ELSEVIER

Food Chemistry 70 (2000) 113–121

Food Chemistry

www.elsevier.com/locate/foodchem

Analytical, Nutritional and Clinical Methods Section

# Multi-way chemometrics for mathematical separation of fluorescent colorants and colour precursors from spectrofluorimetry of beet sugar and beet sugar thick juice as validated by HPLC analysis

Dorrit Baunsgaard *, Claus A. Andersson, Allan Arndal, Lars Munck

*Chemometrics Group, Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, 1958 Frederiksberg C, Denmark*

Received 10 September 1999; received in revised form 3 December 1999; accepted 6 December 1999

**Abstract**

In previous analyses of colour impurities in processed sugar, a multi-way chemometric model, CANDECOMP-PARAFAC (CP), has been used to model fluorescence excitation-emission landscapes of sugar samples. Four fluorescent components were found, two of them tyrosine and tryptophan, correlating to important quality and process parameters. In this paper HPLC analyses are used to chemically verify and extend the CP models of sugar. Thick juice, an intermediate in the sugar production, was analysed by size exclusion HPLC. Tyrosine and tryptophan were confirmed as constituents in thick juice. Colorants were found to be high molecular weight compounds. Fluorescence landscapes on collected column fractions were modelled by the CP model and seven fluorophores were resolved. Apart from tyrosine and tryptophan, four of the fluorophores were identified as high molecular weight compounds, three of them possible Maillard reaction polymers, whereas the seventh component resembled a polyphenolic compound. It is concluded that the relevance of CP for mathematical separation of fluorescence landscapes has been justified on two levels by HPLC; firstly as a screening method of fluorophores in complex samples and secondly as a confirmation of peak purity in chromatographic separation. 

## 1. Introduction

White sugar produced industrially from sugar beet contains minute traces of unwanted colorants. Extensive research into the origin and development of the sugar colorants has been carried out for many years (Godshall, 1996). The earliest works date more than 130 years back (Scheibler, 1869). The fact that significant components have not yet been identified reflects the extreme complexity of the sugar streams as they occur in the sugar factory. Approaches that use isolated laboratory experiments tend to diverge from the natural seasonal variations of the streams, whereby the findings become too specialised to have any practical value in the real process streams at the factories. We have chosen a new approach to reach conclusions that adapt to the natural (co)variations of the constituents in the sugar streams. With the use of exploratory data analysis, functional components in the process streams are found by soft adaptive modelling instead of using hard chemical analysis to identify actual chemical substances (Munck, Nørgaard, Engelsen, Bro & Andersson, 1998). Advanced multi-way models, such as the CANDECOMP-PARAFAC (CP) model, can be used to decompose complex excitation-emission fluorescence landscapes into excitation and emission spectral profiles of characteristic components (Leurgans & Ross, 1992). Bro (1999) used the CP model on fluorescence landscapes from 268 sugar samples collected from a factory during a sugar campaign. A model with four fluorescent components was found to capture the variation in that time period. Two of them had pseudo-spectra, which showed a close similarity to pure fluorescence spectra of tyrosine and tryptophan. In addition, the concentrations of the four components estimated from the sugar samples could be correlated to several quality and process parameters. Thus, the four fluorescent components found in the final

* Corresponding author. Tel.: +45-3528-3500; fax: +45-3528-3505.

*E-mail address:* dba@kvl.dk (D. Baunsgaard).

0308-8146/00/$ - see front matter 
PII: S0308-8146(00)00046-7

sugar product are considered as indicator substances of the chemistry in the sugar process.

In the sugar process streams, there are several potential fluorophores. These include colour precursors such as amino acids and polyphenolic compounds (Wolfbeis, 1985). Colour precursors can interact in colour forming reactions such as amino acids with reducing sugars in Maillard reactions or enzymatic oxidation of phenolic compounds to form melanins (Godshall, Clarke, Dooley & Blanco, 1991). Coloured Maillard reaction products have been reported to exhibit fluorescence (Adhikari & Tappel, 1973). One of the preferred methods for analysing colorants and colour precursors has been gel permeation chromatography (GPC) since many of the colorants are considered as high molecular weight compounds (Madsen, Kofod Nielsen, Winstrøm-Olsen & Nielsen, 1978; Reinefeld, Schneider, Westphal, Tesch & Knackstedt, 1973; Shore, Broughton, Dutton & Sissons, 1984).

In this paper, we combine CP modelling of fluorescence excitation-emission landscapes with HPLC size exclusion analysis. After separating the sample on the column, collected fractions are measured as fluorescence landscapes and modelled with the CP model. Thick juice, an intermediate product from the sugar manufacturing process, is analysed instead of sugar since the latter is too pure and not suitable for chromatographic analysis. The purpose of the chromatography is twofold. It can be used to verify the identity of the mathematically modelled fluorophores in sugar with peak identification. Also, the number of identifiable components may be improved by the pre-separation of the components on the column before the fluorescence measurements. The pre-separation is used to reduce quenching and other interactions in the complex sample, which influences the fluorescence, and may violate the assumptions made prior to application of the CP model.

## 2. Materials and methods

### 2.1. Chemicals

L-tyrosine, L-tryptophan and L-phenylalanine were purchased from Sigma (USA). The reagents for the HPLC buffer were obtained from Merck KGaA (Germany). Water was distilled and deionized (Milli-Q, Waters, USA). HPLC eluents were filtered and degassed before use.

