Chemometric Analysis of Dietary Fibre Associated Compounds

Keld Ejdrup ANDERSEN [1], Charlotte BJERGEGAARD [1], Lars NØRGAARD [2], and Hilmer SØRENSEN [1]

[1] Chemistry Department, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Denmark

[2] Department of Dairy and Food Science, Food technology, Rolighedsvej 30, DK-1958 Frederiksberg C, The Royal Veterinary and Agricultural University, Denmark

INTRODUCTION

Dietary fibres (DF) includes a complex group of mainly plant cell wall derived compounds, which are non-digestible in the human small intestine. These compounds include non-starch polysaccharides (NSP), proteins, and lignins, and to a smaller extent various non-carbohydrate compounds such as various lipids and phenolics (Andersen *et al.*, 1997, 1998a). Certain physiological responses have been associated with the consumption of dietary fibres and the physical and chemical properties of each of the compounds of DF appear to be important in determining the physiological response to sources of DF in the diet (Schneeman, 1987). Various methods for the determination of DF exists, however, the analytical procedures are in general very time-consuming. To overcome this problem with time consuming DF determination, chemometric analysis call for attention. It is thus found from various publications concerning spectroscopic measurements on a wide span of foodstuffs for classification and prediction of quality parameters, that the data treatments have benefited form the use of chemometric, which in short can be defined as "the application of mathematical and statistical methods to extract reliable and relevant information from chemical data (Anonymous, 1986). This applies to moister (Forina *et al.*, 1995; Windham *et al.*, 1997), oil (Forina *et al.*, 1995), protein (Williams *et al.*, 1991; Forina *et al.*, 1995), starch (Williams *et al.*, 1991; Baker, 1995), total carbohydrate (Baker, 1995), and fibre (Williams *et al.*, 1991; Baker, 1995; Kays *et al.*, 1996; Windham *et al.*, 1997; Kays *et al.*, 1998).

The present study investigates the possibility of using chemometric methods on data from UV-VIS-spectra and micellar electrokinetic capillary chromatograms (MECC) of supercritical fluid extracts (SFE) from isolated total dietary fibre (TDF) fractions (Andersen *et al.*, 1997, 1998a, 1998b), in order to find relationships within the different plant families or genera from which the various starting materials are derived. Beside the chemotaxonomic use of chemometrics, data from two types of extracts (SFE A and SFE B) are investigated using partial least squares regression (PLS-R) in attempt to identify some of the chemical properties of the compounds analysed by MECC.

MATERIALS AND METHODS

TDF of 13 types of starting materials (Table 1) were supercritical fluid extracted (SFE) using pure CO_2 at 50 MPa, 75°C in 30 minutes (SFE A) followed by extraction using 15 % methanol in CO_2 at the same extraction conditions (SFE B). The extractions were performed in duplicates. The two types of extracts were analysed by UV-VIS-spectroscopy (190-800 nm) and micellar electrokinetic capillary chromatography (MECC). (Details on the procedures for TDF isolation, SFE, UV-VIS-spectroscopy, and MECC are presented elsewhere (Andersen *et al.*, 1997; 1998a; 1998b). UV-VIS-spectra as well as the normalised area of 30 compounds separated by MECC were read into The Unscrambler™ ver. 7.01 (CAMO A/S, Oslo, Norway). Principal component analysis (PCA)

Table 1. Taxonomic grouping of TDF-fractions.

FAMILY	GENUS	SPECIES	VARIETY	FRACTION	TDF
Cruciferous (Brassicáceae)	Rape (*Brássica*)	*B. napus* L. ssp. napus	Apex	Seed	RpS
				Protein	LIPRO
					PRM
				Hull	RpH
Legume (Fabáceae)	Lupine (*Lupinus*)	*L. lúteus* L.	Juno	Seed	LS
				Protein	LM
				Hull	LH
	Pea (*Pisum*)	*P. sativum* L.	Bodil	Seed	RP
			Kelwo	Seed	WP
Cereal (Poáceae)	Wheat (*Triticum*)	*T. aestivum* L.	Kraka	Seed	W
	Barley (*Hórdium*)	*H. vulgare* L.	Vega	Seed	B
	Rye (*Secále*)	*S. cerále* L.	Petrus	Seed	R
	Oat (*Avéna*)	*A. satíva* L.	Hedvig	Seed	O

(Martens and Næs, 1989) were performed on the spectra (52 × 1201 data matrix) and the chromatographic results (52 × 30 data matrix) respectively, followed by Partial Least Square Regression (PLS-R) (Martens and Næs, 1989) relating the chromatographic data to the spectrophotometric analysis (52 × 1201 X matrix and 52 × 30 Y matrix).

