Biochemical and Biophysical Research Communications **286,** 150–155 (2001)
doi:10.1006/bbrc.2001.5350, available online at http://www.idealibrary.com on IDEAL®

Automated Mode-of-Action Detection by Metabolic Profiling

Nelly Aranìbar,[1] Bijay K. Singh,[2] Gerald W. Stockton, and Karl-Heinz Ott[3,4]
BASF Agro Research, P.O. Box 400, Princeton, New Jersey 08543-0400

Received July 2, 2001

Rapid classification and identification of the mode-of-action of bioactive compounds applied to plants can be achieved by a robust and easy-to-use metabolic-profiling method. This method uses artificial neural network analysis of one-dimensional proton NMR spectra of aqueous plant extracts to rapidly classify changes in the total metabolic profile caused by application of crop protection chemicals.

Key Words: **herbicide; artificial neural network; NMR spectroscopy; metabonomics; metabolic profiling; mode-of-action; *Zea mays*.**

The term "metabolome" has recently been coined to describe the chemical profile or fingerprint of the metabolites in an organism. The metabolome reflects the life history of each individual plant, including age and environmental factors such as soil type and moisture content, temperature, stress factors, and exposure to applied fertilizers and crop protection chemicals. With the expectation that, following exposure to a herbicide, the herbicide's mechanism-of-action might be recognizable in the plant metabolome, we investigated whether such characteristics can be reliably detected in the NMR spectrum of a plant extract.

The gross chemical composition of various biological fluids has been investigated by a variety of chromatographic and spectral techniques, notably gas and liquid chromatography (1–3), NMR spectroscopy (4–11), mass spectrometry (12, 13), and infrared spectrophotometry (14). In animal/human fluids, much of the NMR research has been directed towards disease characterization and diagnosis (4–11, 15–18). NMR has provided information on biosynthesis (19), and on the effects of herbicides on metabolism (20) and mode-of-action (21), or used in investigations of whole plants (22, 23). A variety of computational methods have been applied for the statistical analysis of spectral data (24–28), including artificial neural networks (29–32). In many cases, however, it was found that environmental factors contribute significant "noise" to the metabolite profile and reproducibility has often limited the applicability. Furthermore, in many reports only two states (e.g., normal vs treated) are simultaneously distinguished. A robust NMR method able to simultaneously detect multiple treatment groups has not yet been described. In the search for new pharmaceuticals and crop protection chemicals, it is sometimes desirable to have a fast and reliable means to detect the mode-of-action of a new active compound, or pinpoint unusual phenotypes by an altered metabolic profile.

There are currently over twenty established biochemical mechanisms for the numerous commercial herbicides used in agriculture. We describe in this paper the automated neural network analysis of ^{1}H NMR spectra of raw, aqueous plant extracts that can simultaneously, and with high reliability, detect the modes-of-action of the various herbicides. The computational classification utilizes artificial neural network methods that are shown to produce robust assignments under conditions where changes in sample characteristics are very small and often close to the statistical variation between samples.

EXPERIMENTAL

Zea mays seeds (Pioneer 3514) were set to germinate in paper towel rolls in tap water for 5 days in the growing chamber. The environment was adjusted to "summer conditions" (day/night ratio of 14/10 h, regulated temperature of 27°C and humidity of 70%). After germination the seedlings were visually inspected. Seedlings that were homogeneous in size and appearance were selected, set in 50-ml amber bottles in 25-ml Hoagland nutrient solution (12 ml micronutrients stock solution, 12 ml FeEDTA (5 g/100 ml), 2.4 ml KH_2PO_4 (1 M), 24 ml $MgSO_4$ (1 M), 60 ml KNO_3 (1 M), 60 ml $Ca(NO_3)_2$ (1 M), and 60 ml MES buffer (200 mM), diluted to 12 litre with deionized water) and grown for five more days, after which they reached the three-leaf stage. At this point, 20 μl of a stock solution of technical grade

[1] Present address: Chemistry Department, Princeton University, Princeton, NJ 08544. E-mail: aranibar@princeton.edu.
[2] Present address: BASF Plant Science, 26 Davis Drive, Research Triangle Park, NC 27709.
[3] Present address: Bristol–Myers Squibb Company, 311 Pennington-Rocky Hill Road 3A-005, Pennington, NJ 08534.
[4] To whom correspondence should be addressed. Fax: (609) 818-3100. E-mail: karl-heinz.ott@bms.com.

