



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

“DETERMINACIÓN DEL GRUPO FARMACÓFORO
DE LA FAMILIA DE LOS ÉSTERES ANTRANÍLICOS
CON ACTIVIDAD ANTITUMORAL”

T E S I S

QUE PARA OBTENER EL TÍTULO DE
QUÍMICA FARMACÉUTICA BIÓLOGA

PRESENTA
MAYRA CATILINA GONZÁLEZ RENDÓN



MÉXICO, D.F.

Febrero 2009.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Jurado asignado:

Presidente	Prof. Elia Brosla Naranjo Rodríguez
Vocal	Prof. Francisco Hernández Luis
Secretario	Prof. Fernando Cortés Guzmán
1er. Suplente	Prof. Blas Flores Pérez
2º. Suplente	Prof. Claudia Ivonne Araiza Saldaña

Sitio en donde se desarrolló el tema:

Departamento de Química Orgánica en el cubículo 203 Edificio F. Facultad de Química y en el Departamento de Fisicoquímica en el Instituto de Química ambos ubicados en Ciudad Universitaria México D.F.

Sustentante

Mayra Catilina González Rendón

Asesor

Dr Fernando Cortés Guzmán

Agradecimientos

A Dios por darme la capacidad de entendimiento.

A mis abuelos: María Carmona Flores, Emilio Rendón Piza, por su crianza, consejos y ser mas que abuelitos.

A mi madre: Sonia Rendón Carmona por su apoyo dado en la vida y su ejemplo de lucha.

A mi inseparable pareja Angel Padilla Salazar, que ha estado a mi lado desde iniciado este camino. Porque entre sonrisas y amor me ha enseñado mucho.
Gracias rey.

A mi tía Uvelina Rendón Carmona por su cariño, comprensión y su apoyo incondicional. Sobre todo su confianza en mí.

Al Dr. Fernando Cortés Guzmán:

Por su infinita paciencia y apoyo, tanto para mi crecimiento académico como personal.

A mi amigo Roberto Johnson (Bob) por escucharme en momentos difíciles y a Paty porque sin conocerme me apoyo tantos años económicamente.

A mis amigos de la facultad en especial a: Joseelyne Hernández, Jorge Reyes, Azucena Sánchez y Carolina Pérez por su amistad, compañerismo, apoyo, confianza y por darme una mano cuando lo necesitaba.

A toda mi familia política, Georgina Salazar, Kikis Salazar y Ángeles Juárez, y demás por las porras

A la UNAM por darme cobijo en sus alas, así como por las experiencias y momentos que he tenido en ella.

Agradecimientos

A DGAPA-UNAM (PAPIIT: IN216205 y IN207107) y CONACyT (52619)
por el financiamiento de los respectivos proyectos.

Al proyecto PAPIIT-IXTLI IN500107 por la beca de tesis de licenciatura que
me otorgó

A DGSCA-UNAM por el tiempo de cómputo que se requirió para los cálculos
necesario para la tesis.

Al Departamento de Química Orgánica en el cubículo 203 Edificio F. Facultad
de Química y en el Departamento de Fisicoquímica en el Instituto de Química
ambos ubicados en Ciudad Universitaria México D.F., donde se realizó esta
tesis.

ÍNDICE

LISTADO DE FIGURAS.....	i
LISTADO DE TABLAS.....	ii
LISTADO DE GRÁFICAS.....	iii
ABREVIATURAS.....	iv
1. INTRODUCCIÓN	3
2. ANTECEDENTES	4
2.1 LOS ANTIINFLAMATORIOS NO ESTEROIDALES (AINE) Y EL CÁNCER	4
2.2 LA ENZIMA CICLOOXIGENASA Y AINE	6
2.3 AINES Y MECANISMOS CONTRA EL CÁNCER	9
2.3.1 ANGIOGENESIS	9
2.3.2 APOPTOSIS.....	10
2.4. EL ÁCIDO FLUFENÁMICO	10
2.5. ANÁLISIS CUANTITATIVO ESTRUCTURA-ACTIVIDAD (QSAR)	11
2.6. QUÍMICA COMPUTACIONAL	14
2.7. MÉTODOS DE MECÁNICA CUÁNTICA	14
2.7.1. FUNCIÓN DE ONDA Ψ	15
2.7.2. HAMILTONIANO ELECTRÓNICO	16
2.8. GAUSSIAN	17
2.8.1. BASES	19
2.9. TEORÍA DE ÁTOMOS EN MOLÉCULAS (AIM)	22
2.10. ESTUDIO QTMS	26
2.10.1. PROPIEDADES DE ENLACE	27
2.10.2. PROPIEDADES ATÓMICAS	28
2.11. ANÁLISIS DE DATOS MULTIVARIABLE	29

2.11.1 PCA	30
2.11.1.1 ESPACIO MULTIDIMENSIONAL K	30
2.11.1.2 EL PRIMER COMPONENTE PRINCIPAL	31
2.11.1.3EXTENDIENDO EL MODELO	31
2.11.2 PLS	33
2.11.2.1ESPACIO MULTIDIMENSIONAL K CON UNA RESPUESTA $M=1$	34
2.11.2.2 EL PRIMER COMPONENTE PLS ($M=1$)	35
2.11.2.3 EXTENDIENDO EL MODELO	36
2.11.2.4 INTERPRETANDO UN MODELO PLS	37
3. PLANTEAMIENTO DEL PROBLEMA	39
3. HIPÓTESIS	39
5. OBJETIVO	39
5.1. METAS	39
6. METODOLOGÍA	40
6.1. OPTIMIZACIÓN	41
6.2. ANÁLISIS DE LA FUNCIÓN DE ONDA	41
6.3. ANÁLISIS ESTADÍSTICO	43
7. RESULTADOS Y DISCUSIÓN	49
7.1. ESTUDIO DE LAS PROPIEDADES LOCALES	46
7.2. ESTUDIO DE LAS PROPIEDADES INTEGRADAS	53
8. CONCLUSIONES	59
9. REFERENCIAS	60
10. ANEXO.....	63

LISTADO DE FIGURAS

- Figura 1.** Síntesis de prostaglandinas.¹⁵ Las estructuras de cada prostaglandinas se muestran en la Figura 2. Pág.6
- Figura 2.** El ácido de Araquidónico primero es liberado de glicerofosfoípidos de la membrana por la acción de la fosfolipasa A₂ (PLA₂). Las enzimas ciclooxigenasas (COX) entonces catalizan la biosíntesis del endoperoxido bicíclico PGG₂ intermedio, seguida por la reducción a PGH₂. PGH₂ se convierte posteriormente, a una de varias prostaglandinas estructural relacionadas (PGs), incluyendo PGE₂, PGD₂, PGF₂α, PGI₂ y tromboxano A₂ (TxA₂), por la actividad de las PG sintasas específicas.¹⁴ Pág. 7
- Figura 3.**Figura que describe lo básico para obtener un modelado de QSAR.²² Pág. 12
- Figura 4.** Gráfico molecular del cubano, donde se señala las trayectorias de enlace y los diferentes puntos críticos que se pueden presentar.
Código de colores: C = negro, H = gris, punto crítico de enlace (PCE) = rojo, punto crítico de anillo (PCA) = amarillo, punto crítico de caja (PCC) = verde.²⁶ Pág. 24
- Figura. 5** a) Se ilustran las superficies interatómicas (SIA), los puntos críticos de enlace (PCE) y las líneas interatómicas (LIA) en una molécula del eteno, b) La misma molécula pero ilustrando las trayectorias del campo vectorial de $\nabla\rho$, los núcleos están indicados por los círculos en blanco y las líneas negras son las líneas de $\nabla\rho$. Pág. 26
- Figura 6:** a) Un espacio K-dimensional, sólo se muestra 3 ejes que representan 3 variables. El largo de cada uno de los ejes representa la escala de la variación. Las observaciones (renglones) en la matriz de datos X se presentan como un “enjambre” de puntos. b) En la calibración del espacio multidimensional K, primero se extraen los promedios de las variables. Esto resulta en un punto que representa el promedio de todos los valores en el conjunto de datos (punto gris) y está situado en el centro del “enjambre” de datos. c) La calibración del espacio multidimensional K resulta en un reposicionamiento del origen de los ejes coordenados de manera que el promedio de los datos coincida en el centro de estos ejes (punto gris).^{30,31} Pág. 30
- Figura 7:** a) El primer componente principal, PC1, está representado por una línea que mejor se ajusta en el “enjambre” de datos. Cada observación puede ser proyectada hacia esta línea para obtener una distancia, conocida como score. b) El segundo componente principal, PC2, se genera de tal manera que refleje la segunda fuente de variación principal dentro de la matriz de datos X. Esta línea es ortogonal al primer componente principal y también cruza el conjunto de datos por el promedio. c) Dos componentes principales forman un plano. Cada observación se puede proyectar hacia este plan resultando en una nueva distancia “score”. Pág. 32
- Figura 8:** Un conjunto de datos con N observaciones y K variables forman la matriz X y el mismo número de observaciones N con M variables de respuesta forman la matriz Y. Las observaciones son las moléculas a analizar en el análisis PCA. Las variables son los descriptores electrónicos obtenidos de la densidad electrónica para capturar las propiedades de las observaciones y se utiliza el análisis PLS para relacionar X con Y. PLS permite predecir X a partir de Y. Pág. 34

- Figura 9:** Un conjunto de datos con $K=3$ variables X y $M=1$ variable Y . La longitud de cada eje coordenado ha sido ajustada de acuerdo a la media de los datos. Pág. 34
- Figura 10:** En un problema de regresión PLS, las observaciones se distribuyen en dos regiones del espacio multidimensional. Uno con los predictores (el espacio X) y otro con las variables de respuesta (el espacio Y). Los datos se han ajustado a la media de todo el conjunto de datos y los sistemas de coordenadas de ambos espacios se han ajustado a este valor (punto oscuro). Pág. 35
- Figura 11:** Con una sola variable en Y , el espacio se reduce a un solo vector de una dimensión. El primer componente en el modelo PLS entonces se orienta de manera que describa lo mejor posible los puntos en el espacio X mientras que dé una buena correlación con el vector de Y . Las proyecciones de las observaciones hacia la línea del espacio en X da los valores “score” de cada observación.³⁰ Pág. 36
- Figura 12:** La segunda proyección en el espacio X es ortogonal a la primera. Si se proyectan las observaciones en este plano, se tiene el vector “score” t_1 . Los dos componentes forman un plano en el espacio X .³⁰ Pág. 36
- Figura 13.** Estructura esqueleto del ácido flufenámico.¹ Pág. 40
- Figura 14.** El patrón de numeración que se otorga a los átomos de cada compuesto. Pág. 42
- Figura 15.** Etiquetas que se otorgaron a cada enlace de cada compuesto. Pág. 43

LISTADO DE TABLAS

- Tabla 1.** Métodos para cálculo de estructura electrónica. Pág. 18
- Tabla 2.** Conjuntos básicos más utilizados en Gaussian.²³ Pág. 21
- Tabla 3.** Sustituyentes del esqueleto del ácido flufenámico.¹ Pág. 40
- Tabla 4.** Datos obtenidos del National Cancer Institute (NIC) en vitro orientados a la enfermedad humana tumor de células.¹ Las actividades de cada uno de los compuestos, pGI_{50} , es $-\log$ de la concentración molar que causa el 50% de inhibición del crecimiento neto celular. El tipo de tejido celular usado In Vitro para la prueba de actividad; a) K-562 b) NCIH23 c) SW-620 d) SF-268 e) UACC-62 f) SN12C g) MCF-7. Pág. 41
- Tabla 5.** Se muestra cada una de las graficas moleculares con los puntos críticos de cada enlace
Código de colores: carbonos esferas negras, hidrógenos esferas grises, nitrógenos esferas azules, oxígenos esferas rojas, flúor esferas café, cloros esferas verdes, puntos críticos de enlace esferas rojas pequeñas y puntos críticos de anillo esferas amarillas pequeñas. Pág. 45
- Tabla 6.** Correlación y predicción del modelo generado para cada descriptor utilizando PCA. Pág. 47
- Tabla 7.** Resultados de PLS total y de cada una de las actividades biológicas. Pág. 49
- Tabla 8.** Resultados de PLS de los átomos 21-26. Pág. 53
- Tabla 9.** Análisis PCA de cada una de las propiedades de enlace. Pág. 54
- Tabla 10.** Resultados de PLS total y de cada una de las actividades biológicas. Pág. 56

LISTADO DE GRÁFICAS

- Gráfica 1:** Gráfica de scores (t1/t2) del análisis PCA del modelo generado para el descriptor λ . Se muestran tres grupos de compuestos que influyen en la respuesta. Pág. 48
- Gráfica 2.** Resumen de valores de R^2Y y Q^2Y para los compuestos 17 y del 20 al 25 (Grupo II), solo para las respuestas biológicas de leucemia, pulmón y SNC. Pág. 50
- Gráfica 3:** Gráfica de variables importantes en la proyección (VIP) para la regresión. Pág. 50
- Gráfica 4:** Gráfica de datos de actividad biológica observados contra datos de actividad biológica de SNC predichos por el modelo, para los compuestos del grupo II.(ver anexo iv). Pág. 51
- Gráfica 5.** Coeficientes de regresión para los descriptores electrónicos con mayor influencia en la predicción del grupo II. Pág. 51
- Gráfica 6.** Gráfica de datos de actividad biológica observados contra datos de actividad biológica predichos por el modelo, para los compuestos del grupo II usando solo las propiedades K y Vol. Pág. 53
- Gráfica 7.** Gráfica de scores (t1/t2) del análisis PCA del modelo generado para el descriptor V. Se muestran tres grupos de compuestos que influyen en la respuesta. (ver anexo v para el resto de las gráfica de scores). Pág. 55
- Gráfica 8.** Resumen de valores de R^2Y y Q^2Y para los compuestos todos los compuestos excepto la molécula 10. Pág. 56
- Gráfica 9:** Gráfica de variables importantes en la proyección (VIP) para la regresión PLS de los compuestos excepto del 10. Pág. 57
- Gráfica 10:** Gráfica de datos de actividad biológica observados contra datos de actividad biológica predichos (SNC) por el modelo, para los compuestos del grupo II y III. (ver resto de graficas en anexo vi). Pág. 58
- Gráfica 11.** Coeficientes de regresión para los descriptores electrónicos con mayor influencia en la predicción del grupo II. Pág. 58

ABREVIATURAS

AIM	Teoría de átomos en moléculas	PGH2	Prostaglandina H2
AINE	Antiinflamatorios no esteroidales	PGI2	Prostaglandina I2
CoMFA	Comparative molecular field analysis	pGI ₅₀	Menos el logaritmo de la concentración molar que causa el 50% de inhibición del crecimiento neto celular
CoMSIA	Comparative molecular similarity indexes analysis		
COX	Enzima ciclooxigenasa	PLS	Mínimos Cuadrados Parciales
COX 1	Enzima ciclooxigenasa 1	q	Carga atómica
COX 2	Enzima ciclooxigenasa 2	Q ²	Coefficiente de correlación cruzada
ELEP (ϵ)	Elepticidad		
FAP	Polipos adenomatosos familiar	QSAR	Quantitative Structure-Activity Relationship (Análisis Cuantitativo Estructura-Actividad)
G	Densidad de energía cinética		
\hat{H}	Hamiltoniano	QSAR-2D	QSAR bidimensional
H _e	Hamiltoniano electrónico	QSAR-3D	QSAR tridimensional
K	Energía potencial de un átomo	RHO (ρ)	Densidad electrónica
LAP ($\nabla^2\rho$)	Laplaciano	SIA	Superficies interatómicas
LIA	Líneas interatómicas	T _n	Energía cinética de los núcleos
LI	Índice de localización	TxA ₂	Tromboxano A ₂
LM(TK ⁻)	Líneas celulares de fibroblasto de ratón	V	Densidad de energía potencial
LNCaP	Células de cáncer de próstata humano	VEGF	Factor de Crecimiento Endotelial Vascolar
MC	Mecánica cuántica	VIP	Variable Influence on Projection
MM	Mecánica molecular	Vol	Volumen atómico
N	Población atómica de electrones		
NIC	National Cancer Institute		
PC	Punto Crítico		
PCA	Análisis de Componentes Principales		
PCE	Punto Crítico de Enlace		
PC2	Componente principal 2		
PLS	Mínimos cuadrados parciales		
PGs	Prostaglandinas		
PGD2	Prostaglandina D2		
PGE2	Prostaglandina E2		
PGF2	Prostaglandina F2		
PGG2	Prostaglandina G2		

1. INTRODUCCIÓN

El ácido flufenámico es un fármaco del grupo ácido *N*-fenilantranílico con efecto antiinflamatorio y analgésico mediado por la inhibición de la síntesis de la prostaglandina. Como varios de los antiinflamatorio no esteroideos, el ácido flufenámico ha mostrado un efecto antiproliferativo en líneas celulares de fibroblasto. Se ha divulgado la notable actividad anticancerígena de algunos ésteres aromáticos del ácido antranílico que poseen al grupo aza en el esqueleto.

En este trabajo se busca encontrar el sitio farmacóforo del ácido *N*-fenilantranílico por medio de correlacionar los descriptores de la densidad electrónica de las moléculas, tanto en su versión local como integrada, con las actividades biológicas antiproliferativas de estos. Además de localizar cuales descriptores influyen más en la respuesta biológica.

Se usan herramientas computacionales que permiten el análisis estructura-actividad y el análisis estadístico. Se utilizan descriptores de la densidad electrónica como son la densidad de carga (ρ), laplaciano ($\nabla^2\rho$), índice de localización electrónica (λ), que surgen de la teoría de átomos en moléculas. Finalmente la correlación de los descriptores electrónicos con la respuesta biológica se logra utilizando un análisis de datos multivariable. De esta manera se observa que existe una relación entre la densidad electrónica de las moléculas y la actividad biológica que presentan. El sitio de la molécula más importante para la actividad biológica es la región del fenilo del éster.

2. ANTECEDENTES

2.1 LOS ANTIINFLAMATORIOS NO ESTEROIDALES (AINE) Y EL CÁNCER.

Los antiinflamatorios no esteroideos (AINE) se encuentran entre los fármacos más utilizados en todo el mundo. Se consideran eficaces contra la inflamación, además de ser antipiréticos y analgésicos. Adicionalmente la aspirina es eficaz en la prevención de las enfermedades cardiovasculares. Así en los últimos años hay una evidencia creciente, que sugiere una correlación entre el uso de AINE y la menor incidencia de cáncer.

Estudios epidemiológicos y de laboratorio, han aportado pruebas de que los (AINE) juegan un papel importante en la prevención, el desarrollo de cáncer y que podrían ser una terapia alterna contra este mal que aqueja a la humanidad.

Los estudios epidemiológicos han divulgado que la gente que utilizan regularmente los AINE tiene una incidencia más baja de polipos adenomatosos y de muertes por cáncer colorectal comparado con los no usuarios enfermos. Los experimentos en roedores, los estudios clínicos de polipos adenomatosos familiar (FAP) y los estudios epidemiológicos sugieren que los tumores reanudan su crecimiento después de finalizar el tratamiento con AINE.^{2,3,4}

La evidencia clínica comienza con la hipótesis de que los AINE pueden inhibir la ocurrencia o el crecimiento de cáncer colorectal cuando a mediados de los 70s Bennett, Del Tacca⁵ y Jaffe⁶ divulgaron que la concentración de la prostaglandina E2 era más alta en tejido fino colorectal humano con tumor que en la mucosa normal circundante. Esto ha estimulado numerosos estudios en roedores en los cuales se ha demostrado que la aspirina piroxicam,⁷ sulindac,⁸ indometacina,⁹ y celecoxib,¹⁰ inhiben la carcinogénesis inducida en ratas y ratones.

De estos estudios en roedores se derivan varios puntos importantes. Uno de ellos es que los AINE no selectivos suprime crecimiento del tumor a un mayor grado y en dosis más bajas cuando el tratamiento se comienza antes o coincide con la exposición al agente carcinógeno, que cuando se retrasa hasta la fase promoción/progresión del tumor.^{7,8} Otro es que tanto los AINE selectivos como los no selectivos inhiben con eficacia los inicios del desarrollo del

tumor, mientras que solamente los selectivos a la COX-2 son eficaces cuando el tratamiento es retrasado.¹⁰ Finalmente el tratamiento de AINE se debe continuar sin interrupción para prevenir la reaparición del crecimiento del tumor.⁹

2.2 LA ENZIMA CICLOOXIGENASA Y AINE

La enzima ciclooxigenasa (COX) cataliza la formación de las prostaglandinas (PGs) (**Figura 1**). La Prostaglandina es formada por la ciclación oxidativa de los cinco carbonos centrales dentro de un ácido graso poliinsaturado de 20 carbonos tal como el ácido araquidónico.¹⁵

El ácido araquidónico, precursor de prostaglandinas, se encuentra casi exclusivamente como éster en la posición 2 de los fosfolípidos de membrana. El primer paso en síntesis de la prostaglandina es la hidrólisis de los fosfolípidos para producir el ácido araquidónico libre en una reacción catalizada por la fosfolipasa A₂. Después, en una reacción catalizada por la COX, el oxígeno molecular se inserta en el ácido araquidónico para producir un intermediario inestable, la prostaglandina G₂ (PGG₂), que es convertida rápidamente a la prostaglandina H₂ (PGH₂) por la actividad de la peroxidasa de la COX (**Figura 1**). Las isomerazas específicas entonces convierten a la PGH₂ a las diversas prostaglandinas y tromboxanos por la actividad específica de PG sintasa.^{14, 15}

Las PGs tienen funciones importantes en casi cada sistema del organismo y regula diversos procesos fisiológicos tales como inmunidad, reproducción, mantenimiento de la integridad y tono vascular, crecimiento y desarrollo del nervio, y metabolismo del hueso. Las PGs se sintetizan en una amplia gama de diferentes tipos de tejido fino. La COX 1 se expresa constitutivamente en la mayoría de los tejidos finos y parece mediar la producción de las PGs que controlan funciones fisiológicas normales, tales como mantenimiento de la mucosa gástrica y de la regulación del flujo sanguíneo renal. La COX 2, por otra parte, es imperceptible en la mayoría de los tejidos finos normales. Es inducida por los estímulos pro-inflamatorios y mitogénicos, aumentando la síntesis de PGs en tejidos finos inflamados y neoplásicos.^{14, 15}

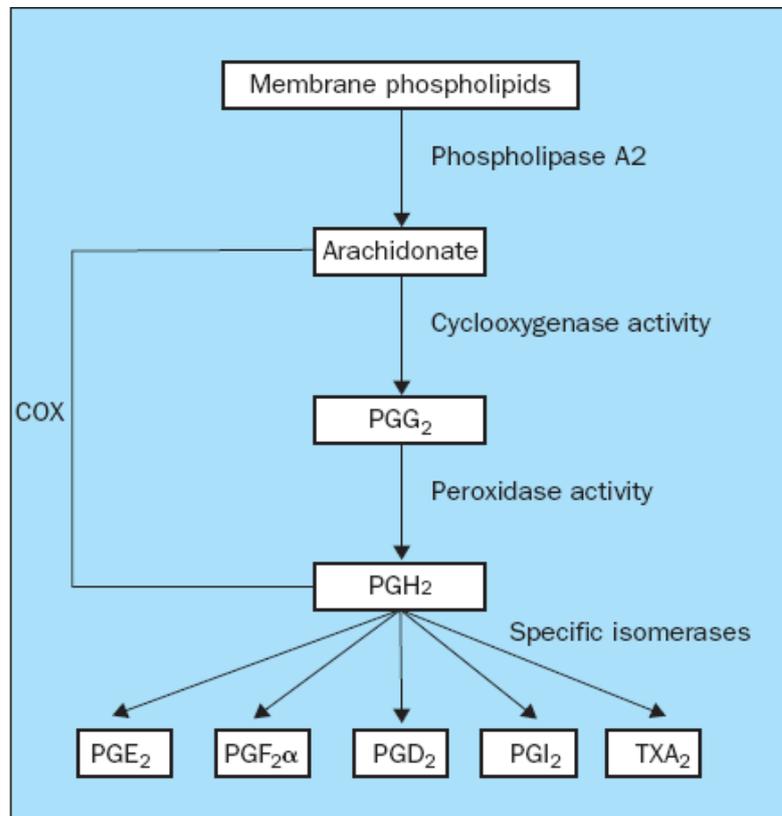


Figura 1. Síntesis de prostaglandinas.¹⁵ Las estructuras de cada prostaglandinas se muestran en la Figura 2.