### 2.2. Samples

Beet sugar samples and beet sugar thick juice samples were all provided by Danisco Sugar A/S, Denmark. Ten thick juice samples from five different sugar factories, two from each, were dissolved in water 1:500 (v/v) and used to measure fluorescence landscapes. Five sugar samples collected from one of the sugar factories were prepared by dissolving 7 g sample in 15 ml water for the fluorescence measurements. For the HPLC analyses thick juice samples from one of the five factories was prepared by diluting 100 μl thick juice with 100 μl 0.2 M ammonium buffer, pH = 8.9 and 300 μl water. Due to the high viscosity of the thick juice sample, a pipette designed for viscous samples (Microman 250, Gilson, USA) was used to take samples of the thick juice.

### 2.3. HPLC analyses

The HPLC size exclusion analyses were performed on a Gilson system with a Gilson 170 UV–VIS diode array detector (range: 210–550 nm) and a Jasco FP-920 fluorescence detector (excitation/emission wavelengths: 280/325 nm). A Waters 250 Ultrahydrogel column (range 1–80 kDa) was used equipped with a guard column of the same material and thermostatted at 30°C. The mobile phase consisted of 0.2 M ammonium buffer ($NH_4Cl/NH_3$), pH = 8.9 and water (20:80 v/v) at a flow rate of 0.5 ml/min. All sample solutions were filtered through a 0.22 μm hydrophilic PVDF membrane filter (Millipore, USA) before injecting an aliquot of 100 μl onto the column. In this publication, two representative HPLC runs of thick juice were selected for fluorescence landscape measurements of 41 collected fractions of 750 μl (1.5 min) from 10 to 71.5 min in each run.

### 2.4. Amino acid standards

Tyrosine and tryptophan were identified by peak identification of spiked thick juice samples with amino acid standards. The spiked thick juice samples were prepared by mixing 50 μl thick juice and 100 μl 0.2 M ammonium buffer, pH = 8.9 with 350 μl tyrosine solution (192 mg/l) or 350 μl tryptophan solution (43 mg/l). The two amino acid solutions and a phenylalanine solution (1.4 g/l) were used to establish the size exclusion range of the column in a 1:2 (v/v) dilution with the ammonium buffer.

The fluorescence spectra of the tyrosine and tryptophan standards were measured with the same parameters as with the other samples using a tyrosine concentration of 1.6 mg/l and a tryptophan concentration of 0.3 mg/l.

### 2.5. Fluorescence landscape measurements

A Perkin–Elmer LS50 B fluorescence spectrometer was used to measure fluorescence landscapes using excitation wavelengths between 230–300 nm with 5 nm intervals and 310–460 nm with 10 nm intervals. The emission wavelength range was 288–700 nm. Excitation

and emission monochromator slit widths were set to 10 nm, respectively. Scan speed was 1500 nm/min. A micro quartz cuvette with the dimensions 5×5 mm was used to avoid dilution and to reduce any concentration quenching effects of the sample solution.

### 2.6. The CANDECOMP-PARAFAC model

The CANDECOMP-PARAFAC (CP) model was proposed in 1970 (Carroll & Chang, 1970; Harshman, 1970) and fits the premises of fluorescence spectroscopy for resolving pure excitation and emission spectra from measured net signals of mixtures. To allow for a discussion of the CP model, we consider a fluorescence data set with elements denoted by $x_{ijk}$, where $x_{ijk}$ is the intensity of the $i$th sample excited by light at the $j$th excitation wavelength and measured at the $k$th emission wavelength. The resulting data set thus spans a three-dimensional table structure, where each entry represents an observation that depends on discrete levels of the three parameters, (sample number×excitation wavelength×emission wavelength). The three-way data array can be approximated by

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk} \quad (1)$$

In (1) it is assumed that the measured net signal is a sum of $F$ individual contributors, or fluorophores. For fluorophore number $f$, $a_{if}$ is the concentration in the $i$th sample, $b_{jf}$ is the relative amount of light absorbed at excitation wavelength $j$, and $c_{kf}$ is the relative intensity emitted at wavelength $k$. This tri-linear structure of the light intensity model is similar to the tri-linear CP model for which solution algorithms have been devised (see Carroll & Chang, 1970; Harshman, 1970). Under the assumption of tri-linearity in the signal/concentration ratio and additivity of the intensities, the CP model parameters will be estimates of the underlying excitation spectra, i.e. the $b_{jf}$ parameters, and the emission spectra, i.e. the $c_{kf}$ parameters of each of the $f$ contributing fluorophores. However, based on the observations or a priori knowledge, the task of defining the correct number of fluorophores, $f$, remains. The mathematical uniqueness of the CP-model will provide parameters in A, B and C of the individual fluorophores contributing to the net signal. Not only will the parameter estimates be unique to the individual fluorophores, but since the fundamental mechanistic model of the net signal of a single fluorophore is in exact accordance with the CP–model for $F=1$, the resolved parameters will be relative estimates of concentration level, excitation ability (absorbance spectrum) and emission ability (emitted spectrum).

Furthermore, the CP model allows for simultaneous presence of many such single contributors to the overall observed emitted intensity, $x_{ijk}$. Thus, by estimating the CP parameters, the collection of net signal can be separated mathematically into a number of characteristic profiles for each of the fluorophores/contributors. See Leurgans and Ross (1992) for an in-depth discussion of multi-linear models in spectroscopic contexts. For a more thorough presentation of the model, the reader is referred to a tutorial on the CP model (Bro, 1997).

The CP results have been obtained with the use of the $N$-way Toolbox for MATLAB (Andersson & Bro, 1998) running MATLAB 5.3 under Microsoft Windows NT 4 SP5 on a dual 450 MHz Intel PII Xeon PC. For the tri-linear CP model to be valid, infeasible measurements (i.e. Rayleigh scatter and emission wavelengths less than excitation wavelengths) have to be treated as missing values. To circumvent the scaling ambiguities of the CP model and to enhance the interpretability of the model, the profiles were constrained to non-negativity while minimising the sum of squared errors, i.e. the constrained model parameters were estimated from a total least squares optimisation of Eq. (1).

