



## VISUALIZACIÓN DE MODELOS MOLECULARES CON SMILES™

*ChemSketch*© Y Jmol

Ángel Lucas de la Cruz, Col·legi Sant Gabriel de Viladecans,

angel.lucas@gmail.com

Carlos Giménez Esteban, Col·legi Sant Gabriel de Viladecans,

carlos.gimenez@gmail.com

**Palabras clave:** TIC, modelos moleculares, SMILES

**Resumen:** La presencia masiva de las nuevas tecnologías (TIC) en nuestras aulas representa un reto para el profesorado, dado que la dinámica del aula se ve necesariamente modificada, especialmente por el hecho de que el nuevo escenario supone una relativa autonomía del alumno en el acceso a la información y, por tanto, un inevitable cambio de rol del profesor.

En esta comunicación, presentamos una secuencia didáctica apoyada en el uso de software gratuito (SMILES™ – *ChemSketch*© – Jmol) que permite generar y manipular de forma sencilla modelos tridimensionales virtuales de cualquier molécula, con el objetivo de facilitar la comprensión por parte de los estudiantes de algunos aspectos de química estructural.

### Objetivos

Los objetivos que persigue la experiencia de aula descrita en esta comunicación son:

- Dar a conocer las normas básicas de escritura del lenguaje químico SMILES™, caracterizado por su gramática sencilla, compacta y comprensible.
- Enumerar algunos recursos *online* del ámbito de la educación y de la industria química, donde se utiliza de manera habitual el lenguaje químico universal SMILES™.
- Presentar herramientas de *software* gratuito que permiten generar y manipular modelos moleculares, así como describir las opciones que deben utilizarse en cada una de ellas para obtener y manipular la representación tridimensional de cualquier molécula.
- Proponer una secuencia didáctica, experimentada con nuestros alumnos de las materias de química y biología de primero de bachillerato, para aplicar estas herramientas y, por tanto, para mejorar la comprensión de estos de determinados aspectos de química estructural.

### Descripción de la experiencia

#### Normas básicas de escritura

SMILES™ significa, en inglés, *Simplified Molecular Input Line Entry System* (sistema simplificado de entrada en línea de moléculas).

## Difusión e impacto de la Química en nuestra sociedad

A finales de los años ochenta David Weininger comenzó el desarrollo de este lenguaje. Para representar una molécula utilizó un gráfico con nudos en lugar de átomos y aristas en lugar de enlaces. En la actualidad SMILES™ está considerado un lenguaje químico muy sencillo, en el cual las moléculas y reacciones pueden ser expresadas mediante caracteres ASCII que representan enlaces y átomos.

SMILES™ aporta algunas ventajas con respecto a la información que nos dan otros lenguajes. Una línea escrita en SMILES™ se comprende muy con mucha facilidad, es muy compacta y muestra de forma intuitiva un patrón que representa un identificador universal para una estructura química específica.

Algunos ejemplos sencillos de SMILES™	
NOMBRE	SMILES
Etanol	CCO
Ácido acético	CC(=O)O
Ciclohexano	C1CCCCC1
Cloruro de sodio	[Na+].[Cl-]

La notación de SMILES consiste en una serie de caracteres que no contienen espacios entre ellos. Los átomos de hidrógeno pueden ser omitidos. Los paréntesis son utilizados para indicar puntos de bifurcación y etiquetas numéricas que designan localizaciones específicas de funciones. La gramática básica de SMILES™ también incluye información isotópica, la configuración de enlaces dobles y el mundo de los estereoisómeros.

Se pueden establecer 4 categorías diferentes: átomos, enlaces, radicales y estructuras cíclicas.

### *Átomos:*

- Los átomos se representan con sus símbolos atómicos.
- Cada elemento tenemos que especificarlo con su símbolo atómico (si hace falta segunda letra, ésta ha de ser una minúscula) entre corchetes, [ ], con la excepción del hidrógeno.
- Los elementos más habituales de la Química Orgánica: B, C, N, O, P, S, F, Cl, Br y I se pueden escribir sin corchetes cuando trabajamos con su valencia más pequeña: B (3), C (4), N (3,5), O (2), P (3,5), S (2,4,6) y 1 para los halógenos.

