Quantitative, Qualitative and Exploratory Analysis of Food Using Spectroscopy and Chemometrics

Ph.D. thesis by

Elisabeth Micklander

Supervisors:

Associate Professor Lars Nørgaard

Associate Professor Søren Balling Engelsen

Food Technology
Department of Food Science
The Royal Veterinary and Agricultural University
Denmark

Preface

This Ph.D. thesis is submitted as a partial fulfillment of the requirements for a Ph.D.-degree at the Royal Veterinary and Agricultural University, Copenhagen, Denmark. The research which is the basis of the thesis has mainly been performed at Food Technology, Institute of Food Science under supervision of Associate Professor Lars Nørgaard and Associate Professor Søren Balling Engelsen.

I am grateful to my supervisors for giving me the opportunity to do a Ph.D. in the Food Technology group and for sharing their enthusiasm and great expertise on spectroscopy and chemometrics. I would also like to thank Professor Lars Munck for inspiring input and all the other people at Food Technology for providing a unique working environment, professionally as well as socially. I especially wish to thank Jakob, Vibeke, Jesper, Birthe, Lisbeth, Dorrit, Henrik, Christian and Gilda.

Professor Kim Esbensen and the rest of ACABS, Ålborg University, Esbjerg, Denmark are thanked for having me in their group in November-December 2002 and for introducing me to the world of multivariate image analysis.

Henrik J. Andersen and the Muscle Biology and Meat Science group at the Danish Institute of Agricultural Sciences, Foulum, Denmark are greatly appreciated for letting me draw on their experience on carrying out measurements early post mortem on porcine meat and for letting me use the research abattoir at Foulum.

Finally, Erhvervsfremmestyrelsen is much appreciated for the required financial support during the period of the Ph.D. through the center contract “Center for New Sensor Systems for the Measurement of Food Quality”.

Summary

Spectroscopy in combination with chemometrics has proven to be an outstanding tool for rapid analysis of foods, which can be utilized in food research and industry. The non-invasive spectroscopic techniques can measure the individual food components simultaneously *in situ* in the food matrix, and chemometrics can effectively extract the quantitative information and the underlying qualitative features (latent structures) from the multivariate, covariate spectral data. An extraordinary feature obtained by combining spectroscopy and chemometrics is exploratory analysis and/or pattern recognition. This is central in food research, since novel hypotheses about the food systems under observation may be generated using this inductive analytical approach. For the food industry it is an additional advantage that the fast, non-invasive, remote sensing nature of the spectroscopic methods allows on-line measurements. In this way spectroscopy in combination with chemometrics meets the high throughput needs for quality control, process control and monitoring. In this Ph.D. project the exceptional possibilities provided by spectroscopy and chemometrics have been utilized to improve the analysis and understanding of different foods. The work is presented in six papers which make up the core of the thesis.

Two of the papers demonstrate the potential of using spectroscopy for fast, non-invasive quantitative determinations. In PAPER V, a method for quantification of total nitrogen in grass was developed using near infrared spectroscopy and partial least squares regression with correlation coefficients of 0.97-0.98 depending on pre-processing of the data. However, independent test set validation revealed that more samples have to be included in the model before near infrared spectroscopy can be implemented for routine measurements. The other study (PAPER I) was a feasibility study, which showed that the toxic cyanogenic compound amygdalin in bitter almonds could be quantified *in situ* using Raman spectroscopy and partial least square regression. Correlation coefficients of 0.99 and 0.94 were obtained on a standard addition series and a data set consisting of bitter almonds and sweet almonds spiked with amygdalin, respectively. Unexpectedly it was not the C≡N triple bond of the nitrile group alone that carried the best quantitative information.

In the latter study, the possibility of obtaining vibrational spectra with high spatial resolution using Raman microspectroscopy was utilized to investigate the distribution of amygdalin in bitter almonds. Amygdalin was not present in

measurable amounts in the center of the almonds. It was present in small quantities 0.5 mm from the center and from this point on the amygdalin content increased towards the surface but with great local variation.

In PAPER III, other applications of infrared and Raman microspectroscopy for elucidating the structure of foods were presented and the advantages and disadvantages of the two methods were compared. The main conclusion relates to the relative poor resolution (10x10 μm) of Fourier transform infrared microspectroscopy in relation to food heterogeneity. Raman microspectroscopy using visual lasers has much better spatial resolution but the visual lasers can give problems with heat inducement and fluorescence emission.

Another study (PAPER II) demonstrated the high potential of exploratory studies. Dynamic low-field nuclear magnetic resonance (LF-NMR) measurements and principal component analysis were utilized to explore the changes taking place in meat during cooking. The temperatures at which the major changes in meat occur were identified and it was shown that just above 40ºC the resulting changes in the water properties gave rise to the development of an additional water component. The study also revealed that principal component analysis on entire relaxation curves gives a much more contrasted/detailed picture of the changes occurring in meat during cooking as compared to classic exponential curve fitting.

More examples of the advantages of applying chemometrics analysis to LF-NMR data were presented in PAPER IV. Among other things it was shown that it may be beneficial to generate tri-linear LF-NMR data since application of multi-way chemometrics offer other possibilities for extraction of characteristic relaxation times than bi-linear models.

Furthermore, a study was conducted to elucidate whether on-line ultrasonic measurements performed early post mortem on pig-carcasses could be used for prediction of the water-holding capacity of porcine meat (PAPER VI). The employed ultrasound method could differentiate high and low water-holding capacity. However, classification was not obvious before 85 min post mortem, which is not sufficiently early for on-line use in a commercial abattoir.

Resumé

Spektroskopi kombineret med kemometri har vist sig at være et fremragende værktøj til analyse af levnedsmidler, som kan udnyttes i levnedsmiddelforskning og –industri. De non-invasive spektroskopiske teknikker kan måle de enkelte levnedsmiddelkomponenter på samme tid *in situ* i levnedsmiddelmatricen, og kemometrien kan effektivt uddrage den kvantitative information og de underliggende kvalitative egenskaber (latente strukturer) fra de multivariate, kovariate spektrale data. Et ekstraordinært analytisk redskab, som opnås ved at kombinere spektroskopi og kemometri, er eksplorativ analyse og/eller mønster genkendelse. Dette er centralt i levnedsmiddelforskning, da nye hypoteser om det levnedsmiddel, der undersøges, kan udvikles ved at bruge denne induktive analytiske fremgangsmåde. For levnedsmiddelindustrien er det en yderligere fordel, at de spektroskopiske metoder er kendetegnet ved hurtige, non-invasive, "remote sensing" målinger, hvilket muliggør on-line analyser. På denne måde kan de spektroskopiske teknikker kombineret med kemometri afhjælpe storproduktionens behov for kvalitetskontrol, proceskontrol og –overvågning. I dette Ph.D. projekt er de enestående muligheder, som spektroskopi og kemometri giver, udnyttet til at forbedre analysen og forståelsen af forskellige levnedsmidler. Arbejdet er præsenteret i seks artikler, som udgør kernen i afhandlingen.

To af artiklerne demonstrerer potentialet i at anvende spektroskopi til hurtige, non-invasive kvantitative bestemmelser. I ARTIKEL V blev der udviklet en metode til kvantificering af det totale kvælstofindhold i græs baseret på nærinfrarød spektroskopi og partial least squares regression med korrelationskoefficienter på 0,97-0,98 afhængigt af forbehandlingen af data. Validering med et uafhængigt testsæt viste imidlertid, at der skal inkluderes flere prøver i modellen, før nærinfrarød spektroskopi kan implementeres som rutine analyse. I den anden undersøgelse (ARTIKEL I) blev det vist, at det er muligt at måle indholdet af det toksiske stof amygdalin *in situ* i bitre mandler ved hjælp af Raman spektroskopi og partial least squares regression. Der blev opnået korrelationskoefficienter på 0,99 og 0,94 for henholdsvis en standardrække og et datasæt, der bestod af bitre mandler og søde mandler tilsat ekstra amygdalin. Imod forventning var det ikke kun C≡N triple bindingen i nitrilgruppen, der indeholdt den bedste kvantitative information.

I sidstnævnte undersøgelse blev muligheden for at måle vibrationsspektre med høj spatial opløsning ved hjælp af Raman mikrospektroskopi udnyttet til at undersøge

fordelingen af amygdalin i bitre mandler. Amygdalin fandtes ikke i målelige mængder i midten af mandlerne. Det var til stede i små mængder 0,5 mm fra midten og fra dette punkt steg indholdet hen imod overfladen, men med stor lokal variation.

I ARTIKEL III blev andre anvendelser af infrarød- og Raman mikrospektroskopi til at belyse strukturen af levnedsmidler præsenteret, og fordele og ulemper ved de to metoder blev sammenlignet. Hovedkonklusionen er relateret til den relativt lave opløsning (10x10 μm) af Fourier transform infrarød mikrospektroskopi i forhold til levnedsmiddelheterogenitet. Raman mikrospektroskopi udstyret med visuelle lasere har en meget højere spatial opløsning, men de visuelle lasere kan give problemer med opvarmning af prøven og fluorescens emission.

En anden undersøgelse (ARTIKEL II) demonstrerede det store potentiale, der er i eksplorativ analyse. Dynamiske lavfelts kernemagnetiske resonansmålinger (LF-NMR) og principal komponent analyse blev brugt til at udforske ændringerne i kød under kogning. De temperaturer, hvor de største ændringer i kød foregår, blev bestemt, og det blev vist, at der opstår en ny vand komponent lige over 40ºC. Undersøgelsen viste yderligere, at principal komponent analyse på hele relaxationskurver giver et meget mere detaljeret billede af de ændringer, der sker i kød under kogning, end klassisk eksponentiel kurvetilspasning.

Andre eksempler på de fordele, der er ved at anvende kemometri på LF-NMR data, blev præsenteret i ARTIKEL IV. Det blev blandt andet vist, at det kan være gavnligt at generere triliniære LF-NMR data, da anvendelse af multivejs kemometri giver andre muligheder end biliniære modeller for at ekstrahere de karakteristiske relaxationstider.

Endvidere blev der udført et forsøg, der skulle belyse om on-line ultralydsmålinger udført tidligt post mortem på slagtekroppe af svin kunne bruges til prædiktion af vandbindindsevnen af svinekød (ARTIKEL VI). Den anvendte ultralydsmetode kunne skelne mellem høj og lav vandbindingesvne. Klassifikationen var imidlertid ikke tydelig før end 85 minutter post mortem, hvilket ikke er tidligt nok til, at metoden kan bruges on-line på et kommercielt slagteri.

List of publications

PAPER I

E. Micklander, L. Brimer and S.B. Engelsen (2002). Noninvasive assay for cyanogenic constituents in plants by Raman spectroscopy: Content and distribution of amygdalin in bitter almond (*Prunus amygdalus*), *Applied Spectroscopy*, 56, 1139-1146.

PAPER II

E. Micklander, B. Peshlov, P. Purslow and S.B. Engelsen (2002). NMR-cooking: Monitoring the changes in meat during cooking by low-field H^1-NMR, *Trends in Food Science and Technology*, 13, 341-346.

PAPER III

L.G. Thygesen, M.M. Løkke, E. Micklander and S.B. Engelsen (2002). Vibrational microspectroscopy of food. Raman vs. FT-IR. *Trends in Food Science and Technology*, 14, 50-57.

PAPER IV

E. Micklander, L.G. Thygesen, H.T. Pedersen, F. van den Berg, R. Bro, D.N. Rutledge and S.B. Engelsen (2003). Multivariate analysis of time domain NMR signals in relation to food quality. In *Magnetic Resonance in Food Science: Latest Developments*. Eds. P.S. Belton, A.M. Gil, G.A. Webb and D.N. Rutledge. The Royal Society of Chemistry, Cambridge, UK, pp. 239-254.

PAPER V

R. Gislum, E. Micklander and J.P. Nielsen (2004). Quantification of nitrogen concentration in perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and chemometrics. *Field Crops Research*, in press.

PAPER VI

E. Micklander, H.C. Bertram, H. Marnø, L.S. Bak, H.J. Andersen, S.B. Engelsen and L. Nørgaard. Testing of an on-line ultrasound method for early discrimination of high and low water-holding capacity in pig carcasses – a study including LF-NMR. *Food Science and Technology*, submitted.

Table of contents

Abbreviation list

ATR	attenuated total reflection
CCD	charge-coupled device
CPMG	Carr-Purcell-Meiboom-Gill
FID	free induction decay
FT	Fourier transform
iPLS	interval partial least squares regression
IR	infrared
LF-NMR	low-field nuclear magnetic resonance
MLR	multiple linear regression
MSC	multiplicative signal correction
NIR	near infrared
PC	principal component
PCA	principal component analysis
PCR	principal component regression
PLSR	partial least squares regression
RF	radio frequency
RMSECV	root mean square error of cross validation
RMSEP	root mean square error of prediction
T_1	longitudinal or spin-lattice relaxation time constant
T_2	transverse or spin-spin relaxation time constant
WHC	water-holding capacity

1. Introduction

Food quality can be difficult to assess, because the complex composition of food complicates quantitative and qualitative determination of the individual constituents. When chemical methods are used, destruction of the food matrix is usually necessary prior to measurements in order to isolate the compound of interest. The destruction step is problematic, because it is laborious, it often requires a large amount of chemicals, and information is lost when the individual food components are removed from their context. With spectroscopy it is possible to measure non-invasively. In this way, the individual food components can be assessed simultaneously *in situ* in the food matrix. Thus, spectroscopy has opened up new possibilities within food analysis. However, the possibilities of the spectroscopic methods are only fully exploited when appropriate data analysis is applied. Therefore, parallel to the rapid development of the spectroscopic techniques during recent decades, mathematical tools which can handle the multivariate, covariate spectroscopic data optimally have been developed. In the context of chemical analysis, the multivariate data analytical methods are referred to as chemometrics, originating from *chemeia* (Greek, chemistry) and *metros* (Greek, measure).

A synergy is achieved by merging the non-invasive spectroscopic measurements with chemometrics. In particular, exploratory analysis becomes possible, which opens up for investigation of unknown aspects. Exploratory analysis is based on an inductive analytical approach, by which as much information as possible is collected (using spectroscopy which measures the entire food system) and pragmatically evaluated (using chemometrics which explore the collected data for underlying influences based on a minimum of *a priori* assumptions). This may lead to the development of hypotheses *a posteriori*, which eventually may lead to a better understanding of the food system under investigation (Munck *et al.*, 1998). It is important to note that the chemometric methods are essential for exploratory food analysis with spectroscopy. If the spectroscopic data are analyzed using data reduction based on *a priori* knowledge, the multi-informative and potential exploratory feature of the spectroscopic data is lost.

Quantitative determinations using absorbance spectroscopy can be performed without the use of chemometrics. Classically, the data analysis is performed on selected variables that are known to contain information about the functional group to be quantified. However, it is highly advantageous to use chemometrics when

spectral information is available. Regressions on spectral data usually give better prediction ability and the calibrations are more robust, as the multivariate chemometric methods can cope with interferences and allow identification of deviating samples. Thus, there is no reason to use knowledge based variable selection for quantitative analysis, as the computer power and data analytical tools able to handle full spectra are available.

Qualitative analysis, i.e. compositional and structural analysis of single molecules, is the area of spectroscopic analysis in which chemometrics has had the least impact. This is because all variables are important for interpretation of the spectra. Therefore, chemometric methods are not advantageous as they are based on reduction of the complexity of data. However, chemometrics is important for a special type of qualitative analysis, specifically when the goal of the analysis is to detect qualitative differences such as quality changes or quality classes. Exploratory chemometric methods are used for such "qualitative analysis".

The possibility of using remote sensing in the food industry is an additional advantage made possible by combining spectroscopy and chemometrics. The fast, non-invasive nature of the spectroscopic methods allows measurements of food products before, after and during production. In this way the spectroscopic techniques in combination with chemometrics meet the need for quality control, process control and monitoring in the food industry. Measurements can be performed in-line where the sensor is in contact with foodstuff passing by, e.g. inside a pipe, or on-line where non-contact measurements are performed on the moving product, e.g. on a conveyor belt. Furthermore, measurements can be performed on sub-samples transferred from the production line to an instrument situated close by (at-line) or in a different location (off-line).

The aim of this thesis is to utilize and explore the exceptional possibilities for analysis of food, which are provided by the spectroscopic techniques in combination with chemometrics. Six papers (PAPERS I-VI) present the work involving different spectroscopic and chemometric techniques, ranging from laboratory studies on bitter almonds, meat and grass to studies with on-line equipment on pig carcasses. Chapter 2 provides an introduction to the chemometric methods used in the thesis. In Chapters 3-5, vibrational spectroscopy, low-field nuclear magnetic resonance (LF-NMR) relaxometry and ultrasonic techniques are presented. Furthermore, the application of the spectroscopic methods to quantitative, qualitative and exploratory food analysis in research and in industry is

discussed, and the major findings in the six papers are presented. In Chapter 6, the conclusions of the project are summarized and the perspectives for the use of spectroscopic techniques and chemometrics in food science are discussed.

2. Chemometrics on covariate spectroscopic data

2.1 Covariate spectroscopic data

Spectroscopy is a general term for the science that deals with the interactions of various types of radiation with matter. The interaction between radiation and a sample depends on 1) the energy, i.e. the frequency of the radiation, 2) the type and amount of atoms and molecules which are present in the sample and 3) the influence of the sample matrix, i.e. the surrounding microenvironment on the individual molecules. An investigation of the response from a sample upon irradiation at specific wavelengths will therefore provide information about the qualitative and quantitative, chemical and physical state of the sample.

Spectroscopic data often consist of several hundreds or thousands of variables and the individual variables are generally strongly correlated. Chemometric methods are designed to treat such multivariate, covariate data which may be problematic to model with statistical methods such as multivariate linear regression (MLR). The bilinear chemometric methods - principal component analysis (PCA), principal component regression (PCR) and partial least squares regression (PLSR) - overcome the problems of the collinear data by extracting linearly independent latent variables from the original data. An introduction to PCA and PLSR is given below.

2.2 Introduction to PCA

PCA (Hotelling, 1933; Wold *et al.*, 1987) is a fundamental chemometric method which can be used for exploratory analysis of multivariate data (e.g. covariate spectroscopic data). Essentially, PCA is a compression method which models the original multivariate data using a limited number of latent variables (principal components, PC's). First the original data matrix (X) is centered by subtracting the mean spectrum $\bar{x}$ from each sample spectrum, where after the centered data matrix, X_c, is decomposed to a score matrix (T) and a loading matrix (P) by consecutive orthogonal subtraction of the largest variation in data until the variation left is unsystematic. The residuals are collected in a matrix (E):

$$X_c = TP^T + E \qquad \text{(Eq. 2.1)}$$

The loading vectors can be considered as hidden spectral profiles that are common to all the measured spectra, i.e. the loadings contain qualitative information. The scores are the amounts of the hidden profiles in the individual spectra, i.e. the scores contain quantitative information. An important feature of PCA is the graphic interface, especially plots of the scores and loadings. Two-dimensional or three-dimensional scatter plots of the score vectors (in different combinations) depict the covariance between samples, providing a data overview. Patterns and clusters of objects as well as outliers are easily identified in the score plots which enable the data analyst to explore the expected and unexpected trends in the data. Line plots of the loadings illustrate the importance of the original variables for each PC and may be used to deduce, which qualitative differences cause the clusters or trends in the data.

2.3 Introduction to PLSR

PLSR (Martens & Næs, 1989) is a multivariate regression method which is commonly used in quantitative spectroscopy to correlate spectroscopic data (X) with related chemical/physical data (y). Like PCA, PLSR is based on latent variables, but in PLSR the decomposition of X during regression is guided by the variation in y, i.e. the co-variation between X and y is maximized, so that the variation in X directly correlating with y is extracted. As mentioned, an important feature of PLSR is that it is based on latent variables and can therefore handle the (usually) highly collinear spectroscopic data, in contrast to MLR.

The purpose of PLSR is to build a linear model to enable the prediction of y (the chemical/physical variable) from x (a measured spectrum). The linear model between the vector y_c containing the centered reference data and the matrix X_c containing the centered spectral data can be described by:

$$y_c = X_c b + e \qquad \text{(Eq. 2.2)}$$

where b is a vector which contains the regression coefficients that are determined during the calibration and e is the residual. In order to obtain a good estimation of b, the PLSR model needs to be calibrated on samples which span the variation in y well and in general are representative of the future samples. Depending of the

complexity, i.e. the chemical/physical rank, of the future samples, this may require a huge number of samples. In PAPER V, near infrared (NIR) spectra and total nitrogen concentrations of 837 grass samples were available, yet a robust calibration could not be made. Thus, calibration of a quantitative spectroscopic method may be labor-intensive, even for a seemingly simple application as prediction of protein, for which NIR is used more or less routinely.

2.4 Validation of chemometric models

Validation of the chemometric models is very important in order to determine the correct number of PC's so that noise is not included in the models, to detect outliers and to obtain reliable estimates of the prediction error. In this project cross-validation or test-set validation were applied, depending on the data set.

Cross-validation is used when the number of samples in a data set is limited and all samples therefore are needed for both calibration and validation. The data matrix is divided into a number of segments containing one or more samples. One by one the segments are left out and a model is calibrated on the remaining samples and used to predict the samples in the left out segment. The prediction error is estimated as the root mean square error of cross validation (RMSECV) which is calculated by comparing the predicted values with the reference values as given below:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{N}(y_{i,pred} - y_{i,ref})^2}{N}} \qquad \text{(Eq. 2.3)}$$

where y_{pred} is the predicted value, y_{ref} is the reference value and N is the number of samples.

Test-set validation is used when the data set is large enough to be divided into two subsets: a calibration set which is used for calibration of the model, and a validation set which is used to estimate the prediction error. Since different samples are used for calibration and validation, test-set validation is a stronger test of the model than cross-validation and the estimate of the prediction error, root

mean square error of prediction (RMSEP) (calculated using Eq. 2.3), will usually be more conservative and closer to the true value than RMSECV.

A test set may be a dependent test set, when it consists of samples which are selected randomly or orderly from the original data set, or it may be independent, when a new series of samples is analyzed and used as a test set. Dependent test sets are usually used for model validation as described above, whereas independent test sets are primarily used to test the applicability of a calibrated model to new samples letting the independent test set represent future samples. The applicability depends on the robustness of the model, i.e. whether the variations (presumed) to be found in the future are sufficiently represented in the model. In PAPER V, validation with an independent test set (samples from a separate year) was used to test whether a model on samples from two years was sufficiently robust to be used for quantitative measurements. The RMSEP was considerably higher than the RMSECV and it was concluded that more samples representing a greater variation were needed in the calibration before it could be applied in practice.

2.5 Variable selection

Even though PLSR automatically gives high priority to variables that co-vary with the *y* variable, the predictive ability often increases when the regression is based on the significant variables or variable intervals only. The reason for this is that PLSR is labile when a lot of noise is present in the spectra. If large spectral regions are noisy, exclusion of the problematic noisy variables is an advantage.

There are many different methods for "exploratory" variable selection, where the selection is not based on *a priori* knowledge, but where the full spectra are explored for important variables, e.g. principal variables (Höskuldsson, 1994), forward stepwise selection (Höskuldsson, 1996), significance testing by jack-knife estimation of parameter uncertainty (Martens & Martens, 2000), interval PLS (iPLS) (Nørgaard *et al.*, 2000) and genetic algorithms (Leardi, 2001).

In PAPER I, iPLS was applied on Raman spectra in order to improve the prediction of amygdalin content, since full spectrum models on the noisy Raman spectra gave a high prediction error. A practical feature of iPLS is the graphic interface, which gives an overview of the performance of the selected spectral regions compared to the full-spectrum model. The iPLS plot in Fig. 2.1 illustrates efficiently that the information on amygdalin content is primarily found in the vibrational bands of the

nitrile group (2242 cm^{-1}) and the aromatic ring (3060 cm^{-1}, 1600 cm^{-1} and 1000 cm^{-1}) of the amygdalin molecule. Upon pre-processing (see section 2.6) and optimization of the interval selections considerable improvement in the prediction performance (up to 83%) with respect to RMSECV was obtained by using the spectral sub-regions selected by iPLS compared to full spectrum models.

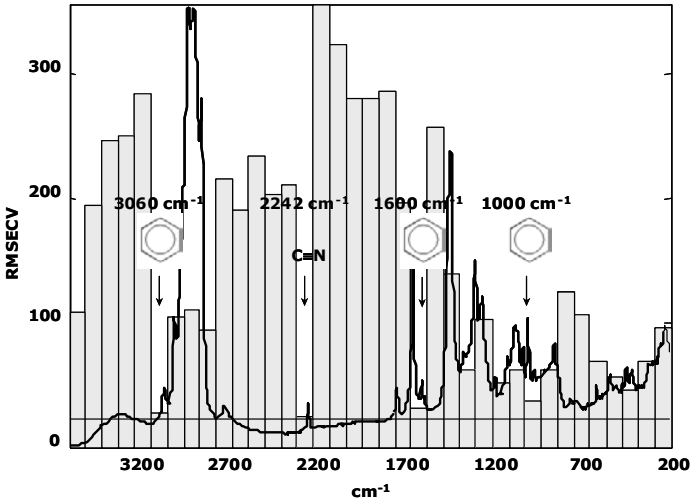


Fig. 2.1 Cross-validated prediction of performance (RMSECV) for 37 PLSR interval models (bars) and for full-spectrum model (line) plotted together with the mean spectrum of seven standard samples (amygdalin added to ground sweet almond in concentrations of 0, 58, 117, 175, 350, 525 and 700 nmol/mg). The full-spectrum model (line) was made with 2 PC's; the sub-interval models (bars) were made using 1 PC (PAPER I).

2.6 Pre-processing of spectral data

NIR, infrared (IR) and Raman spectra often contain systematic variation like an additive or multiplicative offset which may be caused by scatter effects, chemical interferences, or instrument drift. Such variation may complicate the data analysis and interpretation, for which reason mathematical pre-processing may be employed. The pre-processing eliminates or reduces the impact of the non-relevant spectral information and often leads to simpler and more robust regression models and better interpretability of the data.

