

GENERACIÓN ESTOCÁSTICA DE ESTRUCTURAS MOLECULARES PROMEDIOS A PARTIR DE DATOS DE CARACTERIZACIÓN DE QUÍMICA ANALÍTICA DE SISTEMAS ORGÁNICOS COMPLEJOS

STOCHASTIC GENERATION OF AVERAGE MOLECULAR STRUCTURES FROM ANALYTICAL CHEMISTRY CHARACTERIZATION DATA OF COMPLEX ORGANIC SYSTEMS

Iván Machín^{1*}, Gledys Gonzalez²

(1) Castillo Max Oil & Gas, Av. Casanova, Centro Comercial El Recreo, Torre Sur, piso 8, ofic. 8-5, Caracas - Venezuela

(2) Haltermann Carless UK Limited, Schlegendeich 17, 21107 Hamburg - Germany

(*e-mail: machin_ivan@outlook.com)

Recibido: 20/04/2020 - Evaluado: 29/05/2020 - Aceptado: 11/06/2020

RESUMEN

La conversión de un compuesto dado durante un proceso químico es un reflejo de su estructura molecular. Esta es la razón por la cual en las últimas décadas, se ha desarrollado todo un despliegue de metodologías para convertir los datos de caracterización del petróleo y sus derivados en estructuras moleculares. Esto ayudaría enormemente en el diseño y optimización de los procesos de conversión del petróleo y sus derivados. En este trabajo se presenta un algoritmo que permite convertir los datos de caracterización de Química Analítica (RMN ¹H, VPO y análisis elemental) de sistemas orgánicos complejos, en estructuras moleculares generadas de manera estocástica. Este algoritmo es una herramienta de apoyo para el desarrollo de programas computacionales de generación de estructuras moleculares promedios. La metodología propuesta en ese estudio fue aplicada a los datos de caracterización de fracciones asfálticas, reportadas en la literatura, generando estructuras moleculares promedio similares a estas.

ABSTRACT

The conversion of a given compound die during a chemical process is a reflection of its molecular structure. This is the reason why in the last decades, a whole array of methodologies has been developed to convert the characterization data of oil and its derivatives into molecular structures. This would help enormously in the design and optimization of the conversion processes of oil and its derivatives. In this work, an algorithm is presented that allows converting analytical chemistry characterization data (¹H NMR, VPO and elemental analysis) of complex organic systems into stochastically generation molecular structures. This algorithm is a support tool for the development of computational programs for generating average molecular structures. The methodology proposed in that study was applied to the characterization data of asphalt fractions, reported in the literature, generating average molecular structures similar to these.

Palabras clave: parámetros estructurales, RMN protones, estructura molecular asfaltenos, generación estocástica
Keywords: structural parameters, NMR protons, asphaltenes molecular structure, stochastic generation

INTRODUCCIÓN

La historia del análisis de los constituyentes del petróleo tiene su origen en 1885, cuando LaRue y Miller identificaron ciertos hidrocarburos aromáticos en un petróleo de Burma (Speight, 1980). Actualmente se han identificado 4 familias químicas distintas representativas de las diferentes fracciones del petróleo (desde las más livianas hasta las más pesadas), y éstos son: Saturados, Aromáticos, Resinas y Asfaltenos (SARA) (Speight, 1980). El componente que genera más problemas en una gran variedad de procesos en la industria petrolera son los asfaltenos, debido a su capacidad de floculación y precipitación. Esta característica ha generado una amplia gama de investigaciones dedicadas a la caracterización de las resinas y asfaltenos, sobre las cuales se han propuesto estructuras moleculares promedios con la finalidad de entender sus propiedades funcionales y su capacidad de formar agregados de acuerdo a las condiciones experimentales en que se encuentren (Ali *et al.*, 2004; Alí *et al.*, 2006; Artok *et al.*, 1999; Allen *et al.*, 1984; Bansal, 2004; Christopher *et al.*, 1995; Coelho *et al.*, 2006; Dickinson, 1980; Alhreez & Wen, 2019; Liu & Li, 2015). Una de las primeras aproximaciones al conocimiento de estructuras moleculares de asfaltenos es debida a Williams (1958), quien propuso una metodología para la generación de estructuras moleculares de familias del petróleo que tienen un componente aromático partiendo del cálculo de un grupo de atributos moleculares. Estos atributos son el número de anillos aromáticos y saturados, número y tipo de sustituciones en los anillos aromáticos, número de átomos de carbono e hidrógeno, entre otros, que en promedio representan la forma esquelética o geometría de la estructura molecular de la sustancia en estudio. Los datos experimentales utilizados por Williams para la generación de los parámetros moleculares promedios son los provenientes de: RMN ^1H , análisis elemental y pesos moleculares determinados por VPO. Otro gran aporte al campo de aproximación a estructuras moleculares lo realizó Dickinson (1980), quien al igual que Williams, postuló correlaciones "interdependientes" para la determinación de parámetros moleculares promedios, utilizando además datos experimentales de RMN ^{13}C . Donde las correlaciones fueron aplicadas a materiales provenientes de la separación del crudo de diferentes procedencias.

Posteriormente, Allen *et al.* (1984), presentaron un método de caracterización estructural para productos hidrotratados derivados del carbón, definido como análisis de grupos funcionales. En este método se calculan las concentraciones de diferentes grupos funcionales utilizando datos de análisis elemental, RMN ^1H y RMN ^{13}C y se proponen las estructuras moleculares como la suma de grupos funcionales a los cuales, mediante diferentes correlaciones, se les asigna un valor de probabilidad. Michon *et al.* (1996), postularon parámetros moleculares promedio "independientes", utilizando como información experimental de partida la proveniente de RMN ^1H y RMN ^{13}C .

Otros investigadores basados tanto en los trabajos de Williams, como en metodologías propias han propuesto diferentes programas para la automatización del proceso de generación de estructuras moleculares. Basados en los lineamientos postulados por Speight y Moschopedis (1981) y Yen (1984), sobre la estructura general de asfaltenos y otros componentes del petróleo. Por otro lado, Savage y Klein (1989) y también Neurock *et al.* (1989), establecen una metodología para la construcción estocástica de estructuras moleculares de asfaltenos. Estos autores proponen una distribución de estructuras generadas estocásticamente para representar el componente asfáltico de un crudo particular. Los algoritmos de estos autores se fundamentan en distribuciones de tipo estadístico para establecer la probabilidad de existencia de un dado atributo molecular en la molécula a generar. Esta metodología basada en la distribución estadística de los datos de los parámetros moleculares ha sido aplicada por varios autores recientemente (De León *et al.*, 2017; Law *et al.*, 2019).

Es importante mencionar que estas distribuciones estadísticas se basan en la experiencia adquirida en los estudios de química analítica de una gran cantidad de tipos de crudos y sus componentes SARA, sin embargo, los detalles de estas correlaciones y sus formas funcionales específicas no están disponibles en la literatura, así como tampoco, los programas asociados con los cálculos de estos parámetros para generar las estructuras.

