INTRODUCTION

The “natural” and “healthy” image of fruit juices and derived products is one of the reasons leading to their dynamic evolution in the global food sector. The fruit juice industry purchases its raw materials in a world-wide market. Despite a general trend in the food industry to install suitable traceability systems, the need for analytical control of authenticity and legal conformity has increased over the last years. Due to the relatively large natural variability and more-and-more ingenious adulterations observed on the market, sophisticated control has become an important economic factor in the fruit juice industry. Therefore the European fruit juice industry runs an own independent and global advanced quality control system for raw materials, SGF/IRMA (SGF International e.V. - Sure Global Fair - International Raw Material Assurance). SGF/IRMA co-operates with national and regional control organisations for market finished fruit juice products organised in the European organisation EQCS (European Quality Control System). Such control systems, official control bodies and individual companies have to analyse a very large number of products. The whole range of possible quality deviations can not be tested by routine analyses for cost reasons in every case.

Main deviations and adulterations of fruit juices are the addition of sugar, water, flavour compounds (ref. 1) or cheaper fruit types than those declared on the label (ref. 2). Furthermore non-permitted production techniques can also lead to quality deviations. Consequently a need exists for a robust high throughput screening method checking the products under multiple relevant aspects simultaneously. Taking into account this requirement the present paper describes the SGF-Profiling (Spin Generated Fingerprint Profiling) as a new approach based on 1H-NMR-techniques combined with advanced statistical evaluation. The initial intention was to create a low cost screening method which is able to pre-select samples for submission to conventional target analysis for specific quality aspects. The outcome of one single SGF-Profiling experiment may be a series of analytical data, characterising the sample with respect to different individual quality aspects.

ANALYTICAL PRINCIPLE

Nuclear magnetic resonance (NMR) is one of the principal techniques used to obtain physical, chemical, electronic and structural information about one molecule or different substances in a complex mixture such as fruit juices. If the magnetic properties of protons are used we speak about proton-NMR or 1H-NMR.

The SGF-Profiling is based on a full 1H-NMR-spectrum. In principle all NMR spectrometers are suitable for this analysis. In order to achieve the best compromise between information content and analytical costs a 400 MHz spectrometer with flow injection cell and automatic sample preparation unit was used for the work described here. Automatic sample preparation with a Gilson Liquid Handler into NMR tubes is also possible, however with reduced speed and increased costs. In this case a standard sample changer can be used to handle the NMR tubes.

The interpretation of data is carried out by statistical comparison between a reference spectral database of authentic juices containing 1H-NMR spectral information and the samples to be reviewed. For a suitable calibration of the system the reference database should contain spectra of representative samples covering the natural variability of the examined sample type. The quality of any analytical answer improves with the number of representative reference spectra. All spectra have to be acquired under identical conditions to ensure comparability.

Depending on the individual case, a variety of statistical tools can be used for a chemometric evaluation to perform spectrum characterisation and comparison. A non-exhaustive list of possible mathematical approaches is given in table 1. Furthermore the identification and quantification of individual components can be per-
formed by using information from a reference compound NMR spectral base.

As an example of the multitude of possible exploitations of one unique sample spectrum the following possibilities are summarised. An orange juice or orange juice concentrate could be checked for the addition of different foreign fruits such as mandarin or grapefruit, the addition of different types of sugar syrup, the use of pulp wash (a low quality by-product from the juice production), any added preservatives or the presence of microbiological metabolites as indicator for product spoiling. Information about the geographical origin and/or the employed production technique could also be obtained. For each of these aspects the information in the corresponding reference spectra should be evaluated in different statistical models. Every model applies one or more dedicated mathematical approach and each of the models is based on the same proton NMR experiment which means all information is extracted from the same raw data. For each sample only one single proton NMR analysis is carried out. This is true for reference samples during construction of the models as well as for the sample to be tested.

Additional analyses using the same spectra could be carried out to identify and quantify individual compounds in the total juice matrix. This mixture analysis can be supported by a data bank which contains the original NMR spectra of the metabolites, measured under exactly the same conditions in terms of pH, temperature and NMR acquisition and processing parameters.

