ANALYSIS OF COLOR IMPURITIES IN SUGAR PROCESSING USING FLUORESCENCE SPECTROSCOPY AND CHEMOMETRICS

Dorrit Baunsgaard

Ph.D. Thesis

August, 2000

Supervisors:

Professor, Fil. Dr. Lars Munck
Associate Professor, Ph.D. Lars Nørgaard
The Royal Veterinary and Agricultural University
Department of Dairy and Food Science
Chemometrics Group, Food Technology
Denmark

Preface

This thesis is submitted as a partial fulfillment of the requirements for a Ph.D.-degree at the Royal Veterinary and Agricultural University, Denmark. The research work leading to the results presented in the thesis was carried out at the Food Technology research unit at the Department of Dairy and Food Science under the supervision of Associate Professor Lars Nørgaard and Professor Lars Munck.

I am very grateful for the opportunity given to me by my supervisors to commence a Ph.D.-study in their group. It was their enthusiasm and dedication to the discipline of chemometrics that sparked my interest in the first place and I thank them both. I am also indebted to Rasmus Bro and Claus A. Andersson for their help and inspiration during my project. Without their Multiway Toolbox, much of my research would have been impossible.

During my stay at the Sugar Processing Research Institute, Inc. (SPRI) in New Orleans, I had the pleasure of working with Ms. Mary An Godshall and Dr. John R. Vercellotti, who managed to pass on some of their vast knowledge of sugar processing to me. I also thank Xavier, Bequi, Debbie, Frank and Ronnie for making my first stay in the USA a real pleasure.

My research was mainly based on samples provided by Danisco Sugar A/S and I thank Ole Hansen, Lars Bo Jørgensen and John Jensen for their interest and help.

Finally, I would like to thank all the people at Food Technology for their help and inspiration and for providing a good working environment. I especially wish to thank Jesper P., Per, Søren, Allan, Jesper W., Elisabeth, Henriette, Karen, Kristine, Kirsten, and Peter.

Summary

For many years the sugar industry has done extensive research into the chemical reactions taking place during the production of sugar. The extraction of sucrose from sugar cane or sugar beet is a well-defined process, but impurities formed in side reactions during processing, e.g. colorants, lower the production yield and the sugar quality. The complexity of these impurities makes it difficult to analyze the chemical composition of process samples and model systems have often been used instead. The development of multivariate statistical methods (chemometrics) has made it possible to analyze large amounts of spectral data containing multiple components. It has been demonstrated that these methods can be used to extract information about the chemical variation from fluorescence measurements of sugar beet process samples.

In this thesis five papers are presented that further exploit the use of fluorescence spectroscopy and chemometric data analysis of sugar process samples for a better comprehension of the identity and formation of impurities and their influence on the composition of the sugar crystal. In the first paper separation chromatography is combined with fluorescence measurements and subsequent three-way modelling to verify the identity of the color precursors, tyrosine and tryptophan, in beet sugar and beet sugar thick juice. The paper also demonstrates the use of multi-way data analysis to mathematically separate high molecular weight compounds in the chromatograms.

In the second paper fluorescence in cane sugar is studied using a data set of raw sugars. Three principal fluorophores are found in raw sugar, a color precursor and two colorants that are all related to color reactions in the process.

The origin of fluorescence of process-derived colorants is subsequently studied by comparing model colorants of known color reactions with sugar syrups from the cane and beet sugar processes. The data analyses show that some similarity in fluorescence profiles is found between real and synthesized colorants and the overall conclusion is that the same color reactions occur in beet and cane sugar processing but with varying degree of development of color due to the different physical conditions of the two processes.

The subject of the two remaining papers is fluorescence measurements performed directly on crystal samples. The first paper contains a study of the general applicability of solid-phase fluorescence to the analysis of sugar quality. Two different solid-phase fluorescence methods applied to beet sugar samples are compared with measurements of solutions. Chemometric classification of the solid-phase data is not as successful as with the solution data but prediction of chemical parameters such as color and ash is obtained with almost as good model errors. The fact that the solid-phase fluorescence models are highly influenced by the crystal sample and solid-phase methods is elaborated on in the second paper where sieved fractions of beet sugar samples are used to study the influence of the crystals on the measurements. It is shown that solid-phase transmission fluorescence is closer to the fluorescence that is obtained with solutions whereas reflection fluorescence is less precise due to the crystal transparency. However, the dominating physical effects in the solid-phase spectra prevent the study of colorant distribution in the sugar crystal.

Resumé

I sukkerindustrien har man i mange år arbejdet med at klarlægge de kemiske processer, der foregår under produktionen af melis. Udvinding af sucrose fra enten sukkerrør eller sukkerroer er en veldefineret proces, men de urenheder der dannes under forløbet, f.eks. farvede komponenter, mindsker udbyttet og forringer kvaliteten på sukkeret. Kompleksiteten af disse urenheder har gjort det svært at analysere den kemiske sammensætning i procesprøver og man har ofte måttet ty til modelsystemer.

Med udviklingen af kemometrien er det blevet muligt at håndtere store mængder spektrale data med mange indgående komponenter. Dette har bl.a. vist sig at være anvendeligt til at få information om den kemiske variation fra fluorescensmålinger på procesprøver i roesukkerproduktionen.

Formålet med dette projekt har været at udnytte fluorescensspektroskopi og kemometri til at opnå en bedre forståelse af urenhedernes kemiske sammensætning i sukkerfremstillingen og på den baggrund er fem artikler blevet udarbejdet. I den første artikel anvendes en kombination af traditionel kromatografi og kemometrisk modellerede fluorescensmålinger til at påvise tilstedeværelsen af de to fluorescerende aminosyrer, tryptofan og tyrosin i roesukker og tyksaft samt at matematisk separere farvede højmolekylære komplekser dannet i processen. Den næste artikel tager udgangspunkt i de opnåede resultater fra roesukkerfremstillingen til at undersøge forekomsten af fluoroforer i rørsukkerfremstilling. Fra kemometriske modeller af fluorescenslandskaber af rørsukker (raw sugar) påvises tre fluorescerende komponenter, der alle kan relateres til farvedannelse i processen.

I et forsøg på at verificere oprindelsen af de farvede fluoroforer i såvel roe- som rørsukkerprocessen, sammenlignes fluorescens af syntetiserede farvekomponenter fra kendte farvereaktioner med tyksaft fra roesukkerprocessen og ”evaporator syrup” fra rørsukkerprocessen i den tredje artikel. Den kemometriske databehandling viser at der er visse ligheder mellem farvede fluoroforer fra de syntetiserede prøver og procesprøverne og at det er snarere en varierende grad af farvedannelse i samme reaktionsmønster end forskellige kemiske reaktioner, der ligger til grund for de forskelle der er mellem fluorescens af urenheder i roe- og rørsukkerproduktionen.

De to sidste artikler handler om fluorescensmålinger på faste sukkerkrystaller. Den ene artikel undersøger anvendeligheden af faststof-fluorescensmålinger til at analysere kvaliteten af roesukkerprøver ved at sammenligne to forskellige faststofmålemetoder med fluorescens af opløsninger. Resultaterne viser at klassifikation af prøver fra forskellige fabrikker er mindre entydig ved at bruge faststofmålingerne, men at kvalitetsparametre relateret til kemiske fænomener (farve, aske) kan prædikteres fra faststofmålingerne med omtrent samme fejl som ved at benytte opløsninger. Krystalprøvernes indvirkning på fluorescensmålingerne har stor betydning for de kemometriske modeller, hvilket uddybes i en efterfølgende artikel. Udfra sigtede krystalprøver af roesukker påvises sammenhængen mellem krystalprøve og faststofmålemetoder. Det fastslås at fluorescensrefleksion er en mere usikker metode anvendt på de gennemsigtige sukkerprøver end fluorescenstransmission. Karakteristisk af farvefordelingen i sukkerkrystaller kan dog ikke lade sig gøre netop pga. faststofmetodernes dominerende indvirkning på de målte fluorescensdata.

List of publications

PAPER I

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 and Lars Munck, Food Chemistry, 70 (1), 2000, 113-121.

PAPER II

Fluorescence of raw cane sugars evaluated by chemometrics. Dorrit Baunsgaard, Lars Nørgaard and Mary An Godshall, Journal of Agricultural and Food Chemistry, accepted for publication.

PAPER III

Specific screening for color precursors and colorants in beet and cane sugar liquors in relation to model colorants using spectrofluorometry evaluated by HPLC and multi-way data analysis. Dorrit Baunsgaard, Lars Nørgaard and Mary An Godshall, Journal of Agricultural and Food Chemistry, submitted.

PAPER IV

Evaluation of the quality of solid sugar samples by fluorescence spectroscopy and chemometrics. Dorrit Baunsgaard, Lars Munck and Lars Nørgaard, Applied Spectroscopy, 54 (3), 2000, 438-444.

PAPER V

Analysis of the effect of crystal size and color distribution on fluorescence measurements of solid sugar using chemometrics. Dorrit Baunsgaard, Lars Munck and Lars Nørgaard, Applied Spectroscopy, 54 (11), 2000, in press.

Table of Contents

Abbreviation list

HPLC	High performance liquid chromatography
ICUMSA	International Commission for Unified Methods of Sugar Analysis
MLR	Multiple linear regression
PARAFAC	Parallel factor analysis
PCA	Principal component analysis
PLS	Partial least squares regression
PV	Principal variables
SIMCA	Soft independent modelling of class analogy
UV-VIS	Ultraviolet-visible absorption

1. Introduction

White sugar is produced from sugar cane and sugar beet with a purity of more than 99.8 % sucrose (Schiweck and Clarke, 1994). However, the remaining 0.2 % or less contains trace amounts of other carbohydrates, such as polysaccharides, and inorganic and organic non-sugar compounds including colorant polymers. The impurities contribute to a reduction in yield, i.e. influence the crystallization rate, and lower the sugar quality due to incorporation in the crystal. Color impurities have received much attention due to their complex composition and multitude of chemical pathways but despite extensive research for more than a century (Scheibler, 1869), there are still many unknown aspects concerning the identity, origin, and formation of the undesirable compounds. Traditional separation chromatography coupled with UV-VIS detection has mainly been the analysis technique used to study the color composition of sugar process samples, but the complexity of the process samples and the large variety of chemical species makes it difficult to obtain a clear identification of the more composite compounds and to comprehend their interactions in the processing (Madsen *et al*., 1978; Shore *et al*., 1984, Papageorgiou *et al*., 1999).

Fluorescence spectroscopy compared with absorption spectroscopy has the advantage of selectivity and sensitivity, i.e. the spectral information is simplified (fewer fluorophores) and the detection of trace compounds is improved. It has been known for years that commercial sugar exhibits fluorescence. In 1972 Carpenter and Wall measured fluorescence emission spectra at several excitation wavelengths of cane sugar process samples. Excitation-emission contour plots were inspected and they found a characteristic emission peak pattern in the samples. However, they could only base their conclusions about the chemistry in the samples on visual comparisons of dominant fluorescence peaks.

With the use of multivariate mathematical and statistical methods (chemometrics), it has become possible to extract relevant chemical information hidden in the spectral data. Fluorescence spectroscopy has proved to be an effective screening method in the beet sugar manufacturing process when applying appropriate chemometric models to the multivariate data (Munck *et al.*, 1998). In a study of fluorescence from beet sugar

samples, SIMCA classification of the samples according to factory and PLS models predicting quality parameters showed that information about the chemistry in the samples with regard to the composition of the raw material as well as the influence of the process is present in the fluorescence spectra (Nørgaard, 1995b).

In another study of beet sugar fluorescence the three-dimensional structure of fluorescence excitation-emission landscapes was utilized to resolve spectral excitation and emission profiles of fluorophores in sugar with a multi-way chemometric model, PARAFAC (Bro, 1999). Four fluorescent components were found that captured the variation in the fluorescence data of sugar samples collected from a factory during a three-months sugar campaign. The concentrations of the four components estimated from the sugar samples could be correlated to several quality and process parameters, e.g. color, ash and pH, and they were characterized as potential indicator substances of the chemistry in the sugar process stream. The authenticity of the estimated pseudo components was partly verified by the similarity of the spectral appearance of two of the components with the pure fluorescence spectra of the amino acids, tyrosine and tryptophan, which are known to be present in sugar beet process syrups (Drewnowska *et al*., 1979).

To exploit these promising results from fluorescence screening of beet sugar, a study has been undertaken to gain further insight into the identification and characteristics of the fluorophores in sugar process samples. The fluorescence of sugar impurities has proved to be closely related to the color development during processing and in this context five papers have been prepared (cf. List of Publications).

The aim of this thesis is to discuss the essential findings of the five publications in relation to the established research. The papers are referred to in the text by their Roman numerals. In Chapter 2, as an introduction to the experimental work in the five papers, the concept of using fluorescence screening and chemometric modelling on complex samples is addressed. Chapter 3 discusses the fluorescence results of characterizing colorants and color precursors in sugar processing (PAPER I-III), and in Chapter 4 the results of measuring fluorescence directly on sugar crystal samples is discussed (PAPER IV-V). Conclusions and perspectives are presented in Chapter 5. The papers in full length are found in the back of the thesis.

2. Fluorescence screening and chemometric modelling

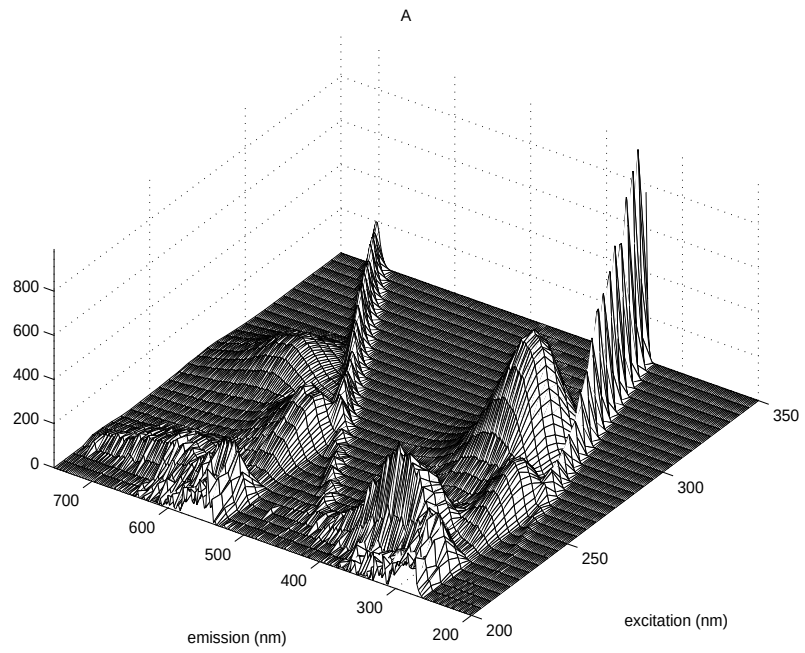
2.1 The acquisition of fluorescence data

Fluorescence spectroscopy is widely used as an analytical technique in many fields of science including chemistry, biology, biochemistry, medicine, environmental science, and food science (Munck, 1989; Lakowicz, 1999; Rettig, 1999). The simplest way to measure fluorescence is with instruments recording a response at a pre-set excitation wavelength and emission wavelength, e.g. in a quantitative measurement of an added fluorescent probe using a filter-based instrument (Georgi and Burba, 1977). With just one measurement it is not possible to detect interferences.

In reported fluorescence analyses it is usually an emission spectrum (or excitation or both) that is presented, measured at the excitation (or emission) maximum of the fluorophore of interest (Adhikari and Tappel, 1973; Pongor *et al*., 1984). Many authors often tend to report only the fluorescence excitation and emission wavelength maxima of a known analyte despite the use of a scanning fluorometer, e.g. in kinetic analysis (Hayase *et al*., 1996; Morales *et al*., 1996).

Measuring several emission spectra at different excitation wavelengths (or vice versa) creates an excitation-emission matrix (EEM) or a landscape that covers the total area of fluorescence (Fig. 2.1). The landscape structure has the advantage that analytes or interferences peaking in different areas are discovered. In the fluorescence screening analysis of complex samples of unknown origin, a landscape should always be measured.

(a)



(b)

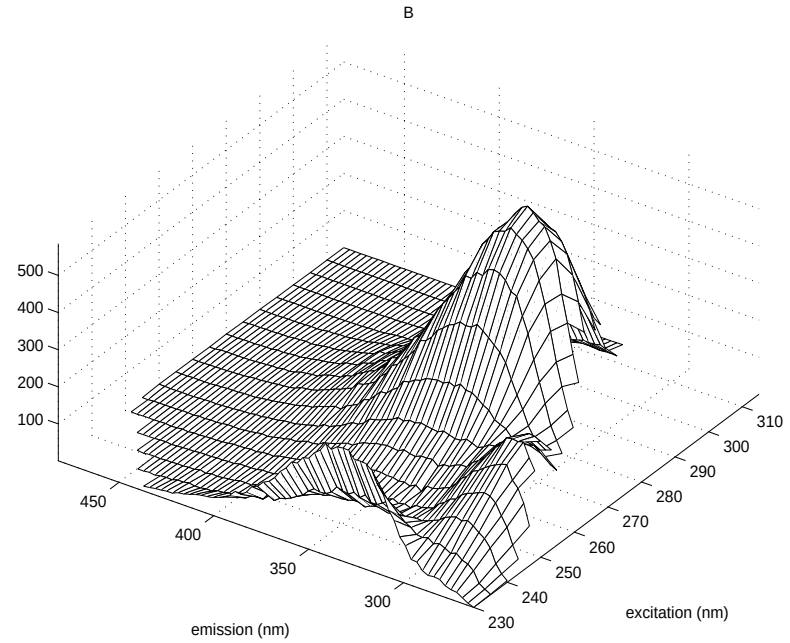


Fig. 2.1 Fluorescence landscape of a sample of mixed fluorophores before (a) and after (b) replacement of all areas except the first order fluorescence data with missing (white areas). From the right: first order Rayleigh scattering and fluorescence, and then second order Rayleigh scattering and fluorescence.

2.2 Fluorescence data adjustment

The fluorescence landscape shown in Fig. 2.1a consists of 31 emission spectra (range 200-750 nm, 0.5 nm interval) measured at 200-350 nm excitation wavelengths with a 5 nm interval. Before modelling fluorescence landscape data, it is important to remove all the areas not conforming to true fluorescence. These areas include the Rayleigh scattering peak, first and second order, but also all variables lower than the Rayleigh peak. The second order fluorescence is removed as well to keep the modelling simple and is generally not measured at all. The Raman peaks are usually hidden below or are slightly visible in the measured fluorescence spectra. Subtracting a blank spectrum is the best way to try to remove the contribution from the Raman peak, but because the Raman peaks behaves differently from the spectral response, bilinear calibration models can treat the Raman peak as interference (Nørgaard, 1995a). In the modelling of three-way fluorescence data from the sugar samples using PARAFAC (cf. Section 2.4), it was observed that Raman peaks clearly visible in the measured fluorescence landscape were hardly influencing the model and were thus found in the residuals (Fig. 2.2).

When all unwanted areas have been removed, the wavelength area has been reduced to include only data with reasonable intensity, and the landscape has been reduced to every third emission wavelength, it is ready to be used as one sample entry in a chemometric model (Fig. 2.1b). The white areas denote the missing data areas.

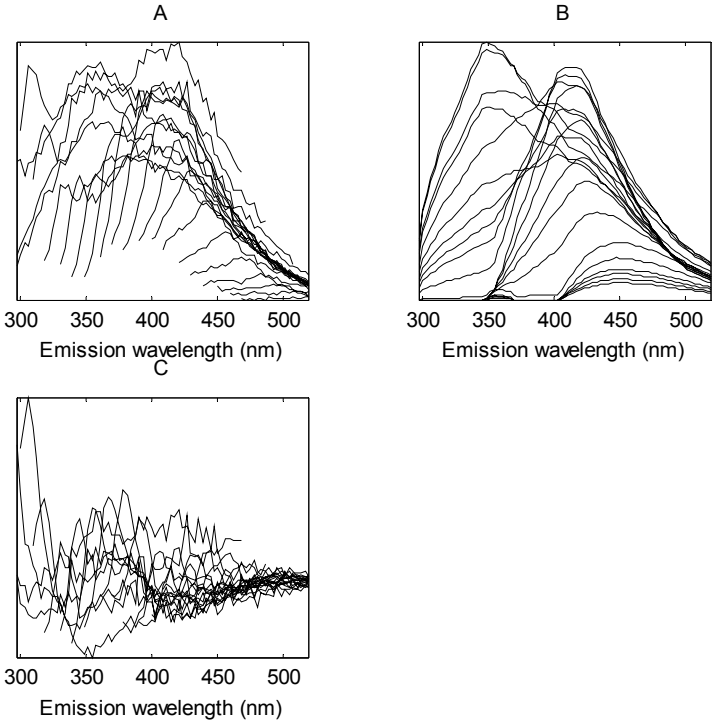


Fig. 2.2 Plots of the fluorescence landscape of a raw sugar sample taken from a three-component PARAFAC model of 47 raw sugar samples (PAPER II) demonstrating that the Raman peaks that are clearly visible in the raw data (A) are not present in the modelled fluorescence of the sample (B) but are visible in the residual plot (C) demonstrating that the Raman peaks in the landscape are treated as non-trilinear data.

2.3 Modelling of fluorescence landscape data – bilinear modelling

The bilinear chemometric methods, principal component analysis (PCA), partial least squares regression (PLS) and principal variables (PV) were used in the analysis of fluorescence data in this project. These methods are well established as statistical methods for the analysis of spectral and highly co-linear data structures (Höskuldsson, 1994; Martens and Næs 1993). The bilinear models are two-dimensional models and the fluorescence landscape cannot be used in these models directly. Instead the landscapes are "unfolded" into a two-dimensional data set (samples x wavelength variables). In an unfolded landscape the emission spectra are arranged as one vector in ascending order of the excitation wavelengths and the missing data areas are removed (Fig. 2.3).

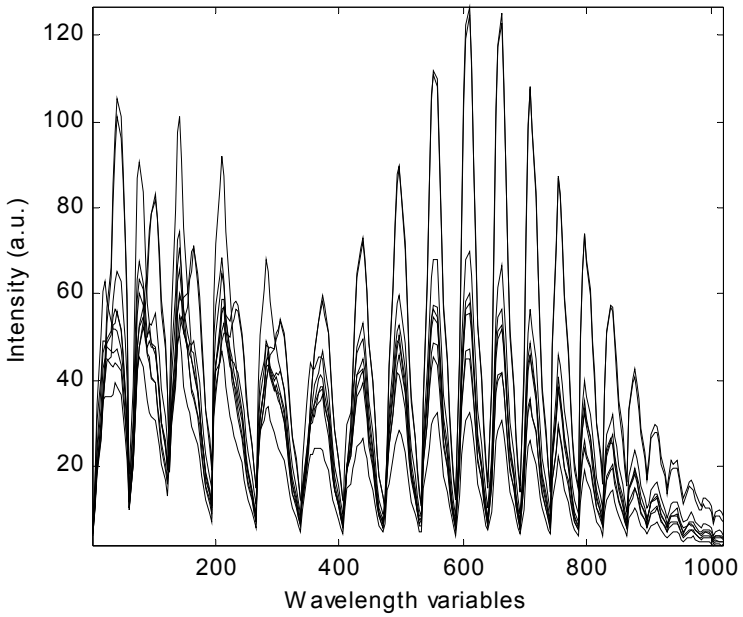


Fig. 2.3 Unfolded fluorescence landscapes of nine samples with 21 emission spectra arranged after excitation wavelength in ascending order.

According to Bro (1998), unfolding a three-way data structure, e.g. a sample set of fluorescence landscapes, to a two-way data structure may lead to models that are less robust and less predictive. However, when the fluorescence landscape data sets from PAPER II-IV were modelled with a three-way regression model known as N-PLS (Bro, 1996), the predictive ability of the PLS models presented in the said papers was not improved. The reason for lack of improvement is probably due to the fact that fluorescence is correlated to chemical parameters (color, amino-N, etc.) with a few excitation and emission wavelength pairs as shown by Nørgaard (1995b). Thus there is no advantage in maintaining the three-way structure in the fluorescence data for predictive purposes contrary to the advantage of the three-way structure in making it possible to resolve individual fluorophores with PARAFAC.

2.4 The trilinearity of fluorescence and the PARAFAC model

The fluorescence intensity in a sample *i* measured at emission wavelength *j* at excitation wavelength *k* can be expressed as (Leugans and Ross, 1992):

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

where a_{if} is the concentration of the *f*th analyte in the *i*th sample, b_{jf} is the relative emission emitted at wavelength *j* of analyte *f*, and c_{kf} is the relative amount of light absorbed at the excitation wavelength *k* of analyte *f*. Under the assumption that the absorbance of the samples is very small and no energy transfer takes place between fluorophores, trilinear fluorescence data as described in equation (1) can be approximated by the PARAFAC (parallel factor analysis) model (Harshman, 1970; Bro, 1997). An important advantage of the PARAFAC model when it comes to fluorophore identification is the uniqueness of the solution. Uniqueness means that the scores and loadings cannot be rotated as in PCA and if the proper number of components is used, it is possible to estimate pure excitation and emission spectra of the analytes by decomposing the fluorescence data with PARAFAC (Bro, 1997).

In reality fluorescence data of complex samples does not often meet the terms of a trilinear model. The assumptions stated above are rather strict when considering the chemical conditions of natural samples and the physical conditions of the fluorescence measurements. Factors such as concentration quenching, interactions between analytes and light scattering may influence the model results. In this project these effects have made it difficult to estimate the number of fluorophores in the models (PAPER I and II) and have lead to perturbations in spectral profiles (PAPER III-V). The physical conditions of the measurements (Rayleigh peaks, etc.) leading to missing areas in the fluorescence data (cf. Section 2.2) may also influence the PARAFAC model results. In Fig. 2.4 some artifacts introduced by missing data areas are demonstrated. In the estimated excitation spectrum of tyrosine (dash-dot), an extra band at 340-400 nm is caused by model instability in the missing data area at the higher excitation wavelengths and the 330-360 nm wavelength area in the estimated colorant emission spectrum (dash)

is due to the missing data area at the lower excitation wavelengths (Fig. 2.1b). Despite the above mentioned factors that has to be considered when validating the PARAFAC model results, the models presented in this thesis together with the information that was gained from them demonstrate the usefulness of the PARAFAC model when dealing with fluorescence screening analyses.

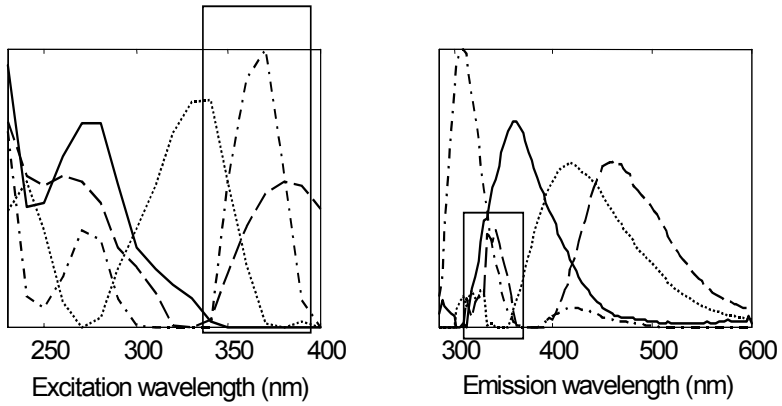


Fig. 2.4 Example of PARAFAC model results of beet sugar where there are modelled unwanted intensity bands in the spectra (see black boxes) due to the missing data areas. Unwanted band in dash-dot excitation spectrum at 340-400 nm and in dash emission spectrum at 330-360 nm.

3. Colorants and color precursors in sugar processing

The development of color is dependent on the chemical composition of the sugar cane and sugar beet process streams and a sensible interpretation of the fluorescence results therefore demands a basic understanding of the actual conditions in the sugar factories.

3.1 The processing of white sugar from sugar cane and sugar beet

The process of extracting sugar from either sugar cane or sugar beet involves many production steps and a thorough presentation of every step in the production is beyond the scope of this thesis. More detailed information concerning sugar production is found in Schiweck and Clarke (1994) and DDS (1985).

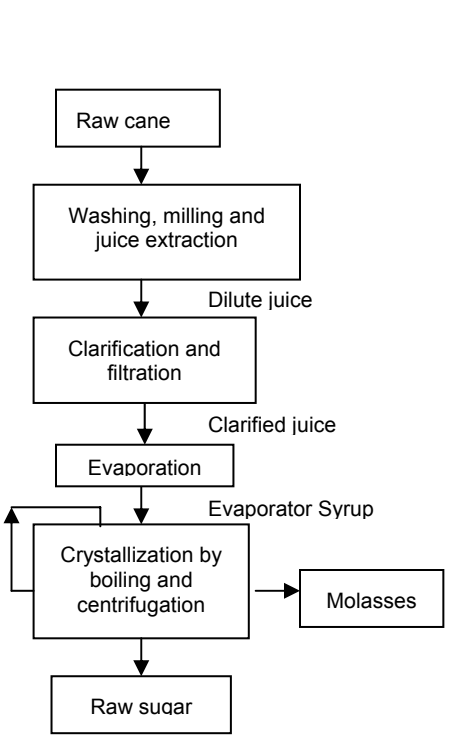


Fig. 3.1 Typical process flowchart for the production of raw sugar at a cane factory (adapted from Clarke *et al*., 1997)

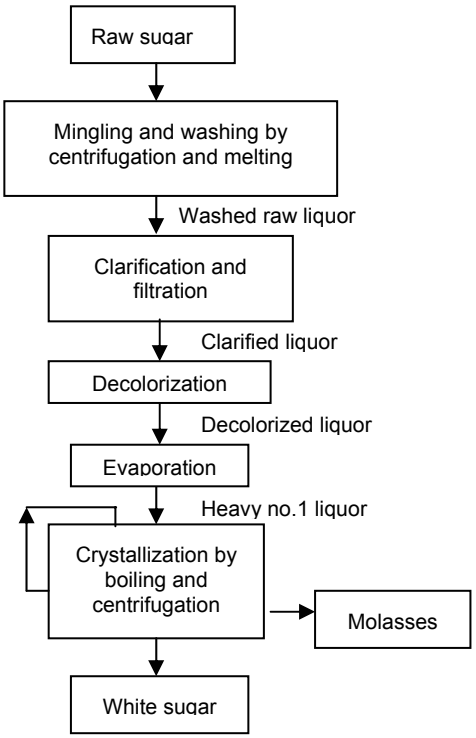


Fig. 3.2 Typical process flowchart the for production of white cane sugar at a cane refinery (adapted from Clarke *et al*., 1997)

Outlines of the typical unit processes for the manufacture of sugar from sugar cane and sugar beets are illustrated in Figs. 3.1-3.3. One of the most pronounced differences between beet and cane sugar processing is that white cane sugar is generally produced in two separate processes and beet sugar in one continuous process. Nevertheless, both white sugar end products fulfill the same quality criteria according to international standards in terms of sucrose content, color, invert sugar and ash content (Schiweck and Clarke, 1994).

In a cane sugar factory, *dilute juice* from washed, milled cane is clarified to *clarified juice*, evaporated to *evaporator syrup* (62-69% solids), and crystallized to *raw sugar* (Fig. 3.1). *Molasses* is run-off syrup from the centrifugal separation of crystals and syrup. Raw sugar is a golden yellow/brown sugar with crystals containing approximately 97-98 % of sucrose and the rest is moisture, invert sugars, and non-sugar compounds including colored compounds (Clarke *et al*., 1997). The produced raw sugar is transported to refineries where the sugar is washed to remove the syrup coating and then melted (*washed raw liquor*), clarified to *clarified liquor*, decolorized to *decolorized liquor* and then crystallized to *white sugar* with *molasses* as a by-product (Fig. 3.2). Refined cane sugar has not been examined in this project but the process outline is mentioned to illustrate the difficulties with producing a white cane sugar in comparison with the beet sugar process.

