



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

La Universidad Católica de Loja

TITULACIÓN DE INGENIERO QUÍMICO

Diseño QSAR para inhibidores de acetilcolinesterasa

Trabajo de fin de titulación.

AUTORA: Vanessa del Rocío Jiménez Soto

TUTORA: Silvia González Perez, PhD.

CO-TUTOR: Piercosino Tripaldi Cappelletti, Dr.

LOJA - ECUADOR
2013

Certificación

PhD.

Silvia González

DIRECTORA DEL TRABAJO DE FIN DE TITULACIÓN

CERTIFICA:

Que el presente trabajo, denominado: **“Diseño QSAR para inhibidores de acetilcolinesterasa”** realizado por el profesional en formación: Vanessa del Rocío Jiménez Soto; cumple con los requisitos establecidos en las normas generales para la Graduación en la Universidad Técnica Particular de Loja, tanto en el aspecto de forma como de contenido, por lo cual me permito autorizar su presentación para los fines pertinentes.

Loja, junio de 2013

f)

CI:.....

Cesión de derechos

Yo Vanessa del Rocío Jiménez Soto declaro ser autor (a) del presente trabajo y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales.

Adicionalmente declaro conocer y aceptar la disposición del Art. 67 del Estatuto Orgánico de la Universidad Técnica Particular de Loja que en su parte pertinente textualmente dice: “Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través, o con el apoyo financiero, académico o institucional (operativo) de la Universidad”.

f.

Vanessa del Rocío Jiménez Soto

1104876170

Agradecimientos:

Este presente proyecto de fin de carrera quisiera agradecer primeramente a Dios por haberme bendecido y permitido cumplir una etapa en mi vida profesional.

También agradezco a mi familia, mis padres César y Rosa, por haberme apoyado y dado ánimo en la culminación de mis estudios superiores, a mis hermanas Cristina, Cecibel, Viviana y Daniela, por su apoyo incondicional en los momentos más difíciles de mi vida.

A la directora de mi proyecto de tesis, la Ph.D Silvia González, por la dedicación y el esfuerzo puesto en mí, ya que con sus conocimientos y experiencia, ha logrado que culmine con el presente proyecto. Al Dr. Tripaldi Piercosimo por toda su paciencia, y rectitud para que pueda comprender este tema, que al inicio parecía tan complicado, pero que gracias a ellos pude ver que era un campo muy interesante.

Y como no agradecer a mis amigos de universidad con quienes he compartido tantos desvelos, tristezas, alegrías ya que sin ellos no hubiera podido seguir adelante. Fueron un gran soporte a lo largo de mi vida estudiantil, Huguito Silva, Jessica Cuesta, John Montoya, Joana Bustamante, Diego Maza y Jorge Ontaneda gracias por todo. A Richard Rivera por su inmensa paciencia y apoyo incondicional dándome animo cuando no podía seguir, no se qué hubiera hecho sin ti.

Y por último agradecer a los miembros GETNANO por abrirnos paso a este mundo tan pequeño de las moléculas pero un campo inmenso por descubrir.

Dedicatoria

A mis padres, hermanas y amigos

ÍNDICE DE CONTENIDOS

Portada y contraportada	i
Certificación de cesión de derechos	ii
Certificación de revisión del tutor	iii
Agradecimientos y dedicatorias	iv
Índice	vi
Resumen	viii
Abstract	ix
Abreviaturas	x
Fin, propósito y componentes	
*Fin de proyecto	2
*Propósito de proyecto	2
*Componentes de proyecto	2
Capítulo I	
1. Introducción	4
Capítulo II	
2. Antecedentes	7
Capítulo III	
3. Marco teórico	
3.1 SAR	10
3.2 QSAR	11
3.2.1 Descriptores moleculares	11
3.2.2 Obtención del diseño QSAR	12
3.2.3 Métodos de clasificación	13
Capítulo IV	
4. Metodología	
4.1 Conjunto molecular	15
4.2 Selección de descriptores	16
4.3 Diseño molecular	16
4.4 Validación del diseño molecular	17
Capítulo V	
5. Resultados	
	vi

5.1 Base de datos	19
5.2 Diseño molecular	23
5.3 Validación del diseño molecular	28
5.4 Discusión	29
Capítulo VI	
6. Conclusiones	31
Capítulo VII	
7. Recomendaciones	32
Bibliografía	33
Artículo	36
ANEXOS	48

RESUMEN

En este trabajo se presentan los resultados del análisis cuantitativo QSAR de la correlación entre la estructura de una serie de moléculas y su actividad inhibidora de acetilcolinesterasa, sustancia que ha sido asociada con la enfermedad de Alzheimer. Para el análisis de los datos obtenidos, se han empleado diferentes métodos de clasificación multivariable, con el fin de organizar las moléculas según su estructura, y de esta forma establecer ciertas pautas que permitan diseñar un modelo, que permita predecir la actividad biológica de moléculas diferentes a las utilizadas dentro del modelo. Se seleccionaron los mejores descriptores moleculares para 3 diferentes clases de moléculas, una de ellas está formada por estructuras con grupos carbamatos, otra con compuestos con grupos piridínicos, y una última clase formada por moléculas con estructuras diversas. La fiabilidad del modelo calculado se confirmó con el método de los vecinos más cercanos *k-nn* obteniendo un error de clasificación del 5% *leave-more-out*.

ABSTRACT

This paper shows the results of QSAR quantitative analysis about the correlation between the structure of a set of molecules and their acetylcholinesterase inhibitor activity, a substance that has been associated with Alzheimer's disease.

For analyze the obtained data, some multivariate classification methods have been used, for organize different molecules according to their structure, and thus establishing some guidelines to propose a model that allows predicting the biological activities of molecules that are different to the ones used in the training set. The best molecular descriptors for three different kinds of molecules were chosen. The first one of them is formed by molecules with a diversity of structures, the second one contains pyridine groups, and the last one consists in structures containing carbamate groups. The reliability of the method is confirmed through the KNN (k-nearest neighbors) method, obtaining a classification error of 5% leave-more-out.

ABREVIATURAS

AG	Algoritmos genéticos
IC50	Inhibitory concentration 50
Knn	K nearest neighbors algorithm
NER	Non-error rate
QSAR	Quantitative Structure-Activity Relationships
SAR	Structure-Activity Relationships
NCI	National Cancer Institute
NSC	Number structure chemistry



FIN, PROPÓSITO Y COMPONENTES



Fin del proyecto

- El objetivo principal de esta investigación es diseñar un modelo computacional (QSAR) que relacione la estructura de una serie de moléculas con su actividad inhibidora de acetilcolinesterasa para incluir en el modelo, en un futuro, la estructura de moléculas prometedoras de las cuales se desconozca su actividad.

Propósito del proyecto

- Con este proyecto se espera tener un modelo que correlacione las actividades biológicas de inhibición de acetilcolinesterasa de una serie de moléculas que permita predecir la actividad de otras moléculas, cuya actividad sea desconocida.

Componentes del proyecto

- Obtener una base de datos de la actividad inhibidora de acetilcolinesterasa para una serie de moléculas.
- Elaborar de un modelo QSAR con la información obtenida.
- Validar el modelo encontrado, el cual permitirá predecir las actividades biológicas de nuevas estructuras.



INTRODUCCIÓN



CAPÍTULO I

1. INTRODUCCIÓN

Hace casi un siglo se reservó el término de enfermedad de Alzheimer (EA) a demencias seniles dadas antes de los 65 años. Actualmente se sabe que es una de las causas principales en la demencia de adultos mayores [1].

La posibilidad de padecer esta enfermedad aumenta con la edad, y se sabe que es la demencia más frecuente en la población anciana, representando un 50 al 60 % de los casos de demencia. Se calcula que en el mundo hay 22 millones de personas que la sufren y que en tres décadas habrá el doble [2].

Entre una de las consecuencias tempranas de la EA se encuentran la pérdida específica de neuronas cerebrales acompañada de una disminución de la neurotransmisión colinérgica¹, da por sustancias como la acetilcolina [1]. Clínicamente se expresa como una demencia de comienzo insidioso y que progresa lentamente, que habitualmente se inicia con fallas de la memoria reciente y termina con un paciente postrado en cama, totalmente dependiente [3].

El tratamiento farmacológico está basado en medicamentos cuyo objetivo es mejorar los defectos cognitivos y otros destinados a corregir los trastornos conductuales. Anteriormente se usaba ácido nicotínico, pero actualmente los medicamentos cuyo efecto es el mejor comprobado, son aquellos que inhiben la acetilcolinesterasa [4]. Estos medicamentos intentan prolongar la vida útil de la acetilcolina en la corteza cerebral del paciente con EA con el uso de los inhibidores mencionados, tratando de detener la pérdida de neurotransmisores [5]. Algunos de esos medicamentos son la *tacrina*, *galantamina*, *rivastigmina*, *donopezil* y *huperzina*, entre otros [6].

Varias investigaciones han demostrado que la *huperzina A* posee una fuerte actividad de inhibición de acetilcolinesterasa [7] y es considerada también como un compuesto menos tóxico [8] que los mencionados anteriormente. Cabe mencionar que las estructuras moleculares del presente proyecto están son muy similares a la estructura de la *huperzina*

¹Las neuronas se conectan a otras por medio de neurotransmisores, por ejemplo en este caso la acetilcolina



Fig 1. *Huperzia serrata* (Thunb. ex Murray) Trevis colectada de Shillong Peak. 1981.

En la actualidad existen numerosos métodos experimentales para determinar estructuras moleculares. No obstante a esto, la química computacional contribuye complementando la información obtenida experimentalmente con datos como: sitios activos, hidrofobicidad, distribución de densidad electrónica, entre otros; empleando datos como la geometría y la energía de las moléculas.

Por otra parte, hay métodos que relacionan la estructura química con la actividad biológica, asistidos por computadoras, los cuales son llamados métodos QSAR, término que proviene de su expresión en inglés: *Quantitative Structure Activity Relationship* [9]. Estos métodos se utiliza para la predicción de actividades biológicas de nuevas moléculas, con ecuaciones matemáticas que relacionan la estructura química de las moléculas con su actividad [10].

Esta técnica está basada en el uso de series de compuestos derivados de una estructura en común, en donde se observan variaciones de actividad biológica en función de la presencia de sustituyentes. Para utilizar esta metodología, el primer requisito es cuidar el diseño de exploración, y de esta manera conseguir un modelo fiable [11]. En donde los descriptores seleccionados caracterizan a la molécula cuantitativamente por medio de un procedimiento matemático que transforma la información química codificada en un número [12]. Con esta información se construye el modelo QSAR, que permite predecir actividades de moléculas no incluidas en el modelo.

En la presente investigación, y por medio del método computacional QSAR, a partir de un conjunto de moléculas se diseñó un modelo molecular capaz de relacionar su actividad inhibitoria de acetilcolinesterasa con su estructura, a través de una correlación matemática multivariable.



ANTECEDENTES



CAPÍTULO II

2. ANTECEDENTES

Las investigaciones con la metodología QSAR en la mayoría de estudios que se mencionan a continuación se refieren a actividades biológicas de diversas sustancias, sobretodo farmacológicas.

En 2011 L. Shin- Hua, y W. Josephine [9] realizaron un estudio buscando identificar varios compuestos inhibidores de acetilcolinesterasa, por medio de la metodología 3D QSAR y métodos estadísticos, con el fin de analizar diversas estructuras y elaboraron un modelo matemático con los cinco mejores descriptores, entre los cuales destacaban el número de átomos de hidrogeno y el descriptor de la hidrofobicidad. Este modelo partió de 62 estructuras moleculares, algunas de las cuales fueron obtenidas de la base de datos del NCI (*National Cancer Institute*). Con ese modelo se obtuvo una correlación de $R=0,85$ (coeficiente de correlación entre la variable dependiente e independiente), lo cual indica que es un buen modelo para localizar compuestos que tengan estructuras diferentes y determinar si son eficaces como inhibidores.

En 2010 A. Zabala, J. González, y L. Chacon [13] utilizaron la metodología QSAR para obtener un modelo por medio de algoritmos genéticos (AGs), (métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización) de una serie de 27 compuestos reconocedores de secuencias del ADN. En donde hay una correlación entre las características estructurales y la temperatura a la cual un biopolímero se desnaturaliza en un 50% (T_m) del complejo que forman el ligando ADN, en este estudio modelaron las estructuras por medio de mecánica molecular y luego optimizaron la geometría, y posteriormente obtuvieron los descriptores por medio del software DRAGON, los descriptores de helicoicidad, algunos geométricos y de electronegatividad son los que más influyeron en el modelo. Se obtuvo una capacidad predictiva de $Q^2=80,98; 80,79; 91,31^2$. Con este modelo fue posible predecir la T_m de compuestos que no habían sido incluidos en el modelo QSAR.

En 2012 D. Vitorovic-Todorovic y otros [14], presentaron un estudio 3D-QSAR basado en cierto tipo particular de descriptores independientes (GRIND-2), el cual es un paquete de descriptores específicos. Utilizando un conjunto de 110 estructuras de moléculas inhibidoras

² Q^2 es una correlación cruzada, en donde se extrae configuraciones moleculares que no son aptas para el modelo.



reversibles³ de acetilcolinesterasa, se obtuvieron tres modelos basados en tres conformaciones: 1) configuración de energías mínimas, 2) configuración más similar a la conformación del ligando co-cristalizada, 3) configuración acoplada. Independientemente de cada modelo, obtuvieron una buena capacidad de predicción. Con estadísticas de $R^2=0.9$ y $Q^2=0.7$, este estudio hizo hincapié en el beneficio de utilizar descriptores independientes.

En 2010 J. Granados, y otros, [15] realizaron un estudio 3D-QSAR en un conjunto de 40 moléculas de tipo *benzimidazol* e *imidazolina* usando el programa GOLD 3.0. El mejor modelo que obtuvieron fue a partir de descriptores de campos estéricos y electrostáticos de CoMFA (*Comparative Molecular Field Analysis*), que proporcionan un conjunto de datos estructuralmente diversos con predicciones elevadas. Los autores obtuvieron parámetros de $Q^2=0,659$ y $R^2=0,997$. Con estas características usaron el módulo de LeapFrog de SYBYL, el cual mejora la energía de enlace con relación a la precursora⁴, y con el que se pueden generar moléculas nuevas, las cuales mostraron teóricamente mejor actividad biológica que la molécula precursora.

³ Unión no covalente de un inhibidor y una enzima, la cual puede regresar a su condición inicial.

⁴ Molécula a partir de la cual se obtienen las demás.



MARCO TEÓRICO



CAPÍTULO III

3. MARCO TEÓRICO

3.1 SAR (*Structure-Activity Relationships*)

Los análisis que relacionan la estructura de una molécula con su actividad (SAR) se refieren a la relación estructura-actividad, es decir, al uso de información de la estructura molecular de las sustancias químicas para predecir características importantes en materia de persistencia, distribución, captación-absorción y toxicidad de diversas sustancias; en donde el objetivo general de este análisis es comprender la mecánica de la acción farmacológica o toxicológica, para sugerir la síntesis de nuevos compuestos con una actividad biológica definida [16].

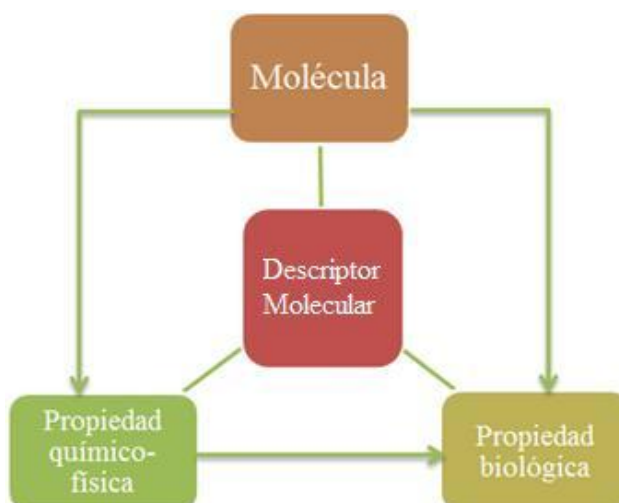


Fig. 2 Relación estructura-actividad (SAR)

El enfoque de Hanch [17] proporciona una explicación matemática mediante la siguiente ecuación:

$$\ln A = f_h(X_h) + f_e(X_e) + f_s(X_s) + \text{cte} \quad (1)$$

Donde A es la actividad y $f_h(X_h) + f_e(X_e) + f_s(X_s)$ son funciones de índices o descriptores que caracterizan a cada molécula. Esta ecuación también establece que:

- La actividad biológica es función de la estructura molecular.
- La estructura implica ciertas propiedades que pueden ser detalladas por descriptores moleculares.



- Estas propiedades pueden ser cuantificadas por diferentes parámetros.
- Siempre existirá una relación entre la actividad biológica y las propiedades moleculares.

3.2 QSAR (*Quantitative Structure-Activity Relationship*)

Hace más de cuarenta años que empezaron los estudios QSAR, y ahora tiene utilidad en un amplio campo de la química farmacéutica, de la agroquímica, de la toxicología, entre otros [16]. Esta metodología se basa en postulados que indican que la actividad de las moléculas está relacionada con: su estructura, características electrónicas, hidrofobicidad, propiedades estéricas; esta relación se refleja con valores numéricos llamados descriptores que se construyen con la ayuda de técnicas computacionales. El objetivo de QSAR es el análisis de variables independientes (descriptores) para predecir la actividad de una serie de moléculas la cual en este caso es la variable dependiente; por ello es necesario desarrollar descriptores moleculares adecuados, pues de ello depende la capacidad de predicción [18].

Fujita y Hanch realizaron un estudio de estructura actividad de reguladores de crecimiento de plantas, las cuales eran dependientes de la constante de hidrofobicidad [19, 20, 21]. En la ecuación siguiente π es la hidrofobicidad de un sustituyente.

$$\pi_x = \log P_x - \log P_H \quad (2)$$

P_x y P_H son los coeficientes de partición de un derivado de la molécula.

3.2.1 Descriptores moleculares

La clasificación más común de los descriptores moleculares de acuerdo con el enfoque de Hanch, hace referencia a tres grupos de propiedades fundamentales:

- a. Parámetros de hidrofobicidad
- b. Parámetros electrónicos
- c. Parámetros estéricos (forma y dimensión) [22].
 - **Hidrofóbicos:** Describen la solubilidad que tienen en el agua usando el coeficiente de partición [23]. Pueden provenir de medidas experimentales, de cálculos teóricos o una simple operación de suma de factores [22].
 - **Electrónicos:** Usan la constante de Hammett, constantes de campo y de resonancia. Dentro de estos descriptores electrónicos estarían el potencial de ionización,



el momento dipolar, los puentes de hidrógeno [23].

- **Estéricos:** Son descritos por la refractividad molar. Dentro de este grupo se encuentra el volumen molecular, la forma molecular, área superficial, radios de Van der Waals, etc. [23]. Se pueden dividir en descriptores monodimensionales (1D), bidimensionales (2D), tridimensionales (3D) a nivel microscópico y macroscópico.
- Los descriptores a nivel 1D son los más sencillos, son derivados directamente de la fórmula empírica de la molécula.
- En el nivel 2D se encuentran los descriptores moleculares derivados del gráfico molecular y de la matriz de conectividad, principalmente los descriptores topológicos.
- Los descriptores moleculares 3D son todos los descriptores que provienen de la estructura tridimensional, y del conocimiento de las coordenadas espaciales de la molécula [16].

Así también existen otros descriptores:

- **Generales:** Como punto de fusión, ebullición, presión, constantes de disociación, etc.
- **Químico-cuánticos:** Se obtienen mediante cálculos cuánticos como son cargas, densidades electrónicas, energía de orbitales.
- **Estructurales:** Enumeran la cantidad de fragmentos de átomos, enlaces, o subestructuras, el número de átomos de un elemento, número de anillos, entre otros [23].

3.2.2 Obtención del diseño QSAR

Para obtener un diseño deseado se debe seguir ciertas pautas:

- Primero se debe elegir un conjunto de estructuras moleculares con una actividad biológica definida y a su vez conocer esta actividad.
- Como segunda etapa está la elección de los descriptores, con los cuales se sabrá si existe una correlación con las variables seleccionadas.
- Por último, se pueden utilizar modelos discretos para clasificación o continuos para regresión, y así saber si existe una relación entre estructura y actividad. Una vez obtenido el modelo, este se puede utilizar para predecir la actividad de otros compuestos sólo con saber su estructura [24].

3.2.3 Métodos de clasificación

Dentro de los métodos de clasificación se encuentran:

- **Métodos de clasificación no paramétricos:** aquí se encuentran
 - El **análisis discriminante:** Utiliza algoritmos que buscan funciones para



combinaciones lineales (ADL) o cuadráticas (ADC), que maximizan la varianza⁵ entre categorías.

- **El análisis de redes neurales:** Es un procedimiento sistemático de datos, que es capaz de predecir con precisión variables categóricas y de escala [25].

• **Métodos de clasificación paramétricos:** Se basan en la proximidad del objeto nuevo a los objetos del conjunto de prueba. El más importante y utilizado en el presente estudio es el:

- **Método *knn*:** o método de los k vecinos, que clasifica al objeto nuevo en la categoría a la que pertenecen sus k vecinos más próximos, siendo k un valor arbitrario. Para aplicar este método se hallan las distancias entre el objeto nuevo y todos los objetos del conjunto de prueba.

Las ventajas de este método son:

- Un solo modelo puede clasificar nuevos objetos independientes del número de clases que exista. En cambio en un análisis discriminante se requiere establecer una frontera para cada categoría y así puede clasificar los objetos independientes.
- La calidad de la predicción casi no depende de la forma que tenga la categoría. Por esta razón, el método puede clasificar correctamente, aunque las categorías no sean lineales o cuadráticas y establece fronteras irregulares entre las categorías [25].

Debido al tamaño de las bases de datos, primero se debe realizar una selección de las variables más importantes para la construcción del modelo.