0006-291X/01 $35.00

TABLE I

Herbicides Used in the Experiments Listed with Their Mode-of-Action, Chemical Class, and Number of Spectra Recorded

Treatment	Mode-of-action	Chemistry	# Spectra
Imazethapyr	AHAS inhibitor	Imidazolinone	30
Imazapyr	AHAS inhibitor	Imidazolinone	12*
Imazamethabenz	AHAS inhibitor	Imidazolinone	12*
Chlorsulfuron	AHAS inhibitor	Sulfonylurea	12*
Sulfometuron methyl	AHAS inhibitor	Sulfonylurea	12*
Sethoxydim	ACCase inhibitor	Cyclohexenone	12
Glyphosate	EPSPS inhibitor	*N*-phosphonoglycine	12
Diuron	PS II inhibitor	Chlorophenylurea	12
Untreated	—	(Acetone)	30

Note. AHAS, acetohydroxy acid synthase enzyme, also known as acetolactate synthase (ALS); ACCase, acetyl CoA carboxylase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; PSII, photosystem II. The spectra were recorded for at least two fully independent batches of plants. The spectra marked by * were used in a blind test of the NN procedure (see text).

herbicide in acetone (see Table I) was added to the hydroponic solution or applied to the second leaf (with similar results). The control group of "Untreated Plants" received 20 μl acetone only and all of the plants were returned to the growing chamber. Twenty-four hours posttreatment, the plants were harvested by excising between the coleoptile and the first leaf collar. At this time, the plants show only slight growth stunting in response to the treatments (see Fig. 1). The first leaf sheet was separated and the meristematic tissue (approximately 250 to 300 mg per plant) was collected, flash frozen in liquid nitrogen in a cryogenic 3 ml tube, and stored in a liquid nitrogen freezer until further use. The plant meristems were each pulverized in a mortar (under liquid N_2), suspended in 2.4 ml of HCl solution (0.25N) and centrifuged at 14000*g*, 4°C, for 60 min. The NMR samples were prepared from 0.8 ml of the supernatants and 0.2 ml D_2O (with TSP 0.05 w/v) and kept on ice.

The 500 MHz ^{1}H NMR spectra of plant extracts were recorded at 295K using a Bruker AMX 500 NMR spectrometer equipped with a TXI 5 mm probe with 1 sec solvent presaturation (60 dB), 256 transients of 16K complex points per sample. The data were multiplied with an exponential function (LB = 0.5 Hz), Fourier transformed, and manually phase- and baseline-corrected. The processed spectra were exported in the JCAMP file format, automatically processed using a package of Perl scripts that presented the data to the Stuttgart Neural Network Simulator (33) and reported the results. In short, the NN analysis as applied here consists of: (i) Definition of a NN topology: three layers, comprising one input layer with 1080 nodes, one hidden layer with six or twelve nodes, one output layer with one node for each class (six classes in the examples presented here). The NN units were represented by a logistic activation function, and all units were fully connected with the adjacent layer. The input layer represents the spectral information and is initialized with the pattern created as described above. For training the NN, the output layer is initialized with a corresponding vector that describes the desired answer of the NN for a given input vector. For the examples presented here, the definition of the output nodes is as follows: 1st node, untreated; 2nd node, AHAS inhibitor; 3rd node, ACCase inhibitor; 4th node, EPSPS inhibitor; 5th node, PSII inhibitor; 6th node, dead plant. Note that the enzyme abbreviations are defined in the legend to Table I. The hidden layer and all connections are initialized using random values in the range of $[-1, 1]$. (ii) Presentation of a training set (a subset of the pattern, with known assignments for the output nodes) to this NN, and the training, i.e., initialization and adjustment of the weights of the connections in an iterative manner using a learning function until convergence or a step limit is reached. During this step, a validation set (a subset of the patterns different from those used as the training set) can, optionally, be periodically presented to the NN to gauge the performance of the NN and detect possible "overtraining." (iii) A test set (a pattern for which the output nodes are not defined, i.e. the mode-of-action unknown) can then be presented to the NN for classification.

