Advances in near infrared spetroscopy in phytochemistry

C.W. Huck

Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University Innsbruck, Innrain 52a, 6020-Innsbruck, Austria.
e-mail: christian.w.huck@uibk.ac.at

Introduction

During recent years the pyhtopharmaceutical industry has continued to be an area of growth—between 1994 and 1997 sales of phytopharmacia in the United States increased by 93%, in Europe by15% and in Asia by 15%. Due to this increase in demand, new analytical methods for quality control are required and the analysis of plants, their extracts and natural compounds has become a demanding challenge. Therefore, near-infrared (NIR) spectroscopy, supported by sophisticated statistical software, offers a fast, reliable and non-destructive method for controlling not only chemical but also physical parameters at once. As a continuation of our earlier developed method for the quality control of *Flos Primulae veris* extracts[1,2] and its phenolic ingredients, an NIR method for the quantitative analysis of hypericin in *Hypericum perforatum* L. extracts was established. *St. John´s Wort* extract is used for the treatment of skin injuries, burns, neuralgia, antibacterial activity and as a treatment of mild to moderate depression.[3–9] In recent years, scientific studies to examine ingredients and quality of red wine also have become very important. The reason for this can be found in its health benefits.[10-12] Some papers have already discussed the problem of finding out the geographical origin of a wine without sensorial tests, for example with high-performance liquid chromatography (HPLC),[13] gas chromatography (GC) or pyrolysis mass spectrometry.[14] These classifications are only possible by the use of multivariate analyses. Due to the high demand for a fast analytical tool for the analysis of red wine, an NIR method to distinguish between different varieties and age groups was established. Correlation with results received from HPLC and HPLC coupled to mass spectrometry (HPLC-MS) also enabled the content of antioxidative phenolic ingredients to be predicted. Finally, based on these experiments, a qualitative model for the differentiation between *arabica* and *robusta* beans was established. This method has proved to be an interesting and powerful analytical tool for the coffee producing industry as quality and prices are fixed by the type of beans.

Materials and methods

The near infrared spectra used in this work were collected with a FT-NIR universal spectrometer from Büchi Labortechnik AG (Uzwil, Switzerland), which is equipped with a tungsten–halogen source and a PbS-detector. The absolute wavelength accuracy is ± 2 cm^{-1} over a wavelength range from 1000 to 2500 nm. The sample measurement was performed using fibre optics with ten scans for one average spectrum to eliminate inhomogenities. Software from Büchi Labortechnik was used: BCAP 5.0 for controlling the spectrometer and NIRCAL 3.0 for the processing of spectral data and chemometrics. All samples were heated to 23°C and scanned without any sample pretreatment. For reference analyses HPLC, HPLC-MS and capillary electrophoresis (CE) were used. All individual parameters can be found in the cited literature.[15–17] The determination of the ethanol/water content in red

Near Infrared Spectroscopy: Proceedings of the 10th International Conference

wine was carried out by gas chromatography with flame ionisation detection (GC-FID) and Karl–Fischer titration. GC-FID: HP1 fused silica (50 × 0.32 mm I.D.); gas, hydrogen 0.7 bar; injector temperature, 30°C; temperature from 50°C (4 min) up to 220°C (8°C min^{-1}) and up to 300°C (20°C min^{-1}); detector temperature, 300°C; split, 35 mL min^{-1} 1 : 120; sample size, 5 µl. Karl-Fischer titration: 20 µl of each extract were titrated on a 684 KF Coulometer (Metrohm, Filterstadt, Germany).

Quantitative analysis of hypericin in *Hypericum perforatum* L. extracts

The otpical pathlength used in the transflectance mode was 3 mm. Eighty three samples were measured over a wavelength range from 4500–9996 cm^{-1}. First derivation was calculated and the 332 spectra were divided into a learning (60%) and a validation set (40%). The smoothed average of the spectra was three points. For partial least square regression (PLSR) 15 factors were selected.

Qualitative wine analysis

The otpical pathlength used in the transflectance mode was 0.5 mm. After performing a second derivation of the recorded 172 spectra they were randomly partitioned into a calibration set (70% of total samples) and a validation set (30% of total samples). The smoothed average of all spectra was three points and six factors were used. With these data sets a principal component analysis (PCA) over a wavelength range from 4000 to10000 cm^{-1} was carried out.

