MEDINA PREDICTS A NEW ERA FOR NATURAL PRODUCTS IN MODERN DRUG DISCOVERY. HERE, THE ORGANISATION DISCUSSES NEW TRENDS IN DEREPLICATION, FRACTIONATION AND STRUCTURAL ELUCIDATION

Modern drug discovery

Natural product (NP) research continues to be one of the most prolific sources for obtaining new classes of compounds and new drugs. Although discontinued at many pharmaceutical companies over the last two decades, NP drug discovery still offers a wide range of opportunities for the discovery of new leads. Technical difficulties inherently associated with the isolation and structural characterisation of bioactive constituents in extracts and the high rate of rediscovery of new compounds have been frequently cited as two of the main reasons for the gradual abandonment of NP discovery programmes by most pharmaceutical companies in favour of other alternatives such as combinatorial chemistry, which has not turned out to be the panacea that it was expected to be, with only a few drugs discovered in combination with high throughput screening (HTS) programmes.

Dereplication in NP research

This situation and development of new dereplication strategies and improved analytical tools have partially helped to overcome the aforementioned problems inherent to NP research and have led to a renewed interest in NP in drug discovery. Dereplication can be defined as the early identification of molecules present in an extract or mixture of compounds without the need to perform a separation. It constitutes nowadays an essential step in the drug discovery process from natural sources as it allows the selection of extracts containing the most interesting/novel bioactive molecules for fractionation and determination of its structure, and is crucial to the fast discovery of novel molecules.

The last two decades have seen a revolution in the development of new dereplication strategies. These usually consist of the combination of a separative analytical technique – i.e. thin-layer chromatography (TLC), gas chromatography (GC), capillary electrophoresis (CE), solid phase extraction (SPE), column chromatography (CC), flash column chromatography (FCC), high performance liquid chromatography (HPLC) or ultra-high performance liquid chromatography (uHPLC) – with a detection method – i.e. ultraviolet/visible spectroscopy (UV-vis), mass spectrometry (MS) in the low (LRMS) or high resolution (HRMS) mode, or nuclear magnetic resonance (NMR). The remarkable advances in analytical instrumentation alongside the development of suitable databases have enabled the development of the fast dereplication processes required in current drug discovery programmes based on NPs.

Among the separative techniques, liquid chromatography approaches such as HPLC or uHPLC in the reversed-phase (RP) mode are preferred as they suit the polarity and features of most drug-like secondary metabolites, but other techniques such as hydrophilic interaction chromatography (HILIC) can also provide the ideal solution in the separation of small or very polar compounds. Regarding the detection methods, UV-vis approaches have proved to be useful only when used in combination with in-house databases or as complementary data after an MS-based search. HRMS is the preferred approach when applying MS-based dereplication strategies. It allows the assignment of molecular formulae (MFs) for most of the peaks in an HPLC trace and the direct search of these MFs in most databases.

Additionally, MS/MS fragmentation used in combination with computational methods for predicting fragmentation patterns offers a robust alternative for the identification of molecules. Finally, the use of NMR via the hyphenated techniques HPLC-NMR or HPLC-SPE-NMR offers a much more powerful tool for dereplication and is essential in some cases to undoubtedly establish the structure of a given compound. Integrated with UV and HRMS, data provides in most cases the definite answer for the dereplication of a given molecule, even if only a proton NMR spectrum has been acquired. The development of low volume NMR probes and cryoprobes over the last decade allows the acquisition of spectra with just micrograms of sample obtained by analytical or semi-preparative HPLC. The use of NMR diffusion ordered spectroscopy (DOSY) even allows the dereplication of molecules present in low complexity mixtures, eliminating the need to perform one or several time-consuming step(s) of chromatographic separation.

The use of appropriate databases also plays a major role in all the dereplication processes mentioned above. Ideally, they must contain as much information as possible on individual compounds and be searchable by substructure, structure identity/similarity and spectroscopic identity/similarity. Desirable query fields include the trivial name, the molecular weight (MW), the accurate mass, the molecular formula (MF), the taxonomy of the producing organism, the biological activity of the compound, and the UV absorption maxima. Those allowing experimental accurate mass searches and/or containing actual NMR spectra or chemical shifts (either experimental or calculated) are preferred.

