

### **Información Importante**

La Universidad Santo Tomás, informa que el autor ha autorizado a usuarios internos y externos de la institución a consultar el contenido de este documento a través del catálogo en línea, página web y Repositorio Institucional del CRAI-USTA, así como en las redes sociales y demás sitios web de información del país y del exterior con las cuales tenga convenio la Universidad.

Se permite la consulta a los usuarios interesados en el contenido de este documento, para todos los usos que tengan **finalidad académica**, siempre y cuando mediante la correspondiente cita bibliográfica se le dé crédito al trabajo de grado y a su autor, nunca para usos comerciales.

De conformidad con lo establecido en el Artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, la Universidad Santo Tomás informa que “los derechos morales sobre documento son propiedad del autor, los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.”

**Centro de Recursos para el Aprendizaje y la Investigación, CRAI-USTA  
Universidad Santo Tomás, Bucaramanga**

**Construcción de Modelos Relación Cuantitativa Estructura-Actividad para  
Predecir la Concentración Inhibitoria Mínima de Mono Fenoles frente a *Streptococcus  
Sobrinus* y *Porphyromonas Gingivalis*.**

**Danilo Andrés Blanco García**

**Trabajo de grado para optar el título de Químico Ambiental**

**Director**

**Ciro Eduardo Rozo Correa**

**Msc. en Química**

**Universidad Santo Tomás, Bucaramanga**

**División de Ingenierías y Arquitectura**

**Facultad de Química Ambiental**

**2021**

**Contenido**

	Pág.
Introducción .....	10
1. Construcción de Modelos Relación Cuantitativa Estructura-Actividad para Predecir la Concentración Inhibitoria Mínima de Mono Fenoles frente a Streptococcus Sobrinus y Porphyromonas Gingivalis.....	11
1.1 Planteamiento del Problema.....	11
1.2 Justificación.....	12
1.3 Objetivos .....	14
1.3.1 Objetivo General .....	14
1.3.2 Objetivos Específicos .....	14
2. Marco Referencial.....	14
2.1 Marco Teórico.....	14
2.1.1 S. Sobrinus.....	15
2.1.2 Inflación de las Encías en la Raíz del Diente .....	15
2.1.3 Porphyromonas Gingivalis .....	16
2.1.4 Uso de Herramientas Computacionales en el Estudio de las Bio actividades.....	16
2.1.5 Modelamiento QSAR .....	17
2.2 Marco de Antecedentes .....	20
2.2.1 Estudio QSAR Sobre Crecimiento Bacteriano Utilizando Índices Topológicos Tipo Conectividad.....	20
3. Diseño Metodológico.....	21

3.1 Selección de Compuestos con Bioactividad frente a Streptococcus Sobrinus y Porphyromonas Gingivalis.....	22
3.2 Construcción del Modelo Matemático para la Predicción de la Concentración Inhibitoria Mínima Frente a Streptococcus Sobrinus Y Porphyromonas Gingivalis.....	22
3.2.1 Cálculo de los Descriptores Moleculares.....	22
3.2.2 Construcción y Validación del Modelo Matemático.....	24
4. Resultados y Análisis.....	25
4.1 Optimización y Selección de Compuestos con Bioactividad Frente a Streptococcus Sobrinus y Porphyromonas Gingivalis para la Construcción del Modelo QSAR.....	25
4.1.1 Optimización.....	26
4.1.2 Selección.....	26
4.1.3 Selección de los Descriptores Moleculares.....	28
4.2 Construcción del Modelo Relación Cuantitativa Estructura Actividad para la Predicción de Concentración Inhibitoria Mínima de Monofenoles Frente a Streptococcus Sobrinus y Porphyromonas Gingivalis.....	31
4.2.1 Modelo QSAR para Predecir la MIC de Monofenoles de Porphyromonas Gingivalis	32
4.2.2 Modelo QSAR para Predecir la MIC de Monofenoles de S. Sobrinus.....	34
4.3 Validación de los Modelos Relación Cuantitativa Estructura Actividad para la Predicción de Concentración Inhibitoria Mínima de Monofenoles Frente a Streptococcus Sobrinus y Porphyromonas Gingivalis.....	36
4.3.1 Validación Interna.....	36
4.3.2 Validación Externa.....	40
5. Conclusiones.....	46

Referencias..... 49

Apéndices..... 55

**Lista de Tablas**

	Pág.
<b>Tabla 1.</b> <i>Correlación de Descriptores Vs LOG(1/MIC) del Total de Monofenoles.</i> .....	28
<b>Tabla 2.</b> <i>Correlación de Descriptores Vs LOG(1/MIC) de Monofenoles Seleccionados.</i> .....	29
<b>Tabla 3.</b> <i>Correlación de Descriptores Vs LOG(1/MIC) de Monofenoles Excluidos.</i> .....	29
<b>Tabla 4.</b> <i>Descriptores Seleccionados y su Valor de R.</i> .....	30
<b>Tabla 5.</b> <i>Modelos QSAR para Porphyromonas Gingivalis.</i> .....	32
<b>Tabla 6.</b> <i>Hipótesis Nula y Alternativa, Prueba F y Z. P. Gingivalis.</i> .....	33
<b>Tabla 7.</b> <i>Modelo QSAR para S. Sobrinus.</i> .....	34
<b>Tabla 8.</b> <i>Hipótesis Nula y Alternativa, Prueba F y Z S. Sobrinus.</i> .....	35
<b>Tabla 9.</b> <i>Validación Interna.</i> .....	36
<b>Tabla 10.</b> <i>Residuales Compuestos Limite. P. Gingivalis.</i> .....	37
<b>Tabla 11.</b> <i>Residuales Compuestos Limite. S. Sobrinus.</i> .....	38
<b>Tabla 12.</b> <i>Parámetros Compuestos Límite. S. Sobrinus.</i> .....	39
<b>Tabla 13.</b> <i>Parámetros Validación Datos Periféricos a la Curva.</i> .....	40
<b>Tabla 14.</b> <i>Selección de Moléculas Validación Externa.</i> .....	41
<b>Tabla 15.</b> <i>Resultados Validación Externa.</i> .....	42
<b>Tabla 16.</b> <i>Error Modelos P.Gingivalis.</i> .....	44
<b>Tabla 17.</b> <i>Error Modelos S. Sobrinus.</i> .....	45

**Lista de Figuras**

	Pág.
<b>Figura 1.</b> $N_{max}$ Vs $(\Delta E_e - \Delta E_n)^2$ .....	27
<b>Figura 2.</b> <i>Correlación Valores Calculados Vs Valores Experimentales. P Gingivalis</i> .....	34
<b>Figura 3.</b> <i>Correlación Valores Calculados Vs Valores Experimentales. S. Sobrinus</i> .....	35
<b>Figura 4.</b> <i>Correlación Validación Externa P Gingivalis.</i> .....	43
<b>Figura 5.</b> <i>Correlación Validación Externa S. Sobrinus.</i> .....	43

### Resumen

Las patologías periodontales de mayor relevancia a nivel mundial son la caries y la gingivitis siendo *Streptococcus Sobrinus* y *Porphyromonas Gingivalis*, los microorganismos de mayor patogenicidad. Estas se encuentran en el puesto 11 de enfermedades de prevalencia humana, y tiene serias consecuencias para los afectados, incluyendo problemas de masticación y del habla, deteriorando el bienestar y la calidad de vida. Mediante QSAR se establecen modelos que predicen bioactividades de monofenoles con  $Q^2$  de 0,96, y 0,94 y valores de  $R^2$  de 0,89 y 0,94 para *P. Gingivalis* y *S. Sobrinus* respectivamente.

*Palabras Clave:* QSAR, *S. Sobrinus*, *P. Gingivalis*, monofenol, periodontal, Validación cruzada.

**Abstract**

The most relevant periodontal pathologies worldwide are caries and gingivitis, *Streptococcus Sobrinus* and *Porphyromonas Gingivalis* being the most pathogenic microorganisms. These are ranked 11th in human prevalence diseases, and have serious consequences for those affected, including chewing and speech problems, deteriorating well-being and quality of life. Using QSAR, models are established that predict bioactivities of monophenols with  $Q^2$  of 0.96, and 0.94 and  $R^2$  values of 0.89 and 0.94 for *P. Gingivalis* and *S. Sobrinus* respectively.

Keywords: QSAR, *S. Sobrinus*, *P. Gingivalis*, monophenol, periodontal, Cross validation.

### Glosario

*Bio -film*: Estructura de bacterias constituida poliméricamente de proteínas, en la que los microorganismos se organizan en comunidades que comparten nutrientes, metabolitos y elementos genéticos (Zhu et al, 2019).

*Descriptor molecular*: Número que describe la estructura química o una propiedad de la molécula o fragmento de esta (Carrasco, 2003)

*Distribución de fisher-snedecor*: Define la relación entre la varianza explicada ( $R^2$ ) e inexplicada ( $1-R^2$ ) para un determinado número de grados de libertad (Gujarati y Porter, 2009, pag 136).

*Error estándar de predicción*: Medida de las diferencias entre los valores observados y los valores predichos que se obtienen por medio de una ecuación de regresión (Gujarati y Porter, 2009 pag 127, 136).

## Introducción

Modelamiento de la relación cuantitativa estructura -actividad QSAR por sus siglas en inglés, es ampliamente utilizado por la academia, industria e instituciones gubernamentales a nivel mundial, enfocados a determinar los impactos de compuestos químicos, materiales y nano materiales sobre la salud humana y el medioambiente, pero tal vez el mayor enfoque se encuentra en la química médica, en la que los químicos computacionales afirman que la principal aplicación es el estudio y/o diseño de compuestos con bioactividad, permitiendo al investigador enfocarse en un grupo de moléculas específicas evitando el gasto de recursos y reduciendo el número de experimentaciones, (Cherkasov and coworker, 2014).

La estrategia básica es valorar la actividad biológica como variable dependiente del conjunto de descriptores moleculares que constituyen las variables independientes, mediante una ecuación matemática tratada estadísticamente por mínimos cuadrados, métodos multivariados y análisis por componente que pueda ser usada para la predicción de propiedades de las moléculas, incluyendo aquellas que todavía no han sido sintetizadas, (Cherkasov and coworker, 2014). Las relaciones se fundamentan en la hipótesis de la existencia de una dependencia continua entre las propiedades experimentales y algunos factores estructurales intrínsecos de la molécula.

En este trabajo se obtienen relaciones cuantitativas estructura actividad para algunos monofenoles mediante el estudio de una serie de moléculas que presentan bioactividad contra los patógenos orales *Streptococcus Sobrinus*, y *Porphyromonas Gingivalis*. La propiedad en estudio es la concentración inhibitoria mínima, MIC, cuyos datos son recolectados de la literatura. Dentro de los aspectos importantes esta que los modelos locales nos permiten descubrir qué tipo de propiedades moleculares tienen mayor influencia en el comportamiento germicida de las

moléculas estudiadas, lo cual es de suma importancia para la aproximación al tratamiento de afectaciones periodontales.

## **1. Construcción de Modelos Relación Cuantitativa Estructura-Actividad para Predecir la Concentración Inhibitoria Mínima de Mono Fenoles frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis*.**

### **1.1 Planteamiento del Problema**

Las dos patologías orales que representan el mayor índice de consultas odontológicas a nivel mundial y son el principal causal de la pérdida de dientes, son las afecciones periodontales: caries y gingivitis, ubicándole en el puesto 11 de enfermedades de prevalencia humana, (Vos et al, 2017).

La caries dental es una enfermedad crónica, compleja y multifactorial que resulta de un proceso dinámico de desmineralización y remineralización del esmalte dental en función a las variaciones del pH en la cavidad oral en contacto con los depósitos microbianos y que a lo largo del tiempo puede o no generar pérdida del diente en concomitancia con procesos inflamatorios de las encías, (Núñez y García, 2010). El patógeno oral se adhiere a las estructuras solidas de la cavidad oral, mediante el bio film, que consiste una estructura de bacterias constituida poliméricamente de proteínas, en la que los microorganismos se organizan en comunidades que comparten nutrientes, metabolitos y elementos genéticos, adquiriendo resistencia a los cambios medioambientales y procesos de erradicación, (Araujo et al, 2017), considerando a *Streptococcus Sobrinus*, como el de mayor patogenicidad, (Ojeda, et al., 2013).

La Organización Mundial de la Salud, OMS, (2016), asegura que la mitad de la población mundial padece de caries y a causa de ello en algunos países asiáticos y del pacífico es alta la incidencia de cáncer oral. En nuestro territorio el Estudio Nacional de Salud Bucal, ENSAB (2014), muestra que en Colombia el 33% de los niños de 1, 3 y 5 años han presentado antecedentes de la enfermedad, y el 92% de las personas entre 12 y 79 años ha tenido caries dental en algún momento de su vida, y en algunas regiones del país la situación es aún más crítica, por ejemplo, en la Región Atlántica y Oriental la prevalencia se encuentra por encima del valor promedio Nacional.

Por su lado gingivitis, es la inflamación de los tejidos de soporte del diente, con pérdida de hueso y ligamento periodontal, desde la encía, con *Porphyromonas gingivalis* su principal representante (Melendes et al, 2016).

El Estudio Internacional sobre la carga de las Enfermedades realizado en 291 países, arroja que los tratamientos para las afecciones periodontales no están disponibles de manera generalizada; y tampoco son asequibles para mucha gente y esta tiene serias consecuencias para los afectados, incluyendo problemas de masticación y del habla, deteriorando el bienestar y la calidad de vida. Representa una carga global en las enfermedades bucodentales con impacto significativo a nivel social, económico y en los sistemas sanitarios, (FDI, 2015).

## **1.2 Justificación**

La forma ideal de contrarrestar la acción de la patogenia oral es deteniendo su desarrollo (Araujo et al, 2017), y en este sentido la investigación de nuevos productos que interfieran con la adhesión, comunicación o el establecimiento de la matriz polisacárida, es relevante.

Los modelos QSAR mediante descriptores moleculares pueden ser utilizados para entender la relación entre la estructura y la actividad biológica de la molécula estudio, y predecir el comportamiento del compuesto mediante un algoritmo, lo cual es de suma importancia en la racionalización del mecanismo de acción (Todeschini and Consonni, 2000), y la identificación de la presencia o ausencia de propiedades peligrosas de las sustancias el cual en ciertos casos evita la experimentación con animales (Gramatica, 2007).

El algoritmo en general es de la forma:

$$\text{Actividad biológica} = f(\text{descriptor molecular})$$

Dónde la actividad es expresada como Log (1/concentración mínima), y representa el valor mínimo de una molécula para generar una respuesta específica sobre un patógeno seleccionado.

En este trabajo se muestran relaciones cuantitativas estructura propiedad para algunos monofenoles que presentan bioactividad contra los patógenos orales *Streptococcus Sobrinus* y *Porphyromonas gingivalis*. La estrategia básica es valorar la actividad biológica como variable dependiente del conjunto de descriptores moleculares que constituyen las variables independientes, mediante una ecuación matemática tratada estadísticamente por mínimos cuadrados, (Cherkasov and coworker, 2014). La propiedad en estudio es la Concentración inhibitoria mínima, MIC, cuyos datos son recolectados de Shapiro y Guggenheim (1998), y Joshi y colaboradores (2012). Dentro de los aspectos importantes esta que los modelos permiten la predicción de la bioactividad frente a *S. sobrinus* y *Porphyromonas gingivalis*.

### 1.3 Objetivos

#### 1.3.1 Objetivo General

Desarrollar modelos de relación cuantitativa actividad de monofenoles que predigan la Concentración inhibitoria mínima frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis*.

#### 1.3.2 Objetivos Específicos

Construir modelos QSAR de monofenoles que predigan la Concentración inhibitoria mínima frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis* empleando la metodología estadística de la regresión lineal múltiple (MLR).

Validar el modelo QSAR por medio del método de validación cruzada de tipo *leave-one-out* o interna y externa.

Predecir la Concentración inhibitoria mínima frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis* empleando compuestos no incluidos en la construcción del modelo QSAR.

## 2. Marco Referencial

### 2.1 Marco Teórico

Las enfermedades orales son el mayor problema de salud pública en el mundo. Al rededor del planeta se cuenta con una prevalencia aproximada con un 12% a 27% en niños de 2 a 3 años de edad, del 27% al 48% en niños entre los 4 y 6 años, con más del 76% en el Oriente Medio, y para las comunidades indígenas de Australia, Estados Unidos y Canadá, del 60% a más del 90%, (Seow, 2018). Se estima que un tercio de la población geriátrica mundial padece de caries y

gingivitis, un 62% en china, 10% en Canadá y 71 % en Estados Unidos, (Zhang et al., 2019). En el departamento de Santander el 51,5% y 39,8% de diagnósticos de atenciones en los servicios de consulta externa y de urgencias, respectivamente, corresponden a estas patologías. Mediante la adición de indicadores relacionados con una mejor caracterización del estado de salud oral, incluyendo el índice de placa bacteriana (IPB) como factor de riesgo de caries dental se determinó que el 2,7% de los niños de 12 años prevaleció la enfermedad, (Otero et al., 2014).

### **2.1.1 S. Sobrinus**

Es una bacteria gran positiva anaerobia, según Zhang y sus colaboradores (2019) y Seow (2018), *S. sobrinus* pertenece a los grupos bacterianos ácido génicos y acidúricos, conocidos como *mutans streptococci* asociados a la caries. Estudios mediante la técnica de polimerasa reactiva, PCR, muestran su presencia en la placa dental entre un 50% a 80% (Igarashi et al, 2000), produciendo desmineralización por su capacidad de disminuir el pH y resistencia a los medios ácidos, utilizando la sucrosa para producir grandes cantidades de polisacáridos extracelulares que le permiten una fuerte adhesión (Pepperney y Chikindas, 2011). Es el principal indicador de la presencia de raíces dentales cariosas y se le atribuye una gran proporción de la pérdida de los mismos, (Lundgren et al, 1997).

### **2.1.2 Inflación de las Encías en la Raíz del Diente**

La especie bacteriana *S. Sobrinus*, es el principal causal de las rices cariosas en los dientes, por tener una mayor capacidad de resistir a medios ácidos y la habilidad de sintetizar a partir de la sucrosa  $\alpha$ -glucanos insolubles, que le permiten la adherencia en la formación de la placa. Al estar cerca de la encía, el  $\alpha$ -glucano genera una respuesta pro inflamatoria lo que se asocia con el

desarrollo de enfermedad periodontal, (Okamoto et al, 2007). De igual modo Berlutti y colaboradores (2010), demostraron la capacidad de *S. sobrinus*, de invadir las encías y evadir los procesos inmunitarios de las mismas lo que le permite vivir en ausencia de regiones duras en la boca.

