

Universidad Católica de Temuco Facultad de Ingeniería Departamento de Ciencias Matemáticas y Física

# Empleo de métodos de la matemática estadística en la evaluación de descriptores moleculares para el diseño de nuevos fármacos anticancerígenos.

por

Oriana Ramírez Zumelzu

Profesor guía

Dr. Karel Mena Ulecia

Actividad Formativa Equivalente, para optar al grado de Magíster en Matemáticas Aplicadas

Temuco, 10 de enero de 2023

Universidad Católica de Temuco Facultad de Ingeniería Departamento de Ciencias Matemáticas y Física

## COMISIÓN EVALUADORA

Profesor guía:	
	Dr. Karel Mena Ulecia
Profesor informante:	
Profesor informante:	
Profesor informante:	
Director del Programa (Mini	stro de fe):
	Dr. Jacobo Hernández Montelongo

Temuco, 10 de enero de 2023

## Perfil de egreso

Magíster en Matemáticas Aplicadas. Universidad Católica de Temuco.

El egresado del Magíster en Matemáticas Aplicadas es un profesional postgraduado que posee la competencia de aplicar la matemática al análisis de sistemas dinámicos y evolutivos. Específicamente:

Formula ecuaciones diferenciales como modelos matemáticos para obtener una relación cuantitativa entre las variables relevantes de sistemas dinámicos y evolutivos.

Resuelve ecuaciones diferenciales como modelos matemáticos, utilizando técnicas numéricas y analíticas, para obtener valores cuantitativos de la variable respuesta del sistema.

Desarrolla y/o utiliza programas computacionales en la resolución, análisis y aplicación de ecuaciones diferenciales en sistemas dinámicos y evolutivos.

Comunica información científico-matemática con rigurosidad técnica y claridad.

## Agradecimientos

A mi profesor guía, a todos los que estan siempre ahí y a mi padrino que ya no esta.

### Abstract

Endothelial Growth Factor Receptor 2 (VEGFR2) is considered the most important marker for the development of endothelial cells. In particular, this receptor is directly related with the regulation of tumor angiogenesis, that is why several inhibitors have been developed of VEGFR2 and many of them are now in clinical trials. For the design of new inhibitors, the maximum mean inhibitory concentration (IC50) is a central step in the investigation pharmacological.

In this study, statistical tools will be used to generate models that can determine variables that can be used for the design of new anticancer drugs capable of to inhibit VEGFR2. For this, a total of 117 molecules with activities were selected from the literature. inhibitory properties against VEGFR2, and all of them were calculated for a total of 500 descriptors, where 13 of them were obtained from quantum mechanics.

#### Resumen

El Receptor 2 del Factor de Crecimiento Endoteliar (VEGFR2) se considera el marcador más importante para el desarrollo de células endoteliales. En particular, este receptor está directamente relacionado con la regulación de la angiogénesis tumoral, es por eso que se han desarrollado varios inhibidores de VEGFR2 y muchos de ellos se encuentran ahora en ensayos clínicos. Para el diseño de nuevos inhibidores, la concentración inhibitoria media máxima ( $IC_{50}$ ) es un paso central en la investigación farmacológica.

En este estudio se utilizarán herramientas estadísticas para la generación de modelos que puedan determinar variables que consigan ser usadas para el diseño de nuevos fármacos anticancerígenos capaces de inhibir el VEGFR2. Para esto se seleccionó de la literatura un total de 117 moléculas con actividades inhibitorias experimentales contra el VEGFR2, y a todas ellas se le calcularon un total de 500 descriptores, donde 13 de ellas fueron obtenidas de la mecánica cuántica.

## Índice

1.	Intr	roducción	1
	1.1.	Diseño de Fármacos	1
		1.1.1. Diseño de Fármacos basado en Estructura	2
		1.1.2. Diseño de Fármacos basado en Ligandos.	3
2.	Hip	ótesis y Objetivos	5
	2.1.	Hipótesis	5
	2.2.	Objetivo General	5
	2.3.	Objetivos específicos	5
3.	Mat	teriales y Métodos	6
	3.1.	Optimización de Geometría y Cálculos de Descriptores Moleculares	6
	3.2.	Modelos derivados de la Matemática Estadística	8
		3.2.1. Random Forest (RF)	8
		3.2.2. Modelo de Regresión Lineal Multivariable $\hdots$	8
	3.3.	Árbol de Desición	9
4.	Res	ultados y Discusión	11
	4.1.	Pre-Tratamiento de los Datos	11
	4.2.	Modelos de Regresión Lineal Multivariable.	13
		4.2.1. Validación de los Modelos de Regresión Lineal Multivariable	15
	4.3.	Árbol de Decisión	16
5.	Con	nclusiones	19
	5.1.	Proyección Futura	19
6.	Bib	liografía	20
7.	Ane	exos	30

#### 1. Introducción

La angiogénesis es la formación de vasos sanguíneos nuevos[1]. Este proceso consiste en la migración, crecimiento y diferenciación de células endoteliales que recubren las paredes internas de los vasos sanguíneos[2]. El proceso de angiogénesis está controlado por señales químicas, como el factor vascular de crecimiento endotelial (VEGF). Algunas veces estas señales pueden desequilibrarse causando un crecimiento descontrolado de vasos sanguíneos que puede resultar en estados anormales o enfermedades.

Debido a que los tumores no pueden crecer más allá de un tamaño determinado ni extenderse sin un suministro de sangre, los científicos han formulado fármacos llamados inhibidores de la angiogénesis. El objetivo de estos fármacos es impedir o hacer lento el crecimiento del cáncer al suprimir el suministro de sangre que necesita.

En este proceso de angiogénesis, el Receptor-2 del Factor de Crecimiento Endotelial Vascular (VEGFR2) tiene una parte importante en el proceso de progresión de tumores sólidos[3, 4], es por esto que se ha convertido en un objetivo terapéutico atractivo en el tratamiento del cáncer[5, 6, 7].

#### 1.1. Diseño de Fármacos

El enfoque tradicional del diseño de fármacos en una primera etapa identifica la diana farmacológica que se quiere tratar, se obtienen los compuestos activos generalmente de las plantas que son moléculas que poseen una actividad biológica y se aíslan sus componentes por técnicas experimentales y se realizan ensayos preclínicos de cada uno de estos componentes en ratones, perros, otras especies y en una última instancia en humanos lo cual se puede demorar en la construcción de un fármaco efectivo de 10 a 15 años[8].

El diseño racional de fármacos, constituye una herramienta casi indispensable en el desarrollo actual de nuevos medicamentos[9]; si se toma en cuenta que el paso limitante en el descubrimiento y desarrollo de nuevos fármacos continúa siendo la identificación y la optimización de compuestos líderes de un modo efectivo, este contribuye a un aumento de las posibilidades de éxito y a una disminución de los costos[10]. Este procedimiento engloba todas las técnicas asistidas por ordenadores usadas en el diseño descubrimiento y optimización de compuestos con propiedades deseadas y ha jugado un rol fundamental en el desarrollo de un número de fármacos que están ahora en el mercado[11]. Este tipo de procedimiento *in silico* evita los procesos actuales de síntesis y bioensayos, los cuales se hacen solamente después de la exploración de los conceptos iniciales con modelos computacionales[12].

En el diseño asistido por computadora ya esta identificada la diana que es la molécula específica del organismo sobre la que un medicamento está diseñado para actuar, en algunos casos están en exceso durante una enfermedad por lo cual el objetivo del medicamento sería bloquear su actividad, o puede existir una deficiencia de esta diana y se requiere aumentarla o reemplazarla, desde este punto si existe los descriptores moleculares de esta se puede realizar un diseño nuevo o una optimización de la estructura y así mediante una simulación molecular se modela, imita o predice el comportamiento de las moléculas y se puede tener las primeras formulaciones de fármacos en un tiempo de 5 a 10 años[13]. Dentro del diseño de fármacos asistido por computadoras podemos encontrar en la literatura varios enfoques, sin embargo los más usados actualmente son[14, 15, 16, 17]:

- Diseño de Fármacos basado en Estructura.
- Diseño de Fármacos basado en Ligandos.

#### 1.1.1. Diseño de Fármacos basado en Estructura.

El diseño de fármacos basado en la estructura se suele aplicar en las primeras etapas de la investigación.Este enfoque se basa fundamentalmente en el conocimiento de las estructuras tridimensionales de las dianas (proteínas) sobre las que se quere actuar[18]. Dentro de las fuentes más comunes, para obtener estructuras tridimensionales de moléculas, se encuentra el DrugBank (http://drugbank.ca/), ChemBL (https://www.ebi.ac.uk/chembl/), entre otros, lo cual nos permite saber cual es el centro activo o lugar de unión del ligando[19].

Dentro de los métodos sobre diseño de fármacos basado en la estructura, se destaca el diseño de *novo*, que consiste en construir moléculas directamente en el receptor (proteína)[20]. El construir las moléculas en el mismo sitio de unión implica a su vez utilizar el método de Acoplamiento Molecular (*Docking*), el cual, es el más empleado dentro de las investigaciones en el diseño de fármacos[18]. El principal objetivo de este método (*Docking*) es encontrar el modo de unión más favorable de un ligando dentro del receptor (Proteína)[21].

Para lograr buenas poses en el acoplamiento molecular se requiere de dos componentes que pueden ser caracterizados como la etapa de búsqueda y la etapa de evaluación[22]; etapa de búsqueda se refiere a la exploración del espacio configuracional accesible para el ligando dentro del receptor. El objetivo de esta exploración es encontrar la orientación y conformación del ligando que corresponda al mínimo global de la energía libre de unión[23]. La etapa de evaluación se refiere a la asignación de un valor numérico a cada una de las configuraciones generadas durante la etapa de búsqueda conformacional[19], esto permite establecer un orden entre las diferentes posiciones y configuraciones encontradas[24]. Otra aplicación del acoplamiento molecular es la evaluación *in silico* de una gran cantidad de moléculas organizadas en lo que se conocen como Virtual Screening[25, 26].

#### 1.1.2. Diseño de Fármacos basado en Ligandos.

El diseño de fármacos basado en ligandos depende de la información experimental disponible para una serie de estructuras químicas con actividad biológica conocida, el método más utilizado en esta área son los estudios de relaciones estructura-actividad, más conocidos como *Métodos QSAR*[27, 28, 18].

El QSAR (*Quantitative Structure-Activity Relationship*) puede definirse como la aplicación de métodos matemáticos y estadísticos en la predicción de relaciones de alguna propiedad biológica[29]. Este método puede analizarse desde dos puntos de vista, el primero es a través de parámetros estructurales que se utilizan para caracterizar a las moléculas, desde fórmulas químicas a estructuras tridimensionales[30]; y por otro lado, los procedimientos matemáticos que se emplean para obtener las relaciones cuantitativas entre los descriptores moleculares y la actividad biológica[31].

De acuerdo con el origen de los descriptores moleculares usados en los cálculos, el método QSAR pueden dividirse en tres grupos. El primero usa un número relativamente pequeño de propiedades fisicoquímicas y parámetros que describen efectos hidrófobos, estéricos y electrostáticos. Estos métodos son conocidos con el nombre de análisis de Hansch[32, 33]. El segundo grupo se basa en las características cuantitativas de las fórmulas estructurales bidimensionales. Estos métodos se conocen como estudios QSAR en dos dimensiones o QSAR 2D[34]. Finalmente el tercer método se basa en los descriptores obtenidos de la representación tridimensional de las estructuras químicas a estudiar (3D QSAR)[35].

Después de desarrollar un modelo que correlacione la actividad biológica con los descriptores moleculares, la siguiente etapa de un estudio QSAR es la validación de los modelos, la cual consiste en evaluar la capacidad que tiene el modelo para predecir con exactitud la propiedad de interés de los compuestos que no fueron utilizados en el desarrollo del modelo[36, 12]. Los descriptores moleculares antes mencionados son el resultado de un procedimiento matemático que transforma la información química de una molécula en un número que las describe, estos pueden ser de diferentes tipos y ello depende de la complejidad de la información. Así, algunos descriptores son bastantes globales y por ende muy simples y otros pueden ser muy complejos[24, 37].

Los modelos *in silico* capaces de predecir con precisión la concentración inhibitoria  $IC_{50}$  (Actividad Biológica) pueden ser de gran beneficio, permitiendo centrar el proceso experimental en los compuestos más prometedores. Los modelos de aprendizaje automático se han utilizado ampliamente en la quimioinformática y el descubrimiento de fármacos[38]. Estos procedimientos se realizaron exitosamente en la predicción de la concentración inhibitoria media para la enzima humana P450 2D6[39]; y recientemente, se evaluaron varios modelos de aprendizaje automático, descriptores moleculares para la enzima convertidora de angiotensina (ACE) en el diseño de agentes antihipertensivos[40]. Las predicciones del modelo QSAR se validaron empíricamente con éxito en el aprendizaje automático para descubrir inhibidores duales de los receptores del factor de crecimiento epidérmico y de fibroblastos[41], en los cuales se estudiaron a través máquinas de vectores de soporte, árbol de decisión dentro de otros métodos[41].

En este estudio se utilizarán herramientas de la *Matemática Estadística* en la generación de modelos para determinar las variables que pueden ser usadas para el diseño de nuevos fármacos anticancerígenos capaces de inhibir el Receptor del Factor de Crecimiento Endoteliar. Para lograr lo planteado se trazaron los objetivos que se describen a continuación.

## 2. Hipótesis y Objetivos

#### 2.1. Hipótesis

El cálculo de descriptores moleculares basados en la Teoría del Funcional de la Densidad (DFT) influyen en el comportamiento biológico de las moléculas estudiadas y se podrían utilizar estas variables para el diseño de nuevos fármacos anticancerígenos contra el Receptor 2 del Factor de Crecimiento Endoteliar.

#### 2.2. Objetivo General

Desarrollar modelos basados en las relaciones cuantitativas estructura-actividad de nuevos inhibidores del Receptor 2 del factor de crecimiento endotelial para el diseño de nuevos fármacos anticancerígenos empleando métodos de la matemática estadística.

#### 2.3. Objetivos específicos

- 1. Calcular descriptores moleculares mecanocuánticos y del modelado molecular en el estudio de nuevos inhibidores del factor de crecimiento endotelial para el diseño de nuevos fármacos anticancerígenos.
- 2. Seleccionar las variables más relevantes a través de la técnica de clasificación Random Forest.
- 3. Generar y validar modelos basados en la estructura-actividad utilizando métodos de la matemática estadística: Regresión lineal múltiple (MLR) y Árbol de decisión.

#### 3. Materiales y Métodos

El crecimiento endotelial es un factor angiogénico (VEGFR2) importante ya que este promueve el crecimiento de vasos sanguíneos nuevos que los tumores necesitan para crecer, por lo tanto, una sobre expresión de esa proteína está asociado a la aparición de tumores malignos, por lo cual hay que tratar de inhibirla. La búsqueda de nuevos inhibidores con la menor cantidad de reacciones secundarias es un tema de actualidad[42, 43] y la generación de modelos QSAR con descriptores moleculares derivados de la Teoría del Funcional de la Densidad (DFT) es de mucha utilidad[44, 45].

Para lograr los objetivos planteados en este estudio, se partió con un total de 117 moléculas las cuales fueron obtenidas de la literatura (Anexo-1)[46, 47, 1, 4]. Estos compuestos tuvieron valores experimentales de inhibición de la enzima VEGFR2 (Anexo-1) ( $IC_{50}$ ), las cuales fueron reunidas para cada una de ellas y corresponde a la variable a predecir en los modelos estadísticos para el diseño de nuevos fármacos anticancerígenos. Debido al amplio rango numérico de actividades inhibitorias ( $IC_{50}$ ) de las moléculas objeto de estudio, se ejecutó la operación  $log(1000/IC_{50})$  para normalizar los datos y obtener valores menos variables. En este estudio, los valores de  $IC_{50}$  superiores a 1.000 se consideran moléculas inactivas debido a la alta concentración de estos compuestos (mM) necesarios para ejercer su acción inhibitoria[48]. Todas las moléculas se dibujaron utilizando el paquete de software Avogadro 1.2.0 [49], obteniéndose la estructura tridimensional de cada una.

#### 3.1. Optimización de Geometría y Cálculos de Descriptores Moleculares

Una vez obtenida la estructura tridimensional de los compuestos objeto de estudio se procedió a la optimización de la geometría de cada molécula. Esta tiene como objetivo fundamental obtener el mínimo de energía de cada inhibidor resolviendo la *Ecuación de Schrödinger*[50] independiente del tiempo(ecuación-1):

$$i\hbar\frac{\partial\phi(x,t)}{\partial t} = -\left[\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\phi(x,t) = H\phi \tag{1}$$

donde V(x) es la energía potencial a la que está sometida la partícula de masa m (que supondremos independiente del tiempo),  $\hbar$  es la constante de Planck reducida y H es el operador Hamiltoniano, que es hermítico.

Dado que la *Ecuación de Schrödinger* solo tiene solución para una partícula (el átomo de hidrógeno), para obtener el mínimo de energía de cada compuesto es necesario utilizar aproximaciones. En este caso cada estructura molecular se optimizó usando el funcional híbrido de Perdew-Burke-Ernzerhof (PBE0)[51] con el conjunto de bases  $6 - 311^{++}g(d, p)$  usando Orca 4.2.1 [52, 53]. Las estructuras optimizadas se comprobaron contando el número de frecuencias imaginarias.