RESULTS AND DISCUSSION

To indicate the differences of the UV-VIS-spectra of the SFE extracts of TDF from the various plant materials a line plot of the UV-VIS data (X) is shown in Figure 1.

After initial analysis including correction for additive and multiplicative effects by use of multiplicative scatter correction (MSC) (Martens and Næs, 1989), adjustment of the wavelength area, and removal of outliers, PCA on UV-VIS-spectra divided the SFE B extracts in the PC 1 direction into groups related to the plant families from which the TDF fractions were derived. SFE extracts were further divided into 2 main groups in the PC 2 direction according to the extraction

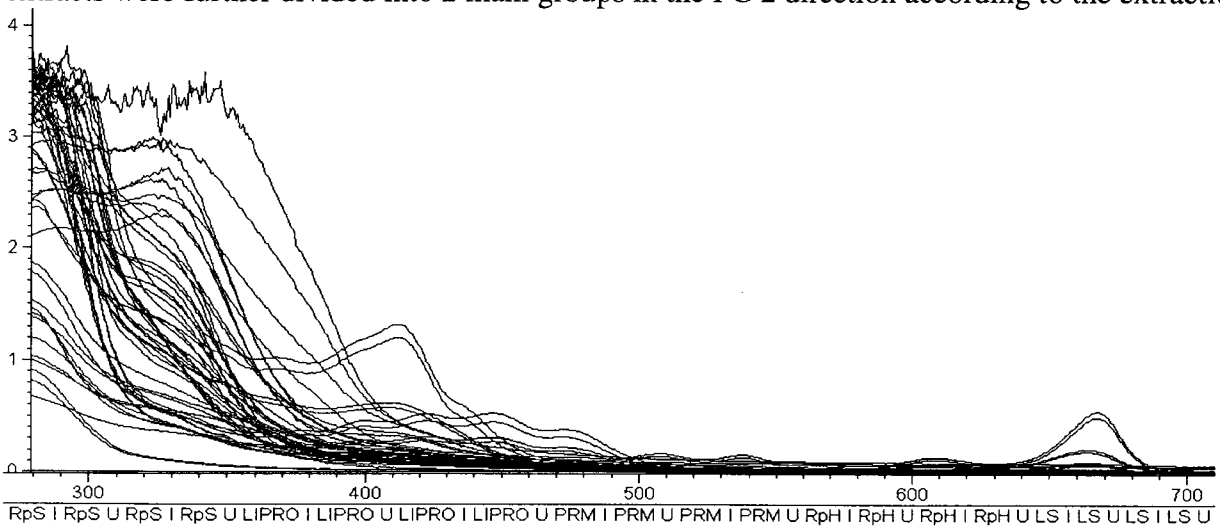

Figure 1. Line plot of UV-VIS spectra from SFE extracts of TDF.

procedure (Figure 2). Replicates are situated in the score plot indicating the reproducability of the SFE/UV-VIS method. From Figure 3 it is seen, that PC 1 primary describes variations in the content of cinnamic acid derivatives (absorption at approximately 325 nm), whereas PC 2 relates to a mixture of compounds. Spectra of chlorophyll-derivatives ($\lambda \approx 667$ nm), carotenoides ($\lambda \approx 425$-478 nm), and probably flavonoid-like compounds ($\lambda \approx 345$-350 nm) can thus be distinguished from the loading plot of PC 2. PC 3 and PC 4 primarily describe further variations in the content of chlorophyll-derivatives.