The CAS Registry (accessible through SciFinder) PubChem, ChemSpider, AntiBase, MarinLit or the Dictionary of Natural Products

are databases commonly used by researchers involved in the dereplication of NPs. Versions of these databases containing the so-called 'structural features' searchable field allow the introduction of certain features (number of methyl groups, alkenes, carbinol protons, etc.) that can be read directly from NMR spectra and are impressively effective at discriminating between alternative candidate structures and isomers. Additionally, some research groups have also developed their own databases, tools and dereplication strategies. These are usually based in the classes of natural products under study in their laboratories and include HPLC or uHPLC retention times in a given chromatographic system, UV, HRMS, MS/MS fragmentation and NMR data in most cases.

Fractionation and structural elucidation tools

The traditional HTS of extracts in NP research had inherent problems associated with the low concentration of some components in the matrix, the presence of interference or nuisance compounds or the additive or synergistic effects of several compounds that disappeared after separation. In some enzymatic assays, the use of extracts is not recommended due to the high rate of false positives generated during the HTS campaigns. The use of fractions of reduced complexity has enormously simplified the process, and pre-fractionation HTS strategies have been shown to give better hit rates and yields in terms of the final number of bioactive compounds isolated. Several pre-fractionation strategies have been applied for the generation of collections of fractions ready for HTS, all of them including one of several steps of chromatographic separations, mainly RP HPLC, which can be done in a fully automated mode in most current research laboratories.

On the other hand, recent developments in instrumentation and technology have greatly facilitated the structural elucidation of natural products. The advent in the last decade of microcoil and microtube low volume NMR cryoprobes has allowed the measurement of NMR spectra using extremely small amount of compounds. NP drug discovery has greatly benefitted from this fact, and the determination of structures of molecules isolated in low amounts is nowadays possible in reasonable timeframes.

Apart from the aforementioned advantages of the application of such technologies for the NMR-based dereplication of compounds, structures of new compounds isolated as part of an HTS programme can now be determined using a few micrograms of material. This allows the assessment of their novelty and their interest as possible candidates to enter the drug development phase without the need to spend the time and resources required to isolate the milligram amounts needed for structural elucidation just a decade ago. Even the amount of compound isolated from small extracts, and hence its yield, can be accurately calculated using NMR quantitation approaches based on ^{13}C NMR satellites. When a few micrograms are isolated, gravimetric determination of this amount on a standard analytical balance is just an impossible task. With the improved capabilities of NMR in terms of elucidating the structure of novel molecules, bottlenecks in the drug discovery process are now moving from small-scale structure elucidation challenges to the development of robust and reliable nanoscale biological assays.

In conclusion, NP-based drug discovery programmes have experienced a re-emergence due in part to the development of new technologies and analytical tools that have reduced, if not eliminated, some of the traditional barriers associated to screening NPs in HTS assays against biological targets. The continuous improvement of these technologies makes it hard to predict what the future might bring in terms of the discovery of new drugs from natural products, but a new era in NP drug discovery is certainly to come.

About MEDINA

MEDINA is a non-profit research organisation established in the Health Sciences Technology Park, Granada (Spain), as a public private partnership between the government of Andalucía, the University of Granada, and Merck Sharp and Dohme de España S.A. Leveraging our experience in NP research we have developed a robust platform for the dereplication of natural products, including in-house built LR and HRMS databases, HRMS combined with commercial database searches, and NMR and DOSY strategies used in combination with searches in proprietary or commercial NMR structural features databases. Additionally, our chemistry laboratories are fully equipped with instrumentation for the generation of collections of NP fractions, as well as the isolation and structural elucidation of molecules, taking advantage of the use of an HR mass analyser and a 500MHz spectrometer equipped with a low volume 1.7mm microcryoprobe. This platform is used within the joint discovery efforts to accelerate the discovery of novel bioactive natural products developed under the collaborative public private innovation models implemented at MEDINA.

Fernando Reyes
fernando.reyes@medinaandalucia.es

Ignacio Pérez-Victoria
ignacio.perez-victoria@medinaandalucia.es

Jesús Martín
jesus.martin@medinaandalucia.es

Olga Genilloud
olga.genilloud@medinaandalcuia.es

Fundación MEDINA

+34 958 993 965

www.medinadiscovery.com