**EXPERIMENTAL SECTION**

**INSTRUMENTATION**

$^1$H NMR measurements were made using two Bruker AV 400, operating: (I) in a flow injection mode (120 µl sample volume) and (II) in 5 mm tube mode at 400 MHz proton resonance frequency. Spectra were recorded at 300 K. For automatic sample transfer in flow mode a GILSON 215 liquid handler was used.

**COLLECTION OF SAMPLES**

Within the scope of the present study, authentic fruit juice reference samples from main cultivation areas, varieties as well as different processing technologies were included. It was identified as extremely important for the study that only authentic samples are used for the establishment of the reference database and that analysed samples have been processed as practised in the industry. Therefore authentic samples were taken by qualified SGF/IRMA inspectors in different countries from the run-

---

**TABLE 1:** Overview of important multivariate statistical methods applied in the context of metabonomics data analysis. Often, methods are combined, e.g. PCA-DA or PLS-DA. Input to the statistical methods can be any parameter derived from NMR spectra, as metabolite concentrations, bucket intensities, coefficients from wavelet decomposition of spectra, etc. (ref. 3,4)

<table>
<thead>
<tr>
<th>Unsupervised methods</th>
<th>Supervised methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Principal components analysis (PCA)</td>
<td>• Discriminant analysis (DA)</td>
</tr>
<tr>
<td>• Factor analysis (FA)</td>
<td>• linear DA (LDA)</td>
</tr>
<tr>
<td>• Independent component analysis (ICA)</td>
<td>• quadratic DA (QDA)</td>
</tr>
<tr>
<td>• Exploratory projection pursuit</td>
<td>• reduced rank LDA</td>
</tr>
<tr>
<td>• Multidimensional scaling (MD)</td>
<td>• regularized DA (RDA)</td>
</tr>
<tr>
<td>• Cluster analysis (CA)</td>
<td>• k-nearest neighbour method (K-NN)</td>
</tr>
<tr>
<td>• hierarchical CA</td>
<td>• Soft independent modeling of class analogies (SIMCA)</td>
</tr>
<tr>
<td>• non-hierarchical CA</td>
<td>• UENQ, DASCO, CLASSY</td>
</tr>
<tr>
<td>• Self-organising maps</td>
<td>• Classification trees</td>
</tr>
<tr>
<td></td>
<td>• Regression/calibrations methods</td>
</tr>
<tr>
<td></td>
<td>• partial least squares (PLS)</td>
</tr>
<tr>
<td></td>
<td>• principal components regression (PCR)</td>
</tr>
<tr>
<td></td>
<td>• reduced rank regression (RRR)</td>
</tr>
<tr>
<td></td>
<td>• ridge regression (RR)</td>
</tr>
<tr>
<td></td>
<td>• orthogonal signal correction PLS (O-PLS)</td>
</tr>
<tr>
<td></td>
<td>• Canonical correlation analysis (CCA)</td>
</tr>
<tr>
<td></td>
<td>• Neural Networks (NN)</td>
</tr>
<tr>
<td></td>
<td>• Genetic algorithms (GA)</td>
</tr>
<tr>
<td></td>
<td>• Support Vector Machines (SVM)</td>
</tr>
</tbody>
</table>
ning production. Each of the samples was accompanied by full signed documentation including origin and condition of processed fruits as well as the applied technology in order to guarantee traceability. Results from this paper include juice and juice concentrate or puree samples derived from different fruit types as listed in Table 2.

Sterile samples were stored in aluminium bags at ambient temperature and non sterile samples were stored at -18 °C before sample preparation for NMR measurements.

SAMPLE PREPARATION

After thawing, the juice samples are centrifuged at 1200 x g for 6 - 10 min. Concentrates are diluted (e.g. 360 mg + 2.0 ml distilled H2O → juice at Brix of between 11.2 ° - 11.8 °) before doing centrifugation. For tube samples 675 µl juice is transferred into 5 mm NMR tubes and 75 µl buffer is added to have a final sample pH between 3.1 and 3.4. For flow injection mode, 300 µl of sample material (10 % buffer added as in case of tubes) are filled into cryovials (1.8 ml). The buffer used contained 0.1 % of TSP (sodium salt of 3-(trimethylsilyl)-propionate acid-d4) and 0.013 % of sodium azide. The purees are handled in the same way, the supernatant after centrifugation is used for the NMR preparation.

