

DEPARTAMENTO DE QUÍMICA FÍSICA

APLICACIÓN DE MÉTODOS TOPOLÓGICOS Y DE
INTELIGENCIA ARTIFICIAL A LA SELECCIÓN DE
NUEVOS ANTIBACTERIANOS

MIGUEL MURCIA SOLER

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2002

Aquesta Tesi Doctoral va ser presentada a València el dia 21 de gener de 2003 davant un tribunal format per:

- Dr. D. Jose Luis Moreno Frigols
- Dr. D. Benito del Castillo García
- Dr. D. Wladimiro Díaz Villanueva
- Dr. D. Javier Hernández Haba
- Dra. Dña. M^a Teresa Salabert Salvador

Va ser dirigida per:

Prof. Dr. D. Facundo Pérez Jiménez – Francisco José García March

©Copyright: Servei de Publicacions
Miguel Murcia Soler

Depòsit legal:

I.S.B.N.:84-370-5701-9

Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115

UNIVERSITAT DE VALÈNCIA

FACULTAT DE FARMÀCIA

DEPARTAMENT DE QUÍMICA FÍSICA

**"APLICACIÓN DE MÉTODOS TOPOLÓGICOS Y DE
INTELIGENCIA ARTIFICIAL A LA SELECCIÓN DE
NUEVOS ANTIBACTERIANOS"**

Memoria que, para optar al Grado
de Doctor en Farmacia, presenta:

Miguel Murcia Soler

Directores:

Prof. Dr. D. Facundo Pérez Giménez

Prof. Dr. D. Francisco José García March



UNIVERSITAT D VALÈNCIA
DPTO. QUÍMICA FÍSICA
FACULTAD DE FARMACIA
Av. Vicente Andrés Estellés s/n
Tf. 963 86 48 94
Fax 963 86 48 92
46100 BURJASSOT (Valencia)
ESPAÑA

Facundo Pérez Giménez y Francisco José García March, Doctores en Farmacia y Profesores Titular y Asociado del Departamento de Química Física de la Facultad de Farmacia de la Universidad de Valencia,

CERTIFICAN:

Que la Tesis Doctoral presentada con el título "Aplicación de métodos topológicos y de inteligencia artificial a la selección de nuevos antibacterianos" ha sido realizada bajo nuestra dirección por el Licenciado D. **Miguel Murcia Soler** y mediante este escrito, autorizamos su presentación para optar al Grado de Doctor en Farmacia.

Burjassot, 22 de octubre de 2002

Facundo Pérez Giménez

Francisco José García March

Quiero expresar mi más sincero agradecimiento a todas aquellas personas que han colaborado en la realización del presente trabajo, así como a las que me han apoyado en todo momento. Vayan por delante mis disculpas si queda alguna persona sin citar:

En primer lugar quiero mostrar un profundo agradecimiento a mis directores, Dr. D. Facundo Pérez Giménez y Dr. D. Francisco José García March, quienes depositando en mí toda su confianza, han puesto a mi servicio todos sus conocimientos, experiencia y consejo tanto a nivel profesional como personal. A ellos me une una gran amistad y sin los cuales no hubiera sido posible la realización del presente trabajo de Tesis Doctoral.

A la Dra. Dña. María Teresa Salabert Salvador, excelente profesora y mejor persona, quien en todo momento me ha mostrado su apoyo y estímulo, así como una acertada orientación en el ámbito docente.

Al profesor D. Angel Villanueva Pareja, debo agradecerle su interés y ayuda que hicieron posible el ingreso en el departamento y línea de investigación de la que formo parte.

A mi tutor Dr. D. Ramón García Doménech, así como a los profesores Dr. D. Jorge Gálvez Álvarez y Dr. D. José Luis Moreno Frígols, por su excelente trato.

A Ana Pérez Solsona, por su colaboración en la realización de los ensayos farmacológicos y a todos los compañeros del Departamento de Química Física de la Facultad de Farmacia por el interés mostrado.

A los profesores Dra. Dña. María José Castro Bleda y Dr. D. Wladimiro Díaz Villanueva, gracias a los cuales pude disfrutar de una estancia en el Departamento de Sistemas Informáticos y Computación de la Universidad Politécnica de Valencia, ampliando conocimientos en el campo de las redes neuronales artificiales.

AGRADECIMIENTOS

En especial a mis padres Miguel y Desamparados a los que les debo todo y a mi hermana Estrella, quienes en todo momento me han apoyado y animado a seguir adelante.

Mención especial merece mi novia, Ana, por su paciencia y por haber compartido conmigo todos esos momentos buenos y no tan buenos a lo largo de la licenciatura y del desarrollo del presente trabajo, ofreciéndome siempre su apoyo, estímulo y cariño.

A mi abuelo Jesús, con quien me hubiera gustado compartir momentos destacables de mi vida como éste.

Este trabajo ha sido realizado gracias a la concesión de una Beca Predoctoral de Formación de Personal Investigador de la Generalitat Valenciana.

A mis padres.

A Ana.

**"APLICACIÓN DE MÉTODOS TOPOLÓGICOS Y DE
INTELIGENCIA ARTIFICIAL A LA SELECCIÓN DE
NUEVOS ANTIBACTERIANOS"**

ÍNDICE

1. INTRODUCCIÓN.....	1
2. ANTECEDENTES BIBLIOGRÁFICOS.....	9
2.1. Topología molecular.....	10
2.1.1. La matriz topológica.....	11
2.1.2. La matriz de distancia.....	12
2.1.3. Valencia topológica (grado de vértice).....	14
2.1.4. Índices topológicos.....	15
2.1.4.1. Índices basados en la matriz de adyacencia.....	15
2.1.4.1.1. Índices del grupo de Zagreb.....	16
2.1.4.1.2. Índice de conectividad de Randic	17
2.1.4.1.3. Índices de Kier y Hall	18
2.1.4.1.4. Índice de Platt	23
2.1.4.1.5. Índice de Gordon-Scantlebury	23
2.1.4.1.6. Otros índices topológicos basados en la matriz de adyacencia.....	24
2.1.4.2. Índices basados en la matriz de distancia	25
2.1.4.2.1. Índice de Wiener.....	25
2.1.4.2.2. Índice de Altenburg.....	26
2.1.4.2.3. Índice de Hosoya.....	26
2.1.4.2.4. Índice de Bonchev.....	27
2.1.4.2.5. Índice de conectividad suma de distancias medias (Balaban).....	27
2.1.4.2.6. Índices Kappa de forma	28
2.1.4.2.7. Índices de carga.....	30

2.1.4.3. Índices topológicos centrales.....	32
2.1.4.4. Índices basados en la teoría de la información.....	33
2.2. Redes Neuronales.....	35
2.2.1. Panorama histórico.....	35
2.2.2. Modelo biológico.....	36
2.2.2.1. Estructura de la neurona.....	37
2.2.2.2. Comunicación entre neuronas.....	37
2.2.3. Red neuronal artificial.....	39
2.2.3.1. Estado de activación.....	42
2.2.3.2. Función de salida.....	42
2.2.3.3. Conexiones entre neuronas.....	43
2.2.3.4. Función de activación.....	44
2.2.3.5. Regla de aprendizaje.....	46
2.2.3.6. Capas de neuronas y formas de conexión.....	46
2.2.4. El Perceptrón.....	47
2.2.4.1. Regla de aprendizaje del Perceptrón.....	49
2.2.5. El Perceptrón multicapa.....	53
2.2.5.1. Algoritmo de retropropagación del error (<i>backpropagation</i>).....	56
2.2.6. Aplicaciones de las redes neuronales.....	57
2.3. Terapéutica antibacteriana.....	59
2.3.1. Evolución histórica.....	59
2.3.2. Clasificación de los antibacterianos.....	61
2.3.3. Mecanismos de acción.....	63

2.3.4. Resistencia bacteriana.....	66
2.3.5. Farmacocinética.....	68
2.3.6. Elección del antimicrobiano	71
3. MATERIAL Y MÉTODOS	75
3.1. Índices topológico-estructurales.....	76
3.1.1. Obtención de una estructura de datos.....	76
3.1.2. Construcción del grafo a partir del formato SMILES	80
3.1.2.1. Formato SMILES.....	80
3.1.2.2. Obtención del grafo	82
3.1.3. Cálculo de índices.....	82
3.2. Tratamiento informático.....	88
3.2.1 Red Neuronal Artificial.....	89
3.2.1.1. Características de la red neuronal.....	90
3.2.2. Análisis Lineal Discriminante.....	93
3.2.2.1. Programa P7M.....	94
3.2.3. Diagramas de distribución farmacológica.....	96
3.3. Ensayos microbiológicos.....	98
3.3.1. Material y reactivos.....	98
3.3.2. Preparación de los productos.....	100
3.3.3. Cepas de microorganismos.....	101
3.3.3.1. Apertura de liófilos.....	104
3.3.3.2. Viabilidad de los liófilos.....	105
3.3.3.3. Glicerinado de microorganismos.....	105
3.3.4. Protocolo de dispensación en placas de microtitulación.....	106

3.4. Ensayos de toxicidad.....	110
3.4.1. Material y reactivos	110
3.4.2. Preparación de los productos.....	111
3.4.3. Protocolo de ensayo.....	111
3.5. Esquema de trabajo	114
4. RESULTADOS	117
4.1. Cálculo de índices topológico-estructurales.....	118
4.2. Discriminación	120
4.2.1. Discriminación mediante red neuronal artificial	120
4.2.2. Discriminación mediante análisis lineal discriminante	141
4.3. Búsqueda guiada y selección de estructuras.....	161
4.4. Ensayos microbiológicos.....	171
4.5. Ensayos de toxicidad.....	179
5. DISCUSIÓN.....	184
5.1. Obtención de índices topológico-estructurales.....	185
5.2. Discriminación y selección de nuevas estructuras	
con teórica actividad antibacteriana.....	186
5.2.1. Discriminación mediante red neuronal artificial	186
5.2.2. Discriminación mediante análisis lineal discriminante	188
5.2.3. Selección de nuevas estructuras	190
5.3. Ensayos farmacológicos.....	192
6. CONCLUSIONES.....	196
7. BIBLIOGRAFÍA.....	201
8. ANEXO	215

1. INTRODUCCIÓN

1. INTRODUCCIÓN

Desde la Prehistoria el hombre primitivo utilizó remedios terapéuticos porque conoció enfermedades e intentó combatirlas apelando a dos vías intelectuales: el empirismo y la magia. El empleo de estos remedios terapéuticos animales, minerales pero sobre todo vegetales a lo largo de las culturas arcaicas fue sufriendo modificaciones en cuanto a la ampliación de su número, racionalización de su utilización y comprensión de su acción terapéutica.

Durante la Antigüedad Clásica (800 a.C.-395 d.C.) y concretamente en Grecia el vocablo *phármakon* pierde ya su carácter mágico para convertirse en "remedio", con plena significación médica y definiéndose en el *Corpus Hippocraticum* como una substancia exterior al cuerpo, capaz de producir en él una acción favorable o desfavorable, difícil de delimitar de la noción de alimento; como puro medicamento capaz de modificar el estado presente; como medicamento purgante, en el sentido que la "purgación" *-kátharsis-* significó para el hipocrático, acción evacuante y en consecuencia purificadora. De la segunda mitad del siglo IV al primer tercio del siglo III a.C. surge la medicina helenística, lo que supuso por parte de la medicina griega de una base científico-natural en la que el médico deja de ser filósofo para convertirse en médico científico lo cual supuso una mayor valoración de los fármacos en tanto que se subraya especialmente la misión práctica y curativa de la medicina; pero es en Roma donde la medicina griega va a ampliarse con remedios exóticos llegados de Chipre, Capadocia, Siria, Egipto, África y España, a los que se unen nuevas drogas vegetales y animales de dichas regiones. Más tarde Galeno de Pérgamo (131-200 d.C.) en su obra *Sobre el método terapéutico* emplea hasta 473 remedios vegetales, y en menor cantidad minerales y animales.

En la Edad Media comienzan a introducirse medicamentos compuestos surgiendo ciertos problemas sobre todo en lo referente a proporciones o cantidades en que se deben mezclar los medicamentos simples para obtener los compuestos, llegando a pensar la mayoría de autores de la época que las propiedades de un fármaco compuesto son el resultado de la mezcla de las propiedades de los simples. La materia médica se enriqueció considerablemente durante los siglos XVI y XVII gracias a la incorporación de sustancias vegetales no conocidas por los europeos y descubiertas en los grandes viajes a tierras lejanas, lo que llevó consigo la aparición de nuevos herbarios y la utilización de nuevas vías de administración.

Con la llegada de Paracelso (1493-1541) comienzan a introducirse cada vez más los remedios minerales llegando a descubrirse durante el Renacimiento un preparado de mercurio activo contra la sífilis.

Con la constitución de la química como disciplina científica a finales del siglo XVIII y principios del XIX se desarrolla el conocimiento de la composición química de los seres vivos lo que era imprescindible para el estudio del efecto de los fármacos sobre los mismos así como el estudio analítico de las distintas sustancias y productos químicos utilizados en farmacia, surgiendo sustancias como la narcotina, morfina, cinconina, cafeína, colchicina, codeína y atropina entre otras. Ya en el siglo XIX Wöhler sintetiza la urea (1828), lo que demostró que la materia orgánica puede ser producida artificialmente en el laboratorio; pero la evolución de la química iba a poner de manifiesto la capacidad del hombre para obtener, también por síntesis, sustancias no existentes en la naturaleza y capaces de acciones no logradas o alcanzadas de peor modo por las sustancias naturales, de esta forma nació la quimioterapia sintética, destacando como primer paso en este nuevo campo la introducción del hidrato de cloral en medicina como hipnótico y anestésico por

Liebreich (1869), sintetizado por Liebig en 1832. Unas décadas más tarde surgieron otras sustancias como el ácido salicílico, salipirina, aspirina, antipirina, acetanilida, etc., pero uno de los hechos más importantes fue poner en relación la estructura de la molécula del medicamento con su operación terapéutica en el organismo, y buscar en el laboratorio, realizando modificaciones en esa estructura elementos capaces de actuar conforme a lo que de ellos se esperaba. Esta quimioterapia se limitaba a obtener sustancias de acción sintomática, por lo que fue necesario pasar del tratamiento quimioterápico sintomático al tratamiento quimioterápico etiológico, destacando a principios del siglo XX el descubrimiento del dioxidiamidoarsenobenzol, al que luego se denominó salvarsán - arsénico que salva- para el tratamiento de la sífilis, consiguiendo el ideal de que con la inyección de un medicamento en sangre podían destruirse distintos gérmenes patógenos sin lesión del organismo enfermo, surgiendo a partir de este momento la era moderna de la farmacoterapia.¹

La introducción de nuevos y diversos productos a mediados del siglo XX se debió en su mayoría a la modificación parcial y sistemática de la estructura química con el propósito de obtener derivados más potentes y menos tóxicos respecto al producto original.

La historia del descubrimiento de los fármacos, se encuentra íntimamente relacionada con el desarrollo de las ciencias experimentales, en general, y de la química orgánica en particular, sobre todo hasta la aparición de los métodos biológicos. Gracias al desarrollo de métodos de síntesis, aparición de técnicas instrumentales, aplicación de la informática, desarrollo de la biología molecular, se ha podido comprender la acción, el metabolismo y la toxicidad de los fármacos, estructura de los receptores, e incluso se han establecido relaciones entre la estructura química de un fármaco y su actividad biológica.

Una relación estructura-actividad, que designaremos S.A.R. ("Structure-Activity-Relationship")² se establece cuando un conjunto de propiedades de una serie de compuestos explica su actividad o respuesta biológica, a nivel cualitativo. Estas relaciones desde el punto de vista cuantitativo quedan englobadas en lo que denominamos estudios Q.S.A.R. (Relación cuantitativa estructura-actividad)^{3,4} y buscan la relación entre una propiedad determinada y la estructura química de una serie de moléculas, observando la variación de esa propiedad en la serie estudiada.

Actualmente no podemos hablar del diseño de fármacos sin nombrar los estudios Q.S.A.R. Fueron Hansch⁵ y sus colaboradores quienes en 1962 realizaron un trabajo pionero desarrollando el concepto de Q.S.A.R. Este trabajo demostró que la actividad biológica de una molécula puede ser cuantitativamente relacionada con parámetros fisicoquímicos (p.e. descriptores moleculares), por lo que introdujo la idea de que la actividad podía ser descrita por más de un parámetro. Mediante este método puede establecerse una ecuación que relaciona la respuesta biológica con parámetros descriptores de la estructura de una serie de análogos en función de ciertos factores fisicoquímicos, definidos como parámetros de solubilidad, electrónicos y estéricos.

Una variante del modelo de Hansch fue el propuesto por Free y Wilson⁶ en 1964, según el cual para una serie de compuestos considerados como derivados de una estructura madre común, se puede asumir que el efecto de un sustituyente en la actividad biológica es aditivo e independiente de la presencia o ausencia de otros sustituyentes en otras posiciones.

Además de los modelos Q.S.A.R. tradicionales anteriormente citados (Hansch y Free-Wilson), existen otros métodos de diseño molecular encaminados a la obtención de

nuevas estructuras químicas, como son la modelización molecular, métodos mecano-cuánticos y métodos convencionales.

En los últimos años y debido en parte al auge de la Informática se han desarrollado nuevos y potentes métodos que estudian las relaciones estructura-actividad mediante descripciones moleculares que se basan en las posiciones relativas en el espacio de los átomos que forman las moléculas. Estas descripciones moleculares incluyen a menudo cálculos energéticos o de propiedades moleculares. Nos referimos a estos métodos como **Modelización Molecular**. Se parte de la representación de la estructura tridimensional de la molécula mediante espectros de difracción de rayos X, bombardeando con partículas subatómicas (electrones o neutrones) o mediante resonancia magnética nuclear (RMN), especialmente para el caso de moléculas biológicas, pues la geometría de estas que nos interesa es aquella en las que las biomoléculas se encuentran en disolución o interactuando con otras, como ocurre en la interacción ligando-receptor biológico. A partir de esta geometría molecular inicial se debe hallar aquella que reduce al mínimo la energía de la molécula, es decir una "optimización geométrica", y como la energía de un sistema molecular es una función de su geometría, habrá que buscar el mínimo absoluto de dicha función.⁷

Los métodos **Mecano-Cuánticos** intentan describir en términos matemáticos la naturaleza de la materia en los niveles atómico y molecular. Se basan en la obtención de la energía asociada a una determinada geometría a partir de la resolución de la ecuación de Schrödinger, obteniendo la energía del sistema, así como una descripción de su estructura electrónica, ya que la función de onda molecular nos permite obtener una descripción de la densidad electrónica en cada uno de los puntos del entorno molecular. Los métodos

utilizados son cálculos *ab initio* (HF-SCF, DFT y EHT), métodos semiempíricos (MNDO y AM1) que incorporan más aproximaciones y los llamados métodos mixtos (QM/MM).^{7,8}

Métodos convencionales o de "replicación moduladora", que consisten en la modificación sistemática de determinados grupos estructurales de la estructura modelo. La actividad farmacológica del prototipo se mantiene, pero en algunos casos se descubre en un análogo un nuevo perfil farmacológico.⁹

El presente trabajo de Tesis Doctoral se ha desarrollado en la Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular de la Universitat de València. Fue durante los años 80 cuando este equipo investigador desarrolló un nuevo método de diseño de fármacos asistido por ordenador¹⁰, derivado de un método basado en la topología molecular llamado "Conectividad Molecular".¹¹

La conectividad molecular consiste en caracterizar estructuralmente las moléculas a través de una serie de números llamados índices, para posteriormente regresionarlos frente a propiedades físicas, químicas, biológicas o farmacológicas de las moléculas objeto de estudio, obteniendo una serie de ecuaciones denominadas funciones de conectividad.

El objeto del trabajo es utilizar la topología molecular para establecer la descripción numérica de la estructura química de las moléculas y posteriormente realizar el tratamiento matemático de los índices numéricos aplicando un modelo basado en redes neuronales combinado con análisis lineal discriminante para la discriminación de la actividad, que permite establecer las relaciones entre la molécula y sus propiedades.

Como índices numéricos utilizaremos un conjunto de Índices Topológico-Estructurales desarrollados en nuestra unidad de investigación, caracterizados por ser números enteros sencillos y cumplir una de las características fundamentales de los índices topológicos que es la de ser invariantes de grafo.

A partir de la descripción en formato SMILES de las moléculas obtendremos sus correspondientes índices topológico-estructurales que posteriormente serán tratados matemáticamente mediante análisis lineal discriminante y redes neuronales, obteniendo una función discriminante y una arquitectura de la red neuronal que sean capaces de discriminar de entre un conjunto de compuestos químicos aquellos que posean la actividad deseada, en nuestro caso la antibacteriana.

Aquellos compuestos que muestren teórica actividad antibacteriana serán sometidos a ensayos farmacológicos frente a diversos microorganismos para poner de manifiesto su actividad y la capacidad predictiva del método propuesto.

2. ANTECEDENTES BIBLIOGRÁFICOS

2. ANTECEDENTES BIBLIOGRÁFICOS

2.1. TOPOLOGÍA MOLECULAR

La topología molecular es un término general que engloba la información estructural concerniente a ramificaciones, conexiones entre átomos, forma y tamaño de la molécula.¹¹

El primer paso en los métodos Q.S.A.R. es la representación de la estructura molecular mediante números. Estos números se denominan descriptores moleculares¹² que son calculados mediante la aplicación de la teoría de grafos.

La topología molecular, es la aplicación de la teoría de grafos a la caracterización de estructuras químicas y a las correlaciones estructura-propiedad y estructura-actividad¹³, de forma que permite representar una molécula mediante un conjunto de índices numéricos denominados índices topológicos, que a pesar de no describirla con la misma precisión que lo hacen los métodos mecano-cuánticos, sí que sirven para llevar a cabo los estudios Q.S.A.R.

Estos estudios Q.S.A.R. permiten relacionar la estructura de una molécula con su actividad, caracterizando numéricamente la estructura molecular, por lo que las diferencias estructurales se cuantifican y permiten diferenciar numéricamente unas estructuras de otras.¹⁴

Hacia los años 70 surgió un nuevo enfoque de la evaluación cuantitativa de la estructura de la molécula, denominado Conectividad Molecular¹¹, basado en la caracterización de la estructura molecular mediante unos índices topológicos llamados Índices de Conectividad que proporcionan información referente al número de átomos, enlaces, insaturaciones y ensamblaje topológico de la molécula.¹⁵

2.1.1. LA MATRIZ TOPOLÓGICA

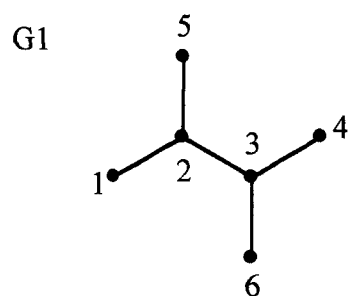
El conjunto de conexiones entre pares de átomos adyacentes puede ser representado en forma de matriz, denominada Matriz Topológica o Adyacente.

Para obtener dicha matriz, en primer lugar se dibuja el grafo de hidrógenos suprimidos. Es el grafo que se obtiene suprimiendo los átomos de hidrógeno de la molécula objeto de estudio, de forma que los átomos quedan simbolizados como vértices (●) y los enlaces como ejes (—) del grafo. Posteriormente se numeran los átomos o vértices de forma aleatoria y se construye la matriz topológica. El elemento a_{ij} de la matriz toma valor uno si los vértices i y j son adyacentes y toma valor cero en cualquier otro caso.

$$(A)_{ij} = \begin{cases} a_{ij} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases} \quad (2.1)$$

Se obtiene así una matriz de n filas por n columnas, siendo n el número de vértices del grafo. Esta matriz es simétrica respecto a la diagonal principal y presenta valor cero para todos los elementos de esta diagonal.

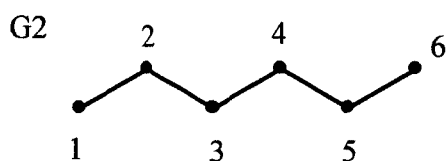
A continuación se muestran los grafos y las matrices topológicas o de adyacencia para el 2,3-dimetil-butano (G_1) y para el n -hexano (G_2).



$$A(G1) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

G1.- 2,3-dimetil-butano

Matriz de adyacencia para G1



$$A(G2) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

G2.- *n*-hexano

Matriz de adyacencia para G2

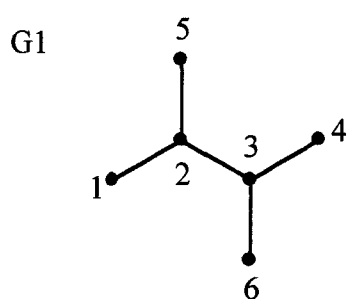
2.1.2. LA MATRIZ DE DISTANCIA

La Matriz de Distancia (D) también puede caracterizar un grafo molecular. Sus elementos d_{ij} , denominados distancias, son iguales al número de ejes que enlazan los vértices i y j por el camino más corto, siendo $d_{ij} = 1$ para vértices vecinos, y por definición $d_{ii} = 0$. La matriz de distancia puede derivarse de la matriz de adyacencia.

$$(D)_{ij} = \begin{cases} d_{ij} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases} \quad (2.2)$$

Mientras que en la matriz de adyacencia los elementos de la misma son iguales a cero o uno, en la matriz de distancia sus elementos toman valor dos, tres, etc., dependiendo del número de ejes entre los vértices considerados.

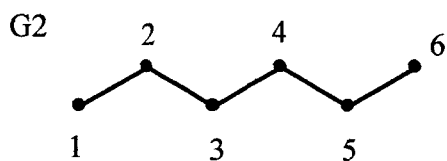
Las matrices de distancia para el 2,3-dimetil-butano (G1) y para el *n*-hexano (G2) se muestran a continuación:



$$D(G1) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 3 \\ 1 & 0 & 1 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 & 1 \\ 3 & 2 & 1 & 0 & 3 & 2 \\ 2 & 1 & 2 & 3 & 0 & 3 \\ 3 & 2 & 1 & 2 & 3 & 0 \end{pmatrix} \end{matrix}$$

G1.- 2,3-dimetil-butano

Matriz de distancia para G1



$$D(G2) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 4 & 3 & 2 & 1 & 0 & 1 \\ 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix} \end{matrix}$$

G2.- *n*-hexano

Matriz de distancia para G2

La matriz de distancia, al igual que la matriz de adyacencia, es simétrica respecto a la diagonal principal y presenta valor cero para todos los elementos de la misma.

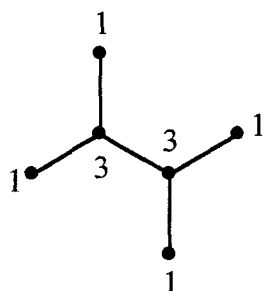
2.1.3. VALENCIA TOPOLÓGICA (Grado de vértice)

Gutman *et al.*¹⁶ fueron los primeros interesados en la caracterización cuantitativa del grado de ramificación. Definieron grado de vértice (δ_i) como el número de enlaces σ que posee cada átomo i , excluyendo los enlaces con átomos de hidrógeno.

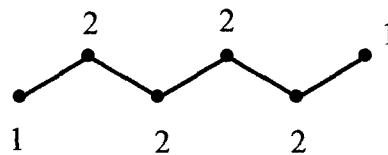
$$\delta_i = \sigma_i - H_i \quad (2.3)$$

El valor del grado de vértice δ_i para un átomo i es igual a la suma de todos los elementos de la fila i o columna i de la matriz de adyacencia, para moléculas que no posean dobles o triples enlaces.

El grado de vértice también se denomina valencia topológica, que es el valor del número de ejes conectados a ese vértice.



Grado de vértice para G1



Grado de vértice para G2

Para aquellas moléculas con dobles y triples enlaces, se modifica la matriz de adyacencia de forma que cada elemento a_{ij} de la matriz toma valor uno, dos o tres, si entre los vértices i y j figura un enlace simple, doble o triple respectivamente, tomando valor cero en cualquier otro caso.

2.1.4. ÍNDICES TOPOLÓGICOS

La teoría de grafos ha sido ampliamente utilizada para la caracterización de estructuras químicas, tanto en correlaciones estructura-actividad como en estructura-propiedad mediante los denominados índices topológicos.

La característica fundamental de un índice topológico es que sea invariante de grafo, es decir, su valor no se modifica sea cual sea el orden en que se numeren los átomos.¹⁷⁻¹⁹

Han sido propuestos un gran número de índices topológicos, pero sólo unos pocos han sido ampliamente utilizados en análisis Q.S.A.R. La matriz de adyacencia y la matriz de distancia han sido las más usadas para la formulación de índices topológicos. Existen numerosos criterios de clasificación de índices topológicos, nosotros utilizaremos el propuesto por Balaban *et al.*²⁰, que se basa en la clasificación en función del origen de los índices en lugar de obedecer a una clasificación cronológica.

2.1.4.1. Índices basados en la Matriz de Adyacencia

Dentro de este grupo trataremos a parte el **Índice de Adyacencia Total** por ser el más simple, ya que se trata del sumatorio de todos los elementos de la matriz de adyacencia, pero como la matriz es simétrica respecto a la diagonal principal, el índice puede ser reducido al sumatorio de los elementos de la mitad superior de dicha matriz, siendo la ecuación correspondiente:

$$A = (1/2) \sum_{i,j=1}^n a_{ij} \quad (2.4)$$

Podemos comprobar que para los compuestos 2,3-dimetil-butano (G1) y *n*-hexano (G2) el valor del índice de adyacencia total es 5.

2.1.4.1.1. Índices del grupo de Zagreb

Los primeros índices topológicos basados en la matriz de adyacencia fueron introducidos por el grupo de Zagreb. Partiendo de la definición de grado de vértice fueron obtenidos los índices M_1 y M_2 .¹⁶

$$M_1 = \sum_{i=1}^n \delta_i^2 \quad (2.5)$$

donde δ_i es el grado de vértice o valencia topológica para el vértice i . Para los compuestos 2,3-dimetil-butano (G1) y n -hexano (G2) los valores del índice son:

$$M_1(G1) = 4 \cdot 1^2 + 2 \cdot 3^2 = 22 \quad M_1(G2) = 2 \cdot 1^2 + 4 \cdot 2^2 = 18$$

$$M_2 = \sum \delta_i \cdot \delta_j \quad (2.6)$$

donde δ_i y δ_j son las valencias topológicas de los vértices i y j del eje ij .

El valor de M_2 para los compuestos G1 y G2 es:

$$M_2(G1) = 4(1 \cdot 3) + 1(3 \cdot 3) = 21 \quad M_2(G2) = 2(1 \cdot 2) + 3(2 \cdot 2) = 16$$

Los dos índices de Zagreb son aplicables para el cálculo de la energía de los electrones π en sistemas conjugados.

2.1.4.1.2. Índice de Conectividad de Randić.

Un índice de conectividad, χ_R , similar a M_2 fue introducido por Randić^{21,22} para caracterizar la ramificación molecular:

$$\chi_R = \sum (\delta_i \cdot \delta_j)^{-1/2} \quad (2.7)$$

donde δ_i y δ_j son las valencias topológicas de los vértices i y j del eje ij , siendo el valor del índice para los compuestos G1 y G2:

$$M_2(G1) = 4(1 \cdot 3)^{-1/2} + 1(3 \cdot 3)^{-1/2} = 2.642$$

$$M_2(G2) = 2(1 \cdot 2)^{-1/2} + 3(2 \cdot 2)^{-1/2} = 2.914$$

El índice de conectividad es enlace-aditivo, depende del tipo de enlace C-C en la molécula. Según este índice los ejes de un grafo pueden ser clasificados en varios tipos, dependiendo de las valencias topológicas de los vértices. Para el caso de los hidrocarburos hay diez tipos diferentes de ejes que son (1,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4), (3,3), (3,4) y (4,4). La contribución de cada eje se calcula como el inverso de la raíz cuadrada del producto de las valencias topológicas de los vértices que determinan el eje ($1/(\delta_i \cdot \delta_j)^{1/2}$), de manera que a una mayor ramificación de la molécula le corresponde un menor valor de dicho índice.

Sirva como ejemplo el eje de tipo (2,4):

- valencias topológicas de cada vértice: 2 - 4
- contribución de cada vértice: $\sqrt{2}$, 2
- contribución del eje: 0.354

2.1.4.1.3. Índices de Kier y Hall.

A partir del modelo de Randić, Kier y Hall^{11,23} estudiaron la posibilidad de desarrollar un método más general para describir las estructuras moleculares orgánicas mediante índices topológicos.

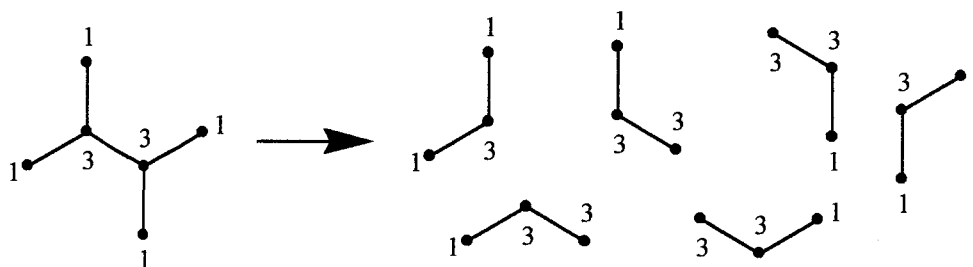
Asumiendo que $\chi_R = {}^1\chi$, donde el prefijo 1 indica que el índice ${}^1\chi$ es una disección de la estructura molecular en fragmentos de dos átomos contiguos, lo que equivale a un eje del grafo.

En el caso del índice de orden 2 (${}^2\chi$) se corresponde con la disección de la estructura molecular en fragmentos de tres átomos contiguos, lo que equivale a dos ejes del grafo. Kier y Hall establecieron la siguiente expresión general para calcular disecciones de mayor orden:

$${}^m\chi = \sum (\delta_{i_1} \cdot \delta_{i_2} \cdots \delta_{i_{m+1}})^{-1/2} \quad (2.8)$$

Como ejemplo se muestra el cálculo del índice de orden 2 para la molécula 2,3-dimetilbutano (G1), como disección de dos ejes (tres átomos contiguos):

$${}^2\chi = \sum (\delta_{i_1} \cdot \delta_{i_2} \cdot \delta_{i_3})^{-1/2} \quad (2.9)$$



$${}^2\chi(G1) = \sum (\delta_i \cdot \delta_j \cdot \delta_k)^{-1/2} = (1 \cdot 3 \cdot 1 + 1 \cdot 3 \cdot 3 + 1 \cdot 3 \cdot 3 + 3 \cdot 3 \cdot 1 + 3 \cdot 3 \cdot 1 + 1 \cdot 3 \cdot 1)^{-1/2}$$

$${}^2\chi(G1) = 2.488$$

$${}^2\chi(G2) = 1.707$$

El índice de orden 0 se corresponde con la disección de la estructura molecular en fragmentos de átomos.

$${}^0\chi = \sum (\delta_i)^{-1/2} \quad (2.10)$$

$${}^0\chi(G1) = 5.155$$

$${}^0\chi(G2) = 4.828$$

Los índices de conectividad de Kier y Hall, en general pueden ser expresados como la suma de los términos de los distintos subgrafos conectados, de acuerdo con la expresión:

(2.11)

$${}^m\chi_t = \sum_{j=1}^{n^m} {}^mS_j$$

donde n^m es el número de subgrafos de tipo "t" y orden "m" (número de ejes del fragmento molecular) y mS_j es una cantidad calculada para cada subgrafo.

Seguidamente se detallan algunos tipos de subgrafos con la correspondiente identificación del índice:

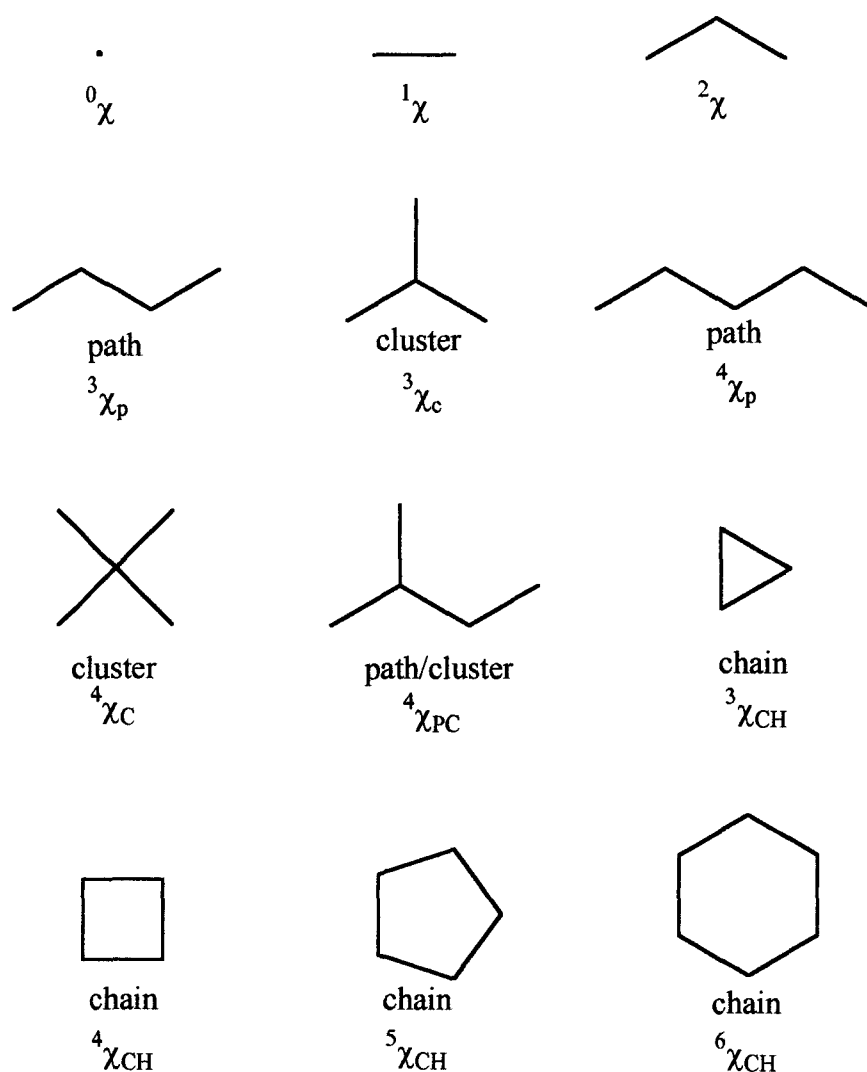


Figura 2.1. Tipos de subgrafos.

Con posterioridad Kier y Hall desarrollaron una extensión de esta aproximación para moléculas con dobles o triples enlaces, así como aquellas moléculas que contienen heteroátomos. En el primer caso el procedimiento tiene en cuenta la valencia de cada átomo y por consiguiente la valencia topológica δ_i es reemplazada por la conectividad del átomo δ_i^v :

$$\delta_i^v = Z_i^v - H_i \quad (2.12)$$

donde Z_i^v y H_i son, respectivamente, el número de electrones de valencia del átomo i y el número de átomos de hidrógeno suprimidos. El cálculo del correspondiente índice ${}^m\chi_i^v$ (chi de valencia) se realiza utilizando la ecuación (2.8). Para el caso de las moléculas con heteroátomos (O, N, F) el cálculo se basa en la ecuación (2.12), teniendo en cuenta los átomos adyacentes enlazados más los electrones π y los electrones desapareados. Para los heteroátomos (P, S, Cl, Br y I), deben considerarse los electrones de no-valencia. Kier y Hall propusieron la siguiente ecuación que considera los electrones de valencia y los de no-valencia:

$$\delta_i^v = (Z_i^v - H_i)/(Z_i - Z_i^v - 1) \quad (2.13)$$

donde Z_i es el número de los electrones totales del átomo i . Valores de δ_i^v para C, N y O en varios estados híbridos y para los heteroátomos P, S, Cl, Br y I se muestran en las tablas 2.1 y 2.2 respectivamente.

Los índices de conectividad ${}^m\chi_i$ y ${}^m\chi_i^v$ han sido ampliamente utilizados en análisis Q.S.A.R. en series de moléculas biológicamente activas (inhibición enzimática, actividad anestésica, etc.), aunque la interpretación de los resultados en ocasiones no siempre es sencilla. De hecho, aunque Kier y Hall señalaron que la información topológica está ligada a los índices de conectividad, estos pueden ser correlacionados con otras propiedades fisicoquímicas como refractividad molar, polarizabilidad y/o lipofilia.

Tabla 2.1. Valores de δ^v para los átomos C, N, O y F en el grafo de hidrógenos suprimidos.

Átomo	Estado híbrido	Átomos δ adyacentes	Electrones de valencia	Electrones π	Electrones desapareados	δ^v
C	sp^3	1	4	0	0	1
		2	4	0	0	2
		3	4	0	0	3
		4	4	0	0	4
	sp^2	1	4	1	0	2
		2	4	1	0	3
		3	4	1	0	4
	sp	1	4	2	0	3
		2	4	2	0	4
	N	sp^3	1	5	0	2
2			5	0	2	4
3			5	0	2	5
sp^2		1	5	1	2	4
		2	5	1	2	5
sp		1	5	2	2	5
O	sp^3	1	6	0	4	5
		2	6	0	4	6
	sp^2	1	6	1	4	6
F		1	7			

Tabla 2.2. Valores de δ^v para los átomos P, S, Cl, Br y I.

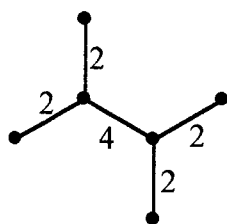
Átomo	$Z_i^v - H_i$	$Z - Z^v - 1$	δ^v
P	3	9	0.33
	4	9	0.44
	5	9	0.56
S	5	9	0.56
	6	9	0.67
Cl	7	9	0.78
Br	7	27	0.26
I	7	47	0.16

2.1.4.1.4. Índice de Platt

Platt^{24,25} fue el primero en usar un índice topológico en correlaciones con ciertas propiedades moleculares y en conjunción con otros índices topológicos. El índice de Platt se define como:

$$F = \sum_{j=1}^A e_j \quad (2.14)$$

donde e_j es el grado del eje j , es decir, el número de ejes adyacentes al eje j . Los valores del índice para los grafos G1 y G2 son:



$$F(G1) = 2 + 2 + 4 + 2 + 2 = 12$$



$$F(G2) = 1 + 2 + 2 + 2 + 1 = 8$$

2.1.4.1.5. Índice de Gordon-Scantlebury

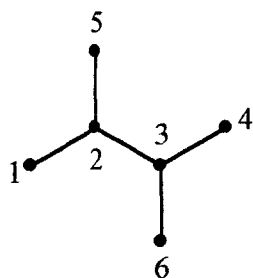
El índice de Gordon-Scantlebury²⁶ está relacionado con el valor de F , y se representa por la ecuación:

$$N_2 = \sum (P_2)_i \quad (2.15)$$

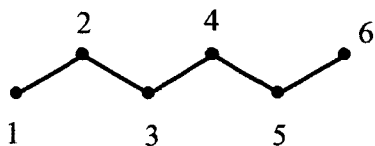
N_2 se define como el número de distintas formas en que un enlace triatómico C—C—C puede ser superpuesto en el grafo de hidrógenos suprimidos. En esta ecuación P_2 son los fragmentos de longitud 2 en el grafo. Una comparación entre N_2 y F muestra que:

$$N_2 = \frac{F}{2} \quad (2.16)$$

El valor del índice N_2 para los grafos G1 y G2 es:



$$N_2(G1) = (125) + (123) + (523) + (234) + (236) + (436) = 6$$



$$N_2(G2) = (123) + (234) + (345) + (456) = 4$$

2.1.4.1.6. Otros índices topológicos basados en la matriz de adyacencia.

Gutman y Randić²⁷ propusieron un procedimiento para la comparación (código de comparación) de estructuras químicas.

Randić^{28,29} transformó la matriz de adyacencia del grafo en una notación binaria por permutaciones de filas y columnas hasta llegar al resultado de un número binario más pequeño, (SBN-Smallest Binary Notation).

2.1.4.2. Índices basados en la Matriz de Distancia

2.1.4.2.1. Índice de Wiener

Harold Wiener³⁰⁻³³ propuso el primer índice topológico basado en la matriz de distancia. Definió para hidrocarburos saturados el índice w como el sumatorio del número de enlaces entre todos los pares de vértices. Este índice el cual refleja la ramificación de la molécula, puede ser calculado a partir de la suma de los elementos de la semimatriz de distancia que se encuentran fuera de la diagonal principal, de acuerdo con la siguiente ecuación:

$$w = \frac{1}{2} \sum_{i,j} d_{ij} \quad (2.17)$$

Así, para el caso de los grafos G1 y G2, el valor del citado índice es $w(G1) = 29$ y $w(G2) = 35$.

Además de este índice Wiener también formuló el denominado número de polaridad p . Es el número de pares de vértices, separados por tres ejes, es igual a la suma de los elementos con valor tres en la semimatriz de distancia sin tener en cuenta la diagonal principal. Para los grafos G1 y G2 su valor es $p(G1) = 4$ y $p(G2) = 3$.

Los índices w y p han sido utilizados para correlacionar valores de puntos de ebullición y otras propiedades de hidrocarburos saturados acíclicos.