### ***2.1.3 Porphyromonas Gingivalis***

Gran negativa y anaerobia bacteria, principal causante de periodontitis. Puede realizar simbiosis con las especies *streptococci*, para colonizar las cavidades orales cercanas a la raíz de los dientes, (Amorin et al, 2008). Se adhiere a los tejidos blandos y adquiere cationes hierro de la hemoglobina de la sangre, mediante procesos anabólicos o catabólicos lo que le permite la co agregación con otras bacterias como *T. denticola* y *A. actinomycetemcomitans*, en la formación del bio film, (Zhu et al, 2019) y (Gerits et al, 2019). Su principal fuente de energía bajo condiciones anaerobias está en la fermentación de aminoácidos que le proporciona los tejidos gingivales, (Khalaf et al, 2017).

### ***2.1.4 Uso de Herramientas Computacionales en el Estudio de las Bio actividades***

Dentro de los procedimientos para contrarrestar la formación de la placa está el uso de antisépticos y/o antibióticos, pero recientemente se ha demostrado que entre un 71,4% y un 74,2% de los patógenos orales son resistente, (Gerits et al, 2019).

Los monofenoles presentan bio actividad frente a los patógenos bucales y su uso a nivel global ha demostrado ser efectivo en este campo, (Chinsembu, 2015), (Almeida et al, 2015) y (Seow, 2018). Estudios realizados muestran el gran potencial que tienen frente a los antibióticos clásicos, actuando como preventivo en la adherencia y formación de bio film, inhibición de enzimas glicolíticas, pH y efecto buffer, (Abachi et al , 2016) y (Wang et al, 2016).

Según Nantasenamat y colaboradores (2010, 2015), las herramientas computacionales permiten el manejo adecuado de la gran información que se da a conocer en torno al tratamiento de enfermedades, para el diseño de nuevos fármacos y la selección de compuestos candidatos en nuevos tratamientos, y en este campo la relación estructura química –actividad, QSAR, por sus siglas en inglés, provee un camino rápido y seguro en la selección de moléculas, y el entendimiento de las propiedades de los compuestos que prevalecen en función de seleccionar posibles candidatos y, predecir su capacidad y efectividad en el tratamiento de enfermedades, a bajo costo.

Mediante QSAR se ha clarificado las características estructurales de eugenol, magnolol, honokiol, y xanthorrhizol en su acción frente a patógenos orales causantes de halitosis, (Maguna et al, 2017) siendo muy atractivo el uso de esta alternativa en concomitancia con los enjuagues y pastas dentales, tal como sucede en el mercado con los compuestos mentol, metilsilicato, eucaliptol y thymol, (Chinsebu, 2015). Cherkasov (2013), afirma que este tipo de modelamiento junto con la actividad experimental guía a futuros estudios para la predicción del comportamiento de nuevos fármacos y tratamientos abriendo la posibilidad del estudio de uso de algunos monofenoles como antibacterianos frente a *S. sobrinus* y *P. Gingivalis*.

### **2.1.5 Modelamiento QSAR**

El modelamiento QSAR comprende esencialmente 5 pasos: i) selección de moléculas con actividades biológicas conocidas y/o estudiadas, ii) cálculo de los descriptores moleculares y selección de los más relevantes, iii) conformación de los subconjuntos de entrenamiento, prueba y comprobación, iv) establecimiento de la relación estructura actividad, y v) validación (Nantasenamat, et al, 2010).

**2.1.5.1 Selección de Moléculas con Actividades Biológicas.** Las propiedades biológicas pueden ser obtenidas experimentalmente o de la literatura, (Nantasenamat, et al, 2010). Para efectos de la presente investigación son recolectados de Shapiro y Guggenheim (1998), y Joshi y colaboradores (2011).

**2.1.5.2 Cálculo de los Descriptores Moleculares y Selección de los más Relevantes.** Existen una gran variedad de paquetes computacionales que permiten el cálculo de los diferentes descriptores moleculares previa optimización de los compuestos. La importancia de estos es la de expresar las propiedades constitucionales, electrónicas, geométricas, hidrofóbicas, lipofílicas, solubilidad, esterificidad, mecano cuánticas y topológicas en forma numérica. Se clasifican según la dimensión de la propiedad, 1D, propiedades globales como la masa; 2D, consideran la presencia o ausencia de átomos, grupos funcionales y fragmentos moleculares y 3D, basados en la conformación espacial de la molécula global (por ejemplo orbital molecular ocupado de mayor energía, HOMO; orbital molecular desocupado de menor energía, LUMO y momento dipolar), y parámetros locales, (Nantasenamat, et al, 2010). Para su selección se realiza generalmente el uso de las técnicas de normalización y/o estandarización.

**2.1.5.3 Conformación de los Subconjuntos de Entrenamiento, Prueba y Comprobación.** A partir de los compuestos seleccionados se crean 2 subconjuntos con el 80% del total para la construcción del modelo, y un 20% del total para la comprobación o validación externa del mismo (Rodríguez, et al, 2019).

**2.1.5.4 Establecimiento de la Relación Estructura Actividad.** El subconjunto de construcción se aborda bajo dos líneas principalmente, la primera aplicando la técnica dejando uno afuera, LOO, por sus siglas en inglés, en donde se construye el modelo mediante regresión lineal múltiple, MLR, dejando a cada iteración una única molécula afuera que sirve para la validación interna y seguimiento del error de predicción, lo que permite los ajustes y posterior adecuación del modelo para ser comprobado; la segunda, conocida como validación cruzada, en el que al 80% se le subdivide en un 60% de entrenamiento y un 20% de prueba, y al igual que en LOO, ese 20 % permite los ajustes para ser sometido a validación externa. Es de aclarar que los subconjuntos internos se forman mediante la selección aleatoria simple y se realizan tantas iteraciones de entrenamiento hasta que el error de predicción interno se sostenga en el mínimo valor, antes de realizar la prueba de validación, (Rodríguez, et al, 2019)

**2.1.5.5 Validación.** La OECD (2004), establece la necesidad de realizar una validación interna con los compuestos que se tuvieron en cuenta para construir el modelo, y una externa con aquellas moléculas que no hicieron parte del establecimiento de la relación estructura actividad. La capacidad predictiva interna se evalúa a través de parámetros estadísticos como lo son: coeficiente de determinación,  $R^2 > 0,6$ ; y el coeficiente de validación cruzada,  $q^2 > 0,5$  y el error cuadrático medio, RMS (Nantasenamat, et al, 2010). En la prueba externa se acepta lo propuesto por Tropsha y colaboradores (2007),  $(R^2 - R_o^2) / R^2 < 0,1$  siendo  $R^2$ , coeficiente de determinación que pasa por el origen observado vs calculado y,  $R_o^2$ , coeficiente de determinación que pasa por el origen calculado vs observado;  $0,85 < k < 1,15$  con  $k$  la pendiente de la regresión lineal que pasa

por el origen de los datos observados vs calculados, y un coeficiente de validación externa,  $Q^2 > 0,6$ .

## 2.2 Marco de Antecedentes

Utilizando combinaciones de palabras claves: Phenol, Porphyromonas, Gingivalys, QSAR, Streptococcus, Sobrinus, oral care y computational chemistry, se encuentran una gran variedad de artículos en los que se relaciona el uso de polifenoles de origen vegetal en la prevención de la formación del bio film en la superficie del diente, así como el uso de QSAR para predecir bioactividades de flavonas y flavonoides frente a *P. Gingivalys*, sin embargo, es escasa o nula la información en lo referente a *S. Sobrinus*, y el uso de la relación cuantitativa actividad de monofenoles frente a *S. sobrinus* y *P. Gingivalys*, en los últimos 20 años siendo muy atractivo esta línea de investigación en pro de aportar al conocimiento.

De los artículos encontrados que referencian el uso de poli fenoles para el control de patógenos orales humanos solo Joshi y colaboradores, (2011), aplican la metodología QSAR para predecir la actividad biológica de fenoles simples frente a los microorganismos mencionados en la presente investigación.

### ***2.2.1 Estudio QSAR Sobre Crecimiento Bacteriano Utilizando Índices Topológicos Tipo Conectividad.***

Joshi y colaboradores, (2011), utilizando algunos fenoles simples que inhiben el crecimiento de *S. sobrinus* y *P. Gingivalis*, mediante QSAR, generaron un modelo que predice la Concentración inhibitoria mínima,  $-\text{LogMIC}(\text{nm})$ , de las moléculas estudio. Los MIC son

obtenidos de la literatura y los cálculos de los índices mediante el software DRAGON, que para el caso son del tipo conectividad extendida. El algoritmo es hallado mediante mínimos cuadrados y soportado estadísticamente con el error estándar menor a 0,25, coeficiente de correlación y coeficiente de correlación ajustable mayor a 0,8, y Fisher.

Para *P. Gingivalis*, el  $-\text{LogMIC}(\text{nm})$ , puede ser estimado con un modelo mono parámetro usando el índice Harary,  $H_{\text{ar}}$ , mostrado a continuación:

$$-\text{LogMIC} = -3,9818 + 0,2478 H_{\text{ar}}$$

$$\text{SE} = 0,209; R^2 = 0,9114; R^2A = 0,9068; F = 195,5103$$

Dónde, SE, error estándar;  $R^2$ , coeficiente de correlación;  $R^2A$ , coeficiente de correlación ajustable y F, Fisher.

Y en el caso *S. sobrinus*, es necesario el uso de multiparámetros así:

$$-\text{LogMIC} = -10,5513 - 0,4185\text{Pol} + 0,7151H_{\text{ar}} + 171,7819\text{Jhet}_z - 170,5441\text{Jhet}_e$$

$$\text{SE} = 0,2487; R^2 = 0,8777; R^2A = 0,8471; F = 28.7080$$

Dónde Pol, número de polaridad;  $\text{Jhet}_z$ , índice Balaban vía número atómico y  $\text{Jhet}_e$ , índice Balaban vía electronegatividad.

### 3. Diseño Metodológico

En esta investigación se construyeron modelos QSAR para predecir la Concentración inhibitoria mínima de monofenoles frente a *S. sobrinus* y *P. gingivalis*. El diseño de este modelo consistió en calcular descriptores moleculares y aplicar un tratamiento estadístico con el fin de encontrar la relación cuantitativa de la estructura química de compuestos con bio actividad frente a *S. sobrinus* y *P. gingivalis*. A continuación se describe la metodología.

### **3.1 Selección de Compuestos con Bioactividad frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis***

Se realizó una búsqueda en la literatura de compuestos con bioactividad frente a *S. sobrinus* y *P. gingivalis*, expresada en términos de la Concentración inhibitoria mínima, MIC. Los compuestos seleccionados corresponden a mono fenoles reportados y estudiados frente a los patógenos estudiados por Shapiro y Guggenheim (1998) y Joshi y colaboradores (2012).

Los compuestos seleccionados se dibujaron sus estructuras químicas moleculares mediante el visualizador Avogadro 1.2, y las optimizaciones se hicieron con el método de los funcionales de la densidad (DFT), utilizando el funcional híbrido de intercambio y correlación B3LYP con la función base cc-pVDZ mediante Gaussian 0.9.

### **3.2 Construcción del Modelo Matemático para la Predicción de la Concentración Inhibitoria Mínima Frente a *Streptococcus Sobrinus* Y *Porphyromonas Gingivalis***

#### **3.2.1 Cálculo de los Descriptores Moleculares**

Se calculó un total de 4885 descriptores moleculares de tipo electrónicos, geométricos, mecanocuánticos, topológicos, electrotopológicos, entre otros, a partir de la geometría molecular optimizada de los compuestos mediante el programa DRAGON 6. A continuación describimos algunos de ellos.

**3.2.1.1 Descriptor de Anillo.** El análisis de las formas fruncidas de un anillo de N miembros se lleva a cabo de forma única calculando el plano medio y los parámetros de fruncido del anillo junto a los ángulos internos de este (Cremer, 1984).

**3.2.1.2 Topológicos o de Conectividad.** Se basa en la información que puede obtenerse de la conectividad mutua entre todos los pares de átomos en una molécula (Carrasco, 2007).

**3.2.1.3 Molecular.** Los índices de conectividad molecular se identifican como componentes de la accesibilidad molecular. Se calcula a partir de los radios de van der Waals y enlaces covalentes de los átomos y ángulo de superposición entre las circunferencias de van der Waals de los átomos enlazados (Estrada, 2002).

También obtuvieron las energías en estado neutro,  $E_N$ ; con carga +1,  $E_{+1}$  y con carga -1,  $E_{-1}$  para obtener el potencial de ionización,  $I$ ; la afinidad electrónica,  $A$ ; carga máxima de almacenamiento,  $N_{\max}$ ; electrofugalidad,  $\Delta E_e$ , y nucleofugalidad,  $\Delta E_n$ , en procura de seleccionar un subconjunto, mediante la gráfica:  $N_{\max}$  Vs  $(\Delta E_e - \Delta E_n)^2$ .

**3.2.1.4 Potencial de Ionización, I.** Es definido como la energía que se requiere para que la molécula libere un electrón y se convierta en un sistema molecular de carga +1 (Vidhya et al, 2020) y (Sandoval y Martínez, 2018). Se calcula mediante la ecuación:

$$I = E_{-1} - E_N \quad (\text{Sanoval y Martinez, 2018})$$

**3.2.1.5 Afinidad Electrónica, A.** Energía para convertir una molécula neutra en una molécula cargada negativamente por la adición de un electrón. (Vidhya et al, 2020) y (Sandoval y Martínez, 2018). Se calcula mediante la ecuación:

$$A = E_N - E_{+1} \quad (\text{Sanoval y Martinez, 2018})$$

**3.2.1.6 Carga Máxima de Almacenamiento, N<sub>máx.</sub>** Carga máxima que un sistema electrófilo puede aceptar (Meenakshi, 2017). Se calcula mediante la ecuación:

$$N_{max} = \frac{-I - A}{I - A} \quad (\text{Meenakshi, 2017})$$

**3.2.1.7 Electrofugalidad, ΔE<sub>e</sub>.** Capacidad electrofílica, (Meenakshi, 2017). Se obtiene mediante la expresión:

$$\Delta E_e = \frac{2I^2}{I - A} \quad (\text{Meenakshi, 2017})$$

**3.2.1.8 Nucleofugalidad, ΔE<sub>n</sub>.** Para Meenakshi (2017), es la capacidad nucleofílica, y se puede hallar mediante la expresión matemática:

$$\Delta E_n = \frac{2A^2}{I - A} \quad (\text{Meenakshi, 2017})$$

### 3.2.2 Construcción y Validación del Modelo Matemático

Se realizó el tratamiento estadístico aplicando el Método de Regresión Multilineal, MLR, para construir el modelo QSAR, a partir de las moléculas seleccionadas mediante las técnica dejando uno afuera y validación cruzada, teniendo en cuenta los parámetros estadísticos de los

coeficientes de determinación,  $R^2$ ; coeficiente de determinación ajustado,  $R^2_{aj}$ ; la significancia estadística,  $p$ ; la distribución de Fisher-Snedecor,  $F$ ; estadísticamente por ANOVA. Como variables experimentales se tomaron los valores de la Concentración inhibitoria mínima en milimoles tomados de Shapiro y Guggenheim (1998), y Joshi y colaboradores (2012).

El modelo de predicción se validó empleando validación interna, CV-LOO, mediante el coeficiente de predicción,  $q^2$ , para evaluar su capacidad predictiva y robustez y por validación externa, CV-Ext, junto con parámetros de tropsha (2007), para la cual se utilizó un conjunto de compuestos de prueba o *test set* con bio actividad reportada en la literatura, no incluidos en el modelo con el objetivo de predecirles su actividad.

#### **4. Resultados y Análisis**

La construcción del modelo QSAR de la Concentración inhibitoria mínima frente a *S. sobrinus* y *P. gingivalis*, se realizó en tres etapas: en la primera optimización y selección de compuestos con bio actividad reportada en la literatura; en la segunda se realizó el tratamiento estadístico para obtener el modelo y la tercera etapa es la validación cruzada interna y externa del modelo.

A continuación se discutirá el trabajo de investigación y los resultados obtenidos.

##### **4.1 Optimización y Selección de Compuestos con Bioactividad Frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis* para la Construcción del Modelo QSAR**

#### **4.1.1 Optimización**

Los monofenoles son derivados del fenol sustituido en las posiciones orto, meta o para, que presentan una importante actividad biológica como antibióticos, anti parasitarios y citotóxicos (Peñarrieta et al, 2014), de solubilidad moderada en agua. Rastija y colaboradores (2013), mediante QSAR identifican que los fenoles simples se inclinan hacia la lipofilidad en el balance hidrofiliidad- lipofiliidad, con una mayor tendencia en comparación a los polifenoles con más de un anillo aromático, y esto les permite atravesar con mayor facilidad la membrana celular por medio de la difusión, siendo uno de los factores más importante en la bioactividad de los compuestos, colocándoles como moléculas de excelentes cualidades para el presente estudio.

Con este panorama se optimizan 100 monofenoles que presentan bioactividad frente a *S. Sobrinus* y *P. Gingivalis*, anexo 1, estudiadas por Shapiro y Guggenheim (1998), con el método DFT, la función de base cc-pVDZ, y el funcional B3LYP, con el paquete computacional Gaussian 09.

#### **4.1.2 Selección**

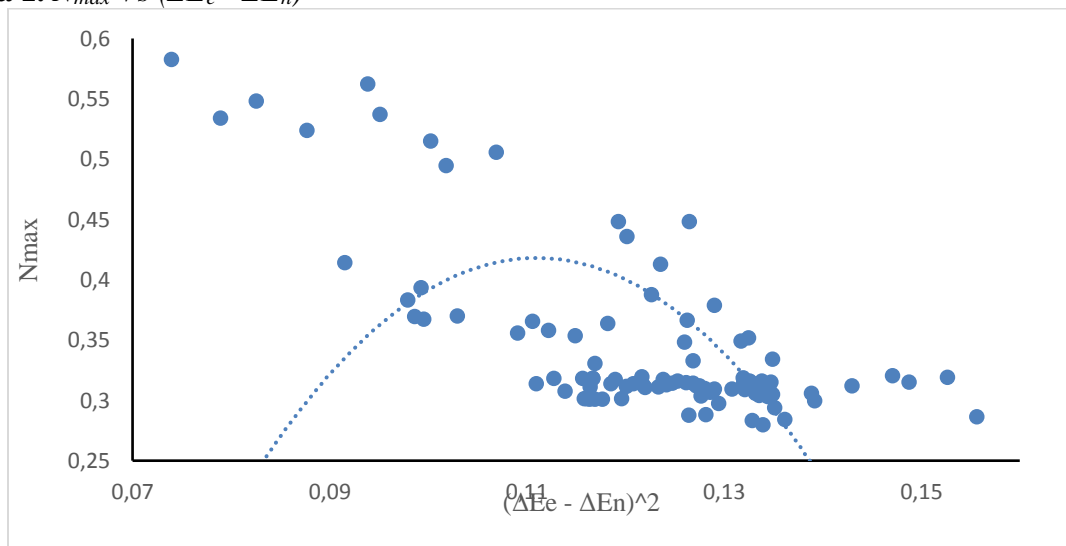
Campodónico y colaboradores (2007), ilustran que una molécula simple R-GS, sistemáticamente se divide en el agua en dos fragmentos  $R^+$  y  $GS^-$ , con el fragmento  $R^+$ , el anillo, y  $GS^-$ , grupo saliente, y aseguran que la electrofugalidad es la medida de la estabilización de la carga positiva en el agua y que está directamente relacionada con la electrofiliidad,  $W$ , que en concepto de Vijayaraj y su grupo de trabajo (2009), es una medida de la reactividad global. De igual manera electrofugalidad y nucleofugalidad son dos conceptos cinéticos que indican la habilidad de átomos o grupos moleculares de abandonar el anillo aromático (Meenakshi, 2017) y

son base en los procesos de activación y desactivación, inmersos con la aceptación o donación de un electrón en los medios acuosos.

