# Fluorescence of Raw Cane Sugars Evaluated by Chemometrics

Dorrit Baunsgaard,[†] Lars Nørgaard,*[,†] and Mary An Godshall[‡]

Chemometrics Group, Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, 1958 Frederiksberg C, Denmark, and Sugar Processing Research Institute, Inc., New Orleans, Louisiana 70124

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 (maximum excitation wavelength, maximum 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 by use of HPLC analysis combined with fluorescence measurements on thick juice samples and evaluation 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

* Corresponding author: telephone +45 35 28 32 67; fax +45 35 28 32 45; e-mail lan@kvl.dk.

[†] The Royal Veterinary and Agricultural University.

[‡] Sugar Processing Research Institute, Inc.

10.1021/jf000131j CCC: $19.00 
Published on Web 09/07/2000

**Table 1. Raw Sugar Sample Set**

| no. | origin | year | color | no. | 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.

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 such as 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 collinear 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 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.

**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.

**Fluorescence Measurements.** The fluorescence measurements were performed on a Perkin-Elmer LS50 B fluorescence spectrometer. Nonsmoothed 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. Figure 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.** A two-way structure of the fluorescence landscapes (samples × 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 Figure 2. The missing areas shown in Figure 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

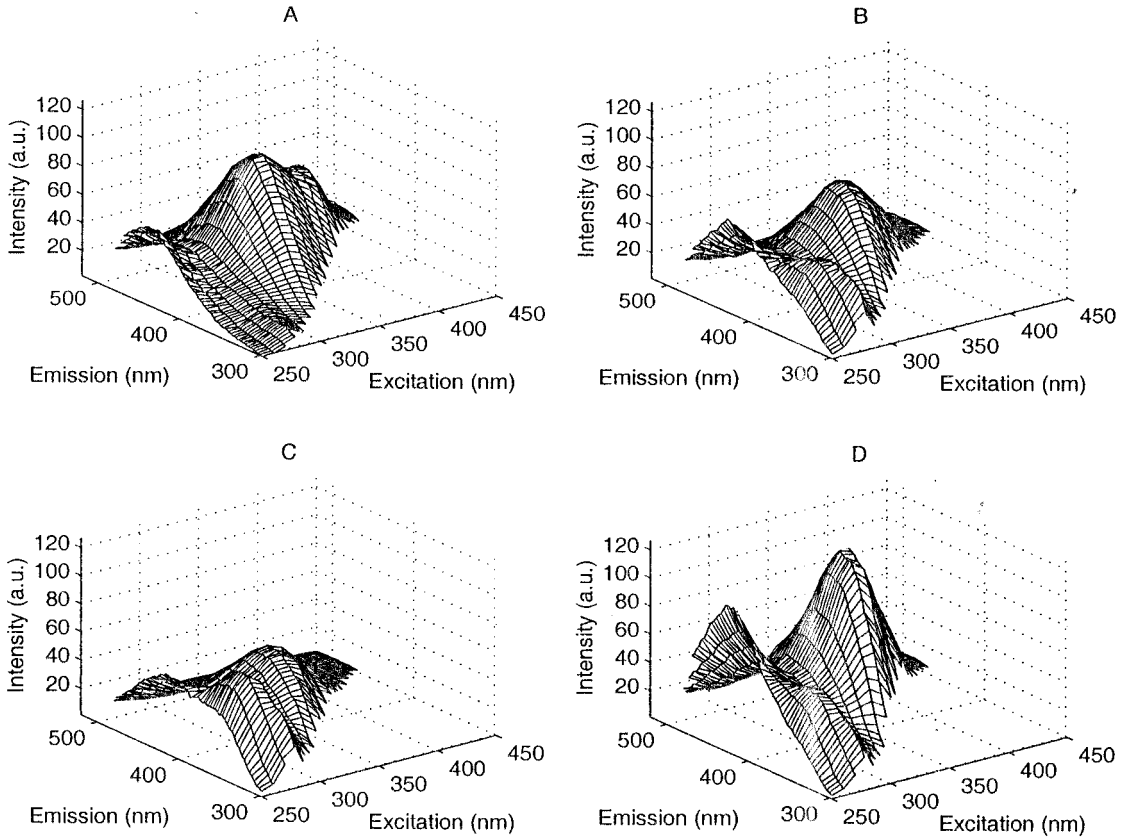


**Figure 1.** Plots of the excitation−emission fluorescence landscapes of four raw sugar samples from Table 1, which are each very different from one another. (A) Sample 1; (B) sample 20; (C) sample 42; (D) sample 43. The white areas in the landscape denote missing data areas due to Rayleigh scattering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