Los fármacos usados para disminuir la inflamación son los AINE y su mecanismo de acción, está dada por su habilidad para inhibir la actividad de PG G/H sintasa de la enzima COX de tal modo que bloquea la biosíntesis de las prostaglandinas. Los AINE previenen la formación de la PG H₂, el primer paso dado en el metabolismo del ácido araquidónico en una compleja cascada de señalización para la formación de PGs, tales como: PGI₂, PGD₂, PGE₂, PGF₂, y tromboxano (**Figura 2**).^{2, 16}

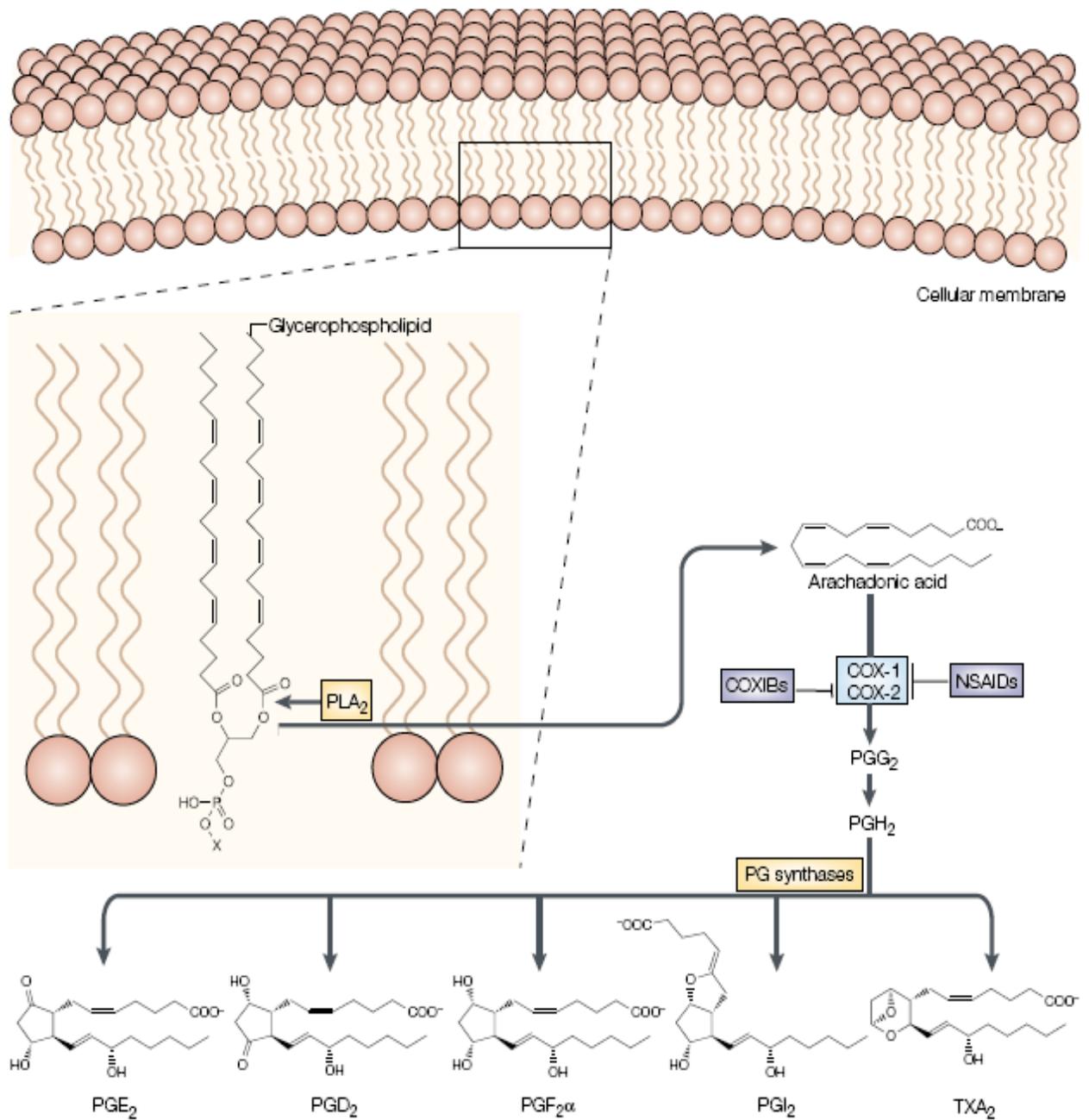


Figura 2. El ácido de Araquidónico primero es liberado de glicerofosfoípidos de la membrana por la acción de la fosfolipasa A₂ (PLA₂). Las enzimas ciclooxigenasas (COX) entonces catalizan la biosíntesis del endoperoxido bicíclico PGG₂ intermedio, seguida por la reducción a PGH₂. PGH₂ se convierte posteriormente, a una de varias prostaglandinas estructural relacionadas (PGs), incluyendo PGE₂, PGD₂, PGF₂α, PGI₂ y tromboxano A₂ (TxA₂), por la actividad de las PG sintasas específicas.¹⁴

2.3 AINE Y MECANISMOS CONTRA EL CÁNCER

Los AINE que inhiben las COXs y la formación de PG se utilizan con frecuencia como agentes anti-inflamatorio, antipirético y analgésico. Publicaciones recientes² revelan entre 40-50% de reducción en la mortalidad de cáncer colorrectal en personas que toman AINE, y la evidencia indica que también afectan la incidencia y la progresión de otros tipos de cáncer.

El debate actual se centra en una amplia gama de los mecanismos de la acción antitumoral de los AINE, en algunos no están relacionados con la inhibición de la actividad COX y la posterior formación de PG, y por otro lado se ha demostrado experimentalmente, que los AINE por su capacidad para inhibir la enzima COX y en especial la COX-2¹⁴ estimulan la apoptosis e inhiben la angiogénesis, dos mecanismos que ayudan a suprimir la transformación maligna y el crecimiento del tumor.¹⁷ Numerosos estudios han demostrado la sobre expresión de la COX-2 en tumores malignos sólidos; de mama, páncreas, próstata y colon, con lo cual apoya una relación causal entre la sobre expresión de la COX-2 y carcinogénesis.¹⁸

2.3.1 ANGIOGÉNESIS

El crecimiento de los tumores depende de un aumento en la formación de nuevos vasos sanguíneos de proveer los nutrientes y oxígeno que necesita y así crece el tumor. Las células tumorales aseguran su propio crecimiento al secretar factores de crecimiento como el factor de crecimiento endotelial vascular (VEGF) que estimulan la angiogénesis.¹⁸ La COX-2 también ha sido implicada en este aspecto de la carcinogénesis por la sobre expresión de la COX-2 en las células del cáncer de colon con lo que aumenta la producción de factores de crecimiento vascular y la migración de las células endoteliales a través de una matriz de colágeno¹⁴ y también se ha visto que generalmente la expresión de la COX-2 induce vasculatura angiogénica en el adenomatoso pólipo de colonrrectal y en carcinomas de colon, pulmón, mama, esófago y de próstata.² Con lo cual el mecanismo clave parece implicar el incremento en la expresión del VEGF. La producción de PGs que pueden estimular directamente la migración de células endoteliales y al factor de crecimiento inductor de angiogénesis.¹⁹

2.3.2 APOPTOSIS

El tamaño de una población de células depende del equilibrio entre la proliferación celular y muerte celular. La disminución de la apoptosis se han observado en células premalignas y tumores malignos. De los muchos factores que regulan la apoptosis, hay un relación inversa entre los niveles de BCL 2 y la apoptosis (**Bcl-2** es una familia de proteínas que regulan procesos de permeabilización mitocondrial y constituyen un punto clave en la vía intrínseca de apoptosis celular.) por ejemplo, en células epiteliales intestinales de rata que sobre expresan COX-2 estable, aumentado las cantidades de BCL 2 y son resistentes a butirato un estimulante de la apoptosis.¹⁵

Por otra parte se trata de la habilidad de los AINE para restablecer la apoptosis en células con deficiencia de APC (la proteína APC, uno de los supresores de tumor, controla la expresión de genes críticos en el proceso de la división celular.). Estas células deficientes de APC se acumulan en los adenomatos pólipos con lo cual se ve suprimida la apoptosis y esto genera mutaciones genética y cambios fenotipicos de estas celulas.²

La supresión de la apoptosis permite a las células APC-deficientes se acumulen en pólipos adenomatosos. Más represión de la apoptosis se produce en el desarrollo de estas células mutaciones genéticas y los cambios fenotípicos adicionales. Ambos AINEs tanto los selectivos como los que no actúan sobre las células deficientes de APC pero en especial los inhibidores de la COX-2 que estimulan la apoptosis y suprime el crecimiento en muchos carcinomas incluyendo cultivo de cáncer humano; de estómago, esófago, lengua, cerebro, pulmón y páncreas².

2.4 EL ÁCIDO FLUFENÁMICO^{1, 2, 11, 12, 13}

El ácido flufenámico es un fármaco de tipo AINE del grupo ácido N-fenilantranílico (fenamates) con efectos antiinflamatorios y potentes analgésicos mediados por la inhibición de la síntesis de la PG. Este mecanismo de acción es responsable de su uso terapéutico pero también tiene efectos secundarios adversos, por ejemplo, problemas gastrointestinales (diarreas). Diversos productos que contienen el ácido flufenámico están disponibles en el

mercado y se administran vía oral o cutánea. Además, el uso cutáneo del fármaco combina la ventaja de una terapia de fármaco local con reducción del metabolismo del primer paso.¹¹

Pero además, como muchos de los AINE también el ácido flufenámico ha mostrado un efecto antiproliferativo en líneas celulares LM(TK⁻) de fibroblasto de ratón con una IC₅₀ de 120mM, y la proliferación retorna cuando es retirado el ácido flufenámico,¹² este mismo efecto lo tiene con células de cáncer de próstata humano (LNCaP).¹³

Aunque, por otra parte se ha reportado que en particular el ácido flufenámico ha reportado que exhibe propiedades inhibitoras de crecimiento en una línea celular de fibroblasto de ratón LM (TK⁻) a través de un bloqueo de un canal catiónico no selectivo 28-pS.² En otro estudio ha mostrado que el ácido flufenámico inhibe la expresión del receptor androgénico en células de cáncer de próstata humano LNCaP.² En trabajos recientes se ha divulgado la notable actividad anticancerígena de algunos esteres aromáticos del ácido antranílico que poseen al grupo aza en el esqueleto.¹

2.5 ANÁLISIS CUANTITATIVO ESTRUCTURA-ACTIVIDAD (QSAR) ^{20, 21}

En la industria farmacéutica, uno de los objetivos más importantes es identificar estructuras químicas de compuestos que tengan una mayor potencia terapéutica con el menor número de efectos colaterales. Una herramienta para lograr este objetivo son los estudios QSAR (Quantitative Structure-Activity Relationship) que intentan correlacionar la estructura o propiedades descriptoras de una serie de compuestos con sus respectivas actividades biológicas o propiedades fisicoquímicas, con la ventaja de predecir la actividad de compuestos aun no sintetizados.

La historia del QSAR comienza hace más de un siglo, en 1869 cuando Crum-Brown y Fraser³³ expresaron la idea de que la acción fisicoquímica de una sustancia fuera una función de su composición y constitución química. Pocas décadas después, en 1893, Richest²¹ demostró que la citotoxicidad de una serie de compuestos químicos era inversamente proporcional a sus correspondientes solubilidades acuosas. En el siglo XX Meyer y Overton

independientemente sugirieron que la acción narcótica de un grupo de compuestos orgánicos estaba paralelamente relacionada con su coeficiente de partición en aceite de olivo/agua.²¹

Pero fue Hansch a principio de los años 60's quien introduce la metodología de los estudios QSAR, la cual está fundamentada en que las diferencias estructurales de los compuestos son responsables de las diferencias en la actividad biológica, donde los parámetros se obtienen a través de un análisis de regresión múltiple.²⁰

En el análisis QSAR es imperativo que los datos a utilizar, ya sean biológicos o de otro origen sean adecuados y confiables para poder desarrollar un modelo con significado real. La selección de los parámetros es un paso fundamental para el éxito de cualquier QSAR, ya que si la relación entre los parámetros seleccionados y la actividad es buena, entonces será posible la predicción de actividades biológicas o propiedades experimentales. Los tres parámetros que fueron propuestos en un inicio, y que aun siguen dominando, son de naturaleza hidrofóbica, estérica y electrónica. En extensos estudios se revela que los atributos electrónicos de las moléculas están íntimamente relacionados a sus reactividades químicas y actividades biológicas.^{20, 21}

Hoy en día existen varios métodos QSAR que pueden dividirse en varios grupos. Uno donde se usa un número relativamente pequeño de propiedades fisicoquímicas y parámetros como los efectos hidrofóbicos, estéricos y electrostáticos. Este tipo de método es conocido con el nombre de análisis de Hansch. Entre los métodos más recientes esta los QSAR bidimensionales (QSAR-2D), y el QSAR tridimensional (QSAR-3D).^{20, 21}

El bidimensional es llamado así por que las graficas moleculares o fórmulas estructurales son bidimensionales. Generalmente, toma la forma de una ecuación lineal, sin embargo, existen comportamientos que no lo son. La información que se obtiene de este tipo de modelos ya sea lineal o no es de gran ayuda para calcular actividades biológicas, para explicar comportamientos biológicos, encontrar similitudes químicas entre moléculas a partir de descriptores. Este tipo de método es bueno para analizar una gran cantidad de compuestos.²¹

El otro método se basa en los descriptores obtenidos de la contribución tridimensional de las estructuras de las moléculas. Este tipo de análisis se le conoce como QSAR-3D que permite establecer la correlación entre una serie de estructuras y sus interacciones biológicas en un sitio blanco; tiene la ventaja de mostrar sus resultados en forma gráfica que pueden interpretarse fácilmente en términos de las interacciones estéricas y electrostáticas importantes para que el ligando se una al receptor, pero se dificulta en el análisis de bases de datos grandes. Probablemente los ejemplos más conocidos de estos son los CoMFA (Comparative Molecular Field Analysis) y el CoMSIA (Comparative Molecular Similarity Indices Analysis).²¹

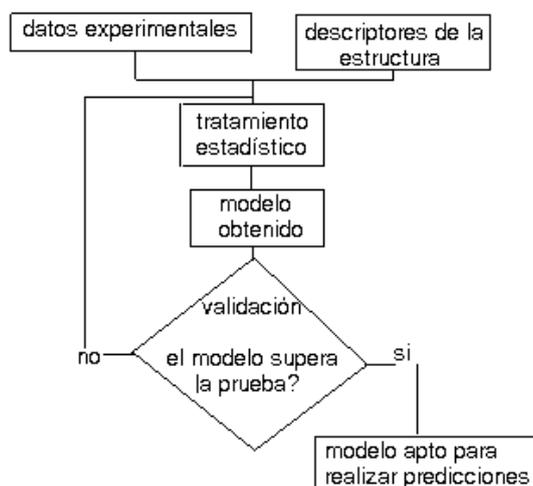


Figura 3. Figura que describe lo básico para obtener un modelado de QSAR.²²

El desarrollo de un modelo QSAR se puede simplificar de la siguiente manera: se monitorea las actividades biológicas para cierto tipo de compuestos, esta actividad constituye la matriz de datos $Y = (N \times M)$, donde M son las variables de respuesta, por otro lado se obtienen descriptores que reflejan las propiedades estructurales y químicas, este nuevo conjunto de datos es la matriz $X = (N \times K)$, donde K es el número de descriptores y en ambas matrices N es el mismo número de compuestos.

Después, viene el tratamiento estadístico para obtener el modelo predictorio. La siguiente etapa es la validación del modelo. Esta etapa es indispensable debido a la posibilidad de que las correlaciones encontradas sean producto de la casualidad. La validación consiste en evaluar la

capacidad que tiene el modelo QSAR para predecir las propiedades de interés con exactitud (por ejemplo la actividad biológica) de los modelos que no fueron utilizados en el desarrollo del modelo QSAR.²² (**Figura 3**)

Al grupo de compuestos con el que se desarrolla el modelo se le llama grupo de entrenamiento, mientras que al grupo con el que se valida se le llama grupo de prueba, este último debe contener al menos 5 compuestos. Otros dos términos muy utilizados en la validación QSAR, son "*training set*" y "*test set*" (conjunto de entrenamiento y de prueba, respectivamente). Para la validación del modelo, al conjunto original se le hace una partición en dos subconjuntos, tal que el primero sirve para determinar el modelo y luego aplicarlo al segundo, y viceversa, operación denominada "*Cross Validation*" o validación cruzada. La concordancia entre las predicciones y los datos experimentales sirven para cuantificar la bondad del modelo original. La validación cruzada más común es la llamada "*leave one out*", que consiste en extraer una molécula del conjunto inicial de entrenamiento y hacer el cálculo con el conjunto de prueba.²²

La correlación obtenida se utiliza para predecir la actividad/propiedad de la molécula no incluida en la determinación del modelo. Este proceso se repite con todas las moléculas restantes y se calcula la bondad de las predicciones, la cual sirve para tener una medida de la bondad del modelo estudiado. Así, si se tiene N compuestos, entonces se derivan N modelos y las N predicciones se comparan con los N valores experimentales. Como resultado se tiene un coeficiente de correlación cruzada (Q^2).²²

Hoy en día, una de las aplicaciones más importante que tienen los estudios QSAR es la predicción de una gran cantidad de compuestos en forma rápida. Esto se debe al creciente desarrollo de la química combinatoria y a las pruebas biológicas de alto rendimiento. De esta manera se puede predecir en forma virtual la actividad de compuestos que podrán ser sintetizados posteriormente. Todo esto hace que en la era moderna de la química farmacéutica, los estudios QSAR constituyan una de las herramientas más importantes del diseño de fármacos asistido por computadora.

2.6 QUÍMICA COMPUTACIONAL²³

La química computacional estudia estructuras y reacciones químicas, basándose en las leyes fundamentales de la física. Algunos métodos no sólo permiten visualizar moléculas estables, sino también estudiar estados de transición e intermediarios de vida media corta, difíciles de aislar experimentalmente.

Existen dos áreas principales en la química computacional orientadas al estudio de las moléculas y su reactividad, estas son la mecánica molecular (MM) y la mecánica cuántica (MC).²³ Ambas tienen en común los siguientes aspectos.²³

- Se calcula energías de una estructura particular, es decir, para un arreglo espacial de átomos.
- Se optimizan geometrías y se encuentra el mínimo de energía para una estructura molecular cercana a la estructura de partida. La optimización de geometría depende de la primera derivada de la energía con respecto a las posiciones de los núcleos.
- Cálculos de frecuencia vibracionales, resultado del movimiento interatómico en la molécula. Las frecuencias dependen de la segunda derivada de la energía con respecto a las posiciones de los núcleos.

2.7 MÉTODOS DE MECÁNICA CUÁNTICA²³

Los métodos de cálculo de estructura electrónica están basados principalmente en las leyes de la mecánica cuántica. Los estados cuánticos, energías y otras propiedades relacionadas, se obtienen resolviendo la ecuación de Schrödinger:

$$\hat{H}\Psi = E\Psi \quad \text{Ec. 1}$$

Donde:

Ψ es la Función de onda para una partícula en un estado estacionario. \hat{H} es el operador de Hamilton o hamiltoniano que está relacionado con la energía total de un sistema y E es la energía electrónica.

2.7.1 FUNCIÓN DE ONDA Ψ^{23}

La mecánica cuántica considera una serie de postulados, con base en los cuales se puede obtener información del sistema en estudio. Antes de hablar sobre el postulado principal es preciso definir algunos términos. Una variable dinámica es cualquier magnitud física asociada a un sistema, por ejemplo posición, momento lineal, momento angular y energía, y un observable es cualquier variable dinámica que puede ser medida. A continuación se describe el primer postulado base de la mecánica cuántica:

Postulado I

El estado de un sistema dinámico de N partículas está descrito por una función $Y(q_1, q_2, \dots, q_N, t)$.

*La cantidad $Y^*Y(q_1, q_2, \dots, q_N, t)$ es proporcional a la probabilidad de encontrar a la partícula 1, entre q_1 y $q_1 + dq_1$; a la partícula 2, entre q_2 y $q_2 + dq_2$,... y a la partícula N entre q_N y $q_N + dq_N$ simultáneamente para un tiempo específico t .*

Este postulado afirma que toda la información acerca de las propiedades de un sistema está contenida en una función Ψ que es una función de las coordenadas (q_i) de las N partículas y del tiempo (t). Si la función de onda incluye explícitamente al tiempo, se denomina función de onda dependiente del tiempo. Si las propiedades observables del sistema no cambian con el tiempo, se dice que el sistema se encuentra en un estado estacionario. La segunda parte del postulado sostiene que la cantidad $\Psi^*\Psi dx$ es la probabilidad de encontrar a la partícula entre x_i y $x_i + dx_i$ al tiempo t dado. La función Ψ es compleja, y no reside en el espacio real. El problema para visualizarla se resuelve introduciendo el producto de Ψ por su complejo conjugado Y^*Y o $|\Psi|^2$ e integrando sobre todos los electrones menos uno. A la cual se le denomina también como densidad de probabilidad.

$$\rho(x_i) = \int \Psi^* \Psi dx_2 \dots dx_N \quad \text{Ec. 2}$$

Para que estas funciones tengan una interpretación física, deben sujetarse a las restricciones siguientes:

- La función debe ser continua, al igual que su primera y segunda derivada.
- La función debe ser univaluada.
- El cuadrado de la función debe ser integrable.

Estas restricciones surgen del postulado que dice que $\Psi^*\Psi dt$ es proporcional a una probabilidad, por lo que ésta debe ser finita. Si la función cumple con la condición:

$$\int \Psi^* \Psi dt \leq a \quad \text{Ec. 3}$$

Se dice que la función está normalizada. En este caso $\Psi^*\Psi$ es directamente una densidad de probabilidad y la ecuación 12 indica que todo el espacio debe ser igual a una constante “a”.

2.7.2 HAMILTONIANO ELECTRÓNICO ²³

Todas las moléculas están formadas de electrones y núcleos; para su estudio se analiza la interacción entre estos obteniendo el operador Hamiltoniano, el cual está relacionado con la energía total de un sistema. El operador Hamiltoniano está conformado por el Hamiltoniano electrónico y el Hamiltoniano nuclear, bajo la aproximación de Born-Oppenheimer.

$$H = H_e + H_n \quad \text{Ec. 4}$$

El hamiltoniano electrónico es el resultado de la suma de las contribuciones de la energía cinética electrónica y las energías potenciales debidas a las interacciones entre electrones y entre los electrones y los núcleos.

$$H_e = T_e + V_{ne} + V_{ee} \quad \text{Ec. 5}$$

Donde H_e es el Hamiltoniano electrónico. T_e es la energía cinética de los electrones. V_{ne} es la energía resultante de la interacción núcleo-electrón y V_{ee} es la energía resultante de la interacción electrón-electrón.

El hamiltoniano nuclear representa la energía cinética de los núcleos:

$$H_n = T_n \quad \text{Ec. 6}$$

Donde T_n es la energía cinética de los núcleos.

La resolución del hamiltoniano de una molécula es muy complicada, para simplificarla se efectúa la aproximación de Born-Oppenheimer que se basa en el hecho de que los núcleos son mucho más pesados que los electrones. Por tanto, los electrones se mueven más rápido que los núcleos y se puede considerar que estos últimos se encuentran fijos en el espacio, mientras que los electrones llevan a cabo su movimiento. De esta manera se puede omitir el término de energía cinética de los núcleos y considerar a V_{nn} como una constante, así el hamiltoniano se simplifica para obtener la ecuación 7.

$$H = T_e + V_{nn} + V_{ne} + V_{ee} \quad \text{Ec. 7}$$

Donde V_{nn} es la energía resultante de la interacción núcleo-núcleo.

La ecuación de Schrödinger (**Ec. 1**) puede ser resuelta exactamente sólo para pocos sistemas. Los métodos de cálculo de estructura electrónica se caracterizan por las aproximaciones matemáticas para resolver esta ecuación, estos pueden ser métodos semiempíricos y métodos ab-initio:

Los métodos ab-initio (a primeros principios), no utilizan parámetros experimentales en el cálculo, están basados únicamente en las leyes de la mecánica cuántica y valores de las constantes físicas fundamentales (velocidad de la luz, masas y cargas de electrones y núcleos, constante de Plank, etc). Los métodos ab-initio, predicen con mayor exactitud desde el punto de vista cuantitativo y cualitativo, pero el costo computacional aumenta sustancialmente, son capaces de manejar cualquier tipo de átomos, y toma en cuenta tanto electrones de valencia como de *core* sin necesidad de ningún parámetro.

2.8 GAUSSIAN²⁴

Existen varios programas para el cálculo de la estructura electrónica, entre ellos Gaussian. Este programa tiene la ventaja de ser el que utiliza los algoritmos matemáticos más actuales. Gaussian es un sistema de programas capaces de ejecutar una variedad de cálculos de estructura electrónica, capaz de predecir diversas propiedades moleculares y reacciones químicas, incluyendo:

- Energías y estructuras moleculares.
- Energías y estructuras de estado de transición.
- Frecuencias vibracionales.
- Espectro infrarrojo y Raman.
- Propiedades termodinámicas.
- Energías de reacción y de enlace.
- Trayectorias de reacción.
- Orbitales moleculares.
- Cargas atómicas.

Gaussian esta diseñado para sistemas en fase gaseosa y en solución, en estado basal y estados excitados; lo que permite explorar áreas de interés químico, como lo son: los efectos del

sustituyente, mecanismos de reacción, superficies de energía potencial, energías de excitación, etcétera.

La filosofía teórica detrás de Gaussian está definida por el siguiente principio:

Un modelo teórico deberá ser uniformemente aplicable a sistemas moleculares de cualquier tipo y tamaño, hasta un tamaño máximo determinado solamente por la habilidad práctica de los recursos computacionales.

Un modelo ab-initio en Gaussian debe ser definido únicamente por una configuración de núcleos y electrones, por lo que solamente se requiere la especificación de la geometría molecular inicial, y ningún otro parámetro es necesario para especificar el problema. El resultado depende en gran parte del método, por esto la elección de este depende del sistema que se quiera estudiar, por ejemplo, el método Hartree-Fock, reproduce la solución exacta a la ecuación de Schrödinger, para problemas de un electrón (átomo de hidrógeno), esto permite calcular exactamente moléculas como H_2^+ . Para otros sistemas se necesitan métodos más completos como los que se muestran en la **Tabla 1**.

Un modelo químico consiste en la combinación de un método teórico con un conjunto base. Cada par único método/conjunto-base representa una aproximación diferente a la ecuación de Schrödinger. Los resultados para sistemas diferentes pueden ser comparados y evaluados cuando los cálculos se han realizado por el mismo modelo químico; es decir, para hacer una comparación formal de diferentes geometrías moleculares (conformaciones, efectos del sustituyente, etc.), se debe hacer utilizando el mismo par método/conjunto-base.

