

## Simulación Bidimensional de RMN Acoplada con Reconocimiento de Patrones vía Redes Neuronales: una poderosa herramienta de Validación Estructural

Zanardi, María Marta<sup>†,‡</sup>; Sarotti, Ariel Marcelo<sup>‡</sup>

<sup>†</sup>Facultad de Química e Ingeniería del Rosario, Pontificia Universidad Católica Argentina, Av. Pellegrini 3314, Rosario 2000, Argentina.

<sup>‡</sup> Instituto de Química Rosario (CONICET), Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, Rosario 2000, Argentina

**RESUMEN.** *El aislamiento, caracterización y síntesis total de nuevos productos naturales con potencial actividad biológica, es una de las áreas más apasionantes de la química. Dentro de esta, la elucidación estructural de moléculas orgánicas complejas es una tarea sumamente desafiante. Entre las distintas técnicas espectroscópicas para llevar a cabo tal fin se destaca la Resonancia Magnética Nuclear (RMN) la cual ha ido evolucionando con el advenimiento de espectrómetros cada vez más sofisticados y nuevas secuencias de pulsos. Sin embargo en la era dorada de la RMN la publicación de estructuras incorrectas continúa siendo una situación común. La gran complejidad molecular, los errores humanos, la ambigüedad de señales y las impurezas de las muestras pueden considerarse como las fuentes más comunes de las asignaciones erróneas. Cientos de revisiones estructurales se han publicado en las últimas décadas, que van desde errores groseros en la conectividad a sutiles (pero no menos importante) errores estereoquímicos. Teniendo en cuenta que las discrepancias se detectan a menudo después de la síntesis total de la estructura originalmente propuesta (errónea), no es ilógico suponer que la arquitectura molecular real de muchos de los productos naturales reportados sigue siendo desconocida. La química computacional moderna ha contribuido significativamente a prevenir estos errores. En este contexto hemos desarrollado herramientas que permiten interpretar la correlación entre datos experimentales y calculados de RMN que contribuyen significativamente en el proceso de elucidación estructural.*

**Palabras Clave:** *Química computacional, RMN, Elucidación estructural, Redes Neuronales Artificiales.*

**ABSTRACT.** *The isolation, characterization and total synthesis of new natural products with potential biological activity is one of the most exciting areas of chemistry. Regarding this, structural elucidation of complex organic molecules is a highly challenging task. Among the different spectroscopic techniques to carry out this purpose stand out the Nuclear Magnetic Resonance, which has evolved with the advent of increasingly sophisticated spectrometers and new pulse sequences. However in the golden age of NMR, the publication of incorrect structures remains a common situation. The large molecular complexity, human errors, signal ambiguity and sample impurities can be considered as the most common sources of erroneous assignments. Hundreds of structural revisions have been published in recent decades, ranging from severe errors in connectivity to subtle stereochemical errors. Taking into account that these discrepancies are often detected after the total synthesis of the originally proposed (erroneous) structure, it is not illogical to assume that the actual molecular architecture of many of the reported natural products remains unknown. Modern computational chemistry has contributed significantly to preventing these errors. In this context we have developed some tools to interpret the correlation between experimental and calculated NMR data that contribute significantly to structural elucidation process.*

**Keywords:** *Computational chemistry, NMR, Structural elucidation, Artificial Neural Network.*

### Introducción

La síntesis total de los productos naturales es una de las áreas de química más bellas y apasionantes, situada en el límite entre un arte y una ciencia dura. Gracias a la síntesis total, aparte de emular la capacidad de la naturaleza para construir complejas arquitecturas moleculares han sido posibles el desarrollo de innumerables estrategias, metodologías y nuevas transformaciones químicas. Estas investigaciones a menudo son difíciles, llenas de dificultades y desvíos. Frecuentemente después de una considerable inversión de tiempo, dinero y recursos humanos, se logra el objetivo sintético planteado. Sin embargo, sorprendentemente en ocasiones este momento emocionante cambia rápidamente a un sentimiento de frustración cuando se detecta una incongruencia entre los datos de RMN del compuesto sintetizado y el

producto natural previamente reportado. A pesar de lo enormes progresos tecnológicos llevados a cabo en el área de la espectroscopia, para la elucidación estructural de compuestos orgánicos, los productos naturales incorrectamente asignados son más frecuentes de lo que creeríamos y desearíamos.

