprotective agents in rotenone-induced cell death. *Anal Biochem* 407:205–210. doi:[10.1016/j.ab.2010.08.011](https://doi.org/10.1016/j.ab.2010.08.011)
70. Li J, Huang W, Zhang H, Wang X, Zhou H (2007) Synthesis of andrographolide derivatives and their TNF-[alpha] and IL-6 expression inhibitory activities. *Bioorg Med Chem Lett* 17:6891–6894. doi:[10.1016/j.bmcl.2007.10.009](https://doi.org/10.1016/j.bmcl.2007.10.009)
71. Marquez-Martin A, Puerta RDL, Fernandez-Arche A, Ruiz-Gutierrez V, Yaqoob P (2006) Modulation of cytokine secretion by pentacyclic triterpenes from olive pomace oil in human mononuclear cells. *Cytokine* 36:211–217. doi:[10.1016/j.cyto.2006.12.007](https://doi.org/10.1016/j.cyto.2006.12.007)
72. Papi C, Aratari A, Moretti A, Mangone M, Margagnoni G, Koch M, Capurso L (2010) Oral beclomethasone dipropionate as an alternative to systemic steroids in mild to moderate ulcerative colitis not responding to aminosalicylates. *Digest Dis Sci* 55:2002–2007. doi:[10.1007/s10620-009-0962-6](https://doi.org/10.1007/s10620-009-0962-6)
73. Keller R (2004) Potential use of budesonide in inflammatory bowel diseases: extended ulcerative colitis. Topical steroids in gastroenterology and hepatology. In: Proceedings of the Falk Workshop—new findings on pathogenesis and progress in management of inflammatory bowel diseases, Part III. Berlin, Germany, June 14, 2003, pp 39–46
74. Kang B-S, Chung E-Y, Yun Y-P, Lee MK, Lee YR, Lee K-S, Min KR, Kim Y (2001) Inhibitory effects of anti-inflammatory drugs on interleukin-6 bioactivity. *Biol Pharm Bull* 24:701–703. doi:[10.1248/bpb.24.701](https://doi.org/10.1248/bpb.24.701)





# Molecular topology: A strategy to identify novel compounds against ulcerative colitis.

María Gálvez-Llompарт <sup>a,b\*</sup>, Maria C. Recio <sup>b</sup>, Ramón García-Domenech <sup>a</sup>, Jorge Gálvez <sup>a</sup>

<sup>a</sup> *Molecular Connectivity and Drug Design Research Unit, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, Avd, V.A. Estellés, s/n 46100-Burjassot, Valencia, Spain*

<sup>b</sup> *Department of Pharmacology, Faculty of Pharmacy, University of Valencia, Avd, V.A. Estellés, s/n 46100-Burjassot, Valencia, Spain*

E-mail: [maria.galvez@uv.es](mailto:maria.galvez@uv.es) Tel: 34-6-3543311

## ABSTRACT

In the present paper, a strategy to identify novel compounds against ulcerative colitis (UC) by Molecular topology (MT) is presented. Several quantitative structure–activity relationship (QSAR) models based on molecular topology have been developed to predict the iNOS and TNF- $\alpha$  mediated anti-ulcerative colitis (UC) activity and the protective activity against a DSS-induced UC model. Each one has been used for the screening of four previously selected compounds as potential therapeutic agents for UC: Alizarin-3-methylindiacetic acid (AMA), Calcein, (+)-Dibenzyl-L-tartrate and Ro 41-0960. The four compounds were then tested *in vitro* and *in vivo*, confirming AMA and Ro 41-0960 as the best candidates to become potential lead compounds against UC.

## KEYWORDS

Quantitative structure–activity relationship (QSAR) - Molecular topology - Virtual screening - RAW 264.7 - Ulcerative colitis – iNOS- TNF- $\alpha$

## INTRODUCTION

Inflammatory bowel disease (IBD), which mainly includes ulcerative colitis and Crohn's disease, is a chronic disabling gastrointestinal disorder. It affects many aspects of patient's life. As Burisch J. *et al.* reported an estimated 2.5–3 million people in Europe are affected by IBD. This condition leads to a direct healthcare cost of 4.6–5.6 billion Euros/year, related to: disease course, risk for surgery and hospitalization, cancer risks, patients' disability and work impairment [1].

Ulcerative colitis (UC) is characterized by periods of relapse and remission and involves the colon and rectum. Patients with UC usually experience episodes of bloody diarrhea, seldom with mucous, abdominal pain, fever and weight loss. The chronic inflammation, in addition to a decrease in the quality of life, ends in an increasing risk of development of colorectal cancer [2].

UC seems to be caused by environmental factors combined with a genetically predisposition. As ECCO e-Guide reported, common environmental factors are: family history of IBD, infectious gastroenteritis, use of antibiotics in childhood and probably use of NSAID. Curiously, appendectomy and tobacco smoking seems to decrease the risk of developing UC [3].

The onset usually occurs in late adolescence or early adult hood and affects equally both genders [3]. To date, no cure has been identified. The available treatments are aimed to control the symptomatology by inducing the disease remission. First-line drugs for UC include: amino salicylates, such as 5-aminosalicylic acid (5-ASA) or sulfasalazine; corticosteroids such as prednisone or budesonide; immunosuppressive agents such as azathioprine or 6-mercaptopurine (6-MP) and finally, the metronidazole [3]. Nevertheless, 20% to 40% of patients do not respond to conventional medication and receives secondary drug treatment, based on monoclonal

antibodies such as infliximab and adalimumab. When each of these approaches fail, colectomy represents the last resort [2].

In UC, several inflammatory mediators such as enzymes, cytokines or transcription factors are overexpressed [4]. Particularly, pro-inflammatory cytokines including IL-6, IL-12, IL-1 $\beta$  and tumor necrosis factor TNF- $\alpha$  have been recognized as key factors in mediating the immune response [5].

The nitric oxide synthases (NOS) transforms the amino acid L-arginine into L-citrulline and NO. High levels of NO in tissues are present when inducible NOS (iNOS) are expressed. Due to the fact that cytokines and different pathological agents can cause the expression of iNOS, an overexpressed production of NO in the gastrointestinal tract can be crucial in the onset of various physiological and pathological conditions [6-7]. Patients with UC have shown a marked increase in NO synthesis and iNOS expression in colonic biopsy specimens [6].

Among the pro-inflammatory cytokines, IL-1 is gaining importance in UC due to the fact that it is predominantly produced by stimulated macrophages and monocytes and has a major role in inflammation and immunological response. Gionchetti P *et al.* reported high levels of IL-1 $\beta$  in the colonic mucosa of patients with UC, correlating the IL-1 $\beta$  level in tissues with the disease activity, and relating the ratio of the endogenous IL-1 receptor antagonist to IL-1 with inflammation [8]. Therefore, IL-1 is thought to play an important role in the pathogenesis of UC [9].

Results from experimental colitis models in animals suggest that a down-regulation of IL-1 has an inhibitory effect on the inflammation and consequently on the tissue damage [9]. The IL-1 receptor antagonism reduces the infiltration of inflammatory cells and suppresses the myeloperoxidase (MPO) activity in the large intestine, while prevents edema and necrosis of the large intestine.

Another relevant pro-inflammatory cytokines linked to UC pathogenesis is TNF- $\alpha$ . It is produced by macrophages, and activates T-cells in response to intestinal injury [10]. Excessive production of TNF- $\alpha$  leads to a further stimulation with recruitment of immune cells in the gut, and to the activation of fibroblasts and endothelial cells [5;11-12]. Biological samples from patients with UC, such as blood, stool and urine shows increased levels of TNF- $\alpha$ . Marked levels of this cytokine can also be found in the colon tissue samples. Consequently, TNF- $\alpha$  is a very promising target for biological agents [5;12]. It is not surprising that the U.S. Food and Drug Administration recently introduced golimumab (anti-TNF agent) and vedolizumab (anti-adhesion therapy) as new treatment for ulcerative colitis [2].

The activation of the nuclear factor kappa B (NF- $\kappa$ B) by TNF- $\alpha$  and IL-1 has a significant pro-inflammatory effect. When the NF- $\kappa$ B is activated by the degradation of the I $\kappa$ B $\kappa$ , it translocate to the nucleus. Here, NF- $\kappa$ B interacts with the target genes and begins the transcription [11]. This factor is also involved in the transcription of genes coding for enzymes such as iNOS, COX-2, 5- and 12-lipoxygenase, chemokines, and cytokines (IL-1, IL-6 and TNF- $\alpha$ ) [11-12]. Moreover, seems like constitutive NF- $\kappa$ B activation leads to the development of colitis-associated cancer in patients with IBD [12].

On the basis of the above considerations, it is understandable the importance of inhibiting the production of NO, IL-6, IL-1 $\beta$ , TNF- $\alpha$  and the activation of NF-kB in order to manage UC. Therefore, a potential drug able to modulate these mediators may be a key drug in UC treatment. In this context, the computer-aided drug design (CADD) techniques play an essential role in the process of research and development of new drugs. CADD methodologies can be divided in two main approaches: Bioinformatic, focused on target identification and Chemoinformatic, based on *lead* finding and optimization. This paper deals with QSAR and virtual screening, which is a branch of the Chemoinformatic CADD techniques [13].

QSAR methodology applied to drug design could be defined as ‘mathematical relationships linking chemical structure and pharmacological activity in a quantitative manner’ [13]. When this mathematical relationship is characterized by the use of topological and topochemical descriptors, it is called Molecular Topology (MT).

MT can be described as a branch of mathematical chemistry, which describes chemical structures in a pure mathematical way. Under the rules of the graph-theory, this description deals with the connectivity of the atoms inside the molecule and it is characterized by numerical descriptors, which are independent of the chemical structure’s conformation or in general of three-dimensional (3D) features [14].

Molecular descriptors have been successfully used for the identification of new potential drugs in different pharmacological areas [15-20]. Many of these drugs have been patented in fields like cancer, malaria and Alzheimer disease [21-24]. MT also allowed the discovery of new potential *lead* compounds against UC [25-26]. This work shows the latest application of MT in IBD area [25-26], for the identification of new potential compounds as inhibitors of some of the main mediators in UC and as protective agents against acute UC by dextran sulfate sodium (DSS).

Finally, the potential *hit* and *lead* compounds are tested in different UC experimental models through *in vitro* and *in vivo* assays (macrophages exposed to lipopolysaccharide (LPS) and DSS induced colitis respectively).

Lipopolysaccharide (LPS), also known as lipoglycan and endotoxin, is an outer membrane element of *Gram-negative* bacteria. The LPS stimuli triggers the signaling pathways of the inflammatory cell (COX-2 and LOX) and consequently, promotes the production of pro-inflammatory cytokines like TNF- $\alpha$ , IL-1 $\beta$ , IL-6 and IL-8 along with the recruitment of leucocytes [11,27].

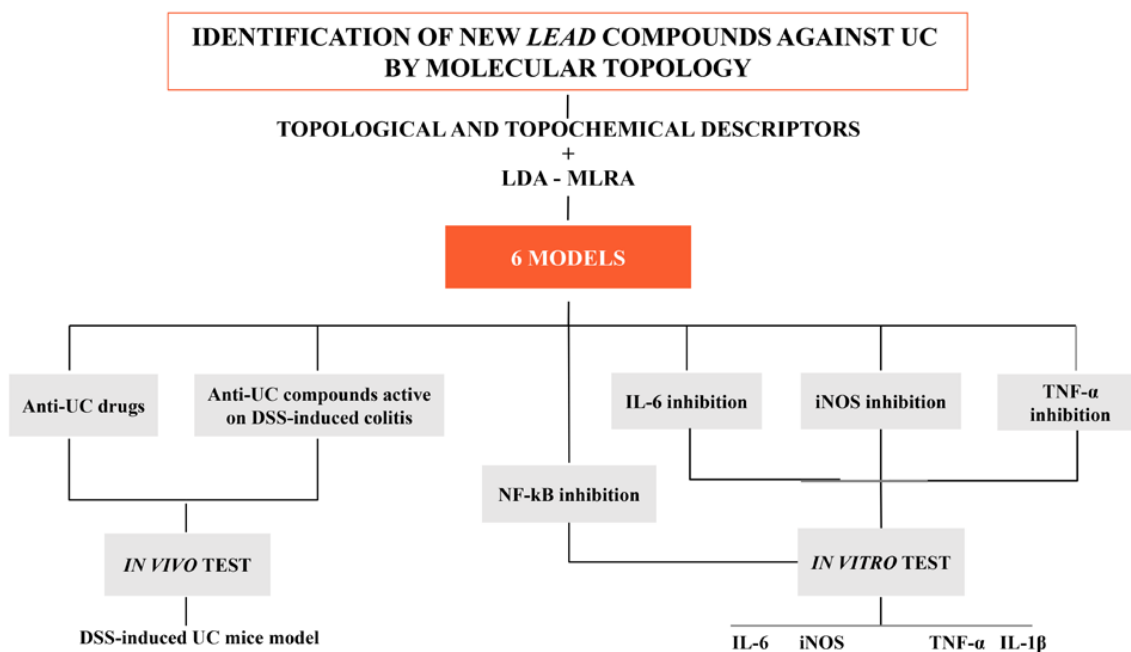
In the present work, in order to study the effect of the compounds (Alizarin-3-methylinodiacetic acid (AMA), Calcein, (+)-Dibenzyl-L-tartrate (DLT) and Ro 41-0960) on the cytokines related to the acute phase of UC, an *in vitro* model was carried out, in which murine RAW 264.7 macrophages were exposed to lipopolysaccharide (LPS).

On the other hand, dextran sulfate sodium (DSS)-induced colitis in mice has been used as an experimental model that exhibits similar symptomatology to the UC in human, such as diarrhea, bloody feces, body weight loss and mucosal ulceration. So, this model represents a reliable tool for testing new drugs as UC treatment [9].

## MATERIALS AND METHODS

### *In silico* study

The strategy for the selection of new chemical compounds against UC by MT is represented in Scheme 1.



**Scheme 1.** Topological virtual screening strategy to identify potential compounds against UC [25-26].

On previous publications, several models for the identification of molecules as inhibitors of NF-kB and IL-6 in UC have been developed. In the present work, new brand models are presented. The objective is to determine whether the selected compounds could be active on a DSS-induced colitis assay and as inhibitors of iNOS and TNF- $\alpha$ .

The data set for building up the models was collected from literature and several databases (US Drugs, DBbinding and Tocris Bioscience), and then it was split into *training* set and *test* set. Calculation of the topological descriptors (TIs) was performed using *Desmoll* and *Dragon* software [28-29].

In order to obtain the predictive equations, two statistical techniques were employed: Linear Discriminant Analysis (LDA) and Multi-linear Regression Analysis (MLRA) [30-31]. Once the models are developed, internal and external validations are carried out [32].

Finally, the models were used to conduct a virtual screening of four preselected compounds with potential activity against UC (AMA, Calcein, DLT and Ro 41-0960), which were previously published by our Research group [25-26]. These compounds are of natural (AMA and Calcein) and synthetic (DLT and Ro 41-0960) origin.

### Molecular descriptors

The 2D structure of each chemical compound was drawn using the ChemDraw Ultra package (version 10.0). Each compound was characterized by a set of TIs. Among them stand out connectivity indices, topological descriptors, topological molecular descriptors derived from the edge adjacency matrix (Spectral moments,  $\mu$ ), walk and path counts, eigenvalue-based indices, and 2D autocorrelations descriptors (Broto-Moreau, Moran y Geary autocorrelations) [33]. Other graph-theoretical indices were also used, but due to their lack of effectiveness are not depicted. All the TIs used in the present study along with their values for every compound (*training*, *test*, and *virtual screening set*) are shown in the supporting information section, Annex I-IX.



## Modeling techniques

### *Linear discriminant analysis (LDA)*

LDA is a method of pattern recognition to find a linear combination of variables (TIs in this case) to qualitatively discriminate two or more category or group of objects (molecules) and enable their classification. The compounds of the database were allocated into the active or inactive group, according to their iNOS, TNF- $\alpha$  inhibitory activity or their anti-DSS-induced UC activity. LDA was then used to obtain three discriminant functions: DF<sub>1</sub>-DF<sub>3</sub>. The compounds were classified either as active or inactive based on their iNOS, TNF- $\alpha$  inhibitory activity values or their anti-DSS-induced UC activity values, respectively. TIs were the independent variables while the activity was the discriminatory property. The percentage of correct classification of the compounds in their respective sets determined the discriminant power of the model. The Mahalanobis distance (distance of each case to the mean of all the cases in a category) was the main classification indicator, while the Wilks parameter,  $\lambda$ , was used to determine the robustness and value of the discriminant function. The Wilks  $\lambda$  is calculated by multivariate analyses of variance that verify the fairness of group means for the variable in the model. Finally, the strategy to choose the indices followed the Fischer-Snedecor parameter (F) that evaluates the weight of each variable. The software for the LDA study is the software package Statistica version 9.0 [34]. A stepwise strategy was used for the selection of the variables of the LDA function: at each step, the variable that makes the largest contribution between the groups is added to the discriminant equation (or the variable that makes the smallest contribution is removed). Subsequently, a test set was creating, in order to perform an external validation of the

discriminant functions. Test set's compounds were meticulously balanced (structurally heterogeneous, similar weight, number of heteroatoms).

It is worth mentioning that another relevant criterion was also used, that usually offers a balanced evaluation of the model's predictive power: the Matthews correlation coefficient (MCC) [35].

This parameter focuses on the outcomes of the prediction process, which are usually four:

- TP (True positive): Active compounds correctly classified or predicted.
- FP (False positive): Inactive compounds misclassified.
- TN (True negative): Inactive compounds correctly classified.
- FN (False negative): Active compounds misclassified.

Therefore, any outcome that represents the predictive capability of the method must account for all the possibilities listed above. MCC fulfills these requests and is defined as shown in Eq. (1):

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TN + FN) \times (TN + FP) \times (TP + FP) \times (TP + FN)}} \quad (1)$$

The MCC ranges from  $-1 \leq MCC \leq 1$ . A value of  $MCC = 1$  indicates the best possible prediction (every compound of the model is correctly classified).  $MCC = -1$  represents the anti-correlation (no correlation at all, no compound is correctly classified). Lastly, a Matthews correlation coefficient value = 0 represents a random prediction.

#### *Pharmacological-activity distribution diagrams*

A pharmacological distribution diagram (PDD) is a plot of expectancy of activity versus the numerical outputs of the discriminant function (DF) for a particular pharmacological activity [36]. Expectancy of activity is defined as:  $E_a = a/(i + 1)$ , where "a" and "i" are the number of active and inactive compounds in a particular interval of DF values. Similarly, it can define  $E_i$ ,

or expectancy of inactivity, as:  $E_i = i/(a + 1)$ . The use of such diagrams eases the visualization of the DF intervals where there is a maximum probability of activity or inactivity.

#### *Multi-linear regression analysis (MLRA)*

The general objective of MLRA is to define the relation between two or more independent variables and a dependent variable by providing a linear equation to observe the data. The regression function was calculated by correlating the experimental values of iNOS IC<sub>50</sub> with the TIs, using the software package Statistica version 9.0 [34].

The Furnival–Wilson algorithm was employed to find the best subsets of descriptors and the selection of the regression equation was determined with the minimal Mallows' Cp parameter [37].

#### Models' validation

The predictive quality of the models along with their measures of goodness-of-fit and robustness must be validated using different criteria [32]. As already mentioned, LDA models were tested with an external validation (*test set*), which consists in a set of compounds that were not used for the construction of the QSAR model.

On the contrary, MLR models were submitted to internal validation or cross-validation with leave-one-out (LOO) and data randomization or Y-scrambling. The L-O-O consists in eliminate one case from the original data set and launch again the regression analysis with the ( $N-1$ ) remaining cases and the previous descriptors (the TIs selected during the first regression). The new property is then calculated. The process is replicate, as many times as there are cases in the

data. The value of the prediction coefficient,  $Q^2$ , expresses the quality of the prediction function ( $Q^2$  value  $> 0.7$  indicates goodness-of-fit of the regression equation obtained).

In the test of randomness, the property values for each compound were randomized and then were linearly correlated with the TIs. Then,  $R^2$  and  $Q^2$  values were calculated for the models. The process is repeated several times, finding that the quality of the regression is, in all cases, much lower than that obtained when each compound is assigned its correct experimental value.

Presence of outliers has been detected for compounds with standardized residuals greater than two standard estimation errors.

### Topological virtual screening

The topological models resulting from  $DF_1$ - $DF_3$  functions and  $IC_{50}$  iNOS equation were used to predict whether the pre-selected compounds (AMA, Calcein, DLT and Ro 41-0960) could be active on a DSS-induced colitis assay and potential iNOS and TNF- $\alpha$  inhibitors. The aforementioned compounds were selected after a topological virtual screening on Sigma-Aldrich and the Micro-Source Pure Natural Products Collection databases [25-26].

### **In vitro study**

#### Material

Unless specified otherwise, all chemicals including products and biochemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA). ELISA kit for TNF- $\alpha$  and IL-1 $\beta$  cytokines was purchased from eBioscience (San Diego, CA, USA).

#### Assays

##### *Cell culture and treatment*

The murine macrophage cell line RAW 264.7 (ECACC, Salisbury, UK) between passages 15 and 30 was maintained in Dulbecco's Modified Eagle's Medium (DMEM) glucose concentration 1 g/L supplemented with 10% fetal bovine serum, penicillin (100 U/mL), and streptomycin sulfate (100  $\mu$ g/mL) in an atmosphere of 5% CO<sub>2</sub>/ 95% air (v:v) at 37 °C.

Macrophage RAW 264.7 was plated at a density of  $1 \times 10^6$  cells per mL in a 96-well plates for determine nitrite production. Then, 24 h after seeding the cells, the medium was replaced with fresh DMEM 0.5 % FBS, and the cells were incubated with graded amounts of compound and with or without lipopolysaccharide LPS (1  $\mu$ g/ml) for 24 h at 37 °C.

Macrophage RAW 264.7 was plated at a density of  $5 \times 10^5$  cells per mL in 24-well plates for ELISA assays. Then, 24 h after seeding the cells, the medium was replaced with fresh DMEM 0.5 % FBS. After this, cells were treated with the four selected compounds at different concentrations (10, 25, 50, and 100  $\mu\text{M}$ ) and then, 2 h later cells were stimulated with 1  $\mu\text{g/mL}$  LPS. Finally, 24 h after stimulation, supernatants were collected for ELISA analysis.

#### *Cytotoxicity assay*

The effect of compounds on cell viability was evaluated with the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl-tetrazolium bromide (MTT) [38] assay as reported on a previously publication [26].

#### *Nitrite ( $\text{NO}_2$ ) Production in RAW 264.7 cells*

The NO production in culture supernatant was determined by spectrophotometry. The level of nitrite (an oxidative product of NO) was calculated using the Griess reaction [39]. The reaction consists in mixing 100  $\mu\text{L}$  of culture supernatant with 100  $\mu\text{L}$  of Griess reagent solution. The absorbance was measured at 540 nm with a microplate reader against a calibration curve with sodium nitrite standards. Compounds were assayed at 10, 50 and 100  $\mu\text{M}$ . In this experiment, nitrite production was assessed as the index of iNOS activity.

#### *TNF- $\alpha$ and IL-1 $\beta$ production in RAW 264.7 cells*

The supernatants media were collected and TNF- $\alpha$  and IL-1 $\beta$  production was evaluated using a specific enzyme-linked immunosorbent assay kit (eBioscience, San Diego, CA, USA) in accordance with the manufacturer's instructions. TNF- $\alpha$  and IL-1 $\beta$  secretion was quantified in pg/mL using the standard provided with the kit with sensitivity limits of 8 pg/mL. Results were subsequently expressed in relative terms to the negative control (untreated cells) to facilitate comparison between groups.

### Statistical analysis

Statistical analysis one-way analysis of variance (ANOVA) and Dunnett's *t* test were performed. GraphPad Prism 4.0 software (GraphPad Software Inc., San Diego, CA, USA) was used for all the calculations. Results of three or more independent experiments (each performed in triplicates) were calculated as mean  $\pm$  SEM of percent NO and cytokine levels relative to control (DMSO vehicle).

### ***In vivo* study**

#### Materials

Unless specified otherwise, all chemicals and biochemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA). Dextran sulfate sodium (36–50 kDa) was purchased from MP Biomedical (Illkirch Cedex, France).

#### Animals

The Institutional Ethics Committee of the University of Valencia (Spain) approved all animal care and experimental protocols. Female BALB/c mice (Harlan Interfauna Iberica, Barcelona, Spain), 6–8 weeks of age weighing 18–20 g were acclimatized under a 12 h light/dark cycle at 22 °C and 60% humidity, for 7 days before the experiments, and fed with a standard laboratory rodent diet and water *ad libitum*.

### Induction of DSS colitis and treatment.

Induction of DSS colitis is an experimental model reproducing some of the characteristic symptoms of the disease in humans such as weight loss, diarrhea, blood in stool and histological changes similar to those experimented in UC patients [40]. Therefore, this model allows evaluating the effect of new drugs for UC pharmacotherapy.

Acute colitis was induced by supplementing drinking water with 3 % (w/v) DSS for 7 days. Mice were randomly assigned to one of the four treatment groups: control (drinking water), DSS group (3% DSS in drinking water), DSS+AMA group; DSS+Calcein group; DSS+DLT group; DSS+Ro 41-0960 group and Sulfasalazine. During the experiment, animals consumed three oral doses (50 mg/kg) of the tested compounds in a volume of 0.2 mL of water. As compounds were tested at one single dosage, reference drug (Sulfasalazine)'s dosage was used to compare potency and activity.

At day 8 mice were sacrificed by cervical dislocation and their colons removed. Three colon samples were prepared for histological analysis as described later, and the rest were rinsed with cold PBS, blotted dry and weighed. Their length was measured. Then, samples were frozen immediately in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use.



### *Disease Activity Index (DAI)*

The extent of colitis was assessed using a composite disease activity index (DAI) [41]. DAI is determined by scoring changes in the animal weight (body weight was monitor three times a week), gross rectal bleeding (measured by visible fecal blood and macroscopic examination of the anus) and stool consistency (Table 1). Values of DAI score range from 0 to 4. Zero represents no symptoms of UC while four represent full-blown disease.

**Table 1.** Scoring system to calculate the Disease Activity (DAI) [41].

DAI <sup>a</sup>			
Score	Weight loss	Stool consistency	Visible blood in feces
0	None		
1	1-5%	None	None
2	6-10%		
3	10-20%	Loose	Slight bleeding
4	>20%	Diarrhea	Gross bleeding

<sup>a</sup>DAI score were calculated as the mean of these three factors.

### *Histological Analysis.*

Colon samples were used for the histopathological study, to identify microscopically signs of the disease progression. The intestines were slice along the vertical axis and rolled up in a spiral, beginning from the duodenal margin. Subsequently, guts were fixed in 4% buffered formaldehyde, paraffin embedded, sectioned and stained with hematoxylin and eosin. The gravity of colitis was measured using a histological disease score. Eight randomly selected fields (magnification  $\times 100$ ) in each section were observed and graded as above in a blinded fashion by an expert pathologist.

## RESULTS AND DISCUSSION

### Mathematical modeling

#### Modeling iNOS inhibition

A mathematical model was built from a *training set* of 433 compounds (iNOS inhibitors and no inhibitors). The *training set* consisted in a group of heterogeneous molecular structures. Active and inactive groups were balanced so that both had the same weight, regardless the number of compounds.

The equation corresponding to  $DF_1$ , discriminates compounds that were predicted as iNOS inhibitors. It is comprised of six variables:

$$DF_1 = (1.40 * PHI) - (156.05 * X1A) + (90.70 * X3A) + (9.63 * ATS1p) - (15.76 * ESpm02u) + (1.74 * ESpm13x) + 51.71 \quad (2)$$

Statistical parameters for the significance of the equation were:

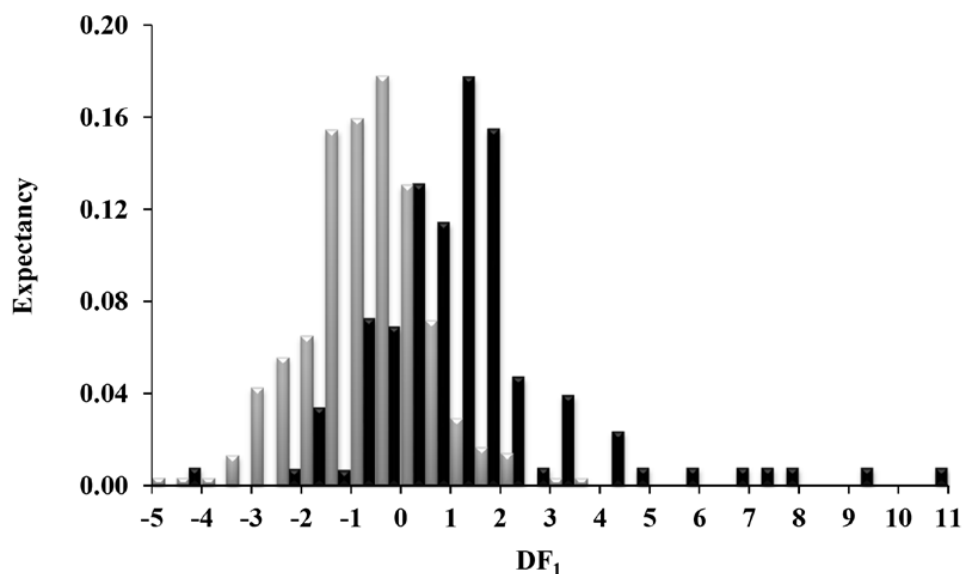
$$N = 433, F = 33.68, \lambda = 0.68, p < 0.00001$$

Where  $N$  is the number of data compounds;  $F$ , Fisher–Snedecor parameter;  $\lambda$ , Wilks' lambda;  $p$ , statistical significance.

Among the topological descriptors stand out: Kier flexibility index (PHI); average connectivity index of order 1 (X1A) and order 3 (X3A); Broto-Moreau autocorrelation of lag 1 (log function) weighted by polarizability (ATS1p); spectral moment 02 from edge adjacency matrix (Espm02u) and spectral moment 13 from edge adjacency matrix weighted by edge degrees (Espm13x).

Between descriptors of the Eq.2, some are particularly noteworthy in terms of activity, as the molecular flexibility (measured by the PHI index), which is important for the proper conformation for interaction with the enzyme, as well as the polarizability (index ATS1p), which highlights the importance of intramolecular electron mobility. Of particular interest is the influence of some purely topological descriptors, like ESpm02u and ESpm13x, which represent spectral moments of adjacency matrices. In previous studies it has been highlighted [42] that certain molecular scaffolds and "frameworks" related to these descriptors, appear much more frequently in drug-like compounds.

According to  $DF_1$ , a compound is classified as iNOS inhibitor if  $DF_1$  takes values from 0.5 to 11.5; on the other hand, a compound will be labeled as inactive if  $DF_1$  takes values from  $<0.5$  to  $>-5$  (see Figure 1). Finally, molecules scoring  $DF_1 > 11.5$  or  $DF_1 < -5$  will be considered as outliers.



**Figure 1.** Pharmacological distribution diagram for iNOS inhibitors obtained using the  $DF_1$  (the black color represents iNOS inhibitors and the grey color, the compounds without iNOS inhibition activity).

This classification criterion was applied to the *training set*. Of the 433 compounds, 98 out of 127 iNOS inhibitors were correctly classified (77% sensitivity) along with 257 out of 306 inactive compounds (84% specificity) (see Table 2). The best way to evaluate the quality of a DF is to perform an external validation (*test set*). In this case, the *test set* was comprised 108 compounds (32 active and 76 inactive), which were not obviously included in the *training set*. *Test set* was randomly built, using the 20% of the whole data set. The model gave a correct classification of 75% for the active group (24 out of 32 compounds) and 83% for the inactive set (63 out of 76 compounds) (see Table 2). Furthermore, although the strict application of  $DF_1$  leads to the loss of some actives (25%), the point is that the 83% of the inactive compounds were correctly classified. This way, the risk of including “false active” is significantly reduced, which means high specificity.

**Table 2.** Classification matrix obtained through the selected DF<sub>1</sub>, DF<sub>2</sub> and DF<sub>3</sub> for *training* and *test set*.

Set	A	I	TP	FP	TN	FN	Sensitivity (%)	Specificity (%)
<b>DF<sub>1</sub> Eq. (2)</b>								
<i>Training</i>	127	306	98	49	257	29	77	84
<i>Test</i>	32	76	24	13	63	8	75	83
<b>DF<sub>2</sub> Eq. (4)</b>								
<i>Training</i>	27	162	19	17	145	8	70	90
<i>Test</i>	18	35	12	8	27	6	67	77
<b>DF<sub>3</sub> Eq. (5)</b>								
<i>Training</i>	43	78	39	16	62	4	91	79
<i>Test</i>	18	18	14	3	15	4	78	83

A: number of active compounds, I: number of inactive compounds, TP: true positive, FP: false positive, TN: true negative, and FN: false negative cases.

Matthews's correlation coefficient considers values between  $-1$  and  $+1$ . It was used here for the evaluation of the predictive quality of the model. The equation shows a value of  $0.59$ , what ensures its reliability. Furthermore, in this specific case, MCC was calculated in a slightly different way (i.e., by adding  $+1$  to each scales value, so that the outcome can be expressed as % accuracy). In other words,  $0$  would have mean no correlation at all,  $1$  represents  $50\%$  of correlation and  $2$  stands for maximum correlation ( $100\%$ ). The model's yield was  $80\%$  (MCC modified =  $1.59$ ).

### Modeling iNOS IC<sub>50</sub>

In order to predict the IC<sub>50</sub> value of the new potential anti-UC compounds, a logIC<sub>50</sub> multi-linear regression model for iNOS inhibition has been developed.

The best linear equation, along with its statistical parameters was:

$$\log\text{IC}_{50} = (0.49 \times \text{nC}) - (2.57 \times \text{nR06}) - 0.06 \quad (3)$$

$$N=10 \quad r^2=0.95 \quad F=70.44 \quad SEE=0.51 \quad p<0.00002 \quad Q^2=0.88$$

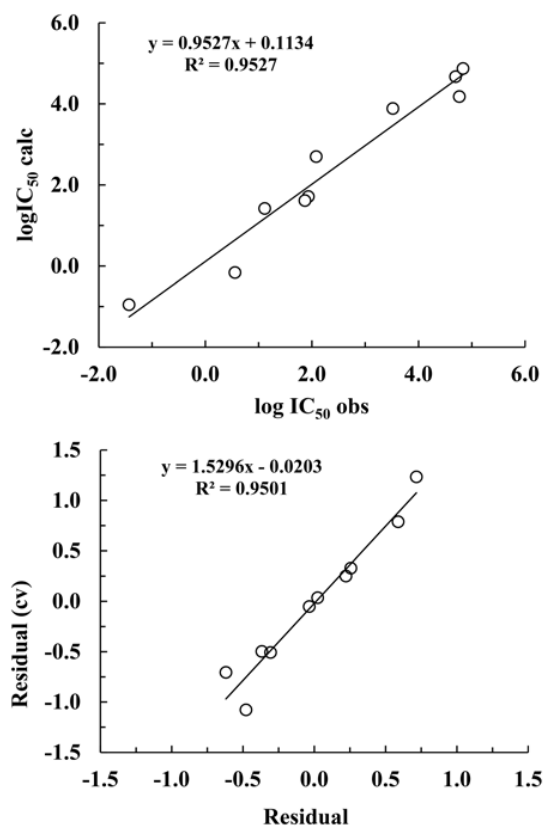
Where  $N$  is the number of data compounds;  $r^2$ , coefficient of determination;  $F$ , Fisher–Snedecor parameter;  $SEE$ , standard error estimation;  $p$ , statistical significance;  $Q^2$ , prediction coefficient. Conformational descriptors evaluating number of carbon atoms (nC) and number of 6-membered rings (nR06) are present.

Values of  $r^2$  and  $Q^2$  (prediction coefficient for the L-O-O cross-validation) next to 1, as well as low values of  $SEE$  (less than 8% of the average values of the property) determine the quality of the models.

Table 3 and Figure 2 (above) show the logIC<sub>50</sub> activity predicted for each compound analyzed. The predictive power was evaluated through cross-validation, using the leave-one-out test ( $Q^2 = 0.88$  for logIC<sub>50</sub> iNOS activity). Results from L-O-O test are shown in Table 3 (column 5) and Figure 2 (below).

**Table 3.** Experimental, predicted and leave-one-out validation test values for multi-linear regression analysis and descriptors value for *training set* compounds.

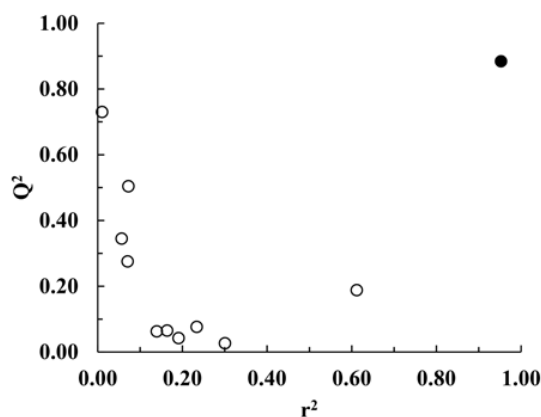
Compound	nC	nR06	log IC <sub>50</sub> obs	logIC <sub>50</sub> cal	logIC <sub>50</sub> cal (cv)
(S)-Ethylisothiourea	3	0	1.11	1.42	1.62
2-[2-(4-Methoxy-2-pyridinyl)ethyl]-1H-imidazo[4,5-b]pyridine	14	2	1.93	1.71	1.68
2-Amino-5,6-dihydro-6-methyl-4H-1,3-thiazine	5	1	0.56	-0.16	-0.68
5-[(4'-Amino-5'8'-difluorospiro[piperidine-4,2'(1'H)-quinaxolin]-1-yl)carbonyl]-2-pyridinecarbonitrile	19	4	-1.43	-0.95	-0.35
BDBM106699	19	2	4.77	4.18	3.98
BDBM111495	20	2	4.7	4.67	4.66
BDBM152702	10	0	4.84	4.87	4.89
BDBM36399	16	2	2.08	2.7	2.78
BDBM50148167	19	3	1.87	1.61	1.54
N6-(1-Iminoethyl)-L-lysine	8	0	3.52	3.89	4.02



**Figure 2.** Prediction of  $\log IC_{50}$  iNOS inhibition. (above) Graphical representation of the observed and calculated  $\log IC_{50}$  values. (below) Graphical representation of the calculated residuals for the training series versus the cross-validation (residual (cv)).