## 3. Results and discussion

### 3.1. Fluorophores in sugar and thick juice

Since thick juice is used instead of sugar in the HPLC separations, it is important to know the differences and similarities between fluorophores found in sugar and in thick juice. In addition, the changes in the properties of an intermediate sugar product to the properties of the final product can be useful, e.g. in process control.

The CP analysis on fluorescence landscapes of sugar samples previously made by Bro (1999) was repeated by making a CP model using five sugar samples from another sugar factory. In addition, fluorescence landscapes were measured on ten thick juice samples from five different sugar factories, two samples from each factory, and modelled with the CP model. A four-component model was generated from the sugar data and a five-component model from the thick juice data. The CP modelling estimates excitation and emission spectra of measured fluorophores as well as a sample profile relating the concentration of each fluorophore in the samples measured. Fig. 1 (rows 1–4) and Fig. 2 present the excitation and emission spectra of the modelled components in the sugar and thick juice samples, respectively. The components are displayed in the same order as they are modelled depending on their contribution in the sample profile. The resolved spectra show reasonable spectral shapes, but they are dependent on the appearance of the measured fluorescence data and the premises of the model. Therefore some of the spectra may display artefacts such

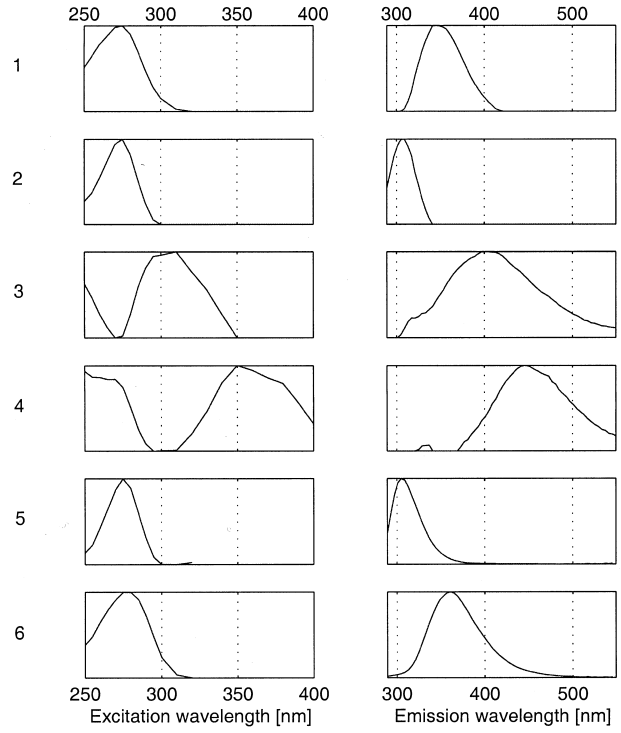


Fig. 1. The results of a four-component CP model of fluorescence landscapes of five beet sugar samples. Rows 1–4 contain the excitation and emission spectra of the four resolved components. The left column shows the excitation profiles and the right column shows the emission profiles. Rows 5 and 6 show the fluorescence excitation and emission spectra of pure tyrosine and pure tryptophan, respectively, for a comparison. All profiles have been normalised to unit length.

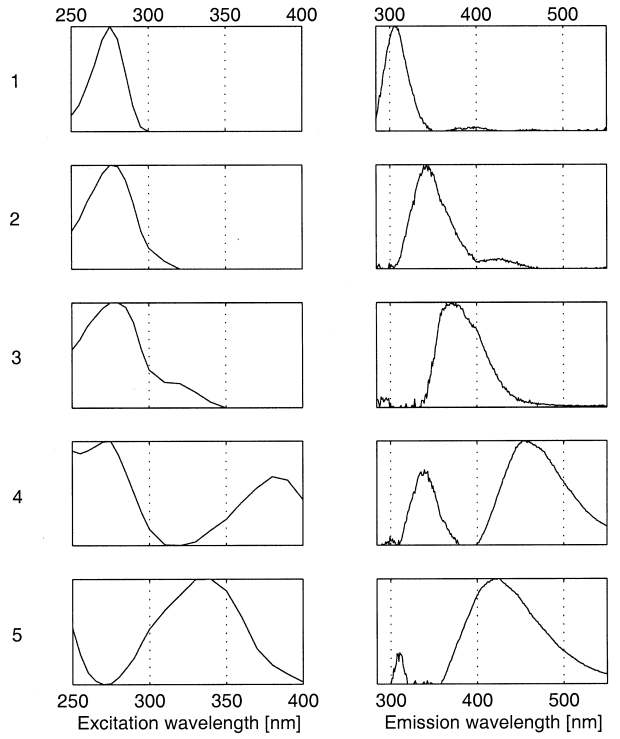


Fig. 2. A five-component CP model of fluorescence landscapes of 10 thick juice samples. The left column shows the resolved excitation profiles and the right column shows the resolved emission profiles of the five components. All profiles have been normalised to unit length.