Una valiosa herramienta con la que contamos en la actualidad para complementar los datos experimentales, muchas veces confusos, es la química computacional. Antiguamente la ciencia química era considerada como un campo netamente experimental alejado de consideraciones matemáticas. Sin embargo con la llegada de la Mecánica Cuántica los pioneros en ese campo comenzaron a darse cuenta que tenía el potencial de ser una teoría predictiva en química [1]. Actualmente, la investigación científica necesita cada vez más de una integración multidisciplinaria que permita aprovechar las distintas herramientas que ofrece cada área científica en post de un objetivo común. Así surge la química computacional como una disciplina que se extiende más allá de los límites tradicionales que separan la química, la física, la biología y la ciencia de la computación; permitiendo la investigación del comportamiento de la materia a nivel molecular mediante un sistema de ordenadores, cuando la investigación de laboratorio sea inapropiada, poco práctica o imposible. Por este motivo, la química computacional se encuentra estrechamente relacionada con la síntesis orgánica ya que permite establecer puntos de partida para trabajos sintéticos, determinar estructuras moleculares y mecanismos de reacción, estudiar especies muy reactivas que no pueden ser aisladas, interpretar y predecir datos experimentales, así como identificar correlaciones entre estructuras químicas y sus propiedades [2]. La química computacional se encarga de incorporar la química teórica a un software informático con el objeto de calcular las estructuras y propiedades de moléculas, ya sea empleando la mecánica molecular o las teorías estructurales electrónicas basadas en la mecánica cuántica [3]. Dada la gran precisión en las simulaciones alcanzada con el modelado molecular, es importante poder combinarlas con métodos de inteligencia artificial (redes neuronales, algoritmos genéticos, etc.) que permitan analizar bases de datos y detectar patrones de comportamiento entre variables calculables o medibles con propiedades o actividades químicas o biológicas de interés, y en base a ellos poder realizar predicciones sobre nuevos casos.

En función de la gran utilidad que ha demostrado tener la química computacional, como aporte tanto para la interpretación, predicción y corroboración de datos experimentales en el área de la química orgánica, este trabajo se enfoca en abordar una de las problemáticas que surgen durante las investigaciones en el campo de la química orgánica, con el objetivo de optimizar tiempo, esfuerzos y recursos.

### **Aplicación de cálculos teóricos de RMN en la validación estructural de compuestos orgánicos**

Actualmente la espectroscopia de resonancia magnética nuclear (RMN) constituye la herramienta más valiosa que disponemos los químicos orgánicos para poder determinar la naturaleza estructural y estereoquímica de los compuestos que diariamente sintetizamos o aislamos de fuentes naturales. Sin embargo, aún con el advenimiento de espectrómetros cada vez más sofisticados y nuevas secuencias de pulsos, la publicación de estructuras incorrectas continúa siendo una situación común. Esto se evidencia en los cientos de revisiones estructurales que han sido publicadas en las últimas décadas, la mayoría de las cuales son detectadas luego de la síntesis total de la estructura originalmente propuesta [4]. Este tipo de situaciones afectan notablemente a gran parte del sistema científico, perjudicando la credibilidad del grupo que asigna erróneamente el producto natural, generando serias confusiones sobre los esquemas biosintéticos del compuesto en cuestión, y afectando en términos económicos, de tiempo y recursos a los grupos sintéticos que deciden aventurarse hacia la síntesis total de la estructura propuesta. En los últimos años se ha potenciado notablemente el empleo de cálculos de RMN para prevenir, al menos en parte, estas problemáticas [5]. Por ello, la generación de nuevas metodologías computacionales para ser empleadas en casos de asignación o validación estereoquímica y/o estructural es de sumo interés para diversas áreas de la química orgánica.

El cálculo del comportamiento magnético de las moléculas se inició formalmente en 1937, cuando London reportó el tratamiento de la anisotropía diamagnética de sistemas aromáticos [6]. Sin embargo, la determinación rutinaria de los tensores magnéticos (y por lo tanto, de los desplazamientos químicos) no resultó práctica hasta que no se resolvió el problema del origen del *gauge* (*gauge problem*). Numerosos métodos fueron desarrollados para solucionar este inconveniente, entre ellos, GIAO (*Gauge Including Atomic Orbitals*) e IGLO (*Individual Gauges for Localized Orbitals*) [5.a]. Las notables mejorías que se fueron implementando con el correr de los años a los métodos de cálculo y a las metodologías auxiliares de correlación, han provocado un auge en el uso de estas herramientas en el campo de la química orgánica. Las propiedades espectroscópicas de una muy extensa lista de productos naturales y no naturales han sido calculadas con fines de revisión, validación o reasignación estructural y/o estereoquímica. De particular importancia es el desarrollo de los métodos estadísticos efectuado por Goodman. El método CP3 fue diseñado para asignar un par de estructuras con dos conjuntos de datos espectroscópicos [7.a], mientras que el DP4 permite asignar dos (o más) estructuras a un solo *set* de datos experimentales [7.b]. Estos métodos son probablemente los más sofisticados y confiables para la asignación estereoquímica de moléculas orgánicas sobre la base de cálculos de RMN. En los últimos años, DP4 ha sido extensivamente empleado para confirmar o corregir la arquitectura molecular de numerosos compuestos orgánicos. En nuestro grupo de trabajo hemos desarrollado una nueva probabilidad mejorada (DP4+), erigida sobre los cimientos matemáticos del DP4, que garantiza un nivel de predicción y confianza muy superiores a los establecidos por el formalismo original [8].

