Anal. Chem. **2008,** *80,* 1978−1987

Combining PARAFAC Analysis of HPLC-PDA Profiles and Structural Characterization Using HPLC-PDA-SPE-NMR-MS Experiments: Commercial Preparations of St. John's Wort

Bonnie Schmidt,[†] Jerzy W. Jaroszewski,[†] Rasmus Bro,[‡] Matthias Witt,[§] and Dan Stærk*,[†]

Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg, Denmark, and Bruker Daltonik GmbH, Fahrenheitstrasse 4, D-28359 Bremen, Germany

Herbal preparations represent very complex mixtures, potentially containing multiple pharmacologically active entities. Methods for global characterization of the composition of such mixtures are therefore of pertinent interest. In this work, chemometric analysis of high-performance liquid chromatography with photodiode-array detection (HPLC-PDA) data from extracts of commercial preparations of *Hypericum perforatum* (St. John's wort) that originate from several continents is described. The spectral HPLC profiles were aligned in the elution mode using correlation optimized warping in order to remove peak misalignment caused by retention time shifts due to matrix effects. Furthermore, the warping was assisted by HPLC-PDA-SPE-NMR-MS (SPE = solid-phase extraction) experiments that yielded ^{1}H NMR and ^{13}C NMR data (from ^{1}H-detected heteronuclear correlations), as well as ESI-MS and HRMS data, which enabled the identification of all major mixture constituents. The preprocessed HPLC-PDA data were subjected to parallel factor analysis (PARAFAC), a chemometric method that is a generalization of principal component analysis (PCA) to multi-way data arrays. PCA of the peak areas obtained from the PARAFAC analysis was used to facilitate sample comparison and allowed straightforward interpretation of constituents responsible for the differences in composition between individual preparations. In addition, loadings from the PARAFAC analysis provided pure elution profiles and pure UV spectra even for coeluting peaks, thus enabling the identification of chromatographically unresolved components. In conclusion, PARAFAC analysis of the readily accessible HPLC-PDA data provides the means for unsupervised and unbiased assessment of the composition of herbal preparations, of interest for assessment of their pharmacological activity and clinical efficacy.

Extracts of St. John's wort (*Hypericum perforatum*) are sold as natural remedies for the treatment of mild to moderate depression. The major secondary metabolites of St. John's wort extracts include phenylpropanoids, flavonoids (flavonols, biflavones, and proanthocyanidins), xanthones, phloroglucinols, and naphtodianthrones.[1] The mechanism underlying the antidepressant activity of St. John's wort is still not fully understood, but several compounds belonging to different structural classes and having different mechanisms of action seem to be responsible for the observed activity.[1−4] Thus, on the basis of recent results, it seems likely that phloroglucinols, naphtodianthrones, and flavonoids contribute to the antidepressant activity of St. John's wort. The phloroglucinol hyperforin has been shown to inhibit the reuptake of dopamine, serotonin, noradrenaline, γ-aminobutyric acid, and L-glutamate.[5−12] While initial studies showed that hypericin inhibits the monoamine oxidases (MAO) type A and B,[13] this result has been a matter of debate, and several studies have shown that the MAO-A and MAO-B inhibitory effects are attributable to flavonoids.[14−19] Furthermore, flavonoids were recently shown to be opioid receptor ligands.[20] The activity of St. John's wort flavonoids, as well as of solubilized hypericin and

* Corresponding author. E-mail: ds@farma.ku.dk. Fax: +45 3533 6040.
[†] Faculty of Pharmaceutical Sciences, University of Copenhagen.
[‡] Faculty of Life Sciences, University of Copenhagen.
[§] Bruker Daltonik GmbH.

(1) Butterweck, V. *CNS Drugs* **2003**, *17*, 539−562.
(2) Barnes, J.; Anderson, L. A.; Phillipson, J. D. *J. Pharm. Pharmacol.* **2001**, *53*, 583−600.
(3) Rodriguez-Landa, J. F.; Contreras, C. A. *Phytomedicine* **2003**, *10*, 688−699.
(4) Verotta, L. *Curr. Top. Med. Chem.* **2003**, *3*, 187−201.
(5) Chatterjee, S. S.; Bhattacharya, S. K.; Wonnemann, M.; Singer, A.; Muller, W. E. *Life Sci.* **1998**, *63*, 499−510.
(6) Chatterjee, S. S.; Noldner, M.; Koch, E.; Erdelmeier, C. *Pharmacopsychiatry* **1998**, *31* (Suppl. 1), 7−15.
(7) Muller, W. E.; Rolli, M.; Schafer, C.; Hafner, U. *Pharmacopsychiatry* **1997**, *30* (Suppl. 2), 102−107.
(8) Muller, W. E.; Singer, A.; Wonnemann, M.; Hafner, U.; Rolli, M.; Schafer, C. *Pharmacopsychiatry* **1998**, *31* (Suppl. 1), 16−21.
(9) Muller, W. E.; Singer, A.; Wonnemann, M. *Pharmacopsychiatry* **2001**, *34* (Suppl. 1), 98−102.
(10) Singer, A.; Wonnemann, M.; Muller, W. E. *J. Pharmacol. Exp. Ther.* **1999**, *290*, 1363−1368.
(11) Wonnemann, M.; Singer, A.; Muller, W. E. *Neuropsychopharmacology* **2000**, *23*, 188−197.
(12) Wonnemann, M.; Singer, A.; Siebert, B.; Muller, W. E. *Pharmacopsychiatry* **2001**, *34* (Suppl. 1), 148−151.
(13) Suzuki, O.; Katsumata, Y.; Oya, M.; Bladt, S.; Wagner, H. *Planta Med.* **1984**, *50*, 272−274.

10.1021/ac702064p CCC: $40.75
Published on Web 02/09/2008

pseudohypericin, has been demonstrated in the forced swimming test.[21,22] Clinical studies support the antidepressant activity of hyperforin.[23,24]

Although the above-mentioned studies show that St. John's wort extracts must be considered as multicomponent pharmaceutical preparations with multiple modes of action, commercial extracts are standardized using methods that are unable to monitor variations in the content of many components. Thus, despite recent publication of high-performance liquid chromatography (HPLC) methods for the quantification of the content of hypericins, hyperforin, and flavonoids,[25] current standardization according to the European Pharmacopoeia (Ph. Eur.) monograph is solely based on the total content of hypericin and pseudohypericin using measurement of the absorbance at 590 nm.[26] According to the United States Pharmacopeia (USP),[27] preparations of St. John's wort should be standardized based on the content of hypericin, pseudohypericin, and hyperforin as determined by HPLC-UV chromatograms monitored at 270 nm. These standardization schemes allow other constituents to be highly variable, and pronounced variability in the flavonoid content has recently been shown for commercial preparations of St. John's wort marketed in Denmark.[28] The age of the plant, time of harvest, the part of the plant used and the manufacturing process, including extraction and drying method, and soil, climate, and nutrient conditions, are all factors that can influence composition of the secondary metabolites. Furthermore, genetic factors have been shown to affect plant yield and content of secondary metabolites of *H. perforatum*.[29]

Previous studies on metabolite profiling of St. John's wort extracts included the discrimination of commercial preparations of St. John's wort based on HPLC-UV fingerprints.[30−32] Several flavonoids, naphtodiathrones, and hyperforin were determined, and significant differences in the qualitative and quantitative composition of the products were observed. The discrimination of 35 dry extracts of St. John's wort of different origins was obtained using near-infrared spectroscopy in combination with multivariate analysis.[33] Here, discrimination was based on the content of hyperforin and 3,8″-biapigenin. Moreover, correlation between ^{1}H NMR spectroscopic data from extracts of St. John's wort and IC_{50} values derived from nonselective binding to opioid receptors has been demonstrated.[33]

Knowledge of all detectable constituents of the extracts of St. John's wort, i.e., their metabolomic profiles, would be highly valuable in connection with animal and clinical studies and could provide the basis for a more comprehensive standardization of commercial preparations in the future. A wide variety of chromatographic and spectroscopic techniques, including HPLC with photodiode-array detection (HPLC-PDA), one- and two-dimensional thin-layer chromatography, gas chromatography, ultraviolet spectroscopy, infrared spectroscopy, mass spectrometry (MS), and NMR spectroscopy can be used for assessment of the metabolomic profiles of plants and plant products.[34] MS is widely used in plant metabolomics due to its sensitivity, speed, and broad applicability.[35,36] However, for extending the coverage of the plant metabolome, other techniques, particularly ^{1}H NMR spectroscopy, may be useful. The wealth of metabolites detected by these techniques is advantageous, but may complicate identification of individual components. Thus, problems with overlapping signals and the potential concentration span of 5 orders of magnitude complicate the application of MS and NMR in metabolomic analysis of plant extracts. Minor MS signals can be difficult to distinguish from noise, and the restricted chemical shift range in ^{1}H NMR spectra can be an obstacle due to signal overlap.[35] Therefore, metabolomic profiles based on separation methods coupled to spectroscopic techniques (e.g., HPLC-PDA) and combined with hyphenated structure elucidation tools, such as HPLC coupled to solid-phase extraction, NMR, and MS (HPLC-PDA-SPE-NMR-MS), offer a valuable opportunity for the detection and identification of individual metabolites that allow discrimination between samples.

