Comprehensive extraction and NMR-based Metabolomics : novel approaches to natural products lead finding in drug discovery

Yuliana, N.D.

Chapter 9

Concluding remarks and perspectives

Nancy Dewi Yuliana[1,2]

[1]Dept of Pharmacognosy, Section.Metabolomics, Institut of Biology, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.
[2]Dept of Food Science and Technology, Bogor Agricultural University, IPB Darmaga Campus, Bogor 16680, Indonesia.

A number of scientific papers introducing new methods to narrow the gaps between natural products and the demand of modern drug discovery have been published recently. Some studies adapted new technologies previously applied in the development of new synthetic drugs, such as generating a high quality natural products library and applying virtual screening to improve the 'hit' quality (350, 364, 375), while other focused on improving the extract fractionation methods (351), or developing the online bioassay and biochemical detection (366-368, 467). The work described in this thesis is a contribution focusing on the improvement of dereplication and identification steps since these two are the major hurdles in natural product-based drug discovery. We introduced a new approach by integrating a new extraction method, namely comprehensive extraction, with NMR metabolomics. The supervised multivariate data analysis, i.e. partial least square (PLS) and orthogonal-partial least square (OPLS) analysis were used as a statistical method to study the correlation of metabolite profile of the extracts with their bioactivities. Several bioassays related to obesity were chosen as a screening tool since this metabolic syndrome has become a serious global health problem in developed and developing countries.

The first three chapters (**Chapter 1, 2, 3**) focused on the classical method of drug discovery from plant sources, starting with a literature study on plants having bioactivities related to obesity (**Chapter 1**). Thirty eight plants reported to have anti-obesity activity were reviewed and categorized based on the targeted biochemical pathways. It is clear from this chapter that a huge effort is going on in anti-obesity compounds prospecting from plants. However, this extensive work is not yet counterbalanced by the success of the active compounds/extracts discovered to go

further to clinical trials. In the next chapter (**Chapter 2**) we screened 39 medicinal plants and spices for four obesity-related bioactivities: the binding activity to adenosine A1 receptor and CB1 receptor, TNF-α inhibition, and induction of lipolysis in 3T3-L1 adipocyte. Nine plants were found to be very active (higher than 75% binding activity) in one or more of the tested bioassays. Several reference compounds present in the active spices were tested in the respective bioassays but they did not show any activity. The high activity could be due to yet unidentified compounds or to synergy between different compounds present in the spices. Further studies are needed to confirm this.

The binding activity of flavonoids to the adenosine A1 receptor has been well-studied (300, 325, 326). After considering the chemical composition of plants which have medium to high binding activity, we followed up the screening work with two plants and chose the adenosine A1 receptor bioassay as a screening tool: *Orthosiphon stamineus* (**Chapter 3**) and *Morus alba* stem bark (result not shown). These two plants have been reported to have many flavonoids and flavonoid derivatives. Seven active flavonoids could be identified from *O. stamineus*, while from *M. alba* stem bark two compounds are isolated from one of the active fractions but they are inactive in a pure form. From the first three chapters it is obvious that to fully utilize the benefit from plants as a drugs source is not a simple task as synergy may play an important role.

Being inspired by the remarkable development of metabolomics in plant science, we hypothesized that metabolomics, particularly untargeted metabolomics, could be applied in the study of drug discovery in plants (**Chapter 4**). For the initial study, we recorded the chemical profile of four *O. stamineus* fractions obtained from single solvent extraction followed by liquid-liquid partition (*n*-hexane, choloroform, butanol, and water fractions) with thin layer chromatography (TLC), tested the adenosine A1 receptor binding activity of each fraction, and eventually correlated both data by using orthogonal partial least square analysis (OPLS) method (**Chapter 5**). Orthogonal-partial least square analysis was found to be a suitable tool in mapping the chemical profile and bioactivity matrices to identify the active compounds. Thin layer chromatography is a simple and inexpensive analytical tool which separation depends mainly on adsorption and partitioning characteristics of adsorbent and sample. Other benefits of TLC include the inert character of adsorbent material which allows the use of stronger corrosive reagents, and the capability to run standards in parallel under

identical conditions (468). However, due to the low resolution of TLC, overlapping of spots hampers the application as metabolomics tool using crude extracts. To solve the problem of the complexity of crude extracts, fractionation is a possible solution. Another problem in metabolomics is the extraction method, the choice of the extraction solvent limits the metabolites extracted on the basis of polarity. We thus developed an approach that could solve both problems in a simple way. The method is based on extraction with a solvent mixture with a continuous increasing polarity gradient. To fully exploit the potential of this approach it is necessary to choose a detection tool with a higher resolution than TLC. Nuclear magnetic resonance spectroscopy (NMR) has been considered to be one of analytical methods which allows the robust detection of a broad range of compounds in a single run (13). A comprehensive extraction method was thus developed, and NMR was used to measure the chemical profile of the obtained fractions. The reproducibility of the comprehensive extraction was assessed by hierarchical clustering anaysis (HCA) and partial least square-discriminant analysis (PLS-DA) (**Chapter 6**), and the statistical and chemical validity were also assessed (**Chapter 7**). The results so far satisfied our expectation.

However, to apply this method successfully it is very important to have a good resolution in each fraction that means good spreading of the metabolites over the fractions. The solvents combination should thus be chosen carefully. In testing different combinations we found that a wider range of polarity gives better resolution as compared to a narrow one (**Chapter 7 and 8**). The solubility of the targeted active compounds should also be taken into consideration. The other essential point is to eliminate or to reduce the availability of false positive compounds from the extracts, such as unsaturated fatty acids in the case of adenosine A1 receptor binding bioassay. Separating defatting step from the extraction did not give the expected result. It is not possible to completely eliminate these compounds from the extract, rather the solvents choice and the gradient scheme should be able to localize compounds giving false positive results into certain fractions which do not overlap with the targeted active compounds.

In conclusion, the studies in this thesis show that comprehensive extraction in a combination with NMR metabolomics has the potential to improve the dereplication and identification steps of drugs discovery from plant extract. With comprehensive

extraction we are able to extract the total metabolome of a sample (i.e. plants material). Certain compounds or class of compounds are present in several fractions with different concentration. Qualitatively this can then be measured with any of the metabolomics tools (e.g. NMR, GC-MS, and LC-MS). By using this approach in combination with supervised multivariate analysis one may identify the signals that correlate with the activities measured for the different fractions obtained. A major advantage of this approach is that in principle after the first round of bioassays, signals due to compounds correlating with activitiy can be used for monitoring further fractionation. Particularly with complex *in-vivo* assays (cell cultures, whole organisms, and clinical trials) this very much simplifies the leadfinding. Even in such systems pro-drugs and synergism can be discovered.

The here described approach is so far only applied in combination with an *in-vitro* assay, but these experiments show that the method is very useful to identify (groups of) active compounds. But in my view it would particularly be very useful in studying medicinal plants in a systems biology type of approach. The use of the highly reproducible NMR metabolomics analysis in such studies has the advantage of building up a public database for rapid identification of known compounds, and of the metabolic fingerprints of the medicinal plants.