Rastija y su grupo de trabajo (2013), resalta la importancia de la electronegatividad atómica para la lipofilia de los compuestos y este punto asociado a lo expuesto en el párrafo anterior son la base para desarrollar una curva de selección, dónde a partir de descriptores electrónicos se tome un grupo de moléculas que interaccionen de la misma manera. Al paquete de 100 compuestos optimizados se les calculan los descriptores electrónicos enunciados en el título 5.2, y se desarrolla la curva de las parejas  $N_{max}$  Vs  $(\Delta E_e - \Delta E_n)^2$ , con línea de tendencia polinómica de orden 2.

Los puntos que están dentro de la curva hacen referencia a 75 compuestos, y tal como se esperaría muestran valores de los descriptores electrónicos cercanos pudiendo decir tal como lo afirman Karelson y colaboradores (1996), que interaccionan molecularmente de la misma manera, lo que nos permite interpretar que la homogeneidad electrónica es suficiente para la selección.

**Figura 1.**  $N_{max}$  Vs  $(\Delta E_e - \Delta E_n)^2$



En el apéndice 2 se aprecian de 1 a 75 los compuestos ubicados al interior de la curva de selección, y de 76 a 100 los que están afuera de la misma. De los datos rechazados para la

construcción del modelo apreciamos que, el 76 es radicalmente diferente a los demás y, de 77 a 100 su electrofugalidad es negativa o cero. 57 a 75, son monofenoles que están sobre la curva o muy cerca de la misma por lo que se descartan seleccionando finalmente los primeros 56.

### 4.1.3 Selección de los Descriptores Moleculares

Una vez calculados los 4885 descriptores moleculares a partir del total de 100 moléculas optimizadas, se generó una matriz de variables dependientes dejando de lado aquellos cuyos valores sean iguales a 0, y se correlaciona con la variable independiente LOG(1/MIC). Los datos se comparan con los generados realizando el mismo procedimiento con los 56 monofenoles seleccionados, y el conjunto de los excluidos. Para la elección se tienen en cuenta los valores de correlación de Pearson de mayor cuantía superior a 0,7.

Para mejor comprensión se muestra como ejemplo a continuación los resultados correlacionales de la matriz generada con los descriptores de anillo frente a *Porphyromonas gingivalis*.

**Tabla 1.** *Correlación de Descriptores Vs LOG(1/MIC) del Total de Monofenoles.*

**Valores de correlación de Pearson, R.**

<b>-LOG(MIC)</b>											-		
<b>mmol</b>	0,56	0,44	0,57	0,60	-0,08	0,22	0,47	-0,23	0,43	0,77	0,57	0,07	
<b>-LOG(MIC)</b>											-		
<b>µg/mL</b>	0,54	0,42	0,54	0,57	-0,08	0,21	0,46	-0,21	0,39	0,74	0,55	0,05	
	nCIC	nCIR	TRS	Rperim	MCD	RCI	NRS	NNRS	ARR	D/Dtr06	nR06	nBnz	

Para las condiciones dadas de  $R > 0,7$ , en la tabla 1, solo el descriptor D/Dtr06 aplica y se prefiere como variable independiente  $\text{LOG}(1/\text{MIC})$  en mili moles.

**Tabla 2.** *Correlación de Descriptores Vs LOG(1/MIC) de Monofenoles Seleccionados.*

<b>Valores de correlación de Pearson, R.</b>													
<b>-LOG(MIC)</b>													
<b>mmol</b>	0,71	0,60	0,71	0,73	0,29	0,42	0,57	-0,45	-0,58	0,82	0,67	0,00	
<b>-LOG(MIC)</b>													
<b>µg/mL</b>	0,71	0,60	0,71	0,73	0,30	0,42	0,57	-0,44	-0,58	0,81	0,67	-0,01	
	nCIC	nCIR	TRS	Rperim	MCD	RCI	NRS	NNRS	ARR	D/Dtr06	nR06	nBnz	

Se observa en la tabla 2 una mejora en las dependencias lineales en comparación a la tabla 1, y este comportamiento al igual que en este ejemplo se repite en los 4885 descriptores calculados.

**Tabla 3.** *Correlación de Descriptores Vs LOG(1/MIC) de Monofenoles Excluidos.*

<b>Valores de correlación de Pearson, R.</b>													
<b>-LOG(MIC)</b>													
<b>mmol</b>	0,55	0,46	0,00	0,22	0,43	-0,21	-0,44	0,54	0,54	0,53	0,26	0,64	
<b>-LOG(MIC)</b>													
<b>µg/mL</b>	0,49	0,43	-0,06	0,22	0,36	-0,20	-0,42	0,47	0,48	0,47	0,12	0,58	
	nCIC	nCIR	TRS	Rperim	MCD	RCI	NRS	NNRS	ARR	D/Dtr06	nR06	nBnz	

Es de resaltar la pobre correlación, y al igual que en las tablas 1 y 2 el valor más alto esta para  $\text{LOG}(1/\text{MIC})$  en mili moles y D/Dtr06 sin que  $R$  supere los 0,7, siendo muy prometedor la exclusión de las moléculas.

Como se menciona en el título 6.1.2, y se puede ver en el anexo 2, los primeros 56 monofenoles guardan homogeneidad en sus descriptores electrónicos, lo que les capacita para interactuar con otras moléculas de la misma forma, mejorando el valor de correlación en comparación con las 100 moléculas agrupadas y los monofenoles excluidos.

El procedimiento permite al final la selección de 14 descriptores con los valores más altos de R, como variables dependientes frente a LOG(1/MIC) en mmol como variable independiente, las cuales se muestran a continuación.

**Tabla 4.** Descriptores Seleccionados y su Valor de R.

Descriptor	Tipo	<i>P. GINGIVALIS</i>	<i>S. SOBRINUS</i>
		R	R
G2	Geométrico	0,9	0,9
G1	Geométrico	0,9	0,89
AMR	Molecular	0,92	0,93
ALOGP	Molecular	0,9	0,92
Vx	Molecular	0,92	0,93
VvdwMG	Molecular	0,92	0,93
VvdwZAZ	Molecular	0,92	0,93
X0v	Conectividad	0,91	0,91
X1Mad	Conectividad	0,92	0,93
X1MulPer	Conectividad	0,93	0,94
X1Per	Conectividad	0,93	0,94
D/Dtr06	Anillo	0,82	0,83
Psi_i_0	Pseudoconectividad	0,93	0,93
P_VSA_v_1	P_VSA	0,9	0,9

La tabla 4 no indica que no existieran otro conjunto que cumplieran con el parámetro R > 0,7, sino que de cada paquete se tomaron solo los valores más altos, es decir, para *P. Gingivalis* en geométricos 7 cumplen con el requisito pero G1 y G2 son los valores cumbre; en molecular 9 superan el 0,7 pero los mostrados en la tabla 4 son los más altos; conectividad, de un conjunto de 25 se seleccionaron los señalados; en anillo se toma D/Dtr06 de 4 candidatos, etc., y a su vez, estos

paquetes de descriptores son los que contienen los valores  $R$  más altos. De igual forma para *S. Sobrinus*, el comportamiento es similar.

#### **4.2 Construcción del Modelo Relación Cuantitativa Estructura Actividad para la Predicción de Concentración Inhibitoria Mínima de Monofenoles Frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis*.**

Básicamente consistió en el tratamiento estadístico de regresión lineal múltiple MLR, una vez seleccionando los descriptores con correlaciones  $R > 0,8$  respecto a la variable dependiente MIC. La validación del modelo es del tipo cruzada interna y externa.

El modelo fue evaluado por los parámetros coeficiente de determinación  $R^2$ , coeficiente de determinación ajustado,  $R^2_{aj}$ , el F de Fisher, prueba F para varianzas de dos muestras, prueba z para medias de dos muestras y el error de estimación, RMS. El coeficiente de determinación representa el grado de ganancia que podemos obtener al predecir una variable basándonos en el conocimiento que tenemos de otra u otras variables y se considera válido si el coeficiente es mayor a 0,50; cuanto más cercana a uno, mejor será el ajuste de los datos al modelo, en otras palabras expresa la proporción de varianza de la variable dependiente que está explicada por la variable independiente; El F de Fischer determina la significancia estadística, correlacionando la varianza explicada  $R^2$ , y el número de grados de libertad con la varianza no explicada,  $1 - R^2$ , y el número de variables del modelo; cuanto más alto sea el valor de la varianza explicada el valor F será mayor, mientras que la existencia de variables que tienen bajo aporte a la varianza explicada, tenderá a disminuir el valor F; las pruebas de hipótesis nulas determina si existen diferencias estadísticamente entre los datos experimentales y los calculados y por último RMS, que es la

medida de cuanto se aleja los valores predichos del modelo, cuya la tendencia a cero supone mayor calidad en la predicción.

**4.2.1 Modelo QSAR para Predecir la MIC de Monofenoles de Porphyromonas Gingivalis**

A continuación se muestran los modelos que obtuvieron los mejores parámetros:

**Tabla 5. Modelos QSAR para Porphyromonas Gingivalis.**

<b>Modelo 1</b>				
LOG(1/MIC) = -2,35 + 0,092 X1MAD + 0,49X1MulPer				
<b>Estadística</b>				
R <sup>2</sup>	R <sup>2</sup> aj	RMS	F	Estadístico F
0,87	0,87	0,17	238	3,3E-17
<b>Modelo 2</b>				
LOG(1/MIC)= -2,3 + 0,02ALOGP + 0,7X1PER				
<b>Estadística</b>				
R <sup>2</sup>	R <sup>2</sup> aj	RMS	F	Estadístico F
0,87	0,87	0,5	244	2,3E-17

En la tabla 5 se muestran dos modelos de regresión lineal. Los descriptores vinculados al modelo 1 son de tipo conectividad y el modelo 2, ALOGP molecular, y X1PER conectividad, juntos muestran una buena correlación R<sup>2</sup> aj = 0,87. En cuanto a los errores de predicción el primero tiende a ser más confiable por su valor menor al segundo modelo, sin embargo los dos presentan buenos parámetros estadísticos.

Con el fin de asegurar la confiabilidad de los modelos se plantea la hipótesis nula: “Ho =Los valores experimental y calculado no presentan diferencias estadísticamente significativas” y la hipótesis alternativa, “H1 = Existen diferencias estadísticamente significativas entre los valores experimentales y calculados”. A continuación podemos ver los resultados de aplicar la prueba F

para varianzas de dos muestras y la prueba z para medias de dos muestras, con un nivel de significancia del 0,05.

**Tabla 6.** Hipótesis Nula y Alternativa, Prueba F y Z. P. Gingivalis.

<i>Ho = Los valores experimental y calculado no presentan diferencias estadísticamente significativas</i>			
<i>H1 = Existen diferencias estadísticamente significativas entre los valores experimentales y calculados</i>			
<b>Modelo 1 <i>Porphyromonas Gingivalis</i></b>			
$LOG(1/MIC) = -2,35 + 0,092 X1MAD + 0,49X1MulPer$			
Varianzas		Medias	
F	1,14	z	-0,24
P(F<=f)	0,35	P(Z<=z)	0,40
V. Crítico	1,74	V. Crítico	1,96
<b>Modelo 2 <i>Porphyromonas Gingivalis</i></b>			
$LOG(1/MIC) = -2,3 + 0,02ALOGP + 0,7X1PER$			
Varianzas		Medias	
F	1,06	z	-1,12
P(F<=f)	0,43	P(Z<=z)	0,13
V. Crítico	1,74	V. Crítico	1,96

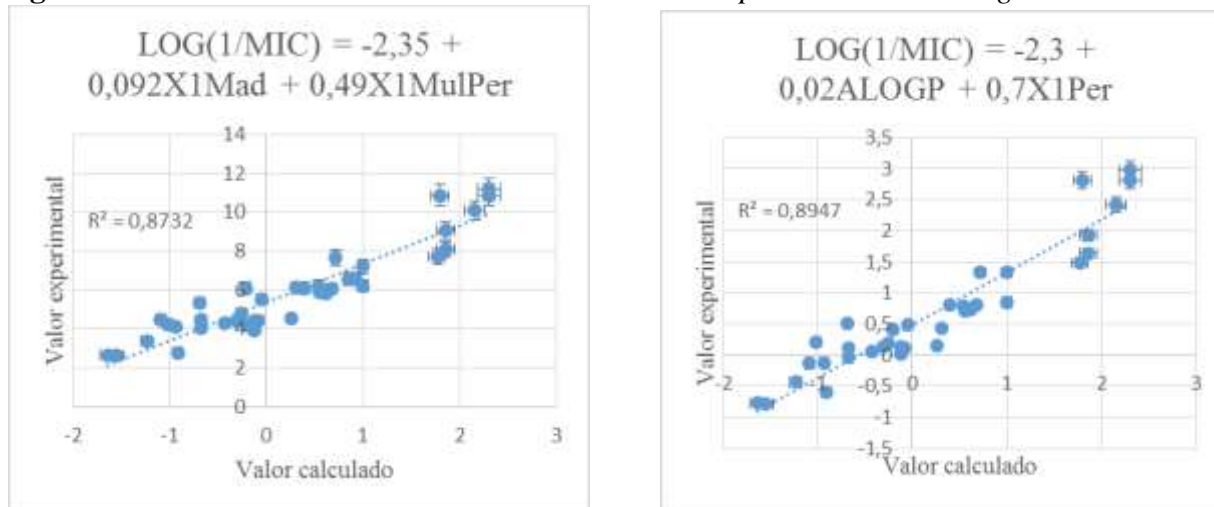
El conjunto de datos se sometieron a las pruebas pertinentes para definir comportamiento normal previo a las pruebas de las hipótesis y de esta manera garantizar su valides. Para que pueda ser aceptada Ho, los valores de F y |Z| han de ser menores que el valor crítico y el nivel de significancia  $p > 0,05$ .

La comparación de los resultados evidencian que ese mayor error de predicción para la propuesta 2 se focaliza principalmente en la comparación de las medias, y este no es una condición suficiente pero si necesaria en la toma de la decisión en la aprobación de la ecuación. Tanto para el modelo 1 como el 2, se acepta la hipótesis nula, es decir, no existen diferencia estadísticamente significativas entre los valores experimentales y calculados.

Para una mejor comprensión en la figura 3 podemos apreciar las gráficas con un ajuste aceptable de los valores calculados versus los valores experimentales para los modelos, de las moléculas utilizadas junto a su porcentaje de error individual, apreciando de manera gráfica la

idea de párrafo anterior, en el que un valor de 0,5 es bueno si se cumplen con las demás condiciones y parámetros.

**Figura 2.** Correlación Valores Calculados Vs Valores Experimentales. *P. Gingivalis*



**4.2.2 Modelo QSAR para Predecir la MIC de Monofenoles de S. Sobrinus.**

A continuación se muestran los modelos que obtuvieron los mejores parámetros:

**Tabla 7.** Modelo QSAR para *S. Sobrinus*.

Modelo 1 <i>S. Sobrinus</i>				
LOG(1/MIC) = -2,23 + 0,0076D/DTr06 + 0,55ALOGP				
Estadística				
R <sup>2</sup>	R <sup>2</sup> aj	RMS	F	Estadístico F
0,89	0,88	0,88	306	3,5E-19
Modelo 2 <i>S. Sobrinus</i>				
LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0				
Estadística				
R <sup>2</sup>	R <sup>2</sup> aj	RMS	F	Estadístico F
0,88	0,88	1,58	260	4,5E-18

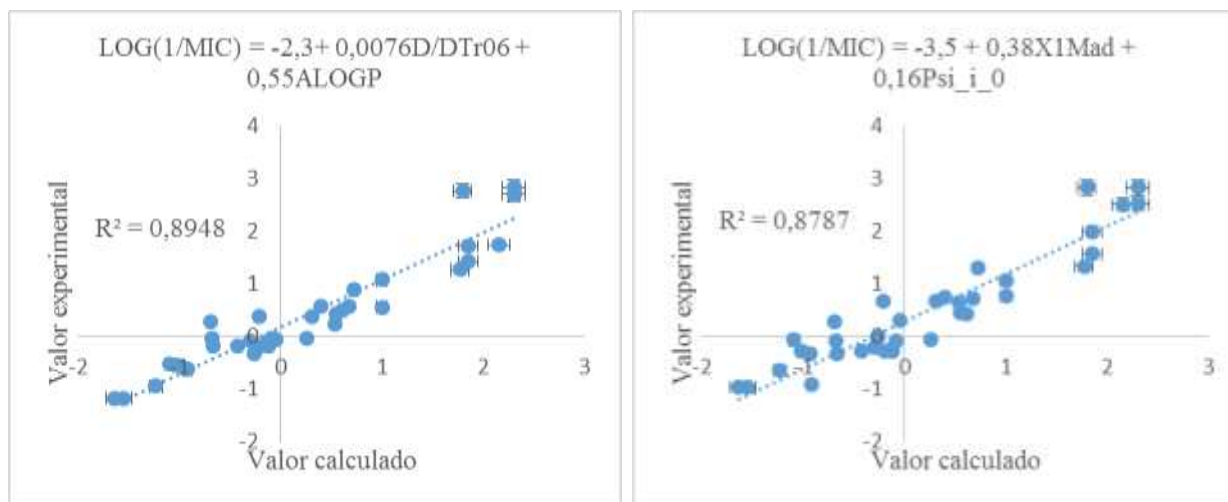
Los errores de predicción sugieren al igual que el tratamiento efectuado para *P. Gingivalis*, realizar la prueba de la hipótesis nula para tomar la decisión de aceptar o rechazar.