In this model study, the possible use of chemometric methods combined with structure elucidation based on the hyphenation of separation techniques and spectroscopic methods was explored using 24 commercial preparations of St. John's wort originating from several continents as a test material. The assessment of the preparations was based on HPLC-PDA profiles analyzed by the three-way chemometric method, parallel factor analysis (PARAFAC),[37,38] which enabled the deconvolution of overlapping peaks. In order to correct for variations in chromatographic retention times, preprocessing of the data was performed using correlation optimized warping (COW).[39,40] Unambiguous identification of individual constituents was based on UV, MS, and 1D

(14) Bladt, S.; Wagner, H. *J. Geriatr. Psychiatry* **1994**, *154*, 125−134.
(15) Chimenti, F.; Cottiglia, F.; Bonsignore, L.; Casu, L.; Casu, M.; Floris, C.; Secci, D.; Bolasco, A.; Chimenti, P.; Granese, A.; Befani, O.; Turini, P.; Alcaro, S.; Ortuso, F.; Trombetta, G.; Loizzo, A.; Guarino, I. *J. Nat. Prod.* **2006**, *69*, 945−949.
(16) Cott, J. M. *Pharmacopsychiatry* **1997**, *30* (Suppl. 2), 108−112.
(17) Demisch, L.; Hölzl, J.; Gollnik, B. *Pharmacopsychiatry* **1989**, *22*, 194−196.
(18) Sparenberg, B.; Demisch, L.; Hölzl, J. *Pharm. Ztg. Wiss.* **1993**, *6*, 50−54.
(19) Thiede, H. M.; Walper, A. *J. Geriatr. Psychiatry Neurol.* **1994**, *7* (Suppl. 1), 54−56.
(20) Katavic, P. L.; Lamb, K.; Navarro, H.; Prisinzano, T. E. *J. Nat. Prod.* **2007**, *70*, 1278−1282.
(21) Butterweck, V.; Petereit, F.; Winterhoff, H.; Nahrstedt, A. *Planta Med.* **1998**, *64*, 291−194.
(22) Butterweck, V.; Jurgenliemk, G.; Nahrstedt, A.; Winterhoff, H. *Planta Med.* **2000**, *66*, 3−6.
(23) Laakmann, G.; Schule, C.; Baghai, T.; Kieser, M. *Pharmacopsychiatry* **1998**, *31* (Suppl. 1), 54−59.
(24) Zanoli, P. *CNS Drug Rev.* **2004**, *10*, 203−218.
(25) *Pharmeuropa*; Council of Europe: Strassbourg, 2006; Vol. 18, pp 295−298.
(26) *European Pharmacopoeia*, 5th ed.; Council of Europe: Strassbourg, 2005; Vol. 2, pp 2485−2486.
(27) *United States Pharmacopeia*; USP 30, The National Formulary 25, The United States Pharmacopeial Convention: Rockville, MD, 2007; pp 978−979.
(28) Rasmussen, B.; Cloarec, O.; Tang, H. R.; Stærk, D.; Jaroszewski, J. W. *Planta Med.* **2006**, *72*, 556−563.
(29) Büter, B.; Orlacchio, C.; Soldati, A.; Berger, K. *Planta Med.* **1998**, *64*, 431−437.
(30) Ganzera, M.; Zhao, J.; Khan, I. A. *J. Pharm. Sci.* **2002**, *91*, 623−630.
(31) Bergonzi, M. C.; Bilia, A. R.; Gallori, S.; Guerrini, D.; Vincieri, F. F. *Drug Dev. Ind. Pharm.* **2001**, *27*, 491−497.
(32) Rager, I.; Roos, G.; Schmidt, P. C.; Kovar, K. A. *J. Pharm. Biomed. Anal.* **2002**, *28*, 439−446.
(33) Roos, G.; Röseler, C.; Büter, K. B.; Simmen, U. *Planta Med.* **2004**, *70*, 771−777.
(34) Holmes, E.; Tang, H. R.; Wang, Y. L.; Seger, C. *Planta Med.* **2006**, *72*, 771−785.
(35) Krishnan, P.; Kruger, N. J.; Ratcliffe, R. G. *J. Exp. Bot.* **2005**, *56*, 255−265.
(36) Hall, R. D. *New Phytol.* **2006**, *169*, 453−468.
(37) Harshman, R. A. *UCLA Work. Pap. Phonetics* **1970**, *16*, 1−84.
(38) Carroll, J. D.; Chang, J. *Psychometrika*, **1970**, *35*, 283−319.
(39) Nielsen, N. P. V.; Carstensen, J. M.; Smedsgaard, J. *J. Chromatogr., A* **1998**, *805*, 17−35.
(40) Tomasi, G.; van den Berg, F.; Andersson, C. *J. Chemom.* **2004**, *18*, 231−241.

and 2D NMR data from HPLC-PDA-SPE-NMR-MS analysis of the extracts. The study described here may be regarded as a prototype methodology that is generally applicable to herbal products.

EXPERIMENTAL SECTION

Chemicals. Acetonitrile-d_3 (99.8 atom % of deuterium) was obtained from Cambridge Isotope Laboratories. All other solvents were of analytical or HPLC grade. The water used for HPLC separations was purified by deionization and 0.22 μm membrane filtration (Millipore).

Sample Preparation. Twenty-four different samples of St. John's wort from several continents (Africa, Asia, Europe, and North America) were obtained from retail stores. Thirteen preparations were formulated as tablets (preparations 1−4, 11, 12, 14, 16, 17, 21−24), and the remaining were formulated as capsules (preparations 5−10, 13, 15, 18−20). Preparations 1−10, 23, and 24 originated in Europe, preparations 11 and 17 originated in Asia, preparations 12, 13, 15, 16, 18, and 20−22 originated in North America, and preparations 14 and 19 originated in Africa. For one of the brands, two different batches were obtained (preparations 7 and 8). An amount of powder corresponding to 300 mg of plant extract was extracted with 10 mL of methanol/pyridine (6:4) by sonication for 45 min. The extracts were centrifuged at 18000*g* for 8 min, filtered, and the solids were re-extracted as above. The combined extracts were evaporated to dryness at temperatures below 40 °C, and the residual extraction solvents were removed by evaporation with toluene. Light was excluded as much as possible during all operations. Samples for HPLC-PDA and HPLC-PDA-NMR-MS were reconstituted in acetonitrile/water (25:75) to concentrations of 10 and 50 mg/mL, respectively.

HPLC-PDA Experiments. Separations were performed at 40 °C on a 150 × 4.6 mm i.d. Phenomenex C_{18}(2) Luna column (3 μm, 100 Å) using a Shimadzu system consisting of a SCL-10A system controller, a SIL-10AD autoinjector, a LC-10AT pump, and a SPD-M10A photodiode array detector, controlled with Shimadzu *Class-VP*, version 6.10, software. Chromatographic separations were performed using mixtures of water/acetonitrile 95:5 + 0.1% formic acid (eluent A) and acetonitrile/water 95:5 + 0.1% formic acid (eluent B) with the following linear gradient elution profile: 0 min, 10% B; 10 min, 20% B; 20 min, 100% B; 30 min, 100% B; 32 min, 10% B, and 10 min conditioning with 10% B. Injection volumes of 20 μL and flow rates of 0.8 mL/min were used. UV spectra were acquired from 190 to 620 nm with a resolution of 1 nm/point at a sampling rate of 1.56 s^{-1} and exported as a 3D ASCII file. The HPLC profiles were preprocessed using an extended algorithm of COW, correcting for shifts in retention time mode according to criterion working on the whole spectrum from 260 to 520 nm. Two regions in the profiles were warped separately; region 1 (t_R = 12.0−15.5 min) containing peaks 2, 3, 4, and 5/6 and region 2 (t_R = 15.5−23.7 min) containing peaks 7, 8, 9, 10, 11, and 12. All preprocessing operations were performed using *MATLAB*, version 7.0.1, software (MathWorks, Natick, MA).