De todos modos, todos estos enfoques (CP3, DP4, DP4+, etc.) fallan en casos de validación estructural (determinar si la

estructura propuesta es correcta o no), pues la intrínseca ausencia de al menos otra segunda estructura hace que la comparación sea simplemente imposible. En este trabajo hemos abordado esta problemática demostrando que la utilización de cálculos de RMN combinado con reconocimiento de patrones vía redes neuronales artificiales, representa un método rápido, sencillo y confiable de validación estructural empleando un único *set* de datos experimentales y calculados [9]. Inicialmente el método se desarrolló en base al análisis exclusivo de los desplazamientos químicos de RMN  $^{13}\text{C}$ , demostrando ser útil en la detección de errores de conectividad y regioquímica [9.b]. Luego a través de la utilización de niveles de teoría más precisos y empleando patrones reconocimiento multidimensional mediados por redes neuronales de correlación de datos experimentales y calculados de RMN de  $^1\text{H}$ ,  $^{13}\text{C}$  y 2D C-H COSY, hemos logrado mejorar la precisión de esta metodología, obteniendo redes capaces de detectar errores estructurales tan sutiles como los de tipo estereoquímico [9.a].

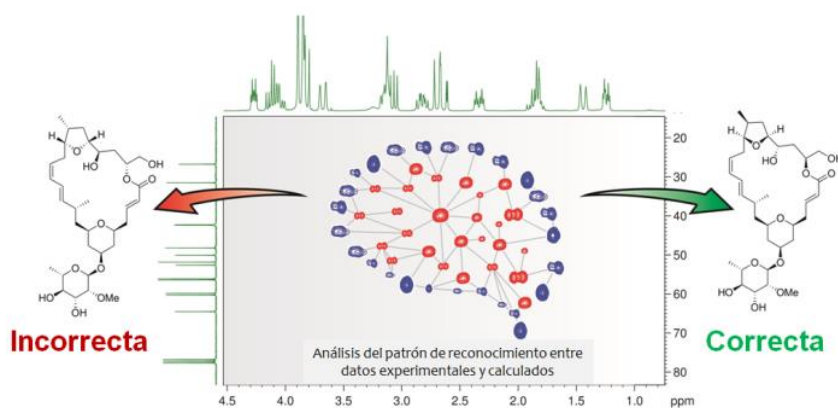


Figura 1

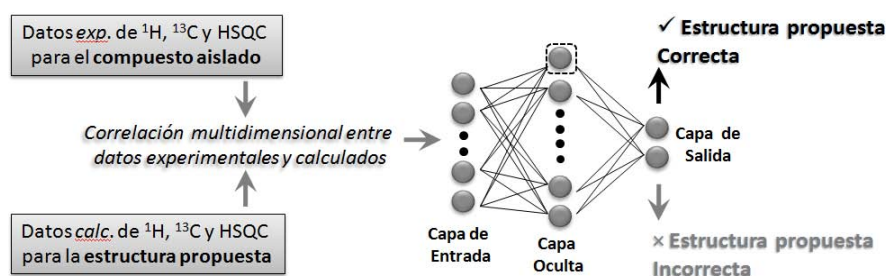


Figura 2

## Resultados y discusión

El cálculo teórico de desplazamientos químicos es una tarea compleja, especialmente la evaluación de la contribución paramagnética. Los desplazamientos químicos son muy sensibles a ligeras variaciones en la densidad electrónica aún en regiones relativamente remotas al núcleo de interés. Por ello, es frecuente encontrar en la literatura cálculos de desplazamiento químico no corregido RMN con errores promedio de hasta 0,4 ppm o más para  $^1\text{H}$  y hasta 10 ppm o más de  $^{13}\text{C}$ ; con valores extremos que muestran aún mayores desviaciones, incluso cuando se realizaron con algunos de los mejores métodos computacionales. Estos errores son claramente demasiado grandes para muchas aplicaciones. Hay muchas fuentes para este tipo de errores, y la comprensión de estas fuentes, es esencial para remediar efectivamente su efecto. Entre ellas podemos destacar: correlación electrónica, efectos del solvente y otras interacciones intermoleculares (como puentes de hidrógeno por ejemplo), efectos estéricos debidos a la libertad conformacional, efectos rovibracionales y efectos de átomos pesados.