Tabla 1. Métodos para cálculo de estructura electrónica.

Palabra clave	Método
Métodos autoconsistentes	
HF	Hartree-Fock auto-consistente
CASSCF	Espacio activo completo auto-consistente
GVB	Enlace valencia generalizado

Métodos de correlación electrónica

MPn	Teoría de perturbación de Moller-Plesset de orden n
QCISD	Interacción de configuración cuadrática
BD	Dobletes de Brueckner
CI, CISD	Método de interacción de configuraciones

Teoría del funcional de la densidad

B3LYP	Combinación de los funcionales de gradiente corregido de Becke, Lee, Yang y Parr
HFS	Intercambio de densidad de spin local de Slater
Xa	Intercambio Xa

Energías de gran exactitud

G1	Teoría gaussian-1
G2,G2MP2	Teoría gaussian-2 y variación
CBS-x	Método de conjunto de base completo.

2.8.1 BASES^{23, 24}

Un conjunto base es una descripción matemática de los orbitales en un sistema molecular, los conjuntos bases más grandes constituyen una aproximación más exacta de los orbitales por imponer pocas restricciones sobre la localización de los electrones. De esta manera las bases son parte del modelo químico junto con el método teórico. La mayoría de los métodos basados en la función de onda, los métodos semiempíricos y los métodos DFT usan conjuntos base. Los conjunto base estándar para cálculos de estructura electrónica son combinaciones lineales de funciones gaussianas para formar cada orbital, de allí el nombre del paquete computacional Gaussian.

La expansión de una función desconocida, como un orbital molecular, en un conjunto de funciones conocidas, no es una aproximación si se tiene una base completa; sin embargo, para

el espacio vectorial que contiene a los orbitales moleculares, una base completa significa el uso de un número infinito de funciones base.

Gaussian ofrece un amplio espectro predefinido de conjuntos base, los cuales pueden clasificarse por el número y tipo de funciones base que contienen. Los conjuntos base asignan un grupo de funciones base a cada átomo en una molécula para aproximar sus orbitales, estas funciones base son una combinación lineal de funciones gaussianas, a cuyas funciones bases se les refiere como funciones contraídas, y las funciones gaussianas componentes son referidas como primitivas. Un conjunto base que consiste en una simple función gaussiana es llamada no-contraída.

Los tipos de funciones base son:

- *Base mínima.* Comprende un número mínimo de funciones base necesaria para cada átomo, este conjunto sólo incluye los orbitales internos y de valencia, por ejemplo para el átomo de carbono, sólo se toman en cuenta los conjunto base de los orbitales: 1s, 2s y 2p (x,y,z). El conjunto base llamado STO-nG corresponde a una base mínima que tiene orbitales tipo Slater representado cada uno por n gaussianas, donde n es usualmente 3 ó 6.
- *Base de valencia.* Este conjunto sólo describe cada átomo utilizando bases para los orbitales de valencia, por ejemplo, para el átomo de carbono, tendríamos que describir los orbitales 2s y 2p (x,y,z).
- *Base extendida.* Este conjunto toma en cuenta los orbitales internos, los orbitales de valencia y además un conjunto de orbitales no ocupados llamados orbitales virtuales, por ejemplo, para el átomo de carbono tendríamos los orbitales 1s, 2s, 2p (x,y,z), 3s, 3p (x,y,z).

Entre los conjuntos base de tipo extendido, tenemos:

- *Conjunto base de valencia dividida.* Dos conjuntos de funciones en la región de valencia da una representación más exacta de orbitales.
- *Conjunto base polarizada.* Añade orbitales con un momento angular mayor que el requerido para la descripción de cada átomo, por ejemplo una base polarizada añade una función d a un carbono ó una función f a un metal de transición, y en algunos casos, añade una función p al átomo de hidrógeno.

- *Funciones difusas*. Hace que la función orbital decaiga mucho más suave, para esto se agregan contribuciones espaciales al orbital. Las funciones difusas son muy importantes para sistemas donde los electrones están relativamente lejos del núcleo como moléculas cargadas, sistemas en estado excitado, sistemas con bajos potenciales de ionización, descripción de acidez absoluta, etc.

Mientras mayor sea el conjunto se requerirá de mayores recursos computacionales. Moléculas grandes son imprácticas de calcularlas en conjuntos de base grandes por el tiempo de cómputo que puede ser de años. La siguiente tabla resume los tipos de conjuntos básicos más utilizados²³:

Tabla 2. Conjuntos básicos más utilizados en Gaussian.²³

Base	Características
STO-NG	Base mínima con N gaussianas en combinaciones fijas para cada una de las funciones de la base; útil para datos cualitativos de sistemas grandes
3-21G y 6-31G	Son doble zeta con tres o seis funciones primitivas en cada una de las funciones de la base mínima que describen el <i>core</i> de los átomos y dos o tres primitivas en la primera de las dos contraídas en la doble zeta de valencia. En la región de valencia da una representación más exacta de orbitales.
3-21G*, 3-21G(*), 31G*, 6-31G(d),	6- Igual que la anterior pero incluyendo funciones de polarización en los átomos de la primera fila (excepto los hidrógenos) cuando el asterisco está sin paréntesis o aparece como (d). El asterisco entre paréntesis indica que se usan funciones de polarización desde el Na hacia arriba.
6-31G**, 31G(d,p), 31G(2d,2p)	6- Igual que la anterior, pero incluyendo también funciones de polarización en los hidrógenos (segundo asterisco). Cuando se incluyen varias funciones de polarización (en lugar de sólo una) se indica como (nd, mp) con n y m el número de funciones de polarización <i>d</i> y <i>p</i> .
6-311++G(3df,3pd)	Indica una base mínima de seis componentes para el <i>core</i> , tres funciones contraídas a partir de cinco primitivas para cada número angular en la capa de la valencia, funciones difusas <i>sp</i> en los átomos de la primera fila (el primer +), funciones difusas en los H (segundo +), tres funciones de polarización <i>d</i> y una <i>f</i> en los átomos de la primera fila y tres funciones de polarización <i>p</i> y una <i>d</i> en los hidrógenos.

2.9 TEORÍA DE ÁTOMOS EN MOLÉCULAS (AIM) ²⁵

La teoría de átomos en moléculas (AIM) fue desarrollada por el profesor Richard Bader²⁵ (Universidad de MacMaster) y ha sido utilizada ampliamente para encontrar la explicación de varios problemas químicos. Esta teoría es una extensión de la química cuántica para un átomo en una molécula. Se define a un átomo dentro de una molécula en el espacio real donde la densidad electrónica es el medio para predecir las propiedades atómicas.

Para extraer información a la densidad electrónica se necesita estudiar el gradiente de la densidad ($\nabla\rho$) y no la densidad directamente (ρ). $\nabla\rho$ es un campo vectorial y tiene tres características:

- $(\nabla\rho)$ apunta en la dirección en la cual crece ρ .
- $(\nabla\rho)$ es perpendicular en cualquier lugar a una isosuperficie de ρ .
- Al $(\nabla\rho)$ se le asocian líneas de flujo que tiene su origen y fin en puntos determinados.

Cada característica topológica de la $\rho(r)$, ya sea un máximo, un mínimo o un punto silla, esta asociado con un punto crítico (PC) donde $\nabla\rho(r_c)=0$. La descripción matemática para identificar a un punto crítico está basada en el espectro de eigenvalores del Hessiano de la matriz de $\nabla\nabla\rho$. El Hessiano es una matriz de segundas derivadas de funciones con respecto a todas las posibles combinaciones de coordenadas. El Hessiano de 3x3 de la densidad electrónica tiene la siguiente expresión (**Ec. 8**):

$$\nabla\nabla\rho = \begin{pmatrix} \frac{\partial^2 \rho}{\partial x^2} & \frac{\partial^2 \rho}{\partial x \partial y} & \frac{\partial^2 \rho}{\partial x \partial z} \\ \frac{\partial^2 \rho}{\partial y \partial x} & \frac{\partial^2 \rho}{\partial y^2} & \frac{\partial^2 \rho}{\partial y \partial z} \\ \frac{\partial^2 \rho}{\partial z \partial x} & \frac{\partial^2 \rho}{\partial z \partial y} & \frac{\partial^2 \rho}{\partial z^2} \end{pmatrix} \quad \text{Ec. 8}$$

Esta matriz puede ser evaluada en cualquier punto del espacio, pero para la teoría AIM es necesario evaluar dicha matriz en un PCE. Cada elemento de la matriz puede ser interpretado como una curvatura, la estructura intrínseca de ρ no depende del sistema de coordenadas seleccionado porque no importa como una molécula es rotada, su función de onda, y por lo

tanto su densidad electrónica, permanece inalterada. Para eliminar las arbitrariedades de la representación del Hessiano se encuentran los eigenvalores y sus correspondientes eigenvectores de $\nabla\nabla\rho$ en un PCE dado. Es una propiedad bien conocida que los eigenvalores sean invariables con respecto a la rotación y de este modo es ideal revelar la estructura interna del Hessiano. La ecuación de eigenvalor tiene tres soluciones posibles, cada una obedece la siguiente ecuación (**Ec. 9**):

$$(\nabla\nabla\rho) \mathbf{u}_i = \lambda_i \mathbf{u}_i \quad (i = 1, 2, 3) \quad \text{Ec. 9}$$

Dónde λ_i es el eigenvalor i y \mathbf{u}_i es una columna de vectores con tres componentes representando el eigenvector asociado.

Los tres eigenvectores son mutuamente ortogonales y coinciden con los llamados ejes principales de curvatura. Este nombre se refiere al hecho de que las magnitudes de las tres segundas derivadas de la ρ calculadas con respecto a estos ejes son extremizadas (localizar máximos y mínimos de dicha función). Además, los tres eigenvalores son siempre reales y son las únicas curvaturas que no desaparecen, comparadas con los nueve elementos del Hessiano. El Hessiano se convierte en una matriz diagonal cuando se expresa con respecto al sistema de ejes constituido por los ejes principales de curvatura. Cada eigenvector representa un eje y el correspondiente eigenvalor determina el perfil de la densidad electrónica a lo largo de este eje. La curvatura, la segunda derivada de $\rho(r)$, es negativa en un máximo y positiva en un mínimo. Por ejemplo, si uno de los eigenvectores en un PCE está localmente dirigido a lo largo de la trayectoria de enlace y si su eigenvalor asociado es positivo, entonces se sabe que la ρ es un mínimo a lo largo de la trayectoria de enlace.

El número de eigenvalores diferentes a cero y sus signos asociados son suficientes para definir el tipo de PC. La clasificación de un PC se da por dos números:

- ω : número de curvaturas diferentes de cero
- σ : la suma algebraica de los signos de las curvaturas. El número -1 es asignado a un eigenvalor negativo y +1 a un eigenvalor positivo. σ es definida como la suma de estos tres números. Por ejemplo en un PCE el Hessiano posee un eigenvalor positivo y dos eigenvalores negativos: $(+1)+(-1)+(-1) = -1$.

De esta manera existen cuatro tipos de PC (no singulares) posibles en la densidad electrónica, los cuales son descritos a continuación e ilustrados en la **Figura 4**:

- 3,-3 Todas las curvaturas de $\rho(r)$ en el PC son negativas y por tanto $\rho(r)$ es un máximo local en r_c . Este punto se asocia a un atractor, a un núcleo.
- 3,-1 Dos curvaturas son negativas y una positiva. $\rho(r)$ es un máxima en r_c en el plano definido por dos ejes asociados a las curvaturas negativas y es un mínima en el eje perpendicular a ese plano. Este punto se asocia a un punto crítico de enlace.
- 3,+1 Dos curvaturas positivas y una negativa. $\rho(r)$ es un mínima en r_c en el plano definido por los ejes asociados a las curvaturas positivas y es un máximo en r_c a lo largo del eje perpendicular a este plano. Este punto se asocia a un punto crítico de anillo.
- 3,+3 Todas las curvaturas son positivas y $\rho(r)$ es un mínimo local en r_c . Este punto se asocia a un punto crítico de caja.

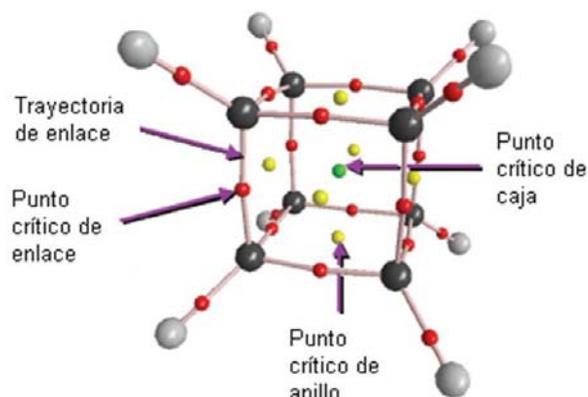


Figura 4. Gráfico molecular del cubano, donde se señala las trayectorias de enlace y los diferentes puntos críticos que se pueden presentar. Código de colores: C = negro, H = gris, punto crítico de enlace (PCE) = rojo, punto crítico de anillo (PCA) = amarillo, punto crítico de caja (PCC) = verde.²⁶

Cuando se analiza la topología de la densidad de una molécula es necesario encontrar todos los puntos críticos de los diferentes tipos para satisfacer la relación de Poincaré-Hopf, $n - b + r - c = 1$, donde n es el número de núcleos, b es el número de enlaces, r es el número de anillos y c es el número de cajas en la molécula.

Como la densidad electrónica es mayor en los núcleos que en sus alrededores las líneas de $\nabla\rho$ se originan en el infinito para terminar en el núcleo. Se puede decir que las líneas del $\nabla\rho$ son atraídas al núcleo, y es la razón por lo que se les denominan atractores nucleares y dominan una porción del espacio. Estas regiones dividen a las moléculas en segmentos que pueden identificarse con el concepto químico de átomo.

La región dominada por el núcleo se conoce como contenedor atómico. Con lo anterior se puede definir al átomo como: “La unión de un atractor (núcleo) y su contenedor atómico asociado”. Cada átomo refleja las características del ambiente químico particular.

Existe un conjunto de líneas de $\nabla\rho$ que inician en el infinito y terminan en un PC entre dos átomos, este conjunto de líneas constituyen una superficie interatómica (SIA), también llamada *superficie de flujo cero*. Esta superficie se distingue de cualquier otra en que en cada punto el vector n que caracteriza a la SIA (ortogonal a la superficie) es ortogonal al $\nabla\rho$ para cada punto en la SIA, en otras palabras no hay líneas del $\nabla\rho$ que crucen la SIA, como los muestra la **Ec. 10**.

$$\nabla\rho(r) \cdot n(r) = 0 \quad \forall r \in S(r) \quad \text{Ec. 10}$$

Además, hay líneas del $\nabla\rho$ que se originan en este PC y terminan en los). A este punto crítico se conoce como punto crítico de enlace (PCE). La densidad en el PCE está relacionada con la estabilidad del enlace núcleos. Este PCE es un punto de silla en la densidad electrónica, ya que es un máximo en la dirección de la SIA y un mínimo en la dirección de los núcleos. Las LIA son conocidas como líneas de interacción atómica y se encuentran en cada par de núcleos que comparten una SIA (**Figura 5**). Entonces un enlace se podría definir de la siguiente manera: “Dos átomos están enlazados cuando comparten una SIA, exista un PCE y una LIA entre ellos”¹⁵. Las LIA entre los átomos reproduce la conectividad encontrada experimentalmente.

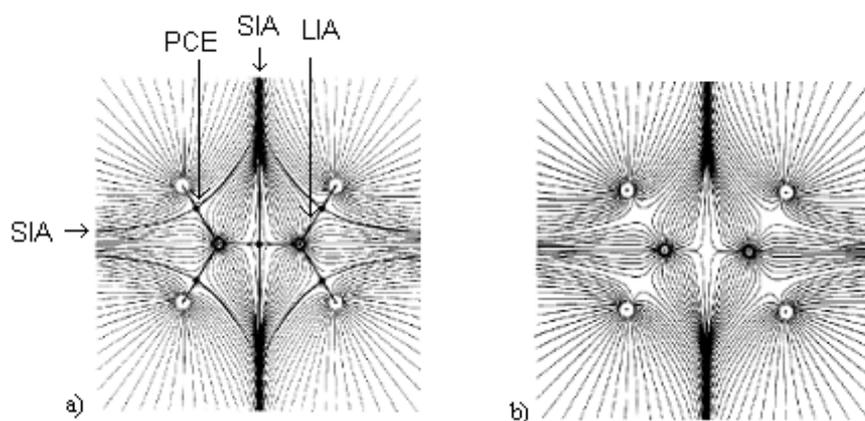


Figura. 5 a) Se ilustran las superficies interatómicas (SIA), los puntos críticos de enlace (PCE) y las líneas interatómicas (LIA) en una molécula del eteno, b) La misma molécula pero ilustrando las trayectorias del campo vectorial de $\nabla\rho$, los núcleos están indicados por los círculos en blanco y las líneas negras son las líneas de $\nabla\rho$.

2.10 ESTUDIO QTMS²⁷

En uno de sus trabajos, Popelier²⁸ demostró que se puede hacer una correlación entre los valores sigma (σ) de una serie de derivados del ácido benzoico, con sus similitudes. La similitud propuesta resulta de un cálculo que toma en cuenta parámetros descriptores de la densidad electrónica. Esta similitud es una distancia matemática, $d(A, B)$ **Ec. 25**, entre dos moléculas A y B, la cual se define como la suma de las distancias de los parámetros de los puntos críticos de dichas moléculas. Entre más bajo sea el valor $d(A, B)$ más similares serán estas dos moléculas. Obteniendo una mejor correlación al tomar como molécula referencia (A) a aquel que presente la menor actividad. Popelier tomó tres propiedades de los puntos críticos (ρ , $\nabla^2\rho$ y ε).

$$d(A, B) = \sum_{i \in A} \sum_{j \in B} dij \quad \text{Ec. 11}$$

En esta tesis se propuso realizar dicho cálculo tomando en cuenta cinco propiedades (ρ , $\nabla^2\rho$, ε , G y V).

Donde:

- $\nabla^2\rho$, es el Laplaciano
- ρ es la densidad electrónica
- ε es la electicidad

- G es la densidad de energía cinética
- V es la densidad de energía potencial

Además, de las propiedades locales de los puntos críticos también se estudiaron tres propiedades integradas o atómicas que son: q , K , N , LI y Vol .

2.10.1 PROPIEDADES DE ENLACE²⁹

- $\nabla^2\rho$, es el Laplaciano, el cual define la acumulación de densidad electrónica. Este parámetro nos indica el carácter del enlace analizado es decir nos ayuda a distinguir dos tipos de enlace ya que es una medida de que tan concentrada está la densidad electrónica en el punto crítico de enlace (PCE): si $\nabla^2\rho > 0$ se tratará de un enlace de interacción compartida (i.e. enlace covalente), mientras que $\nabla^2\rho < 0$ describe a los enlaces de capa cerrada (i.e. enlace iónico, puente de hidrógeno). El Laplaciano se expresa mediante la **Ec. 11**.

$$\nabla^2\rho = \nabla \cdot \nabla\rho(r) = \frac{\partial^2\rho(r)}{\partial x^2} + \frac{\partial^2\rho(r)}{\partial y^2} + \frac{\partial^2\rho(r)}{\partial z^2} \quad \text{Ec. 12}$$

$$\text{Donde: } \frac{\partial^2\rho(r)}{\partial x^2} = \lambda_1; \quad \frac{\partial^2\rho(r)}{\partial y^2} = \lambda_2; \quad \frac{\partial^2\rho(r)}{\partial z^2} = \lambda_3$$

El Laplaciano de un PCE es la suma de las tres curvaturas de la densidad, λ_1 , λ_2 y λ_3 (**Ec. 12**). Las dos curvaturas perpendiculares al enlace, λ_1 y λ_2 , son negativas y por convención $|\lambda_1| > |\lambda_2|$, mientras que la tercera curvatura, λ_3 , se encuentra a lo largo del enlace y es positiva. En un enlace covalente las dos curvaturas negativas son dominantes y $\nabla^2\rho < 0$. En contraste una interacción de capa cerrada se caracteriza por una disminución de densidad en la región de contacto de los dos átomos y $\nabla^2\rho > 0$.

- La densidad de carga o densidad electrónica denotada por r es la probabilidad de encontrar a un electrón en un cierta posición mientras que la posición de los demás electrones está promediada, la densidad es máxima en las posiciones nucleares y presenta puntos de silla entre dos núcleos enlazados. La densidad electrónica está

relacionada con el orden de enlace y con la energía. La fuerza de un enlace u orden de enlace (OE), es reflejada en la densidad electrónica en un PCE (**Ec 13**).

$$OE = \exp[A(\rho_b - B)] \quad \text{Ec. 13}$$

Donde A y B son constantes que dependen de la naturaleza de los átomos enlazados. La densidad electrónica ha demostrado una fuerte correlación con la energía de enlace para una gran cantidad de interacciones de enlace y con la distancia de enlaces S-S.

- La elepticidad (ε) es una medida de la asimetría de la densidad electrónica en el plano ortogonal al enlace, está relacionada a la susceptibilidad de un anillo a romperse y es una medida del carácter π de un enlace (**Ec. 14**).

$$\varepsilon = \frac{\lambda_1}{\lambda_2} - 1 \quad \text{Ec. 14}$$

Donde $|\lambda_1| \geq |\lambda_2|$

La elepticidad mide el grado en el que densidad se acumula preferencialmente en un plano dado que contiene la trayectoria de enlace. Si $\lambda_1 = \lambda_2$, entonces $\varepsilon = 0$, y el enlace es cilíndricamente simétrico, tal sería el caso del enlace sencillo en el etano o el triple enlace en el acetileno. De esta manera ε es una medida del carácter π de un enlace en los límites del doble enlace, en el cual la elepticidad alcanza un máximo

- G y V denotan a las densidades de la energía cinética y potencial respectivamente. Dichas energías, junto con la total, son usadas para resumir los mecanismos de una interacción de enlace.

2.10.2. PROPIEDADES ATÓMICAS ²⁹

El promedio de una propiedad O sobre una base atómica Ω , $O(\Omega)$, es calculado por:

$$O(\Omega) = \left\langle \hat{O} \right\rangle_{\Omega} = \frac{N}{2} \int_{\Omega} dr \int d\tau \left[\Psi^* \hat{O} \Psi + \left(\hat{O} \Psi \right)^* \Psi \right] \quad \text{Ec. 15}$$

Donde \hat{O} es un operador de electrón o la suma de operadores de electrón.

- La población atómica de electrones (N), como su nombre lo indica representa la población total de electrones en un átomo. Dicho valor es obtenido estableciendo que $\hat{O} = \hat{1}$ en la **Ec. 14**. Resultando en la **Ec. 15**.

$$N(\Omega) = \int_{\Omega} \rho(r) dr \quad \text{Ec. 16}$$

- La población electrónica de un átomo puede estar localizada en el átomo o deslocalizada en la molécula $N = \lambda + \delta$, λ es la medida de localización en el átomo, mientras que δ es la cuantificación de la deslocalización.
- La carga atómica es obtenida restando $N(\Omega)$ de la carga nuclear $Z\Omega$. **Ec. 17**.

$$q(\Omega) = Z\Omega - N(\Omega) \quad \text{Ec. 17}$$

- El volumen atómico (Vol), dentro de la teoría de átomos en moléculas, es aquel espacio que abarca una isosuperficie que toma en cuenta el volumen de van der Waals en la fase gaseosa, y usualmente encierra mas del 99% de la población de electrones de la molécula.
- Energía atómica total (E) es la energía asociada a la población electrónica y está descrita por la **Ec. 18**:

$$E(\Omega) = E_e(\Omega) = T(\Omega) + V(\Omega) = -T(\Omega) = \frac{1}{2}V(\Omega) \quad \text{Ec. 18}$$

Donde T es la energía cinética atómica y V es la energía potencial de un átomo.

2.11. ANÁLISIS DE DATOS MULTIVARIABLE ^{30,31}

El análisis multivariable comprende un conjunto de técnicas que permiten el tratamiento simultáneo de numerosas observaciones y variables, por lo que es necesaria una correcta organización de los datos de manera que se facilite encontrar interrelaciones y tendencias para generar conclusiones que expliquen un fenómeno.

Se trabaja con un conjunto de datos constituido por N observaciones y K variables. Las observaciones son las moléculas a analizar mientras que las variables son los descriptores de

cada molécula. Para encontrar correlaciones se utilizan dos métodos estadísticos: análisis de componentes principales (PCA) y mínimos cuadrados parciales (PLS).

2.11.1 PCA³¹

2.11.1.1 ESPACIO MULTIDIMENSIONAL K³¹

Supongamos una matriz X con N observaciones y K variables. Construimos para esta matriz, un espacio con tantas dimensiones como variables. Cada variable se representa con un eje coordenado. En la **Figura 6** se muestra un conjunto de datos con 3 variables. El largo de los ejes se ajusta de acuerdo a la varianza de los datos.