En este estudio se dan los detalles de los fundamentos y algoritmos necesarios para crear un programa de generación de estructuras moleculares promedios de sistemas orgánicos complejos constituidos por

<http://www.executivebs.org/publishing.cl/>

componentes aromáticos, resinas y asfaltenos, a partir de datos de parámetros moleculares promedios obtenidos de RMN ^1H , análisis elemental y pesos moleculares. Las estructuras moleculares son codificadas en la forma de una serie de atributos moleculares denominados índices moleculares (los cuales contienen información detallada a nivel molecular). También se obtuvieron un conjunto de expresiones matemáticas que relacionan los parámetros moleculares promedios de RMN ^1H con los índices moleculares. Es importante mencionar que, en este trabajo, no se están haciendo correlaciones estocásticas entre datos experimentales versus teóricos, sino que se están generando de manera estocástica, las estructuras moleculares promedios de asfaltenos a partir de datos de caracterización de química analítica de sistemas orgánicos complejos, mediante un algoritmo que los autores de este trabajo han desarrollado. El algoritmo en sí no es una correlación, sino que el mismo es un conjunto de ecuaciones deducidas a partir de la geometría molecular. Esto marca la diferencia entre el presente estudio con la mayoría de las propuestas que existen en la literatura.

Los algoritmos desarrollados en este estudio fueron codificados en lenguaje de programación FORTRAN-77 para generar el programa GENEMOL con el objetivo de aplicar el conjunto de ecuaciones desarrolladas en este trabajo, a la data de caracterización de asfaltenos de origen kuwaití que fueron obtenidos por Alí *et al.* (2006), y confrontar sus estructuras moleculares contra las estructuras propuestas en este trabajo. Es importante recalcar que la elección del trabajo experimental de Alí *et al.* (2006) fue realizado por conveniencia técnica, ya que proporciona de forma compacta datos de RMN ^1H , composición elemental y distribución de pesos moleculares directamente transferibles al algoritmo que se han propuesto en este estudio.

Es importante mencionar que la interpretación de los resultados obtenidos de las técnicas analíticas no es siempre trivial en el contexto de los asfaltenos. Por ejemplo, existe consenso en un amplio sector de la industria acerca de que en las condiciones de aplicación de la técnica de cromatografía por permeación de gel (GPC), utilizado por Alí *et al.* (2006), los asfaltenos se encuentran agregados en supra-estructuras no covalentes, por lo que los pesos moleculares observados corresponderían a agregados y no a moléculas de asfalto. Sin embargo, el objetivo de este trabajo no es la determinación exacta de la estructura molecular de los asfaltenos, sino que nuestro objetivo central es probar el algoritmo propuesto en este trabajo, a manera de una herramienta de apoyo a la investigación de las estructuras de los sistemas orgánicos complejos como lo son los asfaltenos. Obviamente, si la data de VPO es de mala calidad, entonces, la herramienta propuesta en este trabajo daría una estructura no realista. Por lo tanto, el objetivo de este trabajo es dar los algoritmos necesarios para que cualquier investigador construya su propio programa de generación de estructuras moleculares complejas.

Durante décadas ha existido una fuerte controversia sobre la estructura cualitativa (isla o archipiélago) de los asfaltenos (Alvarez & Ruiz 2013; Schuler *et al.*, 2015; Velásquez 2016; Niles *et al.*, 2020). Los autores han propuesto que la estructura de los asfaltenos puede ser desde tipo isla (solamente), o, estructura tipo archipiélago (solamente), o, hasta la posibilidad de la coexistencia de las estructuras tipo isla y archipiélago. Afortunadamente, el algoritmo propuesto es flexible, ya que, en sus premisas de construcción, no se asumió ninguna estructura a priori para los asfaltenos. El programa comienza con un número de unidades policondensadas aromáticas igual a 1, y a partir de este primer valor, el algoritmo intenta construir una estructura asfáltica tipo archipiélago. El algoritmo luego convierte la estructura molecular en parámetros de caracterización estructurales de RMN ^1H y peso molecular, y compara con los datos de caracterización experimental. Si al comparar ambos valores teóricos versus experimentales, hay una diferencia muy grande, entonces, el algoritmo vuelve a construir la estructura asfáltica usando dos unidades policondensadas aromáticas unidas por una cadena alifática (estructura tipo archipiélago).

Finalmente, el algoritmo que se propone en este trabajo genera una estructura molecular sin los heteroátomos, sin embargo, usando información de espectroscopía infrarroja junto con la información del análisis elemental, se puede incorporar los heteroátomos a la estructura generada por nuestro algoritmo. Nótese que el objetivo de este algoritmo es apoyar el trabajo del investigador en la elaboración de la estructura molecular de materiales complejos.

FUNDAMENTOS Y METODOS

En esta sección se va a resumir las premisas, los protocolos y algoritmos de la metodología propuesta en este estudio para la generación estocástica de estructuras moleculares.

Es conveniente definir conceptos de parámetros moleculares, atributos moleculares e índices moleculares. Un atributo molecular es un índice de información que da cuenta de una parte, o, de toda la estructura molecular de un sistema en estudio. Los parámetros moleculares son atributos moleculares experimentales, los cuales se correlacionan directamente con las técnicas de química analítica usadas para la caracterización de un sistema en estudio (ver Tabla 1), para una lista de estos parámetros derivados de RMN ¹H. El índice molecular es un atributo molecular que contiene información detallada a nivel de la molécula asociada al sistema en estudio, el cual no puede obtenerse directamente de la experimentación (vía química analítica), sin embargo, es posible establecer correlaciones tanto fenomenológicas (vía estadística) como de fundamento (deductivas vía estequiometría) con los parámetros moleculares.

Tabla 1: Lista de los parámetros moleculares obtenidos de RMN ¹H

Símbolo Índices Moleculares	Descripción
PC	Porcentaje en peso de carbono por molécula promedio
PH	Porcentaje en peso de hidrógeno por molécula promedio
MW	Peso Molecular por fórmula molecular teórica
HA	Corresponde al Índice Molecular harT
n	Número de átomos de carbono por sustitución alquímica
r	Número de anillos nafténicos por sustitución
F	Razón en peso C/H en grupos alquílicos p/p
Cs	Porcentaje en peso de carbonos saturados
Ca	Porcentaje en peso de carbonos aromáticos
fa	Factor de Aromaticidad
C1	Porcentaje en peso de carbonos aromáticos no puente
%AS	Porcentaje de sustitución en carbonos aromáticos no puente
#Ca	Número de carbonos aromáticos
#C1	Número de carbonos aromáticos no puente
#C1	Número de carbonos aromáticos puente
RA	Número de anillos aromáticos por molécula
#C	Número total de carbonos por molécula
#H	Número total de hidrógenos por molécula
RT	Número total de anillos por molécula
RS	Número de sustituyentes alquílicos por molécula
RN	Número de anillos nafténicos por molécula
CN	Porcentaje en peso de carbonos nafténicos
CN'	Porcentaje en peso de carbonos nafténicos, para la fracción de saturados
CP	Porcentaje en peso de carbonos parafínicos
#Cmet	Número de carbonos metílicos por molécula promedio
b	Número de ramificaciones por sustituyente alquílico
m	Fracción mínima de sustituciones metílicas (solo si m es mayor que cero)
Car,ar,ar	Número de carbonos aromáticos cuaternarios
h	Índice de condensación de anillos
fc	Factor de compactibilidad en anillos aromáticos

Protocolos, correlaciones y algoritmos

A continuación, se resumen los procedimientos y etapas de la metodología que genera los índices moleculares, así como una breve discusión de cada uno de ellos. En la Figura 1 se resume el protocolo del ensamblado de los distintos componentes que constituyen la molécula representativa de uno de los componentes SARA del crudo (específicamente el componente asphalténico). La descripción de las distintas partes de la Figura 1 será realizada a medida que se vayan desarrollando las distintas secciones de este artículo.