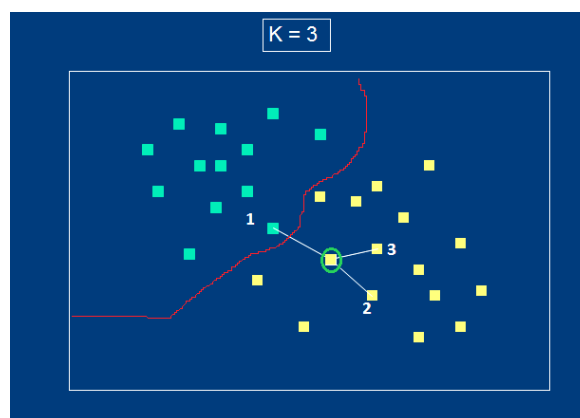


Fig. 3. Clasificación del método *knn* para 2 clases, la línea roja es la frontera entre categorías o clases. En la figura 3 se muestra la relación de distancias entre dos categorías con un punto central, estas categorías se encuentran divididas por la línea roja llamada frontera.

⁵ Es una medida de dispersión definida con la variación de una variable respecto a su media.



METODOLOGÍA



CAPÍTULO IV

4. METODOLOGÍA

Mediante la base de datos NCI (*National Cancer Institute*) se obtuvo un conjunto de estructuras moleculares con su actividad biológica respectiva, estas estructuras fueron ingresadas en el *software* DRAGON para obtener los descriptores que serán utilizados en el modelo, con estos datos y las actividades se realizó una selección de los mejores descriptores, por medio de MATLAB Y SCAN y así se obtuvo un diseño QSAR con una alta predictividad.

4.1 Conjunto molecular

Se utilizó la base de datos con su respectivo número NSC (*Number structure chemical*), de la cual se escogió un conjunto de 145 moléculas con diversas actividades de inhibición de acetilcolinesterasa, y con una estructura similar a moléculas como la *huperzina*, *licodina* y *licopodina* las cuales se caracterizan por tener una elevada actividad biológica y son utilizadas en el campo farmacológico. El formato que se muestra a continuación es el que lee el *software* DRAGON, para la estructura de una molécula.

NSC11857 4-(trimethyl-lambda~5~-azanyl)-5,6,7,8-tetrahydro-1-naphthalenyl methylcarbamate												
0.0000	0.0000	0.0000	C	0	0	0	0	0	0	0	0	0
1.1600	0.3000	-0.9600	N	0	3	0	0	0	0	0	0	0
2.5000	-0.2000	-0.4200	C	0	0	0	0	0	0	0	0	0
0.8800	-0.1800	-2.3800	C	0	0	0	0	0	0	0	0	0
1.8600	6.6400	-0.7800	C	0	0	0	0	0	0	0	0	0
0.7800	5.9000	-0.5000	O	0	0	0	0	0	0	0	0	0
0.8600	4.6000	-0.6000	C	0	0	0	0	0	0	0	0	0
-0.2400	3.8600	-1.1000	C	0	0	0	0	0	0	0	0	0
-0.1400	2.4400	-1.2000	C	0	0	0	0	0	0	0	0	0
1.0400	1.7800	-0.8400	C	0	0	0	0	0	0	0	0	0
1.8400	7.9200	-0.3600	N	0	0	0	0	0	0	0	0	0
2.9800	8.8000	-0.6200	C	0	0	0	0	0	0	0	0	0
2.8200	6.2200	-1.3800	O	0	0	0	0	0	0	0	0	0
2.0200	3.9200	-0.2200	C	0	0	0	0	0	0	0	0	0
2.1200	2.5200	-0.3400	C	0	0	0	0	0	0	0	0	0
-1.5000	4.5200	-1.5200	C	0	0	0	0	0	0	0	0	0
-2.7400	3.7800	-1.0000	C	0	0	0	0	0	0	0	0	0
-2.6400	2.2400	-1.0600	C	0	0	0	0	0	0	0	0	0
-1.3400	1.7200	-1.7200	C	0	0	0	0	0	0	0	0	0
0.3747	-0.5064	0.8650	H	0	0	0	0	0	0	0	0	0
-0.7194	-0.6202	-0.4927	H	0	0	0	0	0	0	0	0	0
-0.4638	0.9175	0.2968	H	0	0	0	0	0	0	0	0	0
3.2977	0.2500	-0.9732	H	0	0	0	0	0	0	0	0	0

Fig. 4. Geometría molecular de la NSC118574 tomada del conjunto molecular.

Con la información de las moléculas con este formato se obtuvieron los descriptores para el modelo QSAR.



4.2 Selección de descriptores

Por medio del uso del software DRAGON 6.0 [26], que es capaz de calcular 4885 descriptores moleculares y con la geometría de cada estructura molecular, se obtuvieron los descriptores que mejor caracterizaron al conjunto de moléculas, eliminando los que son constantes para no considerar información redundante. En este proyecto se utilizó la selección de algoritmos genéticos (AG), el cual consiste en un ciclo interactivo de creación de conjuntos de variables originales que generan poblaciones viables las cuales son óptimas para la construcción de un modelo predictor [27], estos conjuntos son producidos con el mayor intervalo sin error NER (Not error rate) [16].

4.3 Diseño molecular

Se realizó una clasificación del conjunto de moléculas separándolo en tres clases según su estructura: la clase 1 con estructuras que no tienen relación entre sí, la clase 2 con grupos piridínicos y clase 3 con grupos carbamatos, mediante esta clasificación se seleccionaron los mejores descriptores obtenidos de DRAGON a partir de parámetros de los algoritmos genéticos.

Se utilizó MATLAB para relacionar los descriptores y las clases, mediante un conjunto de librerías llamadas toolbox [28], en particular con el método no paramétrico *k*-vecinos o *knn* (*k-nearest neighbors*) [29] que consiste en extraer información de un conjunto de datos conocidos, y usarlos para clasificar nuevos datos o para agrupar datos existentes.

Previamente se realizó un auto-escalado a todas las variables, de donde se extrajo al azar y de forma proporcional el 20% de las moléculas, y con estos datos se construyó un subconjunto de prueba (*test set*) para validar el método con las moléculas restantes.

Tabla 1. Condiciones del algoritmo genético para la selección de los descriptores moleculares (variables independientes)

Método:	knn
Escalado	Autoscaling
Grupos de cross-validación	5
Tipo de cross-validación	Ventanas venecianas
Número de cromosomas	30
Probabilidad de mutación	0.01
Probabilidad de cruce	0.5
Repeticiones	100
Número de Evaluaciones	250
Kernel	Lineal



En la tabla 1 se muestran las condiciones para la selección de las variables más apropiadas para el modelo predictivo.

Posteriormente a aplicar las condiciones de los algoritmos genéticos, se analizó la validación de la clasificación, para ver si el NER es satisfactorio, caso contrario se deben reformular nuevamente los parámetros, hasta obtener un resultado deseado.

Para determinar qué moléculas no tienen relación con las demás se usó el programa estadístico SCAN, con el método *knn*, para poder realizar las gráficas correspondientes y analizar el porcentaje de predicción del modelo molecular y así localizar las moléculas que no pertenecen a la clase asignada.

4.4 Validación del diseño molecular

Para validar el modelo se utilizaron 2 moléculas que no fueron utilizadas para el diseño molecular, se obtuvieron las estructuras y actividades biológicas siendo ingresadas al modelo para saber en qué clase las ubica y compararlas con su respectiva actividad.



RESULTADOS Y DISCUSIONES



CAPÍTULO V

5. RESULTADOS Y DISCUSIONES

5.1 Base de datos

En la tabla 2 se presenta la lista de las moléculas con las que se realizó el diseño QSAR, cuyas actividades de inhibición de acetilcolinesterasa se obtuvo de la base de datos del NCI. La búsqueda de moléculas basó en estructuras que tengan cierta similitud con la *huperzina* (que presenta una elevada actividad biológica), la *licodina* y la *licopodina*, que son las más utilizadas en el campo farmacéutico para inhibir la acetilcolinesterasa.

Tabla 2. Nombre IUPAC y actividad inhibidora de acetilcolinesterasa de las moléculas utilizadas en el modelo computacional.

NSC	Nombre IUPAC	Actividad IC ₅₀ μM
NSC11857	4-(trimethyl-lambda~5~-azanyl)-5,6,7,8-tetrahydro-1-naphthalenyl methylcarbamate	0,82
NSC35355	1-(4-chlorobenzyl)-5-methyl-1lambda~5~-pyridin-3-yl methylcarbamate	0,74
NSC634577	Huperzine A, 6-.beta.-hydroxy-	0,83
NSC85799	6-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline	0,29
NSC296533	2,9-dimethyl-2,3,4,9-tetrahydro-1H-beta-carboline-1-one	0,28
NSC600645	1-benzyl-1,4-dihydro-4-pyridinecarboxamide	0,21
NSC406295	N~1~-(5,6-dimethoxy-8-quinolinyl)-N~6~,N~6~-diethyl-1,6-hexanediamine	0,21
NSC30782	1,3a,8-trimethyl-1,2,3,3a,8,8a-hexahydropyrrolo[2,3-b]indol-5-yl methylcarbamate	0,99
NSC167720	2-bromo-N-(1,2,3,4-tetrahydro-9-acridinyl)acetamide	0,53
NSC381408	2-ethoxy-9-acridinamine	0,53
NSC170548	3-((dimethylamino)methyl)bicyclo[2.2.1]hept-2-yl hydroxy(diphenyl)acetate	0,13
NSC91546	7-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline	0,28
NSC52463	4-amino-3-(2-phenylethyl)-1lambda~5~-pyridin-1-ol	0,12
NSC26899	1-benzyl-1,4-dihydro-3-pyridinecarboxamide	0,19
NSC82413	(E)-(1-methyl-1,4-dihydro-3-pyridinyl)(phenyl)methanone oxime	0,13
NSC167756	8-(4-oxo-4-phenylbutyl)-1-phenyl-1,3,8-triazaspiro[4.5]decan-4-one	0,17
NSC167769	3-methyl-1-(3-oxo-3-phenylpropyl)-4-phenyl-4-piperidinyl propionate	0,24
NSC19896	4-(3-(1,1-dimethyl-1lambda~5~-piperidin-3-yl)-2-propynyl)-4-methyl-1,4lambda~5~-oxazinane	0,14
NSC14726	4-((4-(dimethylamino)phenyl)imino)-3-methyl-2,5-cyclohexadien-1-one	0,15
NSC12488	4-((7-chloro-4-quinolinyl)amino)phenol	0,20
NSC11860	1,7,7-trimethyl-3-(trimethyl-lambda~5~-azanyl)bicyclo[2.2.1]hept-2-yl methylcarbamate	0,42
NSC18615 .mol	1-(3-(diethyl(methyl)-lambda~5~-azanyl)propyl)-1lambda~5~-azabicyclo[2.2.2]octane	0,38
NSC21725 .mol	3,4,5,6-tetrahydrospartein-2-one	0,13
NSC27512	2-(2-(2-pyridinyl)ethyl)-1H-isoindole-1,3(2H)-dione	0,16
NSC33096	N,N-diethyl-1-methyl-1,2,3,4-tetrahydro-3-quinolinecarboxamide	0,23
NSC33104	N,N-diethyl-1-methyl-1lambda~5~-quinoline-3-carboxamide	0,26



NSC116048	9-benzyl-3,9-diazaspiro[5.5]undecan-2-one	0,27
NSC95475	2-(3-azabicyclo[4.2.1]non-3-ylmethyl)-4-methylphenol	0,17
NSC58259	3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-one	0,20
NSC87603	3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-ol	0,19
NSC10569	1-(((6-methoxy-4-quinolinyl)methyl)amino)-2-pentanol	0,14
NSC21340	1-methyl-4-(3,4,5-trimethoxyphenyl)-4-piperidinecarbonitrile	0,29
NSC23029	N-(4-(3-(1-piperidinyl)propanoyl)phenyl)acetamide	0,28
NSC26312	1-(1,2-diethoxy-2-(1-lambda~5~-pyridin-1-yl)ethyl)-1-lambda~5~-pyridine	0,14
NSC32204	2,5-di(1-piperidinyl)benzo-1,4-quinone	0,19
NSC36773	2-methyl-1-(4-(2-methyl-1-lambda~5~-pyridin-1-yl)butyl)-1-lambda~5~-pyridine	0,15
NSC116048	9-benzyl-3,9-diazaspiro[5.5]undecan-2-one	0,27
NSC15289	N-phenyl-N-(2-quinolinyl)amine	0,18
NSC20129	4-(1-lambda~5~-pyridin-1-ylmethyl)quinoline	0,29
NSC25850	4-methyl-2-(1-piperidinyl)quinoline	0,16
NSC26195	8,9,10,11-tetrahydro-7H-pyrido[4,3-a]carbazole	0,31
NSC433	N-(4'-(dimethylamino)[1,1'-biphenyl]-4-yl)-N,N-dimethylamine	0,13
NSC2674	N~4~,N~4~-diethyl[1,1'-biphenyl]-4,4'-diamine	0,16
NSC4222	2-((3-methyl-1-lambda~5~-pyridin-1-yl)methyl)quinolone	0,19
NSC13709	4-((2-quinolinylmethylene)amino)phenol	0,16
NSC15305	benzo[a]phenazine	0,21
NSC17781	1-methyl-3-(2-(4-pyridinyl)ethyl)-1H-indole	0,23
NSC17782	3-(2-(1-methyl-1-lambda~5~-pyridin-4-yl)ethyl)-1H-indole	0,16
NSC19008	2-cyclohexyl-N-methyl-N-(2-(2-pyridinyl)ethyl)ethanamine	0,26
NSC19853	2-(4-methylphenyl)-3-quinolinamine	0,26
NSC24030	N~1~-butyl-N~4~-phenyl-1,4-benzenediamine	0,20
NSC37155	N~1~--(2-naphthyl)-1,4-benzenediamine	0,18
NSC36774	3-methyl-1-(4-(3-methyl-1-lambda~5~-pyridin-1-yl)butyl)-1-lambda~5~-pyridine	0,15
NSC36858	4-methyl-1-(4-(4-methyl-1-lambda~5~-pyridin-1-yl)butyl)-1-lambda~5~-pyridine	0,19
NSC221627	7-methyl-6,7,7a,8,9,10,10a,10b,11,12-decahydrocyclopenta[4,5]pyrimido[1,6-a]quinoline	0,21
NSC45136	2-(2-(diethyl(methyl)-lambda~5~-azanyl)ethyl)-2-methyl-2,3,3a,4,7,7a-hexahydro-1H-2lambda~5~-isoindole	0,32
NSC112670	12-methyl-2,3,4,12-tetrahydro-1H-5lambda~5~-indolo[2,3-a]quinolizine	0,35
NSC41616	No Name	0,31
NSC41615	2-(2-(diethyl(methyl)-lambda~5~-azanyl)ethyl)-2-methyl-2,3-dihydro-1H-2lambda~5~-isoindole	0,44
NSC175803	2,6-ditert-butyl-4-(trimethyl-lambda~5~-azanyl)pyridine	0,75
NSC102	1-(butoxy(phenyl)methyl)piperidine	0,14
NSC25122	4-butyl-2-(1-piperidinylmethyl)phenol	0,16
NSC56222	3-(dimethylamino)-3-phenyl-4-octanone	0,14
NSC401930	2-(dipropylamino)-2-methyl-1-phenyl-1-propanone	0,19
NSC3684	1-(1-piperidinylmethyl)-2-naphthol	0,24
NSC9542	N,N-dimethyl-2-phenoxy-2-phenylethanamine	0,17



NSC12375	2-(propylamino)-9H-fluoren-9-one	0,16
NSC12498	3-(benzylamino)-1-phenyl-1-propanone	0,22
NSC12829	1-methyl-3-phenyl-2(1H)-quinolinone	0,27
NSC15194	3a,5-dimethyl-2-phenyl-3,3a,7,7a-tetrahydro-4H-indol-4-one	0,17
NSC19224	3-cyclopentyl-7-methoxy-4-methylquinoline	0,13
NSC21065	4-(2-naphthylamino)phenol	0,14
NSC26147	4-anilino-2-tert-butylphenol	0,14
NSC58259	11H-indeno[1,2-b]quinolin-11-one	0,24
NSC125581	1-benzyl-4,5,6,7,8,8a-hexahydro-4(1H)-quinolinone	0,25
NSC140638	2-benzyl-1,3,4,7,8,8a-hexahydro-6(2H)-isoquinolinone	0,30
NSC145681	2-(4-(trimethyl-lambda~5~-azanyl)benzylidene)cyclohexanone	0,64
NSC3329	1-phenyl-2-(2-quinolinyl)ethanone	0,26
NSC11392	2-allyl-4-tert-butyl-6-((dimethylamino)methyl)phenol	0,27
NSC17947	N,N-dibutyl-3-methylbenzamide	0,13
NSC20076	N,N-dibutyl-4-methylbenzamide	0,14
NSC27191	N,N-dimethyl-2-((1-phenylcyclohexyl)oxy)ethanamine	0,19
NSC35523	N-(2-(2,3-dihydro-1H-inden-5-ylmethoxy)ethyl)-N,N-diethylamine	0,23
NSC52524	12-methyl-12H-benzo[a]phenoxazine	0,17
NSC111846	1-((phenylimino)methyl)-2-naphthol	0,26
NSC670854	2-(1-piperidinylmethyl)-1-naphthol	0,22
NSC369293	N-benzyl-N-methyl-2-phenoxyethanamine	0,28
NSC121201	1-(9-ethyl-9H-carbazol-2-yl)ethanone	0,23
NSC88889	benzyl 1-methyl-1H-indol-5-yl ether	0,30
NSC78614	2,3,3a,5,6,11,12,12a-octahydro-1H-4lambda~5~-cyclopenta[5,6]pyrido[2,1-a]isoquinolin-8-ol	0,22
NSC 11847	2-methyl-5-(trimethyl-5-azanyl)phenyl methylcarbamate (ACD/Name 4.0)	0,88
NSC 11859	lambda~5-(diethyl methyl)-5-azanyl)-2-methylphenyl dimethylcarbamate (ACD/Name 4.0)	0,85
NSC 15409	lambda Ammonium, (3-hydroxy-p-tolyl)dimethylethyl-, iodide, methylcarbamate	0,87
NSC 18833	4-(trimethyl-lambda~5~-azanyl)benzaldehyde	0,74
NSC 6789	N,N'-diphenylimidoforamide	0,77
NSC 7571	9-acridinamine	0,74
NSC 11846	No Name	0,74
NSC 11869	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 1-piperidinecarboxylate	0,72
NSC 11872	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl phenylcarbamate	0,79
NSC 11875	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 4-methoxyphenylcarbamate	0,75
NSC 15409	5-(ethyl(dimethyl)-lambda~5~-azanyl)-2-methylphenyl methylcarbamate	0,87
NSC 72108	1,2,3,4-tetrahydro-9-acridinamine	0,83
NSC 125821	2-(3-(hydroxy(oxido)amino)phenyl)-1,4-dihydro-2H-3,1-benzoxazine	0,81
NSC 127880	4-(trimethyl-lambda~5~-azanyl)phenyl methylcarbamate	0,96
NSC 128175	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl methylcarbamate	0,86
NSC 168373	3-pyridinyl dimethylcarbamate	0,84
NSC 5047	trimethyl(phenyl)-lambda~5~-azane	0,95



NSC 19699	2-ethylhexyl N,N,N',N'-tetramethyldiamidophosphate	0,72
NSC 17508	3-(diethylamino)phenyl methylcarbamate	0,70
NSC 33764	No Name	0,75
NSC 16853	1-naphthyl aminocarbonylcarbamate	0,49
NSC 21643	trimethyl(6-(trimethyl-lambda~5~-azanyl)hexyl)-lambda~5~-azane	0,50
NSC 38590	1,2,3,4-tetrahydrophenazine	0,46
NSC 96380	2-((2-pyridinylimino)methyl)phenol	0,42
NSC 101483	(4-((hydroxy-lambda~5~-azanylidene)methyl)-3,5-dimethylphenyl)(trimethyl)-lambda~5~-azane	0,44
NSC 108213	N-(4-chlorobenzylidene)-2-pyridinamine	0,50
NSC 114218	(2-(dimethyl(propyl)-lambda~5~-azanyl)ethyl)(dimethyl)propyl-lambda~5~-azane	0,46
NSC 12031	ethyl 4-(((4-(ethoxycarbonyl)anilino)methylene)amino)benzoate	0,53
NSC 205522	4-(((3-(hydroxy(oxido)amino)anilino)carbonyl)oxy)-1,1'-biphenyl	0,57
NSC 632382	N-(4-(dimethylamino)benzylidene)-1-adamantanamine	0,54
NSC 74685	3-methylphenyl 1-naphthylcarbamate	0,50
NSC 132951	1-naphthyl benzoylcarbamate	0,55
NSC 140786	1-(4-(trimethyl-lambda~5~-azanyl)phenyl)-1-nonen-3-one	0,55
NSC 164541	4-(benzyloxy)phenyl butylcarbamate	0,53
NSC 191428	2-naphthyl 2-methylphenylcarbamate	0,56
NSC 191448	2-naphthyl 2-methoxyphenylcarbamate	0,55
NSC 95514	3-(3-azabicyclo[4.2.1]non-3-yl)-1-(2-fluorophenyl)-1-propanone	0,60
NSC 118795	(1-benzyl-1H-indol-3-yl)(oxo)acetic acid	0,52
NSC 133584	1-benzyl-5-ethylideneazonane	0,53
NSC 204269	2-naphthyl cyclohexylcarbamate	0,56
NSC 204733	4-(((tert-butylamino)carbonyl)oxy)-1,1'-biphenyl	0,54
NSC 16856	1-naphthyl 2-cyclopenten-1-yl(methyl)carbamate	0,69
NSC 59666	2-naphthyl dipropylcarbamate	0,70
NSC 91873	(4-(1H-inden-1-ylidenemethyl)phenyl)(trimethyl)-lambda~5~-azane	0,74
NSC 100058	Galanthamine	0,79
NSC 51928	9H-fluoren-9-yl(trimethyl)-lambda~5~-azane	0,87
NSC 176000	phenyl(4-(trimethyl-lambda~5~-azanyl)phenyl)methanone	0,76
NSC 176008	2-phenyl-1-naphthalenamine	0,62
NSC 191481	2-(((propylamino)carbonyl)oxy)-1,1'-biphenyl	0,60
NSC 204566	2-benzoyl-5-methoxyphenyl methylcarbamate	0,64
NSC 527673	triethyl(4-((triethyl-lambda~5~-azanyl)methyl)benzyl)-lambda~5~-azane	0,62
NSC 605449	4a-methyl-9-(2-(1-piperidinyl)ethyl)-2,3,4,4a,9,9a-hexahydro-1H-carbazole	0,62
NSC 169029	N'-(9-anthryl)-N,N-dimethylimidoforamide	0,75
NSC 231606	No Name	0,85
NSC 215562	4-anilinophenyl methylcarbamate	0,77

En esta tabla se nombra todas las moléculas usadas en el diseño QSAR, con el valor de actividad biológica la cual está dada en proporción de IC₅₀ (µM) de cada molécula.