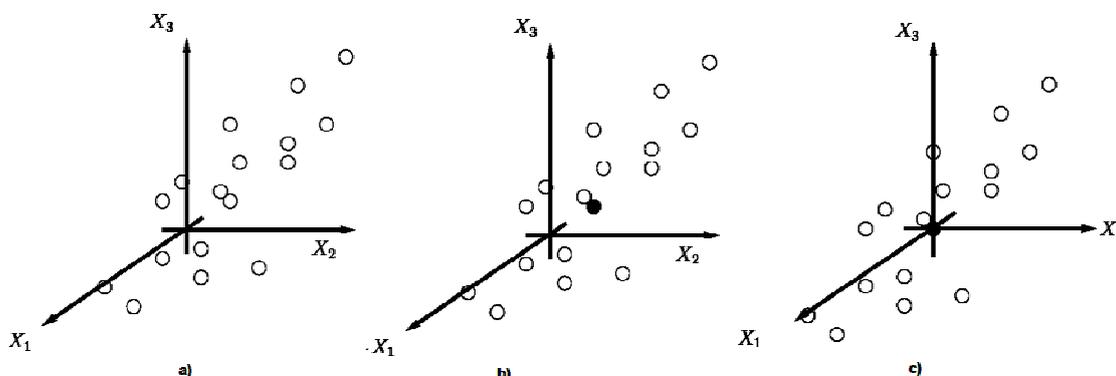


Figura 6: a) Un espacio K-dimensional, sólo se muestra 3 ejes que representan 3 variables. El largo de cada uno de los ejes representa la escala de la variación. Las observaciones (renglones) en la matriz de datos X se presentan como un “enjambre” de puntos. b) En la calibración del espacio multidimensional K, primero se extraen los promedios de las variables. Esto resulta en un punto que representa el promedio de todos los valores en el conjunto de datos (punto gris) y está situado en el centro del “enjambre” de datos. c) La calibración del espacio multidimensional K resulta en un reposicionamiento del origen de los ejes coordenados de manera que el promedio de los datos coincida en el centro de estos ejes (punto gris).^{30,31}

En el siguiente paso, cada observación (cada renglón en nuestro conjunto de datos) de la matriz X se grafica en el espacio multidimensional K. Los renglones de la tabla de los datos forman un "enjambre" de puntos en el espacio (**Figura 6a**).

Es necesario en este punto realizar un ajuste a los datos para asegurar que los rangos de los valores sean homogéneos. Tomando de partida el promedio de todos los valores de la matriz de datos X, se mueve el espacio multidimensional K para que los ejes coordenados coincidan en el centro. Esto se hace restando los promedios de las variables de los datos. Este vector de

promedios corresponde a un punto en el espacio K (**Figura 6b**). Restando ahora los promedios de los datos originales resulta en un reposicionamiento de los ejes coordenados del espacio multidimensional K. Ahora el punto de origen de los ejes coordenados coincide con el promedio del “enjambre” de datos (**Figura 6c**).

2.11.1.2 EL PRIMER COMPONENTE PRINCIPAL

Una vez ajustado el centro de acuerdo a la media de los datos, se calcula el primer componente principal (PC1). Este componente es la línea en el espacio multidimensional K que mejor se ajusta a los datos. La línea cruza los promedios del "enjambre" de datos (**Figura 7a**). Se puede generar una distancia de cada observación hacia la línea del componente principal. Esta distancia se conoce como “score” (t_1).

2.11.1.3 EXTENDIENDO EL MODELO

Generalmente, un componente principal no es suficiente para modelar variaciones sistemáticas dentro del conjunto de datos. Es necesario calcular el segundo componente principal, PC2. Este segundo componente principal también es representado por una línea dentro del espacio multidimensional K, y que es ortogonal al primero (**Figura 7b**). Esta línea cruza también los promedios de las observaciones y ayuda a mejorar la aproximación de los datos de la matriz X lo más posible.

Dos componentes principales definen un plano dentro del espacio multidimensional K (**Figura 7c**). Proyectando todas las observaciones hacia este nuevo sub-espacio, es posible visualizar la estructura del conjunto de datos que se está analizando. Los valores de las coordenadas de las observaciones hacia este plano son los “scores”.

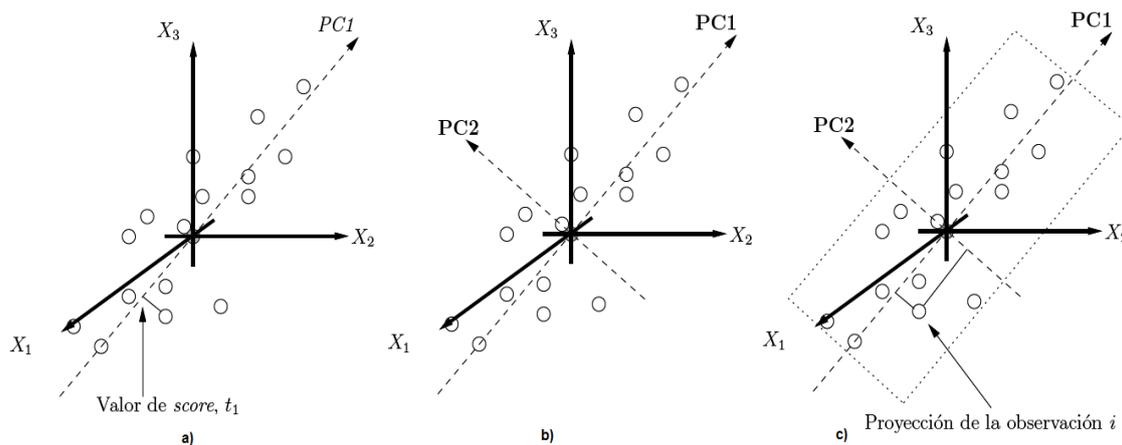


Figura 7: a) El primer componente principal, PC1, está representado por una línea que mejor se ajusta en el “enjambre” de datos. Cada observación puede ser proyectada hacia esta línea para obtener una distancia, conocida como score. b) El segundo componente principal, PC2, se genera de tal manera que refleje la segunda fuente de variación principal dentro de la matriz de datos X. Esta línea es ortogonal al primer componente principal y también cruza el conjunto de datos por el promedio. c) Dos componentes principales forman un plano. Cada observación se puede proyectar hacia este plan resultando en una nueva distancia “score”.

Muchas veces son necesarios más de dos componentes principales para generar un modelo que explique correctamente la variación dentro del conjunto de datos. Estos componentes principales se crean de igual manera, generando planos en el espacio multidimensional K, ortogonales con respecto al componente principal anterior. Para cada una de las observaciones se genera una sub-coordenada con cada uno de los componentes principales. Estos valores explican diferentes porcentajes de variación del modelo y ayudan a descubrir las observaciones o valores dentro de la matriz de datos X que influyen positiva o negativamente en la variación.

El ajuste nos dice qué tan bien podemos reproducir matemáticamente los datos del conjunto. Una medida cuantitativa del ajuste esta dada por el parámetro R^2X . Más importante que el ajuste del modelo, es el poder predictivo del mismo. Esto puede ser estimado con qué tanto podemos predecir los datos en X, ya sea internamente con los datos existentes o con datos externos como si fuera una validación de las observaciones. El poder predictivo del modelo esta resumido con el poder de predicción Q^2X . En este caso, usamos validación cruzada para estimar la habilidad de predicción del modelo, la validación cruzada se explicó en el tema anterior de QSAR.

Los parámetros R^2X y Q^2X muestran comportamientos diferentes del modelo en la medida que van aumentando su valor. La correlación del modelo, R^2X , varía entre 0 y 1.0, donde 1.0 significa un modelo perfectamente ajustado, y el 0 significa ningún tipo de correlación. R^2 es inflacionario y se va acercando a la unidad al mismo tiempo que la complejidad del modelo (número de parámetros, número de componentes...) aumenta. Por otro lado, el índice de predicción, Q^2X , no se acercará a 1.0 mientras más complejo sea el modelo.

Las proyecciones se pueden adaptar a diferentes objetivos de análisis de datos, por ejemplo: a) resumen y visualización de un conjunto de datos, b) clasificación y discriminación de datos, y c) buscar relaciones cuantitativas entre las variables.

Resumen de los datos: Al inicio de un proyecto, cuando no se conoce mucho del problema, uno requiere un resumen simple de la información que se está analizando. Este resumen se puede obtener con la ayuda del análisis de componentes principales, PCA. PCA produce un resumen mostrando como las observaciones se relacionan y si hay relaciones o grupos de observaciones en los datos. De especial interés en el procesamiento de los datos es la habilidad del análisis de componentes principales, PCA, de descubrir tendencias suaves o cambios bruscos entre las variables. PCA describe la correlación entre la estructura en X.

Clasificación y/o discriminación: Es frecuente que el análisis inicial con PCA revele grupos dentro de las observaciones. La presencia de grupos indica la necesidad de un análisis PCA adicional de cada uno de los grupos para ajustar finamente el análisis y entender las características de cada grupo.

2.11.2 PLS

En su forma más simple, el análisis de mínimos cuadrados parciales (projections to latent structures by means of partial least squares), PLS, es un método para relacionar dos matrices de datos, X y Y, por medio de un modelo multivariable lineal (**Figura 8**). Utilizando parámetros que estén relacionados con las observaciones (moléculas, muestras, compuestos, etc), la precisión del método PLS aumenta mientras más número de variables X se tengan. Se puede

considerar la método PLS como una regresión particular para modelar relaciones entre X y Y. En QSAR, PLS se usa para modelar la relación entre, por un lado, variables que caracterizan la variación estructural de un conjunto de N compuestos, por el otro lado, las respuestas biológicas o experimentales de las mismas N sustancias. Lo primero forma la matriz X de variables de predicción y lo segundo forma la matriz Y con las variables de respuesta. Estas matrices tienen dimensiones (N x K y N x M) respectivamente.

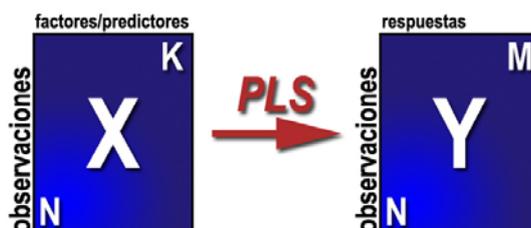


Figura 8: Un conjunto de datos con N observaciones y K variables forman la matriz X y el mismo número de observaciones N con M variables de respuesta forman la matriz Y. Las observaciones son las moléculas a analizar en el análisis PCA. Las variables son los descriptores electrónicos obtenidos de la densidad electrónica para capturar las propiedades de las observaciones y se utiliza el análisis PLS para relacionar X con Y. PLS permite predecir X a partir de Y.

2.11.2.1 ESPACIO MULTIDIMENSIONAL K CON UNA RESPUESTA M=1

Consideremos una aplicación con N observaciones, K=3 variables X, y M=1 variable Y. Así como en PCA, cada observación puede ser representada gráficamente (**Figura 9**).

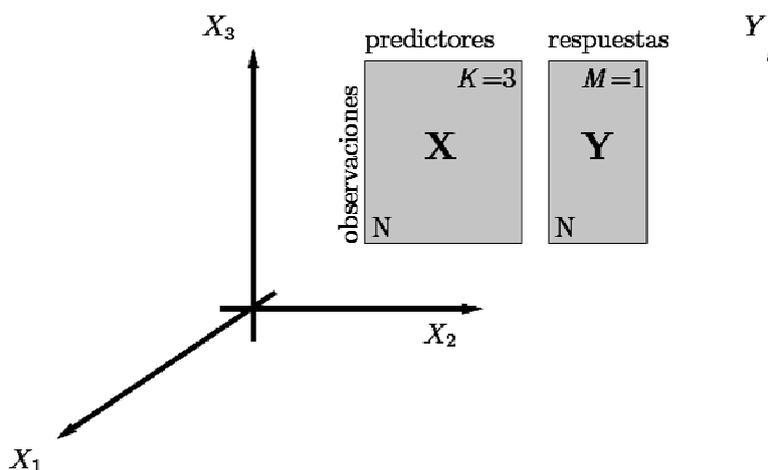


Figura 9: Un conjunto de datos con K=3 variables X y M=1 variable Y. La longitud de cada eje coordenado ha sido ajustada de acuerdo a la media de los datos.

Sin embargo, la gran diferencia en PLS es que cada renglón de datos en la tabla corresponde a dos puntos en lugar de uno. Uno en el espacio X y otro en el espacio Y. La **(Figura 10)** muestra cómo se podrá ver el espacio multidimensional cuando $K=3$ y $M=1$. La tarea del análisis de datos es describir la relación entre las posiciones de las observaciones en el espacio predictor (X) y sus posiciones en el espacio de las variables de respuesta (Y).

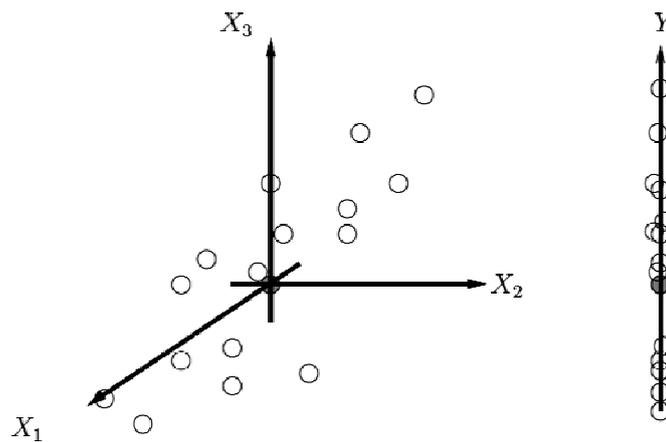


Figura 10: En un problema de regresión PLS, las observaciones se distribuyen en dos regiones del espacio multidimensional. Uno con los predictores (el espacio X) y otro con las variables de respuesta (el espacio Y). Los datos se han ajustado a la media de todo el conjunto de datos y los sistemas de coordenadas de ambos espacios se han ajustado a este valor (punto oscuro)

2.11.2.2 EL PRIMER COMPONENTE PLS ($M=1$)³⁰

Después de ajustar los centros de los ejes coordenados, se calcula el primer componente PLS. Este componente es una línea en el espacio X que cruza todas las observaciones de la mejor manera posible, y que tiene una buena correlación con el vector en el espacio Y. La coordenada de alguna observación se obtiene proyectando la muestra en esta línea. Esta coordenada se conoce como “score”, t_i , de la observación i . Los “scores” de todas las observaciones forman el primer vector de “scores” en X, t_1 (**Figura 11**).

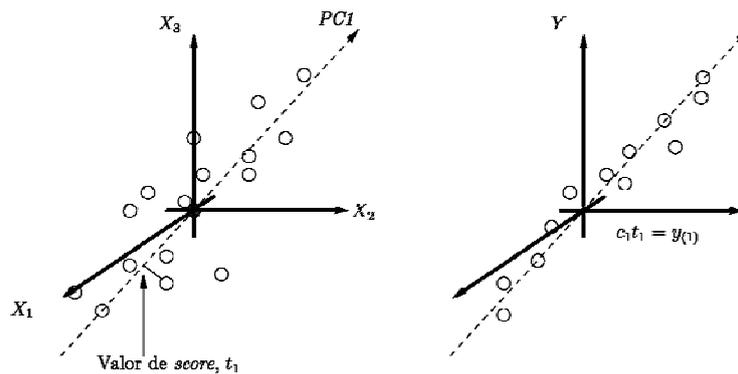


Figura 11: Con una sola variable en Y , el espacio se reduce a un solo vector de una dimensión. El primer componente en el modelo PLS entonces se orienta de manera que describa lo mejor posible los puntos en el espacio X mientras que de una buena correlación con el vector de Y . Las proyecciones de las observaciones hacia la línea del espacio en X da los valores “score” de cada observación.³⁰

2.11.2.3 EXTENDIENDO EL MODELO³⁰

Usualmente, no es suficiente con un solo componente PLS para adecuar correctamente un modelo a la variación de los datos en Y . El poder de la herramienta PLS se incrementa expandiéndolo con un segundo componente. El segundo componente también es una línea en el espacio X , que cruza por el origen y es ortogonal al primer componente (**Figura 12**). Este componente encuentra la dirección en el espacio X y mejora la descripción de los datos lo mejor posible, mientras que provee de una buena correlación con los residuales en Y que quedaron del componente anterior.³⁰

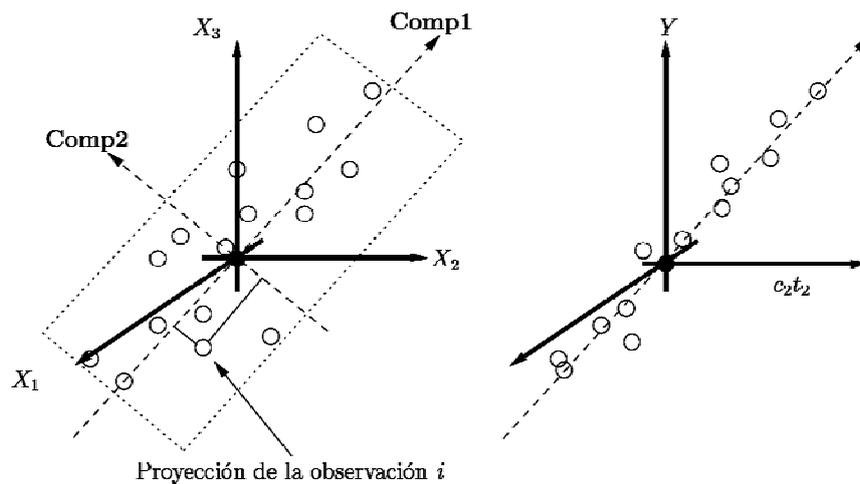


Figura 12: La segunda proyección en el espacio X es ortogonal a la primera. Si se proyectan las observaciones en este plano, se tiene el vector “score” t_1 . Los dos componentes forman un plano en el espacio X .³⁰

Creación de un modelo: El objetivo de crear un modelo de datos complejos utilizando PLS es tener predicciones certeras, rápidas y cuantitativas de respuestas complejas (por ejemplo, calidad de un producto, sabor de un vino, impurezas en una muestra, etc.) basadas en una colección de datos X. Con datos apropiados y un modelo PLS funcional es posible descubrir cómo los factores influyen en las respuestas, cómo las respuestas correlacionan unas con otras y cómo se tienen que ajustar los factores para obtener una respuesta en específico.

2.11.2.4 INTERPRETANDO UN MODELO PLS³⁰

PLS consta de muchos diagnósticos que ayudan a la interpretación de los datos que se están analizando. Entre las herramientas que más se utilizan son los pesos, los coeficientes, los valores VIP (Variable Influence on Projection) así como el valor Q^2 .

Pesos Los pesos en el análisis PLS dan información de cómo las variables X se combinan para formar los índices de “scores”. Estos índices son la base de la relación cuantitativa entre X y Y.

Diagrama de influencia El diagrama de influencia o pesos muestra la relación entre variables. Indican qué tanto influyen las variables X sobre Y. Se podría decir que el diagrama de influencias y los pesos son complementarias y se pueden superponer o encimar, lo que significa que un patrón identificado en la gráfica de los pesos se puede interpretar también observando en la misma dirección en el diagrama de influencias.

Coefficientes de regresión Los coeficientes son de especial interés pues simplifican la interpretación del modelo cuando hay varios componentes. La ventaja es que se obtiene un solo vector con la información del modelo por cada respuesta, a diferencia de muchos vectores utilizando los pesos. La desventaja de los coeficientes es que la información que corresponde a la correlación de la estructura de las respuestas se pierde. Esa información está guardada por los valores de los pesos.

VIP's Interpretar un modelo PLS con muchos componentes y una multitud de respuestas puede ser un trabajo complicado. Un parámetro que hace un resumen de la importancia de las variables X, para ambos modelos X y Y, se llama VIP (Variable Influence on Projection). Los valores VIP son la suma de los cuadrados de los pesos PLS, tomando en cuenta la cantidad de variación explicada en cada una de las dimensiones de Y. Para cada modelo y problema sólo existirá un solo vector VIP, explicando todos los componentes y todas las variables Y. Los

valores VIP mayores a uno se consideran que tienen mucha influencia en la creación del modelo.

Q² Este valor nos dice la fracción de la variación total de las variables Y que se pueden predecir por un componente después de aplicar validación cruzada al modelo.

3. PLANTEAMIENTO DEL PROBLEMA

¿Es posible encontrar el sitio farmacóforo utilizando descriptores locales y/o integrados de la densidad electrónica de una molécula?

4. HIPÓTESIS

- Existe una relación entre la densidad electrónica de las moléculas y la actividad biológica que presentan.

5. OBJETIVO

- Encontrar la relación entre la estructura electrónica de ésteres antranílicos con su actividad farmacológica.

5.1. METAS

1. Optimizar la estructura de los diferentes derivados de ésteres antranílico por métodos computacionales, es decir obtener las estructuras de menor energía.
2. Obtener descriptores locales e integrados de cada molécula, para establecer una relación de estos descriptores con la actividad biológica de cada compuesto.
3. Identificar los sitios activos o los posibles lugares que influyan directamente en la respuesta biológica.
4. Buscar el farmacóforo en la familia de los compuestos estudiados.

6. METODOLOGÍA

Se contó con 10 compuestos que fueron seleccionados por tener reportada actividad biológica significativa¹ y un esqueleto en común que se muestra en la **Figura 13**. Los sustituyentes se enlistan en la **Tabla 3** cuya numeración se respetó del artículo¹ del cual se obtuvieron, ya que son los únicos que reportaron con actividad. En la **Tabla 4** se concentran las actividades biológicas para cada compuesto y el tejido en el cual se probó. Son valores de menos el logaritmo de la concentración molar que causa el 50% de inhibición del crecimiento neto celular (pGI₅₀).¹

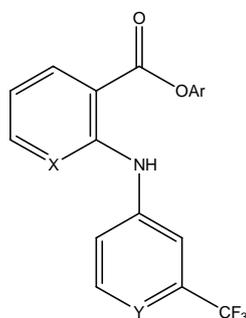


Figura 13. Estructura esqueleto del ácido flufenámico.¹

Tabla 3. Sustituyentes del esqueleto del ácido flufenámico.¹

Compuesto	X	Y	OAr
10	C	C	2-clorofenil
17	C	N	3-metoxifenil
20	C	N	2-clorofenil
21	C	N	3-clorofenil
22	C	N	4-clorofenil
23	C	N	2,4-diclorofenil
24	C	N	2,4,6-triclorofenil
25	C	N	piridin-3-il
27	N	N	2-clorofenil
31	N	N	piridin-3-il

Tabla 4. Datos obtenidos del National Cancer Institute (NIC) In Vitro orientados a la enfermedad humana tumor de células.¹

Molécula	Leucemia ^a	Pulmón ^b	Cólon ^c	SNC ^d	Melanoma ^e	Renal ^f	Mama ^g
10	5.57	4.7	5.09	4.69	4.77	4.63	4.47
17	6.13	5.69	5.39	5.33	5.51	5.27	5.14
20	7.29	6.59	6.44	5.65	6.51	6.48	6.55
21	5.49	5.62	5.36	5.52	5.69	5.53	5.32
22	5.44	5.55	5.18	5.36	5.58	5.53	5.43
23	5.52	5.45	5.34	5.51	5.64	5.52	5.44
24	5.62	5.59	5.34	5.41	5.76	5.51	5.33
25	7.37	7.44	7.39	6.48	6.54	7.15	7.34
27	5.34	5.06	5.43	5.00	5.21	5.39	5.44
31	5.29	5.2	5.26	5.25	5.24	5.19	5.57

Las actividades de cada uno de los compuestos, pGI_{50} , es $-\log$ de la concentración molar que causa el 50% de inhibición del crecimiento neto celular. El tipo de tejido celular usado In Vitro para la prueba de actividad; a) K-562 b) NCIH23 c) SW-620 d) SF-268 e) UACC-62 f) SN12C g) MCF-7.

6.1. Equipo y programas de Cómputo

Este trabajo se realizó en un servidor con procesador opteron y sistema operativo Linux. Las optimizaciones se realizaron con el programa Gaussian 03.²⁴ El análisis de la densidad electrónica se realizó con el paquete AIMPACK³². Y finalmente el análisis estadístico se realizó utilizando el paquete SIMCA-P.^{30, 31}

6.2. OPTIMIZACIÓN²⁴

Se dibujó la estructura de cada uno de los 10 compuestos con el programa Gauss View, así se encontró la conformación de menor energía de cada una, con el nivel de teoría RB3LYP/6-31G (d,p); este fue elegido por que describe bien la densidad electrónica a un bajo costo computacional.

6.3. ANÁLISIS DE LA FUNCIÓN DE ONDA³²

Una vez obtenida la función de onda se calcularon las propiedades de la densidad electrónica, tanto locales como integradas (atómicas). En esta etapa, se extraen todos los parámetros que van a formar la descripción cuántica de cada molécula y que van a intervenir en el estudio QSAR.

El trabajo se dividió en dos partes, la primera relacionada con las propiedades integradas de la densidad electrónica y la segunda con las propiedades locales como se ve en los resultados. Para tener más control de la cantidad de datos a manejar, los átomos de cada uno de los compuestos se nombraron de acuerdo a la **Figura 14**

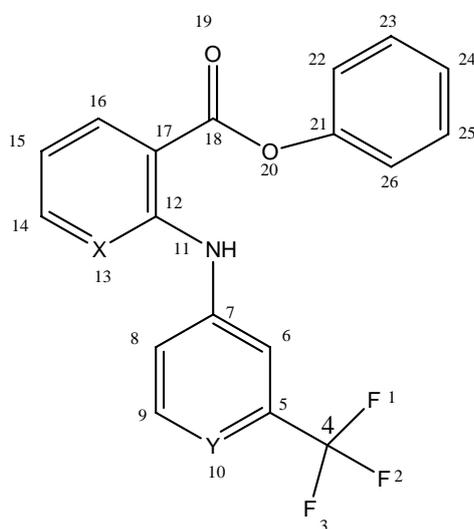


Figura 14. El patrón de numeración que se otorga a los átomos de cada compuesto.

En la **Figura 15** se muestra el esqueleto con las etiquetas para identificar cada uno de los enlaces.

