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

Yuliana, N.D.

General Introduction

Nancy Dewi Yuliana [1, 2]

[1]Div. Pharmacognosy, Section Metabolomics, Institut of Biology, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

[2]Dept. Food Science and Technology, Bogor Agricultural University, IPB Darmaga Campus, Bogor 16680, Indonesia.

Obesity has become a serious epidemic disease in both developed and developing countries. Currently, only one drug has been approved worldwide to be used for long time obesity treatment: Orlistat, a lipase inhibitor. Two appetite suppressants, Sibutramine and Rimonabant have been withdrawn from the European Union market due to their side effects. Few years before, the herbal *Ephedra sinica* was banned by the FDA for the same reason. This indicates that there is a need to find new effective and safe anti-obesity drugs. Whether the present advanced drug discovery technology, including high throughput screening and combinatorial chemistry, is able to address this challenge is questionable, as indicated by the decreasing number of new drugs resulting from this new technology during the last two decades. Sorafenib, a multikinase inhibitor indicated for advanced renal cancer, is the only FDA approved drug for clinical use released from combinatorial chemistry so far (1).

Nature is already well known as a distinguished drugs source. The failure of aforementioned new technologies to deliver new drugs has renewed the interest of pharmaceutical companies and academia to look back to this source. In fact, non-prescription slimming preparations available in the market are predominantly from natural sources. Orlistat is also a natural product derivative (it is a saturated derivative of lipstatin, a microbial lipase inhibitor). Among those commercial preparations, plant-based ingredients are predominant. This indicates that plants could be an interesting source for novel anti-obesity drugs. Different approaches to explore biodiversity can be considered. One of them is the screening of traditionally used plants which could be a successful way to uncover lead candidates. Conventional drug discovery which is based

on bioassay guided fractionation is widely applied to study traditional medicines. This method has several advantages, but it has a long time-line and it is quite elaborate to discover active compounds, besides that possible synergism or pro-drugs in herbal medicines cannot be detected. In order to compete with modern drug discovery (which is predominantly applied to synthetic drugs), a new method for studying natural sources for rapid identification of active constituents while also synergism and prodrugs can be observed, is much desired. For example, by linking the metabolite profile of a herbal preparation with its bioactivity. Some studies introducing this method have been reported, in most cases metabolomics combined with projection-based multivariate data analysis (PLS-DA, PLS) was used (2-5).

Obesity results from the imbalance between energy intake and energy expenditure, consequently several strategies can be applied for obesity drug development: reduction of energy intake by appetite suppression; inhibition of nutrient absorption; increase of energy expenditure; and modulation of fat (6). Several G-protein coupled receptors have been used as targets in anti-obesity drug development, for example the adenosine A1 receptor, and the cannabinoid CB1 receptor. Blocking the adenosine A1 receptor by an antagonist or inverse agonist has been reported to correlate with lipolytic activity (7, 8), while blocking the cannabinoid CB1 receptor reduces the appetite and stimulates lipid metabolism (9). At another cellular level, an increase level of TNF-α in adiposity was found to correlate with obesity related hyperleptinemia (10).

Metabolic profiling by nuclear magnetic spectroscopy (NMR) was for the first time applied by Schripsema *et al*. in 1988 (11) to plant cell cultures to study cell line stability. Beginning of this century metabolomics really took off and an ever increasing number of papers reported applications in different fields of plant sciences, including quality control of herbal materials and phylogenic studies. Several attempts to statistically correlate the metabolite profile of a plant extract with its bioactivity have also been reported. Theoretically, this approach allows the identification of active compounds from crude extracts. Therefore, this approach could potentially overcome the above mentioned problems in drug discovery from natural sources, particularly from plant extracts.