5.2 Diseño molecular

Se decidió usar el método *K-means* que es un algoritmo de clusterización que permite agrupar a las moléculas presentes en la base conforme a su estructura, de esta forma se dividió al conjunto molecular en: clase 1 estructuras diferentes entre sí, clase 2 con grupos piridínicos⁶ y clase 3 con grupos carbamatos⁷.

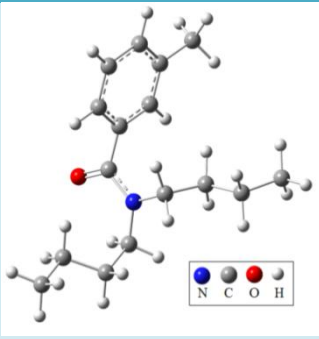
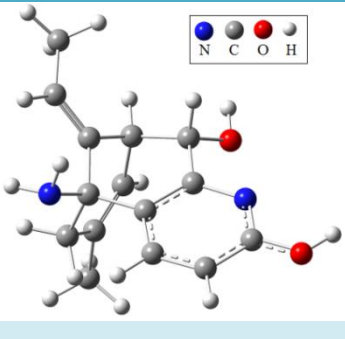
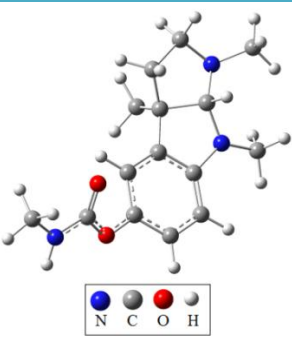
Tabla 3. Las actividades biológicas de cada clase.

Clases	# de Moléculas	Rango de actividad biológica	Media
1	92	0,128-0,945	0,31
2	14	0,122-0,832	0,41
3	39	0,169-0,986	0,66

En la tabla 3 se muestra la cantidad de moléculas que conforman cada clase y su actividad biológica, desde la mínima hasta la máxima, con esto se cuenta con una media de actividades de cada clase.

A continuación, en la tabla 4 se muestra una molécula representativa de cada clase, su nombre IUPAC y al grupo que pertenece.

Tabla 4. Representación de cada clase en la que se dividió el conjunto de moléculas, se detalla la estructura, el nombre IUPAC y la clase a la que pertenece.

Clase 1	Clase 2	Clase 3
		
<p>Ejemplo: N,N-dibutyl-3-methylbenzamide</p>	<p>Ejemplo: Huperzine A, 6-β-hydroxy</p>	<p>Ejemplo: 1,3a,8-trimethyl-1,2,3,3a,8,8a-hexahydropyrrolo[2,3-b]indol-5-yl methylcarbamate</p>

Los descriptores necesarios para el modelo computacional se obtuvieron partir de las geometrías

⁶Proveniente de la piridina, el cual es un compuesto aromático heterocíclico y su estructura es la del benceno con la única diferencia de un CH por un N.

⁷Los carbamatos son compuestos orgánicos derivados de la estructura del ácido carbámico (R₁-O-CO-NR₃-R₂) Son grupos funcionales que se encuentran interrelacionados estructuralmente y pueden ser inter-convertidos químicamente.



utilizando el *software* DRAGON, y con MATLAB se calcularon las relaciones entre la estructura de las moléculas y sus respectivas actividades.

Se obtuvo un modelo con un NER elevado por lo tanto con una predictividad elevada, por medio de método *knn* y la aplicación de algoritmos genéticos. Los resultados que se obtuvieron fueron muy favorables, un error del 5% *leave-one-out* (moléculas tomadas de una en una al azar), y una selección de 12 variables.

En el gráfico de la figura 5, se presenta el número de descriptores que usó el modelo contra el NER cross-validado internamente con 5 grupos de extracciones, obtenido en MATLAB.

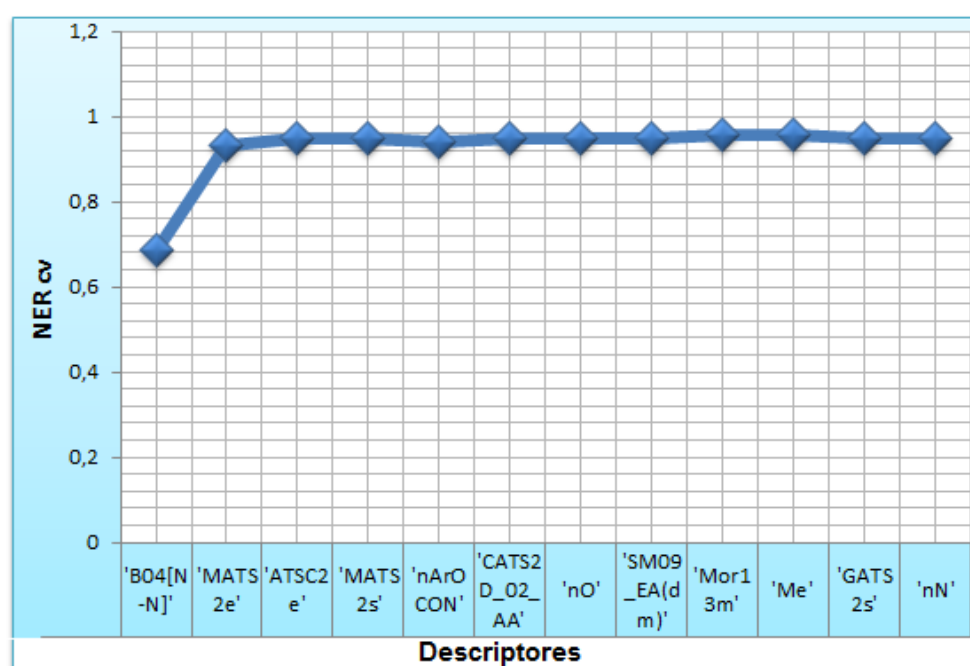


Fig. 5 Cambio del NER conforme aumentan los descriptores

La gráfica anterior muestra que el NER es 0,95 con los 12 primeros descriptores, así que se puede decir que incrementar el número de descriptores no proporciona una mejor predicción.

La gráfica que se presenta muestra el *test* interno de casi el 80% de las moléculas, con el cual se obtuvo la predicción del 0,95 *live-one-out*.

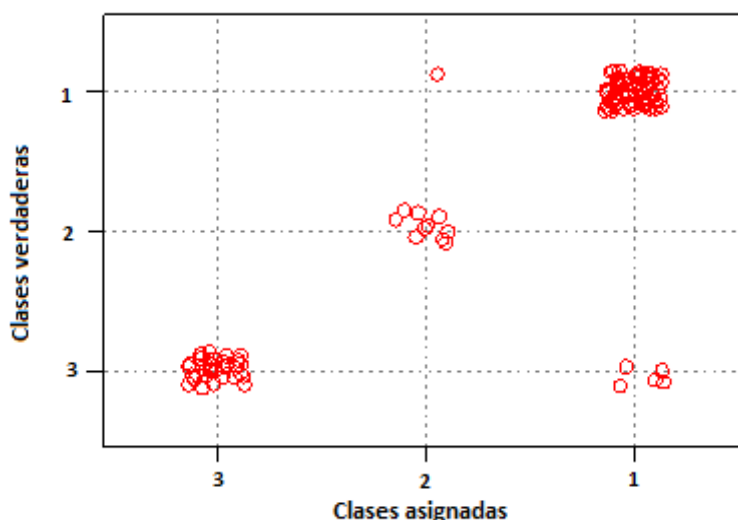


Fig. 6. Clasificación de las clases verdadera contra las asignadas.

En la figura 6 se muestra cuáles son las moléculas que pertenecen a la clasificación asignada y cuáles son las que no se ajustan a la predicción, de esta manera se analizó a las moléculas que no se ajustaron a la clase asignada.

Mediante la matriz de confusión (Tabla 5) se puede saber la relación de las moléculas con las clases asignadas del modelo y determinar el NER.

Tabla 5. Matriz de confusión con el método *knn*.

		clase 1	clase 2	clase 3
clase 1	69	61	2	6
clase 2	11	5	5	1
clase 3	35	4	4	27
	115	70	11	34

Con la ayuda de MATLAB se obtuvo la relación que existía entre clases y moléculas, de las cuales las que se encuentran en la fila diagonal están bien clasificadas, y las que se encuentran fuera de esta diagonal son los que deberían pertenecer a una clase diferente, teniendo un NER del 0,80 de predicción.

También se calcularon algunos parámetros de calidad con el mismo *software* como fueron de sensibilidad y la especificidad, para obtener el NER de las clases.

Tabla 6. Parámetros de calidad con la clasificación del método *knn*.

Clase	1	2	3
Sensibilidad	0,88	0,45	0,77
Especificidad	0,80	0,94	0,91
Total	0,84	0,70	0,84
	NER		0,79



Estos parámetros de calidad indican que el NER es bastante bueno y que a la vez hay una deficiencia en la sensibilidad de la clase 2, ya que estos calores son bajos, por lo que se puede decir que las moléculas de esta clases no se encuentran bien clasificadas, mientras que las otras dos clases muestran una buena especificidad y sensibilidad.

La figura 7 muestra la relación que existe entre clases, en ella se puede diferenciar un grupo de otro con claridad.

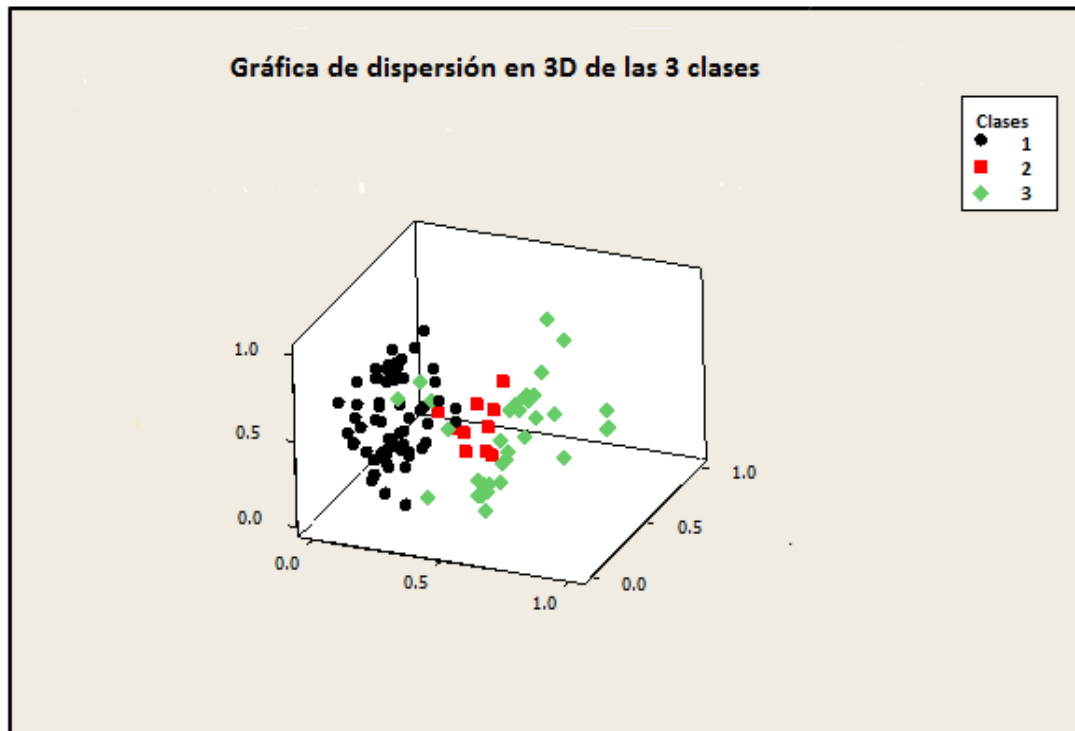


Fig. 7. Gráfica de dispersión de cada agrupación.

Con esta gráfica tridimensional se puede distinguir la relación de las moléculas de cada clase que conforma el modelo computacional.

Mediante el gráfico de los centroides (obtenido de SCAN) de las tres clases se puede visualizar los valores de los descriptores de cada clase y así saber cuál será su influencia en la actividad de una nueva molécula.

Los descriptores que forman el modelo son los que se presentan en la siguiente tabla:



Tabla 7. Descriptores que caracterizan la actividad del conjunto de moléculas.

Descriptores	Descripción
'B04[N-N]'	Descriptor bidimensional, indican la ausencia o presencia de enlaces N-N a una distancia topológica de 4 átomos.
'MATS2e'	Auto correlación bidimensional calculada según Moran, con respecto a la electronegatividad de Sanderson. A una distancia de cada dos enlaces habrá una situación de electronegatividad.
'ATSC2e'	Autocorrelación centrada, con referencia al centro de la molécula. Habrá una distribución de electronegatividad cada 2 átomos.
'MATS2s'	Autocorrelación bidimensional de Moran. Describe un estado de ionización cada dos enlaces.
'nArOCON'	Número de tio-carbamatos aromáticos y grupos funcionales.
'CATS2D_02_AA'	Cada 2 átomos habrá una situación de un átomo aceptor- aceptor.
'nO'	Número de oxígenos en la molécula.
'SM09_EA(dm)'	Matriz de orden de adyacencia de enlaces pesados para un momento dipolar.
'Mor13m'	Descriptor Morse tridimensional referente a la masa.
'Me'	Electronegatividad promedio con respecto al Carbono
'GATS2s'	Autocorrelación Geary bidimensional referente a estados de ionización
'nN'	Número de nitrógenos en la molécula.

Estos descriptores seleccionados fueron los que mejor caracterizaron a cada clase, y mejoraron la predictividad del modelo.

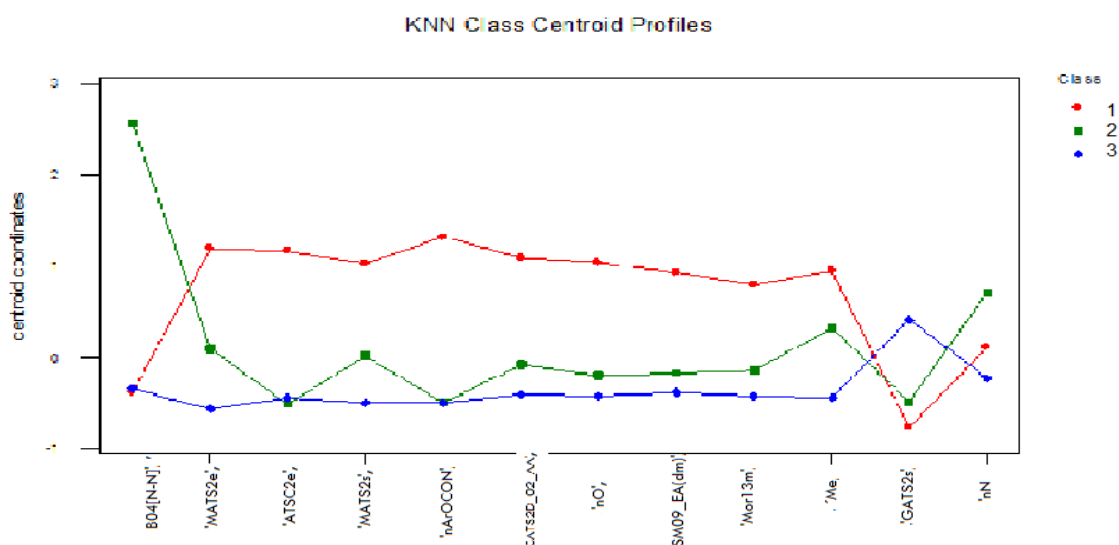


Fig. 8. Representación de valores de los descriptores según su clase.

En la figura 8 se muestran los descriptores que definen al modelo. En este caso el MATS2s, Me, GATS2s y nN son los que mejor definen a cada clase, y se puede ver que existe una buena



clasificación entre la clase 1 y 3, en cambio la clase 2 tiene valores similares de los descriptores de las otras dos clases.

Finalmente el modelo fue construido con 12 descriptores los cuales fueron aplicados al *test set* externo que fue aproximadamente del 20% de moléculas. Se presenta a continuación la matriz de confusión de este *test set*:

Tabla 8. Matriz del *test-set* externo

	No	Clase 1	Clase 2	Clase 3
Clase 1	19	18	0	1
Clase 2	3	2	1	0
Clase 3	8	3	0	5
Suma total	30	23	1	6

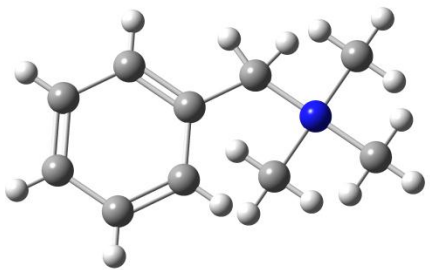
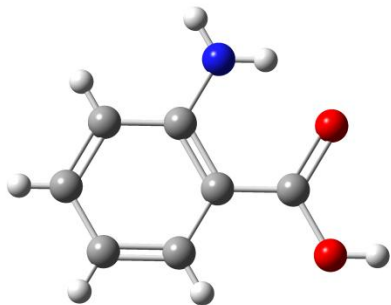
La matriz del *test set* externo proporciona un *NER* del 0.80 de predicción, estas fueron la cantidad de las moléculas que se extrajeron del conjunto molecular de cada clase.

El error del modelo es del 5% incluido el *test externo* (este test incluye las moléculas previamente excluidas).

5.3 Validación del diseño molecular

Las moléculas que fueron utilizadas para validar el modelo fueron las siguientes:

Tabla 9. Estructura, nombre IUPAC y actividad biológica de las moléculas usadas en la validación.

Estructura		
Nombre	benzyl(trimethyl)-lambda~5~-azane	2-aminobenzoic acid
Actividad	0,391	0,135

Las moléculas se ingresaron en el modelo y éste las ubicó en la clase 1 que son estructuras diferentes entre sí con una actividad inhibidora media de 0,31.



5.4 Discusiones

En este modelo QSAR obtenido se ha utilizado una base de datos muy amplia, la cual presentó un buen poder predictivo, con un NER interno relativamente elevado, y un NER externo bajo debido al pequeño número de moléculas utilizadas.

En un comienzo se trató de relacionar directamente la actividad biológica con su estructura, para que este modelo nos dé un valor exacto de actividad con respecto a nuevas moléculas, pero el porcentaje de error era demasiado elevado, por lo que se tuvo que utilizar métodos de clasificación ya sea por índice de actividad o por estructura, en este caso es por estructura, ya que por índice el error también fue elevado.

El método de clasificación utilizado fue el de clusterización *K-means*, el cual clasificó por estructuras al conjunto molecular.

Para la obtención de un NER elevado se realizó una clasificación del conjunto molecular, según Cruz [30] quien señala que estos métodos son apropiados para modelar las respuestas del modelo QSAR con compuestos activo y no activos, compuesto de toxicidad baja o alta, según el estudio que se realice.

Para una buena selección de descriptores se utilizaron los algoritmos genéticos, según Todeschini [31] quien señala que estos algoritmos tienen una gran fiabilidad en la búsqueda de las mejores variables para construir modelos de clasificación y regresión.

Con la mejor clasificación y selección de descriptores se obtuvo el modelo con una predictividad elevada de 0,95, para validar este diseño molecular se utilizaron dos moléculas nuevas que no habían sido incluidas en el modelo, el cual las ubicó en la clase 1 la cual tiene una actividad inhibidora media de 0,31, dichas moléculas de prueba tenían actividad de 0,13 y 0,39, con lo que se puede decir que el modelo las ubica en una clase adecuada.

Los descriptores que mejor describen a cada clase fueron MATS2s; el cual describe que habrá un estado de ionización cada dos enlaces, es decir habrá una situación de enlaces iónicos, Me; electronegatividad promedio del carbono con respecto a los átomos que se encuentran a su alrededor; GATS2s; describe con qué frecuencia habrán estados de ionización en cada molécula, mediante la constante de Geary y nN; es el número de nitrógenos que presentan las moléculas.

El motivo de no haber utilizado moléculas con actividades biológicas media y alta para la validación es porque todas las moléculas de la base de datos se usaron dentro del modelo.

Para finalizar se puede destacar un punto interesante que se desprende de este estudio radica en que las moléculas con más actividad para inhibir la acetilcolinesterasa tienen, en su mayoría, compuestos aromáticos y nitrógeno en su estructura; por otra parte, la actividad biológica es menor para aquellas moléculas con gran número de sustituyentes y con radicales grandes.



