



## LA INTELIGENCIA ARTIFICIAL EN LA IDENTIFICACIÓN DE NUEVAS MOLÉCULAS

*En las últimas décadas se han introducido varios métodos computacionales para reducir los tiempos y costes del descubrimiento de fármacos, así como para mejorar la calidad del proceso de desarrollo y elevar su tasa de éxito.*



**Shin-Ho Kim Lee**, CTO de Altenea Biotech en Parque Científico de Madrid

**Roi Naveiro**, Research Scientist de Altenea Biotech en Parque Científico de Madrid

### **EL DISEÑO Y DESARROLLO TANTO DE NUEVOS FÁRMACOS, ASÍ COMO DE MOLÉCULAS QUÍMICAS ES UN PROCESO ENORMEMENTE COMPLEJO.**

En el mundo farmacéutico, los números ponen de manifiesto esta realidad: de media, descubrir un nuevo fármaco y llegar a comercializarlo requiere entre 12-15 años y, dependiendo del tipo de patología, el coste de este proceso puede elevarse hasta los 2.500 millones de euros. A todo esto, hay que sumarle una elevadísima tasa de fracaso: alrededor del 90-95% de potenciales medicamentos que se testean durante el proceso de descubrimiento de un fármaco nunca verán la luz. Además, no debemos olvidarnos de los animales que utilizamos, por ejemplo, para determinar la toxicidad de las moléculas. Estas cifras, en mayor o menor medida, son extensibles a otros sectores industriales donde se desarrollan nuevas moléculas, como es el diseño de materiales.

El descubrimiento y desarrollo de fármacos puede verse como una carrera de obstáculos en la que muchas moléculas competidoras comienzan, pero muy pocas llegan al final. A modo de ejemplo, entre 2002 y 2012, la tasa de fracaso en el desarrollo de nuevos fármacos para el tratamiento de la enfermedad de Alzheimer alcanzó el 99,6 %: lo que supone un 0,4% de éxito: de 244 compuestos evaluados se aprobó sólo la comercialización de un fármaco.

En las últimas décadas se han introducido varios métodos computacionales para reducir los tiempos y costes del descubrimiento de fármacos, así como para mejorar la calidad del proceso de desarrollo y elevar su tasa de éxito. Sin embargo, aún queda mucho trabajo por hacer. Una de las vías más prometedoras para agilizar este proceso consiste en la adopción de

herramientas basadas en inteligencia artificial (IA). De hecho, se están aplicando diferentes técnicas de IA como el machine learning (aprendizaje automático) o deep learning (aprendizaje profundo), en todas las etapas del proceso de desarrollo de fármacos, incluida la identificación y validación de dianas terapéuticas, la búsqueda y optimización de cabezas de serie, la síntesis de moléculas orgánicas, la predicción de propiedades tan importantes como las conocidas como ADME (Absorción, Distribución, Metabolismo y Excreción) y por supuesto la toxicidad e, incluso, en ensayos clínicos.

Las herramientas y estrategias empleadas a lo largo del proceso del desarrollo de los fármacos son diversas variando en función del objetivo al que nos enfrentemos. En este artículo nos vamos a centrar con un poco más de detalle en dos aproximaciones empleadas en la identificación y desarrollo de nuevas moléculas claves en diferentes sectores químicos, no solo en el descubrimiento de fármacos, como es el desarrollo de modelos QSAR (de las siglas en inglés Quantitative Structure Activity Relationship) y diseño de novo.

### **MODELOS PREDICTIVOS QSAR**

Contar con una estimación del valor de ciertas propiedades de las moléculas, ya sean fisicoquímicas, farmacológicas o toxicológicas, sin la necesidad de sintetizarlas; tiene un sinfín de aplicaciones en diversas áreas. Por ejemplo, en el diseño de fármacos, el cribado virtual de miles de estructuras acelera enormemente las primeras fases de este proceso.

Tradicionalmente, la estimación de las propiedades de moléculas se realiza mediante técnicas de química computacional. No obstante, los cálculos involucrados

# THE ONE-HOUR MYCOPLASMA TEST ANYONE CAN USE ANYWHERE.



2020 bioMérieux, Inc. • BIOMÉRIEUX and the BIOMÉRIEUX logo and BIOFIRE are used, pending and/or registered trademarks belonging to bioMérieux, or one of its subsidiaries, or one of its companies.

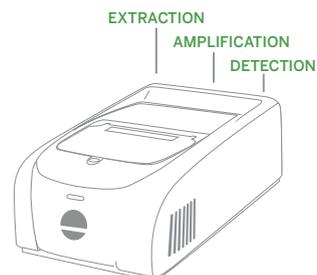
## Introducing BIOFIRE® *Mycoplasma*

With just two minutes of hands-on time, BIOFIRE® *Mycoplasma* delivers accurate results in under an hour, with no PCR lab or specialized skills required. The compact, easy-to-use interface means nearly anyone can test for Mycoplasma anywhere, anytime, and at any point in the manufacturing process.

### The countertop molecular biology lab

The automated FilmArray® instrument performs extraction, amplification, and detection of nucleic acids in a single machine, providing a simple and rapid *mycoplasma* detection in less than an hour.