In this project, different pre-processing methods were used. In PAPER I, first and second derivation, multiplicative signal correction (MSC) (Geladi *et al.*, 1985) and

standard normal variate (Barnes *et al.*, 1989) was tested on Raman spectra which were affected by varying levels of fluorescence and differences in the physical and chemical state of the sample matrix. In PAPER V, MSC and extended inverted signal correction (Martens *et al.*, 2003; Pedersen *et al.*, 2002a) were applied to NIR spectra which were affected by scatter due to different particle size of the samples.

3. Measuring food quality with vibrational spectroscopy

3.1 Vibrational spectroscopy

IR, NIR and Raman spectroscopy are commonly referred to as vibrational spectroscopy as the energy contained in the electromagnetic waves in the IR region (4000-200 cm^{-1}) and the NIR region (13000-4000 cm^{-1} corresponding to 780-2500 nm) corresponds to vibration of chemical bonds. The detection of NIR light dates back to 1800 when it was discovered by Herschel. Developments in electronic amplification and double-beam spectrometers during the Second World War led to the development of commercial IR instrumentation from 1945, but the most important advancement in IR occurred in the 1970's with the development of Fourier transform (FT) instruments. Custom-made NIR instruments appeared in the 1930's, but it was not until 1960's with the development of computers that commercial instrumentation appeared (Sheppard, 2002; McClure, 2004). Raman scatter was observed by Raman and Krishnan and almost concurrently by Mandelstamm in 1928. The use of Raman spectroscopy was boosted when visible lasers were developed in the mid 1960's, but the breakthrough of the technique for routine chemical analysis first came around 1986 with the development of NIR lasers, FT-Raman and charge-coupled devices (CCD) (Hendra *et al.*, 1991; McCreery, 2000).

IR and NIR spectroscopy are based on the measurement of absorbance (A) or transmittance (T) as a function of wavenumber ($\bar{\nu}$). $A(\bar{\nu})$ and $T(\bar{\nu})$ are dependent on the frequency-dependent absorptivity coefficient (a), the path length (b) and the concentration (c) of an absorbing analyte, as described by the Bouguer-Beer-Lambert law (Griffiths, 2002):

$$A(\bar{\nu}) = -\log_{10} T(\bar{\nu}) = a(\bar{\nu}) \cdot b \cdot c \qquad \text{(Eq. 3.1)}$$

For a mixture of N components the total absorbance is expressed as (Griffiths, 2002):

$$A(\bar{\nu}) = \sum_{i=1}^{N} a_i(\bar{\nu}) \cdot b \cdot c_i \qquad \text{(Eq. 3.2)}$$

Since the absorbance of a food sample in this way is linearly related to the concentration of the individual food components, NIR and IR data fulfill the criteria for application of the bilinear chemometric models. The Bouguer-Beer-Lambert law does not apply to Raman spectroscopy, as Raman spectroscopy involves scattering, not absorption. However, the intensity of Raman scattering (I_{Raman}) can be expressed by an equation analogous to the Bouguer-Beer-Lambert law (Coates, 2002):

$$I_{Raman} = (I_L \cdot \sigma \cdot \bar{\nu}^4 \cdot K) \cdot P \cdot C \qquad \text{(Eq. 3.3)}$$

where I_L is the intensity of the incident laser radiation, σ is the absolute Raman scattering cross-section (in cm^2 per molecule), $\bar{\nu}$ is the wavenumber of the incident laser light, K is a constant that includes the measurement parameters, P is a pathlength equivalent term and C is the number of scattering centers per unit volume (concentration).

The wavelengths in the IR region correspond to the frequencies of the fundamental vibrations of the chemical bonds. When the frequencies of incident IR light correspond to the frequencies of the bonds present in a sample, the IR radiation is absorbed by the sample and absorption as a function of frequency gives the IR spectrum (Fig. 3.1). In this project, IR spectra were collected using an FT instrument. FT instruments have been preferred to dispersive instruments since they were developed, because they are more robust and considerably faster, which apart from faster acquisition leads to increased sensitivity.

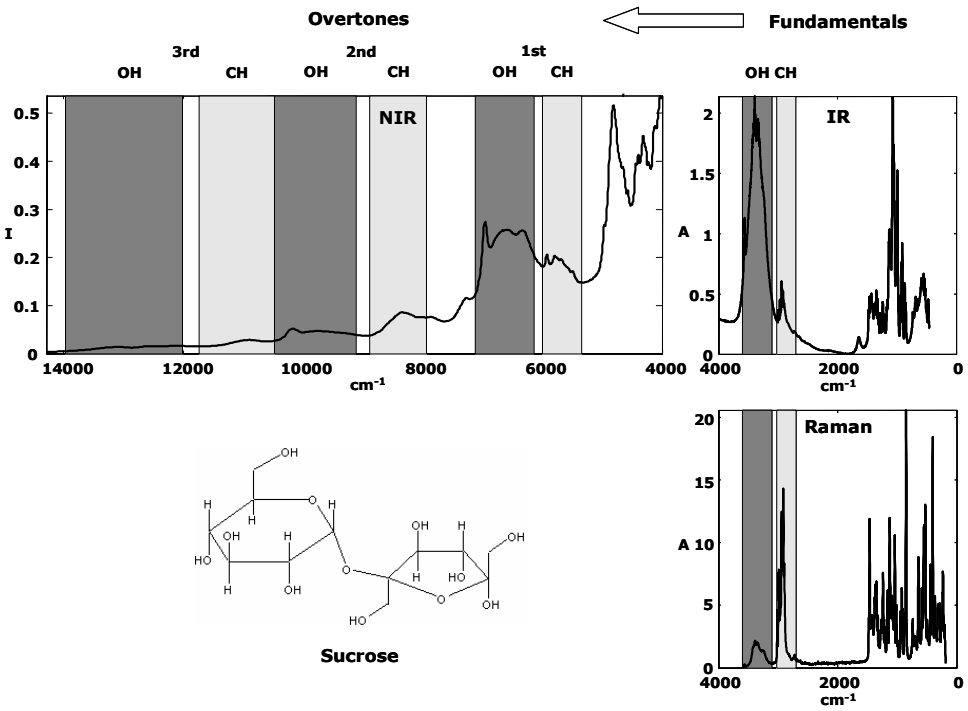


Fig. 3.1 NIR, IR and Raman spectra of sucrose. Fundamental CH and OH stretching bands are highlighted in the IR and Raman spectra, illustrating the different intensities of the bands in the two spectra. The fundamental vibrations in the IR region are repeated as first, second and third overtones in NIR (highlighted areas) as well as combination bands.

Raman spectra cover the same frequency region as IR spectra, but the band intensities are different (Fig. 3.1). The absorption bands in the IR spectra are caused by vibrations which give rise to a change in dipole moments as the motion occurs, while the bands in the Raman spectra depend on vibrations which cause a change of dipole polarizability. Bonds that connect two identical or nearly identical parts of a molecule generally result in more intense bands in Raman spectroscopy than in IR, whereas bonds between two different parts are usually weaker in Raman than in IR. According to these simple rules, the results of IR and Raman spectroscopy complement one another and the complete picture of the molecules' vibrational pattern can only be obtained using both techniques. Raman spectra are obtained by measuring the inelastic scattering of light as a function of the displacement from the Rayleigh line called the Raman shift. It occurs upon irradiation of the sample with a monochromatic source. When CCD-Raman equipment is used, an ultraviolet or visual laser is used for excitation and a CCD

for detection, while FT-Raman equipment uses a NIR laser for excitation and an interferometer for detection. The advantage of using NIR lasers is that they are lower in energy than the visual lasers, which means that excitation of the electronic transitions responsible for fluorescence rarely occurs. This is of immense importance to food analysis, since fluorescent compounds are often present in foods. The drawback of NIR lasers is that the Raman scattering is weaker than for visual lasers, because the Raman scatter intensity depends on the fourth power of the wavenumber of the light source as seen in Eq. 3.3. However, the ratio of Raman scattering to fluorescence is almost always drastically increased when NIR lasers are used rather than visual lasers. Both CCD-Raman and FT-Raman were used in this thesis (PAPER I, PAPER III).

NIR spectroscopy employs the same principles as IR spectroscopy, but NIR spectra are made up of broad bands corresponding to overlapping peaks of overtones and combination bands of the fundamental vibrations. NIR spectra are less distinctive than IR spectra. On the other hand, the chemical information is abundant, since the chemical information contained in one fundamental vibration band in IR is repeated as the first, second and third overtone in the NIR spectrum (Fig. 3.1) and as combination tones.

All four major food ingredients (water, carbohydrates, proteins and fats) are assessed with vibrational spectroscopy. In IR C-H, O-H, N-H, C-O and C=O stretching bands dominate the spectra, whereas only the very anharmonic vibrations (i.e. bonds including the light hydrogen atom such as C-H, O-H, N-H) give intensive bands in NIR. While IR and NIR thus focus on vibration of side groups, Raman spectroscopy is more sensitive to vibrations of the more rigid bonds like C=C stretching which make up the molecular skeleton.

3.2 Advantages and disadvantages of NIR, IR and Raman spectroscopy

A disadvantage of IR spectroscopy is that water absorbs strongly and the intense water bands may obscure other spectral bands of interest. Thus, application of IR to high-moisture foods is problematic. It is possible, but the light path has to be very short in order to decrease the intensity of the water signal. For example, equipment designed for analysis of milk and wine uses a path length of 30-50 μm (Andersen *et al.*, 2002). Another possibility is to use the attenuated total reflection (ATR) principle which was developed in the 1960's (Sheppard, 2002). In Raman

spectroscopy the O-H stretching of water is very weak and therefore it is directly applicable to high-moisture foods. Another disadvantage of IR spectroscopy is that IR radiation cannot be transmitted through quartz glass and that the weak IR radiation only measures the surface of a sample or can only be transmitted through very small amounts of sample, which complicates sample presentation. Development of ATR crystals has solved this problem partly. However, lack of alternative materials to quartz for production of optic fibers for remote on-line measurements remains to be a problem. Disadvantages of Raman spectroscopy are the inherent low signal-to-noise ratio, interference from fluorescence of food components, and heating of the sample. The problems with fluorescence are minimized by using NIR lasers as described above, whereas heating of the sample can only be reduced by reducing the laser power and/or the acquisition time.

3.3 Quantification of food components using vibrational spectroscopy

Quantification of one or more of the major food components is frequently part of quality control in the food industry. The fact that NIR spectroscopy in combination with chemometrics has proven to be an excellent tool for quantitative determination of all four major food constituents has had a tremendous impact on the industry, since this has paved the way for replacement of the laborious standard measurements with fast NIR measurements. NIR spectroscopy in combination with chemometrics has been implemented off-line, at-line, on-line and in-line for quantitative measurement of the major constituents in numerous foods and there is a vast amount of literature available concerning feasibility studies on on-line applications (Li-Chan *et al.*, 2002).

As reviewed in Williams & Norris (1987) and Osborne *et al.* (1993), quantification using NIR was already widespread prior to the development of chemometrics. The methods were generally based on MLR models on selected variables. As mentioned in the introduction, it is advantageous to use chemometrics when spectral information can be measured. Methods based on spectroscopic measurements covering a large spectral area combined with chemometrics are therefore usually preferred today. However, quantification using MLR on a limited number of variables is still used when measurements are performed using the cheaper and more robust filter instruments which may be preferred to scanning and FT instruments for some applications.

When used for quantification, FT-IR and Raman spectroscopy are generally used for determination of minor constituents or subgroups of major constituents where the higher degree of structural information (compared to NIR) can be utilized. For example, both methods have been applied for quantification of the lipid components in oils, margarines and butter such as the content of *cis* and *trans* isomers (Bailey & Horvat, 1972; van de Voort *et al.*, 1995) and degree of unsaturation (Sadeghi-Jorabchi *et al.*, 1990). Furthermore, IR has shown potential for quantification of sucrose, glucose and fructose in juices (Kemsley *et al*., 1992b) and as reviewed in Li-Chan *et al*. (2002) Raman measurements have been correlated with various parameters of carbohydrates such as degree of esterification and amidation of amidated pectins, percent of acetylation of modified wheat, and amylose content of maize starch. The fact that information about the major constituents is also present in the IR and Raman spectra may be exploited for simultaneous quantitative determinations. Hence, FT-IR is, for example, widely used for quantification of both minor and major constituents in dairy products (Andersen *et al.*, 2002).

A study was carried out to investigate whether FT-Raman and chemometrics could be used for quantification of the microconstituent amygdalin in bitter almonds (PAPER I). Amygdalin is a cyanogenic glycoside, i.e. a glycosidic compound with a nitrile group, and the idea behind the study was to utilize the fact that nitrile groups are strongly Raman active, giving rise to an intense vibrational band in an area of the Raman spectrum which is free from interference from other chemical information. A preliminary study was performed on ground samples in order to allow for fake samples with additional amygdalin added to be included in the sample set. This was necessary to ensure a large span in amygdalin concentration. Measurements were performed on a standard addition series in which amygdalin was added to ground sweet almond tissue in seven concentrations between 0-700 nmol/mg. As expected, a distinct band due to the C≡N stretching of the nitrile group was observed in the Raman spectra at 2242 cm^{-1} (Fig. 3.2)

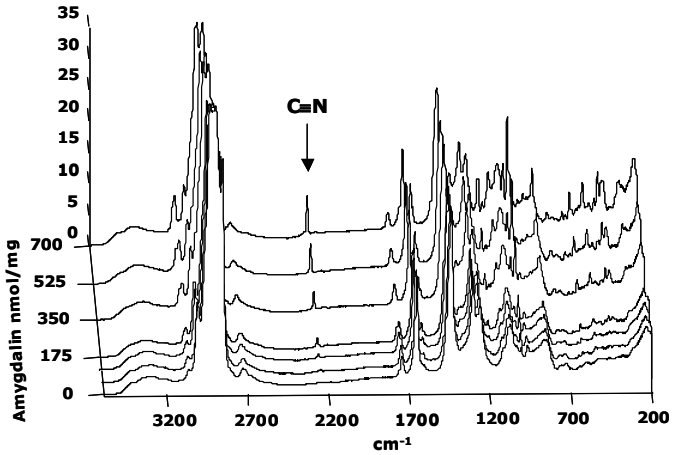


Fig. 3.2 Raman spectra of seven samples of ground sweet almond with amygdalin added in seven concentrations. The C≡N stretching of the amygdalin molecule gives rise to a distinct band at 2242 cm^{-1} which enables quantitative measurements of amygdalin using Raman spectroscopy (PAPER I).

An excellent iPLS calibration could be developed based on the C≡N band and based on different vibrational bands due to stretching of the aromatic ring in amygdalin. The best predictions were obtained using the 3060 cm^{-1} band, probably because it was less influenced by fluorescence and because the nitrile band shifted slightly with increasing amygdalin content. Subsequently, measurements were performed on 18 ground bitter almonds where endogen amygdalin was present. PLSR models were developed for the 18 samples and 4 samples from the standard addition series (0-175 nmol/mg). In this case, the best iPLS model was obtained based on an aromatic stretching band; however, an acceptable model could also be made based on the C≡N band. Since nitrile groups are rare in natural compounds, the C≡N signal is specific to cyanogenic glycosides and cyanogenic compounds in general. Therefore, the model based on the C≡N band shows that Raman spectroscopy has potential for quantification of cyanogenic components in biological material utilizing the unique ability to assess nitrile groups. Thygesen *et al.* (2004) recently showed that the sensitivity of the Raman method for quantification of cyanogenic compounds could be drastically increased using surface-enhanced Raman spectroscopy.

3.4 Qualitative analysis of food using vibrational spectroscopy

FT-IR and Raman spectroscopy are exceptional tools for structural analysis and both methods have been used extensively for qualitative analysis of major as well as minor food constituents (Ozaki *et al.*, 1992; Downey, 1998; Li-Chan *et al.*, 2002). These applications range from quality control to basic research where FT-IR and Raman spectroscopy may be used for compositional identification, detection of adulteration, determination of molecular structures and exploration of structural or conformational changes that occur during processing of foods.

By combining FT-IR and Raman spectroscopy with microscopy, spatially resolved investigations of the chemical composition of heterogeneous food can be made. This was utilized to study the distribution of amygdalin in bitter almonds (PAPER I). Raman microscopic measurements were performed on the transversely sectioned slice of two different bitter almonds where Raman spectra were obtained every 33.3 μm (81 points) along a straight line from the epidermis to the center of the almonds. Since Raman spectra of bitter almonds contain quantitative information about amygdalin via the C≡N stretching band (PAPER I), the variation in the nitrile band was used as an expression of the variation in amygdalin content. In PAPER I, the amygdalin content was estimated using the first score from a PCA on the nitrile peak of the 81 measurements. A simpler approach is to use the area under the nitrile peak as an expression of the amygdalin content, which was used in PAPER III. Either way it was found that amygdalin was not present in measurable amounts in the center of the two almonds, 0.5 mm from the center it was present in small quantities and from this point on the content increased towards the epidermis but the local variation was large (Fig. 3.3). Since cyanogenic glycosides are present in plant tissue to defend the plant against insects and fungi, it makes sense that amygdalin is primarily accumulated close to the surface. Tissue printing was used to investigate the distribution of cyanogenic glycosides in mature seeds from species of the *prunus* family other than bitter almonds (Poulton & Li, 1994), but these measurements only indicated whether or not the cyanogenic glycosides were present in the tissue or not. With Raman microspectroscopy qualitative and quantitative information was obtained with a high spatial resolution through measurements directly on amygdalin. Other successful applications of FT-IR and Raman microspectroscopy to the study of food are described in PAPER III.

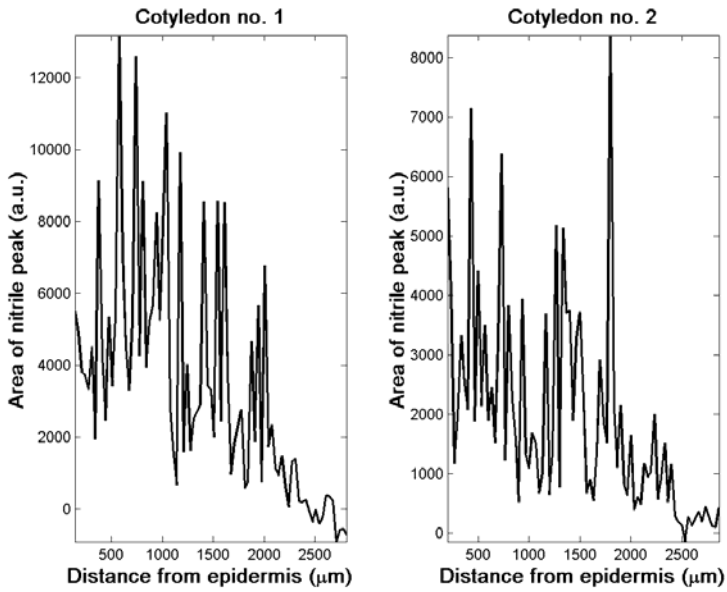


Fig. 3.3 Distribution of amygdalin in two bitter almond cotyledons measured with Raman microspectroscopy. The plots show the area of the nitrile band at approximately 2242 cm^{-1} for 81 positions (every 33 μm) along a line from the epidermis to the center of the almonds (PAPER III).

NIR is not as widely used for qualitative analysis as FT-IR and Raman spectroscopy, since the resolution of the spectra is low. However, NIR is useful for the study of hydrogen bonding. This has been exploited in the study of the state of water in foods and monitoring of structural changes of carbohydrates during processing (Osborne *et al.*, 1993).

3.5 Possibilities of on-line analysis using vibrational spectroscopy

NIR is very applicable for on-line analysis due to the easy sampling, availability of fiber optics and robust instrumentation, for which reason NIR has been applied for on-line analysis in many different food industries. The application of Raman spectroscopy to on-line analysis has only been possible since the development of FT instruments using NIR lasers, as this has significantly resolved the problems of poor resolution and interference due to fluorescence. With the development of optic fibers for remote analysis, FT-Raman is now a qualified method for on-line analysis. Development of FT-IR instruments has also made it possible to perform

on-line IR measurements and FT-IR has been used for in-line compositional analysis of milk (Luinge *et al.*, 1993), as well as on-line for fruit concentrate (Kemsley *et al.*, 1992a) and olive oil (Küpper *et al.*, 2001). However, sample presentation and strong water absorption is still a problem. ATR probes have eased the sampling, but as long as suitable optic fibers are not available, FT-IR cannot be used for remote on-line analysis.

As mentioned above, NIR is already widely used for analysis of the macroconstituents of food and suitable on-line instrumentation has been developed. Thus, FT-Raman and FT-IR spectroscopy is probably going to be implemented mainly as at-line or diverted side stream analysis for monitoring of the minor constituents and quantitative and qualitative analysis for which NIR is not well suited.

It is striking that there is a lot of literature on *potential* on-line applications of vibrational spectroscopy, but very few articles on *actual* on-line applications. This is presumably because the successful on-line implementations are not described in literature since they take place in industry, not in the academic environment, and because the number of on-line implementations is small compared to the number of feasibility studies. Furthermore, in some cases the implementation stops at the off-line or at-line level, since it is very costly to make on-line applications because standard instrumentation and probes have to be adjusted to individual situations and the analyses have to be incorporated in the overall process control system.

4. Exploring foods with LF-NMR relaxometry and chemometrics

4.1 LF-NMR

NMR is based on sample absorption of radio frequencies (RF) in the presence of a static magnetic field. The first experiments were carried out in 1946 by Purcell and Bloch using continuous radio waves, and in 1950 the pulsed echo technique was developed by Hahn. NMR is used to detect the spin resonance of a given nucleus possessing a spin angular momentum. The proton ^{1}H is the most frequently probed nucleus, as the signal is strong due to the high abundance of this nucleus. Other commonly used nuclei in food science are ^{13}C and ^{31}P. Only ^{1}H-NMR was used in this project. Upon excitation with an RF pulse the protons return to their equilibrium via various relaxation processes. In LF-NMR the relaxation decay as a function of time is probed (hence LF-NMR is also called time domain NMR) and the characteristic relaxation time constants T_1 and T_2 may be determined from the decay curves. T_1 describes the relaxation mechanism normally referred to as longitudinal or spin-lattice relaxation and T_2 describes the transverse or spin-spin relaxation. Different pulse experiments are used to measure the different relaxation mechanisms. The most important are inversion recovery (T_1 determination), free induction decay (FID, T_2^* determination), and Carr-Purcell-Meiboom-Gill (CPMG, T_2 determination) (Carr & Purcell, 1954; Meiboom & Gill, 1958). The T_2 determined with FID is normally much smaller than the true T_2, as FID experiments are influenced by the inhomogeneity of the magnetic field, for which reason it is denoted T_2^*. In this project, only CPMG experiments were used (PAPER II, PAPER VI). In Fig. 4.1, decay curves from two CPMG experiments are shown which illustrate the exponential shape of transverse relaxation processes. The intensity of the NMR signal immediately after the RF pulse has been applied is proportional to the amount of protons that is present in the sample and the rate with which they relax depends on the mobility of the protons. In practice, data acquisition cannot begin immediately after application of the RF pulse, so the initial signal intensity is not perfectly proportional to the concentration of the proton populations. The deviations are larger for CPMG data than FID data since the delay before data acquisition can begin is longer. Therefore, LF-NMR data generally do not fulfill the bilinear conditions of PCA and PLSR. Still the chemometric methods have proven to be useful for analysis of LF-NMR decay curves (Bechmann *et al.*, 1999).

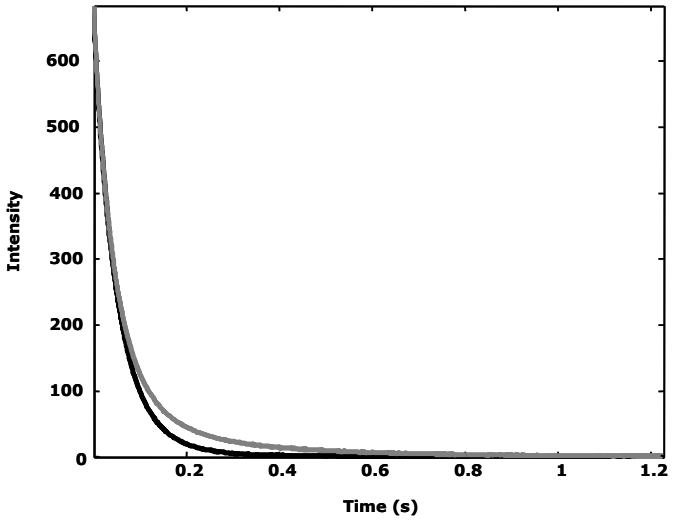


Fig. 4.1 Exponential decaying CPMG curves. The black curve is characterized by a faster relaxation rate than the grey curve, indicating that the black sample contains protons which are bound more tightly than the protons in the grey sample.

4.2 Advantages and disadvantages of LF-NMR

The main advantage of LF-NMR is that the entire sample matrix is probed through the protons. Protons in a solid phase will relax fast, as the energy can easily be transferred to the surrounding molecules, whereas nuclei in more freely moving molecules will be characterized by slower relaxation. In high moisture foods, the major part of the proton signal is due to water. Therefore, the mobility of the protons mainly reflects the state of the water populations. However, the state of the water populations is also an indirect measurement of the state of the molecules making up the surrounding food matrix, i.e. proteins, carbohydrates and ions, as these influence the mobility of the water protons. Another advantage of LF-NMR is that the method is insensitive to the state of the surface or the opacity of the sample, which are problematic to optical methods.

Disadvantages of LF-NMR are that the equipment is complex to optimize compared with, for example, NIR instruments, and that measurements are time-dependent. Therefore, NIR might be preferred for fast, non-invasive quantitative measurements on samples which are relatively homogeneous. However, LF-NMR might be a better choice for heterogeneous samples, as a larger quantity of a sample is probed with LF-NMR than with NIR reflectance or transmission.

4.3 Curve fitting of LF-NMR data

The traditional way of retrieving qualitative and quantitative information from the exponentially decaying FID and CPMG curves is curve fitting where a number (N) of mono-exponentials are fitted to the decay curves:

$$Rmag(t) = \sum_{i=1}^{N} M\left(T_{2,i}\right) \cdot \exp\left(\frac{-t}{T_{2,i}}\right) \qquad \text{(Eq. 4.1)}$$

where *Rmag(t)* is the residual magnetization at a certain time t after application of the first RF pulse which is described as a sum of N mono-exponentials with initial signal intensities (amplitudes) $M(T_{2,i})$ and corresponding relaxation time constants $T_{2,i}$. The signal amplitudes provide quantitative information about the protons, i.e. how many of the protons that belong to each of the N proton populations characterized by the N different relaxation time constants, whereas the relaxation time constants of the N components give qualitative information about the protons as the rate of relaxation depends on the state of the protons.