CONCLUSIONES, RECOMENDACIONES Y BIBLIOGRAFÍA



CAPÍTULO VI

6. CONCLUSIONES:

- Los resultados obtenidos contienen un error del 5%, lo que significa que presentan un factor de correlación de NER de 0.95. Conforme a la información obtenida se puede afirmar que los mejores descriptores moleculares de la actividad de inhibición de la acetilcolinesterasa de las moléculas incluidas en el modelo son: 'B04[N-N]', 'MATS2e', 'ATSC2e', 'MATS2s', 'nArOCON', 'CATS2D_02_AA', 'nO', 'SM09_EA(dm)', 'Mor13m', 'GATS2s', 'nN'.
- Los descriptores que mejor representan al modelo son los 'MATS2s', 'Me', 'GATS2s', 'nN', ya que diferencian de una mejor forma a las tres clases.
- El modelo molecular se realizó con el método *knn* y con la clusterización de *k-means*, se separó la base de las 145 moléculas en tres clases diferentes, según su estructura: una que incluye moléculas con grupos carbamatos, otra con moléculas que contienen grupos piridínicos, y la última con estructuras diferentes entre sí.
- El modelo es muy efectivo para determinar la actividad biológica de moléculas que tengan cierto parecido a la clase 1 y 3, que son las clases que el modelo describe mejor.
- Los centroides fueron de gran utilidad para saber qué descriptores son los que diferencian a cada clase, y también identificar qué clase es la que presenta mayor dificultad de clasificación.
- El diseño molecular se validó con dos moléculas que no se encontraban en el conjunto de moléculas con las que se elaboró el modelo, el cual las ubica en la clase 1 que es a la que pertenecen según su actividad biológica.



CAPÍTULO VII

8. RECOMENDACIONES

- Para un estudio más a fondo de las moléculas con actividad inhibidora de acetilcolinesterasa, se debería crear un conjunto molecular más específico según la familia o la especie a la que pertenecen los compuestos
- Para mejorar la predicción del modelo se puede considerar una base de datos mucho más grande que la actual.
- Probar nuevos métodos de clasificación apoyados en otros algoritmos para ver si mejora la capacidad predictiva del modelo.
- Se recomienda utilizar mayor cantidad de moléculas con estructuras similares entre sí, para incrementar el NER o la predictividad.
- Comprobar experimentalmente la predictividad del modelo.

**BIBLIOGRAFIA**

- [1] Donoso, A.; Quiroz, M.; Yulis, J. Demencia de tipo Alzheimer, experiencia clínica en 71 casos. *Rev. Med. Chile*. **1990**; 118, 139-145.
- [2] Eisendrath, S.; Lichtmacher, J. Trastornos psiquiátricos. Delirio, demencia y otros trastornos cognitivos. En: Tierney L, McPhee S, Papadakis M. Diagnóstico clínico y tratamiento. 41ª edición. México: Editorial El Manual Moderno, **2006**, 955-956
- [3] Donoso, A.; La enfermedad de Alzheimer y otras demencias .3ª ed. Santiago, *Ed. Universitaria*, **2001**.
- [4] Grundman, M.; Thal, L. Treatment of Alzheimer's disease: rationale and strategies. *Neurol Clin N Am*, **2000**, 18, 807-27.
- [5] DeKosky, S.; Ikonovic, M.; Styren, S.D.; Beckett, L.; Wisniewski, S.; Bennett, D.A. Up regulation of choline acetyltransferase activity in hippocampus and frontal cortex of elderly subjects with mild cognitive impairment. *Ann Neurol*, **2002**, 51, 45-55
- [6] Trujillo, J.; Correa, J.; Espinosa, J.; Garcia, J. Solvent-free Synthesis of Arylamides and Arylimides, Analogues of Acetylcholine. *Syn. Commu.* **2005**, 35, 1-7.
- [7] Ma, L.; Wu, F. Chinese Traditional drug to enhance memory-Qiancengta. *Plants.*, **2000**, 3-151.
- [8] Wang, H.; Tang, X. Anticholinesterase effects of huperzine A, tacrine in rats. *Zhongguo Yao Li Xue Bao*, **1998**, 19, 27-30.
- [9] Lu, J.; Wu, W.; Liu, H.; Zhao, J.; Liu, K.; Chuang, C.; Lin, H.; Tsai, W.; Ho, Y. The discovery of potential acetylcholinesterase inhibitors: A combination of pharmacophore modeling, virtual screening, and molecular docking studies. *Journal of Biomedical Science*, **2011**, 12-8.
- [10] Duardo, A.; Patlewicz, G.; Lopez, A. Current topics on software use in medicinal chemistry: Intellectual property, taxes, and regulatory issues. *Curr. Top. Med. Chem.* **2008**, 8, 1666-1675.
- [11] Pastor, M.; Alvarez, J. Técnica QSAR en diseño de fármacos. Departamento de química. *Universidad de Alcalá*
- [12] Todeschini, R.; Consonni, V. Manual de descriptores moleculares, Wiley VCH. **2000**, <http://www.moleculardescriptors.eu/books/handbook.htm>
- [13] Zabala, A.; González, J.; Chacon, L. Estudio QSAR por algoritmos genéticos de reconocedores de surco del DNA. *Revista Biológicas*, Diciembre **2010**, 12(2), 108 – 115
- [14] Vitorovic, D.; Vvijetic, N.; J uranic, O.; Drakulic. J. The 3D-QSAR study of 110 diverse, dual binding, acetylcholinesterase inhibitors based on alignment independent descriptors (GRIND-2). The effects of conformation on predictive power and interpretability of the models. *Journal of Molecular Graphics and Modelling*, Agosto, **2012**. JMG 6198



- [15] Granados, J.; Arias, E.; Pacheco, D.;Valdiris, V.; Vivas, R. Acoplamiento Molecular, 3D-QSAR Y Diseño De Novo De Benzimidazoles E Imidazolinas Derivados De (S)-Isotiazolidinonas Como Inhibidores De La Proteína Ptp 1b. Vitae, *Revista De La Facultad De Química Farmacéutica*, **2010**,17-3
- [16] Todeschini, R. Introducción a la Quimiométrica, Edises, Universidad Milano Bicocca, **2003**, 80-220.
- [17] Escalona, J.; Carrasco, R. Introducción al diseño de Fármacos. Padrón Folleto para la docencia de la asignatura de Farmacia, *Universidad de Oriente*
- [18] Livingstone, D. *J. Chem. Inf. Comput. Sci.*, **2000**, 40, 195.
- [19] Hansch, C.; Maloney, P.; Fujita, T.; Muir, R. *Nature*, **1962**,194, 178.
- [20] Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.*, **1964**,86, 5175.
- [21] Hansch, C.; Leo, A.; Heller, R. Ed., Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, *American Chemical Society*, Washington, DC, **1995**.
- [22] Cardamone, L. Chemistry, Toxicology and QSAR, Politécnico de Milano, Istituto di Ricerche Farmacologiche “Mario Negri”. *Chem. Inf. Model.* **2010**, 45, 1565-1554
- [23] Andres J, Beltran J.; Química Teórica y Computacional. Universidad Jume, **2000**.
- [24] Maroto, E. Cálculos computacionales de estructuras moleculares. Alicante, **2011**.
- [25] García, M.; Ramis, G. Quimiometria. Madrid, **2001**.
- [26] Talete. DRAGON for Windows (Software for Molecular Descriptor Calculations). Version 5.4.Copyright TALETE srl, **2006**.
- [27] Frank, I.; Todeschini, R. The Data Analysis Handbook. Elsevier Science, **2006**, 334-350.
- [28] Franca, C.; Pis, R. Estudio teórico de propiedades fisicoquímicas de compuestos de interés farmacológicos. Universidad Nacional de la Plata, 134
- [29] Ballabio, D.;Todeschini, R. Multivariate Classification for Qualitative Analysis,en *Infrared Spectroscopy for Food Quality analysis and Control* Edited by Da-Wen Sun Elsevier Inc, **2009**, 94.
- [30] Cruz, J. Investigación cualitativa. México, **2009**
- [31] Todeschini, R., Introducción a la Quimiométrica, Edises, Universidad Milano Bicocca, **2003**, pp. 80-220.



ARTÍCULO



Acetylcholinesterase Inhibitors Qsar Study

Vanessa Jiménez¹, Silvia González^{1*}, Piercosimo Tripaldi² y Cristian Rojas²

* Corresponding author: sgonzalez@utpl.edu.ec

¹Departamento de Química, Universidad Técnica Particular de Loja. San Cayetano Alto. Postal Code 1101608, Loja- Ecuador.

² Laboratorio de Análisis Ambiental, Universidad del Azuay. Av. 24 de Mayo 7-77 and Hernán Malo. Postal code 0101981. Cuenca-Ecuador



Abstract:

This paper shows the results of QSAR quantitative analysis about the correlation between the structure of a set of molecules and their acetylcholinesterase inhibitor activity, a substance that has been associated with Alzheimer's disease.

For analyze the obtained data, some multivariate classification methods have been used, for organize different molecules according to their structure, and thus establishing some guidelines to propose a model that allows predicting the biological activities of molecules that are different to the ones used in the training set. The best molecular descriptors for three different kinds of molecules were chosen. The first one of them is formed by molecules with a diversity of structures, the second one contains pyridine groups, and the last one consists in structures containing carbamate groups. The reliability of the method is confirmed through the KNN (k-nearest neighbors) method, obtaining a classification error of 5% leave-more-out.

Keywords: acetylcholinesterase, Alzheimer, QSAR, molecular descriptors, KNN

Introduction:

Alzheimer's disease is a kind of dementia whose study has become recently of a particular relevance. The probability of being affected by this disease increases with the age: it may occur in

1% to 2% in people aged 60, from 3% to 5% in those 70 years, 15% to 20% in 80 years persons and one third to half in people over 85 years [1]. This neurodegenerative disease is progressive

and is characterized by a series of clinical and pathological features [2] that include memory loss. The cause has been attributed to the degradation of acetylcholine, which is a substance that acts as a neurotransmitter a to acetylcholinesterase.

One of the treatments for this affection consist in the use of cholinergic medicaments, such as the



huperzine A [3], which block the acetylcholinesterase produced by the breakdown of acetylcholine.

Now a days, there are methods that relate chemical structure with the biological activity, through the use of specialized software. They are so called QSAR methods (*Quantitative Structure Activity Relationship*) [4], used for the prediction of biological activities of molecules through the use of mathematical equations that establish a relationship between them chemical structure of a molecular with its activity [5].

This methodology was originated in the toxicology field about 40 years ago, in order to statistically relate the chemical structure of certain molecules with their toxicological activity. Currently it is widely used in agro-chemistry, design of drugs and industrial and environmental chemistry. It has become a basic tool for the design of medicines and pesticides [6].

The purpose of the present research was to design of a molecular model that allows the prediction of acetylcholinesterase inhibitory activity of molecules from their structure. In order to get the model, information obtained from the database of the National Cancer Institute of the United States (NCI database) for a set of molecules has been used. Considered data are their structure and their acetylcholinesterase inhibitory activity, which have been

obtained experimentally. The present computational model will permit predicting the activity of molecules from which there is no information available.

Models and methodology:

Materials:

In the present investigation, it has been used information of 145 molecules including carbamates and pyridine groups, similar to the *huperzine A*, *licodyne* and *lipocodyne*, which are molecules with some acetylcholinesterase inhibitory activity.

Metodology:

1. Molecular set:

From the NCI database, which contains around 127000 structures with their respective CAS number (Chemical American Society), a set of 145 molecules were chosen due their similarities with the previously mentioned ones, and their inhibitory activity.

2. Molecular descriptors:

DRAGON 6.0 is a software able of calculate 4885 different descriptors, which are a mathematical procedure that transforms the encoded chemical information, in the molecular structure, into a useful number [7]. The geometry of each molecular structure was used through the use the mentioned software in order to get the descriptors that characterize the set of molecules. Not all the descriptors were used; the constants ones were eliminated trying to eliminate redundant information.



The selection of genetic algorithms (AG) has been used, which consists in an interactive cycle of original variables set creation, which generates viable population, appropriated for the construction of a predictive model [8].

These sets are produced with the minimal NER (Not Error rate) [9].

3. QSAR model:

The set of molecules was divided into three kinds, according to their structure: a group with structures different between them, pyridine groups, and carbamate groups. The previous classification was done manually, and then corroborated with the k-means method, as implemented in the statistical software SCAN. Through this classification, and the best descriptors provided by DRAGON, a set of libraries, a toolbox [10], implemented in the MATLAB software were used; particularly, the non-parametric knn method (k-nearest neighbours) [11], which consists in extracting information of a set of known data and their use to classify new data or to group existing information. The main difficulty in this method is that if the group of data is very small, then the results will be inaccurate [12].

Previously, an auto-scaling in all the

variables was performed, extracting randomly and proportionally 20% of the molecules. With these data, a test subset was constructed in order to validate the method with the remaining molecules.

Table 1. Condition of the genetic algorithm for select the molecular descriptors (independent variables)

Method:	knn
Scaling	Auto scaling
Cross-validation	5
Cross-validation Type	Venetian Windows
Number of	30
Probability of Mutation	0.01
Probability of crossing	0.5
Repetitions	100
Number of evaluations	250
Kernel	Lineal

Table 1 shows the conditions used to select the most suitable variables for the predictive model. After this, the classification is validated through the NER. If results are not satisfactory, then the parameters must be reformulated in order to obtain a desired result.

Results and discussion:

The next table shows the 145 molecules used for the QSAR design, with their IUPAC name and their inhibitor activity index.



Table 2. Molecules used within the QSAR design with their respective biological activities

NSC	Nombre IUPAC	Actividad
NSC11857	4-(trimethyl-lambda~5~-azanyl)-5,6,7,8-tetrahydro-1-naphthalenyl methylcarbamate	0,82
NSC35355	1-(4-chlorobenzyl)-5-methyl-1lambda~5~-pyridin-3-yl methylcarbamate	0,74
NSC634577	Huperzine A, 6-.beta.-hydroxy-	0,83
NSC85799	6-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline	0,29
NSC296533	2,9-dimethyl-2,3,4,9-tetrahydro-1H-beta-carboline-1-one	0,28
NSC600645	1-benzyl-1,4-dihydro-4-pyridinecarboxamide	0,21
NSC406295	N~1~-(5,6-dimethoxy-8-quinoliny)-N~6~,N~6~-diethyl-1,6-hexanediamine	0,21
NSC30782	1,3a,8-trimethyl-1,2,3,3a,8,8a-hexahydropyrrolo[2,3-b]indol-5-yl methylcarbamate	0,99
NSC167720	2-bromo-N-(1,2,3,4-tetrahydro-9-acridinyl)acetamide	0,53
NSC381408	2-ethoxy-9-acridinamine	0,53
NSC170548	3-((dimethylamino)methyl)bicyclo[2.2.1]hept-2-yl hydroxy(diphenyl)acetate	0,13
NSC91546	7-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline	0,28
NSC52463	4-amino-3-(2-phenylethyl)-1lambda~5~-pyridin-1-ol	0,12
NSC26899	1-benzyl-1,4-dihydro-3-pyridinecarboxamide	0,19
NSC82413	(E)-(1-methyl-1,4-dihydro-3-pyridinyl)(phenyl)methanone oxime	0,13
NSC167756	8-(4-oxo-4-phenylbutyl)-1-phenyl-1,3,8-triazaspiro[4.5]decan-4-one	0,17
NSC167769	3-methyl-1-(3-oxo-3-phenylpropyl)-4-phenyl-4-piperidinyl propionate	0,24
NSC19896	4-(3-(1,1-dimethyl-1lambda~5~-piperidin-3-yl)-2-propynyl)-4-methyl-1,4lambda~5~-oxazinane	0,14
NSC14726	4-(((4-(dimethylamino)phenyl)imino)-3-methyl-2,5-cyclohexadien-1-one	0,15
NSC12488	4-((7-chloro-4-quinoliny)amino)phenol	0,20
NSC11860	1,7,7-trimethyl-3-(trimethyl-lambda~5~-azanyl)bicyclo[2.2.1]hept-2-yl methylcarbamate	0,42
NSC18615 .mol	1-(3-(diethyl(methyl)-lambda~5~-azanyl)propyl)-1lambda~5~-azabicyclo[2.2.2]octane	0,38
NSC21725 .mol	3,4,5,6-tetradehydrospartein-2-one	0,13
NSC27512	2-(2-(2-pyridinyl)ethyl)-1H-isoindole-1,3(2H)-dione	0,16
NSC33096	N,N-diethyl-1-methyl-1,2,3,4-tetrahydro-3-quinolinecarboxamide	0,23
NSC33104	N,N-diethyl-1-methyl-1lambda~5~-quinoline-3-carboxamide	0,26
NSC116048	9-benzyl-3,9-diazaspiro[5.5]undecan-2-one	0,27
NSC95475	2-(3-azabicyclo[4.2.1]non-3-ylmethyl)-4-methylphenol	0,17
NSC58259	3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-one	0,20
NSC87603	3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-ol	0,19
NSC10569	1-(((6-methoxy-4-quinoliny)methyl)amino)-2-pentanol	0,14
NSC21340	1-methyl-4-(3,4,5-trimethoxyphenyl)-4-piperidinecarbonitrile	0,29
NSC23029	N-(4-(3-(1-piperidinyl)propanoyl)phenyl)acetamide	0,28
NSC26312	1-(1,2-diethoxy-2-(1lambda~5~-pyridin-1-yl)ethyl)-1lambda~5~-pyridine	0,14
NSC32204	2,5-di(1-piperidinyl)benzo-1,4-quinone	0,19
NSC36773	2-methyl-1-(4-(2-methyl-1lambda~5~-pyridin-1-yl)butyl)-1lambda~5~-pyridine	0,15
NSC116048	9-benzyl-3,9-diazaspiro[5.5]undecan-2-one	0,27
NSC15289	N-phenyl-N-(2-quinoliny)amine	0,18



NSC20129	4-(1-lambda-5-pyridin-1-ylmethyl)quinoline	0,29
NSC25850	4-methyl-2-(1-piperidinyl)quinoline	0,16
NSC26195	8,9,10,11-tetrahydro-7H-pyrido[4,3-a]carbazole	0,31
NSC433	N-(4'-(dimethylamino)[1,1'-biphenyl]-4-yl)-N,N-dimethylamine	0,13
NSC2674	N-4~,N-4~-diethyl[1,1'-biphenyl]-4,4'-diamine	0,16
NSC4222	2-((3-methyl-1-lambda-5-pyridin-1-yl)methyl)quinoline	0,19
NSC13709	4-((2-quinolinylmethylene)amino)phenol	0,16
NSC15305	benzo[a]phenazine	0,21
NSC17781	1-methyl-3-(2-(4-pyridinyl)ethyl)-1H-indole	0,23
NSC17782	3-(2-(1-methyl-1-lambda-5-pyridin-4-yl)ethyl)-1H-indole	0,16
NSC19008	2-cyclohexyl-N-methyl-N-(2-(2-pyridinyl)ethyl)ethanamine	0,26
NSC19853	2-(4-methylphenyl)-3-quinolinamine	0,26
NSC24030	N-1~butyl-N-4~-phenyl-1,4-benzenediamine	0,20
NSC37155	N-1~(2-naphthyl)-1,4-benzenediamine	0,18
NSC36774	3-methyl-1-(4-(3-methyl-1-lambda-5-pyridin-1-yl)butyl)-1-lambda-5-pyridine	0,15
NSC36858	4-methyl-1-(4-(4-methyl-1-lambda-5-pyridin-1-yl)butyl)-1-lambda-5-pyridine	0,19
NSC221627	7-methyl-6,7,7a,8,9,10,10a,10b,11,12-decahydrocyclopenta[4,5]pyrimido[1,6-a]quinoline	0,21
NSC45136	2-(2-(diethyl(methyl)-lambda-5-azanyl)ethyl)-2-methyl-2,3,3a,4,7,7a-hexahydro-1H-2lambda-5-isoindole	0,32
NSC112670	12-methyl-2,3,4,12-tetrahydro-1H-5lambda-5-indolo[2,3-a]quinolizine	0,35
NSC41616	No Name	0,31
NSC41615	2-(2-(diethyl(methyl)-lambda-5-azanyl)ethyl)-2-methyl-2,3-dihydro-1H-2lambda-5-isoindole	0,44
NSC175803	2,6-ditert-butyl-4-(trimethyl-lambda-5-azanyl)pyridine	0,75
NSC102	1-(butoxy(phenyl)methyl)piperidine	0,14
NSC25122	4-butyl-2-(1-piperidinylmethyl)phenol	0,16
NSC56222	3-(dimethylamino)-3-phenyl-4-octanone	0,14
NSC401930	2-(dipropylamino)-2-methyl-1-phenyl-1-propanone	0,19
NSC3684	1-(1-piperidinylmethyl)-2-naphthol	0,24
NSC9542	N,N-dimethyl-2-phenoxy-2-phenylethanamine	0,17
NSC12375	2-(propylamino)-9H-fluoren-9-one	0,16
NSC12498	3-(benzylamino)-1-phenyl-1-propanone	0,22
NSC12829	1-methyl-3-phenyl-2(1H)-quinolinone	0,27
NSC15194	3a,5-dimethyl-2-phenyl-3,3a,7,7a-tetrahydro-4H-indol-4-one	0,17
NSC19224	3-cyclopentyl-7-methoxy-4-methylquinoline	0,13
NSC21065	4-(2-naphthylamino)phenol	0,14
NSC26147	4-anilino-2-tert-butylphenol	0,14
NSC58259	11H-indeno[1,2-b]quinolin-11-one	0,24
NSC125581	1-benzyl-2,3,4,7,8,8a-hexahydro-4(1H)-quinolinone	0,25
NSC140638	2-benzyl-1,3,4,7,8,8a-hexahydro-6(2H)-isoquinolinone	0,30
NSC145681	2-(4-(trimethyl-lambda-5-azanyl)benzylidene)cyclohexanone	0,64
NSC3329	1-phenyl-2-(2-quinolinyl)ethanone	0,26