Qualitative coffee analysis

Eighty five samples were measured in the reflectance mode over a wavelength range from 4500 to 9996 cm^{-1}. A first derivation and three points average smoothing were performed. Five factors were used for the calculation. 67% of the recorded 255 spectra were used for the calibration and 33% for the validation. With these data sets a PCA was carried out.

Results and discussion

The strategy scheme depicted in Figure 1 for the quantitative analysis of hypericin in *St. John´s Wort* can also be used as a general scheme. At the beginning, the extraction procedure and reference

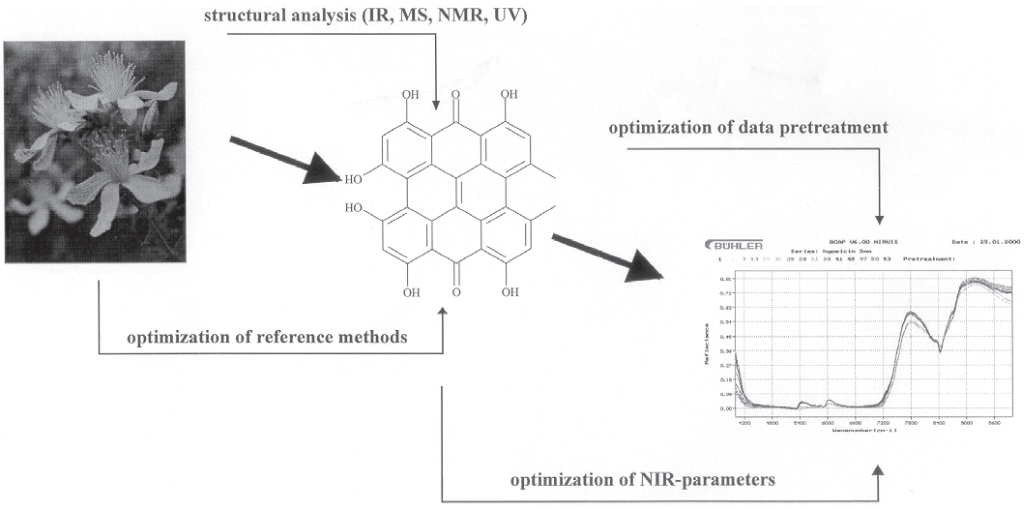

Figure 1. Strategy scheme for the analysis of hypericin in *Hypericum perforatum* L. extracts.