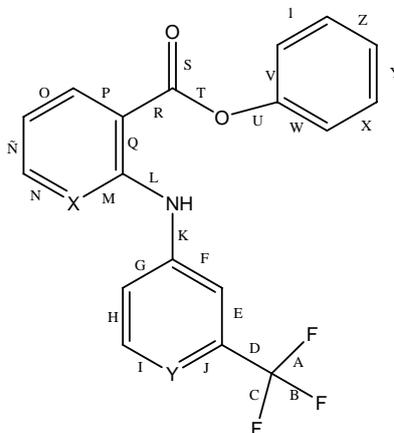


Figura 15. Etiquetas que se otorgaron a cada enlace de cada compuesto.

6.4. ANÁLISIS ESTADÍSTICO^{30,31}

El siguiente paso es realizar el análisis estadístico de los descriptores locales e integrales obtenidos, y la propiedad que se va a utilizar como variable de respuesta. El análisis estadístico permite correlacionar observaciones, factores de respuesta y descriptores; usando mínimos cuadrados parciales (Partial Least Squares, PLS), así como el análisis de componentes principales (PCA).³⁰

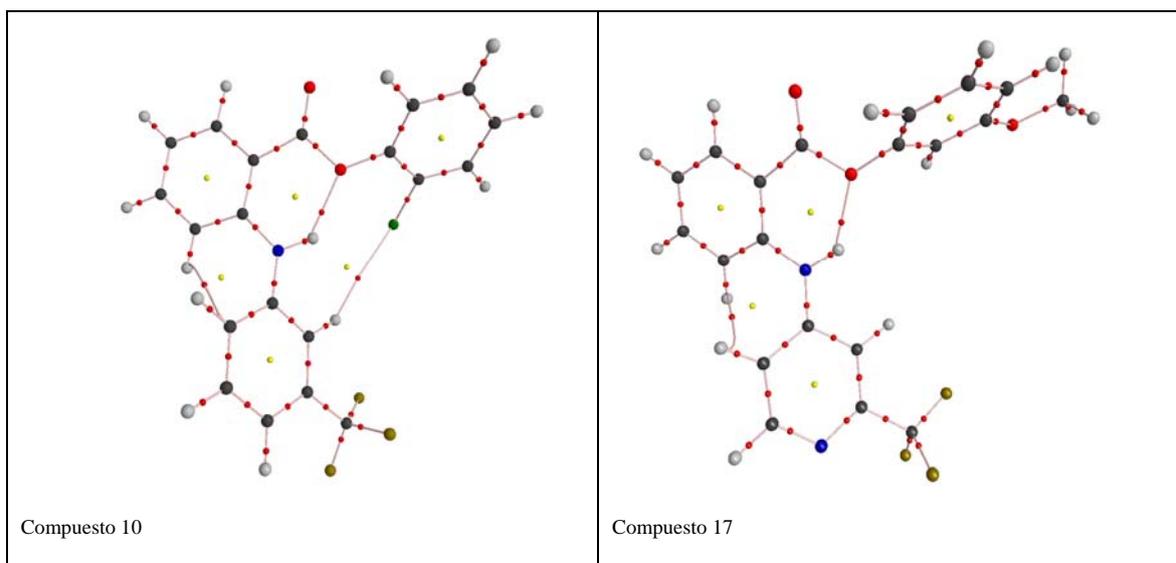
Al inicio del análisis se sabe muy poco o nada de si una familia de moléculas se relacionara entre sí, es por esto que se requiere un resumen inicial de la información del conjunto de datos. Este resumen lo proporciona el análisis de componentes principales, PCA, y muestra cómo las observaciones están relacionadas y si hay alguna observación o grupos de observaciones en los datos que difieran mucho. PCA también descubre tendencias suaves o cambios bruscos en la información analizada. PCA describe la correlación dentro de la estructura de los datos (X). Es común que un análisis inicial con PCA muestre agrupaciones dentro de las observaciones. Con lo que es necesario hacer un segundo análisis por grupo, para así obtener mejores detalles del análisis.

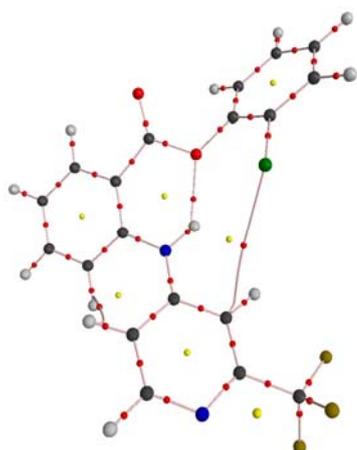
El análisis de mínimos cuadrados relaciona dos o más grupos de datos y extrae la relación de la variación entre los elementos. El paso final para el análisis de datos es un modelado de regresión entre dos bloques de datos, usualmente llamados X y Y, con el objetivo de predecir Y a partir de X. Este tipo de modelo se denomina PLS y se puede considerar como una extensión del análisis PCA por que este segundo análisis se le aplica a las familias o grupos encontrados en el PCA. Las variables X son los factores o predictores, y las variables Y son las respuestas. Además, es posible con este análisis encontrar cómo los factores influyen en las respuestas, cómo las respuestas se correlacionan unas con otras y cómo los factores se ajustan para obtener un modelo que pueda predecir bien las respuestas.

7. RESULTADOS

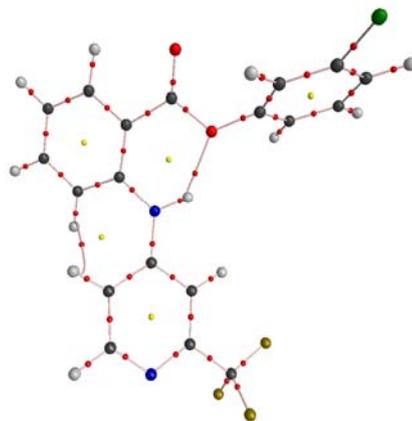
Para cada átomo se obtuvieron 5 propiedades locales (ρ , $\nabla^2\rho$, ε , G , V) y 5 propiedades integradas o atómicas (q , K , Vol , LI , N). Esto después de optimizadas los compuestos derivados del ácido antranílico y obtenidas las graficas moleculares. En la **Tabla 5** se presentan las gráficas moleculares de los puntos críticos. Para el estudio de las propiedades integradas, el conjunto de datos fue de 10 observaciones ($N=10$), es decir, los 10 derivados del ácido antranílico; 26 átomos por 5 descriptores electrónicos que producen 130 variables ($X=130$) (ver anexo i). Para el caso de las propiedades locales se tiene un conjunto de 10 observaciones ($N=10$), 28 enlaces por 5 descriptores electrónicos que producen 140 variables ($X=140$) (ver anexo ii). Las variables de respuestas están representada por 7 actividades biológicas por cada molécula ($Y=7$). Las actividades biológicas se muestran en la **Tabla 4**.

Tabla 5. Se muestra cada una de las graficas moleculares con los puntos críticos de cada enlace

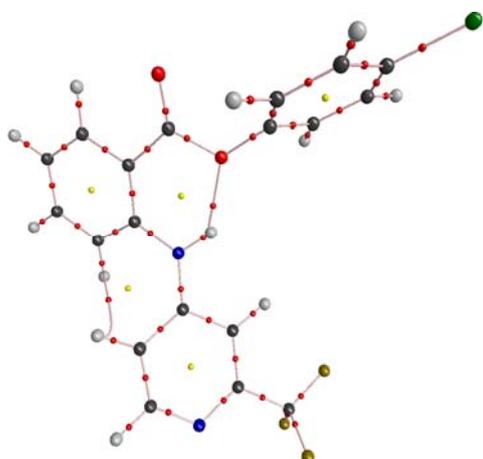




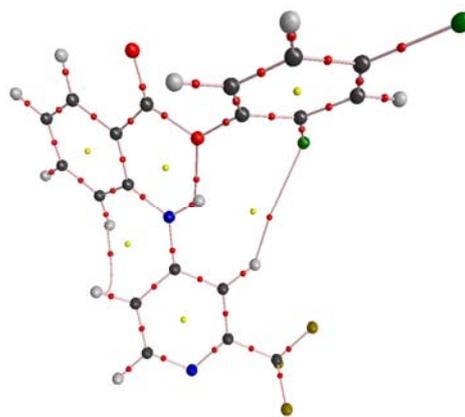
Compuesto 20



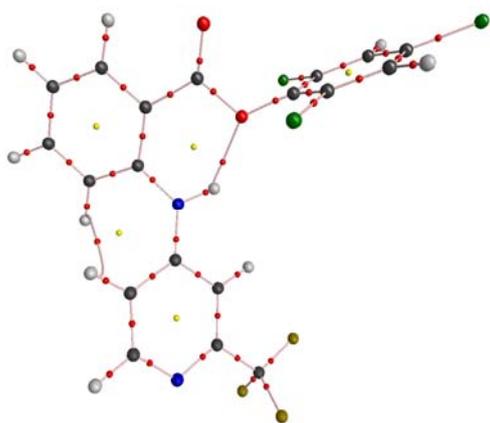
Compuesto 21



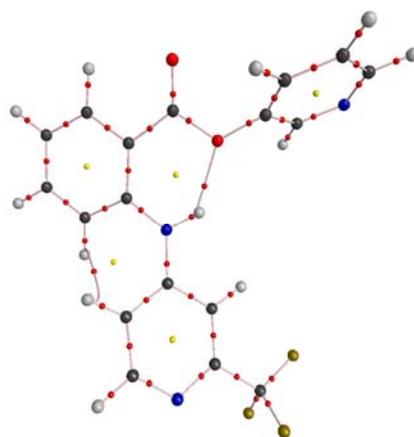
Compuesto 22



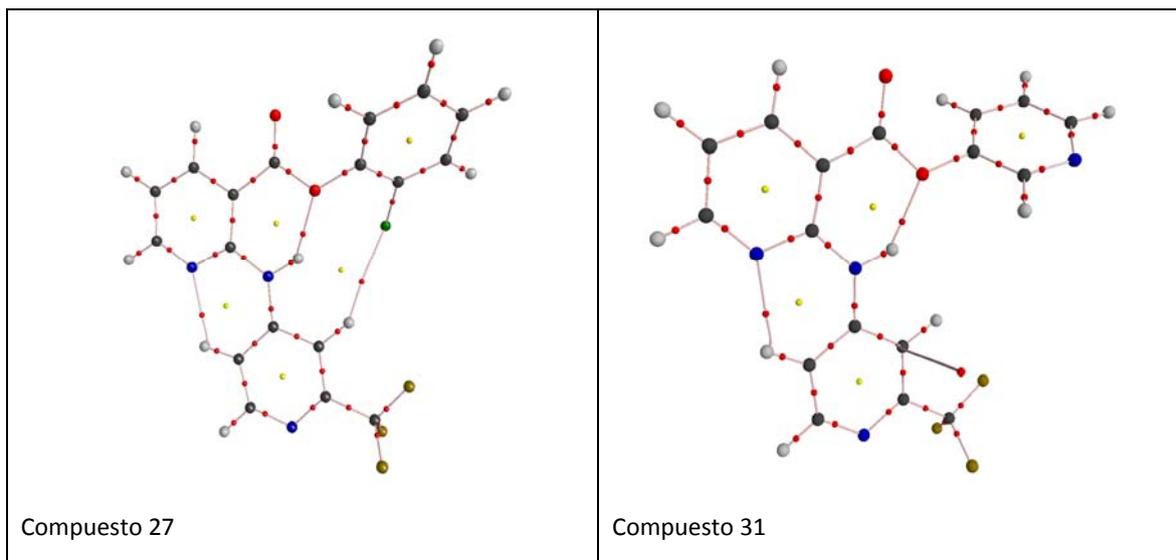
Compuesto 23



Compuesto 24



Compuesto 25



Código de colores: carbonos esferas negras, hidrógenos esferas grises, nitrógenos esferas azules, oxígenos esferas rojas, flúor esferas café, cloros esferas verdes, puntos críticos de enlace esferas rojas pequeñas y puntos críticos de anillo esferas amarillas pequeñas.

7.1 ESTUDIO DE LAS PROPIEDADES LOCALES

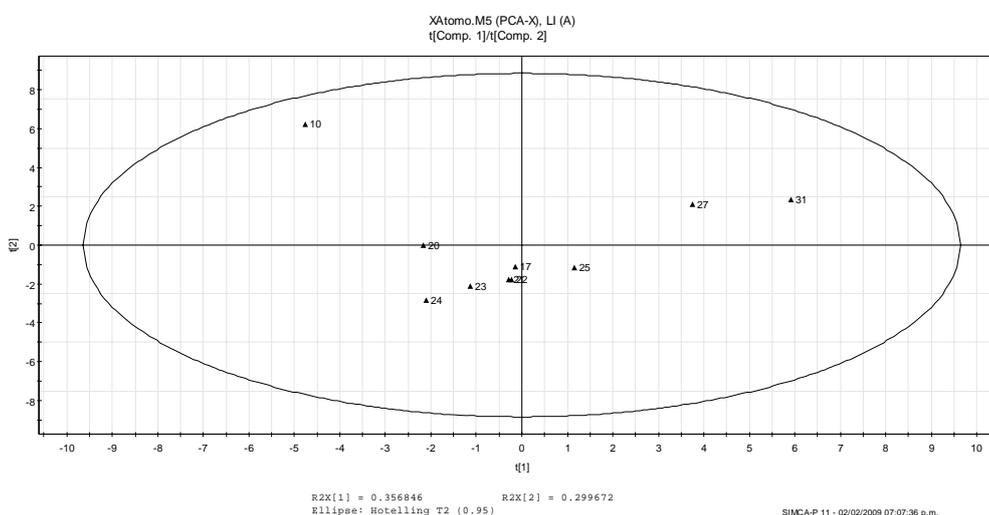
Se inició el estudio con un análisis de componentes en el conjunto completo de datos (130) para generar un modelo, que dio como resultado unas R^2_x y Q^2 muy bajas, por lo que se tomó la decisión de hacer el análisis de componentes por cada una de las variables para generar el modelo. La R^2_x y la Q^2 de estos modelos se muestran en la **Tabla 6**. El mejor ajuste y la mayor predicción se da en el caso de λ , que son de 0.903 y 0.552 respectivamente. Esto muestra que la localización electrónica es un buen descriptor. Por otro lado q y N muestran un ajuste moderado y una predicción nula, mientras que K y Vol presentan un menor ajuste pero mayor predicción en el modelo generado.

Tabla 6. Correlación y predicción del modelo generado para cada descriptor utilizando PCA

Propiedad	R^2_x	Q^2
q	0.805	0.0116
K	0.617	0.575
N	0.797	0.0341
LI	0.903	0.552
Vol	0.646	0.204

Desafortunadamente, los análisis PLS del conjunto completo de datos para cada descriptor no generaron ningún modelo por que los datos están demasiado dispersos. Para estos casos, hay que regresar al análisis PCA para encontrar grupos de datos con características similares y hacer el análisis PLS por grupos.

Al graficar los “scores” (t) de los componentes de cada uno de estos modelos PCA, generados para cada descriptor, se encontró que se podía dividir a las moléculas en tres grupos, según se muestra en la **Gráfica 1** para el descriptor λ . El lector puede consultar el anexo iii para ver las gráficas de “scores” para el resto de descriptores. Para todos los descriptores se observó el mismo patrón de agrupamiento. Por lo que la división se puede considerar general.



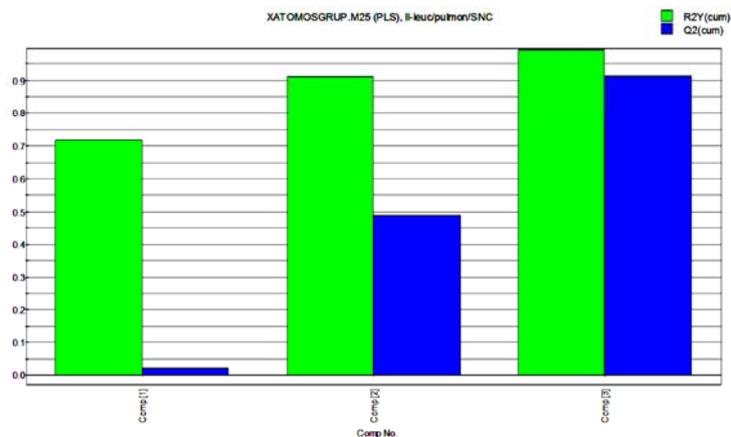
Gráfica 1: Gráfica de scores (t1/t2) del análisis PCA del modelo generado para el descriptor λ . Se muestran tres grupos de compuestos que influyen en la respuesta.

El compuesto 10 forma el primer grupo. Los compuestos 17 y del 20 al 25 forman el segundo grupo. El tercer grupo está constituido por los compuestos 27 y 31. Esta división va de acuerdo con las tres posibles estructuras químicas de los compuestos. El primer grupo (compuestos 10) es el 2-clorofenil éster derivado del ácido flufenámico. Los compuestos 17, y del 20 al 25 son la serie del antranilato de *N*-(trifluorometilpiridinio). Por otro lado, los ésteres nicotínicos del 2-clorofenilo y piridin-3-ilo, compuestos 27 y 31 respectivamente, forman el tercer grupo.

El análisis PLS se realizó solo para el grupo II ya que los grupos I y III no tienen suficientes datos. Además, se utilizaron todos los descriptores atómicos. El total de las variables de respuesta fue usado en una primera exploración y después se examinó cada variable de respuesta por separado. Los resultados se muestran en la **Tabla 7**. Los ajustes y la predicción obtenidos con el modelo PLS generado con todas las variables de respuesta es bueno al igual que en el caso de tumores en el sistema nervioso central (SNC). En el caso de leucemia y tumores en tejido de pulmón se tiene un ajuste regular con X, un buen ajuste con Y y una buena predicción. En el caso de melanoma y tumores en el colon, tejido renal y mama no se pudo generar un modelo PLS. Por lo que el modelo generado con todas las variables de respuesta que si genero modelo, se debe al ajuste con las respuestas para leucemia, tumores en pulmón y SNC. Con lo cual, la siguiente opción era hacer el análisis incluyendo solo las actividades biológicas para leucemia, tumores en pulmón y SNC. Este estudio mostró un modelo con 3 componentes principales (ver **Gráfica 2**) capaz de predecir el 82% ($R^2X = 0.828$) de la variación en X para describir el 99% ($R^2Y = 0.991$) de la variación de Y y predecir 91% ($Q^2 = 0.913$). Este resultado supera los ajustes y la predicción que mostró el total de las actividades.

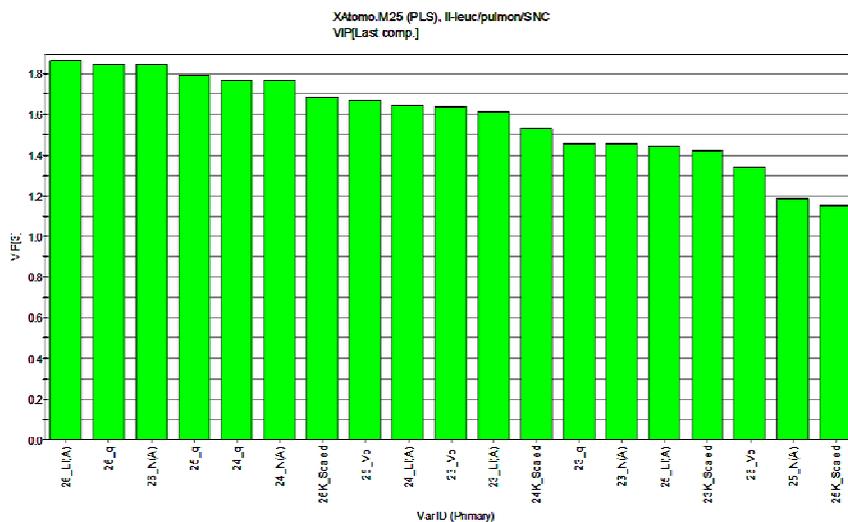
Tabla 7. Resultados de PLS total y de cada una de las actividades biológicas.

Propiedad	R²x	R²y	Q²
Total Actividades	0.829	0.989	0.826
Leucemia	0.62	0.99	0.899
Pulmón	0.617	0.99	0.899
Colon	0	0	0
SNC	0.827	0.995	0.883
Melanoma	0	0	0
Renal	0	0	0
Mama	0	0	0



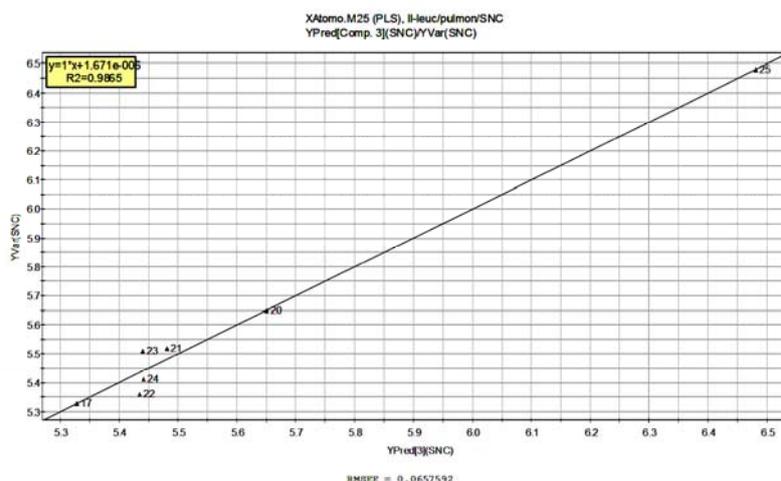
Gráfica 2. Resumen de valores de R²Y y Q²Y para los compuestos 17 y del 20 al 25 (Grupo II), solo para las respuestas biológicas de leucemia, pulmón y SNC.

La **Grafica 3** muestra los valores VIP para el grupo II. Sólo se muestran aquellos valores mayores de uno, que son los que más influyen en el modelo. Podemos ver que los VIP's de mayor influencia involucran a los átomos 26, 25, 24 y 23. Esto centraría al grupo farmacóforo en el fenilo del éster.



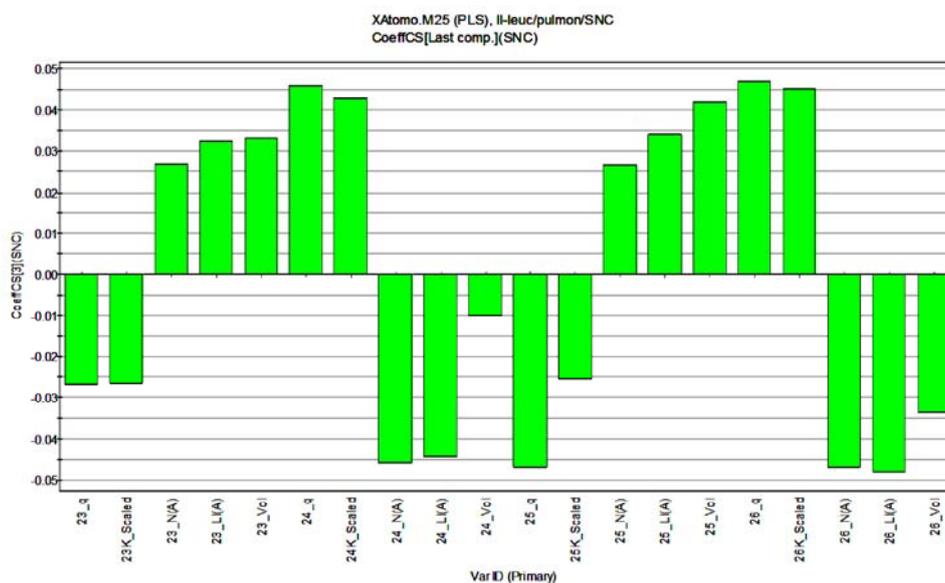
Gráfica 3: Gráfica de variables importantes en la proyección (VIP) para la regresión PLS de los compuestos 17y del 20 al 25 (Grupo II). Sólo se muestran los valores VIP mayores a uno.

Se realizó la comparación de los datos de actividad biológica observados y los datos de actividad biológica calculados con el modelo generado y se trazó la grafica mostrada en la **Grafica 4**. Obteniéndose una regresión lineal de 0.9865.



Gráfica 4: Gráfica de datos de actividad biológica observados contra datos de actividad biológica de SNC predichos por el modelo, para los compuestos del grupo II.(ver anexo iv)

El modelo generado se puede expresar como una ecuación del tipo $Y = Y_{avg} + BX + F$ donde B son los coeficientes de respuesta para cada descriptor y se muestran en las **Gráfica 5**.



Gráfica 5. Coeficientes de regresión para los descriptores electrónicos con mayor influencia en la predicción del grupo II.

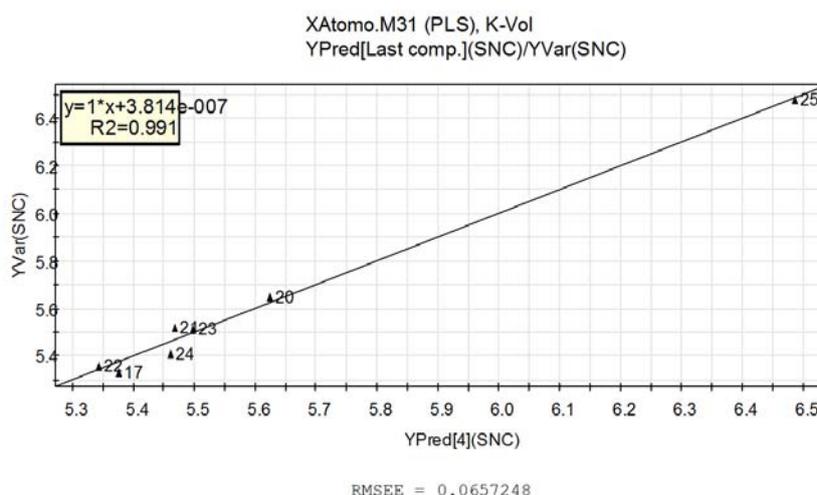
Los coeficientes con valores positivos describen la influencia positiva del descriptor en la respuesta biológica, es decir, al aumentar el valor del descriptor aumenta el valor de la respuesta. Un coeficiente con valor negativo produce el efecto contrario. De esta manera, al aumentar la población electrónica de los átomos 23 y 25 aumenta el pGI_{50} de las moléculas en

estudio. Mientras que al aumentar la población electrónica de los átomos 24 y 26 disminuye el pGI_{50} . La localización electrónica tiene una influencia positiva en los átomos 23 y 25 y negativa en 24 y 26. El volumen, la carga y la energía atómicas tienen el mismo comportamiento.