NSC11392	2-allyl-4-tert-butyl-6-((dimethylamino)methyl)phenol	0,27
NSC17947	N,N-dibutyl-3-methylbenzamide	0,13
NSC20076	N,N-dibutyl-4-methylbenzamide	0,14
NSC27191	N,N-dimethyl-2-((1-phenylcyclohexyl)oxy)ethanamine	0,19
NSC35523	N-(2-(2,3-dihydro-1H-inden-5-ylmethoxy)ethyl)-N,N-diethylamine	0,23
NSC52524	12-methyl-12H-benzo[a]phenoxazine	0,17
NSC111846	1-((phenylimino)methyl)-2-naphthol	0,26
NSC670854	2-(1-piperidinylmethyl)-1-naphthol	0,22
NSC369293	N-benzyl-N-methyl-2-phenoxyethanamine	0,28
NSC121201	1-(9-ethyl-9H-carbazol-2-yl)ethanone	0,23
NSC88889	benzyl 1-methyl-1H-indol-5-yl ether	0,30
NSC78614	2,3,3a,5,6,11,12,12a-octahydro-1H-4lambda~5~-cyclopenta[5,6]pyrido[2,1-a]isoquinolin-8-ol	0,22
NSC 11847	2-methyl-5-(trimethyl-5-azanyl)phenyl methylcarbamate (ACD/Name 4.0)	0,88
NSC 11859	lambda~5-(diethyl(methyl)-5-azanyl)-2-methylphenyl dimethylcarbamate (ACD/Name 4.0)	0,85
NSC 15409	lambda Ammonium, (3-hydroxy-p-tolyl)dimethylethyl-, iodide, methylcarbamate	0,87
NSC 18833	4-(trimethyl-lambda~5~-azanyl)benzaldehyde	0,74
NSC 6789	N,N'-diphenylimidoforamamide	0,77
NSC 7571	9-acridinamine	0,74
NSC 11846	No Name	0,74
NSC 11869	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 1-piperidinecarboxylate	0,72
NSC 11872	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl phenylcarbamate	0,79
NSC 11875	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 4-methoxyphenylcarbamate	0,75
NSC 15409	5-(ethyl(dimethyl)-lambda~5~-azanyl)-2-methylphenyl methylcarbamate	0,87
NSC 72108	1,2,3,4-tetrahydro-9-acridinamine	0,83
NSC 125821	2-(3-(hydroxy(oxido)amino)phenyl)-1,4-dihydro-2H-3,1-benzoxazine	0,81
NSC 127880	4-(trimethyl-lambda~5~-azanyl)phenyl methylcarbamate	0,96
NSC 128175	3-(diethyl(methyl)-lambda~5~-azanyl)phenyl methylcarbamate	0,86
NSC 168373	3-pyridinyl dimethylcarbamate	0,84
NSC 5047	trimethyl(phenyl)-lambda~5~-azane	0,95
NSC 19699	2-ethylhexyl N,N,N',N'-tetramethyldiamidophosphate	0,72
NSC 17508	3-(diethylamino)phenyl methylcarbamate	0,70
NSC 33764	No Name	0,75
NSC 16853	1-naphthyl aminocarbonylcarbamate	0,49
NSC 21643	trimethyl(6-(trimethyl-lambda~5~-azanyl)hexyl)-lambda~5~-azane	0,50
NSC 38590	1,2,3,4-tetrahydrophenazine	0,46
NSC 96380	2-((2-pyridinylimino)methyl)phenol	0,42
NSC 101483	(4-((hydroxy-lambda~5~-azanylidene)methyl)-3,5-dimethylphenyl)(trimethyl)-lambda~5~-azane	0,44
NSC 108213	N-(4-chlorobenzylidene)-2-pyridinamine	0,50
NSC 114218	(2-(dimethyl(propyl)-lambda~5~-azanyl)ethyl)(dimethyl)propyl-lambda~5~-azane	0,46
NSC 12031	ethyl 4-(((4-(ethoxycarbonyl)anilino)methylene)amino)benzoate	0,53
NSC 205522	4-(((3-(hydroxy(oxido)amino)anilino)carbonyl)oxy)-1,1'-biphenyl	0,57



NSC 632382	N-(4-(dimethylamino)benzylidene)-1-adamantanamine	0,54
NSC 74685	3-methylphenyl 1-naphthylcarbamate	0,50
NSC 132951	1-naphthyl benzoylcarbamate	0,55
NSC 140786	1-(4-(trimethyl-lambda~5~-azanyl)phenyl)-1-nonen-3-one	0,55
NSC 164541	4-(benzyloxy)phenyl butylcarbamate	0,53
NSC 191428	2-naphthyl 2-methylphenylcarbamate	0,56
NSC 191448	2-naphthyl 2-methoxyphenylcarbamate	0,55
NSC 95514	3-(3-azabicyclo[4.2.1]non-3-yl)-1-(2-fluorophenyl)-1-propanone	0,60
NSC 118795	(1-benzyl-1H-indol-3-yl)(oxo)acetic acid	0,52
NSC 133584	1-benzyl-5-ethylideneazonane	0,53
NSC 204269	2-naphthyl cyclohexylcarbamate	0,56
NSC 204733	4-(((tert-butylamino)carbonyl)oxy)-1,1'-biphenyl	0,54
NSC 16856	1-naphthyl 2-cyclopenten-1-yl(methyl)carbamate	0,69
NSC 59666	2-naphthyl dipropylcarbamate	0,70
NSC 91873	(4-(1H-inden-1-ylidenemethyl)phenyl)(trimethyl)-lambda~5~-azane	0,74
NSC 100058	Galanthamine	0,79
NSC 51928	9H-fluoren-9-yl(trimethyl)-lambda~5~-azane	0,87
NSC 176000	phenyl(4-(trimethyl-lambda~5~-azanyl)phenyl)methanone	0,76
NSC 176008	2-phenyl-1-naphthalenamine	0,62
NSC 191481	2-(((propylamino)carbonyl)oxy)-1,1'-biphenyl	0,60
NSC 204566	2-benzoyl-5-methoxyphenyl methylcarbamate	0,64
NSC 527673	triethyl(4-((triethyl-lambda~5~-azanyl)methyl)benzyl)-lambda~5~-azane	0,62
NSC 605449	4a-methyl-9-(2-(1-piperidinyl)ethyl)-2,3,4,4a,9,9a-hexahydro-1H-carbazole	0,62
NSC 169029	N ¹ -(9-anthryl)-N,N-dimethylimidofornamide	0,75
NSC 231606	No Name	0,85
NSC 215562	4-anilinophenyl methylcarbamate	0,77

The classification of the molecules was made through the k-means method, according to their structures. The first one kind claser is formed by molecules with a

diversity of structures, the second one contains pyridine groups, and the last one consists in structures containing carbamate groups.

Table 3. Biological activity of each group

Classes	# Molecules	Biological activity range	Mean
1	92	0.128 – 0.945	0.31
2	14	0.122 – 0.832	0.41
3	39	0.169– 0.986	0.66

Table 3 shows the number of molecules in each class, along with their range of

biological activity and the mean in every group.



Table 4. Structures with its respective name, belonging to the second classification

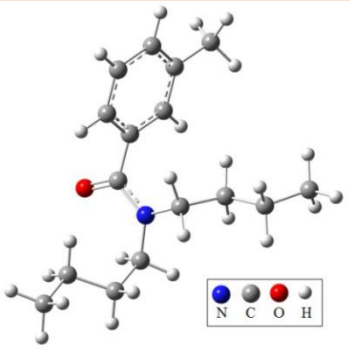
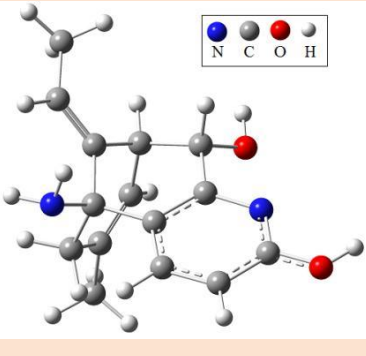
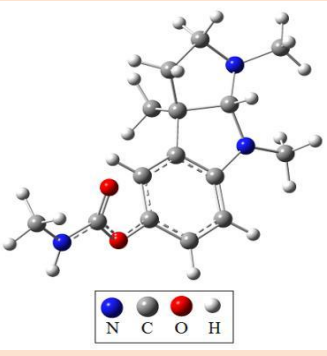
Class 1	Class 2	Class 3
		
N,N-dibutyl-3-methylbenzamide	Huperzine A, 6- -hydroxy	1,3a,8-trimethyl-1,2,3,3a,8,8a-hexahydropyrrolo[2,3- b]indol-5-yl methylcarbamate
Structures different of each other	Mostly pyridine groups.	Carbamate groups at most

Table 4 shows the structure of one molecule of each class, the UIPAC name and the characteristic that characterize the class.

Through the KNN method and the use of the genetic algorithms, along with the previously mentioned parameters, a model with a high NER and prediction was achieved. The results were highly favorable, with an error of 5% *leave-one-*

out (every one of them, randomly chosen) and a selection of 12 variables.

Figure 1, obtained in MATLAB software, shows the number of descriptors used in the model, versus the cross-validated NER, with 5 groups of extractions.

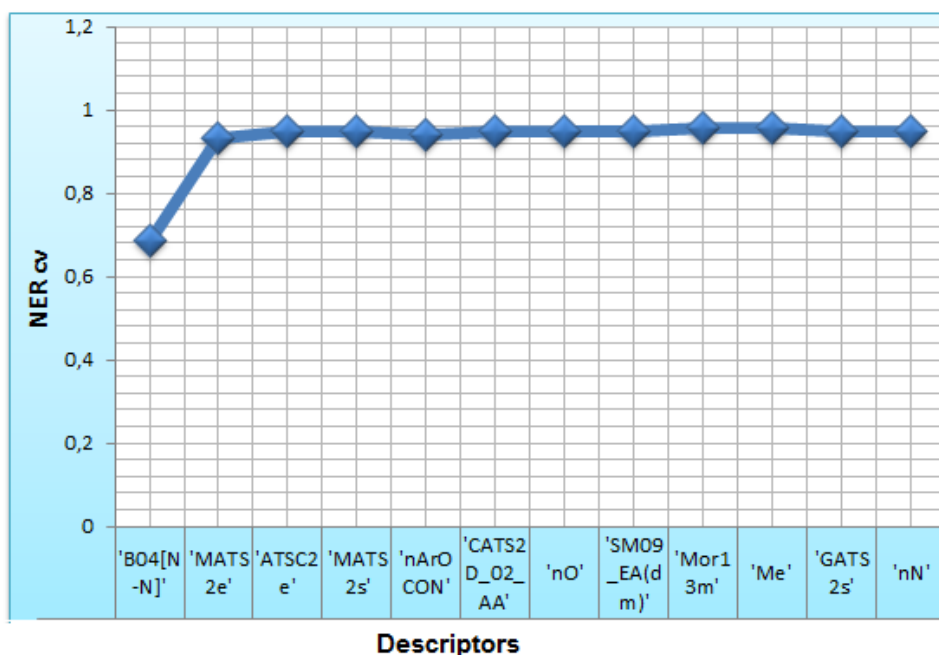


Fig 1. NER vs. number of descriptors
Source: Author

The previous figure shows that the NER is 0.95 with the first 12 descriptors. Increasing the number of descriptors does not improve the prediction.

The next figure, fig. 2, shows the internal

test, which includes almost 80% of the molecules, and that provides the prediction of 0.95 *leave-one-out* (obtained through SCAN software).

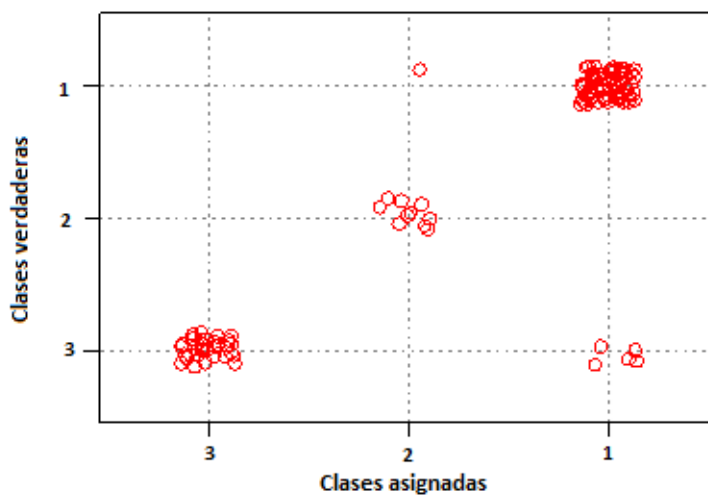


Fig 2. Classification of molecules according to their structures in the respective classes
Source: Author



Table 5. Confusion matrix (knn method)

		class 1	class 2	class 3
class 1	69	61	2	6
class 2	11	5	5	1
class 3	35	4	4	27
	115	70	11	34

Table 5 shows the confusion matrix, which is presenting the molecules with their assigned class and allows to determinate the NER. The elements in the diagonal row are

the ones that are correctly assigned, and the remaining molecules should be in a different class. The NER is 0.8 of prediction.

Table 6. Descriptors

Descriptors	Description
'B04[N-N]'	2D descriptor. Presence/absence of N - N at topological distance of 4 atoms.
'MATS2e'	2D Moran Autocorrelation, regard to Sanderson's electronegativity. Every two bonds distance there will be an electronegativity situation.
'ATSC2e'	Centered autocorrelation, referring to the center of the molecule. There will be a distribution of electronegativity every 2 atoms.
'MATS2s'	2D Moran autocorrelation. It describe an ionization state every 2 bonds.
'nArOCON'	Number of (thio-) carbamates (aromatic)
'CATS2D_02_AA'	Acceptor-Acceptor every two atoms
'nO'	Number of oxygen atoms.
'SM09_EA(dm)'	Adjacency matrix weighted by dipole moment.
'Mor13m'	3d Morse descriptor weighted by mass
'Me'	Mean Sanderson electronegativity (regarding C)
'GATS2s'	2D Geary autocorrelation weighted by ionization state.
'nN'	Number of nitrogen atoms.

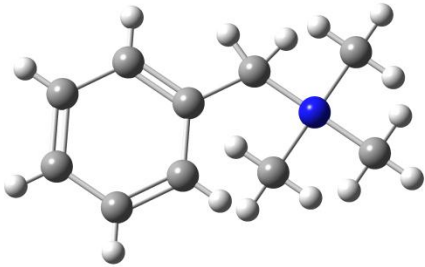
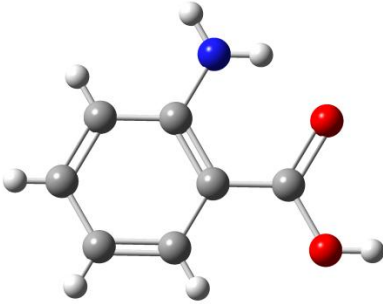
Table 6 shows the 12 descriptors that have been used for the characterization of each class and that enhance the predictive model. These descriptors were the ones used in the formerly mentioned external test subset that used up 20% of the molecules. The global error in the model is 5%,

including the external test (these results were obtained including the molecules that were previously excluded).

In order to validate the developed model, it was tested with two molecules that were not part of the research. These molecules are depicted in table 6.



Tabla 6. Structure, IUPAC name and biological activity of the molecules used for the validation.

	
benzyl(trimethyl)-lambda-5-azane	2-aminobenzoic acid
0.391	0.135

The model predict the molecules to belong to class 1. Considering their inhibitory activity (0.391 and 0.135, respectively), the

assignment was the proper one, which means the model is appropriate.

Conclusions:

- The obtained results have an error of 5%, which means that the NER correlation factor is 0.95. According to the obtained data, the molecular descriptors that best describe the acetylcholinesterase inhibitory activity of the studied molecules are: “B04[N-N]”, “MATS2e”, “ATSC2e”, “MATS2s”, “nArOCON”, “CATS2D_02_AA”, “nO”, “SM09_EA(dm)”, “Mor13m”, “Me”, “GATS2s”, “nN”.
- The best classification for the model was obtained through the K-MEANS method, separating the set of 145 molecules in three different classes according to their structure.
- The present model is highly effective to determine the biological activity of molecules similar to the ones in class 1 and class 3, which are the better classified.
- The present work shows the utility of the QSAR method to predict biological activities without expensive experimental techniques.
- It is interesting that the molecules with the larger inhibitory activity have, mostly, aromatic compounds and nitrogen in their structure; on the other hand, the biological activity is smaller for molecules with a large number of substituents and large radicals.
- The molecules used for the validation of the model were properly assigned to their corresponding classes, which mean that the model works properly.



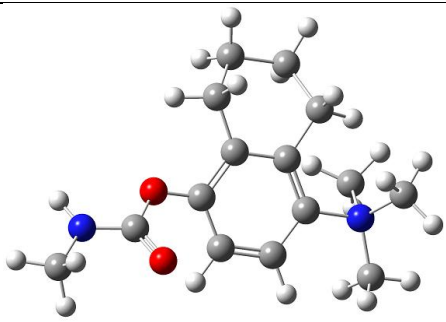
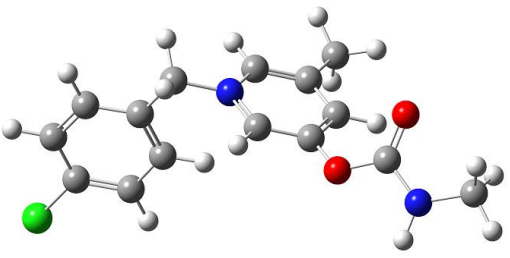
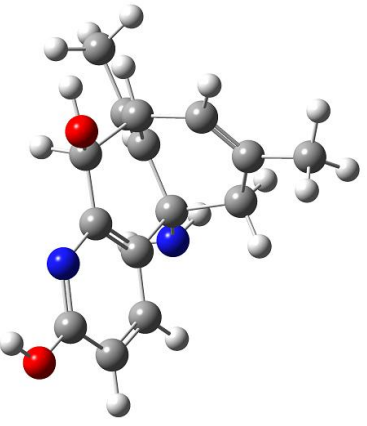
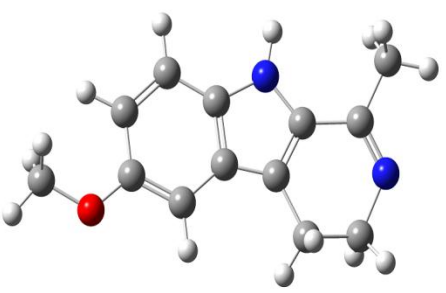
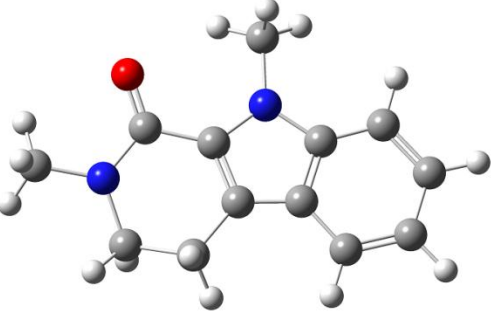
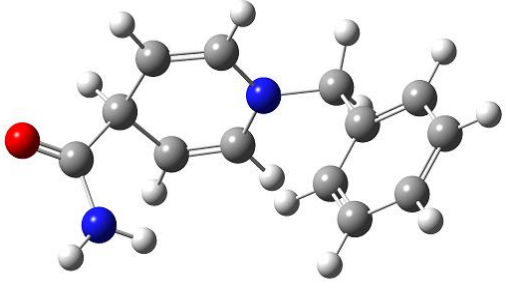
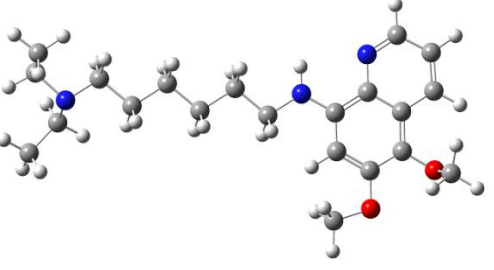
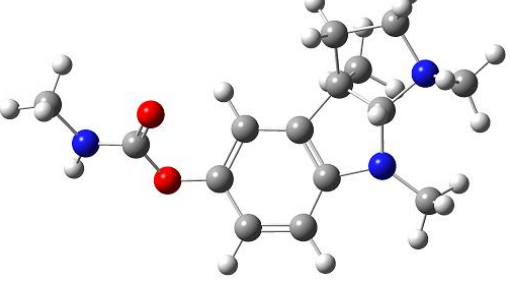
Bibliography

- [1] Donoso, A. Revista Chilena de neuro-psiquiatría **2003**, 41, 13-22.
- [2] Peña-Casanova, J. Enfermedad de Alzheimer; 1st ed.; Fundación “la Caixa”, Barcelona, **1999**; p. 121.
- [3] Bai, D. L.; Tang, X. C.; He, X. C. Current medicinal chemistry **2000**, 7, 355,74.
- [4] Lu, S.; Wu, W.; Liu, H.; Zhao, J.; Liu, K.; Chuang, C.; Lin, H.; Tsai, W.; Ho, Y. The discovery of potential acetylcholinesterase inhibitors: A combination of pharmacophore modeling, virtual screening, and molecular docking studies. Journal of Biomedical Science **2011**,12-8.
- [5] Duardo, A.; Patlewicz, G.; Lopez-Diaz, A. Current topics on software use in medicinal chemistry: Intellectual property, taxes, and regulatory issues. Curr. Top. Med. Chem. 2008, 8, 1666-1675.
- [6] Gramatica, P. A short history of QSAR evolution [http://www.qsarworld.com/qsar- archives](http://www.qsarworld.com/qsar-archives) (accessed May 2010).
- [7] Todeschini, R.; Consonni, V. Manual de descriptores moleculares, Wiley-VCH. <http://www.moleculardescriptors.eu/books/handbook.htm> (2000)
- [8] Frank, I.; Todeschini, R. The Data Analysis Handbook.ElsevierScience, 2006, 334-350.
- [9] Todeschini, R. Introducción a la Quimiométrica, Edises, Universidad Milano Bicocca, 2003, 80-220.
- [10] Franca, C.; Pis, R. Estudio teórico de propiedades fisicoquímicas de compuestos de interés farmacológicos. Universidad Nacional de la Plata, 134
- [11] Ballabio, D.;Todeschini, R. Multivariate Classification for Qualitative Analysis, en Infrared Spectroscopy for Food Quality analysis and Control Edited by Da-Wen Sun Elsevier Inc, 2009, 94.
- [12] Hansch, A.; Leo, S.; Heller, Ed.Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, American Chemical Society, Washington, 1995.