In the sugar beet factory (Fig. 3.3), the beets are washed, sliced and the sugar is extracted with a diffusion technique using 70 °C hot water (*diffusion juice*). After several clarification steps with carbonation and sulfitation, *thin juice* is evaporated to *thick juice* (68-74% solids). Thick juice is mixed with backfeeded remelted raw sugar crystallized from run-off syrups and this mixture is called *standard liquor*. *White sugar* is crystallized from standard liquor and *molasses* is produced as a by-product.

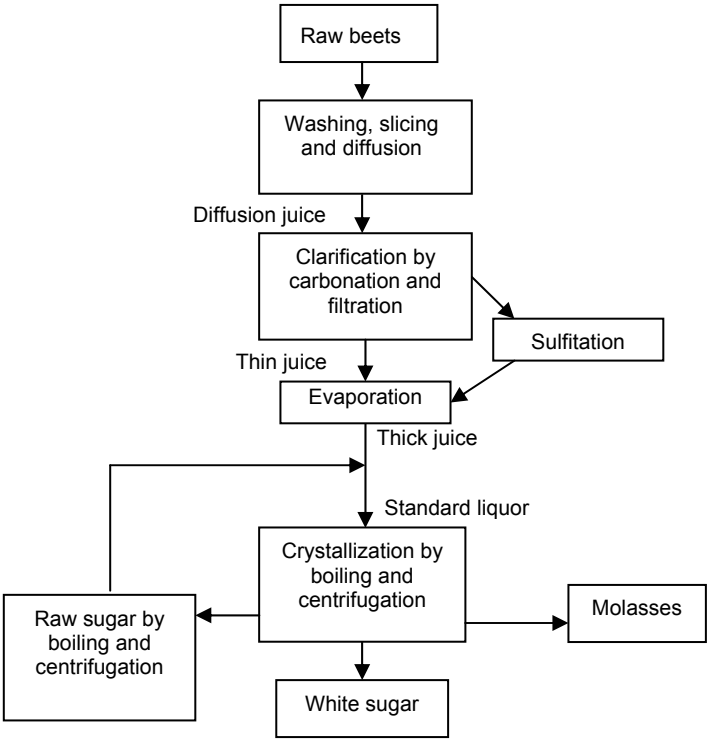


Fig. 3.3 Typical process flowchart for the production of white sugar at a beet factory (adapted from Clarke *et al*., 1997)

3.2 Potential fluorophores in beet and cane sugar processing

There are several potential fluorophores in sugar processing. Chemical compounds having fluorophore characteristics both originate from the plant and are formed in the process. An important group of plant-derived fluorophore candidates are amino acids and polyphenolic compounds (Wolfbeis, 1985). This group includes the three fluorescent amino acids, phenylalanine, tyrosine, and tryptophan, but also catecholamines such as dopa and dopamine, and phenolic acids such as ferulic acid, 4-coumaric acid, and 4-hydroxybenzoic acid have fluorophore characteristics. Chemical structures and fluorescence excitation and emission maxima of the named fluorophores are listed in

Appendix A. Catecholamines have been isolated from beet diffusion juice (Winstrøm-Olsen *et al*., 1979) and the three amino acids were found in beet process samples from diffusion juice to molasses, but they were not determined in beet sugar (Drewnowska *et al*., 1979). The organic acids have been found in beet standard syrup and raw cane sugar (Godshall, 1996; Papageorgiou *et al*., 1999).

Flavanoids and anthocyanins are pigments that are found in the sugar cane plant (Godshall, 1996), and they are known to fluoresce (Wolfbeis, 1985).

All these plant-derived fluorophores are either classified as potential colorants or color precursors (Winstrøm-Olsen *et al*., 1979; Godshall, 1996). Color precursors are participating in color forming reactions during processing and it is therefore as important to characterize color precursors as colorants.

The process-derived colorants are divided into different categories. Melanoidins are products from the Maillard reaction of amino acids and reducing sugars, e.g. fructose and glucose (cf. Section 3.3). Melanin is formed in the enzymatic oxidation of phenolic compounds, and degradation reaction products are formed in the degradation of sucrose (caramel), fructose, and glucose (Madsen *et al*., 1978; Shore *et al*., 1984). Also ferrous iron (Fe^{2+}) can form colored complexes with polyphenolic compounds or caramels to darker colorants (Godshall and Chou, 2000).

It is known that fluorophores are formed both in the Maillard reaction (Adhikari and Tappel, 1973) and in sucrose and fructose degradation (Carpenter and Roberts, 1974). Melanin formed in the enzymatic reactions of polyphenols is a dark pigment insoluble in aqueous solutions (Robinson and Smyth, 1997). It is formed from phenolic compounds such as tyrosine and dopa in beet diffusion juice and from chlorogenic acid in cane dilute juice, i.e. in the initial processing steps when the enzymes are still active (Gross and Coombs, 1976). Melanin in itself is not considered to be a potential fluorophore in sugar processing but a recent report reveals that soluble intermediates in the melanin reaction are fluorescent (Mosca *et al*., 1999). It is not known if iron complexes in sugar processing show fluorescence, but a complex of L-dopa and Tb (III) has been reported to fluoresce at 490 nm and 550 nm emission maxima (Yang *et al*., 1993).

All the color-forming reactions are complex and they involve many intermediate steps. The reaction products consist of polymers of varying molecular sizes (melanin,

melanoidins, and degradation reaction products) and side reactions produce smaller by-products that may also recombine further to other polymers (Gross and Coombs, 1976; Belitz and Grosch, 1987; Dills, 1993). It is difficult to isolate fluorophores from the color reactions but in the Maillard reaction using model systems some fluorophores have been isolated with nitrogen-containing structures from incorporated amino acid residues (Chio and Tappel, 1969; Pongor *et al*., 1984).

3.3 The influence of pH in sugar processing

The pH level highly influences the chemical reactions in sugar processing. One of the most important reactions, which is pH dependent, is the acid hydrolysis of sucrose to equimolar mixture of glucose and fructose also known as invert sugar. The hydrolysis of sucrose to invert sugar is known to occur at pH up to 8.3 (Parker, 1970). In the initial process juices pH is low both in cane dilute juice (4.5-5.5) and in beet diffusion juice (5.8-6.2) and the conditions are favorable for formation of invert sugar (Schiweck and Clarke, 1994). In beet sugar processing pH is raised to 10-11 during the purification step and one of the reasons is to decrease the formation of invert sugar (DDS, 1985). However, the inverts already formed participate in degradation reactions and the Maillard reaction that are more efficient at alkaline conditions (Clarke *et al*., 1997). In raw cane sugar processing pH is maintained at 7-8 and thus the sucrose hydrolysis may continue through the processing. The differences in pH in beet and cane sugar processing have a direct impact on which chemical reactions that will predominate in the process. This together with the fact that beet juice in general contains less invert sugar but more amino acids and other nitrogen-containing compounds than cane juice, are important factors to consider (Clarke *et al*., 1997).

3.4 Colorant and color precursors in beet sugar and beet sugar syrup

The first aim of this project was to examine more closely the chemical identity of the fluorescent components that are found in beet sugar samples (Bro, 1999). As a way to accomplish this, size exclusion chromatography was combined with fluorescence

landscapes measurements of fractions collected from the column (PAPER I). Thick juice from the stage in beet processing sugar where the crystallization of sugar commences (Fig. 3.3) was used in the analysis partly due to the fact that beet sugar is too pure to collect fractions concentrated enough for reasonable fluorescence measurements but also to examine the color formation at a previous stage than the end product. Colorants in sugar processing have been reported to separate into well-defined groups of higher and lower molecular weight using size exclusion chromatography (Shore *et al*., 1984; Godshall *et al*., 1991), and the size exclusion separation was therefore used to establish a connection between fluorophores and colorant polymers.

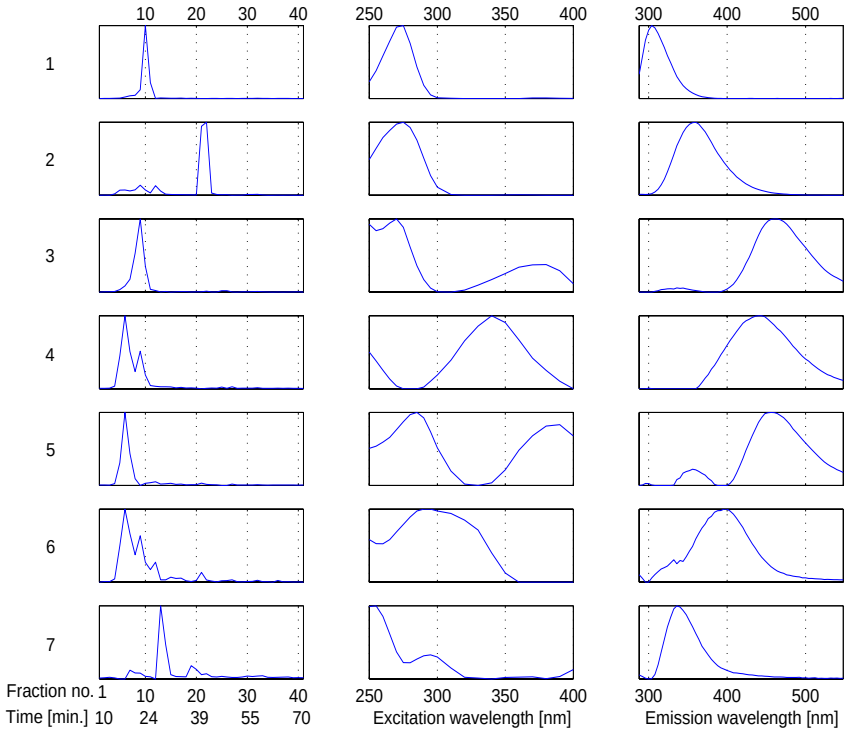


Fig. 3.4 A seven-component PARAFAC model of the measured fluorescence landscapes of collected HPLC fractions of thick juice. The left-most column holds the chromatographic profiles, which show the concentration of each of the seven fluorescent components in the fractions. The centre column holds the excitation profiles and the right-most column holds the emission profiles of the fluorophores. Row 1: tyrosine, row 2: tryptophan, row 3-6: visible fluorophores, row 7: ultraviolet component (PAPER I).

The fluorescence data measured from the collected fractions of a thick juice sample was resolved into excitation and emission spectra and chromatographic profiles of seven individual components with PARAFAC (Fig. 3.4). In PARAFAC models based on comparable fluorescence landscape data measured directly on beet sugar and thick juice samples only four and five components could be resolved, respectively.

Of the seven fluorophores found in the fractions, two of them were confirmed to be tyrosine and tryptophan using HPLC peak identification and their spectral identity as two of the four fluorophores in beet sugar was equally established. The five remaining fluorophores divided themselves into four components in the visible area and one in the ultraviolet. Three of the visible fluorescent components (row 3-5 in Fig. 3.4) show similar emission spectra (approx. 450 nm), but due to their different excitation spectra and/or chromatographic profiles PARAFAC resolved three components. However, as stated in PAPER I it is questionable if the three components are true individual fluorophores. They are more likely average spectra of colorant polymers that fluoresce at slightly different wavelengths. This is supported by the fact that the PARAFAC models of fluorescence from thick juice and beet sugar samples could only resolve one component that has an emission spectrum in the same wavelength area and that the chromatographic profiles of the three components are resolved in the same chromatogram area with overlapping peaks in Fig. 3.4. Pongor *et al*. (1984) described the fluorescence characteristics of an isolated fluorophore from the Maillard reaction of polypeptides and glucose. The fluorophore, an imidazole derivative, had the same fluorescence excitation spectra as the browning mixture itself (excitation/emission max. at 380/440 nm), but their absorption spectra differed. Pongor and co-authors suggested that the fluorescence of the browning mixture was primarily associated with a subclass of the Maillard products that could be chemically isolated. Considering the similarity in fluorescence behavior of the imidazole fluorophore and the three resolved fluorophores in row 3-5 in Fig. 3.4, it is not unlikely that all three fluorophores belong to the same type of Maillard products. Besides, the diode array absorption spectrum of that particular peak area in the chromatogram (peak 3 in Fig. 4, PAPER I) is very similar to the absorption spectrum of the Maillard browning mixture reported by Pongor *et al*. (1984).

The fourth visible component (row 6) is a likely intermediate in the formation of the more colored fluorophores.

In general, using the HPLC column to separate the thick juice sample before performing the fluorescence measurements succeeded in providing more detailed information about the fluorophores. The number of resolved components was increased and despite the fact that some of the extra components are probably representing the same type of colorants, the seventh component in Fig. 3.4 was only modelled as an individual component because of the column separation. This component is low in fluorescence intensity and may very well be one of the many phenolics that are present in sugar beet processing (e.g. p-hydroxybenzoic acid in Appendix A). The chromatographic profiles resolved in the PARAFAC model also helped to characterize the resolved fluorophores. The visible fluorophores that eluted first on the column are higher molecular weight compounds. This supported the identity of these fluorophores as colorant polymers instead of small fluorescent compounds originating from the sugar beet, e.g. ferulic acid (cf. Appendix A). The chromatographic profiles are thus used as a "mathematical separation" that can resolve overlapping peaks. For example, the PARAFAC results in Fig. 3.4 show that tryptophan is present as a fluorescent substructure in the higher molecular weight color compounds, which is not the case for tyrosine. These findings of tryptophan substructures in the colorants support that the colorant structures in beet sugar and thick juice are derived from Maillard reactions of amino acids and carbohydrates. The lack of tyrosine fluorescence in higher molecular weight compounds is also found in proteins dominated by tryptophan fluorescence (Lakowicz, 1983).

3.5 Fluorescence in raw cane sugar

As a result of the success of the combined spectral and chemometric analyses on beet processing samples, similar studies were initiated to analyze color formation in cane sugar processing. Color is more easily co-crystallized in cane sugar crystals than in beet sugar crystals, e.g. a 20-30 ICU (ICUMSA unit) color white sugar is produced from a 2000-3000 ICU beet standard liquor whereas a cane sugar syrup can only have a color about 200-300 ICU to produce the same low color of white sugar (Godshall and Chou,

2000). This implies that cane sugar process samples have a different color composition and/or different factors influencing the color transfer into the crystal compared with the beet sugar process samples. Raw sugar, the end product from sugar cane factories, is an appropriate intermediate product to examine considering the rather large proportion of color impurities occluded in the crystals (cf. Section 3.1). The study (PAPER II) involved a sample set of raw cane sugars from all over the world, which ensured that the chemometric models were not susceptible to local variation of raw material or production methods. The sugar samples were also spread over forty production years and this was used to study the formation of color during storage, which is problem with high storage temperatures or long-term storage (Mak *et al*., 1975; Shore *et al*., 1984). Two-way models (PCA, PLS and MLR on selected variables) based on unfolded fluorescence landscapes of 47 raw sugar samples showed that three significant fluorophores are responsible for the fluorescence in raw sugar. A four-component PARAFAC model turned out to be the best model based on the three-way fluorescence landscape data (Fig. 3.5). However, two of the components had very similar emission spectra (component II and III). Based on the other model results, it was suggested that these two components represent the same type of colorant fluorophore but they are resolved as two components because of small wavelength shifts in the fluorescence. The same explanation was used to account for the three components with similar emission spectra resolved from the PARAFAC model of fractionated beet thick juice (cf. Section 3.4).

These wavelength shifts are either caused by differences in fluorophore size due to darker color or reabsorption of fluorescence at lower wavelengths in samples of darker color. Regardless of the actual reason, it is the gradual darkening of color in sugar samples that makes it very difficult for PARAFAC to decompose colorant fluorescence into the true number of independent colorant fluorophores, if indeed such a number is to be found (cf. Section 2.4).

The “three” principal fluorophores thus found in raw sugar are all indicators of color. One of them (component I) is considered to be a color precursor participating in the formation of color during storage (excitation/emission max. at 275/350 nm). It has been stated that color formed during storage is generated by the Maillard reaction as well as the thermal degradation reactions (Mak *et al*., 1975; Godshall *et al.*, 1991).

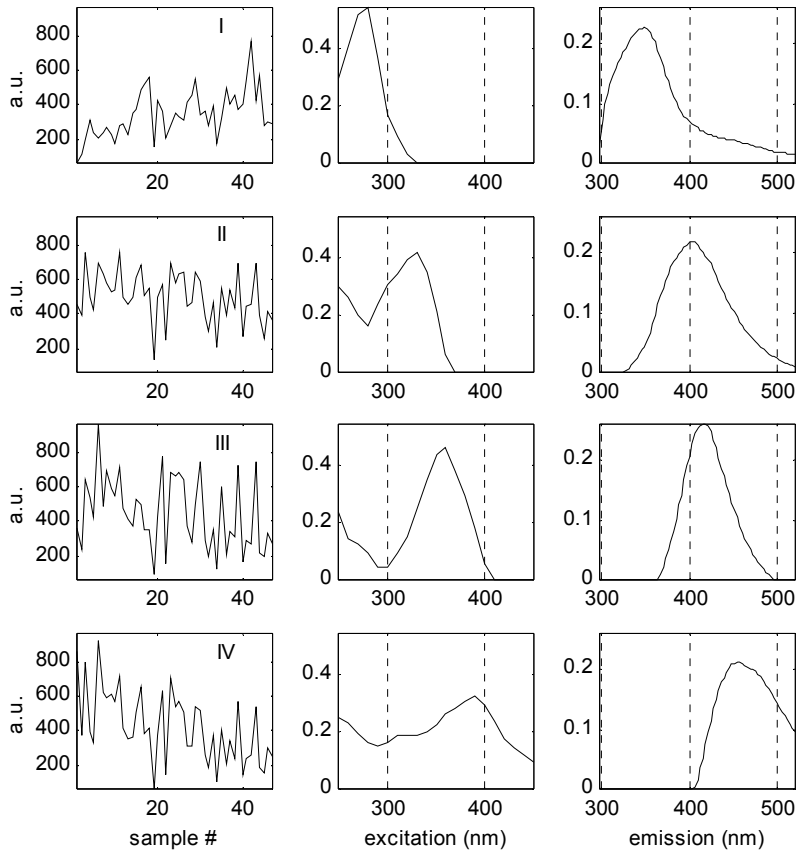


Fig. 3.5 A four-component PARAFAC model of the measured fluorescence landscapes of 47 raw sugar samples. The left column shows the concentration profiles, the middle column shows the excitation profiles, and the right column shows the emission profiles of the fluorophores. Component I: color precursor of storage (possible tryptophan), component II and III: suggested representatives of the same fluorophore, and component IV: storage colorant (PAPER II).

The resolved color precursor fluorophore has a close resemblance with tryptophan fluorescence and the presence of tryptophan in evaporator syrup, from which raw sugar is crystallized, has been established (PAPER III). Tryptophan fluorescence represents the amino acid residues in the raw sugar samples participating in the Maillard reaction, which agrees with the disappearance of this fluorophore with aging of the sample. Also the low water content of the sugar crystals is a favorable condition for the Maillard reaction (Fennema, 1985).

Component II/III and component IV have colorant characteristics and are considered to represent two groups: lighter colorants and darker colorants. The lighter colored fluorophore (II/III) at approx. 340/420 nm has very intense and distinct fluorescence in most of the samples whereas the dark color fluorophore (IV) at the highest wavelengths (390/460 nm) only appears as an individual peak in the fluorescence landscapes in the very old samples. Prey and Andres (1971) reported that the darkest colored fractions of molasses had higher contents of carbonyl groups and lower contents of hydroxyl groups and double bonds. Generally, molecular structures with double bonds have electronic transitions at higher energy (lower wavelengths) than carbonyl containing structures and the results of Prey and Andres can explain the observed change in fluorophores with the development of darker color. Hence, when color darkens in raw sugar (age or heating), the fluorophore structure of II/III is extended to IV presumably in polycondensation reactions. No apparent decrease of component II/III in the old samples is found as more component IV is formed, and this suggests that during storage of raw sugar more of component IV color is formed in browning reactions involving component I with component II/III acting as an intermediate reaction product.

A provisional experiment undertaken to examine the formation of fluorescent Maillard reaction products with time supports this theory (Appendix B). From a gradually browning mixture of L-dopa and glucose samples were taken at 0, 11, and 22 min and their fluorescence landscapes were modelled with PARAFAC. The model results (Fig. B.1 in Appendix B) show that at 0 min only the L-dopa fluorophore is present. After 11 min L-dopa is almost used up in the browning reaction but two new fluorescent components have been formed with excitation/emission maxima at 330/400 nm and 380/480 nm. After 22 min the 380/480 component is still increasing but the 330/400 component is now decreasing indicating that this component is formed from L-dopa and acts as an intermediate in the formation of the other component, the actual colorant. The appearance and behavior of the two browning fluorophores are very close to the raw sugar fluorophores and that confirms the contribution of the Maillard reaction products to storage color formation as well as to the color formation during processing. The spectral properties of the visible fluorophores resolved from the beet thick juice fractions with PARAFAC also suggest the same colorant formation pattern, i.e. the fluorophore at

the lowest visible wavelengths is an intermediate in the formation of the darker colorants (cf. Section 3.4).

3.6 Colorant recognition using model colorant systems

The fluorescence studies of thick juice from sugar beets (PAPER I) and raw cane sugar (PAPER II) showed that specific colorant fluorophore identification in process samples is very difficult to achieve. However, since chemometric analysis can be used to compare likenesses and differences in spectral data without knowing the exact identity of the analytes involved, it is possible to employ model colorant systems as colorant "standards" to be used in colorant recognition. Accordingly, fluorescence landscapes of model colorants and process samples were measured and individual components resolved with PARAFAC were compared (PAPER III). The browning mixtures from the Maillard reactions of glucose-glycine and glucose-lysine as well as degradation products from fructose and glucose were chosen as model systems because nonenzymatic colorants reactions are known to dominate the evaporation and crystallization steps where the color composition has a direct influence on the color occluded in the crystals (Shore *et al*., 1984; Eggleston, 1998). Beet sugar thick juice and cane evaporator syrup represented the factory samples.

The emission spectra of resolved fluorophores from the various model colorant samples were practically all located in the visible wavelength area and their spectral appearances were quite similar (Fig. 3.6). The findings suggested that the different browning reactions yield the same fluorescent products and this is probable conclusion when considering the complex chemical pathways of nonenzymatic browning reactions. In the Maillard reaction between carbonyl containing carbohydrates and amino-nitrogen containing amino acids, the amino group is eliminated again from the initially formed carbohydrate-amino acid structure and some of the intermediary products resemble those in the thermal degradation reactions (Belitz and Grosch, 1987). The amino group can then interact with the formed intermediates in reactions leading to colored melanoidins with nitrogen-containing structures, but other pathways may lead to reaction products without nitrogen (Nursten and O'Reilly, 1983). Which of the various reaction pathways

that dominate are influenced by many physical factors such as the temperature, pH, reactant concentration, reaction time, and the nature of the reactants (Schallenberger and Birch, 1975; Ajandouz and Puigserver, 1999).

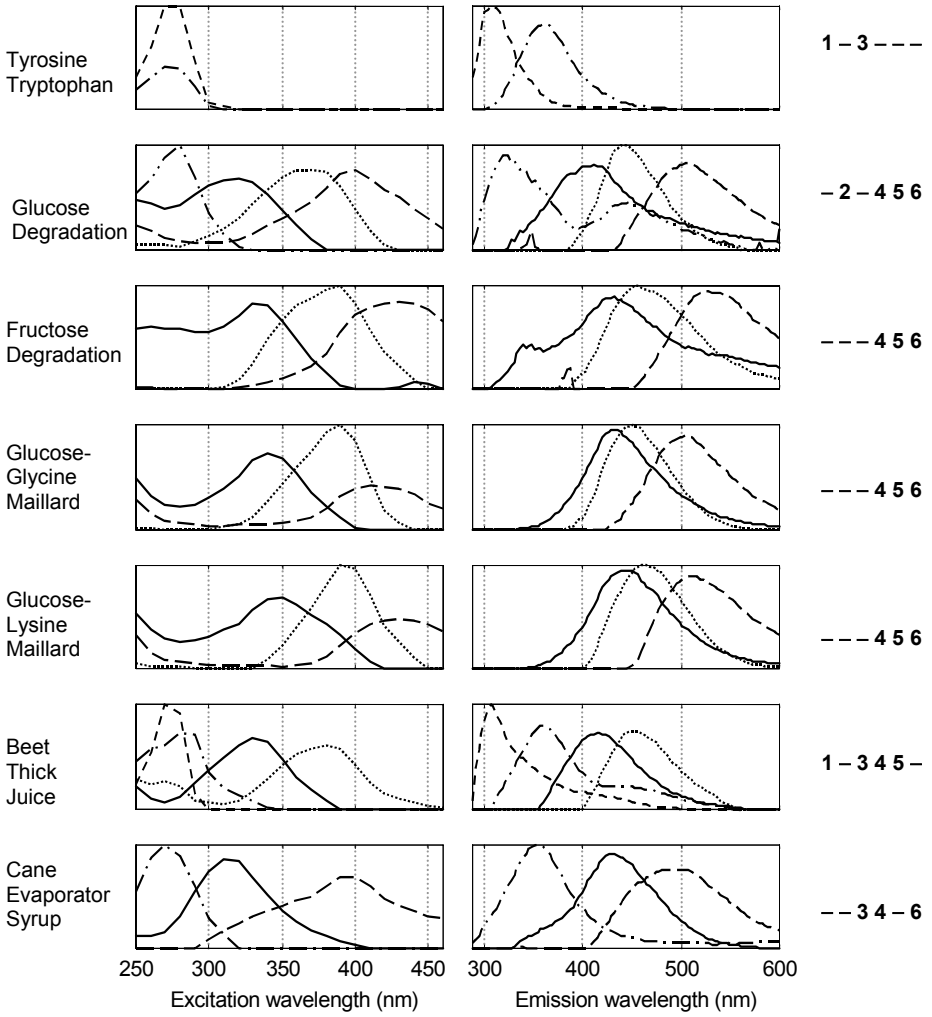


Fig. 3.6 PARAFAC models of the measured fluorescence landscapes of model colorants and beet and cane factory syrups. The emission spectra of the resolved fluorophores are compared by assigning numbers to the spectra as shown to the right in the figure. Components 1 and 3 are tyrosine and tryptophan components, respectively, component 2 is an ultraviolet fluorophore, and components 4-6 are assigned to colorant fluorophores in the visible wavelength area. (PAPER III)

Since it was not possible to differentiate the colorant fluorophores from the Maillard reactions and the thermal degradation reactions, the small changes in the spectra were explained by different reaction rates in the model systems depending on the different reactants. The resolved fluorophores from beet thick juice and cane evaporator syrup (Fig. 3.6) were subsequently matched with the fluorophores from the model browning reactions by looking at PCA score plots and correlation coefficients between the emission spectra from model colorants and factory samples (PAPER III). Spectral likenesses were found between some of the resolved components, e.g. component 5 of the beet sample and all the model colorants. The general conclusion was that beet colorant fluorophores resembled more the model colorant fluorophores at lower wavelengths and cane colorant fluorophores those at higher wavelengths.

The small differences that are found in cane and beet syrup fluorophores in Fig. 3.6 thus seem most likely to be influenced by the different physical conditions of the two processes than by different chemical reactions. As mentioned in Section 3.3, the pH level and the contents of invert sugars and nitrogen containing compounds differ in cane and beet sugar processing. These factors will as mentioned above affect the contribution from the various non-enzymatic browning reactions. For example, the higher amounts of amino acids and alkalinity at the evaporation stages of beet sugar processing would favor the Maillard reaction (Schallenberger and Birch, 1975). Another factor that has to be considered is the contribution of fluorescence from other color reactions to the total fluorescence of the factory processes. For instance, melanin intermediates from enzymatic color reactions taking place in the initial stages of sugar processing have shown fluorescence with maximum at 440 nm when excited at 330 nm (Mosca *et al*., 1999), which is in the wavelength area of the resolved colorant fluorophores from the factory samples. It has been shown that these intermediates are not totally removed in the carbonation and filtration steps in beet sugar processing (Vukov *et al*., 1984) and they may influence the overall fluorescence measured at later processing steps.

Dissimilarities in the spectra of the resolved components from model colorants and factory samples, e.g. that the model colorants showed fluorescence of higher wavelengths than any of the factory samples, demonstrate how difficult it is to obtain the right conditions of model systems to simulate real processes. Apparently the model

systems reached a more advanced stage of color formation than permitted by controlled factory conditions at evaporation level. A new fluorescent study of color development through a cane factory supports this supposition (non-published work). The study has shown that colorant fluorophores resolved with PARAFAC from the molasses end product have the closest resemblance to the model colorant fluorophores as presented in Fig. 3.6, which agrees with the fact that the molasses end product will contain the highest concentration of color impurities and thus the darkest color.

4. Fluorescence of solid sugar

Solid-phase fluorescence measurements are in general more difficult to analyze than fluorescence measurements of solution samples due to extra perturbations introduced in the spectra. These perturbations are closely connected with the progression of light in a sample of particles.

4.1 The propagation of light in a crystal sample

The exact path of the propagating radiation is difficult to describe for particulate samples especially if they contain a heterogeneous distribution of particles. Many theories on light transmitted in heterogeneous samples have been developed and they present a complex system of different scattering and reflection paths of the incident radiation (Kortüm, 1969; Hurtubise, 1981). When monochromatic light interacts with a particle, it may enter the particle or be scattered. Some of the light entering the particle is absorbed and some of it is transmitted through the sample, where the light is refracted, i.e. changes the direction of propagation. Scattering of the incident light is either reflection or diffraction. All the different propagations of the incident light are shown in Fig. 4.1.

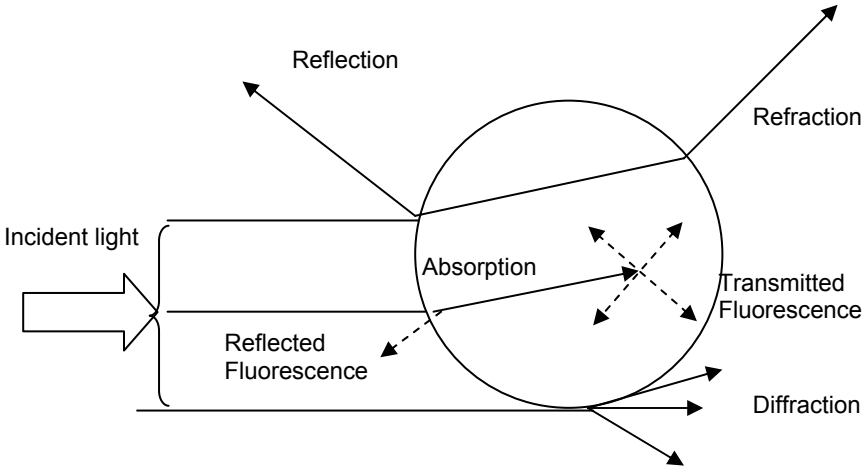


Fig. 4.1 The effects of incident light striking a fluorescent, transparent particle. Propagation of incident light (—), propagation of fluorescence (---).

The fluorescence emitted from the sugar crystals is submitted to the same particle interactions as the incident light. However, as displayed in Fig. 4.1, the two main categories of solid fluorescence is transmitted fluorescence, which is detected on an opposite face of the particle than the face where the incident lift strikes, and reflected fluorescence, which is detected on the same face of the crystal as the incident light. The scattered light as well as the reflected and the transmitted fluorescence may be reabsorbed by other particles creating a confusion of light propagation. All these light-crystal interactions can affect the detected fluorescence spectrum.