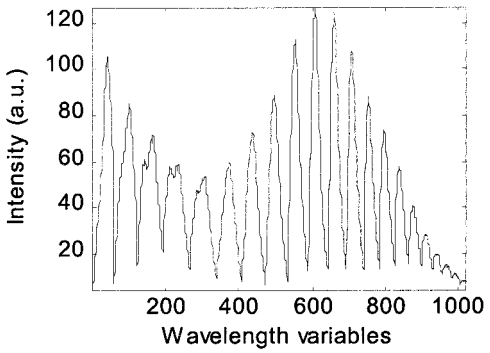


**Figure 2.** Example of an unfolded fluorescence landscape of a raw sugar sample (sample 43, Figure 1D). Emission spectra are arranged in the order of the 21 excitation wavelengths (in total 1021 wavelength variables).

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):

$$\mathrm{RMSECV} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(C_n^{\text{predicted}} - C_n^{\text{reference}})^2}$$

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 × excitation wavelengths × 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 counterparts (Bro, 1998). The multiway 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 nonnegativity constraint to improve the interpretability of the resolved profiles. In the model results presented here, the emission profiles have been estimated under a unimodality constraint to avoid the interference of artificial extra peaks in the spectra due to too many missing variables in the data set.

**Software for 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 Figure 3A a score plot of the first principal component (PC1) against the second principal component (PC2) of all the 47 samples is shown. The numbers in the plot correspond to the sample numbers in Table 1. The two components

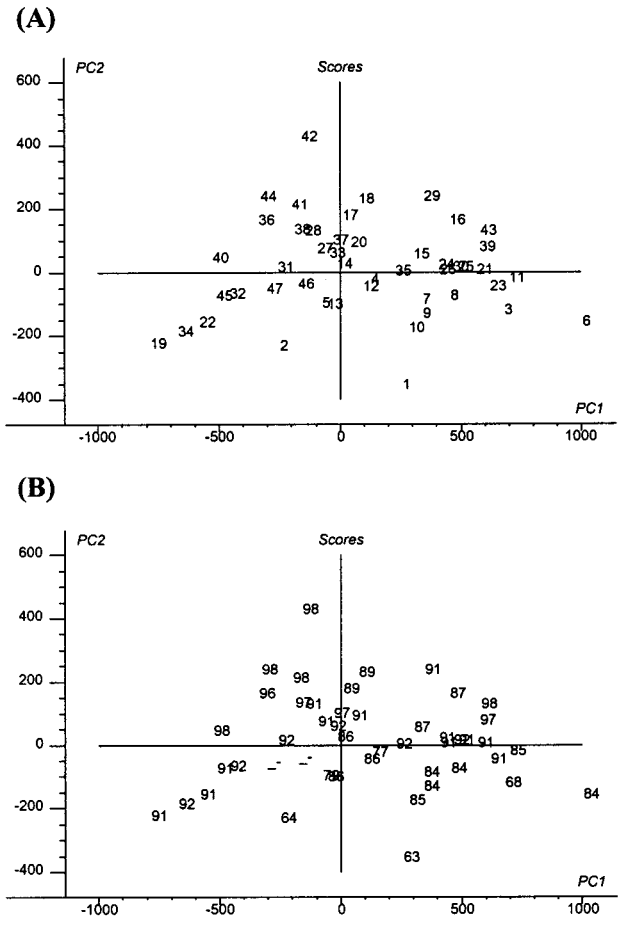


**Figure 3.** (A) Score plot of the first principal component (PC1) against the second principal component (PC2) of all 47 raw sugars. Numbers correspond to the raw sugars in Table 1. (B) Same score plot as in panel A, showing the year of production of each of the 47 raw sugars.