Para interpretar estos errores hemos desarrollado herramientas [9] que significan un gran avance en los aportes de la química computacional a la validación estructural de compuestos orgánicos mediante el uso de redes neuronales artificiales.

## Entrenamiento de las Redes Neuronales

Para entrenar las redes neuronales, se buscaron de la literatura compuestos quirales con al menos dos centros estereogénicos, y cuyos datos espectroscópicos (RMN  $^{13}\text{C}$  y  $^1\text{H}$ ) estén adecuadamente reportados y sean fiables. Sobre estas estructuras, se procedió a invertir deliberadamente al menos uno de los centros estereogénicos para conseguir un conjunto de diastereoisómeros incorrectos, obteniendo así un *set* de entrenamiento con de 100 ejemplos de estructuras correctas y 82 ejemplos de estructuras incorrectas, más 26 combinaciones incorrectas de datos experimentales y

teóricos. Posteriormente, todas las geometrías se optimizaron empleando métodos computacionalmente accesibles (MMFF, HF/3-21G, B3LYP/6-31G\*) y los parámetros de RMN se calcularon al nivel *mPW1PW91/6-31G\** en vacío y *mPW1PW91/6-31G\*\** en solución. Los desplazamientos químicos se calcularon usando como estándar de referencia tanto TMS como el método multi-estándar (MSTD), desarrollado por parte de nuestro grupo, de acuerdo con la ecuación:

$$\delta_{\text{calc}}^x = \sigma_{\text{ref}} - \sigma^x + \delta_{\text{ref}}$$

Donde  $\delta_{\text{calc}}^x$  es el valor del desplazamiento químico calculado para en núcleo  $x$ ,  $\sigma_{\text{ref}}$  es el tensor magnético calculado para el compuesto de referencia al mismo nivel de teoría,  $\sigma^x$  es el tensor calculado para el núcleo  $x$  del compuesto y  $\delta_{\text{ref}}$  es el desplazamiento químico experimental del estándar de referencia.

Los desplazamientos químicos calculados fueron escalados según la siguiente ecuación:

$$\delta_{\text{esc}} = (\delta_{\text{calc}} - b) / m$$

Donde  $m$  y  $b$  son la pendiente y ordenada al origen obtenidos de la gráfica  $\delta_{\text{calc}}$  vs  $\delta_{\text{esc}}$ .

Para cada estructura se calcularon los desplazamientos químicos de  $^{13}\text{C}$  y  $^1\text{H}$  escalados y no escalados. Los sets de datos calculados se correlacionaron con los datos experimentales empleando descriptores estadísticos que sirvieron de capa de entrada para sistemas de redes neuronales posibilitando un adecuado análisis de los datos. Así para cada compuesto se calcularon 18 parámetros de las correlaciones 1D (9 de  $^1\text{H}$  y 9 de  $^{13}\text{C}$ ) y 18 de las correlaciones 2D (que incluyen parámetros estadísticos que tienen que ver con la distribución en cuadrantes de los cross-peaks tomando como centro el cross-peak obtenido experimentalmente Figura 3), por cada estándar de referencia (TMS y MSTD). Obteniendo de este modo una matriz de 72 elementos para cada uno de los dos niveles de teoría empleados.