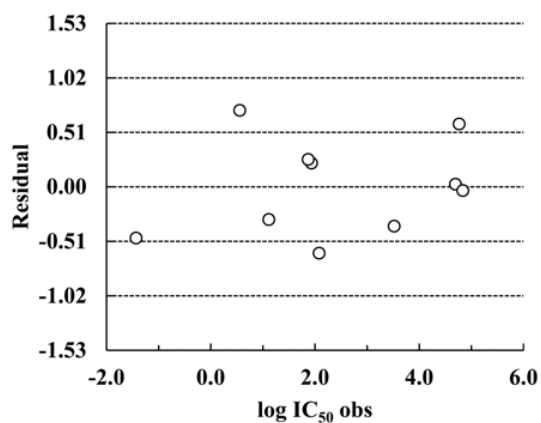
A randomization test was carried out to prevent possible fortuitous regressions. As a result, the property values for each compound were randomized and then were linearly correlated with the TIs. The process is replicate, as many times as there are cases in the data. The usual way to express the results of a randomization test is by comparing the correlation coefficients versus the predicted ones,  $r^2$  and  $Q^2$ . The results, shown in Figure 3, show a sensible stability of the model (all regressions were deficient except for the selected equation (black point) with the real values for each compound).





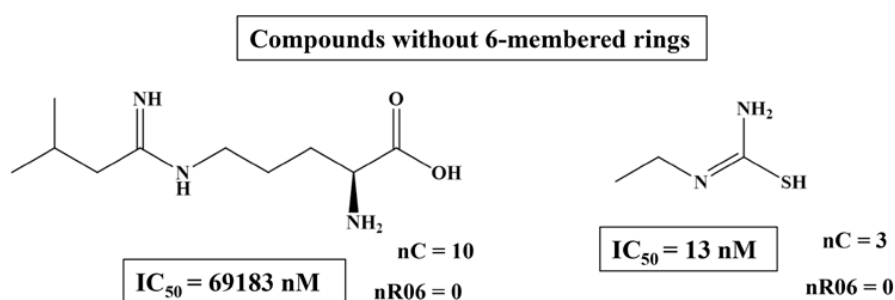
**Figure 3.** Graphic representation of  $Q^2$  (prediction coefficient), versus  $r^2$  (the correlation coefficient) calculated with the randomization test for  $\log IC_{50}$  iNOS inhibition values. Black point: selected model.

Figure 4 depicts the presence of outliers. All the compounds lie within the SEE area: standardized residuals  $\pm$  two standard estimation error. Therefore, the model is applicable to the whole *training set* data.



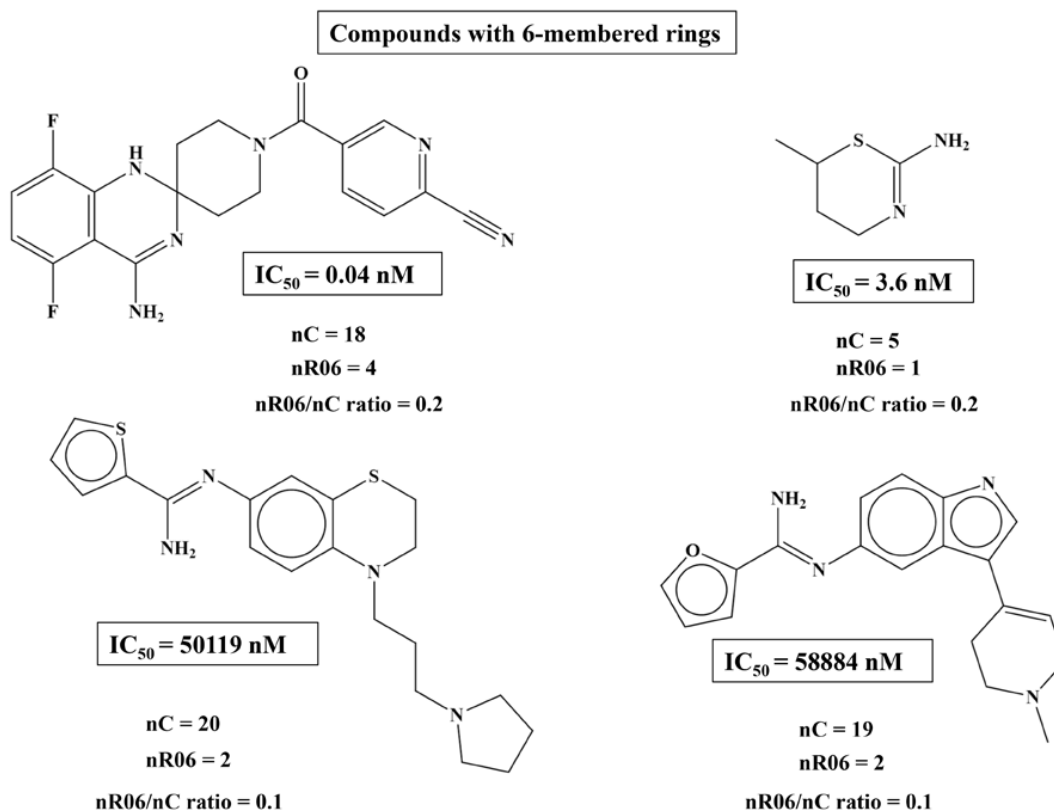
**Figure 4.** Graphic representation of the  $\log IC_{50}$  iNOS inhibition observed values versus residuals values. Outliers are points with standardized residuals greater than two standard estimate error units (SEE=0.51).

Analyzing the results of the equation 3, it is evident that in the absence of 6-member rings (nR06), the smaller is the number of carbons (nC), the greater is the iNOS inhibitory power. This fact is reflected in Figure 5, where (S)-Ethylisothiourea has an  $IC_{50}$ =13 nM while BDBM152702 has an  $IC_{50}$ = 69183 nM.



**Figure 5.** Training set compounds without 6-membered rings and their  $IC_{50}$  iNOS inhibition values.

On the contrary, compounds with presence of 6-member rings and high values of nR06/nC ratio shows improved iNOS inhibitor activity. Again, this fact is reflected in Figure 6, where the most promising iNOS inhibitors has higher nR06/nC ratio.



**Figure 6.** *Training set* compounds with presence of 6-membered rings and their IC<sub>50</sub> iNOS inhibition potency.

### Modeling TNF- $\alpha$ inhibition

A *training set* of 189 compounds with or without TNF- $\alpha$  inhibitory activity was used to build a third topological model. Again, the *training set* consisted in a group of heterogeneous molecular structures. Active and inactive groups were balanced so that both had the same weight, regardless the number of compounds.

The equation corresponding to DF<sub>2</sub>, discriminates compounds that were predicted as TNF- $\alpha$  inhibitors. It is comprised of six variables:

$$DF_2 = (0.39 \times nO) + (0.98 \times piPC10) - (6.79 \times ATS1e) + (11.88 \times MATS1m) - (17.85 \times MATS1e) - (5.41 \times GATS1v) + 18.40 \quad (4)$$

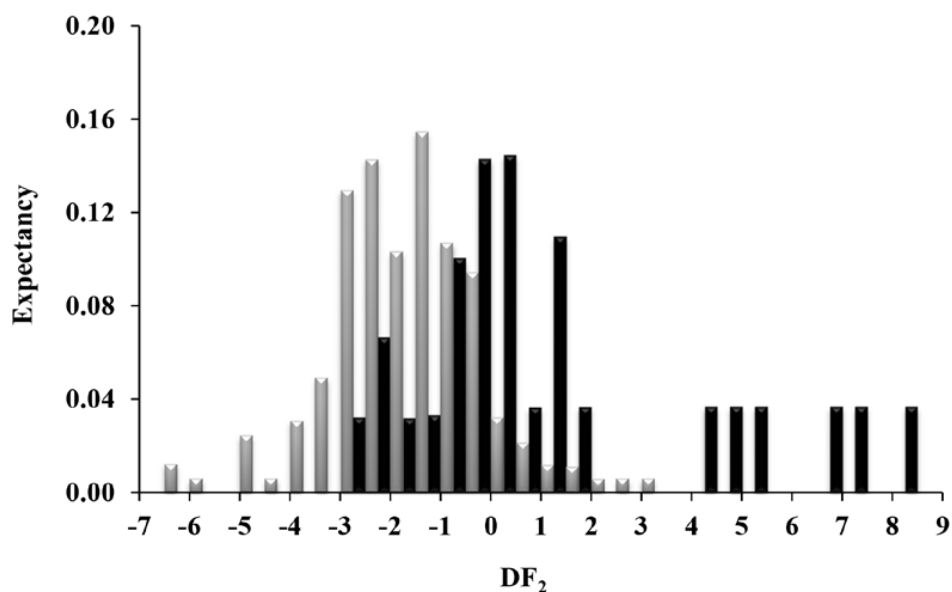
Statistical parameters for the significance of the equation were:

$$N = 189, F = 12.15, \lambda = 0.71, p < 0.00001$$

In the equation above, there are topological descriptors evaluating: molecular multiple path count of order 10 (piPC10); Broto-Moreau autocorrelation of lag 1 (log function) weighted by Sanderson electronegativity (ATS1e); Moran autocorrelation of lag 1 weighted by mass (MATS1m); Moran autocorrelation of lag 1 weighted by Sanderson electronegativity (MATS1e); Geary autocorrelation of lag 1 weighted by van der Waals volume (GATS1v) and number of oxygen atoms (nO).

Unlike what happens in the first discriminant function (DF<sub>1</sub>), in the second (DF<sub>2</sub>) appears indices basically related to the presence of certain heteroatoms in particular locations of the molecule. Thus, factors such as the size and electronegativity of atoms, assessed by indices such as MATS1m or ATS1e, play a decisive role. For example, most molecules exhibit hydrogen donor groups (NH- and OH-), which determine an increase in water solubility as well as a greater capability to interact by means of hydrogen bonds interactions with amino acid residues of TNF- $\alpha$  receptor.

According to DF<sub>2</sub>, a compound is classified as TNF- $\alpha$  inhibitor if DF<sub>2</sub> takes values from 0 to 9; on the other hand, a compound will be labeled as inactive when DF<sub>2</sub> takes values from <0 to >-7 (see Figure 7). Finally, molecules scoring DF<sub>2</sub>>9 or DF<sub>2</sub><-7 will be considered as outliers.



**Figure 7.** Pharmacological distribution diagram for TNF- $\alpha$  inhibitors obtained using the DF<sub>2</sub> (the black color represents TNF- $\alpha$  inhibitors and the grey color, the compounds without TNF- $\alpha$  inhibition activity).

This classification criterion was applied to the *training set*. Of the 189 compounds, 19 out of 27 TNF- $\alpha$  inhibitors were correctly classified (70% sensitivity) along with 145 out of 162 inactive compounds (90% specificity) (see Table 2). The best way to evaluate the quality of a DF is to perform an external validation (*test set*). In this case, the *test set* was comprised 53 compounds (18 active and 35 inactive), which were not obviously included in the *training set*. *Test set* was randomly built, using the 20% of the whole data set. The model gave a correct classification of 67% for the active group (12 out of 18 compounds) and 77% for the inactive set (27 out of 35 compounds) (see Table 2). Furthermore, although the strict application of DF<sub>2</sub> leads to the loss of some actives (33%), the point is that the 77% of the inactive compounds were correctly

classified. This way, the risk of including “false active” is significantly reduced, which means high specificity.

Matthews’s correlation coefficient value for  $DF_2$  was 0.53, what ensures its reliability. And model’s yield was 77% of accuracy (MCC modified = 1.53).

#### Modeling DSS-induced UC protection

Finally, a topological model for the identification of new active compounds against experimental DSS-induced UC was developed. The *training set* consisted in a group of 121 heterogeneous molecular structures. Active and inactive groups were balanced so that both had the same weight, regardless the number of compounds.

The fourth equation, corresponding to  $DF_3$ , discriminates compounds that were predicted to infer protection against DSS-induced colitis (which, as mentioned above, is a model for testing potential UC treatments). It is comprised of six variables:

$$DF_3 = (-1.40 \times nR05) + (1.78 \times piPC03) + (1.73 \times MATS6p) - (4.93 \times GATS4e) - (3.09 \times EEig02d) + (4.48 \times EEig14d) - (3.89 \times EEig14r) + 9.99 \quad (5)$$

Statistical parameters for the significance of the equation were:

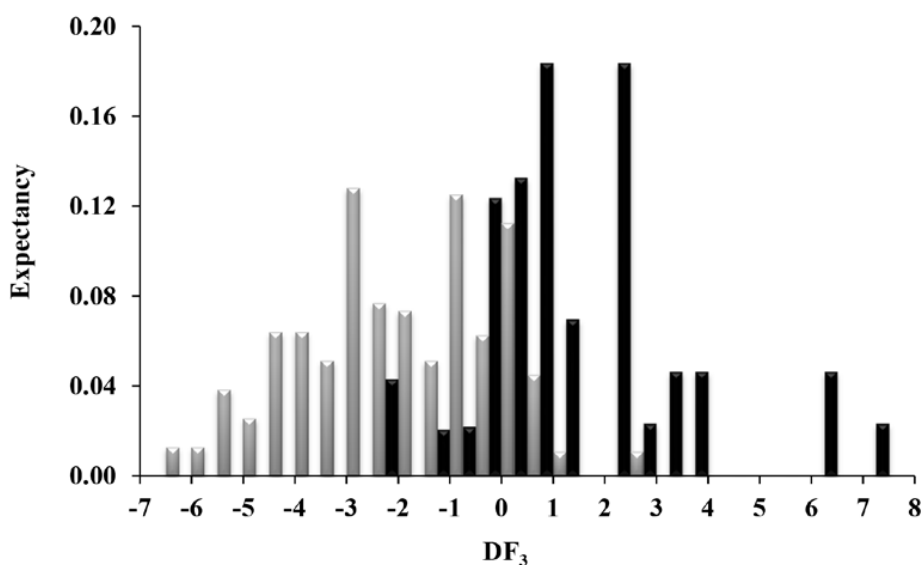
$$N = 121, F = 14.03, \lambda = 0.54, p < 0.00001$$

In the equation above, there are topological descriptors evaluating the number of 5-membered rings (nR05); molecular multiple path count of order 3 (piPC03); Moran autocorrelation of lag 6 weighted by polarizability (MATS6p); Geary autocorrelation of lag 4 weighted by Sanderson electronegativity (GATS4e); eigenvalue 02 and 14 from edge adj. matrix weighted by dipole

moments (EEig02d and EEig14d, respectively) and eigenvalue 14 from edge adj. matrix weighted by resonance integrals (EEig14r).

Multiple factors play a significant role, as for instance molecular polarizability, dipole moments or resonance effects. However, unlike the previous cases, purely structural factors such as the presence of five-membered rings also are important for activity.

According to  $DF_3$ , a compound is classified as active versus the DSS-induced UC if  $DF_3$  takes values from 0 to 8; on the other hand, a compound will be labeled as inactive when  $DF_3$  takes values  $<0$  to  $>-7$  (Figure 8). Finally, molecules scoring  $DF_3 > 8$  or  $DF_3 < -7$  will be non-classified by this model.



**Figure 8.** Pharmacological distribution diagram for compounds active versus DSS-induced UC obtained using the  $DF_3$  (the black color represents active compounds versus a DSS-induced UC model and the grey color, the inactive ones).

This classification criterion was applied to the *training set*. Of the 121 compounds, 39 out of 43 potential UC drugs were correctly classified (91% sensitivity) along with 62 out of 78 inactive compounds (79% specificity) (see Table 2). The best way to evaluate the quality of a DF is to perform an external validation (*test set*). In this case, the *test set* was comprised 36 compounds (18 active and 18 inactive), which were not obviously included in the *training set*. *Test set* was randomly built, using the 20% of the whole data set. The model gave a correct classification of 78% for the active group (14 out of 18 compounds) and 83% for the inactive set (15 out of 18 compounds) (see Table 2). Furthermore, although the strict application of DF<sub>3</sub> leads to the loss of some actives (22%), the point is that the 83% of the inactive compounds were correctly classified. This way, the risk of including “false active” is significantly reduced, which means high specificity.

DF<sub>3</sub> shows a MCC value of 0.68, what ensures its predictive quality. And when, MCC was calculated in a slightly different way it was obtained a model's yield of 84% accuracy (MCC modified = 1.68).

### **Topological virtual screening**

Based on the mathematical models (Eq. 2-5), a virtual screening was performed on a group of 4 molecules with potential activity against UC as inhibitors of NF-kB and IL-6 (Alizarin-3-methyliminodiacetic acid (AMA), Calcein, (+)-dibenzyl-l-tartrate (DLT) and Ro 41-0960) [25-26]. This *in silico* study has been performed to predict their potential inhibitory activity against iNOS, TNF- $\alpha$  and a DSS-induced UC model. As shown in Table 4, all 4 compounds were classified as active under the aforementioned mechanisms.



**Table 4.** Values of  $DF_{1-3}$  and  $\log IC_{50}$  (iNOS): probability of activity and classification of the potential anti UC compounds.

Compound	iNOS LDA model			iNOS MLR model		TNF- $\alpha$ LDA model			DSS-induced UC model		
	$DF_1$	<i>Prob (activ.)</i>	<i>Class.</i>	$\log IC_{50}$	$IC_{50}$ (nM)	$DF_2$	<i>Prob (activ.)</i>	<i>Class.</i>	$DF_3$	<i>Prob (activ.)</i>	<i>Class.</i>
AMA	0.77	0.67	A	1.54	35	1.09	0.75	A	2.25	0.91	A
Calcein	2.1	0.88	A	4.36	22909	0.33	0.58	A	3.64	0.97	A
DLT	3.42	0.97	A	3.62	4169	-1.16	0.24	I	0.09	0.52	A
Ro 41-0960	1.27	0.77	A	1.17	15	0.69	0.67	A	0.23	0.56	A

## In vitro study

### Cytotoxicity assay

In a previous paper, cell viability of RAW 264.7 macrophages treated with the compounds was reported [26]. All compounds were assayed at 100  $\mu$ M (cell viability >100%), except Ro 41-0960 which was tested at 25  $\mu$ M in RAW 264.7 cells (cell viability of 78%) [26].

### Nitrite (NO<sub>2</sub>-), IL-1 $\beta$ and TNF- $\alpha$ production in RAW 264.7 cells

RAW264.7 was exposed to AMA, Calcein, DLT, and Ro 41-0960 to investigate the inhibitory effect on the iNOS activity (Table 5). The assay consisted in the detection of nitrite by Griess reaction [38]. In addition, due to the fact that IL-1 $\beta$  production is linked with NF-kB activity, the 4 previous NF-kB inhibitors selected as potential UC treatments were tested to evaluate their IL-1 $\beta$  inhibitory activity on RAW 264.7 cells as well (Table 5). Finally, as TNF- $\alpha$  is a pro-inflammatory cytokine linked to the NF-kB pathway and therefore related to UC. AMA, Calcein, DLT and Ro 41-0960 were tested to determine their effect on the production of TNF- $\alpha$  in RAW 264.7 macrophages (see Table 5).

**Table 5.** NO, IL-1 $\beta$  and TNF- $\alpha$  secretion in RAW 264.7 cultured medium exposed to the inflammatory stimuli LPS (1  $\mu$ g/mL), 2 h after products administration.

Group	C ( $\mu$ M)	NO determination ( $\mu$ M)		IL-1 $\beta$ determination (pg/mL)		TNF- $\alpha$ determination (pg/mL)	
		Mean $\pm$ SEM	I (%)	Mean $\pm$ SEM	I (%)	Mean $\pm$ SEM	I (%)
<b>Control</b>		25.41 $\pm$ 0.98	-	23.94 $\pm$ 2.36	-	11799 $\pm$ 519	-
	10	21.38 $\pm$ 1.23*	16	43.88 $\pm$ 2.07*	-	10740 $\pm$ 258	9
<b>AMA</b>	50	14.06 $\pm$ 0.64***	45	81.35 $\pm$ 1.85***	-	10200 $\pm$ 192	14
	100	8.63 $\pm$ 0.67***	66	94.29 $\pm$ 7.35***	-	8594 $\pm$ 854	27
<b>Calcein</b>	10	23.41 $\pm$ 1.32	8	17.15 $\pm$ 0.21	28	12787 $\pm$ 750	-
	50	17.50 $\pm$ 0.62***	31	53.28 $\pm$ 2.73***	-	10232 $\pm$ 426	13
	100	14.38 $\pm$ 0.36***	43	64.28 $\pm$ 6.64***	-	8103 $\pm$ 1245	31
<b>DLT</b>	10	24.22 $\pm$ 0.69	5	27.71 $\pm$ 3.93	-	9430 $\pm$ 1483	20
	50	23.54 $\pm$ 0.49	7	31.98 $\pm$ 6.03	-	10921 $\pm$ 898	7
<b>Ro 41-0960</b>	100	24.92 $\pm$ 1.97	2	15.90 $\pm$ 2.25	34	11381 $\pm$ 639	4
	10	13.68 $\pm$ 0.70***	46	22.28 $\pm$ 1.29	7	7929 $\pm$ 1248	33
	25	1.32 $\pm$ 0.40***	95	14.06 $\pm$ 1.30	41	6682 $\pm$ 754*	43

C: assayed concentration of compounds; I: inhibition.

Values were expressed based on the negative control (untreated cells) and present means (NO or IL-1 $\beta$  or TNF- $\alpha$  production)  $\pm$  SEM. Results are representative of at least three independent experiments per group. The differences between the control group (untreated cells) and the other groups were determined by means of one-way analysis of variance (ANOVA) followed by Dunnett's t test (\*p<0.05; \*\*\*p<0.001).

As shown in Table 5, three of the compounds (AMA, Calcein and Ro 41-0960) inhibited NO production of 66, 43 and 95 % respectively at the highest concentration. These compounds inhibited in a concentration-dependent manner.

All compounds except AMA inhibited IL-1 $\beta$  production in RAW 264.7 macrophages at least at one concentration (Table 5). However, none of the compounds have shown statistical significance. Stand out Calcein, with a 28% of inhibition at the lowest concentration (10  $\mu$ M), despite it did not show activity at higher concentrations. This phenomenon could be due to

homologous desensitization of the IL1- $\beta$  receptor. Finally, all the compounds were able to reduce at least the 20% of the TNF- $\alpha$  production in RAW 264.7 cells (AMA, Calcein and Ro 41-0960 in a concentration-dependent way, while DLT in a concentration-independent way).

### ***In vivo* study**

#### **Effect of compounds on DSS-induced colitis**

DAI score was used to determine the severity of the disease. It measures the loss of body weight, the stool consistency and the rectal bleeding. Body weight loss was evaluated at days 0, 4 and 8. Stool consistency and rectal bleeding were both monitored at day 8. The treatment with DSS reduced the body weight, along with the colon length. It also produced severe diarrhea and rectal bleeding in comparison with mice that received regular drinking water (control). Treatment with AMA, Ro 41-0960 and the reference drug Sulfasalazine significantly protect mice from body weight loss (Table 6).

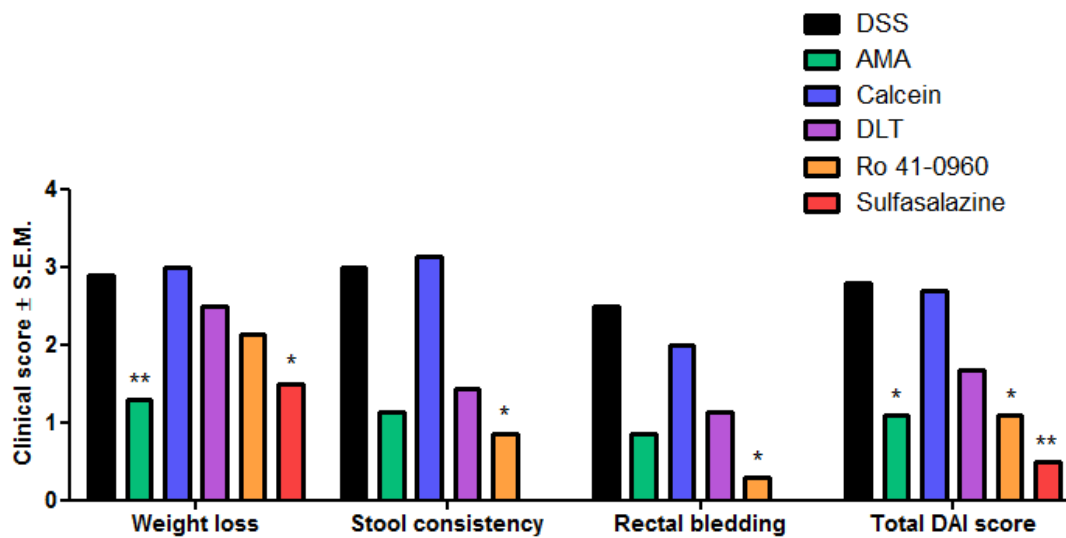
The incidence of death associated with the present experiment could be related to differences in the intestinal microflora between different animals, that is known to alter the outcome of DSS administration [43]. Optimization of the DSS dosage is required in order to avoid high incidence of death in future experiments.

**Table 6.** BALB/c mice's weight loss by DSS-induced UC. Results are expressed based on the negative control and represent means  $\pm$  SEM. Values are representative of an independent experiment per group (n = 5-11 animals).

Group	Body weight evaluation (g)			Body weight loss (%)	N° death animals/total
	Día 0	Día 4	Día 8		
<i>Control</i>	19,04 $\pm$ 0,57	19,24 $\pm$ 0,22	19,68 $\pm$ 0,45 <sup>***</sup>	-	0/6
<i>DSS</i>	18,67 $\pm$ 0,25	18,16 $\pm$ 0,44	15,91 $\pm$ 0,35 <sup>###</sup>	15	2/11
<i>AMA</i>	18,70 $\pm$ 0,28	18,00 $\pm$ 0,58	18,03 $\pm$ 0,42 <sup>*</sup>	3	1/8
<i>Calcein</i>	19,29 $\pm$ 0,23	19,06 $\pm$ 0,24	16,03 $\pm$ 0,64	17	0/7
<i>DLT</i>	18,98 $\pm$ 0,34	18,65 $\pm$ 0,26	16,74 $\pm$ 0,71	12	1/7
<i>Ro 41-0960</i>	18,43 $\pm$ 0,19	18,30 $\pm$ 0,27	17,03 $\pm$ 0,22	8	0/7
<i>Sulfasalazine</i>	17,56 $\pm$ 0,19	17,24 $\pm$ 0,83	16,65 $\pm$ 0,57	5	1/5

Differences between the control group (untreated animals) and the other groups were determined by means of one-way analysis of variance (ANOVA) followed by Dunnett's *t* test. Significant differences with the control group (<sup>###</sup>p<0.001) and with DSS group (<sup>\*</sup>p<0.05; <sup>\*\*\*</sup>p<0.001).

Figure 9 shows the DAI score value for every group. AMA, Ro 41-0960 and the reference compound Sulfasalazine protected to animals of DSS-induced ulcerative colitis (DAI score of 1.1, 1.1 and 0.5 respectively) if compared with DSS group (DAI score value: 2.8). DLT was not able to protect mice against weight loss, but its DAI score value was smaller than the DSS one. Finally, DAI score value for Calcein showed no protective effect against UC.

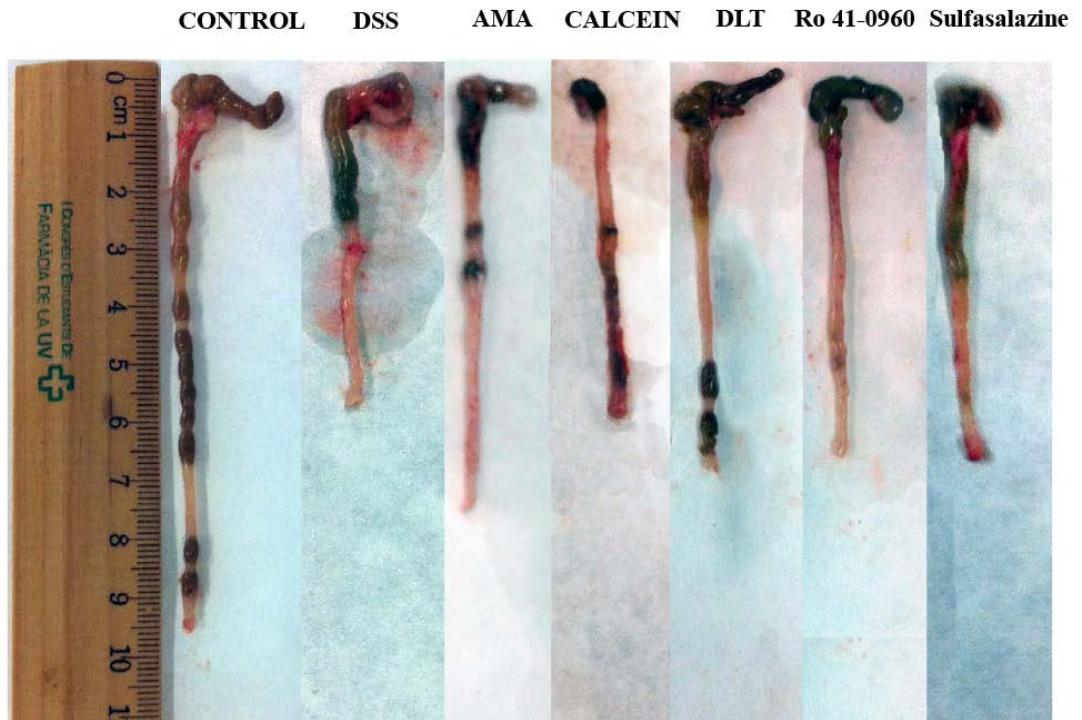


**Figure 9.** DAI score values for the compounds on DSS-induced UC in BALB/c mice. The disease activity index was calculated at the end of the experiment for all the groups according to the Table 1. Significant differences from the DSS group (\* $p < 0.05$ ; \*\* $p < 0.01$ ); one-way analysis of variance followed by Dunnett's *t*-test). Results are representative of an independent experiment per group ( $n = 5-11$  animals) and represent means  $\pm$  SEM.

The macroscopic evaluation of the disease consisted in the measurement of the colon lengths for each mouse of each group (day 8). Table 7 and Figure 10 show how DSS group suffered an evident colon shortening ( $5.4 \pm 0.2$  cm). On the contrary, the 4 compounds prevented this shortening. Stand out AMA, DLT and Ro 41-0960 with a greater protection against colon shortening than the reference drug (Sulfasalazine).

**Table 7.** Effect of compounds on colon length shortening in DSS-induced UC in BALB/c mice. Significant differences from the DSS group (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001) and from control group (####p<0.001); one-way analysis of variance followed by Dunnett's *t*-test). Results are representative of at an independent experiment per group (n = 4-9 animals) and represent means  $\pm$  SEM.

<b>Group</b>	<b>Colon length (cm)</b>
<i>Control</i>	8.8 $\pm$ 0.2 <sup>***</sup>
<i>DSS</i>	5.4 $\pm$ 0.2 <sup>####</sup>
<i>AMA</i>	7.1 $\pm$ 0.4 <sup>**</sup>
<i>Calcein</i>	6.1 $\pm$ 0.3
<i>DLT</i>	6.7 $\pm$ 0.5 <sup>*</sup>
<i>Ro 41-0960</i>	6.8 $\pm$ 0.2 <sup>*</sup>
<i>Sulfasalazine</i>	6.6 $\pm$ 0.5

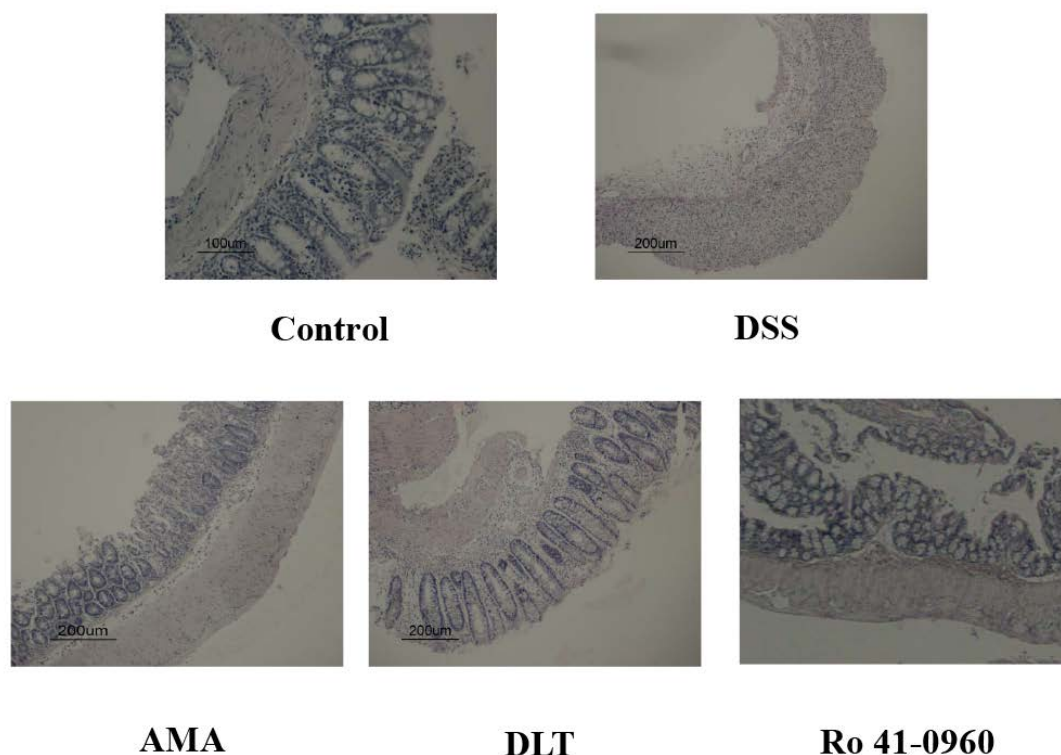


**Figure 10.** Representative image of the effect of compounds on the colon shortening by UC. The control group received only fresh tap water *ad libitum*; the DSS group fresh tap water with 3% DSS for 7 days. The rest of the groups received tap water with 3% DSS and three doses of 50 mg/kg of AMA, Calcein, DLT, Ro 41-0960 and reference drug Sulfasalazine.

## **Histological Evaluation**

The histopathological study of all groups related with DSS-UC are shown in Figure 11; colon cross-sections for control group (only fresh tap water) (Figure control) shows a normal morphology, without inflammation or ulcers, with intact architecture of crypts and physiological cellular infiltration, along with normal *muscularis*. Nevertheless, by analyzing the DSS cross-section, eye-catching changes in colon architecture typically related with the inflammatory process can be observed (Figure 11 DSS), such as total loss of the crypts architecture, loss of the structure of the goblet cells, edema and neutrophil infiltration and ulceration of the mucous membrane. On the contrary, samples from the 3 tested compounds (AMA, DLT and Ro41-0960) showed significant clinical improvement, such as, intact goblet cells, low cellular infiltration, intact crypts architecture and normal *muscularis*.

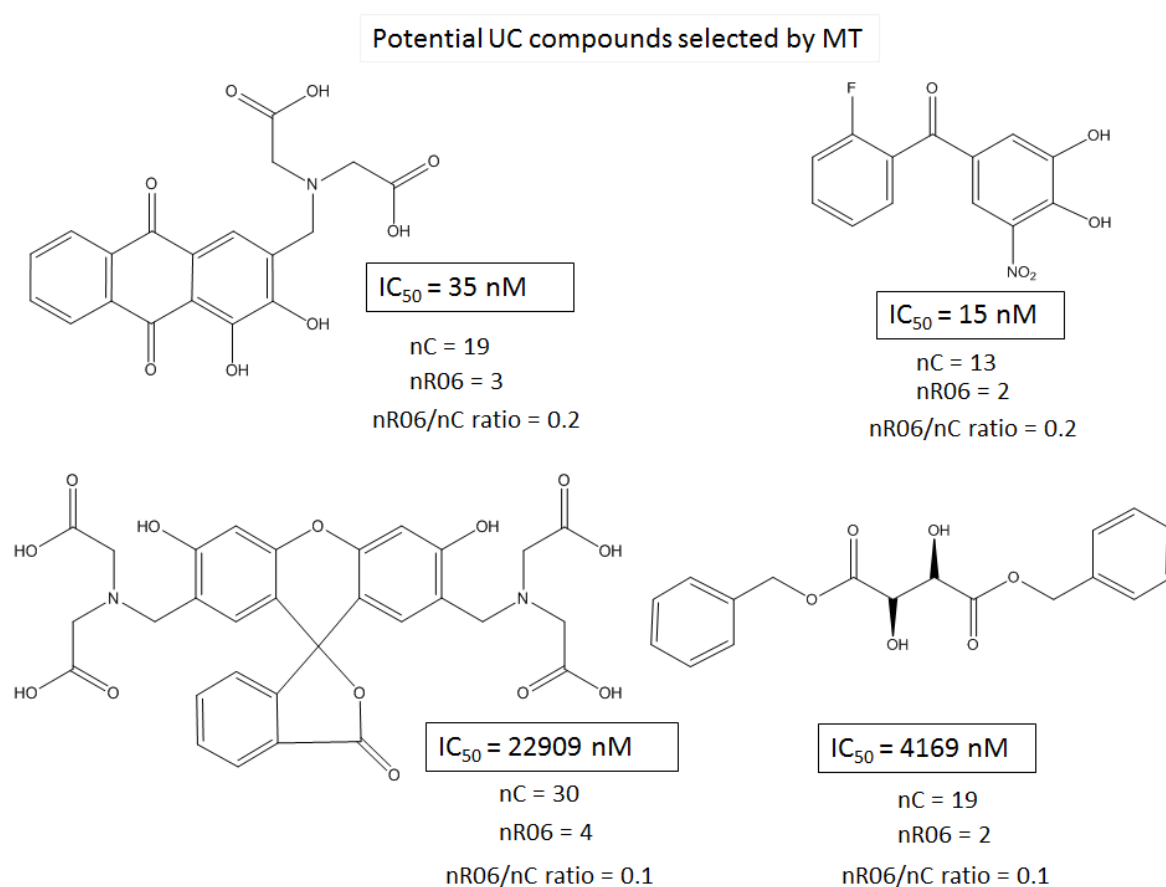




**Figure 11.** Histological evaluation of the protective effect of AMA, DLT and Ro 41-0960 against dextran sulfate sodium (DSS)-induced colitis in mice. Representative histological cross-sections with hematoxylin and eosin staining (200  $\mu\text{m}$ ).

As for the information encoded in the mathematical-topological equations regarding the physicochemical and structural requirements of the selected compounds, because of its high complexity, there are many factors involved. Among the most relevant physicochemical parameters, show up the molecular polarizability, which is a measure of the mobility of the charge within the molecule and the dipole moment, which is a measure of the intramolecular net charge separation. Among the structural requirements, the molecular flexibility as well as the presence of certain moieties, such as 6-member rings, just to mention two significant features.

Indeed, presence of 6-member rings seems to be relevant for the iNOS inhibitory activity in UC. When the nR06/nC ratio for each compound is calculated (Figure 12), as expected, compounds with larger ratios (such as AMA and Ro 41-0960) were predicted as the most potent ones ( $IC_{50}$  values of 35 and 15 nM, respectively). This is also corroborated by *in vitro* results for iNOS inhibition, where AMA and Ro 41-0960 show  $IC_{50}$  values of 53 and 11  $\mu$ M respectively.