We found the "Resilient Backpropagation" (Rprop) learning function for training the NN to be rapidly converging and consistent in producing reliable networks with the following learning parameters: the initial update value $\Delta_0 = 0.1$, the limit for the maximum step size, $\Delta_{max} = 50$, and the weight decay exponent $\alpha = 4$ (a range of 3–9 was found suitable for α; the learning is rather insensitive to changes in other two parameters). The training was done in cycles of 25 steps, after which the network was saved. The validation set was presented and the network error on the validation set was calculated. This procedure was repeated for up to 20 epochs (500 cycles total) and the network that produced the minimum error on the validation set was kept.

RESULTS

In Fig. 2, representative ^{1}H NMR spectra of corn meristematic tissue extracts of plants treated with the herbicides diuron, sethoxydim, glyphosate, and imazethapyr are compared with a spectrum of an untreated plant extract. The effects of the herbicide treatment, although small, are discernible in the spectra even though the variances between spectra of different groups and within the spectra of the treatment groups are comparable. Initial experiments with several batches of untreated and imazethapyr-treated plants demonstrated that an automated neural network procedure can, with almost absolute certainty, distinguish between the two treatment groups. A network trained with a small subset of spectra (two arbitrary spectra of each treatment group) was capable of correctly classifying the other samples with output node activation consistently being >0.99 for the correct class and <0.01 for the other classes. Second, a neural network, trained with spectra of two or more chemically different herbicides can also be used to recognize, with high reliability, the same treatment across different batches of plants. For example, using untreated, diuron-, sethoxydim-, and glyphosate-treated samples from a first batch of plants for training of a NN, the samples in a second batch are all correctly classified by this NN, as illustrated in Fig. 3.

FIG. 1. Corn seedlings in Hoagland solution, 24 h after application of the herbicides (the control received 20 μl acetone). The seedlings are 10 days old and at the three-leaf stage. The only observable symptoms on the treated plants are some damage to the second leaf, where herbicide treatment has been applied, and slight stunting.

The third example demonstrates that a NN can also generalize and recognize the same mode-of-action among chemically different herbicides. Since there are many herbicides from a wide variety of chemistries that act as AHAS inhibitors, we performed a blind test (the analyst performing the neural network calcula-