4.2 Fluorescence of sugar crystal samples

Two studies of beet sugar fluorescence measurements of sugar crystal samples were undertaken to examine the use of solid sugar samples in fast screening analysis of sample quality. The first study (PAPER IV) comprised a data set consisting of 7-15 beet sugar samples from six factories and two extra pure sugars, which had been used in a similar study of the quality characteristics of sugar solutions (Nørgaard, 1995b). The aim of this study was to examine if it was possible to extract the same information from the sugar samples as in the previous solution fluorescence study. Transmission as well as reflection fluorescence of the crystal samples was measured using detection at right angles through the sample and a reflection front surface detection, respectively.

PCA and PLS models of the measured fluorescence spectra showed that despite the evident distortions of the spectra due to scattering and self-absorption, considerable information could be extracted that described the factory imprint and chemical quality of the crystal samples. Groups of samples related to the factories were found in PCA score plots, but classification of the samples according to factory was inferior to the corresponding classification results from sugar solutions. However, the three quality parameters, color, ash and α-amino-N, especially color, were predicted from PLS models of the crystal sample data with practically the same model error as obtained with the solution data (Nørgaard, 1995b). More PLS factors had to be used in the models of the raw fluorescence data from the crystal samples in comparison with the solution models reflecting the influence of scattering effects and self-absorption introduced in the

spectra from the crystal samples. Normalization of the spectra was able to reduce the number of PLS factors in the models by removing the sampling variance introduced by the differences in crystal packing and crystal sizes in the area detected by the instrument.

In the second study of crystal sample fluorescence, the relation between the spectral responses and the crystal sample composition was examined (PAPER V). The sample set consisted of five beet sugars, three from different factories, and two special sugars with high and low color. All the samples were sieved into five different crystal sizes and fluorescence of crystal samples and solutions was measured. The fluorescence data was modelled with PARAFAC (Fig. 4.2).

The results clearly demonstrated the strong correlation between the crystal size, the propagation of light in the crystal sample and the measured response. The resolved fluorophores from the transmission data are strongly influenced by self-absorption at the lowest wavelengths, i.e. the tyrosine fluorescence cannot be detected when compared to the solution data. However, the resolved fluorophores at higher wavelengths have a close resemblance with the solution fluorophores (Fig. 4.2b). The reflection data was difficult to model due to high background scattering from the uneven surface and the deep penetration of the incident light, and only the fluorescence of high intensity was well separated from the scattering in the resolved emission spectra (Fig. 4.2c).

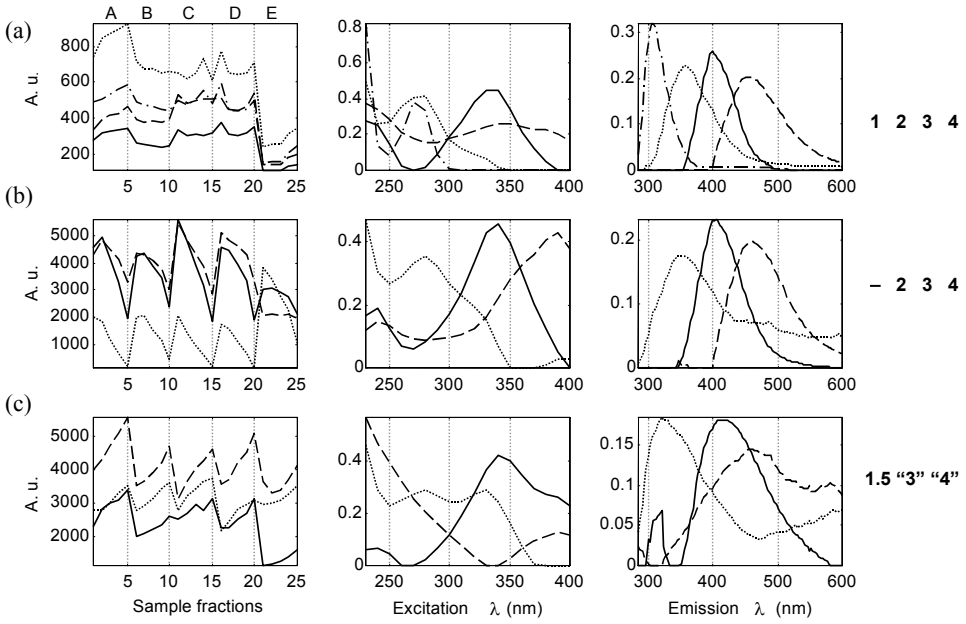


Fig. 4.2 PARAFAC models of the measured fluorescence landscapes of sieved sugar samples. (a) Solution data, (b) transmission data, and (c) reflection data. The emission spectra of the resolved fluorophores are compared by assigning numbers to the spectra as shown to the right in the figure. Components 1 and 2 are tyrosine and tryptophan components, respectively, component 3 is a colorant intermediate, and component 4 is a colorant fluorophore (PAPER V).

4.3 Color incorporation in the crystal

The incorporation of color impurities in sugar crystals has often been studied in sugar color research (Broughton *et al*., 1987; Grimsey and Herrington, 1996). Hence, a relationship between color distribution in the sugar crystal and the resolved fluorophores might be inferred. As Pollak theoretically determined in 1975 considering a solid sample system, fluorescence determinations in the transmission mode yield results that are almost independent from the distribution of the analyzed material with depth whereas fluorescence measurements in the reflection mode are strongly dependent on distribution of concentration. However, in the reflection case Pollak assumed little penetration in the sample depth, which is not the case in transparent sugar crystals, and therefore neither fluorescent technique can actually be used to determine the distribution of the fluorophores in sugar crystals.

5. Conclusions and perspectives

This project has employed fluorescence spectroscopy as a screening tool in cane and beet sugar processing to obtain chemical information about color impurities, their formation, and their influence on the composition of the sugar crystal. The chemical information was extracted from the multivariate spectral data with the use of bilinear and trilinear chemometric methods suited for highly co-linear spectral data.

A few but characteristic fluorophores were found in both beet and cane sugar processing, which simplified the spectral information as compared to absorption measurements. In the beet sugar process samples tyrosine and tryptophan were the only fluorescent color precursors that were chemically defined as simple compounds, and in cane sugar processing tryptophan alone was found as an individual fluorophore. The present results showed that the stable fluorescence properties make tryptophan a valuable indicator of color formation during sugar processing and in storage of sugar.

All the visible fluorophores found in beet sugar and raw cane sugar were colorants or colorant intermediates from complex color forming reactions in the process. No plant derived fluorescent colorants were found. The colorants displayed a continuous fluorescence that appeared at higher wavelengths when the color darkened. The colorant fluorescence was separated into individual fluorophores representing average components of fluorophores with slightly different wavelength positions. Colorant recognition using model colorants was rendered difficult due to the above-mentioned continuity in colorant fluorescence and the fact that the model colorant reactions had formed darker color than was present in the factory syrup samples used for comparison. The apparent differences found in colorant fluorophores from beet and cane sugar processing were ascribed to different process conditions rather than different chemical reactions.

Fluorescence on crystal sugar samples introduced distortions in the measured data but the chemometric methods were still able to extract relevant information about the sample characteristics. The prediction of color in beet sugar was obtained with a model error equaling the solution data demonstrating the possibility of a faster and competitive analysis method to be used directly in the production. The fluorescence spectra of sugar

crystals also contained information about the individual fluorophores, but the distribution of the fluorophores in the crystal could not be inferred.

It is important to bear in mind that the above conclusions are only based on the end product and one semi-product from beet sugar and raw cane sugar productions. It is known that color formation is already taking place in the initial process stages and that these early reactions can influence the composition of the crystal sugar end product. To fully exploit the fluorophores as color forming indicators, it is thus important to expand the fluorescence data with measurements of other process samples to observe the changes of fluorophores in the process. A thorough examination of fluorophores in beet sugar processing is currently under progress where the fluorescence landscapes of collected fractions of all the process samples from diffusion juice to molasses from one factory are measured. Also studies of fluorescence changes in raw cane factory samples and cane refinery samples to establish the color development in cane sugar processing is a future project.

A difficulty in this project has been the time-consuming measurement of the fluorescence landscapes, which was caused by the long scanning period of each fluorescence emission spectrum as well as the time delay when changing excitation wavelengths. Since this project has been about qualitative modelling of the fluorescence data, the focus has been on measuring spectra of high resolution and sensitivity from a limited number of samples. However, if the fluorophores should be used as chemical indicators to model other parameters, e.g. during a campaign, more samples will be needed to make more robust models. There exists faster instrumentation for fluorescence measurements, e.g. diode array fluorescence detectors and fast filter-instruments, but the former has usually low sensitivity, and the latter produce cruder landscapes with large wavelength intervals and consequently low resolution. Nevertheless, the applicability of fast filter instruments for at-line measurements of some quality parameters in the sugar industry has already been demonstrated and will be a potential in-line instrument in the future (Christensen and Nørgaard, 1999).

Fluorescence as a screening method for indicator substances in dynamic processes and experimental studies of composite samples will certainly benefit from developments in

the chemometric discipline. For instance, the premises of the PARAFAC model are rather strict, as discussed in Chapter 2, and it is difficult to obtain three-way data, which fully meets these conditions having non-ideal measurements conditions and samples. The development of model algorithms that are better in handling perturbations from screening instruments are already in progress, e.g. modelling retention time shifts in three-way chromatographic data (Kiers *et al*., 1999, Bro *et al*., 1999). Improvements of algorithms or data preprocessing tailored to handle non-spectral contributions to the measured fluorescence such as scatter and quenching will very likely increase the use of fluorescence as a screening analysis in the future.

References

Adhikari, H. R. and Tappel, A. L. (1973). Fluorescent products in a glucose-glycine browning reaction. *J. Food Sci.* **38**, 486-488.

Ajandouz, E. H. and Puigserver, A. (1999). Nonenzymatic browning reaction of essential amino acids: effect of pH on caramelization and Maillard reaction kinetics. *J. Agri. Food Chem.* **47**, 1786-1793.

Belitz, H. D. and Grosch, W. (1987). *Food Chemistry* (Springer Verlag, Berlin), pp. 213-225.

Bro, R. (1996). Multi-way calibration. Multi-linear PLS, *J. Chemom*. **10**, 47-62.

Bro, R. (1997). PARAFAC: Tutorial and applications. *Chemom. Intell. Lab. Systems* **38**, 149-171.

Bro, R. (1998) *Multi-way analysis in the food industry: models, algorithms and applications*. Ph.D. Thesis, University of Amsterdam, The Netherlands.

Bro, R. (1999). Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemom. Intell. Lab. Systems* **46**, 133-147.

Bro, R., Andersson, C. A. and Kiers, H. A. L. (1999). PARAFAC2 – Part II. Modelling chromatographic data with retention time shifts. *J. Chemom*. **13**, 295-309.

Broughton, N. W., Sargent, D., Houghton, B. J. and Sissons A. (1987). The inclusion of colour and ash components in U.K. beet white sugar. In *Comptes Rendus 18th Assemblée Générale Commission Internationale Technique Sucrerie*, Ferrara, Italy, 101-130.

Carpenter, F. G. and Wall, J. H. (1972). Fluorescence in commercial sugars. In *Proceedings of the 1972 Technical Session Cane Sugar Refining Research*; New Orleans, LA, 47-61.

Carpenter, F. G. and Roberts, E. J. (1974). Colorant formation under refining conditions. In *Proceedings 1974 Technical Session Cane Sugar Refining Research*; Cherry Hill, NJ, 106-115.

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

Christensen, J. and Nørgaard, L. (1999). On-line fluorescence spectroscopy and chemometrics for qualitative and quantitative analysis: application in sugar industry. *Spectr. Europe* **11**, 20-22.

Clarke, M. A., Edye, L. A. and Eggleston, G. (1997). Sucrose decomposition in aqueous solution, and losses in sugar manufacture and refining. *Adv. Carbohydr. Chem. Biochem*. **52**, 441-470.

DDS (1985). *Om sukker og sukkerfabrikation hos De Danske Sukkerfabrikker*. Aktieselskabet De Danske Sukkerfabrikker, Copenhagen.

Dills Jr, W. L. (1993). Protein fructosylation: fructose and the Maillard reaction. *Am. J. Clin. Nutr.* **58** (suppl.), 779S-787S.

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

Eggleston, G. (1998). Changes in chemistry across evaporators. *Int. Sugar J.* **100**, 49-52, 92-95.

Fennema, O. R. (1985). *Food Chemistry* (Marcel Dekker Inc., New York), p. 55-58.

Georgi, B. and Burba, M. (1977) Spezielles Filterfluorometer für die routinemässige Aminosäurebestimmung in Zuckerrüben und Zuckerfabriksprodukten. *Z. Zuckerind*. **27**, 650-654.

Godshall, M. A., Clarke, M. A., Dooley, C. D. and Blanco, R. S. (1991). Progress in beet sugar colorant research. *J. Sugar Beet Res.* **28**, 155-165.

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

Godshall, M.A and Chou, C. C. (2000). The nature of colorant. In *Proceedings of the 2000 Conference on Sugar Processing Research*, Porto, Portugal, in press.

Grimsey, I. M. and Herrington, T. M. (1996). The incorporation of colored compounds in sucrose crystals. *Zuckerind*. **121**, 40-45.

Gross, D. and Coombs, J. (1976). Enzymic colour formation in beet and cane juices. *Int. Sugar J*. **78**, 69-73 and 106-109.

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.

Hayase, F., Shibuya, T., Sato, J. and Yamamoto, M. (1996). Effects of oxygen and transition metals on the advanced Maillard reaction of proteins with glucose. *Biosci. Biotech. Biochem.* **60**, 1820-1825.

Hurtubise, R. J. (1981). Solid Surface Luminescence Analysis: Theory, Instrumentation, Applications (Marcel Dekker Inc., New York), pp. 1-3.

Höskuldsson, A. (1994). The H-principle: new ideas, algorithms and methods in applied mathematics and statistics. *Chemom. Intell. Lab. Systems* **23**, 1-28.

Kiers, H. A. L., Ten Berge, J. M. F. and Bro, R. (1999). PARAFAC2 – Part I. A direct fitting algorithm for the PARAFAC2 model. *J. Chemom*. **13**, 275-294.

Kortüm, G. (1969). *Reflectance Spectroscopy* (Springer-Verlag, Berlin).

Lakowicz, J. R. (1983). *Principles of Fluorescence Spectroscopy* (Plenum Press, New York), Chap. 11.

Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy* (Kulwer Academic/Plenum Publishers, New York), 2nd ed.

Leurgans, S. and Ross, R. T. (1992). Multilinear models: applications in spectroscopy. *Stat. Sci.* **7**, 289-319.

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

Mak, F. K., Kelly, F. H. C. and Gunaseelan, K. (1975). Thermal effects on crystal sugar. *Int. Sugar J.* **77**, 195-197.

Martens, H and Næs, T. (1993). *Multivariate Calibration* (Wiley, New York), 2nd ed.

Morales, F. J., Romero, C. and Jiménez-Pérez, S. (1996). Fluorescence associated with Maillard reaction in milk and milk-resembling systems. *Food Chemistry* **57**, 423-428.

Mosca, L., De Marco, C., Fontana, M. and Rosei, M. A. (1999). Fluorescence properties of melanins from opioid peptides. *Arch. Biochem. Biophys*. **371**, 63-69.

Munck, L. (1989). *Fluorescence analysis in foods* (Longman Scientific and Technical, Harlow).

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

Nursten, H. E. and O'Reilly, R. (1983). Colored compounds formed by the interaction of glycine and xylose. In: *The Maillard Reaction in Foods and Nutrition*, Waller, G. R. and Feather, M. S. (Ed.), ACS Symp. Ser. 215, ACS, Washington D. C., pp. 103-121.

Nørgaard, L. (1995a). A multivariate chemometric approach to fluorescence spectroscopy. *Talanta* **42**, 1305-1324.

Nørgaard, L. (1995b). Classification and prediction of quality and process parameters of thick juice and beet sugar by fluorescence spectroscopy and chemometrics. *Zuckerind*. **120**, 970-981.

Papageorgiou, J., Doherty, W. O. S., Walker, B. and Edye, L. A. (1999). Analysis of organic impurities in raw sugar by a HPLC-UV-mass spectrometer. *Zuckerind*. **124**, 133-138.

Parker, K. J. (1970). Chemical problems in the sucrose industry. *Sucr. Belge* **89**, 119-126.

Pollak, V. (1975). Effect of non-uniform concentration distribution with depth upon quantitative optical analysis of chromatograms. *J. Chrom*. **133**, 195-198.

Pongor, S., Ulrich, P., Bencsath, F. A., and Cerami, A. (1984). Aging of proteins: isolation and identification of a fluorescent chromophore from the reaction of polypeptides with glucose. *Proc. Natl. Acad. Sci. USA* **81**, 2684-2688.

Prey, V. and Andres, H. (1971). Nebenreaktionen der Saftreinigung. *Z. Zuckerind*. **21**, 267-271.

Rettig, W. (1999). *Applied fluorescence in chemistry, biology and medicine* (Springer-Verlag New York).

Robinson, G. M. and Smyth, M. R. (1997). Simultaneous determination of products and intermediates of L-dopa oxidation using capillary electrophoresis with diode-array detection. *Analyst* **122**, 797-802.

Schallenberger, R. S. and G. G. Birch (1975). Nonenzymatic browning reactions. In *Sugar Chemistry* (Avi Publishing Company, Inc.), pp. 169-193.

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.

Schiweck, H. and Clarke, M. (1994). Sugar. In *Ullmann's Encyclopedia of Industrial Chemistry* (VCH Verlagsgesellschaft, Weinheim, Germany), Vol. A 25.

Shore, M., Broughton, N. W., Dutton, J. V. and Sissons, A. (1984). Factors affecting white sugar colour. *Sugar Technol. Rev.* **12**, 1-99.

Vukov, K., Hangyál, K. and Bara-Ányos, É. (1984). Über die Verfärbung der Säfte durch phenolische Substanzen in einigen Zuckerfabriken. *Zuckerind*. **109**, 1089-1092.

Winstrøm-Olsen, B., Madsen, R. F. and Kofod Nielsen, W. (1979). Investigation of phenolic compounds from sugar beet in relation to the formation of colour. *Int. Sugar J.* **81**, 362-367.

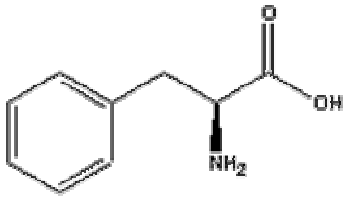
Wolfbeis, O. S. (1985). The fluorescence of organic natural products. In: *Molecular Luminescence Spectroscopy: Methods and Applications – Part I*, Schulman, S. G. (Ed.), John Wiley and Sons, New York, pp. 167-370.

Yang, J., Zhao, W., Jie, N., Zou, H. and Ren, X. (1993). Rare earth as fluorescence probe to determine 3,4-dihydroxyphenylalanine. *Anal. Lett*. **26**, 2291-2306.

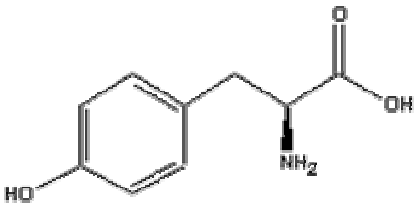
Appendix A

Fluorescent color precursors

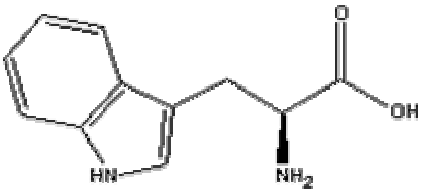
Chemical structures and fluorescence excitation and emission maxima of several color precursor fluorophore candidates in the sugar processes (Lakowicz, 1983; Wolfbeis, 1985). Ex/Em means the excitation and emission maxima of the fluorescence in water at pH 7.



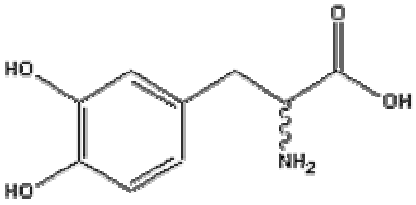
Phenylalanine, Ex/Em = 255/285 nm



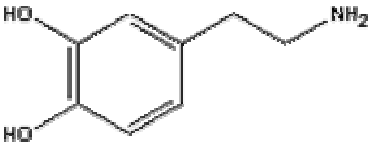
Tyrosine, Ex/Em = 275/305 nm



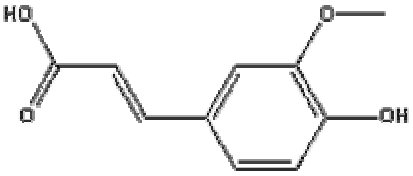
Tryptophan, Ex/Em = 275/360 nm



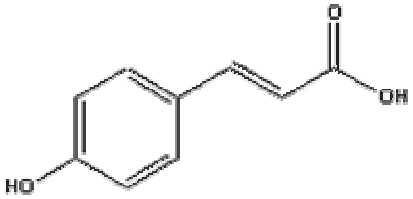
Dopa (3,4-dihydroxy-phenylalanine), Ex/Em = 280/325 nm



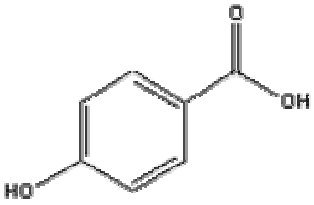
Dopamine (3,4-dihydroxyphenethylamine), Em = 280/325 nm



Ferulic acid, Ex/Em = 345/450 nm



4-Coumaric acid, Ex/Em = 340/450 nm



4-Hydroxybenzoic acid, Ex/Em = 290/330 nm

Appendix B

Time study of formation of Maillard reaction products using L-dopa and glucose

A glucose solution (1%, w/w) was mixed with a 6 x 10^{-5} M L-DOPA solution in the ratio 2:1. The mixture was adjusted to pH 9 and heated to 100 °C. During the development of color, samples were taken from the mixture after 0, 11 and 22 min and the temperature was 45, 90 and 95 °C, respectively. Fluorescence landscapes were measured using a Perkin-Elmer LS50 B fluorescence spectrometer with excitation wavelengths between 230-420 nm with 10 nm intervals and emission wavelengths between 288-606 nm. Excitation and emission monochromator slit widths were set to 10 nm, respectively, and the scan speed was 1500 nm/min. The samples were measured without dilution. A three-component model was made from the fluorescence data using the multi-way decomposition model PARAFAC (Harshman, 1970; Bro, 1997). In Fig. B.1 the resolved sample concentration profiles and excitation and emission spectra of the three components are shown. Component 1 (---) is L-dopa with 280/325 nm as excitation/emission maximum. Component 2 (—) has 330/400 and component 3 (– –) has 380/480 nm as excitation/emission maxima. The sample concentration profiles show that at 0 min only L-dopa fluorescence is measured. After 11 min the L-dopa fluorescence has almost disappeared whereas the fluorescence of component 2 and component 3 is increasing. After 22 min L-dopa is completely used up and the fluorescence of component 3 is still increasing. However, the fluorescence of component 2 is decreasing. The results show that component 2 and component 3 are Maillard reaction products and that L-dopa is quickly used up in the reaction. The fact that the fluorescence of component 2 decreases while the fluorescence of component 3 continues to increase suggests that component 2 is an intermediate used in the formation of component 3.

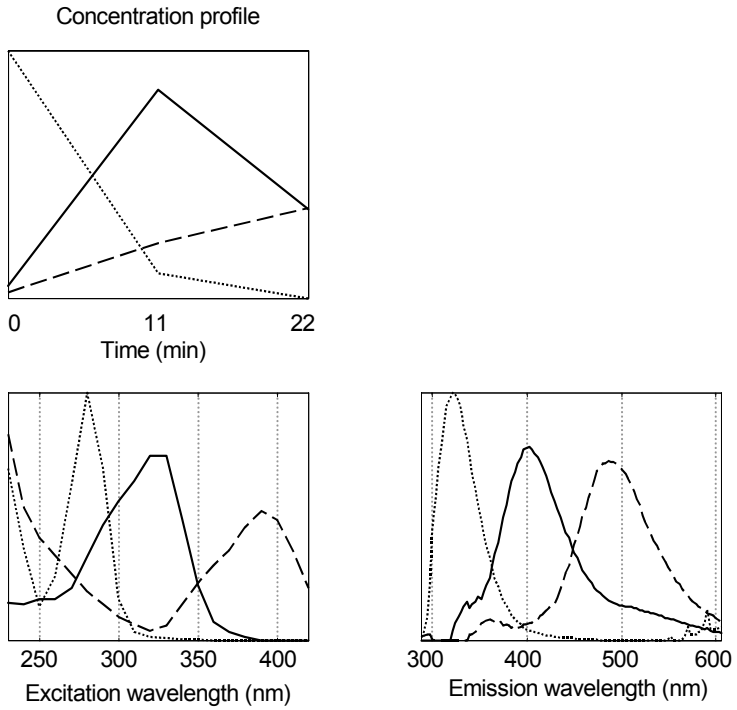


Fig. B.1 A three-component PARAFAC model of fluorescence landscapes of three samples taken from a Maillard reaction mixture of L-dopa and D-glucose at 0, 11 and 22 min. The time-dependent concentration profiles are shown together with the excitation and emission spectra of the three resolved components. Component 1 (---), component 2 (—) and component 3 (– –). All profiles have been normalized to unit length.

Paper I

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

D. Baunsgaard, C. A. Andersson, A. Arndal and L. Munck

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 and 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 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 and 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 (Reinefeld, Schneider, Westphal, Tesch & Knackstedt, 1973; Madsen, Kofod Nielsen, Winstrøm-Olsen & Nielsen, 1978a; 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 mg/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 x 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 (Harshman, 1970; Carroll and Chang, 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

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

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 ith 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 and Chang, 1970) and (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 and 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 Equation (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

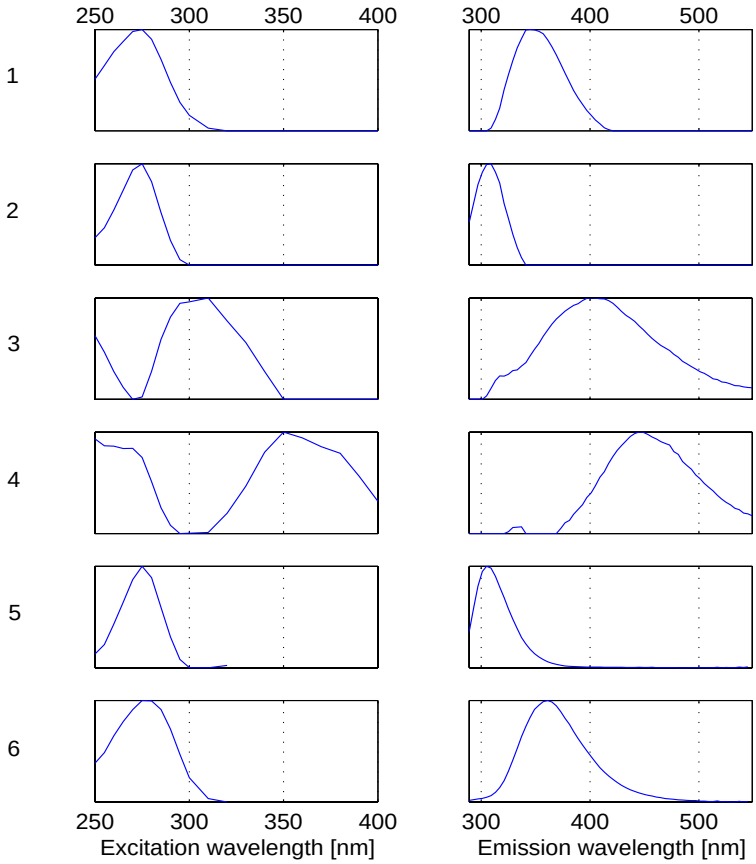


Fig. 1. The results of a four-component CP model of fluorescence landscapes of five beet sugar samples. Row 1-4 contain the excitation and emission spectra of the four resolved components. The left column shows the excitation profiles and the right column shows the emission profiles. Row 5 and 6 show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

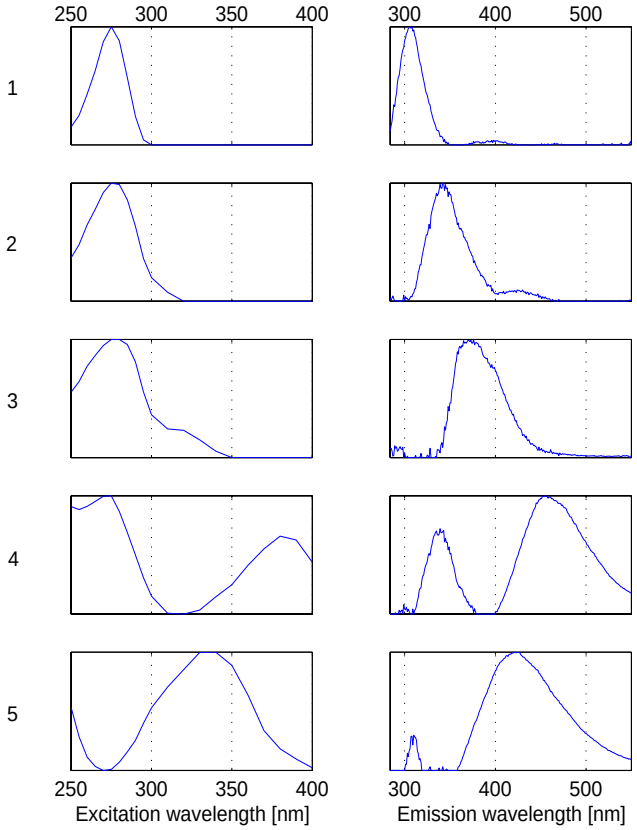


Fig. 2. A five-component CP model of fluorescence landscapes of ten thick juice samples. The left column shows the resolved excitation profiles and the right column shows the resolved emission profiles of the five components. All profiles have been normalised to unit length.

emission spectra of measured fluorophores as well as a sample profile relating the concentration of each fluorophore in the samples measured. Fig. 1 (row 1-4) and in 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 artifacts such as extra bands in the emission spectra, e.g. the emission spectrum of component 4 in Fig. 2. The excitation (1^{st} excited state) and emission wavelength maxima of the spectra in Fig. 1 and Fig. 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 row 5 and row 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.

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

Component[a]	Sugar λ_{max} (nm)		Thick juice λ_{max} (nm)		HPLC fractions of thick juice λ_{max} (nm)	
	Ex.	Em.	Ex.	Em.	Ex.	Em.
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 Fig. 1, 2 and 6 for each of the three CP models.

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 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 min 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.