as extra bands in the emission spectra, e.g. the emission spectrum of component 4 in Fig. 2. The excitation (1st excited state) and emission wavelength maxima of the spectra in Figs. 1 and 2 are presented in Table 1. The shape and maxima of the emission spectra of the sugar model in Fig. 1 are comparable to the previously modelled spectra of the four-component sugar model by Bro (1999). In Fig. 1 the excitation and emission spectra of pure tyrosine and tryptophan standards are displayed in rows 5 and 6, respectively. Comparing the spectra of the two amino acids with the spectra of the modelled components in Fig. 1, there is a close similarity between tyrosine and component 2 and between tryptophan and component 1. The spectral profiles of thick juice fluorophores in Fig. 2 are consistent with the spectra of the sugar components in Fig. 1, although there are some differences in the profiles. This is also evident by comparing the excitation and emission maxima in Table 1. The tyrosine-like fluorophore is component 1 in Fig. 2. Component 2 in Fig. 2 resembles the tryptophan-like component in Fig. 1, but the emission profile is shifted towards lower wavelengths and a fifth component (component 3) is introduced in the thick juice model. The spectral properties of the new component are close to tryptophan. Thick juice contains much more impurities than sugar and the fluorescence data is more difficult to model. If component 2 is tryptophan, component 3 might be another fluorophore or a tryptophan-derived component with somewhat changed fluorescent properties, either of which affecting the estimated tryptophan profile. Using a larger sample set, it will be possible better to solve such ambiguities.

In the modelling of sugar and thick juice fluorescence data, samples from several factories have been used. It is found that sugar models from different factories contain the same four fluorophores, e.g. the similarity of the modelled spectra of the five sugar samples in Fig. 1 with the previously modelled spectra from another factory (Bro, 1999). Furthermore, the thick juice model in Fig. 2 was based on samples from five different sugar factories and HPLC analyses made on the thick juice samples from the five factories all showed the same qualitative chromatographic pattern. Therefore, the modelled fluorophores from the sugar and thick juice fluorescence data are considered to be common constituents of sugar and thick juice and not factory related.

### 3.2. Peak identification using HPLC analyses

It is important to validate the results of the CP modelling of fluorescence landscapes of sugar and thick juice. When comparing the resolved pseudo-spectra

Table 1
The excitation and emission maxima of the modeled spectra of sugar, thick juice and HPLC fractions of thick juice

| Component[a] | Sugar$_{\lambda max}$ (nm) | | Thick juice$_{\lambda max}$ (nm) | | HPLC fractions of thick juice$_{\lambda max}$ (nm) | |
|---|---|---|---|---|---|---|
| | Excitation | Emission | Excitation | Emission | Excitation | Emission |
| 1 | 275 | 350 | 275 | 305 | 275 | 305 |
| 2 | 275 | 305 | 275 | 340 | 275 | 360 |
| 3 | 310 | 400 | 280 | 370 | 375 | 460 |
| 4 | 350 | 450 | 380 | 455 | 340 | 440 |
| 5 | – | – | 335 | 420 | 385 | 460 |
| 6 | – | – | – | – | 290 | 400 |
| 7 | – | – | – | – | 290 | 330 |

[a] The component numbers correspond to the row numbers given in Figs. 1, 2 and 6 for each of the three CP models.

with pure spectra of tyrosine and tryptophan, a level of uncertainty is involved due to the limited number of samples as well as quenching and non-linearities in the measured fluorescence data. Using chromatographic peak identification, it is possible to verify that the two amino acids really are constituents of thick juice. Drewnowska, Walerianczyk, Butwilowicz, Jarzebinski, Fitak and Gajewska (1979) have previously estimated the contents of tyrosine and tryptophan in thick juice with the use of liquid chromatography. Fig. 3 shows a HPLC size exclusion separation of one of the thick juice samples before and after spiking the sample with the two amino acids and monitored by fluorescence detector set at 280/325 nm. The three chromatograms show good overlap and the spiked peaks confirm that the two dominating components eluting at 25 and 42 min are the free amino acids, tyrosine and tryptophan, respectively. In addition, the identities of the peaks were confirmed by comparison with chromatograms of amino acid standards. The tyrosine and tryptophan peaks at 25 and 42 min were also found in the chromatograms of the thick juice samples from the four other factories used in the thick juice model. The corresponding diode array scans of the two spiked peaks in Fig. 3 are also displayed in the figure. The spectra are practically identical with pure spectra of tyrosine and tryptophan, which is an additional certainty of the identification of the peaks.

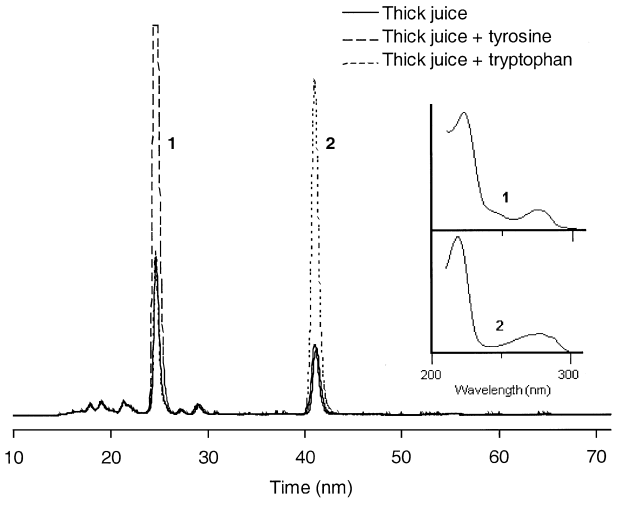


Fig. 3. HPLC size exclusion chromatograms with qualitative standard additions of tyrosine (1) and tryptophan (2) to thick juice monitored by fluorescence detection at 285/325 nm. The chromatograms verify the expected presence of tyrosine and tryptophan. The corresponding UV/VIS absorbance spectra from the diode array scans of the two peaks are also displayed.

It is difficult to analyse the very pure sugar on a HPLC system. However, the similarity of the spectral profiles in the sugar model with the thick juice model and the spectra of the amino acid standards confirm indirectly the identification of the corresponding fluorophores.

