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

Yuliana, N.D.

Chapter 8

Comprehensive extraction integrated with NMR metabolomics as a new way of bioactivity guided identification of biologically active compounds: compounds binding to adenosine A1 receptor from *Morus alba* L. stem bark

Nancy Dewi Yuliana[1,2], Dela Rosa[1], Young Hae Choi[1], Robert Verpoorte[1]

[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 Dramaga Campus, Bogor 16680, Indonesia.

Abstract

Comprehensive extraction coupled to NMR metabolomics was applied to the identification of active compounds from *Morus alba* stem bark binding to the adenosine A1 receptor. Orthogonal partial least square analysis was used to estimate which compounds significantly correlate with the activity. Based on the loading bi-plot and Y-related coefficient plot, unsaturated fatty acid signals strongly correlate to the activity. The characteristic NMR signals of alkaloids and prenylated aromatic compounds found to be abundant in this plant do not positively correlate with the activity. Characteristic signals from two compounds isolated from this plant (betulinic acid and morusin) also have negative Y-related coefficients and indeed they are inactive when tested. Aromatic compounds without prenyl or methoxy units are proposed to be the active ones. To further identify the possible active compounds, the comprehensive extraction needs to be optimized, particularly on the selection of a solvents combination which gives a high resolution of the metabolites over the fractions.

Keywords: *Morus alba*, adenosine A1 receptor, comprehensive extraction, orthogonal partial least square, metabolomics

Introduction

Mullberry or *Morus alba* L. (Moraceae) is cultivated mainly for its foliages as silkworms feed (435), but in several European countries such as Greece and Turkey the fruits are also used as a table fruit, marmalade, juice, liquor, and as natural dyes (for wool or cotton) (436). In traditional Chinese medicine different parts of this plant are used to cure various diseases such as diabetes (leaves stem and root bark), tonic, diuretic and high blood pressure (fruit) (435).

Numerous compounds have been isolated from *M. alba* (303, 428, 429, 437-444). Asano *et al*. (428, 445) reported the presence of sugar-mimic alkaloids and polyhydroxylated alkaloids from the water extract of *M. alba* roots and leaves, which showed a weak to medium glucosidase inhibition activity. Other different types of alkaloids (nortropane alkaloids, nor-ψ-tropine alkaloids) and several new amino acids were also isolated from the fruits of *M. alba* (446) but they did not show significant glucosidase inhibition activity.

Prenylated chalcones having moderate cytotoxic activity against certain human cancer lines were isolated by Yang *et al*. from *M. alba* leaves (442). Flavonoid derivatives were also found to be abundant in the roots and leaves of this plant, such as steppogenin-7,4'-di-*O*-β-D-glucoside, a flavanone glycoside which has significant activity in inhibiting the proliferation of human ovarian cancer cells (443). Prenylated isoflavanes were isolated from the leaves (437). Except the glycosides, all show significant anti-oxidant activities. The authors also tested the compounds for radical scavenging activity but no significant activity was shown. Nomura *et al*. (439) reported the presence of prenylated flavones from the root bark but no bioactivity test was performed. Dat *et al*. (447) reported the cytotoxic activity of several prenylated flavones in which morusin was found as the most active one with IC_{50} value of 0.64 μM. Since kaempferol did not show any activity, the authors concluded that the prenyl or geranyl group as well as their cyclization with the hydroxy group increased the cytotoxicity of these flavones. Du *et al*. (303) found that a prenylated flavanone, leachianone G, has a potent anti-viral activity with IC_{50} 1.6 μg/mL. Farnesyl or prenyl group attached to the flavone skeletone seems not important for antiviral activity since other compounds having such structures were tested but no significant activity was found. However, Sohn

et al. (448) suggested the importance of prenyl substituents for the anti-microbial and cytotoxic activities of several flavonoids and phenolic derivatives including those isolated from *M. alba* root bark (albanol B, kuwanon C, morusin, mulberrofuran G, sanggenon B, and sanggenon D).

Several stilbenes were also isolated from this plant. Oxyresveratrol from the twigs of *M. alba* shows tyrosinase inhibitor activity with Ki 9.1 $x10^{-7}$ M (449). Mulberroside A, a glycosylated form of oxyresveratrol which is a major compound found in the water fraction did not show tyrosinase inhibitor acitivity except when the glucoside was hydrolyzed (yielding oxyresveratrol) (450). Mulberroside A shows significant activity as antitussive and anti asthmatic effect in animals (440). Other stilbenes were also found in *M. alba* but there are no further reports on their bioactivities (440, 451).