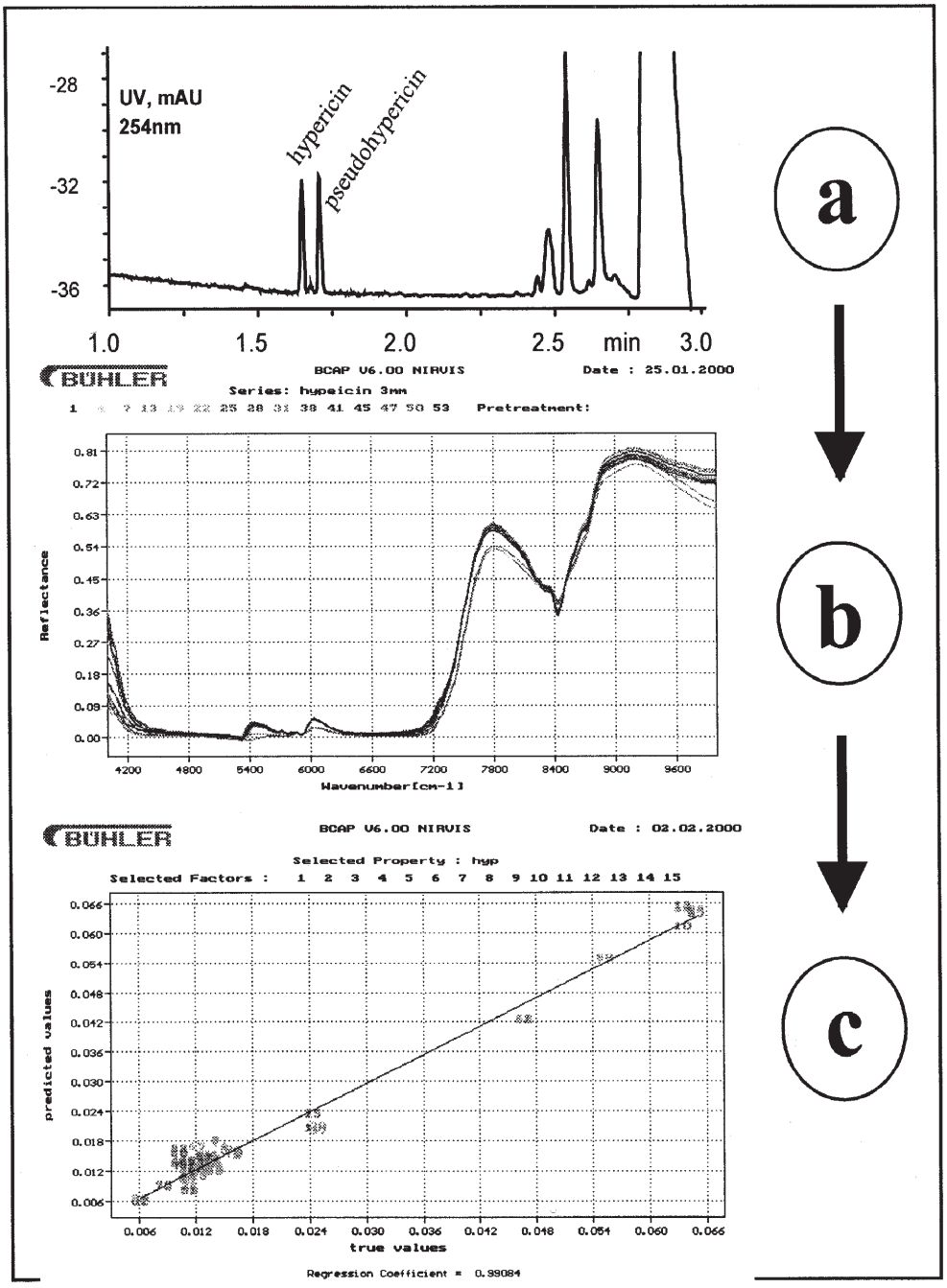

Figure 2. Quantitaive analysis of hypericin in *Hypericum perforatum* L. extracts. (a) Capillary electrophoresis (reference method); (b) NIR-spectra; (c) calibration curve.

method have to be established, optimised, evaluated and validated. After that, NIR parameters such as temperature and optical thin-layer thickness must be optimised. Furthermore, the best data pretreatment must also be found. For the control of the hypericin content in *Hypericum perforatum* L. extracts, CE was used as the reference method. Each extract was measured three-fold and cross-wise by CE [Figure 2(a)]. For the separation, a buffer system consisting of 26 mM phosphate, 0.001% HDB, 4.5% butanol and 20% acetonitrile (pH 2.65) was used. The calibration calculated the content of hypericin. After optimisation of the temperature and the optical thin layer, 332 spectra of 83 extracts were recorded in the transflectance mode with a NIRVIS instrument. Mathematical pretreatment and statistical analysis were carried out by performing PLS. By recording the NIR spectrum [Figure 2(b)] and calculation of its first derivation, characteristic absorption bands were identified. The most intensive band belonged to the vibration of the second overtone of the carbonyl group (5376 cm^{-1}), followed by the C–H stretch and C–H deformation vibration of ethanol (7212 cm^{-1}), the –OH vibration of water and ethanol (4440 cm^{-1}), the $-CH_2$ overtone (5742 cm^{-1}) and the $-CH_2$ / $-CH_3$ overtone (5808 cm^{-1}). All recorded spectra were transformed to their first derivative before being calculated in the linear PLS model. Fifteen principal components were necessary in order to reach the best calibration equation. The multivariate statistical method, PLS, is a full spectrum method. Therefore, the information of the whole recorded spectral range can be used for the calibration. In the course of model optimisation the best statistical results were obtained when the spectral information in the interval between 4500 and 9996 cm^{-1} was used for calculating the PLS. Finally, a correlation coefficient of 0.99084 for the calibration curve of NIR values against CE values helped to assess the linearity of the model [Figure 2(c)]. For a higher robustness of the calibration system, extracts with spiked hypericin were included in the curve. This method was proven to be a highly suitable spectroscopic tool to quantify hypericin in *Hypericum perforatum* L. extracts in the low percentage range.