2.1.4.2.2. Índice de Altenburg

Altenburg^{34,35} modificó el índice de Wiener w expresándolo en forma de polinomio,

$$(2.18)$$

donde g_i es la frecuencia del número de pares de vértices separados por d_i enlaces y a_i ha sido introducido únicamente para la enumeración en la forma polinómica. Para G_1 y G_2 el valor es:

$$P_A(G_1, a) = 5 \cdot a_1 + 6 \cdot a_2 + 4 \cdot a_3 = 5 \cdot 1 + 6 \cdot 2 + 4 \cdot 3 = 29$$

$$P_A(G_2, a) = 5 \cdot a_1 + 4 \cdot a_2 + 3 \cdot a_3 + 2 \cdot a_4 + 1 \cdot a_5 = 5 \cdot 1 + 4 \cdot 2 + 3 \cdot 3 + 2 \cdot 4 + 1 \cdot 5 = 35$$

2.1.4.2.3. Índice de Hosoya

Hosoya³⁶ introdujo un nuevo índice topológico, extendiendo la aplicación de Altenburg para grafos cíclicos, definiéndolo por la siguiente ecuación:

$$Z = \sum_k p(G, K) \quad (2.19)$$

donde $p(G, K)$ representa el número de caminos en los cuales K ejes de todos los enlaces del grafo G pueden ser elegidos para que dos de ellos no sean adyacentes. Para el caso del grafo G_1 :

$$p(G_1, 0) = 1 \quad (\text{por definición}) \qquad p(G_1, 2) = 4$$

$$p(G_1, 1) = 5 \quad (\text{igual al número de ejes}) \qquad p(G_1, 3) = 0$$

$$Z(G_1) = 1 + 5 + 4 + 0 = 10$$

2.1.4.2.4. Índice de Bonchev

Por analogía con el grado de vértice δ_i propuesto por Gutman *et al.*, Bonchev *et al.*³⁷ desarrollaron el denominado índice rango de distancia, el cual se basa en la suma de las distancias del vértice i , que corresponde a la suma de las distancias de la fila i en la matriz de distancia.

$$V_{Di} = \sum d_{ij} \quad (2.20)$$

Para los grafos G1 y G2 el valor de este índice es 58 y 70 respectivamente. Podemos establecer una relación entre este índice y el índice de Wiener:

$$V_{Di} = 2 \cdot w \quad (2.21)$$

2.1.4.2.5. Índice de conectividad suma de distancias medias (Balaban).

De forma similar al índice de conectividad de Randic, Balaban^{38,39} desarrolló el índice de conectividad sumatorio de las distancias medias, definido como:

$$J = \frac{q}{\mu + 1} \cdot \sum \bar{V}_{Di} \cdot \bar{V}_{Dj} \quad (2.22)$$

donde i y j son vértices adyacentes, q es el número de ejes, μ representa el número de ciclos y \bar{V}_{Di} es la suma de distancia media.

$$\bar{V}_{Di} = \frac{V_{Di}}{q} \quad (2.23)$$

2.1.4.2.6. Índices Kappa de forma

Kier⁴⁰ desarrolló el índice ${}^2\kappa$ basado en la cuenta de fragmentos de dos enlaces. El método considera dos estructuras de referencia no cíclicas. Considerando como ejemplo grafos de cinco átomos podemos obtener dos estructuras de referencia, pentano o neopentano.



El grafo del pentano tiene el mínimo número de fragmentos de dos enlaces para cinco átomos; Kier introdujo el término ${}^2P_{\min}$ como el recuento del número de caminos de dos enlaces, que para cualquier molécula no-cíclica es igual a la ecuación siguiente:

$${}^2P_{\min} = A - 2 \quad (2.24)$$

donde A es el número total de vértices en el grafo. Para el caso del neopentano, el grafo tiene el máximo número posible de eslabones de dos enlaces (${}^2P_{\max}$) para cinco átomos y corresponde a una estructura en forma de estrella. El valor ${}^2P_{\max}$ se calcula para cualquier serie isomérica mediante la ecuación:

$${}^2P_{\max} = \frac{(A-1)(A-2)}{2} \quad (2.25)$$

Para el caso del pentano y neopentano obtenemos valores de ${}^2P_{\min} = 3$ y ${}^2P_{\max} = 6$ respectivamente. Obviamente cuando $A > 5$ las estructuras de estrella no son reales por la naturaleza tetravalente del carbono.

$${}^2P_{\min} \leq {}^2P_i \leq {}^2P_{\max} \quad (2.26)$$

Kier obtuvo para el índice ${}^2\kappa$ la siguiente expresión:

$${}^2\kappa = 2 \frac{{}^2P_{\min} \cdot {}^2P_{\max}}{({}^2P_i)^2} = \frac{(A-1)(A-2)^2}{({}^2P_i)^2} \quad (2.27)$$

Posteriormente el índice de forma ${}^2\kappa$ se extendió para el caso de heteroátomos y diferentes estados de valencia del átomo de carbono; el índice se expresó como:

$${}^2\kappa_\varepsilon = \frac{(A+\alpha-1)(A+\alpha-2)}{({}^2P_i + \alpha)^2} \quad (2.28)$$

donde se introduce el factor α , que se define como:

$$\alpha_x = \frac{r_x}{r_{Csp^3}} - 1 \quad (2.29)$$

donde r_x es el radio covalente del átomo x y r_{Csp^3} es el radio covalente para el carbono en el estado de hibridación sp^3 .

El concepto de índice de forma desarrollado a partir del grafo molecular descrito anteriormente, fue extendido⁴¹ para incluir dos índices adicionales derivados de recuento de caminos de un enlace (1P) y de tres enlaces (3P), los cuales se denominaron ${}^1\kappa$ y ${}^3\kappa$, respectivamente:

$${}^1\kappa = 2 \frac{{}^1P_{\min} \cdot {}^1P_{\max}}{({}^1P_i)^2} = \frac{A(A-1)^2}{({}^1P_i)^2} \quad (2.30)$$

$${}^3\kappa = 2 \frac{{}^3P_{\min} \cdot {}^3P_{\max}}{({}^3P_i)^2} \quad (2.31)$$

La ecuación (2.31) tiene dos soluciones dependiendo si A es par o impar:

para A par:

$${}^3\kappa = \frac{A(A-3)(A-2)^2}{({}^3P_i)^2} \quad (2.32a)$$

para A impar:

$${}^3\kappa = \frac{A(A-1)(A-3)^2}{({}^3P_i)^2} \quad (2.32b)$$

Las definiciones anteriores pueden modificarse para diferentes contribuciones de átomos de carbono con estado de hibridación distinto de sp^3 , reemplazando A por $A + \alpha$ y usando ${}^1P_i + \alpha$ o ${}^3P_i + \alpha$ en las ecuaciones (2.30), (2.32a) y (2.32b).

2.1.4.2.7. Índices de carga

Gálvez *et al.*⁴² presentaron alrededor del año 1993 unos nuevos índices topológicos denominados índices de carga, G_k y J_k , los cuales, de alguna forma evalúan las transferencias de carga entre pares de átomos y, por lo tanto las transferencias de carga globales en la molécula.

Teniendo en cuenta que muchas propiedades físicas, químicas y biológicas están relacionadas con la distribución de carga, es necesaria la introducción de un índice topológico capaz de caracterizar esta propiedad.

Como la influencia de la carga decrece con el cuadrado de la distancia, se emplea como primer descriptor de carga la matriz de las distancias inversas al cuadrado, denominada por D^* , siendo las distancias, distancias topológicas, es decir, el número de ejes que separan los vértices i y j por el camino más corto. Sin embargo, para evitar la división por cero, los elementos de la diagonal principal de la matriz se mantienen.

Definimos la matriz M como:

$$M = A \times D^* \quad (2.33)$$

donde A es la matriz de adyacencia ($N \times N$) y D^* la matriz de distancias inversas al cuadrado. Los términos de carga, CT , se definen como:

$$CT_{ij} = m_{ij} - m_{ji} \quad (2.34)$$

donde m_{ij} y m_{ji} corresponden a elementos de la matriz M .

Se define el índice de carga G_k , como la suma en valor absoluto de los términos CT_{ij} definidos para los vértices i y j situados a una distancia D_{ij} igual a k :

$$G_k = \sum_{\substack{i=1, j=i+1 \\ i=N-1, j=N}} |CT_{ij}| \delta \cdot (k, D_{ij}) \quad (2.35)$$

donde N es el número de vértices en el grafo y los valores D_{ij} son los elementos de la matriz de distancia. G_k representa el sumatorio de todos los términos CT_{ij} con $D_{ij} = k$. Este descriptor es invariante de grafo y puede evaluar la transferencia de carga entre átomos localizados a la distancia topológica k . Para el caso de moléculas lineales existen $(N-1)$ valores de G_k desde G_1 hasta G_{N-1} .

Además se introduce el índice J_k como:

$$J_k = G_k / (N - 1) \quad (2.36)$$

donde N es el número de átomos o vértices. Este índice representa el valor medio de la carga transferida por cada eje o enlace del grafo, ya que el número de ejes para compuestos acíclicos es $N - 1$.

También se han descrito los índices de carga de valencia G_k^v y J_k^v . El cálculo es el mismo, salvo que se debe introducir en la diagonal principal de la matriz de adyacencia el

valor de la electronegatividad de Pauling, modificada por un factor de escala tal que el valor asignado al átomo de cloro sea dos.

2.1.4.3. Índices Topológicos Centrales

Balaban propuso un conjunto de cinco índices topológicos clasificados como índices topológicos centrales⁴³ en base a secuencias de números obtenidos por reducción de un grafo acíclico hacia el centro del mismo. Balaban desarrolló un índice céntrico C , un índice cuadrático normalizado Q y dos índices binormalizados C' y Q' respectivamente, definidos como:

$$B = \sum_{i=1}^N d_i^2 \quad (2.37)$$

$$C = (1/2)(B - 2N + U) \quad (2.38)$$

$$Q = 3V_4 + V_3 \quad (2.39)$$

$$C' = \frac{B - 2N + U}{(N - 2)^2 - 2 + U} \quad (2.40)$$

$$Q' = \frac{3V_4 + V_3}{2(N - 2)(N - 3)} \quad (2.41)$$

donde d_{ij} es el número de vértices suprimidos en cada paso, N es el número total de vértices, $U = [1 - (-1)^N]$ y V_3 y V_4 son el número de vértices de grado tres y cuatro respectivamente. El índice B es similar al índice M_1 de Zagreb, mientras que el índice céntrico C está relacionado con el índice N^2 de Gordon-Scantlebury.

2.1.4.4. Índices basados en la Teoría de la Información

De acuerdo con Balaban la cuarta clase de índices puede ser definida sobre cualquier base estructural sea topológica o no-topológica, a partir de la teoría de la información.^{44,45} Esta teoría permite la introducción de índices que contienen información referente a la estructura. Para un sistema con N elementos (átomos) distribuidos en K clases de equivalencia N_1, N_2, \dots, N_K y con una distribución de probabilidad $P (p_1, p_2, p_3, \dots, p_k)$, de acuerdo con la ecuación de Shannon,⁴⁴ pueden ser derivados dos valores interrelacionados; el primero de estos es la información contenida por átomo, definida como

$$i = -\sum p_i \cdot \log p_i \quad (2.42)$$

donde p_i es la probabilidad de seleccionar aleatoriamente un átomo del conjunto. La información contenida en una molécula con N átomos es iN . Bruillon⁴⁶ modificó esta expresión para obtener la redundancia de la molécula.

$$R = 1 - \frac{i}{\log N} \quad (2.43)$$

Aunque otros muchos índices basados en la información han sido concebidos, los más útiles parecen ser los desarrollados por Bonchev y Trinajstić.⁴⁷ Estos índices han sido definidos como información contenida e información contenida promedio en coeficientes polinomiales (I_{pc}, \bar{I}_{pc}), en distancias en el grafo (I_D^w, \bar{I}_D^w) y en la distribución de distancias (I^E, \bar{I}_D^E)

$$I_{pc} = Z \log_2 Z - \sum_K p(G, K) \log_2 p(G, K) \quad (2.44)$$

$$\bar{I}_{pc} = I_{pc} / K \quad (2.45)$$

$$I^E = \frac{N(N-1)}{2} \cdot \log_2 \frac{N(N-1)}{2} \cdot g_i \log_2 g_i \quad (2.46)$$

$$\bar{I}_D^E = I_D^E / \frac{N(N-1)}{2} \quad (2.47)$$

$$I_D^w = w \log_2 w - g_i d_{ij} \log_2 d_{ij} \quad (2.48)$$

$$\bar{I}_D^w = I_D^w / w \quad (2.49)$$

Estos índices se han definido a partir de parámetros como los índices de Wiener w , valor Z de Hosoya y el índice de conectividad de Randic χ ; pues bien, Bonchev y Trinajstić también definen sus valores promedio como:

$$\bar{Z} = Z / K \quad (2.50)$$

$$\bar{w} = \frac{2w}{N(N-1)} \quad (2.51)$$

$$\bar{\chi} = \frac{\chi}{N(N-1)} \quad (2.52)$$

2.2. REDES NEURONALES

2.2.1. PANORAMA HISTÓRICO

La construcción y diseño de máquinas capaces de realizar procesos que requieren un cierto grado de inteligencia ha sido uno de los principales objetivos de los científicos a lo largo de la historia. Inicialmente los esfuerzos se dirigieron a la obtención de autómatas, es decir máquinas que realizaran alguna función típica de los seres humanos, sin embargo, hoy en día se sigue investigando en esta misma línea con herramientas muy sofisticadas, y debido a que la habilidad mecánica ha pasado a convertirse en disponibilidad microinformática, actualmente existen diversas maneras de realizar procesos similares a los inteligentes y que podemos encuadrar dentro de la denominada inteligencia artificial.

El primero en estudiar el cerebro como una forma de ver el mundo de la computación fue Alan Turing en 1936, pero los primeros teóricos que introdujeron fundamentos de computación neuronal fueron el neurofisiólogo Warren McCulloch y el matemático Walter Pitts, los cuales modelaron una red neuronal simple mediante circuitos eléctricos en 1943.⁴⁸

Frank Rosenblatt⁴⁹ comenzó en 1957 el desarrollo del Perceptrón, es la red neuronal más antigua y actualmente se utiliza de varias formas para la aplicación como reconocedor de patrones, la cual es capaz de generalizar el conocimiento adquirido, es decir, después de haber aprendido una serie de patrones era capaz de reconocer otros similares. Dos años más tarde Bernard Widrow y Marcial Hoff desarrollaron el modelo ADALINE, primera red neuronal aplicada a un problema real (eliminación de ecos en líneas telefónicas).⁵⁰

En 1969 surgieron numerosas críticas y resultados pesimistas que frenaron hasta 1980 las investigaciones sobre redes neuronales. Durante los años 80 se produce un resurgimiento con nuevos algoritmos, nuevas arquitecturas y experimentos demostrativos, así, John Hopfield presentó una red a la que dio su nombre en 1982. De 1986 a 1995 se produce un apogeo en el campo de las redes neuronales, editándose multitud de libros, revistas y se realizan congresos donde se pone de manifiesto las innumerables aplicaciones de estos modelos. A partir de 1996 se establece un asentamiento de la materia obteniéndose resultados formales por parte de investigadores como Bishop, Rippley, Ney, Lugosi, etc. En la actualidad se publican numerosos trabajos y surgen nuevas aplicaciones que las empresas lanzan al mercado como nuevos productos, sobre todo en el campo de la simulación.

2.2.2. MODELO BIOLÓGICO

La teoría y modelado de redes neuronales artificiales está inspirada en la estructura y funcionamiento del sistema nervioso, donde la neurona es el elemento fundamental.

El sistema nervioso central del ser humano contiene alrededor de 10^{11} neuronas, también contiene entre 10 y 50 veces este número de células gliales. Las neuronas, que son los bloques básicos de construcción del sistema nervioso, evolucionaron a partir de células neuroefectoras primitivas. Las neuronas cumplen la función de recibir e integrar información y de enviar señales a otros tipos de células excitables a través de contactos sinápticos, una de las características que las diferencian del resto de células vivas, es su capacidad de comunicarse.

2.2.2.1. Estructura de la neurona

Las neuronas se componen básicamente de tres partes: el cuerpo neuronal o soma, una prolongación larga y ramificada llamada axón con ramificaciones denominados terminales axónicos y unas prolongaciones muy ramificadas alrededor del soma llamadas dendritas.

La forma de una neurona depende de la función que cumple, es decir de la posición que ocupa en la red de neuronas y de los contactos que recibe.

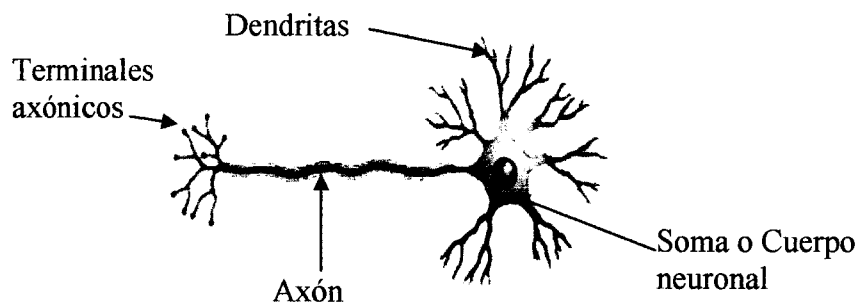


Figura 2.2. Estructura de la neurona.

En forma esquemática, se puede decir que las dendritas actúan como antenas que reciben los contactos de otras células. En el soma se lleva a cabo la integración de toda la información obtenida en las dendritas. Finalmente el axón transmite a otras células el mensaje resultante de la integración.

2.2.2.2. Comunicación entre neuronas

Las neuronas se organizan en redes y sistemas. El contacto entre ellas se realiza a través de contactos funcionales altamente especializados denominados sinapsis, entre

terminales axónicos de una neurona y las dendritas de otra neurona. La mayor de parte de las sinapsis son de tipo químico, es decir, utilizan moléculas llamadas neurotransmisores para comunicarse entre sí.

La información que activa a las dendritas es convertida en el soma en una señal eléctrica llamada potencial de acción. Esta señal viaja a lo largo del axón a una velocidad de centenares de metros por minuto. En la terminación del axón el potencial de acción promueve la liberación de los neurotransmisores. Estos difunden hacia la neurona vecina (post-sináptica) donde activan receptores específicos capaces de modificar la actividad eléctrica. A su vez, esta información podrá generar un nuevo potencial de acción en la segunda neurona.

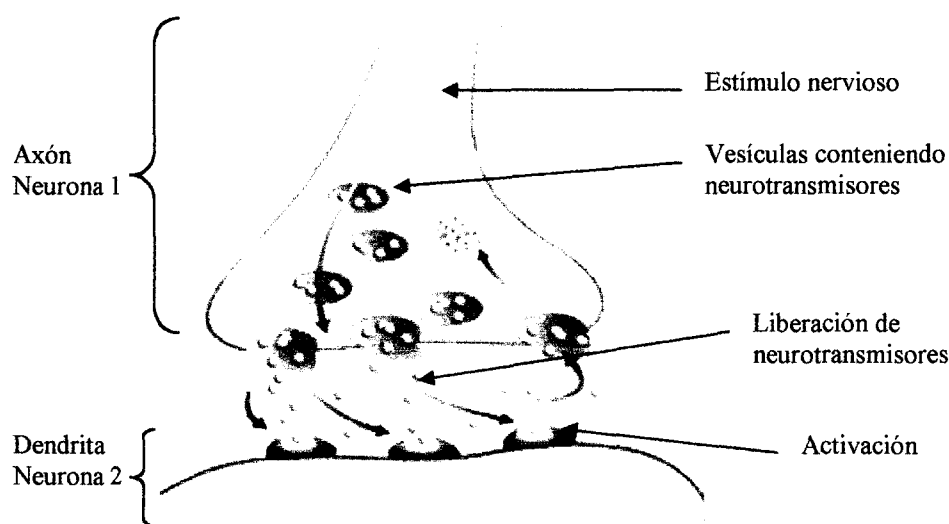


Figura 2.3. Transmisión sináptica.

Existen dos tipos de sinapsis:

- sinapsis excitadoras: los neurotransmisores provocan disminuciones de potencial en la membrana de la célula postsináptica, facilitando la generación de impulsos de mayor velocidad.

- sinapsis inhibitoras: los neurotransmisores tienden a estabilizar el potencial de la membrana, dificultando la emisión de impulsos.

Para establecer una similitud entre la actividad sináptica y las redes neuronales artificiales, nos fijaremos en los siguientes aspectos: las señales que llegan a la sinapsis son las entradas a la neurona; estas entradas son ponderadas (atenuadas o amplificadas), a través de un parámetro denominado peso; estas señales de entrada pueden excitar a la neurona (sinapsis con peso positivo) o inhibirla (peso negativo), siendo el efecto la suma de las entradas ponderadas; si la suma es igual o mayor que el umbral de la neurona, la neurona se activa. Es una situación del *todo o nada*, es decir, cada neurona se activa o no se activa. Hay factores que dificultan la sinapsis como son la fatiga, deficiencia de oxígeno y fármacos como los anestésicos, en cambio hay factores que incrementan el grado de activación. Esta habilidad para ajustar señales es un mecanismo de aprendizaje.

2.2.3. RED NEURONAL ARTIFICIAL

Existen numerosas formas de definir las redes neuronales, desde definiciones genéricas y poco específicas a otras más detalladas. A continuación veremos algunos ejemplos:

- *Una nueva forma de computación, inspirada en modelos biológicos.*
- *Un modelo matemático compuesto por un gran número de elementos procesales organizados en niveles.*
- *Un sistema de computación hecho por un gran número de elementos simples, elementos de proceso muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas.*⁵¹

- *Son redes interconectadas masivamente en paralelo de elementos simples y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico.*⁵²

Las redes neuronales artificiales son modelos que intentan reproducir el comportamiento del cerebro, así, realiza una simplificación, averiguando cuales son los elementos más relevantes del sistema, bien porque la cantidad de la información es excesiva o porque es redundante. El procedimiento convencional utilizado para construir redes capaces de realizar una determinada tarea es una elección adecuada de sus características, más una estructura conveniente.

La unidad elemental de proceso de la red neuronal es la neurona. Estableciendo una relación con el funcionamiento neurobiológico anteriormente mencionado, podemos encontrar tres tipos de neuronas:

1. Aquellas que reciben estímulos externos, relacionadas con el aparato sensorial, que tomarán la información de entrada.
2. Dicha información se transmite a ciertos elementos internos que se ocupan de su procesado. Es en las sinapsis y neuronas correspondientes a este segundo nivel donde se genera cualquier tipo de representación interna de la información. Como no tienen relación directa con la información de entrada ni con la de salida, se denominan unidades ocultas.
3. Finalizado el período de procesado, la información llega a las unidades de salida, cuya misión es dar la respuesta del sistema.

La neurona artificial pretende mimetizar las características más importantes de la neurona biológica. Cada neurona i -ésima está caracterizada en cualquier instante por un valor numérico denominado estado de activación $a_i(t)$; asociado a cada unidad existe una

función de salida, f_i , que transforma el estado actual de activación en una señal de salida, y_i . Esta señal es enviada a través de canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se codifica de acuerdo con la sinapsis (el peso w_{ji}) asociada a cada uno de ellos según una determinada regla. Las señales moduladas que han llegado a la unidad j -ésima se combinan entre ellas, generando así la entrada total, Net_j .

$$Net_j = \sum_i y_i w_{ji} \quad (2.53)$$

Una función de activación F , determina el nuevo estado de activación $a_j(t+1)$ de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación $a_j(t)$.

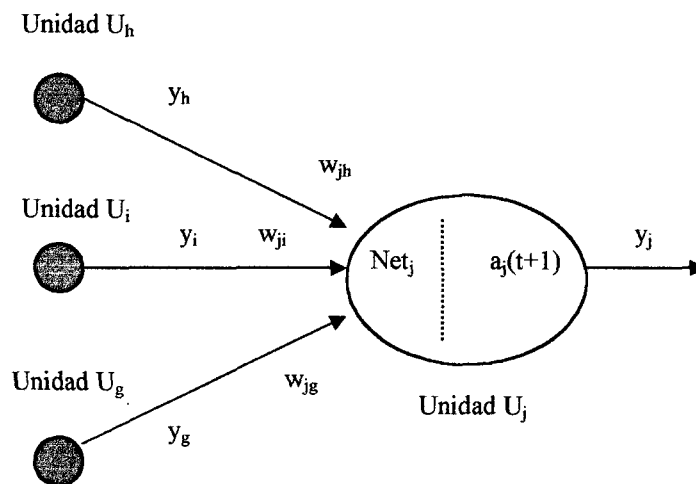


Figura 2.4. Entradas y salidas de una neurona U_j .

Se conoce como capa o nivel a un conjunto de neuronas cuyas entradas provienen de la misma fuente (que puede ser otra capa de neuronas) y cuyas salidas se dirigen hacia

el mismo destino (que puede ser otra capa de neuronas), de forma que encontramos redes monocapa y redes multicapa.

2.2.3.1. Estado de activación

La activación de una unidad U_i en el tiempo t se designa por $a_i(t)$.

Las neuronas que componen la red se hallan en un cierto estado, podemos decir que hay dos estados, reposo y excitado, a los que denominaremos estados de activación. Los valores de activación pueden ser continuos o discretos y limitados o ilimitados. Los discretos suelen tomar un conjunto pequeño de valores o valores binarios. Con notación binaria, un estado activo se representaría por un 1, mientras que un estado pasivo se representaría por un 0, y significaría que la neurona está en reposo. En otros modelos se considera un conjunto continuo de estados de activación, en lugar de sólo dos estados, por lo que se les asigna un valor entre $[0,1]$ o entre $[-1,1]$, siguiendo una función sigmoideal.

2.2.3.2. Función de salida

Entre las neuronas que forman la red neuronal artificial existe un conjunto de conexiones que unen unas con otras. Cada neurona transmite señales a aquellas que están conectadas con su salida, por lo que asociada a cada unidad U_i hay una función de salida $f_i(a_i(t))$, que transforma el estado actual de activación $a_i(t)$ en una señal de salida $y_i(t)$, es decir:

$$y_i(t) = f_i(a_i(t)) \quad (2.54)$$

Existen cuatro funciones de salida típicas que determinan distintos tipos de neuronas: función escalón, función lineal y mixta, sigmoideal y función gaussiana.

La función escalón se utiliza cuando las salidas de la red son binarias (dos posibles valores). La salida de una neurona sólo se activa cuando el estado de activación es mayor o igual que un cierto umbral (θ_i). La función lineal equivale a no aplicar función de salida, se usa muy poco. Las funciones mixta y sigmoideal son las más apropiadas cuando se quiere obtener como salida información analógica.

2.2.3.3. Conexiones entre neuronas

Las conexiones que unen a las neuronas que forman una red tienen asociado un peso, que es el que hace que la red adquiera conocimiento. Consideremos y_i como el valor de salida de una neurona i . Una neurona recibe un conjunto de señales que le dan información del estado de activación de las neuronas con las que se encuentra conectada. Cada conexión entre la neurona i y la neurona j está ponderada por un peso w_{ij} . La entrada neta que recibe una neurona (net_j) es la suma del producto de cada señal individual por el valor del peso de la conexión entre ambas:

$$net_j = \sum_i^N w_{ij} \cdot y_i \quad (2.55)$$

Esta regla muestra el procedimiento para combinar los valores de entrada a una neurona con los pesos de las conexiones que llegan a la misma, y se conoce como regla de propagación.

2.2.3.4. Función de activación

Al igual que es necesaria una regla que combine las entradas a una neurona con los pesos de las conexiones, también se necesita una que combine las entradas con el estado actual de la neurona para producir un nuevo estado de activación.

Dado un estado de activación $a_i(t)$ de la unidad U_i y la entrada total que llega a ella, Net_i , el estado de activación siguiente, $a_i(t+1)$, se obtiene aplicando una función F llamada función de activación.

$$a_i(t+1) = F(a_i(t), Net_i) \tag{2.56}$$

En la mayoría de los casos, F es una función identidad, por lo que el estado de activación de la neurona en $t+1$ coincidirá con el Net de la misma en t , y el parámetro que se le pasará a la función de salida, f , de la neurona será directamente el Net , por lo que consideraremos únicamente la función f , que denominaremos de activación.

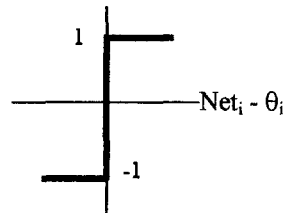
La salida que se obtiene en una neurona para las diferentes formas de la función f será:

a) Función de activación escalón

Si el conjunto de estados de activación es $E = [0,1]$:

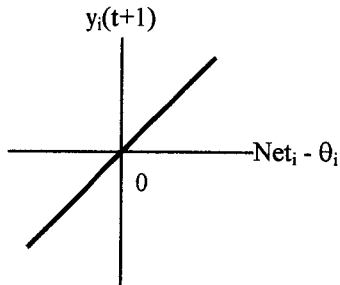
$$y_i(t+1) = \begin{cases} 1 & \text{si } [Net_i > \theta_i] \\ y(t) & \text{si } Net_i = \theta_i \\ 0 & \text{si } [Net_i < \theta_i] \end{cases} \tag{2.57}$$


Si el conjunto de estados de activación es $E = [-1,1]$:

$$y_i(t+1) = \begin{cases} +1 & \text{si } [Net_i > \theta_i] \\ y(t) & \text{si } Net_i = \theta_i \\ -1 & \text{si } [Net_i < \theta_i] \end{cases} \tag{2.58}$$


b) Función de activación lineal o identidad

El conjunto de los estados de activación puede contener cualquier número real; el estado de activación coincide con la entrada total que ha llegado a la neurona.

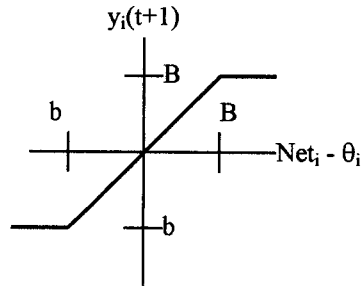


$$y_i(t+1) = Net_i - \theta_i \quad (2.59)$$

c) Función de activación lineal-mixta

El estado de activación de la unidad está obligado a permanecer dentro de un intervalo de valores reales prefijados.

$$y_i(t+1) = \begin{cases} b & Net_i \leq b + \theta_i \\ Net_i - \theta_i & b + \theta_i < Net_i < B + \theta_i \\ B & Net_i \geq B + \theta_i \end{cases}$$

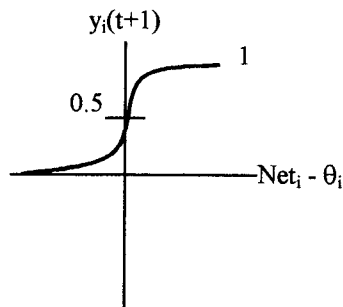


$$(2.60)$$

d) Función de activación sigmoideal

Es una función continua, por tanto el espacio de los estados de activación es un intervalo del eje real.

$$y_i(t+1) = \frac{1}{(1 + e^{-(Net_i - \theta_i)})}$$



$$(2.61)$$

2.2.3.5. Regla de aprendizaje

Una de las definiciones del concepto de aprendizaje podría ser: la modificación del comportamiento inducido por la interacción con el entorno y como resultado de experiencias conducente al establecimiento de nuevos modelos de respuesta a estímulos externos.⁵³

En el caso de las redes neuronales artificiales, se puede considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas, todo proceso de aprendizaje implica cierto número de cambios en dichas conexiones, y en realidad, puede decirse que se aprende modificando los valores de los pesos de la red.

2.2.3.6. Capas de neuronas y formas de conexión

La distribución de neuronas dentro de una red se realiza formando capas de un determinado número de neuronas cada una, de forma que, a partir de su situación dentro de la red, se pueden distinguir tres tipos de capas:

- Capa de entrada: es la capa que recibe directamente la información proveniente de las fuentes externas a la red.
- Capas ocultas: conjunto capas internas de la red que no tienen contacto directo con el exterior. El número de capas ocultas puede estar entre cero y un número elevado.
- Capa de salida: transfieren la información de la red hacia el exterior.

Las conexiones entre los nodos de una red neuronal se realizan de modo que las salidas de las neuronas están canalizadas para convertirse en entradas de otras neuronas. La señal de un nodo puede ser entrada de otro elemento de proceso, o incluso ser conexión de sí mismo (conexión autorrecurrente).

Cuando ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes, la red se describe como de propagación hacia delante o *feedforward*. Cuando las salidas pueden ser conectadas como entradas de neuronas de niveles previos o del mismo nivel, incluyéndose ellas mismas, la red es de propagación hacia atrás o *feedback*.

La arquitectura de la red neuronal consiste en la organización y disposición de las neuronas en la red formando capas o agrupaciones de neuronas más o menos alejadas de la entrada y salida de la red. Como consecuencia serán parámetros fundamentales de la red: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

2.2.4. EL PERCEPTRÓN

Fue el primer modelo de red neuronal artificial desarrollado por Rosenblatt⁴⁹ en 1958, el cual despertó un gran interés durante los años 60 debido a su gran capacidad para aprender a reconocer patrones sencillos: un Perceptrón, formado por varias neuronas lineales para recibir las entradas a la red y una neurona de salida, es capaz de decidir cuándo una entrada presentada a la red pertenece a una de las dos clases que esta es capaz de reconocer.

Pertenece al grupo de redes con conexiones hacia delante (*feedforward*), es decir, presenta arquitectura en niveles o capas y conexiones estrictamente hacia delante entre las neuronas, no existen conexiones hacia atrás, ni autorrecurrentes, ni laterales.

La única neurona de salida del Perceptrón realiza la suma ponderada de las entradas, resta el umbral y pasa el resultado a una función de activación de tipo escalón. Si el patrón presentado pertenece a la clase A la red responde con un +1, si pertenece a la

clase B responde con -1. La salida dependerá de la entrada (suma de las entradas ponderadas) y del valor umbral.

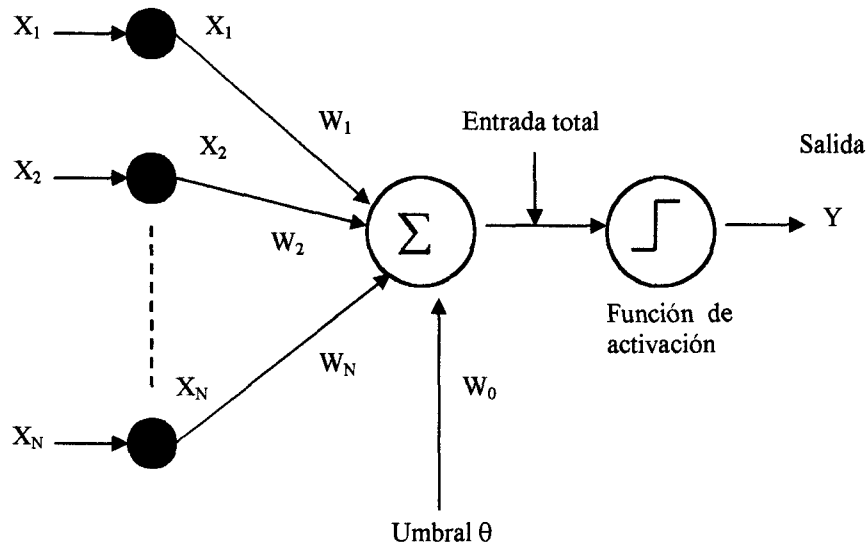
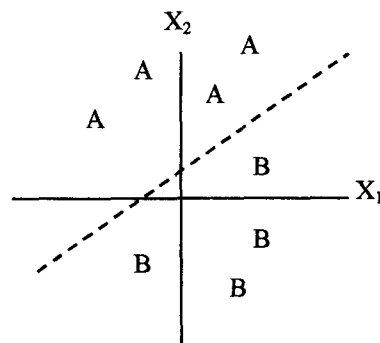


Figura 2.5. El Perceptrón

Para analizar el comportamiento de las redes, se representa en un mapa las regiones de decisión creadas en el espacio multidimensional de entradas a la red, regiones en las que se observa qué patrones pertenecen a una y otra clase. El Perceptrón separa las regiones mediante un hiperplano separador cuya ecuación queda determinada por los pesos de las conexiones y el valor umbral de la función de activación de la neurona. Los valores de los pesos pueden fijarse o adaptarse utilizando diferentes algoritmos de entrenamiento de la red.

$$y = f \left[\sum_{i=1}^N W_i \cdot X_i - \theta \right]$$



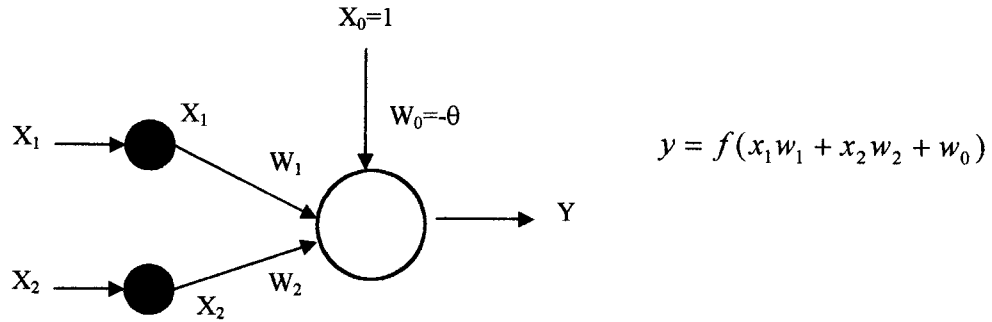
El Perceptrón, al tener sólo una capa de entrada y otra de salida con una única neurona, su capacidad de representación es bastante limitada, es capaz de discriminar patrones muy sencillos, pero presenta problemas para representar la función OR-EXCLUSIVA (XOR). La separabilidad lineal limita a las redes con sólo dos capas a la resolución de problemas en los cuales el conjunto de puntos correspondientes a los valores de entrada sean separables geoméricamente. Para dos entradas la separación se realiza mediante una línea recta, para tres entradas mediante un plano en el espacio tridimensional, y así sucesivamente hasta el caso de N entradas, en el cual el espacio N-dimensional es dividido en un hiperplano.

2.2.4.1. Regla de aprendizaje del Perceptrón

El algoritmo de aprendizaje del Perceptrón es de tipo supervisado, por lo que se requiere que los resultados sean evaluados y se realicen las oportunas modificaciones.

Los valores de los pesos, pueden determinar el funcionamiento de la red, los cuales se pueden adaptar o fijar utilizando diferentes algoritmos de entrenamiento de la red. Por desgracia no se puede aprender a realizar todo tipo de clasificaciones lo cual se debe a que un Perceptrón usa un separador lineal como célula de decisión, por lo que sólo se puede realizar una separación por medio de un hiperplano, es decir, es capaz de abordar aquellos problemas de clasificación que son linealmente separables.

Como ejemplo veamos la resolución de las funciones lógicas OR (O) y AND (Y).⁵⁴



x_1, x_2 : entradas a la neurona, que pueden adoptar los valores 1 ó 0.

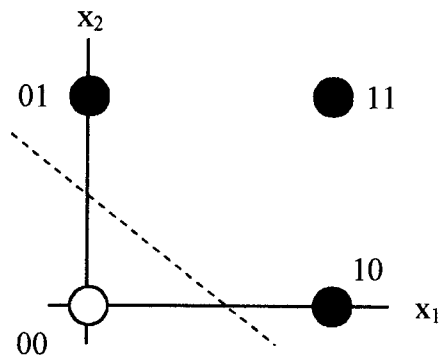
w_1, w_2 : pesos entre las neuronas de la capa de entrada y la de salida

f : función escalón

- Función OR (O):

Si $x_1 w_1 + x_2 w_2$ es mayor que el umbral (0), la salida es 1; en caso contrario la salida es 0.

x_1	x_2	salida
0	0	0
0	1	1
1	0	1
1	1	1

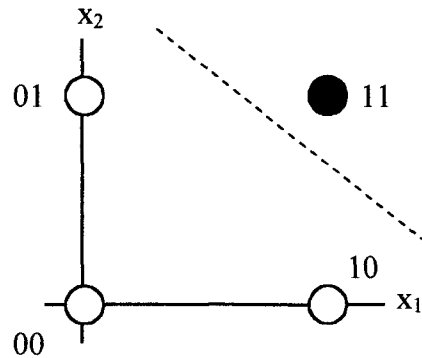


Durante la etapa de aprendizaje se irán modificando los pesos de las conexiones con el objeto de encontrar una recta que divida el plano en dos regiones, separando los valores 01, 10 y 11 del valor 00.

- Función AND (Y):

Para resolver el problema de la función AND se procede de igual forma, obteniendo una recta que sea capaz de separar los valores 00, 01 y 10 de 11.

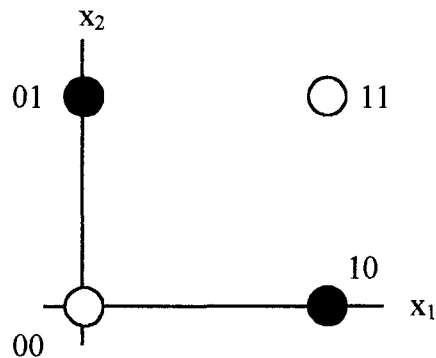
x_1	x_2	salida
0	0	0
0	1	0
1	0	0
1	1	1



- Función OR-EXCLUSIVA (XOR):

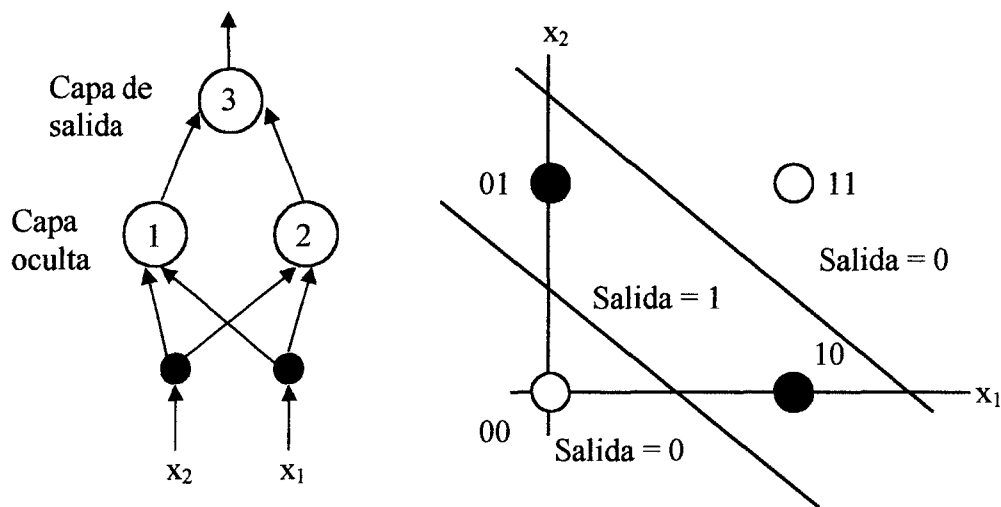
En este caso se pretende que para los valores de entrada 00 y 11 se devuelva clase 0 y para los valores 01 y 10, la clase 1. Como puede comprobarse el problema radica en que no existe ninguna recta que sea capaz de separar los patrones de una clase y de otra.

x_1	x_2	salida
0	0	0
0	1	1
1	0	1
1	1	0



La solución a este problema sería descomponer el espacio en tres regiones, de forma que una de las regiones perteneciese a una clase de salida y las otras dos regiones a la segunda clase de salida. Si en lugar de utilizar una neurona de salida se utilizan dos, se

obtendrían dos rectas y así pueden delimitarse tres regiones, pero para ello es necesario introducir otra capa con una neurona cuyas entradas serán las salidas de las dos neuronas anteriores. Por lo tanto se debe utilizar una red de tres neuronas distribuidas en dos capas. Las dos regiones que contienen los valores 00 y 11 se asocian a una salida de la red nula, mientras que la región central se asocia a una salida de valor 1.



La existencia de esta capa oculta nos proporciona la capacidad de construir redes que puedan resolver problemas complejos, surgiendo así la idea del Perceptrón multicapa.

2.2.5. EL PERCEPTRÓN MULTICAPA

Un Perceptrón multicapa es una red de tipo *feedforward* compuesta de un conjunto de neuronas organizadas por varias capas entre la capa de entrada y la de salida. Cada capa está formada por neuronas de tipo Perceptrón. Esta red permite establecer regiones de decisión más complejas que las obtenidas mediante el Perceptrón de un solo nivel.^{55,56}

Cada neurona tiene asociado un conjunto de pesos. La información que llega a la primera capa se multiplica por un número determinado de pesos. Los pesos asociados a cada conexión distinguen unas neuronas de otras dentro de la misma capa, y dentro de la misma capa todas procesan la misma información de entrada.

En la segunda capa, cada neurona tiene asociado un vector de pesos, comparten la salida de la primera capa, que, aplicando la correspondiente función de activación dan una nueva salida de información.

Solo se le puede suministrar información a la primera capa ($l = 0$), que realmente no es capa si no que es la información numérica que queremos tratar. A la segunda capa ($l = 1$), se le llama capa oculta porque la información que sale de esta no se puede estudiar, sólo se puede estudiar la información que sale de la capa siguiente.

El Perceptrón multicapa se dice que tiene tres capas, pero como la primera es la información que le suministramos, realmente tiene dos capas. El nombre menos ambiguo que se utiliza es el de Perceptrón de una capa oculta.

En la figura 2.6 se representa un Perceptrón multicapa, donde l representa a la capa de neuronas, w_{ij}^l corresponde a los pesos de las conexiones entre la capa 0 y la capa 1, siendo i la neurona destino y j la neurona de origen.

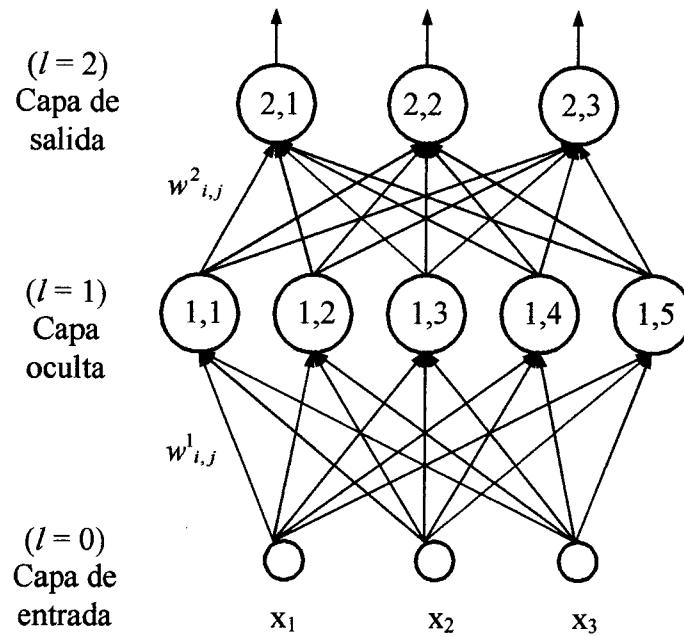


Figura 2.6. Perceptrón multicapa.