HPLC-PDA-SPE-NMR-MS Experiments. Separations were performed with the above-mentioned column at 40 °C using an Agilent 1100 system consisting of a degasser, a quaternary pump, an autosampler, a column oven, and a photodiode array detector. The solvent was delivered at 0.8 mL/min using the following linear gradient elution profile: 0 min, 10% B; 20 min, 15% B; 30 min, 50% B; 40 min, 100% B; 42 min, 10% B, and 10 min conditioning with 10% B (composition of solvents A and B as in HPLC-PDA experiments). Injection volumes were 40 μL. The eluate from the column was split with an Acurate splitter, directing 5% to a Bruker Esquire LC ion trap mass spectrometer with an electrospray ionization (ESI) interface and 95% to the PDA detector. MS spectra were acquired in positive-ion mode, using a drying temperature of 350 °C, a nebulizer pressure of 15 psi, and a drying gas flow of 15 mL/min. The HPLC solvent flow (0.8 mL/min) from the PDA detector was diluted with water delivered by a Knauer K100 Wellchrom pump (2 mL/min) and directed to a Prospekt 2 SPE unit (Spark Holland). GP phase (general-purpose poly(divinylbenzene)-based resin) SPE cartridges, 10 × 2 mm i.d. from Spark Holland, were used for the trapping of selected compounds based on UV absorption levels (340 nm) or predefined time intervals. Prior to use, the SPE cartridges were conditioned with 500 μL of acetonitrile at 6 mL/min and flushed with 500 μL of water at 1 mL/min. After trapping, cartridges were dried with a stream of dry nitrogen gas for 45 min and the analytes eluted to the NMR flow cell with acetonitrile-d_3. The HPLC-PDA-SPE-NMR-MS experiments were controlled with Bruker *HyStar*, version 2.3, software. NMR data acquisition and processing were performed using Bruker *XWINNMR*, version 3.1, software.

A time-slice HPLC-PDA-SPE-NMR-MS experiment of preparation 5 was performed using the same separation conditions as for the HPLC-PDA experiments and five cumulative trappings. Thus, analytes eluted from 1.5 to 24.5 min were trapped on 46 SPE cartridges with trapping bandwidths of 0.5 min.

NMR Experiments. All NMR experiments were performed with a Bruker Avance 600 NMR spectrometer (^{1}H resonance frequency of 600.13 MHz) using standard Bruker library pulse sequences. NMR experiments in the HPLC-PDA-SPE-NMR mode were obtained using an inverse ^{1}H flow probe with an active volume of 30 μL. 1D ^{1}H NMR spectra were recorded using a 1D version of the nuclear Overhauser effect spectroscopy pulse sequence with dual presaturation of residual solvent resonances during the relaxation delay and mixing time (2.4 s and 100 ms, respectively), collecting 128−512 transients as 65 536 data points with a spectral width of 20 ppm. Phase-sensitive, gradient-selected homonuclear correlation spectroscopy (^{1}H−^{1}H COSY) experiments were performed using a spectral width of 7.6 ppm, collecting 2048 × 256 data points (processed with forward linear prediction to 2048 × 512 data points) with 24 transients per increment, and Fourier-transformed by applying squared sine-bell apodization in both dimensions. Residual solvent signals were suppressed using the water suppression enhanced by T_1 effects (WET) sequence with 20 ms sinusoidal pulses for selective excitation, 1 ms dephasing gradients, and ^{13}C decoupling during acquisition using the GARP sequence. Gradient-selected sensitivity-enhanced ^{1}H−^{13}C heteronuclear single-quantum coherence (HSQC) spectra with WET solvent peak suppression were acquired with a spectral width of 7.6 ppm for ^{1}H and 160 ppm for ^{13}C, collecting 1024 × 128 data points (processed with forward linear prediction to 1024 × 512 data points). The HSQC experiments were optimized for $^1J_{C,H}$ = 145 Hz, and the data were Fourier-transformed by applying squared sine-bell apodization in both dimensions. Gradient-selected ^{1}H−^{13}C heteronuclear multiple-bond correlation (HMBC) experiments were optimized for $^nJ_{C,H}$ = 7.7 Hz and acquired with

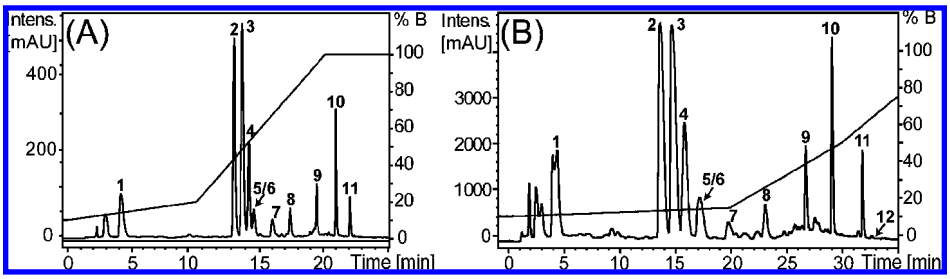


Figure 1. HPLC traces (340 nm) and elution gradient profiles of an extract of a commercial preparation of St. John's wort (preparation 5). (A) Method used for obtaining HPLC-PDA data for three-way chemometric analysis. (B) Method used for HPLC-PDA-SPE-NMR-MS experiments.

spectral widths of 7.6 ppm for ^{1}H and 220 ppm for ^{13}C, collecting 2048 × 128 data points (processed with forward linear prediction to 2048 × 512 data points) and Fourier-transformed by applying squared sine-bell apodization in both dimensions. All spectra acquired in the HPLC-PDA-SPE-NMR mode were referenced to the residual ^{1}H signal of acetonitrile (δ 1.94, CD_2HCN) or the ^{13}C resonance of deuterated acetonitrile (δ 1.32).

High-Resolution MS (HRMS) Experiments. HRMS measurements for exact mass determination were carried out off-line on samples collected after HPLC-PDA-SPE-NMR-MS experiments, using a Bruker APEX Qe Fourier-transform mass spectrometer equipped with a 9.4 T superconducting cryo-magnet and an external electrospray ion source (Apollo II source). The spectra were externally calibrated with arginine cluster in positive and negative ion mode. The samples were dissolved in methanol, further diluted in 50% methanol with 0.2% formic acid (positive ion mode) or 50% methanol (negative ion mode), and introduced into the ion source using a syringe pump with a flow of 2 μL/min.

RESULTS AND DISCUSSION

Identification of Components using HPLC-PDA-SPE-NMR-MS Experiments. Commercial capsules and tablets were extracted with methanol/pyridine,[25,28] the extracts were evaporated, and the residues were redissolved in CH_3CN/H_2O (25:75) for HPLC-PDA analysis and HPLC-PDA-SPE-NMR-MS experiments. A gradient-based reversed-phase HPLC method was developed, enabling the observation of 10 well-defined peaks at 340 nm (Figure 1A). All preparations were analyzed using this method, and associated UV spectra from PDA detection provided data for subsequent chemometric analysis. Preparation 5 was subjected to HPLC-PDA-SPE-NMR-MS analysis using a second HPLC-method with a less-steep gradient and longer separation time (Figure 1B), which allowed higher column loading. This preparation was subsequently used as a reference in the preprocessing of HPLC-PDA profiles used in the multivariate analysis (see discussion in next section), and compounds are numbered according to the peak elution order in this chromatogram. Unambiguous identification of 12 compounds in 11 peaks was based on MS and 1D ^{1}H NMR spectra, as well as the COSY, HSQC, and HMBC spectra acquired for selected peaks. The heteronuclear correlations provided ^{13}C NMR data of selected compounds, as shown in Table 1. Thus, HPLC-PDA-SPE-NMR-MS analysis of preparation 5 identified chlorogenic acid (**1**),[41,42] 3-(6-*O*-α-L-rhamnopyranosyl-*β*-D-glucopyranosyloxy)-3′,4′,5,7-tetrahydroxyflavone (quercetin 3-*O*-rutinoside or rutin, **2**),[43,44] hyperoside (quercetin 3-*O*-*β*-D-galactopyranoside, **3**),[45] isoquercetin (quercetin 3-*O*-*β*-D-glucopyranoside, **4**),[44] miquelianin (quercetin 3-*O*-*β*-D-glucopyranosiduronic acid, **5**),[45,46] astilbin ((2*R*,3*R*)-dihydroquercetin 3-*O*-α-L-rhamnopyranoside, **6**),[45,47] guaijaverin (quercetin 3-*O*-α-L-arabinopyranoside, **7**),[48] quercetrin (quercetin 3-*O*-α-L-rhamnopyranoside, **8**),[49,50] quercetin 3-*O*-*β*-D-(2-*O*-acetyl)galactopyranoside (**9**),[51] quercetin (**10**),[45] and 3,8″-biapigenin (**11**)[52] as major constituents. The amount of material eluted with peak 12 was too low to yield satisfactory ^{1}H NMR spectra in the HPLC-PDA-SPE-NMR mode, but the compound was assigned as amentoflavone (3′,8″-biapigenin, **12**) on the basis of previous work[51] and HRMS data (m/z 537.08240 $[M - H]^-$).