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, either in terms of the place of origin or 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 Figure 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 to 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 PCs 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 are important for each of the principal components. The loading vectors of PC1 and PC2 are shown in Figure 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 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 around (350, 425) nm is the dominating fluorescence in the raw sugar samples at the chosen concentration.

**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

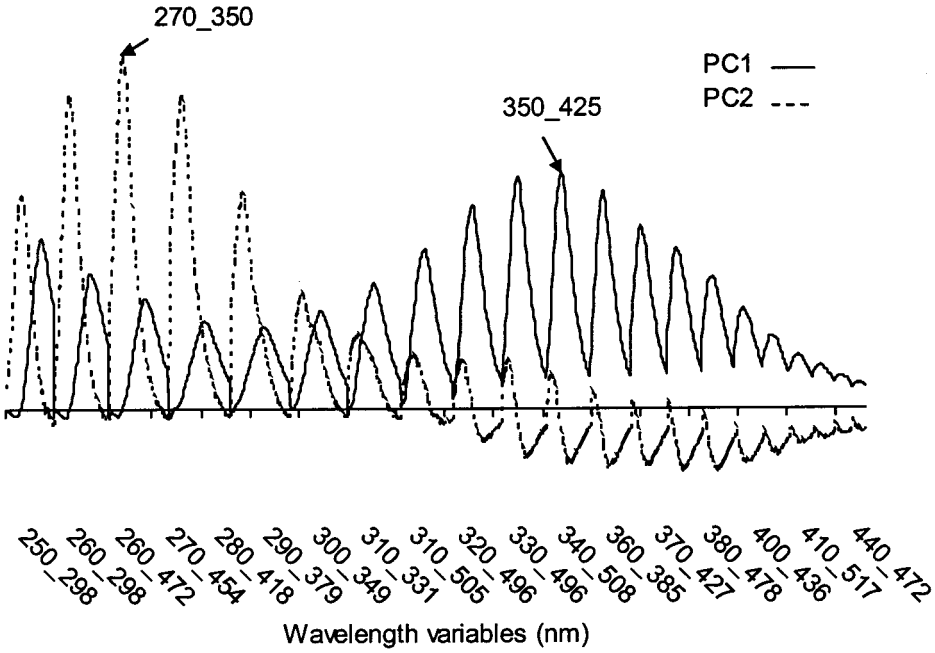


**Figure 4.** Loading vectors of the first two principal components of the 47 raw sugars as a function of wavelength variables. The wavelength variables are shown as excitation wavelength_emission wavelength.