De forma general el comportamiento de una red multicapa se basa en los siguientes puntos:

1. Las actividades de las unidades de entrada son evaluadas de acuerdo con el problema que se quiere resolver.
2. Las señales se propagan de la capa de entrada a la siguiente, de forma que cada unidad realiza una suma ponderada de sus entradas y aplica una función de activación.
3. El proceso se repite en las sucesivas capas hasta que se alcance la capa de salida.
4. Los valores de la salida representan la respuesta de la red a las entradas que recibe.

El significado preciso de los valores de entrada y salida depende de la aplicación para la que fue diseñada la red, y en general, cualquier número de entradas y de salidas puede ser especificado.

La parte más importante de la modulación de la red neuronal es determinar formas para ajustar los pesos, de modo que la red realice un buen ajuste entre entradas y salidas, es lo que se denomina aprendizaje o entrenamiento de la red.

Las redes multicapa usan algoritmos de entrenamiento supervisados, donde el usuario determina qué salidas se esperan de las entradas de entrenamiento. Las muestras de entrenamiento consisten en un conjunto de entradas junto con la salida deseada. La regla de aprendizaje adaptará los pesos de las conexiones de tal modo que el error entre la salida real y la deseada sea mínimo.

La intención es que la red aprenda a generalizar, es decir, habiendo aprendido a clasificar correctamente las muestras de aprendizaje, si se presenta una nueva muestra, la clasificará correctamente también, ya que ha aprendido a reconocer las características que distinguen las diferentes clases de muestras.

También existen algoritmos de aprendizaje no supervisado, donde no hay un conjunto de entrenamiento y no se conocen las salidas deseadas, la red aprende a clasificar entradas basándose únicamente en su similitud con otras entradas. Este tipo de algoritmos son útiles cuando el usuario no conoce exactamente qué clases se desean. El modelo de Kohonen⁵⁷ es un ejemplo de red que utiliza aprendizaje no supervisado.

Existen numerosos algoritmos para llevar a cabo el aprendizaje de una red neuronal multicapa, pero dentro del aprendizaje supervisado uno de los más utilizados es el algoritmo de retropropagación del error (*backpropagation*).

2.2.5.1. Algoritmo de retropropagación del error (*backpropagation*)

En 1986 Rumelhart, Hinton y Williams⁵⁸, basándose en los trabajos de otros investigadores formalizaron un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada y las salidas correspondientes.

El funcionamiento de una red *backpropagation* consiste en un aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo, empleando un ciclo de propagación-adaptación de dos fases: primero se aplica un patrón de entrada como estímulo para la primera capa de neuronas de la red, la cual se va propagando a través de las capas superiores hasta generar una salida, se compara el resultado obtenido con el que se desea obtener y se calcula un valor del error para cada neurona de salida. Seguidamente estos errores se transmiten hacia atrás partiendo de la capa de salida, hacia las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original.

Este proceso se repite capa por capa hasta que todas las neuronas de la red reciban un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, los pesos de conexión de cada neurona se reajustan, de modo que en la siguiente vez que se presente el mismo patrón, la salida sea más próxima a la deseada, con lo cual el error disminuye.

Su importancia radica en la capacidad de adaptar los pesos de las neuronas de las capas intermedias para aprender la relación existente entre un conjunto de patrones dados como ejemplos y sus salidas, y lograr así la capacidad de generalización, es decir, la facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca durante la fase de entrenamiento.

La red debe encontrar una representación interna que le permita generar las salidas deseadas cuando se le dan las entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas según las características que compartan con los ejemplos de entrenamiento.

2.2.6. APLICACIONES DE LAS REDES NEURONALES

Las redes neuronales son una tecnología computacional emergente que puede utilizarse en un gran número y variedad de aplicaciones. Se pueden desarrollar redes neuronales en un período de tiempo razonable y pueden realizar tareas concretas mejor que otras tecnologías convencionales, incluyendo los sistemas expertos. A continuación detallamos algunas aplicaciones.

- Como clasificación de patrones:
 - Detección de fraudes en tarjetas de crédito.
 - Reconocimiento de caracteres impresos para fax.
 - Reconocimiento de caracteres manuscritos en cheques.
 - Reconocimiento del habla y traducción.
 - Control de calidad en la fabricación.
 - Exploración petrolífera.
 - Aplicaciones médicas: detección de células cancerígenas, malformaciones, prevención de problemas con la anestesia, modelado de la esquizofrenia, detección de los ataques de epilepsia, etc.

- Predicción:
 - Predicción financiera.
 - Ayuda a la toma de decisiones de préstamos.
 - Análisis de mercados.
 - Previsión de utilización de vuelos.
 - Evolución de los precios de los solares.
 - Predicción de fallos en motores eléctricos.

- Procesos de control y automatización:
 - Control de la posición de los electrodos en los hornos de arco.
 - Control de la concentración de dopado en semiconductores.
 - Control de procesos químicos.
 - Control en la fabricación de vehículos.
 - Control de fotocopiadoras.
 - Refinado petrolífero.
 - Fabricación de acero.

2.3. TERAPÉUTICA ANTIBACTERIANA

2.3.1. EVOLUCIÓN HISTÓRICA

La utilización de sustancias químicas para tratar procesos infecciosos se conoce desde hace muchos años. Sin embargo, hasta entrado el siglo XX los científicos no se interesaron con firmeza por la terapéutica antimicrobiana.

Pasteur y sus contemporáneos ya observaron fenómenos de antagonismo microbiano y la actividad antibacteriana de ciertos productos, como la sublutina y la piocianasa. Fue Paul Ehrlich (1854-1915), químico alemán quien se dio cuenta en 1882 de la selectividad de ciertos colorantes por determinadas células, entre ellas las microbianas, fue el descubridor del salvarsan, activo contra *Treponema pallidum*, y el creador del término quimioterapia, que significa tratamiento de las enfermedades infecciosas internas por agentes selectivos antiparásitos.

Más tarde, en 1935, Gerhard Domagk se percató del valor terapéutico de las sulfamidas, sintetizadas y patentadas en 1908 por Gelmo, químico de la Bayer. El producto con propiedades antimicrobianas había sido descubierto por Föster en 1923, se llamaba prontosil rojo y estaba formado por la unión de un colorante con una sulfamida. Contrariamente a lo que se suponía, el valor de la sustancia no radicaba en el colorante sino en la sulfamida como Trefonël lo demostró en ese mismo año.

Años más tarde los quimioterápicos fueron perdiendo importancia al ser superados en eficacia por los primeros antibióticos. Los antibióticos son sustancias producidas por microorganismos y activas contra otros microorganismos, este origen biológico los diferencia de los quimioterápicos, que lo tienen químico. El primer antibiótico aislado fue la actinomicetina, en 1924 por Gratia y Dath, aunque nunca fue utilizado en terapéutica.

En 1929 el médico escocés Alejandro Fleming descubrió accidentalmente la inhibición *in vitro* de una cepa de *Staphylococcus* por un hongo que era *Penicillium notatum*, pero por problemas de producción y purificación de la penicilina, le hicieron desistir de su utilización de momento en clínica.

René Dubos, en 1938 observó que el *Bacillus brevis* producía una sustancia que inhibía a los neumococos a la que denominó tirotricina. Un año después descubría dos subproductos la gramicidina y la triocidina, siendo los primeros antibióticos utilizados en clínica, aunque con reservas por su toxicidad.

En 1940 Ernst Boris Chain y Howard Walter Florey consiguieron extraer y purificar la penicilina y al iniciarse la Segunda Guerra Mundial trasladaron su equipo a Norteamérica, donde se desarrolló la producción de este antibiótico a escala industrial. A partir de este momento la búsqueda de antibióticos se generalizó, comenzando lo que podríamos denominar como la era antibiótica.

Waksman junto con investigadores como Schatz, Woodruff y Bugie, se impuso la tarea de buscar nuevos antibióticos efectivos contra las cepas de microorganismos resistentes a la penicilina, aislando a partir de 1939 la actinomicina, la clavacina, la fumigacina, la cetomina y la micromonosporina, ninguna de las cuales se pudo utilizar, las dos primeras por su toxicidad y las restantes por ser poco efectivas. En 1944 obtuvieron la estreptomicina a partir de *Streptomyces griseus*. Waksman y Lechevalier en 1949 aislaron la neomicina (aminoglicósido) a partir de *Streptomyces fradiae*.⁵⁹

Actualmente existe un grave aumento de resistencias en determinados microorganismos entre los que cabe destacar los géneros *Pseudomonas* y *Staphylococcus*, como consecuencia de un uso indiscriminado de los antibióticos.

2.3.2. CLASIFICACIÓN DE LOS ANTIBACTERIANOS⁶⁰

Los antibióticos son sustancias naturales, producidas por hongos y bacterias que tienen acción antimicrobiana, los agentes quimioterápicos son compuestos artificiales obtenidos mediante síntesis química con actividad antimicrobiana. Sin embargo, existen antibióticos sintéticos y otros modificados químicamente que se sitúan entre ambos grupos. Por ello es preferible utilizar el término antimicrobiano que es toda sustancia capaz de inhibir la proliferación o matar a un microorganismo infectante a una concentración no tóxica para el huésped.

La clasificación de los antimicrobianos se puede realizar atendiendo a diferentes propiedades como son: estructura química, toxicidad, espectro de acción, mecanismo de acción y uso terapéutico.

En la tabla 2.3 se muestra la clasificación de los antimicrobianos en función de su estructura química según la doceava edición de The Merck Index.⁶¹

Los antimicrobianos de un mismo grupo tienen un espectro de acción similar, toxicidad semejante y comparten los mismos mecanismos de resistencia, pero debido al gran número de modificaciones realizadas sobre las estructuras generales, se ha llegado a compuestos con propiedades farmacológicas muy variadas.

Dentro de la acción antimicrobiana, la reversibilidad del efecto nos permite distinguir entre bacteriostáticos, que son aquellos que a las concentraciones alcanzables en el hombre inhiben el crecimiento, pero el microorganismo sigue viable y cuando se retira el antimicrobiano reinicia su multiplicación (p. ej., fosfomicina, β -lactámicos, aminoglicósidos, etc.) y bactericidas los cuales poseen efecto irreversible, es decir, ejercen acción letal sobre los microorganismos (p. ej., tetraciclinas, macrólidos, etc.).

Tabla 2.3. Clasificación de los antimicrobianos.

Antimicrobianos	Grupo	Subgrupo	
Antibióticos	Aminoglicósidos		
	Anfenicoles		
	Ansamicinas		
	β -Lactámicos		Carbacefemas
			Carbapenemas
			Cefalosporinas
			Cefamicinas
			Monobactámicos
			Oxacefemas
			Penicilinas
		Otros	
		Lincosamidas	
		Macrólidos	
		Polipéptidos	
	Tetraciclinas		
	Otros		
Sintéticos	2,4-Diaminopiridinas		
	Nitrofuranos		
	Quinolonas		
	Sulfonamidas		
	Sulfonas		
	Otros		
Leprostáticos			
Tuberculostáticos			

Los antimicrobianos que solo ejercen su acción selectivamente sobre alguna especie o género se denominan de corto espectro de acción, mientras que los que son eficaces sobre un mayor rango se llaman de amplio espectro de acción.

El tipo de acción ejercida no solo dependerá de su estructura, lugar de acción, concentración, sino también de la especie microbiana, número de bacterias, tiempo de acción, fase de crecimiento del microorganismo y localización intra o extracelular.

El concepto de actividad antibacteriana exige una normalización o cuantificación que se consigue mediante los métodos utilizados *in vitro* para comprobar la susceptibilidad del microorganismo en relación con el antibiótico (antibiograma y técnicas de dilución).

Con estos métodos se define:

- Concentración Mínima Inhibitoria (CMI): es la menor concentración del antimicrobiano, capaz de inhibir el crecimiento de 10^5 bacterias en 1mL de medio de cultivo, tras 18-24 horas de incubación. Se expresa en $\mu\text{g/mL}$ o mg/L .
- Concentración Mínima Bactericida (CMB): es la menor concentración de antimicrobiano, capaz de destruir o matar (en 1mL de medio de cultivo) a 10^5 bacterias, tras 18-24 horas de incubación. Se expresa en $\mu\text{g/mL}$ o mg/L .
- Concentración Mínima Antibiótica (CMA): es la menor concentración del antimicrobiano capaz de producir *in vitro* alteraciones morfológicas y/o estructurales en una bacteria. Son concentraciones inferiores a la CMI, y se expresan de igual forma.

2.3.3. MECANISMOS DE ACCIÓN

Los agentes antimicrobianos también pueden estudiarse atendiendo a su mecanismo de acción, ya que el conocimiento de dicho mecanismo puede ser de gran importancia en la correcta selección del antimicrobiano o para establecer una combinación racional de los mismos. El conocimiento de los mecanismos de acción posibilita la comprensión del desarrollo y el mecanismo de resistencia bacteriana u otros factores importantes en la terapéutica antiinfecciosa. Los agentes antimicrobianos pueden ejercer su acción en una de las siguientes estructuras o funciones:⁶²

- Acción sobre la pared celular: Se han identificado varios lugares específicos para la inhibición de las biosíntesis de la pared celular, que son las dianas de varios agentes antimicrobianos.
 - Síntesis de precursores: Fosfomicina inhibe la enzima (transferasa), bloqueándola por lo que impide la formación de un precursor de la pared.
 - Transporte de membrana: Bacitracina bloquea el transporte de N-acetilmurámico hacia los lugares de ensamblaje.
 - Ensamblaje: Vancomicina bloquea la elongación del esqueleto del peptidoglicano.
- Acción sobre la membrana citoplasmática: La membrana es una barrera de permeabilidad selectiva que mantiene la integridad funcional de la célula; si es desorganizada, ésta perderá sus componentes como son proteínas, nucleótidos, etc., lo que llevará a la muerte celular. La polimixina provoca pérdida de solutos y fallos en la cadena respiratoria.
- Síntesis proteica: Los ribosomas son los orgánulos celulares que realizan la síntesis proteica y están formados por muchas y diferentes macromoléculas (ARN, proteínas) que interaccionan entre sí. La unión de un antimicrobiano a alguna de ellas podrá, por tanto, interferir con otras y de esta forma afectar varias funciones. Por ejemplo, tetraciclina, cloranfenicol y lincomicina, actúan sobre ribosomas que están realizando la elongación proteica; eritromicina y espectinomicina actúan sólo sobre ribosomas libres y los aminoglicósidos actúan con efectos diferentes sobre ambos tipos de ribosomas.
- Acción sobre ácidos nucleicos: Algunos actuarán como análogos estructurales de los nucleótidos (p. ej., citosín-arabinósido). Otros como la actinomicina y la cloroquina

actúan como agentes intercalantes que bloquean posteriores actividades del ADN. La rifampicina inhibe la acción de las ARN-polimerasas, impidiendo la síntesis de nuevo ARN. La novobiocina y ácido nalidíxico actúan sobre la ADN-girasa, impidiendo la replicación. Los derivados quinoleínicos inducen a la ADN-girasa a producir cortes en el ADN provocando su rápida destrucción.

- Análogos metabólicos: Sulfamidas, trimetoprim y diaminofenil-sulfona, que por diversos mecanismos como inhibición competitiva, incorporación biosintética o competición por los sistemas de transporte, consiguen bloquear el metabolismo o conducen a metabolitos alterados, inútiles para la bacteria.

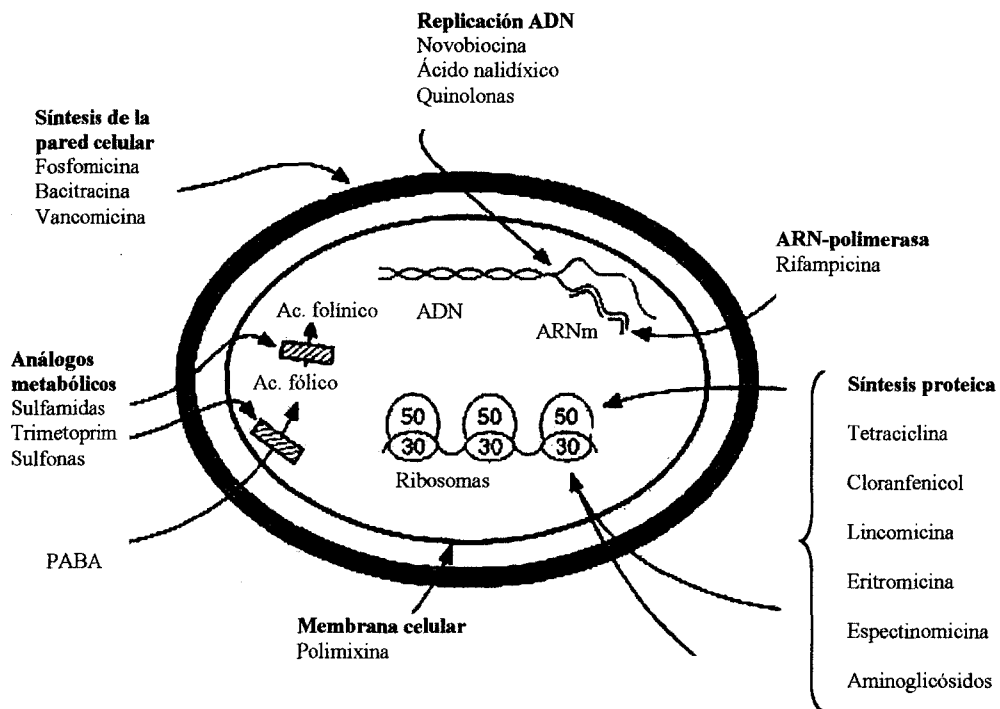


Figura 2.7. Lugares de acción de los diferentes tipos de agentes antibióticos.

(PABA, ácido *p*-aminobenzoico)

2.3.4. RESISTENCIA BACTERIANA

El concepto de resistencia bacteriana es un concepto relativo, que depende del microorganismo (su CMI) y del antimicrobiano (concentración que alcanza en el foco de infección). Definiremos una bacteria como resistente a un determinado antimicrobiano, cuando puede multiplicarse en presencia de las concentraciones que se alcanzan de éste en el lugar donde se está produciendo la infección.

Cuando en una bacteria no existe la diana (punto de acción) de un antimicrobiano, éste no podrá ejercer su acción y entonces hablamos de insensibilidad. En cambio, por resistencia se entiende cuando, poseyendo la diana, la bacteria ha sufrido alguna modificación o adquisición de nuevas propiedades que impiden la acción del antibacteriano.

Hay grupos bacterianos que no son afectados por un antibiótico, bien porque carecen del sitio de acción del antibiótico o porque es inaccesible, esta situación se define diciendo que la bacteria es insensible o presenta resistencia natural.

Otras especies son susceptibles al antibiótico, pero esto no impide que, por diferentes razones, se aíslen ocasionalmente variantes que no lo son y que crecen normalmente en presencia del antibiótico, en este caso se habla de resistencia adquirida.

Hablamos de resistencia cruzada cuando aparece resistencia simultánea a varios antibióticos de un mismo grupo que poseen estructura similar (resistencia cruzada homóloga) o antibióticos que tienen un mecanismo de acción parecido (resistencia cruzada heteróloga) o bien comparten el mismo sistema de transporte. La resistencia cruzada entre dos antibióticos puede ser recíproca, si la resistencia a uno entraña la resistencia a otro, y viceversa, o bien unidireccional si sólo se provoca en un sentido.^{63,64}

Hay dos mecanismos básicos de adquisición de la resistencia.⁶⁵

- Mutación cromosómica. Se produce una mutación espontánea en la información genética de la bacteria. La bacteria resistente (y sus descendientes) será seleccionada, si la población donde surgió es sometida a tratamiento antimicrobiano, encontrándonos al final del proceso con una población uniformemente resistente. Dentro de este tipo de resistencia, podemos encontrar a su vez dos modelos:
 - Resistencia a "un solo escalón", en la que una mutación en un solo *locus* es la responsable de que el microorganismo resista las concentraciones alcanzables por el antimicrobiano en el organismo humano (p. ej., estreptomicina, rifampicina o ácido nalidíxico).
 - Resistencia a "múltiples escalones", en la que para alcanzar la resistencia a dichas concentraciones la bacteria requiere mutaciones sucesivas en varios *loci*, cada una de las cuales le confiere resistencia a una concentración superior a la anterior (p. ej., eritromicina y gentamicina).
- Adquisición de nueva información genética por alguno de los mecanismos de transferencia genética: transformación, transducción, conjugación mediada por plásmidos⁶⁶ (material genético extracromosómico) o transposición, siendo los dos últimos los más frecuentes en la bacterias. Las bacterias donde se han detectado estos fenómenos de conjugación y transposición son *Pseudomonas*, *Yersinia*, *Haemophilus*, *Streptococcus*, *Clostridium*, etc. También es amplia la lista de antibióticos para los que se ha detectado resistencia mediada por plásmidos como betalactámicos, tetraciclina, fosfomicina, aminoglicósidos, sulfamidas, cloranfenicol, trimetoprim, etc.

En cuanto a la prevención de la resistencia, el número de mutaciones espontáneas no puede ser reducido por ningún mecanismo, pero sí podemos intentar prevenir su selección. Para la resistencia cromosómica a un escalón lo más eficaz para prevenirla es el uso simultáneo de dos fármacos sin resistencia cruzada, siendo especialmente efectivo para el caso de la tuberculosis. En el caso de una resistencia a múltiples escalones, la solución es mantener concentraciones farmacológicas suficientemente elevadas desde el inicio del tratamiento y de forma constante para inhibir el primer mutante.⁶⁷

Aquellos pacientes que presenten infecciones graves deben ser monitorizados, con objeto de controlar la aparición de microorganismos resistentes durante el tratamiento. Los antimicrobianos de amplio espectro que favorecen la selección de organismos resistentes deberían ser usados únicamente en situaciones muy concretas, a favor de otros que ejerzan menor presión selectiva. Se ha de evitar el uso inapropiado y masivo de los antibióticos, procurando tener el mejor conocimiento de los mecanismos de resistencia y de sus bases microbiológicas y genéticas, lo cual debe ser tenido muy en cuenta a la hora de determinar el uso de los antibióticos más apropiados en cada caso.

2.3.5. FARMACOCINÉTICA

Conociendo las propiedades farmacológicas de los diferentes antimicrobianos podremos conseguir unas concentraciones eficaces y no tóxicas en el foco de infección. El conocimiento de la farmacocinética permite al clínico establecer la vía de administración y el esquema de dosificación. Los antimicrobianos varían extraordinariamente en su absorción, concentraciones séricas, distribución en el organismo, metabolismo, acumulación y excreción. Asimismo, también es importante la forma galénica del producto, ya que pueden existir diferentes preparaciones de éste.

Absorción. Proceso por el que un fármaco pasa a estar disponible en los líquidos orgánicos, depende de la vía de administración utilizada y de las propiedades del antimicrobiano. En la administración oral, si la absorción es rápida, se consiguen concentraciones séricas más elevadas que si la absorción es lenta. La forma del preparado también influye en la velocidad de absorción, así, las soluciones lo hacen más rápidamente que las formas no disueltas. La absorción depende del grado de ionización de la molécula, pH y constante de disociación (pKa). Los alimentos pueden interferir en la absorción de los antimicrobianos, en cambio otros la favorecen. La vía parenteral consigue una mejor absorción y evita los problemas de absorción del tubo digestivo, pero presenta inconvenientes técnicos en cuanto a su administración, que sólo pueden resolverse si el enfermo está hospitalizado. La absorción por vía intramuscular y subcutánea depende de la solubilidad del preparado y la irrigación del lugar de inyección.

Biodisponibilidad. Es la eficacia relativa de la absorción de una dosis de prueba por vía oral o i.v. en comparación con un preparado control. Es importante ya que diversos preparados de un mismo antimicrobiano administrados por la misma vía pueden conseguir concentraciones séricas muy variadas.

Difusión sérica. Tras la absorción, el fármaco se distribuye por los líquidos orgánicos y tejidos hasta conseguir un equilibrio de distribución entre ellos. Esto se consigue tras un tiempo que varía en función de la capacidad de difusión del antimicrobiano.

Unión a proteínas plasmáticas. La unión de antibióticos a proteínas plasmáticas es un fenómeno reversible. La proporción en que se unen depende fundamentalmente de la naturaleza y concentración de la proteína y de la naturaleza del antibiótico. Los antibióticos se combinan casi exclusivamente con la albúmina sérica, siendo la porción libre, no unida a proteínas la que tiene actividad antimicrobiana. El desplazamiento del fármaco por sustancias que compitan con el sitio de unión de la albúmina es observable sólo cuando la concentración es superior a 100 mg/L. La unión del fármaco depende de la concentración y naturaleza de las proteínas. Desde el punto de vista terapéutico, la unión del antibiótico a proteínas tiene dos consecuencias importantes: el antibiótico unido no ejerce actividad antibacteriana y no es difusible, por lo que la actividad antibacteriana en el suero de antibióticos que tienen una alta afinidad por las proteínas está considerablemente disminuida.⁶⁸

Metabolismo. Los antimicrobianos pueden metabolizarse en proporciones muy variables. El tejido en que se suele producir la metabolización es el hepático, aunque a veces puede existir en riñón e intestino, y en este último se suele producir por las bacterias. Los productos resultantes de la oxidación, reducción, hidrólisis y conjugación suelen carecer de actividad antimicrobiana, aunque ocasionalmente, los productos metabólicos pueden tener una mayor toxicidad, como ocurre con las sulfamidas acetiladas. Los antimicrobianos se clasifican en tres grupos en función de su metabolización: los que no sufren prácticamente metabolización (p. ej. los aminoglicósidos), los que sufren una metabolización parcial (p. ej. los betalactámicos) y los que sufren una intensa metabolización (p. ej. el cloranfenicol).

Eliminación. La mayoría de los antibacterianos se excretan por vía renal mediante filtración glomerular y algunos también por secreción tubular. Unos pocos se excretan por

la bilis y heces como rifampicina, cefoperazona, ceftriaxona, etc. La reabsorción intestinal se puede producir para volver a eliminarse por bilis, constituyendo el llamado círculo enterohepático, que permite conseguir niveles mantenidos.

Vida media ($T_{1/2}$) de un fármaco es el tiempo que transcurre hasta que la concentración plasmática inicial se reduce a la mitad de su valor. Es un índice de tasa de eliminación y se utiliza para determinar el grado óptimo entre dosis; antimicrobianos con vida media corta deben administrarse a cortos intervalos, mientras que en los que la tienen prolongada los intervalos son mayores. La administración simultánea de otros fármacos puede alterar la vida media del antimicrobiano.

2.3.6. ELECCIÓN DEL ANTIMICROBIANO

Una correcta elección del antimicrobiano determina el éxito de la terapéutica. El antimicrobiano elegido debe combinar la mayor probabilidad de éxito con la menor toxicidad de tratamiento.

La elección del antimicrobiano no debe basarse sólo en el microorganismo causal y su sensibilidad, sino también en el mecanismo de acción y características farmacológicas del antimicrobiano, y finalmente en el tipo y localización de la infección y en las características del paciente: edad, estado de su función hepática y renal, es decir, para una correcta elección del antimicrobiano debemos basarnos en fundamentos microbiológicos, farmacológicos y clínicos, tal y como muestra la figura 2.8, en la que se representa la clásica interacción bidireccional de tres vías entre un patógeno microbiano, un agente antimicrobiano y un huésped cuya función inmune es un determinante fundamental en la aparición o no de la infección.

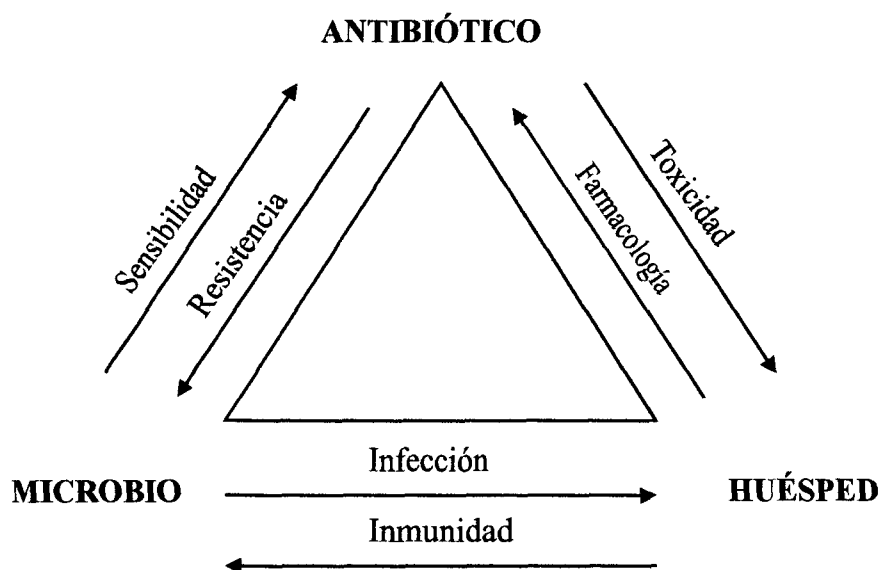


Figura 2.8. Triángulo de Davis.

Los métodos cuantitativos de antibiograma que nos ofrecen las concentraciones mínimas inhibitorias y bactericidas son preferibles a los métodos cualitativos, ya que los primeros facilitan la elección del antimicrobiano más eficaz y menos tóxico. La CMI puede ser suficiente para el tratamiento de la mayoría de las infecciones mientras que la CMB es necesaria para el tratamiento de endocarditis, osteomielitis, infecciones sobre prótesis e inmunodeprimidos.⁶⁹

Los valores de CMI y CMB son de poca utilidad para el clínico si éste no conoce los niveles de antimicrobiano que se alcanzan en suero, orina y líquidos orgánicos, por lo

que es necesario establecer una correlación de datos microbiológicos y farmacológicos, la cual se puede realizar de diferentes maneras y entre ellas destacamos:

- Categoría clínica: es la correlación entre CMI y niveles alcanzables del antimicrobiano en el lugar de infección y se manifiesta como sensible, intermedia o resistente.
- Cociente inhibitorio: expresa el número de veces que se consigue superar la CMI del microorganismo, en el lugar en el que se localiza la infección, con las dosis habituales por vía parenteral del antimicrobiano.

Como norma general se considera que un antimicrobiano es utilizable cuando los niveles superan la CMI de 4 a 8 veces, si embargo en ciertas infecciones como endocarditis y meningitis por gram-negativos, los niveles en sangre y líquido cefalorraquídeo deben ser al menos 8 veces la CMB del microorganismo.

Además de su utilización profiláctica, los antimicrobianos se emplean preferentemente en tres situaciones clínicas:

1. Cuando no se dispone de los resultados del cultivo, o no es posible o necesario el estudio bacteriológico: el tratamiento debe dirigirse contra la causa bacteriana más probable, lo cual es fácil cuando el cuadro clínico es típico y la etiología bien conocida, si la etiología puede ser diversa el tratamiento cubrirá los microorganismos más frecuentemente aislados.
2. Infecciones con compromiso vital: en estos enfermos graves se debe establecer con urgencia un tratamiento eficaz, de forma que los antimicrobianos que se administran inicialmente deben cubrir eficazmente todas las causas más probables de la infección.
3. Si se dispone de resultados microbiológicos y pruebas de sensibilidad: la mejor elección es el antimicrobiano más activo con el espectro antibacteriano más estrecho

posible, que alcance altas y mantenidas concentraciones en el foco de infección y con los menores efectos tóxicos.

3. MATERIAL Y MÉTODOS

3. MATERIAL Y MÉTODOS

3.1. ÍNDICES TOPOLÓGICO-ESTRUCTURALES

El presente trabajo utiliza como expresión de una estructura química un conjunto de índices topológico-estructurales desarrollados en nuestra unidad de investigación,^{70,71} que junto a la condición general de todos los índices topológicos de ser invariantes de grafo, poseen la ventaja de ser números enteros sencillos, los cuales nos permitirán establecer relaciones entre éstos y la propiedad farmacológica objeto de estudio que en nuestro caso es la actividad antibacteriana.

En este apartado veremos una nueva forma de representar las moléculas y de calcular sus índices topológico-estructurales.⁷²

3.1.1. OBTENCIÓN DE UNA ESTRUCTURA DE DATOS

La representación topológica más habitual de la estructura de una molécula se basa en la utilización de matrices, tanto la de distancia como la de adyacencia, como ya se vio en el apartado 2.1. La simetría que presentan ambas matrices se puede aprovechar para condensarlas en una sola matriz. A la matriz resultante se le denomina matriz combinada, de forma que su triángulo superior presenta la semi-matriz de adyacencia y el triángulo inferior la semi-matriz de distancias. Como ejemplo veremos la molécula de 2,3-dimetilbutano, a partir de la cual construiremos la matriz combinada.

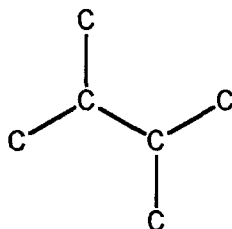


Figura 3.1. 2,3-dimetil-butano

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Matriz de Adyacencia

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 3 \\ 1 & 0 & 1 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 & 1 \\ 3 & 2 & 1 & 0 & 3 & 2 \\ 2 & 1 & 2 & 3 & 0 & 3 \\ 3 & 2 & 1 & 2 & 3 & 0 \end{pmatrix}$$

Matriz de Distancia

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 3 & 0 & 0 \\ 3 & 2 & 1 & 2 & 3 & 0 \end{pmatrix}$$

Matriz Combinada

El uso de matrices para la representación de grafos presenta un inconveniente en lo que a espacio en memoria se refiere, y aunque se intenta evitar mediante una combinación de las matrices, seguimos utilizando una gran cantidad de memoria.

Interesa conocer que el uso de memoria es de orden cuadrático con respecto al número de átomos, y debemos tener en cuenta que existen numerosas moléculas farmacológicamente activas con un número de átomos bastante elevado.

El hecho de combinar ambas matrices presenta algunas complicaciones a la hora de calcular los índices debido a que las filas y columnas ya no contienen la información del átomo que en ellas se encuentran, es decir, la información que nos proporciona una fila o columna ahora depende de la posición de esa fila o esa columna. Por ejemplo, todos los

elementos de la primera columna son valores de distancias, mientras que en la última columna encontramos sólo valores de adyacencia, y la tercera columna sus tres primeros valores son de adyacencia mientras que los tres siguientes son de distancias.

Por todo ello podemos afirmar que aunque la representación matricial de la molécula es sencilla, cuando se pretenden reducir los costes de memoria utilizando una matriz combinada, se pierde gran parte de esa sencillez.

Como solución se diseñará una estructura de datos que se asemeje lo más posible a la estructura real de una molécula (representación bidimensional). En la figura 3.2 observamos una representación bidimensional de la molécula de 2,3-dimetil-butano, donde apreciamos que existen una serie de átomos unidos por enlaces. Los enlaces pueden ser de varios tipos, pero un enlace siempre enlaza dos átomos. Los átomos pueden tener cualquier número de enlaces pero dos átomos no pueden estar enlazados entre sí por más de un enlace.

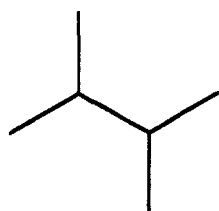


Figura 3.2. Grafo de hidrógenos suprimidos de la molécula de 2,3-dimetil-butano.

La anterior descripción de la estructura de datos podemos representarla en notación UML (Unified Modeling Language)⁷³⁻⁷⁵ como:

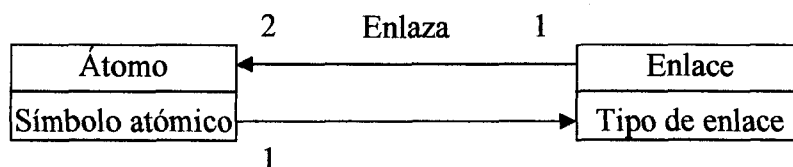


Figura 3.3. Notación UML.

Por lo tanto la molécula mostrada en la figura 3.2, quedaría de la forma siguiente:

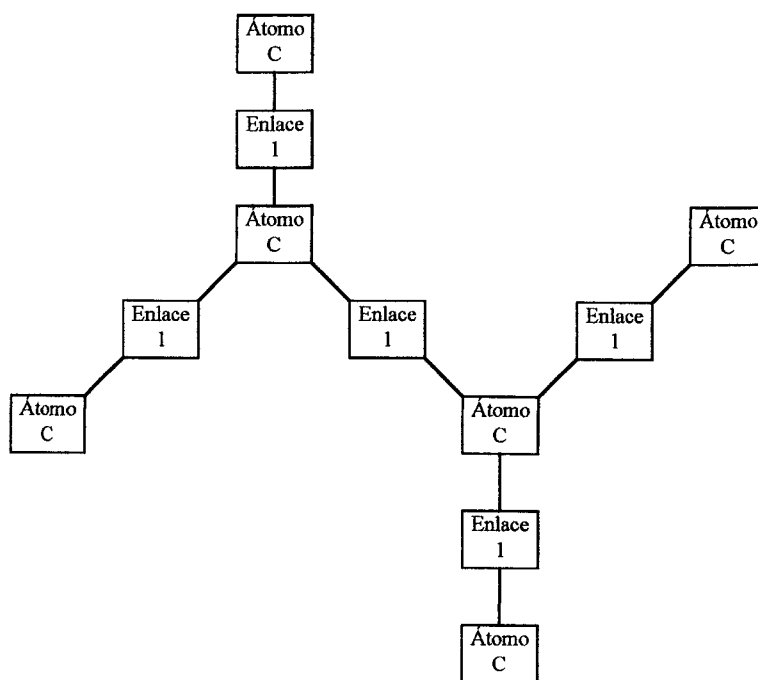


Figura 3.4. Representación UML de la molécula de 2,3-dimetil-butano.

Como podemos observar en la figura anterior, entre cada par de átomos siempre hay un único enlace con un valor que indica el tipo de enlace, y cada átomo puede tener varios enlaces asociados a él. Así la molécula queda almacenada como una estructura de datos en forma de grafo que tiene un coste de memoria lineal con respecto al número de átomos y enlaces de la molécula, es decir, $n + l$ donde n es el número de átomos y l el número de enlaces.

Hemos pasado de un coste de memoria de orden cuadrático a un coste lineal, ahorrando así memoria, lo cual es de vital importancia ya que se van a manejar un número elevado de moléculas en memoria.

3.1.2. CONSTRUCCIÓN DEL GRAFO A PARTIR DEL FORMATO SMILES

Establecida la estructura de datos que vamos a utilizar para representar las moléculas, el siguiente paso es diseñar el algoritmo que nos permita obtener esta estructura a partir de la descripción de la molécula en formato SMILES.

3.1.2.1. Formato SMILES (Simplified Molecular Input File Entry System).⁷⁶

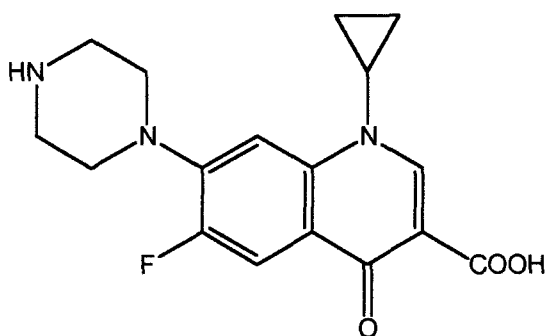
Este formato es un sistema de notación química diseñado para el procesado actual de información química, de modo que basándose en los principios de la teoría de grafos permite una rigurosa especificación estructural mediante el uso de un sencillo vocabulario. El sistema de notación SMILES es el más adecuado para conseguir una alta velocidad de procesamiento.

La notación SMILES nos permite representar la estructura de una molécula mediante una cadena de caracteres de acuerdo con las reglas que a continuación se detallan:

- Átomos: se representan por su símbolo atómico utilizando mayúsculas para átomos en forma alifática; si el símbolo atómico de un átomo se compone de dos letras, la segunda de ellas será minúscula; los átomos de hidrógeno se suelen omitir.
- Enlaces: los enlaces simple, doble y triple se representan mediante los símbolos -, = y # respectivamente, pero los símbolos de los enlaces simples normalmente se omiten.
- Ramificaciones: se especifican entre paréntesis, de modo que un paréntesis abierto significa el comienzo de la ramificación y el paréntesis cerrado su finalización.

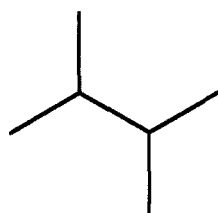
- Estructuras cíclicas: se escoge un enlace en el anillo, se numeran los átomos que une este enlace y queda una estructura sin ciclos la cual se puede expresar usando las reglas anteriormente descritas.
- Aromaticidad: los átomos de anillos aromáticos se pueden escribir con minúscula o también con mayúscula e indicando la posición del enlace correspondiente.

A continuación se detallan las moléculas de 2,3-dimetil-butano y de ciprofloxacino con sus correspondientes descripciones SMILES.



N1CCN(CC1)C2=CC3=C(C=C2F)C(=O)C(=CN3C4CC4)C(=O)O

Figura 3.5. Descripción SMILES de la molécula de ciprofloxacino.



CC(C)C(C)C

Figura 3.6. Descripción SMILES de la molécula de 2,3-dimetil-butano.

3.1.2.2. Obtención del grafo

El análisis de la descripción SMILES es el análisis de una cadena en la que se van identificando una serie de elementos que definen la estructura de la molécula gracias a una función. Esta función recorre la cadena SMILES y nos proporciona información relativa a los átomos y enlaces que va encontrando, así como información referente a ramificaciones y ciclos, irá obteniendo los elementos de la cadena y creando la estructura de grafo que represente a la molécula.

El proceso de creación del grafo es similar al de dibujo de la molécula partiendo del formato SMILES gracias a la similitud de nuestra estructura de datos con la estructura real de la molécula.

Una vez tenemos creado el grafo estamos en disposición de realizar el cálculo de índices.

3.1.3. CÁLCULO DE ÍNDICES

El conjunto de índices topológico-estructurales consta de 62 números enteros sencillos que nos proporcionan información referente a número y tipo de enlace, número y tipo de átomos, grados de vértice, dobles enlaces conjugados, distancias a cada tipo de átomos y distancias generales, por lo que quedan divididos en seis grandes bloques como muestra la tabla 3.1.

Tabla 3.1. Descripción de los 62 índices topológico-estructurales.

Índice	Descripción	Índice	Descripción
Índices de número y tipo de átomo		30	Sumatorio de distancias 6 de O (D^{O6})
1	Número de átomos de C (A^C)	31	Sumatorio de distancias 7 de O (D^{O7})
2	Número de átomos de N (A^N)	32	Sumatorio de distancias 8 de O (D^{O8})
3	Número de átomos de O (A^O)	33	Sumatorio de distancias 1 de S (D^{S1})
4	Número de átomos de S (A^S)	34	Sumatorio de distancias 2 de S (D^{S2})
5	Número de átomos de F (A^F)	35	Sumatorio de distancias 3 de S (D^{S3})
6	Número de átomos de Cl (A^{Cl})	36	Sumatorio de distancias 4 de S (D^{S4})
7	Número de átomos de Br (A^{Br})	37	Sumatorio de distancias 2 de F (D^{F2})
Índices de número y tipo de enlaces		38	Sumatorio de distancias 3 de F (D^{F3})
8	Número de enlaces simples (B^1)	39	Sumatorio de distancias 4 de F (D^{F4})
9	Número de enlaces dobles (B^2)	40	Sumatorio de distancias 5 de F (D^{F5})
10	Número de enlaces triples (B^3)	41	Sumatorio de distancias 6 de F (D^{F6})
Índices de grado de vértice		42	Sumatorio de distancias 2 de Cl (D^{Cl2})
11	Número de átomos con grado de vértice 1 (V^1)	43	Sumatorio de distancias 3 de Cl (D^{Cl3})
12	Número de átomos con grado de vértice 2 (V^2)	44	Sumatorio de distancias 4 de Cl (D^{Cl4})
13	Número de átomos con grado de vértice 3 (V^3)	45	Sumatorio de distancias 5 de Cl (D^{Cl5})
14	Número de átomos con grado de vértice 4 (V^4)	46	Sumatorio de distancias 6 de Cl (D^{Cl6})
Índices de dobles enlaces conjugados		47	Sumatorio de distancias 7 de Cl (D^{Cl7})
15	Número de dobles enlaces a distancia 1 (B^{D1})	48	Sumatorio de distancias 2 de Br (D^{Br2})
16	Número de dobles enlaces a distancia 2 (B^{D2})	49	Sumatorio de distancias 3 de Br (D^{Br3})
Índices de distancias a cada tipo de átomo		50	Sumatorio de distancias 4 de Br (D^{Br4})
17	Sumatorio de distancias 1 de N (D^{N1})	Índices de distancias generales	
18	Sumatorio de distancias 2 de N (D^{N2})	51	Sumatorio de distancias 1 (D^1)
19	Sumatorio de distancias 3 de N (D^{N3})	52	Sumatorio de distancias 2 (D^2)
20	Sumatorio de distancias 4 de N (D^{N4})	53	Sumatorio de distancias 3 (D^3)
21	Sumatorio de distancias 5 de N (D^{N5})	54	Sumatorio de distancias 4 (D^4)
22	Sumatorio de distancias 6 de N (D^{N6})	55	Sumatorio de distancias 5 (D^5)
23	Sumatorio de distancias 7 de N (D^{N7})	56	Sumatorio de distancias 6 (D^6)
24	Sumatorio de distancias 8 de N (D^{N8})	57	Sumatorio de distancias 7 (D^7)
25	Sumatorio de distancias 1 de O (D^{O1})	58	Sumatorio de distancias 8 (D^8)
26	Sumatorio de distancias 2 de O (D^{O2})	59	Sumatorio de distancias 9 (D^9)
27	Sumatorio de distancias 3 de O (D^{O3})	60	Sumatorio de distancias 10 (D^{10})
28	Sumatorio de distancias 4 de O (D^{O4})	61	Sumatorio de distancias 11 (D^{11})
29	Sumatorio de distancias 5 de O (D^{O5})	62	Sumatorio de distancias 12 (D^{12})

Aprovechando que el conjunto de índices queda dividido en seis grupos, vamos a utilizar funciones de cálculo de índices distintas para cada uno de estos grupos.