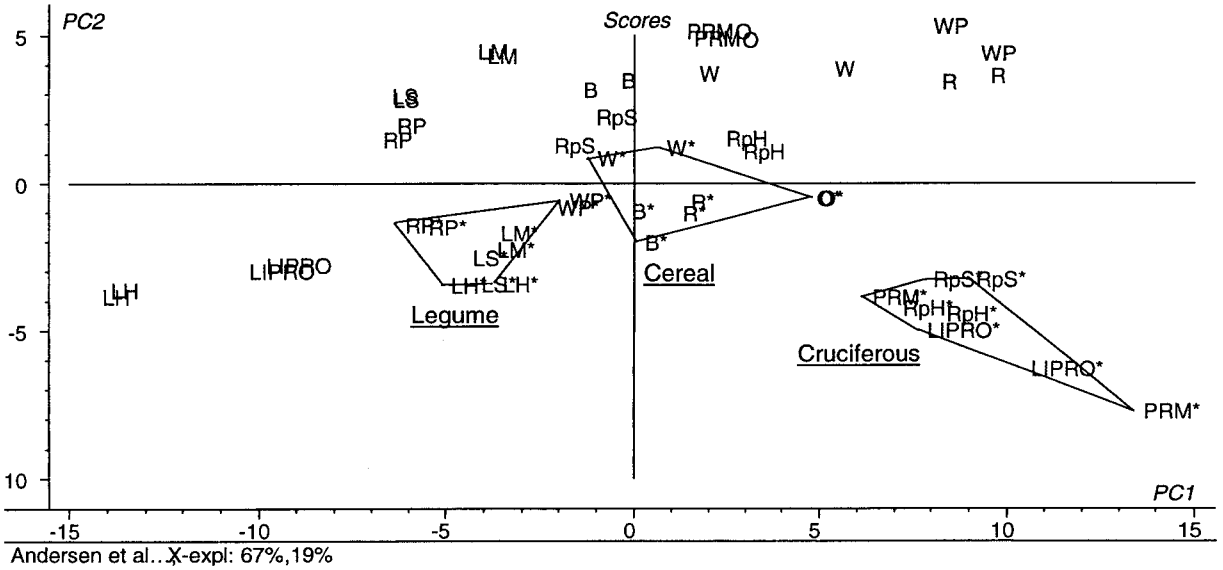

Figure 2. Score plot of PCA on SFE of TDF analysed by UV-VIS-spectroscopy. Unmarked: SFE A; *: SFE B. Their crops are marked Legume, Cereal and Cruciferous.

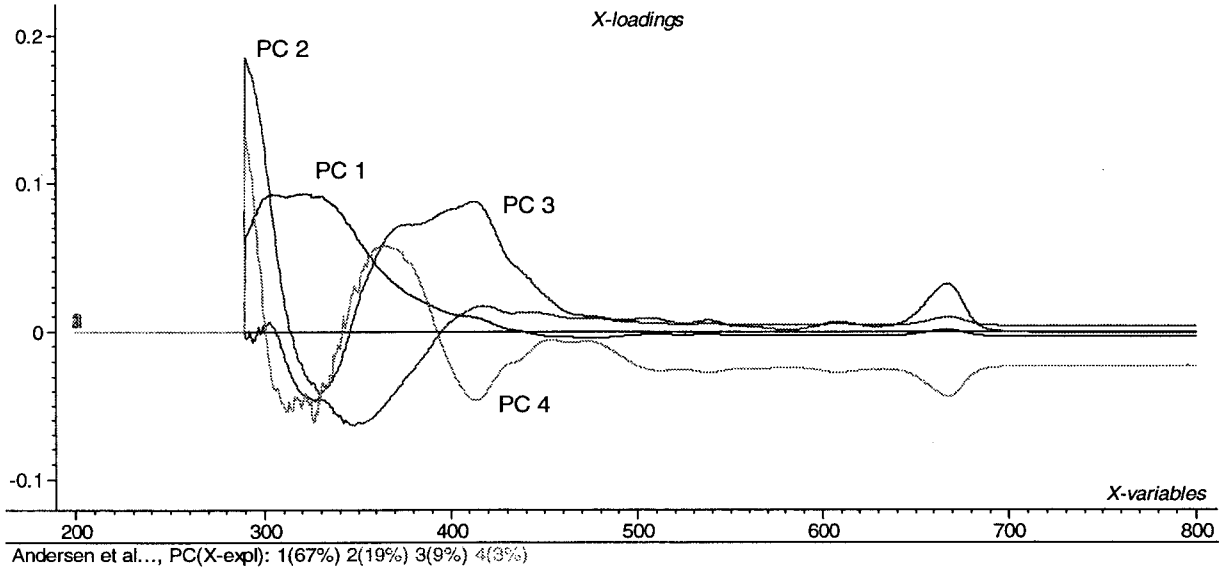

Figure 3. Loading plot from PCA on UV-VIS-spectroscopic data derived from SFE B of TDF.