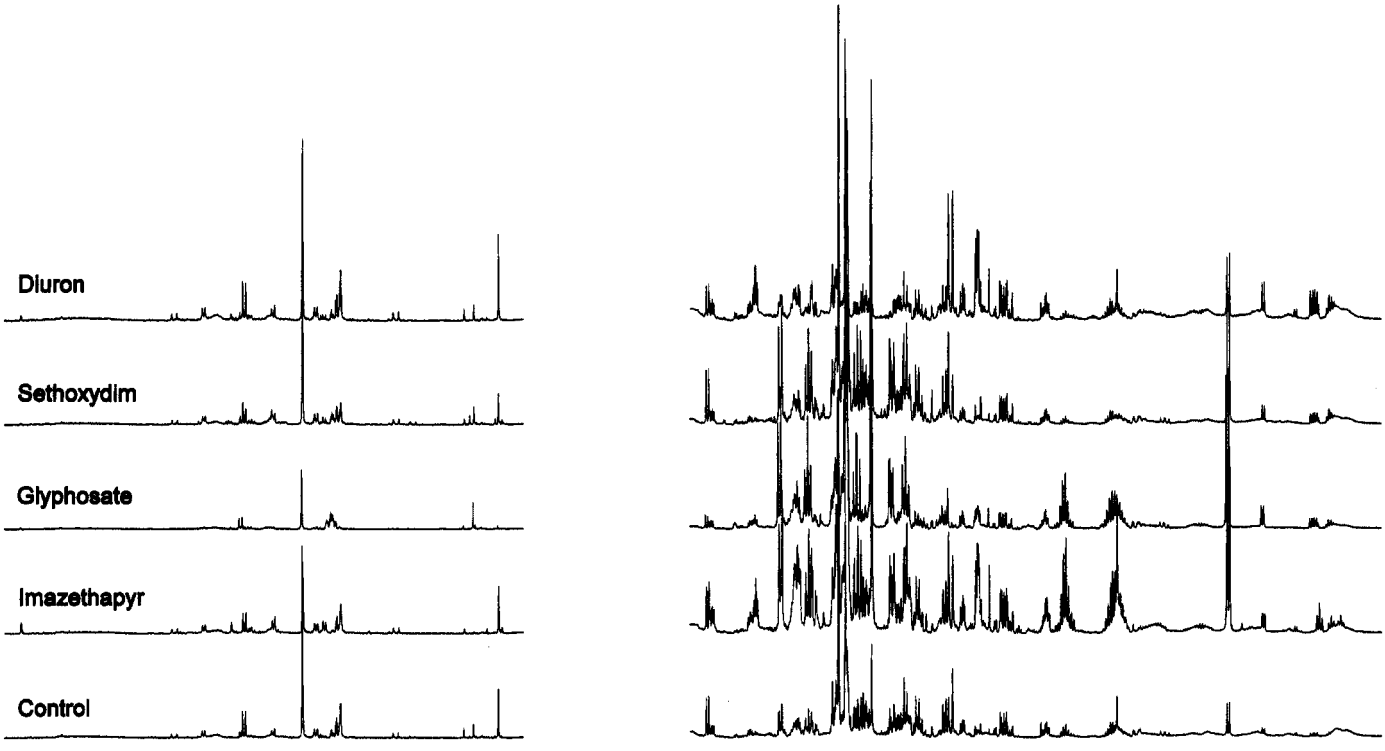

FIG. 2. Sections of ^{1}H NMR spectra (as presented to the NN) of *Zea mays* (corn) meristematic tissue extracts for plants treated with the herbicides diuron (top), sethoxydim, glyphosate, and imazethapyr, and untreated control (bottom).