In order to correlate the peaks in the chromatogram obtained with the HPLC method used for the HPLC-PDA analysis (Figure 1A) with those identified by the HPLC-PDA-SPE-NMR-MS experiments (Figure 1B), a time-slice HPLC-PDA-SPE-NMR experiment was performed with preparation 5 using the former HPLC method. Thus, the HPLC eluate was passed through consecutive SPE cartridges with 0.5 min time intervals, and 1D ^{1}H NMR spectra of trapped material were recorded. These spectra were compared with those obtained after the trapping of individual peaks. The time-slice experiment ensured that no changes in elution order had occurred upon change of the HPLC method. Table 1 shows a summary of chromatographic and spectroscopic data for constituents of preparation 5. The assignment was extended to all preparations based on HPLC-MS and HPLC-PDA data that have been recorded for all samples.

The elution order of St. John's wort constituents determined in this study is identical to that reported by Jürgenliemk and

(41) Gerothanassis, I. P.; Exarchou, V.; Lagouri, V.; Troganis, A.; Tsimidou, M.; Boskou, A. *J. Agric. Food Chem.* **1998**, *46*, 4185−4192.
(42) Pauli, G. F.; Kuczkowiak, U.; Nahrstedt, A. *Magn. Reson. Chem.* **1999**, *37*, 827−836.
(43) Clarkson, C.; Stærk, D.; Hansen, S. H.; Jaroszewski, J. W. *Anal. Chem.* **2005**, *77*, 3547−3553.
(44) Kazuma, K.; Noda, N.; Suzuki, M. *Phytochemistry* **2003**, *62*, 229−237.
(45) Tatsis, E. C.; Boeren, S.; Exarchou, V.; Troganis, A. N.; Vervoort, J.; Gerothanassis, I. P. *Phytochemistry* **2007**, *68*, 383−393.
(46) Bouktaib, M.; Atmani, A.; Rolando, C. *Tetrahedron Lett.* **2002**, *43*, 6263−6266.
(47) Kasai, R.; Hirono, S.; Chou, W. H.; Tanaka, O.; Chen, F. H. *Chem. Pharm. Bull.* **1988**, *36*, 4167−4170.
(48) Hansen, S. H.; Jensen, A. G.; Cornett, C.; Bjørnsdottir, I.; Taylor, S.; Wright, B.; Wilson, I. D. *Anal. Chem.* **1999**, *71*, 5235−5241.
(49) Jayaprakasha, G. K.; Ohnishi-Kameyama, M.; Ono, H.; Yoshida, M.; Rao, L. J. *J. Agric. Food Chem.* **2006**, *54*, 1672−1679.
(50) Zhong, X. N.; Otsuka, H.; Ide, T.; Hirata, E.; Takushi, A.; Takeda, Y. *Phytochemistry* **1997**, *46*, 943−946.
(51) Jurgenliemk, G.; Nahrstedt, A. *Planta Med.* **2002**, *68*, 88−91.
(52) Berghofer, R.; Holzl, J. *Planta Med.* **1987**, 216−217.

Table 1. Structures of St. John's Wort Constituents with ^{1}H and ^{13}C NMR Data from HPLC-SPE-NMR-MS Experiments and HRMS Data from Off-Line Measurements