- *Índices de número y tipo de átomos*

Los índices de este grupo corresponden al número de átomos de un determinado tipo que se encuentran en la molécula, por lo que estos índices están más relacionados con la composición de la molécula que con su topología. Para su cálculo en lugar de utilizar

nuestro grafo utilizaremos la cadena SMILES, recorriéndola, buscando un átomo determinado y contando el número de veces que aparece en la misma.

Es una sencilla función que recorre la descripción SMILES e incrementa un contador cada vez que encuentra un elemento que coincide con el átomo que busca.

- *Índices de número y tipo de enlaces*

En este caso se aplica una función que vaya recorriendo el grafo e incremente un contador cada vez que encuentre un enlace del tipo que estamos buscando. Para realizar un recorrido completo del grafo se utiliza un algoritmo ampliamente documentado en la literatura informática que es el recorrido en profundidad o DFS (Depth First Search).⁷⁷ Al llegar a cada átomo se observa si hemos visitado sus enlaces, si no es así se comprueba el tipo de enlace, se incrementa el contador y se marca el enlace como visitado.

Para lograr una mayor efectividad en lugar de calcular el número de un único tipo de enlace se comprueba todos los tipos de enlace al llegar a cada enlace, es decir, llegamos a un enlace, vemos de qué tipo es e incrementamos la variable que cuenta este tipo de enlace.

De este modo con un único recorrido del grafo contamos el número de enlaces de los tres tipos (simple, doble o triple), en lugar de realizar tres recorridos del grafo. Esto se consigue utilizando dos funciones, una función general que buscará el número de enlaces de un tipo determinado y otra que en una única pasada contará los tres tipos de enlaces del grafo.

- *Índices de grado de vértice*

De igual forma que en el caso anterior se utiliza un recorrido DFS, pero en este caso al llegar cada átomo comprobará el número de enlaces que tiene asociados, incrementando el contador adecuado.

Para lograr una mayor eficiencia sin perder generalidad, se opera como en el caso anterior, es decir, se crea una función general que sólo contará los átomos con un grado de vértice determinado, y otra función que en cada paso comprobará el grado de vértice que tiene el átomo donde se encuentra e incrementará el contador adecuado.

- *Índices de dobles enlaces conjugados*

Estos índices representan las distancias entre los dobles enlaces de la molécula. Su cálculo se realiza mediante el diseño de una función general a la que se le pasa el tipo de enlace T y la distancia D , y esta reconocerá cuantos enlaces de tipo T se encuentran a distancia D .

En este caso se utilizan dos tipos de recorrido del grafo, uno que recorra el grafo en profundidad para encontrar los enlaces de un tipo en concreto y una vez encontrado se utiliza un recorrido en anchura BFS (Breadth First Search)⁷⁷ para reconocer los que se encuentren a una distancia determinada.

El recorrido en anchura se utiliza porque es un algoritmo que permite llegar de un nodo a otro por el camino más corto aunque existan varios caminos alternativos, teniendo en cuenta siempre el que se encuentre a menor distancia.

- *Índices de distancias a cada tipo de átomo*

Es una función semejante a la anterior pero en este caso nos fijamos en el tipo de átomo en lugar del de enlace, por lo que tendremos en cuenta dos parámetros, el tipo de átomo y la distancia.

El grafo se recorre primero en profundidad y cuando se llega al tipo de átomo que se desea, comienza un recorrido en anchura.

- *Índices de distancias generales*

Se utiliza una función idéntica a la anterior pero que no se fijará en el tipo de átomo, sino que calculará las distancias entre ellos, por lo que sólo tendrá en cuenta un

parámetro que será la distancia. Es un recorrido en anchura lanzado desde todos los átomos.

El esquema general seguido para el cálculo de los índices topológico-estructurales queda representado en la figura 3.7.

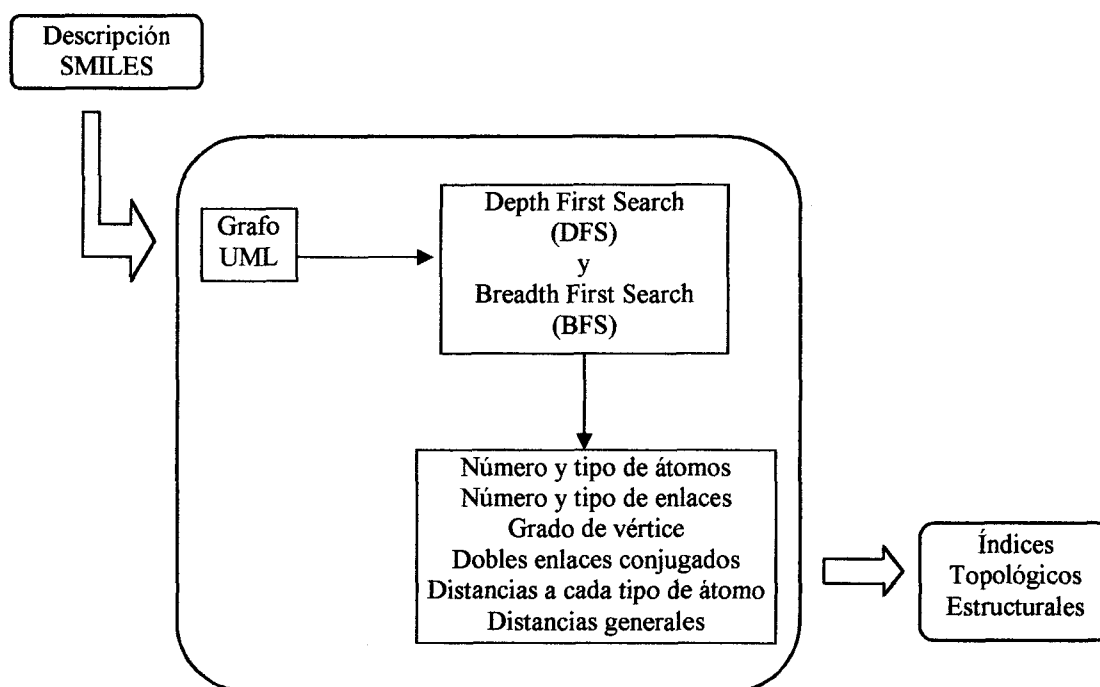


Figura 3.7. Esquema general del cálculo de índices topológico-estructurales.

Como ejemplo del cálculo de índices topológico-estructurales a continuación se muestra una tabla con los valores de los 62 índices para las moléculas 2,3-dimetil-butano y ciprofloxacino.

Tabla 3.2. Valores de los 62 índices topológico-estructurales de las moléculas 2,3-dimetilbutano y ciprofloxacino.

Nº	Índice	2,3-dimetilbutano	Ciprofloxacino	Nº	Índice	2,3-dimetilbutano	Ciprofloxacino
1	A ^C	6	17	32	D ^{O8}	0	4
2	A ^N	0	3	33	D ^{S1}	0	0
3	A ^O	0	3	34	D ^{S2}	0	0
4	A ^S	0	0	35	D ^{S3}	0	0
5	A ^F	0	1	36	D ^{S4}	0	0
6	A ^{Cl}	0	0	37	D ^{F2}	0	2
7	A ^{Br}	0	0	38	D ^{F3}	0	3
8	B ¹	5	21	39	D ^{F4}	0	4
9	B ²	0	6	40	D ^{F5}	0	5
10	B ³	0	0	41	D ^{F6}	0	4
11	V ¹	4	4	42	D ^{Cl2}	0	0
12	V ²	0	10	43	D ^{Cl3}	0	0
13	V ³	2	10	44	D ^{Cl4}	0	0
14	V ⁴	0	0	45	D ^{Cl5}	0	0
15	B ^{D1}	0	6	46	D ^{Cl6}	0	0
16	B ^{D2}	0	4	47	D ^{Cl7}	0	0
17	D ^{N1}	0	8	48	D ^{Br2}	0	0
18	D ^{N2}	0	11	49	D ^{Br3}	0	0
19	D ^{N3}	0	9	50	D ^{Br4}	0	0
20	D ^{N4}	0	8	51	D ¹	5	27
21	D ^{N5}	0	8	52	D ²	6	37
22	D ^{N6}	0	9	53	D ³	4	43
23	D ^{N7}	0	4	54	D ⁴	0	42
24	D ^{N8}	0	5	55	D ⁵	0	39
25	D ^{O1}	0	3	56	D ⁶	0	33
26	D ^{O2}	0	6	57	D ⁷	0	23
27	D ^{O3}	0	8	58	D ⁸	0	15
28	D ^{O4}	0	11	59	D ⁹	0	10
29	D ^{O5}	0	9	60	D ¹⁰	0	5
30	D ^{O6}	0	11	61	D ¹¹	0	2
31	D ^{O7}	0	6	62	D ¹²	0	0

3.2. TRATAMIENTO INFORMÁTICO

Una vez obtenidos los índices topológico-estructurales que describen la estructura de una molécula, el siguiente paso es establecer una relación entre éstos y la actividad o propiedades de la molécula.

Tradicionalmente los estudios QSAR se han desarrollado empleando métodos de regresión lineal^{6,78}. El éxito del análisis por regresión en modelos QSAR radica en la asunción de que las relaciones entre la actividad biológica y los descriptores son lineales, pero cuando el número de descriptores aumenta surgen problemas como redundancia en la información al correlacionar estos descriptores.⁷⁹ Para solucionar el problema se introdujeron técnicas como mínimos cuadrados parciales (PLS, Partial Least Squares), regresión multilineal (MLR, Multiple Linear Regression), análisis lineal discriminante o cuadrático (LDA, QDA), etc., los cuales constituyen una importante herramienta para reducir variables irrelevantes.

Sin embargo todos ellos sólo pueden establecer relaciones lineales, mientras que las redes neuronales artificiales (ANNs, Artificial Neural Networks) son capaces de reconocer relaciones no lineales, lo cual es de vital importancia teniendo en cuenta la cantidad de problemas de esta índole relacionados con el campo farmacéutico.⁸⁰

El uso de las redes neuronales ha experimentado un considerable incremento⁸¹ en el campo de la química desde que se describieron sus primeras aplicaciones en procesos de control⁸² y predicción de la estructura secundaria en proteínas,⁸³ extendiéndose durante la década de los 90 a campos como análisis espectral, desarrollo de productos farmacéuticos, clasificación de compuestos anticancerígenos, predicción de reactividad química, propiedades físicas, potenciales electrostáticos, potenciales de ionización así como análisis QSAR.⁸⁴⁻⁸⁶

Un aspecto interesante del presente trabajo es el empleo simultáneo de Análisis Lineal Discriminante (LDA) y Redes Neuronales Artificiales (ANN) como métodos de discriminación de la actividad farmacológica, en nuestro caso la antibacteriana.

3.2.1. RED NEURONAL ARTIFICIAL

En colaboración con el Departamento de Informática y Electrónica que imparte docencia en la Licenciatura en Ingeniería Informática de la Universidad de Valencia se desarrolló una interfaz gráfica⁷² en entorno Windows basada en el paquete SNNS (Stuttgart Neural Network Simulator),⁸⁷ cuyo objetivo era la creación de un entorno eficiente y flexible para el desarrollo y aplicación de redes neuronales.

El paquete SNNS en el que se basa nuestra interfaz gráfica es un simulador de redes neuronales desarrollado en el Instituto de Alto Desarrollo de Sistemas Distribuidos y Paralelos de la Universidad de Stuttgart (1989). Consta de cuatro componentes principales que son el núcleo o kernel del simulador, la interfaz gráfica de usuario, un lenguaje batch, un compilador de redes neuronales y otras herramientas de diversa índole.

La interfaz gráfica facilitará el diseño, aprendizaje y test de la red neuronal, nos permitirá realizar distintos diseños de la red neuronal, entrenar la red para que aprenda y analizar los resultados que ésta proporcione.

Se creó un entorno gráfico completo que permitía realizar todas las acciones necesarias para simular un experimento completo, es decir, cálculo de índices topológico-estructurales (anteriormente mencionado en el apartado 3.1), creación de redes neuronales (de discriminación o de predicción), y realización de experimento con las redes entrenadas. El lenguaje de programación utilizado es C++.

Por lo tanto en nuestra herramienta se pueden distinguir preferentemente tres casos de uso, que a continuación enumeramos:

- Cálculo de índices topológico-estructurales: se pueden presentar moléculas de varias formas (archivos con varias moléculas o archivos con una sola molécula).
- Creación de red neuronal: es el uso principal de la herramienta, en el que se deben elegir las características de la red, las moléculas participantes, distribuir las en sus respectivos patrones, entrenar la red y analizar los resultados.
- Creación de experimento: mediante redes entrenadas anteriormente se intentará predecir la actividad farmacológica de las nuevas moléculas, por lo que debemos elegir las moléculas, las redes neuronales, ejecutar el experimento y analizar los resultados.

3.2.1.1. Características de la red neuronal

En el presente trabajo de tesis doctoral se empleó una red neuronal de tipo perceptrón multicapa alimentada hacia delante y completamente conectada.

La red se describe como de propagación hacia delante o *feedforward* ya que ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes.

Arquitectura

La arquitectura de la red utilizada es perceptrón multicapa. La red se entrena con distintas unidades en la capa oculta y se analizan los resultados que se obtienen, por lo que en el apartado de resultados se mostrarán las distintas arquitecturas de la red escogidas y cuáles fueron los resultados obtenidos.

Aprendizaje

En cuanto al aprendizaje cabe decir que el perceptrón multicapa se basa en un aprendizaje supervisado, donde la misión de la red neuronal es proporcionar la salida deseada para cada entrada. Cuando se introduce la información de entrada (los 62 índices topológico-estructurales de las moléculas) el supervisor conoce cuál es la salida, por lo que la regla de aprendizaje consistirá en adaptar los pesos para conseguir que el error entre la salida real y la deseada sea mínimo.

Para que la red neuronal aprenda se escoge un conjunto de muestras de entrenamiento donde cada muestra contendrá la entrada y salida deseada. Durante la fase de entrenamiento los pesos sinápticos se adaptan para minimizar el error entre la salida deseada y la calculada. Esta fase se da por finalizada tras un número determinado de iteraciones o bien cuando el error está por debajo de un límite determinado.

Basándose en los pesos determinados durante la fase de entrenamiento o aprendizaje, la red será capaz de encontrar la salida deseada para una muestra de entrada no utilizada en la fase de aprendizaje, por lo tanto la red tendrá capacidad de generalizar el conocimiento adquirido.

Algoritmo de aprendizaje

El algoritmo de aprendizaje utilizado es el algoritmo de retropropagación del error o *backpropagation*, uno de los más extendidos en el aprendizaje supervisado.

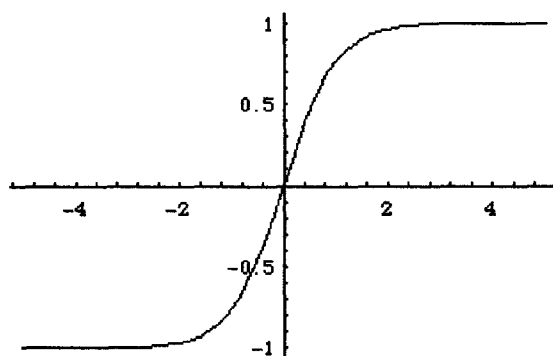
Se realiza en dos pasos, uno hacia adelante y otro hacia atrás. En el paso hacia adelante se presentan las muestras de entrada a la red, lo que provoca las activaciones de las neuronas de la primera capa. Las salidas de las neuronas de la primera capa constituyen las entradas de las neuronas de la segunda capa y así hasta llegar a la capa de salida.

Se compara la salida obtenida con la deseada y se calcula el error para las neuronas de salida, este error sirve para modificar los pesos de las neuronas de salida. Este error se propagará hacia atrás pasando por las capas ocultas hasta llegar a la capa de entrada. El algoritmo de aprendizaje utilizado es una modificación del algoritmo de retropropagación del error pero con mayor eficiencia y sin cambiar su funcionalidad, es el *standard backpropagation*.

Función de activación

Determina como va a ser la salida de una neurona y será la misma para todas las neuronas de la red multicapa.

Utilizando el algoritmo de retropropagación del error se debe utilizar una función continua y diferenciable como es la función tangente hiperbólica. Esta función proporciona una salida continua entre -1 y 1.



$$y = f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Figura 3.8. Función tangente hiperbólica

Elegida la arquitectura de la red, algoritmo de aprendizaje y función de activación nos encontramos en disposición de realizar la experimentación.

3.2.2. ANÁLISIS LINEAL DISCRIMINANTE

El Análisis Lineal Discriminante fue realizado mediante un programa incluido en un paquete informático denominado BMDP (Biomedical Computer Programs) en lenguaje FORTRAN, que se creó bajo la dirección de W. J. Dixon en la Universidad de California.⁸⁸

El primer manual de programas BMDP apareció en 1961, pero fue en 1968 cuando se comenzó a utilizar un tipo de lenguaje más flexible y fácil que el que hasta entonces se había utilizado, surgiendo así el paquete informático BMDP tal cual hoy lo conocemos.

Estos programas BMDP realizan un análisis de los datos que se les suministran, el cual puede variar desde una simple descripción de los mismos hasta las más sofisticadas y complejas técnicas estadísticas. Primeramente se procede a examinar los datos proporcionados con objeto de descubrir los posibles valores erróneos y posteriormente se realiza el análisis que pueda reflejar otras observaciones erróneas o la necesidad de que se sometan a un tratamiento posterior.

Los programas BMDP se identifican mediante un código de tres dígitos, el primer dígito es la letra P de Programa, el segundo es un número del 1 al 9 que indica el orden de creación del programa y el tercero hace referencia al tipo de programa que puede ser:

- D: descripción de datos
- F: tablas de frecuencia
- R: análisis de regresión
- V: análisis de varianza
- M: análisis multivariante
- L: tablas de vida y análisis de supervivencia
- S: mezcla de distintos programas
- T: series de tiempo

3.2.2.1. Programa P7M⁸⁹

Este programa permite realizar un análisis discriminante entre dos o más poblaciones de datos. Las variables empleadas en el cálculo de las funciones de clasificación lineal se eligen paso a paso, aunque también es posible realizar una selección previa o posterior a las mismas. En cada etapa se introduce en la función discriminante la variable que mejor diferencia los grupos, o bien se elimina aquella que menos los diferencia.

El programa evalúa el número de casos correctamente clasificados en cada grupo y permite realizar una validación de Jackknife para reducir las posibles desviaciones en la clasificación de los grupos.

En el resultado final se incluyen parámetros como la F estadística para variables seleccionadas, lambda de Wilks' o U estadística (con una F estadística aproximada), D^2 de Mahalanobis para las medias de los grupos; las funciones de clasificación, matriz de clasificación, clasificación de Jackknife y porcentaje de clasificación correcto; probabilidades posteriores y distancias de Mahalanobis para cada caso asignado dentro de cada grupo.

Dentro de los cálculos realizados se incluyen las medias, desviación estándar y coeficientes de variación de todas las variables en cada grupo y para todos los grupos.

La expresión de la desviación estándar dentro de cada grupo es:

$$s_p = \left(\frac{\sum (N_k - 1) s_k^2}{\sum (N_k - 1)} \right)^{1/2} \quad (3.1)$$

donde s_k^2 es la varianza del grupo K y N_k es el tamaño de muestra del grupo K .

La lambda de Wilks' o U estadística, es un análisis multivariable de la varianza estadística que comprueba la igualdad de las medias de grupo para la variable o variables de la función discriminante.

La F estadística aproximada es una transformación de la lambda de Wilks' que puede ser comparada con la F de distribución.

La matriz F contiene valores de F calculados a partir de la D^2 de Mahalanobis y comprueba la igualdad de las medias de grupo para cada par de grupos, sólo para las variables que se encuentran en la función discriminante.

Las funciones de clasificación para cada grupo son la combinación lineal de variables que mejor discriminan cada caso dentro de los grupos respectivos.

En la matriz de clasificación cada caso se incluye dentro de un grupo de acuerdo con las funciones de clasificación, nos indica el número de casos en cada grupo y el porcentaje de los casos correctamente clasificados.

La matriz de clasificación de Jackknife, en la cual cada caso se clasifica en su correspondiente grupo de acuerdo con las funciones de clasificación calculadas a partir de todos los datos a excepción del caso que se quiere clasificar.

Con los parámetros anteriormente mencionados se obtiene una tabla que resume cada etapa, incluyendo la F inicial para cada variable introducida (o la F final para cada variable eliminada), la lambda de Wilks' y la F estadística aproximada. Para cada caso se muestra la D de Mahalanobis de las medias del grupo.

Otra opción del programa es que calcule las funciones de clasificación para una parte de los grupos estudiados de modo que los datos queden clasificados en los grupos seleccionados. Estos datos pueden ser incluso nuevos, es decir, que no hayan participado en los cálculos de las funciones de clasificación.

Esta característica del programa permite realizar estudios de validación cruzada pudiendo dividir aleatoriamente los casos en cada grupo dentro de dos subgrupos. El primero de ellos se utiliza para estimar la función de clasificación, mientras que el segundo se clasifica de acuerdo con esta función. Observando el porcentaje de clasificaciones correctas para el segundo grupo se tiene una estimación empírica del éxito de la discriminación.

3.2.3. DIAGRAMAS DE DISTRIBUCIÓN FARMACOLÓGICA

Los diagramas de distribución farmacológica fueron introducidos por Gálvez *et al.*⁹⁰ en 1996 constituyendo una importante herramienta en la validación de propiedades limitantes y en la investigación de nuevos fármacos, siendo aplicados posteriormente en distintos grupos terapéuticos como hipolipemiantes,⁹¹ hipoglucemiantes,⁹² antivíricos,⁹³ antihistamínicos,⁹⁴ etc. Denominamos propiedades limitantes a aquellas propiedades que pueden ser utilizadas como funciones de clasificación.

Un diagrama de distribución farmacológica es un diagrama de distribución de frecuencias de variable dependiente, en el que la ordenada representa las espectancias de esta variable para cada intervalo de la función discriminante o valores de salida de la red neuronal.

Las espectancias se definen como la probabilidad de que un compuesto sea activo o inactivo para un determinado valor de la función discriminante o para un valor de salida de la red neuronal artificial. La espectancia para el grupo de moléculas activas (E_a) y para el grupo de moléculas inactivas (E_i) se calculan de acuerdo con las siguientes expresiones:

$$E_a = \frac{\text{Porcentaje de moléculas activas}}{\text{Porcentaje de moléculas inactivas} + 100} \quad (3.2)$$

$$E_i = \frac{\text{Porcentaje de moléculas inactivas}}{\text{Porcentaje de moléculas activas} + 100} \quad (3.3)$$

La principal ventaja de estos diagramas es que permiten determinar de forma visual aquellos intervalos de la propiedad en los que la probabilidad de encontrar nuevos compuestos activos sea máxima frente a la probabilidad de dar con un falso positivo.

En el presente trabajo utilizaremos estas representaciones tanto para los valores obtenidos mediante la función discriminante como para los valores de salida de la red neuronal artificial, comprobando así el mayor o menor grado de solapamiento entre el grupo de compuestos activos y el de inactivos.

3.3. ENSAYOS MICROBIOLÓGICOS

Los productos seleccionados como teóricos antimicrobianos son sometidos a ensayos microbiológicos con el fin de confirmar su actividad antibacteriana. A continuación se detalla el esquema de trabajo.

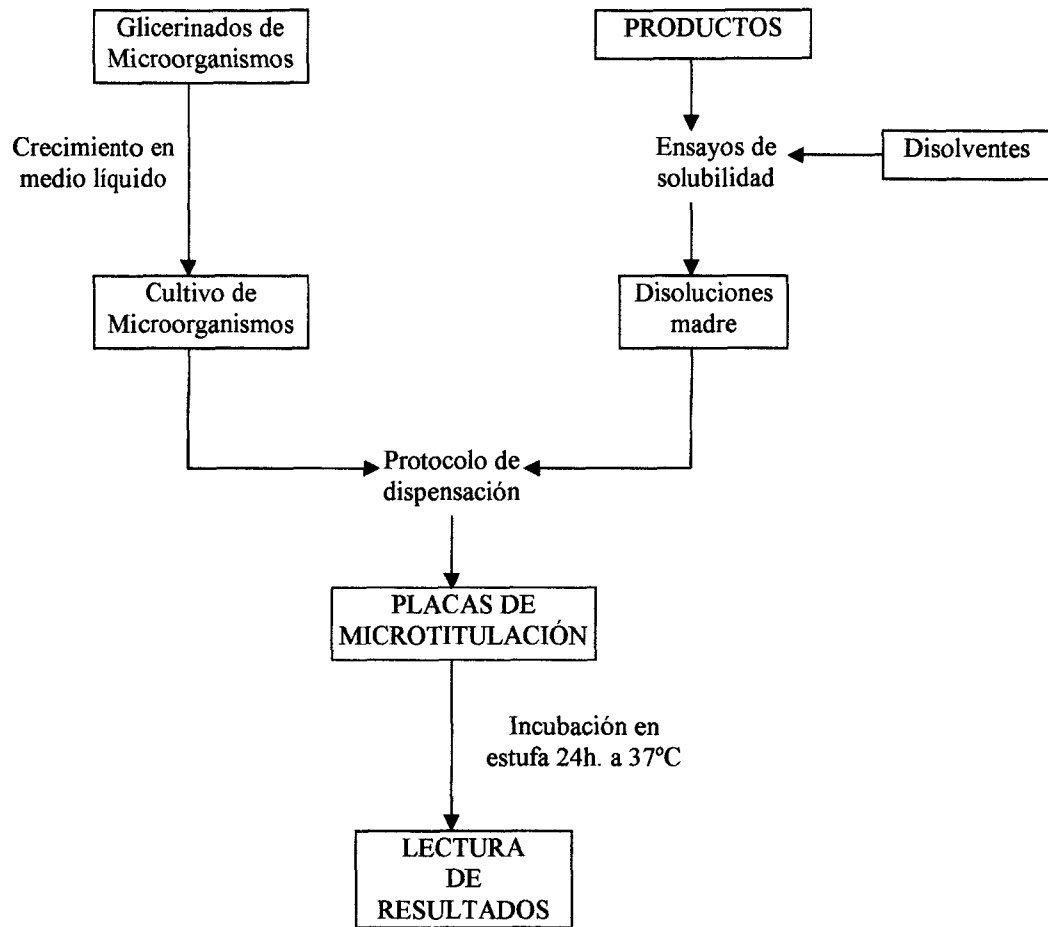


Figura 3.9. Esquema de trabajo.

3.3.1. MATERIAL Y REACTIVOS

a) Empleados en la preparación de las disoluciones de los productos:

- Balanza de precisión Mettler H31
- Baño de ultrasonidos P-Selecta Ultrasons-H

- Viales estériles de 20 mL
- Cabina U.V. P-Selecta
- Micropipetas Nichiryo Nichipet: 100-1000 μ L, 1000-5000 μ L
- Puntas estériles de 1 y 5 mL.
- Agua estéril
- Dimetilsulfóxido
- Tween 80
- Polietilenglicol 400

b) Empleados en la realización de los ensayos microbiológicos:

Medios de cultivo líquidos:

- Caldo TSB (Trypticase-Soja)

Material de microbiología:

- Placa de microtitulación estéril de 96 pocillos con tapa
- Pipetas Pasteur estériles y chupete de goma
- Micropipeta Socorex de 10-100 μ L
- Pipeta multicanal Transferpette-8 Brand
- Puntas estériles de 100 μ L
- Inóculo de los diferentes microorganismos
- Estándar de McFarland
- Mechero Bunsen
- Agua estéril
- Estufa de cultivo
- Nevera

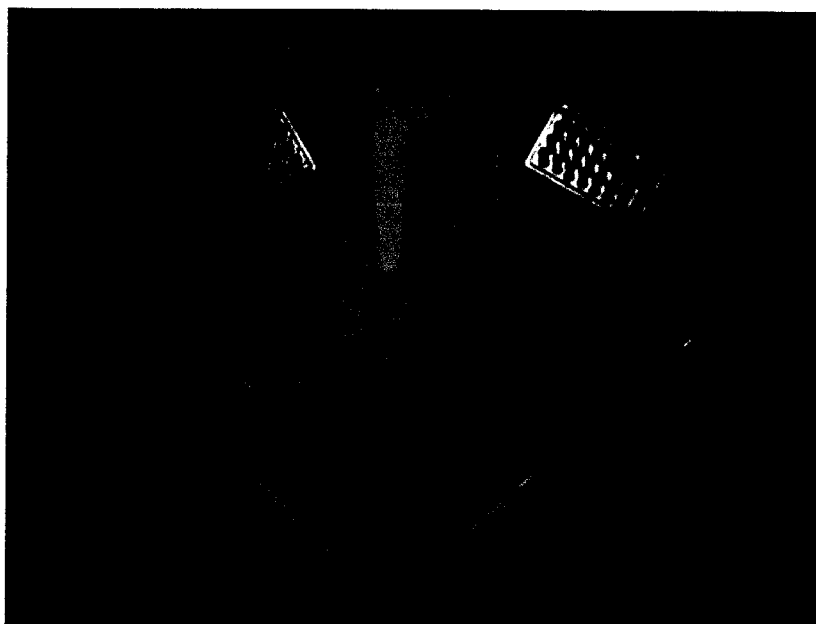


Figura 3.10. Micropipetas.

3.3.2. PREPARACIÓN DE LOS PRODUCTOS

Las disoluciones de los productos se preparan pesando cada uno de ellos en una balanza de precisión y añadiéndoles un determinado volumen de disolvente para conseguir la disolución madre de éstos así como de los productos de referencia. La concentración de las disoluciones madre es distinta dependiendo del producto (2000 $\mu\text{g/mL}$, 1000 $\mu\text{g/mL}$ y 500 $\mu\text{g/mL}$).

Se utilizaron distintos disolventes con el fin de conseguir una mejor disolución de los productos. Dichos disolventes fueron previamente ensayados comprobando que no inhibían el crecimiento de los microorganismos que posteriormente se utilizarían.

Los disolventes utilizados fueron:

- agua estéril
- mezcla de dimetilsulfóxido y agua estéril en proporción 1: 4.
- mezcla de Tween 80, polietilenglicol y agua estéril en proporción 1:1:8.

Una vez añadido el disolvente se agita el vial estéril que contiene el producto y el disolvente y se lleva a un baño de ultrasonidos termostataado a 37°C durante 15-30 minutos, con el propósito de mejorar la disolución de los productos.

3.3.3. CEPAS DE MICROORGANISMOS

Los microorganismos seleccionados para realizar los ensayos microbiológicos pretenden cubrir un amplio espectro bacteriano, incluyendo los microorganismos habitualmente recomendados para la realización de ensayos de susceptibilidad a antibacterianos.

Las cepas de estos microorganismos fueron proporcionadas por la Colección Española de Cultivos Tipo (CECT) dependiente del Departamento de Microbiología de la Facultad de Ciencias Biológicas de la Universidad de Valencia.

A continuación realizamos una descripción de las cepas utilizadas estableciendo dos grupos, microorganismos gram-positivos y gram-negativos.

- Gram-positivos

***Enterococcus faecalis* CECT 795 (ATCC [American Type Culture Collection] 29212)**

Son células ovoideas de menos de 2 μm , no esporulados, inmóviles y que aparecen dispuestas en parejas o en agrupaciones de cadenas cortas. Forman parte de la flora intestinal humana y de muchos animales de sangre caliente, por lo que son eliminados con las heces, pudiendo contaminar aguas y alimentos. Debido a su gran resistencia a la desecación, este microorganismo es utilizado como índice de contaminación fecal en aguas y en alimentos congelados.

Su principal consecuencia patógena son las infecciones del tracto urinario y endocarditis agudas.

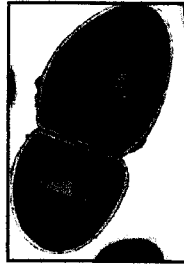


Figura 3.11. *E. faecalis*

***Staphylococcus aureus* CECT 240 (ATCC 6538P)**

Son cocos de 0.5 a 1 μm de diámetro, no esporulados ni capsulados, inmóviles y anaerobios facultativos. Aparecen agrupados en masa o en racimos, aunque en ocasiones también se presentan como elementos únicos, en pares, en tétradas o en cortas cadenas de tres o cuatro células. Esta especie se considera patógena o potencialmente patógena para el ser humano produciendo principalmente infecciones de la piel y mucosas, infecciones metastásicas, síndrome de la piel escaldada, síndrome del shock tóxico, toxiinfecciones alimentarias y enterocolitis estafilocócica.

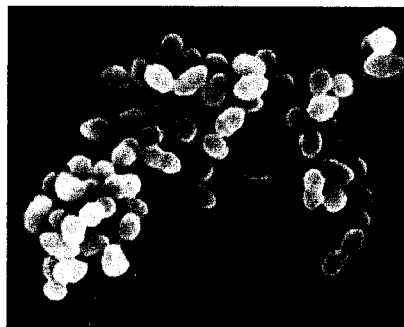


Figura 3.12. *S. aureus*

- Gram-negativos

***Escherichia coli* CECT 405 (ATCC 10536)**

Dentro de este género encontramos bacilos no esporulantes, no móviles o móviles por flagelos peritricos, aerobios facultativos, oxidasa negativos con requerimientos de nutrición relativamente simples. Son habitantes casi universales de las vías intestinales de los humanos y de los animales de sangre caliente. Las cepas enteropatógenicas de *E. coli* se han vuelto más frecuentes en relación a las infecciones disentéricas y las fiebres generalizadas. Estas cepas forman el antígeno K, que permite ataque y colonización del intestino delgado, y la enterotoxina responsable de los síntomas de la diarrea. También pueden ser causantes de infecciones urinarias, septicemias, colecistitis y peritonitis.



Figura 3.13. *E. coli*

***Pseudomonas aeruginosa* CECT 108 (ATCC 27853)**

Bacilos rectos o ligeramente curvos; tamaño de 0.5 - 1.0 μm por 1.5 - 4.0 μm ; móviles por uno o múltiples flagelos polares; sin esporas; metabolismo respiratorio, nunca fermentativo, aunque puede producir pequeñas cantidades de ácido a partir de glucosa aeróbicamente. Esta especie es la más difundida del género. Habita en el suelo y en las aguas, apareciendo en el 10% de muestras de heces humanas. Es un

patógeno oportunista capaz de infectar cualquier órgano o tejido del cuerpo humano produciendo otitis, sinusitis, queratitis, neumonías, enterocolitis, endocarditis, infecciones del tracto urinario, meningitis, etc. ya que se encuentra en el ambiente hospitalario y puede infectar con facilidad a pacientes que reciben tratamientos por otras enfermedades, así como en pacientes que reciben tratamiento de quemaduras graves. Presenta resistencia a numerosos antibióticos.



Figura 3.14. *P. aeruginosa*

3.3.3.1. Apertura de liófilos

Las cepas se suministraron en forma de liófilos, los cuales deben de ser rehidratados y comprobar su viabilidad antes de ser empleados en los ensayos.

El procedimiento comienza quitando la etiqueta identificativa del liófilo la cual debe conservarse para una correcta identificación de las cepas durante todo el proceso. A continuación se deben etiquetar tubos, placas, etc. Seguidamente se toma agua estéril con la pipeta Pasteur, se calienta la parte media del liófilo en la llama del mechero hasta que empiece a oscurecerse el algodón, se echa agua sobre esta zona (de la mitad del algodón hacia la parte estrecha del liófilo), con lo que el cristal se cuartea. Se devuelve la pipeta Pasteur a su envoltorio estéril. Con unas pinzas, previamente mojadas en alcohol y flameadas, se rompe con cuidado el liófilo de la mitad del algodón hacia la parte estrecha

intentando evitar la entrada brusca de aire al interior (que empujaría el algodón hacia el interior), y el liófilo así roto puede manejarse como un tubo cualquiera.

Por último se deben hidratar las células con agua o con el medio donde vayan a crecer y poner los cultivos a la temperatura, agitación, etc., adecuadas para la cepa.

3.3.3.2. Viabilidad de los liófilos

Una vez tenemos hidratados los liófilos, el siguiente paso es comprobar la viabilidad de los mismos, realizando el cultivo de los microorganismos mediante una siembra en medio líquido (caldo TSB) y en medio sólido (TSA).

Tras realizar la siembra en ambos medios partiendo de los liófilos rehidratados, se procede a su incubación en condiciones adecuadas de temperatura durante 24 horas, al cabo de las cuales, se observa la presencia de crecimiento en ambos medios y la uniformidad de colonias aisladas en el medio sólido. Si el crecimiento es adecuado, a partir de cualquiera de los dos medios, se puede tomar una muestra para obtener el inóculo necesario para el ensayo de susceptibilidad a antimicrobianos.

3.3.3.3. Glicerinado de microorganismos

El proceso de glicerinado es una técnica de conservación de cultivos que los mantiene viables durante largos períodos de tiempo y que además permite una rápida y cómoda accesibilidad a los mismos para resiembras posteriores.

En primer lugar se deben etiquetar dos viales eppendorf por cada una de las cepas, de forma que quede perfectamente identificado el nombre del microorganismo, su número de colección, la fecha del glicerinado y la letra A o B. A partir de los cultivos puros de los microorganismos en medio líquido se dispensa a cada vial 0.8 mL junto a 0.2

mL de glicerol, agitando para obtener una mezcla homogénea. Posteriormente los viales se almacenan a -70°C para su conservación.

Para revivir los microorganismos se descongela uno de los dos eppendorf el A o el B, se toma una alícuota del mismo y se siembra en medio líquido (TSB) o en medio sólido (TSA) y después de 24 h. de incubación se comprueba el crecimiento.

3.3.4. PROTOCOLO DE DISPENSACIÓN EN PLACAS DE MICROTITULACIÓN

El principio de la valoración microbiológica de los antimicrobianos se basa en la determinación de su actividad o potencia por comparación de la concentración a la cual se inhibe el crecimiento de un microorganismo susceptible y adecuado con la concentración de una preparación patrón del antimicrobiano que produce la misma inhibición.

Para cada uno de los microorganismos objeto de estudio se prepara una placa de microtitulación, así determinamos la sensibilidad frente a los distintos productos y medimos la CMI correspondiente a cada uno de ellos.

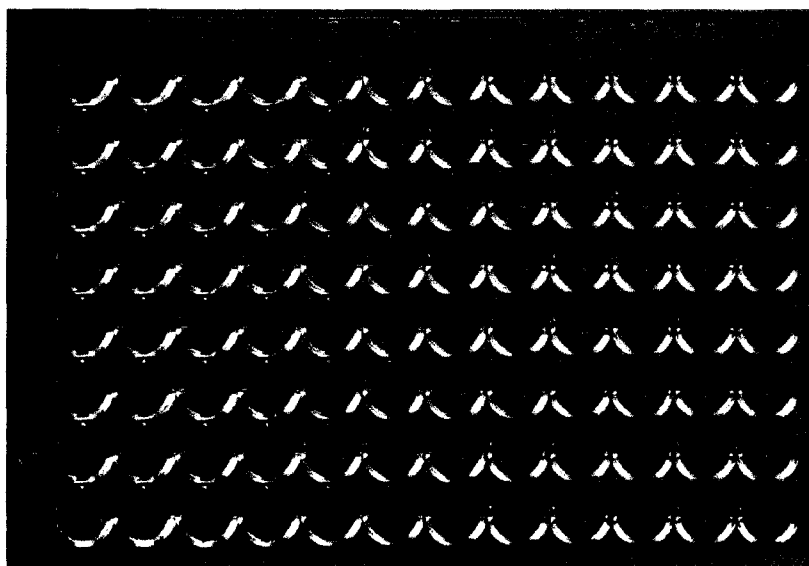


Figura 3.15. Placa de microtitulación.

La placa presenta 96 pocillos de 200 μL de capacidad cada uno dispuestos en 12 columnas y 8 filas, de modo que en las columnas de la 1 a la 9 se sitúan los productos que se van a ensayar, en las columnas 10 y 11 se sitúan productos de referencia como la cefalosporina C y el ácido nalidíxico y las filas de la A a la G contienen diluciones dobles de cada producto hasta alcanzar siete concentraciones diferentes por cada fila.

La mayoría de los productos comienzan en la fila A con una concentración de 1000 $\mu\text{g/mL}$ hasta llegar a una concentración de 15.7 $\mu\text{g/mL}$ mientras que otros comienzan con concentraciones de 500 $\mu\text{g/mL}$ y 250 $\mu\text{g/mL}$ y alcanzan a concentraciones en la fila G de 7.85 $\mu\text{g/mL}$ y 3.9 $\mu\text{g/mL}$ respectivamente.

En cuanto a los productos de referencia, para la cefalosporina C se cubren concentraciones de 500 a 7.85 $\mu\text{g/mL}$, mientras que para el ácido nalidíxico lo son de 250 a 3.9 $\mu\text{g/mL}$.

Los pocillos del H1 al H11 se utilizan como control negativo, que nos permiten determinar si durante el proceso de dispensación se ha producido algún tipo de contaminación por microorganismos. Estos pocillos contienen diluciones muy bajas de los productos que se están ensayando así como de los productos de referencia y caldo estéril. Si las disoluciones, el agua estéril o las placas hubieran sufrido algún tipo de contaminación microbiana se observaría crecimiento en el pocillo o pocillos correspondientes.

La columna 12 es utilizada para albergar los controles positivos de crecimiento, ya que sólo contienen agua estéril e inóculo de microorganismo.

El proceso de dispensación se realiza de la siguiente forma:

1. Se añaden 50 μL de agua estéril en los pocillos B1 a H12.
2. Se añaden 50 μL de disolución madre (2000, 1000 ó 500 $\mu\text{g/mL}$) de cada producto a probar en el pocillo correspondiente, del A1 al B9.

3. Se añaden 50 μL de disolución de los productos de referencia, 1000 $\mu\text{g}/\text{mL}$ para la cefalosporina C y 500 $\mu\text{g}/\text{mL}$ para el ácido nalidíxico en los pocillos A10, A11, B10 y B11.
4. Se mezcla el contenido de los pocillos de la fila B (B1 a B11), mediante la micropipeta multicanal y se pasan 50 μL a los pocillos de la fila C (C1 a C11). Esta operación se repite de la misma forma para las siguientes filas, de la C a la D, de la D a la E,..., de la G a la H y de la fila H una vez mezclado el contenido de los pocillos se toman 50 μL y se desechan.
5. Se añaden 50 μL de inóculo en los pocillos A1 a G12.
6. Se añaden 50 μL de caldo estéril a los pocillos H1 a H11.

Una vez se ha completado la dispensación se procede a la incubación de la placa en estufa de cultivo a 37°C y durante 24 horas. Transcurrido este tiempo se procede a la lectura de los resultados, obteniendo así los valores de Concentración Mínima Inhibitoria (CMI) de cada producto frente al microorganismo utilizado.

A continuación la figura 3.16 muestra un ejemplo de las distintas concentraciones resultantes tras la realización del proceso de diluciones seriadas antes mencionado.

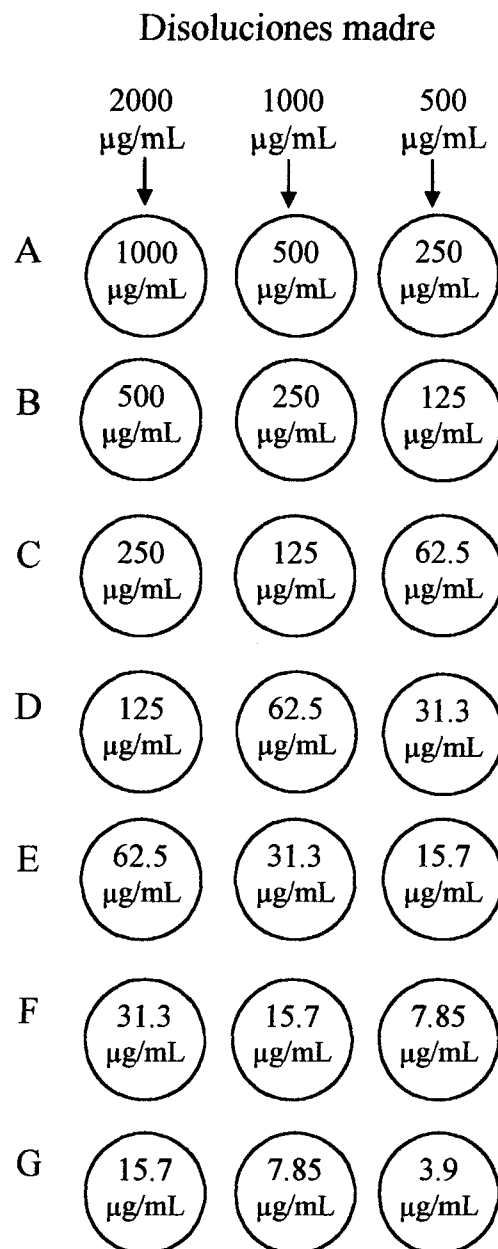


Figura 3.16. Concentraciones resultantes.

3.4. ENSAYOS DE TOXICIDAD

Con objeto de comprobar la toxicidad de los productos que muestren actividad antibacteriana tras los correspondientes ensayos microbiológicos se realizará la determinación de la Dosis Letal 50.

La determinación de la Dosis Letal 50 refleja la toxicidad aguda de un producto, al indicar la dosis que mataría al 50 % de la población en estudio cuando se administra por una vía específica. Por lo tanto, siempre se refiere a un animal de experimentación y a una vía de administración determinada.⁹⁵

El método seguido para la realización de los ensayos es el propuesto por Miller y Tainter.⁹⁶ Consiste en administrar dosis crecientes de los productos en estudio, partiendo de una dosis que no mata a ningún animal hasta la que los mata a todos.

3.4.1. MATERIAL Y REACTIVOS

- Balanza de precisión Mettler H31
- Baño de ultrasonidos P-Selecta Ultrasons-H
- Viales estériles de 20 mL
- Cabina U.V. P-Selecta
- Micropipeta Nichiryo Nichipet: 1000-5000 μ L
- Puntas estériles de 5 mL
- Agua estéril
- Etanol absoluto
- Tween 80
- Jeringas estériles de 1mL PLASTIPAK

- Mortero de cerámica
- Ratones de laboratorio raza Swiss
- Jaulas para ratones de laboratorio

3.4.2. PREPARACIÓN DE LOS PRODUCTOS

Los productos se deben preparar de la forma adecuada para ser administrados a los animales de experimentación. Inicialmente los productos se pesan en una balanza de precisión para posteriormente pulverizarlos en mortero y añadirles el excipiente. El excipiente utilizado es Etanol / Tween 80 / Agua estéril (2 : 2 : 35). Se agita el vial y se lleva a un baño de ultrasonidos durante 15 minutos.