Chemometric methods of near-infrared spectra of dietary fibres have resulted in fast methods for determination of the DF content in various types of foodstuffs (Williams *et al.*, 1991; Baker, 1995; Kays *et al.*, 1996, Windham *et al.*, 1997; Kays *et al.*, 1998). Another possibility is to determine small impurities, or "indicator substances" for the evaluation quality parameters and process

efficiency as provided by Nørgaard (1995). Further studies of the SFE B extractable compounds of TDF might, in the same manner, result in chemometric models that could supplement the methods for DF determination by estimating the fibre composition and perhaps the plant materials from which the DF materials were derived.

PCA on the MECC (Y) data showed the highest variation in the SFE B extracts (the more hydrophilic SFE extracts) as expected when considering the pre-processing water extraction of compounds with proteolytic active groups (Andersen *et al.*, 1997; 1998a). PCA on MECC of SFE B separated the rape extracts (RpS*, LIPRO*, PRM* and RpH*) and the extracts from the wrinkled pea (WP*) from the rest of the samples. From the biplot in Figure 4 it is seen, that the rape extracts were influenced by a high amount of compound 1-6, 10-12, 14, and sinapic acid while the compounds 16-18, 20-22, 24-25 and 28 had impact on WP.

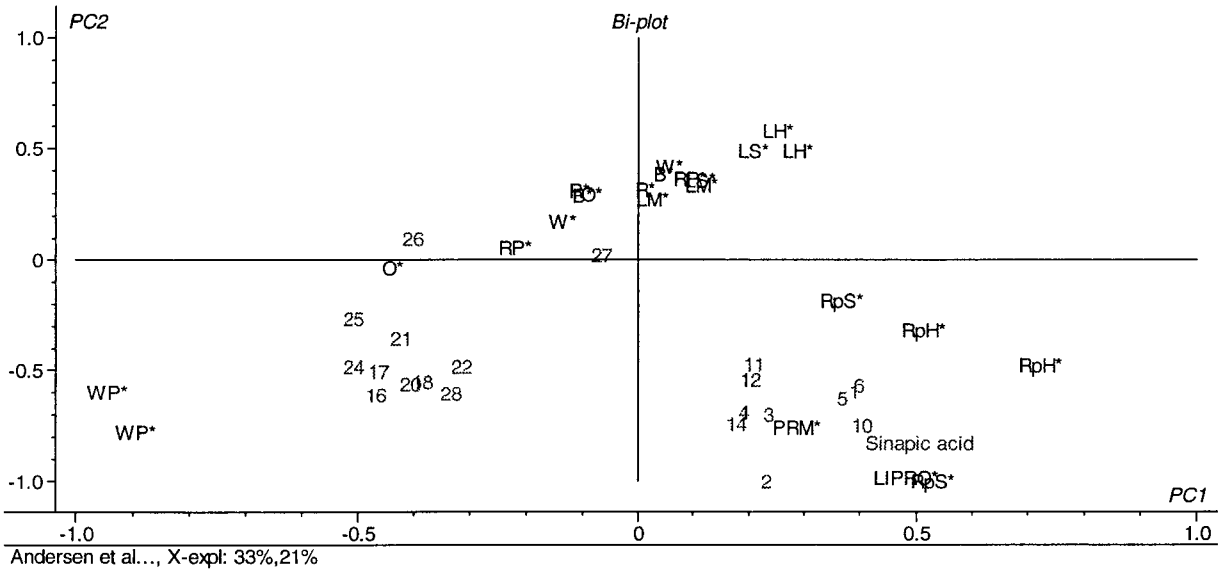

Figure 4. Biplot of PCA on MECC data derived form SFE B of TDF. Names: scores; Numbers: loadings.

The results of the PLS 2 regression model at the UV-VIS-spectra (X-matrix) and MECC results (Y-matrix) is shown in Figure 5 to Figure 9.

From Figure 5 it is seen, that the SFE B extracts of WP and cereals TDF have slightly positive score values, whereas the rest of the legumes TDF extracts have negative, and the SFE B extracts of cruciferous TDF have relatively high positive score values in the PC 1 direction. In the PC 2 direction, the cereals and WP TDF have positive, the remaining legume TDF have neutral, and cruciferous TDF have negative score values. Figure 6 indicates thus, that the score values in the PC 3 direction differentiates the SFE B extracts of the cereals and the WP. Comparing the score plots with the respective loading plots (Figures 6 and 7) it is possible to relate the MECC data to the samples. It is thus seen, that the compounds 1-6, 9-12, 14 and sinapic acid correlate with the rape extracts, the compounds 26-27, 29 and coumaric acid with the cereals, and that the legumes generally have a high content of compound 23 with the exception of WP which correlate with compound 17-18 and 20.