**Figure 12.** Selected UC treatments by molecular topology with their iNOS inhibitory  $IC_{50}$  values, as predicted by Eq. (3).

Of course, a much more detailed analysis should be performed to carry out an in depth study of the structure -activity relationships, but this analysis is out of the scope of the present work.

## Conclusions

The present results, along with the ones of previous publications, depicted a Topological Virtual Screening strategy to identify new compounds against UC by inhibiting NF- $\kappa$ B, IL-6, IL-1 $\beta$ , TNF- $\alpha$  and iNOS.

After testing four potential compounds against UC *in vitro* and *in vivo*: AMA, Calcein, DLT and Ro 41-0960; it was demonstrated that AMA and Ro 41-0960 are the best candidate to become potential lead compounds to treat the disease.

AMA reduce NO and TNF- $\alpha$  production *in vitro* and protect from UC in a DSS-induced acute model, by reducing body weight loss, DAI score and preventing from colon shortening. However, Ro 41-0960 it was the best one. This compound was able to inhibit NO, IL-1 $\beta$  and TNF- $\alpha$  production *in vitro*. In addition, Ro 41-0960 was also capable to protect against DSS-induced colitis *in vivo*, by reducing body weight loss, DAI score and preventing from colon shortening.

According to these results, we have demonstrated that, as in many other pharmacological activities, MT stands as a potent tool for the discovery of novel drugs potentially useful for the treatment of ulcerative colitis.

## ASSOCIATED CONTENT

### **Supporting Information**

All the TIs used in the present study along with their values for every compound (*training, test, and virtual screening set*) are shown in the supporting information section, Annex I-IX.

## AUTHOR INFORMATION

### **Corresponding Author**

\* María Gálvez-Llompart

Molecular Connectivity and Drug Design Research Unit, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, Spain

Tel: 9635 43311 // Fax: 9635 44892

[María.galvez@uv.es](mailto:María.galvez@uv.es)

## ACKNOWLEDGMENT

The authors want to express their acknowledgements to PhD Riccardo Zanni from University of Malaga, for his help and useful comments.

## REFERENCES

1. Burisch J, Jess T, Martinato M, Lakatos PL (2013) The burden of inflammatory bowel disease in Europe. *J Crohns Colitis* 7:322-337.
2. Park SC, Jeen YT (2015) Current and emerging biologics for ulcerative colitis. *Gut & Liver* 9.
3. Dignass A, Eliakim R, Magro F, Maaser C, Chowers Y, Geboes K, Mantzaris G, Reinisch W, Colombel J, Vermeire S (2012) Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. *J Crohns Colitis* 6:965-990.
4. Lawrance IC (2015) Early investigational TNF receptor antagonists for the treatment of ulcerative colitis. *Expert Opin Investig Drugs* 24:761-768.
5. Arora Z, Shen B (2015) Biological therapy for ulcerative colitis. *Gastroenterol Rep (Oxf)* 3:103-109.
6. Kolios G, Rooney N, Murphy CT, Robertson DA, Westwick J (1998) Expression of inducible nitric oxide synthase activity in human colon epithelial cells: modulation by T lymphocyte derived cytokines. *Gut* 43:56-63.
7. Saijo F, Milsom AB, Bryan NS, Bauer SM, Vowinkel T, Ivanovic M, Andry C, Granger DN, Rodriguez J, Feelisch M (2010) On the dynamics of nitrite, nitrate and other biomarkers of nitric oxide production in inflammatory bowel disease. *Nitric Oxide* 22:155-167.
8. Gionchetti P, Campieri M, Belluzzi A, Tampieri M, Bertinelli E, Brignola C, Ferretti M, Boni P, Iannone P, Miglioli M (1992) Interleukin 1 $\beta$  (IL-1 $\beta$ ) release from fresh and cultured colonic mucosa in patients with ulcerative colitis (UC). *Agents Actions* 36:C50-C52.
9. Arai Y, Takanashi H, Kitagawa H, Okayasu I (1998) Involvement of interleukin-1 in the development of ulcerative colitis induced by dextran sulfate sodium in mice. *Cytokine* 10:890-896.
10. Fiorino G, Peyrin-Biroulet L, Repici A, Malesci A, Danese S (2011) Adalimumab in ulcerative colitis: hypes and hopes. *Expert Opin Biol Ther* 11:109-116.
11. Atreya I, Atreya R, Neurath M (2008) NF- $\kappa$ B in inflammatory bowel disease. *J Intern Med* 263:591-596.

12. Delhalle S, Blasius R, Dicato M, Diederich M (2004) A Beginner's Guide to NF- $\kappa$ B Signaling Pathways. *Ann N Y Acad Sci* 1030:1-13.
13. García-Domenech R, Gálvez-Llompарт M, Zanni R, Recio MC, Gálvez J (2013) QSAR methods for the discovery of new inflammatory bowel disease drugs. *Expert Opin Drug Discov* 8:933-949.
14. Gálvez J, Gálvez-Llompарт M, García-Domenech R (2012) Molecular topology as a novel approach for drug discovery. *Expert Opin Drug Discov* 7:133-153.
15. Galvez-Llompарт M, Giner M, Recio C, Candeletti S, Garcia-Domenech R (2010) Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity. *Lett Drug Des Discov* 7:438-445.
16. Galvez-Llompарт M, Zanni R, García-Domenech R (2011) Modeling Natural Anti-Inflammatory Compounds by Molecular Topology. *Int J Mol Sci* 12:9481-9503.
17. Garcia-Domenech R, Zanni R, Galvez-Llompарт M, Galvez, J. (2015) Predicting antiprotozoal activity of benzyl phenyl ether diamine derivatives through QSAR multi-target and molecular topology. *Mol Divers* 19:357-366.
18. Caboni L, Galvez-Llompарт M, Gálvez J, Blanco F, Rubio-Martinez J, Fayne D, Lloyd DG (2014) Molecular Topology Applied to the Discovery of 1-Benzyl-2-(3-fluorophenyl)-4-hydroxy-3-(3-phenylpropanoyl)-2 H-pyrrole-5-one as a Non-Ligand-Binding-Pocket Antiandrogen. *J Chem Inf Model* 54:2953-2966.
19. Garcia-Domenech R, Zanni R, Galvez-Llompарт M, Vicente de Julian-Ortiz J (2013) Modeling anti-allergic natural compounds by molecular topology. *Comb Chem High Throughput Screen* 16:628-635.
20. de Gregorio Alapont C, Garcia-Domenech R, Galvez J, Ros M, Wolski S, Garcia M (2000) Molecular topology: a useful tool for the search of new antibacterials. *Bioorg Med Chem Lett* 10:2033-2036.
21. Gálvez J, Llompарт J, Pal K (2004) N,N-dicyclohexyl-(1S)-isoborneol-10-sulfonamide (MT103) and related compounds for the treatment of cancer. Patent Application Country: Application: US; Patent Country: US; Priority Application Country: US Patent 2004266732.

22. Llompart J, Galvez J, Pal K (2006) Treatment of cancer with MT477 derivatives. US Pat Appl Publ 2005-179796; 2004-60586519: 57. Patent Application Country: Application: US; Patent Country: US; Priority Application Country: US Patent 2006014770.
23. Mazier D, Mahmoudi N, Farhati K, Garcia-Domenech R, Galvez J, Derouin F, Danis M (2009) Monensin derivatives for preventing and treating plasmodium infections. Patent Application Country: Application: WO 2009074649, A1.
24. Galvez J, Llompart J, Land D, Pasinetti G (2010) Compositions for Treatment of Alzheimer's Disease Using Abeta-Reducing and/or Abeta-Anti-Aggregation Compounds. Mount Sinai School of Medicine of New York University, USA, Inc Medisyn Technologies, editors. Patent Application Country: Application: WO PCT Int Appl. 2010-US1041; 2009-61211821: 184.
25. Gálvez-Llompart M, Recio MC, García-Domenech R (2011) Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B. *Mol Divers* 15:917-926.
26. Galvez-Llompart M, Recio MC, Gálvez J, García-Domenech R (2013) Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW 264.7 cell lines. *Mol Divers* 17:573-593.
27. Meng F, Lowell CA (1997) Lipopolysaccharide (LPS)-induced macrophage activation and signal transduction in the absence of Src-family kinases Hck, Fgr, and Lyn. *J Exp Med* 185:1661-1670.
28. García-Domenech R. DESMOL11 software. Unidad de Investigación de Diseño de Fármacos y conectividad Molecular. Facultad de Farmacia. Universitat de Valencia. Spain.
29. Talete S (2006) Dragon for windows (software for molecular descriptor calculations), version 5.4.
30. Izenman AJ (2013) Linear discriminant analysis. In: *Modern Multivariate Statistical Techniques*, Springer, pp 237-280.
31. Chatterjee S, Hadi AS (2015) In: *Regression analysis by example*, John Wiley & Sons, Chapter 3, pp 57-69.
32. Gramatica P (2007) Principles of QSAR models validation: internal and external. *QSAR Comb Sci* 26: 694-701.

33. Todeschini R, Consonni V (2008) Handbook of molecular descriptors. Wiley-VCH Verlag GmbH, Weinheim. doi:[10.1002/9783527613106.oth2](https://doi.org/10.1002/9783527613106.oth2).
34. Statistica (data analysis software system) (2009) version 9.0; Statsoft: Tulsa, OK.
35. Matthews BW (1975) Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Bba-Proteins Struct* 405:442-451.
36. Gálvez J, García-Domenech R, de Gregorio Alapont C, de Julián-Ortiz J, Popa L (1996) Pharmacological distribution diagrams: a tool for de novo drug design. *J Mol Graph* 14:272-276.
37. Hocking R (1972) Criteria for selection of a subset regression: which one should be used? *Technometrics* 14:967-976.
38. Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 65:55-63.
39. Tarpey MM, Wink DA, Grisham MB (2004) Methods for detection of reactive metabolites of oxygen and nitrogen: in vitro and in vivo considerations. *Am J Physiol Regul Integr Comp Physiol* 286:R431-44.
40. Okayasu I, Hatakeyama S, Yamada M, Ohkusa T, Inagaki Y, Nakaya R (1990) A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice. *Gastroenterology* 98, 694-702.
41. Giner E, Andújar I, Recio MC, Ríos JL, Cerdá-Nicolás JM, Giner RM (2011) Oleuropein ameliorates acute colitis in mice. *J Agric Food Chem* 59:12882-12892.
42. Lipkus AH, Yuan Q, Lucas KA, Funk SA, Bartelt Iii WF, Schenck RJ, Trippe AJ (2008) Structural diversity of organic chemistry. A scaffold analysis of the CAS Registry. *J Org Chem* 73:4443-4451.
43. Kim JJ, Shajib MS, Manocha MM, Khan WI (2012) Investigating intestinal inflammation in DSS-induced model of IBD. *JoVE (Journal of Visualized Experiments)* e3678-e3678.



# RESULTADOS Y DISCUSIÓN



## VI. RESULTADOS Y DISCUSIÓN

### VI.1. CAPÍTULO 1

Artículo 1: “**Application of Molecular Topology to the Search of Novel NSAIDs: Experimental Validation of Activity**” (En: *Letters in Drug Design & Discovery*, **2010**, 7, 438-445). JCR [2010, CAT: Medicinal Chemistry; FI: 0, 668; Rank: 48 de 54 (**Q4**)].

Los fármacos antiinflamatorios no esteroideos (AINE) destacan entre los más utilizados en todo el mundo. Sólo en los EE.UU., más de 30 mil millones de dosis de AINE se consumen anualmente, lo que se debe a su triple acción: analgésica, antiinflamatoria y antipirética (Wiegand y Tarabar, 2012). Este grupo terapéutico se emplea en distintas enfermedades asociadas con procesos inflamatorios tal y como se ha comentado en el apartado de Introducción. Además, algunos AINE como Aspirina<sup>®</sup>, sulindaco, indometacina y el inhibidor selectivo de COX-2, celecoxib, se pueden utilizar como agentes preventivos del cáncer colorrectal asociado a CU (Stolfi y cols., 2013). Sin embargo, los antiinflamatorios inhibidores no selectivos de la COX se asocian a los efectos adversos de tipo gastrointestinal, mientras que para los selectivos de COX-2, los efectos negativos son cardiovasculares (Ng y Chan, 2010). Por ello, la búsqueda de nuevos AINE con un perfil más seguro presenta un gran interés en la industria farmacéutica.

Como se ha indicado en la Introducción, la experiencia de la UDF en el descubrimiento de nuevos fármacos mediante modelos matemáticos junto con la dilatada trayectoria del grupo “Antiinflamatorios de origen natural” (NAPRODIN) del Departamento de Farmacología, permitió la gestación del Capítulo 1 de la presente Tesis Doctoral. Dicho artículo se centra en el desarrollo y validación de un modelo QSAR basado en la Topología molecular capaz de identificar nuevos agentes antiinflamatorios, fundamentalmente inhibidores no selectivos de la COX.

## Modelo matemático

Para la obtención del modelo matemático-topológico se aplicó la técnica estadística del ALD. El grupo de moléculas empleadas para la construcción del modelo (grupo de entrenamiento o *training set*), está conformado por 48 compuestos distribuidos en dos subgrupos: activos como agentes antiinflamatorios (21 moléculas) e inactivos (27 moléculas que no presentan o no está descrita actividad antiinflamatoria para ellas en la literatura). Las moléculas pertenecientes al grupo de activas tienen en común su actividad como inhibidoras de la COX, mientras que las pertenecientes al grupo de inactivas presentan diversas actividades farmacológicas: antibióticas, anticancerosas, antidiabéticas, anticolinérgicas, antihipertensivas, vasodilatadoras e incluso radiofarmacéutica, entre otras.

La función discriminante (FD) seleccionada fue:

$$FD_1 = (3,36 \times nCL) - (11,49 \times MATS8e) + (19,67 \times PCR) - 27,99 \quad (1)$$

$$N= 48 \quad F= 32,0 \quad \lambda = 0,314 \quad p < 0,00002$$

Donde:

N: número de compuestos empleados en la elaboración del modelo.

F: parámetro Fisher–Snedecor.

$\lambda$ : lambda de Wilks.

p: valor del p estadístico.

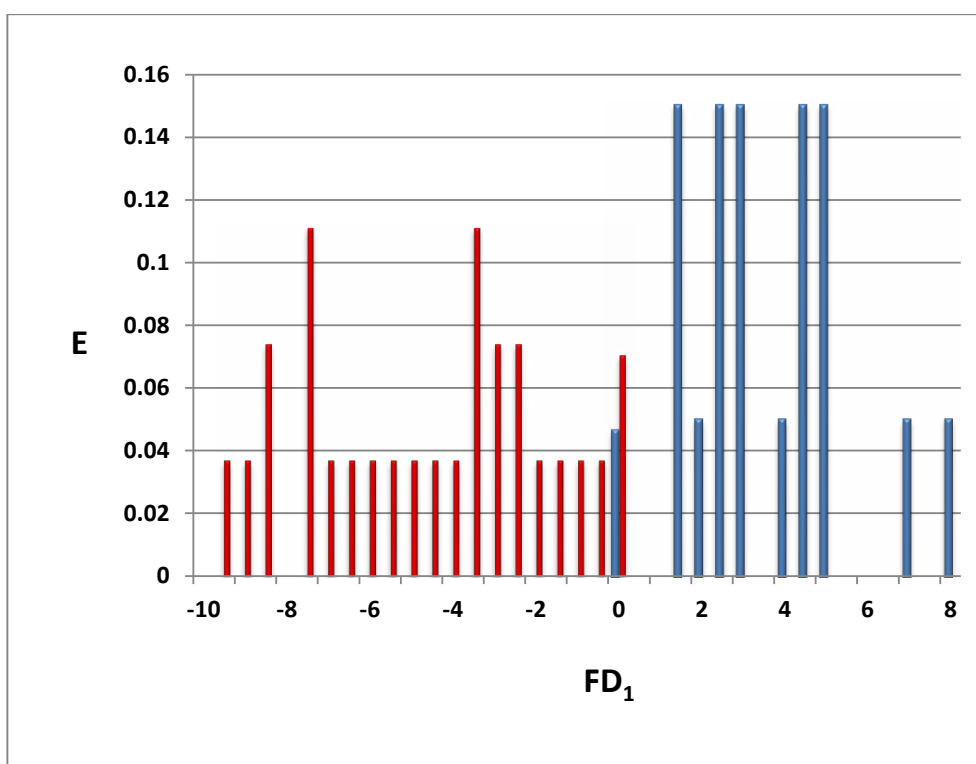
nCL: números de átomos de cloro.

MATS8e: índice de autocorrelación de Moran a distancia 8 ponderado por la electronegatividad de Sanderson.

PCR: cociente entre el número de enlaces múltiples y el número total de enlaces.

Los parámetros estadísticos que acompañan a la Ecuación 1, demuestran la elevada significación estadística del modelo (valor de p) así como de los índices que lo conforman (F). Generalmente valores de  $\lambda$  por debajo de 0,6 se asocian a una buena capacidad discriminante del modelo.

En la aplicación de la  $FD_1$ , un compuesto es clasificado como antiinflamatorio si adopta valores entre 0 y 8, ya que dentro de este rango presenta la mayor probabilidad de actividad antiinflamatoria. Por el contrario, un compuesto será clasificado como no antiinflamatorio si presenta valores de la  $FD_1$  comprendidos entre 0 y -10 (Figura 15). Cabe resaltar que a la hora de aplicar el modelo al rastreo de nuevos compuestos con actividad antiinflamatoria, el intervalo entre 0 - 0,5 podría dar lugar a falsos positivos ya que en este rango se posicionan tanto compuestos activos como inactivos.



**Figura 15.** Diagrama de distribución farmacológica para la actividad antiinflamatoria representando la expectativa de actividad o inactividad (E) frente al valor de la  $FD_1$  para las distintas moléculas que conforman el *training set* (el color azul representa los antiinflamatorios y el rojo los compuestos sin actividad antiinflamatoria).

La matriz de clasificación para la  $FD_1$  obtuvo un porcentaje de acierto medio del 97%, clasificando correctamente al 100% de los compuestos activos mientras un 93% de los compuestos inactivos fueron correctamente clasificados, como se puede ver en la Tabla 4. En la Tabla 1 del **Artículo 1** se encuentra información adicional sobre el modelo (valores de la  $FD_1$  y de los descriptores para todos los compuestos del *training set*). Con objeto de testar la estabilidad del modelo, se procedió a realizar un test de

validación interna de tipo “leave-one-out” o validación cruzada (cv), donde aleatoriamente se deja fuera a uno de los compuestos del *training set* y se le pide al modelo (construido partiendo de los restantes compuestos) que lo clasifique. Cuando se llevó a cabo este tipo de validación, el porcentaje de correcta clasificación fue del 95% y 89% para el grupo de activos e inactivos, respectivamente (Tabla 4).

**Tabla 4.** Matriz de clasificación obtenida a partir de la  $FD_1$  para el *training set*.

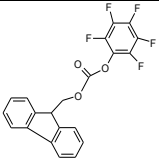
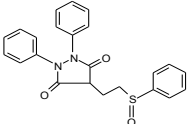
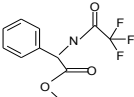

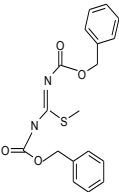

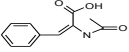
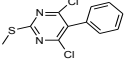
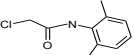
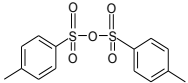
	PORCENTAJE DE CORRECTA CLASIFICACIÓN (%)	CLASIFICACIÓN COMPUESTOS	
		Activos	Inactivos
<b>TRAINING SET</b>			
Grupo activos	100	21	-
Grupo inactivos	93	2	25
<b>CV</b>			
Grupo activos	95	20	1
Grupo inactivos	89	3	24

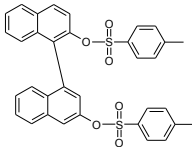
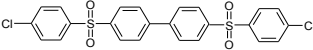
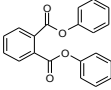
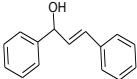
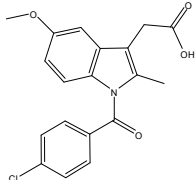
### Cribado virtual

Una vez confirmada la robustez del modelo, se aplicó al rastreo de la base de datos Sigma-Aldrich para buscar potenciales agentes antiinflamatorios. El criterio de selección empleado se basó en el diagrama de distribución farmacológica (DDF), seleccionándose compuestos con un valor de la  $FD_1 > 0$  y descartando los compuestos con valores de  $FD_1$  entre 0 y 0,5, donde se produce solapamiento entre compuestos activos e inactivos, así como a los compuestos con valores de la  $FD_1 > 8,5$ , al encontrarse fuera del campo de aplicabilidad del modelo.

Los compuestos finalmente seleccionados aparecen en la Tabla 5, donde se detalla tanto el n° CAS (*Chemical Abstracts Service*) como la estructura química de todos ellos. Cabe resaltar que de los 14 compuestos seleccionados varios fueron descartados por baja solubilidad (compuestos 6, 10 y 12) o por estar ya descritos en la literatura como AINE (compuesto 4) (Benanti, 1989). El resto de compuestos (**1-3, 5, 7-9, 11, 13 y 14**) pasaron a ser ensayados *in vitro* e *in vivo* según se describe en el apartado correspondiente de Material y Métodos del **Artículo 1**.

**Tabla 5.** Nuevos agentes antiinflamatorios seleccionados por Topología molecular. Resultados del valor de la  $FD_1$  para cada uno de ellos y de los ensayos *in vitro* e *in vivo* a los que fueron sometidos.

N° Compuesto /CAS	Estructura química	<i>In silico</i> Valor $FD_1$	<i>In vitro</i> (inhibición)		<i>In vivo</i> (inhibición)	
			$NO_2^-$ <sup>a</sup>	$PGE_2$ <sup>b</sup>	Carragenina <sup>c</sup>	TPA <sup>d</sup>
1 / 88744-04-1		2.29	- <sup>e</sup>	-	-	-
2 / 57-96-5		0.67	+ <sup>f</sup>	+	-	-
3 / 145513-97-9		10.53	+	+	-	-
4 / 3469-00-9		0.23	N.T. <sup>g</sup>	N.T.	N.T.	N.T.
5 / 25508-20-7		1.88	+	+	+	+
6 / 3516-95-8		1.61	N.T.	N.T.	N.T.	N.T.
7 / 5469-45-4		3.19	+	+	-	-
8 / 65515-11-8		4.02	+	+	+	+
9 / 1131-01-7		1.18	+	+	+	+
10 / 4124-41-8		0.05	N.T.	N.T.	N.T.	N.T.

11 / 128544-06-9		6.56	-	-	-	-
12 / 22287-56-5		10.68	N.T.	N.T.	N.T.	N.T.
13 / 84-62-8		5.54	N.T.	N.T.	-	-
14 / 62668-02-4		1.16	N.T.	N.T.	+	+
Indometacina (fármaco de referencia)		8.3	+	+	+	+

<sup>a</sup> % de inhibición de la producción de NO<sub>2</sub><sup>-</sup> en macrófagos de la línea celular RAW 264.7 a la concentración de 50 μM, tras estimulación con LPS.

<sup>b</sup> % de inhibición de la producción de PGE<sub>2</sub> en macrófagos de la línea celular RAW 264.7 a la concentración de 50 μM, tras estimulación con LPS.

<sup>c</sup> % del volume de inhibición del edema a las 3 horas tras administrar una dosis de producto de 10 mg/kg, en un modelo de inducción del edema por carragenina en ratón.

<sup>d</sup> % del volume de inhibición del edema a las 4 horas tras administrar una dosis de producto de 10 mg/kg, en un modelo de inducción del edema por TPA (13-acetato-12-O-tetradecanoilforbol) en ratón.

<sup>e</sup> - Inhibición < al 50% de lo inhibido por indometacina.

<sup>f</sup> + Inhibición = o > al 50% de lo inhibido por indometacina.

<sup>g</sup> N.T. No testado.



## **Resultado de los ensayos *in vivo* e *in vitro***

A continuación, se comentarán los resultados obtenidos en el modelo agudo del edema plantar inducido por carragenina. La inflamación inducida al inyectar  $\lambda$ -carragenina en la aponeurosis plantar del roedor es de tipo no inmune, es decir, humoral. El desarrollo del edema tras la inyección es un proceso bifásico: en una primera fase (0-1 h), se aprecia enrojecimiento justificado por la liberación de aminas y péptidos vasodilatadores (histamina, serotonina y bradicinina). En la segunda fase (1-6 h) el edema que aparece se relaciona con la producción de PG, alcanzando su máximo a las 3 horas. Este punto hace que el método sea ideal para evaluar fármacos tipo AINE, los cuales reducen el edema inhibiendo la síntesis de productos derivados del ácido araquidónico, fundamentalmente las prostaglandinas procedentes de la vía de la COX. Se sabe que la mayor parte de PGE<sub>2</sub> producida tras la inyección de carragenina procede de la isoforma COX-1, sin embargo, los inhibidores selectivos de COX-2 o los anticuerpos monoclonales de PGE<sub>2</sub> bloquean el edema tanto como los AINE convencionales (Zhang y cols., 1997; King y cols., 2009).

Los resultados obtenidos en el modelo agudo del edema plantar inducido por carragenina reflejados en la Tabla 6, muestran como todos los compuestos ensayados (**5**, **8**, **9** y **14**) demostraron actividad antiinflamatoria, en diferentes momentos del estudio (1, 3 y 5 horas). Destaca la actividad del compuesto **8** que redujo significativamente el edema en un 47% a las 3 h, de modo similar al fármaco de referencia indometacina (58%). Por otro lado, los compuestos **5**, **9** y **14** mostraron una actividad antiedematosa inferior al 40% (Tabla 6).

**Tabla 6.** Efecto antiinflamatorio de indometacina y compuestos (5, 8, 9 y 14) en un modelo agudo de edema plantar inducido por carragenina en ratón. Tanto el fármaco de referencia (indometacina) como los compuestos ensayados, fueron administrados oralmente (10 mg/kg), 1 h antes de la inyección de carragenina. El tamaño del edema se determina midiendo la diferencia entre el volumen de líquido desplazado por la pata del animal mediante un pletismómetro antes y tras la inyección de carragenina ( $\Delta V$ ).

Grupo	1h		3h		5h	
	$\Delta V \pm EEM (\mu L)^a$	% I	$\Delta V \pm EEM (\mu L)$	% I	$\Delta V \pm EEM (\mu L)$	% I
Control	72 $\pm$ 5	-	77 $\pm$ 7	-	80 $\pm$ 8	-
Indometacina	34 $\pm$ 5 ** <sup>b</sup>	52	33 $\pm$ 4 **	58	42 $\pm$ 5 **	48
Compuesto 5	43 $\pm$ 6 **	40	53 $\pm$ 7 *	31	52 $\pm$ 3 **	35
Compuesto 8	48 $\pm$ 4 *	33	41 $\pm$ 5 **	47	36 $\pm$ 4 **	55
Compuesto 9	26 $\pm$ 2 **	64	49 $\pm$ 5 **	36	52 $\pm$ 8 **	35
Compuesto 14	45 $\pm$ 7 **	37	49 $\pm$ 6 **	36	36 $\pm$ 2 **	55

<sup>a</sup> Los valores representan la media  $\pm$  error estándar de la media (EEM) (n=6 animales por grupo).

<sup>b</sup> \*\*p<0.01, \*p<0.05 con respecto al grupo control (test t de Dunnett).

Sin embargo, y a diferencia del AINE de referencia, los compuestos **5** y **9** adoptan su perfil antiinflamatorio máximo en la primera hora tras la administración de la carragenina (40% y 64% de inhibición, respectivamente). El compuesto **14** alcanza su máxima actividad antiinflamatoria a las 5 horas de la inyección de carragenina (55% de inhibición del edema). De nuevo, cabe resaltar el resultado obtenido con el compuesto **8**, el cual a las 5 h sigue siendo activo (55% de inhibición). Este hecho podría atribuirse a los diferentes procesos farmacocinéticos de los compuestos estudiados.

A pesar de su falta de especificidad, un buen modelo en las etapas iniciales del estudio de la actividad antiinflamatoria de un compuesto es el edema auricular inducido por 13-acetato-12-O-tetradecanoilforbol (TPA). La aplicación tópica de una dosis única de TPA induce una reacción inflamatoria en la oreja de ratón caracterizada por la aparición de eritema, edema e infiltración de leucocitos polimorfonucleares que alcanza su máximo a las 4 horas. El edema que se produce está mediado por la activación de la proteincinasa C (PKC) por parte del TPA. La PKC tiene un papel regulador del proceso de la inflamación y en el sistema inmunológico. La mayoría de los efectos del TPA están relacionados con el metabolismo del ácido araquidónico y se ha visto que los

inhibidores de PLA<sub>2</sub>, de COX y de LOX, son capaces de reducir el edema inducido por TPA, así como otros mecanismos de acción como antihistamínicos y antiserotoninérgicos. Los compuesto en estudio (**5**, **8**, **9** y **14**), se administraron por vía tópica (0,5 mg/oreja) disueltos en acetona, y se pudo comprobar como todos los compuestos seleccionados reducían el edema de modo significativo (Tabla 7), destacando el compuesto **14** que lo hizo en un 61%. Sin embargo, ninguno llegó a los valores del fármaco de referencia indometacina, que a la misma dosis inhibió el edema en un 92%.

**Tabla 7.** Efecto antiinflamatorio de indometacina y los compuestos (5, 8, 9 y 14) en el modelo inflamatorio agudo del edema auricular inducido por TPA en ratón. Tanto la indometacina como los compuestos fueron tópicamente administrados, a la dosis de 0,5 mg/oreja en 20 µL de acetona. El edema se determina como la diferencia de peso ( $\Delta P$ ) de la muestra de oreja de ratón obtenida mediante punzón (6 mm de diametro) medida antes y tras 4 horas de la administración del TPA.

Grupo	$\Delta P \pm EEM$ (mg) <sup>a</sup>	% I
Control	18 $\pm$ 1	-
Indometacina	1 $\pm$ 0 <b>**<sup>b</sup></b>	92
Compuesto 5	8 $\pm$ 2 <b>**</b>	52
Compuesto 8	8 $\pm$ 2 <b>**</b>	58
Compuesto 9	7 $\pm$ 2 <b>**</b>	61
Compuesto 14	7 $\pm$ 1 <b>**</b>	59

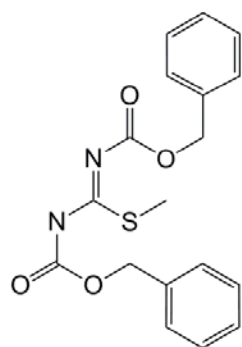
<sup>a</sup> Los valores representan la media  $\pm$  error estandar de la media (E.E.M.). (n=6 animales por grupo).

<sup>b</sup> **\*\***p<0.01 con respect al grupo control (test t de Dunnett).

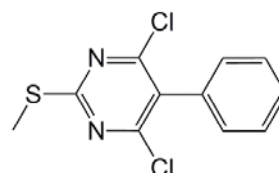
Con el objetivo de elucidar de un modo más específico el mecanismo de acción de los compuestos activos *in vivo*, se realizó el estudio *in vitro*, trabajando con células inflamatorias como son los macrófagos RAW 264.7 estimulados con LPS. A la dosis más alta ensayada (100 µM) ninguno de los compuestos fue tóxico en el test del MTT (viabilidad celular mayor o igual al 90%). La actividad antiinflamatoria detectada *in vivo*, fundamentalmente en el modelo del edema plantar por carragenina, para el compuesto **8**, podría justificarse en base a la inhibición de la producción de prostaglandinas y nitritos de acuerdo a los resultados obtenidos *in vitro* (42% y 51%, respectivamente). Como ocurre muchas veces al trabajar *in vitro*, se ha podido observar

como los factores farmacocinéticos influyen en el efecto final de un fármaco. Así el compuesto **9**, con discreta actividad *in vivo*, resultó ser el más activo a la hora de inhibir de un modo significativo la producción de PGE<sub>2</sub> y la de nitritos (61% y 77%, respectivamente).

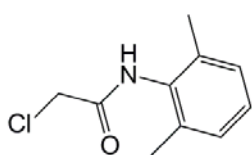
En conclusión, los datos obtenidos en los ensayos farmacológicos corroboran la selección llevada a cabo empleando modelos matemáticos. Tras rastrear la base de datos Sigma-Aldrich, las ecuaciones matemáticas predijeron la actividad antiinflamatoria de catorce compuestos, incluyendo uno descrito previamente como AINE. La predicción es de tipo cualitativo, y los compuestos estudiados (**5**, **8**, **9**, **14**), en mayor o menor medida, han demostrado actividad. Por último, los resultados han permitido sospechar que junto con la inhibición de la vía del ácido araquidónico, puede haber otros mecanismos como la inhibición de la expresión de enzimas proinflamatorias como iNOS.



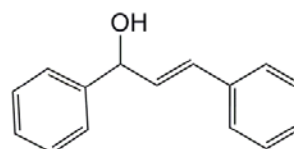
1,3-bis(benciloxycarbonyl)-2-metil-2-tiopseudourea  
Compuesto n° 5



4,6-dicloro-2-metiltio-5-fenilpirimidina  
Compuesto n° 8



2-cloro-2',6'-acetoxilidida  
Compuesto n° 9



trans-1,3-difenil-2-propen-1-ol  
Compuesto n° 14

**Figura 16.** Compuestos ensayados que presentan un mejor perfil antiinflamatorio.

Desde el punto de vista físicoquímico y estructural cabe resaltar la influencia de propiedades electrónicas, del efecto inductivo de los halógenos presentes y de características puramente topológicas, en la actividad antiinflamatoria de estos compuestos. Esto se ve reflejado por los índices empleados en el modelo (FD<sub>1</sub>): nCL, MATS8e y PCR. A continuación, se discuten brevemente cada uno de ellos.

1.- El índice MATS8e.- Este descriptor tiene en cuenta la presencia de heteroátomos electronegativos como O ó S ubicados a distancia topológica = 8 de carbonos aromáticos. Este es el caso de los compuestos **5** y **8** (Figura 16), que contienen S situados a esa distancia de los carbonos bencénicos. Este hecho condicionaría el carácter aromático global de la molécula y la capacidad de movilidad de la carga (polarizabilidad). A este respecto, es significativo que 3 de los 4 compuestos activos (**8**, **9** y **14**) muestren valores positivos o 0 para este índice (Anexo I). Por su parte, compuestos testados con valores más negativos de MATS8e (1,2,3,7,13) resultaron completamente inactivos *in vivo*. Ello es interesante porque, cuanto más negativo es este índice, más alto es el valor de FD<sub>1</sub>. Este hecho ilustra como la FD óptima no es necesariamente la de mayor valor positivo.

2.- La presencia de halógenos.- Por otra parte, los compuestos con F (1 y 3), resultaron inactivos, lo que se relaciona tanto con el índice anteriormente abordado (MATS8e) como con el índice nCl. En efecto, los compuestos que tenían Cl fueron activos (**8**, **9** e indometacina). Uno de los que poseen Cl, el compuesto 12, no fue testado, por lo que podría igualmente ser activo. Esto indica que un elemento demasiado electronegativo como el F podría determinar una excesiva polarización por efecto inductivo que impidiese la movilidad electrónica global en la molécula y determinase la inactividad.

3.- Índice PCR.- Aunque este índice parece jugar un papel menor, dado que toma valores similares entre los compuestos testados (activos e inactivos), sí se aprecia que en general, compuestos que han mostrado actividad antiinflamatoria adoptan valores mayores a 1,4 en este índice (Anexo I). Dado que el PCR es una medida del número de insaturaciones en relación con el total de enlaces, podría interpretarse como una medida de la aromaticidad de la molécula.

Por último, otro factor relevante no tenido en cuenta por la FD<sub>1</sub>, es la presencia de grupos dadores de hidrógeno (fundamentalmente –OH y –NH). Cuatro de los compuestos activos: **5**, **9**, **14** e indometacina, poseen grupos dadores, mientras que gran parte de los inactivos (1, 2, 11 y 13) no poseían estos grupos (Tabla 5). Esto es lógico, dada la gran influencia que la presencia de estos grupos tiene en la mayoría de las interacciones fármaco-receptor, a través de la formación de enlace por puente de hidrógeno (Delgado y cols., 2003).

Este primer artículo permitió adquirir experiencia en la terapia antiinflamatoria, lo que facilitó la consecución del siguiente objetivo planteado en la presente Tesis Doctoral (Capítulo 2): la identificación de compuestos con actividad protectora frente a CU mediante la inhibición de distintos mediadores y vías de señalización implicadas en esta enfermedad (NF-kB, IL-6, iNOS, IL-1 $\beta$  y TNF- $\alpha$ ).

## VI.2. CAPÍTULO 2

Artículo 2: “**Topological virtual screening: a way to find new compounds active in ulcerative colitis by inhibiting NF- $\kappa$ B**” (En: *Molecular Diversity* **2011**, 15:917–926). JCR [2011, CAT: Chemistry Applied; FI: 3,153; Rank: 9 de 71 (**Q1**)].

El Capítulo 2 se centra en la identificación a través de la Topología molecular de nuevos potenciales agentes protectores frente a CU. Este tipo de enfermedad inflamatoria intestinal como se ha comentado en el apartado de Introducción, es una patología que condiciona ampliamente la calidad de vida del paciente, cuya incidencia está creciendo en nuestro país año tras año y cuyo tratamiento actual sigue asociado a numerosos efectos adversos.

En este primer artículo perteneciente al Capítulo 2, se lleva a cabo la búsqueda de nuevos agentes con actividad protectora frente a CU mediante la inhibición del factor de transcripción NF- $\kappa$ B. Este factor de transcripción ha sido ampliamente abordado en el apartado de Introducción de la presente Tesis Doctoral, otorgándosele un papel relevante en el desarrollo de la CU puesto que en biopsias de pacientes con esta enfermedad se encuentra significativamente aumentada su activación (Karban y cols., 2004; Jian y cols., 2005; Ukil y cols., 2006). Además, algunos de los fármacos empleados en la terapia de la CU aguda, como los corticoides, entre otras acciones, inhiben la activación de dicho factor (Venkataranganna y cols., 2007; Piechota-Polanczyk y Fichna, 2014). Este factor de transcripción, también regula la expresión de numerosos mediadores inflamatorios: TNF- $\alpha$ , IL-1 $\beta$ , IL-6 e IL-8 (Lee y cols., 2007). Por último, cabe indicar que este artículo fue el primer trabajo del que tengamos noticia en que se aplicó la Topología molecular a este campo terapéutico.