Para buscar nuevas moléculas con una posible mayor actividad, es necesario introducir sustituyentes que aumenten el valor de los coeficientes en los átomos que más influyen en el modelo.

Para analizar cuál de los descriptores de la densidad electrónica se relacionan más con la respuesta biológica en el grupo II, se analizó por separado cada una de las propiedades integradas. Para esto se realizó un análisis PLS. Las propiedades que más se ajustan a los datos son la energía cinética atómica ($R^2_x = 0.972$, $R^2_y = 0.719$, $Q^2 = 0.456$) y el volumen atómico ($R^2_x = 0.422$, $R^2_y = 0.942$, $Q^2 = 0.63$).

Si se realiza un análisis con la energía K y el volumen como descriptores se encuentra un modelo con 4 componentes principales capaz de predecir el 90% ($R^2X = 0.909$) de la variación en X para describir el 99% ($R^2Y = 0.995$) y predecir 91% ($Q^2Y = 0.914$) de la variación de Y. Superando la correlación de X, Y y predicción que se mostró en análisis con el total de las propiedades atómicas. El ajuste entre la respuesta observada y calculada presentando tiene una correlación de $r^2 0.991$ (ver **Gráfica 6**).



Gráfica 6. Gráfica de datos de actividad biológica observados contra datos de actividad biológica predichos por el modelo, para los compuestos del grupo II usando solo las propiedades K y Vol.

¿Qué pasa si restringimos el análisis a los átomos con un VIP mayor a uno? Hay que recordar que los cambios en las propiedades de los átomos con un VIP mayor a uno correlacionan con los cambios en las repuestas. En otras palabras, estos átomos están relacionados con la sección de la molécula que provee la actividad biológica. Para responder esta pregunta se realizó un análisis PLS tomando en cuenta solo los seis átomos del grupo fenilo, del 21 al 26. Los resultados se muestran en la **Tabla 8**.

Tabla 8. Resultados de PLS de los átomos 21-26.

Propiedad	R²x	R²y	Q²
q	1.00	0.996	0.972
K	0.866	0.725	0.232
N	0.759	0.824	0.594
Li	0.522	0.747	0.601
Vol	0.92	0.946	0.43

Los resultados muestran que el ajuste de los datos y la predicción con la carga atómica es perfecta. El resto de los descriptores dan valores regulares.

Se puede concluir de esta sección de análisis de propiedades atómicas que:

1. La propiedad atómica que mejor ajusta a X es la localización electrónica.
2. El conjunto de moléculas esta dividida en tres grupos, tanto estructural como estadísticamente.
3. Los cambios en las propiedades atómicas del grupo II presentan un buen ajuste y predicción de los PGI₅₀ para leucemia, cáncer de pulmón y SNC.
4. El anillo de fenilo dentro del éster es la parte de la molécula que mejor ajusta con las actividades biológicas.
5. Se estableció el efecto de cada descriptor por cada átomo del punto anterior.
6. La carga atómica es el mejor descriptor para esa porción de la molécula.

7.2 ESTUDIO DE LAS PROPIEDADES INTEGRADAS.

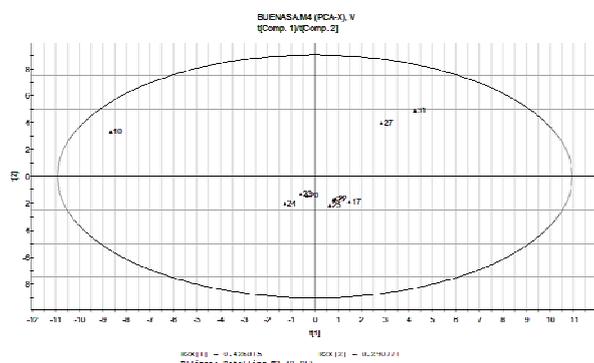
En este caso al igual que en el estudio por propiedades integradas, se inicio con un análisis de componentes principales con el conjunto completo de datos (140) para generar un modelo, pero sin éxito, por lo cual se tomó la decisión de hacer el análisis de componentes principal por cada una de los descriptores. La **Tabla 9** muestra los resultados de R^2x y Q^2 .

Tabla 9. Análisis PCA de cada una de las propiedades de enlace.

Propiedad	R^2x	Q^2
RHO	0.694	0.16
LAP	0.683	0.279
ELEP	0.74	0.32
V	0.859	0.421
G	0.714	0.245

El mejor ajuste y la mayor predicción lo presenta V, que son de 0.859 y 0.421, respectivamente. Esto muestra que la V es un buen descriptor. Por otro lado, el resto de los descriptores muestran un ajuste moderado y una predicción muy baja.

Al graficar los “scores” (t) de los componentes de cada uno de estos modelos generados para cada descriptor se encontró que se podía dividir a las moléculas en los mismos tres grupos: el compuesto 10 como grupo I, los compuestos 17 y del 20 al 25 son el grupo II y el grupo III esta compuesto por 27 y 31 (**Gráfica 7** para el descriptor V). El lector puede consultar el anexo vi para el resto de descriptores.



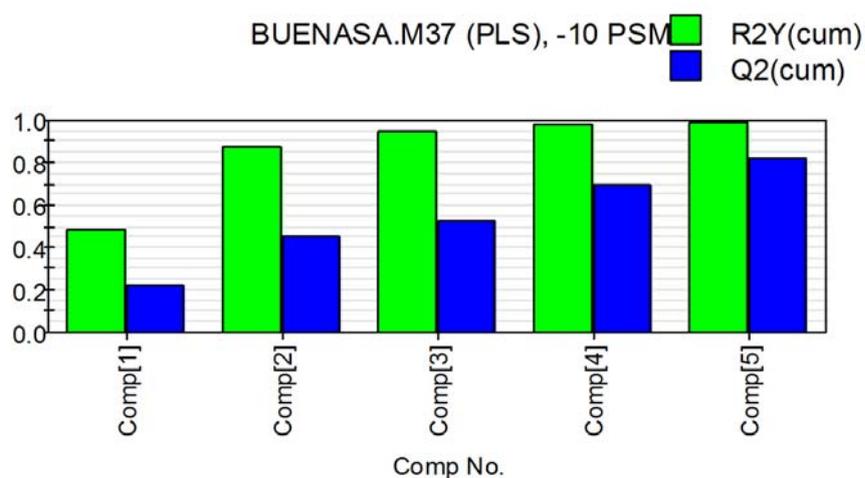
Gráfica 7. Gráfica de scores (t1/t2) del análisis PCA del modelo generado para el descriptor V. Se muestran tres grupos de compuestos que influyen en la respuesta. (ver anexo v para el resto de las gráfica de scores)

Se hicieron pruebas para saber si la unión de dos grupos daba buenos resultados. Cualquier exploración que incluyera al grupo III no producía un modelo. La unión del grupo II con el III dio un modelo PCA con una R^2X de 0.711 y una Q^2 de 0.428. El análisis PLS se realizó con este conjunto (II-III) y dió como resultado un modelo capaz de predecir el 93% ($R^2X = 0.939$) de la variación en X para describir el 98% ($R^2Y = 0.984$) y predecir 71% ($Q^2Y = 0.715$) de la variación de Y.

Se realizó un análisis PLS para todas y cada una de las variables de respuesta Y con el mismo conjunto de datos y se obtuvieron los resultados de la **Tabla 10**. Los ajustes y la predicción son buenos con todas las variables de respuesta al igual que para el caso de tumores en tejido de pulmón. En el caso de melanomas y tumores en el SNC se tiene un ajuste regular con X, un buen ajuste con Y y una regular predicción. En el caso de leucemia y tumores en el colon, tejido renal y mama no se pudo generar un modelo PLS. Por lo que el modelo generado con todos las variables de repuesta se debe al ajuste con las respuestas para melanomas y tumores en pulmón y SNC. Un estudio con sólo tres actividades biológicas (pulmón, SNC y melanoma) mostró un modelo con 5 componentes principales (ver **Gráfica 8**) capaz de predecir el 93% ($R^2X = 0.938$) de la variación en X para describir el 98% ($R^2Y = 0.988$) de la variación de Y y predecir 82% ($Q^2 = 0.824$). Este resultado supera la predicción que mostró el total de las actividades aunque no los ajustes de X y Y que son los mismos (ver **Tabla 10**).

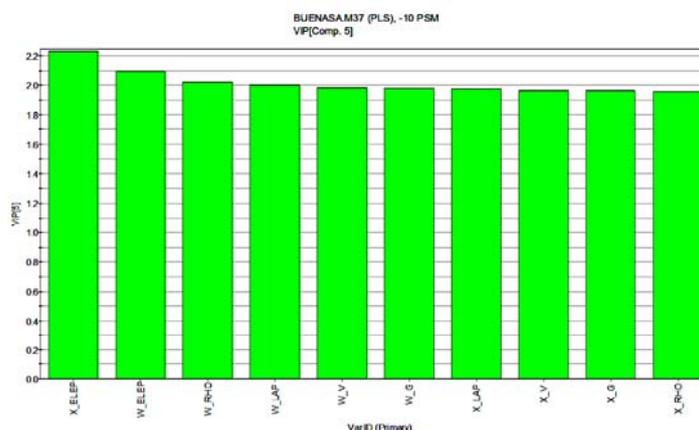
Tabla 10. Resultados de PLS total y de cada una de las actividades biológicas.

Propiedad	R^2x	R^2y	Q^2
Total Actividades	0.939	0.984	0.715
Leucemia	0	0	0
Pulmon	0.86	0.994	0.874
Colon	0	0	0
SNC	0.78	0.975	0.807
Melanoma	0.711	0.985	0.581
Renal	0	0	0
Mama	0	0	0
Pulmón/SNC/Melano	0.938	0.988	0.824



Gráfica 8. Resumen de valores de R2Y y Q2Y para los compuestos todos los compuestos excepto la molécula 10.

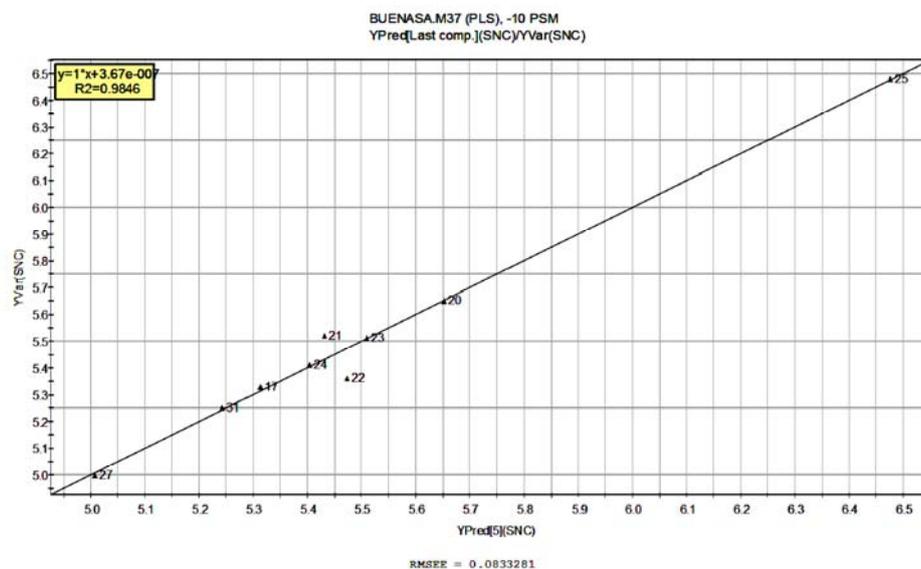
Para este último estudio se obtuvo los valores VIP que se muestran en la **Gráfica 9**. Sólo se muestran aquellos valores mayores de uno, que son los que más influyen en el modelo. Podemos ver que los VIP's de mayor influencia implican a los enlaces W y X. Esto centraría al grupo farmacofórico en el fenilo del éster.



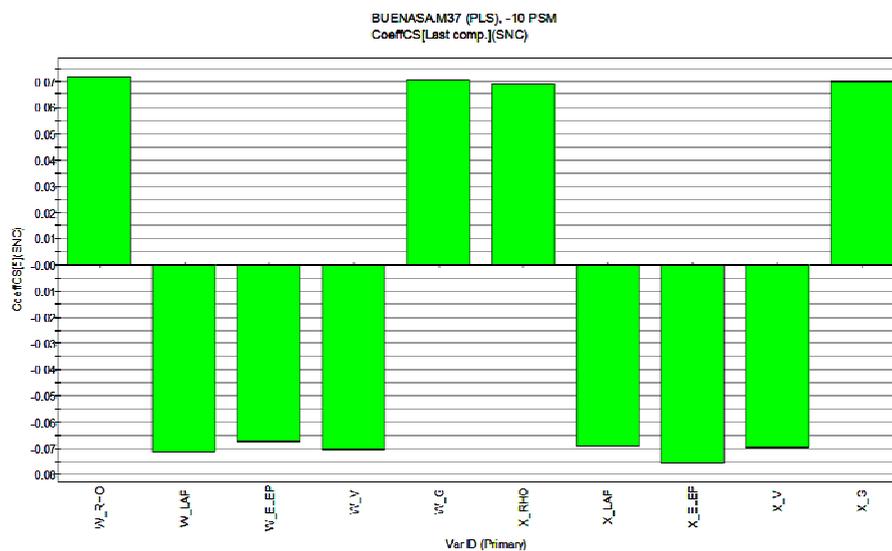
Gráfica 9: Gráfica de variables importantes en la proyección (VIP) para la regresión PLS de los compuestos excepto del 10.

Al igual que en el análisis por propiedades atómicas se realizó la comparación de los datos de actividad biológica observados y los datos de actividad biológica calculados con el modelo

generado y se trazó la gráfica mostrada en la **Gráfica 10**. Obteniéndose una regresión lineal de 0.9846. El modelo generado se puede expresar como una ecuación del tipo $Y = Y_{avg} + BX + F$ donde B son los coeficientes de respuesta para cada descriptor de enlace (**Gráfica 11**). Solo se muestran los coeficientes relacionado a los enlace con mayor VIP.



Gráfica 10: Gráfica de datos de actividad biológica observados contra datos de actividad biológica predichos (SNC) por el modelo, para los compuestos del grupo II y III. (ver resto de graficas en anexo vi)



Gráfica 11. Coeficientes de regresión para los descriptores electrónicos con mayor influencia en la predicción del grupo II.

En la **Gráfica 11** se puede observar que la densidad electrónica y la energía cinética tiene una influencia positiva, esto es, al aumentar estas propiedades en los enlaces W y X aumenta el PGI_{50} . Mientras que cuando el Laplaciano de la densidad, la electicidad o la energía potencial aumentan, el PGI_{50} disminuye.

De esta sección se puede concluir que:

1. La energía potencial en el punto crítico es el mejor descriptor
2. Los datos pueden dividirse en tres grupos aunque la unión del grupo II y III da mejores resultados en el análisis.
3. Que los datos de los grupos II-III ajusta bien a los PGI_{50} de pulmón, melanoma y SNC.
4. Los enlaces con mayor influencia son los X y W que se encuentran el fenilo del éster.
5. Se pudo establecer la influencia que tienen los descriptores de los enlaces X y W en la actividad biológica.
6. Con lo anterior se puede explicar porque la molécula 25 es la más activa ya que, entre otros descriptores, tiene la menor energía potencial tanto en el enlace X como en el W.

8. CONCLUSIONES

La conclusión general de este trabajo fue que es posible relacionar las propiedades de la densidad electrónica, tanto en su versión local como integrada, con las actividades antiproliferativas de los compuestos derivados de ácido antranílico utilizando un análisis de datos multivariable.

Se puede concluir de la sección de análisis de propiedades atómicas que:

1. La propiedad atómica que mejor ajusta a X es la localización electrónica.
2. El conjunto de moléculas está dividida en tres grupos, tanto estructural como estadísticamente.
3. Los cambios en las propiedades atómicas del grupo II presentan un buen ajuste y predicción de los PGI₅₀ para leucemia, cáncer de pulmón y SNC.
4. El anillo de fenilo dentro del éster es la parte de la molécula que mejor ajusta con las actividades biológicas.
5. Se estableció el efecto de cada descriptor de cada átomo del punto anterior.
6. La carga atómica es el mejor descriptor para esa porción de la molécula.

De la sección de propiedades locales:

1. La energía potencial en el punto crítico es el mejor descriptor
2. Los datos pueden dividirse en tres grupos aunque la unión del grupo II y III da mejores resultados en el análisis.
3. Que los datos de los grupos II-III ajusta bien a los PGI₅₀ de pulmón, melanoma y SNC.
4. Los enlaces con mayor influencia son los X y W que se encuentran el fenilo del éster.
5. Se pudo establecer la influencia que tienen los descriptores de los enlaces X y W en la actividad biológica.
6. Con lo anterior se puede explicar porque la molécula 25 es la más activa ya que, entre otros descriptores, tiene la menor energía potencia tanto en el enlace X como en el W.

9. REFERENCIAS

- 1) Cenzo Congiu, Maria Teresa Cocco, Valentina Lilliu, and Valentina Onnis *New Potential Anticancer Agents Based on the Anthranilic Acid Scaffold Synthesis and Evaluation of Biological Activity* *J. Med. Chem.* **2005**, 48, 8245-8252
- 2) Thun, M. J.; Henley, S. J.; Patrono, C. *Nonsteroidal antiinflammatory drugs as anticancer agents: Mechanistic, pharmacologic, and clinical issues.* *J. Natl. Cancer Inst.* **2002**, 94, 252-266.
- 3) Collet JP, Sharpe C, Belzile E, Boivin JF, Hanley J, Abenhaim L. *Colorectal cancer prevention by non-steroidal anti-inflammatory drugs: effects of dosage and timing.* *Br J Cancer* **1999**, 81, 62–8.
- 4) Thun MJ, Heath CW Jr. *Aspirin use and reduced risk of gastrointestinal tract cancers in the American Cancer Society prospective studies.* *Prev. Med.* **1995**, 24, 116–8.
- 5) Bennett A, Del Tacca M. *Proceedings: Prostaglandins in human colonic carcinoma.* *Gut* **1975**, 16, 409.
- 6) Jaffe BM. *Prostaglandins and cancer: an update.* *Prostaglandins* **1974**, 6, 453–61.
- 7) Reddy BS, Maruyama H, Kelloff G. *Dose-related inhibition of coloncarcinogenesis by dietary piroxicam, a nonsteroidal antiinflammatory drug, during different stages of rat colon tumor development.* *Cancer Res* **1987**, 47, 5340–6.
- 8) Reddy BS, Kawamori T, Lubet RA, Steele VE, Kelloff GJ, Rao CV. *Chemopreventive efficacy of sulindac sulfone against colon cancer depends on time of administration during carcinogenic process.* *Cancer Res* **1999**, 59, 3387–91.
- 9) Narisawa T, Sato M, Tani M, Takahashi T, Goto A. *Inhibition of development of methylnitrosourea-induced rat colon tumors by indomethacin treatment.* *Cancer Res* **1981**, 41, 1954–7.
- 10) Li H, Kramer P, Steele V, Kelloff G, Lubet R, Pereira M. *Effect of duration of treatment with piroxicam on azoxymethane-induced colon cancer, aberrant crypt foci, apoptosis and cell proliferation in rats [abstract].* *Proc Am Assoc Cancer Res* **1998**, 39, 23.
- 11) Heike Wagner, Karl-Heinz Kostka, Willi Adelhardt, Ulrich F. Schaefer. *Effects of various vehicles on the penetration of flufenamic acid into human skin.* *Eur. J. Pharm. and Biopharm.* **2004**, 58, 121–129.

- 12) Weiser, T.; Wienrich, M. *Investigations on the mechanism of action of the antiproliferant and ion channel antagonist flufenamic acid.* Naunyn Schmiedebergs. Arch. Pharmacol. **1996**, 353, 452-460.
- 13) Wen Zhu, Aaron Smith, and Charles Y. F. Young *A Nonsteroidal Anti-Inflammatory Drug, Flufenamic Acid, Inhibits the Expression of the Androgen Receptor in LNCaP Cells* Endocrinol. **1999**, 140, 5451-5454.
- 14) Gupta, R. A.; DuBois, R. N. *Colorectal cancer prevention and treatment by inhibition of cyclooxygenase-2.* Nat. Rev. Cancer **2001**, 1, 11-21.
- 15) Dannenberg, A. J.; Altori, N. K.; Boyle, J. O.; Dang, C.; Howe, L. R.; Weksler, B. B.; Subbaramaiah, K. *Cyclo-oxygenase 2: A pharmacological target for the prevention of cancer.* Lancet Oncol. **2001**, 2, 544-551.
- 16) Insel PA . *Analgesic-Antipyretic Agents; Pharmacotherapy of Gout.* In: *Goodman and Gilman's The Pharmacological Basic of Therapeutics* McGraw-Hill, New York, Ch 26:673-697
- 17) Masataka Majima, Hideki Amano and Izumi Hayashi *Prostanoid receptor signaling relevant to tumor growth and angiogenesis* TRENDS Pharmacol. Sci. **2003**, Vol.24 No.10 524-529.
- 18) G. Davies, L.-A. Martin, N. Sacks & M. Dowsett *Cyclooxygenase-2 (COX-2), aromatase and breast cancer: a possible role for COX-2 inhibitors in breast cancer chemoprevention.* Ann. Oncol. **2002**, 13, 669–678.
- 19) Turini, M. E.; DuBois, R. N. *Cyclooxygenase-2: A therapeutic target.* Annu. Rev. Med. **2002**, 53, 35-57.
- 20) David R. Bevan. *QSAR and Drug Design.* Network Science. <http://www.netsci.org/Science/Compchem/feature12.html> consultada 8 /nov/ 2008
- 21) C. D. Selassie. *History of Quantitative Structure-Activity Relationships.* Burger's Medicinal Chemistry and Drug Discovery. Sixth Edition, Volume 1: Drug Discovery, 1-48. **2003**, John Wiley&Sons, Inc.
- 22) <http://www.monografias.com/trabajos41/metodos-qsar-qspr/metodos-qsar-qspr2.shtml> consultada 17/nov/2008
- 23) Cuevas, G.; Cortés, F. *Introducción a la Química Computacional.* Fondo de Cultura Económica, México, **2003**.
- 24) Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.;

- Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**.
- 25) Bader, R.F.W. *Atoms in molecules: a quantum theory*. Clarendon Press, Oxford. **1990**.
- 26) Cherif F. Matta and Russell J. Boyd. *The Quantum Theory of Atoms in molecules, From Solid State to DNA and Drug Design*. Wiley-VCH. Canada, **2007**.
- 27) S. E. O'Brien and P. L. A. Popelier. *Quantum Molecular Similarity. 3. QTMS Descriptors*. J.Chem. Inf. Comp. Sci. **2001**, 41, 764-775.
- 28) P. L. A. Popelier. *Quantum Molecular Similarity 1. BCP Space*. Journal of Physical Chemistry A **1999**, 103, 2883-2890.
- 29) Cherif F. Matta and Russell J. Boyd. *The Quantum Theory of Atoms in molecules, From Solid State to DNA and Drug Design*. Wiley-VCH. Canada, **2007**.
- 30) L Eriksson, E Johansson, N Kettaneh-Wold, J Trygg, C Wikström, and S Wold. *Multi- and Megavariate Data Analysis, Part II, Advanced Applications and Method Extensions*. Umetrics Acad., **2006**.
- 31) L Eriksson, E Johansson, N Kettaneh-Wold, J Trygg, C Wikström, and S Wold. *Multi- and Megavariate Data Analysis, Part I, Advanced Applications and Method Extensions*. Umetrics Acad., **2006**.
- 32) AIMAll (Version 08.05.04), Todd A. Keith, 2008 (aim.tkgristmill.com)
- 33) Miloň Tichý, Iveta Hanzlíková, Marián Rucki, Adéla Pokorná, Rút Uzlová, Jana Tumová *Acute Toxicity of Binary Mixtures: Alternative Methods, QSAR and Mechanisms* Interdisc Toxicol. **2008**,1(1),15–17

10. ANEXO

i. Tablas de datos para el análisis de descriptores atómicos

MOLECULA	Leucemia	Pulmón	Colon	SNC	Melanoma	Renal	Mama	1 q	1K Scaled	1_N(A)	1_L(A)	1 Vol	2_q
10	5.57	4.7	5.09	4.69	4.77	4.63	4.47	-0.63601715	-100.352515	9.63601715	9.05552444	93.9708003	-0.62966844
17	6.13	5.69	5.39	5.33	5.51	5.27	5.14	-0.63902965	-100.492399	9.63902965	9.05868577	93.7127152	-0.62918401
20	7.29	6.59	6.44	5.65	6.51	6.48	6.55	-0.62930941	-100.360962	9.62930941	9.05394616	93.768875	-0.62863365
21	5.49	5.62	5.36	5.52	5.69	5.53	5.32	-0.63903202	-100.342653	9.63903202	9.05859466	93.6956793	-0.62863821
22	5.44	5.55	5.18	5.36	5.58	5.53	5.43	-0.6389443	-100.342848	9.6389443	9.05844192	93.7244619	-0.62882659
23	5.52	5.45	5.34	5.51	5.64	5.52	5.44	-0.63886529	-100.255728	9.63886529	9.05800518	93.7294415	-0.6290058
24	5.62	5.59	5.34	5.41	5.76	5.51	5.33	-0.63881665	-100.197421	9.63881665	9.05830528	93.692068	-0.62886582
25	7.37	7.44	7.39	6.48	6.54	7.15	7.34	-0.6386161	-100.490221	9.6386161	9.05788794	93.6758359	-0.62886413
27	5.34	5.06	5.43	5.00	5.21	5.39	5.44	-0.64018007	-100.341309	9.64018007	9.05863616	93.7240935	-0.62938225
31	5.29	5.20	5.26	5.25	5.24	5.19	5.57	-0.64015318	-100.484669	9.64015318	9.0581966	93.6954132	-0.62858129