### 3.3. HPLC size exclusion analyses of thick juice

When using a size exclusion column (range 1–80 kDa), it is possible to separate the thick juice samples according to molecular weight. This can be used to separate the colorants as high molecular weight compounds from low molecular weight colour precursors. The column dead time was determined to 12.4 min using Blue Dextran 2000. The amino acid standards tyrosine and phenylalanine were used to establish the end of the size exclusion area of the column to 25 min. The fact that tryptophan elutes at 42 min is probably caused by *adsorptive* retention on the column. In Fig. 4 three simultaneously recorded chromatograms of a thick juice sample are shown. The two upper chromatograms are captured from the diode array detector at 280 and 420 nm, whereas the lower chromatogram is from the fluorescence detector at 280/325 nm. 420 nm is the normal wavelength chosen by the sugar industry to represent colour. Many of the known components absorb at 280 nm (amino acids, polyphenols, Maillard reaction products, etc.), which is consistent with the multiple peaks in the chromatogram. The 420 nm chromatogram, on the other hand, shows a limited number of small peaks in the beginning of the run between 15 and 25 min. The estimated size exclusion range was approx. 12–25 min, which means that the colorants are smaller than 80 kDa

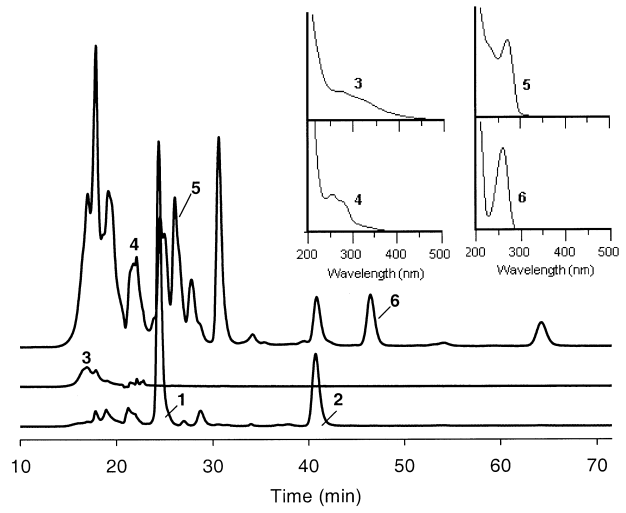


Fig. 4. HPLC size exclusion chromatograms of a thick juice sample. Upper curve: UV detection at 280 nm; middle curve: UV detection at 420 nm; bottom curve: fluorescence detection at 280/325 nm. Diode-array spectra of selected peaks (3–6) are also displayed. The spectra of peaks 1 and 2 are identical with the spectra of the corresponding tyrosine and tryptophan peaks in Fig. 3.

but extend the range down to 1 kDa. Colorants have been estimated to 5 kDa in white beet sugars, though for some sugars up to 40 kDa, and in molasses colorants up to 50 kDa have been found (Godshall et al., 1991). Their findings agree well with the range of the colorants in thick juice separated on the HPLC column. Apart from the two very dominating peaks at 25 and 42 min identified as tyrosine and tryptophan, the fluorescence chromatogram in Fig. 4 also shows a number of smaller peaks in the colorants area in the first 25 min.

A few selected diode array scans with very different spectral appearance are also displayed in Fig. 4. Apart from confirming the findings in the chromatograms, the diode array absorbance spectra can provide more detailed information for the identification of the components in thick juice. The absorbance spectra of peaks 1 and 2 are identical to the displayed spectra of corresponding peaks in Fig. 3. The spectra of peak 3 reveals that the highest molecular weight colorants absorb light up to 500 nm, which is consistent with the golden-orange appearance of the thick juice sample. Peaks 4–6 display different spectra mainly absorbing at 280–300 nm where peaks 4 and 5 appear to be composed of multiple components. The displayed spectra demonstrate the complexity of the thick juice sample and the fact that the separation of thick juice on the HPLC size exclusion column is insufficient to separate the colorants in thick juice.

### *3.4. A CP model of fluorescence landscapes of HPLC fractions of thick juice*

To improve the CP model of thick juice as well as the HPLC separation, 41 fractions of 1.5 min (10–71.5 min) were collected during the HPLC separation shown in Fig. 4. A fluorescence landscape was recorded of each fraction by off-line measurements in a scanning spectrofluorometer. In Fig. 5 the fluorescence landscape of fraction 9 serves as an example of such a landscape. There are clearly multiple overlapping fluorescent peaks in the landscape and a resolution method is required. The 41 landscapes form a three-dimensional data array consisting of the 41 fractions in the first dimension, 31 excitation wavelengths (230–460 nm) in the second dimension, and 431 emission wavelengths (288–700 nm) in the third dimension. The array was modelled by the three-way CP model and seven components were found. The modelling results are shown in Fig. 6. Each component is represented by the estimated excitation and emission spectra as well as a chromatographic profile, which shows the concentration of each component in the 41 collected fractions. The excitation and emission maxima of the seven components are presented in Table 1. The spectral shapes in Fig. 6 are all reasonable. Again extra bands appear in some of the emission spectra. The fluorescence landscape of fraction 9 in Fig. 5 demonstrate that a large part of the landscape has to be treated as missing values due to first and second order Rayleigh scattering (Bro, 1999). In the estimations of these areas, extra bands may appear depending on the condition of the fluorescence data. Components 1 and 2 in Fig. 6 are recognised as the two modelled components tyrosine and tryptophan, which are also found in the sugar and thick juice models (Figs. 1 and 2). In the corresponding chromatographic profiles in Fig. 6 the two components show two dominant peaks in fraction 10 (23.5–25 min) and fraction 21 (40–41.5 min), respectively, which are consistent with the position of the spiked peaks in the fluorescence chromatogram in Fig. 3. The chromatographic profile of tryptophan in row 2 in Fig. 6 also shows contributions in fractions 3–12 similar to the small peaks displayed in the fluorescence chromatogram in Fig. 4. Tyrosine, on the other hand, is only found in fractions 9–11. Tryptophan has very distinct fluorescent properties, which are kept intact even as a functional group in a larger molecule, whereas tyrosine loses the fluorescent properties very easily. For example, in proteins the fluorescence is dominated by the tryptophan residue (Lakowicz, 1983). The fact that the tryptophan fluorophore is modelled in the higher molecular weight fractions in the chromatographic profile could be due to tryptophan residues behaving as individual fluorophores in polymers. This demonstrates that the chromatographic profile from a CP model can be used as a mathematical purification of the overlapping peaks in a chromatogram, provided that the assumptions of linearity and additivity of the model hold. The area of the chromatogram from 15–25 min with many overlapping peaks in Fig. 4 is simplified by the CP model and more information can be obtained.