#### IV Jornadas sobre la Enseñanza de la Química



Se consideran válidas las siguientes notaciones

SMILES	NOM	FÓRMULA	SMILES	NOM	FÓRMULA
C	metano	CH <sub>4</sub>	N	amoniaco	NH <sub>3</sub>
P	fosfina	PH <sub>3</sub>	O	agua	H <sub>2</sub> O

Los átomos que actúan con otras valencias diferentes a las indicadas con anterioridad, y los elementos que habitualmente no forman parte de la Química Orgánica tenemos que escribirlos siempre entre corchetes: [Au].

- Los átomos que pertenecen a anillos aromáticos se escriben en minúsculas.
- Cualquier hidrógeno enlazado y también las cargas formales se deben indicar siempre mediante corchetes. El número de hidrógenos se señala con el símbolo H seguido por un dígito (opcional). Igualmente, una carga formal se marca con los signos + o -, seguidos por un dígito (opcional).
- Algunos ejemplos son: [H<sup>+</sup>] (protón), [Fe<sup>+++</sup>] o [Fe+3] (catión hierro (III)), [OH<sup>-</sup>] (anión hidroxilo) y [NH<sub>4</sub><sup>+</sup>] (catión amonio).

#### **Enlaces:**

- Los enlaces sencillos, dobles y triples se representan mediante los símbolos: -, = y #, respectivamente.
- Los átomos adyacentes se suponen conectados entre sí por un enlace sencillo, y normalmente no se especifica.

Ejemplos de diversos tipos de enlace

SMILES	CC	[H][H]	C=O	C=C	C#N
NOMBRE	etano	hidrógeno molecular	formaldehído	etileno	cianuro de hidrógeno

#### **Radicales:**

Los radicales o ramificaciones se especifican entre paréntesis, y pueden ser anidados. En cualquier caso, un radical escrito entre paréntesis siempre lo consideraremos conectado al átomo de la izquierda.

Ejemplos de notación de radicales

SMILES	CC(C)CC	CC(C)C(=O)O	C=CC(CC)C(C(C)C)CCC
NOMBRE	metilbutano	ácido isobutanoico	3-etil-4-isopropil-1-hepteno

#### **Estructuras cíclicas:**

## Difusión e impacto de la Química en nuestra sociedad

- Las estructuras cíclicas alifáticas se representan como una rotura de un enlace en cada anillo. Los enlaces están numerados en cualquier orden, y se indican los enlaces del principio y del final del anillo con un dígito a continuación del símbolo del átomo. El ciclohexano es un ejemplo típico (C1CCCC1).
- En cuanto a los anillos aromáticos, la alternancia de enlaces sencillos y dobles se indica escribiendo los átomos del anillo en minúsculas. Así, el benceno sería: (c1ccccc1).


En el caso de los alquenos, para distinguir entre las dos formas, en la notación SMILES se utiliza el convenio siguiente:

- **forma *cis***— se colocan dos barras convergentes; una por delante del primer carbono del doble enlace y la otra por detrás del segundo átomo de carbono: /C=C\ ó \C=C/
- **forma *trans***— se colocan dos barras paralelas; una por delante del primer carbono del doble enlace y la otra por detrás del segundo átomo de carbono: /C=C/ ó \C=C\

### Algunos recursos online

El uso de SMILES<sup>TM</sup> puede ir asociado tanto al mundo de la educación como al del comercio de productos químicos y de *software* de este ámbito.

Podemos encontrar referencias a este lenguaje en la enciclopedia *online* más famosa del momento, *Wikipedia*, en su versión en inglés.



The image shows a screenshot of the Wikipedia page for Acetylene. On the left, there is a table of contents with sections like '1 Discovery', '2 Preparation', '3 Bonding', '4 Reactions', '5 Natural occurrence', '6 Safety and handling', '7 References', and '8 External links'. The 'Discovery' section is expanded, showing text about its discovery in 1836 by Edmund Davy and 1860 by Marcellin Berthelot. On the right, there is a table of identifiers for Acetylene, including IUPAC name (Ethyne), Systematic name (Acetylene), CAS number (74-86-2), and others. The 'SMILES' entry in the identifiers table is circled in red.