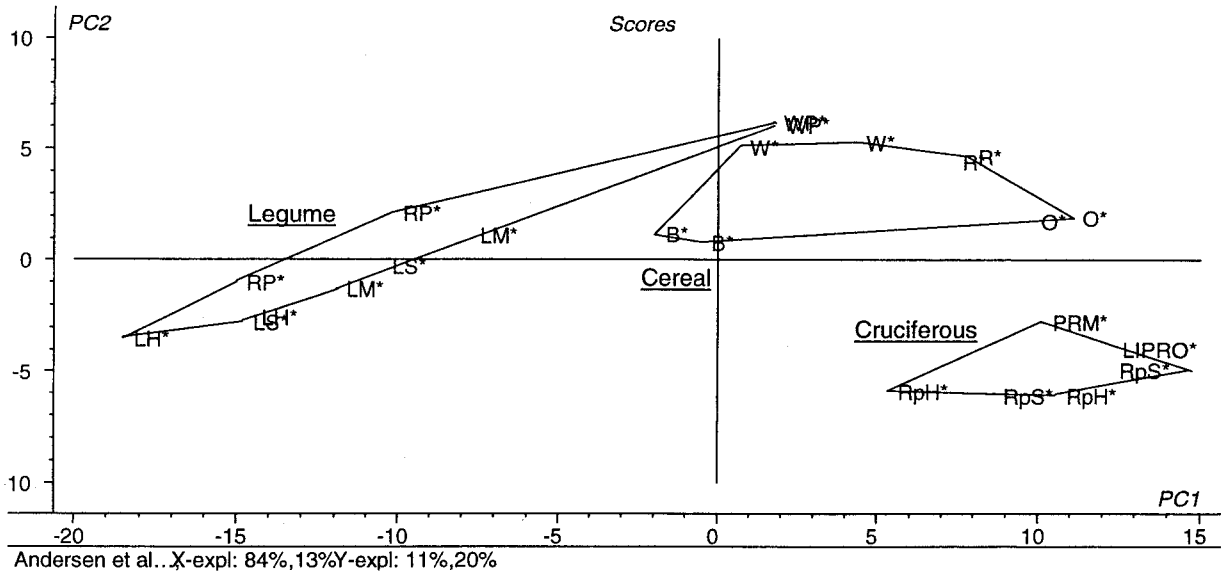

Figure 5. Score plot (PC1 vs. PC2) from PLS 2 regression model on MECC (Y) and UV-VIS-spectra (X) of SFE B extracts of TDF.