## OBTENCIÓN Y VALIDACIÓN DE LOS MODELOS TOPOLÓGICOS

### Modelo de predicción de la inhibición de NF-kB

Para la construcción y validación del modelo para la búsqueda de nuevos compuestos con actividad inhibitoria frente a NF-kB se emplean 121 compuestos (67 activos como inhibidores de NF-kB y 54 inactivos). Los compuestos inactivos se extrajeron de la base de datos Merck Index. Finalmente, el modelo matemático se construyó partiendo de un *training set* formado por 95 moléculas estructuralmente heterogeneas y presentó un buen balance entre moléculas activas (51) e inactivas (44). La  $FD_1$  distingue compuestos con potencial actividad inhibitoria frente a NF- $\kappa$ B y está compuesta por cinco variables independientes:

$$FD_1 = (0,633 \times G_1) + (24,29 \times {}^0C) - (3,62 \times {}^2D) + (2,09 \times {}^4D_p) + (0,288 \times V_3) - 30,94 \quad (1)$$

$$N = 95 \quad F = 9,5 \quad \lambda = 0,354 \quad p < 0,00001$$

Los ITs empleados son:

$G_1$  = índice de carga topológica de orden 1.

${}^0C = {}^0\chi / {}^0\chi_v$  = cociente de índices de conectividad de orden 0.

${}^2D = {}^2\chi - {}^2\chi_v$  = diferencia entre índices de conectividad de orden 2.

${}^4D_p = {}^4\chi_p - {}^4\chi_p^v$  = diferencia de índices de conectividad de orden 4 y tipo path.

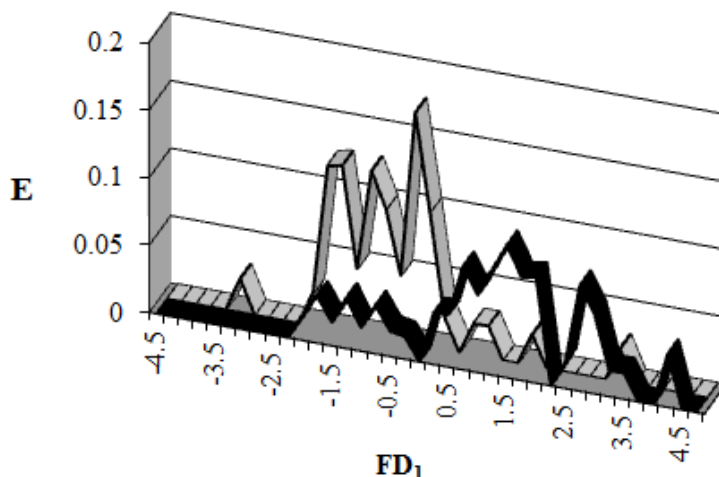
$V_3$  = número de vertices con valencia topológica 3.

En la ecuación anterior, hay descriptores topológicos que evalúan la ramificación molecular ( $V_3$ ), la distribución de la carga intramolecular ( $G_1$ ), el volumen molecular ( ${}^0C$ ) y características topológicas de cada compuesto ( ${}^2D$  y  ${}^4D_p$ ).

De acuerdo con la Ecuación 1, un compuesto debe ser clasificado como activo si la  $FD_1 > 0$ . Mediante la aplicación de este criterio para el *training set* (ver Tabla 1 **Artículo 2**), 40 de los 51 compuestos experimentalmente activos fueron correctamente clasificados como tales (sensibilidad del 78%), mientras que 40 de los 44 compuestos inactivos fueron correctamente clasificados (91% de especificidad). El promedio total de correcta clasificación para el *training set* (activos más inactivos) fue del 84%.



Para establecer el rango adecuado de la actividad inhibitoria respecto al NF- $\kappa$ B, se analizó el DDF obtenido con la  $FD_1$ . En la Figura 17, se puede apreciar que todos los compuestos estudiados muestran valores de la  $FD_1$  en el rango entre 4,5 y -3,8. Además, podemos darnos cuenta de que hay muy pocas zonas de solapamiento entre compuestos con probabilidad de actividad por encima del 40% e inferior al 60%. Estos porcentajes coinciden con los valores de la  $FD_1$  entre 0,4 y -0,4. Por lo tanto, un compuesto se selecciona como inhibidor de NF- $\kappa$ B si adopta valores de la función discriminante entre 4.5 y 0,4 mientras que se considera inactivo si la  $FD_1$  adopta valores entre -0,4 y -3.8. Fuera de estos rangos la clasificación es incierta por lo que cualquier compuesto en esa zona es etiquetado como "no clasificado", N.C.



**Figura 17.** Diagrama de distribución farmacológica de inhibidores de NF- $\kappa$ B obtenida utilizando la  $FD_1$  (el color negro representa los compuestos con actividad inhibitoria y el color gris, los compuestos sin ella).

Una forma sencilla de evaluar la calidad de la  $FD_1$  es aplicarla a un grupo externo (*test set*). En este caso, el grupo externo estaba formado por 26 compuestos (16 activos y 10 inactivos) que no fueron incluidos en el *training set*, lo que representa aproximadamente el 20% de la data total de partida. Los resultados de la clasificación obtenida para cada compuesto del *test set* se muestran en la Tabla 3 del **Artículo 2**. Destaca el hecho de que en el grupo inactivo sólo hay un compuesto clasificado erróneamente, lo que indica que la  $FD_1$  es muy eficiente en la detección de compuestos inactivos (alta especificidad). Por lo tanto, podemos asegurar que el número de falsos activos va a ser mínimo.

## Modelo de predicción de la actividad anti CU

El segundo modelo desarrollado,  $FD_2$ , fue empleado para identificar compuestos activos frente a CU. Cuatro variables independientes conforman la  $FD_2$ :

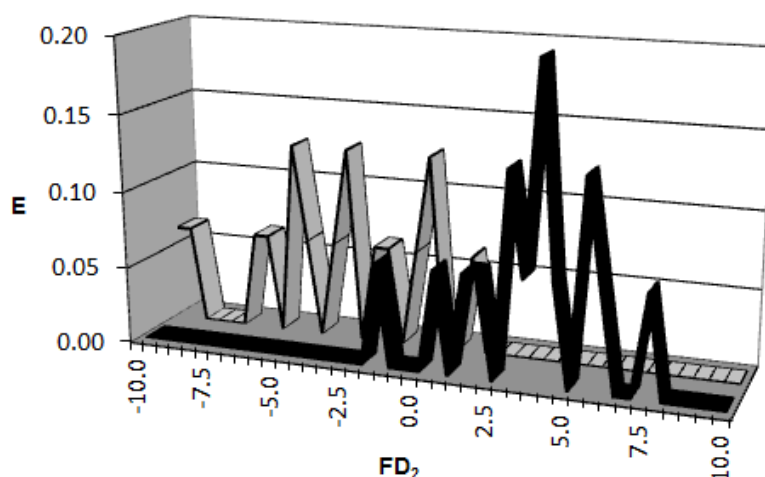
$$FD_2 = (3,15 \times SEige) - (10,28 \times GATS6p) - (261,66 \times X3A) - (0,06 \times D/Dr05) + 52,73 \quad (2)$$

$$N = 31 \quad F = 13,3 \quad \lambda = 0,328 \quad p < 0,00001$$

Los ITs empleados en la Ecuación 2 son:  $SEige$  = suma de valores propios de la matriz de distancia ponderada por electronegatividad que representa una medida de la energía total de enlace en la molécula;  $GATS6p$  = índice de autocorrelación Geary a distancia 6 ponderado por la polarizabilidad atómica. El índice de Geary mide la autocorrelación espacial, por ejemplo, determina como de similares son las moléculas entre sí basándose en la polarizabilidad de sus átomos;  $X3A$  = promedio del índice de conectividad chi-3. Es un índice puramente estructural que evalúa las contribuciones de los átomos a la distancia topológica 3, independientemente de su entorno;  $D/Dr05$  = índice que representa el cociente de la distancia/detour (o distancia por el camino más largo) para anillos pentagonales.

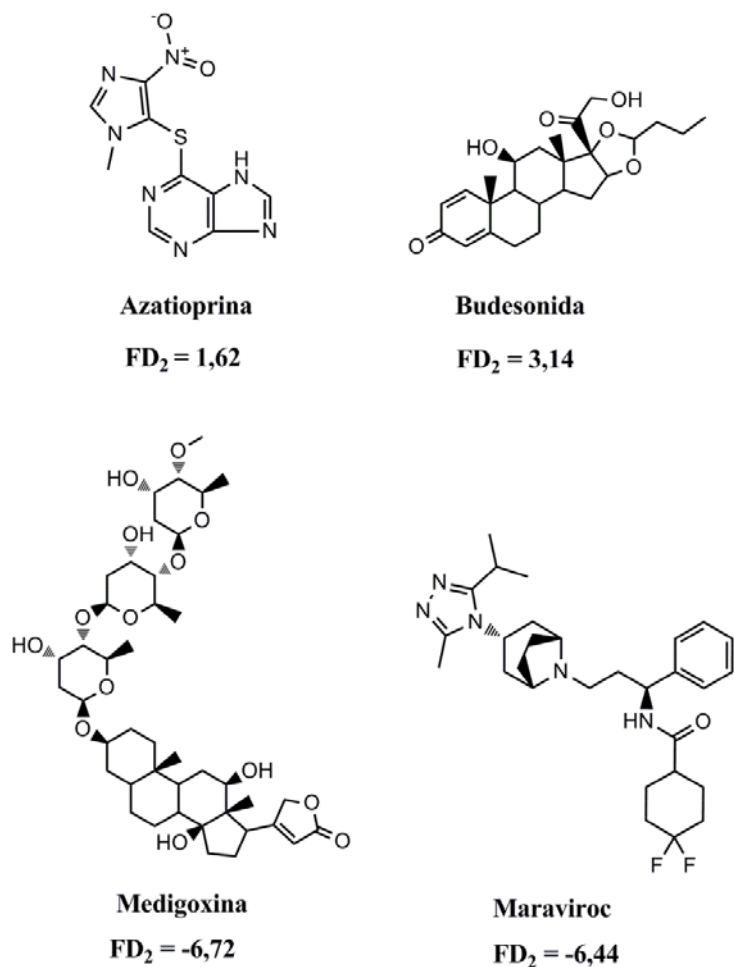
Mediante el empleo de esta función, un compuesto se considera activo si la  $FD_2$  es  $> 0$ , mientras que se marca como inactivo si el valor de la  $FD_2$  es  $< 0$ . El modelo fue capaz de clasificar correctamente a 14 de los 15 compuestos activos y 15 de los 16 compuestos inactivos, lo que implica un porcentaje de acierto del 93% y el 94%, respectivamente. Los valores de la  $FD_2$  para todos los compuestos del *training set* se muestran en la Tabla 2 del **Artículo 2**.

El correspondiente DDF para esta ecuación (Figura 18) muestra que todos los compuestos estudiados presentan valores de la  $FD_2$  en el rango entre 8 y -10. Hay un ligero solapamiento entre los compuestos con probabilidad de actividad por encima de 40% y por debajo de 60%, que coincide con los valores de la función dentro del rango entre 0,89 y -0,49. Por lo tanto, un compuesto se selecciona como activo si está dentro del intervalo: entre 8 y 0,89 y como inactivo si el rango es entre -0,49 y -10. Fuera de estos intervalos, el compuesto analizado será considerado como N.C.

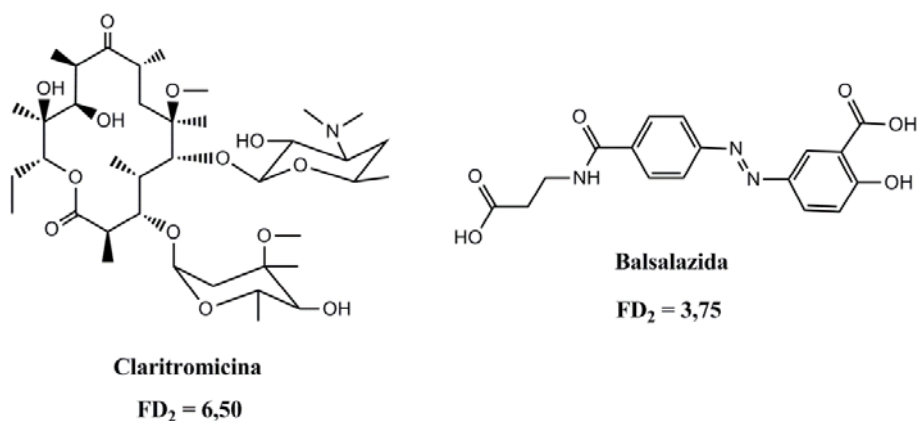


**Figura 18.** Diagrama de distribución farmacológica de compuestos activos frente a CU obtenida utilizando la  $FD_2$  (el color negro representa los compuestos con actividad anti CU y el color gris, los compuestos inactivos).

Aunque no es fácil la identificación de las características estructurales que explican la actividad frente a CU, cabe destacar la importancia del índice más relevante,  $D/Dr05$ . Este índice está relacionado con las trayectorias más largas que unen los átomos ubicados a distancia 5. Codifica información sobre los anillos condensados y se relaciona con la ciclicidad de la molécula (Jain y Agrawal, 2007). En la Ecuación 2, este índice participa con signo negativo, por lo que a mayor valor más baja actividad frente a CU. Por lo tanto, los compuestos aromáticos que muestran anillos pentagonales con heteroátomos (especialmente oxígeno) deben ser los menos activos, por ejemplo, azatioprina o budesonida entre los activos o maraviroc y medigoxina entre los inactivos (Figura 19). Por el contrario, los compuestos que no presentan anillos de cinco miembros ( $D/Dr05 = 0$ ) se espera que sean más potentes, como por ejemplo, balsalazida o claritromicina (Figura 20).



**Figura 19.** Ejemplo de compuestos del *training set* con anillos de cinco miembros con heteroátomos.



**Figura 20.** Ejemplo de compuestos del *training set* (grupo activo) sin anillos de cinco miembros con heteroátomos.

Cuando el modelo se aplica al *test set* (22 compuestos) en la validación externa, 9 de cada 11 compuestos activos fueron clasificados correctamente (82% de precisión) y 7 de 8 compuestos inactivos también fueron adecuadamente clasificados (88% de precisión), mientras que hay tres *outliers* o compuestos no clasificables. Es interesante darse cuenta de que la  $FD_2$  sólo clasifica erróneamente un compuesto perteneciente al grupo de inactivos, lo que indica que la función es altamente específica. Información adicional sobre la validación externa del modelo (valores de la  $FD_2$  para todos los compuestos del *test set*) se encuentra en la Tabla 4 del **Artículo 2**.

Los compuestos pertenecientes al grupo test fueron escogidos mediante la aplicación de un criterio de similitud estructural. Por ejemplo, si el grupo de entrenamiento (*training set*) contenía 7 moléculas con una estructura similar (glucocorticoides), 3 de ellos se asignaban a la validación externa y 4 para el *training set*, con esto se pretendía que tanto el *training* como el *test set* fueran comparables estructuralmente.

Hay que subrayar el importante papel desempeñado por los índices de conectividad y de carga topológica en las diferentes funciones discriminantes obtenidas. Tan solo mirando las funciones discriminantes ( $FD_1$  y  $FD_2$ ), se puede notar como propiedades estructurales y electrónicas de los compuestos determinan su actividad inhibitoria frente a NF- $\kappa$ B y su actividad protectora frente a CU. Dada la alta heterogeneidad estructural de las moléculas utilizadas para construir los modelos, estos pueden ser aplicados a una amplia variedad de compuestos de diferentes bases de datos moleculares.

### **Cribado virtual**

Finalmente, se llevó a cabo un cribado virtual en las bases de datos: Microsource Pure Collection Natural Product (800 productos naturales) y Merck index (aproximadamente 12.000 compuestos).

Aplicando los dos modelos desarrollados se seleccionaron un conjunto de 28 compuestos con potencial actividad frente a CU mediante la inhibición de NF- $\kappa$ B. Casi todos ellos se encontraron comercialmente disponibles y por tanto pueden ser tenidos en cuenta para futuros ensayos *in vitro* e *in vivo*, con el fin de evaluar la capacidad predictiva de los modelos.

La Tabla 8 ilustra los valores de  $FD_1$  y  $FD_2$ , así como la clasificación de cada compuesto. Las moléculas seleccionadas pasan al menos uno de los filtros, ya sea la inhibición de NF- $\kappa$ B o el perfil como agente protector frente a CU, porque ambos representan buenas estrategias para abordar el tratamiento de la CU. Como se muestra en la Tabla 8, la mayoría de los compuestos seleccionados habían sido descritos previamente en la literatura como antiinflamatorios (10-hidroxicamptotecina, purpurina, phycion, orsenillate de metilo, ácido aconítico, genkwanin, uvaol, cromolín sódico, hesperidina, y ácido rosmarínico) y uno de ellos, el ácido ursólico, se describe como antiulceroso, lo que ya supone una pre-validación de nuestros modelos.

**Tabla 8.** Valores de  $FD_1$ - $FD_2$ , probabilidad de actividad, clasificación por parte de los modelos y actividad o categoría terapéutica encontrada en literatura para los compuestos seleccionados como potencialmente activos frente a CU.

Compuesto	$FD_1$	Prob (activ.) %	Clas.	$FD_2$	Prob (activ.) %	Clas.	Actividad/categoría terapéutica
Ácido aconítico	2.90	95	A	2.97	95	A	Antiinflamatorio (Pelzer y cols., 1998)
Ajmalina diacetato	1.16	77	A	4.86	99	A	
Ácido Alizarin-3-metilninodiacético	1.86	87	A	4.33	99	A	Colorante rojo (McLeod,1980)
Apigenina	1.64	84	A	2.6	94	A	Antiinflamatorio (El-Alfy y cols., 2010)
Calceina	2.21	90	A	13.9	100	N.C.	Tinte fluorescente (Parish, 1999)
Carapin	2.51	93	A	3.04	96	A	
Cromolín sódico	0.33	58	N.C.	1.19	78	A	Antiinflamatorio (Zhang y cols., 2001)
(+)-Dibenzil-L-tartrato	0.62	65	A	-12.7	0	N.C.	
Ácido emódico	2.69	94	A	4.44	99	A	
Ácido evérnico	1.42	81	A	7.09	99	A	
Fissinolide	2.38	92	A	4.39	99	A	
Ácido fólico	1.65	84	A	3.9	98	A	
Genkwanina	0.57	64	A	2.59	93	A	Antiinflamatorio (Pelzer y cols., 1998)
Ácido haematómico	2.22	90	A	5.62	100	A	Antioxidante (Marante y cols., 2003)
Hesperidina	2.20	90	A	6.08	100	A	Antiinflamatorio (Benavente-Garcia y Castillo, 2008)
10-Hidroxicamptotecina	0.37	59	N.C.	2.57	93	A	Antiinflamatorio (Cashman y cols., 2004)

Ácido lonchocarpico	1.06	75	A	2.49	93	A	Antimicrobiano (Assreuy y cols., 2005)
Ácido 3-metilorsellínico	0.99	73	A	6.65	100	A	
Metil orsellinato	0.14	54	NC	6.38	100	A	Antiinflamatorio (Marante y cols., 2003)
2-Metoxiresorcinol	0.31	58	N.C.	1.94	88	A	
Parietina	1.17	76	A	2.91	95	A	
Purpurina	1.79	86	A	1.59	84	A	Antiinflamatorio (Mishchenko y cols., 2007)
Ácido pirocatecuico	2.21	90	A	3.24	97	A	
Ro 41-0960	1.75	85	A	-0.46	40	I	
Ácido rosmarínico	1.99	88	A	-8.03	0	I	Antiinflamatorio (Ping y cols., 2008)
Rotenona	0.26	57	NC	1.53	83	A	
Ácido ursólico	2.66	94	A	3.33	97	A	Antiulceroso (Li y cols., 2007)
Uvaol	2.98	95	A	3	96	A	Antiinflamatorio (Marquez-Martin y cols., 2006)

Además de los modelos publicados en el presente trabajo, se emplearon otros modelos enfocados a la predicción de diversos parámetros fisicoquímicos (ADME/Tox, solubilidad en agua, etc.) que influenciaron la selección final de los compuestos presentados en la Tabla 8.

## Capítulo II

Artículo 3: “**Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW264.7 cell lines**” (En: *Molecular Diversity* **2013**, 17(3):573–593). JCR [2013, CAT: Chemistry Applied; FI: 2,544; Rank: 15 de 71 (Q1)].

En este artículo se presenta y valida un modelo para la predicción de la inhibición de IL-6. Este modelo se aplica al rastreo de la actividad inhibitoria frente a IL-6 en compuestos con potencial actividad protectora frente a CU mediante la inhibición NF-κB previamente seleccionados por Topología molecular.

### OBTENCIÓN Y VALIDACIÓN DEL MODELO TOPOLÓGICO

#### Modelo de predicción de la inhibición de IL-6

La función discriminante obtenida, está constituida por tres variables:

$$FD_3 = (3,34 \times ESpm15r) - (8,42 \times SEigv) - (31,34 \times SEige) - 72,25 \quad (3)$$

$$N = 53 \quad F = 54,04 \quad \lambda = 0,232 \quad p < 0,00001$$

De acuerdo con la Ecuación 3, un compuesto se clasifica como activo si  $FD_3 > 0$ , de lo contrario se cataloga como inactivo. Mediante la aplicación de este criterio para el *training set* (53 compuestos), 23 de los 25 compuestos activos fueron clasificados correctamente (92% precisión), y todos los 28 compuestos inactivos fueron clasificado correctamanete (100% de precisión) (Tabla 9). El promedio de clasificación correcta para todo el conjunto de compuestos (activos e inactivos) fue del 96%. En la Tabla 2 del **Artículo 3** se encuentra información adicional sobre el modelo (valores de la  $FD_3$  para todos los compuestos del *training set*).



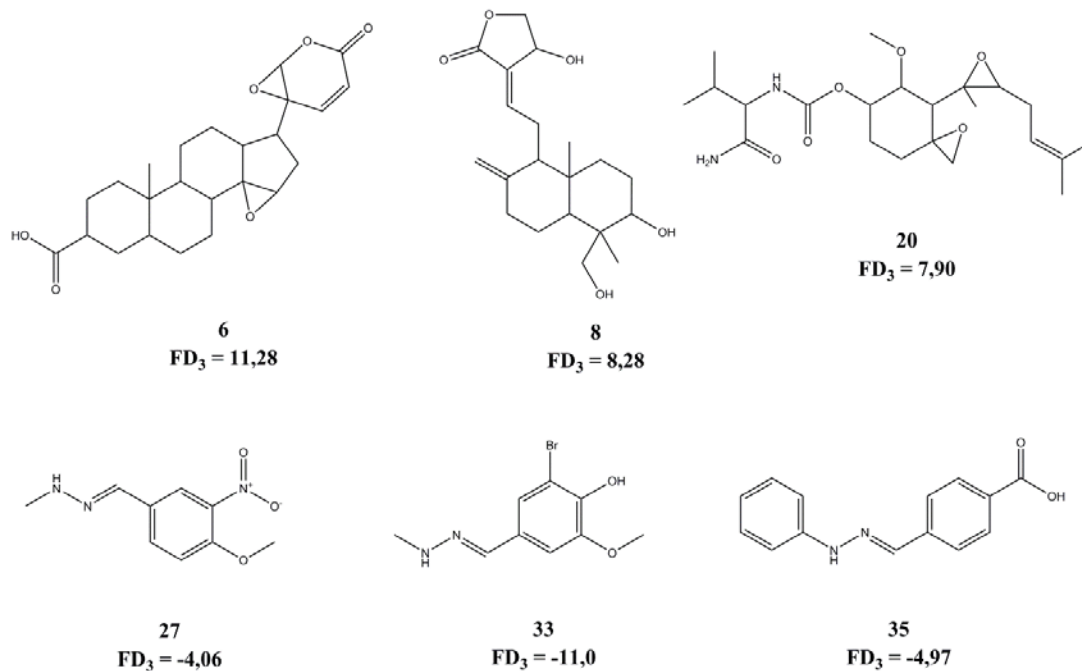
**Tabla 9.** Matriz de clasificación obtenida a través de la FD<sub>3</sub> para el *training* y *test set*.

	Porcentaje de Correcta clas. (%)	A	I	N.C.
<b>Training Set</b>				
Grupo activo	92	23	2	-
Grupo inactivo	100	0	28	-
<b>Test Set</b>				
Grupo activo	75	9	3	-
Grupo inactivo	84	6	36	1

Clas.: clasificación; A: número de compuestos clasificados como activos;  
I: número de compuestos clasificados como inactivos; N.C.: número de  
compuestos no clasificables.

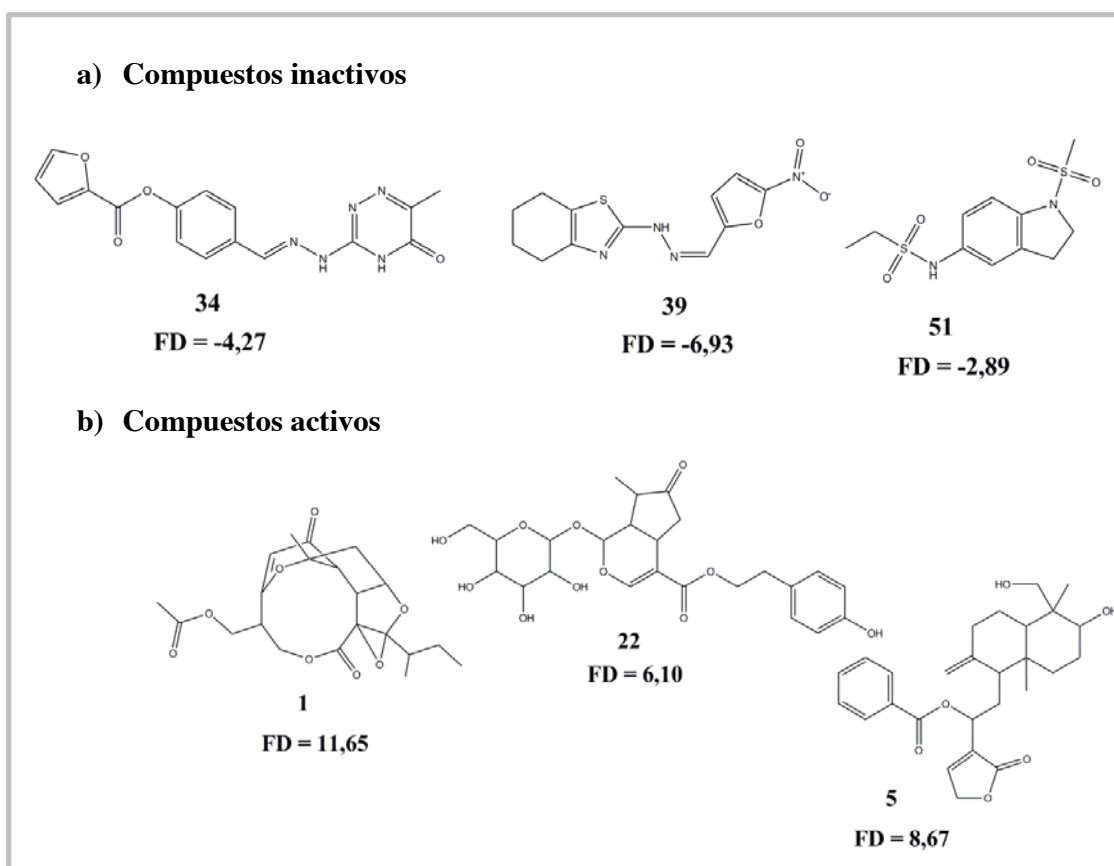
Dada la naturaleza de los índices de la Ecuación 3, no es fácil dar una simple interpretación de la actividad inhibidora de la producción de IL-6 ligada a los descriptores. Eso no es sorprendente, considerando la naturaleza compleja de la propiedad bajo estudio. De hecho, los tres índices del modelo, tienen en cuenta la influencia de las integrales de resonancia, que es una medida de la fuerza de dobles enlaces, relacionada con su grado de conjugación, así como el volumen de Van der Waals y la electronegatividad de los átomos constitutivos de las moléculas. Los índices que aparecen son el ESpm15r, que es el momento espectral 15 de la matriz de adyacencia de aristas ponderada por integrales de resonancia; el SEigv, que corresponde a la suma de autovalores de la matriz de distancia ponderada por volúmenes de van der Waals y el SEige, que es la suma de autovalores de la matriz de distancia ponderada por electronegatividad (Estrada, 1996; Todeschini y Consonni, 2008).

En términos generales, los compuestos activos muestran un bajo grado de conjugación de modo que si contienen anillos aromáticos estos están aislados del resto de la molécula (por ejemplo, compuestos 6, 8, 18, y 20). Por el contrario, los compuestos inactivos muestran generalmente un alto nivel conjugación (véase, por ejemplo, los compuestos 27, 33 y 35) (Figura 21).



**Figura 21.** Relación del grado de conjugación presente en las moléculas con actividad inhibitoria frente a IL-6.

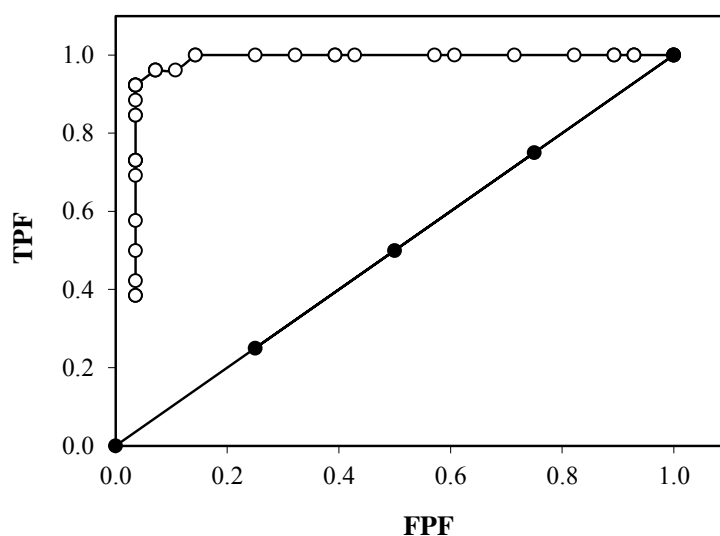
Además, la mayoría de los compuestos que incluyen átomos (como el N) o grupos (como el nitro) fuertemente electronegativos y conjugados con anillos aromáticos son inactivos (compuestos 34, 39 y 51, por ejemplo), mientras que la mayoría de compuestos activos presenta grupos tales como éter, éster o hidroxilo, en compuestos con gran número de anillos alicíclicos (compuestos 1, 5, 8 y 22) (Figura 22).



**Figura 22.** Relación de la presencia de átomos electronegativos y la aromaticidad con la actividad inhibitoria frente a IL-6.

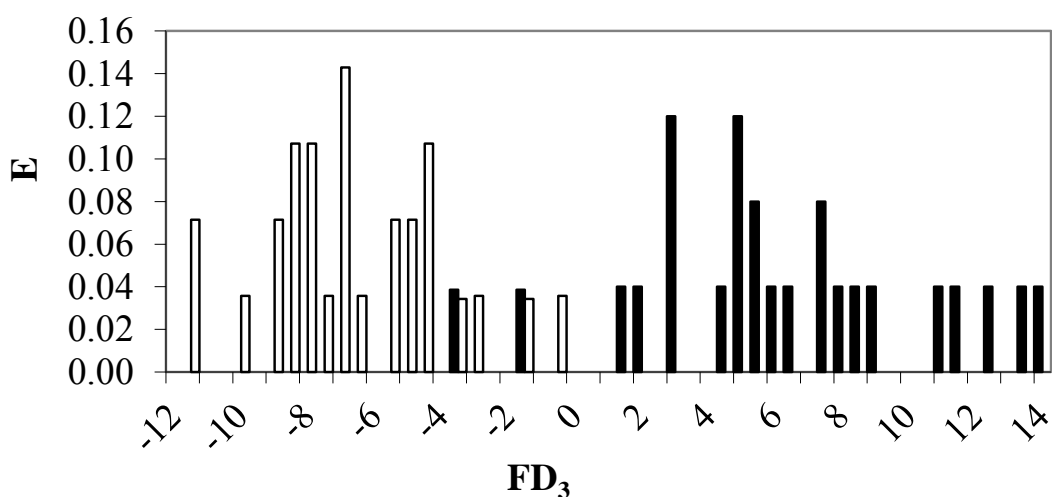
Para evaluar la capacidad predictiva del modelo se usó el coeficiente de correlación de Matthews (MCC), obteniéndose un valor del mismo de 0,93, lo que corrobora su fiabilidad. Además, el valor calculado de MCC se calculó de forma ligeramente distinta, adicionando la unidad al valor de MCC, de modo que éste se pueda expresar como % de precisión. De este modo, 0 significa ninguna correlación en absoluto, 1 representa el 50% y 2 representa la correlación máxima (100%). Al calcular el MCC del modelo, se obtuvo un valor de 1,93 (96,5% de correlación).

Otro método para evaluar la precisión de la  $FD_3$  seleccionada fue el cálculo de la curva ROC, que nos da información sobre la sensibilidad (fracción de verdaderos positivos) y la especificidad (fracción de verdaderos negativos) del modelo, para diferentes umbrales de la FD. La curva ROC para el *training set* se muestra en la Figura 23. El área bajo la curva (AUC) es 0,96 para la Ecuación 3. Este valor elevado de AUC significa una excelente precisión de la función seleccionada.



**Figura 23.** Gráfico de curva ROC para la  $FD_3$ . Cada punto de la curva ROC (puntos blancos) corresponde a un posible punto de corte del modelo, y nos informa de su respectiva sensibilidad (eje y: TPF (Fracción de verdaderos positivos)) y 1-especificidad (eje x: FPF (Fracción de falsos positivos)) para diferentes umbrales de la función de clasificación (entre -10,0 y 10,0). En negro se aprecia la diagonal de referencia o línea de no-discriminación.

Para establecer el rango adecuado de la actividad, se analizó el DDF obtenido con la  $FD_3$ . En cuanto a la Figura 24, podemos apreciar que todos los compuestos estudiados presentan valores entre 14,5 y -12. Fuera de ese rango, cualquier compuesto se clasifica como incierto y está etiquetado como "no-clasificado" (*outlier*), N.C. Además, en la Figura 24 se aprecia cómo no hay solapamiento entre los compuestos activos e inactivos cuando la  $FD_3$  adopta valores  $> 0$ , por lo que al realizar un cribado virtual, se seleccionarán los compuestos como potenciales inhibidores de IL-6 siempre que muestren valores de la  $FD_3$  entre 0 y 15.



**Figura 24.** Diagrama de distribución farmacológica de inhibidores de IL-6 obtenido mediante la  $FD_3$  (el color negro representa inhibidores de IL-6 y el color blanco los compuestos sin actividad inhibitoria frente a IL-6).

Como se ha mencionado anteriormente, el mejor modo de evaluar la calidad de un modelo es validarlo externamente (*test set*). En este caso, el grupo *test set* estaba formado por 55 compuestos (12 activos y 43 inactivos). El *test set* se seleccionó aleatoriamente, representando un porcentaje entre el 30 y 60 de todos los activos e inactivos, respectivamente. En la Tabla 3 del **Artículo 3** se resumen los resultados de la predicción obtenida para cada uno de estos compuestos. Como se aprecia en la Tabla 9, la tasa de éxito para los compuestos activos en el *test set* fue del 75% (9 de 12 compuestos fueron clasificados correctamente), mientras que, en el caso del grupo de inactivos el 84% fue clasificado correctamente (36 de 43). Además, aunque la estricta aplicación de la  $FD_3$  conduce a la pérdida de algunos de los compuestos activos (25%), el punto importante es que de los 43 inactivos del *test set*, el 84% fueron reconocidos con éxito por el modelo, por lo que el riesgo de inclusión de falsos activos se reduce al mínimo. Eso significa que nuestro modelo posee una alta especificidad para predecir la actividad inhibitoria de los compuestos frente a la IL-6.

## Cribado virtual

Empleando el modelo descrito anteriormente (Ecuación 3, Figura 24), se realizó un *screening virtual* de un grupo de 28 moléculas seleccionado previamente con potencial actividad frente a CU (Gálvez-Llompart y cols., 2011). Como se muestra en la Tabla 10, la mayoría de estos 28 compuestos fueron clasificados como potenciales inhibidores de la IL-6.

**Tabla 10.** Valores de la  $FD_3$ , probabilidad de actividad, clasificación por el modelo y actividad biológica descrita en literatura para los compuestos seleccionados como potenciales inhibidores de IL-6 en el proceso de cribado virtual.

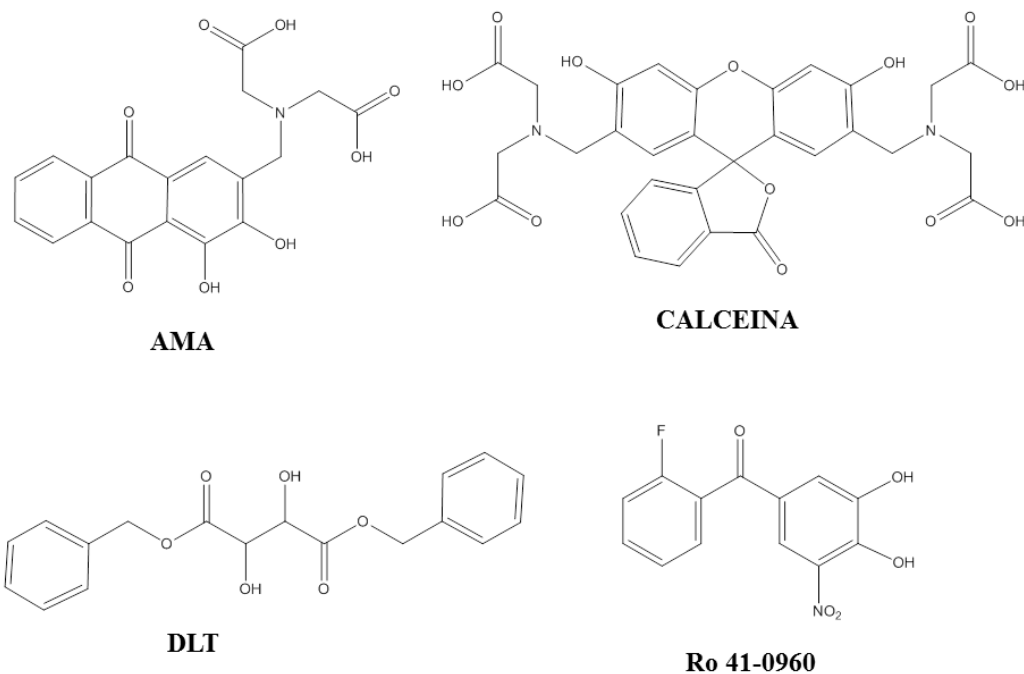
Compuesto	$FD_3$	Prob (activ.)	Clas.	Actividad biológica
Ácido 3-metilorselínico	0,82	0,69	A	
(+)-Dibencil-L-tartrato (DLT)	0,71	0,67	A	
10-Hidroxycampotecina	5,12	0,99	A	Antiinflamatorio (Cashman y cols., 2004)
2-Metoxiresorcinol	-3,23	0,04	I	
Ácido aconítico	-3,45	0,03	I	Antiinflamatorio (Pelzer y cols., 1998)
Ajmalina diacetato	7,80	1,00	A	
Ácido Alizarin-3-metiliminodiacético (AMA)	4,47	0,99	A	Colorante rojo (McLeod, 1980)
Apigenina	1,56	0,83	A	Antiinflamatorio (El-Alfy y cols., 2010)
Calceína	9,28	1,00	A	Tinte fluorescente (Parish, 1999)
Carapina	10,85	1,00	A	
Cromolín sodico	5,17	0,99	A	Antiinflamatorio (Zhang y cols., 2001)
Ácido emódico	4,84	0,99	A	
Ácido evérnico	2,95	0,95	A	
Fissinolide	11,13	1,00	A	
Ácido fólico	-2,75	0,06	I	
Genkwanina	1,61	0,83	A	Antiinflamatorio (Pelzer y cols., 1998)

Ácido haematómico	1,32	0,79	A	Antioxidante (Marante y cols., 2003)
Hesperidina	7,17	1,00	A	Antiinflamatorio (Benavente- Garcia y Castillo, 2008)
Ácido lonchocarpico	5,69	1,00	A	Antimicrobiano (Assrey y cols., 2005)
Metil orsellinato	-0,32	0,42	I	Antiinflamatorio (Marante y cols., 2003)
Parietina	4,07	0,98	A	Antiinflamatorio (Shen y cols., 2011)
Purpurina	4,09	0,98	A	Antiinflamatorio (Mishchenko y cols., 2007)
Ácido pirocatecuico	-1,47	0,19	I	
Ro 41-0960	3,43	0,97	A	Inhibidor de la catecol-O- metiltransferasa (Miller y cols., 1997)
Ácido rosmarínico	0,69	0,67	A	Antiinflamatorio (Ping y cols., 2008)
Rotenona	4,73	0,99	A	Citotóxico (Yoon y cols., 2010)
Ácido ursólico	11,22	1,00	A	Antiulceroso (Li y cols., 2007)
Uvaol	10,77	1,00	A	Antiinflamatorio (Marquez-Martin y cols., 2006)

De los 28 compuestos seleccionados, cuatro se escogieron para las pruebas *in vitro*, de acuerdo con los siguiente criterios:

- Que la molécula sea clasificada como activa por el modelo topológico.
- Que no estuviera descrita en la literatura ninguna actividad antiinflamatoria, antioxidante o antiulcerosa, con el fin de garantizar un enfoque novedoso al tratamiento de la CU.
- Que fuese accesible comercialmente.