Con la estructura optimizada de cada molécula se calculó un total de 13 descriptores moleculares químico-cuánticos derivados de la Teoría del Funcional de la Densidad (DFT) [54, 44, 45]. La tabla 1 muestra los detalles para calcular cada observable químico.

Molecular Descriptor	Formulation	References
$\epsilon_H$	energies of the HOMO orbital	-
$\epsilon_L$	energies of the LUMO orbital	-
$\Delta \epsilon_H - \epsilon_L$	HOMO-LUMO gap	-
Ionization Potential (I)	$I = -\epsilon_H$	[56, 57]
Electroaffinity (A)	$I = -\epsilon_L$	[56, 57]
Chemical Potential $(\mu)$	$\mu = \frac{1}{2}(\epsilon_L + \epsilon_H)$	[56, 58]
Global Hardness $(\eta)$	$\eta = \frac{1}{2}(\epsilon_L - \epsilon_H)$	[59, 56]
Global Smoothness $(S)$	$S = \frac{1}{2\eta} (\epsilon_L - \epsilon_H)$	[57,  60]
Electron egativity $(\chi)$	$\chi = -\frac{1}{2}(\epsilon_L + \epsilon_H)$	[58]
Global electrophilicity ( $\omega$ )	$\omega = \frac{(\epsilon_L + \epsilon_H)^2}{2(\epsilon_L - \epsilon_H)}$	[45,  61]
Electroacceptor power $(\omega^+)$	$\omega^+ = \frac{(\epsilon_L + 3\epsilon_H)^2}{16(\epsilon_L - \epsilon_H)}$	[62]
Electrodonor Power ( $\omega^-$ )	$\omega^- = \frac{(3\epsilon_L + \epsilon_H)^2}{16(\epsilon_L - \epsilon_H)}$	[62]
Net Electrophilicity $(\Delta \omega_{-}^{+})$	$\Delta \omega_{-}^{+} = \omega^{+} + \omega^{-}$	[63]

Tabla 1: Descriptores Moleculares derivados de la Teoría de Koopmans [55, 56].

La Teoría de Koopmans plantea el cálculo de descriptores moleculares basados en la enería de los orbitales moleculares HOMO (Highest Occupied Molecular Orbital) y LUMO (Lowest Unoccupied Molecular Orbital)[55, 56]. Además de estos descriptores moleculares, se calcularon otros 500 utilizando el programa PaDel Descriptor versión 2.21[64](Anexo-2). Todos estos datos hacen una matriz de 514 columnas (descriptores moleculares) y 117 filas (moléculas). Para la generación de los modelos matemáticos predictivos, toda la data se subdividió de manera aleatoria en un 80% para el test de entrenamiento (*training set*) de donde se obtendrá un modelo que indique cual es el descriptor molecular que más está influyendo en el comportamiento biológico de la molécula, y el 20% restante se utilizó para el test de validación (Validación Externa) (*test set*) el cual consiste en evaluar la precisión con la que el modelo predictivo funcionará en la práctica. La división entre el *training set* y el *test set* fue realizada mediante el programa RStudio[65]. Con esta subdivisión realizada se procedió a la preparación de los datos para la generación de los modelos de relación estructura-actividad, utilizando los métodos estadísticos que se describen a continuación, todas ellas ejecutadas en RStudio.

#### 3.2. Modelos derivados de la Matemática Estadística

#### 3.2.1. Random Forest (RF)

El Método de *Random Forest* está basado en árboles de decisión, en el cual se particiona el espacio muestral en hipercubos, asumiendo que cada descriptor se representa como un vector n-dimensional, esto se traduce en reducir la dimensionalidad de un conjunto de datos, de modo que se podría comenzar con 30 variables y solo terminar con 2 o 3 más significativas, lo que se conoce como componentes principales.

En *Random Forest* se ejecutan varios algoritmos de árboles de decisión en lugar de uno solo. Para clasificar un nuevo objeto basado en atributos, cada árbol de decisión da una clasificación y finalmente la decisión con mayor *"votos"* es la predicción del algoritmo.

En este método el árbol de decisión depende de un conjunto de variables aleatorias. Éstos particionan recursivamente el espacio conformado por las p variables independientes x, de forma tal que las regiones generadas por dichas particiones contienen observaciones cuyos valores respecto a la variable dependiente son lo más homogéneas posibles. De este modo los árboles logran capturar con bajo sesgo las relaciones existentes entre las variables predictoras.

#### 3.2.2. Modelo de Regresión Lineal Multivariable

El *Modelo de Regresión Lineal Multivariable* describe de manera predictiva la relación entre la estructura (representada por descriptores numéricos) y la actividad biológica que representa la molécula, utilizando la siguiente ecuación:

$$log(\frac{1}{C}) = a_0 + a_1 x_1 + \ldots + a_p x_p + \varepsilon$$
<sup>(2)</sup>

donde:

 $a_0$ : valor de la variable respuesta cuando las explicativas son cero.

 $a_p$ :coeficientes parciales de regresión.

 $\varepsilon:$ error aleatorio, diferencia entre el valor observado y estimado por el modelo.

#### 3.3. Árbol de Desición

El aprendizaje mediante árboles de decisión es un método de aproximación de una función objetivo de valores discretos en el cual la función es representada mediante un árbol de decisión[66]. Los árboles de decisión se pueden usar para generar sistemas expertos, búsquedas binarias entre otros[67]. Los árboles generalmente son binarios, es decir, cuentan con dos opciones, aunque esto no significa que no puedan existir árboles de tres o mas opciones[68, 69].

El árbol de decisión basado en la búsqueda binaria es un método en el cual se parte el árbol en dos como se muestra en la siguiente figura (figura-1)[70]:



Figura 1: Representación gráfica del método Árbol de Decisión basado en búsqueda Binaria.

El árbol de decisión basado en sistemas expertos son más precisos para poder desarrollar un diagnóstico con respecto a algo, ya que el hombre puede dejar pasar sin querer un detalle, en cam-

bio la máquina mediante un sistema experto con un árbol de decisión puede dar un resultado exacto[71].

Una deficiencia de este método es que puede llegar a ser más lento pues analiza todas las posibilidades, pero esto a su vez es lo que lo vuelve más preciso[72].

La siguiente figura presenta un ejemplo de un sistema experto y de como puede llegar a diagnosticar que se emplee un fármaco determinado en una persona con presión arterial[70]. Este tipo de sistema fue el que se empleo en este estudio para predecir si el fármaco es activo o no.



Figura 2: Representación gráfica del método Árbol de Decisión basado en Sistema Experto.

#### 4. Resultados y Discusión

En este capítulo se abordará los resultados obtenidos en este estudio, iniciando este acápite con la descripción del pre-tratamiento de los datos antes de generar los modelos predictivos derivados de la Matemática Estadística.

#### 4.1. Pre-Tratamiento de los Datos

El proceso de pre-tratamiento de los datos se realiza con el objetivo de eliminar descriptores moleculares (variables) que no tienen significado físico y/o químico, debido a las características estructurales de las moléculas que fueron utilizadas en este estudio. Este procedimiento también es utilizado para la homogenización de los datos que se utilizan en la generación y validación de los modelos[73, 74].

En el procedimiento de pre-tratamiento de los datos, la matriz de 514 columnas (descriptores moleculares) y 117 filas (moléculas), la cual contiene variables de tipo numérica, fue sometida a un proceso de *Normalización de los datos* mediante el comando *"Normalize"* implementado en el software Weka versión 3.8[75, 76, 77, 78]. Este comando normaliza cada una de las variables que están contenidas en la matriz numérica, abarcando cada registro entre los valores de 0 y 1.

Una vez realizado el pre-tratamiento de los datos, la matriz fue sometida a un proceso de reducción de la dimensionalidad con el objetivo de eliminar las variables que no son relevantes en la predicción experimental del  $IC_{50}$ . La reducción de dimensionalidad de los datos se realizó mediante el algoritmo *Random Forest* para un nivel de confiabilidad de un 95 %, implementado en el programa Weka versión 3.8[75, 76, 77, 78]. Los resultados de este modelo se muestran en la siguiente tabla:

Como se observa en la tabla-2, después de 100 iteraciones de cálculo<sup>1</sup>, el mejor modelo obtenido mediante la técnica *Random Forest* tiene un coeficiente de correlación  $R^2 = 0,871$ , lo que se considera bueno en estadística[79, 75].

Es necesario destacar que según el modelo calculado, solo 15 de los 514 descriptores calculados

<sup>&</sup>lt;sup>1</sup>El número de 100 iteraciones considera la construcción de 100 árboles de decisión con las diferentes combinaciones de variables obteniéndose el mejor modelo el cual está representado en la tabla-2

Parámetros del Modelo	Resultados del Modelo				
Instancias	117				
Atributos	514				
Número de Iteraciones	100				
Coeficiente de Correlación $(R^2)$	0,871				
Media del Error Absoluto	0,1146				
Error Cuadrático Medio	0,1532				
Atributos Seleccionados	15 (Ionization-Potential, Global Soft-				
	ness,  Electron-Aceptor, Chemical-Potential,				
	SpMin8-Bhm, SpMax7-Bhv, SpMin8-Bhs,				
	SpMax4-Bhe, SpMin8-Bhe, SpMax7-Bhs, nH-				
	Don,nHDon-Lipinski,n7Ring,JGI10,VR2-D)				

Tabla 2: Resultados obtenidos de la implementación del modelo Random Forest

tienen relevancia en la predicción del comportamiento experimental de las moléculas estudiadas. De estos 15 descriptores relevantes, cuatro de ellos fueron calculados mediante la Teoría del Funcional de la Densidad (DFT)[80] derivados de la resolución de la ecuación de *Ecuación de Schrödinger* [50] independiente del tiempo, los cuales son considerados como descriptores de reactividad global **(Potencial de Ionización, Suavidad Global, Poder Electroaceptor y el Potencial Químico)**. Estos cuatro descriptores están relacionados con la estructura electrónica de las moléculas objeto de estudio, más específicamente, con la densidad electrónica molecular indicando que las atracciones electrostáticas son importantes en la estabilidad de los complejos enzima-inhibidor.

Según los resultados que se muestran en la tabla-2, de los 15 descriptores relevantes en la predicción de la capacidad inhibitoria de las moléculas estudiadas, casi la mitad de estos son descriptores topológicos derivados de la resolución de los valores propios modificados de las matrices carga mínimas y máximas para cada átomo de cada molécula (SpMin8-Bhm, SpMax7-Bhv, SpMin8-Bhs, SpMax4-Bhe, SpMin8-Bhe y SpMax7-Bhs), reafirmando que el componente electrostático (cargas y densidad de cargas) es importante en la estabilidad del complejo enzima-inhibidor.

Otro de los resultados importantes obtenidos mediante la técnica de Random Forest es que dos

de los descriptores relevantes tienen que ver con la cantidad de sitios o grupos de las moléculas estudiadas capaces de donar hidrógenos en las interacciones por puente de hidrógenos(h-bond), estos descriptores son: **nHDon y nHDon-Lipinski**. Las interacciones por puente de hidrógenos son fundamentales en la estabilidad de los complejos enzima-inhibidor[37, 24], es por esta razón que en el diseño de nuevos inhibidores contra VEGFR2, se busca que la molécula tenga varios grupos capaces de donar hidrógenos dentro del centro activo de la proteína para garantizar la estabilidad del complejo.

Los otros tres descriptores relevantes obtenidos mediante la técnica de Random Forest tienen que ver con las características topológicas de las moléculas, por ejemplo el descriptor **n7Ring** se refiere a la cantidad de anillos aromáticos que presenta una molécula. Este descriptor es importante en las interacciones  $\pi - \pi stacking$  que son interacciones estabilizantes entre aminoácidos aromáticos del centro activo de nuestra proteína de interés y el inhibidor a diseñar. Los otros dos descriptores tienen que ver con la topología<sup>2</sup> de las moléculas estudiadas. El descriptor **JGI10** tiene en cuenta la carga topológica en la matríz de distancia atómica de cada molécula y el **VR2-D** es un descriptor que describe la matriz de distancia topológica en cada una de las moléculas estudiadas.

Con estos 15 descriptores se continúa el estudio para ratificar si estas variables pudieran predecir el comportamiento experimental de cada una de las moléculas estudiadas, para tenerlas en cuenta en el diseño de nuevos inhibidores del VEGFR2 como posibles agentes anticancerígenos. Para lograr este objetivo se utilizan los modelos de regresión lineal multivariable cuyos resultados se muestran a continuación.

#### 4.2. Modelos de Regresión Lineal Multivariable.

Con los 15 atributos que se seleccionaron en el acápite anterior, se realizó una corrida de 20 modelos de regresión, con el objetivo de predecir la capacidad inhibitoria  $(IC_{50})$  de las moléculas estudiadas. En la implementación de este método, los datos se dividieron en test de entrenamiento (94 compuestos) para la generación de los modelos y 23 moléculas para la validación de estos (Validación externa).

 $<sup>^{2}</sup>$ La topología de la molécula tiene que ver con la distancia entre los átomos y la densidad de carga local en cada uno de los átomos.

Modelo	Ecuación	Parámetros Estadísticos
Modelo-2	$IC_{50} = 0,0096$ IO-	$R^2 = 0,9873$
	$0,\!0771\mathrm{CP}{+}0,\!0782\mathrm{GS}{+}0,\!6084\mathrm{EA}{+}0,\!0625$	MAE = 0,0665
		RMSE = 0,0778
Modelo-4	$IC_{50}$ =-0,0024SpMin8-Bhm-0,2105SpMax7-	$R^2 = 0,9398$
	$Bhv{+}0,1629SpMin8{-}Bhv{+}0,0506SpMax4{-}$	MAE = 0,0114
	Bhe+0,7949	RMSE = 0,0163
Modelo-6	$IC_{50} = 0.0584 \text{SpMax7-}$	$R^2 = 0,8938$
	$Bhs{+}0,\!3744nHDon{+}0,\!1283nHDon{-}$	MAE = 0,0346
	Lipinski+0.027n7Ring+0.0512	RMSE = 0,0434
Modelo-11	$IC_{50}{=}{-}0{,}0036\mathrm{JGI10}{+}0{,}9984\mathrm{VR2}{-}\mathrm{D}{+}0{,}0223$	$R^2 = 0,84438$
		MAE = 0,0146
		RMSE = 0,0197

Tabla 3: Resultados de los mejores 4 modelos de regresión lineal multivariante obtenidos

IO: Potencial de Ionización, GS: Global Softness (Dureza global), EA: Poder Electroaceptor, CP es el Potencial Químico, SpMin8-Bhm, SpMax7-Bhv, SpMin8-Bhs, SpMax4-Bhe, SpMin8-Bhe, SpMax7-Bhs son descriptores derivados de la resolución de los valores propios modificados de las matrices carga mínimas y máximas para cada átomo de cada molécula y nHDon, nHDon-Lipinski, n7Ring, JGI10,VR2-D son descriptores topológicos Moleculares, MAE es la media del error absoluto y RMSE es el error cuadrático medio.

Los modelos de regresión fueron generados realizando diferentes combinaciones de dos, tres, cuatro y cinco variables. En la tabla-3 se muestran los mejores resultados de los 20 modelos de regresión que se corrieron. El criterio de sección de estos modelos fue el coeficiente de correlación de las variables dentro de los modelos  $(R^2)$ , la media del error absoluto del método (MAE) y el error cuadrático medio (RMSE). Como se puede observar, de todos los modelos generados, los mejores resultados se obtuvieron con la combinación de 2 y 4 variables.

Si se observa el coeficiente de correlación  $(R^2)$ , el mejor modelo que se obtuvo fue combinando 4 variables derivadas de cálculos cuánticos, estas son: el potencial de ionización, la suavidad global, el potencial químico y el poder electroaceptor. Esto revela que los descriptores anteriormente mencionados están seriamente interrelacionados, lo que justifica la inclusión de estas variables en estudio y también la importancia de estas en el diseño de nuevos agentes anticancerígenos, lo que concuerda con los resultados que se obtuvieron mediante el método de *Random Forest*.

#### 4.2.1. Validación de los Modelos de Regresión Lineal Multivariable.

Para la validación de los modelos de Regresión Lineal Multivariable, se tomaron el *set* de moléculas que están en el *test de validación* (23 moléculas, el 20 % de los datos originales). Con estas moléculas y tomando en consideración cada una de las 4 ecuaciones lineales obtenidas con el *Test de Entrenamiento*, la cuales correspondían a los mejores modelos obtenidos, se calculó la  $IC_{50}$ .



Figura 3: Validación externa de los mejores Modelos de Regresión Lineal Multivariable.