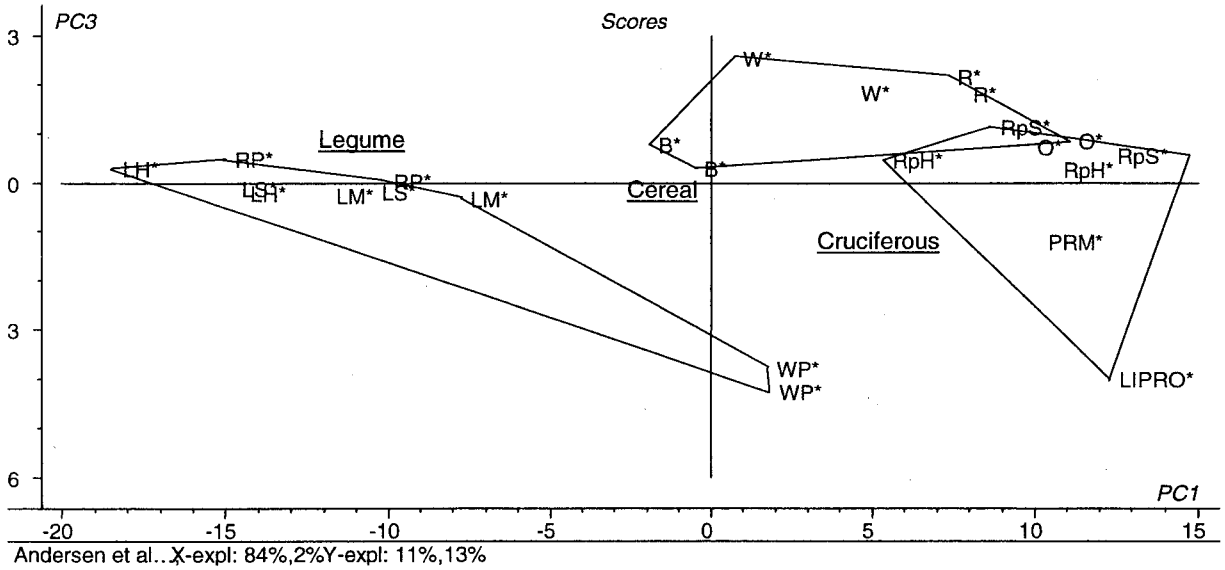

Figure 6. Score plot (PC1 vs. PC3) from PLS2 regression model on MECC (Y) and UV-VIS-spectra (X) of SFE B extracts of TDF.

The loading weights of the PLS 2 regression model (Figure 9) shows the connection between the X and the Y data matrices. The PC 1 have high value at $\lambda \approx 325$ nm, which means that the rape extracts and thus compounds 1-6, 9-12, 14, and sinapic acid is expected to be related to cinnamic acid derivatives. The PC 2 loading weights relate the cereals and WP* and thus compounds 26-27, 29, and coumaric acid to a low content of cinnamic acid derivatives (negative loading weight value at 325 nm) and a small content of chlorophyll derivatives (little positive loading weight value at $\lambda \approx$ 667 nm). The loading weights of PC 3 shows the negative spectrum of chlorophyll, which means that the negative score value of WP* (Figure 6) and the negative loading value of compounds 17-18 and 20 (Figure 8) is related to a high content of chlorophyll and possibly flavonoid ($\lambda \approx 260$ nm and 360 nm) derivatives.

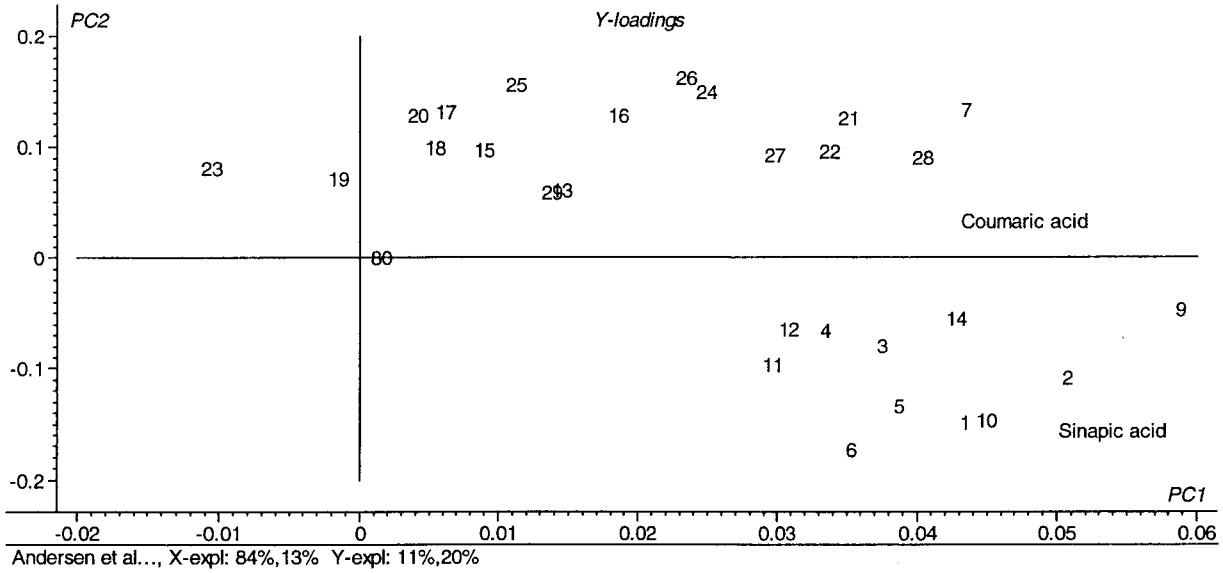

Figure 7. Loading plot (PC1 vs. PC2) from PLS2 regression model on MECC (Y) and UV-VIS-spectra (X) of SFE B extracts of TDF.

