Beginner's guide to the identification of bioactive compounds in plant extracts

Many researchers are daunted by the prospect of the work required to identify the structures of active compounds in hits from natural product screens. However, recent advances in high-performance liquid chromatography (HPLC), mass-spectrometry (MS) and microcoil nuclear magnetic resonance (NMR) techniques, together with ever improving online databases of phytochemical m/z signatures, have dramatically simplified the identification of compounds in natural products.

The following notes offer a basic explanation of the key steps involved in progressing from the identification of hits, to identification of lead compounds in natural product extracts.

Step 1: Perform screen of library using primary assay

The primary screen of the target of interest is ideally performed in a microtitre plate format (typically 96- or 384-wells), but since the library has been optimised to maximise hit rate, it is also feasible to screen each extract in lower throughput assay formats. 'Hits' are sometimes defined by an arbitrary cut-off (such as >80% inhibition of response), and sometimes by an aribitrary definition of a suitable number of hits (e.g. take forward the 10 best inhibitors from the screen).

Hits can also be defined on the basis of a 'Z-prime' calculation, which takes into account the assay variability in positive and negative controls, and the percent inhibition induced by test extracts. A useful introduction to this concept, and hit discovery in general, is provided in this online video, presented by Drs Jim Wells and Michelle Arkin of UCSF.

https://www.youtube.com/watch?v=KQKppAAR3ME

Step 2: Triage hits against appropriate counter-screen assays

Following the identification of hits through the primary screen, the next step is generally to triage those extracts with putative activity using an appropriate counter-screen to weed out those extracts with non-specific activating/inhibitory activity. Such assays could be based, for example, on a related but non-relevant receptor or pathway, or an alternative assay readout (e.g. luminescent reporter if the primary screen is based on fluorescence). Cytotoxicity assays (e.g. MTT) will also help determine whether the observed effects of hits are related to non-specific effects on cell viability.

Step 3: Separate lead extracts into multiple fractions, and re-assay

Next, the active compound(s) present in those hits that survive triage will need to be separated from the crude extract. This can be done in the first instance using very simple and inexpensive low-resolution silica gel chromatography. A simple method for the fractionation of extracts using silica gel chromatography is provided here.

Beginner's guide to the identification of bioactive compounds in plant extracts

(continued)

Fractions containing activity can then be pooled and separated further using HPLC. Alternatively, HPLC can be attempted on unfractionated crude extracts. Because most biological assays require little material, standard C18 analytical columns (e.g. 3 mm inner diameter) are often suitable for this early stage separation.

An appropriate amount of extract to inject on such a column for preliminary separation might be in the range of 300 µg, and standard reverse phase solvent gradients (e.g. 5 to 100% acetonitrile) often separate natural compounds effectively. The delay between detector and fraction collector should be minimised by using the shortest length of small diameter teflon tubing possible, to aid matching of the chromatogram with the activity profile from the bioassay.

Peaks co-inciding with biological activity should then be investigated further, typically by pooling fractions surrounding the peak and performing a second, more focussed round of HPLC separation on the pooled material. Peaks matching biological activity at this stage are often of sufficient purity (>90%) to proceed to structural identification of the hit compound by MS and/or NMR.

Step 4: Identify structure of compound(s) responsible for biological activity in active fractions

Finally, the molecular identify of the compound(s) responsible for the biological activity in active fractions will need to be explored. This can now be achieved with very small amounts of purified compound (e.g. <1 mg), typically using (U)HPLC coupled to high resolution MS and NMR spectroscopy. An accurate m/z value for the candidate molecule is particularly useful, as this information can be used to narrow a search of online databases of natural product structures (see below for links). Those compounds of equal m/z which are commercially available can be retested using the primary bioassay to confirm the identification. The MS and NMR stages of compound identification can be conducted in collaboration with academic partners, or by companies offering such services.

We will be happy to help

If you would like assistance with separation or identification of compounds from your hit extracts, we will be happy to help. Please visit our website help pages, or contact us for further details of the services we offer.