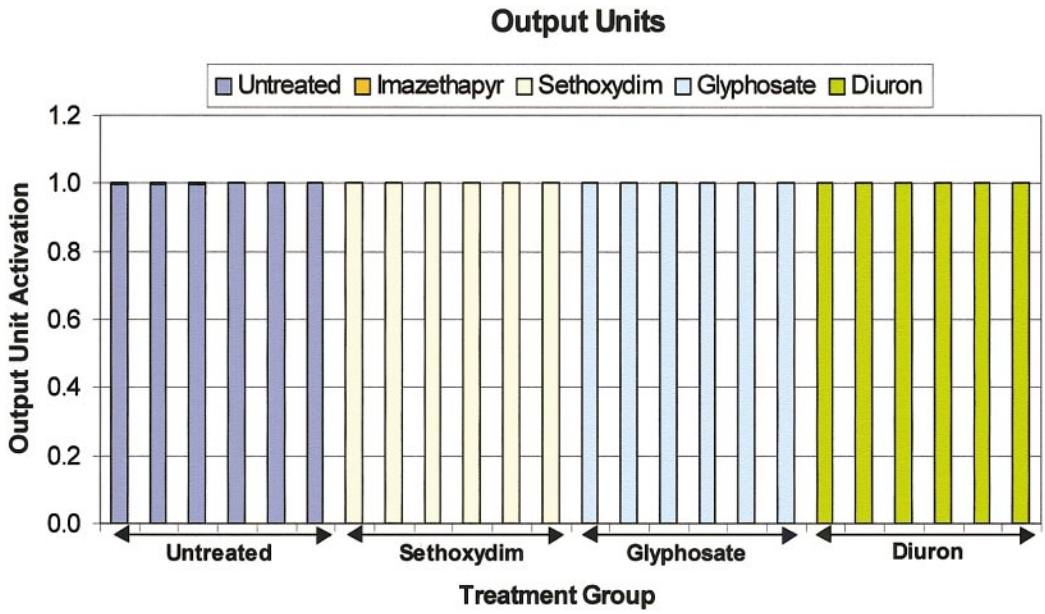

FIG. 3. Graph of output unit activation vs herbicide treatment for sethoxydim-, glyphosate-, diuron-treated, and untreated plants. The NN used was trained with spectra of a first batch of plants that contained the same treatment regimes as that of a second batch. A stacked column representation was chosen to visualize the output values of the output units. The output units are encoded by the color of the column, e.g., the blue column indicates activation of the "Untreated" output unit. Typically, a value of greater than 0.6 for one output unit, with no other unit having a value of greater than 0.4, is considered sufficient for assignments. The graph demonstrates that the correct output unit (color) is activated for each sample.

tions had no knowledge of the actual treatment, and output nodes were not defined) in which the network trained with the samples from the experiments described above was used to classify the mode-of-action of two other imidazolinones (imazapyr and imazamethabenz) and two sulfonylureas (chlorsulfuron and sulfo-

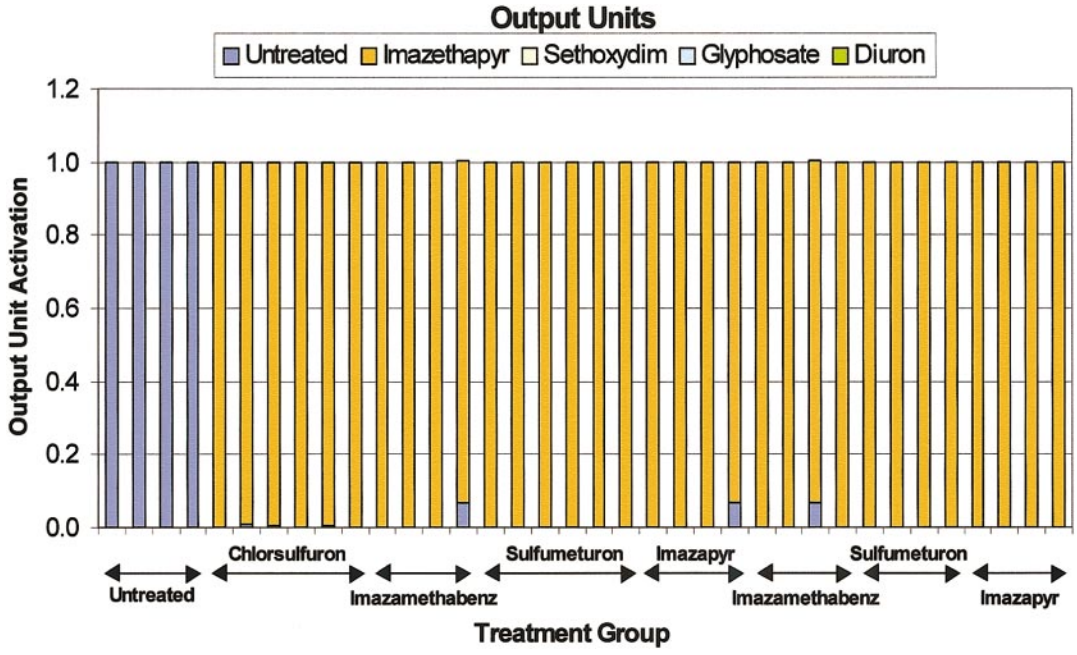

FIG. 4. Graph of output unit activation vs herbicide treatment group for two batches of several chemically different AHAS inhibitors (chlorsulfuron, imazamethabenz, sulfometuron methyl, and imazapyr). The graph demonstrates that all of these herbicides are classified by the neural network as "imazethapyr" due to their mutual mode-of-action (AHAS inhibition).

meturon methyl). The results represented in Fig. 4 show that all samples from plants that were treated with AHAS inhibitors are coclassified with imazethapyr when presented to the trained network.