Identifiers	
CAS number	74-86-2 ✓
ChemSpider	6086 ✓
UNII	QC7TV750B3 ✓
UN number	1001 (dissolved) 3138 (in mixture with ethylene and propylene)
KEGG	C01548 ✓
CHEBI	CHEBI:27518 ✓
CHEMBL	CHEMBL116336 ✓
Jmol-3D images	page 1 
	<b>SMILES</b> <a href="#">[show]</a>
	<a href="#">InChI</a> <a href="#">[show]</a>
Properties	
Molecular	C <sub>2</sub> H <sub>2</sub>

Captura de la versión inglesa de Wikipedia que muestra el uso de la notación SMILES<sup>TM</sup>

#### IV Jornadas sobre la Enseñanza de la Química



También es muy habitual encontrar referencias a este lenguaje en numerosas búsquedas de compuestos químicos en bases de datos, como por ejemplo *Chemspider*, *Chemsynthesis*, etc.

Captura de la base de datos Chemsynthesis con la opción de búsqueda por SMILES™

El uso de SMILES™ como lenguaje químico también aparece en el ámbito de los proveedores de *software* de Química y los gestores de información sobre productos químicos, como por ejemplo la empresa ACDLabs.

- Automatically verify and assign experimental spectra when a structure is attached with one button click.

**Benefits**

- Ease interpretation of spectra for non-routine experiments and complex structures
- Quickly verify and auto assign experimental spectra to a chemical structure
- Reduce instrument time by estimating <sup>15</sup>N chemical shifts and narrowing down the experimental acquisition range for these experiments.
- Visualize and compare predicted and experimental spectra on the same screen
- Use the software as a teaching tool by generating simulated spectra and hiding the associated chemical structure
- Easily build a central fully searchable repository of NMR data, in the form of the User (training) Database

**Additional Resources**

Chemical structures, imported from a variety of formats (including SDfiles, molfiles, SMILES, InChI, etc.) or drawn directly with ACD/ChemSketch, can be used to predict NMR spectra (see [full list of supported import formats](#)).

Add-ins for ISIS and ChemDraw provide access to our predictions directly from alternative drawing packages, and allow for tables of shifts, coupling constants, and entire predicted spectra to be placed into ISIS/Base.

Learn more about: [Internal Training Databases of the Predictors](#)

[Full List of Prediction, Processing, and Databasing Features](#)

**NMR Blogs**

[Ryan's Blog on NMR software](#)

[Philosophy to Chemistry to Education](#)

**Customers Say**

"The software is an extraordinarily practical and convenient compendium of NMR data that effectively replaces the time consuming tasks of looking up individual NMR data from the Web, the primary literature, paper-based libraries of NMR spectra, or similar resources."

Ejemplo de recurso online relacionado con SMILES™

#### Herramientas de *software*: generación de la información tridimensional

Una vez hemos codificada la molécula que nos interese con SMILES™ necesitaremos algún programa que sea capaz de evaluar ese código y aplicar los principios de química

## Difusión e impacto de la Química en nuestra sociedad

estructural necesarios (radio atómico, carga, longitud y ángulo de enlace, etc) para determinar la posición relativa de cada átomo en la molécula correspondiente y generar un fichero de texto que contenga dicha información.

Este fichero de texto será el utilizado por el visor tridimensional para mostrar la imagen de la molécula y permitir un cierto grado de manipulación del modelo generado.

Una búsqueda exhaustiva nos sirvió para descubrir las dos grandes alternativas existentes para cubrir esta necesidad:

- Por una parte existen opciones *online*, como la que corresponde al proyecto CACTUS (*National Cancer Institute*, 2011), que hacen posible generar a partir de una cadena SMILES™, un fichero de texto descargable de extensión .mol que contiene la información tridimensional de la molécula en cuestión.
- Por otra parte, existen programas que se pueden descargar e instalar en local y que igualmente permiten, además de otras muchas posibilidades, crear el citado fichero .mol.