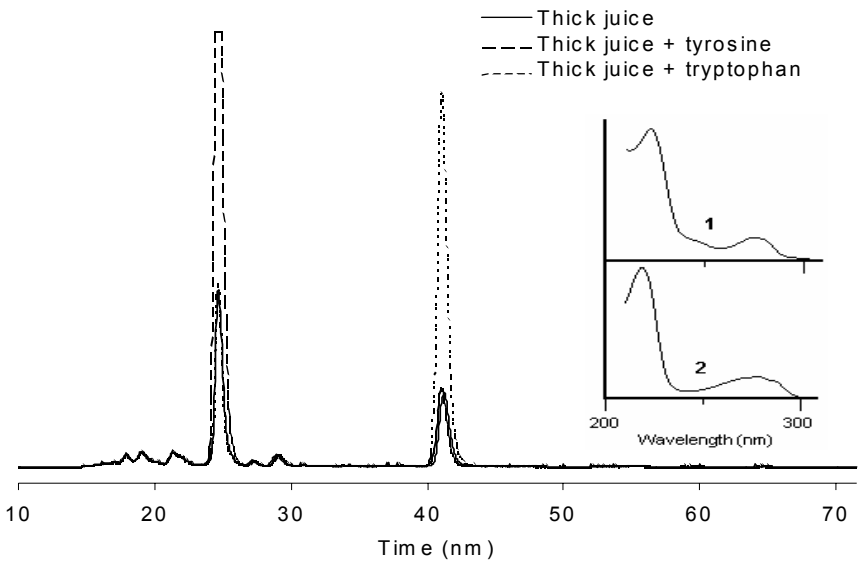


Fig. 3. HPLC size exclusion chromatograms with qualitative standard additions of tyrosine (1) and tryptophan (2) to thick juice monitored by fluorescence detection at 285/325 nm. The chromatograms verify the expected presence of tyrosine and tryptophan. The corresponding UV/VIS absorbance spectra from 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 nm and 420 nm, whereas the lower chromatogram is from the fluorescence detector at 280/325 nm. 420 nm is the normal

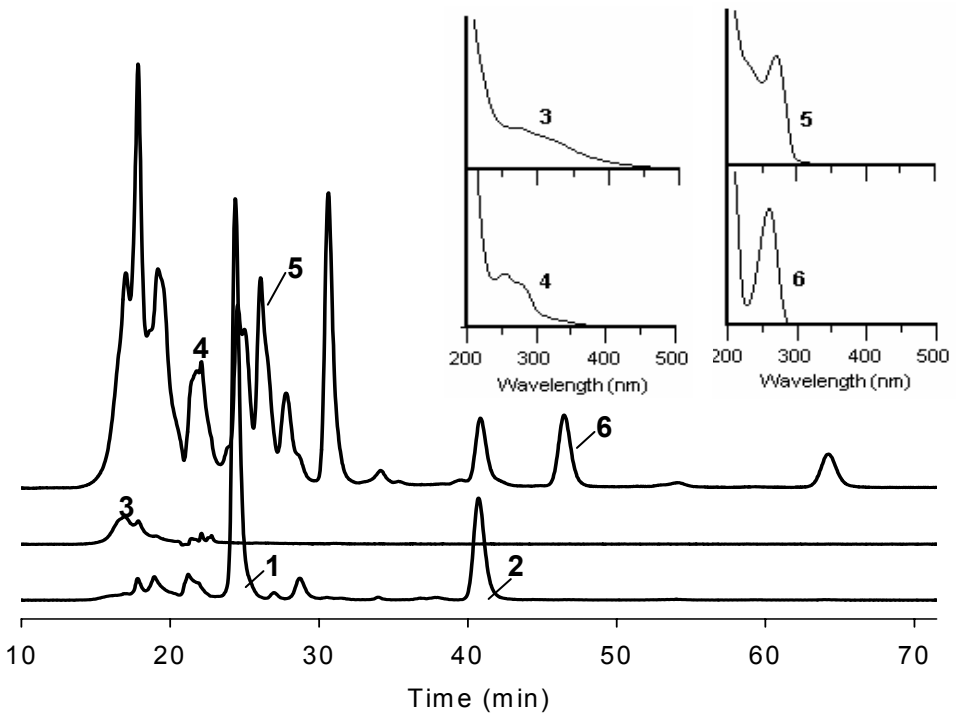


Fig. 4. HPLC size exclusion chromatograms of a thick juice sample. Upper curve: UV detection at 280 nm; middle curve: UV detection at 420 nm; bottom curve: fluorescence detection at 280/325 nm. Diode-array spectra of selected peaks (4-6) are also displayed. The spectra of peak 1 and 2 are identical with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

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 15-25 min.
The estimated size exclusion range was approx. 12-25 min, which means that the colorants are smaller than 80 kDa 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 peak 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. Peak 4, 5 and 6 display different spectra mainly absorbing at 280-300 nm whereas peak 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 3-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.

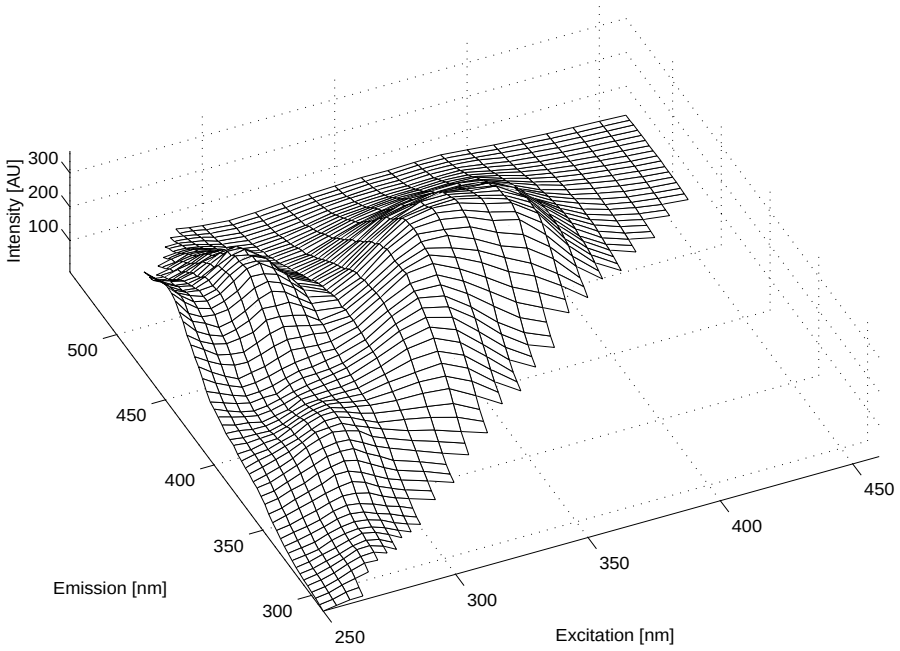


Fig. 5. An example of a fluorescence landscape from HPLC fraction no. 9 (22-23.5 min) of thick juice measured with 31 excitation wavelengths (230-460 nm) and 431 emission wavelengths (288-700 nm). The blank regions hold Rayleigh scatter signals and are thus treated as missing values.

Component 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 (Fig. 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 fraction 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.

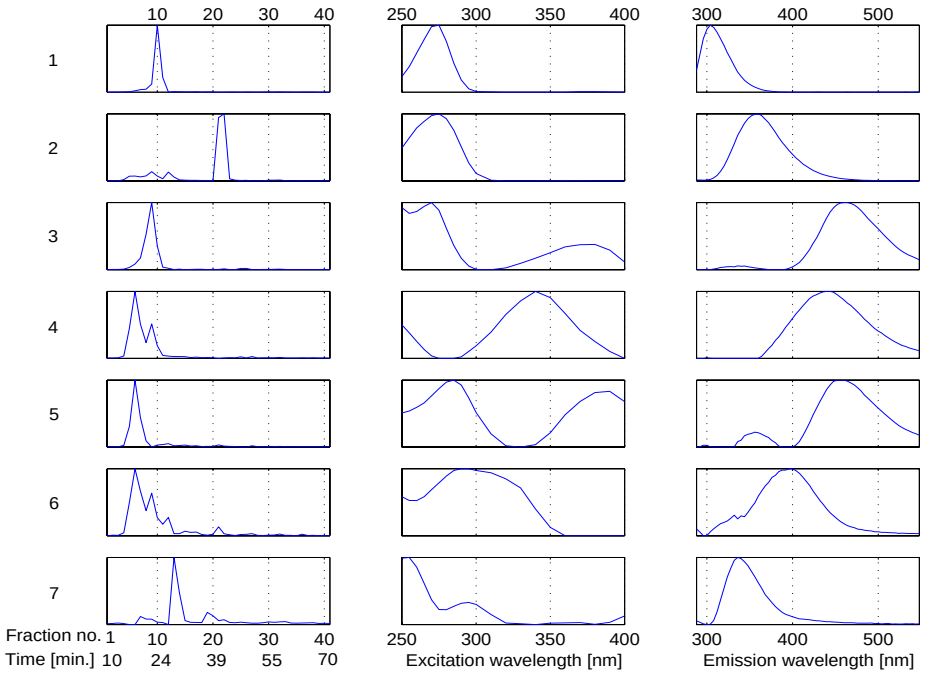


Fig. 6. The results of a seven-component CP model of the measured fluorescence landscapes of 41 collected HPLC fractions (10-71.5 min). The left-most column holds the chromatographic profiles, which show the concentration of each of the seven fluorescent components in the fractions. The centre column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

Component 3, 4 and 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, 1978b). All this suggests that some or all of component 3, 4 and 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 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 ten 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., Clarke, M. A., Dooley, C. D., & Blanco, R. S. (1991). Progress in beet sugar colorant research. *Journal of Sugar Beet Research*, *28*, 155-165.

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.

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* (p. 347). New York: Plenum Press.

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. (1978a). 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. (1978b). 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, (Ed.), *Molecular Luminescence Spectroscopy: Methods and Applications – Part I* (pp. 167-370). New York: John Wiley and Sons.

Paper II

Fluorescence of raw cane sugars evaluated by chemometrics
D. Baunsgaard, L. Nørgaard and M. A. Godshall

Abstract
In a fluorescence study of raw cane sugar samples, two-way and three-way chemometric methods have been used to extract information about the individual fluorophores in the sugar from fluorescence excitation-emission landscapes. A sample set of 47 raw sugar samples representing a varied selection was analyzed, and three individual fluorophores with (275,350) nm, (340,420) nm and (390,460) nm as their approximate excitation and emission maxima were found. The spectral profiles of the fluorophores were estimated with the three-way decomposition model PARAFAC. Two-way principal component analysis (PCA) of unfolded fluorescence landscapes confirmed the PARAFAC results and showed patterns of samples related to time of storage. Partial least squares (PLS) calibration models of color at 420 nm had a high model error due to the very high color range of the raw sugars, but variable selection performed on the fluorescence data revealed that all three fluorophores were correlated to color. The (275,350) nm fluorophore is considered as a color precursor to the color developed on storage and the (340,420) nm and (390,460) nm fluorophores show colorant polymer characteristics.

Keywords: *Raw cane sugar; color; fluorescence; principal component analysis; partial least squares; multiway decomposition*

Introduction

It has been known for many years that commercial sugars exhibit characteristic fluorescence, which can be used to obtain information of minor constituents in the sugar. Carpenter and Wall (1972) described the fluorescence of several raw sugars, raw sugar molasses, and sugar refinery samples. Contour charts of fluorescence landscapes with several excitation and emission wavelengths were used to inspect the fluorescence emission peak pattern in the different process samples. Four peaks were often repeated: (360,430) nm, (280,320) nm, (250,430) nm and (400,600) nm, describing the position of the peaks (max. excitation wavelength, max. emission wavelength). They concluded that the fluorescence measurements seemed very informative, but the contour plots seemed to be very complex with several components in some of the peaks.
In recent years spectrofluorometry has successfully been applied to the beet sugar manufacturing process with the use of multivariate data analysis (Munck et al., 1998). The same approach with multiple excitation and emission wavelengths used by Carpenter and Wall (1972) has been employed, but chemometric evaluation of the excitation-emission landscapes is used to extract the relevant information from the data. In a study of beet sugar samples it was possible to classify white sugar samples according to factory and to predict quality parameters such as α-amino-nitrogen, color and ash from fluorescence data of these samples (Nørgaard, 1995). The fluorescence

data of thick juice samples showed more ambiguous results due to the more complex sample composition. Another study of beet sugar samples utilized the three-dimensional structure of the fluorescence excitation-emission landscapes to resolve spectral excitation and emission profiles of fluorophores in sugar with a multi-way chemometric model, PARAFAC (Bro, 1999). Four fluorescent components were found to capture the variation in the fluorescence data of 268 sugar samples collected from a beet sugar factory during a campaign, where two of them showed spectra with a close similarity to the pure fluorescence spectra of the amino acids, tyrosine and tryptophan. The concentrations of the four components estimated from the sugar samples could be correlated to several quality and process parameters, and they were characterized as potential indicator substances of the chemistry in the sugar process. A recent paper has confirmed these findings using HPLC analysis combined with fluorescence measurements on thick juice samples and evaluated by PARAFAC (Baunsgaard et al., 2000). Seven fluorophores were resolved from thick juice. Apart from tyrosine and tryptophan, four of the fluorophores were identified as high molecular weight compounds, which were related to colorants absorbing at 420 nm. Three of the high molecular weight compounds were found to be possible Maillard reaction polymers. The last of the seven fluorophores indicated a compound with polyphenolic characteristics.

The studies of beet sugar sample fluorescence using chemometric analyses have contributed new information, which may help in the understanding of the chemistry taking place during the manufacturing of beet sugar. Cane and beet sugar production, though with origin in very different plant material, share many production related chemical reactions especially in the development of colorants (Godshall, 1996). The results of raw cane sugar fluorescence by Carpenter and Wall (1972) suggest that similar use of chemometric methods on cane sugar process samples could provide additional knowledge of the chemistry in cane sugar processing.

In this work, 47 raw cane sugar samples were selected as a data set. The data set represented a very wide selection of raw sugars, where few of the samples shared the same origin or production year. Some of the samples had been stored for many years and had darkened during the years due to the formation of additional color. Thus, the data set should amply span the variation in produced raw sugars as normally encountered in the cane sugar industry. From the excitation-emission fluorescence landscapes measured on all the samples, the systematic variation of the fluorescence in the samples was extracted with the use of various two-way and three-way chemometric methods like Principal Component Analysis (PCA), Partial Least Squares Regression (PLSR), Principal Variables (PV), and Parallel Factor Analysis (PARAFAC). These methods are well established as statistical methods for the analysis of spectral and highly co-linear data structures (Martens and Næs 1993; Bro, 1997). The information thus obtained by these methods was used to characterize the various fluorophores in raw sugar and to reveal patterns in the sample set relating the fluorescence to the chemical composition of the samples, especially with regard to color in raw sugar.

Materials and methods

Samples.

A sample set of 47 raw cane sugars was collected representing many different locations and campaign years (Table 1). Some of the samples represent raw sugars with special characteristics. Three of the samples (1-3 in Table 1) had been stored for many years and had darkened during the years due to the formation of more color. Sample 3 was stored cold and therefore developed less color relatively than 1 and 2. Samples 32 and 45 (the same sugar as 27) are lab washed raw sugars where the outer coating of color has been washed off. Samples 46 and 47 have elongated crystals due to a high amount of polysaccharides. The characteristics of these special sugars are related to color in raw sugar (Ravelo B. et al., 1991; Godshall, 1996), and they are used to introduce a larger variation of color in the sample set.

Preliminarily, several of the sugar samples were selected as representative samples and were diluted in ion-exchanged water to different levels of concentration to determine the concentration quenching of the measured fluorescence. A concentration of 9.4 mg/ml was chosen as the concentration level to be used in the fluorescence measurements of all the sugar samples for the chemometric analysis. This concentration showed the least concentration quenching of the fluorescence by inspection of the landscapes combined with an acceptable signal/noise ratio for the purest samples.

TABLE 1. The raw sugar sample set

#	Origin	Year	Color	#	Origin	Year	Color
1	Argentina[a]	1963	37630	25	Nicaragua	1991	6270
2	Florida, USA[a]	1964	51290	26	Panama	1991	4370
3	Louisiana, USA[b]	1968	10020	27	Peru	1991	3830
4	Australia	1977	6740	28	Queensland, Australia	1991	3170
5	South Africa	1979	4290	29	Queensland, Australia	1991	3250
6	Barbados	1984	17460	30	Costa Rica	1992	4490
7	Brazil	1984	10250	31	Guyana	1992	3140
8	Jamaica	1984	8470	32	Louisiana, USA[d]	1992	1470
9	Trinidad	1984	10350	33	Louisiana, USA	1992	2950
10	Brazil	1985	12170	34	Mexico	1992	910
11	Louisiana, USA	1985	5800	35	Panama	1992	2920
12	Dominican Rep.	1986	6580	36	Texas	1996	2210
13	Louisiana, USA	1986	3950	37	Florida, USA	1997	3320
14	Mauritius	1986	4810	38	Louisiana, USA	1997	2360
15	Louisiana, USA	1987	5180	39	Thailand	1997	7510
16	Texas, USA	1987	5880	40	Hawaii, USA	1998	1270
17	Hawaii, USA	1989	3850	41	Costa Rica	1998	2460
18	Hawaii, USA	1989	4560	42	Louisiana, USA	1998	2970
19	Bolivia[c]	1991	340	43	Philippines	1998	5770
20	Cuba	1991	4790	44	Taiwan	1998	2190
21	Dominican Rep.	1991	9400	45	Peru[d]	1991	1750
22	Ecuador	1991	1650	46	Cuba[e]	-	3290
23	Honduras	1991	8300	47	Cuba[e]	-	2570
24	Jamaica	1991	7440				

[a] Stored at room temperature.
[b] Stored cold.
[c] Very light raw sugar.
[d] Lab washed, i.e. the outer coating of color washed off.
[e] Elongated crystals.

Reference color measurements.
The quality parameter color (as the extinction coefficient at 420 nm absorbance) of the raw sugar samples was determined according to the ICUMSA method for raw sugars (ICUMSA Methods Book, 1994). The color values are shown in Table 1.

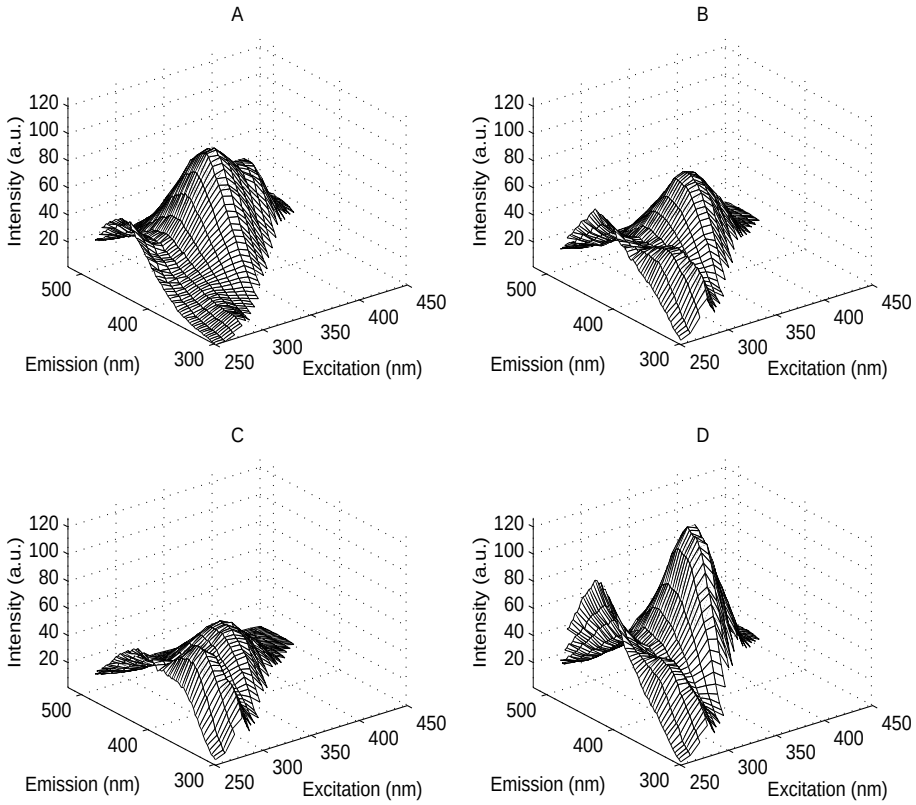


Figure 1. Plots of the excitation-emission fluorescence landscapes of four raw sugar samples from Table 1, which are each very different from one another. A. Sample 1; B. sample 20; C. sample 42; D. sample 43. The white areas in the landscape denote missing data areas due to Rayleigh scattering and other measured areas not conforming to true fluorescence.

Fluorescence measurements.
The fluorescence measurements were performed on a Perkin-Elmer LS50 B fluorescence spectrometer. Non-smoothed emission spectra were recorded at 21 excitation wavelengths in the range 250-450 nm with a 10 nm interval. The emission range was 298-520 nm. Fig. 1 shows plots of the fluorescence landscapes of four of the raw sugar samples, which are very different from one another. Some of the areas in the landscapes do not conform to true fluorescence response, such as the Rayleigh scattering peaks, and they are handled as missing intensity areas, i.e. the white areas in the plot.

Two-way data analysis.

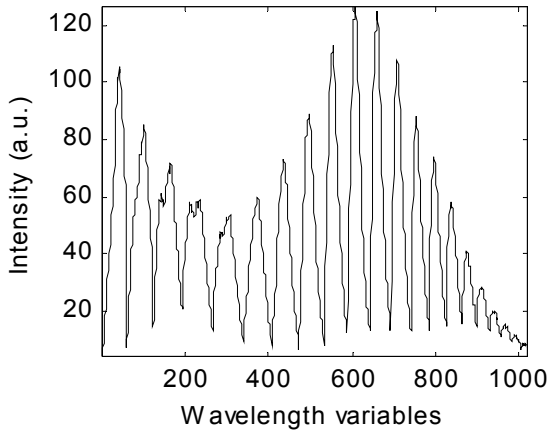


Figure 2. An example of an unfolded fluorescence landscape of a raw sugar sample (sample 43, Fig. 1D). Emission spectra are arranged in the order of the 21 excitation wavelengths (in total 1021 wavelength variables).

A two-way structure of the fluorescence landscapes (samples x wavelengths) is obtained by unfolding the landscape of each sample so that the emission spectra are arranged in the order of the 21 excitation wavelengths. An unfolded landscape is shown in Fig. 2. The missing areas shown in Fig. 1 have been removed. The two-way method principal component analysis (PCA) is used to find the principal directions of variation in the fluorescence data (Wold et al., 1987; Martens and Næs, 1993). For each principal component, a loading common for all the samples is extracted from the unfolded fluorescence data where the scores reflect the contribution of that loading in each sugar sample. Another two-way method, partial least squares regression (PLS), is used to make predictions of the quality parameter color from the unfolded fluorescence data (Höskuldsson, 1988; Martens and Næs, 1993). Full cross validation is used, i.e. one sample is predicted at a time from a calibration model consisting of the rest of the samples, because the very different samples in the data set makes it difficult to chose larger representative subsets for validation. The total prediction error of the model is based on all the individual model predictions for the optimal number of PLS components and it is expressed as the root mean square of cross validation (RMSECV):

$$\mathrm{RMSECV} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(C_n^{\mathrm{Predicted}} - C_n^{\mathrm{Reference}})^2}$$

where $C_n^{\text{Predicted}}$ is the estimated color, $C_n^{\text{Reference}}$ is the measured ICUMSA color and N is the number of samples. RMSECV is given directly as the prediction error of color in ICUMSA units. All PCA and PLS models are based on mean centered data.
The principal variables method (PV) is used to select excitation-emission wavelength pairs, which describe as much of the total variance in the data set as possible in relation to color (Höskuldsson, 1994). Short mathematical descriptions of the three two-way methods are found in Nørgaard (1995).

Three-way data analysis.
Having a three-way structure such as fluorescence landscape data from several samples (samples x excitation wavelengths x emission wavelengths), it can be beneficial to maintain the three-dimensional form while performing the data analysis. The models obtained from three-way data analysis may turn out to be more robust, easier to interpret and more predictive than their unfolded counterpart (Bro, 1998). The multi-way model PARAFAC (Harshman, 1970) fulfills the demand of easy interpretation, because the model decomposes three-way fluorescence data into spectral excitation and emission profiles of fluorophores in the samples. The spectra can be used to identify constituents in the raw sugar samples. Bro (1997) provides a thorough tutorial of the PARAFAC model. All the PARAFAC models of the raw sugar fluorescence data were estimated under a non-negativity constraint to improve the interpretability of the resolved profiles. In the model results presented here, the emission profiles have been estimated under a uni-modality constraint to avoid the interference of artificial extra peaks in the spectra due to too many missing variables in the data set.

Software for the data analyses.
Calculations were performed with Matlab for Windows version 5.3 (The MathWorks, Inc.) and Unscrambler version 7.01 (CAMO ASA). The implementation of the PARAFAC model was obtained from The N-way Toolbox for MATLAB (Andersson and Bro, 1999).

Results and discussion

Qualitative analysis of the raw sugars.
PCA of the fluorescence data is used to establish some common relations between the sugar samples based on the fluorescence information. In Fig. 3A a score plot of the first principal component (PC1) against the second principle component (PC2) of all the 47 samples is shown. The numbers in the plot correspond to the samples numbers in Table 1. The two components explain 96% of the fluorescence variation, 85% and 11%, respectively. The variation explained by PC1 is normally due to the differences in spectral intensity, and the fluorescence landscapes of sample 6 and 19 confirm that they are the two samples with the highest and lowest overall intensity, respectively. Sample 19 is remarkably low in color for a raw sugar (Table 1) and the two lab washed samples (32 and 45), which also have low color values, are situated in the same end of the PC1. There are no other patterns of the samples that seem to be related to PC1, neither in terms of the place of origin nor the year of production. The distribution of the samples

along PC2 shows a pattern related to the age of the samples. This is shown in Fig. 3B where the samples are marked with the year of production. The samples from the sixties are situated in the lower part of the plot and the samples from 1996-1998 are all situated in the upper part of the plot. The samples are not ordered completely by age along PC2, but there is clearly a trend. Higher order PC's were also examined, but they did not reveal any conclusive patterns except that sample 1 was singled out in the PC3 direction. The loadings of the PCA model contain the information about which wavelengths of the fluorescence data that are important for each of the principal components. The loading vectors of PC1 and PC2 are shown in Fig. 4. The most important emission wavelengths for PC1 are approximately 420-430 nm when excited at 340-360 nm. Sample 6 has a high fluorescence contribution from these wavelengths. The excitation and emission wavelengths, which are important contributors for PC2, are centered around 270 nm and 350 nm, respectively. Sample 42 has a particularly intense fluorescence in that area, whereas sample 1 has a low contribution. Thus, the (270,350) nm fluorescence seems to be connected with the changes in the raw sugar fluorescence related to the time of storage and the fluorescence round (350,425) nm is the dominating fluorescence in the raw sugar samples at the chosen concentration.

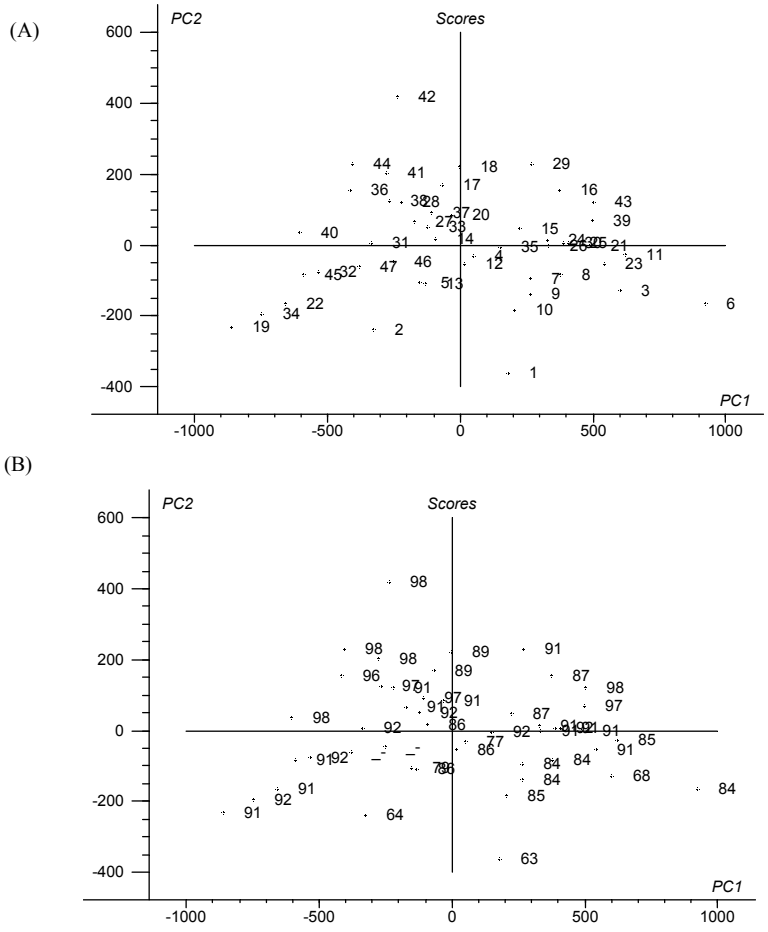


Figure 3. (A) Score plot of the first principal component (PC1) against the second principal component (PC2) of all 47 raw sugars. Numbers correspond to the raw sugars in Table 1. (B) Same score plot as in A, showing the year of production of each of the 47 raw sugars.

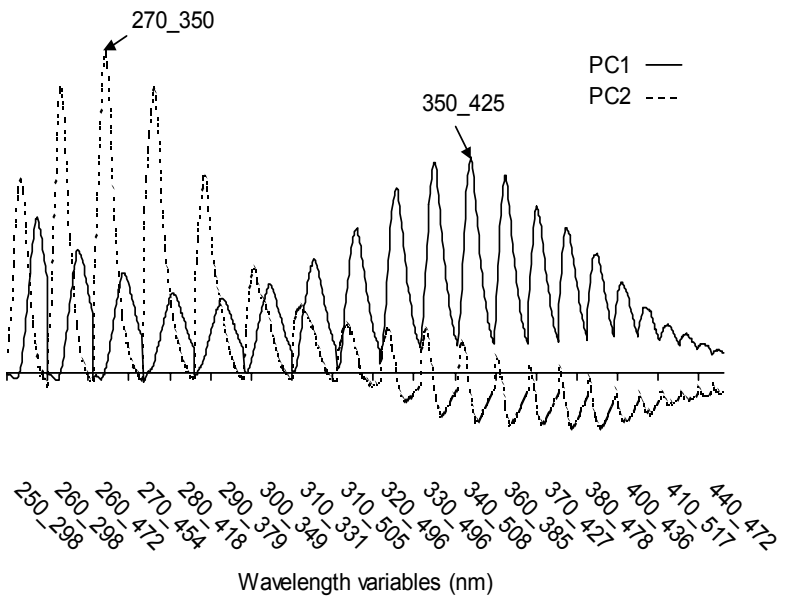


Figure 4. Loading vectors of the first two principal components of the 47 raw sugars as a function of wavelength variables. The wavelength variables are shown as excitation-wavelength_emission-wavelength.

Resolving specific fluorophores by PARAFAC.

It is difficult to extract spectral information of individual fluorophores from the PCA loadings due to the unfolded structure of the fluorescence data. Instead, the three-way model PARAFAC was used to estimate excitation and emission profiles of fluorophores directly from the three-dimensional fluorescence landscapes. PARAFAC may be regarded as a three-way PCA with scores and loadings, but now there are two loadings for each extracted component (excitation and emission profiles) and the fluorescence data used is the raw data, not mean centered as in the PCA model.