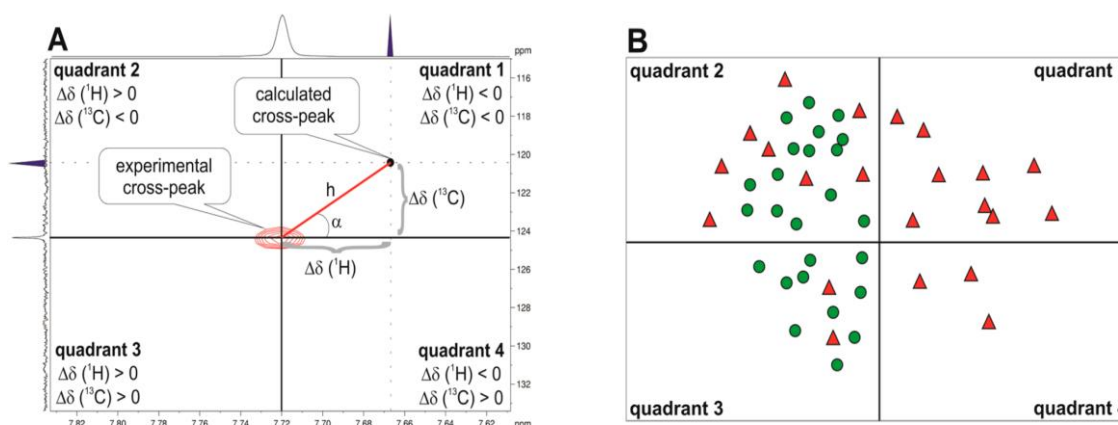


Figura 3

Una vez modelados todos los compuestos para realizar el entrenamiento de las redes neuronales artificiales (RNA) y calculados los 72 parámetros estadísticos de correlación en vacío al nivel *mPW1PW91/6-31G\*\*/B3LYP/6-31G\** y los 72 parámetros estadísticos de correlación en solución al nivel *mPW1PW91/6-31G\*\*/B3LYP/6-31G\** contábamos con una matriz de 144 parámetros de 208 ejemplos de estructuras para realizar el entrenamiento de las RNA.

Brevemente una RNA es un modelo matemático que consiste en un número de neuronas artificiales interconectadas tratando de emular la forma en que funciona un cerebro biológico. De hecho, una de las propiedades más relevantes de las RNAs es su capacidad para aprender de los datos, lo cual las hace muy útiles en el reconocimiento de patrones, clasificación y agrupación.

En la Figura 2 se representa el tipo de red empleada, donde las capas de entrada, ocultas y de salida (los tres componentes principales) están totalmente conectadas, lo que significa que cada neurona está ligada a cada neurona en la capa precedente. En analogía con la neurona biológica, la sinapsis (fuerza de la conexión entre dos neuronas) se mide por un peso número ( $w$ ). Cada conexión lleva un peso asignado, y una función de activación (también conocidas como funciones de transferencia,  $f$ ) que controlan el valor de salida. En este trabajo, se utilizó la transferencia sigmoidea tanto en las capas ocultas como en las de salida, ya que esta arquitectura puede clasificar los vectores arbitrariamente bien. Para el entrenamiento de las redes se empleó un proceso de entrenamiento supervisado utilizando un algoritmo de entrenamiento por retropropagación de gradiente conjugado escalado para determinar los valores óptimos de los pesos ( $w_{ij}$ ) y sesgos ( $b_{ij}$ ).

Se realizó la exploración de la combinación óptima de capas de entrada y capa oculta. Para ello, se desarrollaron y entrenaron más de 400 RNAs diferentes variando el tamaño e identidad de la capa de entrada (número y tipo de

descriptores) y el número de neuronas en la capa oculta (entre 10 y 100). Se realizaron ensayos empleando la matriz completa (con los 144 parámetros), submatrices con los 72 elementos de los cálculos en vacío o sólo los 72 en solución, otras empleando sólo parámetros de la correlación con los desplazamientos calculados con TMS o sólo empleando el método multiestandar, y muchas submatrices más sólo con datos 1D, sólo 2D,  $^1\text{H}$ ,  $^{13}\text{C}$ , etc. La mejor clasificación se logró con todos los datos proporcionados por  $^1\text{H}$ ,  $^{13}\text{C}$ , y las correlaciones C-H, y la eliminación de cualquier subconjunto de parámetros estadísticos resultó en detrimento de la capacidad de reconocimiento de patrones de la red. Mezclar los 36 parámetros de TMS con el 36 parámetros de MSTD (tanto en fase gaseosa o en solución) no mejoró los resultados (de hecho, disminuyó la capacidad de clasificación de los RNAs). Encontramos así que las dos mejores redes fueron: **ANN-TMS<sub>vac</sub>**: construida con los 36 parámetros calculados a *mPW1PW91 / 6-31G\** en vacío con TMS como patrón de referencia y 10 neuronas en la capa oculta, con una clasificación correcta luego del entrenamiento del 97%; y **ANN-MSTD<sub>sol</sub>**: construida con los 36 parámetros calculados a *mPW1PW91 / 6-31G\*\** en solución con MSTD como estándar de referencia y 20 neuronas en la capa oculta con una clasificación del 92%. Se propuso mezclar las capas de entrada de ANN-TMS<sub>vac</sub> y ANN-MSTD<sub>sol</sub> para ver si permitirían tomar la mejor de ambos mundos. De este modo se constituyó la red **ANN-mix**: con una capa de entrada de 72 elementos (36 parámetros derivados de cálculos en fase gaseosa utilizando TMS como estándar y 36 parámetros derivados de cálculos de soluciones utilizando MSTD como estándar) y 14 neuronas en la capa oculta la cual permitió una correcta clasificación luego del entrenamiento del 94%.

Lo más importante de las RNAs es que una vez entrenadas son capaces de predecir datos en casos para los cuales no fue entrenada, por lo que se requiere hacer la validación de las mismas.