Esta nueva  $IC_{50}$  es una predicción del comportamiento experimental de las moléculas contenidas en el proceso de validación externa. Una vez calculada la nueva  $IC_{50}$ , se realizó un proceso de *Regresión Lineal* entre la  $IC_{50}$  experimental y la  $IC_{50}$  calculada, en la cual se sacaron los coeficientes de correlación  $(R^2)$  entre ambas variables como criterio de validación. Como se observa en la figura-3, los 4 modelos predicen muy bien el comportamiento experimental de las moléculas contenidas en el test de validación externa. El modelo que mejor predicción tuvo fue el modelo-11 con dos variables a tener en cuenta derivados de los descriptores topológicos moleculares con un  $R^2 = 0,9880$ . El segundo mejor modelo de validación fue el Modelo-2 con un  $R^2 = 0,9744$ . Este modelo comprende 4 variables, todas ellas derivadas de cálculos mecanocuánticos indicando que éstos descriptores, junto con los topológicos son fundamentales en el diseño de nuevos inhibidores del VEGFR2 como posibles agentes anticancerígenos.

#### 4.3. Árbol de Decisión

Las técnicas de árboles de decisión se han utilizado ampliamente para construir modelos de clasificación, debido a que estos modelos se asemejan mucho al razonamiento humano y son fáciles de analizar[81].

Para comprender mejor el clasificador de árbol de decisión obtenido, los datos numéricos de cada atributo se normalizaron entre [0..1]. Además, el descriptor de clase representa el atributo de destino para predecir  $IC_{50}$ , y sus valores correspondientes significan: [0: activo, 1: inactivo]. En este caso se utilizaron los mismos 15 descriptores moleculares que se usaron en la generación y validación de los modelos de regresión lineal multivariable.

El mejor clasificador obtenido (árbol de decisión para k=15 número de descriptores) se muestra en la figura-4. Como se puede observar de los 15 clasificadores ensayados, 11 de ellos tuvieron influencia en la predicción de compuestos activos e inactivos. Las variables de mayor incidencia de la predicción  $IC_{50}$  para las moléculas estudiadas fueron: Potencial de Ionización, Poder Electroaceptor y la *Global Softness* (Suavidad Global).

El Potencial de Ionización, conceptualmente hablando es un descriptor molecular obtenido mediante cálculos mecanocuántico e indica la energía mínima necesaria para arrancar un electrón a una molécula que está en su estado fundamental[82]. Este descriptor es fundamental en la estabilidad de los complejos enzima inhibidor puesto que átomos dentro de la estructura de posibles inhibidores capaces de ceder electrones[83] y cargarse negativamente son importantes en la interacción con aminoácidos cargados negativamente como son los casos del ácidos glutámicos y aspárticos en el centro activo



Figura 4: Clasificador Árbol de Decisión del mejor modelo obtenido.

de la proteína objeto de estudio[84, 85, 3, 7]. El poder del electroaceptor es un descriptor de reactividad química que describe la capacidad de la molécula para aceptar electrones del medio ambiente.[60].

La mayoría de los compuestos estudiados en este estudio tienen un alto Potencial de Ionización y un alto Poder Electroaceptor. De acuerdo con este hecho, los modelos estudiados predijeron bien la  $IC_{50}$  para el Receptor-2 del Factor de Crecimiento Endotelial Vascular (VEGFR2). Esto sugiere que para el diseño de nuevos inhibidores contra esta proteína, se necesitan moléculas con grupos que tengan esta propiedad.

La mayoria de los ligandos estudiados presentan grupos capaces de atraer electrones porque en el centro activo de la proteína se encuentran aminoácidos clave con densidad de carga negativa como el ácido glutámico-885 (GLU885) y el ácido aspártico-1046 (ASP1046) de la cadena A que podrían estabilizar el complejo[86, 6].

Otro descriptor molecular que se comportó de buena manera fue la Suavidad Global (Global

Softness). Este parámetro tiene relación con los electrones de valencia y juegan un papel central en el establecimiento del principio de ácidos y bases duros y blandos[87]. Los resultados obtenidos de esta variable en todas las moléculas analizadas fueron altos en comparación con otros compuestos de origen orgánico[88, 87]. De acuerdo con la estructura de ligandos (Anexo-3), por lo tanto, se espera que los resultados para este parámetro tengan un comportamiento similar al analizado anteriormente.

De acuerdo con la bibliografía consultada, los resultados concuerdan con los obtenidos por varios autores, quienes han sugerido que los descriptores de reactividad anteriormente mencionados caracterizan la sensibilidad de los ligandos a un ataque nucleofílico[89, 61, 63]. El centro activo del VEGFR2 tiene un marcado carácter nucleofílico[90], por lo que moléculas cuyas estructuras químicas tengan grupos con marcada densidad de carga negativa puede donar una cierta cantidad de densidad de electrones durante la interacción aminoácidos dibásicos quedándose con déficit de  $e^-$  capaces de estabilizarse electrostáticamente con GLU885 y ASP1046[91].

#### 5. Conclusiones

Para el diseño de nuevos inhibidores del Receptor-2 del Factor de Crecimiento Vascular Endoteliar, se han utilizado descriptores de reactividad derivados de la teoría funcional de la densidad (DFT) y otros descriptores tanto topológicos como de matriz de distancias, para predecir las relaciones estructuraactividad ( $IC_{50}$ ) utilizando métodos de la Matemática Estadística. Aplicados los métodos descritos anteriormente se obtuvieron las siguientes conclusiones:

- De los 514 descriptores moleculares calculados, solo 15 de ellos tuvieron relevancia en la predicción del comportamiento experimental de las moléculas estudiadas.
- De todos los descriptores utilizados, el Potencial de Ionización, El Poder Electroaceptor y la Suavidad Global (*Global Softness*) fueron los de mayor incidencia en la predicción de *IC*<sub>50</sub> para todas las moléculas estudiadas en los tres métodos estudiados.
- Estos descriptores están relacionados con el poder nucleofílico de una molécula y concuerdan con el medioambiente del centro activo de la proteína (VEGFR2), el cual presenta aminoácidos de densidad de carga negativa (GLU885 y ASP1046) lo cual pueden estabilizar el complejo enzimainhibidor.
- El centro activo del VEGFR2 tiene un marcado carácter nucleofílico, por lo que moléculas cuyas estructuras químicas tengan grupos con marcada densidad de carga negativa pudieran ser buenos agentes anticancerígenos capaces de inhibir el Receptor-2 del Factor de Crecimiento Vascular Endoteliar.

#### 5.1. Proyección Futura

Como trabajo futuro, se propone estudiar la incidencia de más descriptores en la predicción del comportamiento de  $IC_{50}$ . Otro reto interesante es estudiar la incidencia de la calidad de los datos en los resultados obtenidos aplicando otras técnicas de clasificación. Se sabe que existen criterios de calidad de los datos que influyen negativamente en los resultados al aplicar técnicas de clasificación, como el desequilibrio y la correlación entre atributos. A pesar de no tratarlo en profundidad en este estudio, los resultados obtenidos son alentadores para el diseño de nuevos agentes anticancerígenos.

#### 6. Bibliografía

#### Referencias

- Naoki Miyamoto, Yuya Oguro, Terufumi Takagi, Hidehisa Iwata, Hiroshi Miki, Akira Hori, and Shinichi Imamura. Design, synthesis, and evaluation of imidazo[1,2-b]pyridazine derivatives having a benzamide unit as novel vegfr2 kinase inhibitors. *Bioorganic and Medicinal Chemistry*, 20(24):7051 – 7058, 2012.
- [2] Yumeng Yan, Huanyu Tao, Jiahua He, and Sheng You Huang. The HDOCK server for integrated protein–protein docking. *Nature Protocols*, 15(5):1829–1852, may 2020.
- [3] Jacopo Pizzicannella, Agnese Gugliandolo, Tiziana Orsini, Antonella Fontana, Alessia Ventrella, Emanuela Mazzon, Placido Bramanti, Francesca Diomede, and Oriana Trubiani. Engineered extracellular vesicles from human periodontal-ligament stem cells increase vegf/vegfr2 expression during bone regeneration. *Frontiers in Physiology*, 10:512, 2019.
- [4] Naoki Miyamoto, Nozomu Sakai, Takaharu Hirayama, Kazuhiro Miwa, Yuya Oguro, Hideyuki Oki, Kengo Okada, Terufumi Takagi, Hidehisa Iwata, Yoshiko Awazu, Seiji Yamasaki, Toshiyuki Takeuchi, Hiroshi Miki, Akira Hori, and Shinichi Imamura. Discovery of n-[5-({2-[(cyclopropylcarbonyl)amino]imidazo[1,2-b]pyridazin-6-yl}oxy)-2-methylphenyl]-1,3-dimethyl-1h-pyrazole-5-carboxamide (tak-593), a highly potent vegfr2 kinase inhibitor. *Bioorganic and Medicinal Chemistry*, 21(8):2333–2345, apr 2013.
- [5] Yulin Zou. Benzofuran-isatin conjugates as potent vegfr-2 and cancer cell growth inhibitors. Journal of Heterocyclic Chemistry, 57(1):510-516, 2020.
- [6] Eman Y. Ahmed, Nehad A. Abdel Latif, Mohamed F. El-Mansy, Weam S. Elserwy, and Omaima M. Abdelhafez. Vegfr-2 inhibiting effect and molecular modeling of newly synthesized coumarin derivatives as anti-breast cancer agents. *Bioorganic and Medicinal Chemistry*, 28(5):115328, 2020.
- [7] Markus Krebs, Antonio Giovanni Solimando, Charis Kalogirou, André Marquardt, Torsten Frank, Ioannis Sokolakis, Georgios Hatzichristodoulou, Susanne Kneitz, Ralf Bargou, Hubert Kübler, Bastian Schilling, Martin Spahn, and Burkhard Kneitz. mir-221-3p regulates vegfr2 expression in high-risk prostate cancer and represents an escape mechanism from sunitinib in vitro. *Journal of Clinical Medicine*, 9(3), 2020.

- [8] Sanjay M. Jachak and Arvind Saklani. Challenges and opportunities in drug discovery from plants. *Current Science*, 92(9):1251–1257, 2007.
- [9] Pramodkumar P. Gupta, Virupaksha A. Bastikar, Alpana Bastikar, Santosh S. Chhajed, and Parag A. Pathade. Computational Screening Techniques for Lead Design and Development. In *Computer-Aided Drug Design*, pages 187–222. Springer Singapore, 2020.
- [10] Pramod Katara. Computational Approaches for Drug Target Identification. In Computer-Aided Drug Design, pages 163–185. Springer Singapore, 2020.
- [11] Rajesh Kumar Pathak, Dev Bukhsh Singh, Mamta Sagar, Mamta Baunthiyal, and Anil Kumar. Computational Approaches in Drug Discovery and Design. In *Computer-Aided Drug Design*, pages 1–21. Springer Singapore, 2020.
- [12] Salman Akhtar, M. Kalim A. Khan, and Khwaja Osama. Machine Learning Approaches to Rational Drug Design. In *Computer-Aided Drug Design*, pages 279–306. Springer Singapore, 2020.
- [13] Luisa Quesada-Romero, Karel Mena-Ulecia, Matias Zuñiga, Pedro De-la Torre, Daniela Rossi, William Tiznado, Simona Collina, and Julio Caballero. Optimal graph-based and Simplified Molecular Input Line Entry System-based descriptors for quantitative structure–activity relationship analysis of arylalkylaminoalcohols, arylalkenylamines, and arylalkylamines as σ – 1 receptor ligands. Journal of Chemometrics, (January):13–20, 2014.
- [14] Nishi K. Rao, Arpita Yadav, and Sanjeev Kumar Singh. An ab initio quantum mechanical drug designing procedure: Application to the design of balanced dual ACE/NEP inhibitors. *Journal of Molecular Modeling*, 15(12):1447–1462, 2009.
- [15] P. Buchwald and Nicholas Bodor. Recent advances in the design and development of soft drugs. *Pharmazie*, 69(6):403–413, 2014.
- [16] Markus A Lill and Matthew L Danielson. Computer-aided drug design platform using PyMOL. Journal of Computer-Aided Molecular Design, 25(1):13–19, 2011.
- [17] Aurélien Grosdidier, Vincent Zoete, and Olivier Michielin. SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Research, 39(SUPPL. 2):W270—-W277, jul 2011.

- [18] L. Medina-Franco, José, Eli Fernán-Dezde Gortari, and J. Jesús Naveja. Avances en el diseño de fármacos asistido por computadora. *Educacion Quimica*, 26(3):180–186, 2015.
- [19] Shikha Agnihotry, Rajesh Kumar Pathak, Ajeet Srivastav, Pradeep Kumar Shukla, and Budhayash Gautam. Molecular Docking and Structure-Based Drug Design. In *Computer-Aided Drug Design*, pages 115–131. Springer Singapore, 2020.
- [20] Gisbert Schneider and Uli Fechner. Computer-based de novo design of drug-like molecules. Nature Reviews Drug Discovery, 4(8):649–663, 2005.
- [21] Douglas B. Kitchen, Hélène Decornez, John R. Furr, and Jürgen Bajorath. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery*, 3(11):935–949, 2004.
- [22] Natasja Brooijmans and Irwin D. Kuntz. Molecular recognition and docking algorithms. Annual Review of Biophysics and Biomolecular Structure, 32:335–373, 2003.
- [23] Sander Boonstra, Patrick R. Onck, and Erik van der Giessen. CHARMM TIP3P Water Model Suppresses Peptide Folding by Solvating the Unfolded State. *The Journal of Physical Chemistry* B, 120(15):3692–3698, apr 2016.
- [24] Karel Mena-Ulecia and Desmond MacLeod-Carey. Interactions of 2-phenyl-benzotriazole xenobiotic compounds with human Cytochrome P450-CYP1A1 by means of docking, molecular dynamics simulations and MM-GBSA calculations. *Computational Biology and Chemistry*, 74:253–262, 2018.
- [25] Culo D E Revisi. Diseño de fármacos asistido por computadora: cuando la informática, la química y el arte se encuentran. TIP Revista Especializada en Ciencias Químico-Biológicas, 21(2):124–134, 2018.
- [26] Z. Hao, A. AghaKouchak, N. Nakhjiri, and A Farahmand. Global integrated drought monitoring and prediction system (GIDMaPS) data sets. *figshare* http://dx.doi.org/10.6084/m9. figshare.853801, 2014.
- [27] Karel Mena-Ulecia and Heykel Hernandez. Decentralized peri-urban wastewater treatment technologies assessment integrating sustainability indicators. Water Science and Technology, 72(2):214– 222, 2015.

- [28] Luisa Quesada-Romero, Karel Mena-Ulecia, William Tiznado, and Julio Caballero. Insights into the interactions between maleimide derivates and GSK3β combining molecular docking and QSAR. PLoS ONE, 9(7), 2014.
- [29] Eugene N Muratov, Jürgen Bajorath, Robert P Sheridan, Igor V Tetko, Dmitry Filimonov, Vladimir Poroikov, Tudor I Oprea, Igor I Baskin, Alexandre Varnek, Adrian Roitberg, et al. Qsar without borders. *Chemical Society Reviews*, 49(11):3525–3564, 2020.
- [30] Paola Gramatica. Principles of qsar models validation: internal and external. QSAR & combinatorial science, 26(5):694–701, 2007.
- [31] Paola Gramatica. Principles of qsar modeling: comments and suggestions from personal experience. ce. International Journal of Quantitative Structure-Property Relationships (IJQSPR), 5(3):61–97, 2020.
- [32] Sanya Chabra. Quantitative structure activity relationships (qsar). International Journal of Pharmacy & Life Sciences, 11(7), 2020.
- [33] Bhuvnesh Rai, Medha Srivastava, and Dharmendra Kumar Chaudhary. Methods of applying qsar to predict in vivo and in vitro activity relationship paradigm. International Journal of Applied Biology and Pharmaceutical Technology, 11(3):160–169, 2020.
- [34] Rania Kasmi, Elghalia Hadaji, Oussama Chedadi, Abdellah El Aissouq, Mohammed Bouachrine, and Abdelkrim Ouammou. 2d-qsar and docking study of a series of coumarin derivatives as inhibitors of cdk (anticancer activity) with an application of the molecular docking method. *Heliyon*, 6(8):e04514, 2020.
- [35] Yu-Liang Wang, Fan Wang, Xing-Xing Shi, Chen-Yang Jia, Feng-Xu Wu, Ge-Fei Hao, and Guang-Fu Yang. Cloud 3d-qsar: a web tool for the development of quantitative structure–activity relationship models in drug discovery. *Briefings in Bioinformatics*, 22(4):bbaa276, 2021.
- [36] Ji Yun Kim, Kyu-Bong Kim, and Byung-Mu Lee. Validation of quantitative structure-activity relationship (qsar) and quantitative structure-property relationship (qspr) approaches as alternatives to skin sensitization risk assessment. Journal of Toxicology and Environmental Health, Part A, 84(23):945–959, 2021.
- [37] Yudith Cañizares-Carmenate, Karel Mena-Ulecia, Yunier Perera-Sardiña, Francisco Torrens, and Juan A. Castillo-Garit. An approach to identify new antihypertensive agents using Thermolysin as

model: In silico study based on QSARINS and docking. Arabian Journal of Chemistry, 12(8):4861–4877, 2019.