PARAFAC models of the raw sugar fluorescence data were estimated with one to six components, but the four-component was chosen as the best model based on split half analysis validation (Bro, 1998). The scores and loadings of the model are shown in Fig. 5. The first column presents the scores of samples for each component, the middle column presents the estimated excitation profiles and the last column presents the emission profiles. The excitation and emission maxima of the spectral profiles are shown in Table 2. The maxima of component I and component III are close to the wavelengths that were important for the two first principal components in the PCA. In Fig. 6A a score plot of component I versus component III is shown and the distribution of the samples is very close to Fig. 3A. This confirms that the PCA and PARAFAC models have captured the same two major contributors to the fluorescence in raw sugar. Component I and component III are also recognized as the (280,320) and (360,430) peaks reported by Carpenter and Wall (1972). The last two components in the PARAFAC model seem to

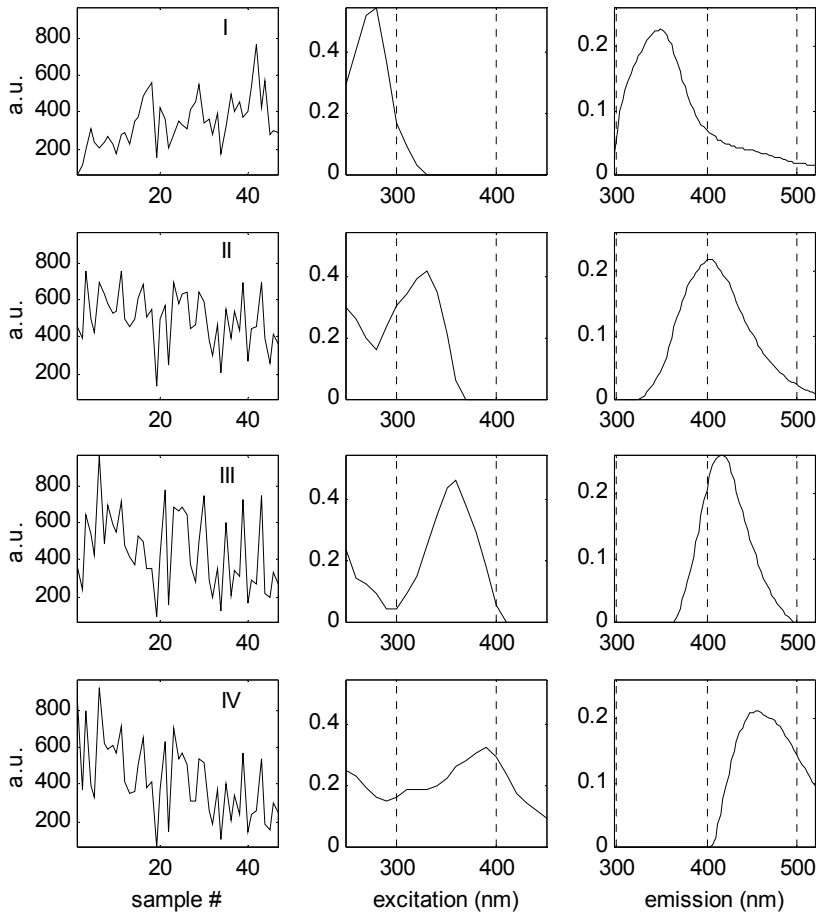


Figure 5. The results of a four-component PARAFAC model of the measured fluorescence landscapes of 47 raw sugar samples. The first column represents the scores of each sample for each component; the second column represents the excitation profile; the third column represents the emission profile. The spectral profiles are normalized, so that all variance is kept in the sample scores.

be correlated to component III. Plotting the scores of the component III versus component IV (Fig. 6B) reveal that the oldest samples (1-3) have a higher contribution from component IV relatively than from component III. Component IV is therefore considered as the component correlating to color development by storage of raw sugars. This agrees with the fact that component IV fluoresces at the highest wavelengths of all the components, which could signify the fluorescence of color polymers with a growing fluorochromic structure. The plot of component II versus component III (plot not shown) has a slight non-linear relationship where the samples with low fluorescence intensity

(e.g. sample 19) apparently have a higher contribution from component II and samples with high fluorescence intensity have a higher contribution from component III (e.g. sample 6). The two components have very similar spectra in Fig. 5 and they may represent two colorants of the same type. The fluorescence of the darker raw sugars samples tends to appear at somewhat higher wavelengths than the fluorescence of the lighter colorants and this is caused either by small size differences in the fluorochromic structure of a colorant polymer (increasing polymerization as color darkens) or self absorption of the fluorescence emission of the darker samples due to high absorbance of the colorants in the 400-420 nm area. Small wavelength shifts for differently colored samples will induce the PARAFAC model to resolve two representative fluorophores such as component II and component III of the many overlapping but slightly different excitation and emission spectra.

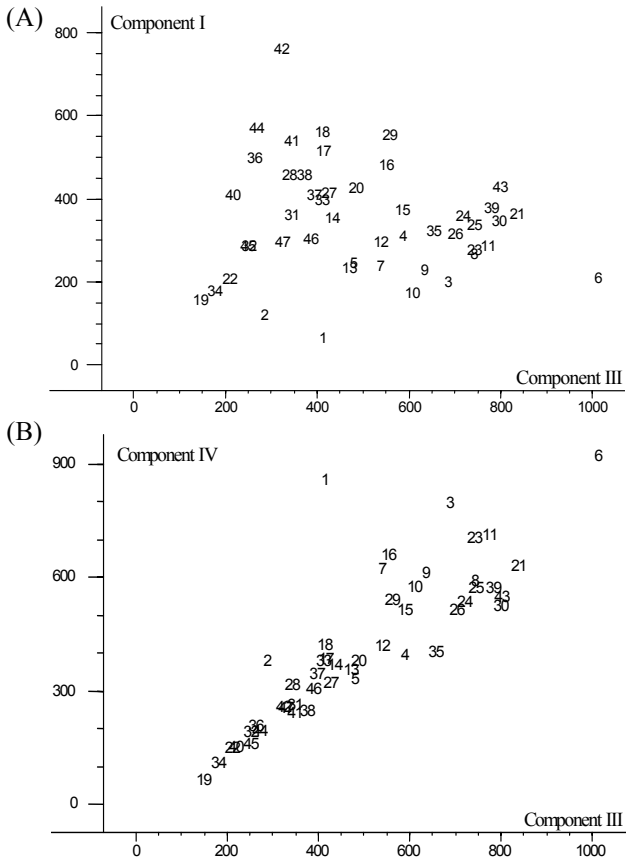


Figure 6. (A) Score plot of component I versus component III from the PARAFAC model. Note the similarity to the plot in Fig. 3A. (B) Score plot of component III versus component IV from the PARAFAC model.

TABLE 2. The excitation and emission maxima of four fluorescence components estimated in the parafac model of the 47 raw sugar samples

Component	λ_{max} (nm)	
	Ex.	Em.
I	275	350
II	330	400
III	360	420
IV	390	460

Prediction of color from fluorescence of raw sugar.

Color is an important parameter for determining the quality of raw sugar and because the color value is defined as the extinction coefficient at 420 nm, color may be related to one or several fluorophores. Carpenter and Wall (1972) reported detection of fluorescence from raw sugar samples when excited at 420 nm. The ICUMSA color of all the raw sugar samples are presented in Table 1. The color range of the samples is very large (340-51,290), which reflects the very different color compositions of the samples. As a way to correlate the color of the raw sugar samples to the measured fluorescence landscapes, PLS calibration models were built from the unfolded fluorescence data set. The predictions of color from the models were validated with full cross validation.

A PLS model of all 47 samples with three PLS components is presented in Table 3. The model error RMSECV, expressed in ICUMSA units, is high compared with the color range of the samples and the correlation coefficient between the measured color and the predicted color is only 0.59. It appeared that the two oldest samples (1 and 2) with a lot of developed color during storage caused the poor model results. A new PLS model omitting the two samples clearly improved the correlation coefficient to 0.88 (Table 3). A plot of the correlation of the predicted color from PLS model of the fluorescence data and the reference ICUMSA color of the 45 remaining raw sugar samples is shown in Fig. 7. The model error of 1622 for the new PLS model is still high when considering that the range of the sample color has been lowered as well. PLS modelling of subsets including only samples with low color gave the same relative model error to the modelled color range with no apparent model improvement. Previously, prediction of color from fluorescence data of white beet sugar samples has been reported with a satisfactory result, where a sample set consisting of 87 beet sugar samples from five different factories was modelled with five PLS components with R = 0.94 and RMSECV = 2.4 (color range = 11-44) (Nørgaard, 1995).

The difficulties in predicting the color of the raw sugar samples from the measured fluorescence are probably explained by the very high concentrations of color in the raw sugar sample set and the very varied color distribution in the samples. Furthermore, the different origins and production years of the samples ensure a global prediction model but the uniformity of the spectral information, which is needed in a good prediction model, decreases. The non-linear effects that influence the fit of the models are consistent with the fact that a linear relationship between fluorescence and color is only valid for samples with very low absorbance, i.e. very low color. The fluorescence is concentration-quenched at high color values and non-linear effects occur. E.g. sample 2 with the highest color displays a moderate fluorescence, which is the reason the sample

did not conform to the PLS model of color (Table 3). Therefore, samples with very low color such as white beet sugars and refined cane sugars and sample sets with more restricted origin should generally produce better PLS models of color.

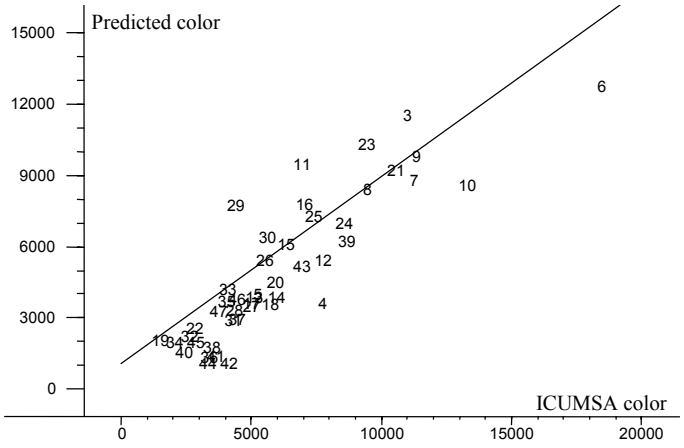


Figure 7. Plot showing the correlation of the predicted color of raw sugar from the PLS model of the fluorescence data using three PLS-components and the ICUMSA color of 45 raw sugar samples. Correlation coefficient = 0.88 and the model error is 1622 in ICUMSA color units.

Table 3. Results of the PLS models of the fluorescence data for the prediction of color using full cross validation

# Samples	# Variables	# PC[a]	R[b]	RMSECV[c]	Range[d]	Mean[e]
47	Full spectrum (1021)	3	0.59	7072	340-51,290	6720
45[f]	Full spectrum (1021)	3	0.88	1622	340-17,470	5040
45[f]	3[g]	3	0.87	1648	340-17,470	5040

[a] # PC is the optimal number of PLS components.
[b] R is the correlation coefficient.
[c] RMSECV is the model error in ICUMSA color units.
[d] Range of the ICUMSA color.
[e] Mean of the ICUMSA color.
[f] Omitting samples 1 and 2 with high storage color.
[g] (270,346), (340,421), (390,457); three excitation-emission (Ex,Em) wavelength variables selected from the unfolded fluorescence data by the PV method.

Selection of significant fluorescence wavelengths.
The above PLS models are based on all the 1021 wavelength variables of the unfolded fluorescence data, but not all variables are important for predicting color. The PV method is used to select the fluorescence variables that are important for color prediction. The three wavelength variables 610, 139 and 844 (Fig. 2) were selected in that order. This is the same number as PLS components used in the unfolded landscape models. Their (excitation,emission) wavelengths are (340,421) nm, (270,346) nm and (390,457) nm, respectively. A three-component multiple linear regression (MLR) model using the three variables for predicting color is made from all samples except 1 and 2 (Table 3).

The prediction results were very close to the model using all 1021 wavelength variables and the plot of the predicted color versus the reference color was almost identical to Fig. 7. Apparently, only the three selected wavelength variables are needed for modelling color at the same level as a full spectrum model. A comparison with the excitation and emission maxima of the four resolved components from the PARAFAC model in Table 2 reveals that the selected variables (340,421), (270,346) and (390,457) corresponds to combined components II/III, component I and component IV, respectively. The fact that component II and component III both correspond to one of wavelength variables supports the supposition that they belong to the same type of fluorophores. Thus, it is also demonstrated that the PARAFAC model on three-way data can extract more precise information such as small spectral changes, than the more crude PV method on the two-way unfolded data.

The characteristics of the resolved fluorophores.

It is interesting to find that all "three" fluorophores estimated in the PARAFAC model are correlated to the ICUMSA color. The choice of using absorbance at 420 nm as the wavelength to determine color was debated for many years before it was made official by ICUMSA (Godshall, 1997). If the three fluorophores are colorants or color precursors, the use of 420 nm for color measurements must be considered as a sensible compromise. The difficult part is to identify the three resolved fluorophores as true constituents of raw sugar. Some indications have already been given. The fluorescence excitation and emission spectrum of component I are located in the ultraviolet region and the component must be considered as a color precursor. The PCA shows that the component is negatively correlated to color development in stored sugars (1 and 2) whereas there is a high contribution of the component in some of the newest raw sugar samples in the data set (42 and 44). Concentration quenching of the ultraviolet fluorescence in the highly colored sugars could be the reason for this, but Fig. 3A shows that sample 43 with high color has a fair contribution of component I, i.e. situated in the upper end of PC2. Therefore, component I is defined as a color precursor which is decreased during storage of the sugar due to participation in color forming reactions, possibly polymerizations.

Component II/III from the PARAFAC model is defined as a fluorophore with (340,420) nm as the approximated excitation and emission maximum. The component has a very intense fluorescence (PCA) and in the selection of important wavelength variables related to color, the (340,421) variable was selected first. The close correlation between this fluorescence component and color imply that the component is a colorant. The fact that the excitation spectra of II/III in Fig. 5 are not extended to the 420 nm wavelength is likely due to the estimation of component IV in the PARAFAC model as an individual component. In some of the samples component IV has an individual peak and in other samples it is only an extension of the peak from component II/III. This is shown in Fig. 1 where the landscape of sample 43 (Fig. 1D), which has a high contribution from component II/III, has one peak at (420,345) whereas sample 1 (Fig. 1A) has an extra peak at (380,450) due to the storage color development. Consequently, component IV is considered as a colorant polymer extending from component II/III during development of additional color and it is not considered as a real individual fluorophore.

Conclusions

It has been demonstrated that chemometric methods applied to multi-wavelength fluorescence data can determine the chemical characteristics of fluorophores in raw cane sugar without knowing their exact chemical structure. Three principal fluorophores are found in raw sugar and one of them is characterized as an ultraviolet colour precursor that participates in color development during storage. The other two fluorophores fluoresce in the visible wavelength area and are potential colorants. The colorants fluorophores show a relation in their fluorescence behavior, perhaps as polymers, where the darker colorant fluoresces at higher wavelengths. Since all the resolved fluorophores correlates to color, they can be used as indicator substances in further studies of the color development in cane sugar processing.

Acknowledgment

Dr. Rasmus Bro and Claus A. Andersson are acknowledged for their expert advice and the programming of the N-way toolbox. The authors thank Deborah L. Harrell for putting together a representative raw sugar sample set and Henriette Werge and Karen Thomsen for technical assistance.

Literature cited

Andersson, C. A.; Bro, R. The *N*-way Toolbox for MATLAB, Ver. 1.03, Food Technology, MLI, KVL, Denmark. http://www.models.kvl.dk/source/nwaytoolbox/

Baunsgaard, D.; Andersson, C. A.; Arndal, A.; Munck L. Multi-way chemometrics for mathematical separation of fluorescent colorants and colour precursors from spectrofluorometry of beet sugar and beet sugar thick juice as validated by HPLC analysis. *Food Chem*. **2000**, *70*, 113-121.

Bro, R. PARAFAC: Tutorial and applications. *Chemom. Intell. Lab. Systems* **1997**, *38*, 149-171.

Bro, R. *Multi-way analysis in the food industry: models, algorithms and applications*. Ph.D. Thesis, University of Amsterdam, The Netherlands, 1998.

Bro, R. Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemom. Intell. Lab. Systems* **1999**, *46*, 133-147.

Carpenter, F. G.; Wall, J. H. Fluorescence in commercial sugars. In *Proceedings 1972 Technical Session Cane Sugar Refining Research*; New Orleans, LA, 1972; pp 47-61.

Godshall, M. A. Recent progress in sugar colorants. In *Proceedings 1996 Sugar Processing Research Conference*; New Orleans, LA, 1996; pp 262-305.

Godshall, M. A. Part 2. Color Analysis. *Symposium on sugar color*. Sugar Industry Technologists, 1997, pp 8-29.

Harshman, R. A. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA working papers in phonetics* **1970**, *16*, 1-84.

Höskuldsson, A. PLS regression methods. *J. Chemom*. **1988**, *2*, 211-228.

Höskuldsson, A. The H-principle: new ideas, algorithms and methods in applied mathematics and statistics. *Chemom. Intell. Lab. Systems* **1994**, *23*, 1-28.

ICUMSA Methods Book. The Determination of Raw Sugar Solution Colour – Official. Method GS1-7, International Commission for Uniform Methods of Sugar Analysis, 1994.

Martens, H.; Næs, T. *Multivariate Calibration*, 2nd ed., Wiley: New York, 1993.

Munck, L.; Nørgaard, L.; Engelsen, S. B.; Bro, R.; Andersson, C. A. Chemometrics in food science – a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemom. Intell. Lab. Systems* **1998**, *44*, 31-60.

Nørgaard, L. Classification and prediction of quality and process parameters of thick juice and beet sugar by fluorescence spectroscopy and chemometrics. *Zuckerind*. **1995**, *120*, 970-981.

Ravelo B., S.; Ramos S., E. L.; Mejias, R. Sugar cane deterioration and its implications in the factory. *Int. Sugar J.* **1991**, *93*, 82-86.

Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Systems* **1987**, *2*, 37-52.

Paper III

Specific screening for color precursors and colorants in beet and cane sugar liquors in relation to model colorants using spectrofluorometry evaluated by HPLC and multi-way data analysis

D. Baunsgaard, L. Nørgaard and M. A. Godshall

Abstract

Model colorants formed in fructose and glucose degradation and glucose-glycine and glucose-lysine Maillard reactions were compared with natural colorants in beet thick juice and cane evaporator syrup. Fluorescence excitation-emission landscapes resolved into individual fluorophores with the decomposition model PARAFAC were used as a screening method to compare the samples, which was validated with size exclusion chromatography using a diode array detector. Fluorophores from the model colorants were mainly located at visible wavelengths. An overall similarity in chromatograms and absorption spectra of the four model colorant samples indicated that the formation of darker color was the distinguishing characteristic, rather than different reaction products. The fluorophores obtained from the factory samples consisted of color precursor amino acids in the ultraviolet region. In the visible region model colorant fluorophores resembled beet colorant fluorophores at lower wavelengths and cane colorants at higher wavelengths. The beet and cane samples showed many similarities in the chromatograms and absorption spectra and therefore the differences in fluorescent colorants of the two factory samples was explained as different ratios of amino acids and invert sugar.

Keywords: *sugar; non-enzymatic browning; size exclusion chromatography; fluorescence; chemometrics; PARAFAC*

Introduction

The development of color impurities while processing sugar is a common problem in both the cane and beet sugar industries. Non-enzymatic color formation in the sugar process is caused by thermal degradation reactions of sugars with or without the participation of amino acids, the former also known as the Maillard reaction (Clarke *et al*., 1997). The color arising during evaporation and crystallization stages and the color developed after the subsequent storage of the processed sugar have been reported to originate mainly from these colorant reactions (Shore *et al*., 1984; Godshall *et al*., 1988). The reaction mechanisms of the degradation of reducing sugars are very complex with many intermediate products, which may react further to produce larger polymer systems (Kelly and Brown, 1978/79). This makes it difficult to characterize the color composition of process samples. Often, model colorant systems focusing on one of the browning reactions at a time have been used to achieve more information (Prey and Andres, 1971; Keramat and Nursten, 1994).

It has been known for years that commercial sugar exhibits fluorescence (Carpenter and Wall, 1972). It is also known that fluorescent components are formed both in the

Maillard reaction (Adhikari and Tappel, 1973) and in fructose degradation (Carpenter and Roberts, 1974), but the contribution of model fluorophores to the fluorescence in sugar process samples has not been clearly elucidated due to the complexity of the measured fluorescence. Spectrofluorometry combined with chemometric methods have been used in the last years as a screening method to obtain chemical information from various beet sugar process samples (Munck *et al*., 1998). Bro (1999) used a multi-way decomposition method PARAFAC (parallel factor analysis) as a form of "mathematical chromatography" to deconvolute, pure excitation and emission spectra from three-way data consisting of the fluorescence excitation-emission landscapes of 268 sugar samples collected from a beet sugar factory during a campaign. Four fluorescent components were found to capture the variation in the fluorescence data and two of them showed spectra with a close similarity to the pure fluorescence spectra of the amino acids, tyrosine, and tryptophan.

Baunsgaard *et al*. (2000a) confirmed the findings by Bro (1999) combining HPLC and fluorescence measurements on beet thick juice samples evaluated with PARAFAC. Seven fluorophores were resolved from HPLC fractions of thick juice. Apart from tyrosine and tryptophan, four of the fluorophores were identified as high molecular weight compounds and three of them were suggested to be Maillard reaction polymers. A recent study by Baunsgaard *et al*. (2000b) estimated fluorophores from raw cane sugar samples with PARAFAC. Three fluorophores were found, an ultraviolet color precursor and two components showing colorant polymer characteristics, one of them related to the development of color on storage.

The found fluorophores in these beet and cane sugar studies are difficult to identify as individual structures, especially the components in the visible wavelength area, due to the evident continuity between the fluorophores, which characterize growing polymers differing slightly in their molecular weight. Because the polymers are suspected to be products of sugar degradation reactions, a comparison of colorants in model systems with colorants in process samples can be used to associate these polymers with specific kinds of colorant reactions without the need for a complete structural identification. Therefore, in the work presented here, fluorescence excitation-emission landscape measurements decomposed by the multi-way PARAFAC model and HPLC size exclusion analyses using a diode array detector were used to extract information of chromophores and fluorophores from sugar colorant solutions. Glucose and fructose degradation as well as Maillard reactions of glucose-glycine and glucose-lysine were selected as model systems. The findings were compared with similar results from beet and cane process liquor samples.

Materials and methods

Materials.

Danisco Sugar A/S, Denmark, provided the beet thick juice and the cane evaporator syrup was collected by Sugar Processing Research Institute in Louisiana, USA.

D-fructose and D-glucose, L-lysine and L-glycine were purchased from Sigma (St. Louis, MO). The model colorants were prepared from heated fructose and glucose solutions, the latter also in mixed solutions with glycine and lysine, as described by Cookson *et al*. (1970). The pH level was 8, which was chosen to reflect a compromise

between the pH levels in the cane and beet sugar processes at the evaporation stage. The pH level was not maintained at 8 during the course of the reactions but was allowed to change with the development of color to imitate the drop in pH, which is normally observed during the color development in the sugar streams (Vercellotti *et al*., 1996). Both the glucose-glycine and the glucose-lysine browning solutions had a final pH at 7.8 ± 0.2 when the reaction was stopped. The final pH of the glucose and fructose browning solutions were 4.5 ± 0.3.

Size exclusion HPLC analysis.
The HPLC method used was originally developed for beet thick juice fractionation (Baunsgaard *et al*., 2000a). A Gilson system with a UV-VIS diode array detector was equipped with a Waters 250 Ultrahydrogel column. The mobile phase consisted of 0.2 M ammonium buffer, pH = 8.9 and water (20:80 v/v) at a flow rate of 0.5 ml/min. The model colorants were diluted 1:1.25 (v/v) and the factory samples 1:10 (v/v) with the mobile phase before a filtered aliquot of 100 µl was injected onto the column. The column dead time was determined as 12 min using Blue Dextran 2000. Adsorptive behavior of the column made it impossible to achieve reliable size calibration of the column using molecular weight standards. Instead, phenylalanine and tyrosine eluting at approximately 25 min were used to establish the size exclusion range of the column.

Fluorescence landscape measurements.
The fluorescence measurements were performed on a Perkin-Elmer LS50 B fluorescence spectrometer. The excitation range was 200-500 nm with an interval of 10 nm and the emission range was 280-700 nm. Excitation and emission slit widths were set to 10 nm and the scan speed was 1500 nm/min. The fluorescence concentration quenching levels of the different colorant samples were determined by measuring the fluorescence of dilution series of each sample diluted with ion-exchanged water. The glucose, fructose and glucose-glycine and glucose-lysine colorant samples had to be diluted 1:2, 1:8, 1:4 and 1:12 (w/w), respectively, to get below the overall concentration quenching level. Thick juice and evaporator syrup had to be diluted 1:200 and 1:320 (w/w), respectively. The fluorescence landscapes were then measured on five dilutions (1:2, w/w) in succession below the found quenching levels for each colorant sample to be used in the PARAFAC modelling. All samples were adjusted to pH 7.

Multivariate data analysis.
The PARAFAC model (Harshman 1970; Bro, 1997) can be used to decompose a three-way structure consisting of excitation-emission fluorescence landscapes of several samples (samples x excitation wavelengths x emission wavelengths) into individual excitation and emission profiles of a number of underlying fluorophores. The third dimension in the three-way structure is decomposed as a concentration profile showing the contribution of the resolved fluorophores to each sample in the data set. Bro (1997) provides a thorough tutorial of the PARAFAC model. For the PARAFAC model to be valid, wavelength areas in the fluorescence landscapes not conforming to true fluorescence, such as the Rayleigh scattering peaks and emission wavelengths less than excitation wavelengths, have to be treated as missing values. An example of a

fluorescence landscape of the glucose browning solution diluted 1:2 is shown in Fig. 1. The white areas in the plot are the missing data areas. All the PARAFAC models were estimated under a non-negativity constraint to improve the interpretability of the resolved spectral profiles. Additionally, in some of the models the emission profiles were estimated under uni-modality constraint to avoid the interference of artificial extra peaks in the spectra due to too many missing variables in the data set. The implementation of the PARAFAC model is obtained from The N-way Toolbox for MATLAB (Andersson and Bro, 1999). Principal component analysis (PCA) was used to compare resolved spectra from PARAFAC models (Martens and Næs, 1993). See Baunsgaard *et al*. (2000b) for details about software used in the data analyses.

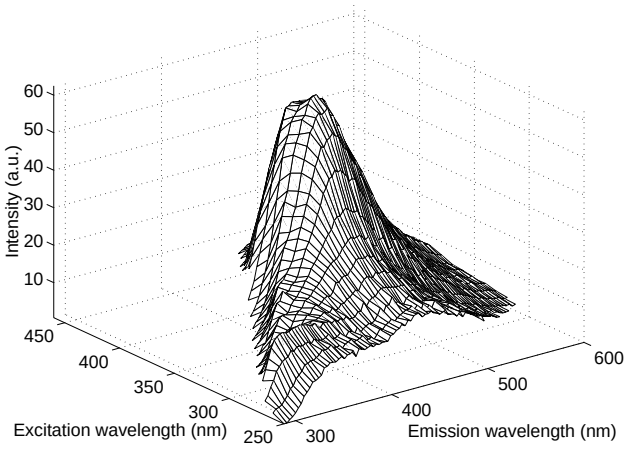


Fig. 1 Fluorescence excitation-emission landscape of the glucose sample (1:2 (w/w) dilution) with excitation range (240-450 nm) and emission range (288-600 nm). The white areas denote missing data due to measurement areas non-conforming to the PARAFAC model.

Results and discussion

Fluorophores in model colorants and factory samples.

The decomposition model PARAFAC was used to resolve fluorophores from three-way fluorescence data of the four model colorant samples and the beet and cane factory samples. Beet thick juice and cane evaporator syrup represent comparatively the same stage in the sugar processes between evaporation and crystallization where considerable color already has been formed. The three-way data array for each sample consisted of excitation-emission landscape measurements of five dilutions, i.e. 5 dilutions x 22 excitation wavelengths (240-460 nm, step 10 nm) x 105 emission wavelengths (288-600 nm, step 3 nm). The data arrays were then modelled successively with PARAFAC and the optimal number of components was determined for each model. The resolved components in a model are represented by estimated excitation and emission spectra and a dilution profile containing the concentration of each component in the five dilutions.

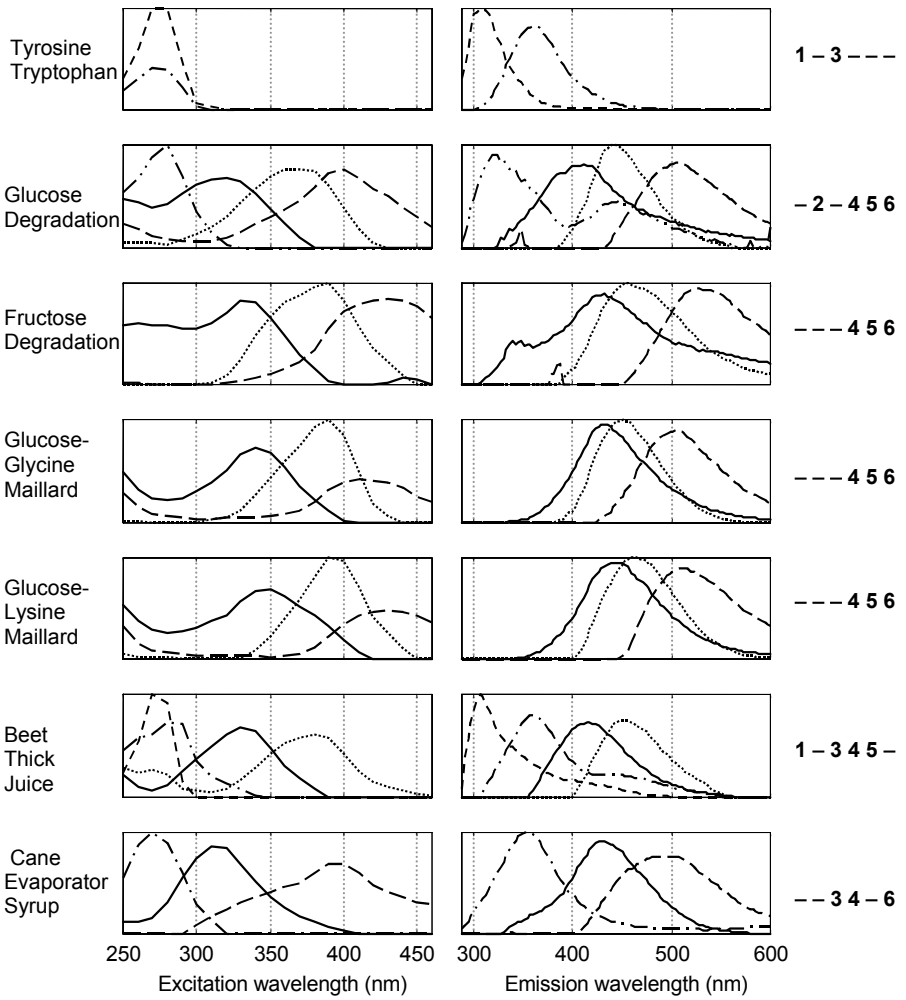


Fig. 2 Estimated excitation and emission spectral profiles from PARAFAC models of the fluorescence data of both model colorants and factory samples. All spectra have been normalized to unit length. The excitation and emission spectra of pure tyrosine and tryptophan are shown in the top row. The numbers to the far right in the figure are assigned to emission spectra in the different models that have comparable spectral profiles and wavelength maxima: 1 (---), 2 (-··-), 3 (-·-), 4 (–), 5 (···) and, 6 (--). Components 1 and 3 are tyrosine and tryptophan components, respectively, component 2 is an ultraviolet fluorophore, and components 4-6 are fluorophores in the visible wavelength area.

In Fig. 2 (row 2-7) the excitation and emission spectra from the PARAFAC models of model colorant samples and factory samples are shown. A few of the emission spectra display more than one peak, e.g. the emission profile farthest to the left in the glucose model in row 2. These extra peaks indicate that the fluorescence data does not totally conform to tri-linearity, which is a premise of the PARAFAC model (Bro, 1999). The dilutions of the samples also makes it more difficult to estimate true spectral profiles

because the dilution of colorants not only decreases the fluorescence intensity but also shifts the fluorescence towards lower wavelengths. In the top row in Fig. 2 the excitation and emission spectra of pure tyrosine and tryptophan are displayed to help the assignment of these amino acid components since they have already been identified in thick juice (Baunsgaard *et al*., 2000a). The numbers shown to the right of the emission spectra in Fig. 2 are used to compare the estimated components in the various models. The same number indicates emission spectra from the different models that show approximately the same spectral profile and wavelength position. The reason why the emission spectra were compared is because the excitation spectra of fluorophores are in general less specific than their emission spectra, e.g. tyrosine and tryptophan have very similar excitation spectra.