Después de analizar las diversas soluciones existentes, decidimos optar por el programa *ChemSketch*®, desarrollado por la empresa *Advanced Chemistry Development*, que ofrece una versión *freeware* para uso educativo que se puede descargar desde su página web (*Advanced Chemistry Development*, 2011). En el momento de cerrar esta comunicación la versión disponible es la 12.0 de 30 de junio de 2010.

Las principales causas de esta elección fueron las siguientes:

- Pese a tratarse de un programa comercial, existe una versión gratuita para uso educativo.
- La interfaz del programa tiene un aspecto plenamente profesional; este hecho permite acostumbrar a los alumnos a trabajar en entornos complejos.
- El programa se encuentra exclusivamente en inglés, por lo que supone una oportunidad para que los alumnos valoren la importancia del conocimiento de esta lengua en el ámbito científico.
- Las numerosas opciones para el trabajo con modelos químicos que ofrece el programa permite a aquellos alumnos con más inquietud explorar libremente algunas de estas opciones.
- La secuencia de manipulaciones necesarias para la generación de los ficheros .mol que contienen la información tridimensional de una molécula a partir de su codificación con SMILES™ resulta extremadamente sencilla.

La secuencia completa necesaria para generar el fichero de texto ( con extensión .mol) que contiene la información tridimensional a partir de la codificación SMILES™ se puede esquematizar de la siguiente manera:

- Una vez iniciado el programa, accederemos a la siguiente opción del menú: *Tools* → *Generate* → *Structure from SMILES*.



#### IV Jornadas sobre la Enseñanza de la Química

- En la ventana emergente que mostrará el programa introduciremos el código SMILES™ correspondiente a la molécula que queramos representar y pulsaremos sobre el botón rotulado como *OK*.
- En la zona de trabajo del programa obtendremos una representación plana de la molécula citada, en este momento deberemos revisar que el resultado obtenido se corresponda con la molécula en cuestión, es decir, comprobaremos si la codificación SMILES™ introducida es correcta o no.
- Accederemos a partir del menú a: *Tools* → *3D Structure Optimization*. Esta acción es la clave del proceso, ya que en ella es donde el programa realiza los cálculos relativos a la disposición tridimensional relativa de cada átomo de la molécula.
- Finalmente, escogeremos la opción del menú: *File* → *Export*, indicaremos el nombre del fichero y la carpeta donde queremos guardarlo, aceptando la opción de formato por defecto de este fichero (*MDL Molfiles [V2000] [\*mol]*).

El fichero de texto obtenido por este proceso contendrá la información tridimensional necesaria para su correcta visualización con el visor correspondiente.

#### Herramientas de *software*: el visor tridimensional

Una vez realizado el proceso indicado anteriormente, el siguiente paso es utilizar un visor que permita generar, visualizar y manipular el modelo tridimensional a partir de la información contenida en dicho fichero *.mol*.

De entre las diversas opciones existentes, nos decantamos por el uso de Jmol, definido por sus creadores como: *un visor Java de código abierto para estructuras químicas en tres dimensiones, con prestaciones para compuestos químicos, cristales, materiales y biomoléculas* (Jmol, 2011).

Jmol fue creado por Dan Gezelter en 1999 como parte del proyecto *Open Science* y posteriormente ha sido desarrollado por otros programadores como Bradley A. Smith y Egon Willighagen, y integrado a partir del año 2002 en el proyecto *The Chemistry development Kit* (The Chemistry Development Kit, 2000).

Los principales argumentos que justifican la elección de este visor son (Herráez, 2006):

- Es un programa de código abierto, liberado bajo una licencia *GNU/GPL* (GNU, 2000). Este hecho garantiza una constante revisión del programa y incorporación de nuevas funcionalidades (en el momento de cerrar esta comunicación la última versión disponible es la 12.2 RC6).
- Se puede descargar de forma gratuita desde la página de los autores (Jmol, 2011).
- Está escrito en *Java* (Oracle, 2011), por lo cual resulta compatible con todos los sistemas operativos habituales, con la excepción de las tabletas, así mismo es compatible con los principales navegadores.