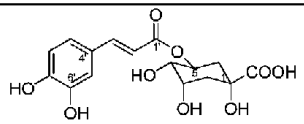
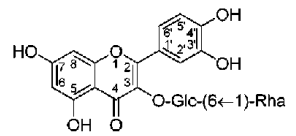
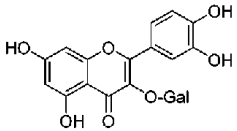
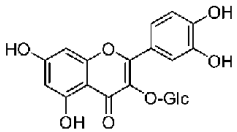
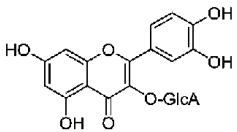
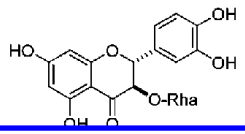
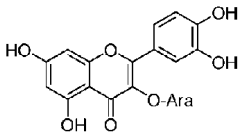
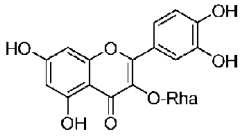
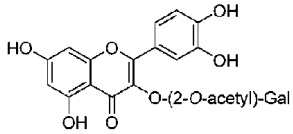
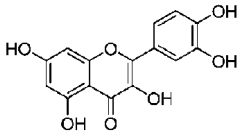
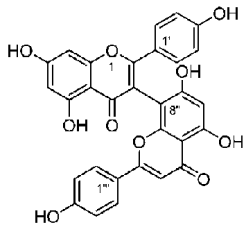
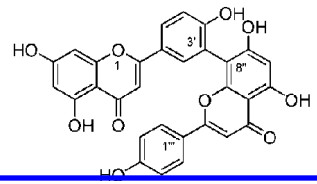
Compound	Peak	T_R (min)[a]	T_R (min)[b]	Structure[c]	^{1}H NMR data[d] (δ, ppm; *J*, Hz)[e,f]	^{13}C NMR data[d] (δ, ppm)[g]	MS (*m/z*)[h]
Chlorogenic acid (**1**)	1	4.44	3.92		1.94 (m, H-6b), 2.06 (m, H-2), 2.18 (m, H-6a), 3.70 (dd, *J* = 9.4, 3.7, H-4), 4.17 (dd, *J* = 3.7, 2.9, H-3), 5.26 (ddd, *J* = 11.2, 9.4, 4.8, H-5), 6.30 (d, *J* = 16.0, H-2'), 6.85 (d, *J* = 8.1, H-8'), 7.04 (dd, *J* = 8.1, 2.0, H-9'), 7.12 (d, *J* = 2.0, H-5'), 7.58 (d, *J* = 16.0, H-3')	39.1 (C-6), 37.7 (C-2), 71.0 (C-5), 71.9 (C-3), 73.4 (C-4), 115.3 (C-5'), 116.1 (C-2'), 116.3 (C-8'), 122.8 (C-9'), 128.5 (C-4'), 146.7 (C-3'), 148.4 (C-6'), 148.7 (C-7')	377.08432/ 377.08430 $(MNa)^+$
Rutin (**2**)	2	13.10	13.92		1.06 (d, *J* = 6.2, H-6'''), 3.19 (dt, *J* = 9.4, 3.5, H-4'''), 3.29 (m, H-3'', H-4''), 3.36 (m, H-5''), 3.37 (m, H-5'''), 3.39 (m, H-2''), 3.42 (m, H-6''B), 3.43 (m, H-3'''), 3.59 (m, H-2'''), 3.65 (d, *J* = 11.4, H-6''A), 4.51 (d, *J* = 1.5, H-1'''), 5.02 (d, *J* = 7.5, H-1''), 6.27 (d, *J* = 2.2, H-6), 6.48 (d, *J* = 2.2, H-8), 6.93 (d, *J* = 8.4, H-5'), 7.66 (dd, *J* = 8.4, 2.2, H-6'), 7.76 (d, *J* = 2.2, H-2'), 12.27 (s, 5-OH)	18.3 (C-6'''), 68.8 (C-6''), 69.6 (C-5'''), 71.1 (C-4''), 72.0 (C-3'''), 72.3 (C-2'''), 74.1 (C-4'''), 76.1 (C-2''), 77.4 (C-3''), 78.4 (C-5''), 95.3 (C-8), 100.0 (C-6), 101.8 (C-1'''), 105.3 (C-1''), 106.6 (C-4a), 116.0 (C-5'), 118.1 (C-2'), 123.5 (C-6'), 124.3 (C-1'), 136.3 (C-3), 145.8 (C-3'), 149.2 (C-4'), 158.5 (C-8a), 159.2 (C-2), 164.5 (C-7), 174.8 (C-4)	611.16081/ 611.16066 $(MH)^+$
Hyperoside (**3**)	3	13.74	14.98		3.48 (m, H-5''), 3.52 (dd, *J* = 9.5, 4.4, H-3''), 3.60 (t, *J* = 5.5, H-6''), 3.64 (dd, *J* = 9.5, 7.7, H-2''), 3.77 (dd, *J* = 4.4, 3.7, H-4''), 5.01 (d, *J* = 7.7, H-1''), 6.28 (d, *J* = 2.2, H-6), 6.48 (d, *J* = 2.2, H-8), 6.94 (d, *J* = 8.4, H-5'), 7.54 (dd, *J* = 8.4, 2.2, H-6'), 7.97 (d, *J* = 2.2, H-2'), 12.23 (s, 5-OH)	62.6 (C-6''), 70.0 (C-4''), 72.6 (C-2''), 74.4 (C-3''), 76.4 (C-5''), 94.8 (C-8), 99.9 (C-6), 105.8 (C-1''), 115.6 (C-5'), 118.1 (C-2'), 122.7 (C-6'), 149.4 (C-4'), 159.2 (C-2)	487.08474/ 487.08470 $(MNa)^+$
Isoquercetin (**4**)	4	14.25	16.03		3.25 (m, H-4'', H-5''), 3.39 (m, H-2'', H-3''), 3.51 (dd, *J* = 12.1, 5.9, H-6''A), 3.68 (dd, *J* = 12.1, 5.5, H-6''B), 5.08 (d, *J* = 7.5, H-1''), 6.27 (d, *J* = 2.2, H-6), 6.5 (d, *J* = 2.2, H-8), 6.94 (d, *J* = 8.4, H-5'), 7.51 (dd, *J* = 8.4, 2.2, H-6'), 7.94 (d, *J* = 2.2, H-2'), 12.23 (s, 5-OH)	62.5 (C-6''), 70.6 (C-4''), 75.0 (C-2''), 77.4 (C-5''), 77.6 (C-3''), 94.7 (C-8), 99.6 (C-6), 105.0 (C-1''), 105.4 (C-4a), 115.7 (C-5'), 118.1 (C-2'), 122.7 (C-6'), 122.9 (C-1') 135.6 (C-3), 144.8 (C-4'), 148.5 (C-3'), 158.1 (C-8a), 158.9 (C-2), 163.1 (C-7), 177.8 (C-4)	487.08476/ 487.08470 $(MNa)^+$
Miquelianin (**5**)	5/6	14.64	17.36		3.43 (dd, *J* = 9.0, 8.8, H-3''), 3.48 (dd, *J* = 9.0, 7.5, H-2''), 3.53 (dd, *J* = 9.7, 8.8, H-4''), 3.76 (d, *J* = 9.7, H-5''), 5.21 (d, *J* = 7.5, H-1''), 6.27 (d, *J* = 2.2, H-6), 6.48 (d, *J* = 2.2, H-8), 6.93 (d, *J* = 8.4, H-5'), 7.61 (dd, *J* = 8.4, 2.2, H-6'), 7.83 (d, *J* = 2.2, H-2'), 12.26 (s, 5-OH)	72.1 (C-4''), 74.6 (C-2''), 75.6 (C-5''), 76.8 (C-3''), 95.2 (C-8), 100.1 (C-6), 104.1 (C-1''), 106.4 (C-4a), 116.1 (C-5'), 117.5 (C-2'), 123.3 (C-6'), 124.5 (C-1'), 136.5 (C-3), 145.8 (C-3'), 149.3 (C-4'), 158.6 (C-8a), 159.8 (C-2), 164.8 (C-7), 174.8 (C-4)	479.08201/ 479.08202 $(MH)^+$, 501.06401/ 501.06396 $(MNa)^+$
Astilbin (**6**)	5/6	14.64	17.36		1.11 (d, *J* = 6.2, H-6''), 3.17 (dd, *J* = 10.6, 9.7, H-4''), 3.43 (m, H-2''), 3.49 (m, H-3''), 3.90 (d, *J* = 1.8, H-1''), 4.04 (dd, *J* = 9.7, 6.2, H-5''), 4.64 and 5.14 (2H, d, *J* = 10.8, H-2, H-3), 5.94 and 5.98 (2H, d, *J* = 2.2, H-6, H-8), 6.86 (d, *J* = 8.2, H-5'), 6.88 (dd, *J* = 8.2, 1.7, H-6'), 6.98 (d, *J* = 1.7, H-2'), 11.81 (s, 5-OH)	Not determined	473.10541/ 473.10543 $(MNa)^+$

Table 1. (Continued)

Compound	Peak	T_R (min)[a]	T_R (min)[b]	Structure[c]	^{1}H NMR data[d] (δ, ppm; *J*, Hz)[e,f]	^{13}C NMR data[d] (δ, ppm)[g]	MS (*m/z*)[h]
Guaijaverin (**7**)	7	16.05	20.44		3.38 (br d, *J* = 11.3, H-5"A), 3.57 and 3.73 (3H, m, H-2", H-3", H-4"), 3.70 (br d, *J* = 11.3, H-5"B), 5.04 (d, *J* = 6.7, II-1"), 6.26 (d, *J* = 2.1, II-6), 6.47 (d, *J* = 2.1, II-8), 6.94 (d, *J* = 8.4, H-5'), 7.63 (br d, *J* = 8.4, H-6'), 7.74 (br, H-2'), 12.4 (s, 5-OH)	Not determined	433.07744 433.07764 (M-H)
Quercetrin (**8**)	8	17.40	23.57		0.87 (d, *J* = 5.9, H-6"), 3.19 (m, H-4", H-5"), 3.58 (m, H-3"), 4.09 (dd, *J* = 4.8, 1.5, H-2"), 5.43 (d, *J* = 1.5, H-1"), 6.24 (d, *J* = 2.2, H-6), 6.43 (d, *J* = 2.2, H-8), 6.96 (d, *J* = 8.3, H-5'), 7.35 (dd, *J* = 8.3, 2.0, H-6"), 7.40 (d, *J* = 2.0, H-2'), 12.62 (s, 5-OH)	17.5 (C-6"), 71.3 (C-2"), 71.4 (C-3"), 72.2 (C-4", C-5"), 94.7 (C-8), 99.3 (C-6), 102.4 (C-1"), 116.4 (C-5'), 117.0 (C-2'), 123.1 (C-6')	471.08983 471.08978 (MNa)$^+$
Quercetin 3-*O*-β-D-(2-*O*-acetyl) galactopyranoside (**9**)	9	19.42	27.17		2.09 (s, 2"-OCOC<u>H</u>$_3$), 3.48 (t, *J* = 5.7, II-5"), 3.55 (t, *J* = 5.7, H-6"), 3.70 (ddd, *J* = 9.7, 6.4, 3.3, H-3"), 3.85 (m, H-4"), 5.09 (dd, *J* = 9.7, 7.9, H-2"), 5.48 (d, *J* = 7.9, H-1"), 6.23 (d, *J* = 2.02, H-6), 6.44 (d, *J* = 2.02, H-8), 6.92 (d, *J* = 8.6, H-5'), 7.57 (dd, *J* = 8.6, 2.0, H-6"), 7.74 (d, *J* = 2.0, H-2"), 12.62 (s, 5-OH)	21.4 (2"-OCO<u>C</u>H$_3$), 62.1 (C-6"), 70.3 (C-4"), 72.5 (C-3"), 74.0 (C-2"), 76.5 (C-5"), 94.5 (C-8), 99.5 (C-6), 100.6 (C-1"), 116.3 (C-5'), 117.6 (C-2'), 122.9 (C-6'), 146.2 (C-4'), 154 (C-2), 172.2 (2"-O<u>C</u>OCH$_3$)	529.09530 529.09526 (MNa)$^+$
Quercetin (**10**)	10	20.92	29.49		6.25 (d, *J* = 2.0, II-6), 6.48 (d, *J* = 2.0, II-8), 6.98 (d, *J* = 8.4, H-5'), 7.65 (dd, *J* = 8.4, 2.0, H-6'), 7.73 (d, *J* = 2.0, H-2'), 11.98 (s, 5-OH)	94.5 (C-8), 98.9 (C-6), 105.0 (C-4a), 115.5 (C-2'), 116.0 (C-5'), 121.5 (C-6'), 124.7 (C-1'), 147.0 (C-3'), 149.1 (C-4'), 158.6 (C-8a), 165.5 (C-7)	303.04997 303.04993 (MH)$^+$
Biapigenin (**11**)	11	21.99	32.21		6.28 (s, H-6"), 6.32 (d, *J* = 2.2, H-6), 6.55 (d, *J* = 2.2, H-8), 6.57 (s, H-3"), 6.70 and 6.83 (d, *J* = 8.8, H3'/H5' and H-3'"/H5'"), 7.39 and 7.58 (d, *J* = 8.8, H2'/H6' and H2'"/H-6'"), 12.79 and 13.06 (2H, 5-OH in 4-hydroxylated B-rings)	94.9 (C-8), 99.8 (C-6"), 99.9 (C-8"), 100.0 (C-6), 104.6 (C-3"), 116.2 and 116.7 (C-3'/C-5' and 3'"/C-5'"), 124.1 and 127.6 (C-1' and C-1'"), 129.0 and 130.7 (C-2'/C-6' and C-2'"/C-6'"), 156.9 (C-8a"), 164.8 and 166.1 (C-2 and C-2")	539.09719 539.09727 (MH)$^+$
Amentoflavone (**12**)	12	22.33	32.93		Not determined	Not determined	537.08240 537.08272 (M-H)