DISCUSSION

The metabolic profile can be a very sensitive indicator of environmental influences, the genotype, phenotype, and condition and developmental state of an organism. Metabolic profiling is used to detect and analyze changes of the total metabolic state of an organism due to patho-physiological stimuli. Other approaches for quantitative metabolite profiling often require complex extraction and purification procedures that precede the analysis, using established separation and spectroscopic methods such as LC, GC, IR, or NMR, or hyphenated methods such as GC-MS, LC-MS, or LC-NMR. Although assignment of the NMR signals is not necessary for the classification of known modes-of-action, assignment of peaks pertaining to altered metabolite concentrations can be useful to identify new modes-of-action. Individual or small sets of signals are also often selected to define biomarkers associated with a treatment or disease, thus providing an intuitive understanding of the response. The method introduced here can complement such approaches, for example, when used in the form of a primary screen, and as a robust, medium-throughput method. It is particularly useful when a moderate number of different outcomes (e.g., modes-of-action, disease states) can be predefined.

Even though the low sensitivity of NMR spectroscopy has been a drawback for many analytical applications, it is actually an advantage for recognizing modes-of-action of crop protection chemicals. The herbicides and their metabolites are usually not visible in the NMR spectra since the herbicides are applied in low concentrations to the leaves or roots of the plant and only a negligible amount is absorbed and translocated to the meristem. What we observe is the spectral manifestation only of the variations in the pool of the plant metabolites.

CONCLUSION

In conclusion, this work has shown the capability of ^{1}H NMR spectroscopy of plant extracts, in combination with artificial neural network analysis, to accurately discriminate the modes-of-action of several different herbicides. The method is reliable when the experimental conditions are well controlled and accurately reproduced under standard conditions. We have optimized growing conditions, extraction procedures, and the bioanalytical methodology to produce highly reproducible conditions, thus creating a robust profiling method that is capable of detecting the many different herbicidal modes-of-action. Using only a small amount of tissue, the method is able to detect minute differences in a plant's metabolic profile even at an early stage of growth, where phenotypic changes are barely visible. The preparation and analysis procedures are simple and fast enough to permit screening of libraries of active compounds, with results being automatically and almost instantaneously reported, whereas traditional biochemical methods for mode-of-action determination require substantial experimental effort. A full paper, that describes the application of the method to over twenty modes-of-action relating to the ca. three hundred commercial herbicides, and to distinguish homozygotic and heterozygotic herbicide resistant strains from wild type plants, will be presented elsewhere.

ACKNOWLEDGMENTS

We thank Dr. J. L. Johnson for helpful comments on the manuscript, Dr. S. Rajan for optimizing the performance of the NMR spectrometer, and Dr. L. A. Wilson for the photography.

REFERENCES

1. Petroff, O. A. C. (1988) *Comp. Biochem. Physiol.* **90B 2,** 249–260.
2. Fiehn, O., Kopka, J., Doermann, P., Altmann, T., Trethewey, R. N., and Willmitzer, L. (2000) *Nat. Biotechnol.* **18,** 1157–1161.
3. Sauter, H., Lauer, M., and Fritsch, H. (1991) Synthesis and Chemistry of Agrochemicals II (Baker, D. R., Fenyes, J. G., and Moberg, W. K., Eds.), *ACS Symposium Series 443,* pp. 288–299, American Chemical Society, Washington, DC.
4. Nicholson, J. K., Sadler, P. J., Bales, J. R., Juul, S. M., MacLeod, A. F., and Sonken, P. H. (1984) *Lancet* **2,** 751–752.
5. Ohsaka, A., Yoshikawa K., and Matsuhashi T. (1979) *Jpn. J. MedSci. Biol.* **32,** 305–309.
6. Nicholson, J. K., and Wilson I. D. (1989) *Prog. NMR Spectr.* **21,** 449–501.
7. Lee, H-S., Chung, Y. H., and Kim C. Y. (1991) *Hepatology* **14,** 68.
8. Bales, J. R., Higham, M., Howe, I., Nicholson, J. K., and Sadler, P. J. (1984) *Clin. Chem.* **30,** 426–432.
9. Haug, H., and Schramm, C. (1975) *Clin. Chem.* **21,** 1025.
10. Rabenstein, D. L., Millis, K. K., and Strauss, E. J. (1988) *Anal Chem.* **60,** 1380A–1391A.
11. Bell, J. D., Brown, J. C. C., Nicholson, J. K., and Sadler, P. J. (1987) *FEBS Lett.* **215,** 311–315.
12. Matsumoto, I., and Kuhara, T. (1996) *Mass Spectrom. Rev.* **15,** 43–57.
13. Wolfender, J. L., and Hostettmann, K. (1996) Applications of Modern Mass Spectroscopy in Plant Science Research (Newton, R. P., and Walton, T. J., Eds.), Oxford Press.
14. Jackson, M., and Mantsch, H. H. (1996) Infrared Spectroscopy of Biomolecules (Mantsch, H. H., and Chapman, D., Eds.), Wiley–Liss, Inc.
15. Nishijima, T., and Fujiwara, K. (1997) *Jpn. J. Clin. Oncology* **27,** 13–17.
16. Somorjai, R. L., Dolenko, B., Nikulin, A. K., Pizzi, N., Scarth, G., Zhilkin, P., Halliday, W., Fewer, D., Hill, N., Ross, I., West, M., Smith, I. C. P., Donnelly, S. M., Kuesel, A. C., and Brière, K. M. (1996) *JMRI* **6,** 437–444.
17. Holmes, E., Foxall, P. J. D., Neild, G. H., Beddell, C., Sweatman,