## Difusión e impacto de la Química en nuestra sociedad

- Se puede utilizar de tres formas diferentes: como programa independiente, como *applet* que se puede incrustar en una página web o en un curso de *moodle* (previa instalación por parte del administrador del filtro correspondiente), o bien como herramienta de desarrollo.
- Las exigencias de *hardware* son moderadas y por tanto se puede trabajar con soltura en ordenadores antiguos o en *netbooks* de escasa potencia.
- Se puede interactuar con los modelos obtenidos incrustados en una página web vía *javascript* (Herráez, 2007).

A esta lista de ventajas técnicas, podemos añadir otras de índole práctica y/o pedagógica que aseguran la eficacia del uso de Jmol en el aula:

- Su funcionamiento es muy simple e intuitivo, por lo que la curva de aprendizaje de su uso básico resulta muy suave para los alumnos. Existen numerosos manuales de uso (Herráez, 2007) y ejemplos de actividades didácticas basadas en el uso de Jmol disponibles de forma gratuita en Internet.
- Los modelos generados pueden imprimirse en color con una calidad más que aceptable, ya que el programa permite la exportación a formatos gráficos de uso común como *JPEG*, *GIF* y *POVRAY* entre otros.
- Es compatible con la mayoría de los formatos utilizados en diversos ámbitos científicos e industriales para la codificación de moléculas, como *MOL*, *XYZ* y *PDB*.
- Permite simular orbitales y vibraciones, y realizar medidas de distancias o ángulos de enlace. De igual manera, los usuarios avanzados pueden realizar animaciones que ayuden a destacar algún aspecto específico de la molécula con la que se trabaje en cada momento.
- **Experiencia de aula**

Describiremos a continuación la concreción en el aula de las posibilidades que nos brindan las herramientas descritas hasta este momento. La gestión de aula se organiza en forma de seminario con dos sesiones de una duración de una hora y media aproximadamente, realizadas en horario extraescolar. En la primera sesión los alumnos reciben un dossier con una pequeña introducción a las normas de escritura de SMILES™, una hoja de ejercicios de ejemplo resueltos para realizar durante la sesión y otra hoja con ejercicios propuestos para realizar en casa y entregar antes de la segunda sesión.

En la primera parte de esta sesión se explica a los alumnos los motivos de la realización del seminario, haciendo hincapié en las características principales del lenguaje SMILES™, así como todos los aspectos prácticos necesarios para la descarga, instalación y manipulación de los diferentes programas informáticos necesarios.

Se explican las principales opciones que nos brinda el menú de Jmol, así como sus prestaciones y opciones más destacadas. El uso de *ChemSketch*©, exige más tiempo, ya





#### IV Jornadas sobre la Enseñanza de la Química

que los alumnos deben aprender la secuencia de codificación, optimización y exportación descrita anteriormente.

En la segunda parte de esta sesión los alumnos ponen en práctica las instrucciones que han recibido anteriormente aplicándolas a los ejercicios que se les proporciona, de forma que se puede hacer una puesta en común de las dudas que vayan surgiendo, especialmente las referentes a las diferentes maneras que a menudo existen de codificar una misma molécula con SMILES™.

Por otra parte, se intenta mostrar a los alumnos la universalidad y versatilidad del lenguaje SMILES™, que le convierten *de facto* en un estándar para la codificación de moléculas. Para ello se les invita a visitar diversas páginas web donde este lenguaje se utiliza de forma habitual, como por ejemplo algunas bases de datos de compuestos como *ChemSpider* (Chemspider, 2011) o bien cualquier artículo sobre un compuesto químico que se pueda encontrar en la versión inglesa de la *Wikipedia* (Wikipedia, 2011).

Entre la primera y la segunda sesión del seminario los alumnos deben aplicar los conceptos y habilidades aprendidos a la resolución de una batería de ejercicios propuestos, y tienen que entregar para ser evaluados de forma individual y por vía telemática los ficheros .mol correspondientes a las moléculas propuestas.

La segunda sesión del seminario sirve para revisar el trabajo realizado, resolver las posibles dudas y afianzar los conocimientos y habilidades adquiridos, así como para realizar una ampliación de las posibilidades que ofrece el procedimiento descrito para trabajar un concepto más complejo como las isomerías.