Component 1 and component 3, which is assigned to tyrosine and tryptophan, respectively, are present in thick juice as expected from the previous results. Component 3 or the tryptophan component is also present in cane evaporator syrup. Component 2 resolved from the glucose fluorescence data is an ultraviolet component, and components 4-6, which are situated in the visible wavelength area, are estimated in all the model colorants with minor spectral differences. Beet thick juice does not fluoresce at the high wavelength of component 6 and cane evaporator syrup lacks an equivalent to component 5.

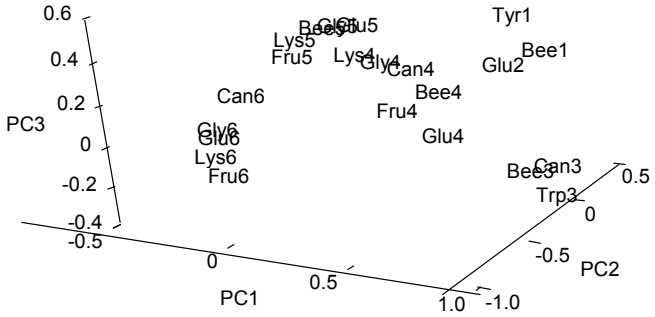


Fig. 3 A three-dimensional PCA score plot of the emission spectra from Fig. 2 showing PC1 vs. PC2 vs. PC3 explaining 48, 32 and 13 % of the total variance, respectively. Tyr = tyrosine, Trp = tryptophan, Glu = glucose, Fru = fructose, Gly = glucose-glycine, Lys = glucose-lysine, Bee = beet thick juice, Can = cane evaporator syrup. The numbers correspond to the numbers assigned to each spectrum in Fig. 2.

It is difficult to compare the likeness of the components in the models only by visual inspection of the emission spectra. Consequently, a principal component analysis (PCA) was performed on all the emission spectra in Fig. 2. The three-dimensional score plot of principal component 1 (PC1) versus principal component 2 (PC2) and principal component 3 (PC3) is shown in Fig. 3. Along PC1 the spectra are divided into a group

containing the ultraviolet spectra and another group containing the visible spectra. PC3 divide tyrosine, component 2 from glucose degradation, and component 1 from the thick juice model in their own group separately from the other ultraviolet components. Components 3 of cane evaporator syrup and beet thick juice are close to pure tryptophan. The visible emission spectra in Fig. 3 are also separated in subgroups especially the group with components 6 along PC2. The score plot thus confirms the approximate assignments made visually in Fig. 2. However, the PCA cannot be used in a more precise match of the individual fluorophores in the visible wavelength area. Another approach to compare the fluorophores is to calculate a correlation coefficient matrix between the emission spectra from the factory samples and all the other emission spectra displayed in Fig. 2. The correlation matrix is shown in Table 1. A visual estimation of the results indicated that spectra with correlation coefficients lower than 0.90 were too different for a spectral likeness and that correlation coefficients above 0.95 represented a spectral match between two components. Coefficients above 0.95 are highlighted in Table 1. The correlations of the ultraviolet components confirm the previous findings. In the visible wavelength area cane component 4 has a close likeness with glucose-glycine component 4 but is also quite similar to glucose-lysine and fructose components 4. In contrast, beet component 4 match glucose component 4. Beet component 5 is correlated strongly to glucose, fructose, glucose-glycine, and glucose-lysine components 5 whereas cane component 6 is weakly correlated to glucose and glucose-glycine components 6. Glucose component 2 and fructose and glucose-lysine components 6 are all three without any reasonable correlations to the factory sample components.

Table 1. Correlation coefficient matrix between the emission spectra from the factory samples and the emission spectra from the model colorant samples and pure amino acids as given in Fig. 2[a]

	Tyr1	Trp3	Glu2	Glu4	Glu5	Glu6	Fru4	Fru5	Fru6	Gly4	Gly5	Gly6	Lys4	Lys5	Lys6
Can3	-	0.95	-	-	-	-	-	-	-	-	-	-	-	-	-
Can4	-	-	-	-	0.90	-	0.95	-	-	0.98	-	-	0.94	-	-
Can6	-	-	-	-	-	0.90	-	-	-	-	-	0.92	-	-	-
Bee1	0.96	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bee3	-	0.97	-	-	-	-	-	-	-	-	-	-	-	-	-
Bee4	-	-	-	0.94	-	-	-	-	-	-	-	-	-	-	-
Bee5	-	-	-	-	0.96	-	-	0.96	-	-	0.99	-	0.93	0.97	-

[a] Tyr = tyrosine, Trp = tryptophan, Glu = glucose degr., Fru = fructose degr., Gly = glucose-glycine, Lys = glucose-lysine, Bee = beet thick juice, Can = cane evaporator syrup. The numbers correspond to the numbers assigned to the emission spectra in Fig. 2. Only coefficients ≥ 0.90 are displayed; highlighted values: >0.95.

Chromophores in model colorants and factory samples.

The fluorophores in the model systems and factory samples show spectral profiles that make them potential colorants or color precursors but these findings have to be supported by more information about the color composition in the various samples. Size exclusion chromatography using UV-VIS diode array detection has proved to be a good way to obtain information about colorless constituents and colorants in complex sugar

browning samples (Baunsgaard *et al*., 2000a). The chromatograms of the four model colorants at 280 and 420 nm are shown in Fig. 4. The chromatogram at 420 nm is used to establish the amount of color formed in the browning reactions, since this is the wavelength normally employed by the sugar industries for color determination (Godshall, 1997). All 420 nm chromatograms in Fig. 4 have peaks in the beginning of the run at 16-21 min. The dead time of the column is 12 min and this implies that the colored components in all the model solutions have molecular weights well below the 80 kDa capacity of the column. Results reported by Hofmann (1998) support these findings. He showed that ultracentrifugation on colorants formed in a glucose-glycine model system produced only trace amounts of compounds with a molecular weight >30 kDa whereas the majority of the compounds had a molecular weight <3 kDa.

The chromatograms of the glucose and fructose samples at 280 nm in Fig. 4a-b resemble each other in the overall peak pattern except for some distinct peaks eluting after 30 min. In Fig. 5a-b the absorption spectra shown at 17.75, 18.87, 22.45, and 45.25 min have rather similar profiles in the two chromatograms. The spectra from the fructose chromatogram are shifted slightly towards higher wavelengths, which implies a more advanced color formation in the fructose sample, which is also supported by the higher absorption intensities. Absorption spectra of peaks that differ in the two chromatograms are also displayed in Fig. 5a-b.

The glucose-glycine and glucose-lysine chromatograms at 280 nm in Fig. 4c-d also show a similar peak pattern except for a lower resolution in the glucose-lysine chromatogram. The absorption spectra at the comparative peaks in Fig. 5c-d display a difference in the spectral profiles where the spectra from the glucose-lysine sample have maxima at higher wavelengths, which along with the higher absorption intensities and lower resolution indicate that lysine forms more color with higher molecular weights than glycine. According to Parker and Williams (1968) darker colorants are formed in glucose-lysine systems due to the cross-linking of lysine. A comparison of all four model colorant chromatograms reveal that the comparative peaks at approximately 19 min have a spectral maximum around 265 nm related to a well-known but unidentified color precursor (Fleming *et al*., 1968). The spectra of the Maillard colorants (Fig. 5c-d) suggest that this color precursor enter into a colorant molecule with a partly intact chromophore even during increased browning.

Size exclusion chromatography was also performed on factory samples from beet and cane sugar processing. Chromatograms of the two factory samples are shown in Fig. 6. The absorption intensity level is four times higher for the cane sample, which would explain the small peaks at higher retention times in the 420 nm chromatogram compared to the beet sample. Otherwise the 420 nm chromatograms of the two factory samples have the same peak pattern as the model colorants. The 280 nm chromatograms of the two factory samples display a close resemblance in the peak distribution, which is supported by several peaks showing quite similar absorption spectra. In Fig. 7 a-b the spectra from peaks at 17.28, 28.35 and 41.20 min in the thick juice chromatogram resemble spectra from comparative peaks at 16.07, 28.02 and 40.32 min in the evaporator syrup chromatogram. The two spectra at 28.35 and 28.02 min have a retention time and a spectral form, which was previously identified as tryptophan (Baunsgaard *et al*., 2000a). Tyrosine was identified as a peak at approximately 25 min in

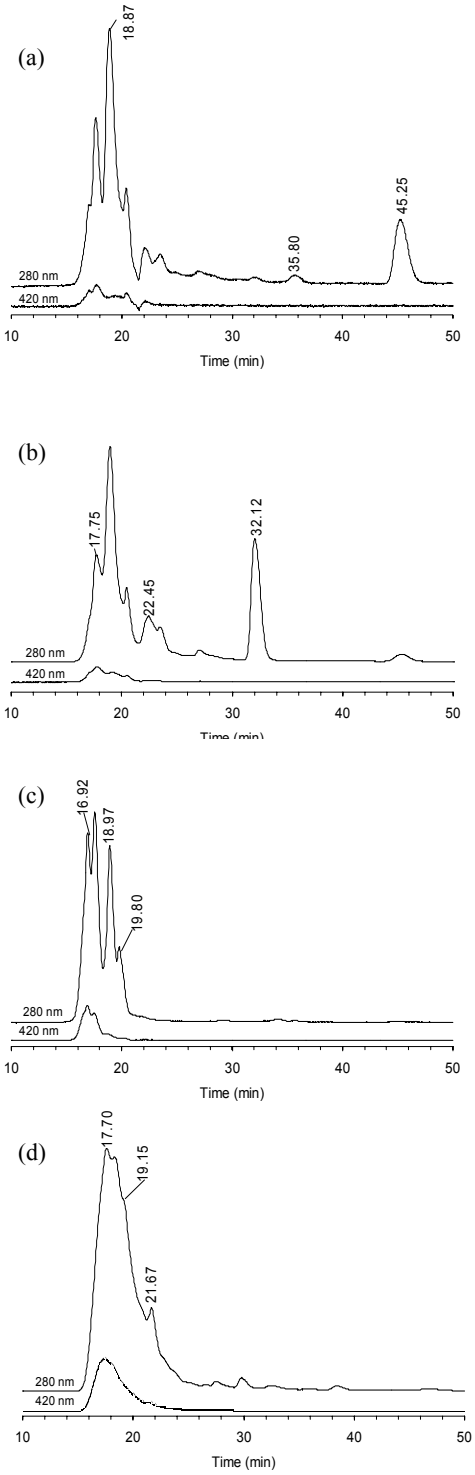


Fig. 4 Chromatograms at 280 and 420 nm of the model colorants diluted 1:1.25 (v/v). (a) Glucose; (b) fructose; (c) glucose-glycine; (d) glucose-lysine.

the beet chromatogram as well (spectrum not shown), but the amino acid was not found in the cane chromatogram. Some spectra from different chromophores in the two chromatograms are shown in Fig. 7c-d. Thick juice has several compounds with absorption maxima round 265 nm, which are well separated late on the column and are consequently not related to the color precursor at 265 nm in the model colorants. The evaporator syrup chromatogram in Fig. 6b has a small peak at 11.97 min that is not retained on the column and the molecular weight has to be higher than 80 kDa. The thick juice chromatogram beginning at 15.5 min does not contain this early peak but has the same chromatogram course as the model colorants.
Bento (1995) separated several groups of colorants from cane and beet factory samples on a sugar crystal column and a colorant group eluated with an alkaline mobile phase had a diode array spectrum very close to the 11.97 min spectrum in Fig. 7d. Bento found this group of colorants in cane syrup but not in beet syrup, Maillard reaction products, or fructose degradation products, which is consistent with our results. The 11.97-min peak obviously represents a category of colorants, which are of very high molecular weight but are not formed in the sugar degradation reactions. These high molecular weight colorants probably originate in the sugarcane plant (Godshall *et al*., 1988).

Colorant and color precursor characteristics in cane and beet liquor samples.
Tyrosine and tryptophan have repeatedly been resolved as fluorophores in beet factory samples (Bro, 1999, Baunsgaard *et al*., 2000a). The separation of a tryptophan peak in the HPLC chromatogram of the evaporator syrup and the close spectral likeness of the cane component 3 with beet component 3 and tryptophan from the PARAFAC models (Fig. 2 and 3) confirm that tryptophan is found in cane evaporator syrup in this study. A tryptophan-like component was also resolved in a PARAFAC model of fluorescence landscapes from raw cane sugars (Baunsgaard *et al*., 2000b). Tyrosine, identified as beet component 1, was not resolved as a component in the cane model in Fig. 2 or found in the cane chromatogram.
A PARAFAC model made on the fluorescence data of tyrosine, tryptophan and the two factory samples together showed a small contribution of the tyrosine component in the resolved cane sample concentration profile (plot not shown). This suggests that tyrosine is present in cane evaporator syrup but in too low a concentration to be resolved in the PARAFAC model.

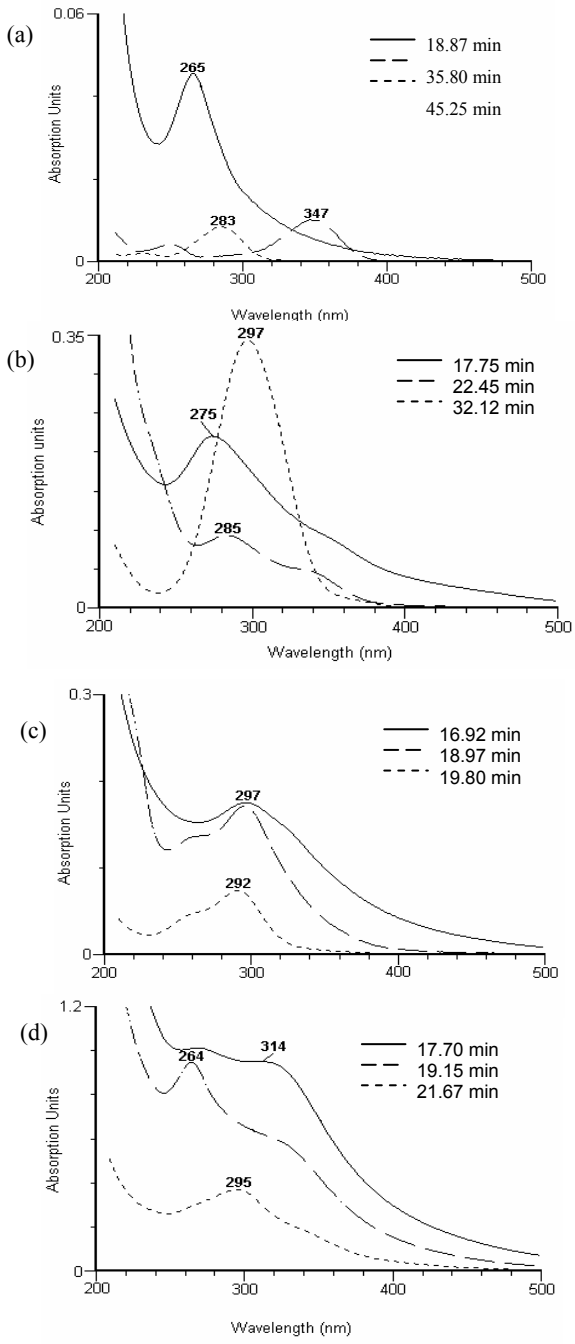


Fig. 5 Diode array absorption spectra of the model colorants corresponding to the peaks marked in the chromatograms in Fig. 4. The bold numbers denote the wavelengths of the absorption maxima of the spectra. (a) Glucose; (b) fructose; (c) glucose-glycine; (d) glucose-lysine.

Clarke *et al*. (1997) stated that there is a higher amount of amino acids in beet juice than in cane juice, which would explain the difficulty in resolving tyrosine from cane factory samples. In general, tryptophan preserves better its fluorophore properties than tyrosine, e.g. in high molecular weight compounds (Baunsgaard *et al*., 2000a). The amino acids are known color precursors in the sugar process samples and as the precursors are important indicators of potential color, e.g. in storage of the produced sugar (Baunsgaard *et al*., 2000b), tryptophan is a promising fluorophore marker of color formation in sugar processing.

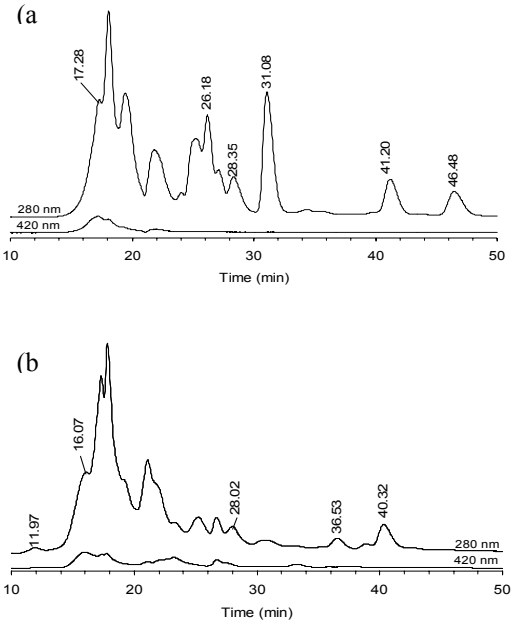


Fig. 6 Chromatograms at 280 and 420 nm of the factory samples diluted 1:10 (v/v). (a) Beet thick juice; (b) cane evaporator syrup.

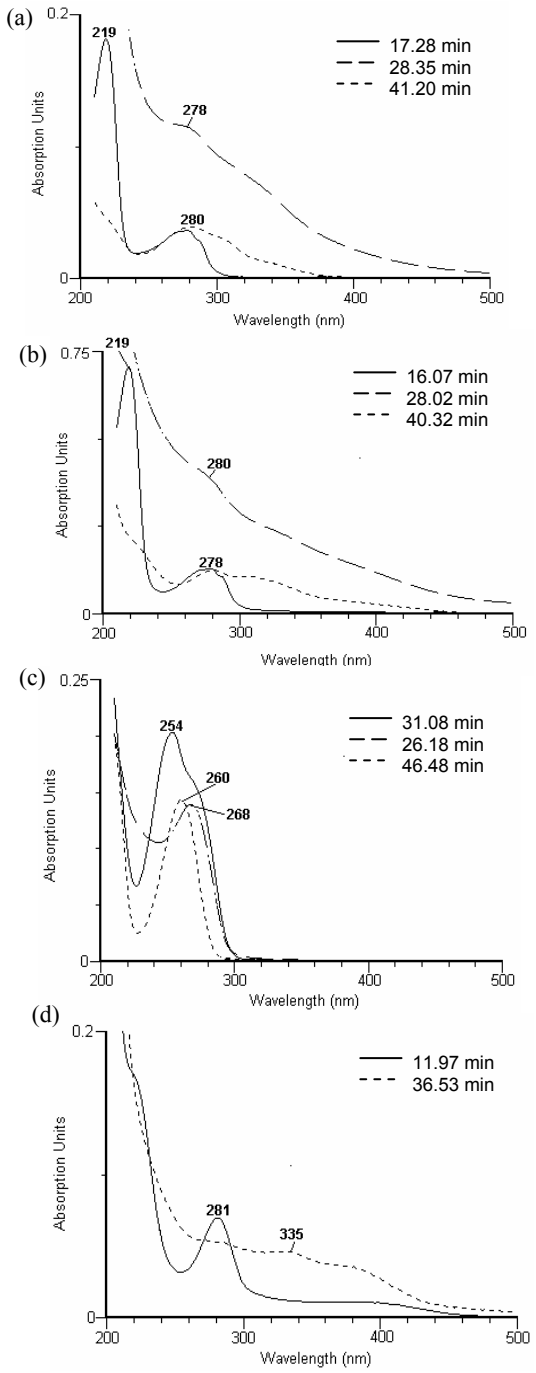


Fig. 7 Diode array absorption spectra of the factory samples corresponding to the peaks marked in the chromatograms in Fig. 6. The bold numbers denote the wavelengths of the absorption maxima of the spectra. Comparison of similar (a) and (b) as well as dissimilar (c) and (d) spectra of beet thick juice and cane evaporator syrup, respectively.

The ultraviolet component 2 from the glucose sample and a corresponding ultraviolet shoulder on the emission spectra of component 4 of the fructose sample (Fig. 2) are suggested to be catechols formed in base-catalyzed degradation of glucose and fructose (Kato et al., 1973). Catechols are fluorescent with excitation and emission maxima at approx. 280 nm and 320 nm, which is consistent with the spectral characteristic of component 2.

The correlations coefficients between the model colorant fluorophores and the factory colorant fluorophores in Table 1 show that that there is a spectral comparability in the visible components. The cane fluorophores resemble the model colorant fluorophores at higher wavelengths more than the beet fluorophores. The two components at the highest wavelengths (Fru6 and Lys6) are too high up in wavelength for a match with cane component 6 and the emission spectrum of the fructose component in Fig. 2 is so different that it almost could be assigned as a component 7 instead. However, glucose and glucose-glycine components 6 give a reasonable correlation with the cane component 6. The beet fluorophores are not represented in that higher wavelength area as shown in Fig. 2, but there are good correlations of beet component 5 with component 5 of all the model colorants in Table 1 at lower wavelengths.

According to Clarke *et al*. (1997) cane juice contains more invert sugar than beet juice. The higher amounts of invert sugar but lower of amino acids in cane juice suggest that the formation of the darker colorants is dependent on the ratio of invert sugar and amino acids. The diode array spectra of higher molecular weight colorants (16-20 min) colorants in the model systems and factory samples (Fig. 5a-d, 7a-b) display almost the same spectral behavior and it is not possible to confirm these specific differences between cane and beet colorants in the fluorophores. Fluorescence is a more specific and sensitive sensor than UV-VIS absorption and it is very likely that fluorophore structures are easier to distinguish than chromophore structures in colorants.

Overall, the chromophores and the fluorophores in the four model colorant samples display very similar characteristics and the color development in the various reactions seems to be more dependent on the reactivity and ratio of the reactants, and thus on the formation of darker color, than on differences in the reaction products. The differences in colorants in the beet and cane factory samples seems to be related to different ratios in invert sugar and amino acids, which would explain the darker and higher molecular weight compounds in the cane sample. The fact that the colorants from the model systems and the two factory samples do not totally match up can be explained by the contribution to color in the factory samples from other sources, e.g. enzymatic browning reactions that occur when the plant cells are disrupted during the extraction of juice from both sugar cane and sugar beet (Gross and Coombs, 1976).

Acknowledgments

Rasmus Bro and Claus A. Andersson are acknowledged for their expert advice and the programming of the N-way toolbox. The authors thank Frank Desimone for synthesizing the model colorants.

Literature Cited

Adhikari, H. R. and Tappel, A. L. Fluorescent products in a glucose-glycine browning reaction. *J. Food Sci*., 1973, *38*, 486-488.

Andersson, C. A. and Bro, R. The N-way Toolbox for MATLAB, Ver. 1.03. Can be downloaded from http://www.models.kvl.dk, 1999.

Baunsgaard, D., Andersson, C. A., Arndal, A. and Munck, L. 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. *Food Chem.*, 2000a, *70*, 113-121.

Baunsgaard, D.; Nørgaard, L.; Godshall, M. A. Fluorescence of raw cane sugars evaluated by chemometrics. *J. Agri. Food Chem*., 2000b, submitted.

Bento, L. S. M. Separation of sugar colorants using chromatographic columns containing sucrose crystals. *Zuckerind*., 1995, *120*, 123-130.

Bro, R. PARAFAC: Tutorial and applications. *Chemom. Intell. Lab. Systems*, 1997, *38*, 149-171.

Bro, R. Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemom. Intell. Lab. Systems*, 1999, *46*, 133-147.

Carpenter, F. G.; Wall, J. H. Fluorescence in commercial sugars. *Proceedings 1972 Technical Session Cane Sugar Refining Research*, New Orleans, LA, 1972; pp. 47-61.

Carpenter, F. G.; Roberts, E. J. Colorant formation under refining conditions. *Proceedings 1974 Technical Session Cane Sugar Refining Research*; Cherry Hill, NJ, 1974; pp. 106-115.

Clarke, M. A.; Edye, L. A.; Eggleston, G. Sucrose decomposition in aqueous solution, and losses in sugar manufacture and refining. *Adv. Carbohydr. Chem. Biochem*., 1997, *52*, 441-470.

Cookson, D.; Parker, K. J.: Williams, J. C. Properties of sugar colorants removable by ion-exhange resins. *Proceedings 1970 Technical Session Cane Sugar Refining Research*; Boston, MA, 1970; pp. 103-113.

Fleming, M.; Parker, K. J.; Williams, J. C. Aspects of the chemistry of the browning fraction of reducing sugars. *Proceedings XIIIth Congress International Society Sugar Cane Technologists*; Taiwan, 1968; pp. 1781-1800.

Godshall, M. A; Clarke, M. A.; Dooley, C. D.; Roberts, E. J. High molecular weight color in refineries. *Proceedings 1988 Sugar Processing Research Conference*; 1988; pp. 75-93.

Godshall, M. A. Part 2. Color Analysis. Symposium on sugar color. Sugar Industry Technologists, 1997; pp. 8-29.

Gross, D. and Coombs, J. Enzymic colour formation in beet and cane juices. *Int. Sugar J*. 1976, 78, 69-73 and 106-109.

Harshman, R. A. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA working papers in phonetics*, 1970, *16*, 1-84.

Hofmann, T. Studies on the relationship between molecular weight and the color potency of fractions obtained by thermal treatment of glucose/amino acid and glucose/protein solutions by using ultracentrifugation and color dilution techniques. *J. Agri. Food Chem.*, 1998, *46*, 3891-3895.

Kato, H.; Mizushima, M.; Kurata, T.; Fujimaki, M. The formation of Alkyl-*p*-benzoquinones and Catechols through Base-catalyzed Degradation of Sucrose. *Agr. Biol. Chem*., 1973, *37*, 2677-2678.

Kelly, F. H. C.; Brown, D. W. Thermal decomposition and colour formation in aqueous sucrose solutions. *Sugar Technol. Rev*., 1978/79, *6*, 1-48.

Keramat, J.; Nursten, H. E. The relationship between the coloured compounds present in the pressed liquor of cane sugar manufacture and those formed in Maillard reactions, in alkaline degradation of sugars, and in caramelisation. *Food Chem*., 1994, *51*, 417-420.

Martens, H.; Næs, T. *Multivariate Calibration*, 2nd ed., Wiley: New York, 1993.

Munck, L.; Nørgaard, L.; Engelsen, S. B.; Bro, R.; Andersson, C. A. Chemometrics in food science - a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemom. Intell. Lab. Systems,* 1998, *44*, 31-60.

Parker, K. J.; Williams, J. C. The isolation and properties of sugar colorants. *Proceedings 1968 Technical Session Cane Sugar Refining Research*; 1968; pp. 117-128.

Prey, V.; Andres, H. Secondary reactions during juice purification. *Z. Zuckerind*., 1971, *21*, 267-271.

Shore, M; Broughton, N. W.; Dutton, J. V.; Sissons, A. Factors affecting white sugar colour. *Sugar Technol. Rev*., 1984, *12*, 1-99.

Vercellotti, J. R.; Clarke, M. A.; Edye, L. A. Components of molasses: I. Sugarcane molasses: factory and seasonal variables. *Proceedings 1996 Sugar Processing Research Conference*; New Orleans, LA, 1996; pp 321-34.

Paper IV

Evaluation of the Quality of Solid Sugar Samples by Fluorescence Spectroscopy and Chemometrics

D. Baunsgaard, L. Munck and L. Nørgaard

Abstract

Its has been shown that fluorescence spectroscopy of sugar in aqueous solution carries important quality and process information related to beet sugar factories, which is accessible by multivariate analyses. A method for measuring crystalline sugar directly on-line in the process should be advantageous. In this paper we compare the solution measurement technique with two methods of fluorescence measurement on solid sugar. Surprisingly, it was possible to measure fluorescence through the sugar crystals by using the same transmission techniques with 90° detection as with the sugar solutions. This method was compared with a 45° front-surface reflection method. Sugar samples from six different sugar factories were examined. The spectral responses were reasonable, but they were influenced by the heterogeneous sample composition and the sample geometry. It was possible with the two methods to separate sugar samples according to factory with the use of principal component analysis (PCA). Seasonal time trends were found in weekly samples from the same factory. Partial least-squares regression (PLS) was used to predict quality parameters, where color (range: 6-41), ash (range: 0.003-0.018) and α-amino-N (range: 0.28-5.07) could be modeled with errors of 2.3-2.6, 0.0015-0.0016 and 0.40-0.42, respectively. Model errors for similar solution data have been determined to 2.4, 0.0012 and 0.266, respectively.

Index Headings: Fluorescence; Sugar; Solid sample; Chemometrics; Multivariate calibration; Principal component analysis (PCA); Partial least squares (PLS).

Introduction

In combination with multivariate methods, fluorescence analysis has proved to be a promising screening method with respect to prediction of quality parameters in beet sugar samples.[1] A study on fluorescence screening of sugar solutions showed the potential usefulness.[2] Fluorescence spectra of 81 sugar samples from six different factories were measured at four characteristic excitation wavelengths. With the use of the classification tool SIMCA (soft independent modeling of class analogy), the sugar samples could be classified according to factory with a 6% misclassification. Multivariate calibration models using partial least-squares regression (PLS) on the fluorescence spectra were developed for predicting sugar quality parameters. With the use of test set validation α-amino-N, color and ash were predicted with a correlation coefficient (R) of 0.98, 0.94 and 0.91, respectively. These results demonstrate that the measured fluorescence spectra reflect the chemistry of the specific sample with regard to the composition of the sugar beet raw material as well as the influence of the process.

Fast analysis is an important objective in screening methods. Measurements performed directly on the sample without any pretreatment are desirable especially when considering "on-line" or "in-line" measurements implemented in the process. Sugar solutions can be used to perform precise and significant laboratory analyses, but direct solid sugar measurements are more advantageous from a process technologist's point of view. Fluorescence measurements on sugar crystals have not previously been reported in the literature, but solid-state fluorescence is likely to differ from that of solutions. Reflectance fluorescence spectra of solid aromatic compounds in potassium bromide pellets[3] have shown fairly good resemblance to the solution spectra but with spectral shifts and changes in the intensity differences of the spectral peaks. In solid materials, such as wheat milling fractions[4] and mechanical pulp,[5] solid-state fluorescence has been used to differentiate chemical components.
In the present work we explore the possibility of measuring fluorescence on sugar crystal samples and of extracting sample characteristics from the data obtained. Two techniques for measuring solid-state fluorescence are examined: transmission detection at right angles through the sample as normally used with solutions, and a reflection front-surface detection, where the emission is detected from the same surface as the incident light.

Experimental

Instrumentation and Software.
Fluorescence spectra were recorded on a Perkin-Elmer LS 50B spectrometer connected to an IBM-compatible 486 / 50 MHz PC through an RS232C interface. Calculations were performed with Matlab for Windows version 5.2 (The MathWorks, Inc.) and Unscrambler version 7.01 (CAMO ASA). See Ref. 6 and 7 for details about software for instrument control.

Measuring conditions.
Uncorrected and unsmoothed emission spectra were recorded with the use of the same parameters as in Ref. 2. The four excitation wavelengths used were 230, 240, 290, and 340 nm. The corresponding emission ranges were 275-560, 275-560, 311-560, and 361-560 nm, respectively (in total 1023 data points or wavelength variables). Excitation and emission monochromator slit widths were set to 10 nm for the transmission measurements and 6 nm for the reflection measurements and the scan velocity used was 1500 nm/min.