**Tabla 8.** Hipótesis Nula y Alternativa, Prueba F y Z S. Sobrinus.

<i>Ho = Los valores experimental y calculado no presentan diferencias estadísticamente significativas</i>			
<i>H1 = Existen diferencias estadísticamente significativas entre los valores experimentales y calculados</i>			
<b>Modelo 1 S. Sobrinus</b>			
$LOG(1/MIC) = -2,23 + 0,0076D/DTr06 + 0,55ALOGP$			
Varianzas		Medias	
F	1,12	z	-1,55
P(F<=f)	0,37	P(Z<=z)	0,06
V. Crítico	1,73	V. Crítico	1,96
<b>Modelo 2 S. Sobrinus</b>			
$LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0$			
Varianzas		Medias	
F	1,07	z	-1,05
P(F<=f)	0,42	P(Z<=z)	0,14
V. Crítico	1,72	V. Crítico	1,95

De la tabla podemos concluir que no existen diferencias estadísticamente significativas entre los valores calculados y experimentales de las ecuaciones predictoras, y al igual que en *P. Gingivalis*, como las propuestas cumplen con las condiciones dadas, el RMS que es una medida de la dispersión es aceptable.

**Figura 3.** Correlación Valores Calculados Vs Valores Experimentales. S. Sobrinus



El comportamiento de los datos observado en la figura 3 para las dos ecuaciones de predicción, muestra una buena tendencia hacia la correlación y esto sumado a que se soporta en pruebas estadísticas, infieren una buena calidad.

### 4.3 Validación de los Modelos Relación Cuantitativa Estructura Actividad para la Predicción de Concentración Inhibitoria Mínima de Monofenoles Frente a *Streptococcus Sobrinus* y *Porphyromonas Gingivalis*.

La OCDE (2004), establece que la validación de los modelos deben contener un procedimiento interno para los datos obtenidos en la construcción del modelo llamado validación interna y uno con datos que no se tuvieron en cuenta o que son nuevos y se desean ingresar para ser testeados, a lo que llama validación externa.

#### 4.3.1 Validación Interna

Las ecuaciones predictoras propuestas luego de asegurar su confiabilidad tal como se indicó en el desarrollo del título 6.2, se las calcula el coeficiente de predicción interna  $q^2_{CV-LOO}$  indicativo de la robustez, obteniendo los siguientes resultados.

**Tabla 9.** Validación Interna.

Patógeno	Modelo	$q^2_{CV-LOO}$
<i>P. Gingivalis</i>	$LOG(1/MIC) = -2,35 + 0,092 X1Mad + 0,49X1MulPer$	0,83
<i>P. Gingivalis</i>	$LOG(1/MIC) = -2,3 + 0,02ALOGP + 0,7X1PER$	0,51
<i>S. Sobrinus</i>	$LOG(1/MIC) = -2,23 + 0,0076D/DTr06 + 0,55ALOGP$	0,94
<i>S. Sobrinus</i>	$LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0$	0,88

Una buena robustez y factor de predictibilidad se presenta cuando  $q^2_{CV-LOO} > 0,7$ , sin embargo, esto no necesariamente es así porque solo representa la capacidad predictiva frente a los valores utilizados para la construcción del modelo, (Gramatica, et al, 2007). Para Tropsha y colaboradores (2007), la validación interna es aceptable cuando su valor supera el 0,5 pero que no es suficiente para determinar su poder en la predicción. Por su parte Roy y Kar (2012), exponen la necesidad de un nuevo indicador en función de que en algunos casos la validación cruzada presenta valores altos en contraposición de la baja correlación y, los tres autores mencionados en el presente párrafo concuerdan que un  $R^2 > 0,7$  y un RMS bajo es indispensable para determinar si el valor de  $q^2_{CV-LOO}$  es aceptable o no, siempre y cuando supere el 0,5.

En función a lo discutido con anterioridad el 0,51 de la segunda fila de la tabla 9 se toma como aceptable en relación a que el  $R^2$  es superior a 0,8 y su RMS calculado es de 0,5. De igual manera las  $q^2_{CV-LOO}$  de las demás filas cumplen con las condiciones entendiéndose que los cuatro modelos tienen la condiciones internas de robustez.

**4.3.1.1 Validación datos límite de la curva de selección.** Es importante el identificar el comportamiento predictivo con los compuestos que se dejaron de lado por estar sobre o muy cerca al límite de selección. En la siguiente tabla se pueden apreciar los resultados:

**Tabla 10.** *Residuales Compuestos Limite. P. Gingivalis.*

Experimental	Modelo 1 <i>P. Gingivalis</i>	Modelo 2 <i>P. Gingivalis</i>	Residual Modelo 1	Residual Modelo 2
1,289	-0,829	-0,816	2,118	2,105
0,342	-0,211	-0,165	0,554	0,507
0,124	0,068	0,133	0,055	-0,009
0,364	-1,001	-0,991	1,365	1,355
0,342	-0,183	-0,117	0,525	0,459
0,292	-0,218	-0,168	0,510	0,460

<b>Experimental</b>	<b>Modelo 1 <i>P. Gingivalis</i></b>	<b>Modelo 2 <i>P. Gingivalis</i></b>	<b>Residual Modelo 1</b>	<b>Residual Modelo 2</b>
1,251	-0,960	-1,002	2,211	2,253
0,435	-0,317	-0,267	0,751	0,702
0,612	-0,497	-0,473	1,108	1,085
0,691	-0,503	-0,477	1,194	1,168
1,289	-0,830	-0,816	2,119	2,105
1,178	-0,823	-0,813	2,001	1,991
1,377	-0,953	-0,996	2,330	2,372
0,342	-0,310	-0,263	0,653	0,606
1,274	-0,746	-0,777	2,020	2,051
0,778	-0,615	-0,585	1,393	1,363
0,702	-0,992	-0,985	1,694	1,686
1,251	-0,960	-1,003	2,211	2,254
1,312	-0,876	-0,949	2,187	2,261

Los residuales tabla 10, indican la poca capacidad predictiva de los modelos de *P. Gingivalis*., frente a los compuestos que se ubican muy cerca o sobre la línea de selección. Los coeficientes calculados para las dos ecuaciones con valores de  $R^2 < 0,55$ ;  $Q^2_{CV\ Externa} < -1$ , y RMS mayor de 4 corroboran esta afirmación.

**Tabla 11.** *Residuales Compuestos Limite. S. Sobrinus.*

<b>Experimental</b>	<b>Modelo 1 <i>S. Sobrinus</i></b>	<b>Modelo 2 <i>S. Sobrinus</i></b>	<b>Residual Modelo 1</b>	<b>Residual Modelo 2</b>
-1,333	-0,935	-1,209	-0,399	-0,124
-0,729	-0,368	-0,545	-0,361	-0,184
-0,493	-0,082	-0,233	-0,411	-0,260
-0,517	-0,791	-1,443	0,274	0,926
-0,124	-0,337	-0,379	0,213	0,255
-0,711	-0,368	-0,553	-0,343	-0,158
-1,251	-1,089	-1,311	-0,162	0,059
-0,785	-0,482	-0,611	-0,303	-0,173
-0,867	-0,652	-0,866	-0,215	-0,001
-0,867	-0,652	-0,874	-0,215	0,007
-1,170	-0,935	-1,209	-0,235	0,039
-1,333	-0,935	-1,201	-0,399	-0,132
-0,590	-0,482	-0,603	-0,108	0,013

<b>Experimental</b>	<b>Modelo 1 <i>S. Sobrinus</i></b>	<b>Modelo 2 <i>S. Sobrinus</i></b>	<b>Residual Modelo 1</b>	<b>Residual Modelo 2</b>
-1,428	-1,180	-0,979	-0,249	-0,450
-1,281	-0,638	-0,919	-0,643	-0,362
-1,063	-0,791	-1,421	-0,272	0,358
-1,251	-1,089	-1,311	-0,162	0,059
-1,488	-0,947	-1,117	-0,541	-0,371

Los residuales de la tabla 11 indican una buena correlación de los datos experimentales y calculados. En principio se puede decir que al ser más robustos los modelos de *S. Sobrinus*, su capacidad predictiva es mayor, pero solo podemos estar seguros hasta la verificación de los parámetros.

**Tabla 12.** *Parámetros Compuestos Límite. S. Sobrinus.*

<b>LOG(1/MIC) = -2,23 +0,0076D/DTr06 + 0,55ALOGP</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	RMS
0,95	0,19	0,72	0,68	1,14
<b>LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	RMS
0,99	0,14	0,91	0,45	0,17

No existe una correlación manifiesta entre el bajo valor de  $R^2$  y el elevado de  $Q^2_{CV\ Externa}$  de la tabla 12. En la literatura ampliamente se ha discutido este fenómeno por diferentes grupos de investigación Gramatica y colaboradores (2007), Tropsha y colaboradores (2007) y Roy (2010, 2012), y para ello Roy (2010), propone el uso del coeficiente de determinación métrico  $rm$ , que no considera diferencias individuales entre los valores predichos y la media del entrenamiento y así evita la sobre estimación, permitiendo aclarar en casos como este, el por qué a pesar de que  $Q^2_{CV\ Externa}$  tiene un valor elevado, los indicadores de Tropsha (2007), no se cumplen en la segunda columna, y el valor de  $R^2$  solo es aceptable en la ecuación 1.

Para que  $r_m$  sea válido su valor calculado ha de ser  $r_m > 0,5$ , así los valores para la ecuación 1 y 2 son 0,45 y 0,19 respectivamente. Esto indica que hay una sobre estimación y explica al igual la disparidad de los parámetros establecidos para la validación externa en *S. Sobrinus*.

En resumen la falla de los 4 modelos para la predicción del Log (1/MIC) de los compuestos límite, permite afirmar que la zona de más alta eficiencia en la predicción se encuentra en la región alejada del límite de la curva y que es factible utilizarlas en la determinación de las bioactividades para compuestos que cumplan con los parámetros electrónicos.

**4.3.1.2 Validación datos periféricos a la curva de selección.** Con respecto a los valores periféricos a la curva  $N_{max} V_s (\Delta E_e - \Delta E_n)^2$ , se calculan los datos mediante los modelos obteniendo los siguientes resultados:

**Tabla 13.** *Parámetros Validación Datos Periféricos a la Curva.*

	$Q^2_{CV\ Externa}$	$R^2$	RMS
Modelo 1 P Gingivalis	-13	0,41	6,19
Modelo 2 P. Gingivalis	-1,19	0,4	2,42
Modelo 1 S. Sobrinus	0,32	0,17	2,39
Modelo2 S. Sobrinus	0,95	0,42	0,65

Según los resultados obtenidos los modelos no aplican para los compuestos ubicados fuera de la curva de selección. En el caso de la última fila de la tabla 13 se realiza el cálculo de  $r_m$  para confirmar lo antes afirmado obteniendo un valor de  $r_m = 0,082$ , menor al límite de 0,5.

### 4.3.2 Validación Externa

Para la validación se utilizaron 20 moléculas tomadas de Shapiro y Guggenheim (1998), y Joshi y colaboradores (2012), no incluidas en la construcción de los modelos las cuales se les somete a igual tratamiento de selección que el conjunto interno.

**Tabla 14.** Selección de Moléculas Validación Externa.

Molécula	A	I	W	$\Delta E_e$	$\Delta E_n$	$N_{max}$
o-metilfenol	-0,299	0,074	-0,034	0,040	-0,333	0,302
m-metilfenol	-0,301	0,069	-0,036	0,033	-0,337	0,313
p-metilfenol	-0,294	0,595	-0,025	0,569	-0,319	-0,169
o-propilfenol	-0,295	0,073	-0,034	0,039	-0,329	0,302
m-propilfenol	-0,298	0,068	-0,036	0,032	-0,334	0,314
p-propilfenol	-0,291	0,069	-0,034	0,034	-0,325	0,309
o-isobutilfenol	-0,294	0,070	-0,035	0,035	-0,329	0,309
o-secbutilfenol	-0,295	0,072	-0,034	0,039	-0,329	0,303
p-butilfenol	-0,290	0,069	-0,034	0,035	-0,324	0,308
o-n-amilfenol	-0,294	0,073	-0,033	0,039	-0,327	0,302
p-n-amilfenol	-0,289	0,068	-0,034	0,033	-0,323	0,310
p-tetrapentilfenol	-0,305	0,066	-0,038	0,028	-0,344	0,322
p-n-heptilfenol	-0,289	0,069	-0,034	0,035	-0,323	0,308
o-terbutilfenol	0,113	0,072	-0,034	0,038	-0,330	0,305
p-ter-octilfenol	-0,286	0,064	-0,035	0,029	-0,321	0,316
m-ciclohexilfenol	-0,294	0,069	-0,035	0,034	-0,329	0,311
m-ciclohexilmetilfenol	-0,295	0,066	-0,036	0,029	-0,331	0,318
p-ciclohexilmetilfenol	-0,287	0,064	-0,035	0,029	-0,323	0,317
(+) ferruginol	-0,271	0,062	-0,033	0,029	-0,304	0,314
p-n-butoxifenol	-0,273	0,068	-0,031	0,037	-0,304	0,301

19 de las 20 moléculas cumplen con las condiciones necesarias para estar dentro de la curva de selección. El p-metilfenol tiene un valor de  $N_{max}$ , atípico y  $(\Delta E_e - \Delta E_n)^2$ , equivale a 0,79 dejándole en la periferia de la curva.

Con los monofenoles seleccionados se validan los modelos mediante lo propuesto por Tropsha y colaboradores (2007),  $(R^2 - Ro^2)/R^2 < 0,1$  siendo  $R^2 > 0,6$ , coeficiente de determinación de los datos observados vs calculados y,  $Ro^2$ , coeficiente de determinación que pasa por el origen de las observaciones calculadas vs observadas;  $0,85 < k < 1,15$ , con  $k$  la pendiente de la regresión lineal que pasa por el origen de los valores observados vs calculados, y un coeficiente de validación externa,  $Q^2_{CV Externa}$  mayor a 0,6.

Los resultados obtenidos se pueden apreciar en la tabla 15.

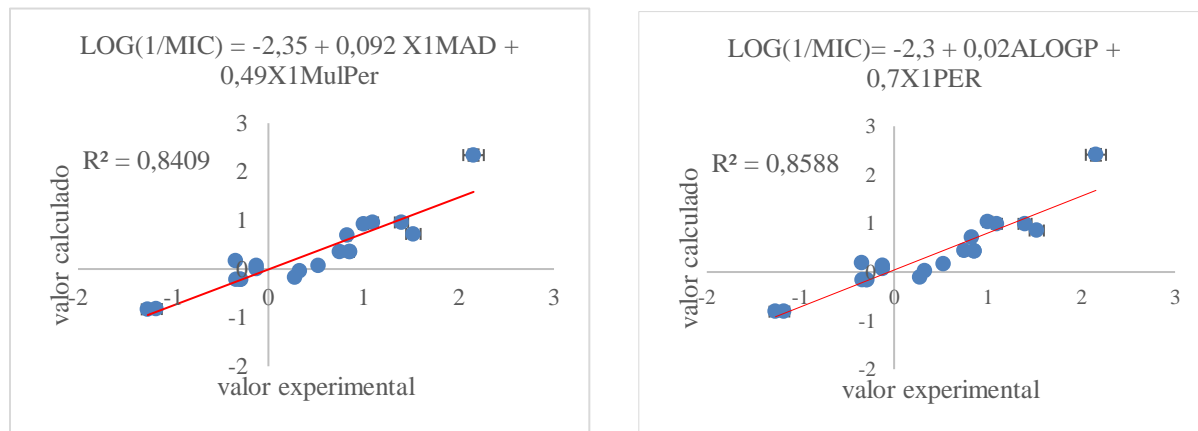
**Tabla 15.** *Resultados Validación Externa.*

<b>LOG(1/MIC) = -2,35 + 0,092 X1Mad + 0,49X1MulPer</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	$RMS$
0,79	-0,007	1,156	0,84	0,37
<b>LOG(1/MIC)= -2,3 + 0,02ALOGP + 0,7X1PER</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	$RMS$
0,96	-0,016	1,12	0,86	0,32
<b>LOG(1/MIC) = -2,23 +0,0076D/DTr06 + 0,55ALOGP</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	$RMS$
0,93	0,009	1,28	0,94	0,27
<b>LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0</b>				
$Q^2_{CV\ Externa}$	$(R^2-Ro^2)/R^2 < 0,1$	$0,85 \leq K \leq 1,15$	$R^2$	$RMS$
0,82	0,04	0,97	0,89	0,28

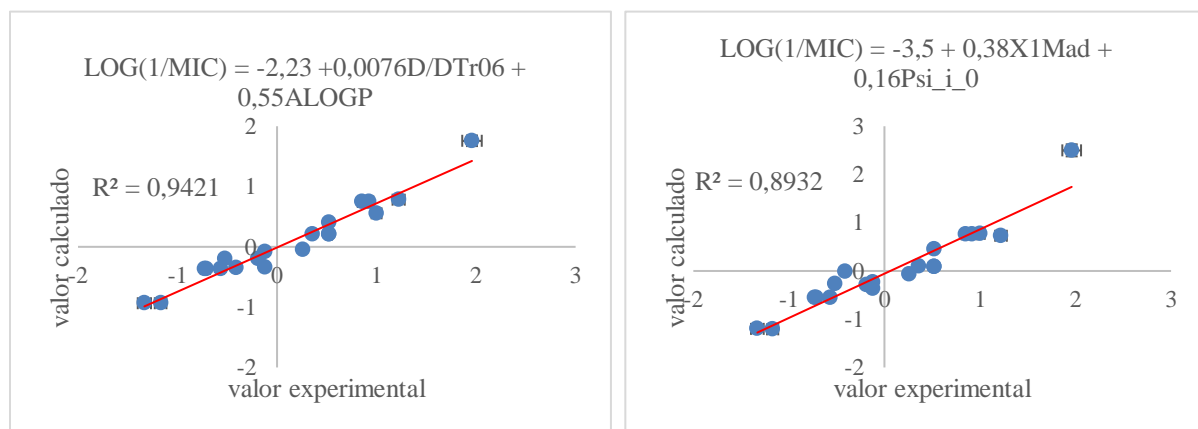
En general las ecuaciones se validan de manera externa satisfactoriamente. Sus RMS, que son una medida del error se mantienen por debajo de 0,37 siendo un buen indicador de la precisión, sus  $Q^2_{CV\ Externa}$  son mayores de 0,78 enmarcándoles en buen carácter predictivo. En cuanto a lo que propone Tropsha (2007), solo dos presentan valores de  $k$  por encima de 1,15, sin embargo, no son una gran diferencia como para no tenerlos en cuenta en relación a los buenos valores en los demás parámetros estadísticos.

En las figuras 4 y 5 que se muestran a continuación se aprecia el buen ajuste de los modelos frente a compuestos externos.