In discrete exponential curve fitting a small discrete number of components (N) is used in Eq. 4.1 and $M(T_{2,i})$ and $T_{2,i}$ can be found using different non-linear curve-fitting algorithms which are based on iterative least squares minimization of the residuals, such as simplex (Nelder & Mead, 1965). If the rank of the system is not known, the number of components must be determined by evaluation of the residuals as a function of model complexity. It is essential to determine and use the correct number of components, as it affects the interpretation of data and it is decisive for correct estimation of the signal amplitudes and relaxation time constants. However, it can be problematic to determine the correct number of proton populations. This was illustrated in PAPER II where the states of water in meat during cooking in an LF-NMR instrument were monitored by consecutive LF-NMR measurements. The least squares residuals (Fig. 4.2) indicate that the system changes from bi-exponential to tri-exponential during the cooking process. The third component appears somewhere between 40ºC and 60ºC, as the error using two components increases from approximately 40ºC. However, it is impossible to determine when the third component is valid based on the plot of the residuals and the plots of $M(T_{2,i})$ and $T_{2,i}$ (PAPER II). This is unfortunate, as

introduction of a third component at 40ºC and 60ºC, respectively, leads to two plausible, but very different interpretations of the changes taking place in meat during cooking. The study underlines that the determination of the number of components is the vulnerable point of discrete exponential fitting and that it especially can be problematic to determine when the rank of the system changes slowly (PAPER II). Part of the problem is that the rank is determined on the basis of a subjective evaluation of the residuals. It has been suggested to use a cross-validation of the exponential fit in the process of selecting the correct number of components (Pedersen *et al.*, 2000). However, this procedure was not sufficient to clarify the rank of the system in the present case.

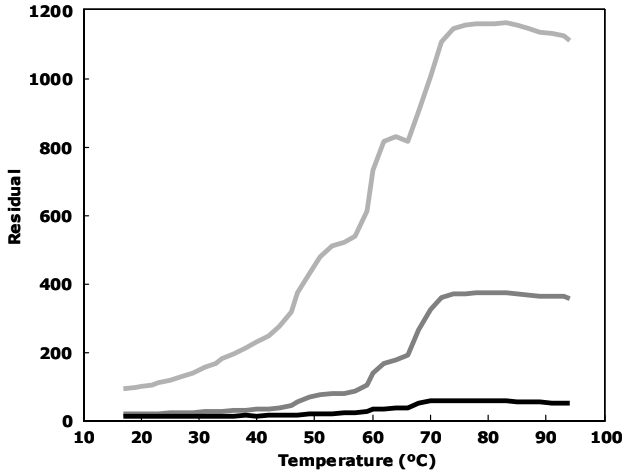


Fig. 4.2 Residuals from discrete exponential fitting of CPMG curves obtained during cooking of meat (16-94ºC). Residuals are shown for calculations using one (light grey), two (dark grey) and three (black) exponentials as a function of temperature (PAPER II).

Despite the fact that it can be difficult to determine the number of components, discrete exponential fitting has been successfully applied for analysis of LF-NMR data from many different foods such as dough and bread (Engelsen *et al.*, 2001), cod (Lambelet *et al.*, 1995) and raw meat (Renou *et al.*, 1985; Fjelkner-Modig & Tornberg, 1986).

Distributed exponential fitting is another approach for analysis of LF-NMR data where the problem with determination of rank is sought to be avoided. Rather than finding a few components (*N)* in Eq. 4.1, a distribution of intensities $M(T_{2,i})$ is

found for a large number of $T_{2,i}$ values sampled linearly or logarithmically in a user-defined interval. This approach is more physically realistic, as the individual protons in a proton population do not experience exactly the same environment and they will therefore posses a distribution of relaxation rates. Accordingly, results from distributed fitting may be easier to interpret. For example, distributed fitting has improved the interpretation of the distribution of water in raw meat compared with discrete fitting (Lillford *et al.*, 1980; Bertram *et al.*, 2001).

The problem with distributed fitting is that the numerical problem and its solution are ill defined. Accordingly, the solution is non-unique and is very sensitive to the choice of parameters in the algorithm and the noise level of the LF-NMR data. The sensitivity of the algorithm was demonstrated in PAPER II, where distributed fitting was performed on LF-NMR data which had been pre-processed using magnitude transformation and principal phase correction (Pedersen *et al.*, 2002b), respectively. The two pre-processing methods result in almost identical data. It is mainly when the system has relaxed to zero intensity that the different handling of noise and negative data points affects the decay curves. Even though the differences between magnitude-transformed and phase-rotated data in this way are very small, distributed fitting of the data described in PAPER II gave very different results, depending on pre-processing method (Fig. 4.3).

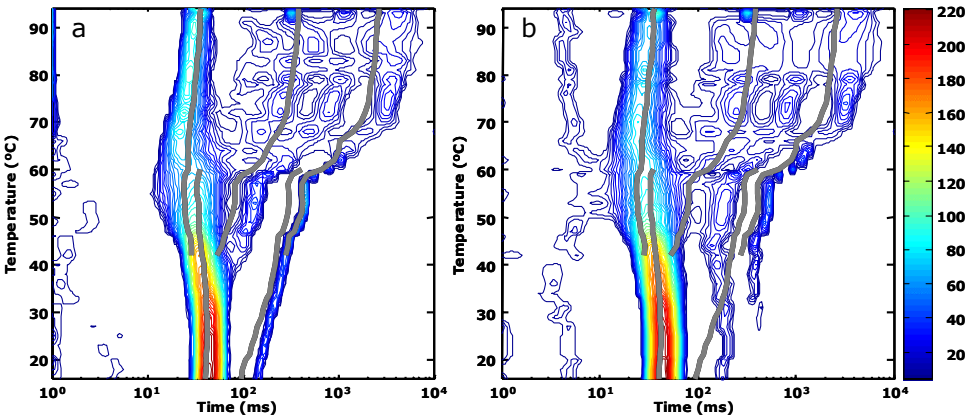


Fig. 4.3 Contour plots of distributed exponential fitting of CPMG data from NMR cooking of meat (16-94ºC): a) magnitude-transformed data and b) phase-rotated data. Discrete exponential fitting is superimposed on the plots, bi- and tri-exponential fitting overlap in the temperature interval 40-60ºC (PAPER II).

Phase-rotation is presumably superior to magnitude transformation, but the result of distributed fitting on magnitude-transformed data agree best with the result of discrete exponential fitting (Fig. 4.3).

4.4 Application of LF-NMR for quantification of food components

Since the intensity of the LF-NMR signal is proportional with the number of protons in a sample the proton density of proton-containing compounds can be quantified with LF-NMR. Protons from the two dominating proton-rich components in the liquid phase of foods, namely water and fats, can be distinguished from protons in the solid phase, i.e. in carbohydrates and proteins. This can be utilized for quantitative determinations, and LF-NMR has accordingly been used for measurement of moisture and fat content in several food systems ranging from relatively simple systems to complex macromolecular foods such as dairy and meat products (Richardson & Steinberg, 1987; Schmidt, 1991). Most of the methods described in the literature are based on one or a few measuring points and this approach for quantitative analysis is also prevalent in industry. However, as reviewed in PAPER IV, it is advantageous to use chemometrics for analysis of LF-NMR data. It has been shown that PLSR on entire decay curves gives increased predictive performance of water and fat content in complex food systems compared with the univariate approach (Gerbanowski *et al.*, 1997; Pedersen *et al.*, 2001). The predictive ability of PLSR is also superior to regressions on signal amplitudes ($M(T_{2,i})$) and time constants ($T_{2,i}$) from curve fitting both for prediction of fat and water (Gerbanowski *et al.*, 1997, Bechmann *et al.*, 1999) and for prediction of water-holding capacity (WHC) in meat (Jepsen *et al.*, 1999; Brøndum *et al.*, 2000).

4.5 Application of LF-NMR and chemometrics for exploratory analysis

LF-NMR is a unique tool for exploratory measurement of foods, because the state of the entire food matrix can be studied non-invasively. In this way LF-NMR works as a window into the intact food matrix, which allows measurement of phenomena not possible with other techniques. This feature in combination with the possibility of dynamic measurements allows observation of the transformations taking place in foods during processing or storage. LF-NMR has, for example, been used for monitoring gelatinization of starch during baking (Engelsen *et al.*, 2001), changes in proteins in meat post mortem (Bertram *et al.*, 2004; PAPER VI)

and heat denaturation of proteins in meat during cooking (Borisova & Oreshkin, 1992; PAPER II).

In PAPER II, the data compression abilities of PCA were utilized to explore the vast amount of data collected during the dynamic study. The PCA reduces the dimensionality of the data and gives an overview of where the major changes in the relaxation decays, i.e. in the properties of the water, occur (Fig. 4.4). The score plots can be used to clearly point out when major changes in the meat take place, namely at 46, 57, 66 and 76ºC. Since LF-NMR does not give a chemical fingerprint, as is the case with vibrational spectroscopy, the loadings from a PCA on LF-NMR data are often of minor use and the PCA must be interpreted using *a priori* knowledge. In this case, the major changes in the scores in Fig. 4.4, i.e. major changes in the meat, can be ascribed to different stages of heat denaturation of the different groups of meat proteins.

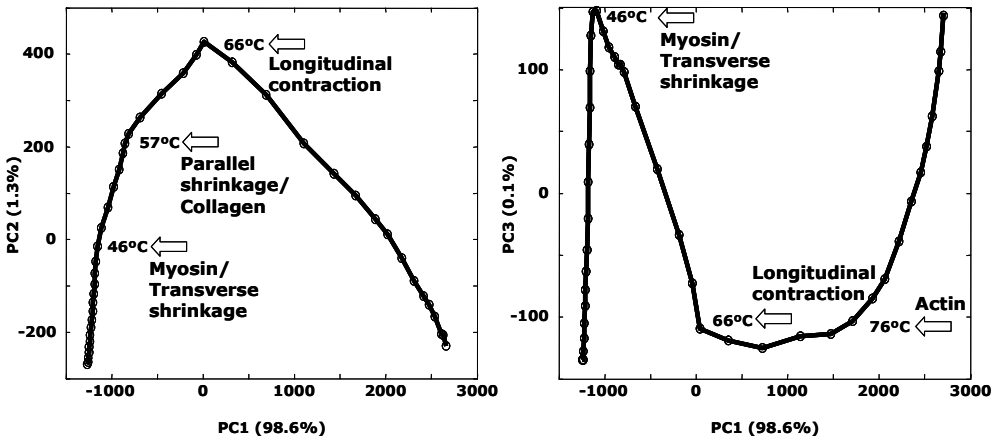


Fig. 4.4 Score plots from PCA on 43 LF-NMR measurements (mean of three experiments) obtained during cooking of meat (16-94ºC). Percentage of explained variance for PC's 1-3 is shown in brackets. The center temperatures are indicated where major changes in the scores are seen and the types of protein denaturation which occur at these temperatures according to the literature are briefly described (PAPER II).

While PCA is the best data analytical method for contrasting the differences in the relaxation curves, curve fitting can help in the interpretation of the LF-NMR data. Estimation of the number of proton populations and the quantitative and qualitative characteristics for each population may elucidate the quantitative and qualitative

state of the system and changes occurring during processing. As mentioned, the results from distributed fitting are more intuitively comprehensible than the results from discrete fitting. In PAPER II, curve fitting revealed that an additional water component is developed in meat during cooking. Based on the two plots from distributed fitting (Fig. 4.3) a dual hypothesis was put forward: the new water population can be either water expelled from the meat matrix or alternatively water trapped in a myosin gel.

4.6 Possibilities of at-line/on-line measurements using LF-NMR

LF-NMR is being used routinely for off-line and at-line determinations of solid-fat and water content in the industry, but it is yet to be applied on-line. The reason for this is primarily that LF-NMR equipment, which is suited for implementation on-line, is not available. A number of challenging problems have to be met in order to develop on-line equipment which can utilize the huge potential of NMR for on-line analysis (Hills *et al.*, 2003). Primarily, the time dependence of the NMR signals is problematic, since it does not allow measurements on moving material. Therefore, the product needs to be stopped during measurement. Furthermore, it is problematic to obtain a stable external magnetic field in a stainless steel environment.

5. Ultrasonic measurements of food quality

5.1 Ultrasound

Ultrasonic energy is sound waves whose frequency by definition is above the upper limit of human audible sound (> 20 kHz). High-intensity ultrasound is utilized to alter the chemical and physical properties of food, whereas low-intensity ultrasound (typically less than 1W/cm^2, McClements, 1997) is non-destructive and can be used for food analysis. In the following, ultrasound refers to low-intensity ultrasound. The theory of ultrasound complies with the theory of sound which was put forward by Lord Rayleigh in 1877. The basis for developing ultrasonic transducers and thereby the basis for ultrasonic measurements was founded in 1880 when the piezoelectric effect was discovered by Jacque and Pierre Curie (Mason & Lorimer, 2002).

The propagation of an ultrasonic wave through a material depends on the compression and extension of the molecular bonds in the material, but the ultrasound does not interact with specific chemical compounds. Each constituent in a sample contributes to the overall ultrasonic signal by affecting the velocity and attenuation, i.e. the decrease in amplitude of the propagating pulse. The classical method for extraction of information about the composition, structure and physical state of foods is therefore to measure the velocity and attenuation coefficient and correlate these to the quantitative or qualitative parameter of interest. The attenuation coefficient of a material depends on several factors, mainly intrinsic attenuation, scattering and reflection. Scattering occurs when an ultrasonic wave hits a discontinuity (e.g. a particle) and is scattered in directions that are different from that of the incident wave. The extent of scattering is highly dependent on the relationship between the wavelength of the ultrasound and the dimensions of the particle. Accordingly, the ultrasonic properties as a function of frequency can be used to determine the size and concentration of particles in a material. Reflection occurs if the food system under investigation consists of layers of different material, since an ultrasonic pulse is partly reflected and partly transmitted when it encounters a boundary between two materials of different acoustic impedance. The larger the difference in impedance, the higher is the proportion of the ultrasound which is reflected.

Ultrasonic measurements are most commonly performed as pulse-echo experiments, where a sample is exposed to a pulse of ultrasound and the echoes of

the sound over time are detected. The signals obtained (amplitude as a function of time) are called A-scans. Most studies on food described in the literature are performed in a measurement cell, where the major echoes are reflections from the cell wall (Fig. 5.1). The echoes are used to determine the ultrasonic velocity and attenuation, which are subsequently correlated with the quality parameter of interest using univariate regression.

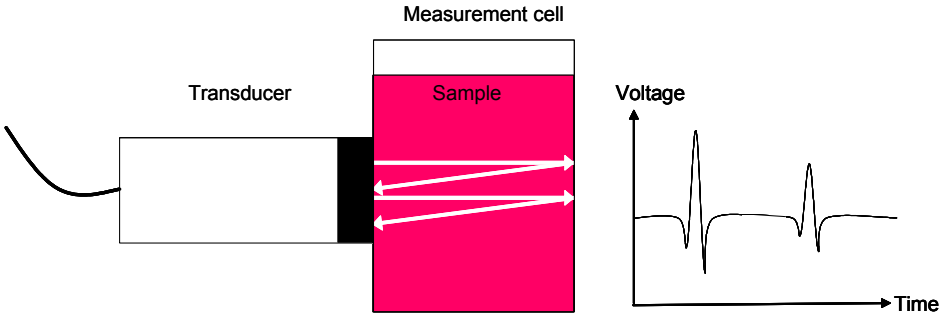


Fig. 5.1 Set-up of ultrasonic pulse-echo experiment. The transducer converts an electrical pulse to an ultrasonic pulse which travels across the sample and is reflected from the inside wall of the measurement cell. The ultrasonic velocity and attenuation coefficient are determined from the time between successive echoes and the decrease in the amplitude of the echoes with distance traveled, respectively, (based on McClements, 1997).

In PAPER VI, an ultrasound instrument developed for on-line analysis of back-fat thickness of pig carcasses was used (Fig. 5.2). Such equipment is also based on the pulse-echo technique, but the echoes which are detected are reflections of ultrasound from the internal structures. All tissue interfaces give rise to echoes, as the ultrasonic pulse is partly reflected when it encounters a boundary between two types of tissue. Therefore, the A-scans are more complex than signals obtained on samples in a measuring cell (Fig. 5.3a). When several A-scans are performed next to each other, the signals make up an ultrasonic image. The instrumentation used contained an array of 64 transducers, but the signal from the three outermost transducers on each side of the array were discarded in order to eliminate noise due to edge errors. Thus an image consisting of 58 A-scans were obtained in a single measurement (Fig. 5.3b). Classical image analysis is usually used to extract information from the ultrasonic images. The locations of specific physical features are determined and used to calculate the parameter of interest e.g. fat thickness.

Fig. 5.2 The Ultrafom 300 pistol (SFK Technology, Herlev, Denmark) used for non-invasive ultrasonic measurements of pig carcasses. When the pistol is triggered, the array of 64 transducers in the front of the pistol emits ultrasound pulses of 3.5 MHz and the ultrasound echoes returning to the transducers are subsequently detected.

5.2 Advantages and disadvantages of ultrasound

The major advantage of ultrasound compared with other spectroscopic techniques is that the ultrasonic instrumentation is inexpensive, very robust and can easily be implemented in-line. Another advantage of ultrasound over the optical spectroscopic methods is that it is insensitive to optical opacity. In addition, it may be possible to carry out measurements on larger samples compared with optical spectroscopy, but the sample thickness depends on the attenuation of the food system. However, a major disadvantage of ultrasonic techniques is that air attenuates the signal considerably. Thus, it is important to secure absolute contact between sample and transducer in order to use the ultrasound technique. In ultrasonic imaging of carcasses, this can be secured by applying oil to the skin where measurements are performed. However, in most cases it will not be practical to apply oil to the surface of the food to be investigated. Furthermore, air within the

food samples also represents a problem. Presence of small gas bubbles in a sample can attenuate ultrasound so much that an ultrasonic wave cannot be propagated through the sample (McClements, 1997).

5.3 Ultrasonic analysis of foods

The most successful application of ultrasound on foods is the assessment of body composition and back-fat thickness of meat from live animals and carcasses. Already in the 1950's, equipment based on A-scans was routinely used in the livestock industry to indicate tissue depths from the skin surface (Fisher, 1997). Today, several commercial instruments based on imaging techniques are available for on-line estimation of the lean percentage and fat thickness in pig carcasses. Some are based on single location measurements, while the Autofom (SFK Technologies, Herlev, DK), which is a three dimensional ultrasound system, generates a full scan of the carcass. Analysis of the Autofom data is based on PLSR models on 127 image features extracted from the ultrasound images (Brøndum *et al.*, 1998) and it is recognized that the prediction ability is higher for the Autofom than for single location instruments (Fortin *et al.*, 2003). This emphasizes that a multivariate approach is superior for investigation of a multivariate system.

Interest in using ultrasound for measurement of foods in other areas than the meat industry has increased since the late 1980's. Applications of ultrasonic pulse-echo measurements to foods based on classical data analysis, i.e. estimation of the velocity of attenuation, were reviewed by McClements (1997). The review article pointed to the following as promising application areas of ultrasound: monitoring the concentration of aqueous solutions and suspensions, determination of droplet size and concentration in emulsions, monitoring crystallization in fats and monitoring creaming profiles in emulsions and suspensions.

Ultrasound is only suggested for quantitative determinations in simple foods such as aqueous solutions (e.g. juice and brine), because measurements are (currently) based on determination of the ultrasonic velocity and/or attenuation and development of empiric or semi-empiric equations for the parameter of interest. This approach works when only one parameter in the systems varies, such as the sugar concentration in fruit juices and drinks (Contreras *et al.*, 1992). However, in multi-component foods the ultrasonic properties are determined by more factors. Hence, the univariate approach does not work. This underlines the need for

multivariate data analysis of the ultrasound data when the system under investigation is multivariate.

One way to obtain more information about the system under investigation is to utilize the fact that the ultrasonic properties of the individual components may have different dependence of temperature and/or frequency. The temperature dependence of the velocity has, for example, been used for determination of water, fat and non-fat-solids in fish and a meat product (Sigfusson *et al.*, 2001; Simal *et al.*, 2003). However, the need for measurements at several temperatures makes the method difficult and inapplicable for on-line use. Thus, NIR or LF-NMR measurements would probably be a better choice for fast non-destructive compositional determinations of the investigated foods than the ultrasound methods reported.

The strength of ultrasound compared to other spectroscopic methods is that physical phenomena can be analyzed. Thus, ultrasound has potential for unique non-invasive investigation of droplets, air bubbles or particles in foods. This feature may also be utilized for monitoring of phase transitions such as fat crystallization in oil and gelation of polymers (Bijnen *et al.*, 2002). In this context, exploratory studies would evidently also be of interest. Therefore, it should be investigated whether chemometrics can be applied to ultrasound data (PAPER VI). If it is successfully implemented, chemometrics may be important for optimal utilization of ultrasound in the future.

5.4 Exploring the applicability of chemometrics on ultrasound data

The study described in PAPER VI was an exploratory study in which it was tested whether a chemometric approach could be used for extraction of information on WHC from ultrasonic images. Chemometric methods cannot be applied to the raw oscillation ultrasonic signals, as they do not satisfy the requirements for bilinear models. Therefore, chemometric models were applied to Fourier transformed ultrasound data originating from the meat part of ultrasonic images of pig carcasses (Fig. 5.3b).

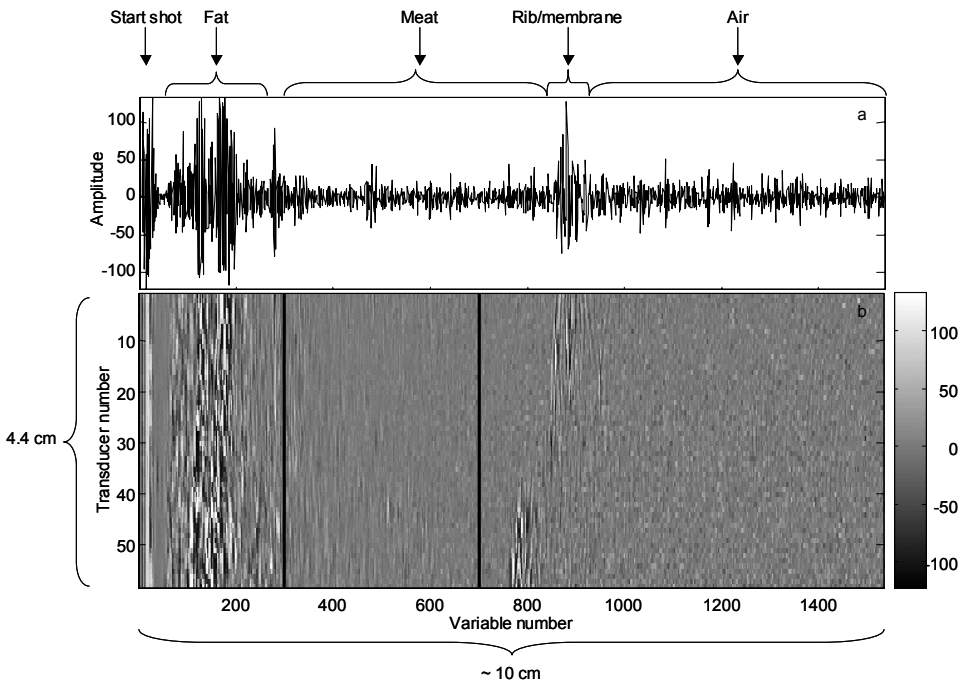


Fig. 5.3 Ultrasound data from a) one transducer b) 58 transducers making up an image. Width of transducer array and approximate traveling distance of the ultrasound are indicated. The area between the two vertical lines in B indicates the part of the images which was used for data analysis (variables 301-700) (PAPER VI).

Frequency analysis using Fourier transformation has been successful in laboratory studies on excised muscles for determination of intramuscular fat and sensory attributes (Whittaker *et al.*, 1992; Park *et al.*, 1994). However, in PAPER VI, a separation according to WHC was not found using PCA on the frequency spectra of ultrasound data obtained on pig carcasses early post mortem. Classification based on histograms, which is an indirect measurement of the attenuation, was more successful. Separation according to WHC was found 85 min post mortem. However, since this was the last measurement out of eight starting 15 min post mortem, further investigations need to be made to verify the result.

In any case, the WHC needs to be predicted earlier than 85 min post mortem in order to be of value in the production. LF-NMR (Bertram *et al.*, 2003; PAPER VI) and vibrational spectroscopy (Forrest *et al.*, 2000; Pedersen *et al.*, 2003) have shown higher sensitivity to WHC earlier post mortem, however, these feasibility studies were performed on excised meat samples. The ultrasonic signals obtained

non-invasively are severely influenced by the tissue layers preceding the meat. Due to the low signal-to-noise ratio the data are not suited for concluding whether multivariate data analysis can be useful for extraction of information from ultrasound data in general. It is suggested that further studies on the applicability of chemometrics to ultrasonic images be performed on signals obtained directly on the material to be investigated. Furthermore, studies on ultrasonic spectra, e.g. attenuation as a function of frequency, which can be analyzed directly with chemometrics, would be interesting.

5.5 Possibilities of on-line applications using ultrasound

As mentioned in section 5.3, ultrasonic imaging is widely used for routine on-line measurements of carcasses. Ultrasound can easily be implemented for in-line measurements, e.g. mounted on a pipe, as the instrumentation is robust and can measure through opaque materials. However, the feasibility studies on ultrasonic measurement of foods have not led to in-line or on-line implementations described in the literature.

6. Conclusions and perspectives

As reviewed in the previous chapters, the potential of the spectroscopic techniques for analysis of foods continues to increase with the advances being made within spectroscopic instrumentation and analysis of spectroscopic data. In this project, innovative applications of various spectroscopic techniques and chemometrics have been explored resulting in improved analysis and understanding of various foods.