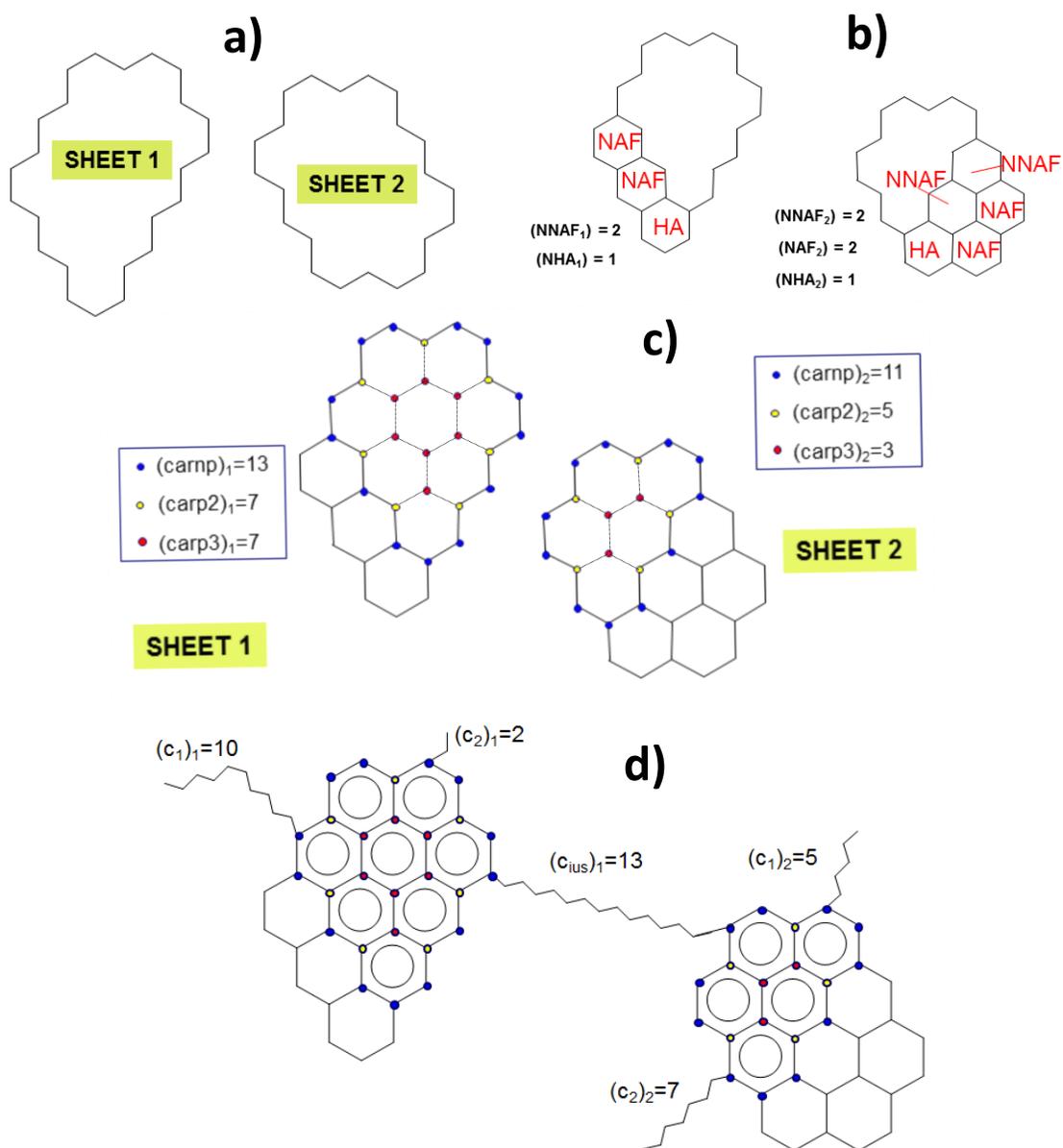


Fig. 1: Etapas propuestas en este estudio del ensamblado de una estructura molecular modelo. (a) Definir la cantidad de Láminas. (b) Definir el núcleo nafténico en cada Lámina. (c) Definir el núcleo aromático en cada Lámina. (d) Definir cadenas alifáticas terminales y puentes. Se ha generado la estructura molecular buscada.

Cálculo del número de Láminas (NSH)

La lámina es un contorno (inicialmente vacío) de trabajo donde se van colocando los distintos atributos moleculares, ver Figura 1a. El número de Láminas (NSH) se genera aleatoriamente, puede ser de 1 a 4. Es importante mencionar que el número óptimo de Láminas se obtiene al confrontar la estructura generada contra los datos analíticos, si la desviación estándar de los parámetros estructurales calculados de la molécula propuesta supera un valor prefijado, entonces, se vuelve a construir nuevamente, la estructura con un nuevo valor del NSH.

Generación del sistema nafténico

Para generar el conjunto de anillos saturados (denominados genéricamente como anillos nafténicos) se usa el criterio más simple de distribuir y organizar los anillos saturados en forma de un núcleo o substrato nafténico. En otras palabras, los anillos saturados no están aislados sino siempre están formando una unidad compacta denominada núcleo nafténico. Este criterio tiene soporte en la literatura (Speight, 1980; Siddiqui & Ali, 1999; Peters & Moldowan, 1993). Se definen tres tipos de anillos saturados dentro del núcleo nafténico, denominados anillos hidroaromáticos, anillos no nafténicos y anillos nafténicos. Los anillos hidroaromáticos (mostrados en la Figura 2a) están unidos a un anillo aromático, los cuáles no se encuentran aislados. Los anillos no nafténicos están unidos a dos o más anillos aromáticos (Ver Figura 2b), los cuáles no están aislados y no pueden flanquear el núcleo aromático. Los anillos nafténicos están unidos solamente a otros ciclos saturados (anillos hidroaromáticos, o, anillos no nafténicos) (Ver Figura 2c). Los índices NHA_i , $NNAF_i$ y NAF_i definen el número de anillos hidroaromáticos, no nafténicos y nafténicos respectivamente. Estos índices son generados aleatoriamente. El subíndice "i" está asociado a una Lámina i donde $i=1, 2, 3$ y 4. Una vez establecido el número de anillos saturados de cada tipo, entonces, se ha generado el núcleo nafténico, ver Figura 1b. El siguiente paso es calcular el número de átomos de carbono y de hidrógeno que existe en dicho núcleo, lo cual es necesario para determinar el balance de masa final. En este punto, hemos definido en este estudio, los índices moleculares $califp_i$ (número de átomos de carbono alifáticos cíclicos no puente), $califp2_i$ (número de átomos de carbono alifáticos cíclicos puente tipo 2), $califp3_i$ (número de átomos de carbono alifáticos cíclicos puente tipo 3). Las correspondientes correlaciones obtenidas en el presente estudio son las siguientes:

$$califp_i = 1.29762 * NHA_i + 1.2789 * NAF_i + 1.4057 * NNAF_i + 1.8168 \quad (1)$$

$$califp2_i = 0.31603 * NHA_i + 1.1887 * NAF_i + 0.82161 * NNAF_i - 0.3761 \quad (2)$$

$$califp3_i = 0.68397 * NHA_i + 0.8113 * NAF_i + 0.17839 * NNAF_i - 0.6239 \quad (3)$$

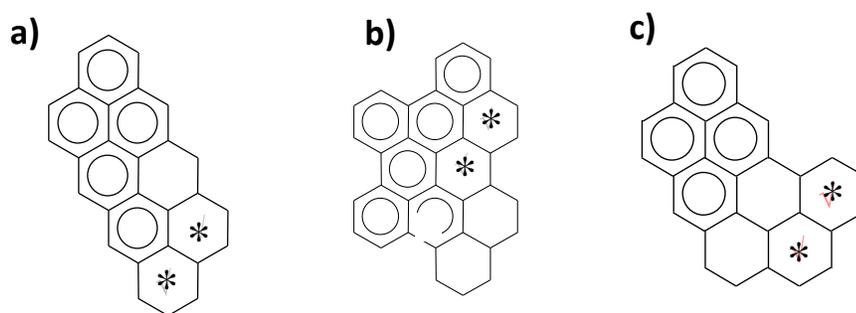


Fig. 2: Definición de los distintos anillos saturados presentes en un Sheet (marcados con "*"). (a) Anillos hidroaromáticos. (b) Anillos No Nafténicos. (c) Anillos Nafténicos.

Las Figuras 3a, 3b y 3c definen los átomos alifáticos cíclicos no puente, y puentes 2 y 3. Es conveniente ver que cuando se genera al azar los valores de estos índices (NHA_i , $NNAF_i$ y NAF_i), no necesariamente, estos números son congruentes desde el punto de vista químico, por lo que es necesario realizar una revisión de la consistencia química de estos valores mediante la siguiente relación propuesta por Savage y Klein (1989):

$$califp2_i + califp3_i = 2 * NAF_i + NHA_i + NNAF_i - 1 \quad (4)$$

Por lo tanto, los números generados para los índices NHA_i , $NNAF_i$, NAF_i , $califp2_i$ y $califp3_i$, deben satisfacer dentro de una determinada tolerancia la ec.(4), de lo contrario, se descarta el proceso realizado hasta este momento, y se procede a repetir todos los pasos que implica generar nuevos valores de los índices NHA_i , $NNAF_i$, NAF_i , $califp2_i$ y $califp3_i$.

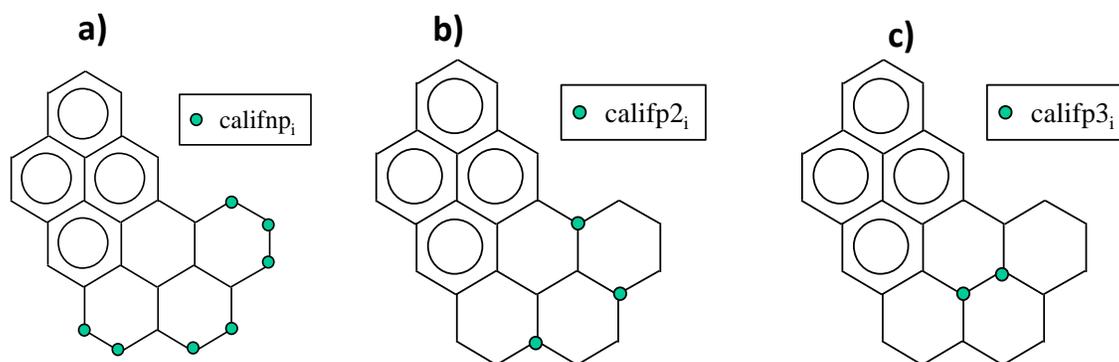


Fig. 3: Definición de los distintos tipos de carbonos asociados al núcleo de anillos saturados. (a) Átomos de carbono alifáticos cíclicos no puente. (b) Átomos de carbono alifáticos cíclicos puente tipo 2. (c) Átomos de carbono alifáticos cíclicos puente tipo 3.

Generación del sistema policondensado aromático

Una de las contribuciones que aporta este trabajo es la obtención de correlaciones fenomenológicas de tipo estadístico entre los distintos índices moleculares, lo cual reduce drásticamente la exploración aleatoria todos los índices moleculares. Por ejemplo, se genera el sistema policondensado aromático mediante los índices: número de carbonos aromáticos no puentes ($carnp$), y número de carbonos aromáticos puentes tipo 2 ($carp2$) y tipo 3 ($carp3$) (tipo 2 significa que el átomo de carbono hace de puente entre dos anillos aromáticos y tipo 3 hace puente entre tres anillos aromáticos). En este estudio se encontraron las siguientes relaciones entre estos índices $carnp$ y $carp2$:

$$carp2_i = carnp_i - 6 \quad (5)$$

$$fc_{i1} = \frac{carp2_i + carp3_i}{carnp_i + carp2_i + carp3_i} \quad (6)$$

$$fc_{iC} = 3.4 \times 10^{-2} carnp_i - 2.0 \times 10^{-2} carp2_i + 2.8 \times 10^{-2} carp3_i \quad (7)$$

Nótese que la molécula a construir puede tener varias Láminas, y por lo tanto, el subíndice "i" en estas ecuaciones, está asociado a una dada Lámina (Lámina $i=1$, Lámina $i=2$, etc.). El índice $carnp$ es generado aleatoriamente, el índice $carp2$ es generado a partir de la ec.(5) (esta ecuación es una relación exacta o matemática). La ec.(6) es también una relación matemática exacta encontrada en este estudio, que permite conectar el parámetro $fc1$ con los con los parámetros $carnp$, $carp2$ y $carp3$. Es importante mencionar que el índice de compacticidad definido originalmente por Williams (1958), se simboliza como fc . La razón por la cual usamos el símbolo fc_1 en la ec.(6) es para recalcar que fc es experimental, y fc_1 es teórico, y en función de índices moleculares. La ec.(7) es otra manera de expresar la compacticidad, pero, en este caso, la ec.(7) es una correlación estadística (factor $R^2 = 0.96$) obtenida en este estudio de una muestra de 157 estructuras aromáticas diferentes, con un número de anillos aromáticos que varían entre 1 y 10, y con un número de

carbonos no puente que varía entre 6 y 24. Por conveniencia de programación hemos definido para el caso de la ec.(7), usar el símbolo fc_c para diferenciarlo de la ec.(6).