- [38] Yu-Chen Lo, Stefano E Rensi, Wen Torng, and Russ B Altman. Machine learning in chemoinformatics and drug discovery. Drug discovery today, 23(8):1538–1546, 2018.
- [39] WB Langdon, SJ Barrett, and BF Buxton. Predicting biochemical interactions-human p450 2d6 enzyme inhibition. In *The 2003 Congress on Evolutionary Computation*, 2003. CEC'03., volume 2, pages 807–814. IEEE, 2003.
- [40] Yutang Wang, Daniel Russo, Chang Liu, Qian Zhou, Hao Zhu, and Yinghua Zhang. Predictive modeling of angiotensin i-converting enzyme (ace) inhibitory peptides using various machine learning approaches. *Journal of Agricultural and Food Chemistry*, 2020.
- [41] Xingye Chen, Wuchen Xie, Yan Yang, Yi Hua, Guomeng Xing, Li Liang, Chenglong Deng, Yuchen Wang, Yuanrong Fan, Haichun Liu, et al. Discovery of dual fgfr4 and egfr inhibitors by machine learning and biological evaluation. Journal of Chemical Information and Modeling, 2020.
- [42] Karel Mena-ulecia, Ariela Vergara-Jaque, Horacio Poblete, William Tiznado, and Julio Caballero. Study of the Affinity between the Protein Kinase PKA and Peptide Substrates Derived from Kemptide Using Molecular Dynamics Simulations and MM / GBSA. *PLoS ONE*, 9(10):e109639, 2014.
- [43] Abel Kolawole Oyebamiji, Oluwatumininu Abosede Mutiu, Folake Ayobami Amao, Olubukola Monisola Oyawoye, Temitope A. Oyedepo, Babatunde Benjamin Adeleke, and Banjo Semire. Dataset on in-silico investigation on triazole derivatives via molecular modelling approach: A potential glioblastoma inhibitors. *Data in Brief*, 34, feb 2021.
- [44] Rifaat Hilal, Ahmed A Abdel Khalek, and Shabaan A K Elroby. Dft investigation of nitrenium ions derived from metabolism of antitumor 2-(4-aminophenyl)benzothiazoles. *Journal of Molecular Structure: THEOCHEM*, 731(1-3):115–121, 2005.
- [45] D R Roy, R Parthasarathi, B Maiti, V Subramanian, and P K Chattaraj. Electrophilicity as a possible descriptor for toxicity prediction. *Bioorganic and Medicinal Chemistry*, 13(10):3405–3412, 2005.
- [46] Masaaki Hirose, Masanori Okaniwa, Tohru Miyazaki, Takashi Imada, Tomohiro Ohashi, Yuta Tanaka, Takeo Arita, Masato Yabuki, Tomohiro Kawamoto, Shunichirou Tsutsumi, Akihiko Sumita,

Terufumi Takagi, Bi-Ching Sang, Jason Yano, Kathleen Aertgeerts, Sei Yoshida, and Tomoyasu Ishikawa. Design and synthesis of novel dfg-out raf/vascular endothelial growth factor receptor 2 (vegfr2) inhibitors: 3. evaluation of 5-amino-linked thiazolo[5,4-d]pyrimidine and thiazolo[5,4b]pyridine derivatives. *Bioorganic and Medicinal Chemistry*, 20(18):5600–5615, sep 2012.

- [47] Yuya Oguro, Naoki Miyamoto, Kengo Okada, Terufumi Takagi, Hidehisa Iwata, Yoshiko Awazu, Hiroshi Miki, Akira Hori, Keiji Kamiyama, and Shinichi Imamura. Design, synthesis, and evaluation of 5-methyl-4-phenoxy-5h-pyrrolo[3,2-d]pyrimidine derivatives: novel vegfr2 kinase inhibitors binding to inactive kinase conformation. *Bioorganic and Medicinal Chemistry*, 18(20):7260–73, oct 2010.
- [48] Artem Cherkasov, Eugene N. Muratov, Denis Fourches, Alexandre Varnek, Igor I. Baskin, Mark Cronin, John Dearden, Paola Gramatica, Yvonne C. Martin, Roberto Todeschini, Viviana Consonni, Victor E. Kuz'min, Richard Cramer, Romualdo Benigni, Chihae Yang, James Rathman, Lothar Terfloth, Johann Gasteiger, Ann Richard, and Alexander Tropsha. Qsar modeling: Where have you been? where are you going to? *Journal of Medicinal Chemistry*, 57(12):4977–5010, 2014. PMID: 24351051.
- [49] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. Journal of Cheminformatics, 4(1):17, 2012.
- [50] Yoshio Tsutsumi. Schrodinger equation. Funkcialaj Ekvacioj, 30:115–125, 1987.
- [51] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, oct 1996.
- [52] Frank Neese. The orca program system. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2(1):73–78, jan 2012.
- [53] Frank Neese. Software update: the orca program system, version 4.0. Wiley Interdisciplinary Reviews: Computational Molecular Science, 8(1):e1327, jan 2018.
- [54] Robert G. Parr, Weitao Yang, and Weitao. Yang. Density-functional theory of atoms and molecules. The international series of monographs on chemistry; 16. Oxford University Press, New York, 1989.

- [55] O. V. Gritsenko. Koopmanstheorem and its density-functional-theory analog assessed in evaluation of the red shift of vertical ionization potential upon complexation. *Chemical Physics Letters*, 691:178–180, jan 2018.
- [56] T. Koopmans. Uber die zuordnung von wellenfunktionen und eigenwerten zu den einzelnen elektronen eines atoms. *Physica*, 1(1-6):104–113, jan 1934.
- [57] Ralph G. Pearson. Absolute electronegativity and hardness correlated with molecular orbital theory. Proceedings of the National Academy of Sciences, 83(22):8440–8441, 1986.
- [58] Robert G Parr, Robert A Donnelly, Mel Levy, and William E Palke. Electronegativity: The density functional viewpoint. *The Journal of Chemical Physics*, 68(8):3801–3807, 1978.
- [59] Robert G Parr and Weitao Yang. Density functional approach to the frontier-electron theory of chemical reactivity. Journal of the American Chemical Society, 106(14):4049–4050, 1984.
- [60] John Faver, K M Merz Jr, and Kenneth M Merz. Utility of the hard soft acid-base principle via the fukui function in biological systems. *Journal of chemical theory and computation*, 6(2):548–559, 2010.
- [61] Pratim Kumar Chattaraj, Utpal Sarkar, and Debesh Ranjan Roy. Electrophilicity index. Chemical reviews, 106(6):2065–2091, jun 2006.
- [62] Cedillo A Gazquez J.L and A Vela. Electrodonating and electroaccepting powers. The Journal of Physical Chemistry, 111(10):1966–1970, 2007.
- [63] Pratim K Chattaraj, Soma Duley, and Luis R Domingo. Understanding local electrophilicity/nucleophilicity activation through a single reactivity difference index. Organic and Biomolecular Chemistry, 10(14):2855–61, apr 2012.
- [64] Chun Wei Yap. Padel-descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of computational chemistry, 32(7):1466–1474, 2011.
- [65] R Core Team. R: A Language and Environment for Statistical Computing, 2019.
- [66] Carl Kingsford and Steven L Salzberg. What are decision trees? Nature biotechnology, 26(9):1011– 1013, 2008.
- [67] Rafael Rivera-Lopez, Juana Canul-Reich, Efrén Mezura-Montes, and Marco Antonio Cruz-Chávez. Induction of decision trees as classification models through metaheuristics. Swarm and Evolutionary Computation, 69:101006, 2022.

- [68] Mikhail Moshkov. On the depth of decision trees with hypotheses. *Entropy*, 24(1):116, 2022.
- [69] Raoul Heese, Patricia Bickert, and Astrid Elisa Niederle. Representation of binary classification trees with binary features by quantum circuits. *Quantum*, 6:676, 2022.
- [70] David I Poole, Randy G Goebel, and Alan K Mackworth. Computational intelligence, volume 1. Oxford University Press New York, 1998.
- [71] Ayman M Mansour. Decision tree-based expert system for adverse drug reaction detection using fuzzy logic and genetic algorithm. International Journal of Advanced Computer Research, 8(36):110–128, 2018.
- [72] Earl Chrysler. Using decision tree analysis to develop an expert system. Director, page 07, 2006.
- [73] Pooja Bhalode, Sonia M Razavi, Andres Roman-Ospino, James Scicolone, Gerardo Callegari, Atul Dubey, Abdollah Koolivand, Scott Krull, Thomas O'Connor, Fernando J Muzzio, et al. Statistical data pre-treatment for residence time distribution studies in pharmaceutical manufacturing. Available at SSRN 4249747, 2022.
- [74] Xingji Yu, Kristian Stenerud Skeie, Michael Dahl Knudsen, Zhengru Ren, Lars Imsland, and Laurent Georges. Influence of data pre-processing and sensor dynamics on grey-box models for space-heating: Analysis using field measurements. *Building and Environment*, 212:108832, 2022.
- [75] Eibe Frank, Mark Hall, Geoffrey Holmes, Richard Kirkby, Bernhard Pfahringer, Ian H Witten, and Len Trigg. Weka-a machine learning workbench for data mining. In *Data mining and knowledge* discovery handbook, pages 1269–1277. Springer, 2009.
- [76] Geoffrey Holmes, Andrew Donkin, and Ian H Witten. Weka: A machine learning workbench. In Proceedings of ANZIIS'94-Australian New Zealnd Intelligent Information Systems Conference, pages 357–361. IEEE, 1994.
- [77] Fatma A Ibrahim and Omar A Shiba. Data mining: Weka software (an overview). Journal of Pure & Applied Sciences, 18(3), 2019.
- [78] Stephen R Garner et al. Weka: The waikato environment for knowledge analysis. In Proceedings of the New Zealand computer science research students conference, volume 1995, pages 57–64, 1995.
- [79] Nasiba Mahdi Abdulkareem, Adnan Mohsin Abdulazeez, et al. Machine learning classification based on radom forest algorithm: A review. *International Journal of Science and Business*, 5(2):128– 142, 2021.

- [80] T Koopmans. Uber die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica*, 1(1-6):104–113, 1934.
- [81] Sotiris B Kotsiantis. Decision trees: a recent overview. Artificial Intelligence Review, 39(4):261– 283, 2013.
- [82] O V Gritsenko. Koopmans' theorem and its density-functional-theory analog assessed in evaluation of the red shift of vertical ionization potential upon complexation. *Chemical Physics Letters*, 691:178–180, 2018.
- [83] Fanny Girard, Vasilica Badets, Sylvie Blanc, Kristaq Gazeli, Laurent Marlin, Laurent Authier, Panagiotis Svarnas, Neso Sojic, Franck Clément, and Stéphane Arbault. Formation of reactive nitrogen species including peroxynitrite in physiological buffer exposed to cold atmospheric plasma. RSC Advances, 6(82):78457–78467, aug 2016.
- [84] Camila Muñoz, Francisco Adasme, Jans H Alzate-Morales, Ariela Vergara-Jaque, Torsten Kniess, Julio Caballero, and Caballero J (Muñoz C, Adasme F, Alzate-Morales JH, Vergara-Jaque A, Kniess T. Study of differences in the VEGFR2 inhibitory activities between semaxanib and SU5205 using 3D-QSAR, docking, and molecular dynamics simulations. Journal of molecular graphics & modelling, 32:39–48, feb 2012.
- [85] Hannah Yang, Won Suk Lee, So Jung Kong, Chang Gon Kim, Joo Hoon Kim, Sei Kyung Chang, Sewha Kim, Gwangil Kim, Hong Jae Chon, and Chan Kim. Sting activation reprograms tumor vasculatures and synergizes with vegfr2 blockade. The Journal of Clinical Investigation, 129(10):4350–4364, 10 2019.
- [86] Melissa F. Adasme, Daniele Parisi, Kristien Van Belle, Sebastian Salentin, V. Joachim Haupt, Gary S. Jennings, Jörg-Christian Heinrich, Jean Herman, Ben Sprangers, Thierry Louat, Yves Moreau, and Michael Schroeder. Structure-based drug repositioning explains ibrutinib as vegfr2 inhibitor. PLOS ONE, 15(5):1–26, 05 2020.
- [87] Y Simon-Manso and P Fuentealba. On the density functional relationship between static dipole polarizability and global softness. *The Journal of Physical Chemistry A*, 102(11):2029–2032, 1998.
- [88] Cedillo A Gazquez J.L, A Vela, JL L Gazquez, A Cedillo, A Vela, Cedillo A Gazquez J.L, and A Vela. Electrodonating and electroaccepting powers. *The Journal of Physical Chemistry*, 111(10):1966–1970, 2007.

- [89] Doh Soro, Lynda Ekou, Bafétigué Ouattara, Mamadou Guy-Richard Kone, Tchirioua Ekou, Nahossé Ziao, and D Soro. DFT Study, Linear and Nonlinear Multiple Regression in the Prediction of HDAC7 Inhibitory Activities on a Series of Hydroxamic Acids Open Access. Computational Molecular Bioscience, 9:63–80, 2019.
- [90] Frank Neese. Software update: the ORCA program system, version 4.0. Wiley Interdisciplinary Reviews: Computational Molecular Science, 8(1):e1327, 2018.
- [91] Masanori Okaniwa, Masaaki Hirose, Takashi Imada, Tomohiro Ohashi, Youko Hayashi, Tohru Miyazaki, Takeo Arita, Masato Yabuki, Kazuyo Kakoi, Juran Kato, Terufumi Takagi, Tomohiro Kawamoto, Shuhei Yao, Akihiko Sumita, Shunichirou Tsutsumi, Tsuneaki Tottori, Hideyuki Oki, Bi-Ching Sang, Jason Yano, Kathleen Aertgeerts, Sei Yoshida, and Tomoyasu Ishikawa. Design and synthesis of novel DFG-out RAF/vascular endothelial growth factor receptor 2 (VEGFR2) inhibitors. 1. Exploration of [5,6]-fused bicyclic scaffolds. *Journal of medicinal chemistry*, 55(7):3452–78, apr 2012.

## 7. Anexos

**Anexo 1**: Compuestos utilizados para la predicción de  $IC_{50}$  incluidos en este estudio y valores de  $log(1000/IC_{50})$  experimentales respectivos.

Anexo 2: Descriptores Moleculares calculados mediante el programa PaDel Descriptor

**Anexo 3:** Valores de los Descriptores Moleculares (ev) derivados de la Teoría del Funcional de la Densidad para cada inhibidor del Factor de Crecimiento Endoteliar.

## Anexo-1: Compuestos utilizados para la predicción de IC<sub>50</sub> incluidos en este estudio y valores log(1000/IC<sub>50</sub>) experimentales respectivos.



Compuesto	Ra	Rb	Rc	Rd	Re	Log(1000/IC₅₀)
hirose_2a	Me	Me	Н	H	(CH <sub>2</sub> ) <sub>2</sub> CN	0.3872
hirose_2b	Me	Me	Н	H	CH <sub>2</sub> C(Me) <sub>2</sub> CN	0.5229
hirose_2c	Me	Me	Н	Н	OC(Me)₂CN	0.2924
hirose_2d	Me	Ме	Н	CI	cPrCN	0.4685
hirose_3	Ме	н	F	н	OC(Me)₂CN	2.0000
hirose_4	Ме	н	F	н	OC(Me)₂CN	1.8539
hirose_5	cPr	н	F	н	OC(Me)₂CN	2.1249
hirose_6	cPr	Н	F	CI	OC(Me) <sub>2</sub> CN	1.1938



2, 5a-d, 6a-b, 8a-b





C

Ĭ



Log(1000/IC<sub>50</sub>) Compuesto Ra Rb Rc Х N N Ph , H miyamoto\_2 Н Н -0.7160 -

Me

Ν

miyamoto_4	<b>0</b>	Ph-5CF₃	-	-	2.1612
miyamoto_5a	Н	N N O Ph	Н	-	1.2840
miyamoto_5b	Н	Ph-3CF <sub>3</sub>	Н	-	1.9208
miyamoto_5c	Н	Ph-4CF <sub>3</sub>	Н	-	1.5376
miyamoto_5d	Н	Н	H H N N O Ph	-	-1.0000
miyamoto_6a	н	N Ph	н	-	1.0969
miyamoto_6b	Н	H N Ph-5CF <sub>3</sub>	Н	-	2.1487
miyamoto_8a	Н	O Ph N	Н	-	0.1938
miyamoto_8b	Н	O Ph−5CF <sub>3</sub>	Н	-	1.7447
miyamoto_10a	• <u> </u>	Ph-5F	-	-	1.5686
miyamoto_10b		Ph-5F	-	-	1.9586

miyamoto_10c	S Me	Ph-5F	-	_	1.7212
miyamoto_10d	F <sub>3</sub> C	Ph-5F	-	-	0.7959
miyamoto_10e	F <sub>3</sub> C	Ph-5F	-	-	0.9208
miyamoto_10f		Ph-5F	-	-	3.0000
miyamoto_10g	Ме	Ph-5F	-	-	2.9586
miyamoto_11		$\rightarrow$	-	-	1.9208
miyamoto_12		C(Me)₃	-	-	0.1871
miyamoto_13		N	-	-	2.8861
miyamoto_14			-	-	2.8861
miyamoto_15		O-N Me	-	-	2.8539
miyamoto_16		Me <sup>-N-N</sup>	-	-	2.8861
miyamoto_17		Me N-Me	_	-	3.0177
miyamoto_18		Me N Me	-	-	1.6021

miyamoto_19		Me <sup>-N-N</sup>	-	-	2.8539
miyamoto_20		Ph <sup>-N-N</sup> Me	-	-	2.9208
miyamoto_21		Me <sup>-N-N</sup> CF <sub>3</sub>	-	-	2.4318
miyamoto_23a	-	-	-	2-Me	3.0223
miyamoto_23b	-	-	-	4-Me	2.7447
miyamoto_23c	-	-	-	6-Me	2.3279
miyamoto_23d	-	-	-	2-Et	2.2924
miyamoto_23e	-	-	-	2-F	2.9586
miyamoto_23f	_	-	-	2-Cl	2.8861