1H-NMR SPECTROSCOPY

1H-NMR spectra of the resulting samples are acquired by using the NOESYPR1D (ref. 6) pulse sequence with continuous wave presaturation of the water resonance during the relaxation delay D1 of 10 sec duration and a mixing time of 10 ms and ns= 64. Time domain data are collected into 64 K data points over an acquisition time of AQ = 4 sec covering a sweep width of SW=20 ppm. An exponential window function with LB = 1 Hz and zero filling is used prior to Fourier transform of the data with zero filling into 64 K frequency domain points.

The data processing was performed using the standard spectrometer software. Baseline and phase correction was performed using an automated processing programme. The complete measurement procedure was performed under full automation controlled by standard automation software, either based on a 60 position sample changer or the liquid handler for automated flow injection.

PREPARATION OF NMR SPECTRA FOR STATISTICS AND CHEMOMETRICAL ANALYSIS

The data analysis is performed mainly by using the software package AMIX 3.6.8. Some particular questions are addressed by using inhouse developed scripts for MatLab 7.0. For chemometrical data analysis, spectra were segmented into chemical shift regions and respective segment integrals are calculated in order to setup so called bucket tables for input into Principal Component Analysis (PCA) and Partial Least Squares Regression (PLS). Bucketing parameters are adjusted to the specific aspects investigated.

<table>
<thead>
<tr>
<th>Fruit Type</th>
<th>number of measured samples*</th>
<th>thereof with complete documentation as authentic samples</th>
<th>Country code of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orange</td>
<td>202</td>
<td>108</td>
<td>Brazil, Cuba, Spain, Mexico, South Africa</td>
</tr>
<tr>
<td>Pineapple</td>
<td>47</td>
<td>47</td>
<td>Brazil, Philippines, Thailand, South Africa</td>
</tr>
<tr>
<td>Apple</td>
<td>61</td>
<td>60</td>
<td>Poland, China</td>
</tr>
<tr>
<td>Citrus reticulata varieties (e.g. mandarin, clementine)</td>
<td>23</td>
<td>21</td>
<td>Brazil, Spain, Italy, Mexico, Paraguay, Uruguay, South Africa</td>
</tr>
<tr>
<td>Peach</td>
<td>12</td>
<td>12</td>
<td>Greece, Italy</td>
</tr>
<tr>
<td>Apricot</td>
<td>5</td>
<td>5</td>
<td>Greece, Italy</td>
</tr>
<tr>
<td>Mango</td>
<td>1</td>
<td>1</td>
<td>India</td>
</tr>
</tbody>
</table>

*Including fruit juice test mixtures with defined adulteration and market samples

TABLE 2: Samples measured by SGF-Profiling (juices, concentrates, purees)
QUALITY CONTROL

RESULTS AND DISCUSSION

The following paragraphs deal with a heterogeneous choice of different items which is indicative for the multi purpose application of the SGF-Profiling. Examples are given for:

- Multi fruit type separation,
- Fruit type differentiation between citrus varieties (e.g. citrus sinensis and citrus reticulata)
- Differentiation of product categories (e.g. orange juice and orange juice made from concentrate)
- Linearity in chemometrical response (e.g. shifts in orange-mandarin mixtures)
- Crosslink between chemometrical analyses and structure specific information (e.g. missing malic acid in an orange juice sample and relative quantification of sucrose, α-glucose and β-glucose)
- Characterisation of compositional differences for two groups of similar products (e.g. apple juice concentrates from Poland and China).