El índice $carp3$ sólo puede tener los valores 0, 2, 4, 6 hasta 8. Estos posibles valores de $carp3$ son introducidos en las ecuaciones (6) y (7) hasta que hay un valor donde fc_1 y fc_c se hacen iguales dentro de una tolerancia previamente definida. Por lo tanto, las ecuaciones (6) y (7) hacen el papel de verificadores de la consistencia de los valores de los parámetros $carnp$, $carp2$ y $carp3$. Es importante recalcar que el valor calculado del parámetro fc_1 se confronta con el valor del parámetro molecular experimental fc . Nótese que los parámetros moleculares calculados tienen el mismo símbolo que los parámetros experimentales, sólo difieren en el índice "1", para diferenciarlos (fc_1 versus fc). Una vez obtenidos los valores de $carnp$, $carp2$ y $carp3$, es necesario calcular el número de anillos aromáticos por cada Lámina, NRA_i , que sea compatible con estos valores. En este estudio se obtuvo la siguiente correlación estadística de una muestra de estructuras:

$$NRA_i = 1.7 \times 10^{-1} carnp_i + 3.3 \times 10^{-1} carp2_i + 5.0 \times 10^{-1} carp3_i \quad (8)$$

Por lo tanto, al tener los valores de los índices moleculares NRA_i , $carnp$, $carp2$ y $carp$, entonces, se ha definido el núcleo aromático, ver Figura 1c.

Generación de las cadenas alifáticas terminales y puentes

Es interesante observar que en una Lámina hay un conjunto de átomos de carbono que son periféricos (estos pueden ser aromáticos o no), además, estos átomos pueden estar unidos o no a átomos de hidrógeno. En este estudio se definen los índices $NSTA_i$, $NSIUA$, $c_{i,j}$ y $cius_k$ como el número de sustituciones terminales, el número de sustituciones puentes, el número de átomos de carbono de las cadenas alifáticas terminales (el subíndice j está asociado a la cadena alifática terminal de la Lámina i) y puentes (el subíndice k está asociado a la cadena puente), respectivamente. Los índices $NSTA_i$, $c_{i,j}$ y $cius_k$ son generados aleatoriamente, mientras que $NSIUA$ se calcula mediante la relación:

$$NSIUA = NSH - 1 \quad (9)$$

En este estudio se asume que las sustituciones alifáticas (terminales y puentes) son, en primera aproximación, distribuidas únicamente sobre el núcleo aromático (siguiendo el criterio más simple para construcción de la molécula). Finalmente, hemos construido la estructura molecular buscada, ver Figura 1d.

Conteo de los átomos de hidrógeno asociados a los átomos de carbono

Una vez generados los índices moleculares previamente expuestos, se tiene el esqueleto básico de la molécula que sólo considera los átomos de carbono, pero se necesita incorporar los átomos de hidrógeno. Los átomos de hidrógeno se incorporan a cada uno de los átomos de carbono de la molécula, a través, de un balance estequiométrico. El número total de átomos de hidrógeno de la molécula viene dado por:

$$h_{Total} = halifct + hT + hiusT \quad (10)$$

Donde $halifct$, hT y $hiusT$ son el número de átomos de hidrógeno de los núcleos nafténicos de todas las Láminas, el número de átomos de hidrógeno alifáticos terminales de todas las Láminas y el número de átomos de hidrógeno de las cadenas alifáticas puente de la molécula, respectivamente. Estos índices se pueden calcular a través de las siguientes relaciones:

$$halifcT = \sum_{i=1}^{NSH} halifc(i) \quad (11)$$

$$hT = \sum_{i=1}^{NSH} \sum_{j=1}^{NSTA(i)} h(j,i) \quad (12)$$

$$hiusT = \sum_{k=1}^{NSIUA} hius(k) \quad (13)$$

Donde $halifc(i)$, $h(j,i)$ y $hius(k)$ son el número de átomos de hidrógeno alifáticos cíclicos asociados a la Lámina i , el número de átomos de hidrógeno alifáticos terminales asociados a la Lámina i y a la cadena alifática terminal j , y el número de átomos de hidrógeno de las cadenas alifáticas puente de la molécula, respectivamente. El índice $halifc(i)$ se calcula por medio de las siguientes relaciones estequiométricas:

$$halifc(i) = halifnp(i) + halifp2(i) + halifp3(i) \quad (14)$$

$$halifnp(i) = 2 * califnp(i) \quad (15)$$

$$halifp2(i) = 2 * califp2(i) \quad (16)$$

$$halifp3(i) = califp3(i) \quad (17)$$

Donde $halifnp(i)$, $halifp2(i)$, y $halifp3(i)$ (hidrógenos asociados a los sistemas cíclicos), son el número de átomos de hidrógeno alifáticos no puente asociados a la Lámina i , el número de átomos de hidrógeno alifáticos puente tipo 2 (une a 2 anillos saturados) asociados a la Lámina i , el número de átomos de hidrógeno alifáticos puente tipo 3 (une a 3 anillos saturados) asociados a la Lámina i , respectivamente. Los índices $h(j,i)$ y $hius(k)$ se calculan mediante las siguientes relaciones estequiométricas:

$$h(j,i) = 2 * c(j,i) + 1 \quad (18)$$

$$hius(k) = 2 * cius(k) \quad (19)$$

Chequeo de la estructura molecular por medio del peso molecular

Una vez generados todos los átomos de carbono e hidrógeno asociados a los distintos fragmentos de la molécula a través de los procedimientos descritos previamente, el siguiente paso es generar la fórmula molecular de la estructura propuesta y posteriormente, calcular el peso molecular. Este valor calculado del peso molecular se compara con el peso molecular experimental reportado para la fracción en estudio. Si existe una diferencia mayor de una tolerancia específica, entonces, se descarta la estructura y se vuelve a construir una nueva molécula.

Chequeo de la estructura molecular por medio del peso molecular

Una vez generados todos los átomos de carbono e hidrógeno asociados a los distintos fragmentos de la molécula a través de los procedimientos descritos previamente, el siguiente paso es generar la fórmula molecular de la estructura propuesta y posteriormente, calcular el peso molecular. Este valor calculado del peso molecular se compara con el peso molecular experimental reportado para la fracción en estudio. Si existe una diferencia mayor de una tolerancia específica, entonces, se descarta la estructura y se vuelve a construir una nueva molécula.