2K_Scaled	2_N(A)	2_L(A)	2_Vol	3_q	3K Scaled	3_N(A)	3_L(A)	3_Vol	4_q	4K_Scaled
-100.357523	9.62966844	9.05600418	94.4403281	-0.63648872	-100.35334	9.63648872	9.05627502	93.8693734	1.72673626	-36.7715919
-100.509809	9.62918401	9.05407656	93.7676233	-0.62863703	-100.508542	9.62863703	9.05459437	93.9478445	1.75178073	-36.8147738
-100.359336	9.62863365	9.05470256	93.8915522	-0.63880898	-100.343747	9.63880898	9.05828007	93.6932264	1.75320988	-36.7598623
-100.36073	9.62883821	9.05331337	93.7501684	-0.62848749	-100.358799	9.62848749	9.05426738	93.9469614	1.75452318	-36.759081
-100.360743	9.62882659	9.0533353	93.759146	-0.62844457	-100.358922	9.62844457	9.05418824	93.9442791	1.75417852	-36.7587811
-100.273089	9.6290058	9.05355625	93.7597733	-0.62851149	-100.271202	9.62851149	9.05441502	93.9401821	1.75501732	-36.7267654
-100.21512	9.62886582	9.05328304	93.7488632	-0.62836831	-100.213023	9.62836831	9.05424572	93.9576551	1.75533403	-36.7053769
-100.507678	9.62886413	9.05336733	93.761042	-0.62834542	-100.50582	9.62834542	9.05413319	93.9711419	1.75559069	-36.8123258
-100.35845	9.62938225	9.05503641	93.9110577	-0.62924224	-100.358606	9.62924224	9.05483872	93.9070912	1.74991184	-36.761748
-100.501999	9.62858129	9.05455124	93.9944665	-0.62918201	-100.503786	9.62918201	9.05387203	93.7828849	1.75245331	-36.8130689

4_N(A)	4_LI(A)	4_Vol	5_q	5K_Scaled	5_N(A)	5_LI(A)	5_Vol	6_q	6K_Scaled	6_N(A)
4.27326174	2.7382172	23.5850239	0.03727144	-38.0293578	5.96272856	3.87132339	65.5699735	0.00647073	-37.9934109	5.99352927
4.24821927	2.72393619	23.2150163	0.59765771	-37.7507856	5.40234229	3.42675505	55.5862147	-0.00294451	-38.0630837	6.00294451
4.24679012	2.72324551	23.1812395	0.59839026	-37.694363	5.40160974	3.42690948	55.5630377	-0.00126473	-38.0062197	6.00126473
4.24547682	2.72217593	23.0031872	0.59950021	-37.6939953	5.40049979	3.4254801	55.4769919	-0.00266497	-38.0070304	6.00266497
4.24582148	2.722107	23.1551011	0.5993162	-37.6941006	5.4006838	3.42562414	55.4982934	-0.00279921	-38.0070191	6.00279921
4.24498268	2.72179621	23.0407199	0.59821663	-37.6612825	5.40178337	3.42683543	55.5522367	-0.00526228	-37.9725978	6.00526228
4.24466597	2.72171868	23.0328954	0.60003411	-37.6389055	5.39996589	3.42531731	55.4400522	-0.00167893	-37.9516728	6.00167893
4.24440931	2.72125339	22.9908747	0.59964434	-37.7490699	5.40035566	3.42551992	55.5059497	-0.00260118	-38.0624687	6.00260118
4.25008816	2.72602126	23.1488554	0.5983022	-37.694743	5.4016978	3.4264152	55.6691322	-0.00872471	-38.0073087	6.00872471
4.24754689	2.72382789	23.0856894	0.60114965	-37.7481173	5.39885035	3.42404448	55.4728291	-0.00607666	-38.0634815	6.00607666

6_LI(A)	6_Vol	7_q	7K_Scaled	7_N(A)	7_LI(A)	7_Vol	8_q	8K_Scaled	8_N(A)	8_LI(A)
3.94013589	78.1298836	0.41555275	-37.7498742	5.58444725	3.57906779	59.8148714	0.00294751	-37.9889168	5.99705249	3.93848048
3.95370187	79.0018342	0.44714161	-37.7880246	5.55285839	3.55519304	58.8239817	-0.01108589	-38.0583438	6.01108589	3.95901909
3.9481535	79.6194626	0.44187203	-37.7363866	5.55832797	3.55874724	59.0245868	-0.01026314	-38.0026867	6.01026314	3.95770656
3.95310518	78.9504379	0.44427873	-37.733724	5.55572127	3.55786186	58.8675066	-0.01012615	-38.0012812	6.01012615	3.95920885
3.95327219	78.986597	0.44429399	-37.733733	5.55570601	3.55784363	58.86934	-0.01057384	-38.0013713	6.01057384	3.95855543
3.9550738	79.6714298	0.44275771	-37.7022926	5.55724229	3.55998003	58.7693871	-0.01080217	-37.9687856	6.01080217	3.95838947
3.9523962	78.9674775	0.44287043	-37.6802555	5.55732957	3.55922004	58.8965406	-0.00944435	-37.9483418	6.00944435	3.95739388
3.95330583	78.9566486	0.44324519	-37.7896202	5.55675481	3.55889393	58.9043307	-0.01050419	-38.0572312	6.01050419	3.95851145
3.95450499	80.1558064	0.42515161	-37.7388923	5.57484839	3.58401621	59.9637408	0.00475182	-37.9995721	5.99524818	3.95803715
3.95372169	78.7955688	0.42533291	-37.7929718	5.57466709	3.58356753	60.0327766	0.00552385	-38.0535677	5.99447615	3.95568877

8_Vol	9_q	9K_Scaled	9_N(A)	9_LI(A)	9_Vol	10_q	10K_Scaled	10_N(A)	10_LI(A)	10_Vol
77.6814916	0.01938482	-37.9696117	5.98061518	3.92531626	79.3579707	0.01253017	-37.9802541	5.98746983	3.93433035	78.9226481
78.511091	0.61197367	-37.671213	5.38802633	3.48792962	67.0109528	-1.20323606	-55.4182153	8.20323606	6.62731553	111.365186
79.3377073	0.61240574	-37.6149757	5.38759426	3.46791345	67.0275173	-1.20362367	-55.3372685	8.20362367	6.62771722	111.273936
79.5207825	0.61295603	-37.6145811	5.38704397	3.46760578	66.9467773	-1.20204078	-55.3359902	8.20204078	6.62618396	111.287782
79.5616106	0.61284818	-37.6147147	5.38715182	3.46762507	66.9640494	-1.20207864	-55.3360284	8.20207864	6.62621431	111.302989
79.3921689	0.6132954	-37.581612	5.3867046	3.46715814	66.9707196	-1.20277874	-55.2886134	8.20277874	6.62683644	111.24824
79.3530693	0.61326817	-37.5599616	5.38673183	3.46744012	66.9331706	-1.20229548	-55.2562293	8.20229548	6.62636831	111.310014
79.4674078	0.61300999	-37.6098257	5.38699001	3.46745067	66.9819117	-1.20222646	-55.4174942	8.20222646	6.62637333	111.310387
80.4721316	0.60808916	-37.6172332	5.39191084	3.47350452	67.0426289	-1.20657238	-55.3400585	8.20657238	6.63149409	111.393896
80.4996679	0.60931552	-37.6712651	5.39168448	3.47362505	67.0871812	-1.20512111	-55.4191084	8.20512111	6.62995383	111.422038

11_q	11K_Scaled	11_N(A)	11_LI(A)	11_Vol	12_q	12K_Scaled	12_N(A)	12_LI(A)	12_Vol	13_q
-1.25262335	-55.3825753	8.25262335	6.55624847	92.3399266	0.4571208	-37.7023221	5.5428792	3.54854215	58.944482	0.00014948
-1.260128	-55.4513724	8.260128	6.56496134	92.7092632	0.43183966	-37.777461	5.56816034	3.56929278	59.7041496	0.00388195
-1.25527458	-55.365476	8.25527458	6.5585456	92.9543368	0.43985613	-37.7164032	5.56014387	3.56205529	59.4474624	0.00403712
-1.25975243	-55.3883134	8.25975243	6.56387624	92.8851689	0.43464315	-37.7188722	5.56535685	3.56731832	59.617991	0.00455222
-1.25979827	-55.3680526	8.25979827	6.5640203	92.7249636	0.43390139	-37.7195792	5.56609861	3.5677433	59.7072865	0.00465285
-1.2555935	-55.3215369	8.2555935	6.55779037	92.2680153	0.44067943	-37.6806658	5.56932057	3.56214705	59.5675474	0.0050492
-1.25516954	-55.2894891	8.25516954	6.55825685	92.6159699	0.44049209	-37.660702	5.55950791	3.56251329	59.4695363	0.00561894
-1.25866564	-55.4480503	8.25866564	6.56260945	92.73448	0.43490186	-37.7743272	5.56509814	3.56678701	59.6149687	0.00495886
-1.26517389	-55.3985977	8.26517389	6.57177026	93.6047825	1.02069311	-37.369219	4.97930689	3.11591438	48.6297761	-1.22530562
-1.27293341	-55.4777868	8.27293341	6.57565105	95.0764238	1.0196439	-37.4243044	4.9803561	3.11703967	48.7015293	-1.22521493

13K_Scaled	13_N(A)	13L(A)	13_Vol	14_q	14K_Scaled	14_N(A)	14_L(A)	14_Vol	15_q	15K_Scaled
-37.9891351	5.99985052	3.94434512	78.6542282	0.01567569	-37.979657	5.98432431	3.92702689	78.8254603	0.0093819	-37.9766925
-38.0451589	5.99611805	3.93914963	77.9939298	0.01638695	-38.0341829	5.98361305	3.92761705	78.0133054	0.01183154	-38.0320307
-37.9894156	5.99596288	3.9391215	77.9159915	0.01716375	-37.9781604	5.98283625	3.92698872	78.9840815	0.01250629	-37.9756244
-37.988207	5.99544778	3.9390204	78.0238153	0.01760227	-37.9776073	5.98239773	3.92682115	79.0029435	0.01335773	-37.9752399
-37.988375	5.99534715	3.93886131	77.9972605	0.01755652	-37.9775891	5.98244348	3.92667335	78.9367494	0.01313161	-37.9752318
-37.9558134	5.9949508	3.93875007	78.0912119	0.01793201	-37.9462825	5.98206799	3.92656164	78.905822	0.01354665	-37.942613
-37.9332908	5.99438106	3.9382762	77.9724251	0.01838935	-37.9227166	5.98161065	3.92629037	78.9427786	0.01417558	-37.9201035
-38.0441314	5.99504114	3.93849264	78.0320266	0.01788374	-38.0332754	5.98211628	3.92671838	78.9596718	0.01352584	-38.0308253
-55.3536528	8.22530562	6.61523106	107.458382	0.62117942	-37.812353	5.37882058	3.47021733	66.4949346	0.00248576	-37.9943999
-55.4330876	8.22521493	6.61557825	107.503071	0.62051803	-37.8668377	5.37948197	3.47113367	66.5448587	0.00357177	-38.0487704

15_N(A)	15_L(A)	15_Vol	16_q	16K_Scaled	16_N(A)	16_L(A)	16_Vol	17_q	17K_Scaled	17_N(A)
5.9906181	3.94003369	81.104745	0.01874513	-37.9941795	5.98025487	3.92380901	76.7995009	-0.01028899	-37.9849477	6.01028899
5.98816846	3.9374239	80.8342837	0.02123533	-38.0481478	5.97876467	3.92385508	77.0206494	-0.01099835	-38.0486702	6.01099835
5.98749371	3.93712474	80.8958419	0.02145837	-37.9915485	5.97854163	3.92321049	76.9427352	-0.00687717	-37.9893817	6.00687717
5.98664227	3.93646593	80.8003961	0.02202129	-37.991362	5.97797871	3.92331049	76.9788057	-0.00936866	-37.9905719	6.00936866
5.98686839	3.93661839	80.7980467	0.02180118	-37.9912569	5.97819882	3.92343824	77.0431284	-0.01039469	-37.9911189	6.01039469
5.98645335	3.93633922	80.8818863	0.02173832	-37.9595771	5.97826168	3.92345649	76.8563009	-0.00665319	-37.9529161	6.00665319
5.98582442	3.93591161	80.8706686	0.02375914	-37.9358801	5.97624086	3.9219836	76.8982294	-0.00309164	-37.9309378	6.00309164
5.98647418	3.93631306	80.7972154	0.02220718	-38.0467753	5.97779282	3.92311826	76.9259222	-0.00902047	-38.0465004	6.00902047
5.99751424	3.95260278	82.0336087	0.02638933	-37.9963181	5.97361067	3.92155572	76.3981827	-0.01217747	-37.998673	6.01217747
5.99642823	3.95183384	82.0029329	0.02686674	-38.0526126	5.97311326	3.92145141	76.3020066	-0.01643219	-38.0543312	6.01643219

17_LI(A)	17_Vol	18_q	18K_Scaled	18_N(A)	18_LI(A)	18_Vol	19_q	19K_Scaled	19_N(A)	19_LI(A)
3.92428944	73.7899771	1.52293659	-36.9845791	4.47708341	2.8518791	39.7257808	-1.20268823	-75.9101819	9.20288823	8.24799439
3.92079308	73.2795924	1.55059103	-37.0034319	4.44940897	2.83023819	39.0839924	-1.19723796	-76.0184877	9.19723796	8.24869606
3.91927957	73.1764135	1.5341692	-36.9594356	4.4658308	2.84332548	39.3956374	-1.19775315	-75.9094366	9.19775315	8.24454121
3.92076804	73.2283332	1.53856364	-36.9564191	4.48143836	2.83975633	39.2891951	-1.19605824	-75.9072865	9.19605824	8.24389675
3.92076464	73.3448431	1.54019165	-36.9556958	4.45980935	2.8384708	39.2265986	-1.19723041	-75.9068152	9.19723041	8.24555339
3.92045241	73.2404879	1.52966359	-36.9275415	4.47033841	2.84809354	39.5919758	-1.19861628	-75.8433657	9.19861628	8.24150064
3.91616224	73.2554105	1.51913449	-36.9183723	4.48086551	2.85400252	40.0268301	-1.19348259	-75.8024447	9.19348259	8.24102604
3.92017488	73.2099614	1.53678322	-37.0120637	4.46321678	2.84140308	39.3380558	-1.19776781	-76.0180688	9.19776781	8.24522469
3.92895329	73.9082251	1.54463639	-36.9507644	4.45536361	2.83297591	39.2679395	-1.19649872	-75.9058787	9.19649872	8.24350964
3.93037896	74.1992542	1.54555306	-37.0032551	4.45444694	2.83186979	39.2368873	-1.19678989	-76.0132433	9.19678989	8.2434815

19_Vol	20_q	20K_Scaled	20_N(A)	20_LI(A)	20_Vol	21_q	21K_Scaled	21_N(A)	21_LI(A)	21_Vol
120.399365	-1.13081854	-75.8447271	9.13081854	7.8912231	89.4494049	0.49372881	-37.695339	5.50627119	3.54839097	61.0583031
121.550449	-1.1332687	-75.9475954	9.1332687	7.90946664	92.2881963	0.4363924	-37.795363	5.5636076	3.60412975	58.941561
120.837114	-1.12065842	-75.8371437	9.12065842	7.89353159	87.9657206	0.49495077	-37.7046974	5.50504923	3.55557701	58.0485773
121.005313	-1.13047784	-75.8312428	9.13047784	7.90667671	92.4524881	0.45374891	-37.7271856	5.54625109	3.59015959	59.0249046
121.06128	-1.13125842	-75.8312813	9.13125842	7.9083436	92.4354216	0.4468492	-37.7316377	5.5531508	3.59689405	59.6224299
120.029082	-1.12943032	-75.7777255	9.12943032	7.89153905	89.1501672	0.50989757	-37.6607138	5.49010243	3.54358199	58.114343
122.262217	-1.11638693	-75.728249	9.11638693	7.87724156	86.5756254	0.56234265	-37.6142924	5.43765735	3.4804449	56.3836831
120.875776	-1.13107169	-75.9424978	9.13107169	7.90854582	92.8828728	0.43829643	-37.8060496	5.56170357	3.60765585	60.1114131
120.88351	-1.12445592	-75.8425132	9.12445592	7.89543366	87.9124591	0.48981795	-37.7081585	5.51018205	3.5609351	58.0318978
120.878598	-1.1319188	-75.9414881	9.1319188	7.91099826	92.3713337	0.43089086	-37.8085957	5.56910914	3.61505842	60.4356448

22_q	22K_Scaled	22_N(A)	22_LI(A)	22_Vol	23_q	23K_Scaled	23_N(A)	23_LI(A)	23_Vol	24_q
0.04802318	-37.9723717	5.95197684	3.90401011	76.3207335	0.02153285	-37.9665871	5.97846715	3.9263311	79.7834584	0.01976611
0.03735272	-38.0287289	5.96264728	3.91449807	77.7159131	0.02484765	-38.0172116	5.97515235	3.91832349	78.8529532	-0.00604384
0.04843034	-37.9723709	5.95156966	3.90308394	76.1928213	0.02222596	-37.9655265	5.97777404	3.9261138	79.8658767	0.0207587
0.02694438	-37.9797625	5.97305562	3.92624359	79.5552496	0.0248725	-37.9625625	5.9751275	3.92253048	79.3446599	0.05937716
0.0495829	-37.969808	5.9504171	3.90179172	76.349031	0.0483375	-37.962817	5.9516625	3.90327422	78.1289147	0.05292668
0.0566407	-37.9341508	5.9433593	3.89634199	75.8939487	0.05205989	-37.9299095	5.94794011	3.9008063	78.1256414	0.06398194
0.09301371	-37.8683542	5.90698629	3.82427058	66.9708266	0.07811113	-37.9084086	5.92188887	3.87831932	76.4994697	0.07483663
0.04338314	-38.0352207	5.95661686	3.90878025	76.3105477	0.0070583	-38.0426427	5.9929417	3.94440555	80.7173848	0.59627165
0.04946631	-37.9709798	5.95053369	3.90251848	76.2628181	0.02371038	-37.9643281	5.97628962	3.92515246	79.729087	0.02173779
0.04402921	-38.0334845	5.95597079	3.90819514	76.3516272	0.008186	-38.0409322	5.991814	3.94377776	80.5641964	0.59762284

24K_Scaled	24_N(A)	24_LI(A)	24_Vol	25_q	25K_Scaled	25_N(A)	25_LI(A)	25_Vol	26_q	26K_Scaled
-37.9667945	5.98023389	3.92814165	79.8359872	0.04332837	-37.9674761	5.95667183	3.90774465	78.1484638	0.06152096	-37.9398556
-38.0342791	6.00604364	3.95400586	80.8120109	0.54255329	-37.7195786	5.45744671	3.51263793	57.9223868	0.03235914	-38.0500237
-37.966029	5.9792413	3.92743798	79.804489	0.04389746	-37.9668135	5.95610254	3.90773117	78.0761202	0.06406354	-37.9382183
-37.9286053	5.94062284	3.85883164	68.6232018	0.02659625	-0.61809151	0.97440375	0.42837982	47.4702527	0.07017479	-37.9718373
-37.9320801	5.94707332	3.86572912	69.1730205	0.04713149	-37.9632906	5.95286851	3.90397485	78.0890941	0.03151659	-37.9777828
-37.8944417	5.93601806	3.86273423	68.7440281	0.07419359	-37.9305537	5.92580641	3.88020471	76.4125665	0.06530397	-37.8999626
-37.8670582	5.92516337	3.84043719	68.0861584	0.07763306	-37.9084507	5.92236694	3.87862828	76.5116485	0.09159694	-37.8692676
-37.6758404	5.40372835	3.48485415	68.0544894	-1.21550422	-55.4099131	8.21550422	6.65186824	114.730358	0.61721814	-37.6857499
-37.9648353	5.97826221	3.92680069	79.8667822	0.0446729	-37.9655111	5.9553271	3.90729788	78.1378915	0.0655685	-37.9373924
-37.6739317	5.40237716	3.48415429	67.891597	-1.212884	-55.4069988	8.212884	6.64900829	114.646996	0.61775387	-37.6827903

<u>26_N(A)</u>	<u>26_L(A)</u>	<u>26_Vol</u>
5.93847904	3.8626504	68.9357449
5.96764086	3.92700291	78.9789421
5.93594646	3.85897781	68.7555615
5.92982521	3.88369497	74.9190856
5.96848341	3.92068428	78.9768242
5.93469603	3.84844549	68.9643717
5.90840306	3.82603156	67.2707589
5.38278186	3.46184425	66.7764369
5.9344115	3.85767085	68.7322006
5.38224613	3.46093813	66.9012171

ii Tablas de datos para el análisis de descriptores por enlace

MOLECULA	A_RHO	A_LAP	A_ELEP	A_V	A_G	B_RHO	B_LAP	B_ELEP	B_V	B_G
10	0.272222038	-0.293172751	0.137009356	-0.732818727	0.329762770	0.272256326	-0.291139139	0.136670759	-0.733440421	0.330327818
17	0.269394422	-0.281258449	0.141964198	-0.723726351	0.326705869	0.275852587	-0.292745949	0.142213555	-0.748433298	0.337623405
20	0.269562274	-0.281392342	0.141700185	-0.724429131	0.327040523	0.278863061	-0.286413244	0.141547830	-0.754341204	0.341368947
21	0.269200767	-0.282134421	0.141847751	-0.722649046	0.326057721	0.275953586	-0.293527298	0.141723008	-0.748664332	0.337641254
22	0.269290990	-0.282071955	0.141748747	-0.723056147	0.326269079	0.275974280	-0.293534432	0.141687442	-0.748751797	0.337684094
23	0.269781226	-0.281396710	0.141348358	-0.725360454	0.327505638	0.275844340	-0.293932880	0.141553316	-0.748082145	0.337299462
24	0.269416875	-0.281991802	0.141615952	-0.723624275	0.326563162	0.275877015	-0.294287839	0.141589516	-0.748138528	0.337283284
25	0.269639082	-0.281873494	0.141385479	-0.724619803	0.327075715	0.275941740	-0.293992202	0.141474209	-0.748489829	0.337495939
27	0.269165870	-0.281294740	0.142465851	-0.722760882	0.326218498	0.276050805	-0.291125543	0.142220692	-0.749705050	0.338461832
31	0.268958210	-0.283296561	0.142287206	-0.721327719	0.325251790	0.277093833	-0.286193485	0.141198904	-0.755386618	0.341919123

C_RHO	C_LAP	C_ELEP	C_V	C_G	D_RHO	D_LAP	D_ELEP	D_V	D_G
0.271956104	-0.323341358	0.135721826	-0.723547430	0.321356045	0.272780416	-0.725589116	0.039655752	-0.301278843	0.059940782
0.276813356	-0.287300465	0.141469152	-0.753921874	0.341048379	0.271775195	-0.728434037	0.056058104	-0.292131118	0.055011304
0.275771889	-0.294069013	0.141695661	-0.747756031	0.337119389	0.271755604	-0.728226398	0.056153030	-0.292004505	0.054973953
0.277388060	-0.285842605	0.140822465	-0.756749342	0.342644345	0.271728087	-0.728072393	0.055763911	-0.291784006	0.054882954
0.277350826	-0.286029234	0.140768544	-0.756538678	0.342515685	0.271722704	-0.728021523	0.055725050	-0.291736538	0.054865679
0.277083801	-0.286603097	0.140918400	-0.755250840	0.341800033	0.271745298	-0.728118591	0.055775135	-0.291727101	0.054848727
0.277398405	-0.285555131	0.140661496	-0.756855111	0.342733164	0.271716248	-0.727956756	0.055739109	-0.291676914	0.054843862
0.277282491	-0.286190368	0.140607977	-0.756195728	0.342324068	0.271706319	-0.727867406	0.055627325	-0.291578318	0.054805733
0.276038819	-0.290694710	0.141978292	-0.749745107	0.338535715	0.271957791	-0.729216512	0.058057411	-0.293027192	0.055381532
0.275869463	-0.294228398	0.141657507	-0.748133297	0.337288099	0.271886625	-0.728850298	0.057559807	-0.292530602	0.055184014