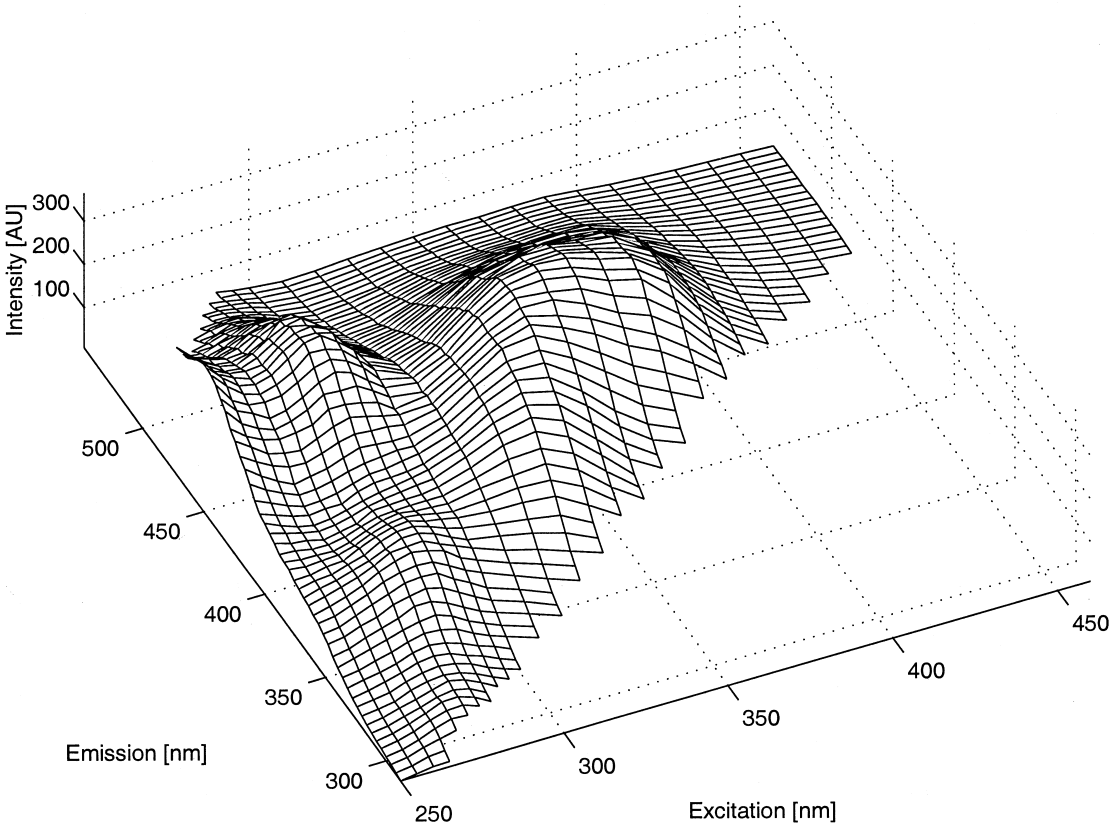


Fig. 5. An example of a fluorescence landscape from HPLC fraction no. 9 (22–23.5 min) of thick juice measured with 31 excitation wavelengths (230–460 nm) and 431 emission wavelengths (288–700 nm). The blank regions hold Rayleigh scatter signals and are thus treated as missing values.

Components 3–5 in Fig. 6 are very similar in their spectral shapes and position of the emission spectra. They are modelled as three individual components because of the differences in the excitation and chromatographic profiles. It can be argued that they are part of the same group of fluorophores, but are resolved individually due to small differences in molecular sizes and/or small differences in the fluorophore environments. The size exclusion on the HPLC column is not good enough in the high molecular weight area and a more refined fraction collection is necessary to obtain a clearer separation. This is supported by the chromatogram at 280 nm in Fig. 4, where there is only one peak at 20–23 min. The three components (3–5) in Fig. 6 have excitation profiles that reach into the visible area above 400 nm, which implies that they contribute to the colorants in thick juice. They are all found in the first fractions of the chromatographic profiles and are thus high molecular weight compounds. The spectral characteristics of these colorants resemble conjugated Schiff bases derived from malonaldehyde and amino acids as reported by Chio and Tappel (1969). The authors ascribed the absorption and fluorescence properties of the Schiff bases to the chromophoric system $-N=C-C=C-N-$. Pongor, Ulrich, Bencsath, and Cerami (1984) isolated a fluorophore from a product of a browning reaction of polypeptides with glucose, which show similar fluorescence spectra. The structure of the isolated fluorophore contained a conjugated system of nitrogen and carbon in an imidazole derivative. Similar compounds isolated from real samples in the sugar processing have not been reported, but quantitative elementary analysis on high molecular weight fractions from GPC separations of thick juice showed an element ratio of carbon and nitrogen as 7:1, which indicated that amino acids were built into the high molecular weight fractions (Madsen, Kofod Nielsen & Winstrøm-Olsen, 1978). All this suggests that some or all of components 3–5 are colorant polymers formed during the sugar processing in Maillard reactions involving amino acids and reducing sugars.