Cálculo de los parámetros moleculares promedios

A continuación, se listan las expresiones que conectan los parámetros moleculares promedios con los índices moleculares definidos previamente y que fueron resumidos en la sección anterior (Protocolos, correlaciones y algoritmos). Es importante mencionar que estas correlaciones son un aporte original de este trabajo y que permite establecer la estructura molecular detallada de los sistemas orgánicos en estudio.

En este estudio se define el parámetro teórico número de átomos de carbono por sustitución alquílica de la siguiente manera:

$$n_1 = \frac{cT + ciusT + califcT}{NSTAT + NSIUA + NSHAT + NSNNAFT} \quad (20)$$

Donde NSHAT y NSNNAFT son el número total de sustituciones hidroaromáticas de la molécula promedio, y el número total de sustituciones no nafténicas de la molécula promedio, respectivamente. En la Figura 4 se definen las sustituciones hidroaromáticas (las que provienen de anillos hidroaromáticos) y las no nafténicas (las que provienen de anillos no nafténicos). Se consideran dos sustituciones hidroaromáticas por anillo hidroaromático y una sustitución no nafténica por anillo no nafténico. Es conveniente mencionar que la ec.(20) define al parámetro número de átomos de carbono por sustitución alquílica en términos de los índice moleculares definidos en este estudio, y que lo hemos simbolizado como n1 para diferenciarlo de su respectivo experimental (se simboliza como n), el cual se calcula como (Williams, 1958):

$$n = \frac{Halfa + Hbeta + Hgamma}{Halfa} \quad (21)$$

Nótese que el parámetro molecular experimental n se calcula en función de parámetros derivados directamente de los datos de RMN ¹H, mientras que el parámetro calculado o teórico n1 se obtiene en función de los índices moleculares definidos en este estudio. La razón por la cual se han generado los índices moleculares es para facilitar la generación de la estructura molecular, ya que los índices moleculares contienen información detallada de cada una de las partes de la estructura molecular y permite un ensamblado más sistemático. Además, los índices moleculares se pueden codificar de manera más eficiente en cualquier lenguaje de computación para convertirlos en estructura molecular.

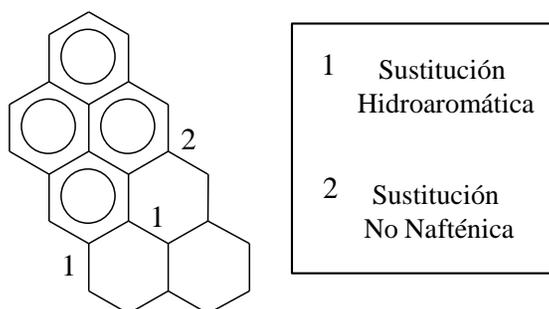


Fig. 4: Sustituciones hidroaromáticas y no nafténicas del núcleo aromático.

A continuación, se definen el resto de los parámetros moleculares teóricos obtenidos en este estudio.

Número de anillos nafténicos por sustituciones totales:

$$r_1 = \frac{NHAT + NAFT + NNAFT}{(NSTAT + NSIUA + NSHAT + NSNNAFT)} \quad (22)$$

Razón en peso C/H en grupos alquílicos (F1):

$$F_1 = \frac{(cT + ciusT + califcT) * 12}{halifct + hT + hiusT} \quad (23)$$

Donde c_T , c_{iust} y $califc_T$ son el número de átomos de carbono de las cadenas terminales alifáticas de la molécula promedio, el número de átomos de carbono de las cadenas alifáticas puente de la molécula promedio, y el número de átomos de carbono alifáticos cíclicos de la molécula promedio, respectivamente.

Porcentaje en peso de carbonos saturados (CS1):

$$C_{s1} = \frac{califT * 12}{MW} * 100 \quad (24)$$

Donde $califT$ y MW son el número de átomos de carbonos alifáticos de la molécula promedio, y el peso molecular de la molécula promedio, respectivamente.

Porcentaje en peso de carbonos aromáticos (CA1):

$$C_{A1} = \frac{carT * 12}{MW} * 100 \quad (25)$$

Donde $carT$ es el número de átomos de carbono aromáticos de la molécula promedio.

Aromaticidad (fa_1):

$$fa_1 = \frac{carT}{cTOT} \quad (26)$$

Donde $cTOT$ es el número de átomos de carbono totales de la molécula promedio.

Porcentaje en peso de carbonos aromáticos no puente (C11):

$$C_{11} = \frac{carnpT * 12}{MW} * 100 \quad (27)$$

Donde $carnpT$ es el número de átomos de carbono aromáticos no puente de la molécula promedio.

Porcentaje de sustitución en carbonos aromáticos no puente (%AS1):

$$\%AS_1 = \frac{NSTAT + 2 * NSIUA + NSHAT + NSNNAFT}{carnpT} * 100 \quad (28)$$

Número de carbonos aromáticos (#CA1):

$$\#C_{A1} = carT \quad (29)$$

Donde $carT$ es el número de átomos de carbono aromáticos de la molécula promedio.

Número de carbonos aromáticos no puente (#C11):

$$\#C_{11} = carnpT \quad (30)$$

Número de anillos aromáticos por molécula promedio (RA1):

$$R_{A1} = NRAT \quad (31)$$

Donde NRAT es el número de anillos aromáticos por molécula promedio.

Número total de anillos por molécula promedio (RT1):

$$R_{T1} = NRAT \quad (32)$$

Número de sustituyentes alquílicos por molécula promedio (RS1):

$$R_{S1} = NSTAT + NSIUA + NSHAT + NSNNAFT \quad (33)$$

Número de anillos nafténicos por molécula promedio (RN1):

$$R_{N1} = NHAT + NAFT + NNAFT \quad (34)$$

Factor de compacticidad en anillos aromáticos (fc1i):

$$fc_{1i} = \frac{carp2_i + carp3_i}{carp_i + carp2_i + carp3_i} \quad (35)$$

RESULTADOS Y DISCUSIÓN

Con la finalidad de aplicar la metodología presentada en este estudio, se procedió a codificar en Fortran-77 los algoritmos descritos previamente, y se ha generado un programa que se bautizó como GENEMOL. El programa GENEMOL usa como datos de entrada los resultados experimentales de caracterización de asfaltenos y resinas determinados por espectroscopia RMN ^1H , análisis elemental y peso molecular, reportados en la literatura para crudos de diferentes orígenes. Los asfaltenos estudiados en este trabajo son provenientes de crudos kuwaitíes analizados por Ali *et al.* (2006). Estos autores reportaron con abundancia de detalles los resultados de la caracterización de los asfaltenos y las resinas de estos crudos por las diferentes técnicas experimentales antes mencionadas. Los datos de caracterización se encuentran reportados en las Tabla 2. Además, estos autores proponen estructuras promedios para los asfaltenos y resinas de los crudos estudiados, esto permite confrontar las estructuras generadas por GENEMOL.