In PAPER II, LF-NMR and chemometrics were used to monitor the changes taking place in meat during cooking. The temperatures at which the major changes in meat occur were identified and it was shown that an additional water population appears above approximately 40ºC. The new water component probably reflects that water is expelled from the meat matrix, but it could also be due to the formation of a porous myosin gel. Such inductive, exploratory analyses of foods are important in food research, as they lead to generation of novel hypotheses about the food systems. It is remarkable in itself that the interior of a food can be assessed on the basis of non-invasive spectroscopic measurements on an intact food matrix; however, the powerful exploratory possibilities are only obtained by combining the spectroscopic techniques with chemometrics. Therefore, it is essential that chemometrics be widely adapted by food scientists. This still lacks within the field of LF-NMR, even though the advantages of applying chemometrics to multivariate LF-NMR data are obvious, as demonstrated in PAPER IV. Among other things it was shown that it may be beneficial to work with tri-linear LF-NMR data since application of multi-way chemometrics offer other possibilities than the bi-linear models for extraction of characteristic relaxation times. NMR is a unique tool for analysis of foods and it will certainly be of great importance in future food research. It is a drawback that the potential of NMR at present cannot be exploited for on-line analysis, because equipment suited for on-line analysis is not available. Since the demand for qualified on-line analysis in the industry is high, an effort should be made to meet the challenges of applying NMR to process analysis.

Ultrasound has been implemented successfully for on-line analysis of meat. However, the perspectives of using ultrasound in semi-solid foods are uncertain, because the data analytical approach is univariate. As for the other spectroscopic techniques, the success of ultrasound in food science and industry may be highly dependent on the possibilities of applying chemometrics to the spectroscopic data. In PAPER VI, a chemometric approach was tested for separation of ultrasonic measurements of porcine meat according to WHC based on ultrasonic Fourier

spectra. This was not successful; however, the signal-to-noise ratio of the ultrasonic signals was low. Therefore, this study cannot be used for making general conclusions about the applicability of chemometrics to ultrasound data. Further studies should be made, as it would be of great interest to the food industry if the great potential of ultrasound for on-line analysis could be exploited. Calculation of histograms from the raw ultrasound signals indicated that the employed ultrasound method could differentiate high and low water-holding capacity. However, classification was not obvious before 85 min post mortem, which is not sufficiently early for on-line use in a commercial abattoir.

One area in which chemometrics has been fully accepted is for quantitative analysis using NIR spectroscopy. During recent decades the combination of NIR spectroscopy and chemometrics has revolutionized quality control, as it has been successfully implemented in some of the major food, agricultural and pharmaceutical industries. The application areas of NIR will presumably increase further in the future, because now production areas with a smaller throughput of food or agricultural products also dare to invest money and time in the development of a useful NIR calibration in order to get faster, cheaper and more frequent measurements in the future. Grass seed production in Denmark is a relatively small-scale production (2-3 % of the total agricultural cultivated area in Denmark), yet there is perspective in developing a NIR method for determination of total nitrogen concentration, which can substitute the routine Dumas method (Hansen, 1989), in grass grown for seed (PAPER V). Furthermore, the Process Analytical Technology initiative (http://www.fda.gov/cder/OPS/PAT.htm) from the U.S. Food and Drug Administration is giving rise to a new boost in the development of the NIR technology originating in the pharmaceutical industry. Apart from on-line analysis, a future issue will presumably be improvement of handheld instrumentation, which will expand the application areas of NIR even further (McClure, 2004).

Raman spectroscopy is not nearly as widely used as NIR. However, the recent developments of instrument technology with increased signal-to-noise ratio, which also allow on-line measurements, may increase the future use of Raman spectroscopy. In PAPER I, it was shown that the toxic cyanogenic compound amygdalin in bitter almonds could be quantified *in situ* using Raman spectroscopy. Furthermore, the vibrational band due to C≡N stretching of amygdalin allowed for qualitative determination of cyanogenic compounds. Accordingly, Raman

microspectroscopy could be utilized to study the distribution of amygdalin in bitter almonds based on the C≡N stretching band. Amygdalin was absent in the center of the bitter almonds, the concentration increased towards the surface, but the variation at the microscopic level was high. Other applications of IR and Raman microspectroscopy were presented in PAPER III and the advantages and disadvantages of the two methods were compared. The resolution of FT-IR microspectroscopy (10x10 μm) is relative poor in relation to food heterogeneity. Raman microspectroscopy using visual lasers has much better spatial resolution but the signal-to-noise ratio is lower and heat inducement and fluorescence emission may be problematic.

Foods are generally heterogeneous at the macroscopic as well as the microscopic level. This variation in the chemical composition cannot be assessed using classical wet chemical analysis, for which reason studies using NIR, IR or Raman imaging for investigation of the macro and microstructure of other food systems appear to be highly relevant. The use of spectroscopic imaging in general is likely to increase in the future as improved imaging equipment emerges and the handling of the large amounts of data contained in images is no longer limited by lack of computer power.

To summarize, the two areas of spectroscopic analysis of foods which will presumably expand the most in the near future are on-line analysis and imaging. Focus should be on improving the spectroscopic instrumentation as well as the chemometric tools for these areas of application. In this connection, important chemometric areas will presumably be process analytical chemometrics, multivariate image analysis (Esbensen & Geladi, 1989; Geladi & Grahn, 1996) and multiway chemometrics (Harshman, 1970; Bro, 1997), the latter of which will be of importance to multivariate image analysis and analysis of other multi-dimensional data structures.

References

Andersen, S.K., Hansen, P.W. & Andersen H.V. (2002). Vibrational spectroscopy in the analysis of dairy products and wine. In *Handbook of Vibrational Spectroscopy*. Eds. J.M. Chalmers and P.R. Griffiths. John Wiley & Sons Ltd, Chichester, UK, pp. 3672-3681.

Bailey, G.F. & Horvat, R.J. (1972). Raman spectroscopic analysis of the *cis/trans* isomer composition of edible vegetable oils. *Journal of the American Oil Chemists' Society*, 49, 494-498.

Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43, 772-777.

Bechmann, I.E., Pedersen, H.T., Nørgaard, L. & Engelsen, S.B. (1999). Comparative chemometric analysis of transverse low-field ^{1}H-NMR relaxation data. In *Advances in Magnetic Resonance in Food Science*. Eds. P.S. Belton, A.M. Gill, I. Delgadillo, The Royal Society of Chemistry, Cambridge, UK, pp. 217-225.

Bertram, H.C., Karlsson, A.H., Rasmussen, M., Dønstrup, S., Pedersen, O.D. & Andersen, H.J. (2001). The origin of bi-exponential T2 relaxation in muscle myowater. *Journal of Agricultural and Food Chemistry*, 49, 3092-3100.

Bertram, H.C., Whittaker, A.K., Karlsson, A.H. & Andersen, H.J. (2003). PH-dependence of the progression in NMR T_2 relaxation times in post mortem muscle. *Journal of Agricultural and Food Science*, 51, 4072-4078.

Bertram, H.C., Schäfer, A., Rosenvold, K. & Andersen, H.J. (2004). Physical changes of significance for early post mortem water distribution in porcine *M. longissimus*. *Meat Science*, 66, 915-924.

Bijnen, F.G.C., van Aalst, H., Baillif, P.-Y., Blonk, J.C.G., Kersent, D., Kleinherenbrink, F., Lenke, R. & vander Stappen, M.L.M. (2002). In-line structure measurement of food products. *Powder Technology*, 124, 188-194.

Borisova, M.A. & Oreshkin, E.F. (1992). On the water condition in pork meat. *Meat Science*, 31, 257-265.

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

Brøndum, J., Egebo, M., Agerskov, C. & Busk, H. (1998). On-line pork carcass grading with the Autofom ultrasound system. *Journal of Animal Science*, 76, 1858-1868.

Brøndum, J., Munck, L., Henckel, P., Karlsson, A., Tornberg, E. & Engelsen, S.B. (2000). Prediction of water-holding capacity and composition of porcine meat with comparative spectroscopy. *Meat Science*, 55, 177-185.

Carr, H.Y. & Purcell, E.M. (1954). Effect of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94, 630-638.

Coates, J. (2002). Classic methods of quantitative analysis. In *Handbook of Vibrational Spectroscopy*. Eds. J.M. Chalmers & P.R. Griffiths, John Wiley & Sons Ltd, Chichester, UK, pp. 2235-2257.

Contreras, I.N., Fairley, P., McClements, D.J. & Povey, M.J.W. (1992). Analysis of the sugar content of fruit juices and drinks using ultrasonic velocity measurements. *International Journal of Food Science and Technology*, 27, 515-529.

Downey, G. (1998). Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics. *Trends in Analytical Chemistry*, 17, 418-424.

Engelsen, S.B., Jensen, M.K., Pedersen, H.T., Nørgaard, L. & Munck, L. (2001). NMR-baking and multivariate prediction of instrumental texture parameters in bread. *Journal of Cereal Science*, 33, 59-69.

Esbensen, K. & Geladi, P. (1989). Strategy of multivariate image analysis (MIA). *Chemometrics and Intelligent Laboratory Systems*, 7, 67-86.

Fisher, A.V. (1997). A review of the technique of estimating the composition of livestock using the velocity of ultrasound. *Computers and Electronics in Agriculture*. 17, 217-231.

Fjelkner-Modig, S. & Tornberg, E. (1986). Water distribution in porcine *M. longissimus dorsi* in relation to sensory properties. *Meat Science*, 17, 213-231.

Forrest, J.C., Morgan, M.T., Borggaard, C., Rasmussen, A.J., Jespersen, B.L. & Andersen, J.R. (2000). Development of technology for the early post mortem prediction of water holding capacity and drip loss in fresh pork. *Meat Science*, 55, 177-185.

Fortin, A., Tong, A.K.W., Robertson, W.M., Zawadski, S.M., Landry, S.J., Robinson, D.J., Liu, T. & Mockford, R.J. (2003). A novel approach to grading pork carcasses: computer vision and ultrasound. *Meat Science*, 63, 451-462.

Geladi, P., McDougall, D. & Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy*, 39, 491-500.

Geladi, P. & Grahn, H. (1996). *Multivariate Image Analysis*, John Wiley & Sons Ltd, Chichester, UK.

Gerbanowski, A., Rutledge, D.N., Feinberg, M.H. & Ducauze, C.J. (1997). Multivariate regression applied to time domain nuclear magnetic resonance signals: Determination of moisture in meat products. *Sciences des Aliments*, 17, 309-323.

Griffiths, P.R. (2002). Introduction to vibrational spectroscopy. In *Handbook of Vibrational Spectroscopy*. Eds. J.M. Chalmers & P.R. Griffiths, John Wiley & Sons Ltd, Chichester, UK, pp. 33-43.

Hansen, B. (1989). Determination of nitrogen as elementary N, an alternative to Kjeldahl. *Acta Agriculturae Scandinavica*, 39, 113-118.

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

Hendra, P., Jones, C. & Warnes, G. (1991). *Fourier Transform Raman Spectroscopy*. Ellis Horwood Limited, Chichester, UK.

Hills, B.P., Meriodeau, L. & Wright, K.M. (2003). NMR studies of complex foods in off-line and on-line situations. In *Magnetic Resonance in Food Science –Latest Developments*. Eds. P.S. Belton, A.M. Gil, G.A. Webb, & D.N. Rutledge, The Royal Society of Chemistry, Cambridge, UK, pp. 186-198.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, 24, 417-441.

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

Höskuldsson, A. (1996). *Prediction Methods in Science and Technology*. Thor Publishing, Copenhagen, Denmark.

Jepsen, S.M., Pedersen, H.T. & Engelsen, S.B. (1999). Application of chemometrics to low-field ^{1}H NMR relaxation data of intact fish flesh. *Journal of the Science of Food and Agriculture*, 79, 1793-1802.

Kemsley, E.K., Wilson, R.H. & Belton, P.S. (1992a). Potential of Fourier-transform infrared-spectroscopy and fiber optics for process-control. *Journal of Agricultural and Food chemistry*, 40, 435-438.

Kemsley E.K., Zhuo, L., Hammouri, M.K. & Wilson, R.H. (1992b). Quantitative analysis of sugar solutions using infrared spectroscopy. *Food Chemistry*, 44, 299-304.

Küpper, L., Heise, H.M., Lampen, P., Davies, A.N. & McIntyre, P. (2001). Authentication and quantification of extra olive oils by attenuated total reflectance infrared spectroscopy using silver halide fiber probes and partial least-squares calibration. *Applied Spectroscopy*, 55, 563-570.

Lambelet, P., Renevey, F., Kaabi, C. & Raemy A. (1995). Low-field nuclear magnetic resonance relaxation study of stored or processed cod. *Journal of Agricultural and Food Chemistry*, 43, 1462-1466.

Leardi, R. (2001). Genetic algorithms in chemometrics and chemistry: a review. *Journal of Chemometrics*, 15, 559-569.

Li-Chan, E.C.Y., Ismail, A.A., Sedman, J. & van de Voort, F.R. (2002). Vibrational spectroscopy of food and food products. In *Handbook of Vibrational Spectroscopy*. Eds. J.M. Chalmers & P.R. Griffiths, John Wiley & Sons Ltd, Chichester, UK, pp. 3629-3662.

Lillford, P.J., Clark, A.H. & Jones, D.V. (1980). Distribution of water in heterogeneous food and model systems. In *Water in polymers* Ed. R.J. Rowland, American Chemical Society, Washington D.C., pp. 177-194.

Luinge, H.J., Hop, E., Lutz, E.T.G., van Hemert, J.A. & de Jong, E.A.M. (1993). Determination of the fat, protein, and lactose content of milk using Fourier transform infrared spectrometry. *Analytica Chimica Acta*, 284, 419-433.

Martens, H. & Martens, M. (2000). Modified jack-knife estimation of parameter uncertainty in bilinear modeling by partial least squares regression (PLSR). *Food Quality and Preference*, 11, 5-16.

Martens, H., Nielsen, J.P. & Engelsen, S.B. (2003). Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Analytical Chemistry*, 75, 394-404.

Martens, H. & Næs, T. (1989). *Multivariate Calibration*. John Wiley and Sons, Chichester, UK.

Mason, T.J. & Lorimer, J.P. (2002). *Applied Sonochemistry: Uses of Power Ultrasound in Chemistry and Processing*. Wiley-VCH Verlag GmbH, Weinheim, Germany.

McClements, D.J. (1997). Ultrasonic characterization of foods and drinks: principles, methods, and applications. *Critical Reviews in Food Science and Nutrition*, 37, 1-46.

McClure, W.F. (2004). 204 years of near infrared technology: 1800-2003. *Journal of Near Infrared Spectroscopy*, 11, 487-518.

McCreery, R.L. (2000). *Raman Spectroscopy for Chemical Analysis*. John Wiley & Sons. Inc., New York, USA.

Meiboom, S. & Gill, D. (1958). Modified spin-echo method for measuring nuclear relaxation times. *The Review of Scientific Instruments*, 29, 688-691.

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.

Nelder, J.A. & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308-313.

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L. & Engelsen, S.B. (2000). Interval partial least squares regression (iPLS): A comparative chemometric study with an example from near infrared spectroscopy. *Applied Spectroscopy*, 50, 413-419.

Osborne, B.G., Fearn, T. & Hindle, P.H. (1993). *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. (2nd ed.), Longman Scientific & Technical, Harlow, UK.

Ozaki, Y., Cho, R., Ikegaya, K., Muraishi, S. & Kawauchi, K. (1992). Potential of near-infrared Fourier transform Raman spectroscopy in food analysis. *Applied Spectroscopy*, 46, 1503-1507.

Park, B., Whittaker, A.D., Miller, R.K. & Bray, D.E. (1994). Measuring Intramuscular fat in beef with ultrasonic frequency analysis. *Journal of Animal Science*, 72, 117-125.

Pedersen, D.K., Martens, H., Nielsen, J.P. & Engelsen, S.B. (2002a). Near infrared absorption and scattering separated by Extended Inverted Signal Correction (EISC). Analysis of NIT spectra of single wheat seeds. *Applied Spectroscopy*, 56, 1206-1214.

Pedersen, D.K., Morel, S., Andersen, H.J. & Engelsen, S.B. (2003). Early prediction of water-holding capacity in meat by multivariate vibrational spectroscopy. *Meat Science*, 65, 581-592.

Pedersen, H.T., Berg, H., Lundby, F. & Engelsen, S.B. (2001). The multivariate advantage in fat determination in meat by bench-top NMR. *Innovative Food Science & Emerging Technologies*, 2, 87-94.

Pedersen, H.T., Bro, R. & Engelsen, S.B. (2002b). Towards rapid and unique curve resolution of low-field NMR relaxation data. Tri-linear SLICING versus two-dimensional curve fitting. *Journal of Magnetic Resonance*, 157, 141-155.

Pedersen, H.T., Munck, L. & Engelsen, S.B. (2000). Low-field ^{1}H nuclear magnetic resonance and chemometrics combined for simultaneous determination of water, oil, and protein content in oilseeds. *Journal of the American Oil Chemists' Society*, 77, 1069-1076.

Poulton, J.E. & Li, G.P. (1994). Tissue level compartmentation of (R)-amygdalin and amygdalin hydrolase prevents large-scale cyanogenesis in undamaged *prunus* seeds. *Plant Physiology*, 104, 29-35.

Renou, J.P., Monin, G. & Sellier, P. (1985). Nuclear magnetic resonance measurements on pork of various qualities. *Meat Science*, 15, 225-233.

Richardson, S.J. & Steinberg, M.P. (1987). Applications of nuclear magnetic resonance. In *Water activity: Theory and Applications to Foods*. Eds. L.B. Rockland and L.R. Beuchat, Marcel Dekker, New York, pp. 235-285.

Sadeghi-Jorabchi, H., Hendra, P.J., Wilson, R.H. & Belton, P.S. (1990). Determination of the total unsaturation in oils and margarines by Fourier transform Raman spectroscopy. *Journal of the American Oil Chemists' Society*, 67, 483-486.

Schmidt, S.J. (1991). Determination of moisture content by pulsed nuclear magnetic resonance spectroscopy. In *Water Relationships in Food*. Eds. H. Levine and L. Slade, Plenum Press, New York, pp. 599-613.

Sheppard, N. (2002). The historical development of experimental techniques in vibrational spectroscopy. In *Handbook of Vibrational Spectroscopy*. Eds. J.M. Chalmers & P.R. Griffiths, John Wiley & Sons Ltd, Chichester, UK, pp. 1-32.

Sigfusson, H., Decker, E.A. & McClements, D.J. (2001). Ultrasonic characterization of Atlantic mackerel (*Scomber scombrus*). *Food Research International*, 34, 15-23.

Simal, S., Benedito, J., Clemente, G., Femenia, A. & Rosseló, C. (2003). Ultrasonic determination of the composition of a meat-based product. *Journal of Food Engineering*, 58, 253-257.

Thygesen, L.G., Jørgensen, K., Møller, B.L. & Engelsen, S.B. (2004). Raman spectroscopic analysis of cyanogenic glucosides in plants. Development of a flow injection surface enhanced Raman scatter (FI-SERS) method for determination of cyanide. *Applied Spectroscopy*, 58, 212-217.

Van de Voort, F.R., Ismail, A.A. & Sedman, J. (1995). A rapid automated method for the determination of *cis* and *trans* content of fats and oils by Fourier transform infrared spectroscopy. *Journal of the American Oil Chemists' Society*, 72, 873-880.

Whittaker, A.D., Park, B., Thane, B.R., Miller, R.K. & Savell, J.W. (1992). Principles of ultrasound and measurement of intramuscular fat. *Journal of Animal Science*, 70, 942-952.

Williams, P. & Norris, K. (1987). *Near Infrared Technology in the Agricultural and Food Industries*. American Association of Cereal Chemists, Minnesota, USA.

Wold, S., Esbensen, K. & Geladi, P. (1987). Principal components analysis. *Chemometrics and Intelligent Laboratory Systems*, 2, 37-52.

PAPER I

Noninvasive Assay for Cyanogenic Constituents in Plants by Raman Spectroscopy: Content and Distribution of Amygdalin in Bitter Almond (*Prunus amygdalus*)

E. Micklander, L. Brimer and S.B. Engelsen

Noninvasive Assay for Cyanogenic Constituents in Plants by Raman Spectroscopy: Content and Distribution of Amygdalin in Bitter Almond (*Prunus amygdalus*)

E. MICKLANDER,* L. BRIMER, and S. B. ENGELSEN

The Royal Veterinary and Agricultural University, Department of Dairy and Food Science, Food Technology, Chemometrics Group, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark (E.M., S.B.E.); and The Royal Veterinary and Agricultural University, Department of Pharmacology and Pathobiology, Section of Pharmacology and Toxicology, Bülowsvej 13, DK-1870 Frederiksberg C, Denmark (L.B.)

The ability of Fourier transform Raman spectroscopy to measure cyanogenic glycoside amygdalin was investigated using a standard addition series in which amygdalin (0–700 nmol/mg) was added to ground sweet almond. As an additional test, the method so developed, and further refined, was used to determine endogenous amygdalin in bitter almonds. Using partial least-squares regression (PLSR), the best correlation between the spectra of the standard addition series and amygdalin content exhibited a value of 0.999, while the root mean square error of cross validation (RMSECV) was found to be 4 nmol/mg using one principal component. The model was based on the first derivative of the aromatic C–H stretching band at 3060 cm^{-1}. A less accurate but still excellent PLSR model could be developed on the vibrational band of the nitrile group, specific for the cyanogenic compound in the system. The best PLSR model obtained for endogenous amygdalin in 17 raw bitter almonds yielded a RMSECV = 13 nmol/mg and r = 0.937. In addition, the distribution of amygdalin in bitter almond was investigated *in situ* by obtaining spectra of an intact cross-section of the cotyledon of bitter almond using Raman microscopy. While amygdalin was found not to be present in measurable amounts in the center of bitter almond, the concentration increased towards the epidermis, but not linearly.

Index Headings: **Raman spectroscopy; Chemometrics; iPLS; Amygdalin; Bitter almonds; Cyanogenic glycosides.**

INTRODUCTION

Cyanogenic glycosides are glycosidic compounds with a nitrile group that upon hydrolysis is liberated as the toxic compound hydrogen cyanide. The ability to accumulate cyanogenic glycosides is widespread among plants, as the compounds work as a defense system against pests and fungi. Cyanogenic glycosides are accumulated in many food and feed plants.[1,2] Some of these are important food crops like cassava[3] and smaller tonnage crops such as sorghum[4] and bitter almonds.[5] Ripe bitter almonds almost exclusively contain amygdalin (Fig. 1) as a cyanogenic constituent, while unripe fruits contain a mixture of amygdalin and the corresponding mono-glucoside prunasin.[6] Contents of around 290 mg HCN per 100 g of tissue (= 107 nmol amygdalin/mg of tissue) have been reported for certain investigated samples.[7]

The liberation of hydrogen cyanide may be a problem in relation to human and animal consumption of the plants. From a food safety point of view it is therefore desirable to know the level of cyanogenic compounds in the plants and their products. Considerable efforts have been devoted to the development of methods for determination of cyanogenic compounds, as reviewed by Brimer.[8] Indirect estimation of the content of cyanogenic compounds is generally accomplished by enzymatic or chemical hydrolysis of the cyanogenic substances to cyanide, which is subsequently measured. The most widely used methods are based on color reactions such as the Guignard[9] reaction and other reactions.[10–12] Several useful methods for field measurements based on color reactions have been developed[13,14] and several biosensors have been tested.[15–17]

Laboratory methods that measure the genuine cyanogenic compounds directly are necessary for research within the field of cyanogenesis. Accordingly, technologies like thin layer chromatography, high pressure liquid chromatography, gas chromatography/mass spectroscopy, and micellar capillary electrophoresis have been applied for the separation and analysis of the cyanogenic glycosides.[18–22]

A major disadvantage of most of the non-chromatographic methods described is that the cyanogenic compounds are measured indirectly and thus do not allow for the identification and quantification of the specific compound responsible for cyanide release. For all methods, including those that separate cyanide prior to the detection, destruction (i.e., comminution and extraction) is part of the procedure.

The present work represents a spectroscopic/chemometric investigation of the qualitative and quantitative performance of FT-Raman spectroscopy for prediction of amygdalin in ripe bitter almonds. The advantage of the method is that it can measure directly *in situ* the cyanogenic compounds present on or near the sample surface without any sample pretreatment; that is, it may be used in a relatively nondestructive and rapid manner.

Nitrile groups give rise to relatively low intensity bands in infrared (IR) spectroscopy; however, they are strongly Raman active. The nitrile vibrational band is found in a part of the infrared and Raman spectrum, which is free from interference from other chemical information. For this reason the nitrile vibrational band is most attractive in relation to measurements of amygdalin, cyanogenic glycosides, and cyanogenic compounds in general, as the nitrile group is seldom found in other natural compounds,[23] rendering the signal nearly specific. In amygdalin and other cyanogenic glycosides with an ar-

Received 18 December 2001; accepted 22 April 2002.
* Author to whom correspondence should be sent.

0003-7028/02/5609-1139$2.00/0

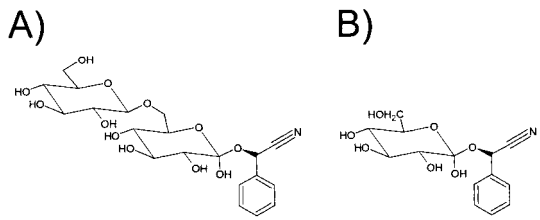


FIG. 1. Molecular structure of amygdalin (D-mandelonitrile-D-gentiobioside) and prunasin (D-mandelonitrile-D-glucoside).

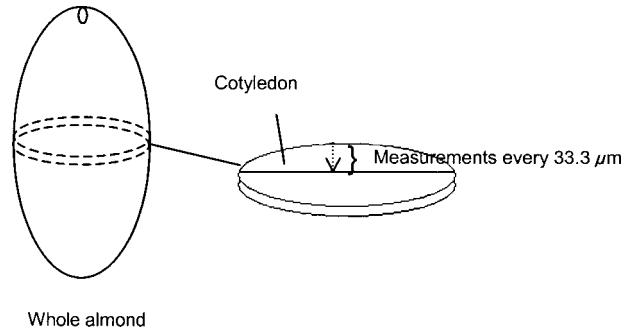


FIG. 2. Description of Raman microscopy measurements. A slice of bitter almond approximately 2 mm thick was created by transverse partition of a bitter almond (in which the hydrolyzing enzymes had been heat inactivated). Measurements were performed approximately every 33.3 μm in a straight line across one of the cotyledons—first measurement at the epidermis, last measurement at the center of the almond.

omatic aglycone, the aromatic ring in the molecule is also Raman active. However, the signal from this moiety in general renders little specificity, as aromatic compounds are ubiquitous in nature.