Another group of compounds isolated from leaves, root bark, and stem bark of *M. alba* have a benzofuran skeleton (452-455). Moracin A, B, E, F, G, and H which have prenyl or methoxy groups were isolated from leaves and all show weak antifungal activity (452, 453). Moracin C is the major active compounds in *M. alba* leaves for oviposition stimulant of mulberry pyralid, *Glyphodes pyloales* (Walker), while moracin N and ώ-hydroxy-moracin N are inactive (454). These three compounds have similar structure (6-hydroxybenzofuran structure with prenyl groups), the difference is only in the presence of hydroxyl substitutes. Yang *et al*. (455) tested moracin N, P, V, W, X, and Y for cytotoxicity against several human cancer cell lines. Moracin X and Y were the most active ones. The authors suggested that an increased number of ring system in these arylbenzofurans, results in stronger cytotoxic activity. The same study reported that moracin N and moracin W, which have prenyl and geranyl groups attached to their benzofurans ring, have only weak cytotoxic activity.

Besides betulinic acid, several prenylated flavones were reported to be present in *M. alba* stem bark and heartwood, those are mulberrin, mulberrochromene, cyclomulberrin, and cyclomulberrochromene (456). There is no report on the bioactivity of these compounds but the stem bark of *M. alba* is traditionally used as purgative and vermifuge, and also to cure chronic bronchitis and and emphysema (457).

The conventional bioassay guided fractionation was applied in these bioactivity-related studies. Although the method has been successful to deliver new

drugs from natural source (458), such as the statins to lower cholesterol level and huperzine A for treating Alzheimer's disease (342), it has some limitations. This aproach is time consuming and thus expensive, which have made it less attractive for pharmaceutical companies. Moreover, this method may not detect compounds giving false-positive results in the early screening assays, and the presence of synergism and pro-drugs in an extract can not be detected (341). To facilitate the exploration of plants as a source for drug development, a new method which can address these problems is needed.

We have recently highlighted the benefits of the use of metabolomics as an alternative tool to identify bioactive compounds from medicinal plants (459). In the present study we used a similar approach to identify compounds from *M. alba* stem bark showing binding activity to the adenosine A1 receptor. We applied a new extraction method we recently developed (chapter 6) to provide several fractions of *M. alba* stem bark which contain fewer overlapping compounds as compared with fractions from single solvent extraction. Before the extraction, a defatting step by *n*-hexane extraction was used to remove unsaturated fatty acids as they bind unspecifically to this receptor. The obtained fractions were then tested for the adenosine A1 receptor binding activity and the metabolite profile was measured by NMR. Orthogonal partial least square (OPLS) analysis with proper validations was used to check which compounds correlate with the adenosine A1 binding activity of the plant. This new method allows the identification of active compounds from unpurified extracts as previously demonstrated in our lab (chapter 7). Further purification steps can then be more focused on the identified active principals, thus the long timeline which is a common hurdle for lead finding in drug discovery from plant extracts, can be reduced.

Materials and methods

Chemicals and reagents

Methanol, *n*-hexane, acetone, ethyl acetate, HCl, NaOH, and DMSO were purchased from Biosolve BV (Valkenswaard, The Netherlands). Tris buffer was purchased from Gibco BRL (New York, NY, USA), [^{3}H]DPCPX (8-cyclopentyl-1,3-dipropylxanthine) was from DuPont NEN, and CPA (N6-cyclopentyladenosine) was

from RBI Inc. (Zwijndrecht, The Netherlands). Kieselguhr (calcined and purified SiO_2) was bought from Fluka Analytical/Sigma Aldrich Chemie GmbH (Steinheim, Germany). All solvents and reagents were of analytical grade.

Plant material

Morus alba stem bark was obtained from Korean Export and Import Federation of Drugs, Seoul, South Korea and identified by Dr. Young Hae Choi (Leiden University).

Comprehensive extraction

Extraction was performed according to ES1 scheme which has been mentioned in chapter 6. The extraction scheme can be seen in Table 1.

Table 1. Extraction scheme: T= time (min), A = ethyl acetate-methanol (1:1), B = methanol-water (1:1). The flow rate was 4 mL/min.

T (min)	Gradient
0 - 10	100% solvent A
10 – 50	100% solvent A to 100% solvent B
50 - 60	100% solvent B

NMR measurement

NMR measurement was performed according to the method described by Kim *et al*. (426). More detail on the NMR measurement procedure can be seen in chapter 7.