3.4.3. PROTOCOLO DE ENSAYO

Se emplearon ratones de laboratorio no consanguíneos de la raza Swiss de pesos comprendidos entre 26 y 30g, elegidos aleatoriamente y distribuidos en lotes de 10 animales.

La administración de los productos se realiza siguiendo un orden aleatorio, de modo que cada ratón recibe un volumen igual de 0.25 mL por vía intraperitoneal. Las dosis administradas han seguido una progresión aritmética en la mayoría de los casos. Al lote testigo se le inyecta el excipiente utilizado para suspender los productos, comprobándose así la inocuidad del mismo.

La figura 3.17 refleja el proceso de administración por vía intraperitoneal de los productos:

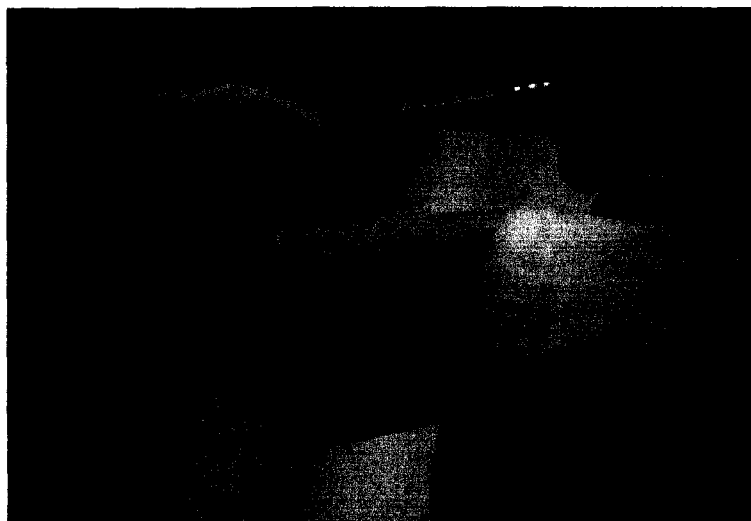


Figura 3.17. Administración de producto.

Los animales se observan a distintos tiempos de 1, 2, 4, 8 y 24 h. anotando el número de animales muertos transcurridas las 24 h. Posteriormente se calcula el porcentaje de muertes para cada una de las dosis, realizando una transformación de este valor en su probit (probabilidad de que haya respuesta para distintos niveles o dosis aplicados)⁹⁷ correspondiente.

Con respecto al porcentaje de muertes debemos realizar una corrección adicional ya que en ocasiones se da la circunstancia de que existe una dosis suficientemente alta o baja para que mueran o sobrevivan respectivamente todos los animales de experimentación, por tanto encontramos que no existe un valor finito del probit para estos datos, por lo que debemos calcular sus valores aproximados mediante la transformación propuesta por el método de Miller y Tainter:

- caso de 0 % de muertes: $100\left(\frac{0.25}{N}\right)$ pasando a ser 2.5 %
- caso de 100 % de muertes: $100\left[\frac{(N-0.25)}{N}\right]$ pasando a ser 97.5 %

donde N es el número de animales del lote experimental. A estos nuevos porcentajes se les calcula su probit, señalado en tablas⁹⁸ que relacionan % de mortalidad con probits. Se representa gráficamente probit en ordenadas frente a logaritmo de dosis en abcisas.

La dosis correspondiente al 50 % de muertes o probit 5 representa la DL_{50} . Como error de la DL_{50} se toma la desviación estándar V , definida por la siguiente expresión:

$$V = \frac{\Delta D}{\sqrt{4N}} \quad (3.4)$$

donde ΔD es la diferencia de las dosis correspondientes a los probits 6 y 4, y N es el número de ratones de un lote.

3.5. ESQUEMA DE TRABAJO

1. Selección del grupo farmacológico

En primer lugar se eligió el grupo farmacológico sobre el que trabajaremos que en nuestro caso fueron los antibacterianos. El interés por dicho grupo terapéutico se basa fundamentalmente en su gran importancia clínica, unido a la creciente aparición de resistencias a los antibióticos habitualmente empleados, lo que hace necesaria la búsqueda de nuevas estructuras químicas que posean dicha propiedad farmacológica. Por otra parte la posibilidad de realizar en nuestra unidad de investigación ensayos microbiológicos “*in vitro*”, permite comprobar la fiabilidad del método de diseño y/o selección propuesto.

2. Distribución de moléculas

Una vez elegido el grupo terapéutico el siguiente paso es seleccionar las moléculas que van a constituir los grupos de discriminación, es decir, activos e inactivos. Las moléculas pertenecientes a ambos grupos fueron obtenidas de la base de datos de la doceava edición en CD-Rom del Merck Index.

El grupo de activos está compuesto por 217 antibacterianos pertenecientes a distintos grupos químicos como son aminoglicósidos, anfenicoles, ansamicinas, β -lactámicos, lincosamidas, macrólidos, polipéptidos, tetraciclinas, nitrofuranos, quinolonas, sulfonamidas, sulfonas y otros.

El grupo de inactivos está compuesto por 216 moléculas pertenecientes a distintas categorías terapéuticas como son analgésicos narcóticos, analgésicos no-narcóticos, antidepressivos, antihistamínicos, hipolipemiantes, antihipertensivos, hipoglucemiantes y sedantes e hipnóticos.

Tanto en el grupo de activos como en el de inactivos se realiza una división estableciendo grupos de aprendizaje o entrenamiento y de test compuestos por el 70 % y el 30 % de moléculas respectivamente.

3. Cálculo de índices topológico-estructurales

A continuación se procede a realizar el cálculo de los 62 índices topológico-estructurales de cada una de las moléculas que participan en el estudio mediante herramientas informáticas desarrolladas en colaboración con el Departamento de Informática y Electrónica de la Licenciatura en Ingeniería Informática de la Universidad de Valencia.^{72,99}

4. Proceso de discriminación y búsqueda guiada de estructuras

El siguiente paso es el procesado de datos mediante la red neuronal artificial, que tras un proceso de entrenamiento determinaremos cuáles serán las características más idóneas de la misma para realizar un adecuado proceso de discriminación.

De forma paralela al procesado de datos mediante la red neuronal artificial, se realiza un análisis lineal discriminante mediante el programa BMDP 7M, tras el cual obtendremos una función discriminante, que permita realizar la clasificación de los compuestos en activos e inactivos.

Posteriormente mediante la función discriminante obtenida realizaremos una búsqueda guiada de estructuras con teórica actividad antibacteriana, que cumplan los requisitos topológicos marcados por dicha función, de entre los compuestos de la base de datos del módulo CS ChemFinder 5.1,¹⁰⁰ perteneciente al paquete CS ChemOffice, una de las herramientas más útiles en el sector de la química computacional. La base de datos

contiene catálogos de moléculas de distintos laboratorios como son Sigma-Aldrich, Fluka, Fisher, Avocado, etc.

Una vez seleccionados distintos compuestos mediante la búsqueda guiada anteriormente citada, debemos calcular los índices topológico-estructurales de los mismos y posteriormente se procede a procesar estas moléculas, introduciéndolas en la red neuronal artificial entrenada, obteniendo una clasificación de estos compuestos en activos e inactivos.

5. Ensayos farmacológicos: microbiológicos y de toxicidad

Los compuestos que la red neuronal clasifica teóricamente como activos y que además presentan mejores valores para la función discriminante son sometidos a ensayos microbiológicos con objeto de determinar la Concentración Mínima Inhibitoria de dichos compuestos frente a diferentes cepas de microorganismos.

Aquellas moléculas que tras realizarse los ensayos microbiológicos muestren actividad antibacteriana serán sometidas a ensayos de toxicidad con el propósito de determinar la Dosis Letal 50 de las mismas.

4. RESULTADOS

4. RESULTADOS

4.1. CÁLCULO DE ÍNDICES TOPOLÓGICO-ESTRUCTURALES

En primer lugar debemos proceder a realizar el cálculo de los 62 índices topológico-estructurales de todas y cada una de las moléculas que forman parte de nuestro estudio. Para ello se han aplicado herramientas informáticas desarrolladas en colaboración con el Departamento de Informática y Electrónica de la Licenciatura en Ingeniería Informática de la Universidad de Valencia.

El cálculo de dichos índices se realiza a partir de la descripción SMILES de las moléculas.

Como ya se ha comentado anteriormente el conjunto de moléculas objeto de estudio (217 con actividad antibacteriana y 216 inactivas) se dividieron en dos grupos, uno de aprendizaje o entrenamiento (70 %) y otro de test (30 %). Una vez calculados los índices se utilizarán como datos de entrada para el desarrollo del proceso de discriminación tanto para el caso de la red neuronal artificial como para el análisis lineal discriminante.

A continuación se muestra la figura 4.1 en la que se representa el valor promedio de cada uno de los 62 índices topológico-estructurales para las moléculas de los grupos de aprendizaje tanto de activas como de inactivas, lo que denominamos espectro molecular del grupo correspondiente.

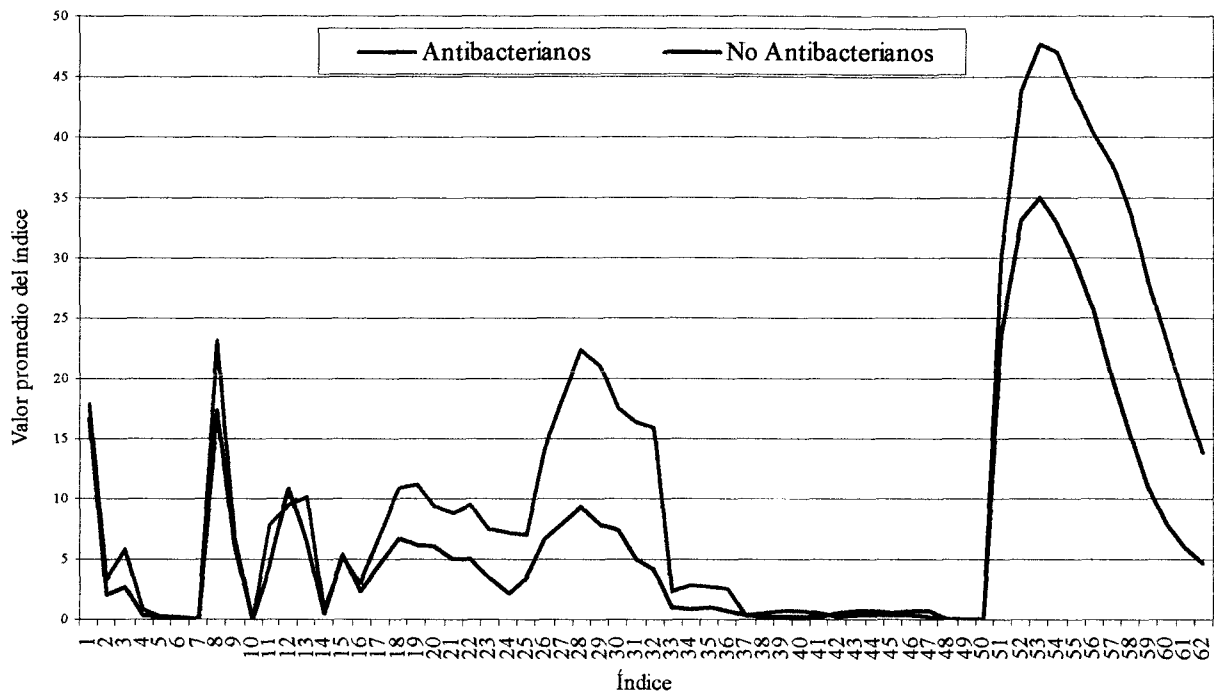


Figura 4.1. Espectro molecular para los grupos activo e inactivo del proceso de aprendizaje

4.2. DISCRIMINACIÓN

El proceso de discriminación se realizó empleando simultáneamente redes neuronales artificiales y análisis lineal discriminante. A continuación pasamos a detallar los resultados obtenidos por ambos métodos.

4.2.1. DISCRIMINACIÓN MEDIANTE RED NEURONAL ARTIFICIAL

Para establecer las características de la red que mejor resultados nos proporcionasen partimos de una red neuronal artificial de tipo perceptrón multicapa alimentada hacia delante (*feedforward*) y completamente conectada.

El algoritmo de aprendizaje utilizado es una modificación del algoritmo de retropropagación del error pero con mayor eficiencia y sin cambiar su funcionalidad, es el algoritmo *standard backpropagation*. A este algoritmo le debe acompañar como función de activación una función continua y diferenciable como es el caso de la función tangente hiperbólica que nos proporciona un valor de salida entre 1 y -1.

Para entrenar la red debemos partir de un número concreto de unidades de entrada que en nuestro caso serán 62, correspondientes a cada uno de los índices topológico-estructurales, una unidad de salida que determinará si la molécula es activa, inactiva o indeterminada y distintas unidades en la capa oculta (2, 4, 6, 8, 10, 16 y 32).

Los datos de entrada a la red vendrán almacenados en ficheros de texto, donde cada fila del fichero representa una molécula. Una molécula estará compuesta de 65 datos separados por comas, el primer dato es el nombre o código de la molécula, el segundo dato es la descripción SMILES de la molécula, a continuación sus 62 índices topológico-estructurales y por último un valor que es el índice de actividad farmacológica de la molécula, el cual toma valor 1 para una molécula antibacteriana y valor -1 para una

molécula no antibacteriana. A continuación se muestra como ejemplo un fragmento del fichero de entrada de la red que contiene tres moléculas:

```

1056, C3 (NC (=O) C (OC1=CC=CC=C1) C2=CC=CC=C2) C (=O) N4C3SC (C) (C) C4C (=O) O, 22, 2, 5
, 1, 0, 0, 0, 24, 9, 0, 6, 13, 10, 1, 6, 1, 5, 9, 10, 6, 9, 7, 4, 4, 6, 12, 16, 23, 20, 12, 11, 13, 2, 5
, 3, 4, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 33, 47, 48, 49, 45, 42, 42, 39, 33, 27, 19, 9, 1
1057, N1 (C) CCN (CC1) C2=C (F) C3=C (C=C2F) C (=O) C (=CN3CCF) C (=O) O, 17, 3, 3, 0, 3, 0, 0,
22, 6, 0, 7, 8, 11, 0, 6, 4, 9, 10, 12, 9, 8, 9, 5, 6, 3, 6, 8, 11, 10, 8, 11, 4, 0, 0, 0, 0, 5, 8, 14, 1
3, 12, 0, 0, 0, 0, 0, 0, 0, 0, 28, 41, 48, 49, 44, 37, 31, 18, 14, 9, 4, 2, 1
1058, CS (=O) (=O) C1=CC=C (C=C1) C (O) C (CF) NC (=O) C (Cl) Cl, 12, 1, 4, 1, 1, 2, 0, 15, 6, 0,
8, 6, 6, 1, 5, 2, 2, 4, 5, 2, 2, 1, 1, 3, 4, 10, 11, 10, 8, 7, 9, 5, 4, 2, 2, 1, 1, 2, 3, 4, 4, 4, 4, 2, 4,
6, 4, 0, 0, 0, 21, 30, 31, 27, 24, 21, 18, 14, 10, 8, 6, 0, 1

```

Figura 4.2. Fragmento de entrada para 3 moléculas

Estos datos de entrada se normalizan con el fin de facilitar el aprendizaje de la red neuronal. Si suponemos que X es el valor del índice topológico, el nuevo valor de éste será su correspondiente valor normalizado $X_{normalizado}$, de modo que la normalización realizada es la siguiente:

$$X_{normalizado} = \frac{X - X_{m\acute{a}xima}}{X_{m\acute{a}xima} - X_{m\acute{i}nima}}$$

Se ha establecido un criterio de clasificación para el análisis de resultados para saber cuando una molécula ha sido correctamente clasificada y evitar, en la medida de lo posible, la aparición de casos falsos positivos y/o falsos negativos. Por este motivo, en lugar de asignar valores positivos para las moléculas activas y negativos para las inactivas, hemos incluido un intervalo de valores a ambos lados del 0 (de 0 a 0.5 y de 0 a -0.5) en que consideramos las moléculas como indeterminadas, tratando así de evitar la aparición de moléculas falsamente activas o inactivas.

- para una molécula con actividad antibacteriana:

Valor de salida de la red	Clasificación
Entre 1 y 0.5	Activa (correctamente clasificada)
Entre 0 y -1	Inactiva (incorrectamente clasificada)
Entre 0.5 y 0	Indeterminada

- para una molécula sin actividad antibacteriana:

Valor de salida de la red	Clasificación
Entre -1 y -0.5	Inactiva (correctamente clasificada)
Entre 0 y 1	Activa (incorrectamente clasificada)
Entre -0.5 y 0	Indeterminada

Dentro de los grupos de antibacterianos y no antibacterianos la distribución entre los grupos de aprendizaje y test fue aleatoria. De este modo un porcentaje de la población se va a reservar para el aprendizaje o entrenamiento de la red y otro porcentaje de la población se utilizará para comprobar el grado de acierto de la discriminación.

La siguiente tabla muestra la distribución de las moléculas en cada uno de los grupos:

Tabla 4.1. Distribución de moléculas en los grupos de aprendizaje y test.

	Muestras Totales		Muestras de Aprendizaje		Muestras de Test	
	Activas	Inactivas	Activas	Inactivas	Activas	Inactivas
número	217	216	153	152	64	64
%	50.1	49.9	70.5	70.4	29.5	29.6

A continuación se muestran los resultados (en porcentaje) obtenidos en distintas pruebas, cada una de ellas con una topología de red neuronal distinta. Manteniendo el número de unidades de entrada (62) y unidades de salida (1), vamos variando el número de unidades en la capa oculta (2, 4, 6, 8, 10, 16 y 32), luego las posibles arquitecturas son 62-2-1, 62-4-1, 62-6-1, 62-8-1, 62-10-1, 62-16-1 y 62-32-1. Las pruebas se realizaron utilizando el número de ciclos como método de parada de la red siendo este valor de 10.000.

Tabla 4.2. Porcentajes de acierto obtenidos con diferentes arquitecturas de red.

Nº de neuronas en capa oculta	Grupo APRENDIZAJE			Grupo TEST		
	Aciertos	Errores	Indeter.	Aciertos	Errores	Indeter.
2	98.69	1.31	0	91.41	8.59	0
4	98.03	1.97	0	92.19	7.03	0.78
6	98.03	1.97	0	91.41	7.81	0.78
8	98.03	1.97	0	91.41	7.81	0.78
10	98.69	1.31	0	91.41	7.81	0.78
16	98.69	1.31	0	92.29	7.03	0.78
32	50.16	49.84	0	50	50	0

En función de los resultados obtenidos mostrados en la tabla anterior se seleccionó la red neuronal con 2 neuronas en su capa oculta, ya que el porcentaje acierto en el proceso de aprendizaje fue de 98.69 % siendo igualado únicamente por modelos de 10 y 16 neuronas en la capa oculta. A pesar de que algunos modelos con más de 2 neuronas en su capa oculta producían un ligero aumento en el porcentaje de aciertos para el grupo test, esto llevaba un aumento en la complejidad de cálculo no compensable con la mejora en los resultados obtenidos (inferior al 1 %).

En la figura siguiente se muestra la arquitectura definitiva de la red neuronal.

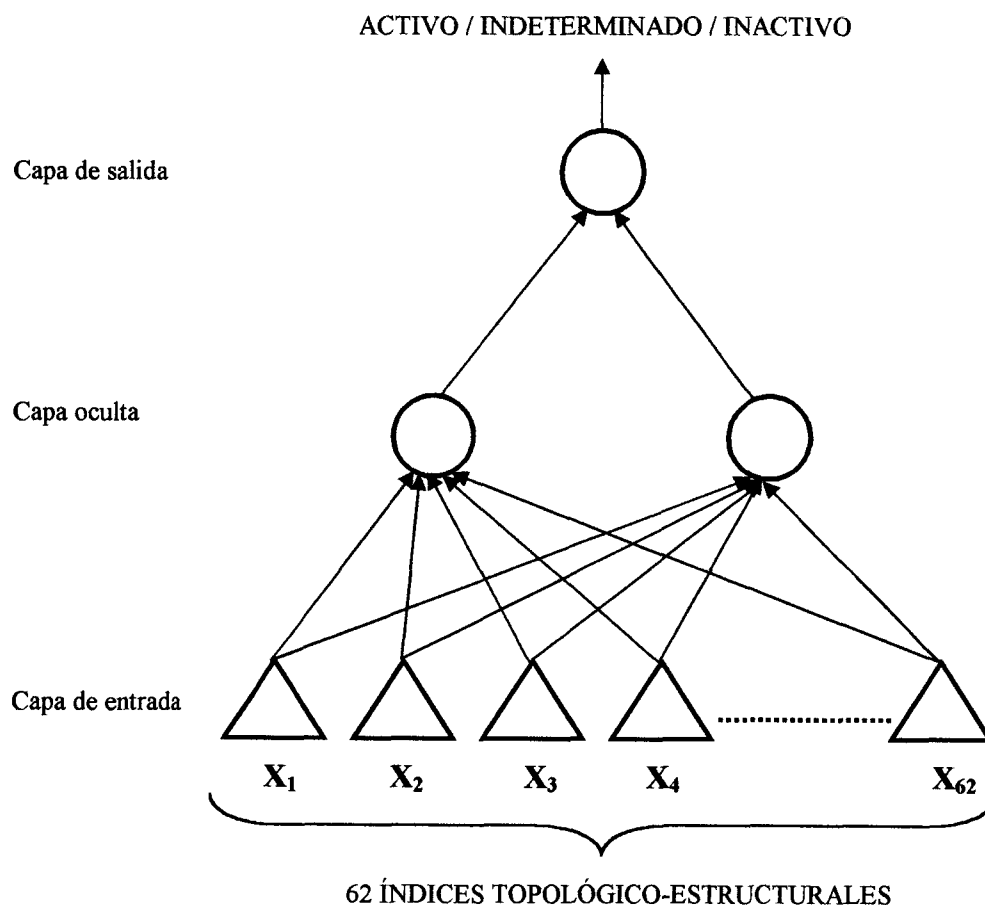


Figura 4.3. Arquitectura de la red neuronal seleccionada

Seguidamente analizaremos los resultados obtenidos mediante la red de arquitectura 62-2-1 de forma que se expresa para cada grupo de moléculas el porcentaje de aciertos obtenidos a partir del valor de salida de la red, con los criterios de clasificación antes expuestos y el porcentaje de acierto ajustado, entendiendo como tal aquellos valores en los que han sido excluidos los casos indeterminados, si los hubiera.

Tabla 4.3. Resultados obtenidos para el grupo de aprendizaje.

APRENDIZAJE			
<i>Grupo Activo</i>			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Activos	151	98.69	98.69
Indeterminados	0	0	
Inactivos	2	1.31	1.31
Total activos	153		
<i>Grupo Inactivo</i>			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Inactivos	150	98.68	98.68
Indeterminados	0	0	
Activos	2	1.32	1.32
Total inactivos	152		

Tabla 4.4. Resultados obtenidos para el grupo test.

TEST			
<i>Grupo Activo</i>			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto</i>	<i>Acierto ajustado (%)</i>
Activos	61	95.31	95.31
Indeterminados	0	0	
Inactivos	3	4.69	4.69
Total activos	64		
<i>Grupo Inactivo</i>			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto</i>	<i>Acierto ajustado (%)</i>
Inactivos	56	87.50	87.50
Indeterminados	0	0	
Activos	8	12.50	12.50
Total inactivos	64		

A continuación se disponen las tablas 4.5, 4.6, 4.7 y 4.8 en las que aparecen reflejados los valores individuales de salida de la red obtenidos para cada molécula en su grupo correspondiente, la clasificación de la misma con arreglo a dicho valor (signo "+" para las moléculas activas y signo "-" para las moléculas inactivas), así como la probabilidad de pertenencia a cada grupo de clasificación.

Tabla 4.5. Resultados obtenidos para los 153 compuestos antibacterianos del grupo de entrenamiento.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Acediasulfone	1.000	+	100
2.	Amdinocillin Pivoxil	1.000	+	100
3.	Amikacin	1.000	+	100
4.	Ampicillin	1.000	+	100
5.	Apalcillin	1.000	+	100
6.	Apramycin	1.000	+	100
7.	Aspoxicillin	1.000	+	100
8.	Azidamfenicol	1.000	+	100
9.	Azlocillin	1.000	+	100
10.	Bacampicillin	1.000	+	100
11.	Benzylsulfamide	0.976	+	98.80
12.	Biapenem	1.000	+	100
13.	Brodimoprim	1.000	+	100
14.	Butirosin	1.000	+	100
15.	Carbenicillin	1.000	+	100
16.	Carbomycin	0.999	+	99.97
17.	Carindacillin	1.000	+	100
18.	Cefaclor	1.000	+	100
19.	Cefatrizine	1.000	+	100
20.	Cefazedone	1.000	+	100
21.	Cefcapene Pivoxil	1.000	+	100
22.	Cefdinir	1.000	+	100
23.	Cefetamet	1.000	+	100
24.	Cefixime	1.000	+	100
25.	Cefodizime	1.000	+	100
26.	Cefonicid	1.000	+	100
27.	Ceforanide	1.000	+	100
28.	Cefpiramide	1.000	+	100
29.	Cefpodoxime Proxetil	1.000	+	100
30.	Cefroxadine	1.000	+	100
31.	Cefteram	1.000	+	100
32.	Ceftibuten	1.000	+	100
33.	Cefizoxime	1.000	+	100
34.	Cefuroxime	1.000	+	100
35.	Cefuzonam	1.000	+	100
36.	Cephalexin	1.000	+	100
37.	Cephalothin	1.000	+	100
38.	Cephapirin Sodium	0.984	+	99.21
39.	Cephradine	1.000	+	100
40.	Chloramphenicol	1.000	+	100
41.	Chlortetracycline	1.000	+	100
42.	Ciprofloxacin	1.000	+	100
43.	Clarithromycin	1.000	+	99.98
44.	Clinafloxacin	0.995	+	99.76
45.	Clindamycin	0.976	+	98.78
46.	Clomocycline	0.991	+	99.53

Tabla 4.5. Continuación.

<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
47. Cloxacillin	1.000	+	100
48. Cyclacillin	1.000	+	100
49. Dapsone	1.000	+	100
50. Demeclocycline	1.000	+	100
51. Dibekacin	1.000	+	100
52. Dichloramine T	0.983	+	99.14
53. Difloxacin	1.000	+	100
54. Dihydrostreptomycin	1.000	+	100
55. Erythromycin	1.000	+	99.98
56. Fenbenicillin	1.000	+	100
57. Fleroxacin	1.000	+	100
58. Florfenicol	1.000	+	100
59. Floxacillin	1.000	+	100
60. Flumequine	1.000	+	100
61. Fortimicin B	1.000	+	100
62. Furaltadone	1.000	+	100
63. Gentamicin	1.000	+	100
64. Hexedine	-1.000	-	0.00
65. Imipenem	1.000	+	100
66. Isepamicin	1.000	+	100
67. Lenampicillin	1.000	+	100
68. Leucomycins	0.968	+	98.39
69. Lomefloxacin	1.000	+	100
70. Lyme cycline	0.999	+	99.96
71. Mafenide	0.942	+	97.08
72. Meclocycline	1.000	+	100
73. 4'-(Methylsulfamoyl)sulfanililide	1.000	+	100
74. Mezlocillin	1.000	+	100
75. Midecamycins	0.999	+	99.96
76. Miloxacin	0.995	+	99.74
77. Minocycline	1.000	+	100
78. Miokamycin	0.999	+	99.97
79. Nadifloxacin	1.000	+	100
80. Nafacillin Sodium	1.000	+	100
81. Neomycin	1.000	+	100
82. Nifuradene	1.000	+	100
83. Nifurfoline	1.000	+	100
84. Nifurpirinol	1.000	+	100
85. Nifurprazine	1.000	+	100
86. Nitrofurantoin	1.000	+	100
87. Norfloxacin	1.000	+	100
88. Ofloxacin	1.000	+	100
89. Oxacillin	1.000	+	100
90. Oxytetracycline	1.000	+	100
91. Paromomycin	1.000	+	100
92. Pazufloxacin	0.973	+	98.63
93. Pefloxacin	1.000	+	100
94. Penicillin G	1.000	+	100
95. Penicillin N	1.000	+	100

Tabla 4.5. Continuación.

<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
96. Penicillin V	1.000	+	99.99
97. Phenethicillin Potassium	1.000	+	100
98. Phthalylsulfacetamide	1.000	+	100
99. Pipacycline	1.000	+	100
100. Piromidic Acid	0.969	+	98.46
101. Pivcefalexin	1.000	+	100
102. Propicillin	1.000	+	100
103. Quinacillin	1.000	+	100
104. Rifamide	1.000	+	99.98
105. Rifamycin SV	0.999	+	99.96
106. Rifapentine	0.998	+	99.90
107. Rosaranicin	0.963	+	98.17
108. Rosoxacin	0.993	+	99.63
109. Salazosulfadimidine	1.000	+	100
110. Sancycline	1.000	+	100
111. Sisomicin	1.000	+	100
112. Solasulfone	0.974	+	98.71
113. Spectinomycin	1.000	+	100
114. Spiramycin	0.989	+	99.47
115. Streptomycin	1.000	+	100
116. Succinylsulfathiazole	1.000	+	100
117. Succisulfone	1.000	+	100
118. Sulfachlorpyridazine	1.000	+	100
119. Sulfachrysoidine	1.000	+	100
120. Sulfadiazine	1.000	+	100
121. Sulfadicramide	1.000	+	100
122. Sulfadoxine	1.000	+	100
123. Sulfaethidole	1.000	+	100
124. Sulfaguanidine	1.000	+	100
125. Sulfalene	1.000	+	100
126. Sulfaloxic Acid	1.000	+	100
127. Sulfamethazine	1.000	+	100
128. Sulfamethizole	1.000	+	100
129. Sulfamethoxazole	1.000	+	100
130. Sulfamethoxypyridazine	1.000	+	100
131. Sulfamidochrysoidine	1.000	+	100
132. Sulfamoxole	1.000	+	100
133. Sulfanilamide	1.000	+	100
134. Sulfanilic Acid	1.000	+	100
135. N4-Sulfanilylsulfanilamide	1.000	+	100
136. Sulfanitran	1.000	+	100
137. Sulfaperine	1.000	+	100
138. Sulfaphenazole	0.959	+	97.95
139. Sulfapyrazine	1.000	+	100
140. Sulfapyridine	1.000	+	100
141. Sulfasymazine	1.000	+	100
142. Sulfathiazole	1.000	+	100
143. Sulfathiourea	1.000	+	100
144. Sulfisoxazole	1.000	+	100

Tabla 4.5. Continuación.

<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
145. Talampicillin	1.000	+	100
146. Taurolidine	1.000	+	100
147. Temafloxacin	1.000	+	100
148. Tetracycline	1.000	+	100
149. Thiazolsulfone	1.000	+	100
150. Tigemonam	1.000	+	100
151. Tosufloxacin	1.000	+	100
152. Trimethoprim	1.000	+	100
153. Xibornol	-1.000	-	0.00

Tabla 4.6. Resultados obtenidos para los 152 compuestos no antibacterianos del grupo de entrenamiento.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Acecarbromal	-1.000	-	0
2.	Acetaminophen	-1.000	-	0
3.	Acetohexamide	-1.000	-	0.01
4.	Adinazolam	-1.000	-	0
5.	Ahistan	-1.000	-	0
6.	Alfentanil	-1.000	-	0
7.	Althiazide	-1.000	-	0
8.	Aminochlorthenoxazin	-1.000	-	0
9.	Amlodipine	-1.000	-	0
10.	Amobarbital	-1.000	-	0
11.	Ampiroxicam	1.000	+	100
12.	Amtolmetin Guacil	-1.000	-	0
13.	Antipyrine	-0.971	-	1.44
14.	Atorvastatin	-1.000	-	0
15.	Azacosterol	-1.000	-	0
16.	Azosemide	-1.000	-	0
17.	Benactyzine	-0.995	-	0.27
18.	Benazepril	-1.000	-	0
19.	Benorylate	-1.000	-	0
20.	Benzylmorphine	-1.000	-	0
21.	Bevantolol	-1.000	-	0
22.	Binifibrate	-0.959	-	2.03
23.	Brallobarbital	-1.000	-	0
24.	Bromfenac	-0.999	-	0.05
25.	Bucetin	-1.000	-	0
26.	Bunitrolol	-1.000	-	0
27.	Butacetin	-1.000	-	0
28.	Butethal	-1.000	-	0
29.	Buthiazide	-1.000	-	0
30.	Butriptyline	-1.000	-	0
31.	Captopril	-0.976	-	1.18
32.	Carbinoxamine	-1.000	-	0
33.	Carteolol	-1.000	-	0
34.	Carvedilol	-1.000	-	0
35.	Chlorhexadol	-1.000	-	0
36.	Chlorpheniramine	-1.000	-	0
37.	Chlorthenoxazin	-1.000	-	0
38.	Clocinizine	-1.000	-	0
39.	Clofenamide	-1.000	-	0
40.	Clomestron	-1.000	-	0
41.	Crotethamide	-1.000	-	0
42.	Cyclothiazide	-1.000	-	0
43.	Cyproheptadine	-1.000	-	0
44.	Delapril	-1.000	-	0
45.	Desipramine	-1.000	-	0
46.	Desomorphine	-1.000	-	0

Tabla 4.6. Continuación.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
47.	Dihydrocodeinone Enol Acetate	-1.000	-	0
48.	Dioxadrol	-1.000	-	0
49.	Diphenylpyraline	-1.000	-	0
50.	Dipipanone	-1.000	-	0
51.	Disulfamide	-1.000	-	0
52.	Dothiepin	-1.000	-	0
53.	Doxylamine	-1.000	-	0
54.	Enalapril	-1.000	-	0
55.	Enfenamic Acid	-1.000	-	0
56.	Eptazocine	-1.000	-	0
57.	Etaqualone	-1.000	-	0
58.	Etodolac	-1.000	-	0
59.	Etonitazene	-0.996	-	0.20
60.	Etymemazine	-1.000	-	0
61.	Febarbamate	-1.000	-	0
62.	Felodipine	-1.000	-	0
63.	Femoxetine	-1.000	-	0
64.	Fenbufen	-1.000	-	0
65.	Feprazone	-1.000	-	0
66.	Flufenamic Acid	-1.000	-	0
67.	Flunitrazepam	-0.976	-	1.19
68.	Fluoresone	-0.967	-	1.64
69.	Flurbiprofen	-1.000	-	0
70.	Furazabol	-1.000	-	0
71.	Glibornuride	-0.998	-	0.12
72.	Gliquidone	-1.000	-	0
73.	Glybuzole	-0.949	-	2.55
74.	Glymidine	-1.000	-	0
75.	Glypinamide	-1.000	-	0
76.	Guanadrel	-1.000	-	0
77.	Guanoxan	-1.000	-	0
78.	Haloxazolam	-1.000	-	0
79.	Hexethal	-1.000	-	0
80.	Hydrochlorothiazide	-1.000	-	0
81.	Hydroxyzine	-1.000	-	0
82.	Ibuprofen	-1.000	-	0
83.	Indeloxazine Hydrochloride	-1.000	-	0
84.	Indoramin	-1.000	-	0
85.	Isomethadone	-1.000	-	0
86.	Isoxicam	1.000	+	100
87.	Ketoprofen	-1.000	-	0
88.	Levcromakalim	-1.000	-	0
89.	Levopropoperane	-1.000	-	0
90.	Lisinopril	-1.000	-	0
91.	Lofentanil	-1.000	-	0
92.	Loxoprofen	-1.000	-	0
93.	Mebhydroline	-1.000	-	0
94.	Meclozamine	-1.000	-	0
95.	Metazocine	-1.000	-	0

Tabla 4.6. Continuación.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
96.	Methyclothiazide	-0.968	-	1.62
97.	Methyl-4-Pyridyl Ketone Thiosemicarbazone	-0.943	-	2.85
98.	Methyl-4-thiazoleethanol-5	-1.000	-	0
99.	Metopon	-1.000	-	0
100.	Metralindole	-1.000	-	0
101.	Mofebutazone	-1.000	-	0
102.	Moveltipril	-1.000	-	0
103.	Moxastine	-1.000	-	0
104.	Muzolimine	-1.000	-	0
105.	Nalbuphine	-1.000	-	0
106.	Nefazodone	-1.000	-	0
107.	Nicardipine	-1.000	-	0
108.	Nicoclonate	-1.000	-	0
109.	Nitrazepam	-1.000	-	0
110.	Nitrendipine	-1.000	-	0
111.	Norpipanone	-1.000	-	0
112.	Opipramol	-1.000	-	0
113.	Oxaflozane	-1.000	-	0
114.	Oxametacine	-1.000	-	0
115.	Parsalmide	-1.000	-	0
116.	Phenobarbital	-1.000	-	0
117.	Phenoperidine	-1.000	-	0
118.	Phenopyrazone	-1.000	-	0
119.	Phentolamine	-1.000	-	0
120.	Phenylbutazone	-1.000	-	0
121.	Phenylbutyramide	-1.000	-	0
122.	Pizotyline	-1.000	-	0
123.	Polythiazide	-1.000	-	0
124.	Prazosin	-1.000	-	0.01
125.	Propacetamol	-1.000	-	0
126.	Propoxyphene	-1.000	-	0
127.	Protizinic Acid	-1.000	-	0
128.	Pyrisuccideanol	-0.999	-	0.05
129.	Pyrrobutamine	-1.000	-	0
130.	Rescimetol	-1.000	-	0
131.	Reserpine	-1.000	-	0
132.	Ritanserin	-1.000	-	0
133.	Ronifibrate	-1.000	-	0
134.	Salacetamide	-0.999	-	0.06
135.	Salicylsulfuric Acid	-0.972	-	1.41
136.	Secobarbital	-1.000	-	0
137.	Sulfinalol	-1.000	-	0
138.	Suprofen	-0.999	-	0.05
139.	Talniflumate	-1.000	-	0.02
140.	Tandospirone	-1.000	-	0
141.	Thenylidiamine	-1.000	-	0
142.	Timolol	-0.999	-	0.07
143.	Tolbutamide	-1.000	-	0

Tabla 4.6. Continuación.

<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
144. Tolfenamic Acid	-1.000	-	0
145. Tolmetin	-0.999	-	0.05
146. Toloxatone	-1.000	-	0
147. Tolpropamine	-1.000	-	0
148. Triamterene	-0.989	-	0.54
149. Triazolam	-1.000	-	0
150. Trimazosin	-1.000	-	0.01
151. Tritoqualine	-0.958	-	2.11
152. Vinbarbital	-1.000	-	0

Tabla 4.7. Resultados obtenidos para los 64 compuestos antibacterianos del grupo test.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Amdinocillin	1.000	+	100
2.	Amoxicillin	0.998	+	99.92
3.	Apicycline	1.000	+	100
4.	Arbekacin	1.000	+	100
5.	Azidocillin	1.000	+	100
6.	Azithromycin	1.000	+	100
7.	Benzylpenicillinic Acid	0.999	+	99.97
8.	Carumonam	1.000	+	100
9.	Cefamandole	1.000	+	100
10.	Cefditoren	1.000	+	100
11.	Cefmenoxime	1.000	+	100
12.	Cefoperazone	1.000	+	100
13.	Cefotiam	1.000	+	100
14.	Cefprozil	1.000	+	100
15.	Ceftazole	1.000	+	100
16.	Ceftriaxone	1.000	+	100
17.	Cephacetrile Sodium	1.000	+	100
18.	Cephalosporin C	1.000	+	100
19.	Clometocillin	1.000	+	100
20.	Diathymosulfone	0.989	+	99.43
21.	Dicloxacillin	0.944	+	97.18
22.	Doxycycline	1.000	+	100
23.	Enoxacin	1.000	+	100
24.	N2-Formylsulfisomidine	1.000	+	100
25.	Fortimicin A	1.000	+	100
26.	Grepafloxacin	1.000	+	100
27.	Hetacillin	-0.409	-	29.55
28.	Lincomycin	1.000	+	100
29.	Meropenem	1.000	+	100
30.	Methacycline	1.000	+	100
31.	Methenamine	1.000	+	100
32.	Methicillin Sodium	1.000	+	100
33.	Micronomicin	1.000	+	100
34.	Nalidixic Acid	1.000	+	100
35.	Nifuratel	1.000	+	100
36.	Nifurtoinol	1.000	+	100
37.	Oxolinic Acid	-1.000	-	0
38.	Penamecillin	-1.000	-	0
39.	Penicillin O	0.996	+	99.82
40.	Phthatylsulfathiazole	1.000	+	100
41.	Pipemidic Acid	0.992	+	99.61
42.	Pivampicillin	1.000	+	100
43.	Ribostamycin	1.000	+	100
44.	Rifampin	1.000	+	99.98
45.	Rokitamycin	0.997	+	99.84
46.	Roxithromycin	0.999	+	99.96
47.	Sparfloxacin	1.000	+	100
48.	Sulbenicillin	1.000	+	100
49.	Sulfabenzamide	1.000	+	100

Tabla 4.7. Continuación.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
50.	Sulfacytine	1.000	+	100
51.	Sulfadimethoxine	1.000	+	100
52.	Sulfaguanole	0.999	+	99.97
53.	Sulfamerazine	1.000	+	100
54.	Sulfamethomidine	1.000	+	100
55.	4-Sulfanilamidosalicylic Acid	1.000	+	100
56.	p-Sulfanilylbenzylamine	1.000	+	100
57.	N-Sulfanilyl-3,4-xylamide	1.000	+	100
58.	Sulfaproxyline	1.000	+	100
59.	Sulfasomizole	1.000	+	100
60.	Sulfisomidine	1.000	+	100
61.	Tetroxoprim	0.943	+	97.16
62.	Thiamphenicol	1.000	+	100
63.	Ticarcillin	1.000	+	100
64.	Trovafloxacin	1.000	+	100

Tabla 4.8. Resultados obtenidos para los 64 compuestos no antibacterianos del grupo test.

	<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Aminopyrine	0.042	+	52.09
2.	Aprobarbital	-1.000	-	0
3.	Aranidipine	-1.000	-	0
4.	Arbutin	1.000	+	100
5.	Bamipine	-1.000	-	0
6.	Bunazosin	-1.000	-	0
7.	Bupropion	-1.000	-	0
8.	Chloralose-alfa	-0.999	-	0.04
9.	Chlorpropamide	-0.997	-	0.14
10.	Chlortalidone	-1.000	-	0
11.	Cicletanine	-1.000	-	0.01
12.	Clemastine	-1.000	-	0
13.	Clinofibrate	-1.000	-	0
14.	Clobenzepam	-1.000	-	0
15.	Clometacin	-1.000	-	0
16.	Dezocine	-1.000	-	0
17.	Dimetacrine	-1.000	-	0
18.	Droxicam	1.000	+	100
19.	Ebastine	-1.000	-	0
20.	Enalaprilat	-1.000	-	0
21.	Ethchlorvynol	-1.000	-	0
22.	Ethylmethylthiambutene	-1.000	-	0
23.	Etozolin	-1.000	-	0
24.	Gemfibrozil	-1.000	-	0
25.	Glafenine	0.959	+	97.97
26.	Glimepiride	-1.000	-	0.01
27.	Guanazodine	-1.000	-	0
28.	Heptabarbital	-1.000	-	0
29.	Isothipendyl	-1.000	-	0
30.	Isoxepac	-1.000	-	0
31.	Lofepamine	-1.000	-	0
32.	Lofexidine	-1.000	-	0
33.	Lormetazepam	-1.000	-	0
34.	L-Tryptophan	-1.000	-	0
35.	Mefenamic Acid	-0.997	-	0.17
36.	Mepindolol	0.252	+	62.61
37.	Meptazinol	-1.000	-	0
38.	Methotrimeprazine	-1.000	-	0
39.	Methyldiphenhydramine-p	-1.000	-	0
40.	Methyprylon	-1.000	-	0
41.	Moclobemide	-1.000	-	0
42.	Nilvadipine	-1.000	-	0
43.	Noxiptilin	-1.000	-	0
44.	p-Bromoacetanilide	-1.000	-	0
45.	PhenazopyridineHydrochloride	1.000	+	100
46.	Phenyltoloxamine	-1.000	-	0

Tabla 4.8. Continuación.

<i>Molécula</i>	<i>Valor salida red</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
47. Pildralazine	1.000	+	100
48. Pirazolac	-1.000	-	0
49. Propallylonal	-1.000	-	0
50. Propizepine	-0.998	-	0.09
51. Quinethazone	-1.000	-	0
52. Rescinamine	-1.000	-	0
53. Salicylamide-Acetic Acid	-1.000	-	0
54. Simfibrate	-1.000	-	0
55. Sufentanil	-1.000	-	0
56. Syrosingopine	-1.000	-	0.01
57. Thonzylamine	-1.000	-	0
58. Tilisolol	1.000	+	100
59. Tofenacin	-1.000	-	0
60. Tolazamide	-0.999	-	0.05
61. Tramadol	-1.000	-	0
62. Tripamide	-0.683	-	15.86
63. Vinylbital	-1.000	-	0
64. Ximoprofen	-1.000	-	0

La tabla 4.9 representa la espectancia de activos de los grupos de aprendizaje (E_a) y test (E_{at}), así como la espectancia de inactivos para los mismos grupos (E_i , E_{it}) para los intervalos de los valores de salida de la red.

Tabla 4.9. Valores de espectancia de actividad e inactividad obtenidos para los grupos de entrenamiento y test.