B. C., Rahr, E., Lindon, J. C., Spraul, M., and Nicholson, J. K. (1994) *Anal. Biochem.* **220,** 284–296.

18. Hahn, P., Smith, I. C. P., Leboldus, L., and Littman, C. (1997) *Cancer Res.* **57**(16), 3398–3401.
19. Lutterbach, R., and Stockigt, J. (1995) *Phytochemistry* **40,** 801–806.
20. Prabhu, V., Chatson, K. B., Abrams, G. D., and King, J. (1996) *J. Plant Physiol.* **149,** 246–250.
21. Hole, S. J. W., Howe, P. W. A., Stanley, P. D., and Hadfield, S. T. (2000) *J. Biomolec. Screening* **5,** 335–342.
22. Schneider, B. (1997) *Planta* **203,** 1–8.
23. Pope, J. M., Jonas, D., and Walker, R. R. (1993) **173,** 177–186.
24. Jackson, M., Mansfield, J. R., Dolenko, B., Somorjai, R. L., Mantsch, H. H., and Watson, P. H. (1999) *Cancer Detect. Prevent.* **23,** 245–253.
25. Shaw, R. A., Kotowich, S., Eysel, H. H, Jackson, M., and Thomson, G. T. D. (1995) *Rheumatol. Int.* **15,** 159–165.
26. Mansfield, J. R., Sowa, M. G., Scarth, G. B., Somorjai, R. L., and Mantsch, H. H. (1997) *Anal. Chem.* **69,** 3370–3374.
27. Mansfield, J. R., Sowa, M. G., Scarth, G. B., Somorjai, R. L., and Mantsch, H. H. (1997) *Comput. Med. Imag. Graph.* **21,** 299–308.
28. Eysel, H. H., Jackson, M., Nikulin, A., Somorjai, R. L., Thomson, G. T. D., and Mantsch, H. H. (1997) *Biospectroscopy* **3,** 161–167.
29. Lisboa, P. J. G, Branston, N. M., El-Deredy, W., and Vellido, A. (1997) Proceedings of the IEEE/INNS International Joint Conference on Neural Networks, Houston, 1385–1390.
30. Lisboa, P. J. G., Kirby, S. P. J., Vellido, A., Lee, Y. Y. B., and El-Deredy, W. (1998) *NMR Biomed.* **11,** 225–234.
31. Nicholson, J. *et al.* (1995) *J. Pharm. Biomed. Anal.* **13,** 205–211.
32. Hiltunen, Y., Heiniemi, E., and Ala-Korpela, M. (1995) *J. Mag. Reson. B* **106,** 191–194.
33. Zell, A. (2000) Simulation neuronaler Netze, R. Oldenbourg Verlag, München.