Data analysis

The ^{1}H NMR spectra were automatically reduced to ASCII files. Bucketing was performed by AMIX software (Bruker). Spectral intensities were scaled to total intensity and reduced to integrated regions of equal width (0.04) corresponding to the region of δ 0.3–10.0. The regions of δ 4.75–4.9 and δ 3.28–3.34 were excluded from the analysis because of the residual signal of H_2O and methanol-d_4, respectively. Orthogonal partial least square analysis (OPLS) was performed with the SIMCA-P

software (version 12.0, Umetrics, Umeå, Sweden) with scaling based on the Pareto method.

Adenosine A1 receptor assay

The assay was performed as previously described by Chang *et al*. (290) except that the volume of the total mixture in the assay was 200 μL. The detail on the bioassay procedure has been described in chapter 2.

Results and discussion

The binding activity profile of *M. alba* fractions is presented in Figure 1. The activity was concentrated in the first 10 fractions. This is confirmed further with the OPLS score plot which shows that these 10 fractions are separated from the inactive ones (Fig. 2A). From the loading bi-plot, some signals in the aliphatic (1) and aromatic (2) region are found to be abundant in the active fractions (Fig. 2B).

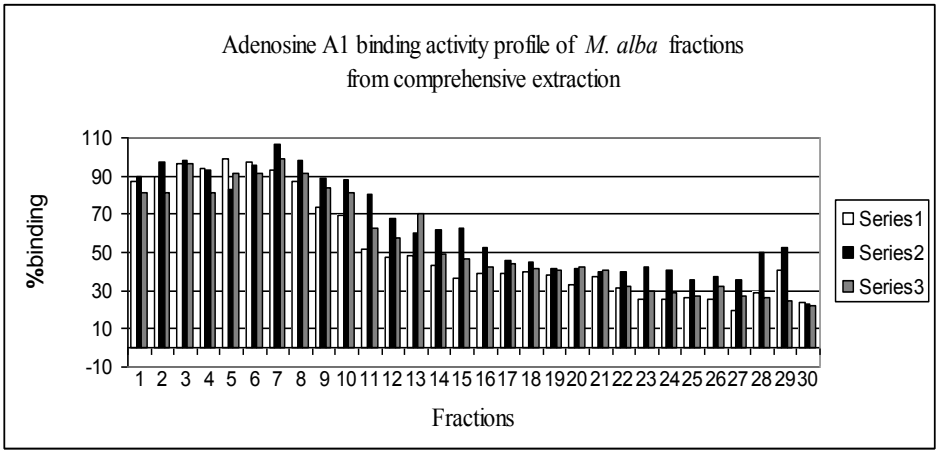
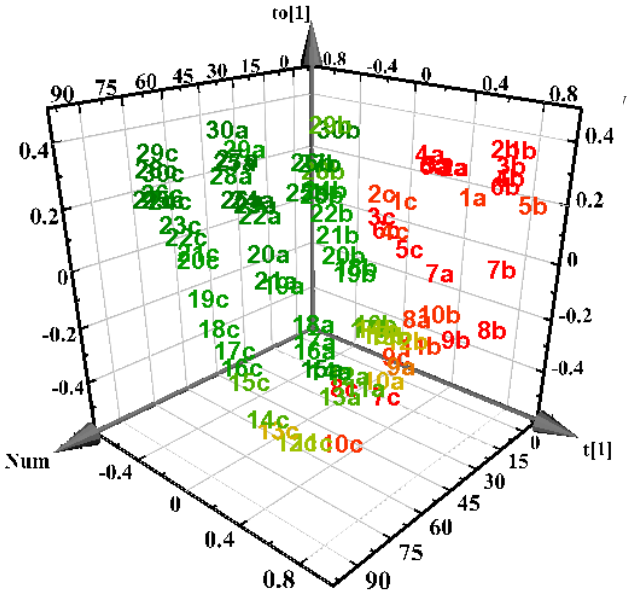


Figure 1. The adenosine A1 binding activity of *Morus alba* fractions obtained from comprehensive extraction.

The model was then statistically validated. From the external validation (in which the data were divided into two: calibration group and test group), the value of root-mean-square errors of prediction (RMSEP) is higher than root-mean-square errors of estimation (RMSEE) indicating a slight overfit. The R^2Y and Q^2 Yvalues after cross validation are close to 1 (0.90 and 0.88 respectively) and the p value from ANOVA

cross-validation is relatively small (2.23. 10^{-36}), showing that the model is statistically acceptable.

A.



B.