MULTI FRUIT TYPE SEPARATION

Fig. 1 represents the 1H-NMR-spectrum of an orange juice concentrate. It shows the vast array and the huge dynamic range of NMR signals that could be used for bucketing. The amount of specific information on samples is high compared to a classical analytical profile obtained by an extended quantification of individual juice compounds, the so called “full compositional analysis”. However SGF-Profiling is not able to replace the full analysis, e.g. parameters in very low concentration and minerals could not be considered. In addition not all obtained NMR signals could be assigned to specific molecules today and interferences could exist as observed for other analytical methods, too. Nevertheless NMR-signals correspond generally to an already characterised or still unknown substance in the sample content. Because identical substances in different samples give identical signals even molecules with unknown structure could be taken into account in any fingerprint comparison or quantitative statistics. Using chemometrical principles the SGF-Profiling compares spectra of known reference samples with the samples to be evaluated.

Because a certain number of signals corresponds to identified substances the chemometrical evaluation can be completed by a specific qualitative conclusion. In addition quantification by calibration is a possible extension of method application.

An example of sample differentiation is shown in fig. 2 for illustration. Multiple samples of several fruit juices, rediluted fruit juice concentrates, fruit purees and rediluted fruit puree concentrates were measured by SGF-Profiling. The applied PCA model defines a specific loca-

Fig 1: 1H-NMR-spectrum of an orange juice concentrate. Some signals are exemplarily assigned to different substances.
The multi fruit model achieves a good differentiation between all fruit types to be separated. It is not surprising that fruit types of less deviating chemical composition are less scattered in multi-dimensional space, e.g. mandarin and orange are closer than e.g. pineapple and orange (see fig. 2).

**FRUIT TYPE DIFFERENTIATION BETWEEN CITRUS VARIETIES**

The SGF-Profiling allows to apply successively different statistical models to one single measurement per sample. It is not necessary to carry out any further NMR experiment to check further questions. The diagram in fig. 3 is based on another PCA model which was applied to spectra from the same database as used for the multi fruit differentiation in fig. 2. In that case a dedicated model was used to differentiate citrus reticulata varieties (mandarin, clementine,...) from citrus sinensis varieties (oranges). The multi-fruit model (fig. 2) shows overlap between both varieties, the dedicated model (fig. 3) shows a better separation. Respective NMR spectral regions which characterise the difference between both groups are taken into account in the specific bucketing.

The double measurement of one orange juice sample from unknown origin situated between the two clusters for citrus reticulata and citrus sinensis varieties were counterchecked by conventional analyses and the presence of citrus reticulata varieties was confirmed (Carotenoid-profile (IFU 59) (ref. 5) with 25 % cryptoxanthinester-fraction).

European legislation does not allow the production of orange juice containing juice from citrus reticulata varieties. On the other hand such illegal juice compositions can give an economic advantage on the market. Lower raw material prices can be obtained and sensorial properties (colour, total acidity, taste) are influenced and can be optimised towards a better customer response compared to legally produced orange juices. To guarantee fair competition extensive analytical screening including a high number of samples is necessary. Conventional analyses and new approaches by DNA analyses are able to confirm the significant presence of citrus reticulata varieties in orange juice, but the
QUALITY CONTROL

15

by SGF-Profiling is a valuable tool for the analyst. This high throughput pre-screening allows definition at a specific target for the conventional analysis of each sample. It is possible and recommended for the SGF-Profiling to develop - beside the general overall classification of samples - different statistical models for each specific quality aspect to be checked.

DIFFERENTIATION OF PRODUCT CATEGORIES

In case of fruit juices an additional control aspect is the separation between direct fruit juice and fruit juice made from fruit juice concentrate. Both juice types are clearly defined by the European Fruit Juice Directive and have to be declared on product labels. The conventional approach to check the quality is a stable isotopic analysis of the oxygen (ref. 1). Fig. 4 shows that the SGF-Profiling is able to differentiate both product types for orange products.