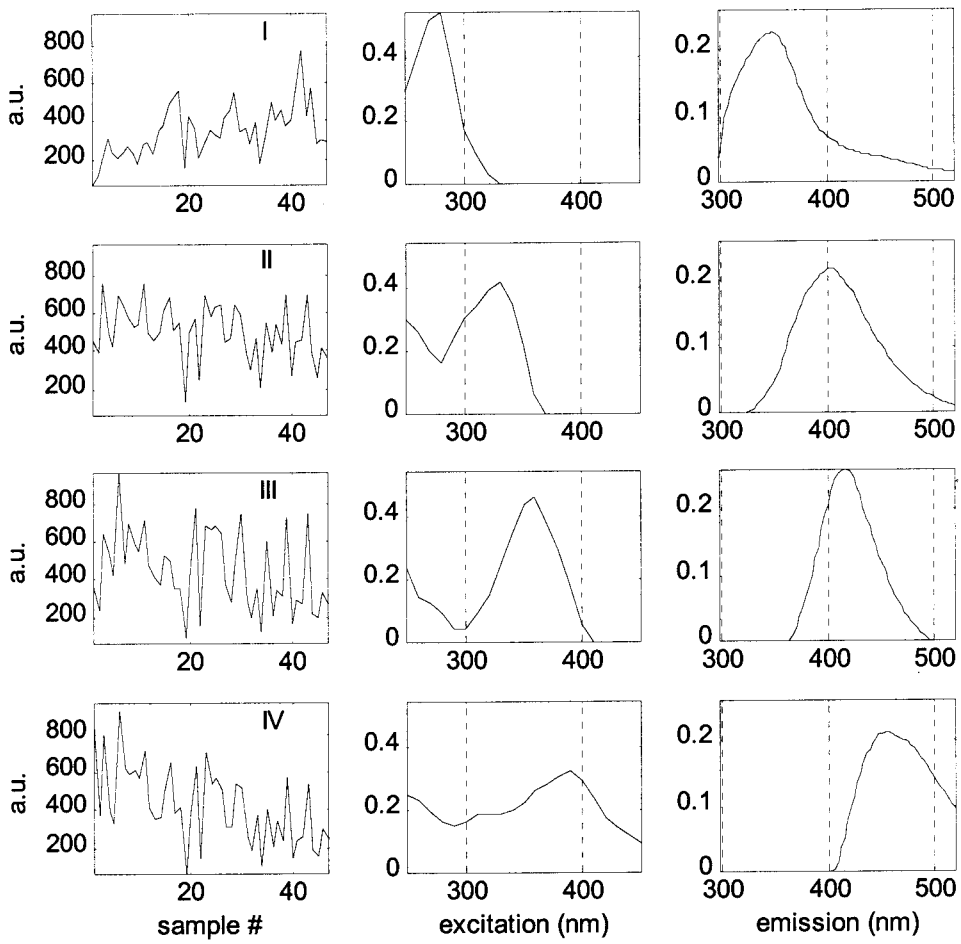


**Figure 5.** Results of a four-component PARAFAC model of the measured fluorescence landscapes of 47 raw sugar samples. The first column represents the scores of each sample for each component; the second column represents the excitation profile; the third column represents the emission profile. The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

**Table 2. Excitation and Emission Maxima of Four Fluorescence Components Estimated in the PARAFAC Model of the 47 Raw Sugar Samples**

| | $\lambda_{max}$ (nm) | |
|---|---|---|
| component | ex | em |
| I | 275 | 350 |
| II | 330 | 400 |
| III | 360 | 420 |
| IV | 390 | 460 |

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 Figure 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 components I and III are close to the wavelengths that were important for the two first principal components in the PCA. In Figure 6A a score plot of component I versus component III is shown, and the distribution of the samples is very close to that in Figure 3A. This confirms that the PCA and PARAFAC models have captured the same two major contributors to the fluorescence in raw sugar. Components I and 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 be correlated to component III. Plotting the scores of component III versus component IV (Figure 6B) reveals that the oldest samples (1−3) have a relatively 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 nonlinear 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 Figure 5 and they may represent two colorants of the same type. The fluorescence of the darker raw sugar samples tends to appear at somewhat higher wavelengths than the fluorescence of the lighter colorants, and this is caused either by small size