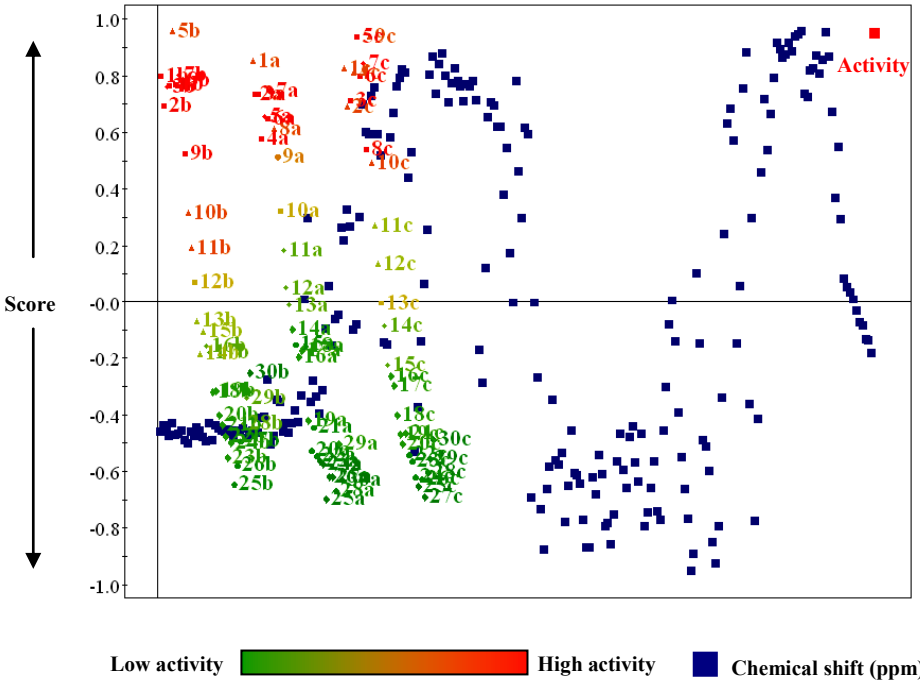


Figure 2. A. The predictive and first orthogonal component of *M. alba* OPLS scores scatter plot, fractions are labeled by number with a-c representing replications. R^2X=48.3%, $R^2XOrthogonal_1$=30.2% **B.** The loading bi-plot: 1=aliphatic region, 2 = aromatic region. Chemical shift with a score value similar to certain fractions are abundant in the respective fraction(s).

Table 1. Statistical validation of *M. alba* OPLS model.

R^2Y	0.90
Q^2Y	0.88
p CV-ANOVA	$2.23.\ 10^{-36}$
RMSEP	12.04
RMSEE	9.12

As can be seen in the OPLS Y-related coefficient plot (Fig. 3), most of aliphatic and aromatic signals have a positive correlation, while most signals between δ 2.40 – 5.16 have a negative correlation. Unsaturated fatty acid signals (δ 1.28 – 1.36) are shown to have the highest positive correlation to the activity. Apparently the defatting steps performed before the extraction did not successfully remove these false positive compounds from the plant material.

From previous work in our lab we also isolated betulinic acid and morusin from *M. alba* stem bark. Both compounds do not show any binding activity to adenosine A1 receptor. Here we checked the correlation of signals present in these two compounds to its adenosine A1 binding activity based on the value of Y-related coefficient from the generated OPLS model (Fig. 3).

Betulinic acid is a pentacyclic triterpenoid. Allthough some of its signals are positive, e.g. methyl protons of C-23 – C-27 (located between δ 0.70 – 1.00), but some signals are located in negative part of the plot, such as two vinyl proton signals at C-29 which are located between 4.50 – 4.60 ppm. Morusin is a prenylated flavonoid, a typical class of compounds found in Moraceae. It has a flavone skeleton with a prenyl group attached to C-3 and a 2,2-dimethylchromene ring connected to A ring. Similar to betulinic acid, some signals of morusin are located in the positive area and some are in the negative area of the Y-related coefficient plot. Signals of methyl protons of the 2,2-dimethylchromene ring, which are located between δ 1.40 – 1.50 (overlap), are

positively correlated with the activity. Some signals from aromatic protons of the flavone skeleton (δ 6.00 – 7.00) are located in the positive coefficient area but few signals are negative. Signals from methylene and vinyl methine protons of the prenyl unit (between δ 3.00 – 5.50) are located in the negative area.