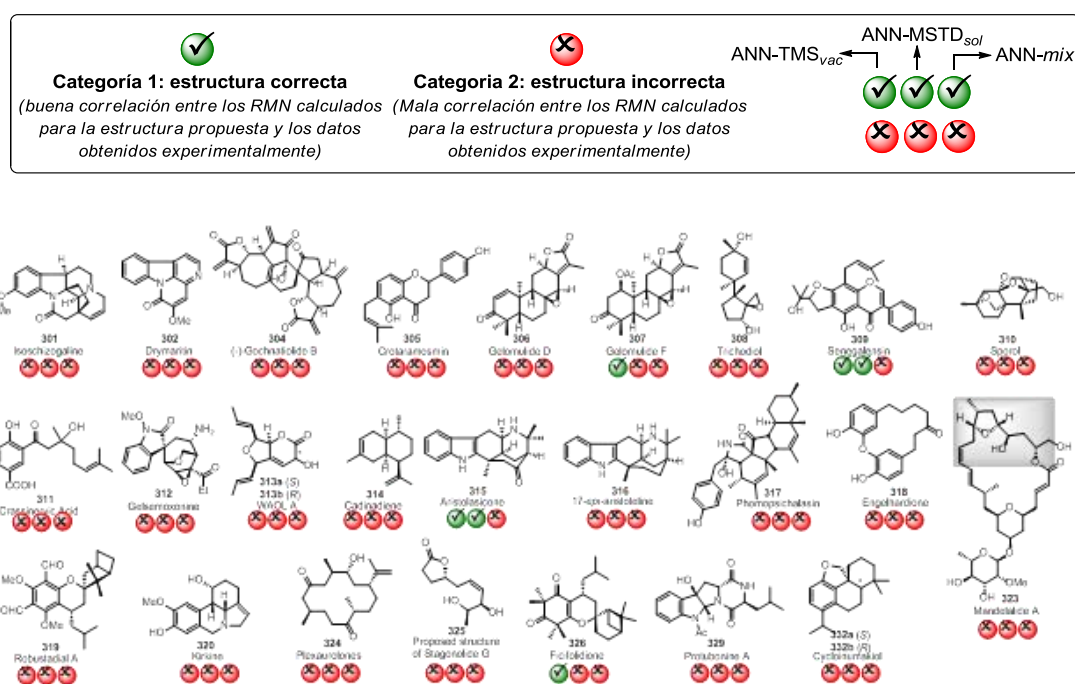
### Validación de las Redes Neuronales Artificiales Entrenadas

Para evaluar capacidad de clasificación de las tres redes entrenadas seleccionadas (ANN-TMS<sub>vac</sub>, ANN-MSTD<sub>sol</sub> y ANNmix) se validaron en 25 casos reales de malas asignaciones estructurales. Las diferencias entre las estructuras originalmente propuestas (Figura 4) y las revisadas (Figura 5) son principalmente regio- y / o estereoquímicas, aunque incluimos algunos ejemplos difíciles de isomerismo constitucional.

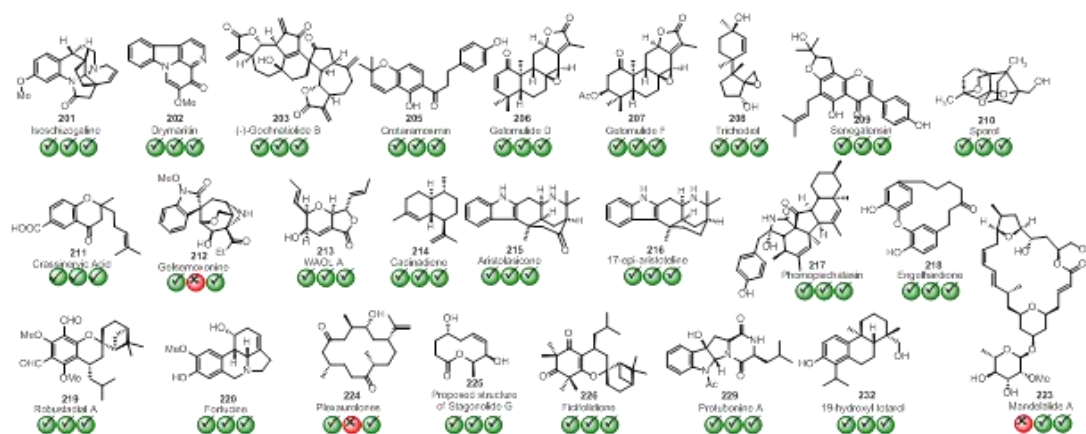
Después de extensas búsquedas conformacionales, se calcularon los tensores de apantallamiento de todas las estructuras mostradas en las Figura 4 y 5 a niveles de teoría en el *mPW1PW91/6-31G\*\*//B3LYP/6-31G\** y *PCM/mPW1PW91/6-31G\*\*//B3LYP/6-31G\**, y los corrimientos químicos resultantes se correlacionaron con los valores de RMN experimental originalmente reportados para aquellos compuestos, usando las redes entrenadas.