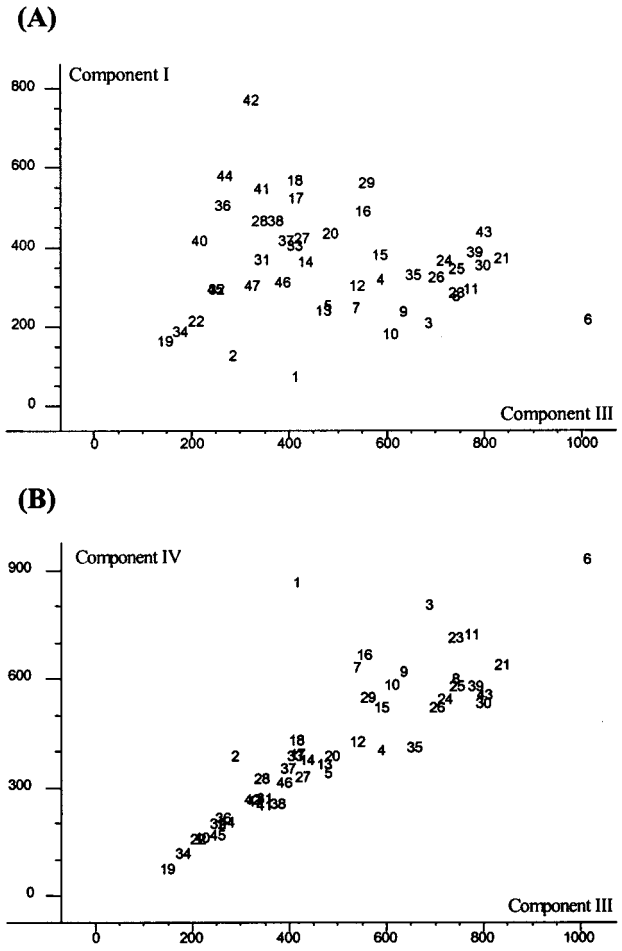


**Figure 6.** (A) Score plot of component I versus component III from the PARAFAC model. Note the similarity to the plot in Figure 3A. (B) Score plot of component III versus component IV from the PARAFAC model.

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 components II and III of the many overlapping but slightly different excitation and emission spectra.

**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 values 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

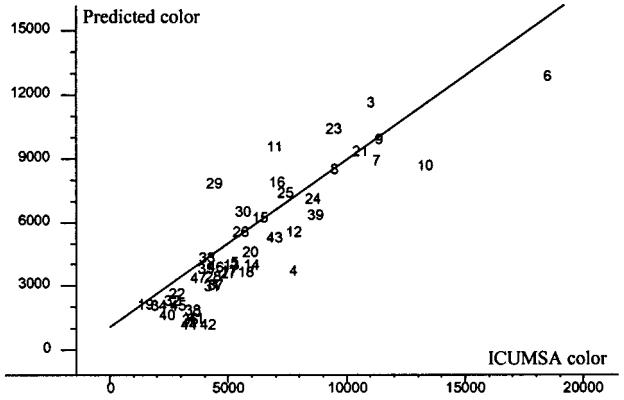


**Figure 7.** Plot showing the correlation of the predicted color of raw sugar from the PLS model of the fluorescence data using three PLS components and the ICUMSA color of 45 raw sugar samples. Correlation coefficient = 0.88 and the model error is 1622 in ICUMSA color units.

compositions of the samples. As a way to correlate the color of the raw sugar samples with 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 Figure 7. The model error of 1622 for the new PLS model is still high when it is considered that the range of the sample color has been lowered as well. PLS modeling of subsets including only samples with low color gave the same relative model error to the modeled 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 modeled 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,

**Table 3. Results of the PLS Models of the Fluorescence Data for the Prediction of Color Using Full Cross Validation**

| no. of samples | no. of variables | no. of 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] Number of 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.

decreases. The nonlinear 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 nonlinear effects occur. For example, 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.

**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 (Figure 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 Figure 7. Apparently, only the three selected wavelength variables are needed for modeling 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) correspond to combined components II/III, component I, and component IV, respectively. The fact that components II and III both correspond to one of the 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.

**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 spectra 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 Figure 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 that 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 Figure 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 Figure 1 where the landscape of sample 43 (Figure 1D), which has a high contribution from component II/III, has one peak at (420, 345) nm whereas sample 1 (Figure 1A) has an extra peak at (380, 450) nm 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 multiwavelength fluorescence data can determine the chemical characteristics of fluorophores in raw cane sugar without knowledge of their exact chemical structure. Three principal fluorophores are found in raw sugar and one of them is characterized as an ultraviolet color precursor that participates in color development during storage. The other two fluorophores fluoresce in the visible wavelength area and are potential colorants. The colorant 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. We 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, version 1.03, Food Technology, MLI, KVL, Denmark. http://www.models.kvl.dk/source/nwaytoolbox/

Baunsgaard, D.; Andersson, C. A.; Arndal, A.; Munck, L. Multiway 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 of the 1972 Technical Session on Cane Sugar Refining Research*; USDA−ARS: New Orleans, LA, 1972; pp 47−61.

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

Godshall, M. A. Part 2. Color Analysis. In *Symposium on sugar color,* Sugar Industry Technologists: New York, 1997; pp 8−29.

Harshman, R. A. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Pap. 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. *Zuckerindustrie* **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.

Received for review January 31, 2000. Revised manuscript received June 26, 2000. Accepted June 29, 2000.

JF000131J