**Figura 4.** *Correlación Validación Externa P Gingivalis.*



**Figura 5.** *Correlación Validación Externa S. Sobrinus.*



Un factor importante a tener en cuenta son los residuales. El RMS nos permite ver de manera general la dispersión de los datos, y el error absoluto nos clarifica sobre la capacidad de predicción.

En la tabla 16 la validación externa para la concentración inhibitoria mínima de monofenoles para *P. Gingivalis* el modelo 1, presenta 8 moléculas por debajo del 50% error, mientras que el modelo 2 tiene 12 compuestos siendo este el de mayor capacidad a la hora del cálculo de bioactividades.

**Tabla 16.** Error Modelos *P. Gingivalis*.

	Modelo 1 LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0			Modelo 2 LOG(1/MIC)= -2,3 + 0,02ALOGP + 0,7X1PER		
	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error
o-metilfenol	-1,18	-1,20	0,02	-1,18	-0,81	0,31
m-metilfenol	-1,27	-1,21	0,05	-1,27	-0,82	0,36
o-propilfenol	-0,34	-0,55	0,59	-0,34	-0,17	0,52
m-proplifenol	-0,29	-0,55	0,89	-0,29	-0,17	0,42
p-propilfenol	-0,30	-0,55	0,86	-0,30	-0,17	0,43
o-isobutilfenol	0,33	-0,29	1,88	0,33	0,03	0,92
o-secbutilfenol	-0,12	-0,27	1,16	-0,12	0,06	1,51
p-butilfenol	-0,12	-0,23	0,88	-0,12	0,13	2,06
o-n-amilfenol	0,75	0,10	0,87	0,75	0,44	0,41
p-n-amilfenol	0,85	0,09	0,90	0,85	0,43	0,49
p-tetrapentilfenol	0,52	-0,07	1,13	0,52	0,16	0,69
p-n-heptilfenol	1,00	0,73	0,27	1,00	1,04	0,04
o-terbutilfenol	0,28	-0,37	2,35	0,28	-0,11	1,41
p-ter-octilfenol	1,52	0,77	0,49	1,52	0,85	0,44
m-cilcohexilfenol	0,82	0,45	0,46	0,82	0,71	0,14
m-ciclohexilmetilfenol	1,10	0,76	0,31	1,10	0,99	0,10
p-ciclohexilmetilfenol	1,40	0,76	0,46	1,40	0,99	0,29
Positivo ferruginol	2,16	2,50	0,16	2,16	2,41	0,12
p-n-butoxifenol	-0,35	-0,01	0,97	-0,35	0,18	1,53

El RMS establece en modelo 2 que alrededor del 32 % de los casos no es satisfactorio el cálculo de la concentración mínima inhibitoria, es decir, que la bioactividad de 6 compuestos presenten diferencias considerables en función de la media de error. La ecuación planteada muestra que 7 moléculas no se ajustan al modelo tal como lo establece el error cuadrático medio de predicción, hecho que confirma que el segundo modelo es el más robusto y confiable en la predicción de la MIC.

De lo anterior junto a los parámetros calculados se puede afirmar que los modelos 1 y 2 para *P. Gingivalis*, estadísticamente son consistentes, pero el número 2 tiene un mejor carácter predictivo.

$\text{LOG}(1/\text{MIC}) = -2,3 + 0,02\text{ALOGP} + 0,7\text{X1PER}$  de *P. Gingivalis*, muestra la dependencia positiva con la hidrofiliidad, ALOGP, factor preponderante en la bioactividad como lo expone Rastija y colaboradores (2013), y con la contribución del monofenol a las interacciones bimoleculares que surgen de encuentros de enlaces entre dos moléculas X1PER, (Estrada, 2002), algo esperado en cuanto a que la selección de los compuestos parte del cálculo de la electrofugalidad y nucleofugalidad y su relación con la Carga máxima de almacenamiento.

Para *S. Sobrinus*, tabla 17, el modelo 1 contienen 15 predicciones por debajo del 50% de error y el modelo 2 un total de 12, siendo el primero el más consistente y el único que se ajusta al valor calculado del RMS.

La ecuación  $\text{LOG}(1/\text{MIC}) = -2,23 + 0,0076\text{D}/\text{DTr}06 + 0,55\text{ALOGP}$ , modelo 1 para *S. Sobrinus*, se selecciona como la de mejor carácter predictivo y en virtud de su calidad estadística y cuenta con una relación positiva con el descriptor de anillo D/DTr06, que indica las deformaciones del anillo en virtud de las elevaciones sobre el plano que sufre cada átomo (Sega, et all., 2010) y, una gran dependencia con la capacidad la relación hidrofiliidad – lipofiliidad ALOGP, calculada por las contribuciones de todos los átomos que conforman la molécula, (Guziałowska, 2017).

**Tabla 17.** Error Modelos *S. Sobrinus*.

	Modelo 1			Modelo 2		
	$\text{LOG}(1/\text{MIC}) = -2,23 + 0,0076\text{D}/\text{DTr}06 + 0,55\text{ALOGP}$			$\text{LOG}(1/\text{MIC}) = -3,5 + 0,38\text{X1Mad} + 0,16\text{Psi}_i_0$		
	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error
o-metilfenol	-1,33	-0,94	0,30	-1,33	-1,20	0,10
m-metilfenol	-1,17	-0,94	0,20	-1,17	-1,21	0,03

	Modelo 1			Modelo 2		
	LOG(1/MIC) = -2,23 + 0,0076D/DTr06 + 0,55ALOGP			LOG(1/MIC) = -3,5 + 0,38X1Mad + 0,16Psi_i_0		
	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error	Experimental -Log(mmol)	Calculado -Log(mmol)	% Error
p-propilfenol	-0,57	-0,37	0,35	-0,57	-0,55	0,02
o-propilfenol	-0,73	-0,37	0,50	-0,73	-0,55	0,25
m-propilfenol	-0,71	-0,37	0,48	-0,71	-0,55	0,23
o-isobutilfenol	-0,19	-0,20	0,03	-0,19	-0,29	0,53
o-secbutilfenol	-0,52	-0,20	0,62	-0,52	-0,27	0,48
p-butilfenol	-0,12	-0,08	0,34	-0,12	-0,23	0,92
o-n-amilfenol	0,36	0,21	0,42	0,36	0,10	0,72
p-n-amilfenol	0,52	0,21	0,61	0,52	0,09	0,83
p-tetrapentilfenol	0,26	-0,05	1,19	0,26	-0,07	1,27
p-n-heptilfenol	1,22	0,78	0,36	1,22	0,73	0,40
o-terbutilfenol	-0,12	-0,34	1,72	-0,12	-0,37	2,08
p-ter-octilfenol	1,00	0,56	0,44	1,00	0,77	0,23
m-ciclohexilfenol	0,52	0,41	0,22	0,52	0,45	0,15
m-ciclohexilmetilfenol	0,85	0,75	0,13	0,85	0,76	0,11
p-ciclohexilmetilfenol	0,92	0,75	0,18	0,92	0,76	0,17
(+) ferruginol	1,96	1,76	0,10	1,96	2,50	0,28
p-n-butoxifenol	-0,41	-0,34	0,17	-0,41	-0,01	0,98

Después de evaluar los modelos propuestos con los diferentes grupos de compuestos: centrales, límites y periféricos a la curva de selección, y realizar la validación externa podemos afirmar que en general muestran una buena calidad, que las capacidades predictivas se ajustan solo para los monofenoles que se ubican en las zonas al interior de  $N_{max}$  Vs  $(\Delta E_e - \Delta E_n)^2$ , y que el modelo 2 de *P. Gingivalis*, y modelo 1 para *S. Sobrinus*, son los de mejor calidad estadística en función de los valores de sus parámetros calculados.

### 5. Conclusiones

Se seleccionaron una serie de 100 monofenoles con bioactividad frente a *S. Sobrinus* y *P. Gingivalis*, para la construcción del modelo QSAR de Shapiro y Guggenheim (1998), y Joshi y colaboradores (2011).

Los descriptores electrónicos tratados estadísticamente no mostraron una buena correlación con la concentración inhibitoria mínima contra *S. Sobrinus* y *P. Gingivalis*, no obstante, los descriptores electrofugalidad, nucleofugalidad y carga máxima se utilizaron para seleccionar 56 monofenoles de los 100 iniciales y así obtener mejores resultados en los modelos QSAR empleando la curva  $N_{max} V_S (\Delta E_e - \Delta E_n)^2$ .

Para *P. Gingivalis* se obtuvo que el modelo 2 presenta el mejor carácter predictivo para la concentración inhibitoria mínima:

$$\text{LOG}(1/\text{MIC}) = -2,3 + 0,02\text{ALOGP} + 0,7\text{X1PER}$$

con parámetros estadísticos  $q^2_{\text{CV-LOO}} = 0,51$ ;  $Q^2_{\text{CV Externa}} = 0,96$ ;  $R^2_{\text{Externa}} = 0,89$  y  $\text{RMS}_{\text{Externa}} = 0,32$ .

La validación externa del modelo 2 de *P. Gingivalis* evidenció un carácter predictivo del 63% de las actividades inhibitorias de las 19 moléculas de prueba. Este conjunto de compuestos corresponde a monofenoles que estabilizan el anillo por medio del efecto inductivo mediante cadenas alifáticas y cíclicas, sin presentar un gran impedimento estérico, dadas las posiciones meta y para principalmente, frente al Hidrógeno hidroxílico.

El 37% de las moléculas de la validación externa del modelo 2 de *P. Gingivalis* que presentaron errores sobre el 50%, hacen referencia a monofenoles sustituidos con cadenas alifáticas en la posición orto principalmente que cuentan con carbonos secundarios y terciarios lo que disminuye la capacidad de solubilización en medio acuoso.

Para *S. sobrinus* se obtuvo que el modelo 1 presenta el mejor carácter predictivo para la concentración inhibitoria mínima:

$$\text{LOG}(1/\text{MIC}) = -2,23 + 0,0076\text{D}/\text{DTr}06 + 0,55\text{ALOGP}$$

con parámetros estadísticos  $q^2_{\text{CV-LOO}} = 0,94$ ;  $Q^2_{\text{CV Externa}} = 0,93$ ;  $R^2_{\text{Externa}} = 0,94$  y  $\text{RMS}_{\text{Externa}} = 0,27$

La validación externa del modelo 1 para *S. sobrinus* evidenció un carácter predictivo del 79% de las actividades inhibitorias de las 19 moléculas de prueba. Este conjunto de compuestos corresponde a monofenoles con cadenas alifáticas cíclicas y alicíclicas de carbonos primarios que estabilizan el anillo por medio del efecto inductivo.

El 21% de los compuestos de la validación externa del modelo 1 para *S. sobrinus* que no presentaron una buena predicción de la concentración mínima inhibitoria, hacen referencias a fenoles simples mono sustituidos con cadenas alifáticas que presentan carbonos secundarios y terciarios de baja hidrofiliidad.

### Referencias

- Abachi. S., Lee. S. and Vasantha. H. P. (2016). *Molecular Mechanisms of Inhibition of Streptococcus Species by Phytochemicals*. *Molecules*, (21), 215; doi:10.3390/molecules21020215
- Almeida. I., Denny. C., Benso. B., Matias de Alencar. S and Rosalen. P. L. (2015). *Antibacterial Activity of Essential Oils and Their Isolated Constituents against Cariogenic Bacteria: A Systematic Review*. *Molecules*, (20), 7329-7358; doi:10.3390/molecules20047329
- Amorin. C., Lamont. R. J. and Demuth D. R. (2008). *Interaction of Porphyromonas gingivalis with Oral Streptococci Requires a Motif That Resembles the Eukaryotic Nuclear Receptor Box Protein-Protein Interaction Domain*. *INFECTION AND IMMUNITY*, (76), 3273–3280. doi:10.1128/IAI.00366-08
- Araújo. J., Albuquerque. R., Turini da Cunha. M., Arruda .T. (2017). *Antibiofilm activity of natural substances derived from Plants*. *African Journal of Microbiology Research*. Vol. 11(26), pp. 1051-1060. doi: 10.5897/AJMR2016.8180
- Berlutti. F., Catizone. A., Riccp. G., Frioni. A., Natalizi. T., Valenti. P. and Polimenp. A. (2010). *Streptococcus Mutans and Streptococcus Sobrinus are Able to Adhere and Invade Human Gingival Fibroblast Cell Line*. *INTERNATIONAL JOURNAL OF IMMUNOPATHOLOGY AND PHARMACOLOGY*. ( 23), (4), 1253-1260.
- Campodómico. P., Andrés. J., Aizman. A., and Contreras. R. (2007). *Nucleofugality index in  $\alpha$ -elimination reactions*. *Chemical Physics Letters*, 439, 177–182.
- Carrasco. R. (2003). Tesis Doctoral. *Nuevos descriptores atómicos y moleculares para estudios de estructura actividad: Aplicaciones*. Editorial Universitaria Cuba. Recuperado el 5 de enero de 2022. <http://revistas.mes.edu.cu>.

- Cherkasov. A., Muratov. E. N., Fourches. D. (2014). *QSAR Modeling: Where Have You Been? Where Are You Going To?*. J. Med. Chem., 57, 4977–5010  
[dx.doi.org/10.1021/jm4004285](https://doi.org/10.1021/jm4004285)
- Chinsembu, K. C. (2015). *Plants and other natural products used in the management of oral infections and improvement of oral health*. Acta Tropica.  
<http://dx.doi.org/10.1016/j.actatropica.2015.10.019>
- Cremer. D. (1984). *On the Correct Usage of the Cremer-Pople Puckering Parameters as Quantitative Descriptors of Ring Shapes - a Reply to Recent Criticism by Petit, Dillen and Geise*. Acta Cryst. B40, 498-500.
- Estrada. E. (2002). *Physicochemical Interpretation of Molecular Connectivity Indices*. J. Phys. Chem. A 106, 9085-9091. Doi 10.1021/jp026238m
- Federación Dental Internacional (FDI). (2015). *EL DESAFÍO DE LAS ENFERMEDADES BUCODONTALES Una llamada a la acción global*. Atlas de Salud Bucodental SEGUNDA EDICIÓN. Recuperado el 6 de mayo de 2019 [www.myriadeditions.com](http://www.myriadeditions.com).
- Gramatica P. (2007). *Principles of QSAR models validation: internal and external*. QSAR CombSci. 26, No. 5, 694 – 701
- Gerits. E., Verstraeten. N. and Michiels. J. (2017) *New approaches to combat Porphyromonas gingivalis biofilms*. Journal of Oral Microbiology, 9:1, 1300366, DOI: 10.1080/20002297.2017.1300366
- Gujarati. D., Porter. D. (2010). *Econometría Quinta Edición*. Mc Garw Hill. México. Pag 127,136,146.

- Guziałowska. J. (2017). *The use of QSAR methods for determination of n-octanol/water partition coefficient using the example of hydroxyester HE-1*. E3S Web of Conferences **19**, 02034.  
DOI: 10.1051/e3sconf/20171902034
- Igarashi. I., Yamamoto. A., Goto. N. (2000). *PCR for Detection and Identification of Streptococcus Sobrinus*. J. Med Microbiol Vol 49, 1069-1074.
- Joshi. S., Yadav. M., Paradkar. L., and Anuraj N. S. (2012). *A QSAR Study on Minimum Inhibitory Concentrations of Some Phenols: Predicting  $-\log$  MIC Using Topological Indices*. Oxidation Communications 35, No 3, 684–693.
- Khalaf. H., Palm. E and Bengtsson .T.(2017) *Cellular Response Mechanisms in Porphyromonas gingivalis Infection. Chapter 3*. INTECH <http://dx.doi.org/10.5772/intechopen.69019>.
- Lundgren. M., Emilson. C.G., and Österberg. T.(1997). *Root Caries and Some Related Factors in 88-Year-Old Carriers and Non-Carriers of Streptococcus sobrinus in Saliva*. Caries Res;32:93–99.
- Maguna. F.P., Okulik. N.B., and Castro. E. A. (2017). *Quantitative Structure–Activity Relationships of Antimicrobial Compounds*. Handbook of Computational Chemistry, DOI 10.1007/978-3-319-27282-5\_38
- Melendez A, Barrón M., Rodríguez R., Morales M. y Elizondo A. (2016). *INHIBICIÓN DE Porphyromonas gingivalis POR Cinnamomum verum E Illicium verum*. Investigación y Desarrollo en Ciencia y Tecnología de Alimentos (1),(1) , 280-285
- Ministerio de salud Min Salud. (2014). *IV Estudio Nacional de Salud Bucal, ENSAB*. Recuperado el 6 de agosto de 2019 <https://www.minsalud.gov.co/Paginas/MinSalud-resultados-cuarto-estudio-nacional-salud-bucal.aspx>

- Nantasenamat .C., Isarankura-Na-Ayudhya. C. Prachayasittikul. V. (2010). *Advances in computational methods to predict the biological activity of compounds*. Expert Opin. Drug Discov. 5(7):633-654.
- Núñez, D. García, L. (2010). *Bioquímica de la caries dental*. Revista Habanera de Ciencias Médicas v.9 n.2 Ciudad de La Habana.
- OECD. 2004. f recuperado el 8 de Julio de 2020.  
<http://www.oecd.org/dataoecd/33/37/37849783.pdf>
- Ojeda, J. Oviedo, E. Salas, L. (2013). *Streptococcus mutans y caries dental*. CES odontol. vol.26 n°.1 Medellín.
- Okamoto. S., Terao. Y., Kaminishi. H., Hamada. S., and Kawabata. S. (2007). *Water-insoluble a-glucans from Streptococcus sobrinus induce infl ammatory immune responses*. Interface Oral Health Science ; 243–248
- Otero. J. A., Uribe J. Ochoa C., M. E. (2014). *Situación de salud bucal, visual y auditiva En Santander*. Revista del Observatorio de Salud Pública de Santander Año 10 · Número 1. Pp. 7-9
- Organizacion Mundial de la Salud. OMS. (2016). Recuperado 25 de Marzo de 2020.  
<https://www.who.int/news-room/fact-sheets/detail/oral-health>
- Peñarrieta. J. M.,Tejeda. L., Mollinedo. P., Vila. J. L., and Bravo. J. A. (2014). *PHENOLIC COMPOUNDS IN FOOD*. REVISTA BOLIVIANA DE QUÍMICA. Vol. 31, No.2, pp. 68-81.
- Pepperney. A., Chikindas. M. L. (2011). *Antibacterial Peptides: Opportunities for the Prevention and Treatment of Dental Caries*. Probiotics & Antimicro. Prot. 3:68–96. DOI 10.1007/s12602-011-9076-5

- Pir. M., Agirbas. H., Budak. F., and Iiter. M. (2016). *Synthesis, characterization, antimicrobial activity, and QSAR studies on substituted oxadiazaboroles*. *Med Chem Res* 25:1794–1812. DOI 10.1007/s00044-016-1603-1.
- Rastija. V., Nikoli. S., and Masand. V. H.(2013). *Quantitative Relationships Between Structure and Lipophilicity of Naturally Occurring Polyphenols*. *Acta Chim. Slov.* 60, 781–789.
- Rodríguez. M., Arcedalia. N., Murillo. M., Arcedalia. M., Ríos. L., and Javier. A. (2019). *Aplicación de la técnica validación cruzada para determinar el número óptimo de datos utilizados en modelos predictivos*. *Revista Electrónica de Divulgación de la Investigación* Vol. 17. Celaya: Instituto Tecnológico de Celaya.
- Roy. K., Kumar. P., Kar. S., Roy. P., and Mitra. I. (2010). *On further application of  $r^2_m$  as a metric for validation of QSAR models*. *J. Chemometrics*, 24. 22–33. DOI: 10.1002/cem.1268
- Roy. K., Kumar. P., Kar. S., Roy. P., and Mitra. I.(2012). *Introduction of  $rm^2$  (rank) metric incorporating rank-order predictions as an additional tool for validation of QSAR/QSPR models*. *Chemometrics and Intelligent Laboratory Systems* 118. 200–210. doi:10.1016/j.chemolab.2012.06.004
- Sega. M., Autieri. E., Pederiva. F.(2010). *Pickett angles and Cremer-Pople coordinates as collective variables for the enhanced sampling of six-membered ring conformations*. *Molecular Physics*, Taylor & Francis.. , <https://doi.org/10.1080/00268976.2010.522208>
- Seow. W. K. (2018). *Early Childhood Caries*. *Pediatr Clin N Am* 65. 941–954. <https://doi.org/10.1016/j.pcl.2018.05.004>

Todeschini, R. Consonni, V. (2000). *HANDBOOK OF MOLECULAR DESCRIPTOR*. WILEY VCH.