Los compuestos seleccionados fueron: ácido Alizarin-3-metiliminodiacético (**AMA**), **Calceína**, (+)-dibencil-L-tartrato (**DLT**) y **Ro 41-0960** (ver Figura 25).



**Figura 25.** Compuestos seleccionados como potenciales inhibidores de IL-6.



## ENSAYOS BIOLÓGICOS

### **Efecto de los compuestos sobre la viabilidad celular**

En primer lugar, se determinó la citotoxicidad de los compuestos seleccionados (**AMA**, **Calceína**, **DLT**, y **Ro 41-0960**) a concentraciones crecientes (25, 50 o 100  $\mu\text{M}$ ) frente a dos tipos de cultivos celulares: células humanas Caco-2 y macrófagos murinos RAW 264.7. El único compuesto que afectó a la viabilidad celular fue el **Ro 41-0960**, por lo que su efecto sobre la producción de IL-6 en cultivos celulares se determinó a una concentración inferior: 50  $\mu\text{M}$  frente a Caco-2 y 25  $\mu\text{M}$  en macrófagos RAW 264.7. El resto de compuestos se ensayaron a una concentración máxima de 100  $\mu\text{M}$  (Tabla 5, **Artículo 3**).

### **Efecto de los compuestos sobre la producción de IL-6**

Las líneas celulares RAW 264.7 y Caco-2 fueron expuestas a los cuatro compuestos (**AMA**, **Calceína**, **DLT** y **Ro 41-0960**) para investigar el efecto inhibitor de estos frente a la producción de IL-6. Los niveles de dicha citocina se determinaron mediante ELISA. Todos los compuestos, excepto **AMA**, inhibieron la producción de IL-6 en las Caco-2, mientras que en los macrófagos, solo **DLT** y **Ro 41-0960** (Tabla 11). La inhibición fue concentración dependiente, destacando los valores  $\text{CI}_{50}$  obtenidos por el compuesto **Ro 41-0960** ( $\text{CI}_{50} = 11,6 \mu\text{M}$  frente a las Caco-2 y 0,4  $\mu\text{M}$  en células RAW 264.7) (Tabla 7, **Artículo 3**).

**Tabla 11.** Efecto de los compuestos sobre la producción de IL-6 en las líneas celulares Caco-2 y Raw 264.7 determinado por ELISA.

Grupo	C ( $\mu$ M)	Caco-2 <sup>a</sup>		RAW 264.7 <sup>b</sup>	
		IL-6 (pg/mL) $\pm$ EEM	% I	IL-6 (pg/mL) $\pm$ EEM	% I
<b>Control</b>		20,56 $\pm$ 0,98		360,0 $\pm$ 10,24	
<b>AMA</b>	10	33,06 $\pm$ 2,85*	-	319,0 $\pm$ 22,39	11
	50	22,11 $\pm$ 4,60	-	681,1 $\pm$ 26,03***	-
	100	16,49 $\pm$ 1,12	20	692,8 $\pm$ 11,13***	-
<b>Calceína</b>	10	22,07 $\pm$ 4,19	-	248,2 $\pm$ 26,07*	31
	50	9,82 $\pm$ 1,48*	52	385,0 $\pm$ 15,64	-
	100	3,55 $\pm$ 1,95***	83	564,10 $\pm$ 14,75 ***	-
<b>DLT</b>	10	13,43 $\pm$ 0,58	35	200,6 $\pm$ 49,05	44
	50	6,24 $\pm$ 2,76***	70	145,2 $\pm$ 33,34***	60
	100	0,47 $\pm$ 0,19***	98	126,5 $\pm$ 23,91***	65
<b>Ro 41-0960</b>	10	11,17 $\pm$ 1,29	46	71,31 $\pm$ 17,79***	80
	25	3,55 $\pm$ 0,81***	83	18,39 $\pm$ 7,83***	95
	50	1,72 $\pm$ 0,93***	92		

<sup>a</sup>Las células Caco-2 fueron expuestas a los siguientes estímulos inflamatorios [IL-1 $\beta$  (25 ng/mL), TNF- $\alpha$  (50 ng/mL) y IFN- $\gamma$  (50 ng/mL)], 2 horas después del tratamiento con los productos; <sup>b</sup>Las células RAW 264.7 fueron estimuladas con LPS (1  $\mu$ g/mL), tras 2 horas de administrar los productos.

Los valores se expresan como la media de la producción de IL-6 (pg/mL)  $\pm$  EEM, siendo representativos de al menos tres experiencias independientes por grupo. Las diferencias entre los grupos frente al control fueron determinadas por medio del análisis de varianza (ANOVA) seguido del test t de Dunnett. No diferencias significativas respecto al control ( $p > 0,05$ ); diferencias significativas respecto al control ( $*p < 0,05$ ;  $***p < 0,001$ ).

% I.: porcentaje inhibido.

El segundo compuesto más activo como inhibidor de IL-6 fue **DLT**, que mostró valores de  $CI_{50}$  de 19  $\mu\text{M}$  en ambas líneas celulares. La actividad inhibidora frente IL-6 exhibida por los compuestos seleccionados (Tabla 7, **Artículo 3**) es comparable con compuestos actualmente empleados en el tratamiento de la CU, tales como dipropionato de beclometasona (Papi y cols., 2010) y budesonida (Dignass y Buhr, 2004), que presentan valores de  $CI_{50}$  frente a IL-6  $>100\mu\text{M}$  y 2,2  $\mu\text{M}$ , respectivamente (Kang y cols., 2001).

Mediante la TM se han seleccionado 28 productos como potenciales inhibidores de IL-6, una citocina proinflamatoria implicada en la colitis ulcerosa. Cuatro de estos compuestos han sido estudiados *in vitro* frente a dos tipos celulares diferentes, obteniendo notables resultados, destacando el compuesto **Ro 41-0960** por su potente actividad inhibidora. Estos datos, junto con el artículo anteriormente presentado (**Artículo 2**) centrado en la identificación de compuestos activos frente a CU a través de la inhibición de la activación del NF- $\kappa$ B, permitió explorar el efecto real de esos compuestos en un modelo de colitis ulcerosa aguda en ratón inducida por DSS (ver artículo siguiente).

Artículo 4: “**Molecular topology: A strategy to identify novel compounds against ulcerative colitis.**” (En: *Molecular Diversity* **2016**, XXX). Enviado.

En este artículo se desarrollan varios modelos de predicción de la actividad inhibitoria frente iNOS y TNF- $\alpha$  relacionada con la CU así como la actividad protectora contra un modelo de CU inducido por DSS. Cada modelo se ha empleado para la predicción de tales actividades farmacológicas por parte de cuatro compuestos previamente seleccionados como agentes terapéuticos potenciales para colitis ulcerosa: **AMA**, **Calceína**, **DLT** y **Ro 41-0960**. Posteriormente, se validaron los modelos mediante el ensayo de estos cuatro compuestos *in vitro* e *in vivo*, lo que confirma a **AMA** y **Ro 41-0960** como los mejores candidatos a convertirse en compuestos *lead* potencialmente activos en CU.

## OBTENCIÓN Y VALIDACIÓN DE LOS MODELOS TOPOLÓGICOS

### Modelo predictivo de la inhibición de iNOS

El modelo se construye partiendo de un *training set* formado por 433 compuestos con estructuras moleculares heterogéneas. Los grupos de moléculas activas e inactivas fueron equilibrados de manera que ambos tuvieran el mismo peso en la ecuación, sin tener en cuenta el número de compuestos que albergaba cada uno.

La ecuación correspondiente a  $FD_4$ , discrimina compuestos con predicha actividad inhibitoria respecto a iNOS y se compone de seis variables:

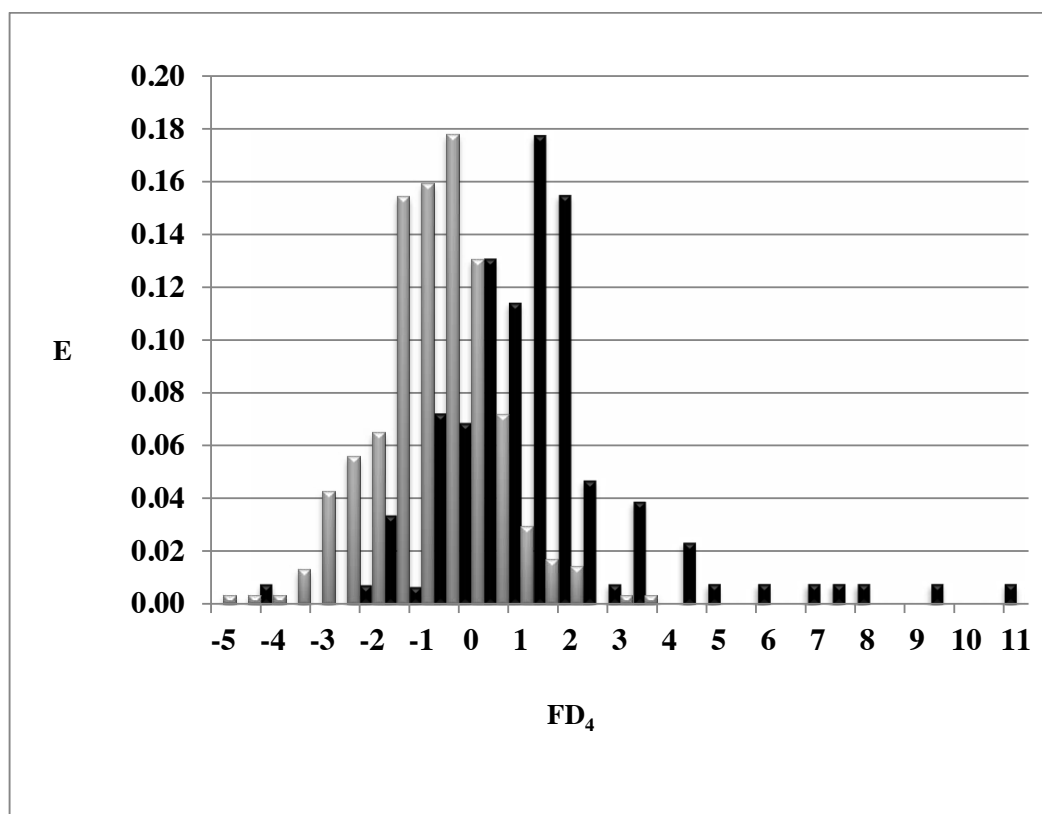
$$FD_4 = (1,40 \times PHI) - (156,05 \times X1A) + (90,70 \times X3A) + (9,63 \times ATS1p) - (15,76 \times ESpm02u) + (1,74 \times ESpm13x) + 51,71 \quad (4)$$

$$N = 433 \quad F = 33,68 \quad \lambda = 0,68 \quad p < 0.00001$$

Los descriptores topológicos empleados son: índice de flexibilidad de Kier (PHI); promedio del índice de conectividad de orden 1 (X1A) y de orden 3 (X3A); índice de autocorrelación de Broto-Moreau a distancia 1 ponderado por polarizabilidad (ATS1p); momento espectral 02 de la matriz de adyacencia (Espm02u) y momento espectral 13 de la matriz de adyacencia ponderado por valencias de los vértices (Espm13x).

Entre los descriptores de la Ecuación 4, algunos son particularmente relevantes en términos de la actividad estudiada, como la flexibilidad molecular (medida por el índice de PHI), importante para poder adoptar una conformación adecuada para facilitar la interacción con la enzima, así como la polarizabilidad (índice ATS1p), lo que pone de relieve la importancia de la movilidad de electrones intramolecular. De particular interés es la influencia de algunos descriptores puramente topológicos, como ESpm02u y ESpm13x, que representan momentos espectrales de matrices de adyacencia. En estudios anteriores se ha puesto de relieve (Lipkus y cols., 2008), que ciertos *scaffolds* moleculares y "frameworks" relacionados con estos descriptores, aparecen con mucha más frecuencia en los compuestos *drug-like*.

Según  $FD_4$ , un compuesto será clasificado como inhibidor de iNOS si  $FD_4$  adopta valores entre 0,5 y 11,5; por otra parte, un compuesto será etiquetado como inactivo si  $FD_4$  toma valores  $< 0,5$  y  $> -5$  (ver Figura 26). Por último, las moléculas con valores de  $FD_4 > 11,5$  o  $FD_4 < -5$  serán considerados como *outliers*.



**Figura 26.** Diagrama de distribución farmacológica para la actividad inhibitoria frente a iNOS empleando el valor de la  $FD_4$  (el color negro representa los inhibidores de la iNOS mientras que el gris, los compuestos sin esta actividad).

Este criterio de clasificación se aplicó al *training set*. De los 433 compuestos, 98 de los 127 inhibidores de iNOS se clasificaron correctamente (sensibilidad 77%), mientras que 257 de los 306 compuestos inactivos (84% especificidad) (ver Tabla 12). El modelo se somete a un grupo test de 108 moléculas (32 activas y 76 inactivas), que no fueron, obviamente, empleados para la construcción del modelo. Estos compuestos se eligieron al azar y representan el 20% total de la data. La FD<sub>4</sub> logró clasificar correctamente al 75% de los activos (24 de los 32 compuestos) y al 83% de inactivos (63 de 76 compuestos) (ver Tabla 12). Por otra parte, aunque la aplicación estricta de FD<sub>4</sub> conlleva una pérdida de algunos potenciales activos (25%), el 83% de los compuestos inactivos fueron clasificados correctamente. De esta manera, el riesgo de incluir falsos activos se reduce significativamente, lo que significa que el modelo posee una alta especificidad. En el Anexo II y III se encuentra información adicional sobre el modelo (valores de la FD<sub>4</sub> para todos los compuestos del *training y test set*).

**Tabla 12.** Matriz de clasificación obtenida a través de la FD<sub>4</sub>, FD<sub>5</sub> y DF<sub>6</sub> para el *training y test set*.

Set	A	I	VP	FP	VN	FN	Sensibilidad	Especificidad
							(%)	(%)
<b>FD<sub>4</sub> Eq. (4)</b>								
<i>Training</i>	127	306	98	49	257	29	77	84
<i>Test</i>	32	76	24	13	63	8	75	83
<b>FD<sub>5</sub> Eq. (6)</b>								
<i>Training</i>	27	162	19	17	145	8	70	90
<i>Test</i>	18	35	12	8	27	6	67	77
<b>FD<sub>6</sub> Eq. (7)</b>								
<i>Training</i>	43	78	39	16	62	4	91	79
<i>Test</i>	18	18	14	3	15	4	78	83

A: n° de compuestos activos, I: n° de compuestos inactivos, VP: verdaderos positivos, FP: falsos positivos, VN: verdaderos negativos y FN: casos de falsos negativos.

El MCC fue utilizado para evaluar la calidad de la FD, adoptando un valor de 0,59 lo cual valida la fiabilidad del modelo. Además, en este caso específico, MCC se calculó de una manera ligeramente diferente (MCC modificado = 1,59), lo que se traduce en un capacidad predictiva del modelo del 80%.

### **Modelo de predicción de la CI<sub>50</sub> de la iNOS**

Con el fin de predecir el valor CI<sub>50</sub> de los nuevos compuestos potenciales anti CU, un modelo de regresión multi-lineal para la predicción del logCI<sub>50</sub> para iNOS se desarrolla. La mejor ecuación lineal, junto con sus parámetros estadísticos fue:

$$\log CI_{50} = (0,49 \times nC) - (2,57 \times nR06) - 0,06 \quad (5)$$

$$N=10 \quad r^2=0,95 \quad F=70,44 \quad EEE=0,51 \quad p<0,00002 \quad Q^2=0,88$$

Donde  $r^2$ , es el coeficiente de determinación;  $F$ , parámetro Fisher–Snedecor;  $EEE$ , error estandar de estimación;  $Q^2$ , coeficiente de predicción.

Dos descriptores constitutivos estan presentes en la ecuación: el número de átomos de carbono ( $nC$ ) y el número de anillos de 6 miembros ( $nR06$ ).

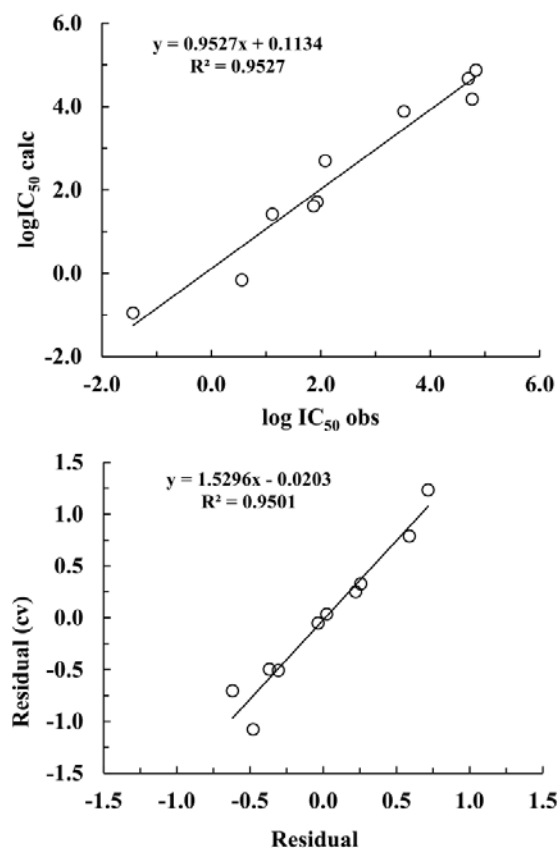
Los valores de  $r^2$  y  $Q^2$  (coeficiente de predicción para el test de validación cruzada L-O-O, leave-one-out) próximos a 1, además de valores bajos de  $EEE$  (menor al 8% de la media de los valores de la propiedad estudiada) avalan la calidad del modelo predictivo.

La Tabla 13 y la Figura 27 (parte inferior) muestran los valores de logCI<sub>50</sub> predichos para cada compuesto analizado. El poder predictivo del modelo como se ha apuntado anteriormente, fue evaluado mediante una técnica de validación cruzada, empleando el test de L-O-O ( $Q^2 = 0.88$  para logCI<sub>50</sub> iNOS). Los resultados del test L-O-O se muestran en la columna 5 de la Tabla 13 y en la Figura 27.

**Tabla 13.** Valor experimental, predicho y test de validación cruzada para el modelo de regresión multineal y valor de los descriptores para los compuestos que conforman el *training set*.

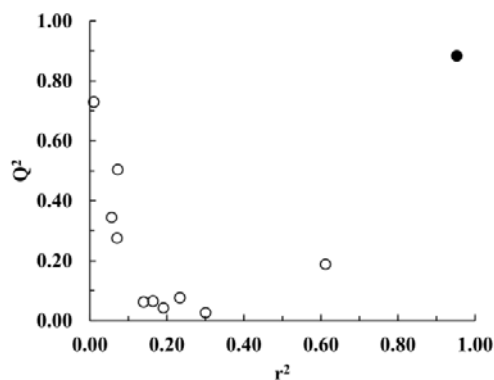
Compuesto	nC	nR06	log CI <sub>50</sub> obs	logCI <sub>50</sub> cal	logCI <sub>50</sub> cal (cv)
(S)-Etilisotiourea	3	0	1,11	1,42	1,62
2-[2-(4-Metoxi-2-piridinil)etil]-1H-imidazo[4,5-b]piridina	14	2	1,93	1,71	1,68
2-Amino-5,6-dihidro-6-metil-4H-1,3-tiazina	5	1	0,56	-0,16	-0,68
5-[(4'-Amino-5'8'-difluorospiro[piperidina-4,2'(1'H)-quinaxolin]-1-yl)carbonil]-2-piridinacarbonitrilo	19	4	-1,43	-0,95	-0,35
BDBM106699	19	2	4,77	4,18	3,98
BDBM111495	20	2	4,70	4,67	4,66
BDBM152702	10	0	4,84	4,87	4,89
BDBM36399	16	2	2,08	2,70	2,78
BDBM50148167	19	3	1,87	1,61	1,54
N6-(1-Iminoetil)-L-lisina	8	0	3,52	3,89	4,02





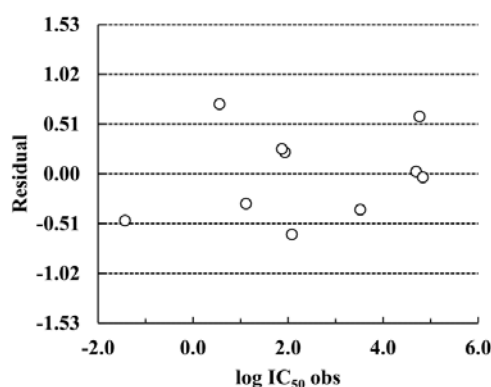
**Figura 27.** Representación gráfica de los valores de logCI<sub>50</sub> iNOS observados y calculados (arriba). Representación gráfica de los residuales calculados pertenecientes al *training set* versus los residuales obtenidos a partir de la validación cruzada (residual (cv) (abajo).

Un test de aleatoriedad fue llevado a cabo para identificar una posible regresión de tipo fortuito. Para ello, el valor de la propiedad de cada compuesto se permuta aleatoriamente y se correlaciona linealmente con los descriptores que conforman el modelo. Este proceso se repite tantas veces como sea necesario. El resultado del test de aleatoriedad se expresa representando gráficamente el coeficiente de correlación frente al de predicción,  $r^2$  vs  $Q^2$ . Los resultados se ilustran en la Figura 28, donde se aprecia que los resultados son significativos y no fortuitos. Así, todas las regresiones obtenidas con las propiedades asignadas aleatoriamente a los compuestos fueron pobres, tan solo el modelo (representado en la figura por un punto negro), obtiene valores de predicción de la propiedad óptimos.



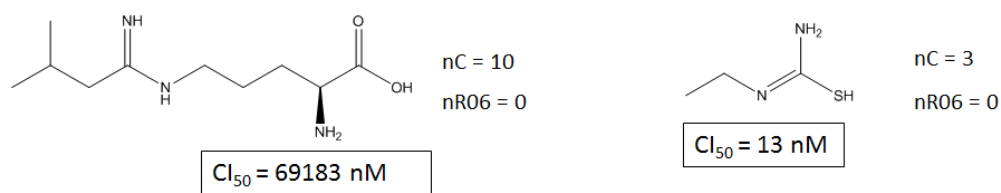
**Figura 28.** Representación gráfica del coeficiente de predicción,  $Q^2$ , frente el coeficiente de correlación,  $r^2$ , obtenidos en el estudio de aleatoriedad del modelo de predicción del  $\log CI_{50}$  para iNOS. Punto negro: modelo seleccionado.

La Figura 29, que refleja la representación de los residuales frente a los valores experimentales, recoge la posible presencia de *outliers*. Puesto que ningún compuesto supera  $\pm 2 \cdot \text{EEE}$ , como se aprecia en la Figura 29, no hay presencia de *outliers* por lo que el modelo es aplicable a todas las moléculas que forman parte del *training set*.



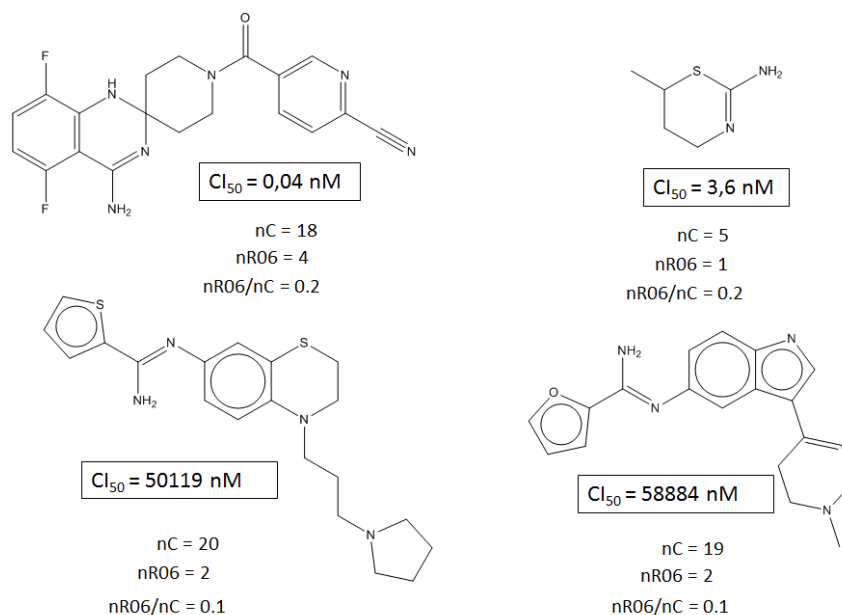
**Figura 29.** Representación gráfica del residual frente al valor observado para cada compuesto, empleando la Ecuación 5. Se considera *outlier* a todo compuesto con un valor residual predicho superior a  $\pm 2 \cdot \text{EEE}$ . ( $\text{EEE}=0.51$ ).

Si analizamos los resultados derivados de la Ecuación 5, es evidente que, en ausencia de anillos de 6 miembros (nR06), a menor número de átomos de carbono (nC) mayor poder inhibidor frente a iNOS. Este hecho se refleja en la Figura 30, donde (S)-Etilisotiourea tiene una  $CI_{50} = 13$  nM, mientras que BDBM152702 tiene una  $CI_{50} = 69183$  nM.



**Figura 30.** Compuesto del *training set* sin anillos hexagonales y sus valores de  $CI_{50}$  para iNOS.

Por el contrario, compuestos con presencia de anillos hexagonales y alto valor de la relación nR06/nC, muestran una mayor actividad inhibitoria frente a iNOS. Esto queda reflejado en la Figura 31, donde las moléculas más prometedoras como inhibidores de iNOS tiene elevada relación nR06/nC.



**Figura 31.** Compuesto del *training set* con anillos hexagonales y sus valores de  $CI_{50}$  para iNOS.

## Modelo de predicción de la inhibición de TNF- $\alpha$

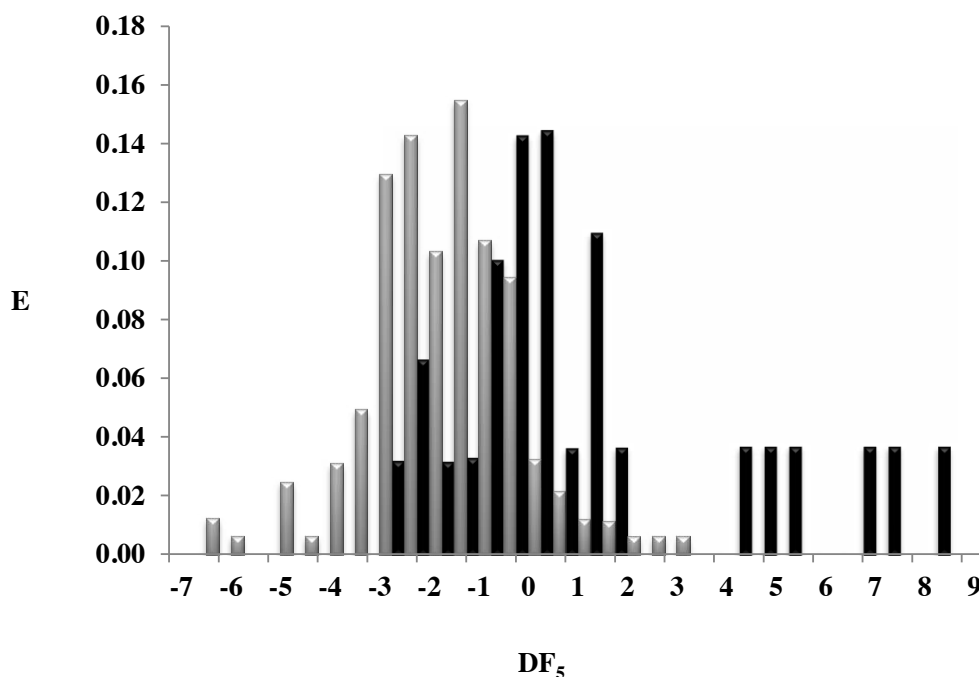
Para la construcción del modelo se emplearon un grupo de 189 moléculas activas e inactivas respecto a la inhibición de TNF- $\alpha$ . Al igual que en la Ecuación 4, el *training set* está constituido por un grupo de estructuras moleculares heterogéneas. El grupo de compuestos activos e inactivos contribuyeron equitativamente en el proceso de construcción del modelo, a pesar de su diferencia numérica. La ecuación correspondiente a FD<sub>5</sub>, discrimina compuestos inhibidores de TNF- $\alpha$ . Se compone de seis variables:

$$FD_5 = (0,39 \times nO) + (0,98 \times piPC10) - (6,79 \times ATS1e) + (11,88 \times MATS1m) - (17,85 \times MATS1e) - (5,41 \times GATS1v) + 18,40 \quad (6)$$

$$N = 189 \quad F = 12,15 \quad \lambda = 0,71 \quad p < 0,00001$$

En la ecuación anterior, hay descriptores topológicos que evalúan: el número de caminos múltiples de orden 10 (piPC10); el índice de autocorrelación de Broto-Moreau a distancia 1 ponderado por la electronegatividad de Sanderson (ATS1e); el índice de autocorrelación de Moran a distancia 1 ponderado por la masa atómica (MATS1m); el índice de autocorrelación de Moran a distancia 1 ponderado por la electronegatividad de Sanderson (MATS1e); el índice de autocorrelación Geary a distancia 1 ponderado por volumen de van der Waals (GATS1v) y el número de átomos de oxígeno (nO). A diferencia de lo que ocurre en la FD<sub>4</sub>, en la FD<sub>5</sub> aparecen índices básicamente relacionados con la presencia de ciertos heteroátomos en determinados lugares de la molécula. Por lo tanto, los factores tales como el tamaño y la electronegatividad de los átomos, evaluados por índices como MATS1m o ATS1e, juegan un papel decisivo. Por ejemplo, la mayoría de las moléculas exhiben grupos dadores de hidrógeno (NH- y OH), que determinan un aumento de la solubilidad en agua, así como una mayor capacidad para interactuar por medio de interacciones de enlaces de hidrógeno con residuos de aminoácidos del receptor de TNF- $\alpha$ .

De acuerdo con FD<sub>5</sub>, un compuesto se clasifica como inhibidor de TNF- $\alpha$  si FD<sub>5</sub> toma valores entre 0 y 9; mientras que un compuesto será etiquetado como inactivo cuando FD<sub>5</sub> adopte valores < 0 hasta > -7 (véase la Figura 32). Por último, las moléculas con valores de FD<sub>5</sub> > 9 o FD<sub>5</sub> < -7 serán consideradas como *outliers*.



**Figura 32.** Diagrama de distribución farmacológica de los inhibidores de TNF- $\alpha$  obtenidos usando la FD<sub>5</sub> (en negro se representan los inhibidores de TNF- $\alpha$  y en gris, los compuestos sin actividad inhibitoria respecto a TNF- $\alpha$ ).

Este criterio de clasificación se aplicó al *training set*. De los 189 compuestos, 19 de los 27 inhibidores de TNF- $\alpha$  se clasificaron correctamente (sensibilidad 70%), mientras que 145 de los 162 compuestos inactivos fueron también correctamente clasificados (90% especificidad) (ver Tabla 12). En este caso, el *test set* se compone de 53 compuestos (18 activos y 35 inactivos). El *test set* fue construido al azar, utilizando el 20% de todo el conjunto de datos disponibles. El modelo logró clasificar correctamente al 67% del grupo activo (12 de los 18 compuestos) y al 77% del conjunto de moléculas inactivas (27 de los 35 compuestos) (ver Tabla 12). Por otra parte, aunque la aplicación estricta de la FD<sub>5</sub> conduce a la pérdida de algunos activos (33%), lo relevante es que el 77% de los compuestos inactivos fueron clasificados correctamente. De esta manera, el riesgo de que se incluyan falso activos se reduce significativamente, lo que se traduce en una alta especificidad del modelo. En el Anexo IV y V se encuentra información adicional sobre el modelo (valores de la FD<sub>5</sub> para todos los compuestos del *training* y *test set*).

El valor del coeficiente de correlación de Matthews para  $FD_5$  fue 0,53, lo que valida su capacidad de predicción mientras que el rendimiento del modelo fue de un 77% de exactitud (MCC modificado = 1,53).

### **Modelo de predicción de la actividad protectora frente a CU inducida por DSS**

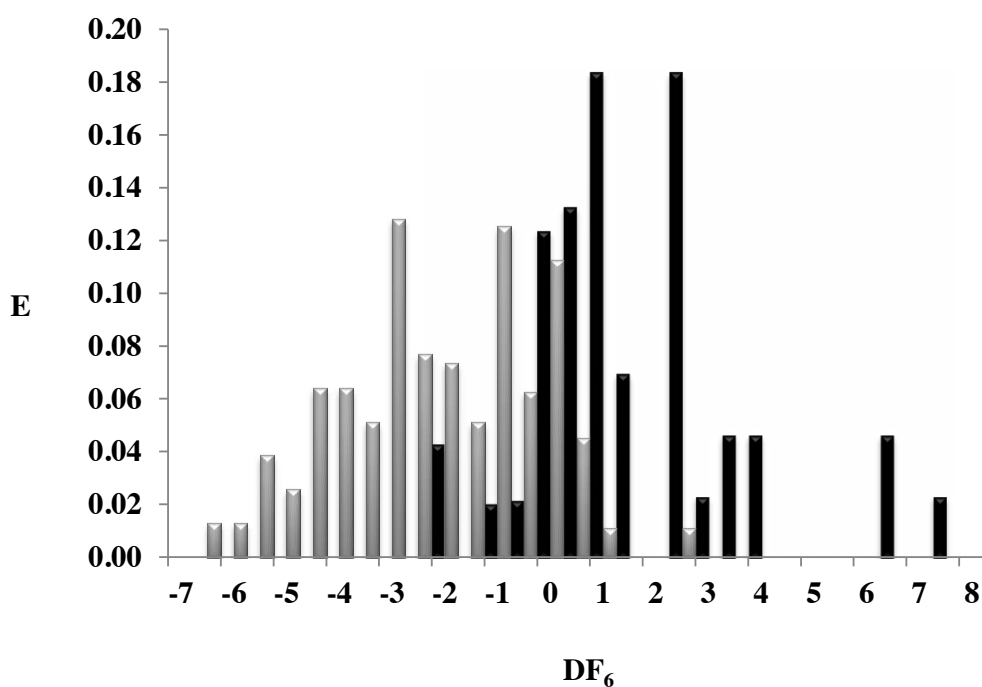
Por último, se desarrolló un modelo matemático para la identificación de nuevos compuestos con actividad protectora frente a CU inducida por DSS. El *training set* está formado por un grupo de 121 estructuras moleculares heterogéneas. Los grupos activo e inactivo contribuyeron equitativamente a la construcción del modelo, sin tener en cuenta el número de compuestos que albergaban. Esta ecuación, correspondiente a  $FD_6$ , discrimina compuestos con actividad protectora frente a CU inducida por DSS. Este modelo está formado por siete variables:

$$FD_6 = (-1,40 \times nR05) + (1,78 \times piPC03) + (1,73 \times MATS6p) - (4,93 \times GATS4e) - (3,09 \times EEig02d) + (4,48 \times EEig14d) - (3,89 \times EEig14r) + 9.99 \quad (7)$$

$$N = 121 \quad F = 14,03 \quad \lambda = 0,54 \quad p < 0.00001$$

En esta ecuación, hay descriptores topológicos que evalúan el número de anillos de 5 miembros (nR05); el número de caminos múltiples de orden 3 (piPC03); el índice de autocorrelación de Moran a distancia 6 ponderado por polarizabilidad (MATS6p); el índice de autocorrelación de Geary a distancia 4 ponderado por la electronegatividad de Sanderson (GATS4e); autovalor número 2 y 14 de la matriz de adyacencia de aristas ponderada por momentos dipolares (EEig02d y EEig14d, respectivamente) y autovalor número 14 de la matriz de adyacencia de aristas ponderada por integrales de resonancia (EEig14r). Múltiples factores juegan un papel importante, como por ejemplo la polarizabilidad molecular, los momentos dipolares o los efectos de resonancia. Además, al igual que en los casos anteriores, factores puramente estructurales tales como la presencia de anillos de pentagonales también son importantes para la actividad.

Según la  $FD_6$ , un compuesto se clasifica como activo frente a la CU inducida por DSS si  $FD_6$  toma valores entre 0 y 8; por otra parte, un compuesto será etiquetado como inactivo cuando  $FD_6$  tome valores  $< 0$  y  $> -7$  (Figura 33). Por último, las moléculas con valores de  $FD_6 > 8$  o  $FD_6 < -7$  no serán clasificables por este modelo.