[a] Retention times with method used for HPLC-PDA analysis. [b] Retention times with method used for HPLC-PDA-SPE-NMR-MS analysis. [c] Glc, β-D-glucopyranosyl; Rha, α-L-rhamnopyranosyl; Gal, β-D-galactopyranosyl; GlcA, β-D-glucopyranosiduronylic acid; Ara, β-D-arabinopyranosyl. [d] ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR spectral data measured in HPLC-SPE-NMR mode (CD_3CN). [e] δ values relative to CD_2HCN signal at δ 1.94. [f] Multiplicity of signals: s, singlet; d, doublet; t, triplet; m, multiplet; br, broad. Coupling constants (apparent splittings) reported as numerical values in hertz. [g] Obtained from HSQC and HMBC spectra; δ values relative to CD_2HCN signal at δ 1.32. [h] HRMS measurements given as experimental/calculated values in atomic mass units.

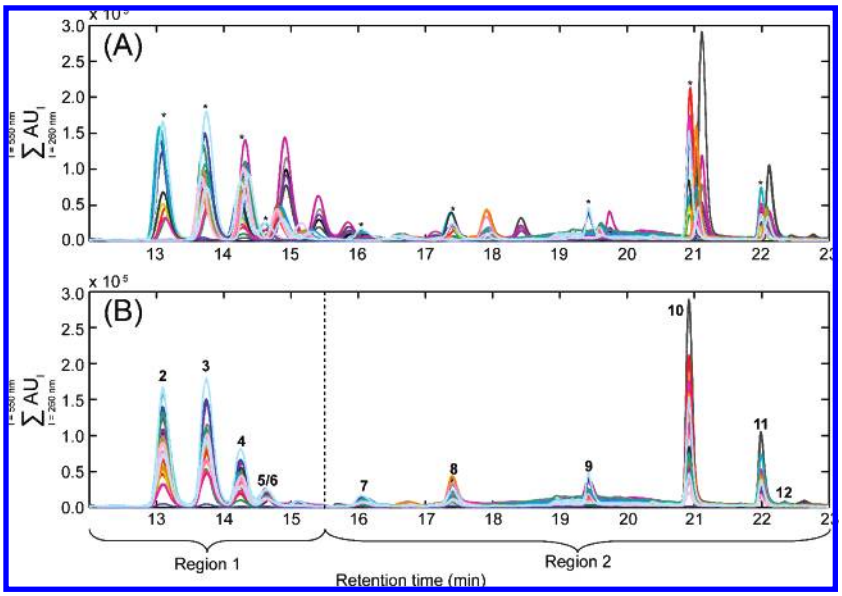


Figure 2. HPLC profiles of extracts of 24 commercial preparations of St. John's wort with 3 or 4 replicates of each sample. Sum of absorbances at all wavelengths (260−550 nm) are shown before (A) and after warping (B) using an extended algorithm of COW. The HPLC profiles are colored according to supplier.

Nahrstedt,[51] where a crude drug material of St. John's wort was investigated. The present study also confirmed the presence of **6** and **9**, reported for the first time in St. John's wort and as a new natural product, respectively, by Jürgenliemk and Nahrstedt.[51]

Previously, direct HPLC-NMR experiments have been used to identify flavonoids, phloroglucinols, and naphtodianthrones in a crude extract of St. John's wort.[48] More recently, HPLC-SPE-NMR coupled with on-line radical scavenging detection was used to identify antioxidants in an extract of St. John's wort,[53] whereas HPLC-SPE-NMR combined with HPLC-MS^2 was used to identify major constituents present in Greek *H. perforatum* extracts.[45] The latter work resulted in the identification of several naphtodianthrones, phloroglucinols, flavonoids, and phenolic acids. Both studies used resin-based SPE stationary phase, as in the present work, in agreement with the recent comparison of SPE phases used for HPLC-SPE-NMR experiments.[54]

Multivariate Analysis of HPLC-PDA Profiles. The HPLC-PDA data generated in this work comprise a two-way matrix of measured absorbance for each sample, and thus for several samples a three-way array, $\underline{X}$, is generated (samples (I) × retention time (J) × measured absorbance (K)). This array can be analyzed using so-called unfold-PCA on the modes of interest or using PARAFAC analysis directly on the three-way array. Unlike PCA, a PARAFAC model can find the underlying individual analyte peaks, providing estimates of their pure UV spectra, pure elution profiles, and their relative concentration; even for partially coeluting components.[55−57] For this reason, PARAFAC modeling is pursued in the following.

In order to perform multivariate analysis, preprocessing of data is often mandatory and can range from simple centering and scaling to more sophisticated methods, such as variable selection and warping.[58] In the case of HPLC-PDA data, variations in retention times due to sample composition or instrumental instability usually have to be corrected before multivariate modeling can be applied. In the HPLC-PDA profiles obtained in this work, shifting of peak positions was observed even though differences in sample preparation and measurement conditions were minimized. When three or four nonconsecutive replicates of each sample were prepared, no differences in chromatographic peak positions were observed. Therefore, the differences observed in retention times of the same constituents in different samples (Figure 2A) are ascribed to differences in matrix composition and possibly also to varying amounts of solvent residues. These unwanted peak shifts can be corrected using algorithms for retention time alignment, e.g., dynamic time warping (DTW) or correlation optimized warping (COW).[40] Since COW has previously been successfully applied as a preprocessing step for chromatographic data,[39,40,59] this method was chosen for the present data set. COW is a piecewise or segmented data preprocessing technique where alignment of a sample chromatogram toward a reference chromatogram is performed by allowing small changes in segment lengths on the sample chromatogram. The reference chromatogram is divided into a number of segments defined by the number of points in the chromatogram (N) and the selected segment length (I), where the number of segments is N/I. The sample chromatogram is divided into an equal number of segments. The slack parameter (t) controls the maximum increase or decrease in sample segment length. Thus, this parameter determines the degree of flexibility allowed to obtain a good correlation between the sample chromatogram and the reference chromatogram. In this study, a version of COW[60]

(53) Exarchou, V.; Fiamegos, Y. C.; van Beek, T. A.; Nanos, C.; Vervoort, J. *J. Chromatogr., A* **2006**, *1112*, 293−302.
(54) Clarkson, C.; Sibum, M.; Mensen, R.; Jaroszewski, J. W. *J. Chromatogr., A.* **2007**, *1165*, 1−9.
(55) Bro, R.; Kiers, H. A. L. *J. Chemom.* **2003**, *17*, 274−286.
(56) Garcia, I.; Sarabia, L.; Ortiz, M. C.; Aldama, J. M. *Analyst* **2004**, *129*, 766−771.
(57) Wiberg, K.; Jacobsson, S. P. *Anal. Chim. Acta* **2004**, *514*, 203−209.
(58) Bro, R. *Crit. Rev. Anal. Chem.* **2006**, *36*, 279−93.
(59) Skov, T.; van den Berg, F.; Tomasi, G.; Bro, R. *J. Chemom.* **2006**, *20*, 484−497.