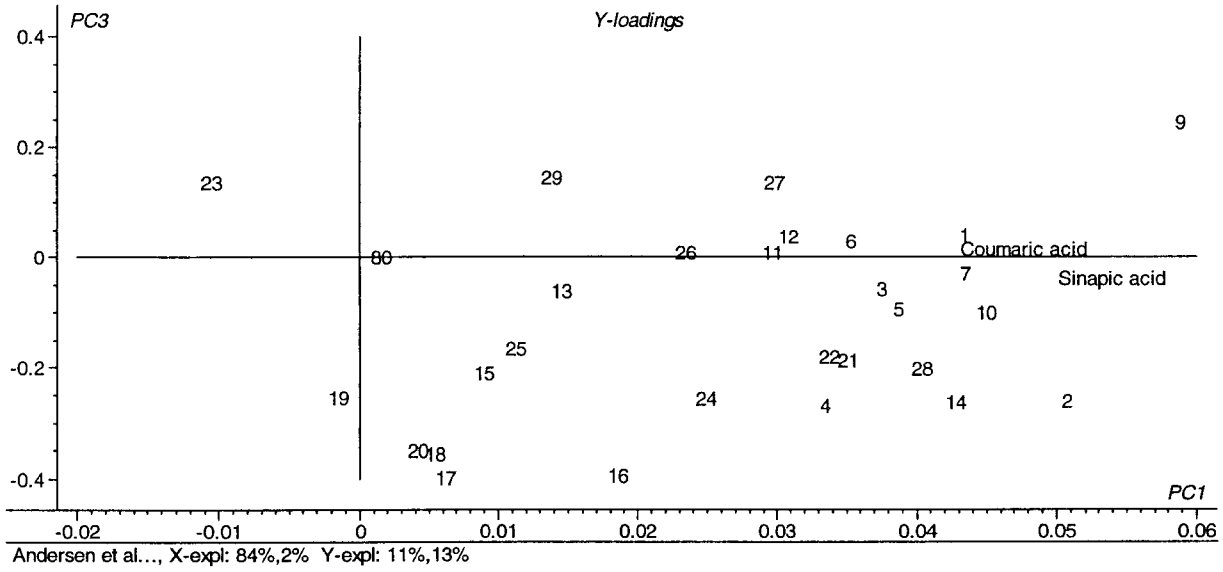

Figure 8. Loading plot (PC1 vs. PC3) from PLS2 regression model on MECC (Y) and UV-VIS-spectra (X) of SFE B extracts of TDF.

The PLS2 model was cross validated and the explained variance of the X and Y matrices after four PLS components were 99 % and 44%, respectively. The Y variance is an average of the explained variance of all Y variables and the PLS model could probably be refined by further optimisation. A higher number of samples and better control of the variation in the migration time of the MECC data might also improve the model. The model is thus not suitable for prediction of MECC detectable compounds based on UV-VIS-spectroscopy, but the information derived from the model reveals some of the physical and chemical properties of the compounds and thus contributes to the final identification.