Learn more about BIOFIRE® *Mycoplasma* at [biomerieux-industry.com/mycoplasma](https://biomerieux-industry.com/mycoplasma)



son muy lentos y requieren un gran poder computacional. En las últimas décadas, la aplicación de modelos de machine learning (ML) ha supuesto una auténtica revolución al reducir drásticamente el tiempo de cómputo de las propiedades y, por consiguiente, permitir la evaluación de muchos más compuestos en fases de cribado.

Los modelos QSAR son modelos matemáticos donde se relaciona cualquier propiedad (físicoquímica, farmacológica, toxicológica...) con descriptores moleculares de una estructura química. Los modelos se entrenan a partir de bases de datos de moléculas con valores de propiedad conocidos. Una vez entrenado el modelo, éste es capaz de predecir el valor de la propiedad de interés en moléculas no vistas durante el entrenamiento.

Estos modelos QSAR son muy interesantes y útiles para cribar grandes quimiotecas de compuestos y sintetizar o estudiar experimentalmente sólo las moléculas más prometedoras (según el modelo), ahorrando tanto tiempo como dinero. Cabe destacar que estos modelos QSAR también tienen aplicaciones tremendamente útiles fuera del sector farmacéutico: en general, en cualquier sector donde se tengan que identificar y/o diseñar nuevas moléculas, como por ejemplo en sectores como la agricultura o el refino, los modelos QSAR son de gran ayuda.

## DISEÑO DE NOVO DE MOLÉCULAS

Si bien el empleo de los modelos QSAR resulta útil para evaluar y cribar librerías de moléculas existentes, éstos carecen de la capacidad de operar de forma inversa para revelar la estructura de las moléculas de máxima actividad.

Por tanto, el usuario debe diseñar posibles nuevas moléculas en base a su experiencia, interaccionando con los modelos existentes para encontrar los candidatos idóneos. Este proceso de prueba y error suele ser

largo y tedioso. Además, a esto se suman los posibles sesgos que podría introducir inconscientemente el usuario en los patrones estructurales que se propone explorar.

El diseño de novo tiene como objetivo generar nuevas moléculas químicas con las propiedades deseadas

de forma completamente automática, sin intervención humana. Existen diferentes estrategias, que difieren en gran medida en el proceso de generación de la estructura química y los métodos de puntuación empleados. En los últimos años, el desarrollo de métodos generativos de moléculas de novo se ha convertido en un ámbito de investigación de intensa actividad.

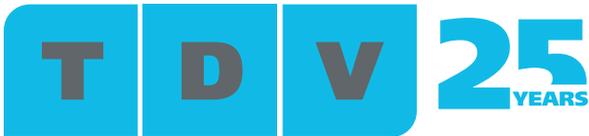
El uso de un método automático de generación de moléculas puede suponer una revolución

sin precedentes en el diseño de fármacos y de materiales. Las grandes librerías de moléculas virtuales generadas utilizando estos métodos permitirían sondear amplias zonas del espacio químico de manera rápida y con apenas intervención humana. Como consecuencia, se facilitaría la identificación de los mejores candidatos moleculares para la aplicación de interés, además de reconocer posibles patrones de estructura-actividad interpretables que guíen a los expertos en química.

Para finalizar se puede concluir, sin ningún género de dudas, que durante los últimos años se ha demostrado la utilidad de las herramientas de IA en el desarrollo de fármacos debido al potencial de los modelos predictivos para navegar en grandes bases de datos químicas. Sin embargo, quedan varios desafíos sobre los que hay que profundizar como son: i) la disponibilidad de conjuntos de datos sólidos y apropiados; ii) una estrategia de novo robusta, precisa e interpretable para explorar el espacio químico; iii) el uso de optimización multiobjetivo para poder optimizar de forma simultáneamente varias propiedades y iv) la reducción de los tiempos de los ciclos de entrenamiento ●



*El uso de un método automático de generación de moléculas puede suponer una revolución sin precedentes en el diseño de fármacos y de materiales*



COMPLIANCE | TECHNOLOGY

+300 CUSTOMERS

+2500 PROJECTS

# MEDICAL DEVICE CONSULTANCY

QUALITY AND REGULATORY COMPLIANCE

## FROM SMALL START-UPS TO BIG COMPANIES

- Quality Management System (**QMS**)
- **Technical Documentation** (CE Mark & FDA)
- Medical device **risk management**
- Management of the **AEMPS License** (companies established in Spain)
- Medical Device Software lifecycle according to **IEC 62304**
- Medical Device Software **functional verification and usability testing**
- **Qualification** and validation activities (equipment, software, processes)
- Compliance with **distribution** requirements
- Internal **audits** & audits of critical **suppliers**
- **Training** services

IVD  
BIOTECH  
MDSW

COMMITTED TO QUALITY TO DRIVE INNOVATION



Ronda General Mitre 147, E-2  
E-08022 Barcelona, Spain.  
Phone: +34 93 417 80 65  
[info@tdvct.com](mailto:info@tdvct.com)

888 Wanhangu Rd, Office 13-I Kaikai  
200041 Shanghai, R.P China  
Phone: +86 21 32200167  
[vwang@tdvct.sh](mailto:vwang@tdvct.sh)

[www.tdvct.com](http://www.tdvct.com)