**Figura 33.** Diagrama de distribución farmacológica de los compuestos activos frente a CU inducida por DSS obtenido utilizando la  $FD_6$  (en negro se representan los compuestos activos frente al modelo de CU inducida por DSS y en gris, los inactivos).

Este criterio de clasificación se aplicó al *training set*. De los 121 compuestos, 39 de los 43 fármacos como potenciales tratamientos frente CU se clasificaron correctamente (sensibilidad 91%), así como 62 de 78 compuestos inactivos (79% especificidad) (ver Tabla 12). También en este caso, se lleva a cabo una validación externa (*test set*) del modelo. El *test set* se compone de 36 compuestos (18 activos y 18 inactivos), escogidos al azar y que representan el 20% de todo el conjunto de datos disponible. El modelo logró clasificar correctamente al 78% del grupo activo (14 de los 18 compuestos) y al 83% del conjunto inactivo (15 de los 18 compuestos) (ver Tabla 12). Por otra parte, aunque la aplicación estricta de la  $FD_6$  conduce a la pérdida de algunos activos (22%), como se ha resaltado anteriormente, lo interesante y buscado es que el 83% de los compuestos inactivos fueron clasificados correctamente. Es decir, nuestro modelo presenta una alta especificidad. En el Anexo VI y VII se encuentra información adicional sobre el modelo (valores de la  $FD_6$  para todos los compuestos del *training* y *test set*).

Por último, la  $FD_6$  muestra un valor de MCC de 0,68, lo que garantiza su calidad predictiva. Y cuando el MCC se calculó de una manera ligeramente diferente, lo que se obtuvo fue un rendimiento del modelo del 84% de precisión (MCC modificado = 1,68).

### Cribado virtual

Sobre la base de los modelos matemáticos (Ecuaciones 4-7), se hizo un cribado virtual sobre un grupo de cuatro moléculas previamente seleccionadas como potenciales agentes activos frente CU mediante la inhibición de NF- $\kappa$ B e IL-6 (**AMA**, **Calceína**, **DLT** y **Ro 41-0960**) (Gálvez-Llompart, 2011; 2013). Este estudio *in silico* se realiza para predecir la potencial actividad inhibidora frente a iNOS, TNF- $\alpha$  y un modelo de CU inducida por DSS de los compuestos anteriormente citados.

**Tabla 14.** Valores de  $FD_{4-6}$  y  $\log IC_{50}$  (iNOS): probabilidad de actividad y clasificación de los potenciales compuestos anti CU.

Compuestos	Modelo iNOS					Modelo TNF- $\alpha$			Modelo CU-DSS		
	$FD_4$	P.A.	Clas.	$\log IC_{50}$	$IC_{50}$ (nM)	$FD_5$	P.A.	Clas.	$FD_6$	P.A.	Clas.
<b>AMA</b>	0,77	0,67	A	1,54	35	1,09	0,75	A	2,25	0,91	A
<b>Calceína</b>	2,10	0,88	A	4,36	22909	0,33	0,58	A	3,64	0,97	A
<b>DLT</b>	3,42	0,97	A	3,62	4169	-1,16	0,24	I	0,09	0,52	A
<b>Ro 41-0960</b>	1,27	0,77	A	1,17	15	0,69	0,67	A	0,23	0,56	A

P.A.: probabilidad de actividad.

Como se muestra en la Tabla 14, todos los compuestos son clasificados como activos bajo los mecanismos anteriormente mencionados, a excepción del DLT, que es predicho como potencialmente inactivo como inhibidor del TNF- $\alpha$ .



## ENSAYOS BIOLÓGICOS

### **Estudio *in vitro***

*Estudio del efecto de los compuestos sobre la producción de nitritos ( $NO_2^-$ ), IL-1 y TNF- $\alpha$  en células RAW 264.7*

Se expuso la línea celular RAW 264.7 a AMA, Calceína, DLT y Ro 41-0960 para investigar el efecto inhibitor sobre la actividad de iNOS (Tabla 15). El ensayo consistió en la detección de nitrito mediante la reacción de Griess (Tarpey y cols., 2004). Además, debido al hecho de que la producción de IL-1 $\beta$  se vincula con la actividad de NF- $\kappa$ B, los 4 anteriores inhibidores de NF- $\kappa$ B seleccionados como posibles tratamientos de la CU (Galvez-Llompart y cols., 2011) se ensayaron para evaluar su actividad inhibitora frente a IL-1 $\beta$  en las células RAW 264.7 (Tabla 15). Por último, como el TNF- $\alpha$  es una citocina proinflamatoria vinculada a la ruta de NF- $\kappa$ B y por lo tanto, relacionada con la CU, AMA, Calceína, DLT y Ro 41-0960 fueron testados para determinar su efecto sobre la producción de TNF- $\alpha$  en macrófagos RAW 264.7 (ver Tabla 15).

**Tabla 15.** Efecto de los compuestos seleccionados sobre la secreción de NO, IL-1 $\beta$  y TNF- $\alpha$  en el medio de cultivos de las células RAW 264.7 estimuladas con LPS (1  $\mu$ g/mL), 2 horas después de la administración de los productos.

Grupo	C ( $\mu$ M)	NO ( $\mu$ M)		IL-1 $\beta$ (pg/mL)		TNF- $\alpha$ (pg/mL)	
		NO $\pm$ EEM	I (%)	IL-1 $\beta$ $\pm$ EEM	I (%)	TNF- $\alpha$ $\pm$ EEM	I (%)
<b>Control</b>		25,41 $\pm$ 0,98	-	23,94 $\pm$ 2,36	-	11799 $\pm$ 519	-
	10	21,38 $\pm$ 1,23*	16	43,88 $\pm$ 2,07*	-	10740 $\pm$ 258	9
<b>AMA</b>	50	14,06 $\pm$ 0,64***	45	81,35 $\pm$ 1,85***	-	10200 $\pm$ 192	14
	100	8,63 $\pm$ 0,67***	66	94,29 $\pm$ 7,35***	-	8594 $\pm$ 854	27
	10	23,41 $\pm$ 1,32	8	17,15 $\pm$ 0,21	28	12787 $\pm$ 750	-
<b>Calceina</b>	50	17,50 $\pm$ 0,62***	31	53,28 $\pm$ 2,73***	-	10232 $\pm$ 426	13
	100	14,38 $\pm$ 0,36***	43	64,28 $\pm$ 6,64***	-	8103 $\pm$ 1245	31
	10	24,22 $\pm$ 0,69	5	27,71 $\pm$ 3,93	-	9430 $\pm$ 1483	20
<b>DLT</b>	50	23,54 $\pm$ 0,49	7	31,98 $\pm$ 6,03	-	10921 $\pm$ 898	7
	100	24,92 $\pm$ 1,97	2	15,90 $\pm$ 2,25	34	11381 $\pm$ 639	4
	10	13,68 $\pm$ 0,70***	46	22,28 $\pm$ 1,29	7	7929 $\pm$ 1248	33
<b>Ro 41-0960</b>	25	1,32 $\pm$ 0,40***	95	14,06 $\pm$ 1,30	41	6682 $\pm$ 754*	43

C: concentración a la que se ensayaron los compuestos; I: inhibición.

Los valores se expresan como la media (producción NO, IL-1 $\beta$  o TNF- $\alpha$ )  $\pm$  EEM y se comparan con los valores obtenidos por el control. Los resultados son representativos de al menos tres experimentos independientes por grupo. La significación estadística de las diferencias de cada grupo respecto al control se determinaron por medio del análisis de varianza ANOVA seguido del test de la t de Dunnett (diferencia respecto al control: \*p<0,05; \*\*\*p<0,001).

Como se muestra en la Tabla 15, tres de los compuestos (**AMA**, **Calceína** y **Ro 41-0960**) inhibieron la producción de NO en un 66, 43 y 95%, respectivamente, a la concentración más alta ensayada. Estos compuestos inhibieron de una manera dependiente de la concentración. Todos los compuestos excepto AMA inhibieron la producción de IL-1 $\beta$  en macrófagos RAW 264.7, al menos, a una concentración (Tabla 15), si bien los resultados no presentaron significación estadística. Cabe destacar a la **Calceína**, con un 28% de inhibición a la concentración más baja (10  $\mu$ M), a pesar de que no mostró actividad en concentraciones más altas. Este fenómeno puede ser debido a la desensibilización homóloga del receptor IL1- $\beta$ . Por último, todos los compuestos fueron capaces de reducir al menos en un 20% la producción de TNF- $\alpha$  en las células RAW 264.7 (**AMA**, **Calceína** y **Ro 41-0960** en manera concentración-dependiente, mientras que el **DLT** en modo concentración-independiente).

### **Estudio *in vivo***

#### *Efecto de los compuestos en un modelo de CU inducida por DSS*

La administración de DSS induce una colitis reproducible caracterizada por diarrea sanguinolenta, ulceraciones de la mucosa intestinal e infiltración de leucocitos. El mecanismo por el que actúa el DSS no está completamente dilucidado, pero se sabe que se comporta como un agente tóxico que afecta a las células epiteliales del colon, afectando por tanto la integridad de la mucosa intestinal. Es un modelo muy utilizado para valorar la participación de la inmunidad innata en la colitis (Wirtz y Neurath, 2007). El grado de gravedad de la enfermedad se establece mediante el índice de actividad de la enfermedad o IAE (Cooper y cols., 1993). Los parámetros que tienen en cuenta dicho índice son: pérdida de peso corporal, consistencia de las heces y sangre fecal visible, asignando la puntuación entre 0 - 4. Cero es la normalidad y cuatro es el máximo de enfermedad con una pérdida de peso > 20%. Los animales hembras BALB/c se distribuyeron a razón de 6 por grupo y se pesaron en los días 0, 4 y 8, vigilando a lo largo del experimento la cantidad de agua ingerida. El octavo día los animales son sacrificados por dislocación cervical y se extrae el colon.

Los animales reciben el DSS (3% p/v) a través del agua de bebida, mientras que los compuestos se administraron por vía oral mediante sonda a una dosis de 50 mg/kg en un volumen de 0,2 mL de agua. El fármaco de referencia utilizado fue sulfasalazina a la misma dosis y vía de administración.

Como era de esperar, los animales no tratados experimentan una reducción del peso alrededor del 15%, mientras que los animales con tratamiento se vieron protegidos a excepción del grupo que recibió Calceína (17% pérdida) (Tabla 16).

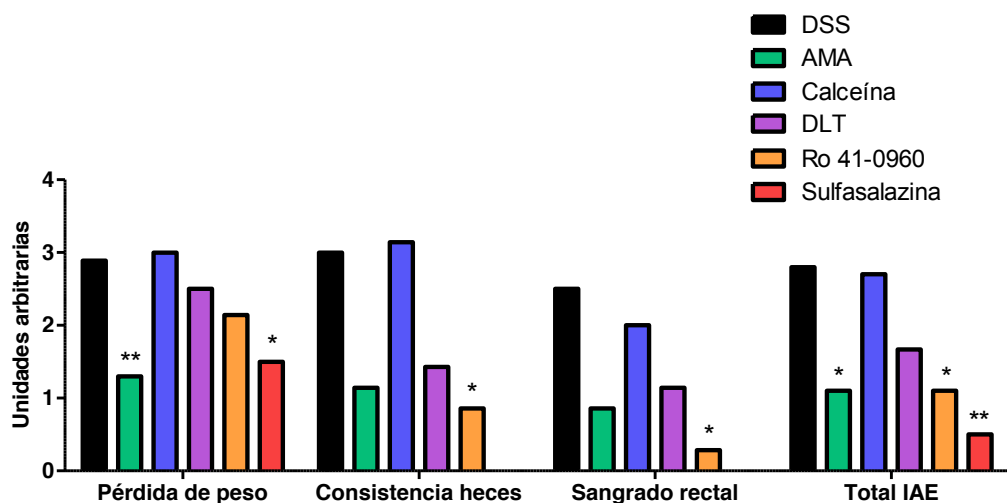
La incidencia de muerte asociada con el presente experimento podría estar relacionada con las diferencias en la microflora intestinal entre diferentes animales, lo que se sabe puede alterar el resultado de la administración de DSS (Kim y cols., 2012).

En la Figura 34 se representa el índice de actividad de la enfermedad para cada uno de los compuestos. En conjunto, se puede decir que los compuestos **AMA**, **Ro 41-0960** y el fármaco de referencia protegieron a los animales de la CU inducida por DSS (IAE = 1,1, 1,1 y 0,5, respectivamente) en comparación con el grupo de DSS (IAE= 2,8). A pesar de no impedir la pérdida de peso, **DLT** evitó la diarrea y la sangre en heces, ofreciendo un IAE inferior al control. De los tres compuestos ensayados, **Calceína** se puede considerar sin actividad *in vivo*.

**Tabla 16.** Pérdida de peso de los ratones BALB/c durante el modelo de CU inducida por DSS. Los resultados se expresan en base al control negativo y representan las medias  $\pm$  EEM. Los valores son representativos de un experimento independiente (n = 5-11 animales por grupo).

Grupo	Evaluación del peso corporal (g)			Pérdida de peso (%)	Nº animales muertos/total
	Día 0	Día 4	Día 8		
<i>Control</i>	19,04 $\pm$ 0,57	19,24 $\pm$ 0,22	19,68 $\pm$ 0,45 <sup>***</sup>	-	0/6
<i>DSS</i>	18,67 $\pm$ 0,25	18,16 $\pm$ 0,44	15,91 $\pm$ 0,35 <sup>###</sup>	15	2/11
<i>AMA</i>	18,70 $\pm$ 0,28	18,00 $\pm$ 0,58	18,03 $\pm$ 0,42 <sup>*</sup>	3	1/8
<i>Calceína</i>	19,29 $\pm$ 0,23	19,06 $\pm$ 0,24	16,03 $\pm$ 0,64	17	0/7
<i>DLT</i>	18,98 $\pm$ 0,34	18,65 $\pm$ 0,26	16,74 $\pm$ 0,71	12	1/7
<i>Ro 41-0960</i>	18,43 $\pm$ 0,19	18,30 $\pm$ 0,27	17,03 $\pm$ 0,22	8	0/7
<i>Sulfasalazina</i>	17,56 $\pm$ 0,19	17,24 $\pm$ 0,83	16,65 $\pm$ 0,57	5	1/5

Las diferencias entre el grupo control (animales no tratados) y los otros grupos se determinaron por medio del análisis de varianza ANOVA seguido del test de la t de Dunnett (diferencia respecto al control: <sup>###</sup>p<0,001 y diferencia respecto al grupo DSS: <sup>\*</sup>p<0,05; <sup>\*\*\*</sup>p<0,001).



**Figura 34.** Valores de puntuación del IAE para los compuestos en el modelo de CU inducida por DSS en ratones BALB/c. Significación estadística de las diferencias entre el grupo control (DSS) y los grupos tratados se determinaron por medio del análisis de varianza ANOVA seguido del test de la t de Dunnett (\* p < 0,05; \*\* p < 0,01) (n = 5-11 animales por grupo).

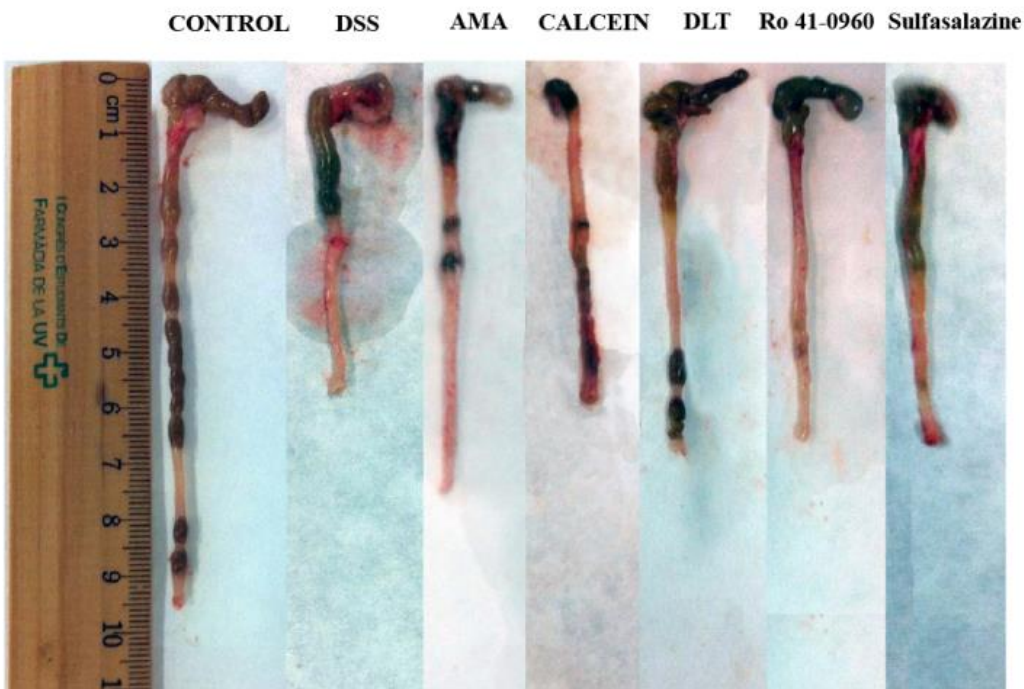
Una vez sacrificados los animales, la evaluación macroscópica del colon nos permite establecer la gravedad de la enfermedad, procediendo a la medida de la longitud del intestino grueso. En la Figura 35 se muestra una imagen representativa del colon procedente de los animales sano, enfermos y tratados con los diferentes compuestos. El DSS provoca un acortamiento del mismo ( $5,4 \pm 0,2$  cm) frente al tamaño de los ratones sanos ( $8,8 \pm 0,2$  cm) (Tabla 17). Como ya se había visto con el IAE, todos los compuestos protegieron a los animales del acortamiento, destacan los compuestos **AMA** y **Ro 41-0960** (7,1 y 6,8 cm vs 5,4 cm del control) y de nuevo **Calceína** resultó el compuesto de menos activo (6,1 cm vs 5,4 cm control).

**Tabla 17.** Efecto de los compuestos sobre el acortamiento de la longitud del colon en la CU inducida por DSS en ratones BALB/c.

Grupo	Longitud del colon (cm)
<i>Control</i>	8,8 ± 0,2 <sup>***</sup>
<i>DSS</i>	5,4 ± 0,2 <sup>###</sup>
<i>AMA</i>	7,1 ± 0,4 <sup>**</sup>
<i>Calceina</i>	6,1 ± 0,3
<i>DLT</i>	6,7 ± 0,5 <sup>*</sup>
<i>Ro 41-0960</i>	6,8 ± 0,2 <sup>*</sup>
<i>Sulfasalazina</i>	6,6 ± 0,5

*Diferencias significativas con el grupo DSS (\*p<0,05; \*\*p<0,01; \*\*\*p<0,001) y del grupo de control (<sup>###</sup>p<0,001); análisis unidireccional de varianza seguido por la prueba t de Dunnett).*

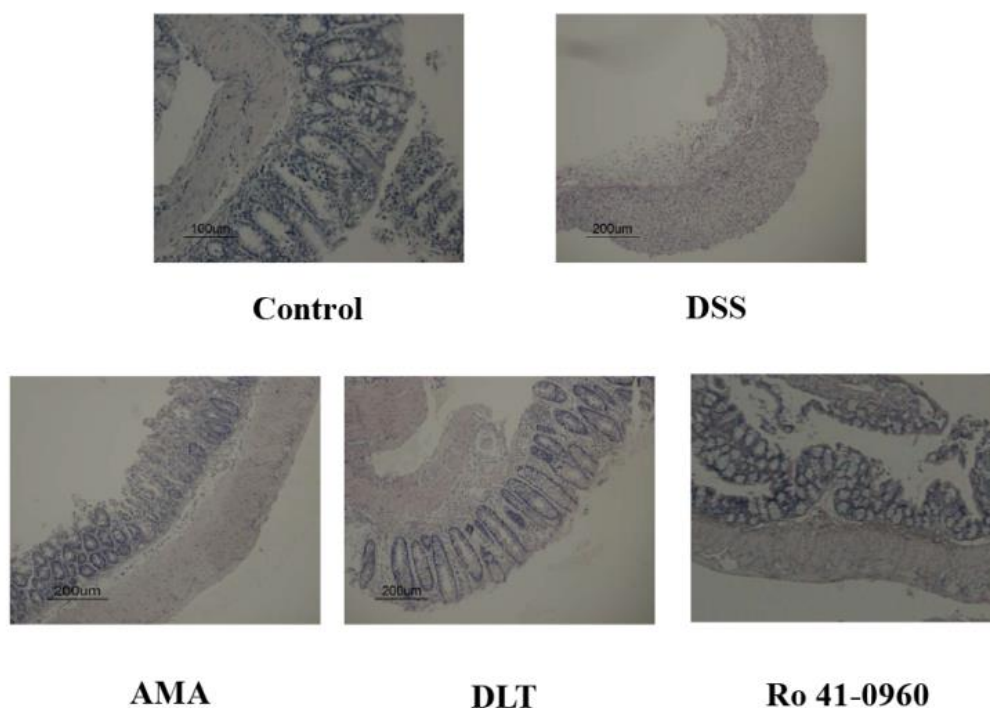
*Los resultados son representativos de un experimento independiente por grupo (n = 4-9 animales) y representan la media ± EEM.*



**Figura 35.** Imagen representativa del efecto de los compuestos sobre el acortamiento del colon por CU. El grupo de control recibió sólo agua *ad libitum*; el grupo DSS, agua con DSS al 3% durante 7 días. El resto de los grupos recibió el agua con DSS al 3% y tres dosis de 50 mg/kg de AMA, Calceína, DLT, Ro 41-0960 y sulfasalazina.

### **Análisis histológico**

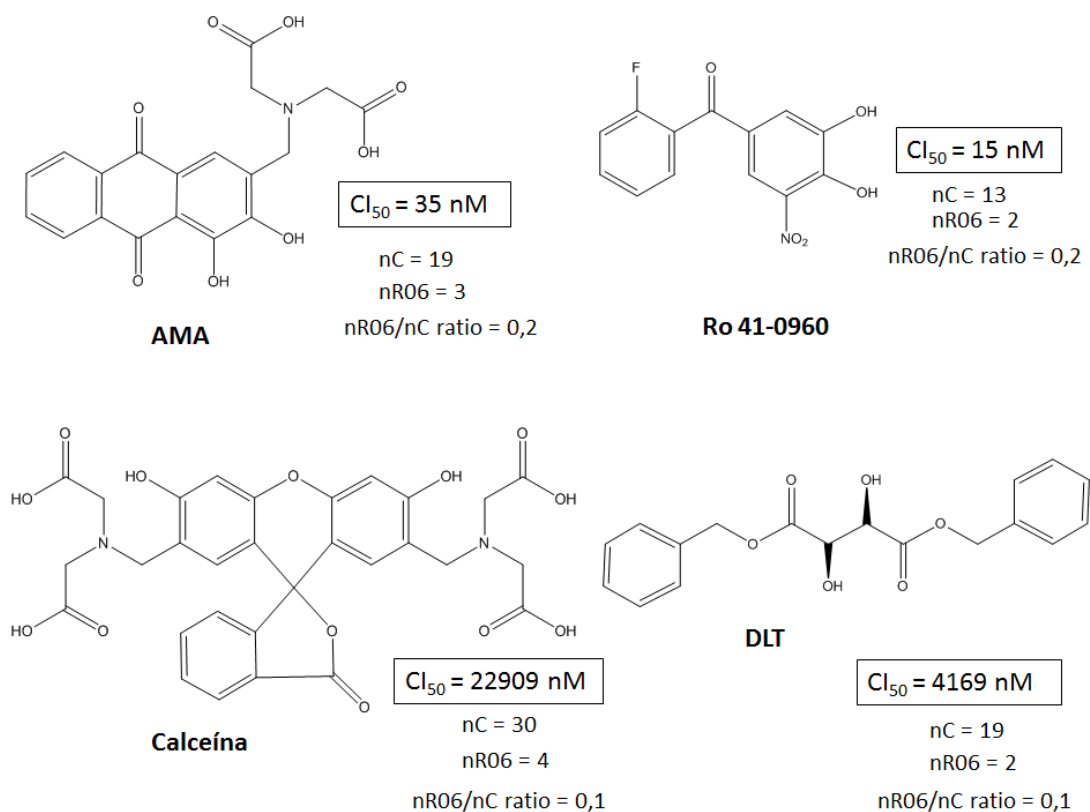
El estudio histopatológico de todos los grupos en relación con CU inducida por DSS se muestran en la Figura 36; las secciones transversales del colon del grupo control (sólo agua) (Figura 36 control) muestra una morfología normal, sin inflamación o úlceras, con arquitectura intacta de criptas e infiltración celular fisiológica, junto con la *muscularis* normal. No obstante, mediante el análisis de la sección transversal de DSS, los cambios llamativos en la arquitectura de colon típicamente relacionadas con el proceso inflamatorio se pueden observar fácilmente (Figura 36 DSS), como la pérdida total de la arquitectura de las criptas, la pérdida de la estructura de las células caliciformes, edema y la infiltración de neutrófilos y ulceración de la membrana mucosa. Por el contrario, las muestras de los 3 compuestos ensayados (**AMA**, **DLT** y **Ro 41-0960**) mostraron una significativa mejoría clínica, tales como, células caliciformes intactas, la infiltración celular baja, arquitectura de las criptas intactas y *muscularis* normal.



**Figura 36.** Evaluación histológica del efecto protector de AMA, DLT y Ro 41-0960 frente a CU inducida por DSS en ratones. Cortes histológicos representativos de secciones transversales teñidas con hematoxilina y eosina (200  $\mu\text{m}$ ).

En cuanto a la información codificada en las ecuaciones matemático-topológicas respecto a los requisitos fisicoquímicos y estructurales de los compuestos seleccionados, debido a su alta complejidad, hay muchos factores que intervienen. Entre los parámetros fisicoquímicos más pertinentes, se muestra la polarizabilidad molecular, que es una medida de la movilidad de la carga dentro de la molécula y el momento dipolar, que es una medida de la separación de la carga neta intramolecular. Entre los requisitos estructurales, la flexibilidad molecular, así como la presencia de ciertos fragmentos, tales como los anillos de 6 miembros, por mencionar sólo dos características importantes. De hecho, la presencia de anillos hexagonales parece ser relevante para la actividad inhibidora de la iNOS en la CU. Cuando se calcula la relación  $nR06/nC$  para cada compuesto (Figura 37), como se esperaba, los compuestos con proporciones más grandes (tales como **AMA** y **Ro 41-0960**) se predijeron como los más potentes (valores de  $CI_{50}$  de 35 y 15 nM, respectivamente). Esto se comprobó por medio de los resultados *in vitro* para la inhibición de iNOS, donde **AMA** y **Ro 41-0960** muestran valores de  $CI_{50}$  de 53 y 11  $\mu\text{M}$ , respectivamente.





**Figura 37.** Compuestos seleccionados por Topología molecular como potenciales agentes protectores frente a CU y sus valores de  $CI_{50}$  del inhibidor de iNOS predicho por la Ecuación 5.

A modo de resumen en la Tabla 18, encontramos reagrupados los modelos construidos y los resultados *in vitro* e *in vivo* obtenidos para los 4 compuestos identificados como potenciales agentes protectores frente a CU: AMA, Calceína, DLT y Ro 41-0960. Como se aprecia todos los modelos logran predecir correctamente la actividad del 75% de los compuestos seleccionados, si bien cuanto se refiere al compuesto **Ro 41-0960** los modelos logran predecir su actividad correctamente en el 100% de los casos.

**Tabla 18.** Sinopsis de los ensayos *in vitro* e *in vivo* llevados a cabo para validar los modelos y las actividades farmacológicas predichas por estos, para los cuatro compuestos seleccionados.

MODELO	COMPUESTO	RESULTADO
		<i>IN VITRO</i>
NF-κB	AMA	-
	Calceína	+
	DLT	+
	Ro 41-0960	+
IL-6	AMA	+
	Calceína	+
	DLT	+
	Ro 41-0960	+
iNOS	AMA	+
	Calceína	+
	DLT	-
	Ro 41-0960	+
TNF-α	AMA	+
	Calceína	+
	DLT	+
	Ro 41-0960	+
		<i>IN VIVO</i>
CU	AMA	A
	Calceína	I
	DLT	A
	Ro 41-0960	A

+: inhibición  $\geq$  al 20% en ensayos *in vitro*.

A: activo en el modelo de CU-inducida por DSS.

I: inactivo en el modelo de CU-inducida por DSS.

Por supuesto, se requeriría un estudio más detallado de las relaciones estructura-actividad más importantes en función de la interacción de los compuestos seleccionados con los correspondientes *targets*, pero este análisis está fuera del objetivo de la presente Tesis Doctoral, necesitándose estudios relacionados con disciplinas tales como Docking molecular, mecánica y dinámica moleculares y mecánica cuántica.

Además, los compuestos identificados en esta Tesis Doctoral se prestan a ser objeto de futuros estudios tanto a nivel *in silico* (optimización de modelos y selección de compuestos más potentes frente a la CU) como a nivel farmacológico (investigar si estos compuestos pudieran afectar a otros mecanismos de acción implicados en la CU).

# CONCLUSIONES



## VII. CONCLUSIONES

1. Se han construido y validado varios modelos, basados en la Topología molecular, para predecir la actividad farmacológica relacionada con la inhibición de enzimas proinflamatorias, tales como COX e iNOS, citocinas (IL-6 y TNF- $\alpha$ ) y factores de transcripción como el NF- $\kappa$ B. Además, se han obtenido modelos matemáticos predictivos de actividad en un tipo concreto de inflamación crónica, como es la colitis ulcerosa. Esos modelos se han aplicado al cribado virtual de diferentes bases de datos (Sigma-Aldrich, Microsource Pure Collection y Merck index), que incluían más de 100000 compuestos, a partir de los cuales se seleccionaron 42 como potenciales antiinflamatorios. De estos, 14 fueron predichos como antiinflamatorios inhibidores de la COX, y 28 como inhibidores de NF- $\kappa$ B e IL-6, potencialmente eficaces frente a colitis ulcerosa.
2. Atendiendo a los criterios expuestos en los Artículos **1** y **3**, algunos de los compuestos seleccionados *in silico* han sido ensayados tanto *in vitro* como *in vivo*, obteniéndose resultados que han corroborado lo determinado por los modelos matemáticos.
3. Los compuestos **5**, **8**, **9** y **14**, ampliamente descritos en el **Artículo 1**, presentaron actividad antiinflamatoria *in vivo* en un test típico para inhibidores de COX, como es el edema plantar por carragenina. Destaca el compuesto **8**, que a la dosis de 10 mg/kg y a las 3 h, inhibió el edema un 47%, es decir, a un nivel similar al de la indometacina. Este mismo compuesto, a una concentración de 50  $\mu$ M, inhibió *in vitro* la producción de nitritos y PGE<sub>2</sub>. Sobre la base de los índices empleados en el modelo, podemos concluir que la presencia de determinados heteroátomos electronegativos (por ejemplo, halógenos), así como la existencia de alta conjugación en las moléculas, parecen ser elementos relevantes en la actividad antiinflamatoria.
4. Por otra parte, se han encontrado también nuevos compuestos activos en colitis ulcerosa mediante la inhibición del NF- $\kappa$ B. De los 28 compuestos inicialmente seleccionados, 13 tenían actividades farmacológicas descritas en la literatura (antiinflamatorios, antioxidantes y antiulcerosos), quedando de este modo validado el modelo. En términos de relación estructura-actividad, se puede

apreciar que propiedades estructurales y electrónicas determinan su actividad inhibitoria.

5. Los compuestos seleccionados en el **Artículo 2** también inhiben la producción de IL-6, lo que respalda igualmente su potencial actividad frente a la colitis ulcerosa. Esta predicción, cuyos resultados experimentales se describen en el **Artículo 3**, permitió la selección de 4 compuestos (el ácido Alizarin-3-metiliminodiacético (**AMA**), el (+)-Dibencil-L-tartrato (**DLT**), la **Calceína** y el **Ro 41-0960**) para su estudio *in vitro* e *in vivo*. Tres de ellos inhibieron la producción de IL-6 de modo significativo en dos líneas celulares, destacando el compuesto **Ro 41-0960** que inhibió la producción de IL-6 con una  $CI_{50}$  en células Caco-2 de 12  $\mu\text{M}$  y de 0,4  $\mu\text{M}$  en RAW 264.7. El compuesto **DLT** presentó valores de  $CI_{50}$  de 19  $\mu\text{M}$  en ambas líneas celulares, mientras que la **Calceína**, un producto natural, solo fue activa en la línea celular Caco-2 ( $CI_{50} = 45 \mu\text{M}$ ). Por último, los descriptores topológicos empleados en el modelo determinan que la actividad inhibitoria de IL-6 guarda relación con el grado de conjugación de la molécula así como con el volumen de Van der Waals y la electronegatividad de los átomos presentes.
6. Atendiendo a los resultados obtenidos y expuestos en el **Artículo 4**, podemos afirmar que los mejores candidatos como cabezas de serie para el tratamiento de la colitis ulcerosa, son **AMA** y **Ro 41-0960**. Ambos productos inhibieron *in vitro* la producción de los mediadores inflamatorios, destacando los resultados *in vivo* del compuesto **Ro 41-0960**, que a la dosis de 50 mg/kg redujo todos los signos de la colitis ulcerosa inducida por DSS en ratón, en un grado similar al de la sulfasalazina. Los descriptores empleados en los modelos sugieren que parámetros químico-físicos tales como la polarizabilidad, el momento dipolar y la flexibilidad molecular, así como la presencia de ciertos grupos tales como anillos hexagonales, son determinantes de la actividad de nuestras moléculas.
7. Por último, se ha demostrado que la Topología molecular representa una herramienta eficaz para el diseño y el descubrimiento de nuevos compuestos cabezas de serie como potencial tratamiento de enfermedades inflamatorias.

# BIBLIOGRAFÍA





## VIII. BIBLIOGRAFÍA

- Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry & Biology*, 10(9), 787-797.
- Arai, Y., Takanashi, H., Kitagawa, H., & Okayasu, I. (1998). Involvement of interleukin-1 in the development of ulcerative colitis induced by dextran sulfate sodium in mice. *Cytokine*, 10(11), 890-896.
- Ardizzone, S., & Porro, G. B. (2005). Biologic therapy for inflammatory bowel disease. *Drugs*, 65(16), 2253-2286.
- Arora, Z., & Shen, B. (2015). Biological therapy for ulcerative colitis. *Gastroenterology Report*, 3(2), 103-109.
- Assreuy, A., Nogueira, N., Cavada, B., & Vale, M. (2005). Anti-inflammatory and antimicrobial effect of lectin from *Lonchocarpus sericeus* seeds in an experimental rat model of infectious peritonitis. *Journal of Pharmacy and Pharmacology*, 57, 919-922.
- Atreya, I., Atreya, R., & Neurath, M. (2008). NF- $\kappa$ B in inflammatory bowel disease. *Journal of Internal Medicine*, 263(6), 591-596.
- Basak, S. C. (2011). Role of mathematical chemodescriptors and proteomics-based biodescriptors in drug discovery. *Drug Development Research*, 72(2), 225-233.
- Benavente-Garcia, O., & Castillo, J. (2008). Update on uses and properties of citrus flavonoids: New findings in anticancer, cardiovascular, and anti-inflammatory activity. *Journal of Agricultural and Food Chemistry*, 56(15), 6185-6205.
- Bodger, K., Yen, L., Szende, A., Sharma, G., Chen, Y. J., McDermott, J., et al. (2014). Medical resource utilization and associated costs in patients with ulcerative colitis in the UK: A chart review analysis. *European Journal of Gastroenterology & Hepatology*, 26(2), 213-221.
- Borm, M., Bodegraven, A., Mulder, C., Kraal, G., & Bouma, G. (2005). A NFKB1 promoter polymorphism is involved in susceptibility to ulcerative colitis. *International Journal of Immunogenetics*, 32(6), 401-405.
- Brophy, S., Taylor, G., Blake, D., & Calin, A. (2003). The interrelationship between sex, susceptibility factors, and outcome in ankylosing spondylitis and its associated disorders including inflammatory bowel disease, psoriasis, and iritis. *The Journal of Rheumatology*, 30(9), 2054-2058.
- Calder, P. C., Albers, R., Antoine, J., Blum, S., Bourdet-Sicard, R., Ferns, G., et al. (2009). Inflammatory disease processes and interactions with nutrition. *British Journal of Nutrition*, 101(S1), 1-45.
- Cashman, J., Burt, H., Springate, C., Gleave, J., & Jackson, J. (2004). Camptothecin-loaded films for the prevention of postsurgical adhesions. *Inflammation Research*, 53(8), 355-362.
- Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., et al. (2014). QSAR modeling: Where have you been? where are you going to? *Journal of Medicinal Chemistry*, 57(12), 4977-5010.

- Clarke, M., Stefanidis, A., & Spencer, S. (2012). Postnatal overfeeding leads to obesity and exacerbated febrile responses to lipopolysaccharide throughout life. *Journal of Neuroendocrinology*, 24(3), 511-524.
- Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, 110(18), 5959-5967.
- Crockett, S. D., Porter, C. Q., Martin, C. F., Sandler, R. S., & Kappelman, M. D. (2010). Isotretinoin use and the risk of inflammatory bowel disease: A case-control study. *The American Journal of Gastroenterology*, 105(9), 1986-1993.
- Delgado, A., Minguillón, C., & Joglar, J. (2003). Introducción a la química terapéutica.
- Delhalle, S., Blasius, R., Dicato, M., & Diederich, M. (2004). A beginner's guide to NF- $\kappa$ B signaling pathways. *Annals of the New York Academy of Sciences*, 1030(1), 1-13.
- Dignass, A., & Buhr, H. (2004). *Topical steroids in gastroenterology and hepatology* Springer Science & Business Media.
- Dignass, A., Eliakim, R., Magro, F., Maaser, C., Chowers, Y., Geboes, K., et al. (2012). Second european evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: Definitions and diagnosis. *Journal of Crohn's and Colitis*, 6(10), 965-990.
- El-Alfy, T., El-Sawi, S., Sleem, A., & Moawad, D. (2010). Investigation of flavonoidal content and biological activities of *chorisia insignis* hbk. leaves. *Australian Journal of Basic and Applied Sciences*, 4(6), 1334-1348.
- Ellinghaus, D., Bethune, J., Petersen, B., & Franke, A. (2015). The genetics of Crohn's disease and ulcerative colitis—status quo and beyond. *Scandinavian Journal of Gastroenterology*, 50(1), 13-23.
- Ellinghaus, D., Jostins, L., Spain, S. L., Cortes, A., Bethune, J., Han, B., et al. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nature Genetics*, 48, 510-518.
- Estrada, E. (1996). Spectral moments of the edge adjacency matrix in molecular graphs. 1. definition and applications to the prediction of physical properties of alkanes. *Journal of Chemical Information and Computer Sciences*, 36(4), 844-849.
- Fedorak, R. N., Wong, K., & Bridges, R. (2010). Canadian digestive health foundation public impact series. inflammatory bowel disease in canada: Incidence, prevalence, and direct and indirect economic impact. *Canadian Journal of Gastroenterology = Journal Canadien De Gastroenterologie*, 24(11), 651-655.
- Fina, D., & Pallone, F. (2008). What is the role of cytokines and chemokines in IBD? *Inflammatory Bowel Diseases*, 14 Suppl 2, S117-8.
- Fiorino, G., Peyrin-Biroulet, L., Repici, A., Malesci, A., & Danese, S. (2011). Adalimumab in ulcerative colitis: Hypes and hopes. *Expert Opinion on Biological Therapy*, 11(1), 109-116.
- Fujita, T. (2012). *The Birth of QSAR-in Memory of Professor Corwin Hansch*.