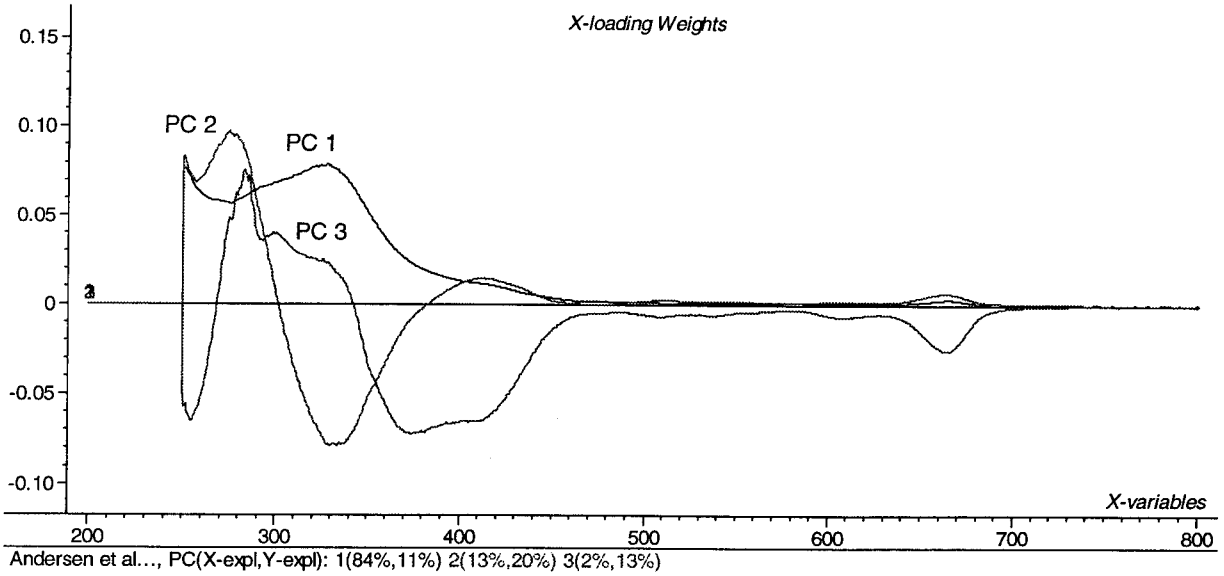

Figure 9. X-loading weights from PLS 2 regression model on MECC (Y) and UV-VIS-spectra (X) of SFE B extracts.

CONCLUSION

The physico-chemical properties and the physiological effects of dietary fibre depend on the composition of the DF fraction. This initial study using chemometric methods on UV-VIS-spectra and MECC chromatograms have indicated that further investigations of amphiphilic TDF associated compounds could result in models for prediction of the composition of the TDF fraction in relation to the plant materials from which the DF are derived.

The study also revealed some of the chemical properties of the amphiphilic compounds that were analysed by MECC. This were done by relating the variation in the relative normalised area of the MECC peaks to the variation of UV-VIS-spectra of the SFE extracts (SFE B) of the samples by the use of PLS 2 regression model. In this way further identification of the compounds is made easier through chemometric analysis.

ACKNOWLEDGEMENTS

This work was financially supported by the Commission of European Union (Contract Nos. FAIR CT 95-0260 and FAIR CT 98-3778).

REFERENCES

Anonymous, (1986). *Chemometrics and Intelligent Laboratory Systems*, 1, cover.

Andersen K.E., Buskov S., Bjergegaard C., Sørensen H., Sørensen J.C. & Sørensen S. (1997). Analyses of dietary fibre associated non-carbohydrate constituents based on supercritical fluid extraction. In: Authenticity and adulteration of food – the analytical approach, Proceedings of Euro Food Chem IX, September 24-26, Interlaken Swizerland (eds. R Amadó og R. Battaglia), Swiss Society of Food and Envionmental Chemistry (SGLUC), vol. 1, 209-214.

Andersen K.E., Bjergegaard C. & Sørensen H. (1998a). Characterisation of dietary fibre associated constituents based on supercritical fluid extraction. *Pol. J. Food Nutr. Sci.*, 7/48 No 4, 617-630.

Andersen K.E., Bjergegaard C. & Sørensen H. (1998b). Dietary fibre associated compounds in different rapeseed products. *GCIRC Bulletin*, 15, 90-95.

Baker D. (1995). The determination of fiber, starch, and total carbohydrate in snack foods by near-infrared reflectance spectroscopy. *Cereal Foods World* 30, 6, 389-392.

Forina M., Drava G., Armanino C., Boggia R., Lanteri S., Leardi R., Corti P., Corti P., Giangiacomo R., Galliena C., Bigoni R., Quartari I., Serra C., Ferri D., Leoni O., and Lazzeri L. (1995). Transfer of calibration function in near-infrared spectroscopy. *Chemometrics and Intel. Lab. Syst.*, 17, 189-203.

Kays S.E., Windham W.R., and Barton F.E. II. (1998). Prediction of total dietary fiber in cereal products using near-infrared reflectance spectroscopy. *J. Agric. Food Chem.*, 44, 2266-2271.

Kays S.E., Windham W.R., and Barton F.E. II. (1996). Prediction of total dietary fiber by near-infrared reflectance spectroscopy in high-fat- and high-sugar-containing cereal products. *J. Agric. Food Chem.*, 46, 854-861.

Martens H., and Næs T. (1989). Multivariate calibration. Wiley, New York.

Nørgaard L. (1995). Classification and prediction of quality and process parameters of thick juice and beet sugar by fluorescence spectroscopy and chemometrics. *Zukerind.*, 120, 11 970-981.

Schneeman B.O. (1987). Soluble vs. insoluble fiber – different physiological responses. *Food Technology*, 2, 81-82.

Windham W.R., Kays S.E., Barton F.E. II. (1997). Effect of cereal product residual content on total dietary fibre determined by near-infrared reflectance spectroscopy. *J. Agric. Food Sci.*, 45, 140-144.

Williams P.C., Cordeiro H.M., and Harnden M.F.T. (1991). Analysis of oat bran products by near-infrared reflectance spectroscopy. *Cereal Foods World*, 36, 571-574.