Tropsha. A. and Globraikh. A. (2007). *Predictive QSAR Modeling Workflow, Model Applicability Domains, and Virtual Screening*. *Current Pharmaceutical Design* 13, 3494-3504.

Vos. T., Alemu. A., Abbafati. C., et all. (2017). *Global, regional and national incidence, prevalence and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016*. *Lancet* (London, England) **390**, 1211–1259, [https://doi.org/10.1016/s0140-6736\(17\)32154-2](https://doi.org/10.1016/s0140-6736(17)32154-2).

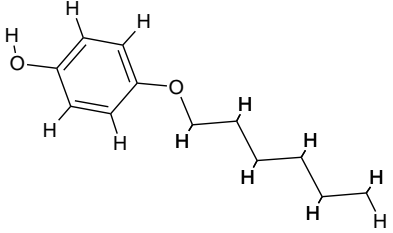

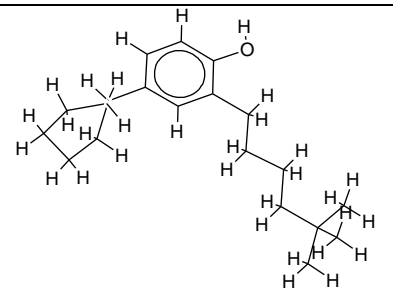
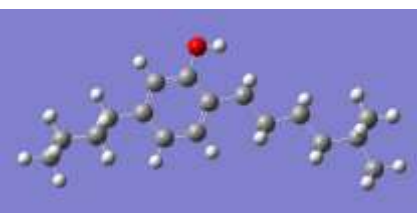
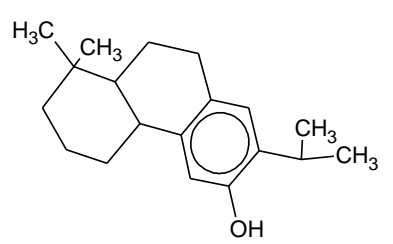
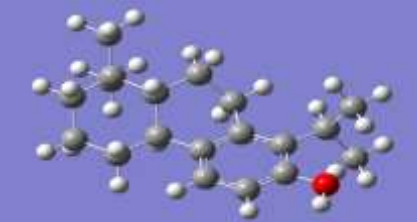
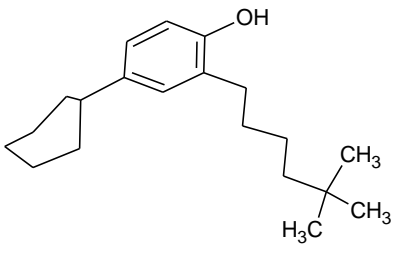
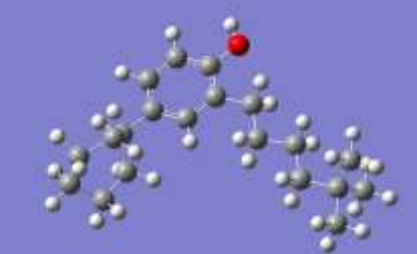
Wang. X., Mitsunaga. T., and Yamauch. K. (2016). *Antibacterial Activity of Selected Essential oils against Streptococcus sobrinus and Porphyromonas gingivalis*. Research Article, *J Pharm Sci Emerg Drugs* (4) Issue: 2

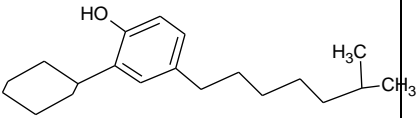

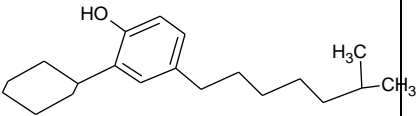
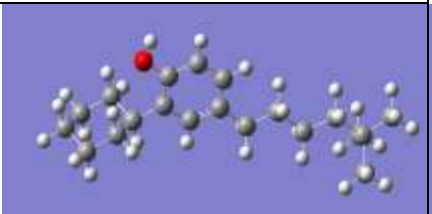
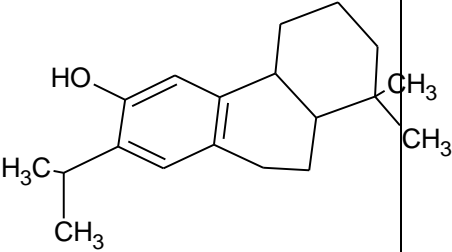
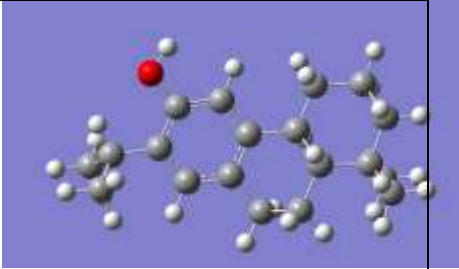
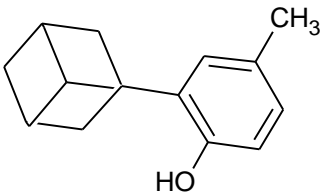
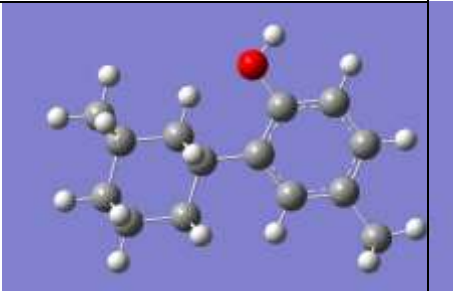
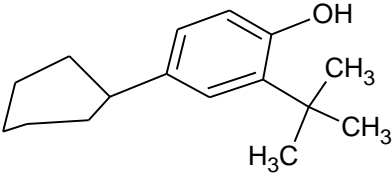
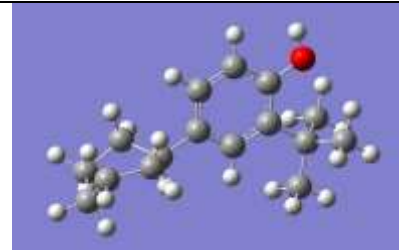
Zhang. J., Leung. K., Sardana. D., Wong .M. C.M., Lo E.C.M. (2019). *Risk predictors of dental root caries: a systematic review*. *Journal of Dentistry*. <https://doi.org/10.1016/j.jdent.2019.07.004>

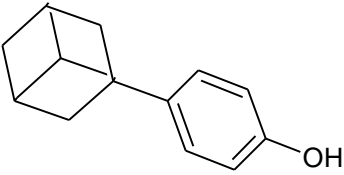
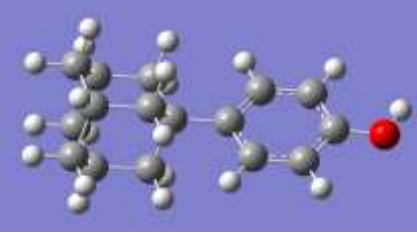
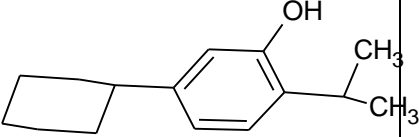

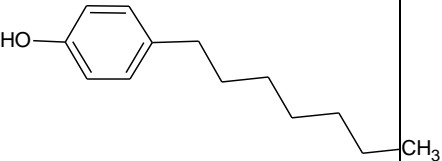
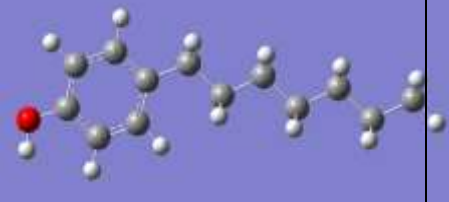
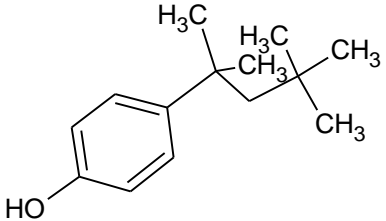
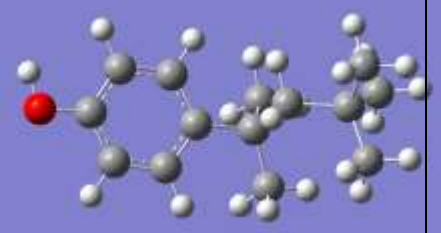
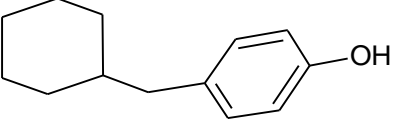

Zhu. B., Macleod. L. C., Newsome. E., Liu. J. and Xu. P. (2019). *Aggregatibacter actinomycetemcomitans mediates protection of Porphyromonas gingivalis from Streptococcus sanguinis hydrogen peroxide production in multi-species biofilms*. *Scientific Reports*. 9:4944. <https://doi.org/10.1038/s41598-019-41467-9>

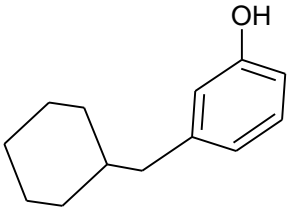

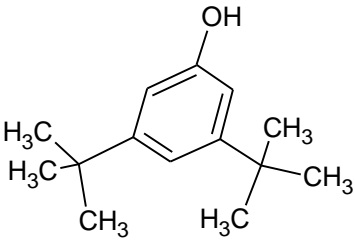

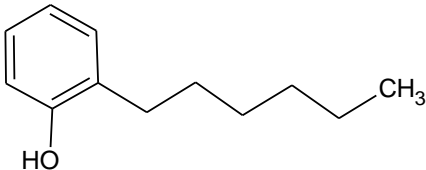

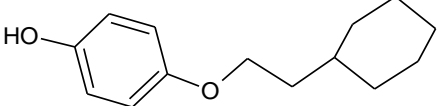

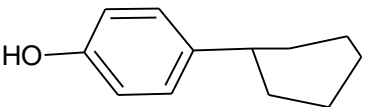
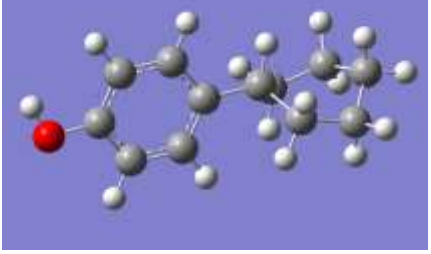
Apéndices

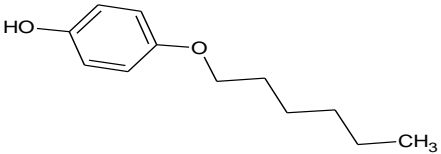

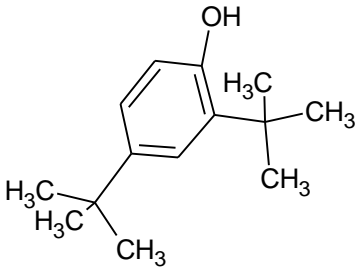
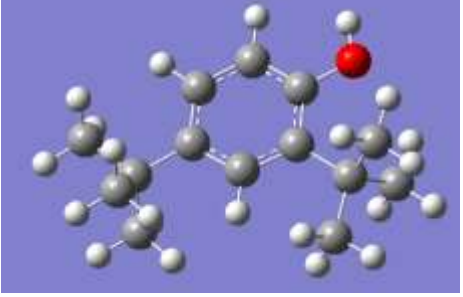
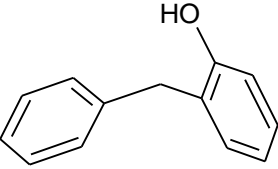

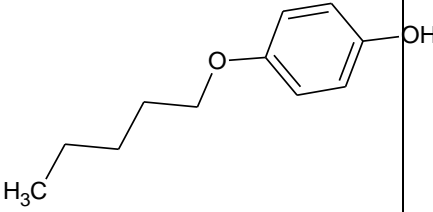

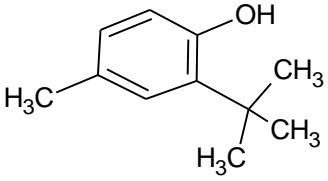

Apéndice A. Monofenoles optimizados

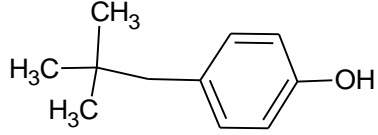
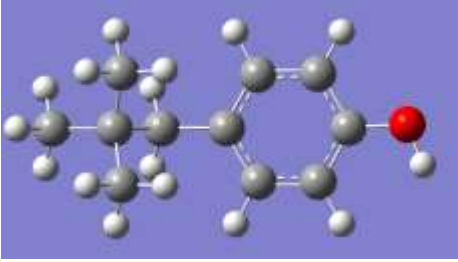
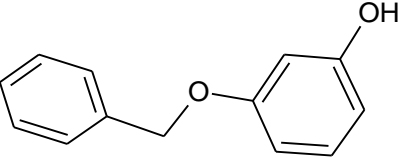

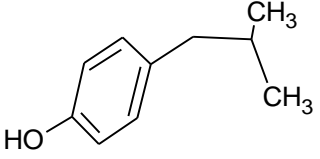
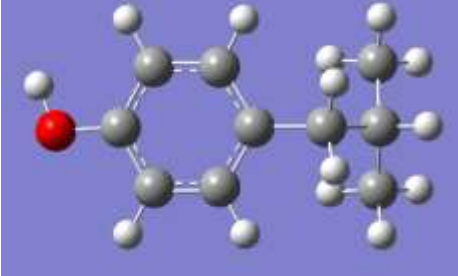
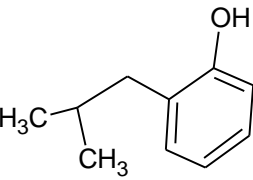
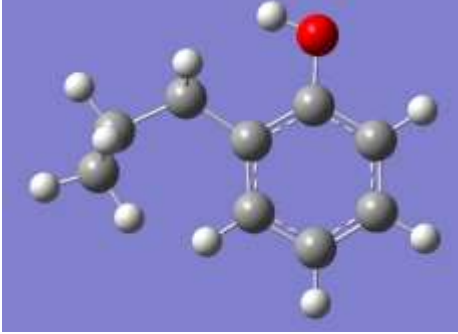
No	Molécula		
1	p-n-octilfenol	 <p>Chemical structure of p-n-octylphenol, showing a benzene ring with a hydroxyl group (-OH) at the para position relative to an octyl chain (-C<sub>8</sub>H<sub>17</sub>).</p>	 <p>3D ball-and-stick model of p-n-octylphenol, with carbon atoms in grey, hydrogen atoms in white, and oxygen atoms in red.</p>
2	2-teroctil-5-ciclohexilfenol	 <p>Chemical structure of 2-teroctil-5-ciclohexilfenol, showing a benzene ring with a hydroxyl group (-OH) at the 2-position, a teroctyl chain at the 1-position, and a cyclohexyl ring at the 5-position.</p>	 <p>3D ball-and-stick model of 2-teroctil-5-ciclohexilfenol, with carbon atoms in grey, hydrogen atoms in white, and oxygen atoms in red.</p>
3	(+) -totarol	 <p>Chemical structure of (+)-totarol, showing a complex polycyclic structure with a benzene ring fused to a bicyclic system, featuring a hydroxyl group (-OH) and two methyl groups (-CH<sub>3</sub>).</p>	 <p>3D ball-and-stick model of (+)-totarol, with carbon atoms in grey, hydrogen atoms in white, and oxygen atoms in red.</p>
4	2-terocti-4-ciclohexilfenol	 <p>Chemical structure of 2-terocti-4-ciclohexilfenol, showing a benzene ring with a hydroxyl group (-OH) at the 2-position, a teroctyl chain at the 1-position, and a cyclohexyl ring at the 4-position.</p>	 <p>3D ball-and-stick model of 2-terocti-4-ciclohexilfenol, with carbon atoms in grey, hydrogen atoms in white, and oxygen atoms in red.</p>

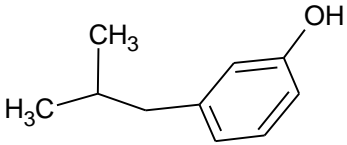

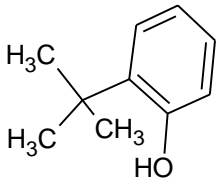

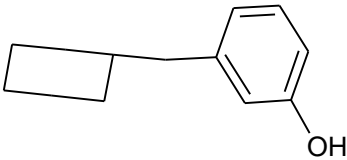

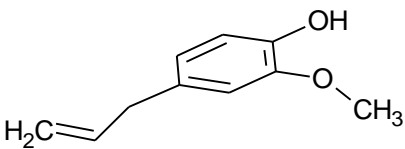

No	Molécula		
5	2-ciclohexil-4-teroctilfenol		
6	2-ciclohexil-4-teroctilfenolsilla		
7	(+ ferruginol		
8	2,1-adamantil-p-cresol		
9	2-terbutil-4-ciclohexilfenol		