A general problem with Raman spectroscopy is the inherently poor signal-to-noise ratio, as the method is a scatter technique based on the relatively weak Stokes bands around the strong Rayleigh line. In addition, fluorescence signals from the samples will in many cases hide some of the Raman information. However, with the use of multivariate data analysis such as principal component analysis (PCA) and partial least-squares regression (PLSR),[24,25] the relevant information can be extracted efficiently from the noisy spectra.

MATERIALS AND METHODS

Materials and Chemicals. Amygdalin, prunasin, and β-glucosidase from almond (G-8625) were purchased from Sigma Chemical Co. (St. Louis, MO). All other chemicals were pro-analysis quality from Merck (Duren, Germany). Thin layer chromatography sheets were from Machery-Nagel (Duren, Germany). Sweet and bitter almonds were provided by Toms Fabrikker A/S (Ballerup, Denmark).

Sample Preparation. A standard addition series of amygdalin in an almond matrix was prepared by adding amygdalin to coarsely milled sweet almond in which the hydrolyzing enzymes were inactivated prior to milling by immersion of the almonds in boiling water for 2 h, as described by Tuncel et al.[26] with the modification that the almonds had been vacuum-packed in polyethylene before the heat treatment. The absence of any remaining hydrolytic activity was checked on special test samples, as described by Tuncel et al.[26] The almond skin (the epidermis (outer skin) and the endosperm (thin layer of storage tissue)) was removed prior to milling by means of a scalpel. A laboratory mill (type A 10, Ika Labortechnik, Staufen, Germany) was used for milling. Samples with amygdalin contents of approximately 0, 58, 117, 175, 350, 525, and 700 nmol/mg were prepared by adding amygdalin to 0.400 g of ground sweet almond. The mixture was finely pulverized and blended in an agate mortar. Two samples of each concentration were prepared independently for double determinations.

The samples of real bitter almonds were also heated and ground prior to Raman measurements in order to make the Raman spectra comparable with the measurements on the standard samples (which had to be ground in order to homogeneously incorporate the amygdalin). Grinding the samples also made it possible to mix the substance from several almonds. Since the amygdalin content *a priori* is assumed to be different in the center than close to the surface, a larger span in amygdalin content could be obtained by making samples of the outer or inner part of several almonds.

Raman Measurements. Fourier transform Raman data were collected on a Perkin Elmer system 2000 Interferometer equipped with a Nd:YAG laser emitting at 1064 nm and an InGaAs detector. Spectra of pure amygdalin and pure prunasin were acquired in the spectral area 3600–200 cm^{-1} with an interval of 1 cm^{-1} between measuring points, and 256 scans were accumulated with a resolution of 4 cm^{-1} using a laser power of 400 mW. Spectra of the standard addition series and samples of bitter almond were acquired in the spectral area 3600–200 cm^{-1} with an interval of 1 cm^{-1} between measuring points, and 64 scans were accumulated with a resolution of 8 cm^{-1} using a laser power of 1000 mW. A resolution of 8 cm^{-1} was used because a higher resolution did not affect the spectra of these complex, condensed phase samples. This was underlined by the fact that the smallest half bandwidth found in spectra obtained at a resolution of 4 cm^{-1} was 10 cm^{-1}. A plane sample surface was obtained by filling a holder to the brim with sample substance and lightly tamping the sample with a stamper. All measurements were carried out as true replicates with independent samplings. A 180° scattering arrangement was used and no correction for the spectral response of the instrument was applied.

Raman Microscopy. Raman microscopy measurements across a bitter almond cotyledon, i.e., one of the two resting and premature leaves of the seed embryo (Fig. 2), were performed on a Labram Infinity dual-laser spectrograph (Jobin-Yvon, Lille, France) equipped with a Peltier cooled charge-coupled device (CCD) detector. The measurements were carried out using a 785 nm diode laser, a 100× objective, and a 600 g/mm grating, resulting in a spectral resolution of approximately 13 cm^{-1}. The spectra were acquired using two accumulations with an integration time of 60 s. Measurements were performed approximately every 33 μm along a line of the transversely sectioned slice of bitter almond (Fig. 2). The first measurement was 0–10 μm from the epidermis and measurements were performed continuously in a straight line towards the other edge of the cotyledon at the center of the almond. The experiment was carried out on two cot-

yledons from two different bitter almonds. The hydrolyzing enzymes in the bitter almonds had been inactivated prior to measurements as described above.

Reference Methods for the Determination of Amygdalin Content. Immediately after the Raman measurements each sample was divided into two fractions and brought to a low pH by suspension in 2.0 mL 0.1 M H_3PO_4 to denature endogenous β-glycosidase and thus avoid any decomposition of the amygdalin. The suspension was homogenized using an Ultra-turrax (Type TP 18/10) for approximately 1 min at 20 000 rotations/min. The Ultra-turrax was rinsed with 2.0 mL of 0.1 M H_3PO_4, which was added to the homogenate. After shaking, 1.5 mL of the homogenate was transferred to an eppendorf tube and centrifuged. The total cyanogenic potential (in this case, amygdalin) in the supernatant was subsequently measured by two different colorimetric methods: a microdiffusion method using solid phase detection[27] (later referred to as ''solid method'') and a wet colorimetric method (later referred to as ''wet method''). For the latter, 5 μL of the supernatant was removed and mixed with 150 μL of 100 mM Mes (buffer) adjusted to pH 6.5 with 6 M NaOH. Five microliters of this mixture was mixed with 200 μL Mes (buffer) pH 6.5 and 0.1 mg emulsin. The diluted amygdalin extracts with emulsin added were shaken for 1 h at 30 °C and the HCN released was collected in 40 μL of 6 M NaOH by aeration with nitrogen. The collected cyanide was measured as described by Halkier and Møller.[28]

Chemometric Methods. Principal component analysis (PCA) and partial least-squares regression (PLSR) are powerful tools for extraction of important variances in a data matrix consisting of many variables.[24,25] In PCA a data matrix is decomposed by consecutive orthogonal extraction of the largest variation (principal components, PCs) in data until the variation left is unsystematic (noise). The PCA model can be written as: $\mathbf{X} = \mathbf{TP}' +$ Noise, where **P** is the loading matrix and **T** is the score matrix. The loading vectors can be considered as pure hidden profiles that are common to all the measured Raman spectra. The scores are the amounts of the hidden profiles in the individual Raman spectra.

The purpose of a PLSR is to build a linear model enabling prediction of a desired characteristic (Y) from a measured spectrum (X). During the regression, X is decomposed as in PCA, but the PCs are found as the underlying structures that co-vary best with the Y variable. The performance of a model is evaluated on the basis of the correlation between measured y values and predicted y values and the root mean square error of cross validation[25] (RMSECV), which is the estimation of the error on the predicted values. Only cross-validated results are reported in this study.

The PLSR algorithm is, however, sensitive to noisy variables when applied to large data sets. The iPLS algorithm[29] is designed to develop local PLSR models on equidistant subintervals of the full-spectrum region. Through a continuous reselection of intervals, the optimal interval(s) of variables is encircled and the best possible model is developed. The advantage of the iPLS approach, in addition to the pedagogical aspects, is that smaller intervals will contain fewer interferences and thus result in simpler, more precise, and more easily interpretable models.

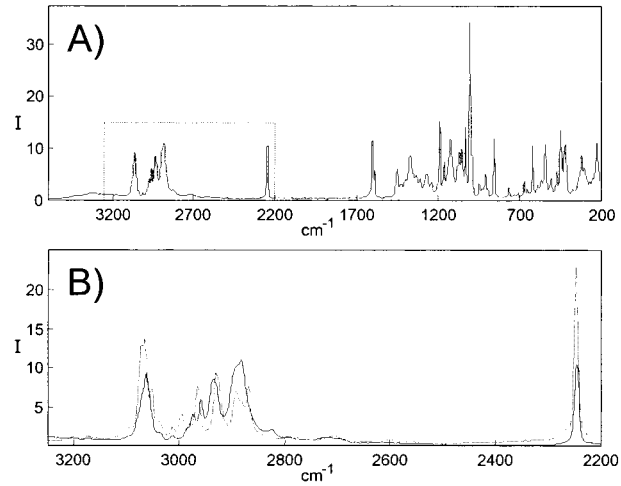


FIG. 3. (**A**) Raman spectrum of pure crystalline amygdalin acquired in the area 3600–200 cm^{-1} with an interval of 1 cm^{-1} between measuring points. The area 3250–2200 cm^{-1} is marked by a dotted line and enhanced in (**B**), where Raman spectra of crystalline amygdalin (-) and prunasin (–) are compared in the area 3250–2200 cm^{-1}.

Programs. Chemometric calculations were performed using Unscrambler ver. 7.6 (CAMO ASA, Trondheim, Norway) and Matlab ver. 5.3 (The MathWorks, Inc., Natick, MA) installed with the PLS Toolbox ver. 1.5.3b (Wise & Gallagher; Eigenvector Research, Manson, WA) and The command line *i*PLS-Toolbox, ver. 2.0 (L. Nørgaard; The Royal Veterinary and Agricultural University, Frederiksberg, Denmark).

RESULTS AND DISCUSSION

Amygdalin and Raman Spectroscopy. Amygdalin consists of a gentiobiose (β-D-GLC*p*(1→6)-β-D-GLC*p*) carbohydrate moiety, a nitrile group, and a benzene ring. The structure of the molecule is shown in Fig. 1 together with the structure of the corresponding mono-glucoside prunasin.

Nitrile groups give rise to Raman bands near 2200 cm^{-1}, and in the spectrum of pure crystalline amygdalin the C≡N stretching is seen as a sharp, intense band at 2244 cm^{-1} (Fig. 3A). The vibrations of the benzene ring also give rise to some very sharp and intense bands in different areas of the spectrum. Most distinct are the characteristic aromatic C–H stretching band at approximately 3060 cm^{-1} and the double band at approximately 1600 cm^{-1} produced by aromatic C–C stretching. The band at 1000 cm^{-1} caused by the in-plane deformation of the aromatic ring is very sharp and intense, but less distinct than the nitrile band, since other bands interfere. Less intensive and less specific and distinct bands from the aromatic ring are also seen in the fingerprint region from 1225 to 950 cm^{-1} and below 900 cm^{-1} where out-of-plane C–H bending vibrations of the aromatic ring produce bands. The pyranosidic sugar rings also exhibit strong bands in the fingerprint region, and intense bands due to CH stretching are observed in the 2800–3000 cm^{-1} region.

An important factor with regard to qualitative measurements is that the spectra of the cyanogenic com-

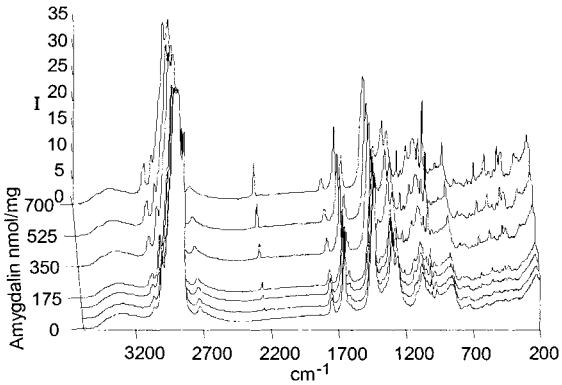


FIG. 4. Raman spectra of seven samples of ground sweet almond with amygdalin added in concentrations of 0, 58, 117, 175, 350, 525, and 700 nmol/mg, respectively. Y axis is amygdalin concentration.

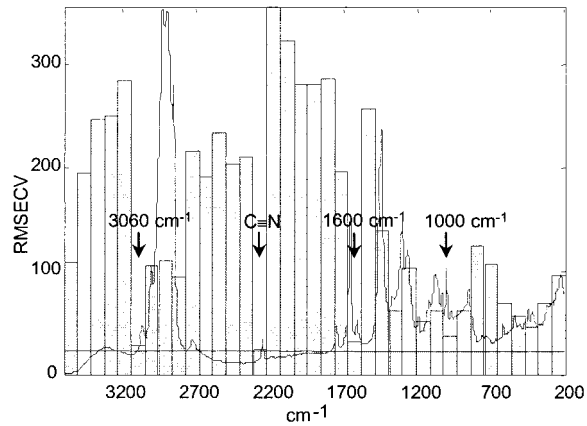


FIG. 5. Cross-validated prediction performance (RMSECV) for 37 PLSR interval models (bars) and for full-spectrum model (line) plotted together with the mean spectrum of seven standard samples (amygdalin added to ground sweet almond in concentrations of 0, 58, 117, 175, 350, 525, and 700 nmol/mg). The full-spectrum model (line) was made with 2 PCs; the sub-interval models (bars) were made using 1 PC.

pounds are distinct. In Fig. 3B the Raman spectra of amygdalin and the corresponding mono-glucoside prunasin are compared in the area 3250–2200 cm^{-1}. The spectra are dissimilar, since some vibration bands are only present in the spectrum of one of the compounds and other bands are shifted; the nitrile band is, however, only slightly shifted (2 cm^{-1}). PLSR models based on the whole spectrum would accordingly be able to distinguish between the contribution to the cyanogenic potential from amygdalin and prunasin, respectively. However, it is unlikely that models specific to the individual cyanogenic compounds can be developed based on the nitrile band alone.

In Fig. 4 the Raman spectra of the seven samples of ground sweet almond with amygdalin added in concentrations of 0 to 700 nmol/mg are compared. The nitrile band is only slightly affected by the surrounding almond matrix; the band is shifted to 2242 cm^{-1}. The intensity of the nitrile band and some of the vibration bands of the aromatic and glucose rings clearly increase with increasing amygdalin concentration; thus, a relationship between Raman spectra and the amygdalin content is observed. The nitrile band is absent in the sample without amygdalin, which confirms that amygdalin is the only measurable compound in the system that contains a nitrile group. Bands corresponding to vibrations of the aromatic ring are also observed in the spectrum of pure sweet almond, indicating that there are other aromatic compounds present in the almond tissue, such as aromatic peptides. The intensity of the aliphatic C–H stretching bands just below 3000 cm^{-1}, corresponding in large part to fat, is much higher in the almond samples (Fig. 4) than in the pure amygdalin (Fig. 3A) since the almonds are rich in oil. This is unambiguously confirmed by the presence of the triacylglyceride carbonyl band centered at 1745 cm^{-1}.

Calibration of PLSR Model of Amygdalin Content in a Standard Addition Series. A PLSR model on the raw Raman spectra of the standard addition series and the nominal amygdalin content yielded a RMSECV of 22–24 nmol/mg using 2 PCs. The fact that 2 PCs are needed to explain the variation of one chemical component (amygdalin) indicates that there is some variation in the data that interferes with the signal correlated to amygdalin content. This could be an offset, introduced by varying levels of fluorescence in the samples. This offset can be removed by differentiation of the data, and, accordingly, a PLSR model based on first-derivative transformed data was calculated. This procedure resulted in a RMSECV that was approximately the same as the model on the raw data, but using only 1 PC.

The prediction error for the full-spectrum PLSR model of 22–24 nmol/mg is relatively high compared with the span of amygdalin concentration (0–700 nmol/mg). This may indicate that the PLSR algorithm has problems extracting the relevant variables in the Raman spectra because of a poor signal-to-noise relationship and the fact that the majority of the 3600 measuring points do not carry information about the amygdalin content. Thus, interval PLS (iPLS) was carried out on raw and transformed spectra. Figure 5 shows an iPLS plot of the raw Raman spectra of the seven levels of amygdalin in a matrix of sweet almond. The horizontal line represents the performance of a two-PC full-spectrum PLSR model (RMSECV = 23.1 nmol/mg) and the vertical bars indicate the performance of each of the 37 one-PC sub-interval PLS models generated.

The vertical bars corresponding to the areas around the bands at 3060 (CH stretching of the aromatic ring) and 2242 cm^{-1} (C≡N stretching of the nitrile group) are close to the horizontal line representing the RMSECV of the full-spectrum model. Thus, the simpler one-PC PLSR models based on relatively narrow spectral intervals yield a quantitative performance equal to the two-PC full-spectrum PLSR model. Upon optimization of the local PLS models by reselection of intervals, the models based on these bands and other local models corresponding to the vibration bands of the aromatic ring prove to be superior to the full-spectrum model, since a lower RMSECV is obtained and fewer PCs are used in the models. In Table I the performance of the PLSR on the full spectrum (3600–200 cm^{-1}) and the best iPLS models based on intervals of the spectrum are compared.

The bands at 3060 cm^{-1}, corresponding to a CH stretching of the aromatic ring ($\nu_1(A_{1g})$), and 853 cm^{-1} correlate best to the amygdalin content. The band at 853 cm^{-1} is not due to α-anomeric carbohydrates (α-glycans), as it is present in amygdalin. In contrast to β-gentiobiose,

TABLE I. Model performance of PLS regressions on Raman spectra of ground sweet almond with amygdalin added in concentrations of 0, 58, 117, 175, 350, 525, and 700 nmol/mg. MSC is multiplicative scatter correction, RMSECV is root mean square of cross validation, and *r* is correlation.

Peak (cm^{-1})	Variables (cm^{-1})	Pretreatment	PC	*r*	RMSECV (nmol/mg)
Whole spectrum	200–3600	raw	2	0.989	26
Whole spectrum	200–3600	1. derivative	1	0.996	23
2242	2299–2257	1. derivative	1	0.999	10
1000	991–986	2. derivative	1	0.999	8
618	627–625	2. derivative	1	0.999	9
Several	2955–2860	2. derivative	1	0.999	6
3060	3084–3012	raw	1	0.999	5
853	852–774	MSC	1	0.999	4
3060	3081–3056	1. derivative	1	0.999	4

prunasin, and mandelonitril, the spectrum of amygdalin exhibits this band, which we hesitate to assign, but apparently the good quantitative performance of the band benefits from the low background in this area. The performance of the models on the two spectral regions (RMSECV = 4 nmol/mg, $r = 0.999$) underlines that the Raman spectrum is highly capable of measuring the amygdalin content. A PLSR model on the untransformed raw data of the band at 3060 cm^{-1} is almost equally as good (RMSECV = 5 nmol/mg, $r = 0.999$), also using one PC, which is in agreement with the fact that this part of the Raman spectrum is little influenced by fluorescence and differentiation is not strictly required to remove spectral artifacts. The signal-to-noise ratio is consequently higher in this spectral area than in the remainder of the spectrum, which may be the reason why a model is better on the vibration band of the aromatic ring than on the band of the nitrile group. Furthermore, it is more stable than the nitrile band, which exhibits a weak shift with increasing amygdalin content. The iPLSR model on the nitrile band specific for amygdalin[6] is, however, more than acceptable, with RMSECV = 10 nmol/mg and $r = 0.999$. The relatively good result obtained on the first derivative of the nitrile band is thus very promising for the development of similar methods for amygdalin and other cyanogens in other matrices, such as other crops and processed products.

Prediction of Amygdalin Content in Bitter Almonds. Above, a PLSR model was made on a standard addition series of amygdalin; the following describes the study on endogenous amygdalin. Table II shows that the amygdalin content found with the solid method and the wet method are consistent and the average content is approximately the same as was found in earlier investigations.[7] The fact that we attain approximately the same results using the two methods originally developed for two different plant materials (raw cassava root and green sorghum shoots), reassures us that no method-specific errors are introduced in the data. In the further calculations only the results from the wet method will be presented, as the precision of this method is highest.

To investigate the Raman method in real quasi-intact almond matrices with natural levels of amygdalin, PLSR models were developed on Raman spectra of bitter almonds and their amygdalin content. Four standard samples with amygdalin concentrations of 0–175 nmol/mg were included in the models in order to stretch the variation in amygdalin content and thereby emphasize variation due to amygdalin in the data and in the models.

TABLE II. Overview of amygdalin content in 18 samples of ground bitter almonds. Measurements performed by two different methods, solid phase detection (solid) and colorimetric determination in solution (wet), are compared.

Sample number	Solid method (nmol/mg)	Wet method (nmol/mg)
1	87	79
2	107	103
3	150	138
4	131	133
5	80	90
6	90	93
7	154	149
8	94	110
9	139	127
10	89	100
11	105	109
12	128	135
13	113	118
14	...	64
15	147	136
16	114	112
17	59	66
18	90	91
Average	110	109
Std. Dev.	27	24

Different mathematical transformations of data were applied in order to reduce the effect of the different matrices surrounding the amygdalin in the standard samples and the samples of bitter almonds. Multiplicative scatter correction[30] (MSC) and standard normal variation[31] (SNV) were applied, as well as first and second differentiation using Savitzky–Golay smoothing and differentiation (averaging over 11 points). MSC is performed to minimize packing and particle size effects; however, in further studies investigations concerning the influence of particle size and packing on the measurements should be performed.

Results of the best PLSR models on Raman spectra of 17 samples of bitter almonds and four standard samples are seen in Table III. It was not possible to make a PLSR model of the amygdalin content based on the whole Raman spectrum, as the residual validation variance increases on the first PC no matter how the data are pretreated. Accordingly, iPLS was performed on raw and pretreated data. As expected, the models on the transformed data were superior to models on the raw data. The models on this mixed data set are not as simple as the model on the standard addition series in which one PC was used to

TABLE III. Results of PLS regressions on Raman spectra of 18 samples of ground bitter almonds and 4 samples of ground sweet almond with amygdalin added in concentrations of 0, 58, 117, and 175 nmol/mg. SNV is standard normal variation, RMSECV is root mean square of cross validation, and *r* is correlation.

Peak (cm^{-1})	Variables (cm^{-1})	Pretreatment	PC	*r*	RMSECV (nmol/mg)
2242	2242–2196	SNV	2	0.871	18
Several	1206–1122	SNV	3	0.877	17
Several	465–297	SNV	3	0.893	16
3060	3079–3031	SNV	2	0.907	15
1000	1138–997	SNV	3	0.916	15
618	617–590	1. derivative	2	0.937	13

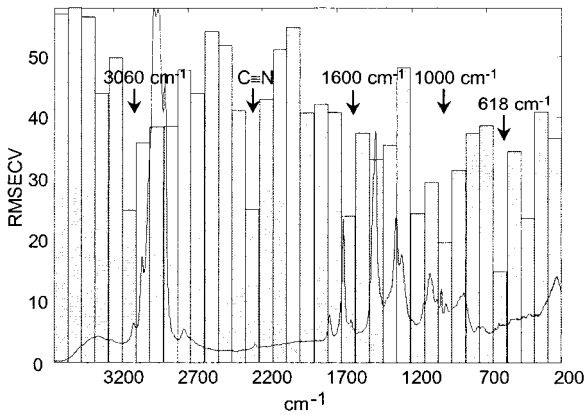


FIG. 6. Cross-validated prediction performance (RMSECV) for 37 PLSR interval models (bars) performed on first-derivative spectra plotted together with the raw mean spectrum of 21 samples (17 bitter almonds and 4 standard samples). The sub-interval models were made using 2 PCs.

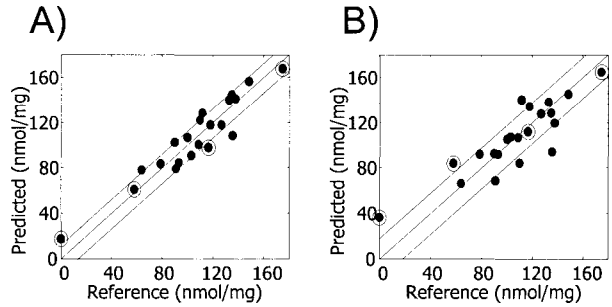


FIG. 7. Predicted vs. measured plots for the iPLS models on the spectral areas (**A**) 617–590 cm^{-1} (first-derivative transformed) and (**B**) 2242–2196 cm^{-1} (Standard Normal Variate transformed) of the Raman spectra of 4 standard samples (◎) (0–175 nmol/mg amygdalin in ground sweet almond) and 17 samples of ground bitter almonds (●) using 2 PLS components. The vertical lines illustrate the target line and the target line ± RMSECV.

explain the amygdalin content. Two to three PCs are used in the iPLS models since the bitter almonds and the standard samples (based on a matrix of sweet almond) obviously have different chemical compositions. Furthermore, the compartmentalization of amygdalin differs between the two kinds of samples. Pretreatment does not remove the variation due to chemical differences, and accordingly, more PCs are needed to explain the variation present in the data.

The best model is obtained on first-derivative transformed data of the spectral area 617–590 cm^{-1}, which corresponds to the right slope of a vibration band of the aromatic ring (ν_{6b}) peaking at 618 cm^{-1}. An iPLS plot illustrating the superior performance of this spectral area is shown in Fig. 6 (the RMSECV of a full-spectrum model is not shown in the plot, as a full-spectrum model could not be developed). The PLSR model results in a RMSECV of 13 nmol/g and a correlation of 0.937. The performance of the model is almost equaled by models based on the two other vibration bands of the aromatic ring, namely, the more characteristic bands at 3060 and 1000 cm^{-1}, having RMSECVs of 15 nmol/g.

Again, a fairly good model is obtained on the nitrile band at 2242 cm^{-1} with a RMSECV of 18 nmol/g and a correlation of 0.871. This is promising with regard to measurements of cyanogens generally using Raman spectroscopy. Predicted amygdalin content against measured values from PLSR models based on the spectral area 616–589 cm^{-1} and based on the nitrile band are shown in Figs. 7A and 7B.

A prerequisite for prediction with a PLSR model is that the samples used for calibration and prediction contain the same variations. This is not the case for the standard addition series and the bitter almonds, and it is therefore not possible to predict the amygdalin content in the bitter almonds based on the models developed on standard samples. The predicted values using, for example, the model on the 3060 cm^{-1} vibration band of raw data were in the interval 72–213 nmol/mg, which is different from the measured values (64–149 nmol/mg) shown in Table II. This indicates that amygdalin mixed into a matrix of sweet almond used for the standard addition series is not a sufficiently good simulation of ''endogenous'' amygdalin in bitter almond tissue. Another reason why the model calibrated on the standard samples cannot be applied for prediction of amygdalin content in the bitter almonds could be that the concentrations of the standard samples are in the range 0–700 nmol/mg, which exceeds the range of the bitter almonds (59–154 nmol/mg) excessively.