As the variety of wine has a big influence on the quality of a red wine, a method to distinguish between different bottles made of different grapes was established. Fifty bottles of red wine of three different wine variety samples were inspected: *Cabernet Sauvignon*, *Lagrein* (both pure wines) and *Sangiovese* (Chianti). After recording the spectra, partitioning into a calibration and a validation set, average smoothing of three points and performance of a second derivation, PCA, over a wavelength range from 1010 to 2222 nm, was carried out. Figure 3(a) shows that

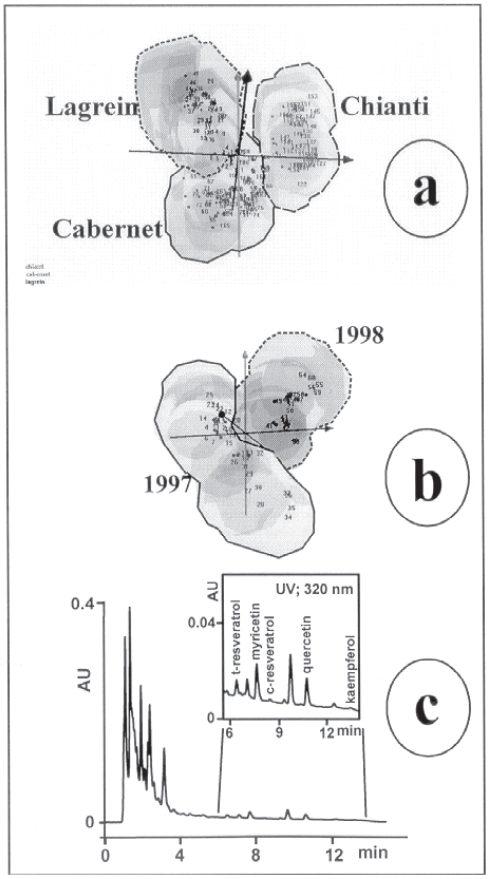

Figure 3. Qualitative analysis of red wine. (a) 3d-scatter plot of 3 different varieties; (b) 3d-scatter plot of two age groups of Cabernet; (c) HPLC-analysis (reference method).

the inspected wines (*Cabernet Sauvignon*, *Lagrein* and *Chianti*) can be classified very clearly by a 3-d scatter-plot. Each property (wine variety) could be assigned to one cluster. Due to the fact that the inspected wines come from very close geographical origins and two of them were produced using the same method in the same wine cellar at the same time, it is possible to estimate that the reason for this separation is the different wine variety. Furthermore, it was possible to distinguish between *Cabernet* 1997 and 1998 [Figure 3(b)]. One reason for this possible classification can be found in the different fingerprint caused by the phenolic ingredients, which was analysed by HPLC [Figure 3(c)]. Correlation of the results received from these HPLC analyses with those from NIR spectroscopy allows us to give a statement about the amount of phenolic ingredients found. If NIR spectroscopy identifies a wine as a *Cabernet Sauvignon* 1997, for example, the content of quercetin is between 18 and 30 μg mL^{-1}.

Based on these results, investigations in the field of coffee analysis were carried out. The development of an NIR spectroscopic method to distinguish between *arabica* and *robusta* green coffee beans is an important quality control tool, as the international coffee trade is conducted almost exclusively with these two varieties. With NIR in reflectance mode it is possible to perform a complete profile of the whole bean. At first, different extracts of roasted *arabica* and *robusta* beans were investigated using HPLC analysis of the three main ingredients which are necessary for the taste of coffee: caffeine, theobromine and theophylline [Figure 4(a)]. By this method, the different patterns of *arabica* and *robusta* could be confirmed. These differences were used in NIR to distinguish between the two varieties. Two hundred and fifty five spectra of 85 samples were recorded in the reflectance mode with a NIRVIS instrument. Mathematical pretreatment and statistical analyses were carried out by performing PCA. Recording the NIR spectrum of a green bean and calculating its first derivative again identified the characteristic absorption bands of xanthines [for structure see, for example, caffeine in Figure 4(a)]. The most intense band belonged to the vibration of the second overtone of the carbonyl group (5268 cm^{-1}), followed by the C–H stretch and C–H deformation vibration at 7116 cm^{-1}. All spectra were transformed to their first derivative before calculating in the PCA. Finally, the *arabica* and *robusta* beans could be classified very clearly by 3-d scatter plot [Figure 4(b)]. Each property (bean variety) could be assigned to one cluster. Due to the fact that *arabica* and *robusta* beans come from different cultivation areas, this method is highly suitable for a fast differentiation of green coffee varieties.