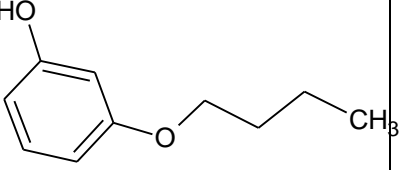

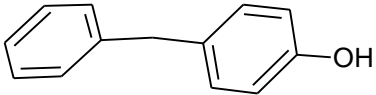
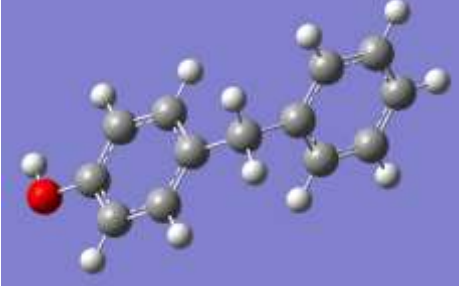
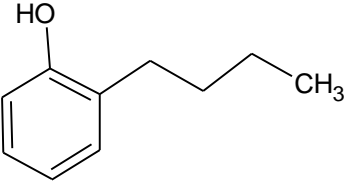

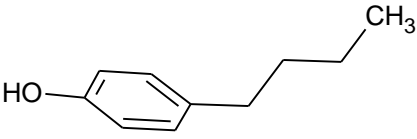

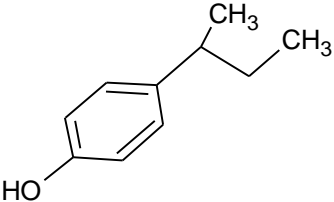

No	Molécula		
11	1-adamantilfenol		
12	2-terbutil-5-ciclohexilfenol		
13	p-n-heptilfenol		
14	p-ter-octilfenol		
15	p-ciclohexilmetilfenol		

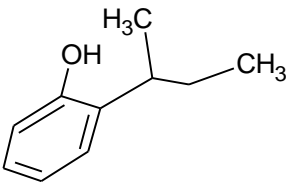

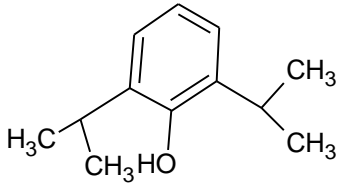

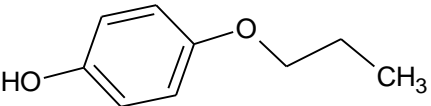
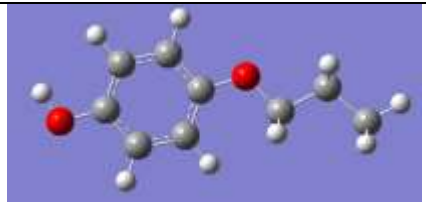
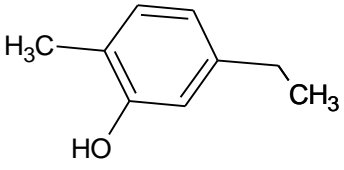
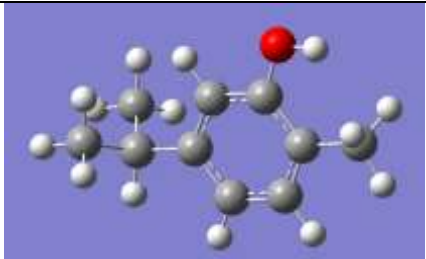
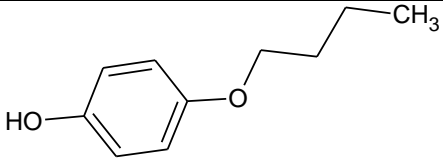
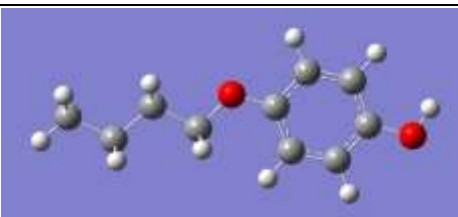
No	Molécula		
16	m-ciclohexilmetilfenolsilla		
17	3,5-diterbutilfenol		
18	o-n-hexilfenol		
19	p-ciclohexilmetiloxifenol silla		
20	p-ciclohexilfenol		

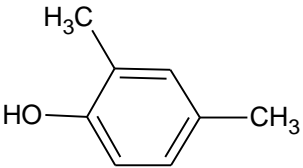
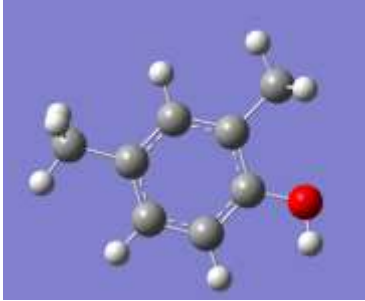
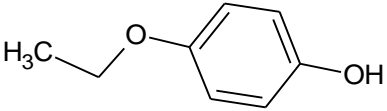

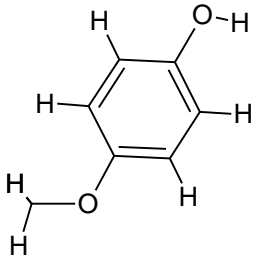

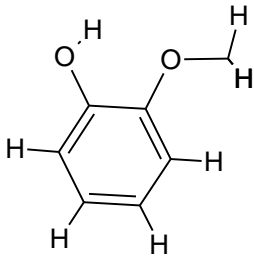

No	Molécula		
21	p-n-hexiloxifenol		
22	2,4-diterbutilfenol		
23	o-bencilfenol		
24	p-n-amiloxifenol		
25	2-terbutil-4-metilfenol		

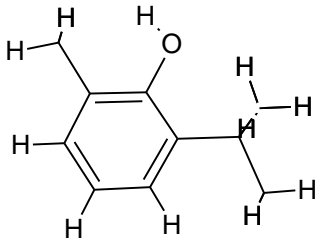

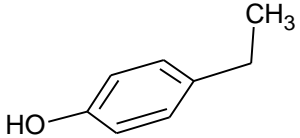
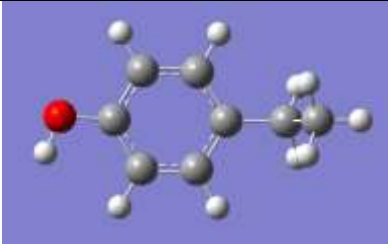
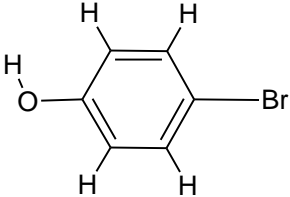
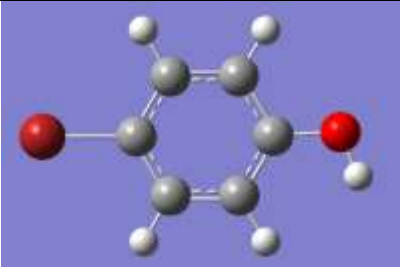
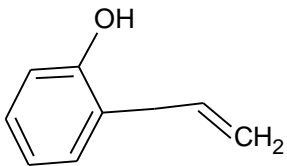

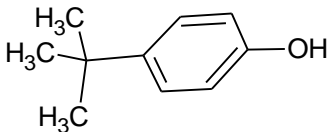
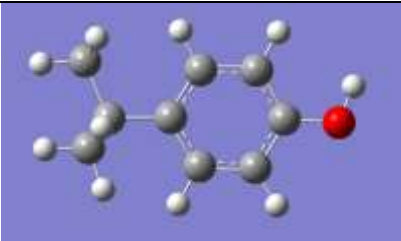
No	Molécula		
26	p-teramilfenol		
27	m-benziloxifenol		
28	p-isobutilfenol		
29	o-isobutilfenol		

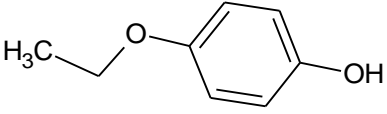

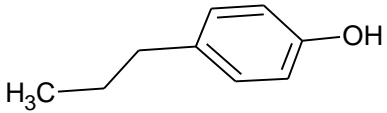
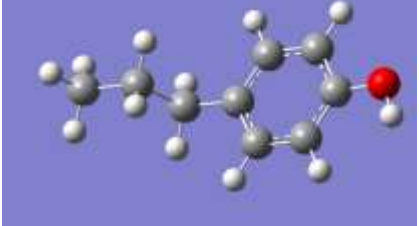
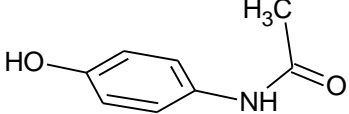

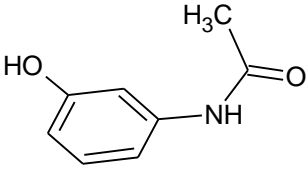
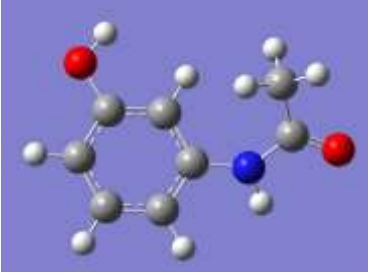
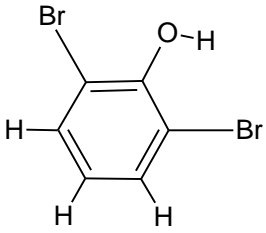

No	Molécula		
30	m-isobutilfenol		
31	o-terbutilfenol		
32	m-ciclohexilmetilfenol		
33	eugenol		

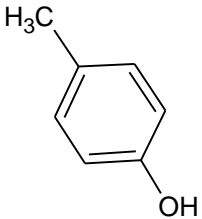
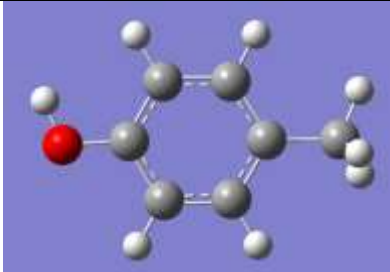
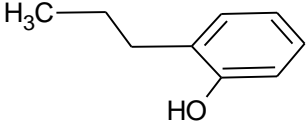

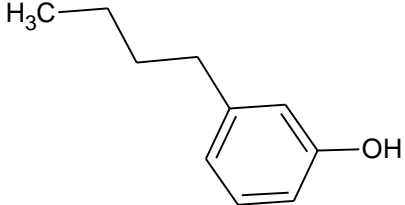

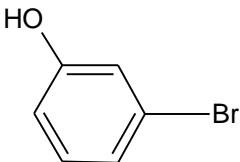

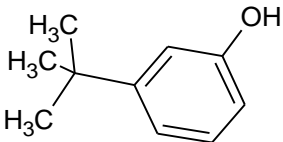

No	Molécula		
34	m-n-butoxifenol		
35	p-bencilfenol		
36	o-n-butilfenol		
37	p-n-butilfenol		
38	p-secbutilfenol		

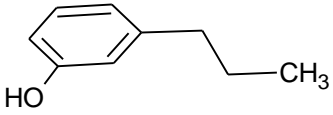

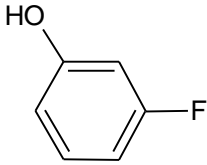
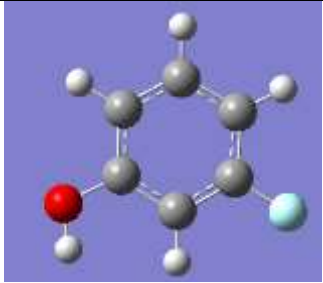
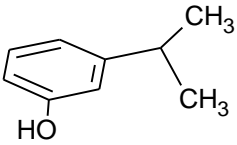
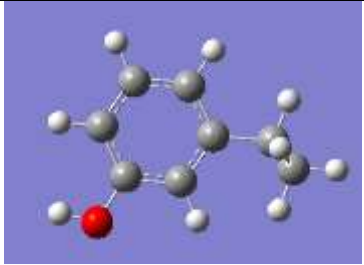
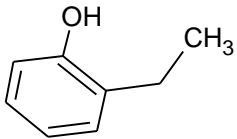

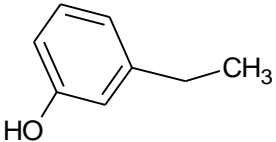
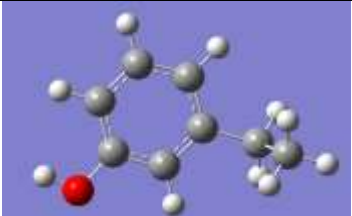
No	Molécula		
39	o-secbutilfenol		
40	2,6-diisopropilfenol		
41	p-n-propoxifenol		
42	carvocrol		
43	p-n-butoxifenol		

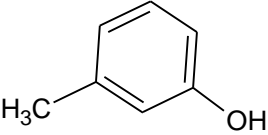

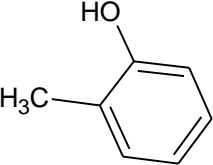

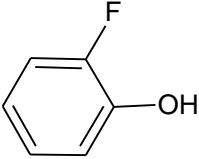

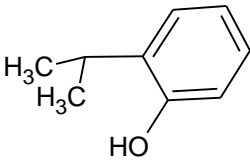

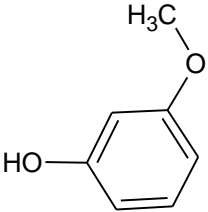
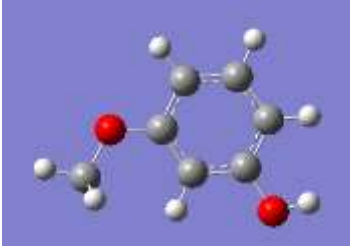
No	Molécula		
44	2,4-xilenol		
45	p- etoxifenol		
46	p- metoxifenol		
47	o-metoxifenol		

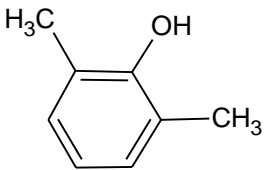

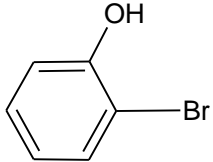
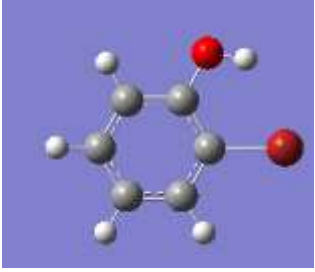
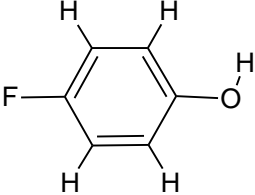

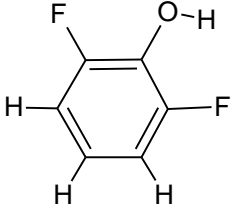

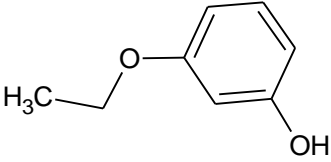

No	Molécula		
48	2-terbutil-6-metilfenol	 <p>Chemical structure of 2-terbutyl-6-methylphenol, showing a benzene ring with a hydroxyl group at position 1, a methyl group at position 6, and a tert-butyl group at position 2.</p>	 <p>3D ball-and-stick model of 2-terbutyl-6-methylphenol, showing the spatial arrangement of atoms (carbon in grey, oxygen in red, hydrogen in white).</p>
49	p-etilfenol	 <p>Chemical structure of p-ethylphenol, showing a benzene ring with a hydroxyl group at position 1 and an ethyl group at position 4.</p>	 <p>3D ball-and-stick model of p-ethylphenol, showing the spatial arrangement of atoms.</p>
50	p-bromofenol	 <p>Chemical structure of p-bromophenol, showing a benzene ring with a hydroxyl group at position 1 and a bromine atom at position 4.</p>	 <p>3D ball-and-stick model of p-bromophenol, showing the spatial arrangement of atoms.</p>
51	o-alilfenol	 <p>Chemical structure of o-allylphenol, showing a benzene ring with a hydroxyl group at position 1 and an allyl group at position 2.</p>	 <p>3D ball-and-stick model of o-allylphenol, showing the spatial arrangement of atoms.</p>
52	p-terbutilfenol	 <p>Chemical structure of p-tert-butylphenol, showing a benzene ring with a hydroxyl group at position 1 and a tert-butyl group at position 4.</p>	 <p>3D ball-and-stick model of p-tert-butylphenol, showing the spatial arrangement of atoms.</p>

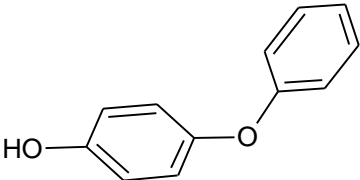
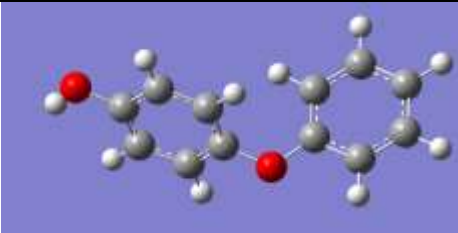
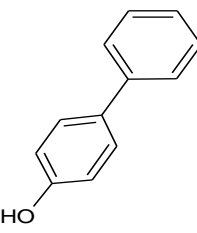
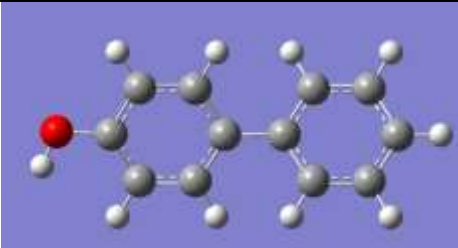
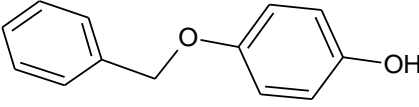
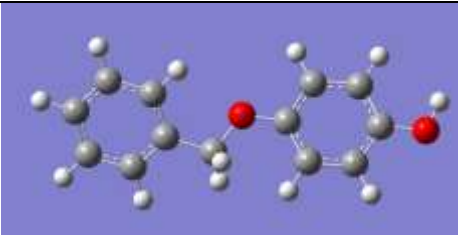
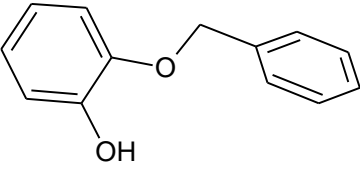
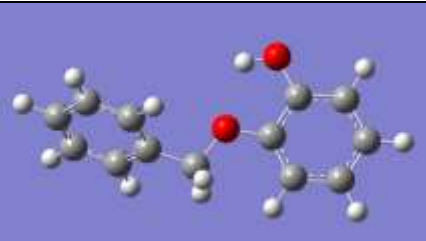
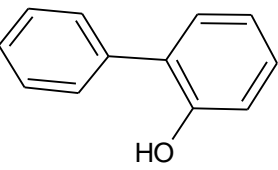
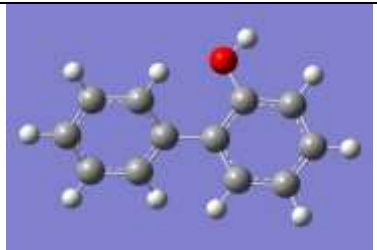
No	Molécula		
53	o-etoxifenol		
54	p-n-propilanol		
55	p-acetoamidofenol		
56	m-acetoamidofenol		
57	2,6-dibromofenol		