Determination of Amygdalin Distribution in Bitter Almonds by Raman Microscopy. The distribution of amygdalin in bitter almond was investigated by performing PCAs on the Raman spectra obtained across two cotyledons from bitter almond. The PCAs were carried out on the nitrile band, as the largest variation in this vibration band is due to variation in amygdalin content, as described above. A PCA on the spectral area from 616 to 589 cm^{-1} would also illustrate the distribution of amygdalin, since this area has the best correlation with amygdalin content in bitter almonds according to the iPLS models mentioned above. This part of the spectrum was, however, not covered with the settings used on the Raman microscope, which were optimized to the nitrile band. The peak of the nitrile band was observed at 2246 cm^{-1} when using the Raman microscope.

For both cotyledons, measurements closest to the epidermis and the center of the almond had to be left out of the calculations because of an immense offset of the spectra in some cases resulting in intensities that were above the detection limit. The offset is partly due to fluorescence and partly due to scatter introduced by the angle of the edge.

Figures 8A and 8B show the scores and loadings from a PCA on the nitrile band of second-derivative-transformed spectra of 81 measurements across a cotyledon for which averaging has been performed on 6 measurements (average over approximately 200 μm), resulting in 14 spectra. Figures 8C and 8D illustrate the scores and loadings from a PCA on similar data from another cotyledon. The loadings (Figs. 8A and 8C) show which variations in the variables are described by each PC. The first loading resembles the nitrile band; accordingly, the first PC describes the variation corresponding to amygdalin. The second and third loadings explain minor variations in the spectra (2–4%), and it seems that the remaining variation is primarily noise and therefore not of relevance to the variation in amygdalin content. The

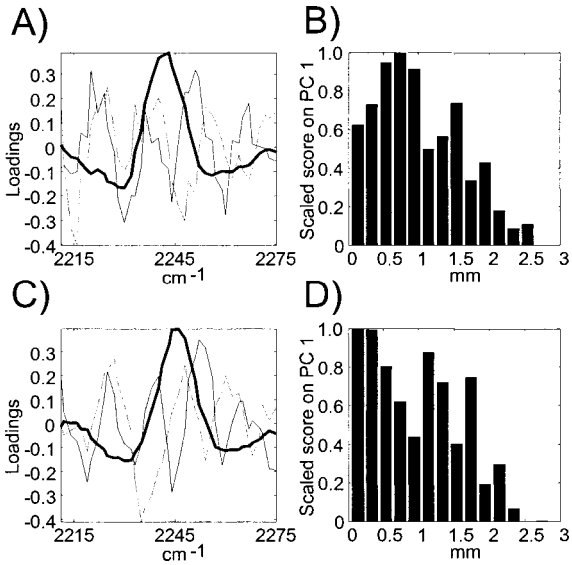


FIG. 8. (**A**) Loadings and (**B**) scores from PCA on the nitrile band (2275–2211 cm^{-1}) of 14 Raman spectra obtained by measuring 81 times across a slice of a cotyledon of a bitter almond using Raman microscopy and averaging on 6 measurements (≈ 200 μm). (**C**) and (**D**) are corresponding plots for a cotyledon from a second bitter almond. In the loading plots (**A**) and (**C**), the curves are first loading (_), second loading (-), and third loading (–). For better interpretation scores have been elevated to positive values by subtraction of the lowest (negative) score value from all scores and subsequently scaled by dividing all values by the value of the maximum content. In (**B**) and (**D**), epidermis is to the left and the center of the seed is to the right.

scores of the first PC (adjusted to be only positive values and scaled to values between 0 and 1; Figs. 8B and 8D) are consequently an index of the amygdalin content.

The score plots show that amygdalin is not present in measurable amounts in the center of the almonds. Amygdalin is present approximately 0.5 mm from the center, albeit in small quantities. From this point the content increases towards the epidermis, but not monotonously. In one of the almonds, the maximum content seems to be 0.5–1.0 mm from the epidermis, whereas maximum is reached at the epidermis in the other almond. The difference between the two samples may be attributed to many factors. First of all, there is a large natural variance between almonds. Secondly, the amygdalin content could vary in various directions within a bitter almond. In cassava root tubers there are both radial and longitudinal gradients in the content of cyanogenic glycosides. The radial gradient is very steep, while the longitudinal gradient is shallower.[10,32] The two slices of cotyledons were prepared from different parts of the almonds (varying distances from the chalaza, the darker spot in the broader end of the seed), which could explain the difference in the amygdalin gradient. Furthermore, the differences could depend on the exact point at which the Raman measurements were performed. A PCA on the 81 measurements individually indicates that the amygdalin content can be five times higher in one point than in a neighboring point of measurement; thus, the variation at the microscopic level seems to be high.

Using tissue printing methods, Poulton and Li[33] found that cyanoglycosides in mature seeds from plum (*Prunus domestica,* var. Stanley) and black cherry (*Prunus serotina* Ehrh.) were localized in the cotyledonary parenchyma (the major, and non-specialized, part of the tissue) and were absent from the thin endosperm and the so-called procambial strands (strands of resting embryogenic cells). The results of the current investigations on *Prunus amygdalus* add to this that a declining gradient in amygdalin concentration seems to exist from seed surface towards the center, or more precisely, from the abaxial and to the adaxial surface of each of the cotyledons. This obviously must be further investigated, combining the use of Raman and light microscopy in order to obtain an even more precise understanding of the tissue distribution of the cyanogenic principle. The advantage of the more sophisticated method used in the current work is that measurements are performed directly on the amygdalin, avoiding the error of an indirect method. Furthermore, the method is quantitative rather than only qualitative as is the tissue printing method, which only indicates whether amygdalin is present or not present. Additionally, measurements are performed with high spatial resolution so that even local fluctuations in amygdalin content may be detected along a concentration gradient.

CONCLUSION

This preliminary study has demonstrated that Raman spectroscopy is a potential rapid method for noninvasive determination of cyanogenic compounds. Initial partial least-squares models on the spectral bands corresponding to vibration of the aromatic ring show that the Raman spectra are highly correlated to amygdalin content. It is most promising that acceptable multivariate calibration models can be made on the C≡N-stretching band, as this band is specific for cyanogenic potential and it is most likely possible to develop similar models for other cyanogens in other seeds and maybe even other tissues and organs from cyanogenic crops. Hence, Raman spectroscopy can be used for noninvasive measurements of other naturally occurring cyanogenic compounds, provided that the fluorescence background of the matrix is limited.

The present study emphasizes that iPLS is an attractive method for providing an overview of interesting spectral areas that can be selected for optimization of predictive ability. The iPLS models in this work represent a case study in the selection of robust group frequencies related to amygdalin. Moreover, the iPLS models on Raman subspectra of almonds contain fewer interferences and thus result in simpler, more precise models that perform better than full-spectrum models.

ACKNOWLEDGMENTS

The authors thank Professor Birger Lindberg Møller, Department of Plant Biology, The Royal Veterinary and Agricultural University, for performing amygdalin determinations using the wet method; the quality laboratory at Toms Fabrikker A/S, Ballerup, Denmark, for providing sweet and bitter almonds; and LMC and SJVF for financing the Raman equipment.

1. D. A. Jones, Phytochemistry **47,** 155 (1998).
2. L. Brimer, "Chemical Hazards and Their Control: Endogenous Compounds", in *Fermentation and Food Safety,* M. R. Adams and M. J. R. Nout, Eds. (Aspen Publishers, Gaithersburg, 2001), Chap. 4, p. 71.
3. M. Bokanga, A. J. A. Essers, N. Poulter, H. Rosling, and O. Tewem, Acta Hortic. **375,** 1 (1994).
4. J. L. Wheeler, Acta Hortic. **375,** 251 (1994).

5. R. Socias i Company, Acta Hortic. **470,** 57 (1998).
6. M. Frehner, M. Scalet, and E. E. Conn, Plant. Physiol. **94,** 28 (1990).
7. E. E. Conn, ''Cyanogenic Glycosides'', in *International Review of Biochemistry, Biochemistry of Nutrition IA, Vol 27,* A. Neuberger and T. H. Jukes, Eds. (University Park Press, Baltimore, 1979), Chap. 2, p. 21.
8. L. Brimer, Cyanogenic Glycosides—Occurrence, Analysis and Removal from Food and Feed. Comparison to Other Classes of Toxic and Antinutritional Glycosides (Department of Pharmacology and Pathobiology, Laboratory of Toxicology, Royal Veterinary and Agricultural University, Copenhagen, 2000), pp. 58–83.
9. M. L. Guignard, C. R. Hebd. Seances Acad. Sci. **142,** 545 (1906).
10. R. D. Cooke, J. Sci. Food Agric. **29,** 345 (1978).
11. S. A. J. A. Essers, M. Bosneld, R. M. van der Grift, and A. G. J. Voragen, J. Sci. Food Agric. **63,** 287 (1993).
12. T. Tatsuma, K. Komori, H. H. Yeoh, and N. Oyama, Anal. Chim. Acta **408,** 233 (2000).
13. L. Brimer and P. Mølgaard, Biochem. Syst. Ecol. **14,** 97 (1986).
14. H. H. Yeoh and S. V. Egan, Food Chem. **60,** 119 (1997).
15. H. H. Yeoh and V. D. Truong, Food Chem. **47,** 295 (1993).
16. T. Tatsuma, K. Tani, T. Ogawa, and N. Oyama, Sens. Actuators, B **49,** 268 (1998).
17. V. Volotovsky and N. Kim, Electroanalysis **10,** 512 (1998).
18. L. Brimer, S. B. Christensen, P. Mølgaard, and F. Nartey, J. Agric. Food Chem. **31,** 789 (1983).
19. L. Dalgaard and L. Brimer, J. Chromatogr. **303,** 67 (1984).
20. R. K. Chadha, J. F. Lawrence, and W. M. N. Ratnayake, Food Addit. Contam. **12,** 527 (1991).
21. D. Chassagne, J. C. Crouzet, C. L. Bayonove, and R. L. Baumes, J. Agri. Food Chem. **44,** 3817 (1996).
22. C. Campa, Ph. Schmitt-Kopplin, T. R. I. Cataldi, S. A. Bufo, D. Freitag, and A. Kettrup, J. Chromatogr., B **739,** 95 (2000).
23. M. Lechtenberg and A. Nahrstedt, ''Cyanogenic Glycosides'', in *Naturally Occurring Glycosides,* R. Ikan, Ed. (John Wiley and Sons, Chichester, 1999), Chap. 5, p. 147.
24. S. Wold, K. Esbensen, and P. Geladi, Chemom. Intell. Lab. Syst. **2,** 37 (1987).
25. H. Martens and T. Næs, *Multivariate Calibration* (John Wiley and Sons, New York, 1993), 2nd ed., Chap. 3, p. 116.
26. G. Tuncel, M. J. R. Nout, and L. Brimer, Food Chem. **63,** 65 (1998).
27. J. D. K. Saka, A. R. K. Mhone, J. Mkambira, L. Brimer, M. Bokanga, N. M. Mahungu, L. Chiwona-Karltun, and H. Rosling, Trop. Agric. **75,** 169 (1998).
28. A. B. Halkier and B. L. Møller, Plant Physiol. **90,** 1552 (1989).
29. L. Nørgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, and S. B. Engelsen, Appl. Spectrosc. **54,** 413 (2000).
30. P. Geladi, D. McDougall, and H. Martens, Appl. Spectrosc. **39,** 491 (1985).
31. R. J. Bjarnes, M. S. Dhanoa, and S. J. Lister, Appl. Spectrosc. **43,** 772 (1989).
32. M. Kojima, N. Iwatsuki, E. S. Data, C. D. V. Villegras, and I. Uritani, Plant Physiol. **72,** 186 (1983).
33. J. E. Poulton and G. P. Li, Plant Physiol. **104,** 29 (1994).

PAPER II

NMR-cooking: Monitoring the Changes in Meat During Cooking by Low-field H^1-NMR

E. Micklander, B. Peshlov, P. Purslow and S.B. Engelsen

ELSEVIER

Trends in Food Science & Technology 13 (2002) 341–346

TRENDS IN FOOD SCIENCE & TECHNOLOGY

Viewpoint

NMR-cooking: monitoring the changes in meat during cooking by low-field ^{1}H-NMR

Elisabeth Micklander[†], Boyan Peshlov[†], Peter P. Purslow[‡] and Søren B. Engelsen[†,*]

[†]Centre for Advanced Food Studies, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark (tel.: +45-35283205; fax: +45-35283245; e-mail: se@kvl.dk)

[‡]Centre for Extracellular Matrix Biology, Department of Biological Sciences, University of Stirling, Stirling FK9 4LA, Scotland, UK

This report illustrates how pulsed low-field nuclear magnetic resonance (LF-NMR) can be used to investigate changes in food during processing, as the technique can non-invasively monitor the behaviour of the water in a product. LF-NMR data obtained during cooking of meat were utilised to identify important temperatures at which major changes in meat structure occur. It was found that a new water population is developed in meat as a result of cooking and a dual hypothesis is presented: The new water component can be either water expelled from the meat matrix or alternatively it can be water trapped in a myosin gel.

Introduction

Conversion of living muscle into 'ready to eat' meat is one of the most longstanding and important food processes known to mankind. In the meat industry, much attention has been directed towards the understanding and optimisation of the quality of meat as a raw material.

Rapid instrumental techniques have been employed for at/on-line screening of meat properties that can be used to improve control and classification of the product in automated processes. One instrumental technique that has proven particularly interesting when applied to foods is low-field nuclear magnetic resonance (LF-NMR) (Berendsen, 1992). There is a major division of NMR techniques into high resolution NMR which probes the chemistry of the sample, and low-field (wide line) NMR which probes the basic physics of the sample, typically the states of water in the sample in ^{1}H NMR.

LF-NMR has been implemented in the investigation of the distribution of water in muscle and has proven to be a powerful tool for identifying water components. The results have been related principally to water holding capacity (Bertram, Karlsson, Rasmussen, Pedersen, Dønstrup, & Andersen, 2001; Brøndum, Munck, Henckel, Karlsson, Tornberg, & Engelsen, 2000b) and sensory properties (Brøndum, Byrne, Bak, Bertelsen, & Engelsen, 2000a; Fjelkner-Modig & Tornberg, 1986). A few investigations have been carried out on the effect of cooking upon the states of water in meat using LF-NMR (Borisova & Oreshkin, 1992; Fjelkner-Modig & Tornberg, 1986; Tornberg, Andersson, Göransson, & von Seth, 1993), but not with continuous or dynamic measurements.

Dynamic LF-NMR measurements have proven to provide valuable non-invasive information about complex food items undergoing processing or storage. Examples of applications to food systems are Rheo-NMR, i.e. performing rheological profiling of complex fluids (Britton & Callaghan, 1997; Callaghan, 1999) and NMR-baking, i.e. monitoring the changes in the states of water during the dough-to-bread baking process (Engelsen, Jensen, Pedersen, Nørgaard, & Munck, 2001). A natural extension would be NMR-cooking of meat to study the transitions in the states of water occurring at different temperatures in the process of cooking meat by cooking inside the NMR magnet.

* Corresponding author.

0924-2244/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0924-2244(02)00163-2

NMR-cooking

Measurements were performed on meat samples of 6 g (three replicates) cut out from a pig *longissimus dorsi* muscle. NMR measurements were performed every 2°C while increasing the temperature linearly from 16 to 100°C over 4 h. Each meat-cooking experiment resulted in 43 relaxation curves containing 4000 data points.

The decaying magnetisation curve of meat changes noticeably during cooking. It is readily observed from the relaxation curves obtained throughout the process that the overall proton relaxation rate increases with temperature (Fig. 1). This is (at least partly) due to the increased mobility of protons with increasing temperature. The figure also reveals that the total NMR signal towards higher temperatures is reduced due to the less favourable Boltzmann distribution between the two spin-states resulting in a decrease in the signal to noise relationship with increasing temperature.

In a first attempt to obtain a global view of the main variation in the multivariate meat-cooking relaxation data, principal component analysis (PCA) (Wold, Esbensen, & Geladi, 1987) was applied. Although the relaxation data as a function of cooking temperature is not a true bilinear data structure, PCA analysis of LF-NMR data has previously proven to be an extraordinarily robust and valuable data enhancement and data reduction method (Bechmann, Pedersen, Nørgaard, & Engelsen, 1999). In Fig. 2 each point represents a full relaxation decay curve with 4000 data points without significant loss of information.

The 3D PCA score plot (Fig. 2) shows that the variation contained in the prime principal component, PC1, is mainly related to the temperature increase, whereas the variations in the data contained in PC2 and PC3 reveal several transition temperatures at which major changes in the states of water in meat take place. Sharp bends in the score plot, corresponding to large changes in the properties of the water, can be identified at 46 and 66°C, while softer bends are found at approximately 42, 57 and 76°C. These transition temperatures identified by the PCA can be explained by changes in the constituents of meat during cooking; namely, denaturation and shrinkage of the myofibrillar proteins and the connective tissue, which affects the water-holding capacity/distribution of water in meat. The lowest transition temperature found in the PCA score plot (42°C) corresponds to the onset of myosin denaturation (Davey & Gilbert, 1974; Martens, Stabursvik, & Martens, 1982). The denaturation of sarcoplasmic and myofibrillar proteins begins at lower temperatures, but the rate and amount of denaturation increases significantly above this temperature (Davey & Gilbert, 1974; Tornberg, Andersson, & Josell, 1997). According to Tornberg *et al.* (1997) and to Bendall and Restall (1983), a slow transverse shrinkage of the myofibrils starts at this temperature, whereas Palka and Daun (1999) found that the transverse shrinkage starts at 45°C, which is where one of the major changes are seen in the PCA score plot. The transition of the states of water at 57°C

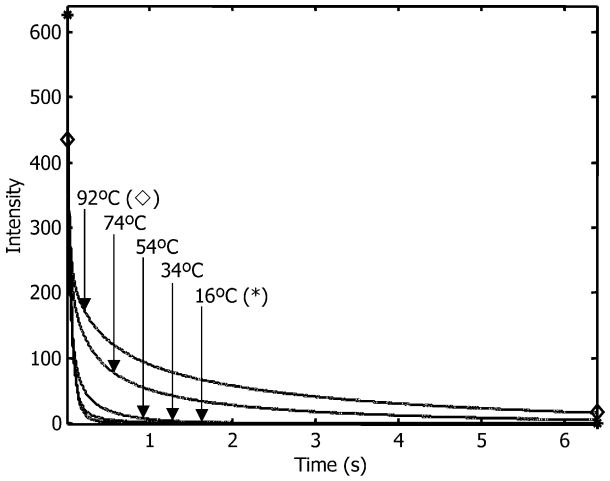


Fig. 1. Five LF-NMR relaxation curves of meat obtained during NMR-cooking. The temperature at the centre of the sample at the time of measurement is indicated. Data were obtained during cooking in a Maran Benchtop Pulsed NMR analyzer (Resonance Instruments, Witney, UK), operating at 23.2 MHz and equipped with an 18 mm variable temperature probe head. A Carr–Purcell–Meiboom–Gill sequence (CPMG) with the following settings was used: number of scans = 4, relaxation delay = 20 s, dwell time = 0.5 s, $\tau = 400$ μs, number of echoes collected = 8000. Only the even-numbered echoes are used, resulting in 4000 data points.

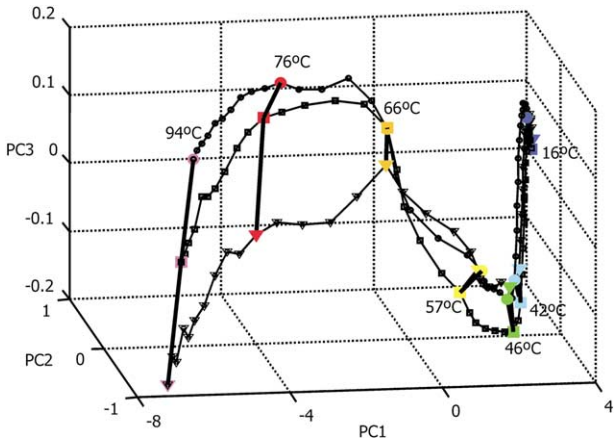


Fig. 2. 3D PCA score plot of three NMR-cooking experiments, each represented by 43 measurements obtained over a period of 4 h. PC1, PC2 and PC3 explain 97, 2.3 and 0.6% of the variance in the NMR data. The approximate centre temperatures are indicated at the beginning and the end of the experiment and where major changes in the scores are seen.

can be explained by the fact that shrinkage parallel to the myofibrils, according to some authors, starts at this temperature (Palka & Daun, 1999), but the onset of collagen denaturation beginning somewhere between 50 and 60°C could also contribute to this transition (Bendall & Restall, 1983; Martens *et al.*, 1982). The longitudinal contraction intensifies (Tornberg *et al.*, 1997) or, according to Bendall and Restall (1983), really starts at 64–65°C, which is where a major change is found in the NMR data as shown by the PCA analysis. This may be related to a major peak in thermal denaturation of the endomysial, perimysial and epimysial collagen around these temperatures (Mohr & Bendall, 1969). The slight transition seen at around 76°C may reflect further gelatinisation of collagen, which occurs up to approximately 80°C (Palka & Daun, 1999), but more likely is due to the onset of actin denaturation, which has been reported to be around 80°C by some authors (Cheng & Parrish, 1979). This straightforward interpretation of the transitions in the PCA results on the basis of denaturation of major components of meat goes a long way to confirm that the information contained in LF-NMR data obtained during cooking of meat is indeed related to important changes of the meat matrix due to heating.

Complementary information about the transitions in the states of water during the cooking process can be extracted by fitting discrete exponentials to the relaxation curves. It is generally accepted that water in meat is contained within inhomogenous compartments whose diffusional exchange is slow on the NMR time scale, giving rise to different observed transverse relaxation times of the water in different compartments. The water molecules in each compartment are made up of structural, multilayer and bulk water, but if the exchange is fast on the NMR time scale between the three possible water phases, they will give rise to only one average T_2 component, which is the most likely situation in protein–water systems. The probability of slowly relaxing bulk water molecules encountering a membrane or insoluble protein structure at the surface of the compartment or pore (whereupon relaxation occurs rapidly) effectively means that multicomponent T_2 relaxation curves give information on the distribution of pore sizes in the inhomogeneous structures.

By fitting a number of monoexponentials to the CPMG decay curves in Fig. 1, the characteristics (i.e. rate of relaxation and quantity) of the water populations present in meat at different stages of cooking can be estimated. One model that can be used is discrete exponential fitting, where the NMR relaxation curves ($Rmag(t)$) are fitted by an equation that comprises a few (N) exponential components, each characterized by a different relaxation constant ($T_{2,i}$) and corresponding intensity ($M(T_{2,i})$):

$$\left\| Rmag(t) - \sum_{1}^{N} M(T_{2,i}) \times \exp\left(\frac{-t}{T_{2,i}}\right) \right\| = 0$$

This equation is normally solved using non-linear curve-fitting algorithms (e.g. Pedersen, Bro, & Engelsen, 2001) such as SIMPLEX (Nelder & Mead, 1965). In Fig. 3a and 3b the relaxation constants and intensities are plotted as a function of the applied temperature. At 16°C, two water populations are present, in agreement with observations in other studies (Bertram *et al.*, 2001; Brøndum *et al.*, 2000b; Fjelkner-Modig & Tornberg,, 1986; Hazlewood, Nichols & Chamberlain, 1969) where it has generally been found that the major fraction of the water (approximately 80–95%) has a T_2 in the range of 35–60 ms (T_{21}) and the minor fraction has a T_2 in the range of 100–150 ms (T_{22}). The most recent studies designate T_{21} water as intramyofibrillar water, which is situated within highly organised protein structures such as the I-band area of the myofibrils, and T_{22} as intermyofibrillar water (Bertram *et al.*, 2001). Some studies (Belton, Jackson, & Packer, 1972; Bertram *et al.*, 2001; Borisova & Dreshkin, 1992) also mention a small, faster relaxing (1–10 ms) water population which is ascribed to water closely associated with macromolecules. As judged by the decrease in the least squares residuals, our data do not support a third relaxation component at 16°C.