Component 6 in Fig. 6 is also a high molecular weight compound with contributions in the first fractions in the chromatographic profile. The emission spectrum is in the visible area, but the excitation spectrum is well below 400 nm and the component is therefore not a colorant. The component is similar to component 3 in the sugar model in Fig. 1 and partly comparable to component 5 in the thick juice model in Fig. 2. At present the component is not associated with any known fluorophore.

Component 7 in Fig. 6 is the only component not comparable to any component in the sugar model in Fig. 1. This component may be the reason that the

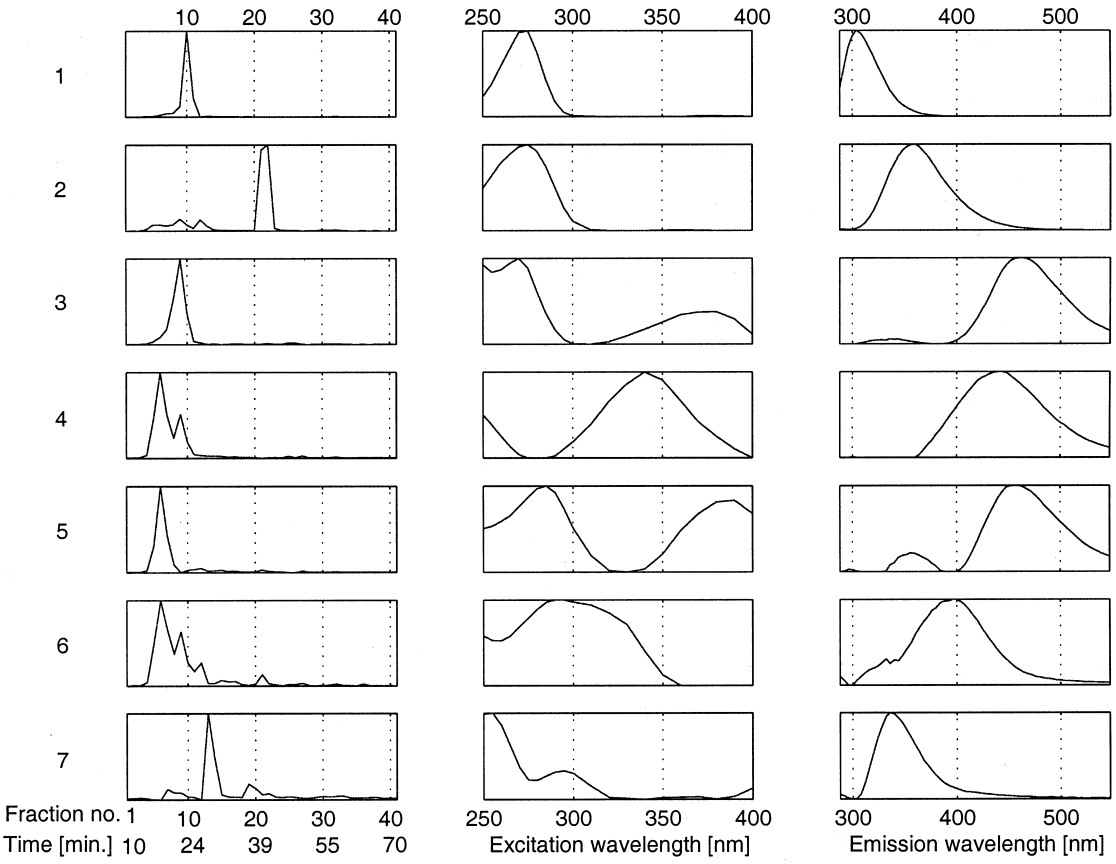


Fig. 6. The results of a seven-component CP model of the measured fluorescence landscapes of 41 collected HPLC fractions (10–71.5 min). The left-most column holds the chromatographic profiles, which show the concentration of each of the seven fluorescent components in the fractions. The centre column holds the excitation profiles and the right-most column holds the emission profiles of the fluorophores. All profiles have been normalised to unit length.

tryptophan component is not estimated as clearly in the thick juice model (Fig. 2) as in the sugar model (Fig. 1), since it has spectral properties close to tryptophan. Its concentration is low in thick juice and therefore the CP model of the 10 thick juice samples in Fig. 2 could not resolve it. Pre-separation on the column made it possible to measure the fluorescence of component 7 without interference like concentration quenching from other fluorophores in thick juice. The component contributes only slightly in the first fractions when looking at the chromatographic profile in Fig. 6, but is spread over several of the later fractions with a dominant peak in fraction 13 (28–29.5 min), which indicates either a low molecular weight compound or a compound with high column affinity. The excitation and emission profiles of this component are suggesting a fluorophore with a polyphenolic group (Duggan, Bowman, Brodie & Udenfriend, 1957).