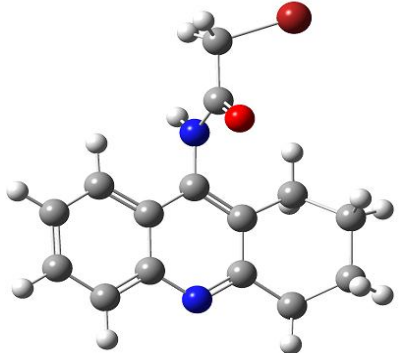
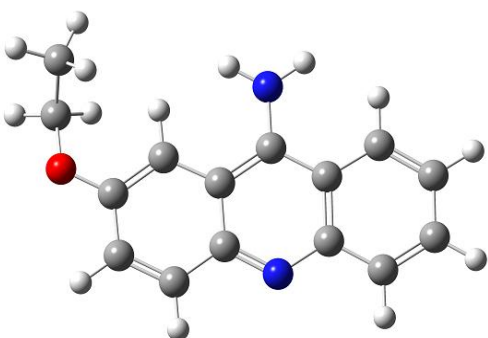
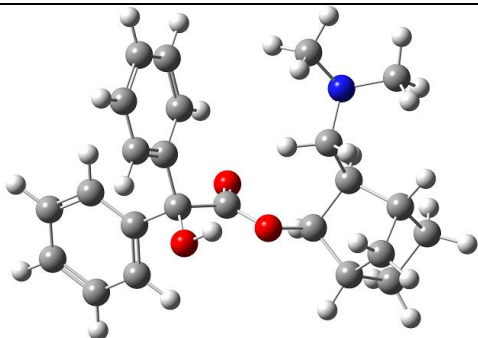
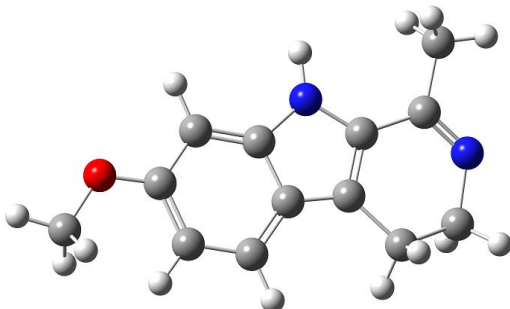
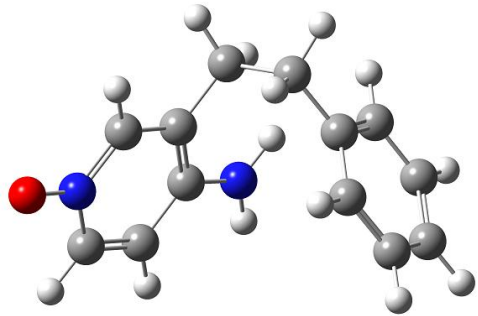
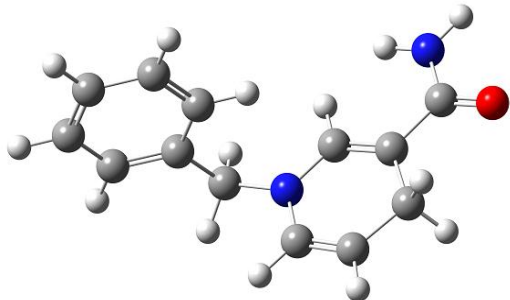
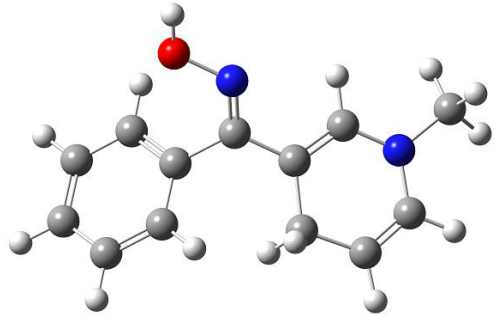
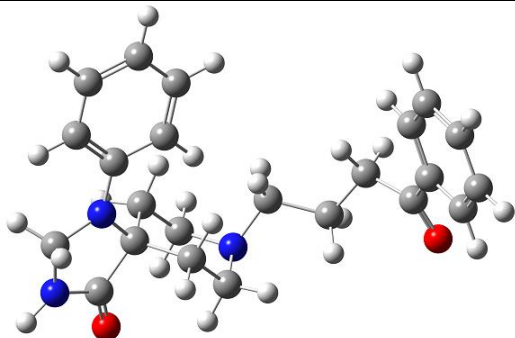


ANEXOS

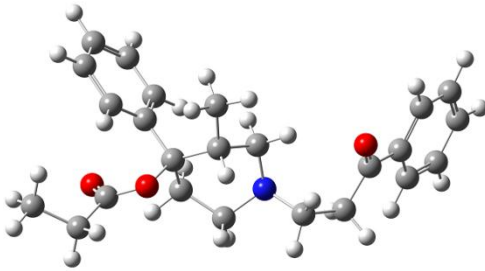
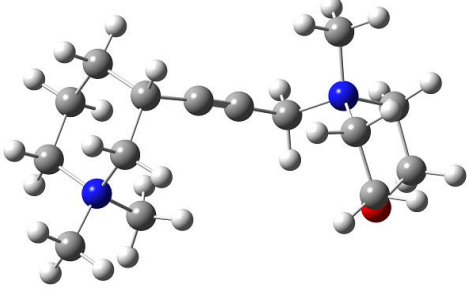
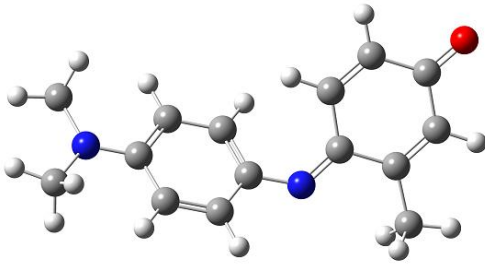
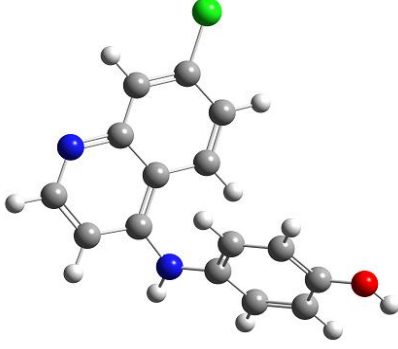
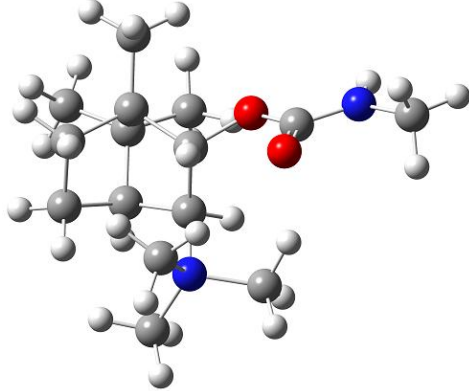
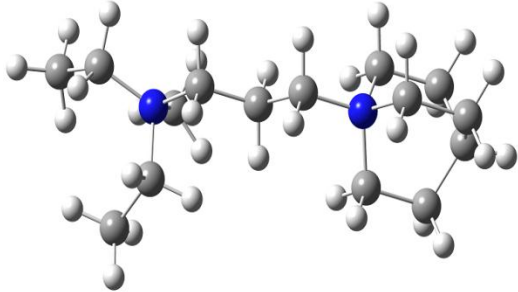
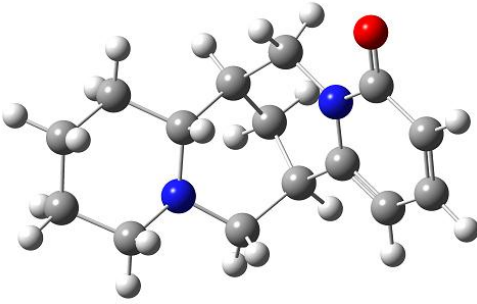
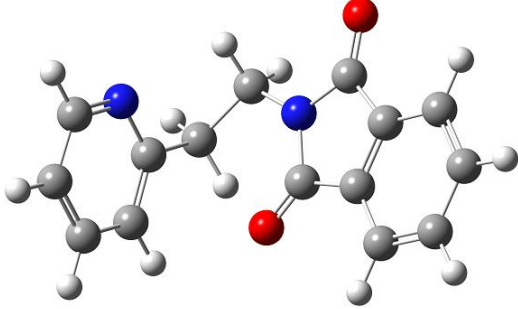
Estructuras del conjunto molecular usado para el diseño del modelo QSAR para inhibidores de acetilcolinesterasa.

	
<p>4-(trimethyl-lambda~5~-azanyl)-5,6,7,8-tetrahydro-1-naphthalenyl methylcarbamate</p>	<p>1-(4-chlorobenzyl)-5-methyl-1lambda~5~-pyridin-3-yl methylcarbamate</p>
	
<p>Huperzine A, 6-.beta.-hydroxy</p>	<p>6-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline</p>
	
<p>2,9-dimethyl-2,3,4,9-tetrahydro-1H-beta-carboline-1-one</p>	<p>1-benzyl-1,4-dihydro-4-pyridinecarboxamide</p>
	
<p>N~1~-(5,6-dimethoxy-8-quinolinyl)-</p>	<p>1,3a,8-trimethyl-1,2,3,3a,8,8a-</p>

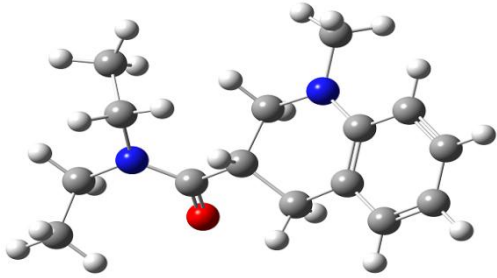
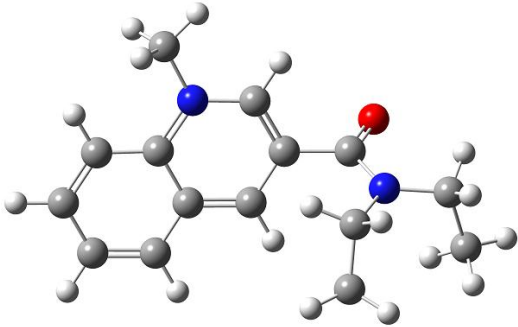
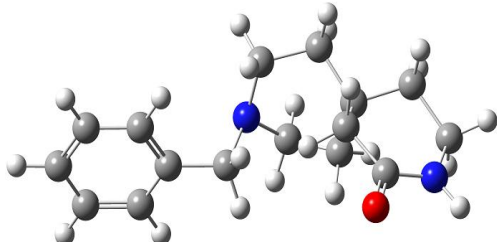
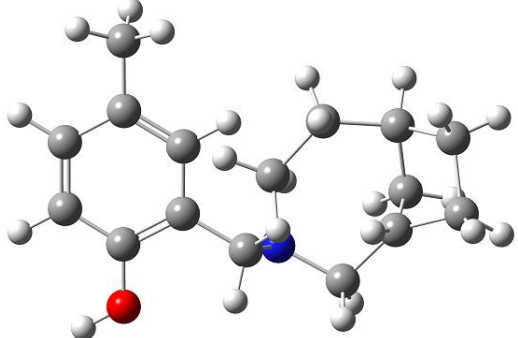
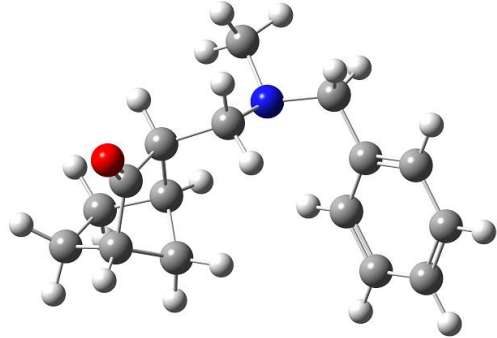
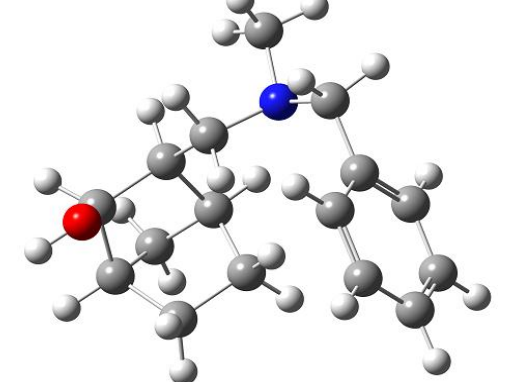
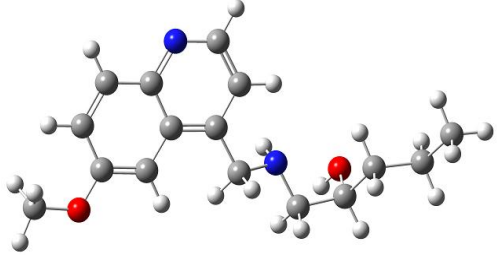
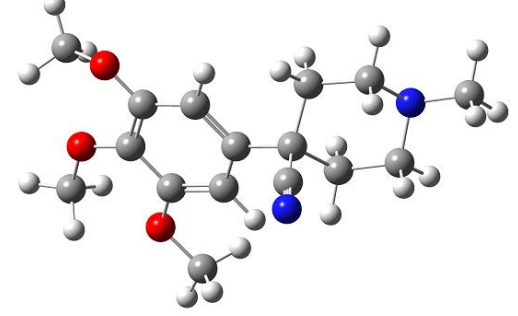


N,N-diethyl-1,6-hexanediamine	hexahydropyrrolo[2,3-b]indol-5-yl methylcarbamate
	
2-bromo-N-(1,2,3,4-tetrahydro-9-acridinyl)acetamide	2-ethoxy-9-acridinamine
	
3-((dimethylamino)methyl)bicyclo[2.2.1]hept-2-yl hydroxy(diphenyl)acetate	7-methoxy-1-methyl-4,9-dihydro-3H-beta-carboline
	
4-amino-3-(2-phenylethyl)-1lambda5-pyridin-1-ol	1-benzyl-1,4-dihydro-3-pyridinecarboxamide
	
(E)-(1-methyl-1,4-dihydro-3-pyridinyl)(phenyl)methanone oxime	8-(4-oxo-4-phenylbutyl)-1-phenyl-1,3,8-triazaspiro[4.5]decan-4-one

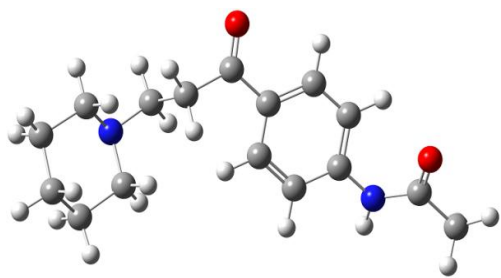
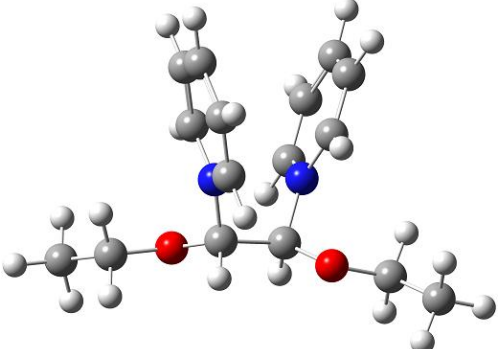
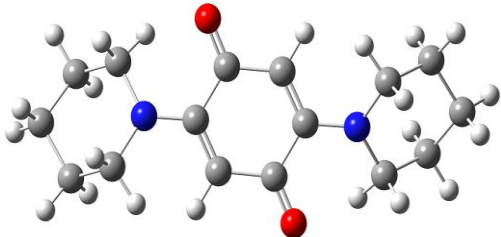
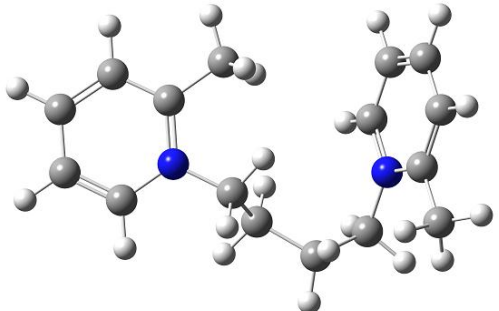
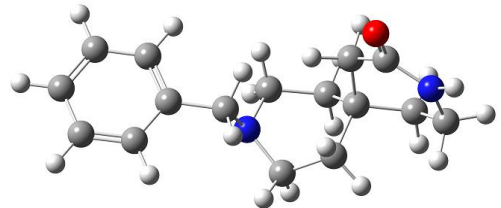
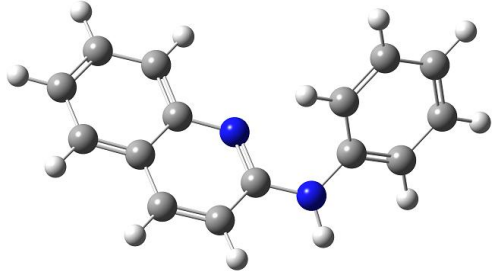
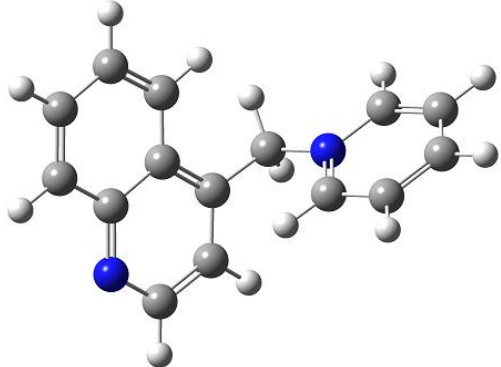
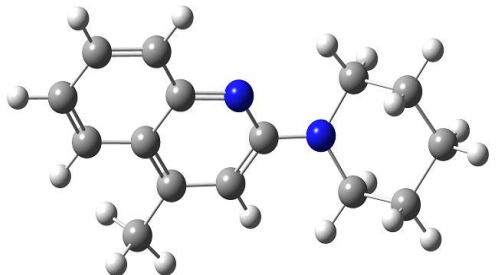


	
<p>3-methyl-1-(3-oxo-3-phenylpropyl)-4-phenyl-4-piperidinyl propionate</p>	<p>4-(3-(1,1-dimethyl-1lambda~5~-piperidin-3-yl)-2-propynyl)-4-methyl-1,4lambda~5~-oxazinane</p>
	
<p>4-((4-(dimethylamino)phenyl)imino)-3-methyl-2,5-cyclohexadien-1-one</p>	<p>4-((7-chloro-4-quinolinyl)amino)phenol</p>
	
<p>1,7,7-trimethyl-3-(trimethyl-lambda~5~-azanyl)bicyclo[2.2.1]hept-2-yl methylcarbamate</p>	<p>1-(3-(diethyl(methyl)-lambda~5~-azanyl)propyl)-1lambda~5~-azabicyclo[2.2.2]octane</p>
	
<p>3,4,5,6-tetrahydrospartein-2-one</p>	<p>2-(2-(2-pyridinyl)ethyl)-1H-isoindole-1,3(2H)-dione</p>

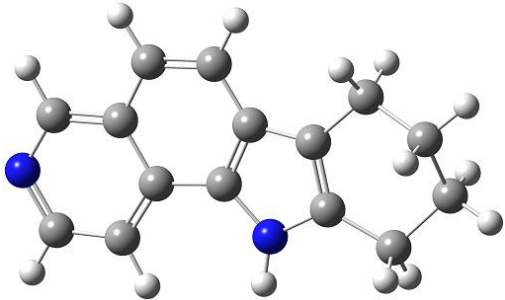
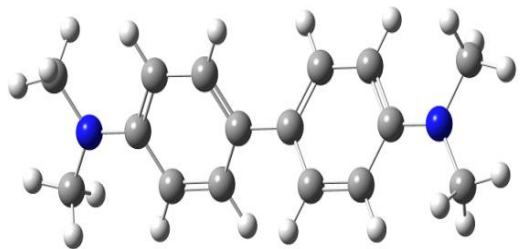
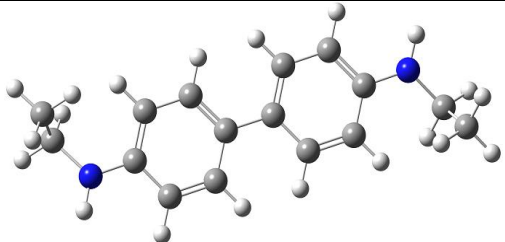
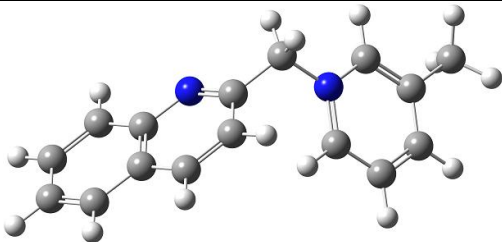
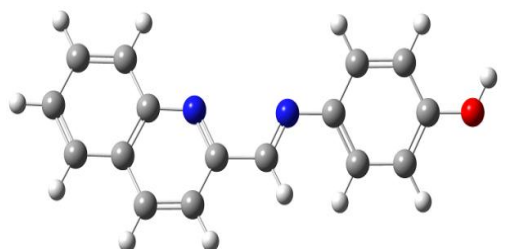
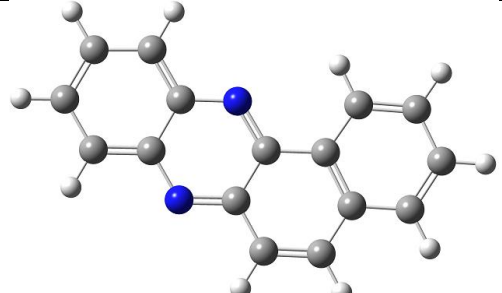
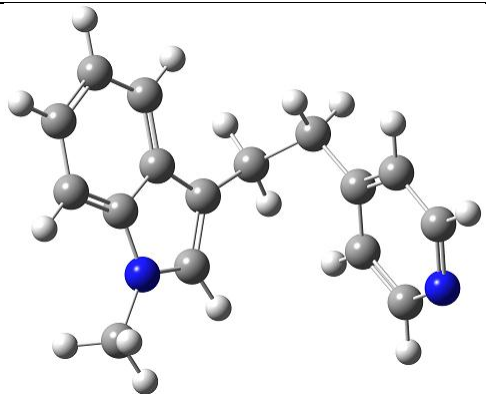
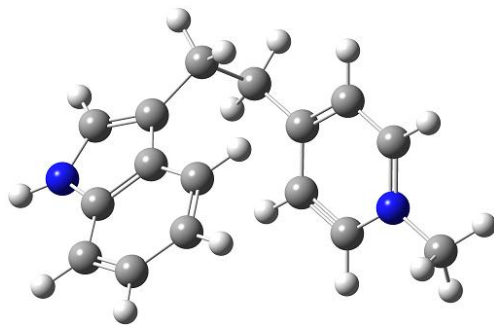