Sample placement.

Transmission: A 10 x 10 mm quartz cuvette was used in a standard cell holder. The measurements were performed directly on the sugar crystal samples by filling the cuvette with the sample and placing it in the cell holder. The light is detected at right angles to the incident light.
Reflection: A Front-Surface Accessory (Perkin-Elmer Part No. 5212 3130) was used. The sample was filled in a Powder Holder (Perkin-Elmer Part No. 5212 3164) with a Silica Window (Perkin-Elmer Part No. 5212 3814) and the holder was mounted in the accessory. It is important to use approximately the same amount of sample and the same

tightening of the holder cap for each sample measurement to ensure uniformity in the sample handling. The emission light was detected at the same surface as the incident light in an angle of approximately 45°.
Solution: The same sample placement and cuvette were applied as with the solid sugar transmission measurements.[2] The sample holder was thermostatted to 24 °C ± 0.1.

Samples.
The sample set (Table I) consisted of 97 white sugar samples divided between six different factories (A-F), and a coarse-grain refined sugar (P) as well as an extra pure sugar (X), both from factory E. All samples are weekly samples, where a weekly sample is a mixture of approximately 90 sugar samples collected during one week of production. The same sample set has been measured as sugar solutions, where each sample was prepared by dissolving 2.25 g sugar in 15.0 ml double ion exchanged water.[2]

Chemical measurements.
Chemical measurements on sugar are ash (measured as conductivity, content in %), color (molar absorbtivity measured at 420 nm) and α-amino-N (ninhydrin-method, content in ppm). All chemical measurements were performed in accordance with the International Commission for Unified Methods of Sugar Analysis (ICUMSA).

Chemometric methods.
For a thorough explanation of principal component analysis (PCA)[8,9] and partial least-squares regression (PLS),[8,10-13] the literature should be consulted. Short descriptions introducing the methods are given in Ref. 2.

Results and discussion

Spectra.
It was quite a surprise that it was possible to take measurements directly through sugar crystals packed in a normal 10 x 10 mm quartz cuvette in the transmission method. Obviously, the bed of sugar crystals was transparent enough to admit radiation of the pulsed xenon lamp through the sample and a signal could be detected on the right-angled side. In Fig. 1 the raw transmission emission spectra of solid sugar samples from three of the factories (A, B and F) are displayed. The intensity of the spectra is high and all spectra have a good signal-to-noise ratio. The spectral outline is very similar when comparing the factories, but intensity differences in spectra from the same factory can be seen, especially for factory A. In Fig. 2a and b the spectra of transmission and reflection samples from factory E are shown, respectively. The spectral shape of the reflection spectra resembles that of the transmission spectra except for the low wavelength side of the emission bands of 230 and 240 nm excitation, where the reflection spectra have a shoulder. The overall intensity of the reflection spectra is also higher than the intensity of the transmission spectra. These spectral differences between transmission and reflection spectra are explained by the two different ways of measuring solid-state fluorescence. The fact that the transmission emission bands lack spectral response in the low wavelength side of the 230 and 240 nm spectra is due to light scattering and self-absorption of the fluorescence passing through the heterogeneous sample.[14]

TABLE I. The sugar sample set consisting of weekly samples from each factory, e.g. A04 means the weekly sample from factory A of the fourth week of the 1993 campaign. The P and X sugars are special sugars from factory E.

	Sugar	# samples
Factory A	A04-A16	13
Factory B	B03-B16	14
Factory C	C02-C16	15
Factory D	D03-D16	14
Factory E	E02-E16	15
Factory F	F03-F14	12
P sugar	P03-P09	7
X sugar	X10-X16	7

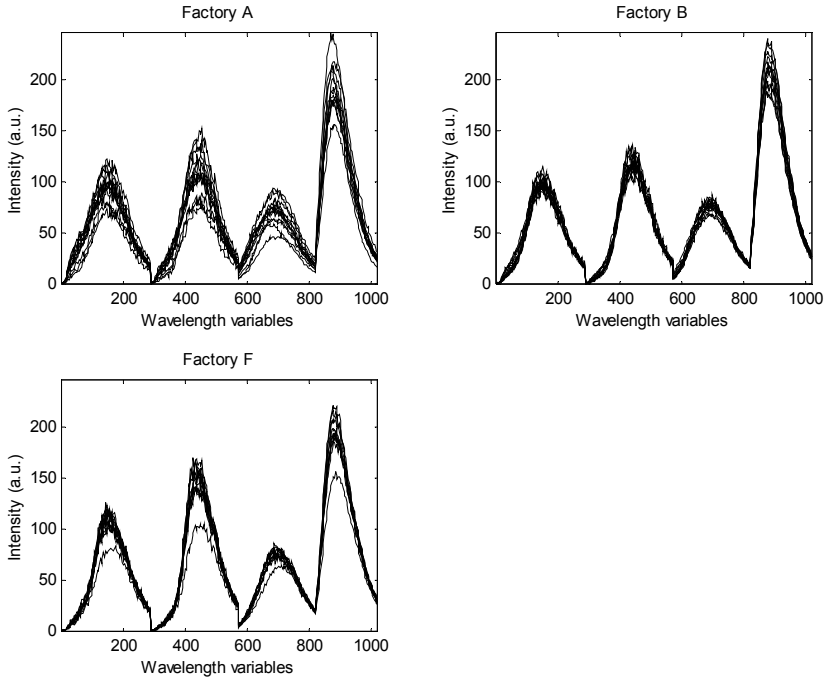


Fig. 1. Raw emission spectra of sugar samples from three different factories A, B and F measured with the transmission method. Wavelength variables 1-286: excitation 230 nm and emission 275-560 nm, variables 287-572: excitation 240 nm and emission 275-560 nm, variables 573-823: excitation 290 nm and emission 311-560 nm and variables 824-1023: excitation 340 nm and emission 361-560 nm.

The fluorescence from the reflection method, on the other hand, is a reflectance fluorescence detected from the same side of the sample as the incident light. Therefore light scattering and self-absorption of the fluorescence emission in the solid medium will be less than those for the transmission method and the low wavelength side of the

reflection spectra will be less intensity quenched. The increase in the overall intensity level in the reflection spectra compared to the transmission spectra is caused by detection of scattering light reflected from the surface of the heterogeneous sample.[15]

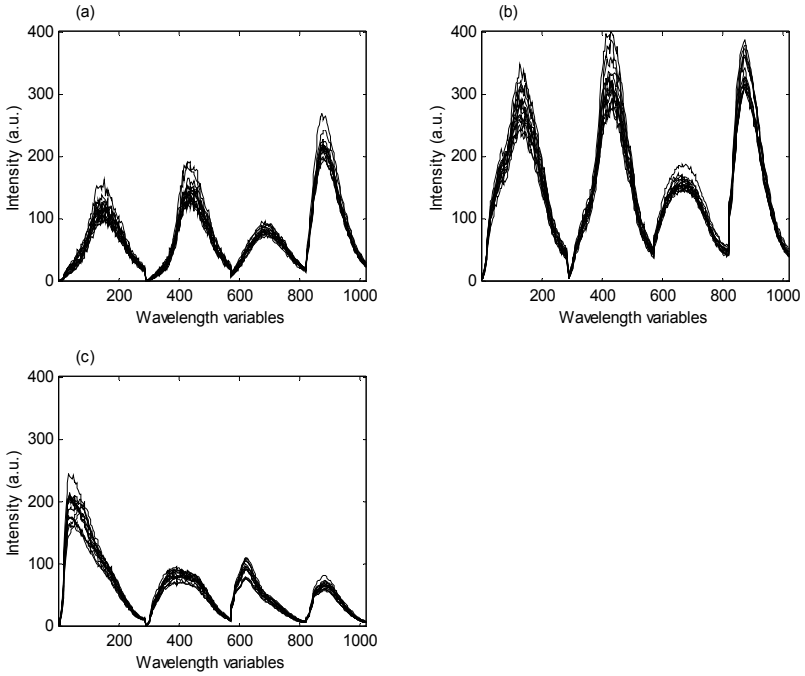


Fig. 2. Raw emission spectra of sugar samples from factory E. (a) Solid samples measured with the transmission method, (b**)** solid samples measured with the reflection method, (c**)** solution samples, 2.25 g sugar in 15.0 mL water (Ref. 2). Wavelength variables are defined in Fig. 1.

In Fig. 2c the fluorescence spectra of sugar solutions[2] from factory E are shown. Compared to the transmission and reflection spectra, the solution spectra have shifted the emission peak towards the low wavelength side of the emission band. The very low concentration of the diluted solutions gives very little light scattering and almost no self-absorption and the solution spectra will have little signal quenching in the low wavelength side. An overall PCA was made on the fluorescence data from the combined 291 samples of the three different methods, transmission, reflection and solution. In Fig. 3 the PCA plot of score vector 1 vs. score vector 2 is shown. The samples separate in two groups in the plot. The solution samples (S) form their own cluster and the samples from the two solid methods form another. The solid sample cluster can almost be divided into a transmission cluster (T) and a reflection cluster (R) except for seven transmission samples, the P-sugar samples, which are situated at the point of the arrow in Fig. 3. The P-sugars are coarse-grain refined sugars from factory E. Due to their larger crystal sizes as well as their higher purity, light scattering and self-absorption are reduced and their

transmission spectra resemble more the reflection spectra than the other transmission spectra. The reflection cluster shows more scattered samples in Fig. 3 than the transmission cluster. This result illustrates well the interference of the reflected scattering light in the reflection spectra, which introduces a considerable background. Also the small size of the sample area measured in the reflection method introduces arbitrariness in the measurements. The area of the surface struck by the incident light is very small (1 x 9 mm) and the fluorescence measurements become very dependent on the composition of the small sample area.

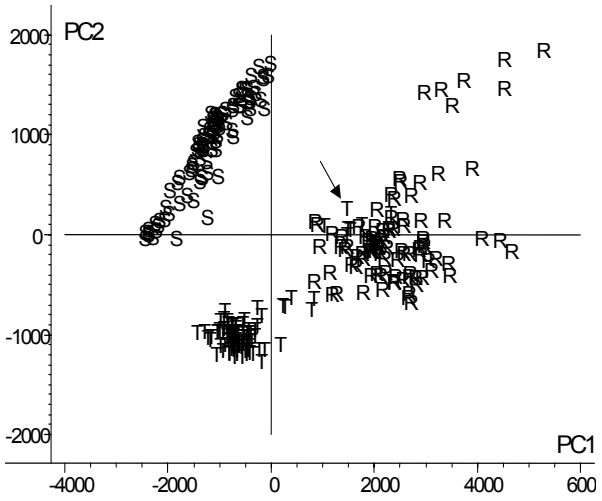


Fig. 3. PCA score plot of the fluorescence data of the total 291 samples from the three methods of measurement; S = solution, T = transmission and R = reflection. Principal component 1 (PC1) vs. principal component 2 (PC2) explains 77% and 20% of the total variance, respectively.

Factory characteristics.

PCA on the fluorescence data of all 97 transmission samples is displayed in Fig. 4. Plotting score vector 1 vs. score vector 3 explains 89% and 3% of the total variance, respectively. It is possible to separate factory F as well as the X and P samples as individual groups from the rest of the samples. A new PCA score plot without the F, X and P samples (Fig. 5) shows that the remaining five factories A-E are more alike, but some grouping of the factories can be found, e.g. A can be divided from D and E. It is not possible to separate the samples from factory B and C into two separate groups with any combination of score vectors. A three-dimensional PCA score plot (Fig. 6) of the two factories confirms that it is not possible to draw a line that separates them completely (B05, C15 and C16 are misclassified). As a characterization of the sugar processing in the fluorescence data, PCA on the weekly samples from sugar factory D was performed. In Fig. 7a a score plot is shown, where the samples have been connected

from the beginning of the campaign to the end. In the plot there is a trend of migration of the weekly samples, which reflects some process and/or sugar beet-related development in the factory.

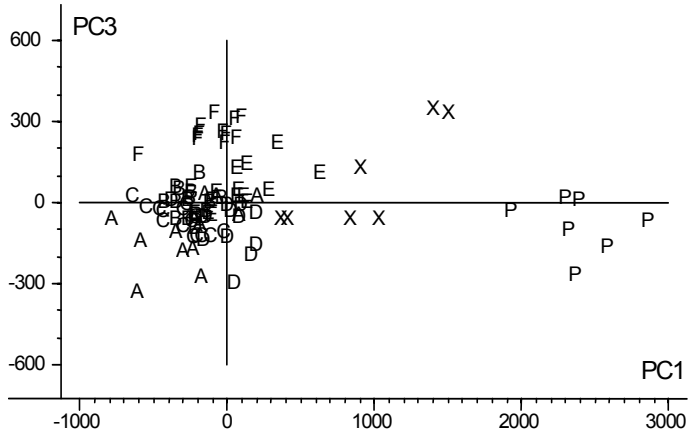


Fig. 4. PCA score plot of the fluorescence spectra of the 97 transmission solid sugar samples (factory A-F, X and P sugars from factory E). PC1 vs. PC3 explains 89% and 3% of the total variance, respectively.

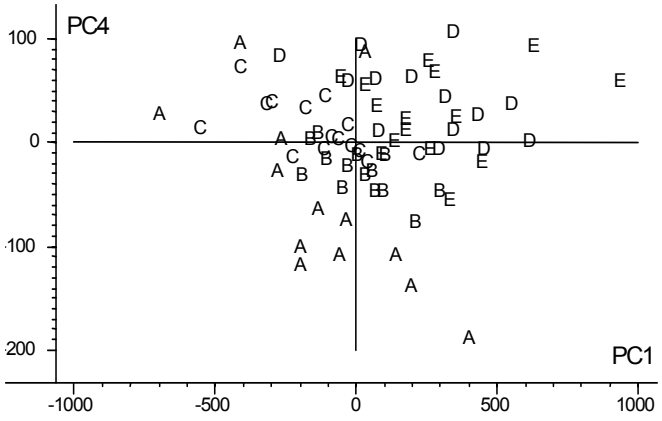


Fig. 5. PCA score plot of the fluorescence spectra of transmission solid sugar samples from 5 sugar factories (A-E). PC1 vs. PC4 explains 68% and 3% of the total variance, respectively.

In Fig. 8 the plot of score vector 2 vs. score vector 3 of a PCA on all 97 reflection samples is shown. The P samples form a separate group (plot not shown), but the X and F samples cannot be separated completely from the other samples in any of the score

plots. PCA on all combinations of factories two and two were performed. All the factories showed individual grouping but not with a complete separation. The distribution of the weekly samples for factory D in a PCA score plot is shown in Fig. 7B. There seems to be less chronological order compared to the transmission samples in Fig. 7A, but without sample D08 and D10 there is a time trend along score vector 1. A comparison of the results from the transmission and reflection fluorescence data analyses in relation to factory characteristics indicates that the transmission data contain more specific information of the individual factories. The sample information is reduced and more random when the reflection method is used.

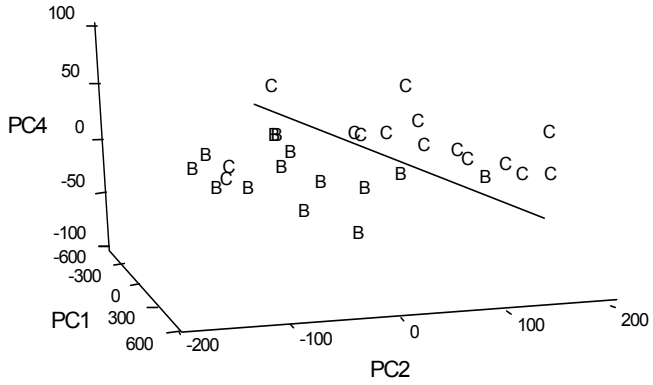


Fig. 6. A 3-dimensional PCA score plot of transmission samples from the factories B and C (PC1 vs. PC2 vs. PC4). One B sample (B05) and two C samples (C15 and C16) are misclassified by the separation line.

In the previous factory characterization made on fluorescence spectra from sugar solutions, a more distinct factory classification was made.[2] SIMCA could classify the sugar solution samples according to factory with a 6% misclassification. SIMCA analyses were also performed on the two solid sugar data sets, but the results were ambiguous with many of the samples misclassified or assigned to several factories, which is a confirmation of the trends in the PCA score plots in Fig. 4-6 and 8.

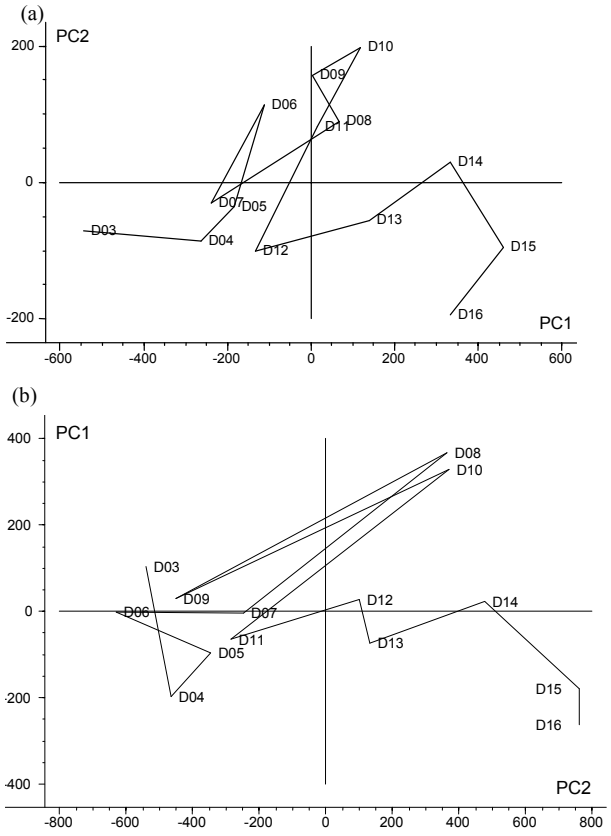


Fig. 7. PCA score plots of fluorescence spectra of weekly samples from factory D (PC1 vs. PC2). (A) Transmission samples, (B) reflection samples. The time trend in factory D is shown by connecting the weekly samples in a chronological order.

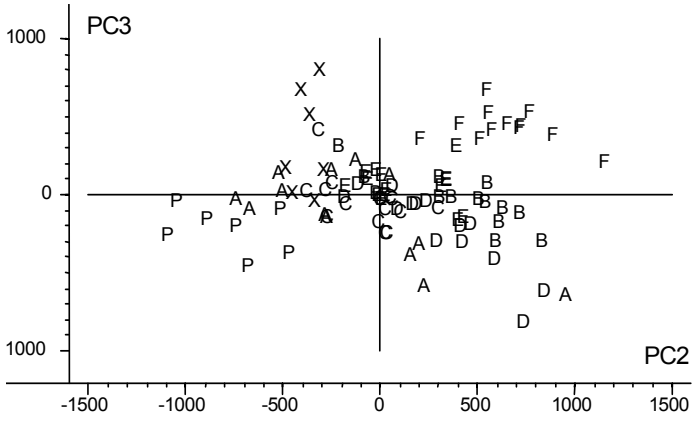


Fig. 8. PCA score plot of the fluorescence spectra of the 97 reflection solid sugar samples (factory A-F, X and P from factory E). PC2 vs. PC3 explains 18% and 7% of the total variance, respectively.

Prediction of quality parameters.

The two solid-state fluorescence data sets were used to build calibration models to predict the quality parameters color, ash and α-amino-N. The previous calibration results performed on the solution fluorescence data[2] were validated by splitting the sample set into two approximate halves using one half to calibrate the model and the other as a prediction set. Since the sample set consist of seven different kinds of sugar (A-F, P and X) with few samples from each (between 7 and 16 samples), it is difficult to divide the sample set into two subsets that span the same variation. Therefore segmented cross-validation has been used to validate the calibration models built on the transmission and reflection fluorescence data. The scores of **X** (the raw fluorescence matrix) plotted against the scores of **y** (the quality parameters) for each PLS factor were used in the initial calibrations to find **X**-**y** relational outliers. The errors in the model are expressed as RMSEP (root mean square error of prediction)[2] using test set validation or RMSECV (root mean square error of cross-validation) using cross validation. RMSECV is calculated as:

$$\text{RMSECV} = \sqrt{\frac{\sum_{n=1}^{N} \text{RMSEP}_n^2}{N}}$$

where N is the number of segments.

The PLS modeling results of color are summarized in Table II. With the use of transmission and reflection fluorescence data without pretreatment of the spectra, it was possible to model color with the same level of RMSECV for both data sets with the use of six PLS factors. The fluorescence spectra were then normalized prior to the calibration and solid sample models with five PLS factors could be made with maintained modeling error (Table II). In Fig. 9A and 9B plots are shown of the predicted color vs. the measured reference color for the 5 PLS factor models based on normalized fluorescence spectra of the transmission and reflection method, respectively. The effect of normalizing the spectra seems to remove a measurement variance. This variance is supposedly related to the heterogeneity of the solid sample. Modeling results of the sugar solutions are also presented in Table II. The model parameters of the normalized solid sample models equal the results of the solution model.

TABLE II. Results of the PLS calibration for color

Sample set	Validation[a]	Spectral pretreatment	Range[b]	Mean[c]	# PC[d]	R[e]	RMSEP/ RMSECV[f]
Solution[g]	Test set	None	11-44	21.8	5	0.94	2.4
Transmission	CV	None	6-36	19.8	6	0.93	2.3
Reflection	CV	None	6-41	20.2	6	0.93	2.5
Transmission	CV	Normalization	6-36	19.9	5	0.93	2.3
Reflection	CV	Normalization	6-41	20.2	5	0.92	2.6

[a] CV = segmented cross-validation with seven segments.
[b] Range of the measured reference color.
[c] Mean of the measured reference color.
[d] # PC is the optimal number of PLS factors.
[e] R is the correlation between predicted color and measured color.
[f] RMSEP/ RMSECV is defined in the text.
[g] PLS results from Ref. 2.

The PLS modeling results of ash and α-amino-N are given in Table III and Table IV, respectively. Only results from normalized fluorescence data are shown. Whereas the transmission data show the best result for color prediction, the reflection data can model ash and α-amino-N with fewer PLS factors. α-amino-N fluorescence is detected in the ultraviolet area of the emission range of the fluorescence measurements.[16] This area is more intensity quenched for the transmission data than the reflection data (see Fig. 2), which means that more PLS factors are needed to obtain an acceptable transmission model of α-amino-N. The ash content has shown correlation to the α-amino-N level,[2] which may explain the need for more PLS factors in the transmission model of ash, as well. The transmission and reflection modeling error for ash is in the range of the solution modeling error (0.0015, 0.0016 and 0.0012, respectively). The α-amino-N modeling error, on the other hand, is almost twice as high (0.40, 0.42 and 0.266, respectively). The lower accuracy of the solid sample models of α-amino-N in comparison to the solution model seems to be related to the intensity quenching of the ultraviolet spectral area.

The models of color, ash and α-amino-N developed for the three different kinds of fluorescence spectra (transmission, reflection and solution) all have reasonable correlation coefficients, but the model errors (RMSEP/RMSECV) are too large when compared to the range and mean of the reference values (Table II-IV). Better models with smaller errors should be obtained if more data material is collected, i.e. more samples from each factory are collected through several years of sugar campaigns.

TABLE III. Results of the PLS calibration for ash[a]

Sample set	Validation	Spectral pretreatment	Range	Mean	# PC	R	RMSEP/ RMSECV
Solution[b]	Test set	None	0.004-0.017	0.011	3	0.91	0.0012
Transmission	CV	Normalization	0.003-0.018	0.010	6	0.90	0.0015
Reflection	CV	Normalization	0.003-0.018	0.011	4	0.90	0.0016

[a] Table explanations are given in Table II.
[b] PLS results from Ref. 2.

TABLE IV. Results of the PLS calibration for α-amino-N[a]

Sample set	Validation	Spectral Pretreatment	Range	Mean	# PC	R	RMSEP/ RMSECV
Solution[b]	Test set	None	0.28-4.91	2.63	4	0.98	0.27
Transmission	CV	Normalization	0.28-5.07	2.71	6	0.94	0.40
Reflection	CV	Normalization	0.28-5.07	2.71	5	0.94	0.42

[a] Table explanations are given in Table II.
[b] PLS results from Ref. 2.

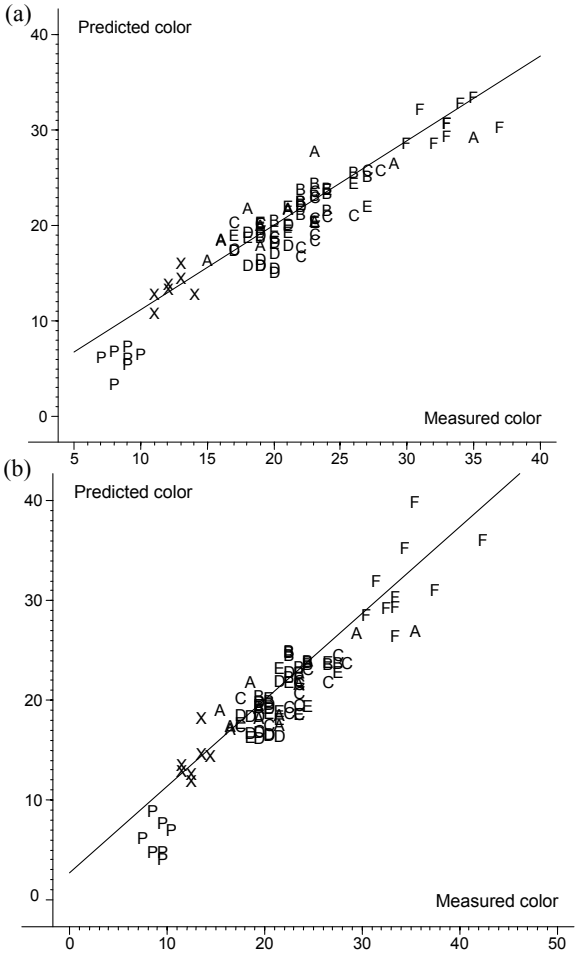


Fig. 9. Predicted color vs. measured color plots of five-factor models with normalized fluorescence data using cross-validation with seven segments. (A) Transmission samples, R = 0.93, RMSECV = 2.3, (B) reflection samples, R = 0.92, RMSECV = 2.6.

Conclusions

It has been shown that fluorescence spectra obtained from solid-state measurements on sugar crystal samples contain specific information concerning the characteristics of the chemical quality as well as a factory imprint of raw material and process conditions. Principal component analysis on the fluorescence spectra can be used to withdraw factory-related information from the samples. Multivariate calibration models are developed for three important sugar quality parameters: color, ash and α-amino-N. When comparing with sugar solution results, it is found that the solid sampling techniques have an important influence on the information obtained from the spectra.

Light scattering and self-absorption in the heterogeneous material affect the transmission method. The reflection method, on the other hand, is influenced by reflected scattering light and the size and composition of the surface area measured. It is important for the further exploitation of the solid sample measurements to have a better understanding of the interactions between the solid sugar and the propagating excitation and emission light and the way these interactions affect the fluorescent response. The size of the sugar crystals and the packing of the crystal sample in the sample cell both seem to have a large effect on the spectral output and we will examine these issues more closely in a future paper.

Acknowledgments

The authors thank Ole Hansen, Lars Bo Jørgensen and John Jensen, Danisco Sugar Development Center, Nakskov, Denmark for providing the sample set and routine data for calibration analyses.

References

1. L. Munck, L. Nørgaard, S. B. Engelsen, R. Bro and C. A. Andersson, Chemom. Intell. Lab. Systems **44**, 31 (1998).

2. L. Nørgaard, Zuckerind. **120**, 970 (1995).

3. B. L. Van Duuren, C. E. Bardi, Anal. Chem. **35**, 2198 (1963).

4. S. Å. Jensen, L. Munck and H. Martens, Cereal Chem. **59**, 477 (1982).

5. J. A. Olmstead and D. G. Gray, J. Pulp Paper Sci. **23**, J571 (1997).

6. L. Nørgaard, Talanta **42**, 1305 (1995).

7. L. Nørgaard, Chemom. Intell. Lab. Systems **29**, 283 (1995).

8. H. Martens and T. Næs, *Multivariate Calibration* (Wiley, New York, 1993), 2nd ed.

9. S. Wold, K. Esbensen and P. Geladi, Chem. Intell. Lab. Systems **2,** 37 (1987).

10. A. Höskuldsson, J. Chemom. **2**, 211 (1988).

11. P. J. Brown, Anal. Proc. **27,** 303 (1990).

12. P. Geladi and B. R. Kowalski, Anal. Chim. Acta **185,** 1 (1986).

13. P. Geladi and B. R. Kowalski, Anal. Chim. Acta **185**, 19 (1986).

14. R. J. Hurtubise, "Luminescence from solid surfaces", in *Molecular Luminescence Spectroscopy: Methods and Applications – Part II*, S. G. Schulman, Ed. (John Wiley and Sons, New York, 1988), Chap. 1, pp. 1-77.

15. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Plenum Press, New York, 1983), Chap. 2, p. 30.

16. O. S. Wolfbeis, "The Fluorescence of Organic Natural Products", in *Molecular Luminescence Spectroscopy: Methods and Applications – Part I*, S. G. Schulman, Ed. (John Wiley and Sons, New York, 1985), Chap. 3, p. 178.

Paper V

Analysis of the effect of crystal size and color distribution on fluorescence measurements of solid sugar using chemometrics

D. Baunsgaard, L. Munck and L. Nørgaard

Abstract

Fluorescence from sugar crystal samples has previously been used to obtain information about factory imprint and sugar quality. Solid-phase fluorescence has potential as a fast screening method but the spectra are highly influenced by the measurement geometry and sugar crystal sample. The aim of the present study was to examine how the fluorescence measurements are related to the sugar crystals for a better understanding of both. Initially, five sugar samples of varied composition were sieved into five crystal size fractions. Fluorescence excitation-emission landscapes of the fractions were measured with solid transmission and reflection techniques and in solution. The transmission fluorescence was quenched at ultraviolet wavelengths and light scatter highly influenced the reflection fluorescence. Principal component analysis (PCA) showed that large crystals favored the transmission fluorescence whereas smaller crystals improved the reflection fluorescence measurements. The multi-way method PARAFAC was used to resolve spectra of individual components from the fluorescence landscapes. Transmission and solution components had similar spectral profiles at higher wavelengths characterizing a colorant and a colorant intermediate. The resolved components of the reflection data were very influenced by scatter. Color predictions based on a few significant wavelength variables equaled the model results of full-spectrum models using partial least squares regression (PLS). The variables corresponded to wavelength maxima of the resolved colorants and ultraviolet wavelengths characterizing colorant precursors.

INDEX HEADINGS: Fluorescence; sugar crystals; particle size; color; principal component analysis (PCA); partial least squares (PLS); multi-way analysis (PARAFAC)

Introduction

Fluorescence spectroscopy is a useful technique in analyzing biological samples and food products due to its selectivity and sensitivity.[1,2] Fluorescence analyses are ideally performed on diluted homogeneous solutions but often it is desirable to measure solid samples directly without changing the environment of the fluorophore, e.g. in conformational studies involving proximity and interactions of molecules in the sample material.[3] Solid sample measurements also have the advantage of a non-destructive sample treatment, which is important when developing fast analysis methods to be used on-line or at-line during sample processing.[4] Fluorescence measurements of solid samples are more difficult to implement than measurements of solution samples, because undesirable effects such as light scattering and concentration quenching are more pronounced in the solid sample matrix.[5] Therefore, papers found in the literature have

mainly dealt with fluorescence measurements of sample solutions solidified in transparent gels (semi-solids)[3,6] or highly ground powders such as flour,[7,8] where light scatter and quenching can be minimized with the use of very narrow quartz cells and front-surface reflection techniques.