11i-p, 14, ib, 5a-c, 10a-c



18a-c, 20a-d



4-5, 13a-b, 8-11, 16

Compuesto	Ra	Rb	Rc	Rd	Re	Rf	Log(1000/IC <sub>50</sub> )
oguro ib	Me	н	CI	н	н	CE	2 4318
	IVIC						2.4310
oguro_5a	(CH <sub>2</sub> ) <sub>2</sub> OH	н	CI	н	Н	CF₃	2.3979
oguro 5b	(CH <sub>2</sub> ) <sub>2</sub> OCH <sub>3</sub>	Н	CI	н	Н	CF₃	2.0458

oguro_5c	(CH <sub>2</sub> ) <sub>2</sub> O(CH <sub>2</sub> ) <sub>2</sub> OH	н	CI	н	Н	CF₃	2.5686
oguro_10a	Ме	CH₂OH	CI	н	Н	CF₃	2.0132
oguro_10b	Ме	CH₂OC H₃	CI	н	н	CF₃	1.8539
oguro_10c	Ме	C(CH <sub>3</sub> ) <sub>2</sub> OH	CI	н	Н	CF₃	1.7447
oguro_11a	N	-	-	-	-	-	0.6778
oguro_11b	N	-	-	-	-	-	0.6383
oguro_11c		-	-	-	-	-	0.1871
oguro_11d	N CF <sub>3</sub>	-	-	-	-	-	2.5686
oguro_11e	N N CF <sub>3</sub>	-	-	-	-	-	2.3979
oguro_11f	N N C(Me) <sub>3</sub>	-	-	-	-	-	2.1192
oguro_11g	O C(Me) <sub>3</sub>	-	-	-	-	-	2.0506
oguro_11h	C(Me) <sub>3</sub>	-	-	-	-	-	1.7212
oguro_11i	н	Ме	CI	н	ОМе	-	2.0269

oguro_11j	н	Ме	CI	н		CF₃	1.2147
oguro_11k	Н	Ме	СІ	н		CF₃	1.7212
oguro_11I	н	Ме	CI	н		CF₃	2.0315
oguro_11m	Н	Ме	CI	н		CF₃	1.9208
oguro_11n	н	Ме	CI	O Me	Н	CF₃	1.6990
oguro_11o	н	Ме	CI	$\langle \rangle$	Н	CF₃	1.6383
oguro_11p	Н	Ме	CI		Н	CF₃	1.9586
oguro_12a	CF <sub>3</sub>	-	_	-	-	-	2.5376
oguro_12b	N CF <sub>3</sub>	-	-	-	-	-	2.4202
oguro_14	н	Ме	CI	N.	Н	CF₃	1.2218
oguro2_4	н	NHCON HPh	-	-	-	-	-0.2553
oguro2_5	Н	NHCOP h	-	_	-	-	0.0269
oguro2_7a	Ph	Н	-	_	-	-	1.4815

	1	1	1		1		
oguro2_7b	Ме	-	-	-	-	-	-0.7993
oguro2_7c	o2_7c Pr		-	-	-	-	-0.5911
oguro2_7d	2-CIPh	-	-	-	-	-	0.4685
oguro2_7e	3-CIPh	_	-	-	-	-	2.3872
oguro2_7f	4-CIPh	-	-	-	-	-	1.4949
oguro2_7g	3-CF₃Ph	-	-	-	-	-	2.2757
oguro2_7h	3-BrPh	_	-	-	-	-	2.3565
oguro2_7i	3-FPh	-	-	-	-	-	1.7212
oguro2_7j	3-MePh	_	-	-	-	-	2.5686
oguro2_8	NHCOPh	н	-	-	-	-	-0.9731
oguro2_9	NHCOCH₂Ph	н	-	-	-	-	-1.0000
oguro2_10	NHCSNHPh		-	-	-	-	-0.8261
oguro2_11		н	-	-	-	-	-0.6232
oguro2_13a	Н	CONHP h	-	-	-	-	0.0605
oguro2_13b	CONHPh	н	-	-	-	-	-1.0000
oguro2_16	CH₂CONHPh	н	-	-	-	-	-0.5185
oguro2_18a	Н	-	s	-	-	-	0.9586
oguro2_18b	Н	-	NH	_	-	-	-0.1461
oguro2_18c	CF₃	-	N(Me)	-	-	-	-1.0000
oguro2_20a	CF₃	2-CI	0	_	-	-	2.4318
oguro2_20b	CF₃	3-CI	0	-	_	-	1.5229
oguro2_20c	CF <sub>3</sub>	2-OMe	0	-	-	-	1.8539
oguro2_20d	CF₃	2-F	0	_	-	-	2.2076



1h-n

1a-g, 2, 3a-b, 4a-b, 5, 6a-e

Compuesto	Ra	Rb	Rc	Rd	x	Y	z	Log(1000/ IC₅₀)
okinawa_1a	0				С	N	N	2.5086
okinawa_1b	H₃C-	н	н	CF₃	с	N	N	2.9586
okinawa_1c	CH3	н	н	CF₃	С	N	N	2.7696
okinawa_1d	H <sub>3</sub> C H <sub>3</sub> C	Н	Н	CF₃	С	N	N	2.3768
okinawa_1e		н	н	CF₃	С	N	N	2.7212
okinawa_1f	$\diamond$	н	н	CF₃	С	N	N	2.3188
okinawa_1g	$\frown$	Н	Н	CF₃	с	N	N	2.0315
okinawa_1h	CF3	-	-	-	-	-	-	2.7212
okinawa_1i	CF <sub>3</sub>	-	-	-	-	-	-	2.8539
okinawa_1j	OCH3	-	-	-	-	-	-	3.5086

okinawa_1k	OC(CH <sub>3</sub> ) <sub>3</sub>	-	-	-	-	-	-	2.7447
okinawa_1m	C(CH <sub>3</sub> ) <sub>2</sub> OH	-	_	-	-	-	-	3.0862
okinawa_1n	C(CH <sub>3</sub> ) <sub>2</sub> CN	-	-	-	-	-	-	1.8539
okinawa_2		Н	Н	C(CH <sub>3</sub> ) <sub>2</sub> C N	С	N	С	2.7447
okinawa_3a	$\triangleright$	Н	н	C(CH <sub>3</sub> ) <sub>2</sub> C N	NH	С	С	2.0915
okinawa_3b		Н	н	C(CH <sub>3</sub> ) <sub>2</sub> C N	NCH	С	С	2.2596
okinawa_4a		н	Н	C(CH <sub>3</sub> ) <sub>2</sub> C N	NH	С	N	2.5086
okinawa_4b		Н	Н	C(CH <sub>3</sub> ) <sub>2</sub> C N	NCH 3	С	N	2.6198
okinawa_5		Н	н	C(CH <sub>3</sub> ) <sub>2</sub> C N	S	С	С	1.8539
okinawa_6a		Н	Н	C(CH <sub>3</sub> ) <sub>2</sub> C N	S	С	N	2.5528
okinawa_6b		Н	Н		S	С	N	1.5086
okinawa_6c		н	CI		S	С	N	2.3979
okinawa_6d		F	CI	CN CN	S	С	N	2.6576
okinawa_6e		CI	CI	[≻⊂ <sub>CN</sub>	S	С	N	2.1249

## Anexo-2: Descriptores Moleculares calculados mediante el programa "PaDel Descriptor": Estos fueron obtenidos del Programa PaDel Descriptor.

Descriptores Moleculares	Descripción		
nAcid	Number of acidic groups. The list of acidic groups is defined by these SMARTS "\$ ([O;H1]-[C,S,P]=O)", "\$([*;-;!\$(*~[*;+])])", "\$([NH](S(=O)=O)C(F)(F)F)", and "\$ (n1nnnc1)" originally presented in JOELib		
ALogP	Ghose-Crippen LogKow		
ALogP2	Square of ALogP		
AMR	Molar refractivity		
apol	Sum of the atomic polarizabilities (including implicit hydrogens)		
naAromAtom	Number of aromatic atoms		
nAromBond	Number of aromatic bonds		
nAtom	Number of atoms		
nHeavyAtom	Number of heavy atoms (i.e. not hydrogen)		
nH	Number of hydrogen atoms		
nB	Number of boron atoms		
nC	Number of carbon atoms		
nN	Number of nitrogen atoms		
nO	Number of oxygen atoms		
nS	Number of sulphur atoms		
nP	Number of phosphorus atoms		
nF	Number of fluorine atoms		
nCl	Number of chlorine atoms		
nBr	Number of bromine atoms		
nl	Number of iodine atoms		
nX	Number of halogen atoms (F, CI, Br, I, At, Uus)		
ATS0m	Broto-Moreau autocorrelation - lag 0 / weighted by mass		
ATS1m	Broto-Moreau autocorrelation - lag 1 / weighted by mass		
ATS2m	Broto-Moreau autocorrelation - lag 2 / weighted by mass		
ATS3m	Broto-Moreau autocorrelation - lag 3 / weighted by mass		
ATS4m	Broto-Moreau autocorrelation - lag 4 / weighted by mass		
ATS5m	Broto-Moreau autocorrelation - lag 5 / weighted by mass		
ATS6m	Broto-Moreau autocorrelation - lag 6 / weighted by mass		
ATS7m	Broto-Moreau autocorrelation - lag 7 / weighted by mass		
ATS8m	Broto-Moreau autocorrelation - lag 8 / weighted by mass		
ATS0v	Broto-Moreau autocorrelation - lag 0 / weighted by van der Waals volumes		
ATS1v	Broto-Moreau autocorrelation - lag 1 / weighted by van der Waals volumes		
ATS2v	Broto-Moreau autocorrelation - lag 2 / weighted by van der Waals volumes		
ATS3v	Broto-Moreau autocorrelation - lag 3 / weighted by van der Waals volumes		

ATS4v	Broto-Moreau autocorrelation - lag 4 / weighted by van der Waals volumes
ATS5v	Broto-Moreau autocorrelation - lag 5 / weighted by van der Waals volumes
ATS6v	Broto-Moreau autocorrelation - lag 6 / weighted by van der Waals volumes
ATS7v	Broto-Moreau autocorrelation - lag 7 / weighted by van der Waals volumes
ATS8v	Broto-Moreau autocorrelation - lag 8 / weighted by van der Waals volumes
ATS0e	Broto-Moreau autocorrelation - lag 0 / weighted by Sanderson electronegativities
ATS1e	Broto-Moreau autocorrelation - lag 1 / weighted by Sanderson electronegativities
ATS2e	Broto-Moreau autocorrelation - lag 2 / weighted by Sanderson electronegativities
ATS3e	Broto-Moreau autocorrelation - lag 3 / weighted by Sanderson electronegativities
ATS4e	Broto-Moreau autocorrelation - lag 4 / weighted by Sanderson electronegativities
ATS5e	Broto-Moreau autocorrelation - lag 5 / weighted by Sanderson electronegativities
ATS6e	Broto-Moreau autocorrelation - lag 6 / weighted by Sanderson electronegativities
ATS7e	Broto-Moreau autocorrelation - lag 7 / weighted by Sanderson electronegativities
ATS8e	Broto-Moreau autocorrelation - lag 8 / weighted by Sanderson electronegativities
ATS0p	Broto-Moreau autocorrelation - lag 0 / weighted by polarizabilities
ATS1p	Broto-Moreau autocorrelation - lag 1 / weighted by polarizabilities
ATS2p	Broto-Moreau autocorrelation - lag 2 / weighted by polarizabilities
ATS3p	Broto-Moreau autocorrelation - lag 3 / weighted by polarizabilities
ATS4p	Broto-Moreau autocorrelation - lag 4 / weighted by polarizabilities
ATS5p	Broto-Moreau autocorrelation - lag 5 / weighted by polarizabilities
ATS6p	Broto-Moreau autocorrelation - lag 6 / weighted by polarizabilities
ATS7p	Broto-Moreau autocorrelation - lag 7 / weighted by polarizabilities
ATS8p	Broto-Moreau autocorrelation - lag 8 / weighted by polarizabilities
ATS0i	Broto-Moreau autocorrelation - lag 0 / weighted by first ionization potential
ATS1i	Broto-Moreau autocorrelation - lag 1 / weighted by first ionization potential
ATS2i	Broto-Moreau autocorrelation - lag 2 / weighted by first ionization potential
ATS3i	Broto-Moreau autocorrelation - lag 3 / weighted by first ionization potential
ATS4i	Broto-Moreau autocorrelation - lag 4 / weighted by first ionization potential
ATS5i	Broto-Moreau autocorrelation - lag 5 / weighted by first ionization potential
ATS6i	Broto-Moreau autocorrelation - lag 6 / weighted by first ionization potential
ATS7i	Broto-Moreau autocorrelation - lag 7 / weighted by first ionization potential
ATS8i	Broto-Moreau autocorrelation - lag 8 / weighted by first ionization potential
ATS0s	Broto-Moreau autocorrelation - lag 0 / weighted by I-state
ATS1s	Broto-Moreau autocorrelation - lag 1 / weighted by I-state
ATS2s	Broto-Moreau autocorrelation - lag 2 / weighted by I-state
ATS3s	Broto-Moreau autocorrelation - lag 3 / weighted by I-state
ATS4s	Broto-Moreau autocorrelation - lag 4 / weighted by I-state
ATS5s	Broto-Moreau autocorrelation - lag 5 / weighted by I-state

ATS6s	Broto-Moreau autocorrelation - lag 6 / weighted by I-state
ATS7s	Broto-Moreau autocorrelation - lag 7 / weighted by I-state
ATS8s	Broto-Moreau autocorrelation - lag 8 / weighted by I-state
AATS0m	Average Broto-Moreau autocorrelation - lag 0 / weighted by mass
AATS1m	Average Broto-Moreau autocorrelation - lag 1 / weighted by mass
AATS2m	Average Broto-Moreau autocorrelation - lag 2 / weighted by mass
AATS3m	Average Broto-Moreau autocorrelation - lag 3 / weighted by mass
AATS4m	Average Broto-Moreau autocorrelation - lag 4 / weighted by mass
AATS5m	Average Broto-Moreau autocorrelation - lag 5 / weighted by mass
AATS6m	Average Broto-Moreau autocorrelation - lag 6 / weighted by mass
AATS7m	Average Broto-Moreau autocorrelation - lag 7 / weighted by mass
AATS8m	Average Broto-Moreau autocorrelation - lag 8 / weighted by mass
AATS0v	Average Broto-Moreau autocorrelation - lag 0 / weighted by van der Waals volumes
AATS1v	Average Broto-Moreau autocorrelation - lag 1 / weighted by van der Waals volumes
AATS2v	Average Broto-Moreau autocorrelation - lag 2 / weighted by van der Waals volumes
AATS3v	Average Broto-Moreau autocorrelation - lag 3 / weighted by van der Waals volumes
AATS4v	Average Broto-Moreau autocorrelation - lag 4 / weighted by van der Waals volumes
AATS5v	Average Broto-Moreau autocorrelation - lag 5 / weighted by van der Waals volumes
AATS6v	Average Broto-Moreau autocorrelation - lag 6 / weighted by van der Waals volumes
AATS7v	Average Broto-Moreau autocorrelation - lag 7 / weighted by van der Waals volumes
AATS8v	Average Broto-Moreau autocorrelation - lag 8 / weighted by van der Waals volumes
AATS0e	Average Broto-Moreau autocorrelation - lag 0 / weighted by Sanderson electronegativities
AATS1e	Average Broto-Moreau autocorrelation - lag 1 / weighted by Sanderson electronegativities
AATS2e	Average Broto-Moreau autocorrelation - lag 2 / weighted by Sanderson electronegativities
AATS3e	Average Broto-Moreau autocorrelation - lag 3 / weighted by Sanderson electronegativities
AATS4e	Average Broto-Moreau autocorrelation - lag 4 / weighted by Sanderson electronegativities
AATS5e	Average Broto-Moreau autocorrelation - lag 5 / weighted by Sanderson electronegativities
AATS6e	Average Broto-Moreau autocorrelation - lag 6 / weighted by Sanderson electronegativities
AATS7e	Average Broto-Moreau autocorrelation - lag 7 / weighted by Sanderson electronegativities
AATS8e	Average Broto-Moreau autocorrelation - lag 8 / weighted by Sanderson electronegativities
AATS0p	Average Broto-Moreau autocorrelation - lag 0 / weighted by polarizabilities
AATS1p	Average Broto-Moreau autocorrelation - lag 1 / weighted by polarizabilities