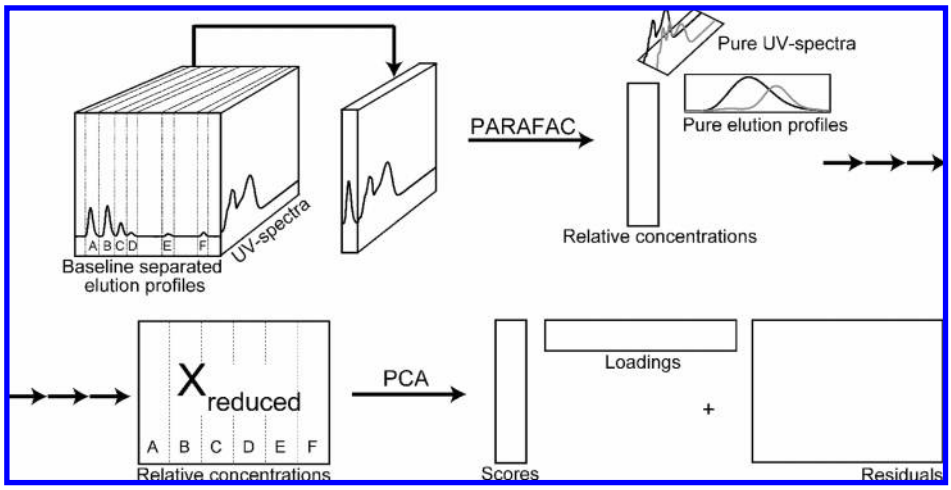


Figure 3. Scheme of the analysis steps for multivariate analysis of HPLC-PDA profiles of the extracts of St. John's wort preparations. Segments around chromatographic peaks are defined based on baseline separation, and PARAFAC analysis is applied to each segment, giving three loading matrices that represent relative concentrations, pure elution profiles, and pure UV spectra, respectively. The matrices comprising relative concentrations are combined to a new matrix, $X_{reduced}$, for subsequent PCA analysis.

that allows the shifting of whole spectra rather than just a single elution profile was used. Thus, alignment was performed in the elution mode but according to a criterion working on the whole spectrum. Segment lengths and slack parameters were individually defined for each sample chromatogram warped and were determined on the basis of the degree of shifting in the chromatographic peaks between the sample chromatogram and the reference chromatogram. Two regions of the three-way array, $\underline{X}$, were chosen for chemometric analysis based on the peaks identified from HPLC-PDA-SPE-NMR-MS experiments: region 1 ($t_R = 12.0-15.5$ min) containing peaks 2, 3, 4, and 5/6 and region 2 ($t_R = 15.5-23.7$ min), with peaks 7, 8, 9, 10, 11, and 12 (Figure 2B). Each region was warped separately using a chromatogram of preparation 5 as reference chromatogram, for which the HPLC-PDA-SPE-NMR-MS analysis was performed.

The use of COW aligning in the elution mode according to a criterion working on the whole spectrum in combination with HPLC-MS measurements for each sample ensured that no misalignment of peaks has occurred during the warping procedure. The original and aligned chromatograms are shown in Figure 2.

After this preprocessing, the data array $\underline{X}$ was divided into segments, each containing one chromatographic peak based on baseline separation (Figure 3). Only peaks 2−11, having well-characterized peak shapes, were selected for further analysis. Thus, segments around peaks 2−11 were defined, and PARAFAC analysis was applied to each segment. The output of the PARAFAC analysis performed on the selected chromatographic peaks is a vector or a matrix of scores (depending on the number of components used to fit the PARAFAC models) that represents the relative concentration of the metabolite represented by the chromatographic peak in each preparation (Figure 3). Furthermore, elution profiles and UV spectra of the individual metabolites are also obtained as loadings from the PARAFAC analysis. The score vectors or matrices derived from the PARAFAC analyses, containing relative concentrations of the metabolites, were combined to a new matrix, $X_{reduced}$ which hence provide an extremely condensed yet comprehensive representation of the chromatographic data in terms of underlying concentrations. PCA was applied to that matrix in order to compare samples of St. John's wort subjected to standardization procedures described in the respective pharmacopoeias. The use of the score matrix, $X_{reduced}$, instead of the full three-way array, reduces the data set size and focuses the analysis on well-characterized constituents.

Figure 4A shows a 3D score plot from the PCA analysis of the score matrix, $X_{reduced}$. For interpretation purposes, 2D score plots of PC1 versus PC2 and PC2 versus PC3 (Figure 4B and 4C, respectively) are also shown. Clustering of replicate samples according to supplier is apparent, indicating that no artifacts were introduced during the measurement or the analysis. Moreover, no discrimination between capsules and tablets is apparent, similarly to the previous work based on ^{1}H NMR profiles.[28] Apart from that, the analysis discloses differences in sample composition over the entire collection of preparations. However, a significant proportion of all preparations (1, 2, 4, 11, 14, 16, 18, 19, 20, and 24) forms a cluster in the plot of PC1 versus PC2, which, however, accounts for only about 56% of the variation in data described by the PCA model.

Because the analysis of HPLC-PDA profiles is limited to well-characterized compounds, interpretation of the derived PCA model in terms of compounds accounting for the differences is straightforward. Preparations 3, 6, 15, and 17 are discriminated from other preparations in the negative direction of the first principal component (Figure 4B). The corresponding loadings (Figure 4D) reveal that the first principal component is negatively influenced by peaks 10 and 11 and positively influenced by peaks 2, 3, 4, and 9. Further analysis of the data reveals that preparations 3, 6, and 15 contain higher levels of **10** and **11**, whereas preparation 17 only contains a higher level of **11**. Furthermore, these preparations contain lower levels of **2**, **3**, **4**, and **9**, with preparation 17 containing the lowest levels of these flavonoid glycosides, which explains the apparent discrimination of the above-mentioned preparations in the negative direction of the first principal component. Preparations 12, 13, and 21 are discriminated to a minor degree in the negative direction of the first principal

(60) The Quality and Technology Website: http://www.models.life.ku.dk/source (warping toolbox), May, 2006.

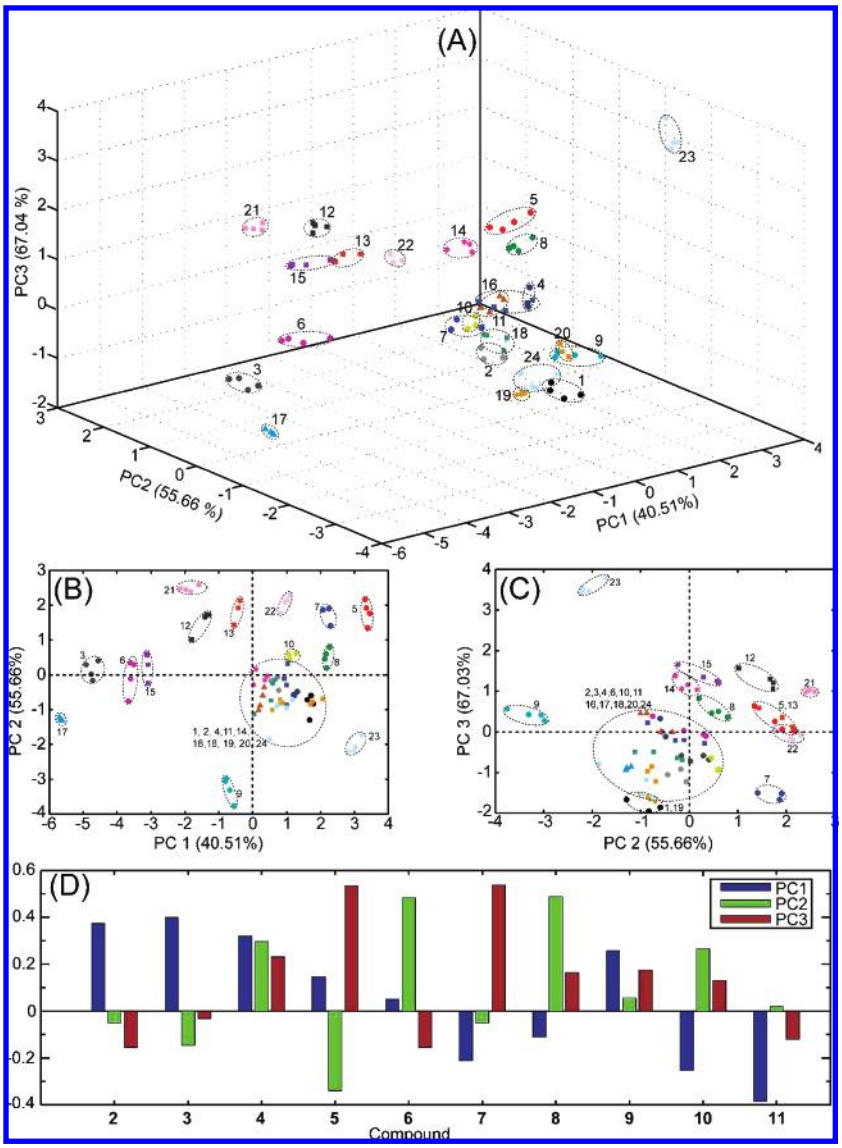


Figure 4. Results from PCA analysis of the reduced matrix, $X_{reduced}$, which contains relative concentrations of constituents identified in the HPLC-PDA profiles. The preparations are colored according to supplier and are represented by circles, squares, triangles, and stars for preparations manufactured in Europe, North America, Asia, and Africa, respectively. (A) 3D score plot of PC1 vs PC2 vs PC3. (B) 2D score plot of PC1 vs PC2. (C) 2D score plot of PC2 vs PC3. (D) Loading bar plot showing the influence of compounds **2**−**11** on PC1, PC2, and PC3.