In a previous study of fluorescence measurements on sugar crystal samples, we measured fluorescence directly on solid sugar samples without sample pretreatment or sample cell optimization.[9] Samples from six different sugar factories were measured using two different optical arrangements; transmission detection at right angles through the sample and a reflection front surface detection. Multivariate data analyses of the fluorescence data were used to extract information about the chemical quality and a factory imprint of raw material and process conditions. Classification of sugar samples according to factory and calibration models of the quality parameters color, ash, and α-amino-N were achieved with fair results but not as conclusive as similar analyses already performed on fluorescence data of the same sugar samples in solution.[10] It was concluded that light scattering and self-absorption in the heterogeneous material affected the transmission method, whereas the reflection method was influenced by reflected scattering light and the size and composition of the surface area measured.

Thus, the model results show that solid sample fluorescence has promise as a way of gaining process information directly in the production but that more understanding of the problems related to the fluorescence measurements techniques contra sugar crystal samples is needed. Solid sample fluorescence may even provide information of the crystal composition, which is lost when the crystals are ground or dissolved. The amount of color in the sugar crystal as well as the distribution of various types of colorants in the inside of the crystal and in the syrup film surrounding the crystals are important issues for the sugar industries[11,12] that could benefit from spectral information obtained from intact crystals. It is known from near- and mid-infrared spectroscopy that the particle characteristics such as size, shape and surface conditions has a direct influence on spectroscopic measurements of sucrose particle samples.[13] Therefore, in this work a sample set of five different sugar samples were separated in several particle sizes by a sieving method and full fluorescence excitation-emission landscapes of the sample fractions were recorded using the transmission and reflection solid fluorescence techniques as well as measuring the fractions in solution. The fluorescence data was afterwards processed using two-way and three-way chemometric models and the effect of the measurement geometry in combination with crystal composition on the spectral information is discussed.

Experimental

Samples.

Five sugar samples were collected from three sugar factories. Three of the samples (A, B, and D) were produced sugar from each of the factories. The two remaining samples were special sugars. The C sample was a sugar processed without added SO_2 from the same factory as the B sample and the E sample was a refined sugar from the factory of the D sample. Sugar without added SO_2 contains more color than normally found in processed sugar and refined sugar has less color. The last two sugars were used to introduce a larger variation of color in the sample set.

Sieve fractions.
A JEL 200/65 horizontal sieve shaker (J. Engelsman Akt. Ges., Ludwigshafen, Germany) was used to separate the sugar samples into fractions according to sieve mesh size using a standard procedure from Danisco Sugar A/S. 100 g of a thoroughly but carefully mixed sugar sample was applied to a sieve stack, which included five sieves of mesh size 800 mm, 630 mm, 500 mm, 315 mm, 160 mm, respectively, and a bottom pan. The total sieve time was set at 10 min. The % weight distributions of the particle size ranges fractionated by the sieves have been listed in Table I.

TABLE I. Particle size distribution[a] of sugar fractions

Fraction #	Fraction size (μm)	Sample				
		A	B	C	D	E
8	> 800	14.8	24.2	45.9	12.4	11.0
6	630-800	29.8	33.2	27.3	27.9	17.7
5	500-630	20.7	19.0	12.1	24.9	16.1
3	315-500	23.3	18.7	11.6	26.6	29.1
1	160-315	10.6	4.7	2.9	7.9	24.1
-[b]	< 160	0.8	0.2	0.2	0.3	2.0

[a] % by weight
[b] The lowest fraction was not used in the data analyses due to inconsistency in the measured spectral data.

Reference color measurements.
The quality parameter color of the sugar sample fractions was determined according to the ICUMSA (International Commission for Unified Methods of Sugar Analysis) standard method. The sugar sample is dissolved in water to a 50% refractometric dry matter solution and the absorbance is measured in a 10 cm quartz cell at 420 nm. The instrumentation used was a Perkin-Elmer Lambda 900 spectrometer. The color value is calculated by multiplying the absorbance with a constant, which depends on the dry matter value. Three replicates for each of the five sugar samples were measured. The average color values are given in Table II.

TABLE II. Color of the sugar fractions.[a]

Fraction #	Sample				
	A	B	C	D	E
8	15.6	15.1	25.1	21.2	6.8
6	13.8	18.3	21.7	18.6	6.1
5	15.7	15.5	21.8	17.2	7.1
3	16.7	17.0	21.6	18.4	6.5
1	18.9	16.6	25.3	23.4	9.4

[a] Average ICUMSA color of three replicate measurements.

Fluorescence measurements.

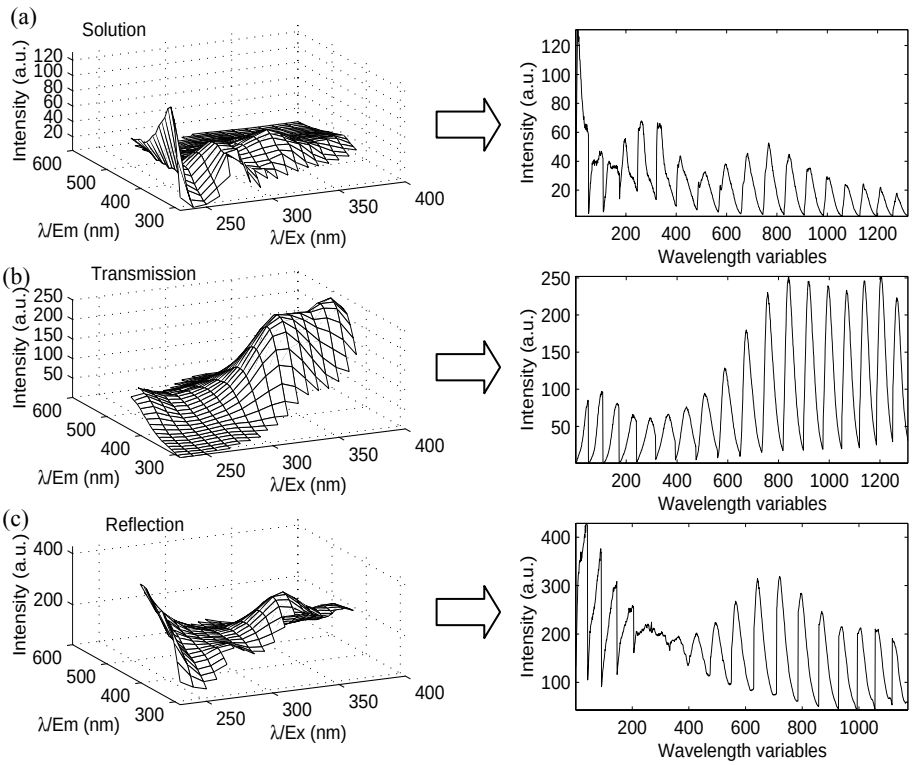


Fig. 1 Left side: fluorescence excitation-emission landscapes with 18 excitation wavelengths (230-400 nm, step 10 nm) and 106 emission wavelengths (285-600 nm, step 3 nm) of sugar sample C, fraction # 8 measured (a) in solution, 2.25 g/15 ml, (b) with solid-phase transmission, and (c) with solid-phase reflection. The blank regions hold Rayleigh scatter signals and are thus treated as missing values. Right side: corresponding unfolded two-dimensional landscapes used in the two-way data analyses. The emission spectra are arranged in ascending order of the 18 excitation wavelengths.

A Perkin-Elmer LS50 B luminescence spectrometer was used for the measurements with a right angle cell for the transmission measurements of solution and crystal samples and a front surface accessory for the reflection measurements of crystal samples. Fluorescence landscapes were recorded with 230-400 nm (10 nm step) as the excitation wavelength range and 280-600 nm as the emission wavelength range. Details and specifications of the fluorescence system and sample cell geometries have previously been published.[9] On the left side in Fig. 1 the fluorescence landscapes of one of the sample fractions (sample C, fraction 8) measured in solution, with solid-phase transmission and solid-phase reflection, respectively, are shown. Some of the areas in the landscapes do not conform to true fluorescence response, such as the Rayleigh scattering

peaks, and they are handled as missing intensity areas, i.e. the white areas in the plot. On the right side in Fig. 1 the corresponding unfolded landscapes are shown. The unfolded landscapes were generated by arranging the emission spectra as one vector in ascending order of the 18 excitation wavelengths and subsequently removing missing data areas.

Multivariate data analysis.
The two-way models principal component analysis (PCA),[14] partial least squares regression (PLS),[14] and the principal variables method (PV)[15] are used on two-way unfolded fluorescence data (samples x wavelength variables). The data was mean centered before modelling. Short mathematical descriptions of the three two-way methods are found in Ref. 10.
Maintaining the three-way structure of the fluorescence landscape data from several samples (samples x excitation wavelengths x emission wavelengths) has the advantage that decomposition models may be used which give an easier interpretation of the resulting loadings. Whereas the orthogonal loadings from PCA and PLS models of the unfolded fluorescence data are difficult to interpret, the multi-way decomposition model PARAFAC (parallel factor analysis)[16] decomposes three-way fluorescence data into spectral excitation and emission profiles in terms of pure components due to the uniqueness of the mathematical solution.[17] PARAFAC performs a tri-linear decomposition of the data array, **X,** and it is regarded as a generalization of the bi-linear PCA. In PARAFAC the tri-linear model of **X** is found to minimize the sum of squares of the residuals in the model, e_{ijk}, in the model:

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

where x_{ijk} is an element in **X** (i = 1,…, I, j = 1,…, J, and k = 1,…, K), a are the scores, b and c are loadings in the model and F is the number of factors. In contrast to PCA, the number of factors to be included in the PARAFAC model has to be determined before the modelling. Under the assumption of tri-linearity of the fluorescence data used in the PARAFAC model, i.e. no spectra are linearly dependent on any of the others and the concentrations of the fluorophores vary independently in the samples, the PARAFAC estimates are directly related to the true underlying spectra. For a more thorough presentation of the PARAFAC model, the reader is referred to a tutorial.[18] In this work no centering or scaling of the fluorescence data were used in the PARAFAC models, which were estimated under a non-negativity constraint to improve the interpretability of the resolved spectra.

Software for the data analysis.
Calculations were performed with Matlab for Windows version 5.3 (The MathWorks, Inc.) and Unscrambler version 7.5 (CAMO ASA). The PARAFAC results have been obtained with the use of the *N*-way Toolbox for MATLAB.[19]

Results and discussion

Effect of the crystal size on the solid sample fluorescence.

The three-dimensional fluorescence landscapes and the corresponding unfolded two-dimensional spectra of one of the sugar fractions measured with the three different methods in Fig. 1 demonstrate the spectral differences caused by the sample composition and choice of optical arrangement.[9] The fluorescence of the solid transmission in Fig. 1b is quenched in the ultraviolet wavelength area in comparison with the fluorescence of the solution samples in Fig. 1a. A high optical density at lower wavelengths (200-240 nm) causes high absorption of the transmitted excitation light and the emitted fluorescence, which strongly reduces the amount of fluorescence getting through the sample to the detector side. The transmission fluorescence at higher wavelengths, on the other hand, has good intensity and some peak maxima are visible in the fluorescence landscape. It is characteristic for transmission through solid samples that since the concentration of absorbing species cannot be controlled, the intensity maximum at higher wavelengths will predominate.[20] The reflection fluorescence in Fig. 1c has more the landscape profile of the solution fluorescence in the ultraviolet wavelength area because of less influence from filter effects but the detected emission is much affected by light scatter. Backscatter from the crystal samples increases the amount of stray light reaching the detector and smoothes the curves and details in the detected fluorescence into a less distinctive landscape.

Since the measured fluorescence from the two geometrical arrangements is so highly influenced by the sugar crystals, the sieved fractions of the five sugar samples were examined to gain more insight into the effect of crystal sizes. In Table I the particle size distributions of the sieved samples are shown. The sugar samples were divided into six fractions. The C sample without added SO_2 has relatively large fractions of the biggest crystals, whereas the refined E sample has relatively large fractions of the smallest crystal sizes. The lowest fraction (<160 μm), which represents less than 1 % of the total crystals except in the refined sugar sample (E), had to be omitted from the subsequent analyses. The spectra and color measurements from the collection of small crystals and chips of bigger crystals, which ended up in the bottom pan of the sieve stack, were not reproducible.

Principal component analyses (PCA) of the unfolded fluorescence data of the five remaining fractions of the sugar samples were performed. Three-dimensional score plots of principal component 1 versus principal component 2 and principal component 3 of the solution, transmission, and reflection data are shown in Fig. 2a-c, respectively. In the score plot of the solution samples in Fig. 2a the sieving fractions are primarily separated in groups of the five sugar samples, A-E, and a distribution according to fraction size is only slightly indicated, e.g. the D sample fractions. In the score plot of the transmission data in Fig. 2b and reflection data in Fig. 2c, the distribution of sample fractions according to size is the dominating factor. The fractions are distributed according to size along PC1 in opposite directions in the two plots. The corresponding loadings confirmed that PC1 explains the differences in spectral intensity of the fractions (plots not shown). It was possible to plot the solid fluorescence fractions in groups of the sugar samples as in Fig. 2a but this involved higher principal components than PC3 explaining less than 1% of the total variation (plots not shown).

The score plots demonstrate that the solid fluorescence measurements are highly influenced by the transparency of the sugar crystals and how compact the sample is. The larger the crystals, the more light will pass through the crystals and between them and consequently more emission will reach the detector in the transmission measurements. Conversely, in the reflection measurements the transparency of the sugar crystal sample reduces the fluorescence emission reflected to the detector and the largest crystal fractions have the lowest fluorescence intensity. In general the reflection technique appears to be a less advantageous way to measure very transparent solid particle samples.

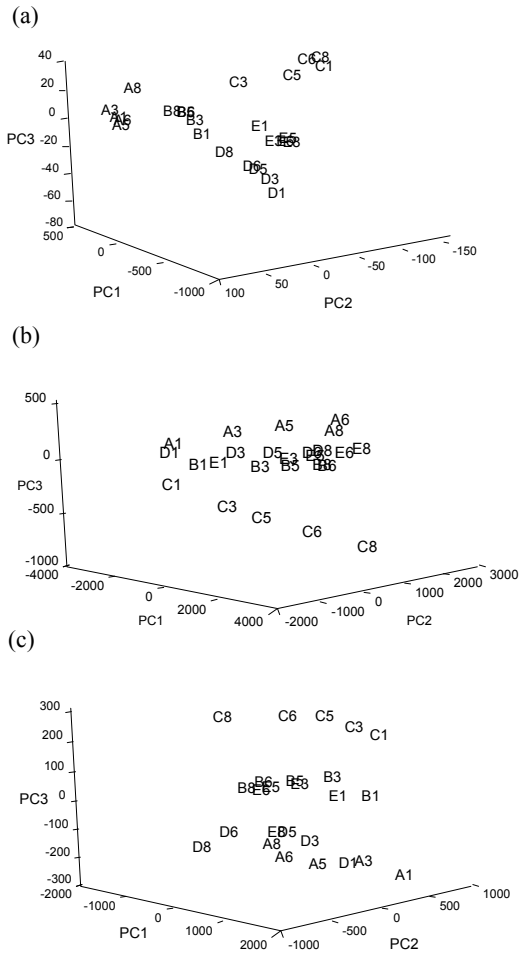


Fig. 2 Three-dimensional PCA score plots of principal component 1 (PC1) versus principal component 2 (PC2) and principal component 3 (PC3) of unfolded fluorescence data of the five sugar samples (A-E) in five sieving fractions (# 8, 6, 5, 3 and 1). (a) Solution, PC's explain 95, 4, and 1% of the total variance; (b) transmission, PC's explain 73, 24, and 2% of the total variance; (c) reflection, PC's explain 83, 12, and 4% of the total variance.

Color in the sugar crystal.

In the previous paper about fluorescence measurements of solid sugar samples, partial least squares regression (PLS) models of quality parameters based on the fluorescence data showed that color could be predicted with better model results than ash and α-amino-N when compared with models based on solution data.[9] The information in the fluorescence related to color is apparently well preserved despite the side effects from the solid sample. Table II lists color of the various sugar fractions. The five samples show different color levels, which as expected is highest in the C sample without added SO_2 and lowest in the refined sugar (E). The major role of SO_2 in the sugar processing is to inhibit both the enzymatic and the non-enzymatic browning reactions thereby decreasing color in the sugar.[21] Consequently, color in the C sample is approximately 7 units higher than color in the B sample from the same factory. PCA of the color fractions could not show any correlation between color and crystal size (plots not shown).

More information about color in the crystals may be obtained from the fluorescence of the solid sample fractions. It has been shown that it is chiefly colorants and color precursors, which causes the fluorescence in sugar.[22] Therefore, the fluorescence landscapes of the sample fractions were resolved into individual spectral components with the use of the decomposition model, PARAFAC. Three-dimensional data arrays are formed consisting of the 25 sample fractions in the first dimension (A-E sugars and fraction # 8, 6, 5, 3 and 1 of each sugar), 18 excitation wavelengths (230-400 nm, 10-nm step) in the second dimension, and 106 emission wavelengths (285-600 nm, 3-nm step) in the third dimension. The arrays were modelled with PARAFAC and suitable models of the solution, transmission and reflection data were fitted with four, three, and three components, respectively. The modelling results are shown in Fig. 3. Each component in a model is represented by the estimated excitation and emission spectra as well as a sample fraction profile, which shows the concentration of each component in the 25 sample fractions.

The four resolved components of the solution data in Fig. 3a show spectral shapes and wavelength maxima that are comparable to previous modelled spectra from other beet sugar samples.[22] The numbers shown to the far right of the emission spectra in Fig. 3 are used to compare the estimated components in the various models. The same number indicates emission spectra from the different models that show approximately the same spectral profile and wavelength position. The reason why emission spectra were compared rather than excitation spectra is because the latter spectra are less specific of individual fluorophores.

Component 1 (dash-dot) and component 2 (dot) in Fig. 3a have been identified as tyrosine and tryptophan, respectively, and component 4 (dash) has characteristics of a colorant from the Maillard reaction. Component 3 (solid) has not been associated with any known fluorophore. The three components resolved from the transmission data in Fig. 3b resemble components 2-4 in Fig. 3a. The PARAFAC model is not able to resolve the tyrosine as an individual component due to the low fluorescence intensities in the ultraviolet wavelength area and it appears only as a shoulder in the tryptophan emission spectrum. The tryptophan emission spectrum also displays a profile at higher wavelengths that is not seen in the corresponding solution emission spectrum. The resolved spectra are dependent on the tri-linearity of the fluorescence data and the premises of the PARAFAC model and the extension of the tryptophan component

demonstrates that the transmission data is not as tri-linear and therefore does not meet the premises of the model as well as the solution data. The resolved spectra in Fig. 3c of the reflection data resemble the solution spectra even less. Especially the emission profiles are affected and it confirms that the reflection landscapes are more influenced by light scatter and thus have less distinctive spectral curves than the other measurements. The resolved components marked 1.5 (dot) and "4" (dash) are not considered as true individual fluorophores but rather a combination of fluorescence and light scatter. However, the last component, "3" (solid), is very distinctive and suggests a combination of component 3 and partly component 4 from the solution and transmission models.

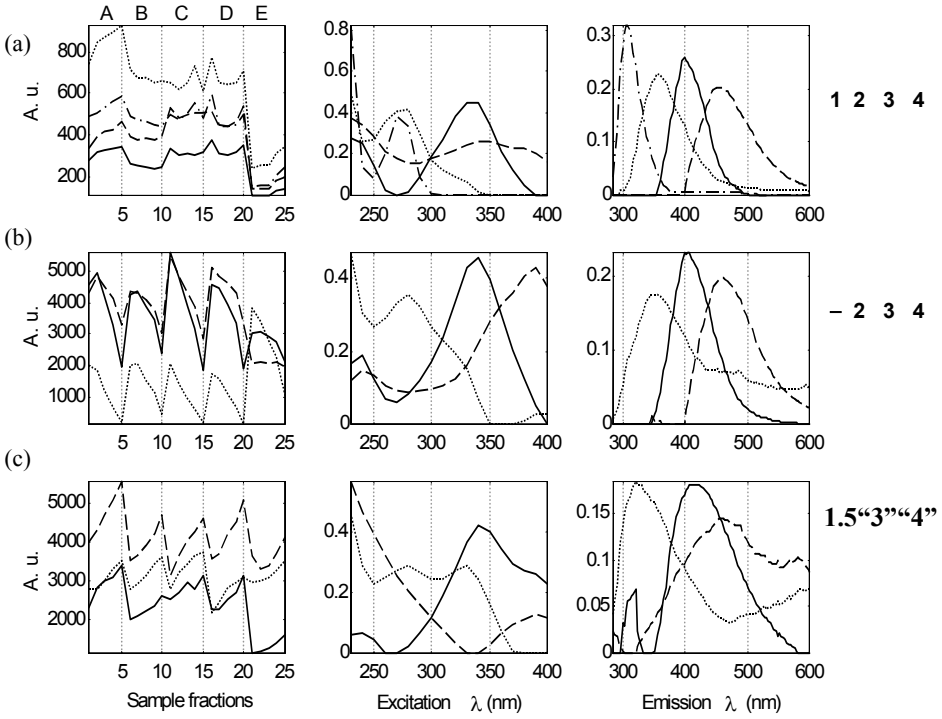


Fig. 3 Results of the PARAFAC models of the fluorescence landscapes of all the 25 sugar sample fractions. (a) Solution data, four components, (b) transmission data, three components, and (c) reflection data, three components. The left column displays the concentration of each resolved component in the five crystal fractions (# 8, 6, 5, 3, 1) of the five sugar samples. Sample A: fraction 1-5, B: fraction 6-10, C: fraction 11-15, D: fraction 16-20, and E: fraction 21-25. The center and right column display excitation and emission profiles, respectively. The line patterns indicate the corresponding profiles of each component in the three columns. The numbers to the far right are used to compare emission spectra of the resolved components in the three models; **1** (-·-), **2** (···), **3** (–), **4** (--). All spectral profiles have been normalized to unit length.

The sample fraction profiles of the three different measurement techniques in the first column in Fig. 3 show that the transmission profiles have decreasing intensity and the reflection profiles have increasing intensity with decreasing crystal size. This confirms the findings in the PCA score plots in Fig. 2. In the solution concentration profiles the ultraviolet components have the highest intensities due to the fact that the solution samples were diluted to a concentration where there was insignificant fluorescence concentration quenching,[9] and thus the highly fluorescent amino acids dominate the landscape (see Fig. 1a). The intensity of component 3 (solid) and component 4 (dash) in Fig. 3a is higher for sample C without added SO_2 than sample B from the same factory, which means that both components are involved in the formation of color. These findings support that component 4 is a likely Maillard colorant and that component 3 is a color reaction intermediate. The lower intensity of all four components in sample E fractions compared to sample D fractions in Fig. 3a shows that the concentration of all four components is reduced in the refining process.

The transmission sample fraction profile of tryptophan component 2 (dot) has a lower intensity than component 3 (solid) and component 4 (dash) in Fig. 3b except in the E sample fractions (21-25) and this is ascribed to the concentration quenching in the transmission measurements. The E sample has a lower concentration of all the fluorophores compared to the other samples, as demonstrated in Fig. 3a, and the concentration quenching in the ultraviolet wavelength area is accordingly diminished.

These dominating sample effects in the transmission model spoil any inference that might have been made about the location of the resolved fluorescent components in the crystal. In the reflection sample fraction profiles the apparent contribution from light scatter renders it difficult to make any conclusive observations except for the very low contribution of component "3" (solid) in the refined sample (E). Since the two other resolved components have intensities as high as in the unrefined sample (D), it is unlikely that the low intensity of component "3" was caused by the transparency of the crystal sample. In the refinery process the remelting and purification steps remove color through the crystal,[23] and this suggests that component "3" represents color, which is easily removed during refining.

Significant fluorescence wavelengths in the color prediction.

The principal variables method (PV) was used to select excitation-emission wavelength pairs (wavelength variables) from the unfolded fluorescence data of the solid sugar fractions to provide information about which excitation and emission wavelengths that are important in describing color to confirm the PARAFAC results. To begin with partial least squares regression (PLS) models of color based on the unfolded landscapes were made to estimate the optimal number of PLS-factors (Table III). The model results are better than PLS models of color made previously.[9,10] However, this data set consists of fewer samples and contains no seasonal variation. Relative comparisons between the models in Table III show that the solution model only use two PLS factors but the model results based on the transmission and reflection data are almost as good with four and five PLS factors, respectively. In the PARAFAC models of the solid fluorescence data in Fig. 3, three-component solutions were suitable. It may therefore seem surprising that the regression models of color based on the same data had to use more than three PLS factors. However, the loadings of the PLS factors reveal that one of the four transmission

PLS factors and two of the reflection PLS factors are used to explain the crystal sample influence on the fluorescence (plots not shown).

TABLE III. Prediction results for color[a] based on the unfolded fluorescence data

Sample set	# PC[b]	R[c]	RMSECV[d]
Solution	2	0.98	1.2
Transmission	4	0.94	1.8
Reflection	5	0.96	1.5

[a] Range: 6.1-25.3, mean: 16.3
[b] #PC is the optimal number of PLS factors
[c] R is the correlation coefficient between predicted color and measured color
[d] RMSECV is the model error using full cross validation

A number of wavelength variables corresponding to the number of PLS factors in the full-spectrum models were selected using PV and multiple linear regression (MLR) models for predicting color were made based on these variables (Table IV). A comparison with the full-spectrum models in Table III shows that the models made on the two solution variables and the five reflection variables had almost the same model errors as the full-spectrum models. The transmission model results were even improved using only the four variables. Some of the chosen variables correspond to wavelength maxima of the resolved excitation and emission spectra in Fig. 3. The (excitation,emission) wavelength pairs at higher wavelengths, e.g. (330, 414) in the solution model, (390,471) in the transmission model, and (340,417) in the reflection model, fit the excitation and emission maxima of components 3, 4, and "3", respectively. This confirms the suggestions of component 3 and 4 being color reaction intermediate and colorant, respectively. The selected wavelength variables in the ultraviolet wavelength area are less obvious. An examination of the loadings of the full-spectrum PLS-models revealed that most of the ultraviolet wavelength variables selected by the PV method were negatively correlated to color, which is characterizing color precursors.[24]

TABLE IV. Prediction results for color[a] based on a selected number of wavelength variables

Sample set	# wavelength variables[a]	R	RMSECV
Solution	2 (230,381), (330,414)	0.97	1.3
Transmission	4 (230,327), (230,402), (360,423), (390,471)	0.97	1.4
Reflection	5 (230,330), (230,333), (240,420), (270,489), (340,417)	0.96	1.5

[a] Excitation-emission (Ex,Em) wavelength variables selected from the unfolded fluorescence data by the PV method.

Conclusion

The sugar crystal structure is both an advantage and a disadvantage for solid fluorescence measurements. The transparency of the crystals favors the transmission measurements because the detected fluorescence contains the combined emission from a

large section of the crystals through the sample and the intensity has low signal-to-noise ratio. However, the high absorption of light in the crystals at the ultraviolet wavelengths quenches the emission, which is unfavorable for the detection of ultraviolet fluorophores. Increasing crystal sizes has the same effect as diluting solutions, i.e. reduced concentration quenching and increased fluorescence intensity. The reflection fluorescence is very influenced by detected scatter from the sugar crystal surface and much of the fluorescence emission, especially from less distinctive fluorophores, does not reach the detector due to the transparency of the crystals. Smaller crystals favor a higher intensity and more detailed fluorescence measurements.
The components resolved from the transmission and reflection fluorescence with PARAFAC do not resemble true individual fluorophores as well as the solution components. The non-spectral influences on the model results render it difficult to obtain information about the distribution of fluorophores in the sugar crystals. However, spectra of fluorescent colorants are resolved well using PARAFAC due to their high fluorescence intensity in the measured landscapes.
Color in the sugar crystal was predicted nearly as well with solid fluorescence as with solution fluorescence using a few selected wavelength variables. The chosen variables consist of spectral maxima of colorants and ultraviolet wavelengths representing color precursors.

Acknowledgments

The authors thank Ole Hansen, Lars Bo Jørgensen and John Jensen, Danisco Sugar Development Center, Nakskov, Denmark for providing the sample set and for valuable discussions. Rasmus Bro and Claus A. Andersson are acknowledged for their expert advice and the programming of the N-way toolbox.

References

1. W. Rettig, B. Strehmel, S. Schrader and H. Seifert, *Applied Fluorescence in Chemistry, Biology and Medicine* (Springer-Verlag, Berlin, 1999).

2. L. Munck, *Fluorescence Analysis in Foods* (Longman Singapore Publishers Ltd., Singapore, 1989).

3. C. Genot, F. Tonetti, T. Montenay-Garestier and R. Drapron, Sci. Aliments **12**, 199 (1992).

4. S. Å. Jensen, L. Munck and H. Martens, Cereal Chem. **59**, 477 (1982).

5. R. J. Hurtubise, *Solid Surface Luminescence Analysis: Theory, Instrumentation, Applications* (Marcel Dekker, Inc., New York, 1981), pp. 1-3.

6. M. del Olmo, J.Laserna, D. Romero, J. Rohand and J. L. Vilchez, Talanta **44**, 443 (1997).

7. C. Genot, F. Tonetti, T. Montenay-Garestier, D. Marion and R. Drapron, Sci. Aliments **12**, 687 (1992).

8. M. Zandomeneghi, J. Agri. Food Chem. **47**, 878 (1999).

9. D. Baunsgaard, L. Munck and L. Nørgaard, Appl. Spectrosc. **54**, 438 (2000).

10. L. Nørgaard, Zuckerind. **120**, 970 (1995).

11. N. W. Broughton, D. Sargent, B. J. Houghton and A. Sissons, *The inclusion of colour and ash components in U.K. beet white sugar*, Comptes Rendus 18th Assemblée Générale Commission Internationale Technique Sucrerie, Ferrara, Italy, 101 (1987).

12. M. A. Godshall, M. A. Clarke, C. D. Dooley and R. S. Blanco, J. Sugar Beet Res. **28**, 155 (1991).

13. J. M. Olinger and P. R. Griffiths, Appl. Spectrosc. **47**, 687 (1993).

14. H. Martens and T. Næs, *Multivariate Calibration* (Wiley, New York,1993), 2nd ed.

15. A. Höskuldsson, J. Chemom. **2**, 211 (1988).

16. R. A. Harshman, UCLA working papers in phonetics **16**, 1 (1970).

17. J. B. Kruskal, Linear Algebra Appl. **18**, 95 (1977).

18. R. Bro, Chemom. Intell. Lab. Systems **38**, 149 (1997).

19. C. A. Andersson and R. Bro, The N-way Toolbox for MATLAB, Ver. 1.04, November 1999, Food Technology, MLI, KVL, Denmark. http://www.models.kvl.dk/source/nwaytoolbox/

20. W. L. Butler and K. H. Norris, Arch. Biochem. Biophys. **87**, 31 (1960).

21. M. Shore, N. W. Broughton, J. V. Dutton and A. Sissons, Sugar Technol. Rev. **12**, 1 (1984).

22. D. Baunsgaard, C. A. Andersson, A. Arndal and L. Munck, Food Chem., **70**, 113 (2000).

23. H. Schiweck and M. Clarke, "Sugar", in *Ullmann's Encyclopedia of Industrial Chemistry* (VCH Verlagsgesellschaft, Weinheim, Germany, 1994), Vol. A 25, p. 384.

24. D. Baunsgaard, L. Nørgaard and Godshall, M. A., *Fluorescence of raw cane sugars evaluated by chemometrics*, J. Agri. Food Chem., accepted (2000).