The least squares residual (the amount of the NMR signal that cannot be explained) when fitting with two components increases at temperatures above 40°C, indicating that another water population emerges; however, a third component cannot be fitted safely before 59°C has been reached. The result of bi-exponential fitting in the temperature interval 16–57°C, followed by tri-exponential fitting above 59°C, is shown in Fig. 3a and 3b. The concentration of the T_{21} water seems to decrease as soon as the heating begins, but the decrease accelerates around 40°C, where the amount of T_{22} water concomitantly increases. Since the amount of T_{21} decreases and the amount of the T_{22} water component increases from 40°C, it is assumed that the initial stages of the denaturation of myosin which begin around 42°C causes small amounts of water to be squeezed out of the T_{21} domain (within the myofibrils) to the T_{22} domain (inter myofibrillar space) due to reduction of the myofibrillar lattice spacing. The expulsion of water is, however, enhanced when the major lateral contraction of the myofibrils occurs around 46°C. This contraction causes a decrease in myofibrillar space, i.e. less bulk water is present within the myofibrils, resulting in a decrease of the average relaxation time T_{21} (from 42 ms at 25°C to 28 ms at 60°C), whereas the T_{22} increases due to more bulk water. Around 60°C, transverse contraction of the myofibrils produces a new water component with a relaxation time of

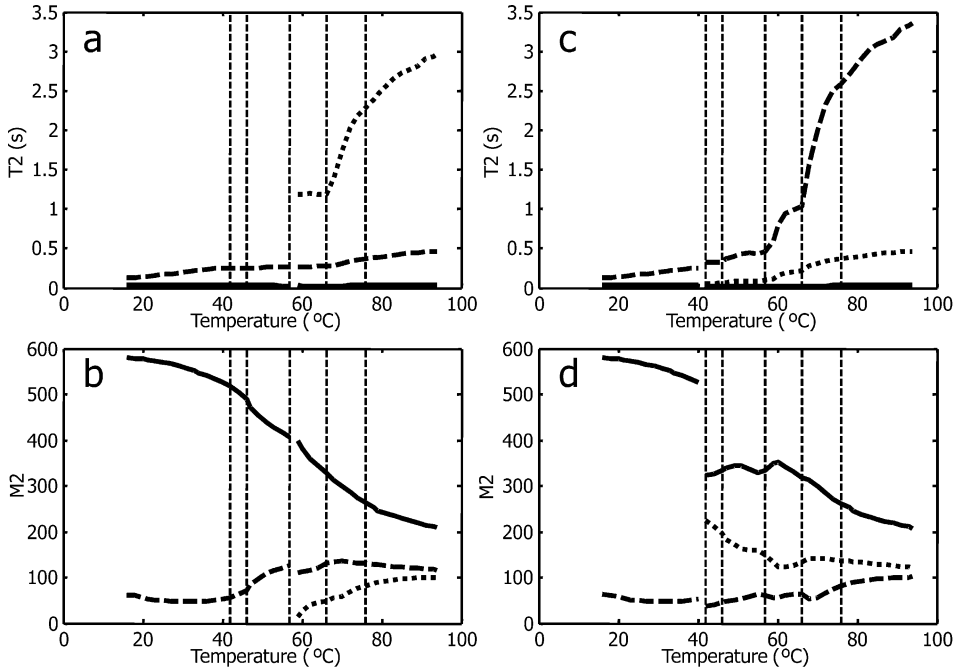


Fig. 3. Evolution in the CPMG transverse time constants (T_{2i}) and amplitudes (M_{2i}) during NMR-cooking estimated by discrete exponential fitting: (a) and (b) bi-exponential fit 16–40°C, tri-exponential fit 42–94°C, solid: T_{21}, M_{21}, dashed: T_{22}, M_{22}, dotted T_{23}, M_{23}; (c) and (d) bi-exponential fit 16–57°C, tri-exponential fit 59–94°C, solid: T_{21}, M_{21}, dashed: T_{22}, M_{22}, dotted T_{2x}, M_{2x}. The curves represent averages of three experiments.

approximately 1 s, which is faster than the relaxation time of totally free water, but slow compared to the water bound by the meat matrix. Thus, the T_{23} water may be assigned to water expelled from the meat matrix mixed with a high concentration of proteins. Some water might be expelled at lower temperatures (as indicated by the increase in residuals from 40°C), but it is not until substantial transverse contraction takes place and water is expelled in significant amounts that this fraction can be modelled using the discrete fitting approach. The increased transverse contraction of the myofibrils is reflected in the T_{23} time and the amount of the T_{23} population which increases further from around 66°C.

It is most interesting to note that if tri-exponential fitting is applied from 40°C, a different interpretation of the data can be deduced (Fig. 3c and 3d). In this case, it would appear that the first major change identified by the PCA on the CPMG data at 42°C is related to the development of a new water population with an intermediate relaxation time, T_{2x}. Thus, it is reasonable to assume that the new component T_{2x} originates from inside the fine architecture of the myofibrils and ends up in a relatively free environment, yet still (on the NMR time scale) being isolated from the slowly relaxing T_{22} component representing intermyofibrillar water and extracellular water. As T_{2x} emerges at temperatures where myosin begins to denature, it is possible that T_{2x} water represents water emerging in small compartments developed inside the myofibrils between denatured

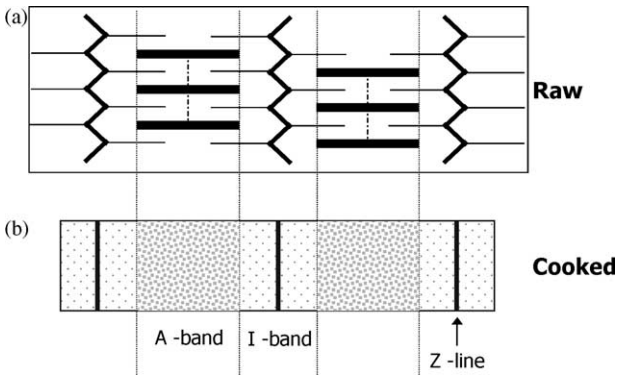


Fig. 4. Schematic diagram of the structure of (a) a raw myofibril (b) a cooked myofibril. There are approximately 800–1000 myofibrils packed laterally together inside one muscle fibre. In raw meat the thick filaments (principally myosin) interdigitate with the thin filaments (principally actin) which are attached to the Z-lines marking the boundaries of the sarcomere. Protein density in the A-band, where thin filaments overlap the thick filaments, is higher than the I-band which contains only thin filaments and the central Z-line. On heating, banding and Z-lines can still be seen, but the filamentous proteins denature and form gel-like structures. Protein density in the A-band gel region is likely to remain higher than in the I-band of heated myofibrils.

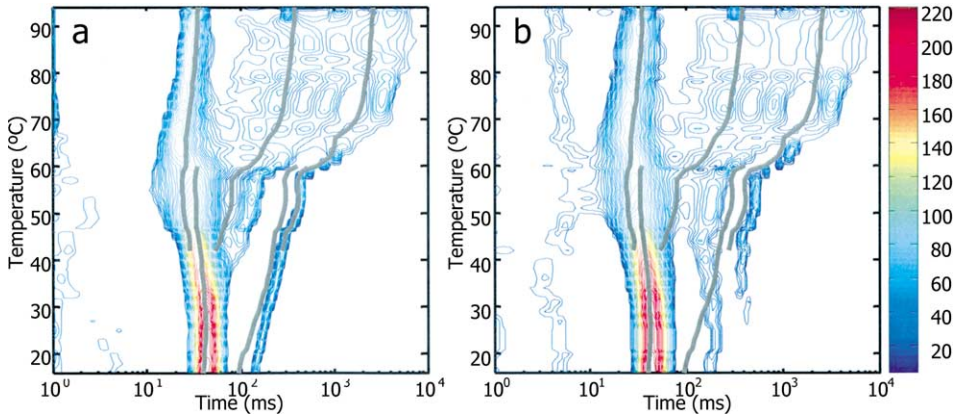


Fig. 5. Contour plots of distributed exponential fitting of CPMG data from NMR-cooking of meat (16–94°C): (a) magnitude-transformed data, (b) phase-rotated data. Discrete exponential fitting is superimposed on the plots, bi- and tri-exponential fitting overlap in the temperature interval 40–60°C.

myosin units where it is trapped in a gel-like structure (Fig. 4). The labile nature of the discrete fitting approach is underlined by the concentration plot. When two T_2 components are as close in time as T_{21} and T_{2x}, the stability of the algorithm deteriorates and it becomes numerically insignificant whether the amplitude is modelled on either M_{21} or M_{2x}.

A more physically realistic analytical approach of the data from this basic low-resolution technique is to assume a continuous distribution, where the water population will be composed of water molecules with a distribution of time constants rather than a single well-defined T_2 constant (or a small series of discrete constants). On the other hand, the numerical problem and its solution in this approach is even more ill-defined and non-unique. The debate concerning which model is most appropriate will not be discussed in this report. In distributed exponential fitting an 'infinite' number of exponentials are used to fit the relaxation curves. Rather than finding a few components N in eqn (1), $M(T_{2,i})$ is a distribution of intensities for each T_2 component, sampled logarithmically in the interval $T_{2,\min}-T_{2,\max}$ (as calculation cannot in practice take place on an infinite number of time constants). In this case the problem is numerically solved by constrained matrix inversion techniques.

Figure 5 displays the contours of the magnitude of distributed exponentials of the NMR transverse relaxation of meat as a function of cooking temperature for data pre-processed by magnitude transformation (5a) and phase rotation (5b), of which the former is inferior. The magnitude of the transformed data is retrieved as $\mathbf{x}_{\text{magnitude}} = \sqrt{a^2 + b^2}$ where a and b are the signals from the real and the imaginary channel of the NMR instrument. Phase rotated data are obtained by Principal Phase Correction (Pedersen *et al.*, 2002), which is based on singular-value decomposition of a matrix containing the signals from the real and imaginary channels. When superimposing the result of discrete fitting, it is surprising, but gratifying to observe that the distributed fitting of magnitude-transformed data agrees very well with the result of the discrete fitting. This strongly implies that the result is independent of the analysis method, and therefore is unlikely to be an artefactual product of a particular analysis. The branching of T_{2x} from the major population T_{21} is unambiguous and supports the hypothesis that a myosin gel (Fig. 4) can be detected by the NMR measurements as it is formed. Distributed fitting of phase-rotated data (Fig. 5a), however, supports the hypothesis that the new water population introduced in the meat system due to cooking is simply 'free' water outside the meat matrix with a high protein content. In this analytical approach the T_{23} population emerges as early as around 30°C, with a relaxation time in the region of 380 ms.

In addition to the NMR cooking of pure meat, NMR measurements have been performed during cooking of meat in water, which is a normal way of cooking meat, but it naturally dilutes the information from protons inside the meat. From these experiments it can be concluded that the major changes in meat which take place as a result of cooking seem to be independent of cooking medium, as the same transition temperatures are found when analysing the NMR data for pure meat and meat in added water. Three components are present throughout the cooking process in added water, which supports the hypothesis that the third component seen during cooking of plain meat is water expelled not only from the myofibrils, but also from the meat matrix.

Conclusion

LF-NMR measurements can serve as a window into a food matrix, as changes can be observed in a non-invasive manner. In this study, NMR measurements have been used to monitor the changes in meat during cooking, with special emphasis on the changes in the distribution of water. Using multivariate data analysis on the relaxation curves, the temperatures where major changes occur in meat are readily identified as 42, 46, 57, 66 and 76°C. Exponential fitting of the NMR curves reveal that the heat denaturation of some major meat proteins lead to the formation of a new water population above approximately 40°C. What is observed is probably expulsion of water from the myofibrillar lattice, but it could also be the formation of a porous myosin gel. Interpretation of hard modelling exponential deconvolution techniques is ambiguous and relatively low-informative, whereas multivariate soft modelling in the form of PCA appears much more informative and sensitive to the changes occurring during thermal processing of meat, albeit interpretation of the results requires external knowledge of the structures potentially involved.

References

Bechmann, I. E., Pedersen, H. T., Nørgaard, L., & Engelsen, S. B. (1999). Comparative chemometric analysis of transverse low-field ^{1}H NMR relaxation data. In P. S. Belton, B. P. Hills, & G. A. Webb (Eds.), *Advances in magnetic resonance in food science* (pp. 217–225). Cambridge: The Royal Society of Chemistry.

Belton, P. S., Jackson, R. R., & Packer, K. J. (1972). Pulsed NMR studies of water in striated muscle. I. Transverse nuclear spin relaxation times and freezing effects. *Biochimica et Biophysica Acta*, *286*, 16–25.

Bendall, J. R., & Restall, D. J. (1983). The cooking of single myofibres, small myofibre bundles and muscle strips from beef *M. psoas* and *M. Sternomandibularis* muscles at varying heating rates and temperatures. *Meat Science*, *8*, 93–117.

Berendsen, H. J. C. (1992). Rationale for using NMR to study water relations in foods and biological tissues. *Trends in Food Science and Technology*, *3*, 202–205.

Bertram, H. C., Karlsson, A. H., Rasmussen, M., Pedersen, O. D., Dønstrup, S. D., & Andersen, H. J. (2001). Origin of multi-exponential T_2 relaxation in muscle myowater. *Journal of Agricultural and Food Chemistry*, *49*, 3092–3100.

Borisova, M. A., & Oreshkin, E. F. (1992). On the water condition in pork meat. *Meat Science*, *31*, 257–265.

Britton, M. M., & Callaghan, P. T. (1997). NMR microscopy and the non-linear rheology of food materials. *Magnetic Resonance in Chemistry*, *35*, 37–46.

Brøndum, J., Byrne, D. V., Bak, L. S., Bertelsen, G., & Engelsen, S. B. (2000a). Warmed-over flavour in porcine meat—a combined spectroscopic, sensory and chemometric study. *Meat Science*, *54*, 83–95.

Brøndum, J., Munck, L., Henckel, P., Karlsson, A. H., Tornberg, E., & Engelsen, S. B. (2000b). Prediction of water-holding capacity and composition of porcine meat with comparative spectroscopy. *Meat Science*, *55*, 177–185.

Callaghan, P. T. (1999). Rheo-NMR: nuclear magnetic resonance and the rheology of complex fluids. *Reports on Progress in Physics*, *62*, 599–670.

Cheng, C. S., & Parrish, F. C. (1979). Heat-induced changes in myofibrillar proteins of bovine longissimus muscle. *Journal of Food Science*, *44*, 22–24.

Davey, C. L., & Gilbert, K. V. (1974). Temperature-dependent cooking toughness in beef. *Journal of the Science of Food and Agriculture*, *25*, 931–938.

Engelsen, S. B., Jensen, M. K., Pedersen, H. T., Nørgaard, L., & Munck, L. (2001). NMR-baking and multivariate prediction of instrumental texture parameters in bread. *Journal of Cereal Science*, *33*, 59–69.

Fjelkner-Modig, S., & Tornberg, E. (1986). Water distribution in porcine *M. longissimus dorsi* in relation to sensory properties. *Meat Science*, *17*, 213–231.

Hazlewood, C. F., Nichols, B. L., & Chamberlain, N. F. (1969). Evidence for the existence of a minimum of two phases of ordered water in skeletal muscle. *Nature*, *222*, 747–750.

Martens, H., Stabursvik, E., & Martens, M. (1982). Texture and color changes in meat during cooking related to thermal denaturation of muscle proteins. *Journal of Texture Studies*, *13*, 291–309.

Mohr, V., & Bendall, J. R. (1969). Constitution and physical chemical properties of intramuscular connective tissue. *Nature*, *223*, 404–405.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.

Palka, K., & Daun, H. (1999). Changes in texture, cooking losses, and myofibrillar structure of bovine M-semitendinosus during heating. *Meat Science*, *51*, 237–243.

Pedersen, H. T., Bro, R., & Engelsen, S. B. (2001). SLICING—a novel approach for unique deconvolution of NMR relaxation decays. In G. A. Webb, P. S. Belton, A. M. Gill, & I. Delgadillo (Eds.), *Magnetic resonance in food science: a view to the future* (pp. 202–209). Cambridge: The Royal Society of Chemistry.

Pedersen, H. T., Bro, R., & Engelsen, S. B. (2002). Towards rapid and unique curve resolution of low-field NMR relaxation data. Trilinear SLICING versus two-dimensional curve fitting. *Journal of Magnetic Resonance*, *157*, 141–155.

Tornberg, E., Andersson, A., Göransson, Å., & von Seth. G. (1993). Water and fat distribution in pork in relation to sensory properties. In E. Poulanne & D.I. Demeyer (Eds.), *Pork quality: genetic and metabolic factors* (pp. 239–258). CAB International

Tornberg, E., Andersson, K., & Josell Å. (1997). The rheological properties of whole and minced meat during cooking as related to sensory and structural characteristics. In *Proceedings 1st International Symposium on Food Rheology and Structure*, Zürich, 16–20 March, 1996

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis—a tutorial. *Chemometrics and Intelligent Laboratory Systems*, *2*, 37–52.

PAPER III

Vibrational Microspectroscopy of Food.
Raman vs. FT-IR

L.G. Thygesen, M.M. Løkke, E. Micklander and S.B. Engelsen

ELSEVIER

Trends in Food Science & Technology 14 (2003) 50–57

TRENDS IN FOOD SCIENCE &TECHNOLOGY

Viewpoint

Vibrational microspectroscopy of food. Raman vs. FT-IR

Lisbeth G. Thygesen, Mette Marie Løkke†, Elisabeth Micklander and Søren B. Engelsen*

Centre for Advanced Food Studies, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark (tel: +45-35-2832-05; fax: +45-35-2832-45; e-mail: se@kvl.dk)

FT-IR and Raman spectroscopy are complementary techniques for the study of molecular vibrations and structure. The combination with a microscope results in an analytical method that allows spatially resolved investigation of the chemical composition of heterogeneous foods and food ingredients. The high spatial resolution makes it possible to study areas down to approximately 10×10 μm with FT-IR microspectroscopy and approximately 1×1 μm with Raman microspectroscopy. This presentation highlights the advantages and disadvantages of the two microspectroscopic techniques when applied to different heterogeneous food systems. FT-IR and Raman microspectroscopy were applied to a number of different problems related to food analysis: (1) *in situ* determination of starch and pectin in the potato cell, (2) *in situ* determination of the distribution of amygdalin in bitter almonds, (3) the composition of blisters found on the surface of bread, (4) the microstructure of high-lysine barley and (5) the composition of white spots in the shell of frozen shrimps.

* Corresponding author.
† Present affiliation: Arla Foods amba.

Introduction

Infrared (IR) and Raman spectroscopy are complementary techniques that provide information on molecular structure. By combining spectroscopy with microscopy molecular information can be obtained with great spatial resolution at the microscopic level. Samples of microscopic size can be analysed directly, in air, at ambient temperature and pressure, wet or dry, and in many cases without destroying the sample.

Both qualitative and quantitative information can be obtained using microspectroscopy. A number of organic compounds and functional groups can be identified by their unique pattern of absorption, and the intensity of the absorption may be used for the calculation of the relative concentration in the sampled entity (Wetzel & LeVine, 1999).

FT-IR and Raman microspectroscopy may be combined with at least three different mapping techniques: point, line and area. With point acquisition several spectra are measured from different places in a sample selected after visual inspection through the microscope, i.e. the spectra are not systematically related to each other spatially. Line mapping defines a series of spectra to be obtained along one dimension (a line) and can be used to investigate changes in a chemical component along a certain direction, i.e. a profile. An area map uses two dimensions, providing a spectroscopic image that can be directly compared to the corresponding visual image, but with an entire spectrum in each pixel instead of a simple colour. By using a series of partly overlapping acquisition areas it is possible to obtain a smoothing of a profile or a map. Using a standard microscope equipped with an XY stage it may take up to several hours to map a few square millimetres, even for an FT-based instrument. However, new instruments that record an entire row (i.e. a line scan) simultaneously by use of a CCD array will make area mapping much less time consuming.

Vibrational microspectroscopy offers many possibilities for the study of biological material, including foods systems. In this presentation we demonstrate the potential of the method by relaying examples from our laboratory where IR or Raman microspectroscopy has been successfully applied to potatoes, bitter almonds, bread, barley kernels and shrimp shells to obtain information about microstructure and chemical composition.

0924-2244/03/$ - see front matter
PII: S0924-2244(02)00243-1

Raman or infrared microspectroscopy for food studies

Both Raman and infrared microspectroscopy may reveal useful information about food samples. In infrared spectroscopy the sample is radiated with infrared light. Different chemical bonds absorb at different infrared wavelengths depending on the atoms connected, the surrounding molecules, and the type of vibration the absorbance gives rise to (for example stretching or bending). In Raman spectroscopy, the sample is radiated with monochromatic visible or near infrared light from a laser. This brings the vibrational energy levels in the molecule to a short-lived, high-energy collision state, which returns to a lower energy state by emission of a photon. Normally the photon has a lower frequency than the laser light (Stokes Raman scattering), and the difference in frequency (given in reciprocal centimetres) between the frequency of the laser and that of the scattered photon is called the Raman shift. The Raman shift corresponds to the frequency of the fundamental IR absorbance band of the bond.

Even though both methods probe molecular vibrations and structure, they do not provide exactly the same information. While IR spectroscopy detects vibrations during which the electrical dipole moment changes, Raman spectroscopy is based on the detection of vibrations during which the electrical polarisability changes (Pistorius, 1995). As a rule of thumb, this implies that bonds that connect two identical (or nearly identical) parts of a molecule tend to be more active in Raman than in IR spectroscopy. For example, the C=C bond is generally more intense in Raman than in IR spectra. Characteristic absorbance bands found in food systems are shown in Table 1 for both IR and Raman. The table shows that all four major food ingredients (water, fat, protein and carbohydrates) absorb both in Raman and IR, but that the intensities of the absorbance bands vary. An important difference to note is that the O–H stretching vibration is very strong in IR, but very weak in Raman, because OH bonds are only weakly polarisable. For this reason, water is practically "invisible" in Raman spectroscopy, while it dominates the IR spectrum, if present. The other three major ingredients are visible with both types of spectroscopy, albeit their Raman and IR spectra are not identical, because the relative intensities differ for the bands making up the spectrum of each ingredient. Unless a sample has too high a water content to be effectively studied using IR, other factors will determine which of the two types of spectroscopy best solves a given problem. Table 2 lists characteristics of the two techniques that are important to consider when analysing food samples. Raman spectroscopy has the potential of a better spatial resolution due to the lower wavelengths used, and furthermore offers confocality, i.e. it is possible to focus on different planes below the sample surface. It is, for example, possible to focus beneath a quartz plate or through a food packaging material to obtain pure spectra of food samples without exposing the food to the atmosphere. On the other hand, the signal-to-noise ratio is much lower, and if the sample fluoresces, measurements may even be impossible. For example, measurements on samples that contain one or more of the three fluorescent amino acids (tyrosine, tryptophan and phenylalanine) or chlorophyll may prove difficult or, in practice, impossible to study by Raman spectroscopy. However, the problem may be overcome if a Raman instrument is equipped with a near infrared laser instead of a laser in the visible range, though the spatial resolution would be poorer. Another problem related to Raman spectroscopy is heating of the sample (West, 1996). The heat generated by the laser may alter or even destroy the sample during measurement. In some cases, the best setting of the Raman instrument is therefore a compromise between destructive heating, which calls for short sampling times, and a high signal to noise ratio, which calls for long sampling times and/or repeated samplings.

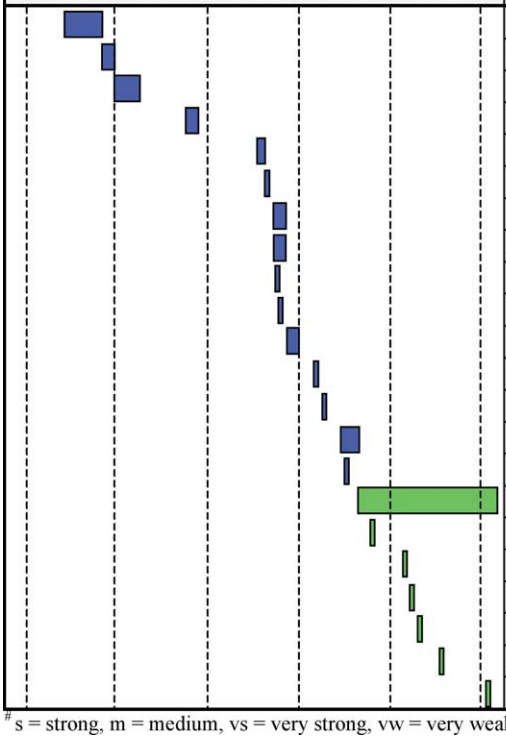
The characteristics of IR and Raman spectroscopy mentioned in Table 2 and discussed above apply both to normal macroscopic IR and Raman spectroscopy and to microspectroscopy. In macroscopic IR or Raman spectroscopy the combination of small sampling areas and heterogeneous food samples is often a problem, because it calls for homogenisation of the samples prior to spectral analysis or for repeated sampling of each sample in order to obtain representative spectra. However, with microspectroscopy this combination is turned into an advantage, because it offers the possibility of studying the food heterogeneity in detail, the only concern being that the heterogeneities to be studied must be on a larger scale than the spatial resolution obtainable with the instrument used. In other words, it must be possible to find homogeneous areas in the sample that are larger than or equal to the smallest possible acquisition area, if characteristic spectra of the feature(s) are to be obtained. In the food industry, microspectroscopy is thus most often called for in situations where the objective is identification of advantageous or disadvantageous microscopic features of food samples. For this reason microspectroscopy remains a laboratory technique for experts, while on-line solutions for quantitative spectroscopy are out of the question, at least with the present technology.

Food applications

In situ starch and pectin in the potato cell

FT-IR microscopy has previously been utilised to characterize pectins in the cell wall of plants (McCann *et al.*, 1997). In this study Raman microspectroscopy was used to measure pectin in raw potato tubers,

Table 1. Infrared and Raman characteristic group frequencies (Naumann *et al.*, 1991; Pistorius, 1995 and Piot *et al.*, 2000)



Group Vibration	Intensity#		Description	Mainly observed in
	IR	Raman		
O-H stretch	vs	vw	Hydroxyl	Liquid phase
=C-H stretch	s-m	m	Unsaturated	Lipids
-C-H stretch	s-m	m	Saturated	Lipids
-C≡N stretch	m	s	Nitrile	
C=O stretch	s	m-w	Ester	Lipids, Amino Acid
C=O stretch	s	w-m	Carboxylic acid	Lipids, Amino Acid
C=O stretch	s	m-s	Amide I	Proteins
C=C stretch	m-w	s	Not conjugated	Lipids
C=C stretch	m	s	*Trans*	Lipids
C=C stretch	m	s	*Cis*	Lipids
N-H bending	s	w	Amide II	Proteins
C-H scissoring	m	m-w	Aliphatic -CH_2	Lipids
C-O stretch	s		Carboxylates	Amino Acids, Lipids
N-H bending	w-m	var	Amide III	Proteins
P=O stretch	vs	m-w	Phosphate ester	Lipids, Nucleic Acids
Fingerprint from skeleton				
C-O stretch	s	m-w	Ether	Carbohydrates
Skeletal mode		m	α-(1→4) linkage	Starch
C-O-C skeletal	m-w	m-w	β-configuration	Glucose, galactose, mannose
C-O-C skeletal	m-w	m	α-configuration	
C-H rocking	w-m	vw	Aliphatic -CH_2	Lipids
Skeletal mode		vs		Starch

s = strong, m = medium, vs = very strong, vw = very weak

generating spectra directly from the "intact" potato cell wall. In the cell wall, pectin shows a unique galactoronic methyl ester peak in the area around 1745 cm^{-1}, where the surrounding and dominating starch granules in the cell do not have signals that interfere (Figure 1).

Using the Raman microscope, it proved difficult to locate the pectin signals and obtain good signals before the open cell wall collapsed, which it did after a few minutes. Characteristic starch spectra were easily recorded by focusing on the individual starch granules in the cell. The pectin spectrum in Figure 1 indicates the presence of some fluorescence emission, but in this case an efficient green laser (532 nm) could be used to measure the potato cells, even though the green light gave rise to fluorescence.