Para nuestro deleite, ANN-TMS<sub>vac</sub>, ANN-MSTD<sub>sol</sub> y ANN-mix detectaron con éxito las malas asignaciones de las estructuras originalmente propuestas en el 85%, 94% y 100% de los casos, respectivamente (Figura 4).

Por otro lado, cuando se trata de la versión revisada (correcta-Figura 5), ANN-TMS<sub>vac</sub> y ANN-MSTD<sub>sol</sub> fracasaron en sólo uno y dos casos, respectivamente, mientras que ANN-mix clasificó correctamente todos los ejemplos, demostrando claramente su habilidad de reconocimiento de patrones.



**Figura 4:** Productos naturales que fueron inicialmente mal asignados. Las redes neuronales eficientemente determinaron que las estructuras propuestas son incorrectas.



**Figura 5:** Estructuras revisadas de los productos naturales 201-226  
Las redes neuronales eficientemente determinaron que las estructuras revisadas son correctas

## Metodología General

Las búsquedas conformacionales de compuestos se llevaron a cabo empleando el método de mecánica molecular MM+ y/o MMFF implementado en Hyperchem y Spartan. Todos los cálculos semiempíricos, *ab initio* o DFT se llevaron a cabo empleando Gaussian 09. Se realizaron cálculos de frecuencia para caracterizar las estructuras optimizadas como mínimos de energía (ausencia de frecuencias imaginarias) o como estructuras de transición (una frecuencia imaginaria). Los cálculos de RMN se hicieron empleando el método GIAO, empleando funcionales DFT (mPW1PW91, B3LYP, etc) combinados con diversas funciones base, tanto en fase gas como solución. Las redes neuronales artificiales se desarrollaron y entrenaron empleando Matlab.

## Conclusión

El cálculo de RMN empleando métodos de bajo costo computacional, acoplado con redes neuronales artificiales entrenadas, representa una poderosa herramienta para la identificación temprana de errores de asignación estructural en forma rápida, simple y confiable. Hemos desarrollado, entrenado y validado redes neuronales artificiales capaces de detectar errores de asignación tan sutiles como los de tipo estereoquímico, incluyendo ejemplos de cambios de un único centro estereogénico en moléculas con múltiples carbonos quirales. Resolviendo en parte el problema de validación estructural (confirmar o rechazar una estructura candidata) utilizando enfoques de química cuántica.

Se han construido y entrenado más de 400 RNAs, y las más eficientes en términos de capacidad de clasificación fueron validadas con éxito en desafiantes ejemplos reales de errores de asignación de productos naturales.

Los mejores resultados se obtuvieron mezclando los descriptores estadísticos calculados al nivel *mPW1PW91 / 6-31G \** (en fase gaseosa) con TMS como estándar de referencia con los obtenidos en el método MSTD al nivel *mPW1PW91/6-31G\*\** (en solución), y se recomienda este enfoque para obtener los resultados más confiables. Además, en un esfuerzo por llevar esta metodología a la comunidad de química, desarrollamos un archivo de Excel que facilita el procedimiento de cálculo está disponible como parte del *Supporting Information*. de la publicación del trabajo [9.a].

## Simplificación del Procedimiento

Desarrollamos un archivo de *Excel* para simplificar el procedimiento de análisis. Con sólo completar los *desplazamientos químicos experimentales* y los *tensores calculados*, indicando la hibridación  $sp^2$  cuando corresponda, la hoja de cálculo calcula los desplazamientos químicos escalados y no escalados, junto con los 72 parámetros estadísticos de correlación que se introducen automáticamente en la red neuronal entrenada para así obtener el patrón de respuesta indicando si la estructura propuesta es **Correcta** o **Incorrecta**.