E_RHO	E_LAP	E_ELEP	E_V	E_G	F_RHO	F_LAP	F_ELEP	F_V	F_G
0.311810219	-0.838193793	0.229721007	-0.413872830	0.102162191	0.309771995	-0.840938882	0.233095096	-0.403344831	0.096555058
0.317468258	-0.865039928	0.259791266	-0.429170085	0.106455052	0.307004261	-0.832985591	0.209076051	-0.396239579	0.093996591
0.317369329	-0.865080643	0.257152150	-0.428666804	0.106200822	0.307525077	-0.835816883	0.209165676	-0.397525248	0.094285516
0.317470481	-0.865243716	0.258978283	-0.429025016	0.106357044	0.307322120	-0.834452055	0.210288302	-0.397159848	0.094273417
0.317448721	-0.865110526	0.258973746	-0.428971536	0.106346952	0.307312982	-0.834423194	0.210250062	-0.397161651	0.094277926
0.317387167	-0.864994400	0.257580952	-0.428759065	0.106255233	0.307273479	-0.834391269	0.209408405	-0.397268395	0.094335289
0.317363992	-0.864844879	0.258056923	-0.428652622	0.106220701	0.307560537	-0.835689216	0.210614134	-0.397839779	0.094458738
0.317365059	-0.864723219	0.258502510	-0.428708189	0.106263692	0.307358676	-0.834662539	0.210357788	-0.397399419	0.094366892
0.317833158	-0.867178823	0.259311861	-0.429810282	0.106507788	0.307160831	-0.832616406	0.210363877	-0.396647914	0.094246906
0.317796115	-0.866912566	0.260063662	-0.429644479	0.106458169	0.307314925	-0.833023103	0.212334039	-0.397035117	0.094389671

G_RHO	G_LAP	G_ELEP	G_V	G_G	H_RHO	H_LAP	H_ELEP	H_V	H_G
0.307786697	-0.832318682	0.227148974	-0.400428187	0.096174258	0.312936073	-0.852711343	0.222888792	-0.414863401	0.100842782
0.307549084	-0.834692323	0.216141754	-0.402740622	0.097033771	0.315164300	-0.864091078	0.240658910	-0.418906070	0.101441650
0.307987739	-0.836440564	0.217322536	-0.403687836	0.097288848	0.315058835	-0.863799372	0.239481088	-0.418522127	0.101286142
0.307707836	-0.835490039	0.218853590	-0.403336016	0.097231753	0.315021011	-0.863512226	0.239694092	-0.418442651	0.101282297
0.307667012	-0.835257738	0.216812939	-0.403227597	0.097206581	0.315040902	-0.863584711	0.239843074	-0.418512809	0.101308316
0.307803550	-0.835820905	0.217475851	-0.403475781	0.097260278	0.315013070	-0.863507509	0.239810318	-0.418423296	0.101273209
0.307853193	-0.836181110	0.217169220	-0.403671580	0.097313141	0.315058085	-0.863802908	0.239438478	-0.418511948	0.101280610
0.307672289	-0.835239135	0.217002372	-0.403271355	0.097230786	0.315071563	-0.863759253	0.239901065	-0.418600732	0.101330460
0.308145187	-0.842532294	0.211528936	-0.407921481	0.098644204	0.314577902	-0.863098568	0.227507765	-0.416196536	0.100210947
0.308121719	-0.842711339	0.211410345	-0.408191937	0.098757051	0.314607817	-0.863294222	0.227273508	-0.416243293	0.100209869

L_RHO	L_LAP	L_ELEP	L_V	L_G	J_RHO	J_LAP	J_ELEP	J_V	J_G
0.312955229	-0.855968423	0.213195445	-0.413591599	0.099799747	0.310403091	-0.835176225	0.215971357	-0.409937499	0.100571721
0.345495390	-1.041135650	0.115188188	-0.874274459	0.306995273	0.347267476	-1.136091320	0.136803971	-0.820816520	0.268396845
0.345517659	-1.038316880	0.115820107	-0.875630904	0.308025841	0.347275077	-1.132501470	0.137179525	-0.823508642	0.270190637
0.345507599	-1.040460710	0.114485713	-0.874705107	0.307294984	0.347435330	-1.135798510	0.136369340	-0.822312119	0.269181246
0.345517767	-1.040796750	0.114515679	-0.874812549	0.307206680	0.347420566	-1.135760340	0.136274294	-0.822239779	0.269149847
0.345662391	-1.040363830	0.115270795	-0.875582940	0.307745991	0.347244570	-1.133561860	0.135999135	-0.822703853	0.269656695
0.345523788	-1.039734380	0.114632514	-0.875114325	0.307590385	0.347447537	-1.134381120	0.136422955	-0.823441285	0.269923002
0.345567412	-1.041017990	0.114598543	-0.874798612	0.307272058	0.347414220	-1.135014260	0.136030146	-0.822788631	0.269517533
0.345432767	-1.019106150	0.115314992	-0.883635223	0.314429343	0.347359873	-1.132272340	0.140513988	-0.823636876	0.270384395
0.345337707	-1.020074900	0.113995401	-0.882770433	0.313875853	0.347641130	-1.133713200	0.139929214	-0.824791283	0.270681491

K_RHO	K_LAP	K_ELEP	K_V	K_G	L_RHO	L_LAP	L_ELEP	L_V	L_G
0.293937499	-0.905290759	0.083486920	-0.610271836	0.191849573	0.305528844	-0.941603992	0.098719289	-0.673097315	0.218848158
0.303320332	-0.933999011	0.077565984	-0.664907840	0.215704044	0.299126097	-0.923277459	0.087536807	-0.639723377	0.204452006
0.301672085	-0.932892871	0.076231581	-0.653310985	0.210043884	0.301093692	-0.933224051	0.089722012	-0.646905800	0.206799894
0.302439079	-0.929371885	0.076451132	-0.662019878	0.214838453	0.299611551	-0.923646670	0.088131162	-0.643394819	0.206241576
0.302498018	-0.929623534	0.076499892	-0.662244161	0.214919139	0.299514207	-0.923682878	0.087966466	-0.642546363	0.205812822
0.301944823	-0.926061172	0.075884484	-0.661067297	0.214776002	0.300763899	-0.929323872	0.090566424	-0.647857657	0.207763344
0.301838709	-0.927074282	0.075884708	-0.659468984	0.213850207	0.300497172	-0.927407400	0.089235965	-0.647381231	0.207764690
0.302231295	-0.927970790	0.076123881	-0.661559668	0.214783485	0.299653172	-0.924899326	0.088232821	-0.642578685	0.205876927
0.295549537	-0.877092183	0.057029972	-0.652675498	0.216701226	0.314109022	-1.000018900	0.153636888	-0.706573887	0.228284581
0.295688408	-0.875934114	0.056598816	-0.654435310	0.217725891	0.313416383	-0.995699224	0.152237621	-0.704678877	0.227877035

M_RHO	M_LAP	M_ELEP	M_V	M_G	N_RHO	N_LAP	N_ELEP	N_V	N_G
0.303842400	-0.815969582	0.216149285	-0.391942370	0.093974987	0.317256270	-0.874195885	0.233656472	-0.427023560	0.104237294
0.306162810	-0.825953212	0.225376079	-0.397972554	0.095742126	0.315890203	-0.868557048	0.226324709	-0.422632161	0.102746449
0.305729238	-0.824392401	0.222800451	-0.396770223	0.095336062	0.316357609	-0.870848434	0.227249109	-0.423936876	0.103112384
0.306056508	-0.825804925	0.224638829	-0.397932321	0.095740545	0.316059265	-0.869537120	0.225952141	-0.423037546	0.102826633
0.306114271	-0.826043954	0.224801839	-0.398053961	0.095771487	0.316022012	-0.869347260	0.225868742	-0.422933407	0.102798296
0.305725429	-0.824514343	0.222713735	-0.396892584	0.095381999	0.316567472	-0.872237430	0.226363116	-0.424391483	0.103168063
0.305916906	-0.825584705	0.223239386	-0.397509077	0.095556450	0.316377322	-0.871206843	0.226166544	-0.423881761	0.103040025
0.306132737	-0.826236064	0.224541690	-0.398062451	0.095751718	0.316104648	-0.869827237	0.225711706	-0.423130145	0.102836668
0.348509305	-1.232740260	0.146327003	-0.752872108	0.222343622	0.345088414	-0.986473548	0.091300485	-0.896221965	0.324801789
0.313416383	-0.995699224	0.152237621	-0.704678877	0.227877035	0.348933637	-1.234491460	0.147494565	-0.755166696	0.223271916

N_RHO	N_LAP	N_ELEP	N_V	N_G	O_RHO	O_LAP	O_ELEP	O_V	O_G
0.310329017	-0.848494740	0.198050699	-0.405374217	0.095525266	0.319179127	-0.887216267	0.225959829	-0.431797147	0.104996540
0.311617266	-0.854255333	0.198930016	-0.408787904	0.097612035	0.317907762	-0.881830449	0.218395306	-0.427432225	0.103487307
0.311335844	-0.853249019	0.197188551	-0.407910701	0.097299223	0.318395924	-0.884203228	0.220655931	-0.429029913	0.103989553
0.311541822	-0.854250096	0.197434487	-0.408450660	0.097444068	0.318229031	-0.883572625	0.218955616	-0.428389794	0.103748319
0.311548381	-0.854217491	0.197606135	-0.408479649	0.097462638	0.318164693	-0.883224781	0.218926503	-0.428212633	0.103703219
0.311364270	-0.853949185	0.195448505	-0.407786983	0.097149843	0.318816422	-0.886453289	0.220489070	-0.430139033	0.104262855
0.311361381	-0.853759162	0.195836820	-0.407832369	0.097196289	0.318682490	-0.885875850	0.220147984	-0.429742167	0.104136602
0.311497539	-0.854076445	0.197089294	-0.408288696	0.097384792	0.318277771	-0.883800169	0.219257680	-0.428554769	0.103802364
0.316643545	-0.880252699	0.221943399	-0.421430943	0.100683884	0.317320269	-0.882501479	0.203389769	-0.426185300	0.102769965
0.316651581	-0.880369999	0.221596100	-0.421406501	0.100657001	0.317326805	-0.882597102	0.202966076	-0.426136766	0.102743755

P_RHO	P_LAP	P_ELEP	P_V	P_G	Q_RHO	Q_LAP	Q_ELEP	Q_V	Q_G
0.304002855	-0.804395054	0.201007495	-0.390231540	0.094566388	0.293069433	-0.749858831	0.217998689	-0.362923292	0.087729292
0.305819902	-0.813817575	0.201792651	-0.395071660	0.095808633	0.295877784	-0.762229018	0.223850303	-0.370163795	0.089803270
0.304961685	-0.809345888	0.202160776	-0.393022661	0.095343095	0.295295081	-0.759590726	0.222962963	-0.368474083	0.089288191
0.305370605	-0.811714835	0.201205521	-0.393970139	0.095520715	0.295436088	-0.760100443	0.223659769	-0.368987320	0.089481105
0.305413058	-0.811875190	0.201446265	-0.394108084	0.095569643	0.295550940	-0.760624936	0.223792970	-0.369279400	0.089561583
0.304830485	-0.808217422	0.199267708	-0.391821236	0.094883440	0.294627410	-0.756467823	0.222315835	-0.366669939	0.088776491
0.304794946	-0.809090924	0.199972369	-0.392429007	0.095078138	0.295270194	-0.759625218	0.223341651	-0.368263957	0.089178826
0.305243276	-0.811090363	0.201132373	-0.393675801	0.095451605	0.295491203	-0.760324014	0.223809917	-0.369058440	0.089488718
0.309733485	-0.832795483	0.205393938	-0.404710338	0.098255734	0.291556011	-0.744864668	0.214242559	-0.358986018	0.086384925
0.309709196	-0.832832039	0.204577012	-0.404591240	0.098191615	0.291773036	-0.745755783	0.214912212	-0.359679895	0.086620475

R_RHO	R_LAP	R_ELEP	R_V	R_G	S_RHO	S_LAP	S_ELEP	S_V	S_G
0.271231106	-0.684621120	0.161842742	-0.310203908	0.069524314	0.416705963	0.167237119	0.118698762	-1.475185760	0.758497518
0.268814345	-0.675311469	0.144704238	-0.301419614	0.066295873	0.417462498	0.178128089	0.118950084	-1.480590760	0.762561392
0.270308517	-0.682722729	0.148598337	-0.306336274	0.067827796	0.417901486	0.185127476	0.120222007	-1.483898140	0.765090005
0.269691166	-0.679344254	0.148349918	-0.304533525	0.067348731	0.417899588	0.181819698	0.119556927	-1.483422410	0.764438666
0.269669268	-0.679343014	0.147695353	-0.304292923	0.067228565	0.417553394	0.179479092	0.118995480	-1.481317410	0.763093593
0.270380228	-0.682003972	0.155416793	-0.307525980	0.068512494	0.417492554	0.176977162	0.119960666	-1.480688940	0.762466613
0.271438908	-0.688729974	0.150592694	-0.310234995	0.069026251	0.420338053	0.212381204	0.122180437	-1.500249330	0.776672316
0.270131300	-0.681643916	0.148849362	-0.305656015	0.067622518	0.417362015	0.177790527	0.119052191	-1.480118630	0.762283131
0.270260589	-0.678406387	0.154002273	-0.305750309	0.068074356	0.417741412	0.177490417	0.120428250	-1.482102570	0.783237588
0.270132415	-0.677478554	0.154118467	-0.305324594	0.067977478	0.417164238	0.170583609	0.119781728	-1.478219650	0.760432778

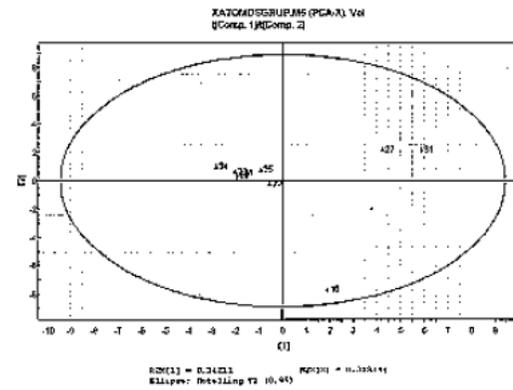
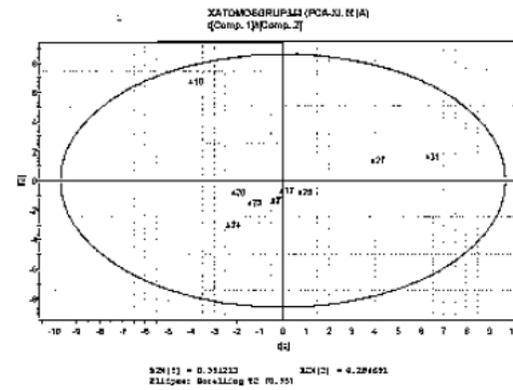
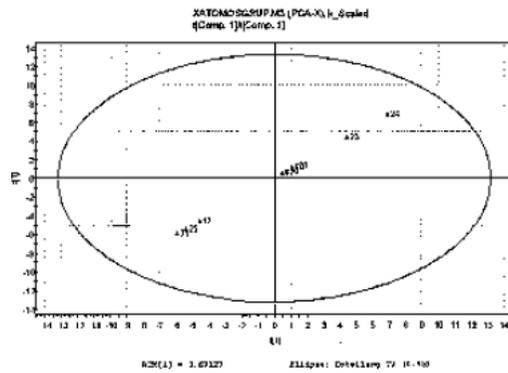
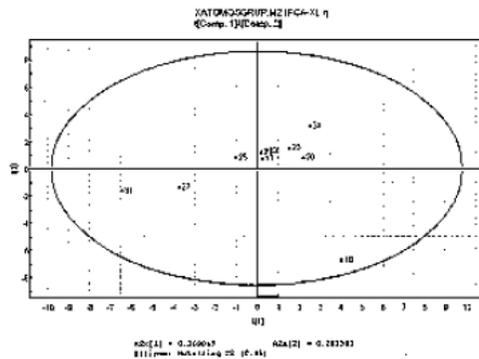
T_RHO	T_LAP	T_ELEP	T_V	T_G	U_RHO	U_LAP	U_ELEP	U_V	U_G
0.270559658	-0.498445385	0.017107188	-0.670232541	0.272810597	0.266541711	-0.319712306	0.010919027	-0.707719659	0.313895791
0.277951877	-0.508734556	0.007492899	-0.703942588	0.288379474	0.265081357	-0.314962178	0.037587121	-0.658458660	0.289858008
0.272355149	-0.507836137	0.012317708	-0.675716402	0.274378684	0.264126785	-0.321172384	0.009999023	-0.696398541	0.308052722
0.274307313	-0.509685904	0.011328044	-0.685043600	0.278811062	0.258988633	-0.323053208	0.028912859	-0.672196374	0.296716536
0.275176108	-0.510866581	0.010741069	-0.689073173	0.280678264	0.257886567	-0.313574727	0.021259596	-0.671082815	0.296344567
0.271127418	-0.496599111	0.018417713	-0.673790467	0.274820344	0.267282078	-0.316766490	0.017026891	-0.711668523	0.316238450
0.265714793	-0.525805323	0.014701760	-0.633043161	0.250820915	0.273925510	-0.370464715	0.018757512	-0.725600672	0.316492247
0.274430501	-0.509552562	0.012857402	-0.685829719	0.279220789	0.259773402	-0.307595408	0.028065022	-0.680501488	0.301801318
0.274607713	-0.505168300	0.010499036	-0.688411228	0.281059577	0.262809135	-0.305918525	0.009721445	-0.693944371	0.308732370
0.276258978	-0.507463360	0.009519452	-0.696036745	0.284585452	0.257750414	-0.294452112	0.027939037	-0.675226739	0.300806856

V_RHO	V_LAP	V_ELEP	V_V	V_G	W_RHO	W_LAP	W_ELEP	W_V	W_G
0.314109383	-0.871194659	0.256767201	-0.422974556	0.102587946	0.312740282	-0.853252067	0.217372490	-0.414615010	0.100650997
0.320427000	-0.897371362	0.297319709	-0.437364670	0.108510916	0.311809927	-0.860682971	0.248622458	-0.407528856	0.098179057
0.314876552	-0.873917962	0.258741419	-0.424114573	0.102817541	0.312866099	-0.854140397	0.217369711	-0.415022875	0.100743886
0.318144825	-0.892287579	0.288125328	-0.429569311	0.103248208	0.311852022	-0.847743073	0.216147726	-0.411341718	0.099702975
0.318314787	-0.891219607	0.269780334	-0.427883624	0.102539361	0.310870368	-0.841633312	0.223361725	-0.411077863	0.100334768
0.314110949	-0.859651139	0.313010482	-0.412778212	0.098932714	0.312390635	-0.851090671	0.254233714	-0.418915174	0.103071253
0.313577823	-0.854103675	0.313297587	-0.412848380	0.099661231	0.312883940	-0.854756679	0.249721257	-0.420090208	0.103200519
0.319450259	-0.898790117	0.281426023	-0.428830340	0.102066405	0.346103205	-1.062302170	0.143876217	-0.862902476	0.298663466
0.314734635	-0.862372326	0.315705312	-0.414856944	0.099631931	0.312622018	-0.857094383	0.241439235	-0.419845292	0.102785848
0.317951623	-0.892267393	0.245864499	-0.427301498	0.102117325	0.313011278	-0.856622569	0.206806127	-0.415745879	0.100795118

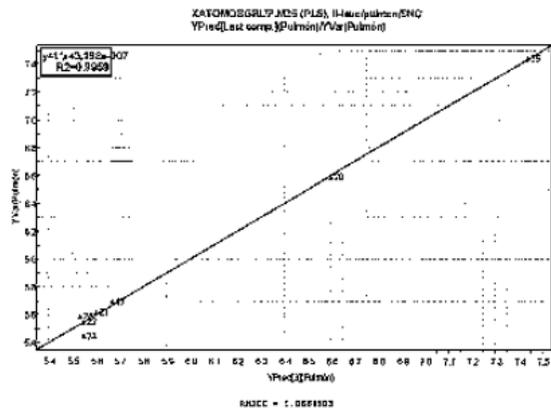
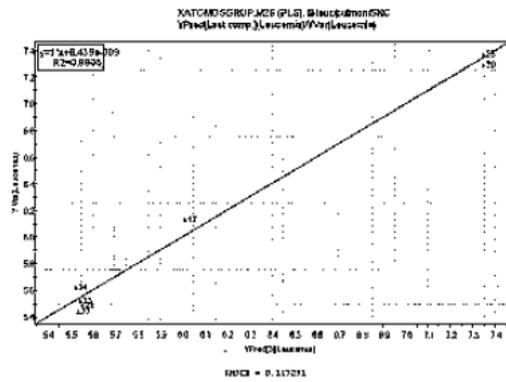
X_RHO	X_LAP	X_ELEP	X_V	X_G	Y_RHO	Y_LAP	Y_ELEP	Y_V	Y_G
0.312451866	-0.854059002	0.211836285	-0.411711428	0.099098339	0.312148972	-0.849883947	0.221348405	-0.414040130	0.100772072
0.311274223	-0.849182380	0.254731013	-0.4114088543	0.100696474	0.309943813	-0.830461516	0.221181363	-0.405245178	0.098814900
0.312399945	-0.853987045	0.210863426	-0.411484155	0.098993697	0.312314963	-0.850908916	0.220797980	-0.414328003	0.100800367
0.311841801	-0.847400841	0.220860431	-0.412690233	0.100420012	0.313163794	-0.859940620	0.237898026	-0.418898701	0.101956773
0.313967094	-0.861410771	0.242642532	-0.419288669	0.101967998	0.313283664	-0.858834313	0.238854363	-0.417316198	0.101303810
0.313694935	-0.856595369	0.251437290	-0.418809878	0.102330518	0.313812852	-0.862964252	0.237753936	-0.419298598	0.101778768
0.313743447	-0.858489530	0.248294936	-0.419459872	0.102418745	0.313726369	-0.858499316	0.247983677	-0.419407904	0.102391538
0.343131216	-1.007813920	0.124707663	-0.874777493	0.311412006	0.314843177	-0.867513866	0.223454116	-0.415594100	0.099357817
0.312270340	-0.850839576	0.219912302	-0.414139367	0.100714737	0.312410649	-0.854116496	0.210356983	-0.411454379	0.096962627
0.314825791	-0.867581909	0.222597343	-0.415461831	0.099283177	0.343238999	-1.008725230	0.123069535	-0.875113165	0.311465928

Z_RHO	Z_LAP	Z_ELEP	Z_V	Z_G	1_RHO	1_LAP	1_ELEP	1_V	1_G
0.312804655	-0.857457312	0.244229600	-0.420261228	0.102958450	0.314130726	-0.859884453	0.313498622	-0.412996589	0.099037738
0.313259867	-0.852432495	0.230156304	-0.417279750	0.102085813	0.314867808	-0.877518787	0.257975383	-0.423759441	0.102190122
0.312565332	-0.856499412	0.242036795	-0.419542623	0.102708885	0.314442971	-0.860867794	0.314517016	-0.414053846	0.099418449
0.313616080	-0.856898800	0.247788041	-0.419099491	0.102437396	0.315480861	-0.874453823	0.266855012	-0.422805818	0.102096181
0.311368304	-0.843444830	0.223812068	-0.412778722	0.100957757	0.316455231	-0.882946223	0.256635516	-0.424201745	0.101732595
0.312159520	-0.847737169	0.223407422	-0.414740790	0.101403249	0.314184339	-0.870292957	0.256906438	-0.422109210	0.102267986
0.312909396	-0.854829371	0.249548194	-0.420125961	0.103209309	0.313380218	-0.853225632	0.312042810	-0.412361920	0.099527756
0.312890630	-0.856358221	0.207062894	-0.415637598	0.100774021	0.317701673	-0.890749913	0.244711521	-0.426456431	0.101884476
0.312795496	-0.853987508	0.216854279	-0.414859139	0.100681131	0.315030761	-0.876088439	0.260031594	-0.425423111	0.103200501
0.346204648	-1.065684970	0.142187842	-0.861913490	0.297746124	0.319529252	-0.899347768	0.282073608	-0.429427863	0.102295460

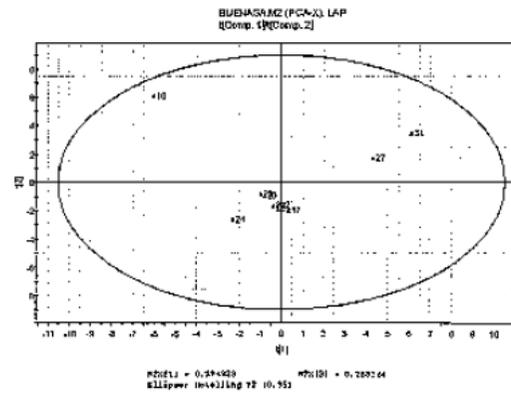
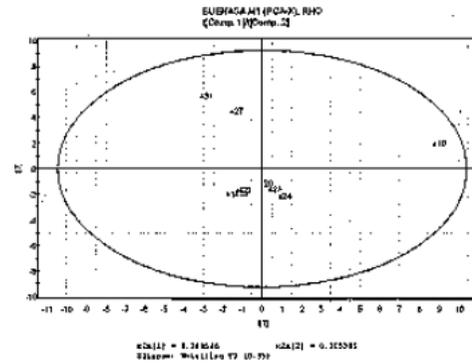
iii. Graficas de los scores para el resto de los descriptores en caso de descriptores atómicos

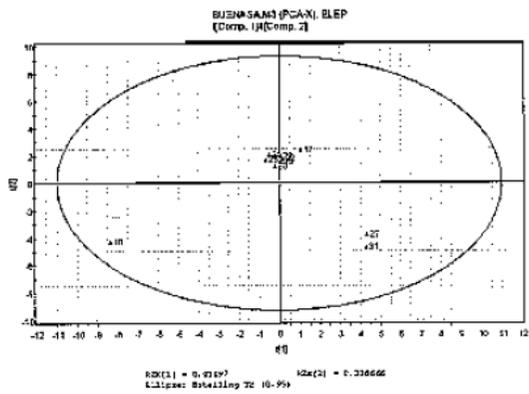


iv. Grafica de predicha contra observada de las actividades leucemia y pulmón



V. Graficas de los scores para el resto de los descriptores en caso de descriptores de enlace





Vi. Grafica de predicha contra observada de las actividades Iucemia y pulmón