LINEARITY IN CHEMOMETRICAL RESPONSE

The principle of sample evaluation in examples of fig. 2, 3 and 4 is to check if the sample to be tested is inside or outside of the expected region of the dedicated model. The test model shown in fig. 5 demonstrates that deviations in any result plot can be interpreted in a semiquantitative way. Three different mandarin juices were added to two orange juice samples in amounts of 10, 20 and 30 % each. The observed deviation from the initial measurement of authentic orange juice shows a linear relationship to the amount of mandarin juice added. Beside the distance of deviation, the orientation of any deviating result in the plot contains additional information about differences in the chemical composition, i.e. in the amino acid composition. In fig. 5 we can identify different
deviating curves for respective mandarin juices added to the orange juices. The exactness and significance of such considerations depends on the robustness of the method.

CROSSLINK BETWEEN CHEMOMETRICAL ANALYSES AND STRUCTURE SPECIFIC INFORMATION

In the examples discussed so far, the ability of NMR to identify discrete signals and to link them to specific chemical compounds was not used. It is clear that the knowledge of molecules which cause differentiations and deviations in the chemometric model is also important for the quality controller using the SGF-Profiling in combination with compositional analyses. On the other hand in many cases the optimisation of statistical models in the context of SGF-Profiling is much easier if the conventional analysis can indicate chemical structures and concentration to be examined. For example possible sugar addition to orange juice can be screened by using the ratio between respective integrals of signals from amino acids and carbohydrates. Experiences from classical analysis allow selecting samples from particular origin with generally low amino acid concentration in order to avoid false positive interpretation and to refine statistical models.

One example of using structural identification by NMR is presented in fig 6. The measured orange juice sample from a suspicious source deviated from the expected PCA region. The examination of corresponding spectra shows the absence of malic acid in the deviating sample. As known very well from the conventional analyses absence indicates a serious quality deviation.

Structure-specific information can be combined with quantification in order to come up with absolute or relative metabolite concentrations. In fig. 7, relative quantification results for sucrose and glucose are given for illustration. Interestingly, NMR can not only quantify the metabolites but can even quantify different isomers and give their ratios, as the ratio between α-glucose and β-glucose, which is not easily accessible by alternative methods. However, this ratio is directly measurable from the very same NMR data used for all the other evaluations discussed here.

CHARACTERISATION OF COMPOSITIONAL DIFFERENCES FOR TWO GROUPS OF SIMILAR PRODUCTS

Fig. 8 shows an image plot of the most differentiating buckets to distinguish two groups of samples. As reference data pool, 30 authentic Chinese apple juice concentrates and 30 authentic Polish apple juice concentrates were re-diluted to juice and measured by SGF-Profiling.

Both populations are clearly separated and the involved NMR-frequencies are attributed to typical chemical structures. For the graphic representation samples and differentiating signals are classified and listed in a chosen order. Typical frequencies for sugar are listed in the upper part, typical frequencies for phenolics are listed in the lower part. Red and yellow highlighted fields represent signals which contribute to differentiation by high signal intensity.

The upper left part of the matrix indicates that sugar associated signals are relatively high for Chinese products. The lower right part of the matrix is indicative at relatively higher concentrations of phenolic compounds. The result of this observation is consistent with observations made by classical analytical methods. Differences could be caused by technology and/or varieties.

On the right side of the matrix a commercial apple juice made from apple juice concentrate with unknown origin of processed fruits was measured and displayed. Comparing the respective pattern with pattern from reference samples it must be assumed that this juice was probably made neither by using a concentrate of pure Chinese nor of pure Polish origin. A blend of concentrate of both origins could be possible.
CONCLUSION