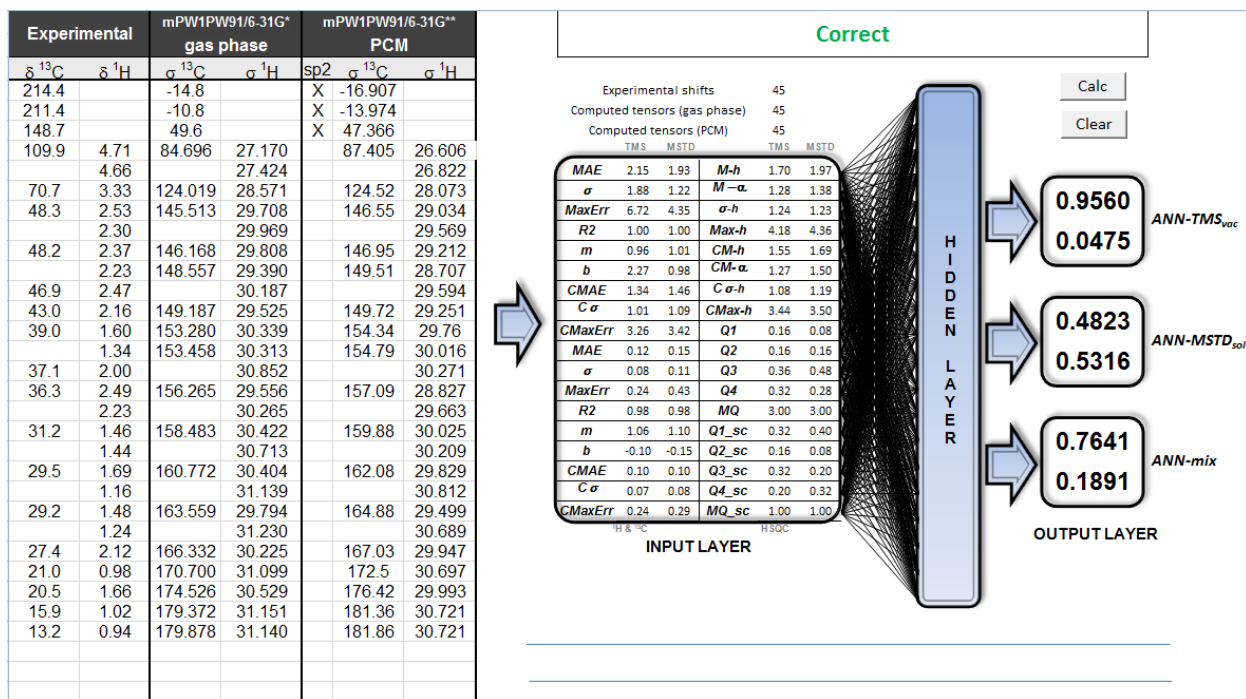


Figura 5

## Referencias

- [1] M. Head-Gordon, E. Artacho, *Chemistry on the computer*. Physics Today, 2008, 61, 4: 58-63.
- [2] Lewars, E. *Computational chemistry. Introduction to the theory and applications of molecular and quantum mechanics*, 2nd ed., Springer, 2011.
- [3] a) Foresman, J. B.; Frisch, A., *Exploring Chemistry with Electronic Structure Methods*, 2<sup>a</sup> Ed., Gaussian Inc., 1996. b) Koch, W.; Holthausen, M. C., *A Chemist's Guide to Density Functional Theory*, 1<sup>a</sup> Ed. Wiley-VCH Verlag, 2002. c) Cramer, C. J., *Essentials of Computational Chemistry, Theories and Models*, 1<sup>a</sup> Ed. Wiley-VCH Verlag, 2002.
- [4] a) Nicolaou, K. C.; Snyder, S. A. *Angew. Chem. Int. Ed.* 2005, 44, 1012. b) Suyama, T. L.; Gerwick, W. H.; McPhail, K. L. *Bioorg. Med. Chem.* 2011, 19, 6675. c) Maier, M. E. *Nat. Prod. Rep.* 2009, 26, 1105.
- [5] a) Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. *Chem. Rev.* 2012, 112, 1839. b) Pisano, P.L.; Sarotti, A.M.; Pellegrinet, S.C. *Tetrahedron Lett.* 2009, 50, 6121. (c) Sarotti, A.M.; Suárez, A.G.; Spanevello, R.A. *Tetrahedron Lett.* 2011, 52, 3116.
- [6] London, F. *J. Phys. Radium* 1937, 8, 397.
- [7] a) Smith, S. G.; Goodman, J. M. *J. Org. Chem.* 2009, 74, 4597. b) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* 2010, 132, 12946.
- [8] Grimblat, N.; Zanardi, M. M. Sarotti, A. M. *J. Org. Chem.* 2015, artículo en prensa.
- [9] a) Zanardi, M. M.; Sarotti, A. M. *J. Org. Chem.* 2015, 80, 9371. b) Sarotti, A.M. *Org. Biomol. Chem.* 2013, 11, 4847.
- [10] a) Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* 2009, 74, 7254. b) Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* 2012, 77, 6059.
- [11] a) Zupan, J.; Gasteiger, J. *Neural networks in chemistry and drug design*. John Wiley & Sons, Inc. 1999. b) Haupt, R. L.; Haupt, S. E. *Practical genetic algorithms*. John Wiley & Sons. 2004.