## 4. Conclusion

It is possible to capture the same fluorescent information from the CP models of fluorescence landscapes of sugar samples and thick juice samples. Four principal components are resolved from the sugar model, where two of them have spectra similar to tyrosine and tryptophan. The tyrosine component is also found in the five-component thick juice model, whereas the estimation of the tryptophan component is less certain due to the more complex sample. However, the presence of the two amino acids in the thick juice model is verified by HPLC peak identification, which also confirms the spectral identification of the model components. The HPLC size exclusion separation of thick juice further confirms that the fluorescent colorants, which are found in the CP analysis of the sugar and thick juice samples, are high molecular weight compounds. Landscape measurements on HPLC collected fractions of thick juice are successfully modelled and seven components are found. The resolved chromatographic profile of the model can be used as a mathematical purification of the not perfectly separated chromatogram. Two of the seven modelled components are identified as the free amino acids, tyrosine and tryptophan, but the latter also appears in higher molecular weight fractions in the chromatographic profile implying intact tryptophan residues in polymers. Four out of the seven modelled components

are identified as high molecular weight components; three of them are suggested to be Maillard reaction polymers of amino acid origin with different molecular weights. The seventh component is of low concentration and has a spectral appearance of a polyphenolic-like compound. It is important to improve further the fluorescence information of the sugar streams by modelling a larger data material to improve the CP model estimations, and it is currently in progress. Future research will also include CP models of fluorescence data from samples taken throughout the sugar process to increase the information of the origin and development of the fluorophores in the sugar streams.

Thus, this paper demonstrates the usefulness of mathematical deconvolution by the CP model of fluorescence data from complex sample matrices as well as for peak purity evaluation in chromatography.

## Acknowledgements

The authors thank Ole Hansen, Lars Bo Jørgensen and John Jensen, Danisco Sugar Development Center, Nakskov, Denmark for inspiring cooperation and in providing the sugar and thick juice samples. We also thank Dr. Rasmus Bro for valuable discussions.

## References

Adhikari, H. R., & Tappel, A. L. (1973). Fluorescent products in a glucose–glycine browning reaction. *Journal of Food Science, 38*, 486–488.

Andersson, C. A. & Bro, R. (1998) The *N*-way toolbox for MATLAB. Can be downloaded from http://models.kvl.dk/source/nwaytoolbox/

Bro, R. (1997). PARAFAC: tutorial and applications. *Chemometrics and Intelligent Laboratory Systems, 38*, 149–171.

Bro, R. (1999). Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemometrics and Intelligent Laboratory Systems, 46*, 133–147.

Carroll, J. D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an *N*-way generalization of "Eckart–Young" decomposition. *Psychometrika, 35*, 283–319.

Chio, K. S., & Tappel, A. L. (1969). Synthesis and characterization of the fluorescent products derived from malonaldehyde and amino acids. *Biochemistry, 8*, 2821–2827.

Drewnowska, W., Walerianczyk, E., Butwilowicz, A., Jarzebinski, J., Fitak, B., & Gajewska, M. (1979). Contents of free amino acids and amides in sugar and sugar industry semiproducts. *Acta Alimentaria Polonica, 5*, 315–321.

Duggan, D. E., Bowman, R. L., Brodie, B. B., & Udenfriend, S. (1957). A spectrophotofluorometric study of compounds of biological interest. *Archives of Biochemistry and Biophysics, 68*, 1–14.

Godshall, M. A. (1996) Recent progress in sugar colorants. In *Proceedings of the 1996 Conference on Sugar Processing Research* (pp. 262–305), New Orleans, USA.

Godshall, M. A., Clarke, M. A., Dooley, C. D., & Blanco, R. S. (1991). Progress in beet sugar colorant research. *Journal of Sugar Beet Research, 28*, 155–165.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics, 16*, 1–84.

Lakowicz, J. R. (1983). *Principles of fluorescence spectroscopy*. New York: Plenum Press p. 347.

Leurgans, S., & Ross, R. T. (1992). Multilinear models: applications in spectroscopy. *Statistical Science, 7*, 289–319.

Madsen, R. F., Kofod Nielsen, W., Winstrøm-Olsen, B., & Nielsen, T. E. (1978). Formation of colour compounds in production of sugar from sugar beet. *Sugar Technology Reviews, 6*, 49–115.

Madsen, R. F., Kofod Nielsen, W., & Winstrøm-Olsen, B. (1978). Juice purification system, sugar house scheme, and sugar quality. *Sugar Journal, 41*, 15–19.

Munck, L., Nørgaard, L., Engelsen, S. B., Bro, R., & Andersson, C. A. (1998). Chemometrics in food science — a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemometrics and Intelligent Laboratory Systems, 44*, 31–60.

Pongor, S., Ulrich, P., Bencsath, F. A., & Cerami, A. (1984). Aging of proteins: isolation and identification of a fluorescent chromophore from the reaction of polypeptides with glucose. *Proceedings of the National Academy of Sciences of the USA, 81*, 2684–2688.

Reinefeld, E., Schneider, F., Westphal, K., Tesch, K., & Knackstedt, H. G. (1973). Model studies on the browning reaction in technical sugar juices. *Zucker, 26*, 581–588.

Scheibler, C. (1869). Über das Betaïn, eine im Safte der Zuckerrüben (Beta vulgaris) vorkommende Pflanzenbase. *Zeitschrift des Vereines für die Rübenzucker-Industrie im Zollverein, 19*, 549–558.

Shore, M., Broughton, N. W., Dutton, J. V., & Sissons, A. (1984). Factors affecting white sugar colour. *Sugar Technology Reviews, 12*, 1–99.

Wolfbeis, O. S. (1985). The fluorescence of organic natural products. In S. G. Schulman, *Molecular luminescence spectroscopy: methods and applications–part I* (pp. 167–370). New York: John Wiley and Sons.