There are three important factors to consider when one wants to apply this approach: extraction, detection, and statistical methods. To find correlations between a

metabolite profile and bioactivity of an extract, an untargeted metabolomics approach is the most suited. This requires an extraction method which can provide as much as possible information about all metabolites present in the extract, from non-polar to polar ones. Next, we need a reliable detection method to record this information. Nuclear magnetic spectroscopy (NMR) is thought to be the most reliable method to study metabolomics due to its excellent reproducibility, though its low sensitivity is a disadvantage if compared to mass spectrometry (1 μM – 1 mM in NMR tube) (12, 13). The simple and fast sample preparation, short measurement time, the availability of automatization and advanced data analysis methods, ease of quantitation plus the possibility to elucidate structures of known or unknown compounds in a complex mixture using advanced two-dimensional (2D) NMR methods, are further advantages (12, 13). The high dimensional data resulting from NMR measurements require advanced chemometric methods such as multivariate data analysis (MVDA) methods for interpretation. More specifically, to see the correlation between the chemical profile of a plant extract, MVDA can be applied to observe a regression correlation between two blocks of data, denoted as X (chemical profile) and Y (bioactivity). The most common MVDA method for this modeling is the partial least square (PLS) method (14). Since a plant extract is a complex mixture of mostly uncharacterized compounds with a large dynamic range (e.g. the difference in the levels of major compounds as compared to that of minor compounds, differences in polarity, in boiling or evaporating points, etc), this regression based MVDA should be applied carefully since its precision increases when more Y-related X are introduced into the equation. A minimum amount of Y-uncorrelated X (e.g. baseline noise, impurities) improves the system. To address this, several extended versions of PLS such as orthogonal signal correction (OSC) and orthogonal-PLS (OPLS) can be applied (15).

After a proper statistical validation, compounds correlated to bioactivity can be elucidated by means of 2D NMR or other methods, and then further confirmation can be performed by chemical validations by testing the reference compounds. The flow chart of this proposed method is presented in Figure 1.

By applying this method, the identification of active compounds becomes easier since they are detected directly in the extract even in the presence of other compounds. Theoretically, synergism or antagonism can also be detected.

Aim of the thesis

This thesis aims to develop a new approach to detect active compounds from crude plant extracts.

This general aim results in several objectives:

- Development of a highly reproducible extraction method that extracts the widest possible range of metabolites.
- Development of fast detection methods for all metabolites.
- Development of chemometric method for identification of compounds related to activity.
- Validation of the total procedure.

Comprehensive metabolite profiling was achieved by developing a continuous extraction method (comprehensive extraction). The metabolite profile of an extract is recorded by means of TLC and NMR. An adenosine A1 receptor binding assay was used to measure the activity of the extracts. Several multivariate data analysis methods were explored and validated. Signals correlating with bioactivity were studied and elucidated by means of 2D NMR. To chemically validate the method, compounds active on the adenosine A1 receptor previously isolated by conventional bioassay guided fractionation were studied. If the proposed model is valid, NMR signals of these compounds should positively correlate with the bioactivity.