component. This is caused by elevated levels of **7** and **8**, which also influence the first principal component in the negative direction (Figure 4D). The discrimination of preparations 5, 7, 8, and 23 from the remaining preparations in the positive direction of the first principal component is due to higher content of all four flavonoid glycosides, which influence the first principal component in that direction (**2**, **3**, **4**, and **9**).

There is no obvious clustering of preparations according to the continent of origin. Only preparations 12, 13, 21, and 22, four out of eight preparations from North America, display a clustering which can be correlated to the sample origin. Clustering of these preparations in the positive direction of the second principal component is due to higher levels of **4** and **8**. As seen in the loading bar plot (Figure 4D), the second principal component is also positively influenced by **6** and **10**. Elevated levels of **10** are observed in preparations 12 and 13, and this explains their discrimination in the positive direction of the second principal component. Preparation 9 and to some extent preparation 23 are discriminated from other preparations mainly in the negative direction of the second principal component. This discrimination is due to a higher content of **5**, with preparation 23 containing the highest level of this constituent. The more clear discrimination of preparation 9 in the negative direction of the second principal component is caused by the low content of **4**. Furthermore, preparation 23 also contains a relatively high level of **4**, which influences the second principal component in the positive direction. The higher levels of **5** and **4** in preparation 23 are also displayed in the analysis of loadings corresponding to the third principal component, since these constituents influence this component in the positive direction. In addition, peak 7 (**7**) influences the third principal component in the positive direction, and further analysis reveals that preparation 23 also contains a relatively high level of this constituent.

An interesting observation from the score plots shown in Figure 4 is that preparations 5, 7, and 8 are quite different. Preparations 7 and 8 are different batches of the same brand,

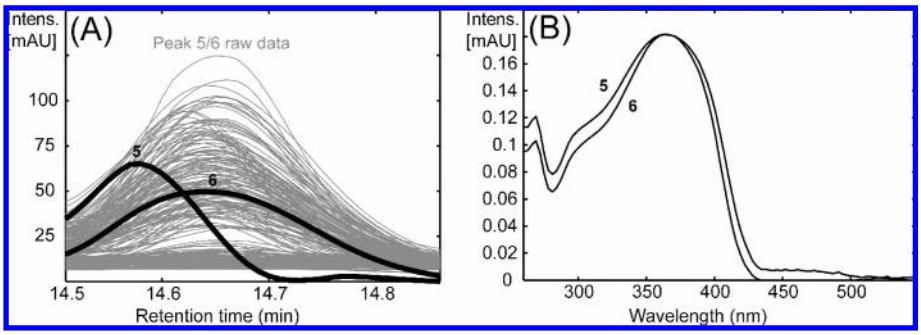


Figure 5. (A) Raw data (HPLC peaks) of peak 5/6 in 24 samples with 3 or 4 replicates of each sample (gray lines) and derived loadings from PARAFAC analysis that display elution profiles for miquelianin (**5**) and astilbin (**6**). (B) Loadings corresponding to pure UV spectra of **5** and **6** obtained by PARAFAC analysis of peak 5/6.

whereas preparation 5 was sold under a different brand name, but with the same label insert as the other two. The discrimination of preparation 5 from preparations 7 and 8 in the first principal component is due to a higher content of **3**, **4**, and **9**. Preparation 8 is discriminated from preparations 5 and 7 in the second principal component due to a lower level of **8**. Lower content of **4** and **10** causes the separation of preparation 7 from preparations 5 and 8 in the negative direction of the third principal component.

The significance of the initial use of PARAFAC for analysis of chromatographic peaks is illustrated by the analysis of peak 5/6. Inspection of this peak at single wavelengths or at the sum of all wavelengths (Figures 1 and 2) shows a symmetrical band that suggests the presence of one compound only. However, to achieve satisfactory fit of the PARAFAC model, two components were necessary, indicating that the peak resulted from two overlapping, different elution profiles. The raw data for peak 5/6 for all preparations together with the two loadings from the PARAFAC analysis, corresponding to individual elution profiles of **5** and **6**, are shown in Figure 5A. Concentrations of miquelianin and astilbin in each preparation are available in the score matrix derived from the PARAFAC analysis.

Thus, the unsupervised PARAFAC analysis of HPLC-PDA data can be used to mathematically resolve coeluting chromatographic peaks. Obviously, the ability to separate the constituents in coeluting peaks by use of PARAFAC only holds for constituents with different UV spectra. However, as seen in Figure 5A, PARAFAC enabled separation of elution profiles of **5** and **6**, which have very similar UV spectra (Figure 5B). The presence of two coeluting compounds was also apparent from the instrumentally much more elaborate HPLC-PDA-SPE-NMR-MS analysis, which allowed structural assignment of the coeluting constituents. The advantage of using PARAFAC as a decomposition method is that it provides unique solutions, which implies that the PARAFAC components will reflect chemically meaningful components, such as estimated pure component spectra, provided that the correct number of components is used and that the signal-to-noise ratio is appropriate. However, derived PARAFAC models can sometimes suffer from numerical problems, which need to be addressed appropriately in order to obtain chemically meaningful components. Components with almost identical but opposite sign or contribution are typical results of such numerical problems.[61] Thus, careful evaluation of the model must be performed to ensure that the right number of components is used.

The warping procedure and the PARAFAC model enabled exploratory and unbiased comparison of samples based on data obtained with simple and common analytical equipment (HPLC instrument with a PDA detector). The analysis demonstrated differences between various preparations in terms of flavonoid content. There are strong indications that flavonoids contribute to the antidepressant activity of St. John's wort.[1,20,21,62,63] Therefore, the product composition differences, such as those demonstrated in this work by chemometric analysis of HPLC-PDA data, may be highly significant for the quality of individual products.

CONCLUSIONS

In conclusion, HPLC-PDA-SPE-NMR-MS experiments in combination with HRMS led to identification of 12 constituents of commercial preparations of St. John's wort. Moreover, 24 different commercial preparations of St. John's wort were compared based on metabolomic analysis of three-way HPLC-PDA profiles. An extended algorithm of COW successfully aligned chromatographic profiles of all preparations, and subsequent, unsupervised PARAFAC analysis yielded relative concentrations of metabolites represented by individual chromatographic peaks, including peaks resulting from coeluting constituents. This work illustrates that chemometric methods, such as PARAFAC analysis, in combination with hyphenated techniques, offer excellent means of analysis of complex mixtures. The interpretation of the chemometric analysis of HPLC-PDA is straightforward in terms of compounds that account for similarities or differences between the samples. Differences in sample composition, revealed by chemometric analysis, may influence the pharmacological activity of St. John's wort preparations and may explain differences in outcomes of clinical trials.

ACKNOWLEDGMENT

The NMR equipment used in this work was purchased via a grant from "Apotekerfonden af 1991" (Copenhagen). The technical assistance of Ms. Birgitte Simonsen and Ms. Dorte Brix is gratefully acknowledged.

Received for review October 05, 2007. Accepted November 29, 2007.

AC702064P

(61) Smilde, A.; Bro, R.; Geladi, P. *Multi-Way Analysis: Applications in the Chemical Sciences*; John Wiley & Sons: Chichester, U.K., 2004.
(62) Butterweck, V.; Hegger, M.; Winterhoff, H. *Planta Med.* **2004**, *70*, 1008−1011.
(63) Nöldner, M.; Schotz, K. *Planta Med.* **2002**, *68*, 577−580.