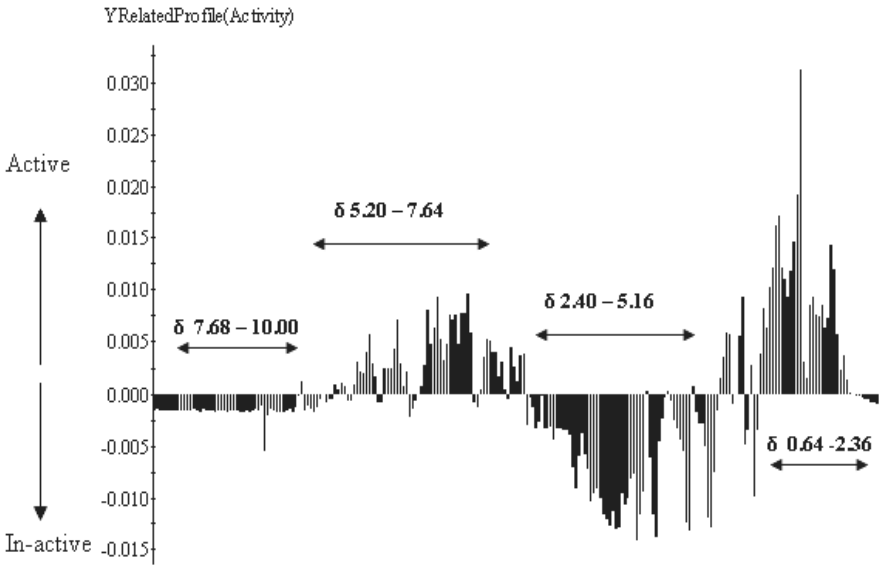


Figure 3. The Y-related coefficient plot.

It has to be noted that certain signals of different compounds may overlap. This may result in a positive correlation with the activity of a signal of non-active compounds, or the reversed. Reducing the bucket size of NMR data from 0.04 into 0.02, 0.01, and 0.001 ppm did not solve the problem. The proton NMR spectra of *M. alba* fractions obtained from comprehensive extraction indeed do not show a good resolution, which can be seen in many overlapping peaks especially in the first three fractions. However, as most compounds have several signals still the active compounds may be traced by this approach. Though, further optimization of the comprehensive extraction should be considered. Particularly, attention should be given to the selection of a

solvents combination which gives a high resolution of the metabolites over the fractions.

However, we further tried to elucidate the possible active compounds. We chose fraction 7b as the most active fraction for 2D NMR analysis. There are several correlations on COSY and HMBC spectra which can be attributed to the skeletons of flavone, flavane, and prenyl units. Unsaturated fatty acids signals are predominant here which can be the reason for the high activity. As previously described, several classes of compounds which have aromatic and non-aromatic structures have been isolated from *M. alba* (303, 428, 429, 437-445, 447, 452, 455, 460-466). From the activity profile, it is clear that the activity cannot be attributed to alkaloids which are abundant in polar fractions. Besides, the areas where alkaloid characteristic signals are located have a negative correlation with the activity. Based on the Y-related coefficient plot, the active principals could be the compounds which have aromatic ring without any groups containing protons which give signals in δ 3.00 – 5.00 (beside alkaloids and sugars, signals within this range can be also from protons of vinyl, methine or methylene groups such as prenyl, geranyl, or farnesyl groups, or methoxy group). This is in accordance with the result from previous work which concluded that compounds isolated from *Morus* species having more complex ring system have stronger cytotoxic activity (455), and that the prenylation is not important for the anti-viral activity (303). Moracin E, moracin M, and moracin X which have a benzofuran ring can be one of them, but not for example moracin A, B, C, F, G, H, N, P, V, W and Y since in the latter group all contain methoxy, formoyl, prenyl or dihydrooxepin substituents and these have signals in the negative area. However, this prediction should be validated further by testing the above mentioned compounds with the respected bioactivity.

Conclusion

Comprehensive extraction integrated with NMR metabolomics is able to identify the correlation of compounds found in *M. alba* stem bark with their adenosine A1 binding activity. Unsaturated fatty acids are found to be strongly correlated with the activity, while alkaloids and prenylated aromatic compounds are predicted not to be active. The chemical validation by using two inactive compounds isolated from this

plant showed that characteristic signals of these two compounds indeed have negative Y-related coefficients. The information obtained from the loading bi-plot and Y-related coefficient plot shows that the active ones should have aromatic structures without any methoxy or prenyl groups attachment. The approach as described here allows to exclude certain compounds for being involved in the activity. By 2D NMR further information on the spin systems of activity related signals can be obtained. This might already result in the identification of the active compound(s). If full structure elucidation of a novel compound is required, isolation will be necessary which also enables final testing of the biological testing. In this approach, no biosssay-guided fractionation is applied, instead fractionation will be monitored on the basis of the targeted signals. This approach is particularly of interest for in combination with complex *in-vivo* bioactivity assays, including clinical trials of medicinal plants.