Intervalos de valores de salida	Moléculas Grupo Aprendizaje		Moléculas Grupo Test		Valores deEspectancia			
	Activos	Inactivos	Activos	Inactivos	Ea	Ei	Eat	Eit
-1 a -0.9	2	150	2	55	0.0066	0.9741	0.0168	0.8333
-0.9 a -0.8	0	0	0	0	0	0	0	0
-0.8 a -0.7	0	0	0	0	0	0	0	0
-0.7 a -0.6	0	0	0	1	0	0	0	0.0156
-0.6 a -0.5	0	0	0	0	0	0	0	0
-0.5 a -0.4	0	0	1	0	0	0	0.0156	0
-0.4 a -0.3	0	0	0	0	0	0	0	0
-0.3 a -0.2	0	0	0	0	0	0	0	0
-0.2 a -0.1	0	0	0	0	0	0	0	0
-0.1 a 0	0	0	0	0	0	0	0	0
0 a 0.1	0	0	0	1	0	0	0	0.0156
0.1 a 0.2	0	0	0	0	0	0	0	0
0.2 a 0.3	0	0	0	1	0	0	0	0.0156
0.3 a 0.4	0	0	0	0	0	0	0	0
0.4 a 0.5	0	0	0	0	0	0	0	0
0.5 a 0.6	0	0	0	0	0	0	0	0
0.6 a 0.7	0	0	0	0	0	0	0	0
0.7 a 0.8	0	0	0	0	0	0	0	0
0.8 a 0.9	0	0	0	0	0	0	0	0
0.9 a 1	151	2	61	6	0.9741	0.0066	0.8714	0.0480
<i>Total moléculas</i>	<i>153</i>	<i>152</i>	<i>64</i>	<i>64</i>				

La figura 4.4 muestra el diagrama de distribución farmacológica obtenido para los valores de salida de la red neuronal entrenada.

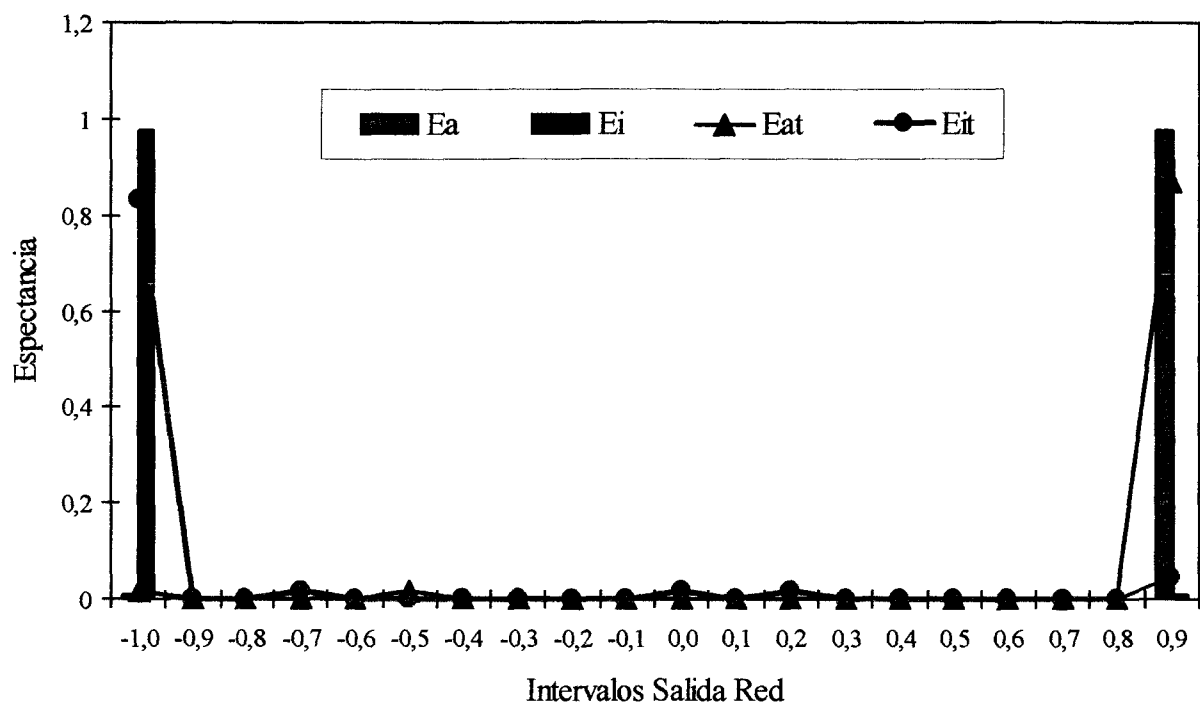


Figura 4.4. Diagrama de distribución farmacológica para los valores de salida de la red neuronal entrenada

4.2.2. DISCRIMINACIÓN MEDIANTE ANÁLISIS LINEAL DISCRIMINANTE

En este apartado se detalla el proceso de obtención de la función discriminante, la cual nos permitirá clasificar correctamente las moléculas que posean la actividad farmacológica deseada, en nuestro caso la antibacteriana, evitando al mismo tiempo la aparición de moléculas inactivas.

Se procedió a realizar un análisis lineal discriminante con la ayuda del programa BMDP 7M descrito anteriormente en el apartado 3.2.2.1 con el propósito de seleccionar aquellos índices topológico-estructurales capaces de distinguir moléculas con actividad antibacteriana de moléculas que no posean dicha actividad, por lo que estos índices se utilizaron como variables independientes.

El análisis discriminante se lleva a cabo con dos grandes grupos de moléculas, uno de antibacterianos de reconocida actividad antibacteriana formado por 217 moléculas y otro con 216 moléculas sin dicha actividad terapéutica.

El estudio se realizó empleando como grupo discriminante un total de 153 moléculas con actividad antibacteriana y 152 moléculas sin dicha actividad, el resto de muestras de cada uno de los grupos (64 del grupo de activos y 64 del grupo de inactivos) se emplearon para realizar el correspondiente test de validación cruzada de la función discriminante obtenida.

La figura 4.5 muestra la distribución de moléculas en los grupos de entrenamiento y test. Para los grupos activo e inactivo, el 70% de las muestras se emplean en el proceso de discriminación mientras que el 30% restante se reserva para la validación cruzada.

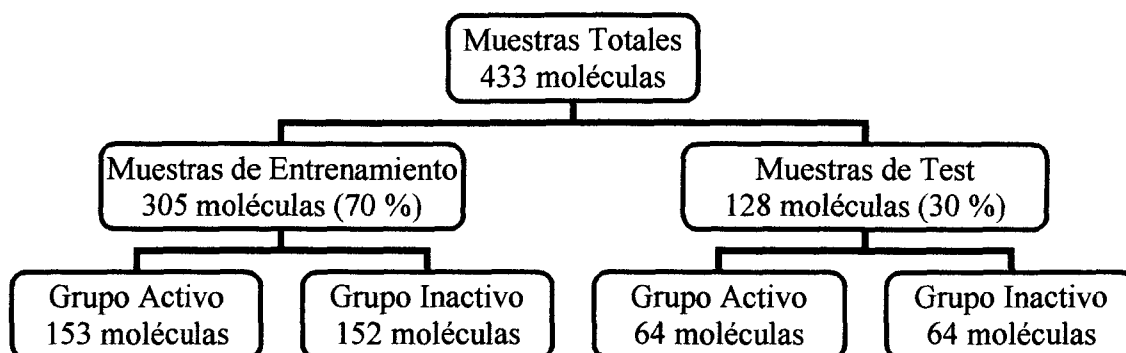


Figura 4.5. Distribución de moléculas en los grupos de entrenamiento y test.

El criterio para la selección de la mejor función discriminante se realizó basándonos en los siguientes aspectos:

- comparación de los valores de F y U estadísticas,
- determinación del porcentaje de moléculas correctamente clasificadas durante el proceso de entrenamiento (grupo de entrenamiento activo y grupo de entrenamiento inactivo),
- predicción de la clasificación de moléculas no incluidas en el proceso de entrenamiento (grupo test activo y grupo test inactivo), a lo que se denomina validación cruzada.

En primer lugar se ejecuta el programa BMDP 7M utilizando como variables independientes los 62 índices topológico-estructurales, cuyo cálculo se ha descrito en capítulos anteriores: $A^C, A^N, A^O, A^S, A^F, A^{Cl}, A^{Br}, B^1, B^2, B^3, V^1, V^2, V^3, V^4, B^{D1}, B^{D2}, D^{N1}, D^{N2}, D^{N3}, D^{N4}, D^{N5}, D^{N6}, D^{N7}, D^{N8}, D^{O1}, D^{O2}, D^{O3}, D^{O4}, D^{O5}, D^{O6}, D^{O7}, D^{O8}, D^{S1}, D^{S2}, D^{S3}, D^{S4}, D^{F2}, D^{F3}, D^{F4}, D^{F5}, D^{F6}, D^{Cl2}, D^{Cl3}, D^{Cl4}, D^{Cl5}, D^{Cl6}, D^{Cl7}, D^{Br2}, D^{Br3}, D^{Br4}, D^1, D^2, D^3, D^4, D^5, D^6, D^7, D^8, D^9, D^{10}, D^{11}, D^{12}$.

El programa BMDP 7M realizará numerosas correlaciones pero sólo seleccionaremos aquella correlación que presente los mejores resultados para los tres criterios anteriormente mencionados.

La función discriminante seleccionada clasificará las moléculas como activas para valores de la misma superiores a 0.5 ($\Delta P > 0.5$), como inactivas para valores inferiores a 0 ($\Delta P < 0$) y el resto serán clasificadas como indeterminadas. Estas condiciones de discriminación intentan minimizar el porcentaje de error, es decir, obtener el menor número posible de falsos positivos, aunque se vean sacrificadas algunas moléculas activas.

Tabla 4.10. Clasificación de moléculas en función del valor de ΔP .

Valor de la función discriminante	Clasificación
$\Delta P > 0.5$	Activa
$\Delta P < 0$	Inactiva
$0.5 > \Delta P > 0$	Indeterminada

La mejor función discriminante fue obtenida con las variables V^2 , D^{N8} , D^{O5} , D^{S2} , D^{F5} y D^{Cl4} .

A continuación se muestra la función discriminante, obtenida como diferencia entre las variables de los grupos de moléculas activas e inactivas, así como los valores para la F estadística y U estadística, y los coeficientes de cada una de las variables.

Tabla 4.11. Parámetros de la función discriminante.

V^2	-0.3680	42.2422	0.6605
D^{N8}	0.2602	99.4499	0.7529
D^{O5}	0.1057	38.4833	0.5234
D^{S2}	0.5739	35.6434	0.5906
D^{F5}	0.4680	19.5497	0.4617
D^{Cl4}	-0.6088	19.1343	0.4919
constante	0.0742	-	-

Función discriminante

$$\Delta P = -0.3680 V^2 + 0.2602 D^{N8} + 0.1057 D^{O5} + 0.5739 D^{S2} + 0.4680 D^{F5} - 0.6088 D^{Cl4} + 0.0742$$

En la siguiente tabla se analizan los resultados obtenidos para los grupos activo e inactivo del proceso de entrenamiento (moléculas que participan en la selección de la función discriminante) como porcentaje de aciertos obtenidos a partir del valor de la función discriminante, con los criterios de clasificación anteriormente nombrados, así como el porcentaje de acierto ajustado, entendiendo como tal aquellos valores en los que han sido excluidos los casos indeterminados.

Tabla 4.12. Resultados obtenidos en la clasificación de los grupos activos e inactivos en el proceso de entrenamiento.

Grupo Entrenamiento Activo			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Activos	119	77.78	81.51
Indeterminados	7	4.58	
Inactivos	27	17.65	18.49
Total activos	153		
Grupo Entrenamiento Inactivo			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Inactivos	132	86.84	89.90
Indeterminados	5	3.29	
Activos	15	9.87	10.20
Total inactivos	152		

La validación de la función discriminante se realizó mediante un conjunto de moléculas no utilizadas en la selección de dicha función y divididas en grupo test activo y grupo test inactivo. Los resultados se indican en la tabla siguiente.

Tabla 4.13. Resultados obtenidos en la clasificación de los grupos activos e inactivos en el proceso de test.

Grupo Test Activo			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Activos	48	75.00	81.36
Indeterminados	5	7.81	
Inactivos	11	17.19	18.64
Total activos	64		
Grupo Test Inactivo			
<i>Clasificación</i>	<i>Nº moléculas</i>	<i>Acierto (%)</i>	<i>Acierto ajustado (%)</i>
Inactivos	58	90.63	93.55
Indeterminados	2	3.13	
Activos	4	6.25	6.45
Total inactivos	64		

A continuación detallamos las tablas 4.14, 4.15, 4.16 y 4.17 en las que figuran los valores de la función discriminante (ΔP) obtenidos para cada una de las moléculas en su grupo correspondiente, la clasificación de las mismas con arreglo a dicho valor (signo "+" para las moléculas activas, signo "-" para las inactivas, "I" para las indeterminaciones), así como la probabilidad de pertenencia a cada grupo de clasificación.

Tabla 4.14. Resultados obtenidos para las 153 moléculas del grupo entrenamiento de activos.

	<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Acediasulfone	0.576	+	53.05
2.	Amdinocillin Pivoxil	2.301	+	62.16
3.	Amikacin	6.298	+	83.29
4.	Ampicillin	3.254	+	67.20
5.	Apalcillin	4.598	+	74.30
6.	Apramycin	4.991	+	76.38
7.	Aspoxicillin	6.543	+	84.59
8.	Azidamfenicol	0.999	+	55.28
9.	Azlocillin	4.604	+	74.33
10.	Bacampicillin	3.197	+	66.90
11.	Benzylsulfamide	-2.354	-	37.56
12.	Biapenem	4.089	+	71.61
13.	Brodinoprim	0.415	I	52.19
14.	Butirosin	5.615	+	79.68
15.	Carbenicillin	3.108	+	66.43
16.	Carbomycin	2.411	+	62.74
17.	Carindacillin	1.841	+	59.73
18.	Cefaclor	-0.745	-	46.06
19.	Cefatrizine	4.751	+	75.11
20.	Cefazedone	3.349	+	67.70
21.	Cefcapene Pivoxil	6.125	+	82.37
22.	Cefdinir	4.981	+	76.33
23.	Cefetamet	5.135	+	77.14
24.	Cefixime	5.082	+	76.86
25.	Cefodizime	9.459	+	100.00
26.	Cefonicid	6.380	+	83.72
27.	Ceforanide	5.690	+	80.07
28.	Cefpiramide	6.574	+	84.75
29.	Cefpodoxime Proxetil	5.605	+	79.62
30.	Cefroxadine	1.745	+	59.22
31.	Cefteram	5.696	+	80.11
32.	Ceftibuten	4.767	+	75.20
33.	Ceftizoxime	4.662	+	74.64
34.	Cefuroxime	3.106	+	66.42
35.	Cefuzonam	8.924	+	97.17
36.	Cephalexin	1.690	+	58.93
37.	Cephalothin	3.261	+	67.24
38.	Cephapirin Sodium	2.785	+	64.72
39.	Cephradine	1.690	+	58.93
40.	Chloramphenicol	-1.042	-	44.49
41.	Chlortetracycline	2.032	+	60.74
42.	Ciprofloxacin	0.987	+	55.21
43.	Clarithromycin	3.726	+	69.70
44.	Clinafloxacin	-2.065	-	39.08
45.	Clindamycin	-0.814	-	45.70
46.	Clomocycline	1.719	+	59.09

Tabla 4.14. Continuación.

	<i>Molécula</i>	ΔP	<i>Clasif.</i>	<i>Probabilidad %</i>
47.	Cloxacillin	1.223	+	56.46
48.	Cyclacillin	2.888	+	65.27
49.	Dapsone	0.890	+	54.70
50.	Demeclocycline	2.324	+	62.29
51.	Dibekacin	3.934	+	70.79
52.	Dichloramine T	-1.326	-	42.99
53.	Difloxacin	1.607	+	58.49
54.	Dihydrostreptomycin	8.080	+	92.71
55.	Erythromycin	3.834	+	70.27
56.	Fenbenicillin	1.315	+	56.95
57.	Fleroxacin	5.832	+	80.83
58.	Florfenicol	1.295	+	56.84
59.	Floxacin	2.491	+	63.17
60.	Flumequine	1.163	+	56.15
61.	Fortimicin B	2.159	+	61.41
62.	Furaltadone	-2.091	-	38.95
63.	Gentamicin	3.980	+	71.04
64.	Hexedine	-3.238	-	32.89
65.	Imipenem	2.312	+	62.22
66.	Isepamicin	6.501	+	84.36
67.	Lenampicillin	3.043	+	66.08
68.	Leucomycins	1.142	+	56.04
69.	Lomefloxacin	4.536	+	73.98
70.	Lymecycline	4.482	+	73.69
71.	Mafenide	-0.406	-	47.85
72.	Meclocycline	3.432	+	68.14
73.	4'-(Methylsulfamoyl)sulfanilamide	1.984	+	60.49
74.	Mezlocillin	6.697	+	85.40
75.	Midecamycins	1.252	+	56.62
76.	Miloxacin	-0.388	-	47.95
77.	Minocycline	4.481	+	73.68
78.	Miokamycin	1.833	+	59.69
79.	Nadifloxacin	1.882	+	59.95
80.	Nafcillin Sodium	2.419	+	62.79
81.	Neomycin	5.456	+	78.84
82.	Nifuradene	-1.300	-	43.13
83.	Nifurfoline	-1.358	-	42.82
84.	Nifurpirinol	-1.931	-	39.80
85.	Nifurprazine	-0.994	-	44.75
86.	Nitrofurantoin	-0.827	-	45.63
87.	Norfloxacin	1.355	+	57.16
88.	Ofloxacin	2.979	+	65.75
89.	Oxacillin	2.575	+	63.61
90.	Oxytetracycline	5.979	+	81.60
91.	Paromomycin	5.785	+	80.58
92.	Pazufloxacin	1.267	+	56.70
93.	Pefloxacin	1.983	+	60.48
94.	Penicillin G	2.106	+	61.13
95.	Penicillin N	3.576	+	68.90

Tabla 4.14. Continuación.

	<i>Molécula</i>	ΔP	<i>Clasif.</i>	<i>Probabilidad %</i>
96.	Penicillin V	1.949	+	60.30
97.	Phenethicillin Potassium	2.317	+	62.25
98.	Phthalylsulfacetamide	0.742	+	53.92
99.	Pipacycline	4.529	+	73.94
100.	Piromidic Acid	-1.872	-	40.11
101.	Pivcefalexin	2.366	+	62.51
102.	Propicillin	1.949	+	60.30
103.	Quinacillin	4.984	+	76.34
104.	Rifamide	4.596	+	74.29
105.	Rifamycin SV	3.304	+	67.46
106.	Rifapentine	4.472	+	73.64
107.	Rosaranicin	0.927	+	54.90
108.	Rosoxacin	-1.979	-	39.54
109.	Salazosulfadimidine	-0.264	-	48.60
110.	Sancycline	2.601	+	63.75
111.	Sisomicin	3.614	+	69.11
112.	Solasulfone	5.403	+	78.56
113.	Spectinomycin	0.915	+	54.84
114.	Spiramycin	0.658	+	53.48
115.	Streptomycin	8.080	+	92.71
116.	Succinylsulfathiazole	1.883	+	59.95
117.	Succisulfone	0.631	+	53.34
118.	Sulfachlorpyridazine	-0.450	-	47.62
119.	Sulfachrysoidine	0.376	I	51.99
120.	Sulfadiazine	-0.209	-	48.90
121.	Sulfadicramide	0.743	+	53.93
122.	Sulfadoxine	0.686	+	53.62
123.	Sulfaethidole	2.455	+	62.98
124.	Sulfaguanidine	1.208	+	56.39
125.	Sulfalene	0.320	I	51.69
126.	Sulfaloxic Acid	0.584	+	53.09
127.	Sulfamethazine	0.527	+	52.79
128.	Sulfamethizole	2.823	+	64.92
129.	Sulfamethoxazole	0.584	+	53.09
130.	Sulfamethoxypyridazine	-0.103	-	49.45
131.	Sulfamidochrysoidine	-0.838	-	45.57
132.	Sulfamoxole	0.846	+	54.47
133.	Sulfanilamide	-0.038	-	49.80
134.	Sulfanilic Acid	0.067	I	50.36
135.	N4-Sulfanilylsulfanilamide	1.779	+	59.40
136.	Sulfantran	0.734	+	53.88
137.	Sulfaperine	0.159	I	50.84
138.	Sulfaphenazole	-1.101	-	44.18
139.	Sulfapyrazine	-0.209	-	48.90
140.	Sulfapyridine	-0.469	-	47.52
141.	Sulfasymazine	0.311	I	51.65
142.	Sulfathiazole	1.621	+	58.57
143.	Sulfathiourea	2.096	+	61.08
144.	Sulfisoxazole	1.058	+	55.59

Tabla 4.14. Continuación.

<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
145. Talampicillin	2.673	+	64.13
146. Taurolidine	0.522	+	52.76
147. Temafloxacin	3.119	+	66.48
148. Tetracycline	4.237	+	72.40
149. Thiazolsulfone	3.553	+	68.78
150. Tigemonam	4.679	+	74.73
151. Tosufloxacin	3.487	+	68.43
152. Trimethoprim	0.153	I	50.81
153. Xibornol	-1.237	-	43.46

Tabla 4.15. Resultados obtenidos para las 152 moléculas del grupo entrenamiento de inactivos.

	<i>Molécula</i>	<i>AP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1.	Acecarbromal	-0.869	-	45.41
2.	Acetaminophen	-1.449	-	42.34
3.	Acetohexamide	-1.040	-	44.50
4.	Adinazolam	-4.366	-	26.92
5.	Ahistan	-0.781	-	45.87
6.	Alfentanil	-2.417	-	37.23
7.	Althiazide	-0.579	-	46.94
8.	Aminochlorthenoxazin	-3.036	-	33.95
9.	Amlodipine	-2.436	-	37.12
10.	Amobarbital	-1.343	-	42.90
11.	Ampiroxicam	1.788	+	59.45
12.	Amolmetin Guacil	-2.974	-	34.28
13.	Antipyrine	-1.922	-	39.84
14.	Atorvastatin	-3.408	-	31.98
15.	Azacosterol	-2.984	-	34.23
16.	Azosemide	-1.751	-	40.74
17.	Benactyzine	-1.506	-	42.04
18.	Benazepril	-3.287	-	32.63
19.	Benorylate	-2.079	-	39.01
20.	Benzylmorphine	-4.023	-	28.74
21.	Bevantolol	-4.031	-	28.69
22.	Binifibrate	-3.999	-	28.86
23.	Brallobarbital	-0.920	-	45.14
24.	Bromfenac	-1.658	-	41.24
25.	Bucetin	-2.447	-	37.07
26.	Bunitrolol	-1.823	-	40.36
27.	Butacetin	-1.817	-	40.40
28.	Butethal	-1.711	-	40.96
29.	Buthiazide	-0.623	-	46.71
30.	Butriptyline	-3.821	-	29.80
31.	Captopril	-0.190	I	49.00
32.	Carbinoxamine	-4.219	-	27.70
33.	Carteolol	-1.614	-	41.47
34.	Carvedilol	-5.139	-	22.84
35.	Chlorhexadol	-2.277	-	37.97
36.	Chlorpheniramine	-4.322	-	27.15
37.	Chlorthenoxazin	-3.876	-	29.51
38.	Clocinazine	-7.221	-	11.83
39.	Clofenamide	-0.324	I	48.29
40.	Clomestron	-4.641	-	25.47
41.	Crotethamide	-0.764	-	45.96
42.	Cyclothiazide	-1.515	-	41.99
43.	Cyproheptadine	-4.557	-	25.91
44.	Delapril	-2.766	-	35.38
45.	Desipramine	-4.557	-	25.91
46.	Desomorphine	-2.709	-	35.68

Tabla 4.15. Continuación.

	<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
47.	Dihydrocodeinone Enol Acetate	-1.345	-	42.89
48.	Dioxadrol	-5.289	-	22.04
49.	Diphenylpyraline	-4.088	-	28.39
50.	Dipipanone	-5.547	-	20.68
51.	Disulfamide	-0.564	-	47.02
52.	Dothiepin	-2.467	-	36.96
53.	Doxylamine	-3.978	-	28.97
54.	Enalapril	-2.449	-	37.05
55.	Enfenamic Acid	-3.919	-	29.29
56.	Eptazocine	-2.447	-	37.07
57.	Etaqualone	-3.183	-	33.18
58.	Etodolac	-2.549	-	36.53
59.	Etonitazene	-2.835	-	35.01
60.	Etymemazine	-0.897	-	45.26
61.	Febarbamate	-1.191	-	43.70
62.	Felodipine	-3.854	-	29.63
63.	Femoxetine	-4.551	-	25.94
64.	Fenbufen	-3.234	-	32.91
65.	Feprazone	-3.284	-	32.64
66.	Flufenamic Acid	-0.007	I	49.96
67.	Flunitrazepam	-0.834	-	45.59
68.	Fluoresone	0.635	+	53.36
69.	Flurbiprofen	-0.575	-	46.96
70.	Furazabol	-2.144	-	38.67
71.	Glibornuride	0.381	+	52.01
72.	Gliquidone	-1.150	-	43.92
73.	Glybuzole	1.674	+	58.85
74.	Glymidine	-2.144	-	38.67
75.	Glypinamide	-1.863	-	40.15
76.	Guanadrel	-2.248	-	38.12
77.	Guanoxan	-1.880	-	40.06
78.	Haloxazolam	-1.050	-	44.45
79.	Hexethal	-1.926	-	39.82
80.	Hydrochlorothiazide	-1.094	-	44.21
81.	Hydroxyzine	-6.060	-	17.97
82.	Ibufenac	-1.711	-	40.96
83.	Indeloxazine Hydrochloride	-3.612	-	30.91
84.	Indoramin	-4.931	-	23.93
85.	Isomethadone	-3.813	-	29.85
86.	Isoxicam	2.060	+	60.89
87.	Ketoprofen	-2.498	-	36.80
88.	Levcromakalim	-0.300	I	48.41
89.	Levopropacetoperane	-3.339	-	32.35
90.	Lisinopril	-2.299	-	37.85
91.	Lofentanil	-3.604	-	30.95
92.	Loxoprofen	-2.235	-	38.18
93.	Mebhydroline	-4.189	-	27.86
94.	Meclozamine	-3.851	-	29.65
95.	Metazocine	-1.711	-	40.96

Tabla 4.15. Continuación.

<i>Molécula</i>	<i>AP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
96. Methyclothiazide	-1.507	-	42.03
97. Methyl-4-Pyridyl Ketone	-0.573	-	46.97
98. Methyl-4-thiazoleethanol-5	0.273	+	51.44
99. Metopon	-1.813	-	40.42
100. Metralindole	-2.400	-	37.31
101. Mofebutazone	-2.603	-	36.24
102. Moveltipril	0.009	+	50.05
103. Moxastine	-3.978	-	28.97
104. Muzolimine	-2.860	-	34.88
105. Nalbuphine	-2.862	-	34.87
106. Nefazodone	-5.009	-	23.52
107. Nicardipine	-0.518	-	47.26
108. Nicoclonate	-2.954	-	34.39
109. Nitrazepam	-2.973	-	34.28
110. Nitrendipine	0.285	+	51.51
111. Norpipanone	-6.021	-	18.18
112. Opipramol	-5.405	-	21.43
113. Oxaflozane	0.361	+	51.91
114. Oxametacine	-2.060	-	39.11
115. Parsalmide	-2.451	-	37.04
116. Phenobarbital	-2.235	-	38.18
117. Phenoperidine	-5.598	-	20.41
118. Phenopyrazone	-3.128	-	33.47
119. Phentolamine	-3.978	-	28.97
120. Phenylbutazone	-3.652	-	30.69
121. Phenylbutyramide	-1.922	-	39.84
122. Pizotyline	-2.728	-	35.58
123. Polythiazide	2.135	+	61.28
124. Prazosin	-2.150	-	38.63
125. Propacetamol	-1.766	-	40.67
126. Propoxyphene	-3.760	-	30.12
127. Protizinic Acid	0.060	+	50.32
128. Pyrisuccideanol	-0.922	-	45.12
129. Pyrrobutamine	-6.162	-	17.43
130. Rescimetol	-2.445	-	37.07
131. Reserpine	-1.077	-	44.31
132. Ritanserin	-2.745	-	35.49
133. Ronifibrate	-4.526	-	26.08
134. Salacetamide	-1.343	-	42.90
135. Salicylsulfuric Acid	0.182	+	50.96
136. Secobarbital	-1.499	-	42.07
137. Sulfinalol	-2.088	-	38.96
138. Suprofen	-0.619	-	46.73
139. Talniflumate	0.192	+	51.01
140. Tandospirone	-3.829	-	29.76
141. Thenyldiamine	-3.194	-	33.12
142. Timolol	-0.315	I	48.33
143. Tolbutamide	-0.516	-	47.27
144. Tolfenamic Acid	-3.664	-	30.63

Tabla 4.15. Continuación.

<i>Molécula</i>	<i>AP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
145. Tolmetin	-1.656	-	41.25
146. Toloxatone	-1.762	-	40.69
147. Tolpropamine	-3.713	-	30.37
148. Triamterene	-1.937	-	39.76
149. Triazolam	-5.977	-	18.41
150. Trimazosin	0.057	+	50.30
151. Tritoqualine	0.500	+	52.64
152. Vinbarbital	-1.343	-	42.90

Tabla 4.16. Resultados obtenidos para las 64 moléculas del grupo test de activos.

<i>Molécula</i>	<i>AP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1. Amdinocillin	1.624	+	58.58
2. Amoxicillin	3.728	+	69.71
3. Apicycline	6.474	+	84.22
4. Arbekacin	5.082	+	76.86
5. Azidocillin	4.185	+	72.12
6. Azithromycin	3.661	+	69.35
7. Benzylpenicillinic Acid	2.106	+	61.13
8. Carumonam	6.128	+	82.39
9. Cefamandole	4.962	+	76.23
10. Cefditoren	6.529	+	84.51
11. Cefmenoxime	8.611	+	95.52
12. Cefoperazone	8.558	+	95.23
13. Cefotiam	7.820	+	91.34
14. Cefprozil	1.899	+	60.04
15. Ceftezole	7.040	+	87.21
16. Ceftriaxone	8.776	+	96.39
17. Cephacetrile Sodium	2.432	+	62.85
18. Cephalosporin C	2.641	+	63.96
19. Clometocillin	1.438	+	57.60
20. Diathymosulfone	-0.802	-	45.76
21. Dicloxacillin	-1.348	-	42.88
22. Doxycycline	5.028	+	76.58
23. Enoxacin	1.355	+	57.16
24. N2-Formylsulfisomidine	0.523	+	52.76
25. Fortimicin A	4.257	+	72.50
26. Grepafloxacin	1.722	+	59.10
27. Hetacillin	3.311	+	67.50
28. Lincomycin	1.541	+	58.14
29. Meropenem	4.671	+	74.69
30. Methacycline	5.028	+	76.58
31. Methenamine	-2.134	-	38.72
32. Methicillin Sodium	3.263	+	67.24
33. Micronomicin	3.247	+	67.16
34. Nalidixic Acid	-0.920	-	45.14
35. Nifuratel	-1.143	-	43.96
36. Nifurtoinol	-0.721	-	46.19
37. Oxolinic Acid	-0.916	-	45.16
38. Penamecillin	1.685	+	58.91
39. Penicillin O	3.624	+	69.15
40. Phthatylsulfathiazole	1.464	+	57.74
41. Pipemidic Acid	-1.459	-	42.29
42. Pivampicillin	3.931	+	70.78
43. Ribostamycin	4.727	+	74.98
44. Rifampin	5.944	+	81.42
45. Rokitamycin	1.039	+	55.49
46. Roxithromycin	2.771	+	64.64
47. Sparfloxacin	5.840	+	80.87
48. Sulbenicillin	4.573	+	74.17
49. Sulfabenzamide	-0.412	-	47.82

Tabla 4.16. Continuación.

<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
50. Sulfacytine	0.525	+	52.77
51. Sulfadimethoxine	0.474	I	52.51
52. Sulfaguanole	0.942	+	54.98
53. Sulfamerazine	0.159	I	50.84
54. Sulfamethomidine	0.525	+	52.77
55. 4-Sulfanilamidosalicylic Acid	0.324	I	51.71
56. p-Sulfanilylbenzylamine	0.522	+	52.76
57. N-Sulfanilyl-3,4-xylamide	0.430	I	52.27
58. Sulfaproxyline	0.214	I	51.13
59. Sulfasomizole	1.989	+	60.51
60. Sulfisomidine	0.787	+	54.16
61. Tetroxoprim	-0.528	-	47.21
62. Thiamphenicol	-0.154	-	49.18
63. Ticarcillin	4.624	+	74.44
64. Trovafloxacin	3.855	+	70.37

Tabla 4.17. Resultados obtenidos para las 64 moléculas del grupo test de inactivos.

<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
1. Aminopyrine	-1.554	-	41.78
2. Aprobarbital	-0.975	-	44.85
3. Aranidipine	0.139	+	50.73
4. Arbutin	-0.916	-	45.16
5. Bamipine	-4.665	-	25.34
6. Bunazosin	-2.408	-	37.27
7. Bupropion	-2.455	-	37.02
8. Chloralose-alfa	-3.571	-	31.13
9. Chlorpropamide	-0.757	-	46.00
10. Chlortalidone	-2.319	-	37.74
11. Cicletanine	-2.901	-	34.67
12. Clemastine	-4.485	-	26.29
13. Clinofibrate	-5.124	-	22.91
14. Clobenzepam	-3.880	-	29.49
15. Clometacin	-2.001	-	39.42
16. Dezocine	-2.815	-	35.12
17. Dimetacrine	-2.933	-	34.50
18. Droxicam	1.747	+	59.24
19. Ebastine	-6.192	-	17.27
20. Enalaprilat	-1.973	-	39.57
21. Ethchlorvynol	-3.224	-	32.96
22. Ethylmethylthiambutene	-0.162	I	49.14
23. Etozolin	0.270	+	51.42
24. Gemfibrozil	-1.973	-	39.57
25. Glafenine	-4.303	-	27.26
26. Glimepiride	-0.894	-	45.27
27. Guanazodine	-2.717	-	35.64
28. Heptabarbital	-2.603	-	36.24
29. Isothipendyl	-1.310	-	43.08
30. Isoxepac	-2.654	-	35.97
31. Lofepramine	-5.906	-	18.78
32. Lofexidine	-6.522	-	15.52
33. Lormetazepam	-5.174	-	22.65
34. L-Tryptophan	-2.079	-	39.01
35. Mefenamic Acid	-2.447	-	37.07
36. Mepindolol	-1.563	-	41.74
37. Meptazinol	-3.288	-	32.62
38. Methotrimeprazine	-0.580	-	46.93
39. Methyldiphenhydramine-p	-3.610	-	30.92
40. Methyprylon	-1.292	-	43.17
41. Moclobemide	-4.373	-	26.89
42. Nilvadipine	1.537	+	58.13
43. Noxiptilin	-2.882	-	34.77
44. p-Bromoacetanilide	-1.554	-	41.78
45. Phenazopyridine Hydrochloride	-2.565	-	36.44
46. Phenyltoloxamine	-3.978	-	28.97
47. Pildralazine	-1.142	-	43.97
48. Pirazolac	-3.480	-	31.60
49. Propallylonal	-0.396	I	47.91

Tabla 4.17. Continuación.

<i>Molécula</i>	<i>ΔP</i>	<i>Clasif.</i>	<i>Probabilidad %</i>
50. Propizepine	-2.297	-	37.86
51. Quinethazone	-2.662	-	35.93
52. Rescinnamine	-2.496	-	36.81
53. Salicylamide-Acetic Acid	-1.499	-	42.07
54. Simfibrate	-4.972	-	23.72
55. Sufentanil	-2.990	-	34.20
56. Syrosingopine	-0.545	-	47.12
57. Thonzylamine	-3.823	-	29.79
58. Tilisolol	-0.880	-	45.35
59. Tofenacin	-3.978	-	28.97
60. Tolazamide	-1.254	-	43.37
61. Tramadol	-2.605	-	36.23
62. Tripamide	-3.159	-	33.30
63. Vinylbital	-1.343	-	42.90
64. Ximoprofen	-2.343	-	37.61

La tabla 4.18 representa la espectancia de activos para los grupos de entrenamiento (E_a) y test (E_{at}), así como la espectancia de inactivos para los mismos grupos (E_i , E_{it}).

Tabla 4.18. Valores de espectancia de actividad e inactividad obtenidos para los grupos de entrenamiento y test.

Intervalos de valores de ΔP	Moléculas Grupo Aprendizaje		Moléculas Grupo Test		Valores de Espectancia			
	Activos	Inactivos	Activos	Inactivos	Ea	Ei	Eat	Eit
-8 a -7	0	1	0	0	0	0.0066	0	0
-7 a -6	0	3	0	2	0	0.0197	0	0.0313
-6 a -5	0	7	0	3	0	0.0461	0	0.0469
-5 a -4	0	13	0	5	0	0.0855	0	0.0781
-4 a -3	1	26	0	10	0.0056	0.1699	0	0.1563
-3 a -2	3	34	1	20	0.0160	0.2194	0.0119	0.3078
-2 a -1	9	33	3	11	0.0483	0.2050	0.0040	0.1642
-1 a 0	14	20	7	9	0.0809	0.1205	0.0959	0.1268
0 a 1	24	11	11	2	0.1463	0.0626	0.1667	0.0267
1 a 2	25	2	10	2	0.1613	0.0113	0.1515	0.0270
2 a 3	20	2	4	0	0.1290	0.0116	0.0625	0
3 a 4	18	0	8	0	0.1176	0	0.1250	0
4 a 5	17	0	7	0	0.1111	0	0.1094	0
5 a 6	11	0	5	0	0.0719	0	0.0781	0
6 a 7	7	0	3	0	0.0458	0	0.0469	0
7 a 8	0	0	2	0	0.0000	0	0.0313	0
8 a 9	3	0	3	0	0.0196	0	0.0469	0
9 a 10	1	0	0	0	0.0065	0	0	0
<i>Total moléculas</i>	<i>153</i>	<i>152</i>	<i>64</i>	<i>64</i>				

La figura 4.6 refleja el correspondiente diagrama de distribución farmacológica para los intervalos de valores de ΔP .

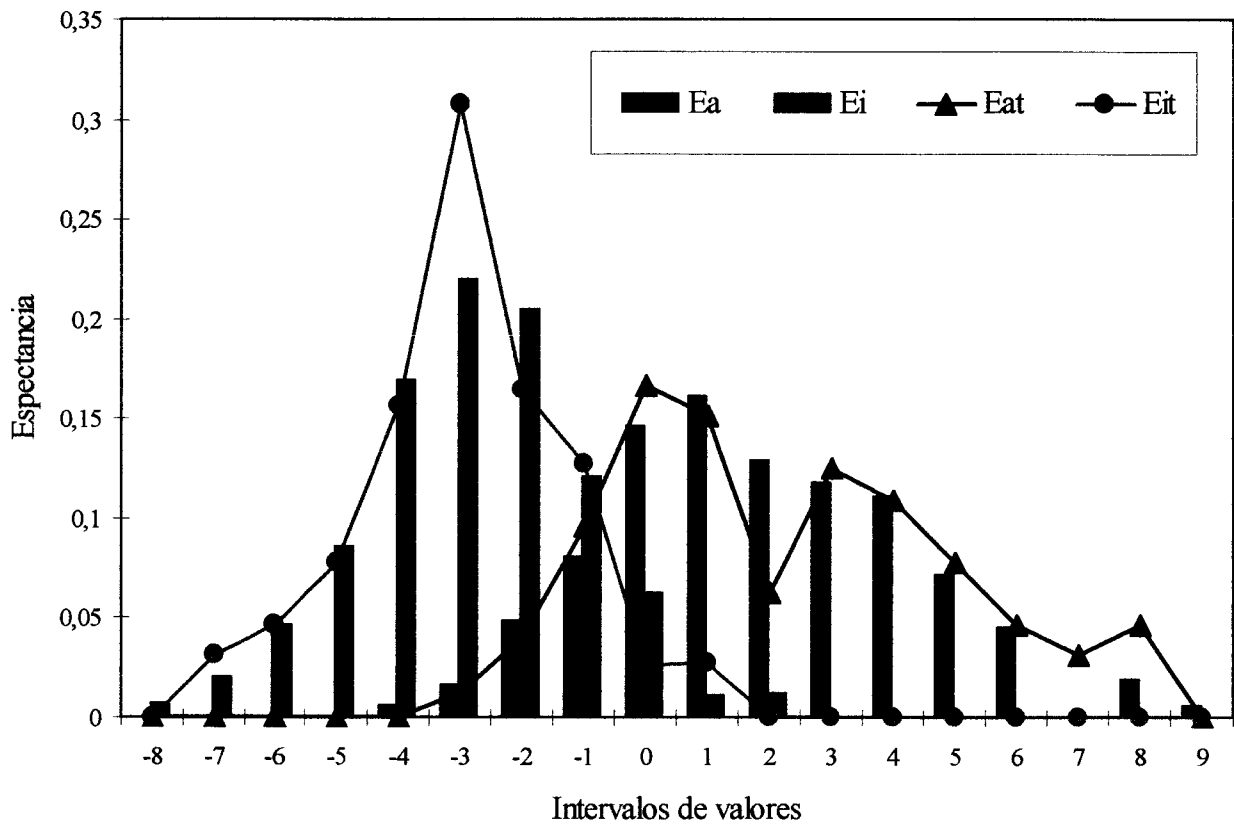


Figura 4.6. Diagrama de distribución farmacológica para los valores de la función discriminante.

4.3. BÚSQUEDA GUIADA Y SELECCIÓN DE ESTRUCTURAS

Una vez obtenida la función discriminante y la arquitectura de la red neuronal artificial más eficaz para llevar a cabo el proceso de discriminación, el siguiente paso es realizar una búsqueda guiada de estructuras con teórica actividad antibacteriana.

En primer lugar hacemos uso de la información topológica que contiene la función discriminante, de modo que utilizamos como parámetros de búsqueda aquellos índices topológico-estructurales que contribuyen de forma positiva en dicha función.

Función discriminante

$$\Delta P = -0.3680 V^2 + 0.2602 D^{N8} + 0.1057 D^{O5} + 0.5739 D^{S2} + 0.4680 D^{F5} - 0.6088 D^{Cl4} + 0.0742$$

Como podemos observar cuatro de los seis índices de la función discriminante contribuyen de forma positiva a la misma, estos índices son:

- D^{N8} : sumatorio de distancias 8 del átomo de N
- D^{O5} : sumatorio de distancias 5 del átomo de O
- D^{S2} : sumatorio de distancias 2 del átomo de S
- D^{F5} : sumatorio de distancias 5 del átomo de F

mientras que la presencia de átomos de Cl influiría de manera negativa.

En la figura 4.7 se muestran los fragmentos estructurales derivados de los cuatro índices que contribuyen de forma positiva a la función discriminante que utilizamos como parámetros de búsqueda de nuevas estructuras con teórica actividad antibacteriana.

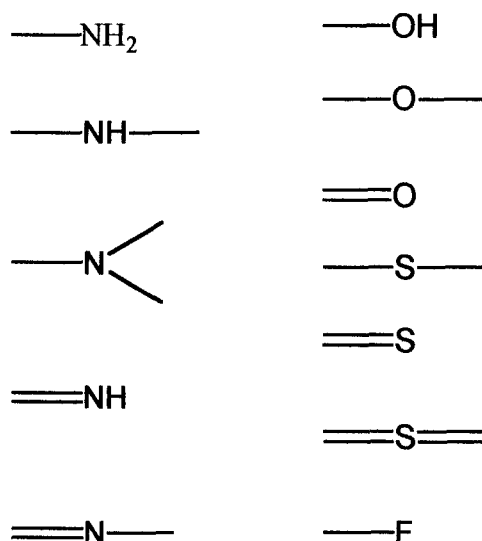


Figura 4.7. Parámetros de búsqueda.

Para realizar la búsqueda disponemos del módulo CS ChemFinder 5.1 del paquete CS ChemOffice, el cual nos permite establecer parámetros de búsqueda constituidos por los fragmentos de estructuras antes expuestos. Con dichos fragmentos la herramienta realiza una selección de estructuras entre bases de datos de amplia utilización en química, como son Sigma-Aldrich, Fluka, Fisher, Avocado, etc., de forma que nos permite obtener un listado de moléculas que contienen estos fragmentos (uno o varios) proporcionándonos su correspondiente descripción SMILES, número C.A.S., peso molecular y algunas propiedades físicas de las mismas.

Como resultado de la búsqueda se obtuvieron numerosas moléculas, de las que fueron excluidas aquellas en cuya composición formaban parte átomos que por sí mismos pudieran presentar un efecto tóxico (Hg, Pb, As, Sb, Bi, Ni, etc.). Asimismo se excluyeron las moléculas constituidas por grandes polímeros agregados de macromoléculas. A partir de la información de dichas moléculas existentes en los catálogos comerciales, también excluimos aquellas estructuras de las que se había descrito efecto teratógeno, mutagénico,

cancerígeno o citotóxico, así como aquellas de las que se hubiese descrito algún tipo de actividad farmacológica.

A cada una de las moléculas restantes se les calculó los correspondientes índices topológico-estructurales a partir del formato SMILES. Posteriormente se les aplicó la función discriminante con el propósito de seleccionar aquellas moléculas que mejores valores de ΔP presentasen para dicha función. El resultado final fueron las siguientes moléculas.

Tabla 4.19. Valores de ΔP para las moléculas seleccionadas.

<i>Nombre molécula</i>	<i>Número CAS</i>	<i>Valor de ΔP</i>
1. Methylthymol blue, sodium salt	1945-77-3	35.657
2. Glycine, N,N'-[3H-2,1-benzoxathiol-3-ylidenebis[(6-hydroxy-5-methyl-3,1-phenylene)methylene]]bis[N-(carboxymethyl)-, S,S-dioxide, tetrasodium salt	3618-43-7	28.230
3. o-Cresolphthalein complexone	2411-89-4	17.953
4. 5-Notroorotic acid potasium, salt monohydrate	60779-49-9	14.517
5. Tetrazolium violet	1719-71-7	11.159
6. 1,1'-(Methylenedi-4,1-phenylene) bismaleimide	13676-54-5	9.726
7. N,N'-(1,3-phenylene)dimalimide	3006-93-7	8.654
8. Ellagic acid	476-66-4	6.189
9. 2',3'-Di-O-benzoyluridine	50408-20-3	6.170
10. Quinoline Yellow	8004-92-0	6.138
11. 2,3,4-Trifluoroaniline	3862-73-5	4.948
12. Ethylenediaminetetraacetic dianhydride	23911-25-3	3.798
13. Direct Yellow 27	10190-68-8	2.343
14. 5-Phenyl-2-(2-((5-phenyl-3-(4-sulfo-butyl)-2(3H)-benzoxazolylidene)methyl)-1-butenyl)-3-(4-sulfo-butyl)benzoxazolium hydroxide, inner salt, sodium salt	35501-76-9	1.672
15. N-[4-(2-Benzoxazolyl)phenyl]maleimide	16707-41-8	1.438
16. 2-Benzofurancarboxylic acid	496-41-3	1.071

Estas 16 moléculas se procesaron posteriormente mediante la red neuronal entrenada, obteniendo los siguientes resultados:

Tabla 4.20. Valores de salida de la red neuronal para las moléculas seleccionadas.