En el trabajo reportado por Ali *et al.* (2006), se caracterizan asfaltenos de un crudo de origen kuwaití. Estos asfaltenos son provenientes de dos tipos de residuos de destilación, uno atmosférico (AR) y otro de vacío (VR). Posteriormente, estos asfaltenos son fraccionados en once subfracciones de acuerdo a su peso molecular, utilizando para tal fin, la técnica de Gel-Permeation Chromatography (GPC) con cloroformo como fase móvil. Todas las once subfracciones fueron caracterizadas por peso molecular promedio, análisis elemental y técnicas de RMN (^1H y ^{13}C). También, se usó difracción de rayos X pero solo a unas subfracciones seleccionadas con la finalidad de derivar los parámetros estructurales promedios. Específicamente, nosotros hemos seleccionado la caracterización de la subfracción N°4 de asfaltenos, la cual contenía información de difracción de rayos X (XRD), que ayudó a los autores Ali *et al.* (2006), a definir mejor sus propuestas de estructuras moleculares de los asfaltenos AR y VR. Los valores reportados para los análisis de peso molecular por GPC, análisis elemental y RMN ^1H se encuentran en la Tabla 2 para los asfaltenos AR y VR de la subfracción N°4. Es importante mencionar que en el presente trabajo sólo se estudian los residuos AR y subfracción N°4. Los datos de la caracterización de AR de la Tabla 2 fueron introducidos como input en el programa GENEMOL con el objetivo de validar sus resultados.

Tabla 2: Caracterización Analítica de la subfracción N° 4 de asfalteno de AR y VR del crudo kuwaití (Alí *et al.*, 2006).

Propiedades	Asfalteno AR	Asfalteno VR
	Valores reportados	Valores reportados
%p/p de C	83.76	86.22
%p/p de H	7.94	8.23
%p/p de N	0.87	0.97
%p/p de S	7.43	4.58
%p/p de O	0.0	0.0
PM (GPC)	5260	5360
H _α	76.9	111.3
H _β	225.9	224.6
H _γ	91.8	76.8
H _A	23.0	28.3
Fórmula Molecular	[C ₃₆₉ H ₄₂₀]	[C ₃₇₆ H ₄₅₀]

Se ha corrido el programa GENEMOL con el máximo rango de exploración de las variables aleatorias a fin de que el programa tenga suficientes grados de libertad para encontrar la estructura con la mejor correspondencia con la data analítica de caracterización. Se destaca que las variables aleatorias como el número de carbonos de las cadenas alifáticas tipo puente pueden variar en un rango de 3 a 10, y los carbonos aromáticos puente varían de 6 a 24 (1 a 10 anillos aromáticos).

En la Figura 6 se presenta la estructura propuestas en este estudio del AR obtenida con el programa GENEMOL. Estos resultados son comparados con la estructura molecular del AR propuesta por Alí *et al.* (2006) (Ver figura 5). La estructura generada por GENEMOL, es similar a la propuesta por dichos autores.

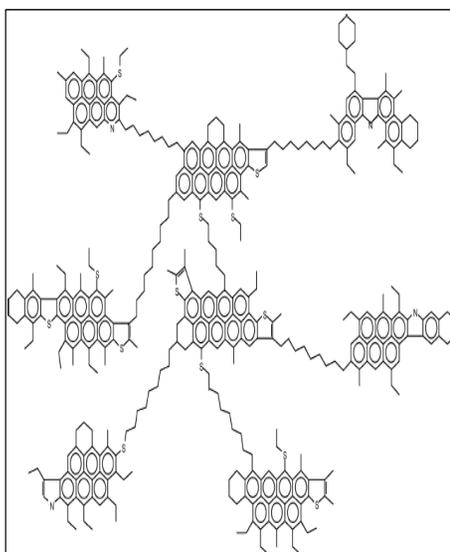


Fig. 5: Estructura molecular para la subfracción N° 4 de asfaltenos del residuo atmosférico (AR) del crudo kuwaití propuesta por Alí *et al.* (2006).

Por otro lado, en la Tabla 3 se muestran los valores de los parámetros moleculares de RMN ¹H experimentales y los parámetros moleculares calculados para la estructura propuesta en este estudio (ver Figura 6). Existe una coincidencia entre ambos valores experimentales vs calculados cuya desviación promedio es igual a 23%. Este resultado nos dice que la idea de representar por una sola molécula al sistema del AR no es del todo herrada, ya

que la desviación promedio tiene un valor sorprendentemente pequeño (tomando en cuenta la complejidad del sistema AR). Por lo tanto, sería interesante (a manera de sugerencia para futuros trabajos) proponer una distribución de estructuras moleculares en lugar de una sola estructura.

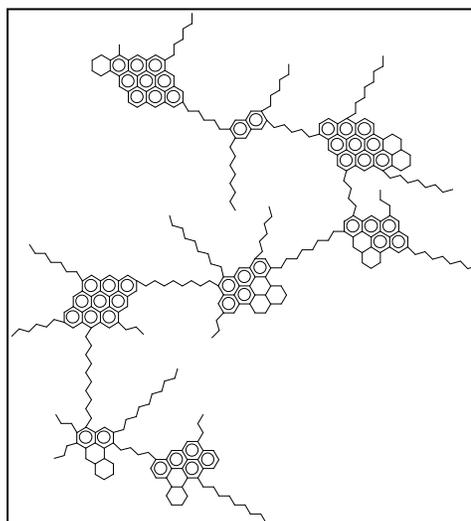


Fig. 6: Estructura molecular generada por GENEMOL de la subfracción N° 4 del asfaltenos del residuo atmosférico (AR) del crudo kuwaití.

Es importante mencionar que la estructura propuesta en este trabajo no presenta la incorporación de heteroátomos, ya que estos se incorporan manualmente, previa identificación de los grupos funcionales presentes.

Tabla 3: Comparación de los parámetros moleculares de RMN ¹H experimentales versus los calculados para la estructura molecular propuesta en este estudio (ver Figura 6) de la subfracción N° 4 de asfalteno AR del crudo kuwaití (Alí *et al.*, 2006).

PARAMETRO	EXPERIMENTAL	CALCULADO	DESV(%)
n	5.1	5.1	0.6
r	0.3	0.4	61.5
F	5.7	6	4.7
CS	43.1	47.7	10.6
CA	40.7	32.9	19.2
fa	0.5	0.4	16
C1	13.7	17.8	30.3
PAS	61.5	56.4	8.3
NCA	178.2	144	19.2
NC1	59.8	78	30.3
RA	60.2	37	38.5
NC	367.1	353	3.9
NH	417.6	445	6.6
RS	36.8	41	11.3
RN1	10	18	79.8
fc	0.7	0.5	30.2
Promedio			23

CONCLUSIONES

En este estudio se dan los detalles de los fundamentos y algoritmos necesarios para crear un programa de generación de estructuras moleculares promedios de sistemas orgánicos complejos constituidos por componentes aromáticos, resinas y asfaltenos, a partir de datos de parámetros moleculares promedios obtenidos de RMN ^1H , análisis elemental y pesos moleculares. Estos algoritmos se codificaron en lenguaje Fortran 77 para generar un programa denominado GENEMOL para la generación automática de estructuras moleculares de las fracciones de aromáticos, resinas y asfaltenos de un crudo, residuo de vacío, etc. A partir de los datos de RMN ^1H , análisis elemental y peso molecular reportados en la literatura es posible proponer estructuras hipotéticas para las fracciones de asfaltenos y resinas. En general, GENEMOL reproduce de manera satisfactoria las moléculas modelo de asfaltenos y resinas reportadas en la literatura. La diferencia promedio entre los parámetros moleculares experimentales y los calculados a partir de los índices moleculares son del orden de 23%.