AATS2p	Average Broto-Moreau autocorrelation - lag 2 / weighted by polarizabilities
AATS3p	Average Broto-Moreau autocorrelation - lag 3 / weighted by polarizabilities
AATS4p	Average Broto-Moreau autocorrelation - lag 4 / weighted by polarizabilities
AATS5p	Average Broto-Moreau autocorrelation - lag 5 / weighted by polarizabilities
AATS6p	Average Broto-Moreau autocorrelation - lag 6 / weighted by polarizabilities
AATS7p	Average Broto-Moreau autocorrelation - lag 7 / weighted by polarizabilities
AATS8p	Average Broto-Moreau autocorrelation - lag 8 / weighted by polarizabilities
AATS0i	Average Broto-Moreau autocorrelation - lag 0 / weighted by first ionization potential
AATS1i	Average Broto-Moreau autocorrelation - lag 1 / weighted by first ionization potential
AATS2i	Average Broto-Moreau autocorrelation - lag 2 / weighted by first ionization potential
AATS3i	Average Broto-Moreau autocorrelation - lag 3 / weighted by first ionization potential
AATS4i	Average Broto-Moreau autocorrelation - lag 4 / weighted by first ionization potential
AATS5i	Average Broto-Moreau autocorrelation - lag 5 / weighted by first ionization potential
AATS6i	Average Broto-Moreau autocorrelation - lag 6 / weighted by first ionization potential
AATS7i	Average Broto-Moreau autocorrelation - lag 7 / weighted by first ionization potential
AATS8i	Average Broto-Moreau autocorrelation - lag 8 / weighted by first ionization potential
AATS0s	Average Broto-Moreau autocorrelation - lag 0 / weighted by I-state
AATS1s	Average Broto-Moreau autocorrelation - lag 1 / weighted by I-state
AATS2s	Average Broto-Moreau autocorrelation - lag 2 / weighted by I-state
AATS3s	Average Broto-Moreau autocorrelation - lag 3 / weighted by I-state
AATS4s	Average Broto-Moreau autocorrelation - lag 4 / weighted by I-state
AATS5s	Average Broto-Moreau autocorrelation - lag 5 / weighted by I-state
AATS6s	Average Broto-Moreau autocorrelation - lag 6 / weighted by I-state
AATS7s	Average Broto-Moreau autocorrelation - lag 7 / weighted by I-state
AATS8s	Average Broto-Moreau autocorrelation - lag 8 / weighted by I-state
ATSC0c	Centered Broto-Moreau autocorrelation - lag 0 / weighted by charges
ATSC1c	Centered Broto-Moreau autocorrelation - lag 1 / weighted by charges
ATSC2c	Centered Broto-Moreau autocorrelation - lag 2 / weighted by charges
ATSC3c	Centered Broto-Moreau autocorrelation - lag 3 / weighted by charges
ATSC4c	Centered Broto-Moreau autocorrelation - lag 4 / weighted by charges
ATSC5c	Centered Broto-Moreau autocorrelation - lag 5 / weighted by charges
ATSC6c	Centered Broto-Moreau autocorrelation - lag 6 / weighted by charges
ATSC7c	Centered Broto-Moreau autocorrelation - lag 7 / weighted by charges
ATSC8c	Centered Broto-Moreau autocorrelation - lag 8 / weighted by charges
ATSC0m	Centered Broto-Moreau autocorrelation - lag 0 / weighted by mass
ATSC1m	Centered Broto-Moreau autocorrelation - lag 1 / weighted by mass
ATSC2m	Centered Broto-Moreau autocorrelation - lag 2 / weighted by mass
ATSC3m	Centered Broto-Moreau autocorrelation - lag 3 / weighted by mass

ATSC4m	Centered Broto-Moreau autocorrelation - lag 4 / weighted by mass
ATSC5m	Centered Broto-Moreau autocorrelation - lag 5 / weighted by mass
ATSC6m	Centered Broto-Moreau autocorrelation - lag 6 / weighted by mass
ATSC7m	Centered Broto-Moreau autocorrelation - lag 7 / weighted by mass
ATSC8m	Centered Broto-Moreau autocorrelation - lag 8 / weighted by mass
ATSC0v	Centered Broto-Moreau autocorrelation - lag 0 / weighted by van der Waals volumes
ATSC1v	Centered Broto-Moreau autocorrelation - lag 1 / weighted by van der Waals volumes
ATSC2v	Centered Broto-Moreau autocorrelation - lag 2 / weighted by van der Waals volumes
ATSC3v	Centered Broto-Moreau autocorrelation - lag 3 / weighted by van der Waals volumes
ATSC4v	Centered Broto-Moreau autocorrelation - lag 4 / weighted by van der Waals volumes
ATSC5v	Centered Broto-Moreau autocorrelation - lag 5 / weighted by van der Waals volumes
ATSC6v	Centered Broto-Moreau autocorrelation - lag 6 / weighted by van der Waals volumes
ATSC7v	Centered Broto-Moreau autocorrelation - lag 7 / weighted by van der Waals volumes
ATSC8v	Centered Broto-Moreau autocorrelation - lag 8 / weighted by van der Waals volumes
ATSC0e	Centered Broto-Moreau autocorrelation - lag 0 / weighted by Sanderson electronegativities
ATSC1e	Centered Broto-Moreau autocorrelation - lag 1 / weighted by Sanderson electronegativities
ATSC2e	Centered Broto-Moreau autocorrelation - lag 2 / weighted by Sanderson electronegativities
ATSC3e	Centered Broto-Moreau autocorrelation - lag 3 / weighted by Sanderson electronegativities
ATSC4e	Centered Broto-Moreau autocorrelation - lag 4 / weighted by Sanderson electronegativities
ATSC5e	Centered Broto-Moreau autocorrelation - lag 5 / weighted by Sanderson electronegativities
ATSC6e	Centered Broto-Moreau autocorrelation - lag 6 / weighted by Sanderson electronegativities
ATSC7e	Centered Broto-Moreau autocorrelation - lag 7 / weighted by Sanderson electronegativities
ATSC8e	Centered Broto-Moreau autocorrelation - lag 8 / weighted by Sanderson electronegativities
ATSC0p	Centered Broto-Moreau autocorrelation - lag 0 / weighted by polarizabilities
ATSC1p	Centered Broto-Moreau autocorrelation - lag 1 / weighted by polarizabilities
ATSC2p	Centered Broto-Moreau autocorrelation - lag 2 / weighted by polarizabilities

ATSC3p	Centered Broto-Moreau autocorrelation - lag 3 / weighted by polarizabilities
ATSC4p	Centered Broto-Moreau autocorrelation - lag 4 / weighted by polarizabilities
ATSC5p	Centered Broto-Moreau autocorrelation - lag 5 / weighted by polarizabilities
ATSC6p	Centered Broto-Moreau autocorrelation - lag 6 / weighted by polarizabilities
ATSC7p	Centered Broto-Moreau autocorrelation - lag 7 / weighted by polarizabilities
ATSC8p	Centered Broto-Moreau autocorrelation - lag 8 / weighted by polarizabilities
ATSC0i	Centered Broto-Moreau autocorrelation - lag 0 / weighted by first ionization potential
ATSC1i	Centered Broto-Moreau autocorrelation - lag 1 / weighted by first ionization potential
ATSC2i	Centered Broto-Moreau autocorrelation - lag 2 / weighted by first ionization potential
ATSC3i	Centered Broto-Moreau autocorrelation - lag 3 / weighted by first ionization potential
ATSC4i	Centered Broto-Moreau autocorrelation - lag 4 / weighted by first ionization potential
ATSC5i	Centered Broto-Moreau autocorrelation - lag 5 / weighted by first ionization potential
ATSC6i	Centered Broto-Moreau autocorrelation - lag 6 / weighted by first ionization potential
ATSC7i	Centered Broto-Moreau autocorrelation - lag 7 / weighted by first ionization potential
ATSC8i	Centered Broto-Moreau autocorrelation - lag 8 / weighted by first ionization potential
ATSC0s	Centered Broto-Moreau autocorrelation - lag 0 / weighted by I-state
ATSC1s	Centered Broto-Moreau autocorrelation - lag 1 / weighted by I-state
ATSC2s	Centered Broto-Moreau autocorrelation - lag 2 / weighted by I-state
ATSC3s	Centered Broto-Moreau autocorrelation - lag 3 / weighted by I-state
ATSC4s	Centered Broto-Moreau autocorrelation - lag 4 / weighted by I-state
ATSC5s	Centered Broto-Moreau autocorrelation - lag 5 / weighted by I-state
ATSC6s	Centered Broto-Moreau autocorrelation - lag 6 / weighted by I-state
ATSC7s	Centered Broto-Moreau autocorrelation - lag 7 / weighted by I-state
ATSC8s	Centered Broto-Moreau autocorrelation - lag 8 / weighted by I-state
AATSC0c	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by charges
AATSC1c	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by charges
AATSC2c	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by charges
AATSC3c	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by charges
AATSC4c	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by charges
AATSC5c	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by charges
AATSC6c	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by charges
AATSC7c	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by charges

AATSC8c	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by charges
AATSC0m	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by mass
AATSC1m	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by mass
AATSC2m	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by mass
AATSC3m	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by mass
AATSC4m	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by mass
AATSC5m	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by mass
AATSC6m	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by mass
AATSC7m	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by mass
AATSC8m	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by mass
AATSC0v	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by van der Waals volumes
AATSC1v	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by van der Waals volumes
AATSC2v	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by van der Waals volumes
AATSC3v	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by van der Waals volumes
AATSC4v	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by van der Waals volumes
AATSC5v	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by van der Waals volumes
AATSC6v	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by van der Waals volumes
AATSC7v	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by van der Waals volumes
AATSC8v	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by van der Waals volumes
AATSC0e	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by Sanderson electronegativities
AATSC1e	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by Sanderson electronegativities
AATSC2e	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by Sanderson electronegativities
AATSC3e	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by Sanderson electronegativities
AATSC4e	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by Sanderson electronegativities
AATSC5e	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by Sanderson electronegativities
AATSC6e	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by Sanderson electronegativities
AATSC7e	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by Sanderson electronegativities

AATSC8e	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by Sanderson electronegativities
AATSC0p	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by polarizabilities
AATSC1p	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by polarizabilities
AATSC2p	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by polarizabilities
AATSC3p	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by polarizabilities
AATSC4p	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by polarizabilities
AATSC5p	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by polarizabilities
AATSC6p	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by polarizabilities
AATSC7p	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by polarizabilities
AATSC8p	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by polarizabilities
AATSC0i	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by first ionization potential
AATSC1i	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by first ionization potential
AATSC2i	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by first ionization potential
AATSC3i	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by first ionization potential
AATSC4i	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by first ionization potential
AATSC5i	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by first ionization potential
AATSC6i	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by first ionization potential
AATSC7i	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by first ionization potential
AATSC8i	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by first ionization potential
AATSC0s	Average centered Broto-Moreau autocorrelation - lag 0 / weighted by I-state
AATSC1s	Average centered Broto-Moreau autocorrelation - lag 1 / weighted by I-state
AATSC2s	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by I-state
AATSC3s	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by I-state
AATSC4s	Average centered Broto-Moreau autocorrelation - lag 4 / weighted by I-state
AATSC5s	Average centered Broto-Moreau autocorrelation - lag 5 / weighted by I-state
AATSC6s	Average centered Broto-Moreau autocorrelation - lag 6 / weighted by I-state

AATSC7s	Average centered Broto-Moreau autocorrelation - lag 7 / weighted by I-state
AATSC8s	Average centered Broto-Moreau autocorrelation - lag 8 / weighted by I-state
MATS1c	Moran autocorrelation - lag 1 / weighted by charges
MATS2c	Moran autocorrelation - lag 2 / weighted by charges
MATS3c	Moran autocorrelation - lag 3 / weighted by charges
MATS4c	Moran autocorrelation - lag 4 / weighted by charges
MATS5c	Moran autocorrelation - lag 5 / weighted by charges
MATS6c	Moran autocorrelation - lag 6 / weighted by charges
MATS7c	Moran autocorrelation - lag 7 / weighted by charges
MATS8c	Moran autocorrelation - lag 8 / weighted by charges
MATS1m	Moran autocorrelation - lag 1 / weighted by mass
MATS2m	Moran autocorrelation - lag 2 / weighted by mass
MATS3m	Moran autocorrelation - lag 3 / weighted by mass
MATS4m	Moran autocorrelation - lag 4 / weighted by mass
MATS5m	Moran autocorrelation - lag 5 / weighted by mass
MATS6m	Moran autocorrelation - lag 6 / weighted by mass
MATS7m	Moran autocorrelation - lag 7 / weighted by mass
MATS8m	Moran autocorrelation - lag 8 / weighted by mass
MATS1v	Moran autocorrelation - lag 1 / weighted by van der Waals volumes
MATS2v	Moran autocorrelation - lag 2 / weighted by van der Waals volumes
MATS3v	Moran autocorrelation - lag 3 / weighted by van der Waals volumes
MATS4v	Moran autocorrelation - lag 4 / weighted by van der Waals volumes
MATS5v	Moran autocorrelation - lag 5 / weighted by van der Waals volumes
MATS6v	Moran autocorrelation - lag 6 / weighted by van der Waals volumes
MATS7v	Moran autocorrelation - lag 7 / weighted by van der Waals volumes
MATS8∨	Moran autocorrelation - lag 8 / weighted by van der Waals volumes
MATS1e	Moran autocorrelation - lag 1 / weighted by Sanderson electronegativities
MATS2e	Moran autocorrelation - lag 2 / weighted by Sanderson electronegativities
MATS3e	Moran autocorrelation - lag 3 / weighted by Sanderson electronegativities
MATS4e	Moran autocorrelation - lag 4 / weighted by Sanderson electronegativities
MATS5e	Moran autocorrelation - lag 5 / weighted by Sanderson electronegativities
MATS6e	Moran autocorrelation - lag 6 / weighted by Sanderson electronegativities
MATS7e	Moran autocorrelation - lag 7 / weighted by Sanderson electronegativities
MATS8e	Moran autocorrelation - lag 8 / weighted by Sanderson electronegativities
MATS1p	Moran autocorrelation - lag 1 / weighted by polarizabilities
MATS2p	Moran autocorrelation - lag 2 / weighted by polarizabilities
MATS3p	Moran autocorrelation - lag 3 / weighted by polarizabilities
MATS4p	Moran autocorrelation - lag 4 / weighted by polarizabilities

MATS5p	Moran autocorrelation - lag 5 / weighted by polarizabilities
MATS6p	Moran autocorrelation - lag 6 / weighted by polarizabilities
MATS7p	Moran autocorrelation - lag 7 / weighted by polarizabilities
MATS8p	Moran autocorrelation - lag 8 / weighted by polarizabilities
MATS1i	Moran autocorrelation - lag 1 / weighted by first ionization potential
MATS2i	Moran autocorrelation - lag 2 / weighted by first ionization potential
MATS3i	Moran autocorrelation - lag 3 / weighted by first ionization potential
MATS4i	Moran autocorrelation - lag 4 / weighted by first ionization potential
MATS5i	Moran autocorrelation - lag 5 / weighted by first ionization potential
MATS6i	Moran autocorrelation - lag 6 / weighted by first ionization potential
MATS7i	Moran autocorrelation - lag 7 / weighted by first ionization potential
MATS8i	Moran autocorrelation - lag 8 / weighted by first ionization potential
MATS1s	Moran autocorrelation - lag 1 / weighted by I-state
MATS2s	Moran autocorrelation - lag 2 / weighted by I-state
MATS3s	Moran autocorrelation - lag 3 / weighted by I-state
MATS4s	Moran autocorrelation - lag 4 / weighted by I-state
MATS5s	Moran autocorrelation - lag 5 / weighted by I-state
MATS6s	Moran autocorrelation - lag 6 / weighted by I-state
MATS7s	Moran autocorrelation - lag 7 / weighted by I-state
MATS8s	Moran autocorrelation - lag 8 / weighted by I-state
GATS1c	Geary autocorrelation - lag 1 / weighted by charges
GATS2c	Geary autocorrelation - lag 2 / weighted by charges
GATS3c	Geary autocorrelation - lag 3 / weighted by charges
GATS4c	Geary autocorrelation - lag 4 / weighted by charges
GATS5c	Geary autocorrelation - lag 5 / weighted by charges
GATS6c	Geary autocorrelation - lag 6 / weighted by charges
GATS7c	Geary autocorrelation - lag 7 / weighted by charges
GATS8c	Geary autocorrelation - lag 8 / weighted by charges
GATS1m	Geary autocorrelation - lag 1 / weighted by mass
GATS2m	Geary autocorrelation - lag 2 / weighted by mass
GATS3m	Geary autocorrelation - lag 3 / weighted by mass
GATS4m	Geary autocorrelation - lag 4 / weighted by mass
GATS5m	Geary autocorrelation - lag 5 / weighted by mass
GATS6m	Geary autocorrelation - lag 6 / weighted by mass
GATS7m	Geary autocorrelation - lag 7 / weighted by mass
GATS8m	Geary autocorrelation - lag 8 / weighted by mass
GATS1v	Geary autocorrelation - lag 1 / weighted by van der Waals volumes
GATS2v	Geary autocorrelation - lag 2 / weighted by van der Waals volumes