In this paper, an established and global approach to analyse mixtures of biological origin has been tailored to the specific case of juice screening. It should be noted that this approach has been successfully used in metabonomics NMR on bodyfluids, tissue and cell extracts for more than a decade, e.g. in toxicity screening, drug efficacy screening, medical or disease screening for example in inborn errors or risk evaluation of coronary heart disease, population health and lifestyle screening, to give just a few examples. Similarly, applications of NMR in combination with pattern recognition to food and feed as well as plant extracts have been documented in a multitude of publications over many years. Also, chemometrical modelling is well established in other fields of spectroscopic data analysis, e.g. infrared spectroscopy, where PCA and PLS are used routinely and are well established in the quality control laboratory. Although only a few chemometrical methods are applied in the context of the illustrative examples given in this paper, it is clear, that no fundamental new aspect would have been added to the statements and results given in this paper if any other data analysis approach, e.g. from the ones listed in Tab.1, would have been used. For example, the clustering shown in fig. 2 can be similarly found when using a cluster analysis or a Multidimensional scaling approach. Classification of direct orange juice versus juices from concentrates through e.g. Discriminant analysis (DA), PLS-DA, Neural Networks or k-NN algorithms would have been similarly valid approaches than the one demonstrated in this work. Any kind of regression approach (e.g. PLS, PCR, RR) could have been applied for quantification, obtaining comparable results as simple regional integration of the spectra.
To summarise, the SGF-Profiling is able to perform a high throughput screening of fruit juices to make a pre-selection for a targeted use of conventional control analyses on suspicious samples. It can be used for general quality control tasks if no specific suspicion of quality deviation for any product exists, saving analytical effort and costs. If new quality aspects arise, the inclusion of additional screening aspects in the SGF-Profiling is simplified because also spectra measured in the past can be used to establish a new statistical model with new specific objectives. Furthermore, based on its properties and performances the SGF-Profiling is able to be used for process control and raw material characterisation. Its introduction in the fruit juice industry - as well as in other food branches - as a standard inward and process control procedure is foreseeable. SGF International e.V. is going to include the SGF-Profiling in its routine control scope on a trial basis whereas they continue to use in parallel classical methods.

REFERENCES

(2) S. Koswig, FRUIT PROCESSING, 2006, 6, 401-407, Determination of foreign fruit types and fruit varieties - Analyses, Evaluation and Practical Problems
(4) T. Hastie, R. Tibshirani, J. Friedman; The Elements of Statistical Learning - Data Mining, Inference and Prediction; Springer 2001
(5) The collected Analyses of the International Federation of Fruit Juice Producers, Loose-Leaf-edition, Zug, Swiss Fruit Association

AUTHORS

Dr Peter Rinke
Sophie Moitrier
SGF International e.V. – Sure-Global-Fair,
D-55268 Nieder-Olm/Germany
www.sgf.org
Eberhard Humpfer
Silke Keller
Monika Moertter
Dr Markus Godejohann
Gudrun Hoffmann
Dr Hartmut Schaefer
Dr Manfred Spraul
Bruker BioSpin GmbH,
D-76189 Karlsruhe/Germany
www.bruker-biospin.de

READERS SERVICE

The Classified Index for all 2006 issues of the international industry magazine FRUIT PROCESSING is available free of charge to download at www.fruit-processing.com.
Magazine copies are on sale at EUR 14.- each (please add 19 % VAT and postage)
The **Enzyme Supplement** to FRUIT PROCESSING issue 4/2006 covers the Panel Discussion during the 46th International Fruit Juice Week 2006 on the Advantages and Risks of Enzymes used in Juice Production.

It is made up of three parts
- introductory statements from several participants
- an account of the Panel Discussion held on April 6, 2006
- a listing of products and services marketed by the companies taking part in the Panel Discussion.

With Professor Rupert Binnig from Trier University of Applied Science in the chair, the Panel Discussion brought together high-ranking experts from industry and research
- Prof Helmut Dietrich, Geisenheim Research Institute, Germany
- Jürgen Helbig, Erbsloh Geisenheim AG, Geisenheim, Germany
- Dr Hans-Jürgen Hofsommer, GfL Berlin, Germany
- Rainer Junker, E.Begerow GmbH, Langenlonsheim, Germany
- Harm Veerkamp, DSM Food Specialties, Delft, the Netherlands

This frank, interdisciplinary debate among committed experts opened up new perspectives. The statements from the panellists and the comments from the auditorium highlighted details that bridge the gap between theory and practice.

**Order your copy of the Enzyme Supplement from Readers Service at info@fruit-processing.com**

As long as available, subscribers are welcome to free-of-charge copies of the Enzyme Supplement. Additional copies of the Enzyme Supplement are available to non-subscribers at euro 10.- each. Please always allow for postage/shipping costs (VAT is included).