<i>Nombre molécula</i>	<i>Número CAS</i>	<i>Valor de salida</i>
1. Direct Yellow 27	10190-68-8	1
2. N-[4-(2-Benzoxazolyl)phenyl]maleimide	16707-41-8	1
3. 1,1'-(Methylenedi-4,1-phenylene) bismaleimide	13676-54-5	1
4. 2',3'-Di-O-benzoyluridine	50408-20-3	1
5. Ellagic acid	476-66-4	1
6. 5-Notroorotic acid potasium, salt monohydrate	60779-49-9	1
7. Quinoline Yellow	8004-92-0	0.99999
8. Methylthymol blue, sodium salt	1945-77-3	0.99960
9. o-Cresolphthalein complexone	2411-89-4	0.99936
10. Glycine, N,N'-[3H-2,1-benzoxathiol-3-ylidenebis[(6-hydroxy-5-methyl-3,1-phenylene)methylene]]bis[N-(carboxymethyl)-, S,S-dioxide, tetrasodium salt	3618-43-7	0.99925
11. Ethylenediaminetetraacetic dianhydride	23911-25-3	0.99568
12. 5-Phenyl-2-(2-((5-phenyl-3-(4-sulfobutyl)-2(3H)-benzoxazolylidene)methyl)-1-butenyl)-3-(4-sulfobutyl)benzoxazolium hydroxide, inner salt, sodium salt	35501-76-9	-1
13. Tetrazolium violet	1719-71-7	-1
14. 2-Benzofurancarboxylic acid	496-41-3	-0.9887
15. 2,3,4-Trifluoroaniline	3862-73-5	-0.96101
16. N,N'-(1,3-phenylene)dimalleimide	3006-93-7	-0.99996

El criterio de clasificación de estas moléculas fue el mismo que el establecido para llevar a cabo el proceso de entrenamiento de la red neuronal descrito en el apartado 4.2.1.

<i>Valor de salida de la red</i>	<i>Clasificación</i>
Entre 1 y 0.5	Activa (correctamente clasificada)
Entre 0 y -1	Inactiva (incorrectamente clasificada)
Entre 0.5 y 0	Indeterminada

Teniendo en cuenta este criterio de clasificación puede observarse que de las 16 moléculas 11 de ellas son clasificadas como teóricos antibacterianos, cuyas estructuras se detallan a continuación.

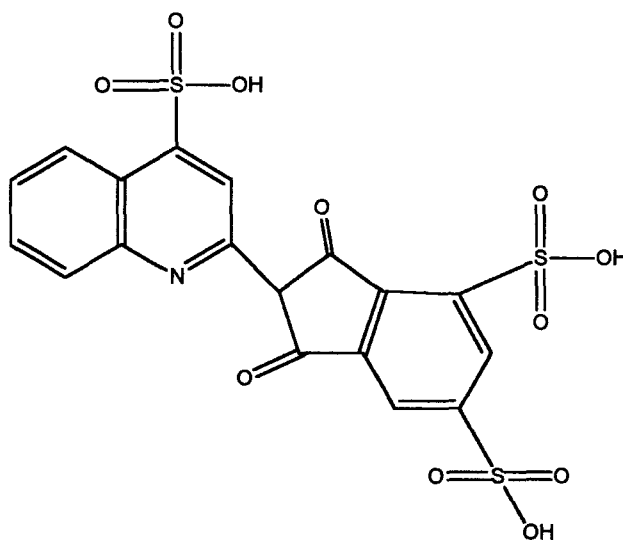


Figura 4.8. N° C.A.S. 8004-92-0.

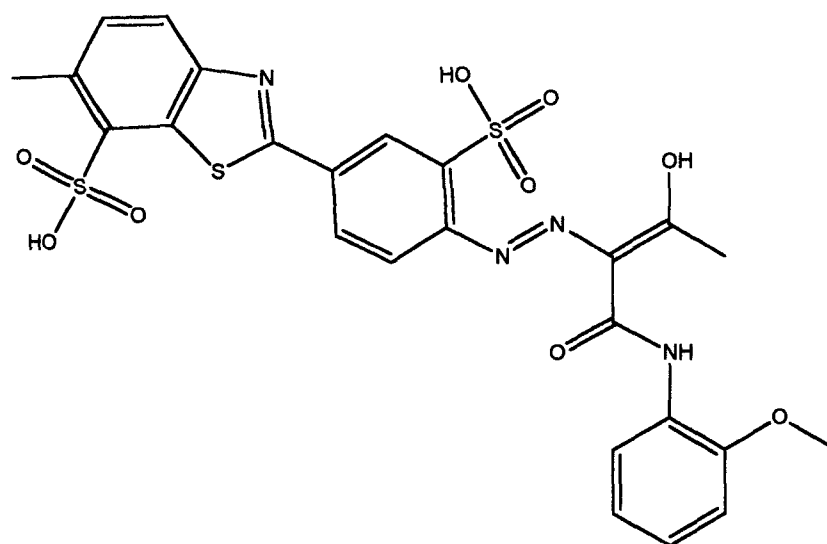


Figura 4.9. N° C.A.S. 10190-68-8.

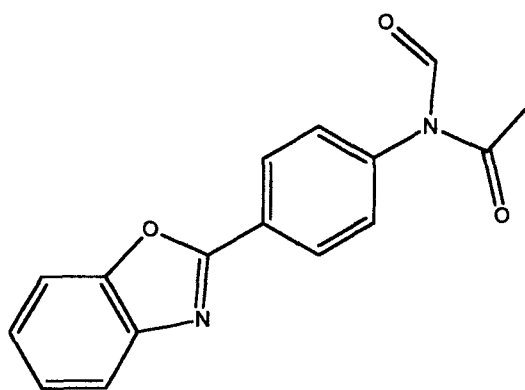


Figura 4.10. N° C.A.S.16707-41-8.

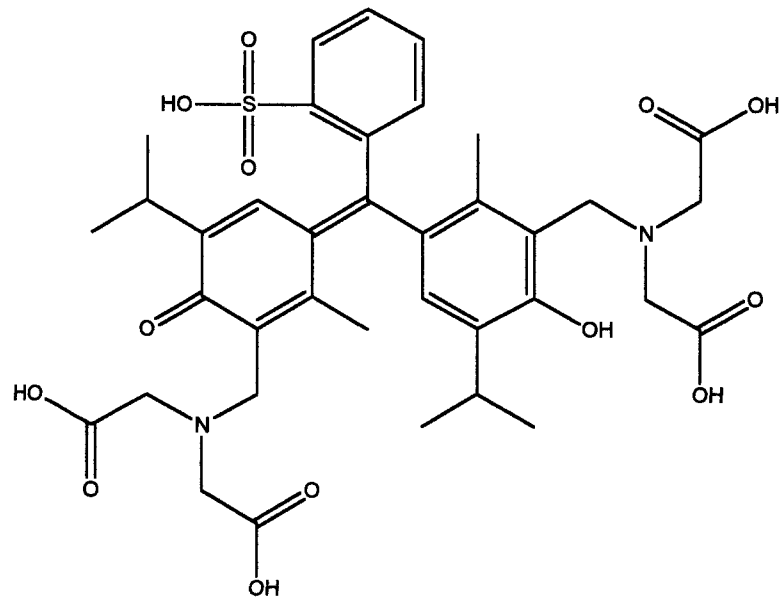


Figura 4.11. N° C.A.S. 1945-77-3.

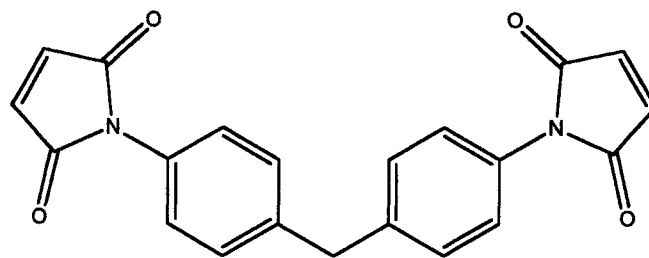


Figura 4.12. N° C.A.S. 13676-54-5.

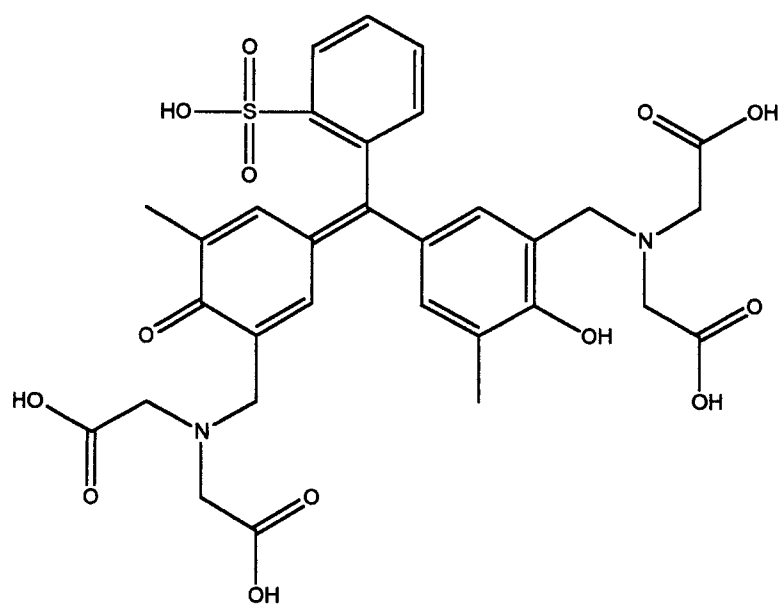


Figura 4.13. N° C.A.S. 3618-43-7.

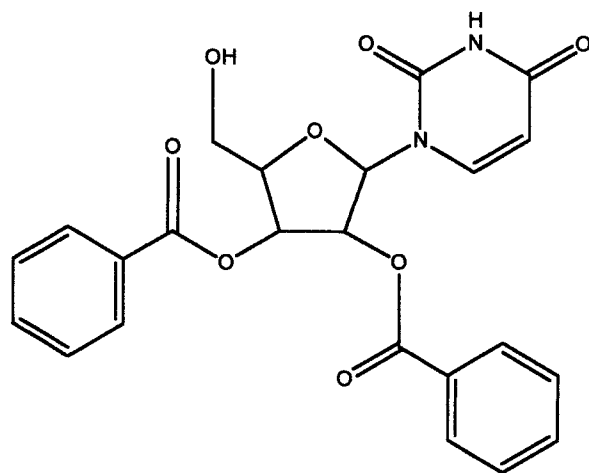


Figura 4.14. N° C.A.S. 50408-20-3.

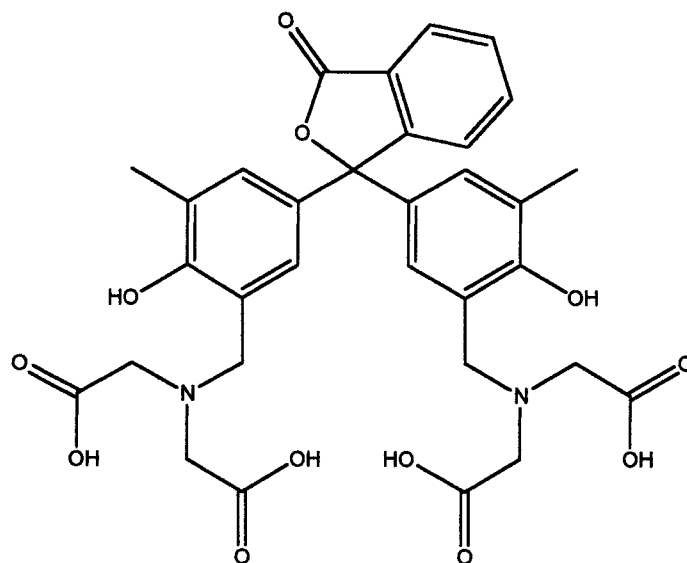


Figura 4.15. N° C.A.S. 2411-89-4.

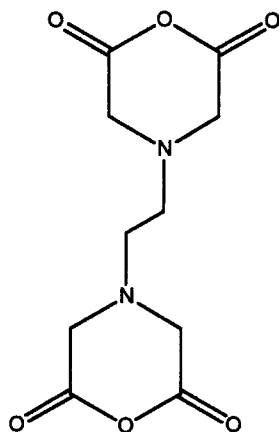


Figura 4.16. N° C.A.S. 23911-25-3.

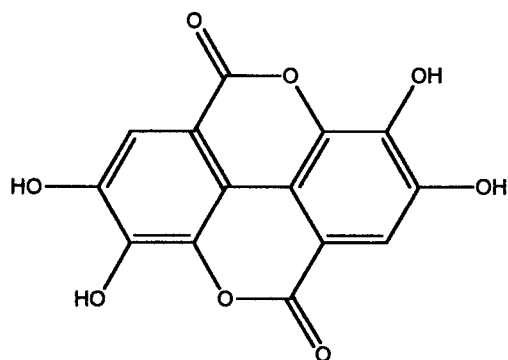


Figura 4.17. N° C.A.S. 476-66-4.

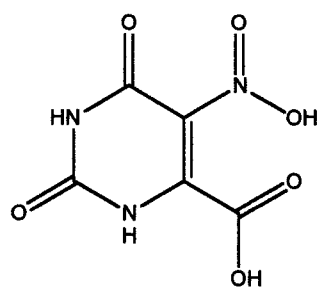


Figura 4.18. N° C.A.S.60779-49-9.

4.4. ENSAYOS MICROBIOLÓGICOS

Las 11 moléculas seleccionadas como teóricos antibacterianos fueron sometidas a ensayos microbiológicos con el fin de confirmar su actividad antibacteriana.

Previo al ensayo microbiológico las moléculas fueron sometidas a ensayos de solubilidad con objeto de determinar cual era el disolvente más idóneo para cada una de ellas. De igual forma se comprobó la solubilidad de los dos productos de referencia que fueron el ácido nalidíxico (C.A.S.: 3374-05-8) y la cefalosporina C (C.A.S.:12567-06-5).

Con objeto de comprobar que los disolventes no inhibiesen el crecimiento bacteriano se relizaron ensayos microbiológicos de combinaciones de estos frente a cepas de *E. coli* y *S. aureus*, con los resultados que se muestran en la tabla 4.21.

La tabla 4.22 muestra los disolventes utilizados para cada una de las moléculas a excepción de la molécula Ellagic Acid (476-66-4) la cual se mostró insoluble, por lo que las moléculas que definitivamente se sometieron a ensayos microbiológicos fueron las 10 restantes.

Como ya se nombró en el apartado de material y métodos se utilizaron cuatro cepas de microorganismos para llevar a cabo los ensayos de susceptibilidad antibacteriana:

- *Enterococcus faecalis* CECT 795 [ATCC 29212]
- *Staphylococcus aureus* CECT 240 [ATCC 6538P]
- *Escherichia coli* CECT 405 [ATCC 10536]
- *Pseudomonas aeruginosa* CECT 108 [ATCC 27853]

Tabla 4.21. Resultados de susceptibilidad de *E. coli* y *S. aureus* frente a distintas combinaciones de disolventes.

	Agua estéril (mL)	Polietilenglicol 400 (mL)	Tween 80 (mL)	HCl 0.1 M (mL)	NaOH 0.1M (mL)	DMSO (mL)	CRECIMIENTO
1	10	-	-	-	-	-	+
2	-	10	-	-	-	-	-
3	-	-	-	-	-	10	-
4	7.5	2.5	-	-	-	-	+
5	5	5	-	-	-	-	-
6	2.5	7.5	-	-	-	-	-
7	7.5	-	-	-	-	2.5	-
8	5	-	-	-	-	5	-
9	2.5	-	-	-	-	7.5	-
10	-	7.5	-	-	-	2.5	-
11	-	5	-	-	-	5	-
12	-	2.5	-	-	-	7.5	-
13	9.5	-	-	-	-	0.5	+
14	9	-	-	-	-	1	+
15	9	0.5	-	-	-	0.5	+
16	8	-	-	-	-	2	+
17	8	1.5	-	-	-	0.5	+
18	8	1	-	-	-	1	+
19	8	0.5	-	-	-	1.5	+
20	7.5	2	-	-	-	0.5	+
21	7	2	-	-	-	1	-
22	7	1	-	-	-	2	-
23	8	-	2	-	-	-	+
24	8	-	-	2	-	-	+
25	8	-	-	-	2	-	+
26	8	2	-	-	-	-	+
27	8	1	1	-	-	-	+

Tabla 4.22. Disolventes de las 10 moléculas sometidas a ensayos microbiológicos.

<i>Molécula (C.A.S.)</i>	<i>Disolvente</i>	<i>Proporción de disolvente</i>
1. 8004-92-0	Agua	10
2. 10190-68-8	Agua	10
3. 16707-41-8	Agua / Tween 80 / PEG	8 : 1 : 1
4. 1945-77-3	Agua	10
5. 13676-54-5	Agua / Tween 80 / PEG	8 : 1 : 1
6. 3618-43-7	Agua	10
7. 50408-20-3	Agua / Tween 80 / PEG	8 : 1 : 1
8. 2411-89-4	Agua / DMSO	8 : 2
9. 23911-25-3	Agua / Tween 80	8 : 2
10. 60779-49-9	Agua	10

La tabla 4.23 refleja los valores de Concentración Mínima Inhibitoria (CMI) para los 10 productos ensayados, frente a las cuatro cepas de microorganismos:

Tabla 4.23. Valores de Concentración Mínima Inhibitoria para los productos ensayados y los de referencia.

Molécula (C.A.S.)	C.M.I. ($\mu\text{g/mL}$)			
	<i>S. aureus</i>	<i>E. faecalis</i>	<i>E. coli</i>	<i>P. aeruginosa</i>
1. 8004-92-0	> 1000	> 1000	> 1000	> 1000
2. 10190-68-8	> 1000	> 1000	> 1000	> 1000
3. 16707-41-8	15.7	31.3	500	> 1000
4. 1945-77-3	> 1000	> 1000	> 1000	> 1000
5. 13676-54-5	125	250	> 500	> 500
6. 3618-43-7	> 1000	> 1000	> 1000	> 1000
7. 50408-20-3	> 1000	> 1000	> 1000	> 1000
8. 2411-89-4	> 1000	>1000	1000	1000
9. 23911-25-3	> 1000	1000	500	> 1000
10. 60779-49-9	> 1000	> 1000	> 1000	> 1000
Cefalosporina C	31.3	> 1000	31.3	> 500
Ácido Nalidíxico	62.5	1000	15.7	500

En las figuras 4.19, 4.20, 4.21 y 4.22 se muestran las placas de microtitulación para cada una de las cepas de microorganismos donde se detalla la CMI de los productos activos.

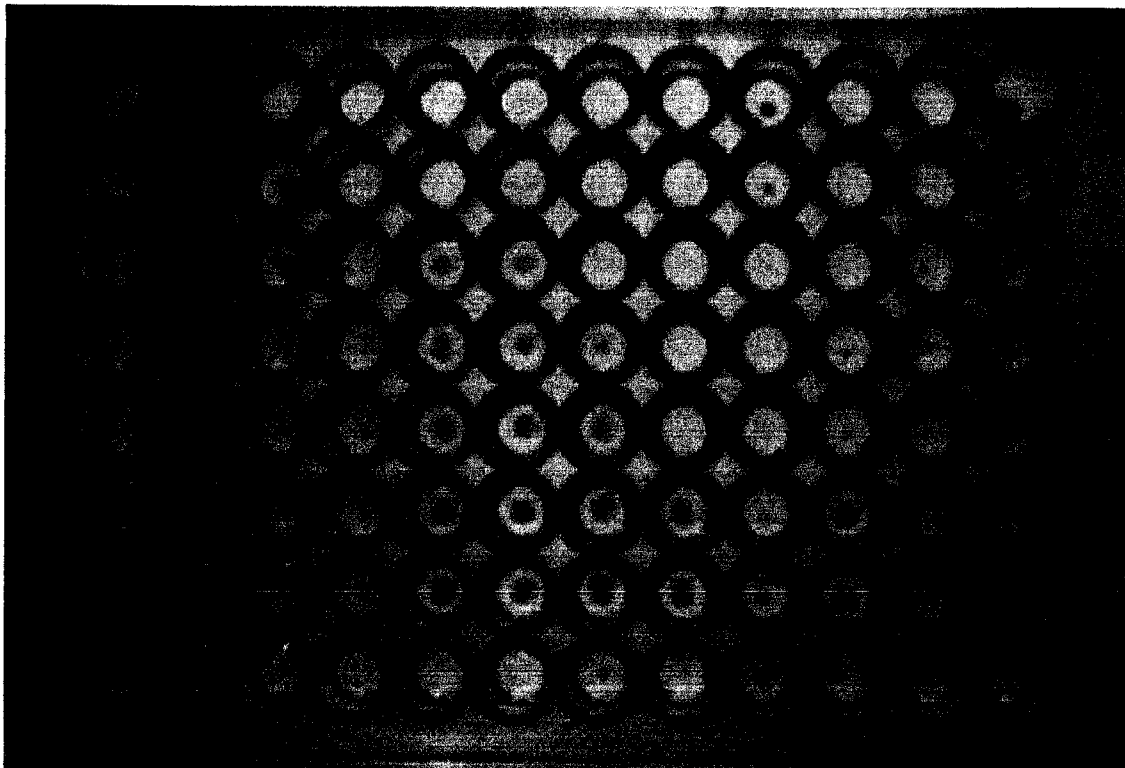
Las columnas de la 1 a la 9 corresponden a productos, la 10 y la 11 para las referencias cefalosporina C y ácido nalidíxico respectivamente y la columna 12 alberga el control positivo de crecimiento.

Las filas de la A a la G contienen diluciones dobles de cada producto hasta alcanzar siete concentraciones diferentes por cada fila. Los pocillos del H1 al H11 se utilizan como control negativo.

Las concentraciones de partida en la fila A son distintas para cada uno de los productos.

En cuanto a los productos de referencia para la cefalosporina C se parte siempre de una concentración de 500 $\mu\text{g}/\text{mL}$, mientras que para el ácido nalidíxico partimos de una concentración de 250 $\mu\text{g}/\text{mL}$ salvo en la placa de *E. faecalis* que se parte de 1000 $\mu\text{g}/\text{mL}$.

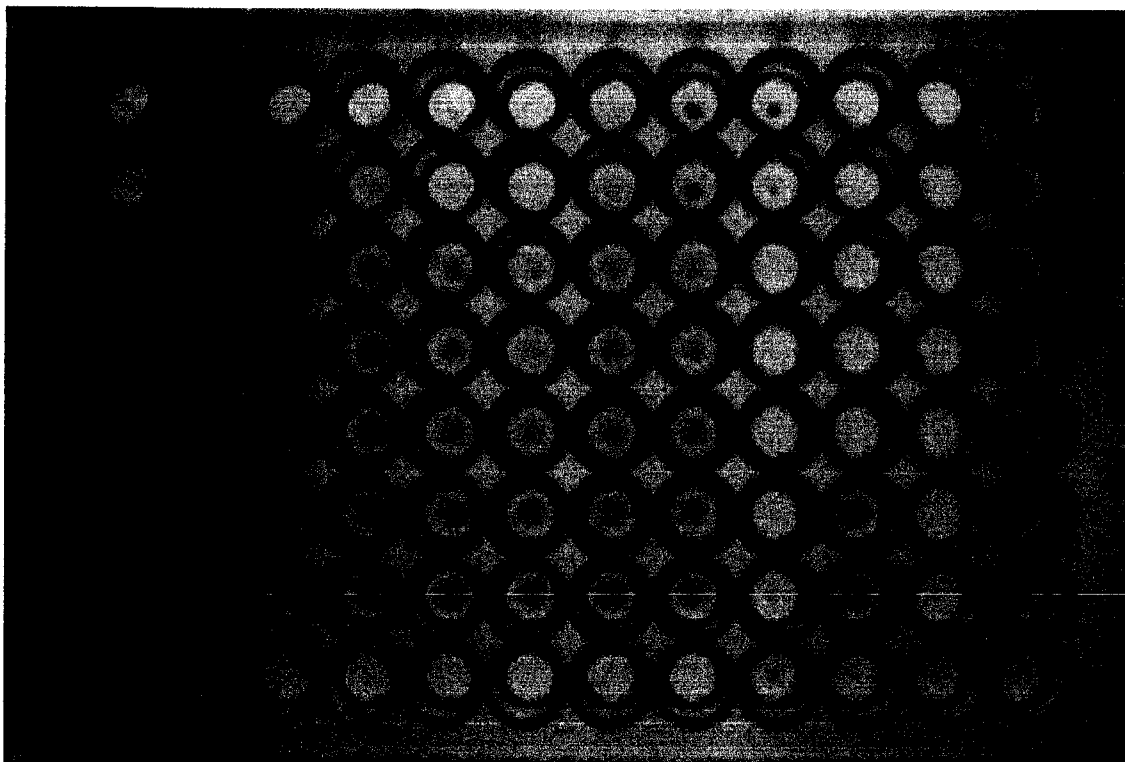
Figura 4.19. Placa de microtitulación para *Staphylococcus aureus* CECT 240.



Para esta cepa de microorganismo resultaron activos los siguientes productos:

- **13676-54-5:** Columna 7; partimos del pocillo A7 con una concentración de 500 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo C7 y corresponde a 125 $\mu\text{g/mL}$.
- **16707-41-8:** Columna 8; partimos del pocillo A8 con una concentración de 250 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo E8 y corresponde a 15.7 $\mu\text{g/mL}$.

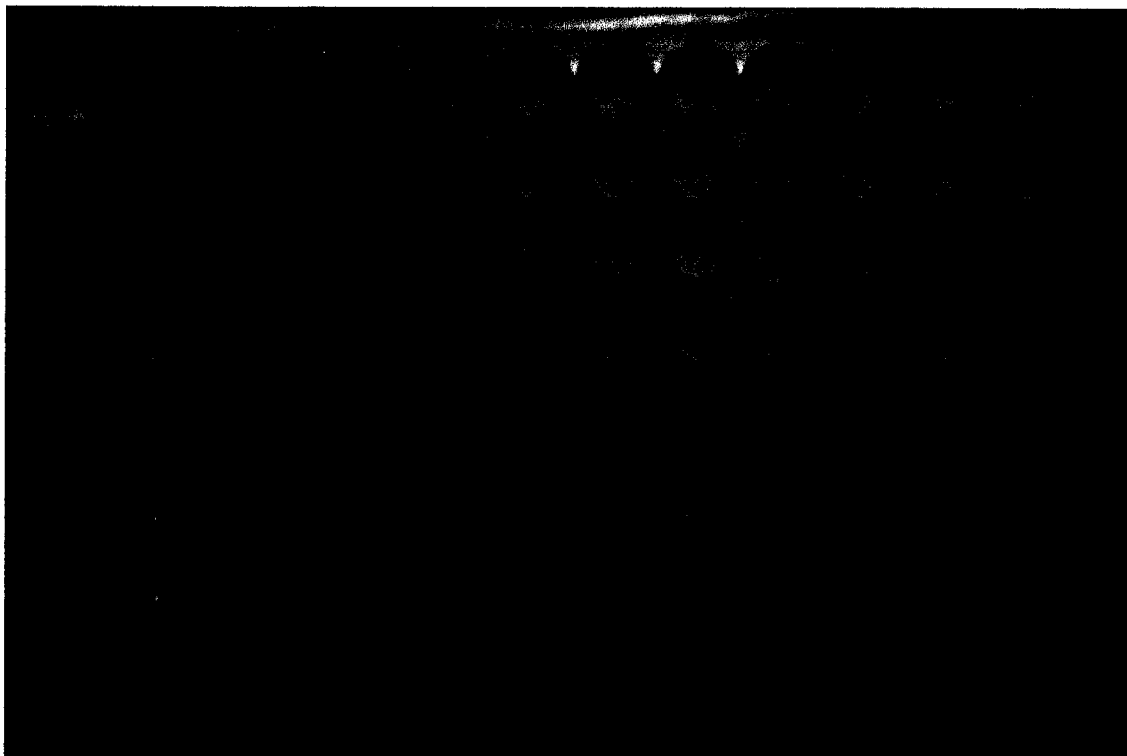
Figura 4.20. Placa de microtitulación para *Escherichia coli* CECT 405.



Para esta cepa de microorganismo resultaron activos los siguientes productos:

- **2411-89-4:** Columna 3; partimos del pocillo A3 con una concentración de 1000 $\mu\text{g/mL}$; su CMI corresponde a este mismo pocillo, luego es 1000 $\mu\text{g/mL}$.
- **23911-25-3:** Columna 6; partimos del pocillo A6 con una concentración de 1000 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo B6 y corresponde a 500 $\mu\text{g/mL}$.
- **16707-41-8:** Columna 8; partimos del pocillo A8 con una concentración de 1000 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo B8 y corresponde a 500 $\mu\text{g/mL}$.

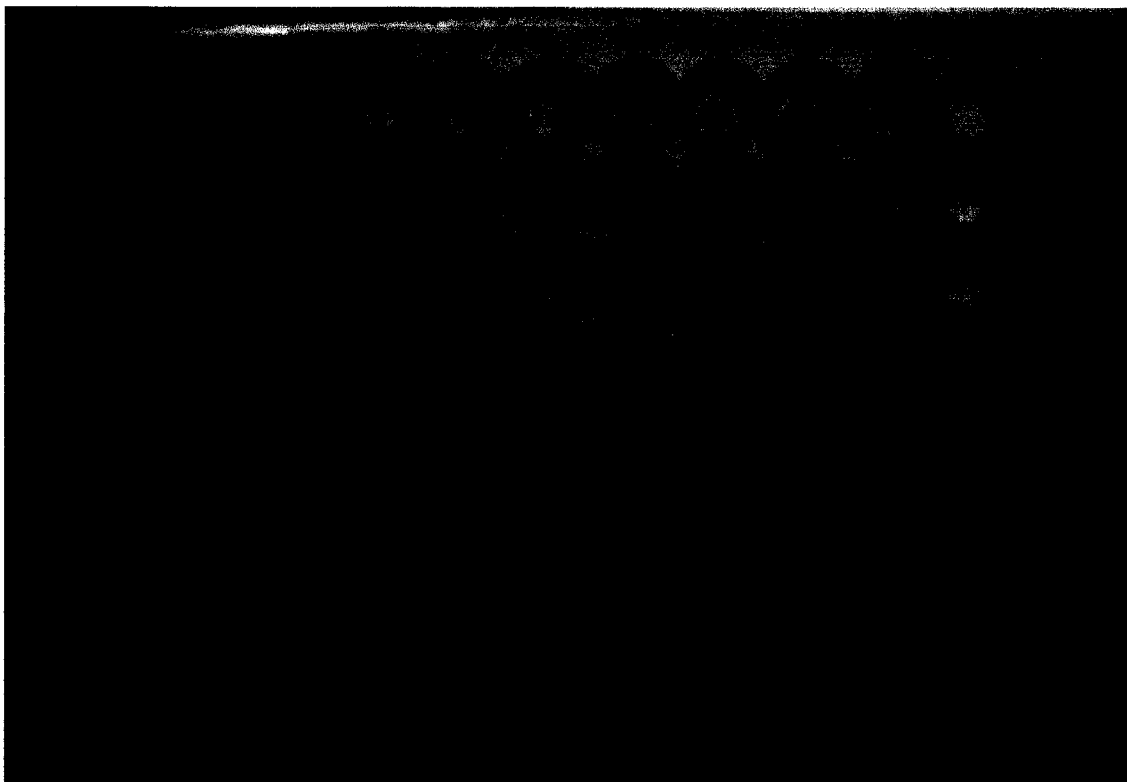
Figura 4.21. Placa de microtitulación para *Enterococcus faecalis* CECT 795.



Para esta cepa de microorganismo resultaron activos los siguientes productos:

- **23911-25-3:** Columna 6; partimos del pocillo A6 con una concentración de 1000 $\mu\text{g/mL}$; su CMI corresponde a este mismo pocillo, luego es 1000 $\mu\text{g/mL}$.
- **13676-54-5:** Columna 7; partimos del pocillo A7 con una concentración de 500 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo B7 y corresponde a 250 $\mu\text{g/mL}$.
- **16707-41-8:** Columna 8; partimos del pocillo A8 con una concentración de 250 $\mu\text{g/mL}$; su CMI se encuentra en el pocillo D8 y corresponde a 31.3 $\mu\text{g/mL}$.

Figura 4.22. Placa de microtitulación para *Pseudomonas aeruginosa* CECT 108.



Para esta cepa de microorganismo resultó activo el siguiente producto:

- **2411-89-4:** Columna 3; partimos del pocillo A3 con una concentración de 1000 $\mu\text{g/mL}$; su CMI corresponde a este mismo pocillo, luego es 1000 $\mu\text{g/mL}$.

4.5. ENSAYOS DE TOXICIDAD

Aquellos productos que mostraron actividad antibacteriana fueron sometidos a ensayos de toxicidad aguda, en ratones, por vía intraperitoneal.

La Dosis Letal 50, DL_{50} , se define como la dosis que mata al 50 % de los animales tratados por vía de administración intraperitoneal.

Hemos comenzado por probar una dosis aproximadamente igual o ligeramente superior a la CMI y en función de los resultados obtenidos para esta dosis, se continúa el estudio con dosis inferiores o superiores.

Se deben tabular de forma adecuada los valores correspondientes a cada dosis (D), indicando el log D, el porcentaje de muertes y los probits correspondientes a este porcentaje de muertes. A continuación se procede a su representación gráfica y al final se determina la dosis correspondiente al probit 5 o la DL_{50} con su correspondiente desviación estándar V, según indica el método.⁹⁶

En las tablas 4.24, 4.25, 4.26 y 4.27 se expresan los resultados obtenidos para los productos 2411-89-4, 23911-25-3, 16707-41-8 y 13676-54-5 respectivamente.

Tabla 4.24. Toxicidad aguda intraperitoneal para el producto 2411-89-4.

Dosis (mg/Kg)	Nº Muertes	% Muertes
375	0	0
750	0	0
1125	0	0
1500	4	40
DL ₅₀ > 1500 mg/Kg		

La DL₅₀ para este producto resultó ser mayor de 1500 mg/Kg, ya que con una dosis de 1500 mg/Kg el porcentaje de muertes es inferior al 50 %, y una mayor dosis dificulta la suspensión de producto y la administración del mismo.

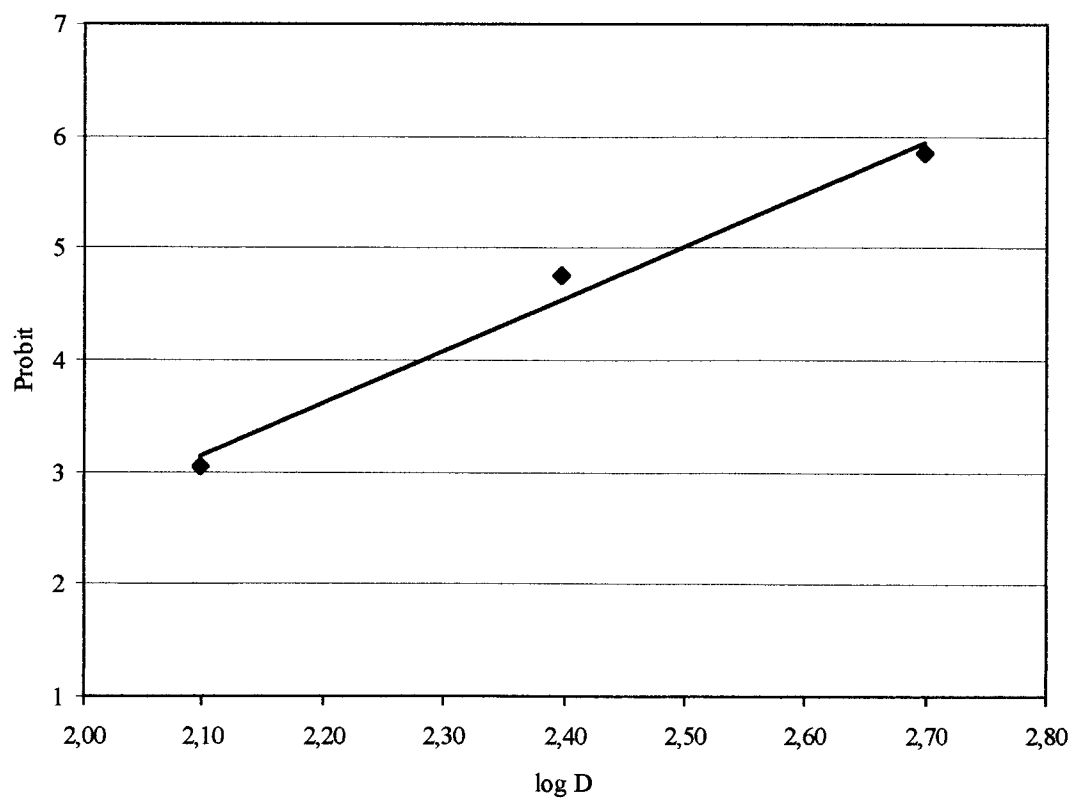
Tabla 4.25. Toxicidad aguda intraperitoneal para el producto 23911-25-3.

Dosis (mg/Kg)	Nº Muertes	% Muertes
1000	10	100
1500	10	100
DL ₅₀ < 1000 mg/Kg		

Tabla 4.26. Toxicidad aguda intraperitoneal para el producto 16707-41-8.

Dosis (mg/Kg)	log D	Nº Muertes	% Muertes	Probit
125	2.097	0	2.5	3.040
250	2.398	4	40	4.747
500	2.699	8	80	5.842

La figura 4.23 muestra la recta de regresión para la toxicidad aguda i.p. del producto 16707-41-8, así como la ecuación de la misma y la DL_{50} con su desviación estándar.

Figura 4.23. Recta de regresión para la toxicidad aguda i.p. del producto 16707-41-8.

Ecuación de la recta y DL_{50} :

$$\text{Probit} = - 6.6171 + 4.654 \log D \quad r = 0.992$$

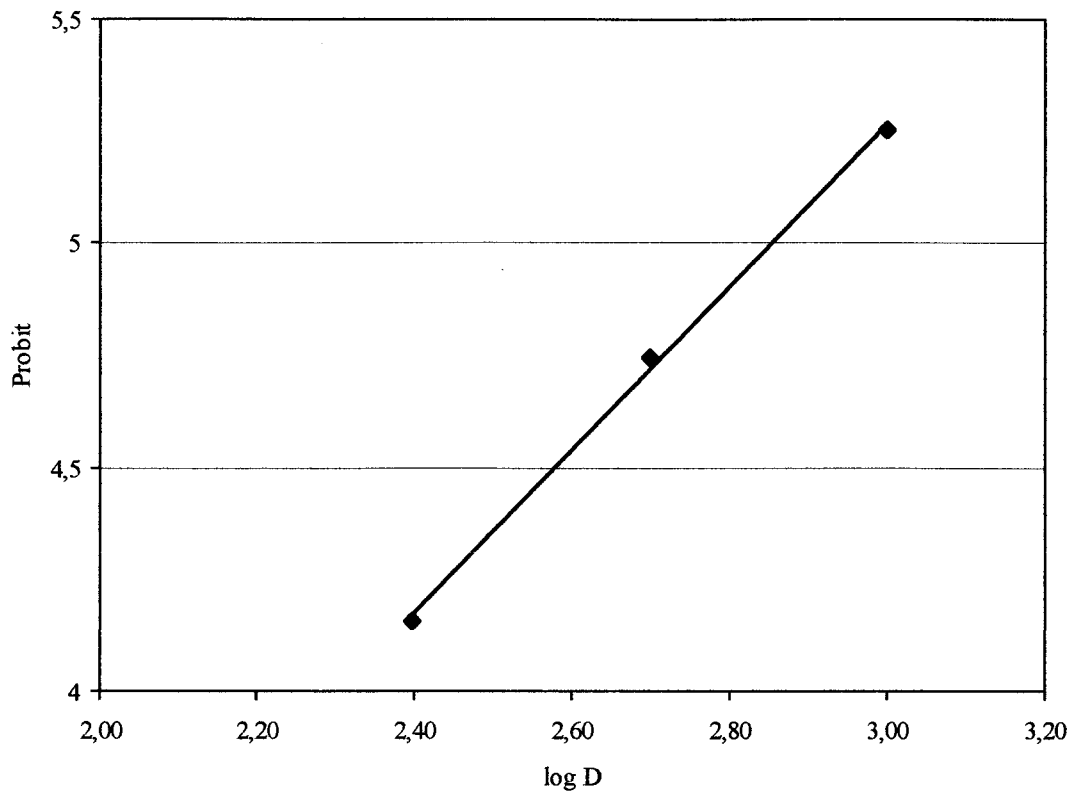
$$DL_{50} = 313 \pm 51 \text{ mg/Kg}$$

Tabla 4.27. Toxicidad aguda intraperitoneal para el producto 13676-54-5.

Dosis (mg/Kg)	log D	Nº Muertes	% Muertes	Probit
250	2.398	2	20	4.158
500	2.699	4	40	4.747
1000	3.000	6	60	5.253

La figura 4.24 muestra la recta de regresión para la toxicidad aguda i.p. del producto 13676-54-5, así como la ecuación de la misma y la DL_{50} con su desviación estándar.

Figura 4.25. Recta de regresión para la toxicidad aguda i.p. del producto 13676-54-5.



Ecuación de la recta y DL_{50} :

$$\text{Probit} = -0.1894 + 1.8188 \log D \quad r = 0.999$$

$$DL_{50} = 713 \pm 368 \text{ mg/Kg}$$

5. DISCUSIÓN

5. DISCUSIÓN

Para realizar la discusión hemos seguido con el esquema secuencial de las secciones anteriores, dividiéndolo en varios apartados para facilitar su seguimiento y concordancia con los apartados correspondientes del capítulo de resultados. Lo hemos estructurado de la siguiente manera:

1. Obtención de índices topológico-estructurales.
2. Discriminación y selección de nuevas estructuras con teórica actividad antibacteriana.
3. Ensayos farmacológicos.

5.1. OBTENCIÓN DE ÍNDICES TOPOLÓGICO-ESTRUCTURALES

El cálculo de los 62 índices topológico-estructurales para todas las moléculas empleadas en el estudio, ha puesto de manifiesto, por una parte la robustez y fiabilidad del programa desarrollado a tal efecto y por otra la comodidad y rapidez de su manejo, lo que siempre es interesante cuando, como es nuestro caso, se manejan gran número de datos y además con diferente grado de complejidad.

Un primer análisis de los datos obtenidos, no por sencillo menos interesante, nos lo puede proporcionar la visión del espectro molecular obtenido (figura 4.1) para los grupos de moléculas con y sin actividad antibacteriana, donde podemos apreciar como determinados grupos de índices poseen valores promedio marcadamente superiores en el espectro molecular para las estructuras activas, como es el caso de los grupos de índices de distancias referidas a los átomos de N, O y S (índices 17 a 36), así como de los de distancias generales (índices 51 a 62), lo que parece sugerir una contribución positiva de

dichos índices a la actividad antibacteriana, que las diferencia de las que no poseen dicha actividad.

5.2. DISCRIMINACIÓN Y SELECCIÓN DE NUEVAS ESTRUCTURAS CON TEÓRICA ACTIVIDAD ANTIBACTERIANA

La utilización simultánea de redes neuronales artificiales y de análisis lineal discriminante, obedece a varias razones. Una de ellas conseguir la validación del nuevo procedimiento de discriminación propuesto con el empleo de una red neuronal artificial, comparándolo con otro como el análisis lineal discriminante, ampliamente empleado en análisis QSAR y con resultados altamente satisfactorios. Además, si los resultados son aceptables por ambos métodos, dispondríamos de un doble criterio de discriminación de la actividad farmacológica buscada, con lo que el posterior proceso de selección de estructuras con teórica actividad antibacteriana sería mucho más restrictivo y presumiblemente con mayores garantías de éxito.

A la vista de los resultados y como iremos analizando en este capítulo, ambas premisas se han logrado plenamente.

5.2.1. DISCRIMINACIÓN MEDIANTE RED NEURONAL ARTIFICIAL

El programa desarrollado a tal efecto está basado en el paquete informático SNNS y permite un gran número de posibilidades de uso y configuraciones diferentes del mismo, de ahí la importancia de realizar una adecuada elección de la arquitectura de la red que permita obtener los resultados más satisfactorios.

A tal efecto se ha repetido el proceso de entrenamiento de la red con el mismo grupo de moléculas de aprendizaje y test, variando únicamente el número de neuronas de

la capa oculta (2, 4, 6, 8, 10, 16 y 32), empleando en todos los casos como método de parada el número de ciclos (10.000). Los resultados obtenidos aparecen en la tabla 4.2 en que puede apreciarse que el mayor porcentaje de aciertos obtenidos en el grupo de aprendizaje se consigue con 2, 10 y 16 neuronas en la capa oculta (98.69 %), mientras que el mayor rendimiento del grupo de moléculas test se obtiene para la red con 16 neuronas en la capa oculta (92.29 %). Sin embargo, dado que el porcentaje de aciertos en el grupo test para la red con dos neuronas en la capa oculta es tan sólo un 0.88 % inferior, razón por la que hemos elegido como arquitectura de la red la de mayor simplicidad (2 neuronas en la capa oculta), ya que las de mayor complejidad no mejoran significativamente los resultados obtenidos, mientras que sí se ve sensiblemente afectado el tiempo de cálculo.

Tanto para las moléculas activas como inactivas que componen los grupos de aprendizaje y test, los resultados obtenidos con esta red neuronal son altamente satisfactorios.