	
N,N-diethyl-1-methyl-1,2,3,4-tetrahydro-3-quinolinecarboxamide	N,N-diethyl-1-methyl-1lambda~5~-quinoline-3-carboxamide
	
9-benzyl-3,9-diazaspiro[5.5]undecan-2-one	2-(3-azabicyclo[4.2.1]non-3-ylmethyl)-4-methylphenol
	
3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-one	3-((benzyl(methyl)amino)methyl)bicyclo[2.2.1]heptan-2-ol
	
1-(((6-methoxy-4-quinolinyl)methyl)amino)-2-pentanol	1-methyl-4-(3,4,5-trimethoxyphenyl)-4-piperidinecarbonitrile



	
<p>N-(4-(3-(1-piperidinyl)propanoyl)phenyl)acetamide</p>	<p>1-(1,2-diethoxy-2-(1-lambda~5~-pyridin-1-yl)ethyl)-1-lambda~5~-pyridine</p>
	
<p>2,5-di(1-piperidinyl)benzo-1,4-quinone</p>	<p>2-methyl-1-(4-(2-methyl-1-lambda~5~-pyridin-1-yl)butyl)-1-lambda~5~-pyridine</p>
	
<p>9-benzyl-3,9-diazaspiro[5.5]undecan-2-one</p>	<p>N-phenyl-N-(2-quinolinyl)amine</p>
	
<p>4-(1-lambda~5~-pyridin-1-ylmethyl)quinoline</p>	<p>4-methyl-2-(1-piperidinyl)quinoline</p>

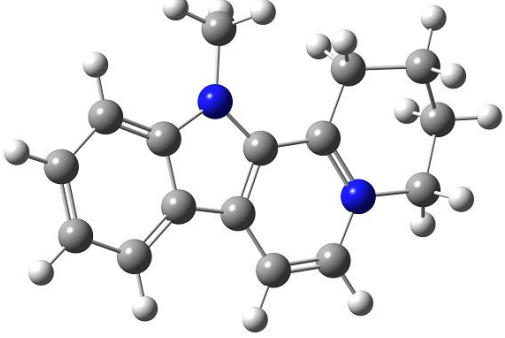
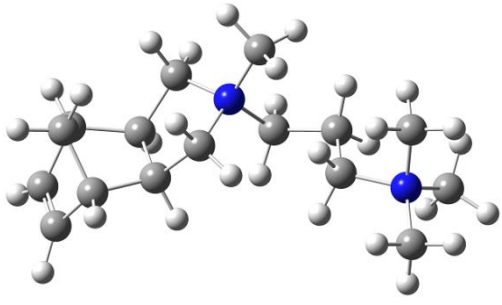
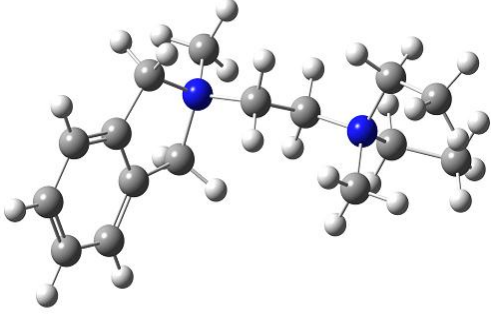
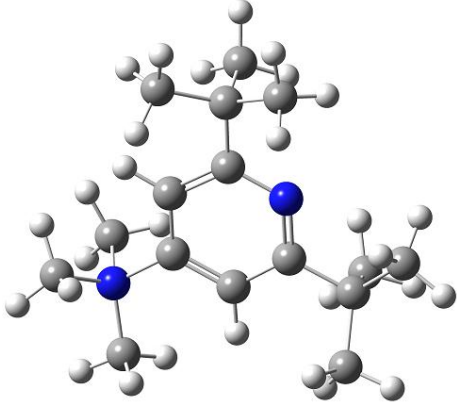
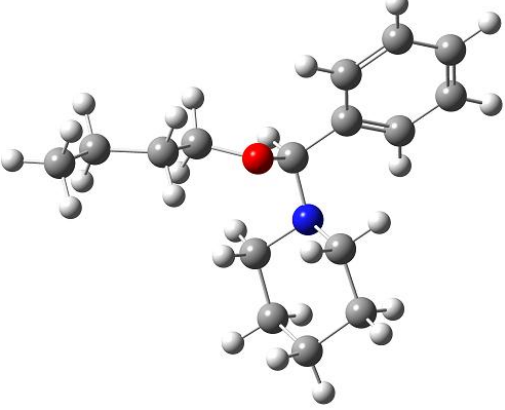
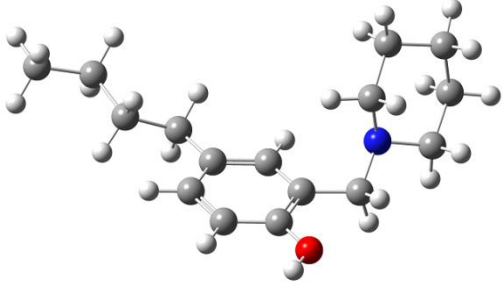
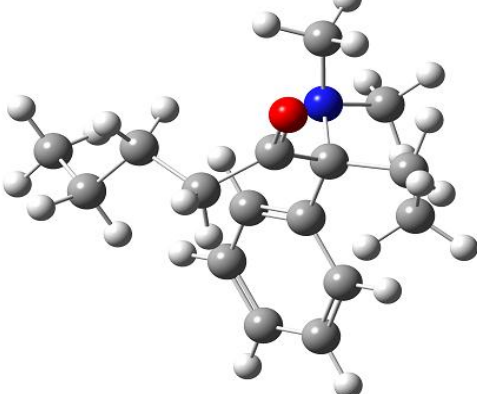
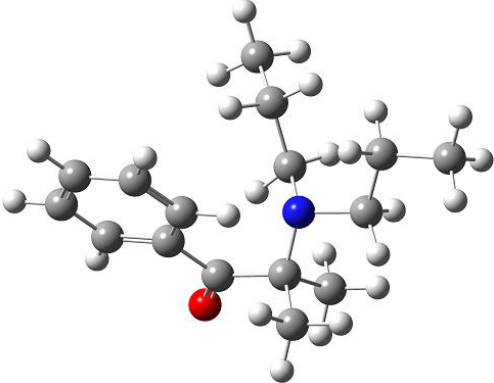


	
8,9,10,11-tetrahydro-7H-pyrido[4,3-a]carbazole	N-(4'-(dimethylamino)[1,1'-biphenyl]-4-yl)-N,N-dimethylamine
	
N,N'-diethyl[1,1'-biphenyl]-4,4'-diamine	2-((3-methyl-1lambda~5~-pyridin-1-yl)methyl)quinoline
	
4-((2-quinolinylmethylene)amino)phenol	benzo[a]phenazine
	
1-methyl-3-(2-(4-pyridinyl)ethyl)-1H-indole	3-(2-(1-methyl-1lambda~5~-pyridin-4-yl)ethyl)-1H-indole

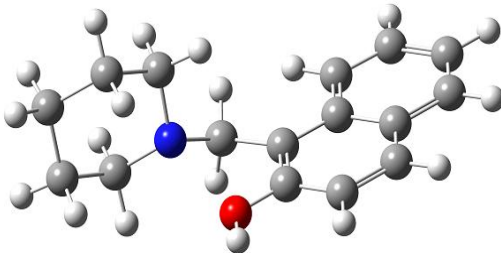
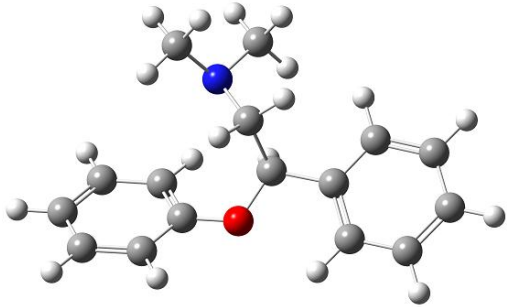
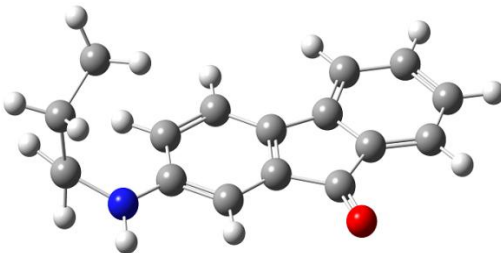
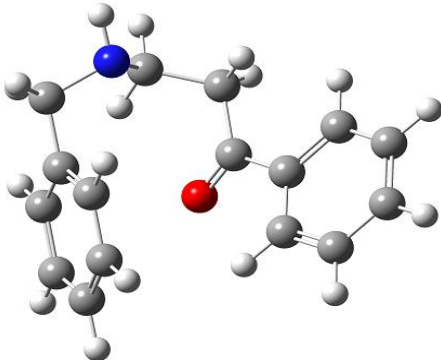
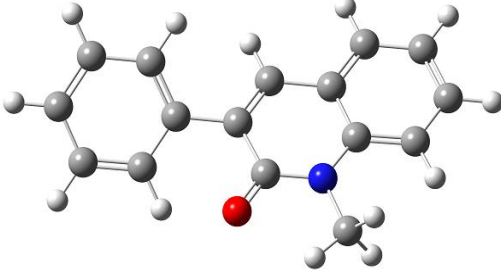
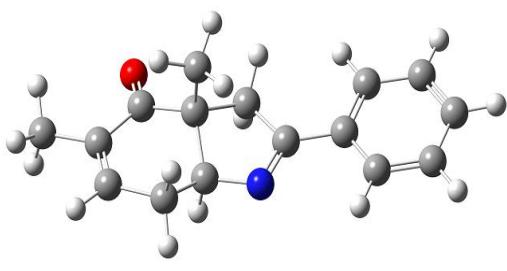
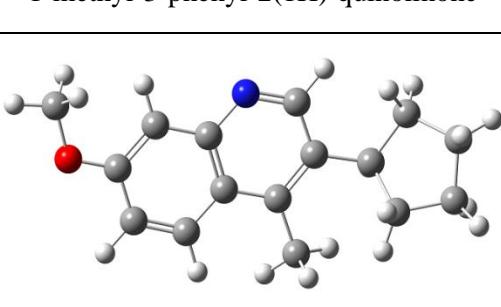
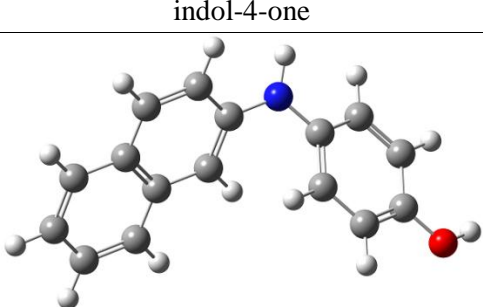


<p>2-cyclohexyl-N-methyl-N-(2-(2-pyridinyl)ethyl)ethanamine</p>	<p>2-(4-methylphenyl)-3-quinolinamine</p>
<p>N-1-butyl-N-4-phenyl-1,4-benzenediamine</p>	<p>N-1-(2-naphthyl)-1,4-benzenediamine</p>
<p>3-methyl-1-(4-(3-methyl-1lambda5-pyridin-1-yl)butyl)-1lambda5-pyridine</p>	<p>4-methyl-1-(4-(4-methyl-1lambda5-pyridin-1-yl)butyl)-1lambda5-pyridine</p>
<p>7-methyl-6,7,7a,8,9,10,10a,10b,11,12-decahydrocyclopenta[4,5]pyrimido[1,6-a]quinoline</p>	<p>2-(2-(diethyl(methyl)-lambda5-azanyl)ethyl)-2-methyl-2,3,3a,4,7,7a-hexahydro-1H-2lambda5-isoindole</p>

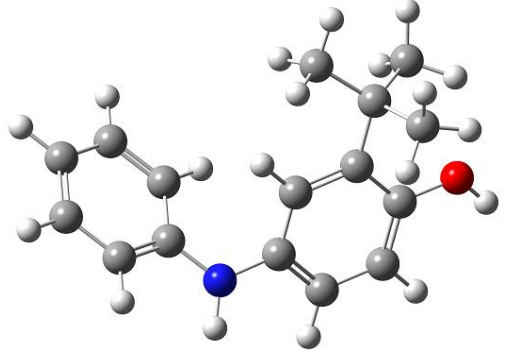
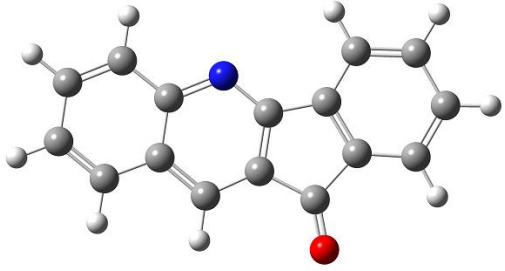
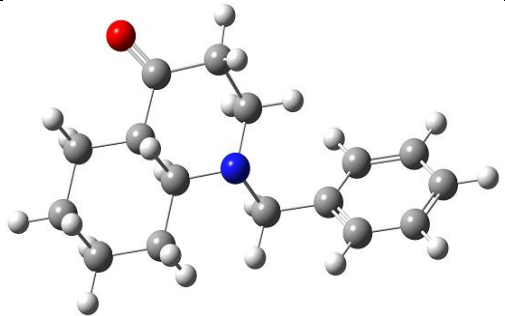
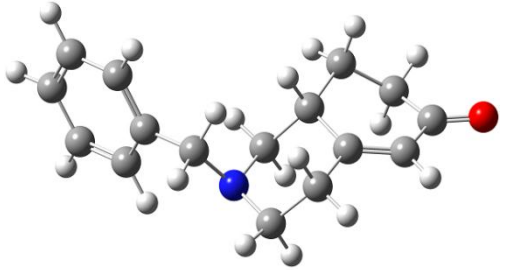
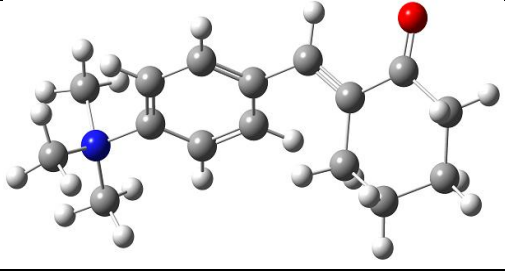
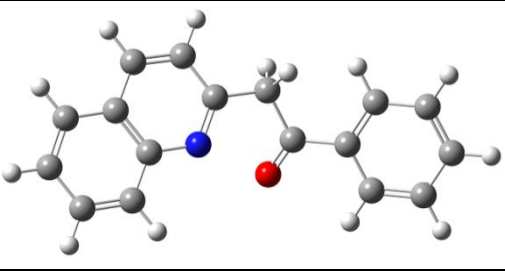
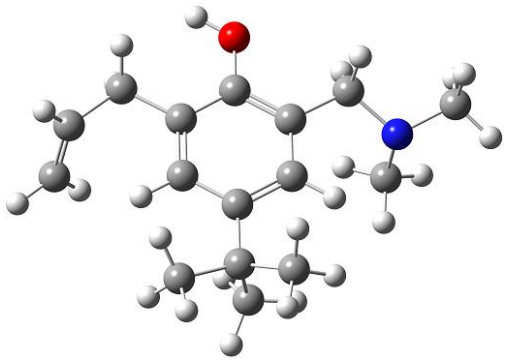
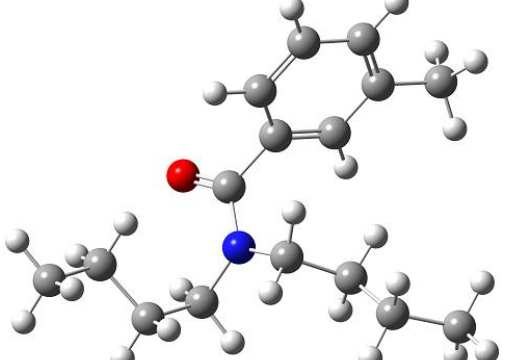


	
12-methyl-2,3,4,12-tetrahydro-1H-5lambda~5~-indolo[2,3-a]quinolizine	No Name
	
2-(2-(diethyl(methyl)-lambda~5~-azanyl)ethyl)-2-methyl-2,3-dihydro-1H-2lambda~5~-isoindole	2,6-ditert-butyl-4-(trimethyl-lambda~5~-azanyl)pyridine
	
1-(butoxy(phenyl)methyl)piperidine	4-butyl-2-(1-piperidinylmethyl)phenol
	

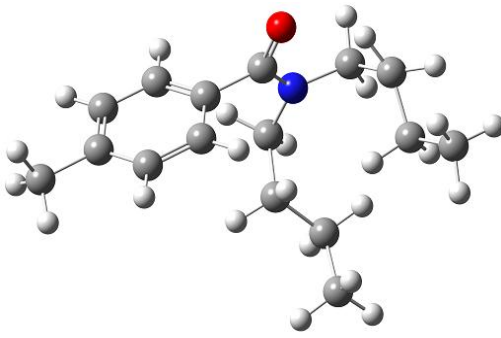
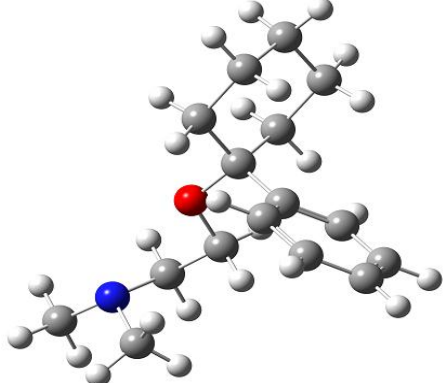
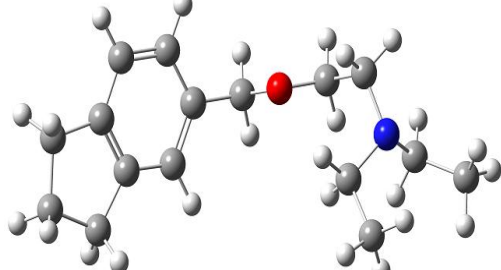
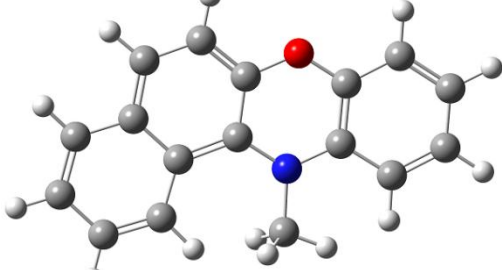
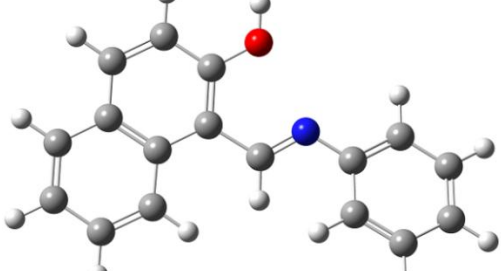
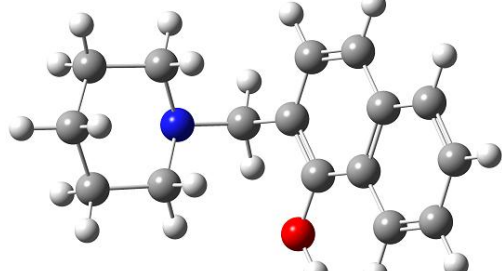
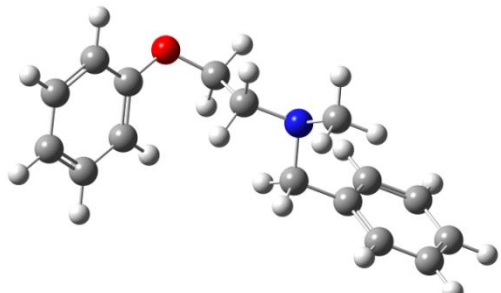
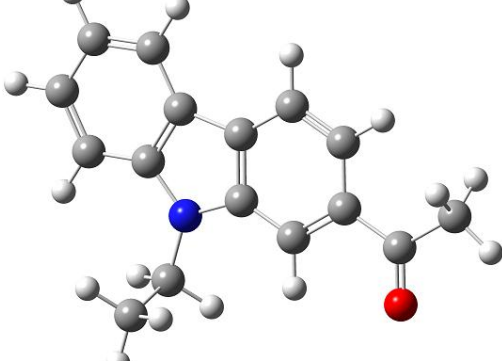