- Fujita, T., Kawazu, K., Mitsui, T., & Katsumi, M. (1967). Studies on plant growth regulators—XX.: Structure/activity relationship of AC-alkyl-hydro-1-naphthoic acids and related compounds. *Phytochemistry*, 6(6), 889-897.
- Furnival, G. M. (1971). All possible regressions with less computation. *Technometrics*, 13(2), 403-408.
- Gálvez, J., Gálvez-Llompert, M., & García-Domenech, R. (2012). Molecular topology as a novel approach for drug discovery. *Expert Opinion on Drug Discovery*, 7(2), 133-153.
- Galvez-Llompert, M., Iglesias, María del Carmen Recio, Gálvez, J., & García-Domenech, R. (2013). Novel potential agents for ulcerative colitis by molecular topology: Suppression of IL-6 production in caco-2 and RAW 264.7 cell lines. *Molecular Diversity*, 17(3), 573-593.
- Garcia, I. (2013). Update of QSAR & docking & alignment studies of the DNA polymerase inhibitors. *Current Bioinformatics*, 8(4), 472-482.
- García-Domenech, R., Gálvez, J., de Julián-Ortiz, J. V., & Pogliani, L. (2008). Some new trends in chemical graph theory. *Chemical Reviews*, 108(3), 1127-1169.
- Gomollón, F., García-López, S., Sicilia, B., & Gisbert, J. P. (2013). Guía clínica GETECCU del tratamiento de la colitis ulcerosa elaborada con la metodología GRADE. *Gastroenterol Hepatol*, 36(8), e1-e47.
- Hansch, C. (1971). Quantitative structure-activity relationships in drug design. *Drug Design*, 1, 271-342.
- Hansch, C., & Steward, A. R. (1964). The use of substituent constants in the analysis of the structure-activity relationship in penicillin derivatives. *Journal of Medicinal Chemistry*, 7(6), 691-694.
- Huang, L. (1985). Pharmacological effect of the organic acid of achillea alpina. *Zhong Yao Tong Bao (Beijing, China: 1981)*, 10(11), 38-40.
- Hulsmans, M., Sinnaeve, P., Van der Schueren, B., Mathieu, C., Janssens, S., & Holvoet, P. (2012). Decreased miR-181a expression in monocytes of obese patients is associated with the occurrence of metabolic syndrome and coronary artery disease. *The Journal of Clinical Endocrinology & Metabolism*, 97(7), E1213-E1218.
- Imbimbo, B. P. (2009). An update on the efficacy of non-steroidal anti-inflammatory drugs in alzheimer's disease. *Expert Opinion on Investigational Drugs*, 18(8), 1147-1168.
- Izenman, A. J. (2013). Linear discriminant analysis. *Modern multivariate statistical techniques* (pp. 237-280) Springer.
- Jain, H. K., & Agrawal, R. K. (2007). Topological descriptors in modeling tumor necrosis factor alpha inhibitory activity of xanthines, pteridinediones and related compounds. *Internet Electron J Mol Des*, 6, 218-228.
- Jess, T., Simonsen, J., Nielsen, N. M., Jorgensen, K. T., Bager, P., Ethelberg, S., et al. (2011). Enteric salmonella or campylobacter infections and the risk of inflammatory bowel disease. *Gut*, 60(3), 318-324.

- Jian, Y. T., Mai, G. F., Wang, J. D., Zhang, Y. L., Luo, R. C., & Fang, Y. X. (2005). Preventive and therapeutic effects of NF-kappaB inhibitor curcumin in rats colitis induced by trinitrobenzene sulfonic acid. *World Journal of Gastroenterology*, *11*(12), 1747-1752.
- Jurjus, A., Eid, A., Al Kattar, S., Zeenny, M. N., Gerges-Geagea, A., Haydar, H., et al. (2016). Inflammatory bowel disease, colorectal cancer and type 2 diabetes mellitus: The links. *BBA Clinical*, *5*, 16-24.
- Kang, B., Chung, E., Yun, Y., LEE, M. K., LEE, Y. R., LEE, K., et al. (2001). Inhibitory effects of anti-inflammatory drugs on interleukin-6 bioactivity. *Biological and Pharmaceutical Bulletin*, *24*(6), 701-703.
- Karban, A. S., Okazaki, T., Panhuysen, C. I., Gallegos, T., Potter, J. J., Bailey-Wilson, J. E., et al. (2004). Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. *Human Molecular Genetics*, *13*(1), 35-45.
- Khalili, H. (2015). Risk of inflammatory bowel disease with oral contraceptives and menopausal hormone therapy: Current evidence and future directions. *Drug Safety*, *39*(3):193-197.
- Khodabandehloo, H., Gorgani-Firuzjaee, S., Panahi, G., & Meshkani, R. (2016). Molecular and cellular mechanisms linking inflammation to insulin resistance and  $\beta$ -cell dysfunction. *Translational Research*, *167*(1), 228-256.
- Kim, J. J., Shajib, M. S., Manocha, M. M., & Khan, W. I. (2012). Investigating intestinal inflammation in DSS-induced model of IBD. *JoVE (Journal of Visualized Experiments)*, (60), e3678-e3678.
- King, J., Dawson, J., Esser, R., Fujimoto, R., Kimble, E., Maniara, W., et al. (2009). Preclinical pharmacology of robenacoxib: A novel selective inhibitor of cyclooxygenase-2. *Journal of Veterinary Pharmacology and Therapeutics*, *32*(1), 1-17.
- Klebe, G., & Abraham, U. (1999). Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *Journal of Computer-Aided Molecular Design*, *13*(1), 1-10.
- Kolios, G., Rooney, N., Murphy, C. T., Robertson, D. A., & Westwick, J. (1998). Expression of inducible nitric oxide synthase activity in human colon epithelial cells: Modulation by T lymphocyte derived cytokines. *Gut*, *43*(1), 56-63.
- Kubinyi, H. (1993). *3D QSAR in drug design: Volume 1: Theory methods and applications* Springer Science & Business Media.
- Lawrance, I. C. (2015). Early investigational TNF receptor antagonists for the treatment of ulcerative colitis. *Expert Opinion on Investigational Drugs*, *24*(6), 761-768.
- Lee, J. Y., Kim, J. S., Kim, J. M., Kim, N., Jung, H. C., & Song, I. S. (2007). Simvastatin inhibits NF- $\kappa$ B signaling in intestinal epithelial cells and ameliorates acute murine colitis. *International Immunopharmacology*, *7*(2), 241-248.
- Li, J., Huang, W., Zhang, H., Wang, X., & Zhou, H. (2007). Synthesis of andrographolide derivatives and their TNF- $\alpha$  and IL-6 expression inhibitory activities. *Bioorganic & Medicinal Chemistry Letters*, *17*(24), 6891-6894.

- Li, Y., de Haar, C., Chen, M., Deuring, J., Gerrits, M. M., Smits, R., et al. (2010). Disease-related expression of the IL6/STAT3/SOCS3 signalling pathway in ulcerative colitis and ulcerative colitis-related carcinogenesis. *Gut*, 59(2), 227-235.
- Lipkus, A. H., Yuan, Q., Lucas, K. A., Funk, S. A., Bartelt Iii, W. F., Schenck, R. J., et al. (2008). Structural diversity of organic chemistry. A scaffold analysis of the CAS registry. *The Journal of Organic Chemistry*, 73(12), 4443-4451.
- Long, M. D., Kappelman, M. D., Martin, C. F., Chen, W., Anton, K., & Sandler, R. S. (2016). Role of nonsteroidal anti-inflammatory drugs in exacerbations of inflammatory bowel disease. *Journal of Clinical Gastroenterology*, 50(2), 152-156.
- Lopez San Roman, A., & Bermejo, F. (2004). Inflammatory bowel disease in Spain: Problems grow. *Revista Espanola De Enfermedades Digestivas : Organo Oficial De La Sociedad Espanola De Patologia Digestiva*, 96(5), 291-296.
- Low, D., Nguyen, D. D., & Mizoguchi, E. (2013). Animal models of ulcerative colitis and their application in drug research.
- Lucke, K., Miehle, S., Jacobs, E., & Schuppler, M. (2006). Prevalence of bacteroides and prevotella spp. in ulcerative colitis. *Journal of Medical Microbiology*, 55(5), 617-624.
- Lyons, C. L., Kennedy, E. B., & Roche, H. M. (2016). Metabolic inflammation-differential modulation by dietary constituents. *Nutrients*, 8(5), 247.
- Macalino, S. J. Y., Gosu, V., Hong, S., & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, 38(9), 1686-1701.
- Marante, F. T., Castellano, A. G., Rosas, F. E., Aguiar, J. Q., & Barrera, J. B. (2003). Identification and quantitation of allelochemicals from the lichen lethariella canariensis: Phytotoxicity and antioxidative activity. *Journal of Chemical Ecology*, 29(9), 2049-2071.
- Marquez-Martin, A., De La Puerta, R., Fernandez-Arche, A., Ruiz-Gutierrez, V., & Yaqoob, P. (2006). Modulation of cytokine secretion by pentacyclic triterpenes from olive pomace oil in human mononuclear cells. *Cytokine*, 36(5), 211-217.
- McLeod, M. J. (1980). Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. *Teratology*, 22(3), 299-301.
- Mease, P. (2012). Update on treatment of psoriatic arthritis. *Bulletin of the NYU Hospital for Joint Diseases*, 70(3), 167.
- Meyer, S. C., & Levine, R. L. (2014). Molecular pathways: Molecular basis for sensitivity and resistance to JAK kinase inhibitors. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 20(8), 2051-2059.
- Mishchenko, N., Fedoreev, S., Bryukhanov, V., Zverev, Y. F., Lampatov, V., Azarova, O., et al. (2007). Chemical composition and pharmacological activity of anthraquinones from rubia cordifolia cell culture. *Pharmaceutical Chemistry Journal*, 41(11), 605-609.
- Molodecky, N. A., & Kaplan, G. G. (2010). Environmental risk factors for inflammatory bowel disease. *Gastroenterology & Hepatology*, 6(5), 339-346.

- Múzes, G., Molnár, B., Tulassay, Z., & Sipos, F. (2012). Changes of the cytokine profile in inflammatory bowel diseases. *World J Gastroenterol*, 18(41), 5848-5861.
- Ng, S. C., & Chan, F. K. (2010). NSAID-induced gastrointestinal and cardiovascular injury. *Current Opinion in Gastroenterology*, 26(6), 611-617.
- Pandurangan, A. K., Mohebbali, N., Hasanpourghadi, M., Looi, C. Y., Mustafa, M. R., & Mohd Esa, N. (2016). Boldine suppresses dextran sulfate sodium-induced mouse experimental colitis: NF- $\kappa$ B and IL-6/STAT3 as potential targets. *Biofactors*, doi: 10.1002/biof.1267.
- Papi, C., Aratari, A., Moretti, A., Mangone, M., Margagnoni, G., Koch, M., et al. (2010). Oral beclomethasone dipropionate as an alternative to systemic steroids in mild to moderate ulcerative colitis not responding to aminosalicylates. *Digestive Diseases and Sciences*, 55(7), 2002-2007.
- Parish, C. R. (1999). Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunology and Cell Biology*, 77(6), 499-508.
- Park, S. C., & Jeon, Y. T. (2015). Current and emerging biologics for ulcerative colitis. *Gut & Liver*, 9(1), 18-27.
- Pedersen, J., Coskun, M., Soendergaard, C., Salem, M., & Nielsen, O. H. (2014). Inflammatory pathways of importance for management of inflammatory bowel disease. *World J Gastroenterol*, 20(1), 64-77.
- Pelzer, L. E., Guardia, T., Juarez, A. O., & Guerreiro, E. (1998). Acute and chronic antiinflammatory effects of plant flavonoids. *Il Farmaco*, 53(6), 421-424.
- Piechota-Polanczyk, A., & Fichna, J. (2014). Review article: The role of oxidative stress in pathogenesis and treatment of inflammatory bowel diseases. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 387(7), 605-620.
- Ping, S., Gennv, W., Dan, W., & Xuefei, S. (2008). Progress in rosmarinic acid biological activities and its sources. *Food and Fermentation Industries*, 12, 055.
- Polanski, J., Bak, A., Gieleciak, R., & Magdziarz, T. (2006). Modeling robust QSAR. *Journal of Chemical Information and Modeling*, 46(6), 2310-2318.
- Prakash, N., & Gareja, D. A. (2010). Cheminformatics. *Journal of Proteomics & Bioinformatics*, 3, 249-252.
- Qin, X. (2012). Etiology of inflammatory bowel disease: A unified hypothesis. *World J Gastroenterol*, 18(15), 1708-1722.
- Rapaka, R. S., & Hawks, R. L. (1993). *Medications development: Drug discovery, databases, and computer-aided drug design* US Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Drug Abuse.
- Rashidian, A., Muhammadnejad, A., Dehpour, A., Mehr, S. E., Akhavan, M. M., Shirkoohi, R., et al. (2016). Atorvastatin attenuates TNBS-induced rat colitis: The involvement of the TLR4/NF- $\kappa$ B signaling pathway. *Inflammopharmacology*, doi: 10.1007/s10787-016-0263-6.

- Reich, K. (2012). The concept of psoriasis as a systemic inflammation: Implications for disease management. *Journal of the European Academy of Dermatology and Venereology*, 26(s2), 3-11.
- Roda, G., Marocchi, M., Sartini, A., & Roda, E. (2011). Cytokine networks in ulcerative colitis. *Ulcers*, 2011. doi:10.1155/2011/391787.
- Saijo, F., Milsom, A. B., Bryan, N. S., Bauer, S. M., Vowinkel, T., Ivanovic, M., et al. (2010). On the dynamics of nitrite, nitrate and other biomarkers of nitric oxide production in inflammatory bowel disease. *Nitric Oxide*, 22(2), 155-167.
- Sanz, Y., & Moya-Pérez, A. (2014). Microbiota, inflammation and obesity. *Microbial endocrinology: The microbiota-gut-brain axis in health and disease* (pp. 291-317) Springer.
- Song, C. M., Lim, S. J., & Tong, J. C. (2009). Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, 10(5), 579-591.
- Stecher, B. (2015). The roles of inflammation, nutrient availability and the commensal microbiota in enteric pathogen infection. *Microbiology Spectrum*, 3(3) doi: 10.1128/microbiolspec.MBP-0008-2014.
- Stolfi, C., De Simone, V., Pallone, F., & Monteleone, G. (2013). Mechanisms of action of non-steroidal anti-inflammatory drugs (NSAIDs) and mesalazine in the chemoprevention of colorectal cancer. *International Journal of Molecular Sciences*, 14(9), 17972-17985.
- Sugimoto, K. (2008). Role of STAT3 in inflammatory bowel disease. *World J Gastroenterol*, 14(33), 5110-5114.
- Sylvester, J. J. (1878). On an application of the new atomic theory to the graphical representation of the invariants and covariants of binary quantics, with three appendices. *American Journal of Mathematics*, 1(1), 64-104.
- Tarpey, M. M., Wink, D. A., & Grisham, M. B. (2004). Methods for detection of reactive metabolites of oxygen and nitrogen: In vitro and in vivo considerations. *American Journal of Physiology.Regulatory, Integrative and Comparative Physiology*, 286(3), R431-44.
- Tarpey, M. M., Wink, D. A., & Grisham, M. B. (2004). Methods for detection of reactive metabolites of oxygen and nitrogen: In vitro and in vivo considerations. *American Journal of Physiology.Regulatory, Integrative and Comparative Physiology*, 286(3), R431-44.
- Todeschini, R., & Consonni, V. (2003). DRAGON software (version 1.11-2001). *Milano, Italy*,
- Todeschini, R., & Consonni, V. (2009). *Molecular descriptors for chemoinformatics, volume 41 (2 volume set)* John Wiley & Sons.
- Tranmer, M., & Elliot, M. (2008). Multiple linear regression. *The Cathie Marsh Centre for Census and Survey Research (CCSR)*.
- Ukil, A., Maity, S., & Das, P. K. (2006). Protection from experimental colitis by theaflavin-3, 3'-digallate correlates with inhibition of IKK and NF- $\kappa$ B activation. *British Journal of Pharmacology*, 149(1), 121-131.

- Venkataranganna, M., Rafiq, M., Gopumadhavan, S., Peer, G., Babu, U., & Mitra, S. (2007). NCB-02 (standardized curcumin preparation) protects dinitrochlorobenzene-induced colitis through down-regulation of NFkappa-B and iNOS. *World Journal of Gastroenterology*, *13*(7), 1103.
- Verma, J., Khedkar, V. M., & Coutinho, E. C. (2010). 3D-QSAR in drug design-a review. *Current Topics in Medicinal Chemistry*, *10*(1), 95-115.
- Waldner, M. J., & Neurath, M. F. (2009). Novel cytokine-targeted therapies and intestinal inflammation. *Current Opinion in Pharmacology*, *9*(6), 702-707.
- Wang, L., Walia, B., Evans, J., Gewirtz, A. T., Merlin, D., & Sitaraman, S. V. (2003). IL-6 induces NF-kappa B activation in the intestinal epithelia. *Journal of Immunology (Baltimore, Md.: 1950)*, *171*(6), 3194-3201.
- Williams, I., & Goh, J. (2011). Investigational new drugs in the treatment of inflammatory bowel disease: A review. *J Exp Pharmacol*, *3*, 13-19.
- Xu, J., & Hagler, A. (2002). Chemoinformatics and drug discovery. *Molecules*, *7*(8), 566-600.
- Yang, X., Chang, Y., & Wei, W. (2016). Endothelial dysfunction and inflammation: Immunity in rheumatoid arthritis. *Mediators of Inflammation*, 2016. doi:10.1155/2016/6813016.
- Yoon, I. S., Au, Q., Barber, J. R., Ng, S. C., & Zhang, B. (2010). Development of a high-throughput screening assay for cytoprotective agents in rotenone-induced cell death. *Analytical Biochemistry*, *407*(2), 205-210.
- Zanni, R., Galvez-Llompart, M., García-Domenech, R., & Galvez, J. (2015). Latest advances in molecular topology applications for drug discovery. *Expert Opinion on Drug Discovery*, *10*(9), 945-957.
- Zhang, Y., Reenstra, W. W., & Chidekel, A. (2001). Antibacterial activity of apical surface fluid from the human airway cell line calu-3: Pharmacologic alteration by corticosteroids and  $\beta$ 2-agonists. *American Journal of Respiratory Cell and Molecular Biology*, *25*(2), 196-202.
- Zhang, Y., Shaffer, A., Portanova, J., Seibert, K., & Isakson, P. C. (1997). Inhibition of cyclooxygenase-2 rapidly reverses inflammatory hyperalgesia and prostaglandin E2 production. *The Journal of Pharmacology and Experimental Therapeutics*, *283*(3), 1069-1075.
- Zundler, S., & Neurath, M. F. (2016). Integrating immunologic signaling networks: The JAK/STAT pathway in colitis and colitis-associated cancer. *Vaccines*, *4*(1), 5. doi: 10.3390/vaccines4010005.

#### **Referencia electrónica:**

- Unidad de Atención Crohn-colitis (UACC). (Torrejón y Borrue). Todo lo que necesitas saber: Colitis ulcerosa. Recuperado el 20 de mayo, 2016, de <http://www.ua-cc.org/es/crohn-colitis/colitis-ulcerosa>



# ANEXOS



## IX. ANEXOS

**Anexo I.** Valor de los índices para los compuestos seleccionados aplicando la FD<sub>1</sub>.

<b>Compuesto n°</b>	<b>nCL</b>	<b>MATS8e</b>	<b>PCR</b>
<b>1</b>	0	-0.158	1.430
<b>2</b>	0	-0.281	1.293
<b>3</b>	0	-1.278	1.212
<b>4</b>	0	0.133	1.381
<b>5</b>	0	-0.138	1.438
<b>6</b>	0	-0.093	1.451
<b>7</b>	0	-0.384	1.361
<b>8</b>	2	0.209	1.408
<b>9</b>	1	0	1.312
<b>10</b>	0	-0.038	1.360
<b>11</b>	0	0.163	1.693
<b>12</b>	2	-0.251	1.478
<b>13</b>	0	-0.236	1.567
<b>14</b>	0	0.067	1.521
<b>Indometacina</b>	1	-0.521	1.370

**Anexo II.** Compuestos empleados en el *training set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de inhibidores de la iNOS (Capítulo 2).

Compuesto	FD <sub>4</sub>	Clas.	P.A.
<b>Grupo Activo</b>			
[(1S 5S 6R 7R)-7-chloro-3-imino-5-methyl-2-azabicyclo[4	1.99	A	0.87
β-caryophyllene	0.76	A	0.67
2-iminopiperidine	1.86	A	0.86
3,5-dicaffeoylquinicacid	4.25	A	0.98
4,5-dicaffeoylquinicacid	4.07	A	0.98
5,7-Dihydroxy-8-methoxyflavone	-0.05	I	0.47
5,7-dimethoxyflavone	-0.61	I	0.34
6-methoxyflavone	0.91	A	0.7
Acacetin	0.41	A	0.59
Amygdalin	0.84	A	0.68
Angelicin	1.85	A	0.86
Apigenin	1.23	A	0.76
Apiin	1.63	A	0.83
Azelaicacid	3.13	A	0.96
Benzylisothiocyanate	3.05	A	0.95
β-escin	2.33	A	0.9
Biochanin A	0.39	A	0.58
Bisabolol	0.31	A	0.56
Brazilein	1.44	A	0.8
Brazilin	1.21	A	0.76
Brossochalcone	1.76	A	0.85
Camptothecin	0.49	A	0.61
Capsaicin	1.39	A	0.79
Carvacrol	1.71	A	0.84
Caryophyllene	0.76	A	0.67
Celastrol	-0.08	I	0.46
Cepharanthine	0.17	A	0.53
Chamazulene	1.47	A	0.8
Chicoricacid	4.23	A	0.98
Chlorogenicacid	1.8	A	0.85
Chrysophanol	0.58	A	0.63
Cineole	1.93	A	0.87
Citrulline	1.49	A	0.81
Coumarin	2.14	A	0.89
Crocetin	4.53	A	0.99
Crustecdysone	-0.68	I	0.32
Cryptotanshinone	0.37	A	0.58
Cyanidin	1.41	A	0.79
Daphnetin	1.64	A	0.83
DHA	10.78	A	1
Dihydroartemisinin	-0.95	I	0.27
Dihydrotanshinone i	0.69	A	0.65
Diosgenin	-0.33	I	0.4
Emodin	0.38	A	0.58
Enoxolone	-0.67	I	0.32
EPA	9.37	A	1
Ergosterol	-0.54	I	0.35
Esculetin	1.83	A	0.85
Euphol	-0.99	I	0.26

Farnesol	0.92	A	0.7
Ferulic acid	0.18	A	0.53
Formononetin	0.87	A	0.69
Gallic acid	1.22	A	0.76
Genistein	1.3	A	0.77
Genkwanin	0.51	A	0.61
Guaiazulene	1.23	A	0.76
Guanylhiazine	6.59	A	1
Hematein	1.39	A	0.79
Heraclenol	0.02	A	0.49
Herniarin	0.53	A	0.62
Hesperetin	-0.06	I	0.47
Hesperidin	1.62	A	0.83
Hinokitiol	1.26	A	0.77
Humulene(alpha)	2.06	A	0.88
Hydroxytyrosol	1.6	A	0.82
Hypericin	-0.01	I	0.48
Imperatorin	-0.32	I	0.41
Indole-3-carbinol	1.69	A	0.84
Irigenin	-1.78	I	0.14
Juglone	1.61	A	0.83
Kaempferol	1.29	A	0.77
Kojic acid	0.78	A	0.67
Kolaviron	1.35	A	0.78
L-abrine	-0.19	I	0.44
Linalool(+)	1.76	A	0.85
Linoleic acid	7.57	A	1
L-N6-(1-iminoethyl)lysine	1.51	A	0.81
L-nitro arginine methyl ester (L-NAME)	-1.79	I	0.14
L-NMMA	1.24	A	0.77
Luteolin	1.19	A	0.76
Luteolin glucoside	1.22	A	0.76
Lycopene	5.54	A	1
Lycorine	0.49	A	0.61
Madecassic acid	-1.59	I	0.16
Myricetin	0.96	A	0.71
Naringin	1.2	A	0.76
Negletein	0.38	A	0.58
Nobiletin	-4.16	I	0.01
Ochnaflavone	2.4	A	0.91
Oleanolic acid	-0.71	I	0.32
Oleic acid	7.31	A	1
Oleuropein	1.33	A	0.78
Ouabain	-0.17	I	0.44
Oxypeucedanin	0.68	A	0.65
Paclitaxel	1.58	A	0.82
Paeoniflorin	-0.47	I	0.37
Parthenolide	0.35	A	0.57
P-coumaric acid	2.34	A	0.91
Phellopterin	-1.57	I	0.16
Phenethylisothiocyanate	3.21	A	0.96
p-hydroxycinnamaldehyde	2.22	A	0.9
Pimpinellin	-1.88	I	0.13
Pinene	3.49	A	0.97

Piperine	1.83	A	0.85
Plumbagin	1.1	A	0.74
Pristimerin	-0.73	I	0.31
Proanthocyanidins	1.15	A	0.75
Psoralen	1.93	A	0.87
Pterostilbene	0.79	A	0.68
Purpurin	0.97	A	0.71
Pyridoxine	-0.86	I	0.29
Quercetin	1.29	A	0.77
Rosmarinic acid	3.11	A	0.95
Rutin	1.96	A	0.87
Safranal	0.33	A	0.57
Shikimic acid	0.05	A	0.5
Shisoflavanonea	-0.86	I	0.28
S-isopropylisothiourea	-1.17	I	0.23
Sphondin	-0.09	I	0.46
Squalene	2.9	A	0.94
Tectorigenin	0.08	A	0.51
Thymol	0.43	A	0.59
Trigonellineok	0.42	A	0.59
Tryptanthrin	1.17	A	0.75
Ursolic acid	-0.93	I	0.27
Veratic acid	-2.08	I	0.11
Grupo Inactivo			
1(2)alpha-epoxydeoxydihydrogedunin	-1.7	I	0.15
1,2alpha-epoxy-7-deacetoxy-7-oxodihydrogedunin	-0.85	I	0.29
1,3-dideacetyldeoxykhivorin	-1.75	I	0.14
1,4,5,8-tetrahydroxy-2,6-dimethylanthroquinone	-0.1	I	0.46
1,7-dideacetoxy-1,7-dioxkhivorin	-1.77	I	0.14
10-hydroxycamptothecin	0.43	A	0.59
11-alpha-acetoxykhivorin	-3.43	I	0.03
11-oxoursolic acid acetate	-1.98	I	0.11
12a-hydroxy-5-deoxydehydromunduserone	-1.13	I	0.23
12a-hydroxy-9-demethylmunduserone-8-carboxylic acid	-1.26	I	0.21
1-methylxanthine	-0.67	I	0.33
1r,9s-hydrastine	-1.81	I	0.13
2,3-dihydroisogedunin	-1.63	I	0.16
2,3-dihydroxy-4-methoxy-4-ethoxybenzophenone	0.02	A	0.49
2,3-methano-7,2-dimethoxyflavanone	-0.27	I	0.42
2,4-dihydroxy-3,4,6-trimethoxychalcone	-0.75	I	0.31
2,5-dihydroxy-3,4-dimethoxy-4-ethoxybenzophenone	-1.36	I	0.2
2-acetylpyrrole	1.26	A	0.77
2-hydroxy-3,4-dimethoxybenzoic acid	-2.82	I	0.05
2-methoxyformonetin	-0.37	I	0.39
2-methoxyresorcinol	-0.18	I	0.44
2-methyl gramine	-0.52	I	0.36
2-methylene-5-(2,5-dioxotetrahydrofuran-3-yl)-6-oxo--	-0.56	I	0.35
10,10-dimethylbicyclo[7: 2: 0]undecane	-0.56	I	0.35
2-propyl-3-hydroxyethylenepyran-4-one	-1.31	I	0.2
3,4,5,6,7-pentamethoxyflavone	-3.22	I	0.04
3,4,5-trimethoxycinnamaldehyde	-3.39	I	0.03
3,4-dimethoxyflavone	-0.37	I	0.39
3,5-dihydroxyflavone	1.57	A	0.82
3-β-acetoxydeoxydihydrogedunin	-2.27	I	0.09

3- $\beta$ -acetoxydeoxyangolensic acid, methyl ester	-2.49	I	0.07
3-deacetylkhivorin	-2.39	I	0.08
3-deoxo-3beta-hydroxymexicanolide-16-enoether	-1.82	I	0.13
3-hydroxyflavone	1.81	A	0.85
3-methoxycatechol	0.6	A	0.63
3-methylorsellinic acid	-0.3	I	0.41
3-pinanone oxime	0.93	A	0.71
4,4-dimethoxydalbergione	-1.2	I	0.22
4-hydroxy-6-methylpyran-2-one	-0.27	I	0.42
4-methoxydalbergione	-0.4	I	0.39
4-o-methylphloracetophenone	-1.87	I	0.13
5,2-dimethoxyflavone	-0.64	I	0.33
5,7-dihydroxyisoflavone	1.34	A	0.78
5,7-dimethoxyisoflavone	-0.46	I	0.37
6,3-dimethoxyflavone	-0.02	I	0.48
6,4-dimethoxyflavone	0.05	A	0.5
6-acetoxyangolensic acid methyl ester	-2.61	I	0.06
6-hydroxyangolensic acid methyl ester	-2.02	I	0.11
7,2-dimethoxyflavone	-0.35	I	0.4
7-deacetylkhivorin	-2.48	I	0.07
7-desacetoxy-6,7-dehydrogedunin	-0.52	I	0.36
7-oxocallitrisic acid, methyl ester	-0.98	I	0.26
7-oxocholesterol	-0.94	I	0.27
7-oxocholesteryl acetate	-1.56	I	0.16
8,2-dimethoxyflavone	-0.5	I	0.36
8- $\beta$ -hydroxycarapin, 3,8-hemiacetal	-1.57	I	0.16
8-hydroxy-15,16-bisnor-11-labden-13-one	-1.06	I	0.25
8-hydroxycarapinic acid	-1.1	I	0.24
8-iodocatechin tetramethyl ether	-0.85	I	0.29
Abietic acid	-0.34	I	0.4
Abrine (I)	-0.19	I	0.44
Abscisic acid (cis,trans; +/-)	-0.5	I	0.36
Acacetin diacetate	-1.18	I	0.22
Acetosyringone	-2.27	I	0.09
Aconitic acid	-1.69	I	0.15
Ajmaline	-0.63	I	0.33
Albizziine	0.26	A	0.55
$\alpha$ -mangostin	-2.01	I	0.11
$\alpha$ -toxicarol	-1.49	I	0.17
Ambelline	-1.27	I	0.21
Andirobin	-1.56	I	0.16
Andrographolide	-0.64	I	0.33
Anethole	-0.1	I	0.46
Angolensic acid, methyl ester	-1.92	I	0.12
Angolensin (r)	0.62	A	0.64
Anhydrobrazilic acid	-0.6	I	0.34
Anthothecol	-1.61	I	0.16
Antiarol	-4.16	I	0.01
Aphyllic acid	-0.9	I	0.28
Apiole	-3.45	I	0.03
Artenimol	-0.95	I	0.27
Arzanol	-1.43	I	0.18
Asarinin (-)	0.45	A	0.6
Asarylaldehyde	-3.88	I	0.02

Asiatic acid	-1.3	I	0.2
Atranorin	-1.39	I	0.19
Austricine	-1.12	I	0.23
Azadirachtin	-2.65	I	0.06
Baccatin III	-1.75	I	0.14
Baeomycesic acid	-1.65	I	0.15
Bergapten	-0.08	I	0.46
Bergenin	-1.76	I	0.14
$\beta$ -Amyrin	-0.35	I	0.4
$\beta$ -Amyrin acetate	-1.21	I	0.22
$\beta$ -Mangostin	-1.58	I	0.16
$\beta$ -Peltatin	-1.77	I	0.14
$\beta$ -Toxicarol	-1.31	I	0.2
Betulinic acid	-1.31	I	0.2
Bicuculline (+)	-0.25	I	0.42
Biochanin a, 7-methyl ether	-0.32	I	0.41
Biochanin a, dimethyl ether	-1.27	I	0.21
Bisabolol acetate	-1.67	I	0.15
Boldine	-1.31	I	0.2
Bovinocidin	-1.11	I	0.24
Bussein	-3.71	I	0.02
Byssochlamic acid	-1.14	I	0.23
Cadinene	-0.21	I	0.43
Cafestol acetate	-0.07	I	0.47
Carapin	-1.32	I	0.2
Carminic acid	0.23	A	0.54
Carnosine	-0.8	I	0.3
Catechin pentaacetate	-1.87	I	0.13
Cedrelone	-0.47	I	0.37
Cellobiose (d[+])	-1.88	I	0.13
Centaurein	-0.68	I	0.32
Cephalotaxine	-0.37	I	0.39
Chelidonine (+)	-0.11	I	0.46
cholestan-3 $\beta$ ,5 $\alpha$ ,6 $\beta$ -triol	-0.81	I	0.3
Cholesteryl acetate	-1.42	I	0.18
Chrysanthemic acid, ethyl ester	-2.43	I	0.08
Chukrasin methyl ether	-5.2	I	0.01
Citropten	-1.49	I	0.18
Colchicine	-2.95	I	0.05
Cortisone	-0.05	I	0.47
Crassin acetate	-1.11	I	0.24
Cytisine	0.71	A	0.66
D,l-threo-3-hydroxyaspartic acid	-0.55	I	0.35
Dalbergione, 4-methoxy-4-hydroxy-	-0.36	I	0.4
Deacetoxy-7-oxisogedunin	-0.85	I	0.29
Deacetylgedunin	-0.81	I	0.29
Deguelin(-)	-1.37	I	0.19
Dehydro (11,12)ursolic acid lactone	-0.78	I	0.3
Demethylnobiletin	-3.32	I	0.03
Deoxyandirobin	-1.38	I	0.19
Deoxygedunol acetate	-2.34	I	0.08
Deoxykhivorin	-2.74	I	0.06
Derrustone	-0.72	I	0.31
Diffraictaic acid	-2.61	I	0.06



Dihydrocelastryl diacetate	-1.54	I	0.17
Dihydrogambogic acid	-2.53	I	0.07
Dihydrojasmonic acid, methyl ester	-1.59	I	0.16
Dihydromyristicin	-1.62	I	0.16
Dimethylcaffeic acid	-1.33	I	0.2
Diosmetin	0.13	A	0.52
Diprotin A	-2.44	I	0.08
Duartin (-)	-1.57	I	0.16
Echinocystic acid	-1.24	I	0.21
Entandrophragmin	-2.83	I	0.05
Epiatzelechin (2r)	1	A	0.72
Epoxygedunin	-1.65	I	0.15
Esculin	0.07	A	0.5
Eudesmic acid	-3.94	I	0.02
Eugenitol	0.27	A	0.55
Eupatorin	-1.48	I	0.18
Euphol acetate	-1.8	I	0.13
Euphorbiasteroid	-1.31	I	0.2
Evernic acid	-1.75	I	0.14
Evoxine	-1.85	I	0.13
Fraxetin	-0.23	I	0.43
Frequentin	-0.43	I	0.38
Friedelin	-0.35	I	0.4
Fumarprotocetraric acid	-0.11	I	0.46
Galacturonic acid	-0.74	I	0.31
Gambogic acid	-1.53	I	0.17
Gangaleoidin	-0.7	I	0.32
Gangleoidin acetate	-1.39	I	0.19
Gedunol	-1.56	I	0.16
Gentisic acid	0.85	A	0.69
Glucitol-4-gucopyanoside	-1.27	I	0.21
Gossypol	-1.7	I	0.15
Griseofulvic acid	-1.35	I	0.2
Haematommic acid	-1.45	I	0.18
Hecogenin	-0.44	I	0.38
Hederagenin	-0.95	I	0.27
Heteropeucenin, methyl ether	-1.92	I	0.12
Hieracin	1.15	A	0.75
Homopterocarpin	-0.63	I	0.33
Homova	-0.78	I	0.3
Hydroquinidine	-0.79	I	0.3
Hymecromone methyl ether	-0.28	I	0.42
Hypoxanthine	0.46	A	0.6
Inosine	-1.79	I	0.14
Iretol	-1.11	I	0.24
Iridin	-0.67	I	0.32
Isoosajin	-0.18	I	0.44
Isopeonol	-1.3	I	0.2
Isosafrole	0.42	A	0.59
Isotectorigenin, 7-methyl ether	-1.55	I	0.17
Juarezic acid	2.89	A	0.94
Karanjin	1.55	A	0.82
Khayanthone	-2.73	I	0.06
Khayasin c	-1.9	I	0.12

Khellin	-2.43	I	0.08
Kinetin	0.39	A	0.58
Koparin	0.4	A	0.58
Koparin 2-methyl ether	-0.6	I	0.34
Lanosterol	-0.99	I	0.26
Lapachol	0.45	A	0.6
Larixol	-1.46	I	0.18
Larixol acetate	-2.63	I	0.06
Lawsone	1.82	A	0.85
Leoidin	-0.27	I	0.42
Leoidin dimethyl ether	-2.18	I	0.1
Leucodin	-0.47	I	0.37
Limonin	-1.26	I	0.21
Linaly	-2.08	I	0.11
Linamarin	-2.03	I	0.11
Lobaric acid	-0.52	I	0.36
Loganic acid	-1.8	I	0.13
Lomatin	0.7	A	0.66
Lonchocarpic acid	-1.24	I	0.21
Lupinine	-0.97	I	0.27
Mandelic acid, methyl ester	0.61	A	0.64
Medicarpin	0.28	A	0.55
Melezitose	-0.58	I	0.34
Menthyl benzoate	-0.74	I	0.31
Metameconine	-1.6	I	0.16
Methyl coclaurine	0.06	A	0.5
Methyl orsellinate	-1.5	I	0.17
Methyl robustone	-0.53	I	0.36
Methylnorlichexanthone	-0.56	I	0.35
Methylxanthoxilin	-3.12	I	0.04
Mundulone	-0.58	I	0.34
Mundulone acetate	-1.47	I	0.18
N-methylanthranilic acid	-0.62	I	0.34
N-methylisoleucine	-4.63	I	0.01
Nomilin	-2.03	I	0.11
Nonic acid	-1.16	I	0.23
Nordihydroguaretic acid	1.96	A	0.87
Norstictic acid	-1.09	I	0.24
Odoratone	-0.95	I	0.27
Oleanolic acid acetate	-1.58	I	0.16
Orsellinic acid	-0.73	I	0.31
Orsellinic acid dimethyl ether	-3.14	I	0.04
Orsellinic acid, ethyl ester	-1.78	I	0.14
Osthol	-0.88	I	0.28
O-veratraldehyde	-2.77	I	0.06
Oxonitine	-3.51	I	0.03
Pachyrrhizin	0.29	A	0.56
Paeonol	-0.88	I	0.28
Peoniflorin	-0.32	I	0.4
Persitol heptaacetate	-3.21	I	0.04
Peucedanin	-1.04	I	0.25
Phloracetophenone	-0.73	I	0.31
Physcion	-0.47	I	0.37
Picroc	-1.78	I	0.14

Plectocomine methyl ether	0.83	A	0.68
Podofilox	-1.91	I	0.12
Podophyllotoxin acetate	-2.52	I	0.07
Pomiferin	-0.22	I	0.43
Prenyletin	0.08	A	0.51
Pristimerol	-0.72	I	0.31
Protoporphyrin ix	-0.2	I	0.43
Ptaeroxylin	0.23	A	0.54
Pterin-6-carboxylic acid	-0.61	I	0.34
Pteryxin	-2.21	I	0.09
Quassin	-2.72	I	0.06
Quebrachitol	-2.83	I	0.05
Quercetin pentamethyl ether	-3.15	I	0.04
Reserpine	-2.06	I	0.11
Retusin 7-methyl ether	-0.24	I	0.43
Retusoquinone	1.02	A	0.72
Rhamnetin	0.32	A	0.56
Rhamnose	-1.52	I	0.17
Rhodinyll acetate	-0.9	I	0.28
Robustic acid	-1.19	I	0.22
Rubescensin A	-0.83	I	0.29
Rutilantinone	-0.82	I	0.29
Safrole	0.36	A	0.57
Salsolidine	-1.64	I	0.15
Salsoline	-0.18	I	0.44
Salvinorin A	-2.15	I	0.1
Sappanone a trimethyl ether	-0.77	I	0.3
Sarmentoside B	-1.14	I	0.23
Scandenin diacetate	-2.14	I	0.1
Selinidin	-0.74	I	0.31
Senecrassidiol 6-acetate	-0.61	I	0.34
Sericetin	-0.32	I	0.41
Sinapic acid methyl ether	-2.97	I	0.05
Sinensetin	-3.27	I	0.03
Spathu	1.47	A	0.8
S-prop	3.28	A	0.96
Stictic acid	-1.89	I	0.12
Stigmasterol	-1.08	I	0.24
Strophanthidin	0.21	A	0.54
Strychnine	0.52	A	0.61
Syringic	-2.89	I	0.05
Tangeritin	56.76	A	0.07
Tetrahydropalmitine	-2.27	I	0.09
Theanine	0.11	A	0.51
Theobromine	-1.88	I	0.13
Totarol	-0.49	I	0.37
Totarol acetate	-1.63	I	0.16
Totarol-19-carboxylic acid, methyl ester	-1.44	I	0.18
Tridesacetoxylkivorin	-0.97	I	0.26
Tropine	0.29	A	0.56
Tubaic acid	-0.38	I	0.39
Ursinoic acid	-1.62	I	0.16
Ursolic acid	-0.93	I	0.27
Utilin	-2.93	I	0.05

Uvaol	-0.8	I	0.3
Vindoline	-3.4	I	0.03
Visnagin	-0.85	I	0.29
Vulpinic acid	0.3	A	0.56
Xanthoxylin	-3.14	I	0.04
Xanthyletin	1.15	A	0.75
Xylocarpus A	-3.04	I	0.04

---

FD: Función discriminante.