Las recomendaciones que se derivan de este trabajo son las siguientes. La posibilidad de generar distribuciones de estructuras modelo para los materiales orgánicos en estudio. Incorporar datos de IR para identificar grupos funcionales con el objetivo de ubicar los heteroátomos en la estructura molecular modelo.

AGRADECIMIENTOS

A las instituciones a las que pertenecen los autores y sus respectivas Direcciones de Investigación por el apoyo otorgado para la realización de la presente investigación.

REFERENCIAS

- Alhreez, M. & Wen, D. (2019). Molecular structure characterization of asphaltene in the presence of inhibitors with nanoemulsions. *RSC Advances*, *34*, 19560-19570. Recuperado de internet: <https://pubs.rsc.org/en/content/articlelanding/2019/ra/c9ra02664a#!divAbstract>
- Ali, F., Khan, Z. & Ghaloum, N. (2004). Structural studies of vacuum gas oil distillate fractions of Kuwaiti crude oil by nuclear magnetic resonance. *Energy & Fuel*, *18*, 1798–1805.
- Ali, F., Ghaloum, N. & Hauser, A. (2006). Structure representation of asphaltene GPC fractions derived from Kuwaiti residual oils. *Energy & Fuels*, *20*, 231-238.
- Alvarez, F. & Ruiz, Y. (2013). Island versus Archipelago Architecture for Asphaltenes: Polycyclic Aromatic Hydrocarbon Dimer Theoretical Studies. *Energy & Fuels*, *27*, 1791–1808.
- Artok, L., Su, Y., Hirose, Y., Hosowaka, M., Murata, S. & Nomura, M. (1999). Structure and reactivity of petroleum – derived asphaltene. *Energy & Fuel*, *13*, 287- 296.
- Allen, D., Petrakis L., Grandy D., Gavalas, G. & Gates, B. (1984). Determination of functional groups of coal-derived liquids by NMR and elemental analysis. *Fuel*, *63*, 803-809.
- Bansal, V., Patel, M. & Sarpal, A. (2004). Structural aspects of crude oil derived asphaltenes by NMR and XRD and spectroscopic techniques. *Petroleum Science and Technology*, *22*, 1401-1426.
- Christopher, J., Sarpal, A., Kapur, G., Krishna, A., Tyagi, B., Jain, M. *et al.* (1995). Chemical structure of bitumen – derived asphaltenes by nuclear magnetic resonance spectroscopy and X – ray diffractometry. *Fuel*, *75* (8), 999 - 1008.
- Coelho, R., Hovell, I., Mello Monte, M., Middea, A. & Lopes, A. (2006). Characterisation of aliphatic chains in vacuum residues (VRs) of asphaltenes and resins using molecular modelling and FTIR techniques. *Fuel Processing Technology*, *87*, 325 – 333.

- De León, J., Velásquez, A. & Hoyos, B. (2017). A stochastic method for asphaltene structure formulation from experimental data: avoidance of implausible structures. *Phys. Chem. Chem. Phys.*, *19*, 9934-9944.
- Dickinson, E. (1980). Structural comparison of petroleum fractions using proton and ¹³C NMR. Spectroscopy. *Fuel*, *59*, 290-294.
- Law, J., Headen, T., Jiménez-Serratos, G., Boek, E., Murgich, J. & Müller, E. (2019). Catalogue of Plausible Molecular Models for the Molecular Dynamics of Asphaltenes and Resins Obtained from Quantitative Molecular Representation. *Energy & Fuels*, *33*, 10, 9779-9795.
- Liu, Y.-J., & Li, Z.-F. (2015). Structural Characterisation of Asphaltenes during Residue Hydrotreatment with Light Cycle Oil as an Additive. *Journal of Chemistry*, 1-8. Recuperado de internet: <https://www.hindawi.com/journals/jchem/2015/580950/>
- Michon, L., Martin, D., Planche, J. & Hanquet, B. (1996). Structural comparison of petroleum fractions using proton and ¹³C NMR spectroscopy. *Fuel*, *76*, 10-15.
- Niles, S., Chacón-Patiño, M., Smith, D., Rodgers, R. & Marshall, A. (2020). Comprehensive Compositional and Structural Comparison of Coal and Petroleum Asphaltenes Based on Extrography Fractionation Coupled with Fourier Transform Ion Cyclotron Resonance MS and MS/MS Analysis. *Energy & Fuels*, *34*, 2, 1492-1505.
- Neurock, M., Libanati, C. & Klein, M.T. (1989). *Modelling asphaltene reaction pathways: Intrinsic Chemistry*. *AICHE Symposium series*, *85* (273), 7-14.
- Peters, K. & Moldowan, M. (1993). *The Biomarker Guide*, Prentice Hall Inc.
- Savage, P. & Klein, M. (1989). Asphaltene reaction pathways. Chemical and Mathematical Modeling. *Chemical Engineering Science*, *44* (2), 393-404.
- Schuler, B., Meyer, G., Peña, D., Mullins, O. & Gross, L. (2015). Unraveling the Molecular Structures of Asphaltenes by Atomic Force Microscopy. *J. Am. Chem. Soc.*, *137*, 9870-9876.
- Siddiqui, M. & Ali, M. (1999). Investigation of chemical transformations by NMR and GPC during aging of arabian asphalt. *Fuel*, *78*, 1407-1416.
- Speight, J. (1980). *The chemistry and technology of petroleum*. Sección 8.4.2.1. 4ta Edición: Marcel Dekker, Inc., USA.
- Speight, J. & Moschopedis, S. (1981). On the molecular nature of petroleum asphaltenes, in chemistry of asphaltenes. *American Chemical Society*, *195*, 1-15.
- Velásquez, A. (2016). *Simulación Molecular del Efecto de la Agregación de Asfaltenos en la Viscosidad de Sistemas Asfalteno-Solvente*. Tesis de Maestría. Universidad Nacional de Colombia. Sede Medellín, Facultad de Minas, Departamento de Procesos y Energía. <http://bdigital.unal.edu.co/52828/>
- Williams, R. (1958). *Characterization of hydrocarbons in petroleum by nuclear magnetic resonance spectrometry*. London: University Microfilms International.
- Yen, T.F., Wu, W.H. & Chilingar, G.V. (1984). A study of the structure of petroleum asphaltenes and related substances by infrared spectroscopy. *Energy Sources*, *7* (3), 203-235.