GATS3v	Geary autocorrelation - lag 3 / weighted by van der Waals volumes
GATS4v	Geary autocorrelation - lag 4 / weighted by van der Waals volumes
GATS5v	Geary autocorrelation - lag 5 / weighted by van der Waals volumes
GATS6v	Geary autocorrelation - lag 6 / weighted by van der Waals volumes
GATS7v	Geary autocorrelation - lag 7 / weighted by van der Waals volumes
GATS8v	Geary autocorrelation - lag 8 / weighted by van der Waals volumes
GATS1e	Geary autocorrelation - lag 1 / weighted by Sanderson electronegativities
GATS2e	Geary autocorrelation - lag 2 / weighted by Sanderson electronegativities
GATS3e	Geary autocorrelation - lag 3 / weighted by Sanderson electronegativities
GATS4e	Geary autocorrelation - lag 4 / weighted by Sanderson electronegativities
GATS5e	Geary autocorrelation - lag 5 / weighted by Sanderson electronegativities
GATS6e	Geary autocorrelation - lag 6 / weighted by Sanderson electronegativities
GATS7e	Geary autocorrelation - lag 7 / weighted by Sanderson electronegativities
GATS8e	Geary autocorrelation - lag 8 / weighted by Sanderson electronegativities
GATS1p	Geary autocorrelation - lag 1 / weighted by polarizabilities
GATS2p	Geary autocorrelation - lag 2 / weighted by polarizabilities
GATS3p	Geary autocorrelation - lag 3 / weighted by polarizabilities
GATS4p	Geary autocorrelation - lag 4 / weighted by polarizabilities
GATS5p	Geary autocorrelation - lag 5 / weighted by polarizabilities
GATS6p	Geary autocorrelation - lag 6 / weighted by polarizabilities
GATS7p	Geary autocorrelation - lag 7 / weighted by polarizabilities
GATS8p	Geary autocorrelation - lag 8 / weighted by polarizabilities
GATS1i	Geary autocorrelation - lag 1 / weighted by first ionization potential
GATS2i	Geary autocorrelation - lag 2 / weighted by first ionization potential
GATS3i	Geary autocorrelation - lag 3 / weighted by first ionization potential
GATS4i	Geary autocorrelation - lag 4 / weighted by first ionization potential
GATS5i	Geary autocorrelation - lag 5 / weighted by first ionization potential
GATS6i	Geary autocorrelation - lag 6 / weighted by first ionization potential
GATS7i	Geary autocorrelation - lag 7 / weighted by first ionization potential
GATS8i	Geary autocorrelation - lag 8 / weighted by first ionization potential
GATS1s	Geary autocorrelation - lag 1 / weighted by I-state
GATS2s	Geary autocorrelation - lag 2 / weighted by I-state
GATS3s	Geary autocorrelation - lag 3 / weighted by I-state
GATS4s	Geary autocorrelation - lag 4 / weighted by I-state
GATS5s	Geary autocorrelation - lag 5 / weighted by I-state
GATS6s	Geary autocorrelation - lag 6 / weighted by I-state
GATS7s	Geary autocorrelation - lag 7 / weighted by I-state
GATS8s	Geary autocorrelation - lag 8 / weighted by I-state

SpAbs_DzZ	Graph energy from Barysz matrix / weighted by atomic number
SpMax_DzZ	Leading eigenvalue from Barysz matrix / weighted by atomic number
SpDiam_DzZ	Spectral diameter from Barysz matrix / weighted by atomic number
SpAD_DzZ	Spectral absolute deviation from Barysz matrix / weighted by atomic number
SpMAD_DzZ	Spectral mean absolute deviation from Barysz matrix / weighted by atomic number
EE_DzZ	Estrada-like index from Barysz matrix / weighted by atomic number (In(1+x))
SM1_DzZ	Spectral moment of order 1 from Barysz matrix / weighted by atomic number
VE1_DzZ	Coefficient sum of the last eigenvector from Barysz matrix / weighted by atomic number
VE2_DzZ	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by atomic number
VE3_DzZ	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by atomic number
VR1_DzZ	Randic-like eigenvector-based index from Barysz matrix / weighted by atomic number
VR2_DzZ	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by atomic number
VR3_DzZ	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by atomic number
SpAbs_Dzm	Graph energy from Barysz matrix / weighted by mass
SpMax_Dzm	Leading eigenvalue from Barysz matrix / weighted by mass
SpDiam_Dzm	Spectral diameter from Barysz matrix / weighted by mass
SpAD_Dzm	Spectral absolute deviation from Barysz matrix / weighted by mass
SpMAD_Dzm	Spectral mean absolute deviation from Barysz matrix / weighted by mass
EE_Dzm	Estrada-like index from Barysz matrix / weighted by mass (ln(1+x))
SM1_Dzm	Spectral moment of order 1 from Barysz matrix / weighted by mass
VE1_Dzm	Coefficient sum of the last eigenvector from Barysz matrix / weighted by mass
VE2_Dzm	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by mass
VE3_Dzm	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by mass
VR1_Dzm	Randic-like eigenvector-based index from Barysz matrix / weighted by mass
VR2_Dzm	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by mass
VR3_Dzm	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by mass
SpAbs_Dzv	Graph energy from Barysz matrix / weighted by van der Waals volumes
SpMax_Dzv	Leading eigenvalue from Barysz matrix / weighted by van der Waals volumes
SpDiam_Dzv	Spectral diameter from Barysz matrix / weighted by van der Waals volumes
SpAD_Dzv	Spectral absolute deviation from Barysz matrix / weighted by van der Waals volumes
SpMAD_Dzv	Spectral mean absolute deviation from Barysz matrix / weighted by van der Waals

	volumes
EE_Dzv	Estrada-like index from Barysz matrix / weighted by van der Waals volumes (ln(1+x))
SM1_Dzv	Spectral moment of order 1 from Barysz matrix / weighted by van der Waals volumes
VE1_Dzv	Coefficient sum of the last eigenvector from Barysz matrix / weighted by van der Waals volumes
VE2_Dzv	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by van der Waals volumes
VE3_Dzv	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by van der Waals volumes
VR1_Dzv	Randic-like eigenvector-based index from Barysz matrix / weighted by van der Waals volumes
VR2_Dzv	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by van der Waals volumes
VR3_Dzv	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by van der Waals volumes
SpAbs_Dze	Graph energy from Barysz matrix / weighted by Sanderson electronegativities
SpMax_Dze	Leading eigenvalue from Barysz matrix / weighted by Sanderson electronegativities
SpDiam_Dze	Spectral diameter from Barysz matrix / weighted by Sanderson electronegativities
SpAD_Dze	Spectral absolute deviation from Barysz matrix / weighted by Sanderson electronegativities
SpMAD_Dze	Spectral mean absolute deviation from Barysz matrix / weighted by Sanderson electronegativities
EE_Dze	Estrada-like index from Barysz matrix / weighted by Sanderson electronegativities (ln(1+x))
SM1_Dze	Spectral moment of order 1 from Barysz matrix / weighted by Sanderson electronegativities
VE1_Dze	Coefficient sum of the last eigenvector from Barysz matrix / weighted by Sanderson electronegativities
VE2_Dze	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by Sanderson electronegativities
VE3_Dze	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by Sanderson electronegativities
VR1_Dze	Randic-like eigenvector-based index from Barysz matrix / weighted by Sanderson electronegativities
VR2_Dze	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by Sanderson electronegativities
VR3_Dze	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by Sanderson electronegativities
SpAbs_Dzp	Graph energy from Barysz matrix / weighted by polarizabilities
SpMax_Dzp	Leading eigenvalue from Barysz matrix / weighted by polarizabilities
SpDiam_Dzp	Spectral diameter from Barysz matrix / weighted by polarizabilities
SpAD_Dzp	Spectral absolute deviation from Barysz matrix / weighted by polarizabilities

SpMAD_Dzp	Spectral mean absolute deviation from Barysz matrix / weighted by polarizabilities
EE_Dzp	Estrada-like index from Barysz matrix / weighted by polarizabilities (In(1+x))
SM1_Dzp	Spectral moment of order 1 from Barysz matrix / weighted by polarizabilities
VE1_Dzp	Coefficient sum of the last eigenvector from Barysz matrix / weighted by polarizabilities
VE2_Dzp	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by polarizabilities
VE3_Dzp	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by polarizabilities
VR1_Dzp	Randic-like eigenvector-based index from Barysz matrix / weighted by polarizabilities
VR2_Dzp	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by polarizabilities
VR3_Dzp	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by polarizabilities
SpAbs_Dzi	Graph energy from Barysz matrix / weighted by first ionization potential
SpMax_Dzi	Leading eigenvalue from Barysz matrix / weighted by first ionization potential
SpDiam_Dzi	Spectral diameter from Barysz matrix / weighted by first ionization potential
SpAD_Dzi	Spectral absolute deviation from Barysz matrix / weighted by first ionization potential
SpMAD_Dzi	Spectral mean absolute deviation from Barysz matrix / weighted by first ionization potential
EE_Dzi	Estrada-like index from Barysz matrix / weighted by first ionization potential (ln(1+x))
SM1_Dzi	Spectral moment of order 1 from Barysz matrix / weighted by first ionization potential
VE1_Dzi	Coefficient sum of the last eigenvector from Barysz matrix / weighted by first ionization potential
VE2_Dzi	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by first ionization potential
VE3_Dzi	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by first ionization potential
VR1_Dzi	Randic-like eigenvector-based index from Barysz matrix / weighted by first ionization potential
VR2_Dzi	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by first ionization potential
VR3_Dzi	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by first ionization potential
SpAbs_Dzs	Graph energy from Barysz matrix / weighted by I-state
SpMax_Dzs	Leading eigenvalue from Barysz matrix / weighted by I-state
SpDiam_Dzs	Spectral diameter from Barysz matrix / weighted by I-state
SpAD_Dzs	Spectral absolute deviation from Barysz matrix / weighted by I-state
SpMAD_Dzs	Spectral mean absolute deviation from Barysz matrix / weighted by I-state
EE_Dzs	Estrada-like index from Barysz matrix / weighted by I-state (In(1+x))

SM1_Dzs	Spectral moment of order 1 from Barysz matrix / weighted by I-state
VE1_Dzs	Coefficient sum of the last eigenvector from Barysz matrix / weighted by I-state
VE2_Dzs	Average coefficient sum of the last eigenvector from Barysz matrix / weighted by I- state
VE3_Dzs	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by I-state
VR1_Dzs	Randic-like eigenvector-based index from Barysz matrix / weighted by I-state
VR2_Dzs	Normalized Randic-like eigenvector-based index from Barysz matrix / weighted by I-state
VR3_Dzs	Logarithmic Randic-like eigenvector-based index from Barysz matrix / weighted by I-state
nBase	Number of basic groups. The list of basic groups is defined by this SMARTS "[\$ ([NH2]-[CX4])]", "[\$([NH](-[CX4])-[CX4])]", "[\$(N(-[CX4])(-[CX4])-[CX4])]", "[\$([*;+;!\$ (*~[*;-])])]", "[\$(N=C-N)]", and "[\$(N-C=N)]" originally presented in JOELib
BCUTw-1I	nhigh lowest atom weighted BCUTS
BCUTw-1h	nlow highest atom weighted BCUTS
BCUTc-1I	nhigh lowest partial charge weighted BCUTS
BCUTc-1h	nlow highest partial charge weighted BCUTS
BCUTp-1I	nhigh lowest polarizability weighted BCUTS
BCUTp-1h	nlow highest polarizability weighted BCUTS
nBonds	Number of bonds (excluding bonds with hydrogen)
nBonds2	Total number of bonds (including bonds to hydrogens)
nBondsS	Number of single bonds (including bonds with hydrogen)
nBondsS2	Total number of single bonds (including bonds to hydrogens, excluding aromatic bonds)
nBondsS3	Total number of single bonds (excluding bonds to hydrogens and aromatic bonds)
nBondsD	Number of double bonds
nBondsD2	Total number of double bonds (excluding bonds to aromatic bonds)
nBondsT	Number of triple bonds
nBondsQ	Number of quadruple bonds
nBondsM	Total number of bonds that have bond order greater than one (aromatic bonds have bond order 1.5).
bpol	Sum of the absolute value of the difference between atomic polarizabilities of all bonded atoms in the molecule (including implicit hydrogens)
SpMax1_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 1 / weighted by relative mass
SpMax2_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative mass
SpMax3_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 3 / weighted by relative mass
SpMax4_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 4 / weighted by relative mass
SpMax5_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 5 / weighted by relative

	mass
SpMax6_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 6 / weighted by relative mass
SpMax7_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 7 / weighted by relative mass
SpMax8_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 8 / weighted by relative mass
SpMin1_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 1 / weighted by relative mass
SpMin2_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative mass
SpMin3_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 3 / weighted by relative mass
SpMin4_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 4 / weighted by relative mass
SpMin5_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 5 / weighted by relative mass
SpMin6_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 6 / weighted by relative mass
SpMin7_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 7 / weighted by relative mass
SpMin8_Bhm	Smallest absolute eigenvalue of Burden modified matrix - n 8 / weighted by relative mass
SpMax1_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 1 / weighted by relative van der Waals volumes
SpMax2_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative van der Waals volumes
SpMax3_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 3 / weighted by relative van der Waals volumes
SpMax4_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 4 / weighted by relative van der Waals volumes
SpMax5_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 5 / weighted by relative van der Waals volumes
SpMax6_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 6 / weighted by relative van der Waals volumes
SpMax7_Bhv	Largest absolute eigenvalue of Burden modified matrix - n 7 / weighted by relative van der Waals volumes

## Anexo-3: Valores de los descriptores moleculares (ev) derivados de la Teoría del Funcional de la Densidad (DFT) para cada inhibidor del Factor de Crecimiento Endoteliar.

Moléculas	номо	LUMO	HOMO- LUMO Gap	I	Α	μ	η	S	x	ω	ω+	ω	Δω⁺
H2a	-5,7547	-1,5562	-4,1984	5,7547	1,5562	-3,6554	2,0992	0,2382	3,6554	3,1827	5,2728	1,6174	6,8902
H2b	-5,7348	-1,5067	-4,2281	5,7348	1,5067	-3,6207	2,1141	0,2365	3,6207	3,1006	5,1753	1,5545	6,7298
H2c	-5,7688	-1,6368	-4,1321	5,7688	1,6368	-3,7028	2,0660	0,2420	3,7028	3,3181	5,4278	1,7250	7,1528
H2d	-5,7696	-1,4463	-4,3233	5,7696	1,4463	-3,6080	2,1617	0,2313	3,6080	3,0109	5,0851	1,4772	6,5623
H3	-5,5830	-1,6961	-3,8869	5,5830	1,6961	-3,6395	1,9434	0,2573	3,6395	3,4079	5,4706	1,8311	7,3017
H4	-5,5941	-1,3606	-4,2335	5,5941	1,3606	-3,4773	2,1168	0,2362	3,4773	2,8562	4,8595	1,3821	6,2416
H5	-5,5691	-1,8019	-3,7671	5,5691	1,8019	-3,6855	1,8836	0,2655	3,6855	3,6056	5,6838	1,9983	7,6821
H6	-5,4551	-1,5617	-3,8934	5,4551	1,5617	-3,5084	1,9467	0,2568	3,5084	3,1614	5,1589	1,6506	6,8095
M4	-6,0583	-2,0253	-4,0330	6,0583	2,0253	-4,0418	2,0165	0,2480	4,0418	4,0507	6,3237	2,2818	8,6055
M10a	-6,0295	-1,8455	-4,1840	6,0295	1,8455	-3,9375	2,0920	0,2390	3,9375	3,7055	5,9357	1,9982	7,9340
M10b	-5,9506	-1,8101	-4,1405	5,9506	1,8101	-3,8803	2,0702	0,2415	3,8803	3,6365	5,8355	1,9552	7,7907
M10c	-5,9098	-1,9644	-3,9454	5,9098	1,9644	-3,9371	1,9727	0,2535	3,9371	3,9288	6,1439	2,2068	8,3508
M10d	-6,1049	-2,6411	-3,4637	6,1049	2,6411	-4,3730	1,7319	0,2887	4,3730	5,5210	7,9240	3,5509	11,4749
M10e	-6,1234	-2,6458	-3,4776	6,1234	2,6458	-4,3846	1,7388	0,2876	4,3846	5,5281	7,9377	3,5531	11,4908
M10f	-5,9465	-1,9418	-4,0047	5,9465	1,9418	-3,9442	2,0024	0,2497	3,9442	3,8845	6,1069	2,1627	8,2696
M10g	-5,9903	-1,9394	-4,0510	5,9903	1,9394	-3,9648	2,0255	0,2469	3,9648	3,8805	6,1161	2,1513	8,2675
M11	-5,8730	-1,8893	-3,9837	5,8730	1,8893	-3,8812	1,9919	0,2510	3,8812	3,7812	5,9708	2,0896	8,0604
M13	-5,9925	-2,6041	-3,3884	5,9925	2,6041	-4,2983	1,6942	0,2951	4,2983	5,4526	7,8136	3,5152	11,3288
M14	-5,9917	-2,0460	-3,9457	5,9917	2,0460	-4,0189	1,9728	0,2534	4,0189	4,0934	6,3494	2,3306	8,6800
M15	-6,0028	-2,1456	-3,8572	6,0028	2,1456	-4,0742	1,9286	0,2593	4,0742	4,3034	6,5816	2,5074	9,0890
M16	-6,0385	-2,0126	-4,0259	6,0385	2,0126	-4,0255	2,0130	0,2484	4,0255	4,0251	6,2895	2,2640	8,5535
M17	-5,8066	-1,7199	-4,0867	5,8066	1,7199	-3,7633	2,0434	0,2447	3,7633	3,4654	5,6025	1,8392	7,4417
M18	-5,9653	-1,9301	-4,0352	5,9653	1,9301	-3,9477	2,0176	0,2478	3,9477	3,8621	6,0882	2,1405	8,2286
M19	-5,9446	-1,8975	-4,0472	5,9446	1,8975	-3,9210	2,0236	0,2471	3,9210	3,7988	6,0123	2,0913	8,1036
M21	-6,0499	-2,4313	-3,6186	6,0499	2,4313	-4,2406	1,8093	0,2764	4,2406	4,9696	7,3161	3,0755	10,3915
M23a	-5,9373	-1,8855	-4,0518	5,9373	1,8855	-3,9114	2,0259	0,2468	3,9114	3,7758	5,9847	2,0734	8,0581
M23b	-5,9100	-1,9187	-3,9913	5,9100	1,9187	-3,9144	1,9957	0,2505	3,9144	3,8389	6,0456	2,1312	8,1768
M23c	-5,8896	-1,8868	-4,0028	5,8896	1,8868	-3,8882	2,0014	0,2498	3,8882	3,7770	5,9712	2,0830	8,0543
M23d	-5,9558	-1,8425	-4,1133	5,9558	1,8425	-3,8991	2,0566	0,2431	3,8991	3,6961	5,9028	2,0036	7,9064
M23e	-5,8379	-1,9277	-3,9103	5,8379	1,9277	-3,8828	1,9551	0,2557	3,8828	3,8555	6,0413	2,1585	8,1998
M23f	-5,8967	-1,8460	-4,0507	5,8967	1,8460	-3,8714	2,0253	0,2469	3,8714	3,7000	5,8888	2,0175	7,9063
M2	-6,0094	-2,0632	-3,9462	6,0094	2,0632	-4,0363	1,9731	0,2534	4,0363	4,1284	6,3932	2,3569	8,7501
M5a	-6,0760	-1,8988	-4,1772	6,0760	1,8988	-3,9874	2,0886	0,2394	3,9874	3,8062	6,0610	2,0736	8,1346