Teniendo en cuenta que se clasifican como activas las moléculas cuyo valor de salida de la red está comprendido entre +0.5 y +1 e inactivas las de valor de salida entre -0.5 y -1, siendo indeterminadas o de clasificación incierta el resto de moléculas, se han alcanzado porcentajes de acierto del orden del 98 % para las moléculas del grupo de aprendizaje (tanto activas como inactivas) y del 95.31 % y 87.50 % respectivamente, para las activas e inactivas del grupo test, no obteniéndose ninguna que pudiera ser indeterminada.

Una visión más detallada de estos resultados nos la proporciona el diagrama de distribución farmacológica para los valores de salida de la red neuronal entrenada (figura 4.4), obtenido a partir de los valores de espectancia para cada uno de los intervalos en que hemos dividido los posibles valores de salida de la red (-1 a +1). Es conveniente destacar

que la mayor parte de moléculas activas presentan valores de salida de la red próximos a 1, repitiéndose idéntica situación para las inactivas, con valores cercanos a -1, no existiendo prácticamente moléculas de ningún grupo con valores alrededor de 0, por lo que no existe ninguna zona de solapamiento entre el grupo de moléculas activas e inactivas.

5.2.2. DISCRIMINACIÓN MEDIANTE ANÁLISIS LINEAL DISCRIMINANTE

Con las mismas moléculas y grupos en que se distribuyeron para realizar la discriminación con redes neuronales artificiales, se procedió a realizar el análisis lineal discriminante empleando el programa 7M del paquete informático BMDP y empleando como variables independientes los 62 índices topológico-estructurales calculados para todas las moléculas.

Como criterio de clasificación de las moléculas se tomó para el valor de la función discriminante (ΔP) la misma restricción que para la red neuronal, es decir, moléculas activas las que obtengan valores superiores a +0.5 e inactivas las de valor inferior a -0.5, siendo el resto indeterminadas o de clasificación incierta. Este criterio puede en algún caso hacer disminuir el porcentaje de acierto dentro de cada grupo, pero presenta la ventaja de evitar la inclusión de moléculas falsamente activas con valores bajos de ΔP (próximos a 0).

La función discriminante obtenida permite, con tan sólo seis variables, realizar una adecuada discriminación tanto en el grupo de entrenamiento como en el grupo test, alcanzando porcentajes de acierto del 81.51 % para las moléculas activas del grupo de entrenamiento y del 89.90 % para las inactivas, mientras que se obtienen aciertos del 81.36 % para las moléculas activas del grupo test y del 93.55 % para las inactivas, quedando el porcentaje de moléculas indeterminadas alrededor del 5 %.

El diagrama de distribución farmacológica obtenido al representar los valores de espectancia para los distintos intervalos de valores de la función discriminante (figura 4.6) pone de manifiesto una clara separación entre las moléculas clasificadas como activas y las inactivas, obteniéndose dos máximos a ambos lados del valor 0 de ΔP . Existe una pequeña zona de solapamiento de los dos grupos de compuestos que correspondería al 5 % de moléculas clasificadas como indeterminadas (-0.5 a +0.5). Se observa además un agrupamiento más homogéneo en la zona de valores negativos de ΔP (moléculas inactivas), que corrobora el porcentaje de acierto (en los grupos de entrenamiento y test) ligeramente superior para el grupo de compuestos inactivos.

Aunque en el grupo de valores positivos de ΔP (moléculas activas) existe algo más de dispersión en los datos, podemos afirmar a la vista del diagrama que para valores de ΔP superiores a 0.5 tan sólo aparecen moléculas activas, con lo que el intervalo de indeterminación (0 a +0.5) permite asegurar la no existencia de compuestos falsamente activos por encima de 0.5, aunque en contrapartida, se vea ligeramente disminuido el porcentaje global de aciertos conseguido.

Por otra parte, la función discriminante nos proporciona información topológica de gran utilidad al conocer qué índices contribuyen positiva o negativamente a la actividad farmacológica buscada. Podemos así observar como los índices D^{N8} , D^{O5} , D^{S2} y D^{F5} contribuyen positivamente a la actividad antibacteriana, información coincidente con la aparecida en el espectro molecular del grupo de compuestos con actividad antibacteriana, en el que los valores promedio de los índices de distancias de N, O y S están por encima de los correspondientes para compuestos inactivos. Algo similar ocurre con los índices V^2 (índice nº 12) y D^{Cl4} (índice nº 44) que contribuyen negativamente a la función discriminante y cuya tendencia también puede apreciarse en el espectro molecular en

idéntico sentido, siendo los valores promedio de dichos índices ligeramente superiores para las moléculas sin actividad antibacteriana.

5.2.3. SELECCIÓN DE NUEVAS ESTRUCTURAS

Con la red neuronal y la función discriminante obtenidas, disponemos de dos herramientas altamente eficaces en la discriminación de la actividad antibacteriana, como así lo demuestran los resultados obtenidos por ambas.

Este hecho permite la validación conjunta de los índices topológico-estructurales propuestos, así como de su posterior procesamiento mediante redes neuronales artificiales, contando como valor de referencia con la discriminación realizada con un método clásico del análisis QSAR como es el análisis lineal discriminante.

Así, por una parte, podemos afirmar que los resultados obtenidos mediante análisis lineal discriminante empleando los índices topológico-estructurales propuestos, son similares a los obtenidos con otros índices topológicos, pero con la ventaja añadida de que estos son números enteros sencillos y por tanto hacen referencia a requisitos topológicos de interpretación más inmediata.

Además, si procesamos estos mismos índices mediante una red neuronal artificial, los resultados mejoran ostensiblemente respecto del análisis lineal discriminante, no sólo por el porcentaje de acierto obtenido en cada grupo sino por la ausencia de zonas de solapamiento que puedan conducir a una clasificación errónea o a la inclusión de compuestos falsamente activos.

De todo ello se deduce la validez y utilidad, tanto de los índices como de los métodos de discriminación empleados, lo que permitiría su uso en la selección de nuevas estructuras que posean la actividad antibacteriana buscada.

Para realizar la búsqueda o selección de nuevas estructuras partimos de la información topológica que nos proporciona la función discriminante, empleando para ello las subestructuras que contengan los índices que presentan contribución positiva a dicha función (D^{N8} , D^{O5} , D^{S2} , D^{F5}).

Las estructuras así seleccionadas a partir de bases de datos fueron sometidas al mismo proceso de cálculo de índices que las empleadas en los procesos de discriminación, calculándose para ellas el valor correspondiente de ΔP , siendo clasificadas como activas con valores altos de la función discriminante las 16 moléculas que aparecen en la tabla 4.19.

Puesto que los resultados del proceso de discriminación realizado mediante la red neuronal artificial se mostraron más eficaces que los obtenidos por análisis lineal discriminante, tanto en el porcentaje de aciertos como en la ausencia de zonas de solapamiento entre los grupos de moléculas activas e inactivas, decidimos emplear dicha red entrenada como segundo criterio de discriminación para estas moléculas, mostrándose, como era previsible, más restrictivo que el primer criterio, pues de las 16 moléculas seleccionadas en el primer paso, tan sólo fueron 11 las que superaron el segundo paso (figuras 4.8 a 4.18), siendo estas por tanto las que posteriormente serán sometidas a ensayos microbiológicos que permitan realizar la comprobación de su actividad antibacteriana.

5.3. ENSAYOS FARMACOLÓGICOS

Resulta evidente que la forma más eficaz de ratificar la validez de un método de predicción teórico, es comprobar si dicha predicción se ve respaldada con resultados satisfactorios en los análisis *in vitro*. Con ese objetivo se sometieron a estas 11 moléculas a ensayos de actividad antibacteriana y posteriormente a ensayos de toxicidad aguda en animales de experimentación.

Tras realizar la pruebas de solubilidad para cada producto y encontrar en todos los casos el disolvente o mezcla de disolventes que permitiesen la correcta disolución del mismo sin afectar el crecimiento del microorganismo en el medio de cultivo, se procedió a realizar la determinación de la Concentración Mínima Inhibitoria de cada producto frente a diferentes cepas de *Enterococcus faecalis*, *Staphylococcus aureus*, *Escherichia coli* y *Pseudomonas aeruginosa*, microorganismos habitualmente empleados como referencia en estos ensayos.

La determinación de la CMI fue llevada a cabo en medio líquido en placas de microtitulación empleando diluciones seriadas de cada producto comparándolas con las obtenidas para la cefalosporina C y ácido nalidíxico, empleados como fármacos de referencia, dió como resultado en el caso de *Staphylococcus aureus* CECT 240, dos productos activos siendo la CMI para el compuesto de número C.A.S. 13676-54-5 de 125 µg/mL y para el 16707-41-8 de 15.7 µg/mL, mientras que para la cefalosporina C y ácido nalidíxico fueron de 31.3 µg/mL y 62.5 µg/mL respectivamente. Por tanto de los dos compuestos activos, uno consigue duplicar la potencia (mitad CMI) del fármaco que mostró más eficacia (cefalosporina C).

Frente al otro microorganismo gram-positivo ensayado, *Enterococcus faecalis* CECT 795, se obtuvieron tres productos activos, dos de ellos coincidentes con los anteriormente citados, como son los compuestos de número C.A.S. 13676-54-5 y 16707-41-8, con valores de CMI de 250 y 31.3 $\mu\text{g}/\text{mL}$ respectivamente, siendo de 1000 $\mu\text{g}/\text{mL}$ para el tercer producto activo (número C.A.S. 23911-25-3). En esta ocasión dos de los compuestos activos, concretamente los que coinciden con el otro microorganismo gram-positivo (*Staphylococcus aureus*) poseen valores de CMI muy inferiores a los que presentan los fármacos de referencia (>1000 $\mu\text{g}/\text{mL}$ para la cefalosporina C y 1000 $\mu\text{g}/\text{mL}$ para ácido nalidíxico).

Los resultados obtenidos frente a los dos microorganismos gram-positivos son por tanto claramente satisfactorios, pues en ambos casos encontramos al menos un compuesto que supera la eficacia mostrada por dos fármacos de referencia de tan amplia utilización en clínica.

En el caso de los microorganismos gram-negativos los resultados no son tan uniformes como en el caso anterior, pues si bien frente a *Escherichia coli* CECT 405 hemos obtenido tres productos activos, sólo encontramos uno frente a *Pseudomonas aeruginosa* CECT 108, el del compuesto con número C.A.S. 2411-89-4, cuya CMI es de 1000 $\mu\text{g}/\text{mL}$, siendo 500 $\mu\text{g}/\text{mL}$ la correspondiente a ambos fármacos de referencia.

Frente a *Escherichia coli*, el producto de número C.A.S. 2411-89-4 presenta una CMI de 1000 $\mu\text{g}/\text{mL}$, siendo de 500 $\mu\text{g}/\text{mL}$ para los compuestos 23911-25-3 y 16707-41-8.

Es de destacar que algunos de los compuestos con actividad antes citados, lo son simultáneamente y con buenos valores de CMI frente a microorganismos gram-positivos y gram-negativos, pues aparecen como activos frente a *Staphylococcus aureus* y/o *Enterococcus faecalis* y además frente a *Escherichia coli*.

Con los productos que mostraron actividad en los ensayos antes citados decidimos realizar pruebas de toxicidad aguda en animales de experimentación, dato que nos permitirá disponer del posible margen terapéutico de estos compuestos, de los que no existe referencia bibliográfica alguna que describa su posible efecto antibacteriano.

Los resultados de este ensayo muestran características bien distintas para los productos analizados, pues nos encontramos con un compuesto (C.A.S. 23911-25-3) que aunque mostró actividad frente a los microorganismos *Enterococcus faecalis* y *Escherichia coli*, su DL_{50} es igual o menor que su CMI, con lo que no existe margen terapéutico alguno pues su actividad puede ser debida a su propia toxicidad.

En otro caso nos encontramos con un producto (C.A.S. 2411-89-4) que mostró actividad frente a *Escherichia coli* y *Pseudomonas aeruginosa*, aunque con una CMI elevada, y que posee una DL_{50} ligeramente superior a dicho valor y por tanto con escaso margen de utilización.

No es así el caso de los dos productos restantes, pues nos encontramos con el producto con número C.A.S. 16707-41-8 que presenta una DL_{50} de 313 mg/Kg que es entre diez y veinte veces superior al valor de CMI obtenido para este compuesto frente a *Enterococcus faecalis* (31.3 $\mu\text{g/mL}$) o frente a *Staphylococcus aureus* (15.7 $\mu\text{g/mL}$), lo que le confiere un amplio margen terapéutico de utilización del producto frente a microorganismos gram-positivos.

En situación similar nos encontramos con el compuesto con número C.A.S. 13676-54-5, para el que hemos obtenido un valor de DL_{50} de 713 mg/Kg, valor entre 3 y 6 veces mayor que la CMI que presenta dicho producto frente a microorganismos gram-positivos como *Enterococcus faecalis* (250 $\mu\text{g/mL}$) y *Staphylococcus aureus* (125 $\mu\text{g/mL}$), por lo

que, al igual que en el caso anterior, disponemos de un notable margen terapéutico que permitiría ratificar la eficacia antibacteriana de estos compuestos.

Estos resultados poseen en sí mismos un notable interés, ya que se han obtenido antibacterianos que pueden considerarse como buenas "cabezas de serie", es decir, no derivados de ninguna estructura con actividad antibacteriana conocida, lo que respalda la eficacia del método, más aún si tenemos en cuenta la sencillez conceptual y de cálculo que presentan los métodos topológicos, a los que, incorporando nuevas metodologías basadas en conceptos de inteligencia artificial, como son las redes neuronales, se ha conseguido mejorar notablemente la capacidad predictiva y/o de selección de estructuras con la actividad farmacológica buscada.

6. CONCLUSIONES

6. CONCLUSIONES

1. Se ha aplicado un paquete informático específicamente desarrollado para el cálculo de 62 índices topológico-estructurales de una serie de moléculas con y sin actividad antibacteriana, mostrándose altamente eficaz, tanto por la sencillez de manejo y rapidez de cálculo como por el ahorro de memoria conseguido en el almacenamiento de datos.
2. El espectro molecular obtenido para los grupos de moléculas con y sin actividad antibacteriana, pone de manifiesto la existencia de grupos de índices que presentan valores superiores para las estructuras con actividad antibacteriana, siendo así el caso de los índices de distancias referidos a los átomos de N, O y S.
3. Se ha empleado una red neuronal artificial de tipo perceptrón multicapa para la discriminación de la actividad antibacteriana, utilizando como algoritmo de aprendizaje una modificación del algoritmo de retropropagación del error y la tangente hiperbólica como función de activación. La arquitectura de la red que obtiene mejores resultados en los grupos de aprendizaje y test es la que emplea una capa de entrada, una de salida y una capa oculta con sólo dos neuronas.
4. La discriminación de actividad antibacteriana conseguida empleando los 62 índices topológico-estructurales, como vectores de entrada de la red neuronal artificial, alcanzó porcentajes de acierto del 98.69 % y 98.68 % para las poblaciones activa e inactiva del grupo de aprendizaje y del 95.31 % y 87.50 % para las mismas poblaciones del grupo test. Mientras que el análisis lineal discriminante realizado empleando las mismas variables y grupos de moléculas, alcanzó porcentajes de

- acierto del 81.51 % y 89.90 % para las poblaciones activa e inactiva del grupo de entrenamiento y del 81.36 % y 93.55 % para las mismas poblaciones del grupo test.
5. Los diagramas de distribución farmacológica obtenidos con ambos métodos de discriminación, ponen de manifiesto que tanto con los valores de salida de la red, como con la función discriminante se consigue una completa separación para las poblaciones de moléculas activas e inactivas, sin embargo, la red neuronal artificial se mostró más eficaz al no existir en su diagrama zonas de solapamiento entre ambos grupos de moléculas.
 6. Se ha realizado una búsqueda guiada de estructuras empleando la información topológica de los índices que aparecen en la función discriminante, encontrando un total de diez compuestos con teórica actividad antibacteriana, al conseguir resultados positivos tanto en dicha función como en la red neuronal entrenada, empleando ambos métodos como filtros sucesivos en el proceso de selección molecular.
 7. Los diez compuestos seleccionados fueron sometidos a ensayos de actividad antimicrobiana frente a cepas de microorganismos gram-positivos, como *Enterococcus faecalis* (CECT 795) y *Staphylococcus aureus* (CECT 240) y gram-negativos, como *Escherichia coli* (CECT 405) y *Pseudomonas aeruginosa* (CECT 108), mostrando actividad cuatro de ellos:
 - N-[4-(2-Benzoxazolil)fenil]maleimida (nº. C.A.S.16707-41-8).
 - 1,1'-(Metilene-di-4,1-fenilen)bismaleimida (nº. C.A.S. 13676-54-5).
 - o-Cresolftaleina complexona (nº. C.A.S. 2411-89-4).
 - Dianhidrido etilendiaminotetraacético (nº. C.A.S. 23911-25-3).

8. El compuesto N-[4-(2-Benzoxazolil)fenil]maleimida mostró actividad simultáneamente frente a *Escherichia coli*, *Staphylococcus aureus* y *Enterococcus faecalis*, obteniendo valores de CMI que mejoran claramente los que presentan Cefalosporina C y Ácido nalidíxico, empleados como fármacos de referencia; el compuesto 1,1'-(Metilene-di-4,1-fenilen)bismaleimida lo hizo frente a *Staphylococcus aureus* y *Enterococcus faecalis*; el producto o-Cresolftaleina complexona se mostró activo frente a *Escherichia coli* y *Pseudomonas aeruginosa*; mientras que el compuesto Dianhidrido etilendiaminotetraacético presentó actividad frente a *Escherichia coli* y *Enterococcus faecalis*.
9. La determinación de la DL₅₀ intraperitoneal en ratón para los cuatro productos que mostraron actividad antibacteriana, ha puesto de manifiesto la escasa toxicidad de los mismos, obteniéndose márgenes terapéuticos de considerable amplitud para dos de ellos, pues el compuesto N-[4-(2-Benzoxazolil)fenil]maleimida presenta valores de DL₅₀ entre diez y veinte veces superior a los de CMI obtenidos frente a *Enterococcus faecalis* y *Staphylococcus aureus*; mientras que el compuesto 1,1'-(Metilene-di-4,1-fenilen)bismaleimida posee valores de DL₅₀ entre tres y seis veces superiores a sus valores de CMI frente a *Enterococcus faecalis* y *Staphylococcus aureus*.
10. Los resultados obtenidos confirman plenamente la validez del método topológico empleado, en la búsqueda y selección de nuevas estructuras con actividad antibacteriana, tanto por la idoneidad de los índices topológico-estructurales como por la utilización combinada de análisis lineal discriminante y redes neuronales artificiales.

7. BIBLIOGRAFÍA

7. BIBLIOGRAFÍA

1. Gracia Guillén, D.; Albarracín, A.; Arquiola, E.; Erill, S.; Montiel, L.; Peset, J. L.; Laín Entralgo, P.; "Historia del Medicamento", Eds. Doyma, Barcelona, **1987**.
2. Manaut, F.; Claramunt, R.; "Parámetros o factores descriptores de las propiedades fisicoquímicas de los compuestos orgánicos y de las relaciones cuantitativas estructura-actividad (QSAR) en el diseño de fármacos", Avendaño, M. C., Introducción a la química farmacéutica, Ed. McGraw-Hill/Interamericana, Madrid, **1997**.
3. Graig, P. N.; " Q.S.A.R. Origins and present: A historical perspective", *Drug Inf. J.*, **1984**, 18, 123-130.
4. Gund, P.; Andose, J. D.; Rhodes, J. B.; Smith, G. M.; "Three-Dimensional molecular modeling and drug design", *Science*, **1980**, 208, 1425-1431.
5. Hansch, C.; Muir, R. M.; Fujita, T.; Maloney, P. P.; Geiger, F; Streich, M.; "The correlation of biological activity of plant growth regulators and chloromycetin derivatives with Hammett constants and partition coefficients", *J. Am. Chem. Soc.*, **1963**, 85, 2817-2824.
6. Free, S. M.; Wilson, J. W.; "A Mathematical Contribution to Structure Activity Studies", *J. Med. Chem.*, **1964**, 7, 395-399.

7. Sanz, F.; "Modelos moleculares tridimensionales y su uso en el estudio de las relaciones estructura-actividad", Avendaño, C., Introducción a la química farmacéutica, Ed. McGraw-Hill/Interamericana, 2ª edición, Madrid, 2001.
8. Gilda H. Loew and Stanley K. Burt; "Quantum Mechanics and the Modeling of Drug Properties", Hansch, C.; Sammes, P.G.; Taylor J.B.; Comprehensive medicinal chemistry, Vol. 4, Quantitative drug design, Pergamon Press, 1990.
9. Menéndez J. C.; Avendaño, C.; "Optimización de un prototipo. Correlaciones cualitativas estructura química-actividad biológica", Avendaño, C., Introducción a la química farmacéutica, Ed. McGraw-Hill/Interamericana, 2ª edición, Madrid, 2001.
10. Gálvez, J.; García-Domenech, R.; Bernal, J. M.; García-March, F. J.; "Desarrollo de un nuevo método de diseño de fármacos por topología molecular. Su aplicación a analgésicos no narcóticos", *An. Real Acad. Farm.*, 1991, 57, 533-546.
11. Kier, L. B.; Hall, L. H.; "Molecular Connectivity in Chemistry and Drug Research", Ed. Academic Press, London, 1976.
12. Randić, M.; "Generalized Molecular Descriptors", *J. Math. Chem.*, 1995, 7, 155-168.
13. Balaban, A. T.; Motoc, I.; Bonchev, D.; Mekenyan, O.; "Topological Indices for Structure-Activity Correlations", *Top. Curr. Chem.*, 1984, 114, 21-71.

14. Lien, E. J.; "SAR: Side Effects and Drug Design", Marcel Dekker Inc., New York, **1987**.
15. Rouvray, D. H.; "De la topología a la química", *Invest. Ciencia*, **1986**, 122, 16-24.
16. Gutman, I.; Ruscic, N.; Trinajstić, N.; Wilcox, C. F.; "Graph Theory and Molecular Orbitals. XII. Acyclic Polyenes", *J. Chem. Phys.*, **1975**, 62, 3399-3405.
17. Stankevich, M. I.; Stankevich, I. V.; Zefirov, N. S.; "Topological Indexes in Organic Chemistry", *Usp. Khimii.*, **1988**, 57, 337-366.
18. Rouvray, D. H.; "Should We Have Designs on Topological Indexes?", *Stud. Phys. Theor. Chem.* **1983**, 28 (*Chem. Appl. Topol. Graph Theory*), 159-177.
19. Balaban, A.; "Chemical Graphs. XXXIV. Five New Topological Indices for the Branching of Tree-like Graphs", *Theor. Chim. Acta*, **1979**, 53, 355-375.
20. Balaban, A. T.; Motoc, I.; Bonchev D.; Mekenyan, O.; "Topological Indices for Structure-Activity Correlations", *Top. Curr. Chem.*, **1984**, 114, 21-71.
21. Randić, M.; "On Characterization of Molecular Branching", *J. Am. Chem. Soc.*, **1975**, 97, 6609-6615.
22. Randić, M.; *Int. J. Quantum Chem., Quantum Biol. Symp.*, **1978**, 5, 245.

23. Kier, L. B.; Hall, L. H.; "Molecular Connectivity in Structure-Activity Analysis", Research Studies Press, Letchworth, UK, 1986.
24. Platt, J. R.; "Influence of Neighbor Bonds on Additive Bond Properties in Paraffins", *J. Chem. Phys.*, 1947, 15, 419-420.
25. Platt, J. R.; "Prediction of isomeric differences in paraffin properties", *J. Phys. Chem.*, 1952, 56, 328-336.
26. Gordon, M. and Scantlebury, G. R.; *Trans. Faraday Soc.*, 1964, 60, 604.
27. Gutman, I. and Randic, M.; *Chem. Phys. Lett.*, 1977, 47, 15.
28. Randic, M.; *J. Chem. Phys.*, 1974, 60, 3920.
29. Randic, M.; *Chem. Phys. Lett.*, 1976, 42, 283.
30. Wiener, H.; "Structural determination of paraffin boiling points", *J. Am. Chem. Soc.*; 1947, 69, 17-20.
31. Wiener, H.; "Correlation of heats of isomerization and differences in heats of vaporization of isomers, among the paraffin hydrocarbons", *J. Am. Chem. Soc.*; 1947, 69, 2636-2638.

-
32. Wiener, H.; "Vapor Pressure-Temperature Relationship Among the Branched Paraffin Hydrocarbons", *J. Phys. Chem.*; **1948**, 52, 425-430.
33. Wiener, H.; "Relation on the Physical Properties of Isomeric Alkanes to Molecular Structure: Surface Tension, Specific Dispersion and Critical Solution Temperature in Aniline", *J. Phys. Chem.*; **1948**, 52, 1082-1089.
34. Altenburg, K.; "Zur Berechnung des Radius verzweigter Molekule", *Kolloid Zeits.*, **1961**, 178, 112-119.
35. Altenburg, K.; *Brennst-Chem.*; **1966**, 47, 100, 331.(Chem. Abstr., 1966, 65, 3714f).
36. Hosoya, H.; "Topological index. A Newly proposed quantity Characterizing the topological nature of structural isomers of saturated hydrocarbons", *Bull. Chem. Soc. Jpn.*, **1971**, 44, 2332.
37. Bonchev, D.; Balaban, A. T.; Mekenyan, O.; " Generalization of the Graph Center Concept and Derived Topological Indexes", *J. Chem. Inf. Comput. Sci.*, **1980**, 20, 106-113.
38. Balaban, A. T. "Highly discriminating distance-based topological index", *Chem. Phys. Lett.*, **1982**, 89, 399-404.
39. Balaban, A. T.; *MATCH*; **1984**, 16, 163.

40. Kier, L. B.; "A Shape Index from Molecular Graphs.", *Quant. Struct. - Act. Relat.*, **1985**, 4, 109-116.
41. Kier, L. B.; "Shape Indexes of Orders One and Three from Molecular Graphs", *Quant. Struct. - Act. Relat.*, **1986**, 5, 1-7.
42. Gálvez, J.; García, R.; Salabert, M. T.; Soler, R.; "Charge Indexes. New Topological Descriptors", *J. Chem. Inf. Comput. Sci.*, **1994**, 34, 520-525.
43. Balaban, A. T.; *Theor.Chim.Acta.*, **1979**, 53, 355.
44. Shanon, C. E. and Weaver, W.; "The Mathematical Theory of Communication", University of Illinois Press, Urbana, IL, **1949**.
45. Bonchev, D.; "Theoretic Information Indices for Characterization of Chemical Structures", Research Studies Press, Chichester, **1983**.
46. Bruillon, L.; "Science and Information Theory", 2nd edn., Academic Press, New York, **1962**.
47. Bonchev, D. and Trinajstic, N.; "Information Theory, Distance Matrix and Molecular Branching", *J. Chem. Phys.*, **1977**, 67, 4517-4533.

-
48. McCulloch, W. S.; Pitts, W. A.; "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Boullletin of Mathematics and Biophysics*, **1943**, 15, 115-133.
49. Rosenblatt, F.; "The Perceptron: A probabilistic model for information storage and organization in the brain", *Psychological Review*, **1958**, 65, 386-408.
50. Widrow, B.; Hoff, M.; "Adaptative Switching Circuits", *IREWESCON Convention Record*, **1960**, 4, 96-104.
51. Hecht-Nielsen, R.; "Neurocomputing: Picking the Human Brain", *IEEE Spectrum*, **1988**, 25, 36-41.
52. Kohonen, T.; "An introduction to neural computing", *Neural Networks*, **1988**, 1, 3-16.
53. Hilera, J. R.; Martínez, V. J.; "Redes neuronales artificiales: fundamentos, modelos y aplicaciones", Ed. Ra-Ma, Madrid, **1993**.
54. Freeman, J. A.; Skapura, D. M.; "Redes neuronales. Algoritmos, aplicaciones y técnicas de programación", Ed. Addison-Wesley, Wilmington, **1993**.
55. Rumelhart, D. E.; Hinton, G. E. and Williams, R. J.; "Learning internal representations by error propagation". In D.E. Rumelhart and J.L. McClelland, ed., *PDP: Computational models of cognition and perception*, I. MIT Press, **1986**.
56. Bishop, C. M.; "Neural networks for pattern recognition", Oxford Univ. Press, **1995**.

-
57. Kohonen, T.; "Associative Memory. A system theoretical approach", *Springer-Verlag*, 1977.
58. Rumelhart, D.; Hinton, G.; Williams, R.; "Learning representations by back-propagating errors", *Nature*, 1986, 323, 533-536.
59. Álvarez, M. V.; Boquet, E.; de Fez, I.; "Manual de técnicas en microbiología clínica", Publicaciones AEFA, Madrid, 1990.
60. Simon, C.; Stille, W.; Perea, E. J.; "Manual de terapéutica antibacteriana", Ed. Salvat, Barcelona, 1987.
61. *The Merck Index*, Twelfth edition on CD-Rom, Budavari, S., Merck & Co., Inc., 1996.
62. Page, C. P.; Curtis, M. J.; Sutter, M. C.; Walker, M. J.; Hoffman, B. B.; "Infecciones bacterianas", *Farmacología integrada*, Ed. Harcourt Brace, Madrid, 1998.
63. Flórez, J.; Mediavilla, A.; García-Lobo, J. M.; "Farmacología de las enfermedades infecciosas: principios generales, selección y asociaciones de antibióticos", *Farmacología humana*, 3ª edición, Ed. Eunsa, Barcelona, 1992.
64. Gold, H. S.; Moellering, R. C.; "Antimicrobial drug resistance", *N. Engl. J. Med.*, 1996, 335, 1445-1454.

-
65. Franklin, T. J. and Snow, G. A.; "Biochemistry of Antimicrobial Action", 2^a ed. Chapman & Hall, London, 1975.
66. Foster, T. J.; "Plasmid-Determined resistance to Antimicrobial drugs and Toxic Metal Ions in Bacteria", *Microbiol. Rev.*, **1983**, 47, 361-409.
67. Levy, S. B.; "Microbial resistance to antibiotics: An evolving and persistent problem", *Lancet*, **1982**, 10, 83-88.
68. Wise, R.; "The clinical relevance of protein binding and tissue concentrations in antimicrobial therapy", *Clin. Pharmacokinet.*, **1986**, 11, 470-482.
69. Shanson, D. C.; "Prophylaxis and treatment of infective endocarditis: current recommendations", *Drugs*, **1983**, 25, 433-439.
70. Jaen-Oltra, José; Salabert-Salvador, M^a Teresa; García-March, Francisco J.; Pérez-Giménez, Facundo and Tomás-Vert, Francisco; "Artificial neural network applied to prediction of fluorquinolone antibacterial activity by topological methods", *J. Med. Chem.*, **2000**, 43, 1143-1148.
71. Tomás-Vert, F.; Pérez-Giménez, F.; Salabert-Salvador, M. T.; García-March, F. J.; Jaén-Oltra, J.; "Artificial neural network applied to the discrimination of antibacterial activity by topological methods", *Journal of Molecular Structure (Teochem)*, **2000**, 504, 249-259.

72. Valls Sánchez, J. M.; "Diseño de fármacos mediante topología molecular: desarrollo de la interfaz gráfica", Proyecto fin de Carrera, Licenciatura de Ingeniería Informática, Universidad de Valencia, **2001**.
73. Fowler, M.; Scott, K.; "UML Gota a gota", Addison Wesley Longman, **1999**.
74. Eriksson, H.; Penker, M.; "UML Toolkit", John Wiley & Sons, **1997**.
75. Booch, G.; Rumbaugh, J.; Jacobson, I.; "The Unified Modeling Language User Guide", Addison-Wesley, Massachusetts, **1999**.
76. Weininger, D.; "SMILES, a Chemical Language and Information System.1. Introduction to Methodology and Encoding Rules", *J. Chem. Inf. Comput. Sci.*, **1988**, 28, 31-36.
77. Black, Paul E.; "Dictionary of Algorithms, Data Structures, and Problems", <http://www.nist.gov/dads/>, National Institute of Standards and Technology, Gaithersburg, Maryland, **2000**.
78. Hansch, C.; "A quantitative approach to biochemical structure-activity relationships", *Acc. Chem. Res.*, **1969**, 2, 232-239.
79. Wikel, J. H.; Dow, E. R.; Heathman, M.; "Interpretative Neural Networks for QSAR", <http://netsci.org/Science/Compchem/feature02.html>, *Network Science*, **2002**.

-
80. Sato, K.; Nakagawa, J.; "An Improvement of Neural Networks Applied to Pharmaceutical Problems", *Chem. Pharm. Bull.*, **1997**, 45, 107-115.
81. Gasteiger, J.; Zupan, J.; "Neural Networks in Chemistry", *Angew. Chem. Int. Ed. Engl.*, **1993**, 32, 503-527.
82. Hoskins, J. C.; Himmelblau, D. M.; "Artificial Neural Network Models of Knowledge Representation in Chemical Engineering", *Comput. Chem. Eng.*, **1988**, 12, 881-890.
83. Qian, N.; Sejnowski, T. J.; "Predicting the Secondary Structure of Globular Proteins Using Neural Network Models", *J. Mol. Biol.*, **1988**, 202, 865-884.
84. Andrea, T. A.; Kalayeh, H.; "Applications of Neural Networks in Quantitative Structure-Activity Relationships of Dihydrofolate Reductase Inhibitors", *J. Med. Chem.*, **1991**, 34, 2824-2836.
85. Aoyama, T.; Suzuki, Y.; Ichicawa, H.; "Neural Networks Applied to Quantitative Structure-Activity Relationships Analysis", *J. Med. Chem.*, **1990**, 33, 2583-2590.
86. Devillers, J.; "Neural Networks in QSAR and Drug Design", Ed. Academic Press, London, **1996**.
87. Zell, A. et al.; SNNS Stuttgart Neural Network Simulator V 4.2, Institute for Parallel and Distributed High Performance Systems (IPVR), University of Stuttgart, **1995**.

88. Dixon, W. J.; "Manual BMD: Biomedical Computer Programs", University of California Press, Berkeley and Los Angeles, **1982**.
89. Dixon, W. J.; Brown, M. B.; Engelman, L.; Hill, M. A.; Jennrich, R. I.; " BMDP statistical software manual", Volume 2, University of California Press, Berkeley, California, USA, **1988**.
90. Gálvez, J.; García-Domenech, R.; de Gregorio Alapont, C.; de Julián-Ortiz, J. V. and Popa, L.; "Pharmacological distribution diagrams: A tool for *de novo* drug design", *J. Mol. Graphics* **1996**, 14, 272-276.
91. Cercós-del-Pozo, R. A.; Pérez-Giménez, F.; Salabert-Salvador, M. T.; García-March, F. J.; Murcia-Soler, M.; "New Hypolipaemic Agents Designed by Molecular Topology: Pharmacological Studies of 2,6-Di-tert-butyl-methylpyridine and 2,6-Di-tert-butylpyridine", *Quant. Struct.-Act. Relat.*, **1999**, 18, 464-473.
92. Murcia-Soler, M.; Pérez-Giménez, F.; Nalda-Molina, R.; Salabert-Salvador, M. T.; García-March, F. J.; Cercós-del-Pozo, R. A.; Garrigues, T. M.; "QSAR Analysis of Hypoglycemic Agents Using the Topological Indices", *J. Chem. Inf. Comput. Sci.*, **2001**, 41, 1345-1354.

93. De Julián-Ortiz, J. V.; Gálvez, J.; Muñoz-Collado, C.; García-Domenech, R.; Gimeno-Cardona, C.; "Virtual Combinatorial Syntheses and Computational Screening of New Potential Anti-Herpes Compounds", *J. Med. Chem.*, **1999**, 42, 3308-3314.
94. Casabán-Ros, E.; Antón-Fos, G. M.; Gálvez, J.; Duart, M. J.; García-Domenech, R.; "Search for New Antihistaminic Compounds by Molecular Connectivity", *Quant. Struct.-Act. Relat.*, **1999**, 18, 35-42.
95. Turner, R. A.; "Quantal response calculation of the ED₅₀", *Screening Methods in Pharmacology*, Second printing, Academic Press, New York, **1972**.
96. Miller, L. C.; Tainter, M. L.; "Estimation of the LD₅₀ and its error by means of logarithmic-probit graph paper", *Proc. Soc. Exp. Biol. Med.*, **1944**, 57, 261-264.
97. Copas, J. B.; "Using regression models for prediction: shrinkage and regression to the mean", *Statistical Methods in Medical Research*, **1997**, 6, 167-183.
98. Pascua, M.; "Metodología bioestadística", Segunda edición, Ed. Paz Montalvo, Madrid, **1974**.
99. Adobes Martín, C.; "Diseño e implementación de herramientas para la predicción de propiedades moleculares", Proyecto fin de carrera, Licenciatura en Ingeniería Informática, Universidad de Valencia, **2000**.
100. CS ChemFinder 5.1, CambridgeSoft Corporation. **1999**.

8. ANEXO

8. ANEXO**ANTIBACTERIANOS****Antibacterial (Antibiotics)****Aminoglycosides**

Amikacin
Apramycin
Arbekacin
Butirosin
Dibekacin
Dihydrostreptomycin
Fortimicin A
Fortimicin B
Gentamicin
Isepamicin
Micronomicin
Neomycin
Paromomycin
Ribostamycin
Sisomicin
Spectinomycin
Streptomycin

Amphenicols

Azidamfenicol
Chloramphenicol
Florfenicol
Thiamphenicol

Ansamycins

Rifamide
Rifampin
Rifamycin SV
Rifapentine

 β -lactams

Carbapenems
Biapenem
Imipenem
Meropenem

 β -lactams (Cont.)

Cephalosporins
Cefaclor
Cefamandole
Cefatrizine
Cefazedone
Cefcapene Pivoxil
Cefdinir
Cefditoren
Cefetamet
Cefixime
Cefmenoxime
Cefodizime
Cefonicid
Cefoperazone
Ceforanide
Cefotiam
Cefpiramide
Cefpodoxime Proxetil
Cefprozil
Cefroxadine
Cefteram
Ceftezole
Ceftibuten
Ceftizoxime
Ceftriaxone
Cefuroxime
Cefuzonam
Cephacetrile Sodium
Cephalexin
Cephalosporin C
Cephalothin
Cephapirin Sodium
Cephradine
Pivcefalexin

β -lactams (Cont.)*Monobactams*

Carumonam

Tigemonam

Penicillins

Amdinocillin

Amdinocillin Pivoxil

Amoxicillin

Ampicillin

Apalcillin

Aspoxicillin

Azidocillin

Azlocillin

Bacampicillin

Benzylpenicillinic Acid

Carbenicillin

Carindacillin

Clometocillin

Cloxacillin

Cyclacillin

Dicloxacillin

Fenbenicillin

Floxacillin

Hetacillin

Lenampicillin

Methicillin Sodium

Mezlocillin

Nafcillin Sodium

Oxacillin

Penamecillin

Penicillin G

Penicillin N

Penicillin O

Penicillin V

Phenethicillin Potassium

Pivampicillin

Propicillin

Quinacillin

Sulbenicillin

Talampicillin

Ticarcillin

Lincosamides

Clindamycin

Lincomycin

Macrolides

Azithromycin

Carbomycin

Clarithromycin

Erythromycin

Leucomycins

Midecamycins

Miokamycin

Rokitamycin

Rosaramicin

Roxithromycin

Spiramycin

Tetracyclines

Apicycline

Chlortetracycline

Clomocycline

Demeclocycline

Doxycycline

Lymecycline

Meclocycline

Methacycline

Minocycline

Oxytetracycline

Pipacycline

Sancycline

Tetracycline

Antibacterial (Synthetic)

2,4-Diaminopyrimidines

Brodimoprim
Tetroxoprim
Trimethoprim

Nitrofurans

Furaltadone
Nifuradene
Nifuratel
Nifurfoline
Nifurpirinol
Nifurprazine
Nifurtoinol
Nitrofurantoin

Quinolones

Ciprofloxacin
Clinafloxacin
Difloxacin
Enoxacin
Fleroxacin
Flumequine
Grepafloxacin
Lomefloxacin
Miloxacin
Nadifloxacin
Nalidixic Acid
Norfloxacin
Ofloxacin
Oxolinic Acid
Pazufloxacin
Pefloxacin
Pipemidic Acid
Piromidic Acid
Rosoxacin
Sparfloxacin
Temafoxacin
Tosufloxacin
Trovafoxacin

Sulfonamides

4'-(Methylsulfamoyl)sulfanilamide
4-Sulfanilamidosalicylic Acid
Benzylsulfamide

Sulfonamides (Cont.)

Dichloramine T
Mafenide
N2-Formylsulfisomidine
N4-Sulfanilylsulfanilamide
N-Sulfanilyl-3,4-xylamide
Phthalylsulfacetamide
Phthalylsulfathiazole
Salazosulfadimidine
Succinylsulfathiazole
Sulfabenzamide
Sulfachlorpyridazine
Sulfachrysoidine
Sulfacytine
Sulfadiazine
Sulfadicramide
Sulfadimethoxine
Sulfadoxine
Sulfaethidole
Sulfaguanidine
Sulfaguanole
Sulfalene
Sulfaloxic Acid
Sulfamerazine
Sulfamethazine
Sulfamethizole
Sulfamethomidine
Sulfamethoxazole
Sulfamethoxy-pyridazine
Sulfamido-chrysoidine
Sulfamoxole
Sulfanilamide
Sulfanitran
Sulfaperine
Sulfaphenazole
Sulfaproxyline
Sulfapyrazine
Sulfapyridine
Sulfasomizole
Sulfasymazine
Sulfathiazole
Sulfathiourea
Sulfisomidine
Sulfisoxazole

Sulfones

Acediasulfone

Dapsone

Diathymosulfone

p-Sulfanilylbenzylamine

Solasulfone

Succisulfone

SulfanilicAcid

Thiazolsulfone

Others

Hexedine

Methenamine

Taurolidine

Xibornol

NO-ANTIBACTERIANOS**Analgesic Narcotic**

Alfentanil
Benzylmorphine
Desomorphine
Dezocine
Dihydrocodeinone Enol Acetate
Dipipanone
Eptazocine
Ethylmethylthiambutene
Etonitazene
Isomethadone
Lofentanil
Meptazinol
Metazocine
Metopon
Nalbuphine
Norpipanone
Phenoperidine
Propoxyphene
Sufentanil

Analgesic Non-Narcotic

Acetaminophen
Aminochlorthenoxazin
Aminopyrine
Amtolmetin Guacil
Antipyrine
Benorylate
Bromfenac
Bucetin
Butacetin
Chlorthenoxazin(e)
Clometacin
Crotethamide
Enfenamic Acid
Etodolac
Flufenamic Acid
Fluoresone
Flurbiprofen
Glafenine
Ibufenac
Ketoprofen
Loxoprofen

Analgesic Non-Narcotic (Cont.)

Methotrimeprazine
Parsalmide
p-Bromoacetanilide
Phenazopyridine Hydrochloride
Phenopyrazone
Propacetamol
Salacetamide
Salicylamide-Acetic Acid
Salicylsulfuric Acid
Suprofen
Talnidflumate
Tolfenamic Acid
Tramadol

Antihistaminic

Ahistan
Bamipine
Carbinoxamine
Chlorpheniramine
Clemastine
Clobenzepam
Clocinizine
Cyproheptadine
Diphenylpyraline
Doxylamine
Ebastine
Etymemazine
Hydroxyzine
Isothipendyl
Mebhydroline
Methyldiphenhydramine-p
Moxastine
Phenyltoloxamine
Pyrrobutamine
Thenylamine
Thonzylamine
Tolpropamine
Tritoqualine

Antihyperlipoproteinemic

Atorvastatin
Azacosterol
Binifibrate
Clinofibrate
Clomestrona
Furazabol
Gemfibrozil
Nicoclonate
Phenylbutyramide
Ronifibrate
Simfibrate

Antihypertensive

Amlodipine
Aranidipine
Benazepril
Bevantolol
Bunazosin
Bunitrolol
Captopril
Carteolol
Carvedilol
Chlortalidone
Cicletanine
Delapril
Enalapril
Enalaprilat
Felodipine
Guanadrel
Guanazodine
Guanoxan
Indoramin
Levcromakalim
Lisinopril
Lofexidine
Mepindolol
Methyl-4-Pyridyl Ketone Thiosemicarbazone
Moveltipril
Muzolimine
Nicardipine
Nilvadipine
Nitrendipine
Phentolamine
Pildralazine
Prazosin

Antihypertensive (Cont.)

Rescimetol
Rescinnamine
Reserpine
Sulfinalol
Syrosingopine
Tilisolol
Timolol
Trimazosin

AINE

Ampiroxicam
Droxicam
Fenbufen
Feprazone
Isoxepac
Isoxicam
Mefenamic Acid
Mofebutazone
Oxametacine
Phenylbutazone
Pirazolac
Protizinic Acid
Tolmetin
Ximoprofen

Diuretic

Althiazide
Arbutin
Azosemide
Buthiazide
Clofenamide
Cyclothiazide
Disulfamide
Etozolin
Hydrochlorothiazide
Methyclothiazide
Polythiazide
Quinethazone
Triamterene
Tripamide

Sedative/Hypnotic

Acecarbromal
Amobarbital
Aprobarbital
Brallobarbital
Butethal
Chloralose-alfa
Chlorhexadol
Etaqualone
Ethchlorvynol
Flunitrazepam
Haloxazolam
Heptabarbital
Hexethal
Lormetazepam
Meclozamine
Methyl-4-thiazoleethanol-5
Methyprylon
Nitrazepam
Phenobarbital
Propallylonal
Secobarbital
Triazolam
Vinbarbital
Vinylbital

Antidiabetic

Acetohexamide
Chlorpropamide
Glibornuride
Glimepiride
Gliquidone
Glybuzole
Glymidine
Glypinamide
Tolazamide
Tolbutamide

Antidepressant

Adinazolam
Benactyzine
Bupropion
Butriptyline
Desipramine
Dimetacrine
Dioxadrol
Dothiepin
Febarbamate
Femoxetine
Indeloxazine Hydrochloride
Levophacetoperane
Lofepamine
L-Tryptophan
Metralindole
Moclobemide
Nefazodone
Noxiptilin
Opipramol
Oxaflozane
Pizotyline
Propizepine
Pyrisuccideanol
Ritanserin
Tandospirone
Tofenacin
Toloxatone