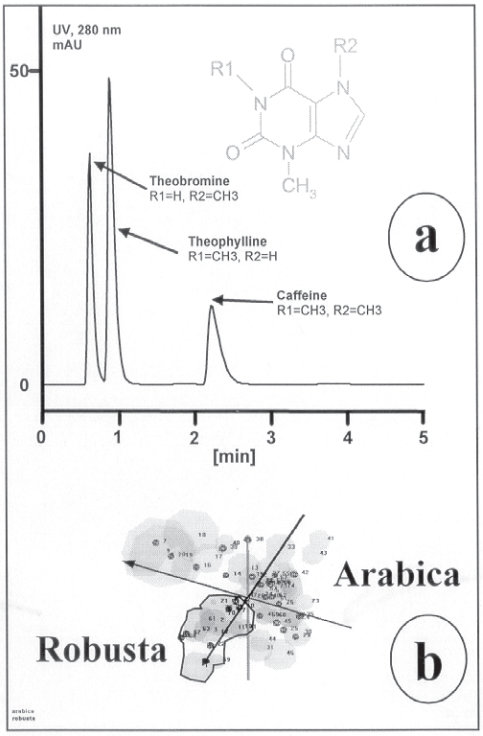

Figure 4. Qualitative differentiation between green coffee varieties. (a) HPLC analysis (reference method); (b) 3d-scatter plot.

References

1. C.W. Huck, R. Maurer, G.K. Bonn, M. Popp and N. Basener, *Near Infrared Spectroscopy: Proceedings of the 9th International Conference*, Ed by A.M.C. Davies and R. Giangiacomo. NIR Publications, Chichester, UK, p. 487 (2000).

2. C.W. Huck, R. Maurer, M. Popp, N. Basener and G.K. Bonn, *Pharm. Pharmacol. Lett.* **9,** 26 (1999).
3. C.M. Hawken, *St. John's Wort*. Woodland Publishing Inc., Pleasant Grove, USA (1997).
4. G. Lavie, Y. Mazur, D. Lavie and D. Meruelo*, Med. Res. Rev.* **15,** 111 (1995).
5. Ch.M. Schempp, K. Pelz, A. Wittmer, E. Schöpf and J.C. Simon, *The Lancet* **353,** 2129 (1999).
6. N. Rosenthal, *St. John's Wort The Herbal Way to Feeling Good*. Harper Collins Publisher Inc., New York, USA (1995).
7. J. Deltito and D. Beyer, *J. Affect. Disord.* **51,** 345 (1998).
8. J. Chang, *Biochem. Pharmacol.* **59,** 211 (2000).
9. E. Celotti, R. Ferrarini, R. Zironi and L.S. Conte, *J. Chromatogr. A* **730,** 47 (1996).
10. F. Fauconneau, P. Waffo-Teguo, F. Huguet, L. Barrier, A. Decendit and J.M. Merillon, *Life Sciences* **61,** 2103 (1997).
11. N. Worm, *Täglich Wein*, 5th Edn. Hallwag AG, Bern, Switzerland (1998).
12. S. Renaud and M. De Lorgeril, *The Lancet* **339,** 1523 (1992).
13. K. Sivertsen, B. Holen, F. Nicolaysen and E. Risvik, *J. Science of Food and Agriculture* **79,** 107 (1999).
14. L. Montanarella and M.R. Bassani, *Rapid Commun. Mass Spectrom.* **9,** 1589 (1995).
15. G. Stecher, C.W. Huck, M. Popp and G.K. Bonn, *Fresenius J. Analyt. Chem.*, in press (2001).
16. N. Schelosky, *Analytische Untersuchungen an Johanniskraut*, Ph.D. thesis, University of Innsbruck (1998).
17. R. Pirker, C.W. Huck and G.K. Bonn, *J. Chromatogr. B*, in press (2001).