3-(dimethylamino)-3-phenyl-4-octanone	2-(dipropylamino)-2-methyl-1-phenyl-1-propanone
	
1-(1-piperidinylmethyl)-2-naphthol	N,N-dimethyl-2-phenoxy-2-phenylethanamine
	
2-(propylamino)-9H-fluoren-9-one	3-(benzylamino)-1-phenyl-1-propanone
	
1-methyl-3-phenyl-2(1H)-quinolinone	3a,5-dimethyl-2-phenyl-3,3a,7,7a-tetrahydro-4H-indol-4-one
	
3-cyclopentyl-7-methoxy-4-methylquinoline	4-(2-naphthylamino)phenol

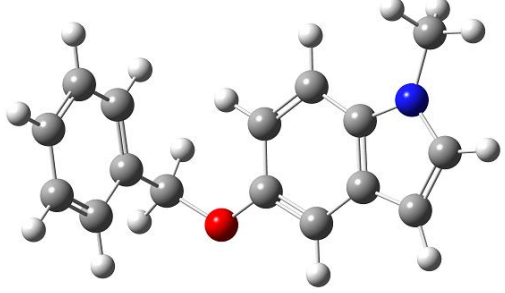
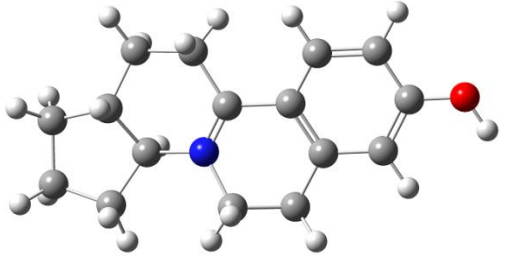
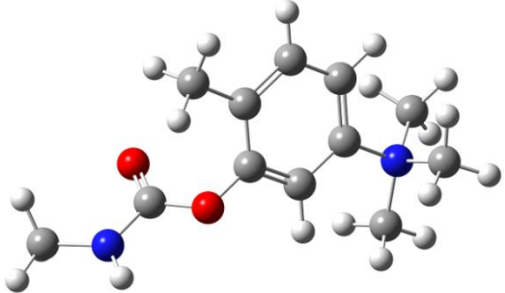
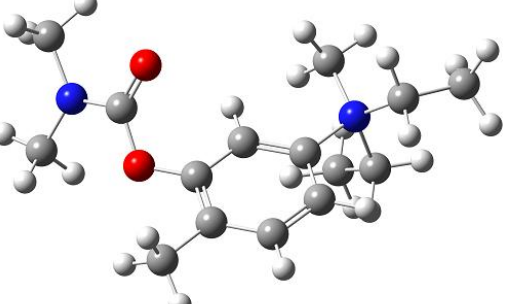
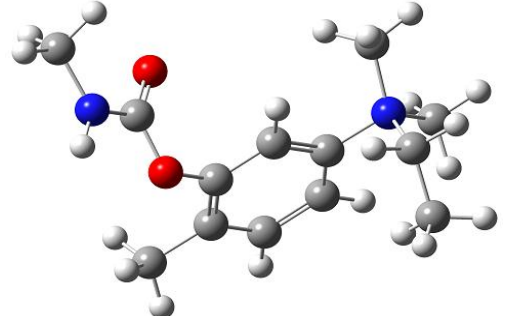
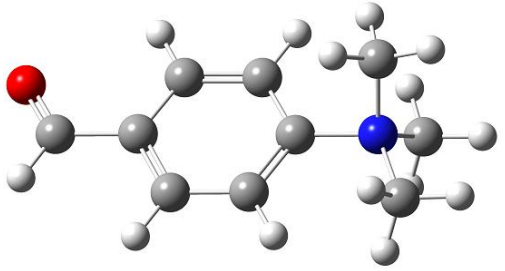
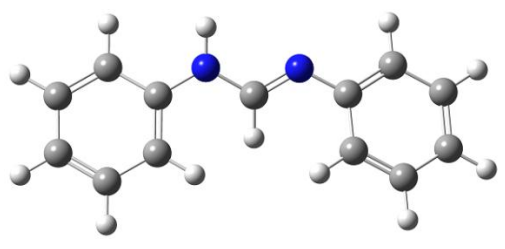
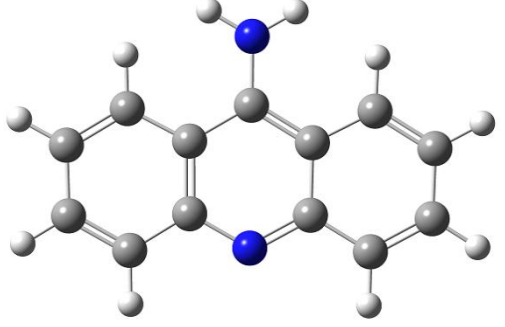


	
4-anilino-2-tert-butylphenol	11H-indeno[1,2-b]quinolin-11-one
	
1-benzyl-4(1H)-quinolinone	2-benzyl-1,3,4,7,8,8a-hexahydro-6(2H)-isoquinolinone
	
2-(4-(trimethyl-azanyl)benzylidene)cyclohexanone	1-phenyl-2-(2-quinolinyl)ethanone
	
2-allyl-4-tert-butyl-6-((dimethylamino)methyl)phenol	N,N-dibutyl-3-methylbenzamide



	
<p>N,N-dibutyl-4-methylbenzamide</p>	<p>N,N-dimethyl-2-((1-phenylcyclohexyl)oxy)ethanamine</p>
	
<p>N-(2-(2,3-dihydro-1H-inden-5-ylmethoxy)ethyl)-N,N-diethylamine</p>	<p>12-methyl-12H-benzo[a]phenoxazine</p>
	
<p>1-((phenylimino)methyl)-2-naphthol</p>	<p>2-(1-piperidinylmethyl)-1-naphthol</p>
	
<p>N-benzyl-N-methyl-2-phenoxyethanamine</p>	<p>1-(9-ethyl-9H-carbazol-2-yl)ethanone</p>

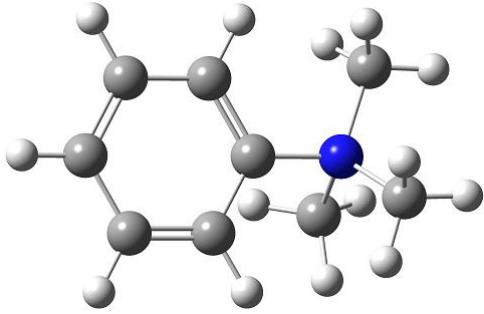
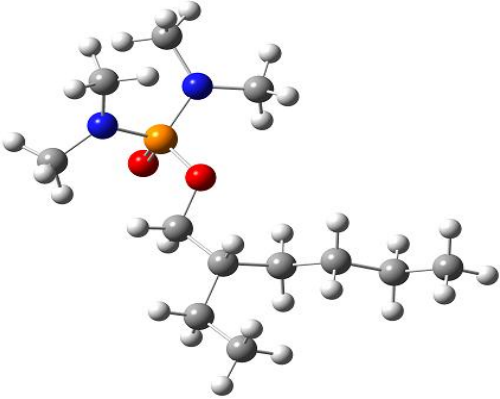
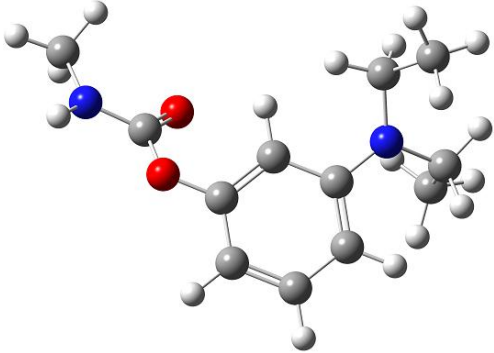
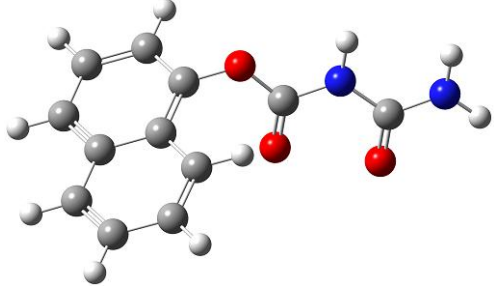
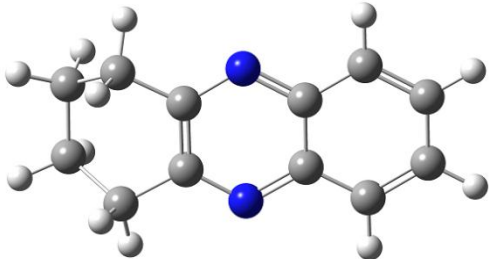



	
benzyl 1-methyl-1H-indol-5-yl ether	2,3,3a,5,6,11,12,12a-octahydro-1H-4lambda~5~-cyclopenta[5,6]pyrido[2,1-a]isoquinolin-8-ol
	
2-methyl-5-(trimethyl-5-azanyl)phenyl methylcarbamate (ACD/Name 4.0)	5-(diethyl(methyl)-5-azanyl)-2-methylphenyl dimethylcarbamate (ACD/Name 4.0)
	
Ammonium,(3-hydroxy-p-tolyl) dimethylethyl-, iodide , methylcarbamate	4-(trimethyl-lambda~5~-azanyl) benzaldehyde
	
N,N'-diphenylimidoformamide	9-acridinamine



<p>No Name</p>	<p>3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 1-piperidinecarboxylate</p>
<p>3-(diethyl(methyl)-lambda~5~-azanyl)phenyl phenylcarbamate</p>	<p>3-(diethyl(methyl)-lambda~5~-azanyl)phenyl 4-methoxyphenylcarbamate</p>
<p>5-(ethyl(dimethyl)-lambda~5~-azanyl)-2-methylphenyl methylcarbamate</p>	<p>1,2,3,4-tetrahydro-9-acridinamine</p>
<p>2-(3-(hydroxy(oxido)amino)phenyl)-1,4-dihydro-2H-3,1-benzoxazine</p>	<p>4-(trimethyl-lambda~5~-azanyl)phenyl methylcarbamate</p>
<p>3-(diethyl(methyl)-lambda~5~-</p>	<p>3-pyridinyl dimethylcarbamate</p>

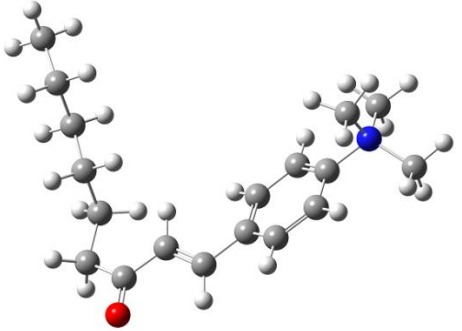
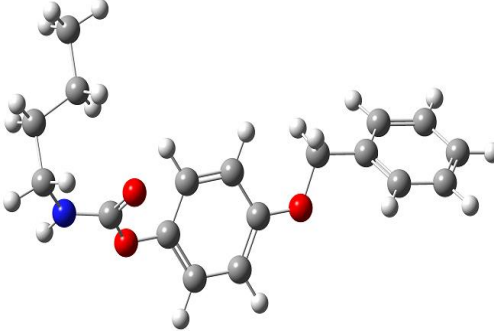
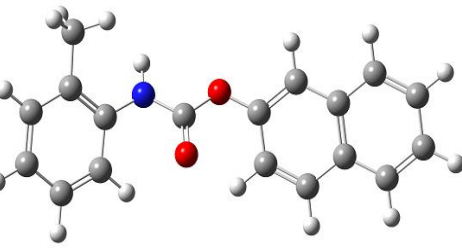
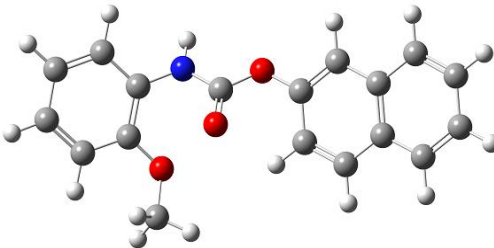
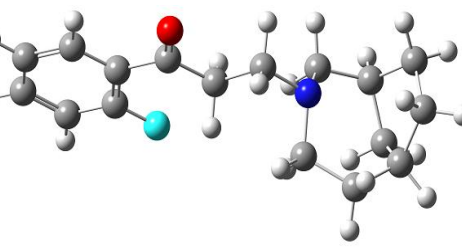
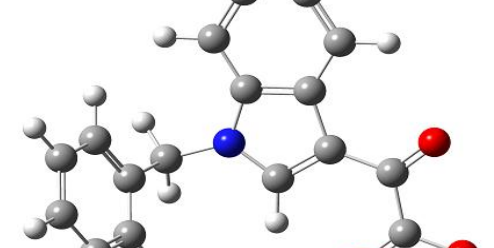
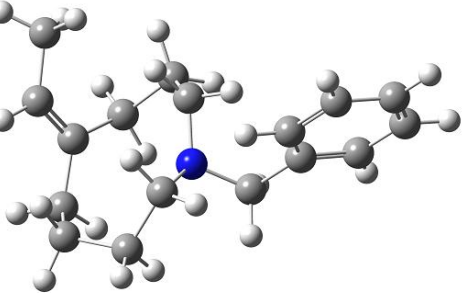
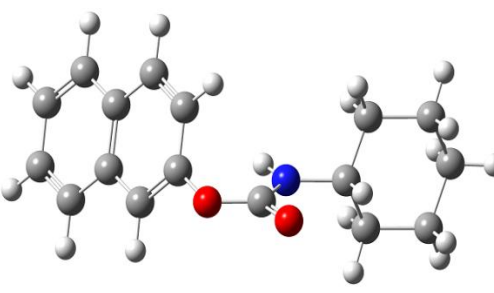
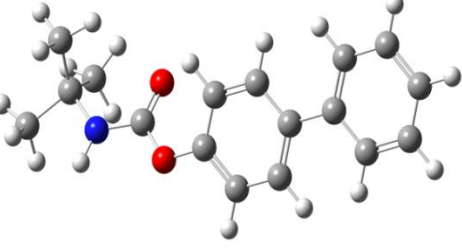
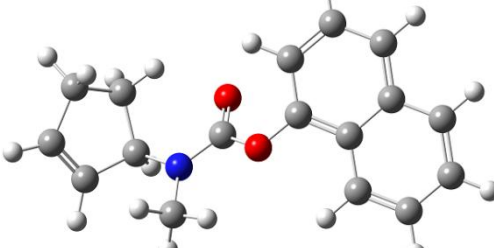


azanyl)phenyl methylcarbamate		
trimethyl(phenyl)-lambda~5~-azane		2-ethylhexyl N,N,N',N'-tetramethyldiamidophosphate
3-(diethylamino)phenyl methylcarbamate		No Name
1-naphthyl aminocarbonylcarbamate		trimethyl(6-(trimethyl-lambda~5~-azanyl)hexyl)-lambda~5~-azane
1,2,3,4-tetrahydrophenazine		2-((2-pyridinylimino)methyl)phenol

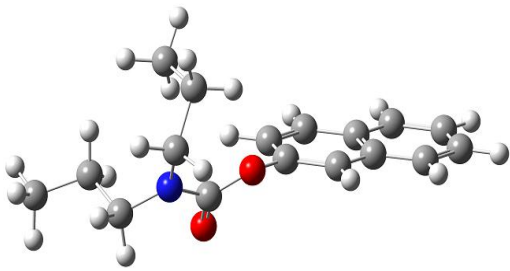
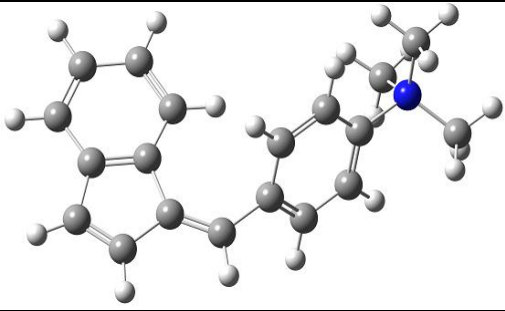
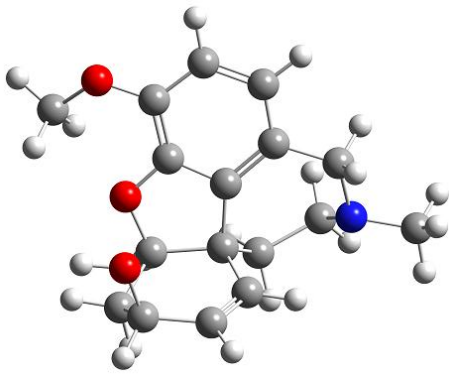
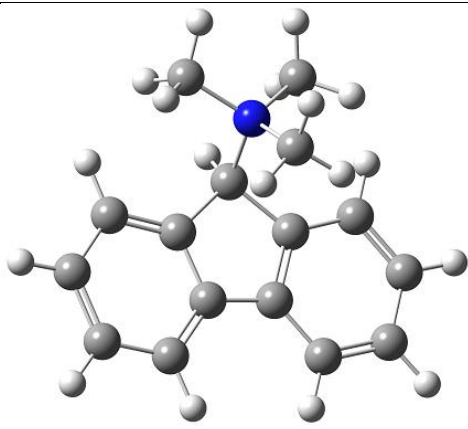
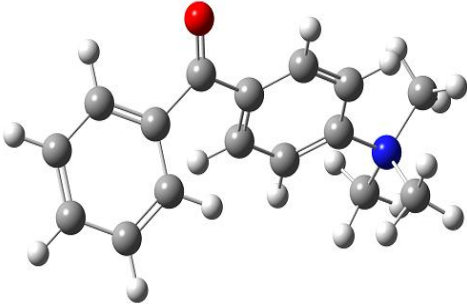
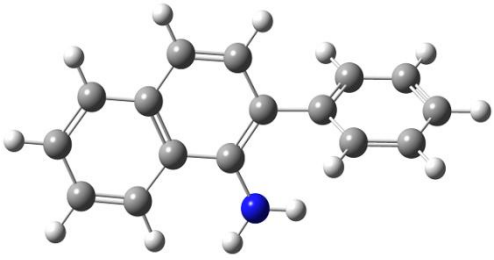
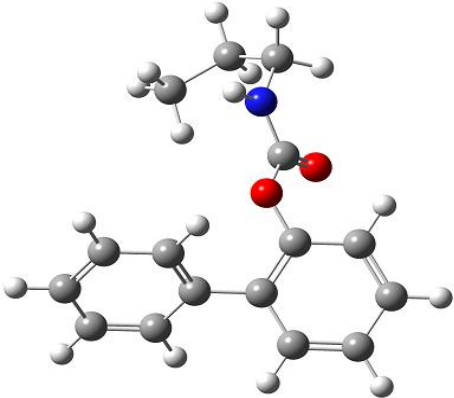
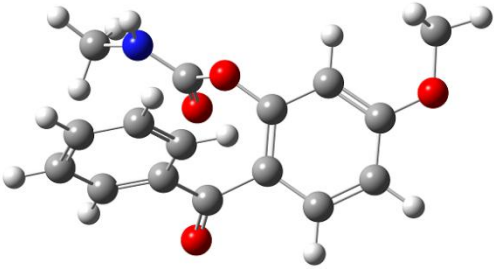


<p>(4-((hydroxy-lambda~5~-azanylidene)methyl)-3,5-dimethylphenyl)(trimethyl)-lambda~5~-azane</p>	<p>N-(4-chlorobenzylidene)-2-pyridinamine</p>
<p>(2-(dimethyl(propyl)-lambda~5~-azanyl)ethyl)(dimethyl)propyl-lambda~5~-azane</p>	<p>ethyl 4-(((4-(ethoxycarbonyl)anilino)methylene)amino)benzoate</p>
<p>4-(((3-(hydroxy(oxido)amino)anilino)carbonyl)oxy)-1,1'-biphenyl</p>	<p>N-(4-(dimethylamino)benzylidene)-1-adamantanamine</p>
<p>3-methylphenyl 1-naphthylcarbamate</p>	<p>1-naphthyl benzoylcarbamate</p>



	
1-(4-(trimethyl-lambda~5~-azanyl)phenyl)-1-nonen-3-one	4-(benzyloxy)phenyl butylcarbamate
	
2-naphthyl 2-methylphenylcarbamate	2-naphthyl 2-methoxyphenylcarbamate
	
3-(3-azabicyclo[4.2.1]non-3-yl)-1-(2-fluorophenyl)-1-propanone	(1-benzyl-1H-indol-3-yl)(oxo)acetic acid
	
1-benzyl-5-ethylideneazonane	2-naphthyl cyclohexylcarbamate
	



<p>4-(((tert-butylamino)carbonyl)oxy)-1,1'-biphenyl</p>	<p>1-naphthyl 2-cyclopenten-1-yl(methyl)carbamate</p>
	
<p>2-naphthyl dipropylcarbamate</p>	<p>(4-(1H-inden-1-ylidenemethyl)phenyl)(trimethyl)-lambda~5~-azane</p>
	
<p>galanthamine</p>	<p>9H-fluoren-9-yl(trimethyl)-lambda~5~-azane</p>
	
<p>phenyl(4-(trimethyl-lambda~5~-azanyl)phenyl)methanone</p>	<p>2-phenyl-1-naphthalenamine</p>
	
<p>2-(((propylamino)carbonyl)oxy)-1,1'-</p>	<p>2-benzoyl-5-methoxyphenyl methylcarbamate</p>



biphenyl	
triethyl(4-((triethyl-lambda~5~-azanyl)methyl)benzyl)-lambda~5~-azane	4a-methyl-9-(2-(1-piperidinyl)ethyl)-2,3,4,4a,9,9a-hexahydro-1H-carbazole
N'-(9-anthryl)-N,N-dimethylimidofornamide	No Name
4-anilinophenyl methylcarbamate	