P.A.: probabilidad de que el compuesto sea activo.

Clas.: clasificación del modelo para el compuesto.

A: activo; I: inactivo.

**Anexo III.** Compuestos empleados en el *test set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de inhibidores de la iNOS (Capítulo 2).

Compuesto	FD <sub>4</sub>	Clas.	P.A.
<b>Grupo Activo</b>			
(S)-Methylisothiourea	10.08	A	1
2-hydroxy-3,4,5-trimethoxychalcone	-0.44	I	0.38
35-dichloro-2-hydroxy-3,4,5-trimethoxychalcone	0.63	A	0.64
5-hydroxyzerumbone	0.99	A	0.72
6-gingerol	1.22	A	0.76
Allicin	3.85	A	0.98
Anisodamine	-0.89	I	0.28
Astaxanthin	1.04	A	0.73
Caffeicacid	2.09	A	0.88
Carnosic acid	-1.17	I	0.23
Carnosol	-0.82	I	0.29
Chrysin	1.19	A	0.75
Citrinin	-1.21	I	0.22
Curcumina	1.96	A	0.87
Diallylsulfide	3.93	A	0.98
Ellagicacid	0.45	A	0.6
Eugenol	-0.64	I	0.33
Flavokawain B	0.26	A	0.55
Harpagoside	0.65	A	0.64
Isoliquiritigenin	2.76	A	0.94
Ligustilide	0.23	A	0.54
Limonene	1.12	A	0.74
N-[[3-(Aminomethyl)phenyl]methyl]-ethanimidamide	0.54	A	0.62
Resveratrol	2.6	A	0.93
Rutoside	1.96	A	0.87
Scopoletin	0.1	A	0.51
Silibinin	0.52	A	0.61
S,S-14-phenylene-bis(1-thymoquinone	5.45	A	1
Umbelliferone	0.2	A	0.53
Usnic acid	1.86	A	0.86
Xanthotoxin	-1.64	I	0.15
	-0.09	I	0.46
<b>Grupo Inactivo</b>			
Biochanin A diacetate	-1.05	I	0.25
Catechin tetramethylether	-2.31	I	0.09
Cholest-5-en-3-one	-0.72	I	0.31
Chondrosine	-1.84	I	0.13
Cinchonidine	0.11	A	0.51
Clovanediol diacetate	-2.33	I	0.08
Colchicine	-2.19	I	0.09
Colforsin	-2.92	I	0.05
Conessine	-0.72	I	0.31
Crinamine	-0.01	I	0.48
Daunorubicin	-0.27	I	0.42
Decahydrogambogic acid	-1.96	I	0.12
Deltaline	-2.82	I	0.05
Deoxygedunin	-1.51	I	0.17
Deoxysappanone b 7,3-dimethyl ether acetate	-0.88	I	0.28

Difucol hexamethyl ether	-4.59	I	0.01
Dihydrogedunic acid, methyl ester	-2.61	I	0.06
Dihydrosamidin	-2.56	I	0.07
Diphenylurea	2.03	A	0.88
Duartin dimethyl ether	-3.36	I	0.03
Epi(13)torulosol	-0.98	I	0.26
Ergosterol acetate	-1.9	I	0.12
Ethyl everninate	-2.7	I	0.06
Eugenyl benzoate	0.12	A	0.52
Eupatoriochromene	0.17	A	0.53
Evernic acid	-1.33	I	0.2
Fissinolide	-2.03	I	0.11
Fraxidin methyl ether	-3.18	I	0.04
Fucostanol	-1.26	I	0.21
Gardenin B	-2.53	I	0.07
Geraldol	0.48	A	0.6
Griseofulvin	-2.27	I	0.09
Hederacoside C	3.23	A	0.96
Hexamethylquercetaletin	2.16	A	0.89
Hydroxyprogesterone	0.52	A	0.61
Ichthyone	-1.26	I	0.21
Isopimpinellin	-1.88	I	0.13
Isotectorigenin trimethyl ether	-2.41	I	0.08
Kainic acid	-3.59	I	0.03
Khayasin	-1.85	I	0.13
Khivorin	-2.86	I	0.05
Kuhlmannin	-1.1	I	0.24
Lagochilin	-1.7	I	0.15
Lappaconitine	-2.79	I	0.05
Lecanoric acid	-0.65	I	0.33
Liquiritigenin dimethyl ether	-0.26	I	0.42
Loganin	-2.18	I	0.1
Marmesin acetate	-0.82	I	0.29
Melatonin	-0.54	I	0.35
Merogedunin	-0.53	I	0.36
Methyl everninate	-2.43	I	0.08
Niloticin	-0.73	I	0.31
Obliquin	-0.08	I	0.47
Osajin	-0.31	I	0.41
Pachyrrhizone	-0.25	I	0.42
Penicillic acid	-2.36	I	0.08
Peucenin	-0.87	I	0.28
Piplartine	-1.65	I	0.15
Podototarín	-1.69	I	0.15
Primuletin	1.58	A	0.82
Pseudo-anisatin	-0.29	I	0.41
Pyrromycin	-1.36	I	0.19
Rhetsinine	0.35	A	0.57
Robustone	0.43	A	0.59
Rotenone	-1.6	I	0.16
Salicin	-0.42	I	0.38
Salvinorin B	-1.4	I	0.19
Scandenin	-1.24	I	0.21
Solidagenone	-0.09	I	0.46

Sumaresinolic acid	-1.13	I	0.23
Totalolal	-0.79	I	0.3
Triacetylresveratrol	-0.16	I	0.44
Veratric acid	-1.98	I	0.11
Vincamine	-1.05	I	0.25
Violastyrene	0.55	A	0.62
Zeorin	1.69	A	0.83

---

**Anexo IV.** Compuestos empleados en el *training set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de inhibidores del TNF- $\alpha$  (Capítulo 2).

<b>Compuesto</b>	<b>FD<sub>5</sub></b>	<b>Clas.</b>	<b>P.A.</b>
<b>Grupo Activo</b>			
Allicin	-0.31	I	0.42
Andrographolide	-0.82	I	0.31
Apremilast	0.66	A	0.66
Aspirin	0.96	A	0.72
Betamethasone	-0.39	I	0.40
Bupropion	0.04	A	0.51
Coumarin	7.40	A	1.00
Cysteamine	8.61	A	1.00
Dexamethasone	-0.39	I	0.40
Genistein	1.79	A	0.86
Griseofulvin	1.65	A	0.84
Honokiol	1.68	A	0.84
Indomethacin	0.75	A	0.68
LA-SS-AZ	-1.75	I	0.15
LMP-420	5.50	A	1.00
LSD	-1.53	I	0.18
Nocodazole	0.81	A	0.69
OPC-6535	0.24	A	0.56
Oxpentyfilline	-2.01	I	0.12
PD98059	0.48	A	0.62
Phosphatidylcoline	7.54	A	1.00
R-doi	2.07	A	0.89
Rolipram	-1.07	I	0.26
Tamatinib	5.73	A	1.00
TCB-2	1.35	A	0.79
Thalidomide	0.22	A	0.55
Thioguanine	4.50	A	0.99
<b>Grupo Inactivo</b>			
Amitriptyline	-3.44	I	0.03
Amlodipine	0.11	A	0.53
Amodiaquine	-0.38	I	0.41
Anethole	-3.26	I	0.04
Aniracetam	-0.42	I	0.40
Apomorphine	1.17	A	0.76
Apramycin	-2.72	I	0.06
Armodafinil	-0.72	I	0.33
Atropineoxide	-1.34	I	0.21
Aztreonam	-1.74	I	0.15
Bendroflumethiazide	-2.73	I	0.06



Benserazide	-0.34	I	0.41
Benurestat	-0.71	I	0.33
Bezafibrate	-0.64	I	0.35
Bifonazole	-2.92	I	0.05
Biotin	-0.85	I	0.30
Bisacodyl	-0.65	I	0.34
Bisoprololfumarate	-2.38	I	0.08
Bromocriptinemesylate	-1.42	I	0.20
Bromperidol	-1.36	I	0.20
Candesartan	-3.23	I	0.04
Capsaicin	-1.96	I	0.12
Carbachol	-2.02	I	0.12
Cefdinir	-2.54	I	0.07
Cefprozil	-1.37	I	0.20
Cefsulodin	-1.21	I	0.23
Chlorotrianisene	-0.44	I	0.39
Chloroxine	2.21	A	0.90
Choline	0.43	A	0.60
Cilostazol	-3.39	I	0.03
Citalopram	-1.09	I	0.25
Clomiphenecitrate	-2.64	I	0.07
Clopidogrel	-0.44	I	0.39
Crotamiton	-0.96	I	0.28
Dantrolene	-0.10	I	0.47
Debrisoquinsulfate	-0.43	I	0.39
Demeclocycline	0.38	A	0.59
Denatoniumbenzoate	-3.22	I	0.04
Dexlansoprazole	-1.60	I	0.17
Dexpropranolol	-0.49	I	0.38
Dibucaine	-2.22	I	0.10
Dicumarol	0.62	A	0.65
Diethyltoluamide	-3.95	I	0.02
Dihydroergotamine	-3.62	I	0.03
Diloxanide	-2.16	I	0.10
Diperodon	-3.54	I	0.03
Dipyridamole	-2.86	I	0.05
Disulfiram	2.96	A	0.95
D-lactitol	-1.71	I	0.15
Dobutamine	-1.00	I	0.27
Domperidone	-2.14	I	0.10
Donepezil	-2.79	I	0.06
Dopamine	-1.48	I	0.19
Doxylamine	-2.49	I	0.08
Droperidol	-2.42	I	0.08

Dyphylline	-1.32	I	0.21
Eletriptan	-2.18	I	0.10
Enalapril	-2.26	I	0.09
Ergonovine	-0.56	I	0.36
Ethamivan	-2.30	I	0.09
Ethinylestradiol	0.71	A	0.67
Ethosuximide	-1.24	I	0.22
Ethoxzolamide	-2.07	I	0.11
Evansblue	-1.97	I	0.12
Ezetimibe	-1.45	I	0.19
Febuxostat	0.47	A	0.62
Felbamate	-0.60	I	0.35
Fexofenadine	-2.62	I	0.07
Fluvastatin	-0.29	I	0.43
Fluvoxamine	-2.06	I	0.11
Folicacid	-1.88	I	0.13
Gallaminetriethiodide	-6.31	I	0.00
Haloperidol	-1.76	I	0.15
Iodipamide	-1.02	I	0.26
Isoniazid	-2.06	I	0.11
Ketanserin	-2.37	I	0.09
Ketoconazole	-1.90	I	0.13
Ketotifen	-0.13	I	0.47
Labetalol	-1.51	I	0.18
Levobunolol	-0.88	I	0.29
Loperamide	-2.87	I	0.05
Meclizine	-4.97	I	0.01
Mefexamide	-2.69	I	0.06
Methacholine	-2.72	I	0.06
Methyleneblue	1.51	A	0.82
Methylphenidate	-1.38	I	0.20
Methylthiouracil	1.71	A	0.85
Montelukast	-2.31	I	0.09
Nalbuphine	-0.35	I	0.41
Naltrexone	-0.12	I	0.47
Neomycin	-2.96	I	0.05
Nicardipine	-2.22	I	0.10
Nicergoline	-0.11	I	0.47
Nifedipine	-0.23	I	0.44
Nilutamide	-1.56	I	0.17
Nimodipine	-0.94	I	0.28
Nithiamide	-6.09	I	0.00
Nitrofurantoin	0.10	A	0.53
Nonoxyno	-1.47	I	0.19

Norgestimate	-1.34	I	0.21
Orlistat	-4.20	I	0.01
Oseltamivir	-2.75	I	0.06
Oxibendazole	-0.61	I	0.35
Oxiconazolenitrate	-2.57	I	0.07
Oxymetazoline	-0.84	I	0.30
Pancuronium	-4.76	I	0.01
Pantoprazole	-0.43	I	0.39
Pararosaniline	0.68	A	0.66
Paromomycin	-2.58	I	0.07
Penteticacid	-2.64	I	0.07
Pentylene-tetrazol	-4.66	I	0.01
Perindopril-erbumine	-2.62	I	0.07
Phenoxybenzamine	-2.38	I	0.08
Phentermine	-1.26	I	0.22
Phenyl	0.85	A	0.70
Phenylephrine	-2.19	I	0.10
Physostigmine-salicylate	-1.50	I	0.18
Pilocarpine	-1.16	I	0.24
Piperidolate	-2.03	I	0.12
Piperine	-0.88	I	0.29
Pirenzepine	-1.97	I	0.12
Pramoxine	-2.91	I	0.05
Prilocaine	-1.23	I	0.23
Probenecid	-0.58	I	0.36
Prochlorperazine	-2.53	I	0.07
Propafenone	-1.44	I	0.19
Quinacrine	-1.24	I	0.22
Quinidine	-1.49	I	0.18
Ramipril	-2.70	I	0.06
Ranitidine	-3.74	I	0.02
Riboflavin	-0.16	I	0.46
Ritanserlin	-3.11	I	0.04
Rolipram	-1.07	I	0.26
Rolitetra-cycline	-0.83	I	0.30
Salicylanilide	1.26	A	0.78
Selegiline	-3.87	I	0.02
Sibutramine	-5.85	I	0.00
Sildenafil-citrate	0.17	A	0.54
Sodium phenylacetate	-0.55	I	0.37
Solifenacin-succinate	-2.71	I	0.06
Strychnine	-1.37	I	0.20
Sulfachlorpyridazine	-3.49	I	0.03
Sulfanitran	-1.64	I	0.16

Sulfinpyrazone	-2.11	I	0.11
Terazosin	-2.25	I	0.10
Tetracaine	-1.95	I	0.12
Thiamylal	-1.95	I	0.12
Tiapride	-0.58	I	0.36
Tilorone	-2.24	I	0.10
Tolazamide	-2.16	I	0.10
Toltrazuril	-1.51	I	0.18
Topiramate	3.43	A	0.97
Tramadol	-1.63	I	0.16
Tranexamicacid	-1.48	I	0.19
Trihexyphenidyl	-2.24	I	0.10
Trimetozine	-1.51	I	0.18
Tripelennaminecitrate	-3.29	I	0.04
Vardenafil	-0.12	I	0.47
Verapamil	-4.87	I	0.01
Yohimbine	-1.22	I	0.23
Zidovudine	-1.28	I	0.22
Zolmitriptan	-1.40	I	0.20

---

**Anexo V.** Compuestos empleados en el *test set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de inhibidores del TNF- $\alpha$  (Capítulo 2).

<b>Compuesto</b>	<b>FD<sub>5</sub></b>	<b>Clas.</b>	<b>P.A.</b>
<b>Grupo Activo</b>			
Daunorubicin	0.10	A	0.52
Diclofenac	0.98	A	0.73
Flurbiprofen	1.57	A	0.83
Meclofenamate	1.16	A	0.76
Mycophenolic	0.15	A	0.54
Naproxen	1.47	A	0.81
Piroxicam	0.78	A	0.69
Prednisone	0.19	A	0.55
Rofecoxib	0.80	A	0.69
Small molecule 35	-3.50	I	0.03
Sulindac	0.69	A	0.67
Suprofen	0.65	A	0.66
TNF molecule 38	-3.56	I	0.03
TNF-alpha 37	-3.90	I	0.02
Tnf-molecule i-2	-1.75	I	0.15
Tnf-molecule ii-12	-2.14	I	0.11
Tnf-molecule iii-21	0.23	A	0.56
Tnf-molecule-iv 22	-0.97	I	0.28
<b>Inactive group</b>			
Aminacrine	1.75	A	0.85
Biperiden	-1.85	I	0.14
Bretyliumtosylate	0.87	A	0.70
Butylparaben	0.11	A	0.53
Carbidopa	-3.78	I	0.02
Cephapirin	-1.30	I	0.21
Citicoline	7.75	A	1.00
Clidinium	-1.93	I	0.13
Clorsulon	-7.79	I	0.00
Cloxacillin	-1.76	I	0.15
Danazol	-1.16	I	0.24
Desoxycorticosterone	-0.15	I	0.46
Dexchlorpheniramine	-1.69	I	0.16
Dicloxacillin	-1.67	I	0.16
Dioxybenzone	1.20	A	0.77
Diphenylpyraline	-2.65	I	0.07
Doxycycline	0.32	A	0.58
Econazole	-1.72	I	0.15
Ethopropazine	-1.44	I	0.19
Gatifloxacin	-0.96	I	0.28

Glycopyrrolate	-2.27	I	0.09
Homosalate	-0.43	I	0.39
Irbesartan	-3.18	I	0.04
Levothyroxine	1.89	A	0.87
Lomefloxacin	-0.75	I	0.32
Medroxyprogesteroneacetate	-1.05	I	0.26
Monobenzone	1.14	A	0.76
Moxalactam	-1.49	I	0.18
Netilmicin	-3.40	I	0.03
Olmesartan	-1.92	I	0.13
Perhexiline	-3.92	I	0.02
Pipamperone	-3.13	I	0.04
Pirenperone	-2.39	I	0.08
Propantheline	-2.07	I	0.11
Simvastatin	-1.94	I	0.13

---

**Anexo VI.** Compuestos empleados en el *training set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de compuestos frente a la CU inducida por DSS (Capítulo 2).

Compuesto	FD <sub>6</sub>	P.A.	Clas.
<b>Grupo Activo</b>			
ABC294640	1.940	0.874	A
Anandamide	2.730	0.939	A
Baicalein	1.330	0.791	A
Balsalazide	2.988	0.952	A
Berberine	0.810	0.692	A
Bortezomib	0.208	0.552	A
Calcitriol	1.360	0.796	A
Chitosan	2.893	0.947	A
Ciprofloxacin	1.888	0.869	A
Clarithromycin	0.989	0.729	A
Corilagin	3.756	0.977	A
Cyclosporine	1.471	0.813	A
Delmitide	1.302	0.786	A
Dolcanatide	4.299	0.987	A
Doxycycline	1.167	0.763	A
Fexofenadine	-1.694	0.155	I
Galacturonic acid	0.166	0.542	A
Hydroxytyrosol	0.200	0.550	A
Ipsalazide	2.864	0.946	A
Lapatinib	0.585	0.642	A
lupeol	-0.537	0.369	I
lycopene	7.555	0.999	A
Melphalan	0.926	0.716	A
Mesalamine	0.210	0.552	A
Metformin	6.537	0.999	A
Methotrexate	3.235	0.962	A
Metronidazole	1.383	0.799	A
Minocycline	1.400	0.802	A
Nicotine	3.706	0.976	A
Nilotinib	0.734	0.676	A
Oleuropein	2.997	0.952	A
Olsalazine	2.993	0.952	A
Oxymatrine	-1.936	0.126	I
Pentamidine	2.504	0.924	A
Plecanatide	4.299	0.987	A
Quinestrol	-0.432	0.394	I
Renzapride	0.518	0.627	A
Roxithromycin	1.450	0.810	A
Sirolimus	2.556	0.928	A
Sivelestat	0.199	0.550	A
Squalene	6.555	0.999	A

Sulfasalazine	0.341	0.584	A
UR-1505	1.745	0.851	A
<b>Grupo Inactivo</b>			
Acebutolol	0.004	0.501	A
Aminohippuric acid	-4.313	0.013	I
Amiodarone	-1.596	0.169	I
Aspartame	-1.409	0.196	I
Azathioprine	-2.846	0.055	I
Benurestat	-3.120	0.042	I
Benzethonium	-0.142	0.465	I
Biotin	-5.327	0.005	I
Bromocriptinemesylate	-3.538	0.028	I
Budesonide	-1.889	0.131	I
Caffeine	-4.772	0.008	I
Candesartan	-2.232	0.097	I
Canrenoic acid	-0.645	0.344	I
Cefoperazone	-2.128	0.106	I
Cefuroxime	-0.502	0.377	I
Chlorophyllide	-3.950	0.019	I
Clavulanate	-3.514	0.029	I
Coumarin	-1.050	0.259	I
Dantrolene	-1.798	0.142	I
Deferoxamine	-1.734	0.150	I
Dehydrocholic acid	0.085	0.521	A
Denatoniumbenzoate	-4.133	0.016	I
Dexpanthenol	-0.456	0.388	I
Dihydroergotamine	-3.500	0.029	I
Diltiazem	0.224	0.556	A
Ergotamine	-3.411	0.032	I
Ethoxzolamide	-4.312	0.013	I
Famotidine	-4.520	0.011	I
Fenretinide	-0.399	0.401	I
Fluorouracil	-2.243	0.096	I
Glipizide	0.415	0.602	A
Indapamide	-0.616	0.351	I
Isoproterenol	-1.334	0.209	I
Isoxsuprine	0.604	0.647	A
Itraconazole	-2.191	0.101	I
Ketoconazole	-2.627	0.067	I
Levamisole	-5.299	0.005	I
Methoxamine	-0.750	0.321	I
Methylthiouracil	-2.737	0.061	I
Miglitol	-2.199	0.100	I
Monensin	-3.106	0.043	I
Nalbuphine	-0.652	0.343	I
Nicardipine	0.513	0.626	A



Noscapine	-0.645	0.344	I
Orlistat	-2.782	0.058	I
Oxethazaine	-1.144	0.242	I
Oxytetracycline	0.196	0.549	A
Pancuronium	-2.533	0.074	I
Phenacetin	0.569	0.639	A
Phytonadione	-2.910	0.052	I
Piracetam	-4.213	0.015	I
Pirenperone	0.101	0.525	A
Proadifen	-2.966	0.049	I
Probucol	-0.835	0.302	I
Propafenone	0.350	0.587	A
Propantheline	-2.957	0.049	I
Proscillaridin	0.445	0.610	A
Ranolazine	0.478	0.617	A
Reserpine	1.129	0.756	A
Ribavirin	-4.221	0.014	I
Risedronate	0.886	0.708	A
Semustine	-3.554	0.028	I
Sisomicin	0.396	0.598	A
Strychnine	-0.985	0.272	I
Sulfachlorpyridazine	-2.831	0.056	I
Sulfamethoxazole	-5.581	0.004	I
Sulfamonomethoxine	-0.083	0.479	I
Terbutalinehemisulfate	-0.799	0.310	I
Terfenadine	-2.353	0.087	I
Theophylline	-2.576	0.071	I
Thiothixene	-0.466	0.385	I
Topiramate	-6.396	0.002	I
Tranexamic acid	-5.338	0.005	I
Triacetin	-1.637	0.163	I
Trisodiummethylenediaminetetra	-0.835	0.303	I
Tubocurarine	2.786	0.942	A
Valsartan	-3.413	0.032	I
Vincristine	-1.600	0.168	I

---

**Anexo VII.** Compuestos empleados en el *test set* y resultados obtenidos al aplicar el análisis lineal discriminante a la búsqueda de compuestos frente a la CU inducida por DSS (Capítulo 2).

<b>Compuesto</b>	<b>FD<sub>6</sub></b>	<b>Clas.</b>	<b>P.A.</b>
<b>Grupo Activo</b>			
Azathioprine	-2.846	I	0.055
Colchicine	1.488	A	0.816
Cromolyn	1.451	A	0.810
Cyclosporine	1.471	A	0.813
Dexamethasone	-1.623	I	0.165
Gallic acid	1.231	A	0.774
Glucosamine	2.204	A	0.901
Levofloxacin	1.131	A	0.756
Loratadine	2.402	A	0.917
Mercaptopurine	2.205	A	0.901
Myricetin	2.146	A	0.895
Novobiocin	2.093	A	0.890
Rofecoxib	-0.806	I	0.309
Rosmarinic	1.030	A	0.737
Sophocarpine	0.428	A	0.605
Tacrolimus	3.110	A	0.957
Tetomilast	1.155	A	0.760
Tixocortol	-2.235	I	0.097
<b>Inactive group</b>			
Acetarsol	-3.466	I	0.030
Capsaicin	-0.433	I	0.393
Cefoxitin	-1.437	I	0.192
Doxazosin	1.668	A	0.841
Ethambutol	-0.752	I	0.320
Fexofenadine	-1.694	I	0.155
Gallaminetriethiodide	-4.791	I	0.008
Hydroxyzinepamoate	4.054	A	0.983
Levodopa	-3.107	I	0.043
Methscopolamine	-4.265	I	0.014
Molsidomine	-2.044	I	0.115
Montelukast	-0.894	I	0.290
Nitrofurantoin	-3.713	I	0.024
Nitromide	-1.097	I	0.250
Oxolinic acid	-0.804	I	0.309
Tamoxifencitrate	1.330	A	0.791
Triamcinolone	-1.355	I	0.205
Vardenafil	-1.927	I	0.127

## Anexo VIII. Autorización experimentación animal



DIRECCIÓN GENERAL DE PRODUCCIÓN AGRARIA Y GANADERÍA

Ciutat Administrativa 9 d'Octubre  
Castán Tobeñas, 77. Edif. B3 P2 46018 VALENCIA



### AUTORIZACION PROCEDIMIENTO 2015/VSC/PEA/00023

Vista la solicitud realizada en fecha **23/01/15** con nº reg. entrada **1147** por D/D<sup>a</sup>. **Pilar Campins Falcó**, Vicerrectora Investigació i Política Científica, centro usuario **ES460780001001**, para realizar el procedimiento:

***"Efecto preventivo de la transformación neoplásica asociada a colitis ulcerosa mediado por productos fenólicos antiinflamatorios"***

Teniendo en cuenta la documentación aportada, según se indica en el artículo 33, punto 5 y 6, y puesto que dicho procedimiento se halla sujeto a autorización en virtud de lo dispuesto en el artículo 31 del Real Decreto 53/2013, de 1 de febrero,

Vista la propuesta del jefe del servicio de Sanidad y Bienestar Animal.

#### AUTORIZO:

la realización de dicho procedimiento al que se le asigna el código: **2015/VSC/PEA/00023** tipo **2**, de acuerdo con las características descritas en la propia documentación para el número de animales, especie y periodo de tiempo solicitado. Todo ello sin menoscabo de las autorizaciones pertinentes, por otras Administraciones y entidades, y llevándose a cabo en las siguientes condiciones:

Usuario: **Universidad de Valencia-Estudio General**

Responsable del proyecto: **Maria del Carmen Recio Iglesias**

Establecimiento: **Sección Producción Animal SCSIE Burjasot**

Necesidad de evaluación retrospectiva:

Condiciones específicas:

Observaciones:

Valencia a, 11 de febrero de 2015

El director general de Producción Agraria y Ganadería

José Miguel Ferrer Arranz



# SUMMARY



## X. SUMMARY

The goal of the present Thesis is the devising of a new strategy to design new *lead* compounds with anti-inflammatory activity. Different *in silico*, *in vitro* and *in vivo* approaches were used. The most innovative aspect, which can be considered as the “trump card”, is the application of Molecular Topology (MT). MT is essentially a method that enable to describe every chemical structure in a pure mathematical way.

Numerous topological models have been developed, along with specific *in vitro* and *in vivo* assays. The most relevant results were published in journals of considerable impact factor and can be considered as the pillars of the present Thesis.

**Article 1** describes the devise of a predictive equation for the identification of new potential NSAIDs as COX inhibitors. The model was used for the virtual screening of the Sigma-Aldrich database, allowing the selection of 14 potential anti-inflammatory compounds. Four of them (compounds **5**, **8**, **9** and **14**) showed a significant anti-inflammatory activity *in vivo* (carrageenan-induced paw edema assay). Subsequently, all the molecules were tested *in vitro*, except the insoluble in water (compounds 6, 10 y 12) and the compound 4, which was been already described as anti-inflammatory. At the concentration of 100  $\mu$ M, none of the products was cytotoxic. The compounds **8** and **9**, which were already active *in vivo*, showed *in vitro* activity by inhibiting the production of prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) and nitrites.

On the basis of the previous results, brand new mathematical models were developed for the inhibition of other pro-inflammatory agents, such as cytokines, transcription factors, enzymes, etc., involved in various inflammatory processes and in particular in ulcerative colitis (UC).

**Article 2** focuses on the prediction of the protective activity against UC by inhibition of NF- $\kappa$ B. In this case, two discriminant functions were developed. The first function focused on the discrimination between inhibitors and no-inhibitors of NF- $\kappa$ B, while the second one distinguished between active and inactive compounds against UC. After performing an external validation of the equations, a virtual screening of the Microsource Pure Collection Natural Products (800 natural products) and Merck index (approximately 12000 compounds) databases were carried out. A group of 28 compounds was finally selected to be tested *in vitro* and *in vivo*.

Moreover, interleukin (IL)-6 is a cytokine widely present in tissues and plasma of patients with chronic inflammation such as arthritis, cancer and IBD. These facts makes IL-6 an attractive target for the treatment of such diseases. Thereupon, in **Article 3**, a QSAR model to predict the IL-6 inhibitory activity has been devised. The model, which resulted statistically

significant and showed good predictive capability, was used to screen the 28 compounds previously selected as potential anti-UC agents (**Article 2**) to determine IL-6 inhibitory activity. Finally, 4 compounds show anti IL-6 activity: Alizarin-3-methylinodiadic acid (**AMA**), (+)-dibenzyl-L-tartrate (**DLT**), **Calcein** and **Ro 41-0960**. The results were experimentally corroborated, since 3 of the molecules inhibited the production of IL-6 in two cell lines (Caco-2 and RAW 264.7): **Calcein** a natural compound, showed an IC<sub>50</sub> value of 45 μM in Caco-2 cell line; **DLT** showed an IC<sub>50</sub> value of 19 μM in both cell lines and **Ro 41-0960** showed an IC<sub>50</sub> value of 12 μM in Caco-2 and 0,4 μM in RAW 264.7.

The crucial role of tumor necrosis factor (TNF)-α and inducible nitric oxide synthase (iNOS) in the development and chronicity of the UC, was a good enough reason to decide to calculate new topological models for the prediction of activity again (**Article 4**). The last step, **Article 4**, consists in the development of 4 topological models to predict the inhibitory activity against iNOS, TNF-α and the protective activity against DSS-induced UC. LDA and multi-linear regression analysis (MLRA) were used to devise the models. After testing 4 compounds *in vitro* and *in vivo*: AMA, Calcein, DLT and Ro 41-0960; **AMA** and **Ro 41-0960** were selected as the best candidates to become potential *lead* compounds. **AMA** reduced NO levels and TNF-α production *in vitro* and protected the animals from UC effects in a DSS-induced acute model by reducing the body weight loss, and preventing the colon shortening (DAI score). **Ro 41-0960** resulted to be the best one. It was able to inhibit NO, IL-1β and TNF-α production *in vitro* and, in addition, it was capable to protect the animals against the DSS-induced colitis effects *in vivo* by reducing the body weight loss and preventing the colon shortening (DAI score).

The overall results of the four articles supports the efficacy of a virtual screening strategy based on MT, which focused on the identification of new brand inhibitory molecules of several pro-inflammatory mediators such as COX, iNOS, IL-6, IL 1β, TNF-α and NF-κB could represent future therapeutic treatments for UC. Deepening the pharmacological characteristics of the selected compounds is out of the reach of the present Doctoral Thesis, although different preliminary tests *in vitro* and *in vivo* confirmed the predictive results of the topological models, demonstrating MT's validity as an effective tool for the design and discovery of new *lead* compounds.



# CONCLUSIONS



## XI. CONCLUSIONS

1. Several mathematical models based on molecular topology have been developed. These models focus on the prediction of the pharmacological activity related to the inhibition of pro-inflammatory enzymes such as COX and iNOS, cytokines (IL-6 and TNF- $\alpha$ ), and transcription factors like NF- $\kappa$ B. In addition, various predictive equations have been calculated, to predict the activity of a particular type of chronic inflammation, the ulcerative colitis. The virtual screening of different databases (Sigma-Aldrich, Microsource Pure Collection and Merck index), which included more than 100000 compounds led to the selection of 42 molecules as potential anti-inflammatory agents. Of these, 14 were predicted as anti-inflammatory (COX inhibitors), and 28 as potential treatments for ulcerative colitis (NF- $\kappa$ B and IL-6 inhibitors).
2. In accordance with the criteria of **Articles 1** and **3**, some of the compounds selected *in silico* have been tested both *in vitro* and *in vivo*, giving results which corroborate the prediction of the topological models.
3. Compounds **5**, **8**, **9** and **14**, widely described in **Article 1**, showed anti-inflammatory activity *in vivo* as COX inhibitors in a carrageenan-induced paw edema assay. Compound **8** stands out, which at the dose of 10 mg/kg at 3 h, inhibited the edema by 47%, i.e. at similar level of the indomethacin. The same compound at a concentration of 50  $\mu$ M inhibited *in vitro* the production of nitrite and PGE<sub>2</sub>. Analyzing the descriptors of the predictive equation, it could be stated that the presence of electronegative heteroatoms, halogens and unsaturation seems to be relevant for the anti-inflammatory activity.
4. Moreover, new active compounds against ulcerative colitis by inhibiting NF- $\kappa$ B have been identified. A group of 28 compounds was finally selected. Thirteen of them have been previously described in literature as anti-inflammatory, antioxidant and anti-ulcerative agents. This fact can be considered as a further validation of the predictive model. In terms of structure-activity relationship, according with the descriptors of the equations, structural and electronic properties determine the inhibitory activity of NF- $\kappa$ B along with the anti UC activity.

5. The compounds selected in **Article 2** also inhibit IL-6 production. This fact supports their potential activity against ulcerative colitis. This prediction, whose experimental results are described in **Article 3**, allowed the selection of 4 compounds: Alizarin 3-methylinodiacetic acid (**AMA**), (+)-dibenzyl-L-tartrate (**DLT**), **Calcein** and **Ro 41-0960**. After *in vitro* and *in vivo* tests, 3 of the molecules inhibited the production of IL-6 in two cell lines (Caco-2 and RAW 264.7): **Calcein** a natural compound, showed an IC<sub>50</sub> value of 45 μM in Caco-2 cell line; **DLT** showed an IC<sub>50</sub> value of 19 μM in both cell lines and **Ro 41-0960** showed an IC<sub>50</sub> value of 12 μM in Caco-2 and 0,4 μM in RAW 264.7. Finally, the indices of the predictive equation explain how the inhibitory activity against IL-6 can be related to the degree of conjugation, the van der Waals volume and the atomic electronegativity of the molecules.
6. Based on the results obtained and described in **Article 4**, it can be state that the best candidates as new *lead* compounds for the treatment of ulcerative colitis are **AMA** and **Ro 41-0960**. Both products inhibited the production of pro-inflammatory mediators *in vitro*, moreover, **Ro 41-0960** *in vivo*, at the dose of 50 mg/kg decreased all the signs of the ulcerative colitis activity (induced by DSS in mice), to an extent similar to the one of sulfasalazine. A deep analysis of the descriptors of the equations, suggested that molecular polarizability, dipole moment, molecular flexibility as well as the presence of certain moieties, such as 6-member rings were crucial for the pharmacological activity of the compounds.
7. Finally, it has been demonstrated that Molecular Topology represents an effective tool for the design and discovery of new *lead* compounds as potential treatment for inflammatory diseases.