M5b	-6,3005	-1,9418	-4,3587	6,3005	1,9418	-4,1212	2,1794	0,2294	4,1212	3,8966	6,2296	2,1084	8,3380
M5c	-6,3258	-1,9486	-4,3772	6,3258	1,9486	-4,1372	2,1886	0,2285	4,1372	3,9104	6,2526	2,1153	8,3679
M5d	-5,9705	-1,9225	-4,0480	5,9705	1,9225	-3,9465	2,0240	0,2470	3,9465	3,8475	6,0737	2,1273	8,2010
M6a	-6,2619	-1,8912	-4,3707	6,2619	1,8912	-4,0765	2,1853	0,2288	4,0765	3,8022	6,1136	2,0371	8,1507
M6b	-6,3588	-2,0564	-4,3024	6,3588	2,0564	-4,2076	2,1512	0,2324	4,2076	4,1148	6,4875	2,2799	8,7674
M8a	-6,2978	-2,1328	-4,1650	6,2978	2,1328	-4,2153	2,0825	0,2401	4,2153	4,2663	6,6342	2,4189	9,0532
M8b	-6,5751	-2,2267	-4,3484	6,5751	2,2267	-4,4009	2,1742	0,2300	4,4009	4,4541	6,9263	2,5254	9,4517
O1b	-6,1884	-1,1412	-5,0472	6,1884	1,1412	-3,6648	2,5236	0,1981	3,6648	2,6611	4,8090	1,1441	5,9531
O5a	-6,2015	-1,1554	-5,0461	6,2015	1,1554	-3,6784	2,5230	0,1982	3,6784	2,6815	4,8361	1,1576	5,9937
O5b	-3,2017	-1,0917	-2,1099	3,2017	1,0917	-2,1467	1,0550	0,4739	2,1467	2,1841	3,3893	1,2426	4,6319
O5c	-6,1683	-1,1804	-4,9878	6,1683	1,1804	-3,6744	2,4939	0,2005	3,6744	2,7068	4,8557	1,1813	6,0370
O10a	-6,1988	-1,2471	-4,9517	6,1988	1,2471	-3,7229	2,4758	0,2020	3,7229	2,7991	4,9700	1,2471	6,2172
O10b	-6,2796	-1,1266	-5,1530	6,2796	1,1266	-3,7031	2,5765	0,1941	3,7031	2,6611	4,8347	1,1316	5,9663
O10c	-6,1609	-1,3121	-4,8488	6,1609	1,3121	-3,7365	2,4244	0,2062	3,7365	2,8794	5,0507	1,3142	6,3649
O11a	-6,0638	-1,0642	-4,9996	6,0638	1,0642	-3,5640	2,4998	0,2000	3,5640	2,5407	4,6351	1,0711	5,7063
O11b	-6,1710	-1,1380	-5,0330	6,1710	1,1380	-3,6545	2,5165	0,1987	3,6545	2,6535	4,7953	1,1409	5,9362
O11c	-6,2439	-1,1799	-5,0640	6,2439	1,1799	-3,7119	2,5320	0,1975	3,7119	2,7208	4,8933	1,1814	6,0746
O11d	-6,2249	-1,8670	-4,3579	6,2249	1,8670	-4,0459	2,1790	0,2295	4,0459	3,7563	6,0516	2,0057	8,0573
O11e	-4,2466	-1,1151	-3,1315	4,2466	1,1151	-2,6809	1,5657	0,3193	2,6809	2,2951	3,8312	1,1504	4,9816
O11f	-5,7160	-1,0980	-4,6180	5,7160	1,0980	-3,4070	2,3090	0,2165	3,4070	2,5135	4,5057	1,0987	5,6043
O11g	-6,1822	-1,1715	-5,0107	6,1822	1,1715	-3,6768	2,5054	0,1996	3,6768	2,6980	4,8496	1,1728	6,0223
O11h	-5,8924	-1,2672	-4,6251	5,8924	1,2672	-3,5798	2,3126	0,2162	3,5798	2,7707	4,8497	1,2699	6,1196
011i	-6,1226	-1,1293	-4,9933	6,1226	1,1293	-3,6259	2,4966	0,2003	3,6259	2,6330	4,7580	1,1321	5,8901
O11j	-6,0899	-1,3445	-4,7454	6,0899	1,3445	-3,7172	2,3727	0,2107	3,7172	2,9118	5,0670	1,3498	6,4168
O11k	-6,1147	-1,3336	-4,7810	6,1147	1,3336	-3,7242	2,3905	0,2092	3,7242	2,9009	5,0618	1,3376	6,3994
0111	-4,7141	-0,9948	-3,7193	4,7141	0,9948	-2,8545	1,8596	0,2689	2,8545	2,1908	3,8505	0,9960	4,8464
O11m	-5,0425	-0,9742	-4,0684	5,0425	0,9742	-3,0084	2,0342	0,2458	3,0084	2,2245	3,9830	0,9746	4,9576
O11n	-6,1226	-1,1293	-4,9933	6,1226	1,1293	-3,6259	2,4966	0,2003	3,6259	2,6330	4,7580	1,1321	5,8901
O11o	-5,4281	-0,9189	-4,5092	5,4281	0,9189	-3,1735	2,2546	0,2218	3,1735	2,2335	4,1021	0,9286	5,0306
O11p	-5,2749	-0,9878	-4,2872	5,2749	0,9878	-3,1314	2,1436	0,2333	3,1314	2,2871	4,1208	0,9894	5,1102
O12a	-6,2935	-1,5380	-4,7555	6,2935	1,5380	-3,9157	2,3777	0,2103	3,9157	3,2243	5,4793	1,5636	7,0430
O12b	-6,3452	-1,4403	-4,9049	6,3452	1,4403	-3,8927	2,4524	0,2039	3,8927	3,0895	5,3424	1,4496	6,7920
O14	-2,4591	-0,7497	-1,7094	2,4591	0,7497	-1,6044	0,8547	0,5850	1,6044	1,5058	2,4148	0,8104	3,2253
O10	-5,8445	-1,7753	-4,0693	5,8445	1,7753	-3,8099	2,0346	0,2457	3,8099	3,5671	5,7264	1,9165	7,6428
O11	-5,4823	-1,0381	-4,4442	5,4823	1,0381	-3,2602	2,2221	0,2250	3,2602	2,3917	4,2995	1,0393	5,3388
O13a	-6,1492	-1,5633	-4,5859	6,1492	1,5633	-3,8563	2,2930	0,2181	3,8563	3,2427	5,4574	1,6012	7,0586
O13b	-6,1634	-2,1870	-3,9764	6,1634	2,1870	-4,1752	1,9882	0,2515	4,1752	4,3839	6,7200	2,5448	9,2648
O16	-6,1310	-1,4082	-4,7228	6,1310	1,4082	-3,7696	2,3614	0,2117	3,7696	3,0088	5,1887	1,4191	6,6079

O18a	-6,0102	-1,2286	-4,7816	6,0102	1,2286	-3,6194	2,3908	0,2091	3,6194	2,7397	4,8482	1,2288	6,0770
O18b	-5,4148	-1,2471	-4,1677	5,4148	1,2471	-3,3309	2,0838	0,2399	3,3309	2,6622	4,5881	1,2572	5,8453
O18c	-3,2648	-1,1309	-2,1339	3,2648	1,1309	-2,1979	1,0670	0,4686	2,1979	2,2637	3,4960	1,2982	4,7942
O20a	-6,1884	-1,1412	-5,0472	6,1884	1,1412	-3,6648	2,5236	0,1981	3,6648	2,6611	4,8090	1,1441	5,9531
O20b	-6,2812	-1,1203	-5,1609	6,2812	1,1203	-3,7008	2,5805	0,1938	3,7008	2,6537	4,8266	1,1259	5,9525
O20c	-3,1282	-0,9320	-2,1962	3,1282	0,9320	-2,0301	1,0981	0,4553	2,0301	1,8765	3,0289	0,9988	4,0276
O20d	-6,1343	-1,1178	-5,0164	6,1343	1,1178	-3,6261	2,5082	0,1993	3,6261	2,6210	4,7476	1,1215	5,8691
O23	-2,8284	-1,8498	-0,9785	2,8284	1,8498	-2,3391	0,4893	1,0219	2,3391	5,5914	6,8221	4,4831	11,3052
04	-5,9348	-1,1674	-4,7674	5,9348	1,1674	-3,5511	2,3837	0,2098	3,5511	2,6451	4,7186	1,1675	5,8861
O5	-6,1852	-1,1535	-5,0317	6,1852	1,1535	-3,6693	2,5158	0,1987	3,6693	2,6758	4,8250	1,1557	5,9806
O7a	-5,8562	-0,9889	-4,8673	5,8562	0,9889	-3,4225	2,4337	0,2055	3,4225	2,4066	4,4221	0,9995	5,4216
O7b	-5,8170	-0,9135	-4,9035	5,8170	0,9135	-3,3652	2,4517	0,2039	3,3652	2,3095	4,2986	0,9334	5,2320
07c	-5,8066	-0,9075	-4,8991	5,8066	0,9075	-3,3571	2,4496	0,2041	3,3571	2,3004	4,2851	0,9280	5,2132
O7d	-5,9441	-1,0041	-4,9400	5,9441	1,0041	-3,4741	2,4700	0,2024	3,4741	2,4432	4,4890	1,0149	5,5039
O7e	-5,9857	-1,1815	-4,8042	5,9857	1,1815	-3,5836	2,4021	0,2082	3,5836	2,6731	4,7652	1,1816	5,9468
O7f	-5,9577	-1,1157	-4,8420	5,9577	1,1157	-3,5367	2,4210	0,2065	3,5367	2,5832	4,6542	1,1175	5,7717
O7g	-6,0336	-1,0324	-5,0012	6,0336	1,0324	-3,5330	2,5006	0,2000	3,5330	2,4958	4,5749	1,0419	5,6168
O7h	-5,9707	-1,0071	-4,9636	5,9707	1,0071	-3,4889	2,4818	0,2015	3,4889	2,4523	4,5070	1,0181	5,5251
07i	-6,0276	-1,2585	-4,7691	6,0276	1,2585	-3,6431	2,3845	0,2097	3,6431	2,7829	4,9025	1,2594	6,1620
O7j	-5,8341	-0,9848	-4,8493	5,8341	0,9848	-3,4095	2,4247	0,2062	3,4095	2,3971	4,4049	0,9955	5,4004
08	-5,9898	-2,1707	-3,8191	5,9898	2,1707	-4,0802	1,9096	0,2618	4,0802	4,3592	6,6380	2,5577	9,1957
<b>O</b> 9	-6,0290	-0,9883	-5,0406	6,0290	0,9883	-3,5086	2,5203	0,1984	3,5086	2,4423	4,5116	1,0030	5,5146
O1a	-6,0578	-2,6893	-3,3685	6,0578	2,6893	-4,3736	1,6842	0,2969	4,3736	5,6785	8,0758	3,7022	11,7780
O1b2	-6,0238	-2,0341	-3,9897	6,0238	2,0341	-4,0289	1,9949	0,2506	4,0289	4,0685	6,3323	2,3034	8,6357
O1c	-5,9958	-2,7388	-3,9617	5,9958	2,0341	-4,0149	1,9809	0,2524	4,0149	4,0688	6,3239	2,3090	8,6329
O1d	-6,0096	-2,7070	-3,3026	6,0096	2,7070	-4,3583	1,6513	0,3028	4,3583	5,7514	8,1370	3,7787	11,9157
O1e	-5,9794	-2,0305	-3,9489	5,9794	2,0305	-4,0050	1,9745	0,2532	4,0050	4,0618	6,3111	2,3061	8,6173
O1f	-5,9881	-2,0221	-3,9661	5,9881	2,0221	-4,0051	1,9830	0,2521	4,0051	4,0445	6,2950	2,2899	8,5848
O1g	-5,9960	-2,7239	-3,2722	5,9960	2,7239	-4,3599	1,6361	0,3056	4,3599	5,8093	8,1938	3,8339	12,0277
O1h	-5,9960	-2,1230	-3,8730	5,9960	2,1230	-4,0595	1,9365	0,2582	4,0595	4,2551	6,5269	2,4673	8,9942
<b>O</b> 1i	-5,9383	-2,0267	-3,9116	5,9383	2,0267	-3,9825	1,9558	0,2556	3,9825	4,0547	6,2904	2,3079	8,5983
O1j	-5,8924	-2,3717	-3,5206	5,8924	2,3717	-4,1321	1,7603	0,2840	4,1321	4,8497	7,1357	3,0037	10,1394
O1k	-5,8774	-2,1810	-3,6964	5,8774	2,1810	-4,0292	1,8482	0,2705	4,0292	4,3919	6,6376	2,6084	9,2459
O1m	-5,8651	-1,7995	-4,0657	5,8651	1,7995	-3,8323	2,0328	0,2460	3,8323	3,6124	5,7826	1,9503	7,7330
O1n	-5,9443	-2,4327	-3,5116	5,9443	2,4327	-4,1885	1,7558	0,2848	4,1885	4,9959	7,3096	3,1211	10,4307
02	-5,8423	-1,8204	-4,0218	5,8423	1,8204	-3,8314	2,0109	0,2486	3,8314	3,6499	5,8170	1,9856	7,8025
O3a	-5,7729	-2,1377	-3,6352	5,7729	2,1377	-3,9553	1,8176	0,2751	3,9553	4,3036	6,5085	2,5532	9,0617
O3b	-5,8714	-2,3309	-3,5405	5,8714	2,3309	-4,1012	1,7702	0,2824	4,1012	4,7507	7,0225	2,9213	9,9439

O4a	-5,8009	-1,7608	-4,0401	5,8009	1,7608	-3,7809	2,0200	0,2475	3,7809	3,5383	5,6813	1,9004	7,5817
O4b	-3,7938	-1,6392	-2,1546	3,7938	1,6392	-2,7165	1,0773	0,4641	2,7165	3,4250	4,9179	2,2014	7,1193
O52	-5,9778	-2,1551	-3,8227	5,9778	2,1551	-4,0665	1,9113	0,2616	4,0665	4,3258	6,5980	2,5315	9,1295
O6a	-6,0099	-1,7902	-4,2197	6,0099	1,7902	-3,9001	2,1098	0,2370	3,9001	3,6047	5,8184	1,9184	7,7368
O6b	-6,0271	-1,7802	-4,2469	6,0271	1,7802	-3,9036	2,1234	0,2355	3,9036	3,5881	5,8053	1,9017	7,7070
O6c	-5,9871	-1,9739	-4,0131	5,9871	1,9739	-3,9805	2,0066	0,2492	3,9805	3,9481	6,1892	2,2087	8,3978
O6d	-6,0676	-1,9451	-4,1225	6,0676	1,9451	-4,0063	2,0613	0,2426	4,0063	3,8934	6,1542	2,1479	8,3021
O6e	-6,1792	-1,8730	-4,3062	6,1792	1,8730	-4,0261	2,1531	0,2322	4,0261	3,7641	6,0463	2,0203	8,0666