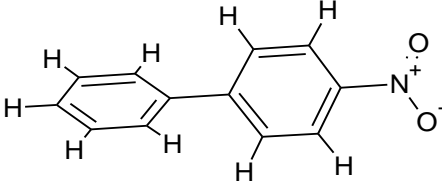
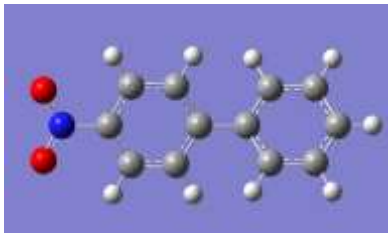
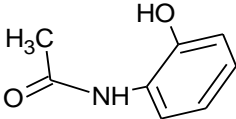

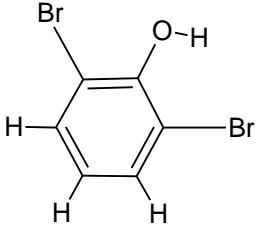

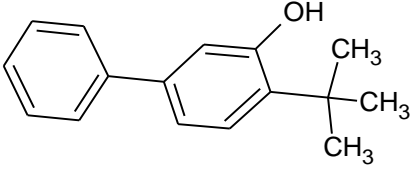
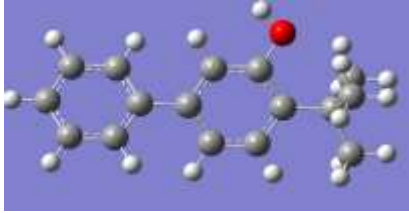
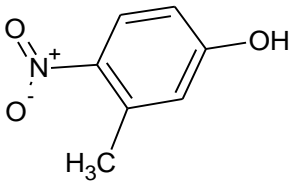

No	Molécula		
58	p-cresol		
59	o-n-propilfenol		
60	m-n-butilfenol		
61	m-bromofenol		
62	m-terbutilfenol		

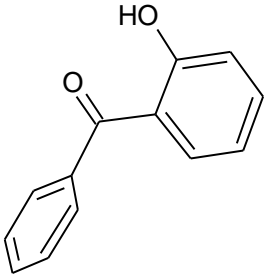
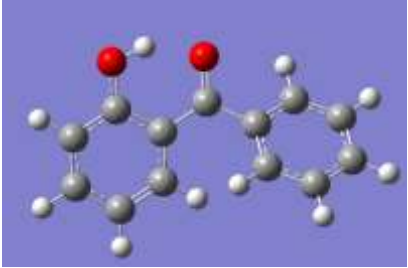
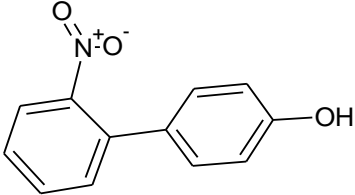

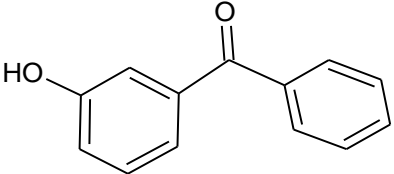

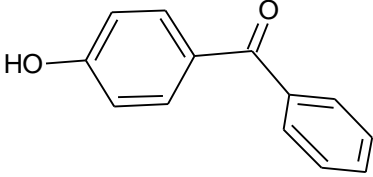
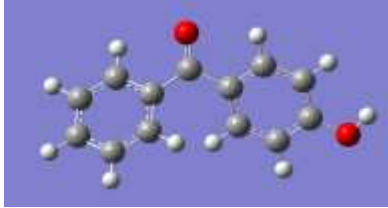
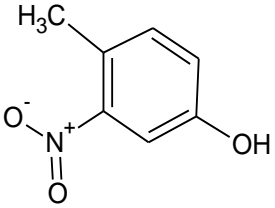

No	Molécula		
63	m-n-propilfenol		
64	m-fluorfenol		
65	m-isopropilfenol		
66	o-etilfenol		
67	m-etilfenol		

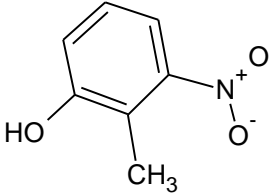

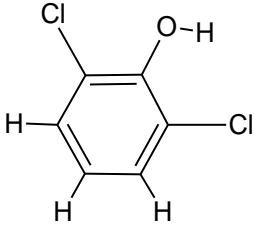
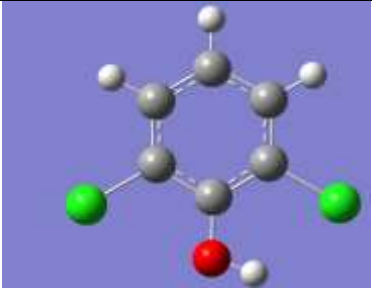
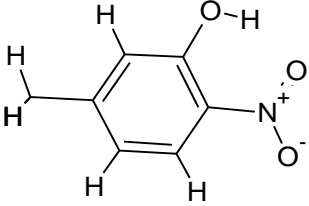
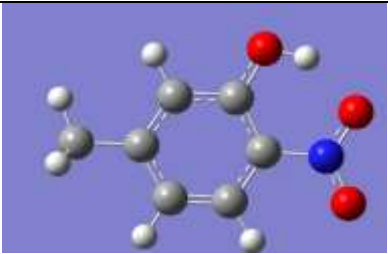
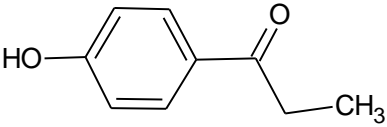
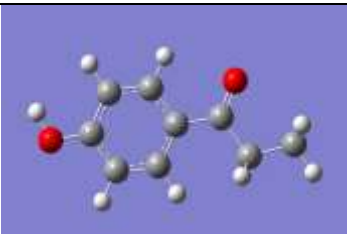
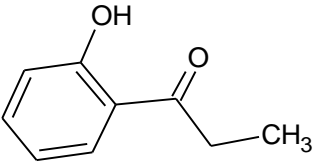
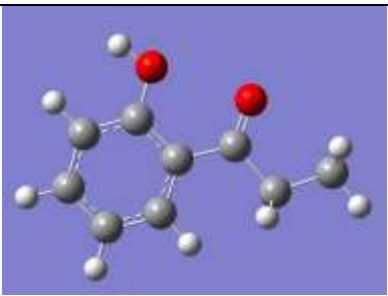
No	Molécula		
68	m-cresol		
69	o-cresol		
70	o-fluorfenol		
71	o-isopropilfenol		
72	m-metoxifenol		

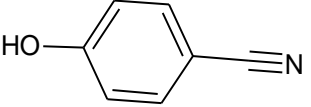

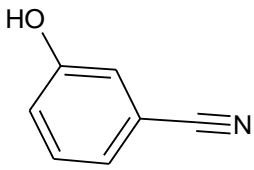

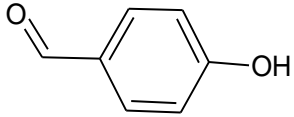
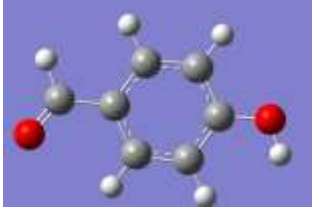
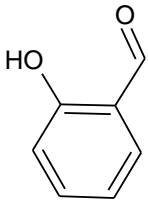

No	Molécula		
73	2,6-xilenol		
74	o-bromofenol		
75	p-fluorfenol		
76	2,6-difluorfenol		
77	m-etoxifenol		

No	Molécula		
78	p-fenoxifenol		
79	p-fenilfenol		
80	p-benziloxifenol		
81	o-benzoiloxifenol		
82	o-fenilfenol		

No	Molécula		
83	4-nitro-4-bifenilfenol	 <p>Chemical structure of 4-nitro-4-biphenylphenol, showing two phenyl rings connected at the para position, with a hydroxyl group and a nitro group on the second ring.</p>	 <p>3D ball-and-stick model of 4-nitro-4-biphenylphenol, showing the spatial arrangement of atoms (C, H, O, N).</p>
84	o-acetamidofenol	 <p>Chemical structure of o-acetamidophenol, showing a benzene ring with a hydroxyl group and an acetamido group in the ortho position.</p>	 <p>3D ball-and-stick model of o-acetamidophenol, showing the spatial arrangement of atoms (C, H, O, N).</p>
85	2,6-dibromofenol	 <p>Chemical structure of 2,6-dibromophenol, showing a benzene ring with a hydroxyl group and two bromine atoms at the 2 and 6 positions.</p>	 <p>3D ball-and-stick model of 2,6-dibromophenol, showing the spatial arrangement of atoms (C, H, O, Br).</p>
86	2-terbutil-5-fenilfenol	 <p>Chemical structure of 2-tert-butyl-5-phenylphenol, showing a benzene ring with a hydroxyl group, a tert-butyl group, and a phenyl group.</p>	 <p>3D ball-and-stick model of 2-tert-butyl-5-phenylphenol, showing the spatial arrangement of atoms (C, H, O).</p>
87	4-nitro-m-cresol	 <p>Chemical structure of 4-nitro-m-cresol, showing a benzene ring with a hydroxyl group, a methyl group, and a nitro group.</p>	 <p>3D ball-and-stick model of 4-nitro-m-cresol, showing the spatial arrangement of atoms (C, H, O, N).</p>

No	Molécula		
88	o-benzoilfenol		
89	2-nitro-4-bifenilfenol		
90	m-benzoilfenol		
91	p-benzoilfenol		
92	3-nitro-p-cresol		

No	Molécula		
93	3-nitro-o-cresol		
94	2,6-diclorofenol		
95	6-nitro-m-cresol		
96	p-propionilfenol		
97	o-propionilfenol		

No	Molécula		
98	p-cianofenol		
99	m-cianofenol		
100	p-acetilfenol		
101	o-acetilfenol		

**Apéndice B. Descriptores moleculares**

<b>Molécula</b>	<b>ΔEe</b>	<b>ΔEn</b>	<b>Nmax</b>	<b>ALOGP</b>	<b>X1PER</b>	<b>D/DTR06</b>
p-n-octilfenol	0,031	-0,324	0,314	5,242	5,044	55,013
2-teroctil-5-ciclohexilfenol	0,032	-0,315	0,311	6,979	7,099	151,941
(+)-totarol	0,027	-0,309	0,318	5,772	5,772	103,837
2-teroctil-14-ciclohexilfenol	0,032	-0,31	0,311	6,979	7,099	145,738
2-ciclohexil-4-teroctilfenol	0,028	-0,314	0,318	7,229	7,331	143,985
2-ciclohexil-4-teroctilfenolsilla	0,028	-0,314	0,318	6,979	7,1	143,587
(+) ferruginol	0,029	-0,304	0,314	5,772	6,566	107,16
2,1-adamantil-p-cresol	0,03	-0,314	0,313	4,458	4,458	196,854
2-terbutil-4-ciclohexilfenol	0,028	-0,313	0,318	5,154	5,479	107,712
2-terbutil-5-ciclohexilfenol	0,032	-0,317	0,311	4,948	5,27	101,906
p-n-heptilfenol	0,035	-0,323	0,308	4,786	4,627	49,961
p-ter-octilfenol	0,031	-0,317	0,313	4,333	4,382	52,967
p-ciclohexilmetilfenol	0,029	-0,323	0,317	4,21	4,581	87,796
m-ciclohexilmetilfenolsilla	0,037	-0,32	0,303	4,124	5,08	109,477
3,5-diter butilfenol	0,027	-0,321	0,319	4,363	4,314	51,143
o-n-hexilfenol	0,038	-0,328	0,304	4,33	4,217	44,992
p-ciclohexilmetiloxifenol	0,033	-0,304	0,307	4,124	5,079	110,39
p-ciclohexilfenol	0,032	-0,321	0,313	3,754	4,192	76,36
p-n-hexiloxifenol	0,037	-0,304	0,301	3,787	4,297	49,961
2,4-diterbutilfenol	0,029	-0,316	0,317	4,363	4,32	51,143
o-bencilfenol	0,005	-0,33	0,358	3,537	3,796	85,829
p-n-amiloxifenol	0,037	-0,304	0,301	3,331	3,881	44,992
2-terbutil-4-metilfenol	0,033	-0,319	0,311	3,449	3,348	38,571
p-teramilfenol	0,03	-0,324	0,316	3,419	3,416	39,61
m-benziloxifenol	0	-0,332	0,365	3,13	3,886	98,132
p-isobutilfenol	0,031	-0,325	0,314	3,213	3,225	35,181
o-isobutilfenol	0,035	-0,329	0,309	3,213	3,232	35,181
m-isobutilfenol	0,031	-0,333	0,315	3,213	3,226	35,181
o-terbutilfenol	0,038	-0,33	0,305	2,963	3,041	34,705
m-ciclohexilmetilfenol	0,029	-0,331	0,318	4,21	4,582	86,711
eugenol	0,021	-0,321	0,33	2,554	3,018	38,81
p-bencilfenol	0,006	-0,324	0,355	3,537	3,79	87,796
o-n-butilfenol	0,038	-0,328	0,302	3,417	3,383	35,363

Molécula	$\Delta E_e$	$\Delta E_n$	Nmax	ALOGP	XIPER	D/DTR06
o-n-butilfenol	0,038	-0,328	0,302	3,417	3,383	35,363
p-n-butilfenol	0,034	-0,324	0,309	3,417	3,377	35,363
p-secbutilfenol	0,033	-0,324	0,312	3,213	3,277	34,943
o-secbutilfenol	0,039	-0,329	0,303	3,213	3,284	34,943
2,6 -diisopropilfenol	0,031	-0,322	0,314	3,951	3,909	42,762
p-n-propoxifenol	0,037	-0,305	0,301	2,419	3,047	35,363
carvocrol	0,035	-0,324	0,308	3,243	3,136	34,381
p-n-butoxifenol	0,037	-0,304	0,301	2,875	3,464	40,119
2,4-xilenol	0,041	-0,319	0,297	2,535	2,373	26
p-etoxifenol	0,037	-0,306	0,301	1,895	2,611	30,752
p- metoxifenol	0,037	-0,308	0,301	1,546	2,13	26,324
o-metoxifenol	0,052	-0,314	0,279	1,546	2,139	26,324
2-terbutil-6-metilfenol	0,036	-0,323	0,307	2,505	3,353	38,571
p-etilfenol	0,035	-0,327	0,309	2,311	2,532	26,324
p-bromofenol	0,011	-0,344	0,348	2,571	1,803	22,133
o-alilfenol	0,02	-0,336	0,333	2,963	2,638	30,752
p-terbutilfenol	0,032	-0,325	0,312	1,895	3,032	34,705
o-etoxifenol	0,046	-0,312	0,288	2,961	2,62	30,752
p-acetoamidofenol	0,007	-0,332	0,353	0,683	2,612	26,324
2,6-dibromofenol	-0,013	-0,363	0,387	3,059	1,873	26
p-cresol	0,037	-0,328	0,306	2,049	2,061	22,133
o-n-propilfenol	0,031	-0,332	0,315	2,961	2,966	30,752
m-n-butilfenol	0,031	-0,333	0,316	3,417	3,378	35,363
m-bromofenol	0,009	-0,355	0,352	2,311	1,804	22,133
m-terbutilfenol	0,029	-0,334	0,318	2,963	3,034	34,705
m-n-propilfenol	0,032	-0,334	0,314	2,961	2,961	30,752
m-fluorfenol	0,03	-0,354	0,32	1,768	1,804	22,133
m-isopropilfenol	0,031	-0,335	0,316	2,757	2,825	30,514
o-etilfenol	0,034	-0,333	0,31	2,505	2,538	26,324
m-etilfenol	0,032	-0,336	0,315	2,505	2,533	26,324
m-cresol	0,038	-0,335	0,306	2,049	2,062	22,133
o-cresol	0,041	-0,332	0,3	2,049	2,066	22,133
o-fluorfenol	0,033	-0,352	0,315	1,768	1,813	22,133
o-isopropilfenol	0,033	-0,332	0,311	2,757	2,831	30,514
m-metoxifenol	0,05	-0,319	0,284	1,546	2,131	26,324
2,6-xilenol	0,044	-0,323	0,293	2,535	2,378	26
p-fluorfenol	0,02	-0,347	0,334	1,768	1,803	22,133
2,6-difluorfenol	0,031	-0,36	0,319	1,974	1,873	26
m-etoxifenol	0,13	0,172	-2,454	1,895	2,612	30,752
p-fenoxifenol	-0,002	-0,316	0,369	3,123	3,543	87,796

Molécula	$\Delta E_e$	$\Delta E_n$	Nmax	ALOGP	X1PER	D/DTR06
p-fenoxifenol	-0,002	-0,316	0,369	3,123	3,543	87,796
p-fenilfenol	-0,009	-0,322	0,383	3,081	3,425	76,36
p-benziloxifenol	-0,001	-0,316	0,367	3,13	3,885	99,124
o-benzoiloxifenol	-0,002	-0,323	0,37	3,13	3,894	97,31
o-fenilfenol	-0,015	-0,33	0,393	3,081	3,433	74,208
o-acetamidofenol	0	-0,355	0,366	0,683	2,62	35,181
2,6-dibromofenol	-0,013	-0,363	0,388	3,059	1,873	26
2-terbutil-5-fenilfenol	-0,026	-0,328	0,414		4,713	110,969
4-nitro-m-cresol	-0,085	-0,412	0,505	1,943	2,442	34,381
o-benzoilfenol	-0,1	-0,387	0,548	2,967	3,789	94,21
m-benzoilfenol	-0,088	-0,384	0,524	2,967	3,781	95,091
p-benzoilfenol	-0,075	-0,376	0,499	2,967	3,78	96,177
3-nitro-p-cresol	-0,1	-0,409	0,537	1,943	2,442	34,381
3-nitro-o-cresol	-0,089	-0,405	0,515	1,943	2,447	34,381
2,6-diclorofenol	-0,008	-0,367	0,378	2,891	1,873	26
6-nitro-m-cresol	-0,116	-0,422	0,562	1,943	2,447	34,381
p-propionilfenol	-0,043	-0,39	0,435	1,969	2,881	34,943
o-propionilfenol	-0,051	-0,383	0,451	1,969	2,889	34,943
p-cianofenol	-0,029	-0,381	0,413	1,442	2,053	26,324
p-acetilfenol	-0,053	-0,409	0,448	1,322	2,086	26,324
o-acetilfenol	-0,076	-0,395	0,494	1,322	2,094	26,324