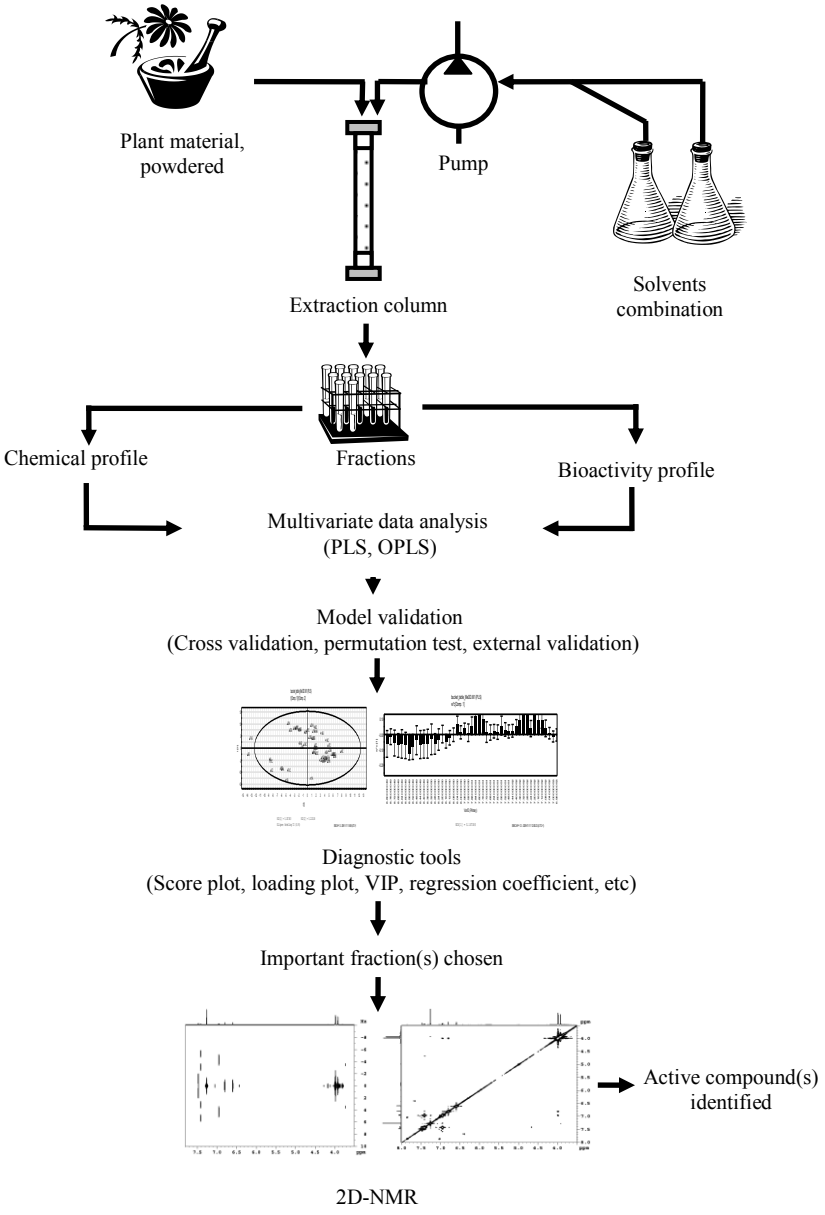

Figure 1. Flow chart of the application of the comprehensive extraction coupled to NMR-based metabolomics to identify active compounds in plant extract.

Outline of the thesis

Commercial slimming preparations from plant based materials are predominant in the market. Therefore, plants are a potential source for new anti-obesity drugs. Traditionally used herbal medicines are an interesting option to explore, as described in **Chapter 1**. To choose a potential plant, screening of various plants including some Asian spices and traditional herbal medicines was performed (**Chapter 2**). Several obesity related bioassays were used: adenosine A1 receptor binding, cannabinoid CB1 receptor binding, TNF-α inhibition, and fat cell lipolysis. *Orthosiphon stamineus* leaves were chosen for a further study with adenosine A1 receptor bioassay. A conventional bioassay guided fractionation approach was used to isolate 7 active compounds from *O. stamineus* (**Chapter 3**). The possibility to use a metabolomics approach coupled to multivariate data analysis for drugs discovery is reviewed in **Chapter 4.** As a preliminary step, *O. stamineus* fractions obtained from a simple extraction method were applied to TLC and then further subjected to OPLS to study the correlation between chemical profile and adenosine A1 receptor binding bioactivity profile (**Chapter 5**). The chemical profiling was then broadened by using comprehensive extraction methods. The obtained fractions were measured by ^{1}H NMR and the results subjected to hierarchical clustering analysis (HCA) and partial least square-discriminant analysis (PLS-DA) to check the reproducibility of the extraction method (**Chapter 6**). The comprehensive extraction method integrated with NMR metabolomics was applied to identify compounds from *O. stamineus* binding to adenosine A1 receptor. Partial least square (PLS) analysis and orthogonal partial least square (OPLS) analysis were used to study the correlation between the chemical profiles of *O. stamineus* and its adenosine A1 binding activity. The model was then statistically validated. The data from chapter 3 were used for chemical validation. The 2D NMR data (J-res, COSY, and HMBC) of one of active fraction obtained from comprehensive extraction were used to identify signals which correlate with bioactivity (**Chapter 7**). A similar approach was applied to identify active compounds of *Morus alba* stem bark extract (**Chapter 8**). Finally, the summary and the future perspective of the results obtained in this thesis are discussed in **Chapter 9.**