Dado que los alumnos ya están familiarizados tanto con el lenguaje SMILES™, como con el *software* utilizado y con la secuencia de acciones necesarias para resolver las cuestiones planteadas, esta segunda sesión resulta esencialmente práctica y más productiva que la anterior, y se pueden plantear situaciones de complejidad creciente en función de las capacidades de los alumnos.

#### Conclusiones

En primer lugar, podemos afirmar que la secuencia de trabajo descrita en esta comunicación resulta fácil de implementar para el profesor y de asimilar para el alumnado; no requiere un aprendizaje complejo ni desembolso alguno, por lo tanto podemos concluir que se trata de una dinámica fácil de incorporar a cualquier aula de química o biología de bachillerato rápida y sencillamente.

En segundo lugar, podemos encontrar con facilidad tanto el *software* descrito en esta comunicación como otras alternativas a explorar, así como manuales de uso y una gran cantidad de recursos y ejemplos de aplicación existentes en Internet que cualquier profesor puede incorporar libremente y sin ninguna complejidad a los materiales que él mismo desarrolle en diversos formatos y soportes.

Finalmente, nuestra experiencia nos indica que nuestros alumnos no sólo se habitúan con facilidad a la dinámica descrita, sino que asimismo experimentan una mejora evidente en la comprensión de los conceptos químicos implicados **en** esta experiencia.

## Difusión e impacto de la Química en nuestra sociedad

### Bibliografía

- Advanced Chemistry Development. (2011). *Freeware*. Recuperado 16 de septiembre, 2011, de <http://www.acdlabs.com/resources/freeware>
- Chemspider. (2011). *Building Community For Chemists*. Recuperado 16 de septiembre, 2011, de <http://www.chemspider.com/simplesearch.aspx>
- Chemical Synthesis Database. (2008). *Chemsynthesis*. Recuperado 16 de septiembre, 2011, de <http://www.chemsynthesis.com>
- Daylight. Chemical Informations Systems, Inc. (2008). *SMILES* Recuperado 16 de septiembre, 2011, de <http://www.daylight.com/smiles/index.html>
- Departament d'Educació (2008). DECRET 142/2008, de 15 de juliol, pel qual s'estableix l'ordenació dels ensenyaments del batxillerat. *Diari Oficial de la Generalitat de Catalunya*, núm. 5183. p. 59042-59401.
- Gnu. (2011). *Gnu Lesser General Public License*. Recuperado 16 de septiembre, 2011, de <http://www.gnu.org/licenses/lgpl.html>
- Herráez, A. (2011). *Biomodel: Complementos De Bioquímica Y Biología Molecular*. Recuperado 16 de septiembre, 2011, de <http://biomodel.uah.es/>
- Herráez, A. (2007). *Cómo utilizar Jmol para estudiar y presentar estructuras moleculares. Volumen 1: aprendiendo a usar Jmol (niveles básico e intermedio)*. Morrisville, North Carolina: Lulu Enterprises.
- Herráez, A. (2006). Biomolecules in the computer: Jmol to the rescue. *Biochemistry and Molecular Biology. Education*, núm, 34. p. 255-261.
- Jmol. (2011). *Jmol: un visor java de código abierto para estructuras químicas en tres dimensiones*. Recuperado 16 de septiembre, 2011, de <http://jmol.sourceforge.net/index.es.html>
- Lucas, A. Giménez, C. (2010). *Codificació de molècules orgàniques amb SMILES™ i representació tridimensional amb Jmol: un exemple de pràctica reflexiva*. *EduQ*, núm. 7. P.15-22.
- National Cancer Institute. (2011). *Online Smiles™ Translator And Structure File Generator*. Recuperado 16 de septiembre, 2011, de <http://cactus.nci.nih.gov/translate/>
- Oracle. (2011). *Java y tú: descargar hoy*. Recuperado 16 de septiembre, 2011, de <http://www.java.com/es/>
- The Chemistry Development Kit. (2000). *The Chemistry Development Kit*. Recuperado 16 de septiembre, 2011, de <http://cdk.sf.net/>
- Wikipedia. (2011). *Wikipedia, the free encyclopedia*. Recuperado 16 de septiembre, 2011, de [http://en.wikipedia.org/wiki/main\\_page](http://en.wikipedia.org/wiki/main_page)