Results. It was found to be ideal to cover the sample with a quartz cover slip to maintain the integrity of the cell wall. The pectin spectrum in Figure 1 is the best pectin spectrum recorded directly from the cell wall through a cover slip with a long-working-distance objective (2.4 mm). The strong Raman band at 474 cm^{-1} in a pectin and/or starch spectrum indicates that starch is measured, and not the pectin. Pectin was characterised by signals near 858 cm^{-1} (α-anomer carbohydrate and indicative of a very low degree of esterification), at 1455 cm^{-1} (ester O–CH_3 stretch) and at 1745 cm^{-1} (ester carbonyl C=O stretch). The signal near 1000 cm^{-1} indicates that aromatic compounds interfere with the pectin spectrum. The starch spectrum in Figure 1 is of remarkable quality, equal to or even better than a typical Raman spectrum obtained from isolated and purified potato starch. This indicates that the high crystallinity of the intact granule is well reflected in the spectrum.

Conclusions. It was possible to measure high-quality starch and pectin spectra directly in the potato cell wall with Raman microscopy using a green 532 nm laser. However, it remains to be elucidated whether the Raman spectra are sufficiently distinct to be able to distinguish between different qualities of the two substances in different potato cultivars and/or in different transgenic potatoes.

The distribution of amygdalin in bitter almonds

Cyanogenic glycosides contain a nitrile group (–C≡N), and upon hydrolysis they liberate the toxic compound hydrogen cyanide (HCN). Due to the relatively rigid triple bond, the nitrile group has special vibrational characteristics. Cyanogenic glycosides are accumulated in a number of plants as a defence system against fungi and pests. Bitter almonds contain the cyanogenic glycoside amygdalin.

Micklander, Brimer, and Engelsen (2002) developed a method for assessment of the content and distribution of amygdalin in bitter almonds. A part of their study

Table 2. Raman spectroscopy vs. Infrared (IR) spectroscopy

	IR spectroscopy	Raman spectroscopy
Water	Strong	Very weak
Spatial resolution	$\geq 10\ \mu m^2$	$\geq 1\ \mu m^2$
Signal-to-noise	High	Low
Fluorescence	None	Can be devastating
Heat induction	Negligible	Strong
Confocality	No	Yes

concerned *in situ* determination of the distribution of amygdalin across the cotyledon within a bitter almond. Raman microspectroscopy was chosen, because nitrile groups (–C≡N) give rise to relatively low intensity bands in IR spectroscopy (very small changes in dipole moment), whereas they are strongly Raman active (large change in polaisability during stretching vibrations), see Table 1. The nitrile vibration band is highly specific, as the nitrile group is rarely found in natural compounds. The band is found near 2240 cm^{-1} in the Raman spectrum (Figure 1) and this area is almost free from interference from other chemical components (Table 1). The aromatic ring found in amygdalin is also strongly Raman active, but the specificity is low, because aromatic compounds are ubiquitous in nature.

Results. Figure 2 shows the content of amygdalin across the cotyledon for every 33.3 μm along a straight line of the transversely sectioned slice of two different bitter almonds. The content is expressed as the area of the nitrile peak in the Raman spectrum. The figure shows that amygdalin was not present in measurable amounts in the centre of the two almonds, but that small quantities were found 0.5 mm from the centre. The content generally increased towards the epidermis, but with large local fluctuations (up to 500% within 33.3 μm). As amygdalin functions as a defence against attacking insects or fungi, high concentrations near the surface of the almond make sense.

Conclusions. Raman microspectroscopy proved to be a rapid and non-destructive method for probing the distribution of amygdalin in bitter almond cotyledons. The advantages of using the nitrile vibration band are that the band is located in a spectral range with no interference from other chemical compounds and that the nitrile group is specific for amygdalin, when studying bitter almonds.

The composition of blisters found on the surface of bread

When dough is stored at low temperatures prior to baking, a phenomenon called blistering can arise on the crust of the bread. Blisters are small white bulges on the crust of the bread with a diameter of 0.5–5 mm. One theory is that oil particles rise to the surface of the bread, creating an increased concentration of lipids. The composition of the blisters and especially the content of starch and gluten were studied and compared with the content of these components in the breadcrumb and crust.

IR microspectroscopy was chosen for this study, as fluorescence from maillard compounds in the crust and from certain amino acids made Raman spectroscopy impossible, at least when using a laser in the visible range. With the microscope it was possible to focus on different locations on the inside of blisters and to obtain good quality spectra from these positions.

Results. The characteristic vibrational bands of pure gluten and starch were found using FT-IR microspectroscopy in reflectance mode. It was found that the bands in the ranges 1663–1630 and 1595–1528 cm^{-1} could differentiate between spectra of blisters and of breadcrumb from comparable positions just below the crust. The two wavelength ranges are characteristic of the amide I band (overlapped with the HOH bending vibration from water) and the amide II band from gluten (Table 1). The spectra of the inside "walls" of the blisters showed higher intensity at the "starch" band and lower intensity at the "gluten" band compared to breadcrumb. This strongly indicates that the ratio of starch to gluten was higher in the breadcrumb just around the blisters than in the ordinary breadcrumb.

Conclusions. The observation that the gluten content is lower in the blister walls than in breadcrumb could explain why blisters are developed. The blisters may be developed in spots on the surface of the bread where the gluten network in the dough/breadcrumb below the developing crust is too weak to withhold the CO_2 produced by the yeast during the baking process. The CO_2 is then instead trapped in the crust, causing blisters. One can speculate that such areas develop because of condensation of water on the surface of the cold dough in the beginning of the baking process. In contrast to the gluten network, small starch grains could diffuse into wetted parts of the surface and part of the surface would therefore become rich in starch. In any case, FT-IR microspectroscopy made it possible to study the microstructure inside blisters, and the technique provided new information about this yet unsolved problem.

The microstructure of high-lysine barley

The nutritional value of barley is limited by a low content of essential amino acids. The primary limitation is lysine, and the nutritional value of barley can be improved significantly by increasing the lysine content (Munck, Karlsson, Hagberg, & Eggum, 1970; Newman & McGuire, 1985; Shewry, Williamson, & Kreis, 1987; Munck, 1992). Albumins and globulins contain high concentrations of lysine and exist primarily in the embryo and the aleurone layer. Since the storage proteins, prolamines and glutelins exist in the endosperm, a larger embryo and a thicker aleurone layer would result in higher concentrations of essential amino acids in barley.

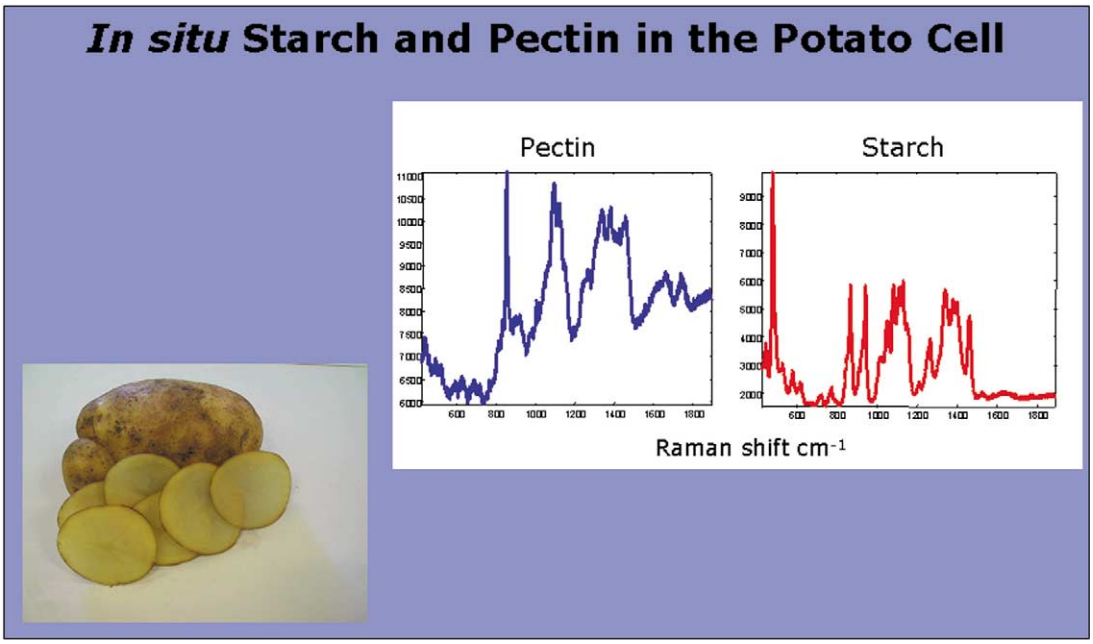


Fig. 1. Raman spectra of pectin and starch obtained directly in the potato cell using a LabRam Infinity instrument equipped with a green laser (532 nm, Nd:YAG), a Peltier-cooled CCD detector and a long working distance 100× Leica objective. Laser power at the sample was approximately 40 mW.

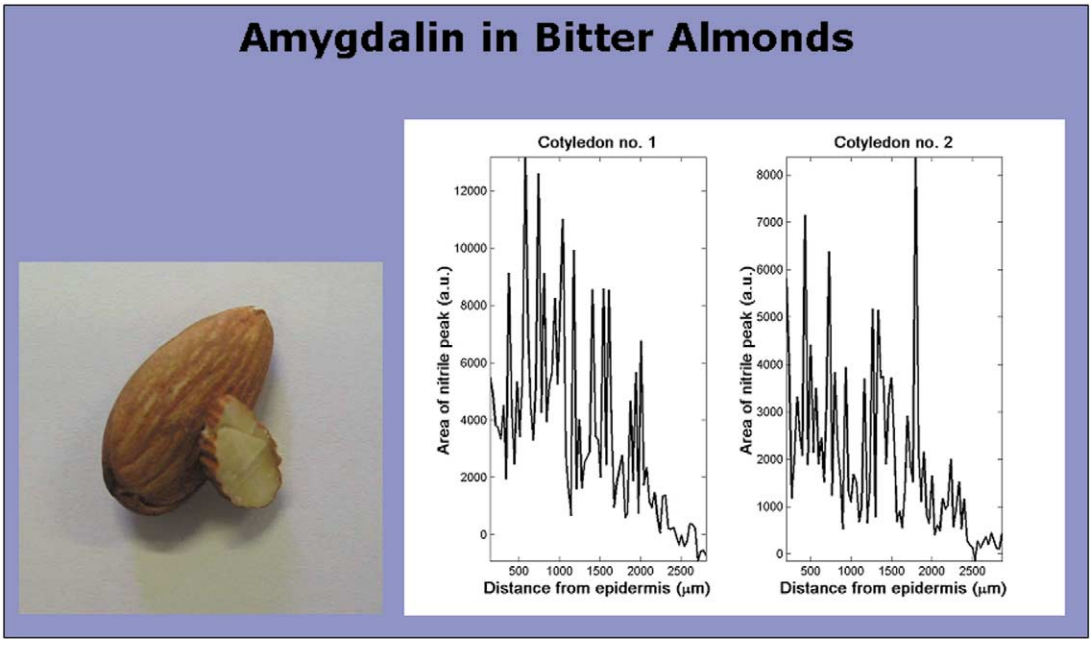


Fig. 2. Raman measurements of amygdalin in bitter almond cotyledons using the same equipment as in Fig. 1. The plots show for two different almonds the area of the nitrile peak at approximately 2242 cm^{-1} for 81 positions along a line from the epidermis to the centre of the almond.

Barley kernels can be forced to mutate by inhibiting the synthesis of prolamines and thereby increase the lysine content of barley. The mutant, *Risø 1508*, and its mother sort, *Bomi*, were investigated here with FT-IR microspectroscopy in order to describe the microstructure of the barley kernels. In addition to the difference in lysine content, it is well known that the *Risø 1508* mutant has lower starch content than *Bomi*. It is well-known that the two cultivars differ with regard to the protein composition and with regard to the ratio between protein-N and amide-N. The amide-N content is approximately 16% of the protein-N content in case of *Risø 1508*, while it is approximately 12% for normal barley (Munck *et al.*, 2001). In a comprehensive study

by Piot, Autran, and Manfait (2000) the content of protein in wheat kernels was measured using Raman microspectroscopy that offers higher spatial resolution.

For measurements in transmittance, slices of kernels had to be 4 μm or less in order to keep the absorbance within an acceptable limit. After slicing of the kernels it was necessary to keep the slices between two cower slips to avoid curling. During measurement the cover slips were removed, and the slice was placed on a ZnSe plate. Area scans were used to map the contents and distribution of macromolecules in the slices. A smoothing of the obtained maps was obtained by making the distance between neighbour acquisitions equal to half the length of the side of the square acquisition area. The area scan generated from a slice of *Risø 1508* consisted of 25×38 scans and of 34×44 scans from a slice of *Bomi* (Figure 3).

Results. Specific wave numbers were chosen to indicate the contents of different macromolecules, i.e. 1543 cm^{-1} (Amide II) was chosen to indicate protein, 1148 cm^{-1} (coupled C–C and C–O vibrations) to indicate carbohydrate, and 1738 cm^{-1} (C = O stretch) to indicate the content of lipid. The ratio between 3100 cm^{-1} (primary amide) and 3300 cm^{-1} (primary amide + secondary amide) was used as an indication of the amount of amide-N and protein-N, respectively.

It was found that protein primarily exists in the borderline of the kernel, but no difference was found between the content of protein between *Risø 1508* and *Bomi*. However, an indication of relatively higher amide-N content in *Risø 1508* than in *Bomi* was found (Figure 3, Left). In *Bomi*, carbohydrates were concentrated in a specific area, whereas in *Risø 1508* they were found over most of the border (Figure 3, Right). The characteristic band of lipids (~1738 cm^{-1}) did not show any specific variations.

Conclusion. With IR miscospectroscopy it was possible to obtain an indication of the composition and distribution of the macromolecules in the two cultivars of barley and the difference between them was evident, especially the spatial distribution of carbohydrate differed between the two cultivars. The results are, however, only indicative as the required thinness of the slices (~4 μm) made it difficult to make slices with approximately the same thickness, which created problems with higher absorbance from thicker slices.

The composition of white spots developing in shrimp shells during frozen storage

Raw shrimps are a valuable food, and they are furthermore used for food decoration in Asian countries. Consequently, the appearance of the shrimps is an important quality parameter. When shrimps of the species *Pandalus borealis* (pink shrimp) are still alive their shell appear transparent, but white spots may develop in the shell during transport, during which the shrimps are frozen. Mikkelsen *et al*. (1997) and Mikkelsen *et al*. (1999) used a number of techniques to identify the substances present in the white spots, among them FT-IR microspectroscopy, Raman spectroscopy, X-ray diffraction, differential scanning calorimetry and electron microscopy.

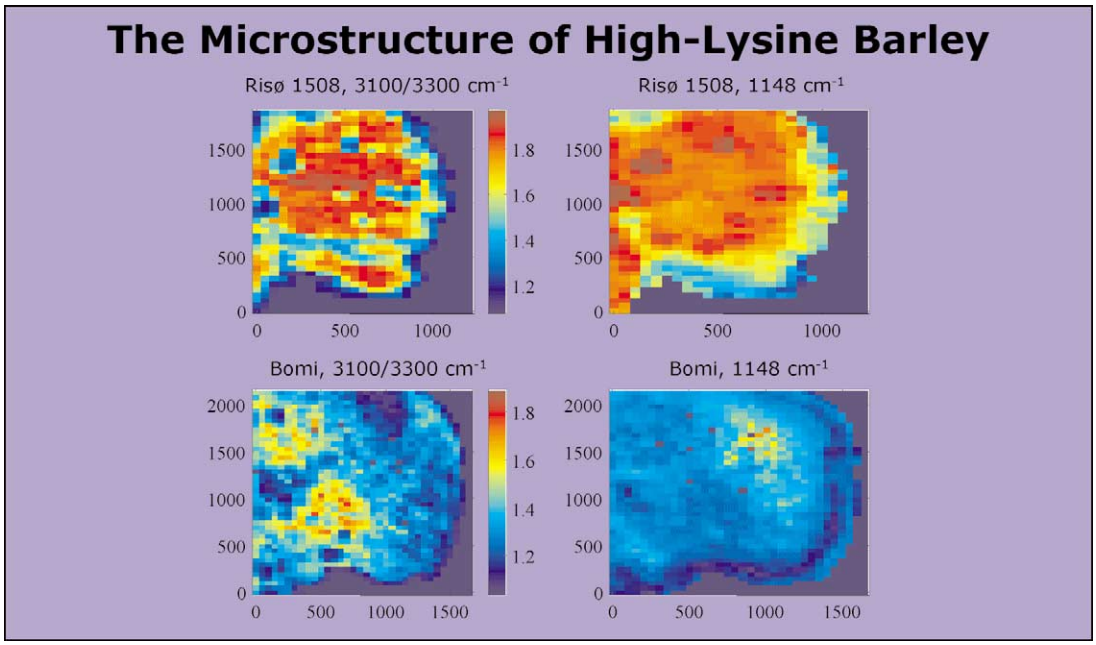


Fig. 3. Microstructure of barley as viewed by a FT-IR microscope. The area scans to the left are coloured according to the intensity ratio between 3100 cm^{-1} and 3300 cm^{-1} to indicate the amount of secondary amides, while the area scans to the right are coloured according to the intensity at 1148 cm^{-1} related to carbohydrate content. The instrument used was a Perkin Elmer System 2000 combined with an AutoImage microscope. The detector was a MCT detector (Mercury Cadmium Telluride, 10,000–700 cm^{-1}) cooled with liquid nitrogen.

Results. By studying the IR and Raman spectra of the white spot areas and by comparing them to spectra of pure α-chitin and of a number of calcium carbonate polymorphs, it was found that the white spots consisted of amorphous α-chitin and $CaCO_3{\cdot}6H_2O$, known as ikaite, from the fjord in Greenland where this inorganic salt was first discovered. Ikaite degrades when brought to room temperature and instead the anhydrous calcium carbonates calcite and vaterite are formed.

Conclusions. The substances present in the white spots in shrimp shells were successfully identified. FT-IR miscrospectroscopy could not have solved the problem alone, but the technique gave a valuable contribution. Raman spectroscopy proved to be an exceptional good probe for monitoring the conversion of $CaCO_3$ hexahydrate to its anhydrous polymorphs due to the insensitivity of Raman spectroscopy to water and the high sensitivity of the technique to different inorganic crystalline arrangements.

Conclusions

Both Raman and FT-IR microspectroscopy offer information on the molecular vibrations and structure of food samples. Raman has an advantage because of its ease of sampling, its higher resolution and the possibility for confocal measurements, but the lower signal to noise ratio, the risk of damaging the sample with the laser, and especially autofluorescence of the sample may hamper its applicability and provide an option for FT-IR microspectroscopy.

In each of the food applications described, new information about the microstructure of the food system was found by the use of microspectroscopy. In the potato cell wall it was possible, but difficult, to record a Raman spectrum of pectin; however, high-quality spectra of individual starch granules in the potato cell were readily acquired. Using Raman microspectroscopy, it was possible to study the distribution of amygdalin in bitter almond cotylodons. FT-IR microspectroscopy showed that blisters in bread crusts contained relatively small concentrations of gluten and large concentrations of starch compared to the ordinary breadcrumb. FT-IR microspectroscopy also showed differences in the distribution of macromolecules between a barley mutant and its mother sort, and it helped identify white spots in shrimp shells as α-chitin and ikaite.

The success of microspectroscopy for such diverse applications as those outlined above shows the vast potential of this technique. Furthermore, future and ongoing development within the field of vibrational microspectroscopy, such as CCD-based imaging systems and NIR microscopy techniques, will no doubt extend the range of possible applications. NIR microspectroscopy may offer the ideal compromise between IR and Raman techniques, as it has the spatial resolution of Raman and the signal-to-noise ratio of IR. Moreover, it is possible to measure through a quartz window, which is a great advantage in many situations. The holographic information carried by NIR spectra (convoluted combinations and overtones) has already proven its worth in numerous macroscopic applications in the food industry, and we believe that microspectroscopy will benefit from it as well.

Acknowledgements

The authors wish to thank the Centre for Advanced Food Studies and the Danish Agricultural and Veterinary Research Council for financial support. Robert Kold Hansen and fellow scientists at Danisco Cultor Denmark (Brabrand) are most gratefully acknowledged for valuable discussions and sample materials. Gilda Kischinovsky is acknowledged for assistance with the manuscript.

References

McCann, M. C., Chen, L., Roberts, K., Kemsley, E. K., Sene, C., Carpita, N. C., Stacey, N. J., & Wilson, R. H. (1997). Infrared microspectroscopy: sampling heterogeneity in plant cell wall composition and architecture. *Physiologia Plantarum*, *100*, 729–738.

Micklander, E., Brimer, L., & Engelsen, S. B. (2002). Non-invasive assay for cyanogenic constituents in plants by Raman spectroscopy: content and distribution of amygdalin in bitter almond (Prunus amygdalus). *Applied Spectroscopy*, *56*, 1139–1146.

Mikkelsen, A., Andersen, A. B., Engelsen, S. B., Hansen, H. C. B., Larsen, O., & Skibsted, L. H. (1999). Presence and dehydration of ikaite, calcium carbonate hexahydrate, in frozen shrimp shell. *Journal of Agricultural and Food Chemistry*, *47*, 911–917.

Mikkelsen, A., Engelsen, S. B., Hansen, H. C. B., Larsen, O., & Skibsted, L. H. (1997). Calcium carbonate crystallization in the α-chitin matrix of the shell of pink shrimp, Pandulus borealis, during frozen storage. *Journal of Crystal Growth*, *177*, 125–134.

Munck, L., Karlsson, K. E., Hagberg, A., & Eggum, B. O. (1970). Gene for improved nutritional value in barley seed protein. *Science*, *168*(3934), 985–987.

Munck, L. (1992). The case of high-lysine barley breeding. In P. R. Shewry (Ed.), *Barley, genetics, biochemistry, molecular bilology and biotechnology* (pp. 573–602). Wallingford, Oxon: CBA International.

Munck, L., Pram Nielsen, J., Møller, B., Jacobsen, S., Søndergaard, I., Engelsen, S. B., Nørgaard, L., & Bro, R. (2001). Exploring the phenotypic expression of a regulatory proteome-altering gene by spectroscopy and chemometrics. *Analytica Chimica Acta*, *446*, 171–186.

Naumann, D., Helm, D., Labischinski, H., & Giesbrecht, P. (1991). The characterization of microorganisms by Fourier-transform infrared spectroscopy (FT-IR). Chapter 3. In W. H. Nelson (Ed.), *Modern techniques for rapid microbiological analysis* (1st ed.). New York: USA:VCH.

Newman, C. W., & McGuire, C. F. (1985). Nutritional quality of barley. In D. C. Rasmusson (Ed.), *Barley* (pp. 403–456). Madison, WI: American Society of Agronomy.

Piot, O., Autran, J. C., & Manfait, M. (2000). Spatial distribution of protein and phenolic constituents in wheat grain as probed by confocal Raman microspectroscopy. *Journal of Cereal Science*, *32*, 57–71.

Pistorius, A. M. A. (1995). Biochemical applications of FT-IR spectroscopy. *Spectroscopy Europe*, *7*, 8–15.

Shewry, P. R., Williamson, M. S., & Kreis, M. (1987). Effect of mutant

genes on the synthesis of storage components in developing barley endosperms. In H. Thomas, & D. Grierson (Eds.), *Developmental mutants in higher plants* (pp. 95–118). Cambridge, UK: Press syndicate of the University of Cambridge.

West, Y. D. (1996). Study of sample heating effects arising during laser raman spectroscopy. *The Internet Journal of Vibrational Spectroscopy*, *1*, 5.

Wetzel, D. L., & LeVine, S. M. (1999). Microspectroscopy — imaging molecular chemistry with infrared microscopy. *Science*, *285*, 1224–1225.

PAPER IV

Multivariate Analysis of Time Domain NMR Signals in Relation to Food Quality

E. Micklander, L.G. Thygesen, H.T. Pedersen, F. van den Berg, R. Bro, D.N. Rutledge and S.B. Engelsen

MULTIVARIATE ANALYSIS OF TIME DOMAIN NMR SIGNALS IN RELATION TO FOOD QUALITY

Elisabeth Micklander[1], Lisbeth G. Thygesen[1], Henrik T. Pedersen[1], Frans van den Berg[1], Rasmus Bro[1], Douglas N. Rutledge[2] and Søren B. Engelsen[1]

[1]Center for Advanced Food Studies, Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, 1958 Frederiksberg C, Denmark
[2]Institut National Agronomique, Laboratoire de Chimie Analytique, 16, rue Claude Bernard, 75005 Paris, Cedex 05, France

1 INTRODUCTION

NMR is one of the most successful analytical techniques of our time. It is directly applicable to amorphous, heterogeneous and opaque systems such as food and foodstuffs. Time domain NMR (TD-NMR) is a relatively inexpensive version of the technique, which still has the potential to serve as a unique window into the complicated food matrix. It allows food engineers and scientists to "get in touch" with the protons in a sample. The major food industries all have problems for which the best method of measurement is TD-NMR. These include, for example, **solid fat** determination in the plant oil industry, **total fat** content in the slaughter, oil seed, fishmeal, confectionary and dairy industry, **droplet size** in the margarine/butter/spread industry, **adsorbed and total water** in the feed, ingredients, snacks and cookie industries and **gel-formation** in the stabilizer/hydrocolloid industry. In addition to these traditional applications TD-NMR offers the possibility to explore end-properties of foods such as texture, mouth feel and slipperiness to the extent that these features depend on the compartmentalization and mobility of water.

The TD-NMR technique is one the most versatile, multivariate, low perturbation, non-destructive analytical techniques on the market. It gives volume measures that in contrast to optical methods are insensitive to surface characteristics and can handle highly opaque samples. Despite the fact that the TD-NMR signals are multivariate, industrial use tends to be strictly application limited, normative and univariate. The reasons for this apparent misuse of TD-NMR are multi-facetted, but one problem is that TD-NMR is associated with a difficult and "dusty" image (poor compatibility with the user), a general lack of university training and instrument manufacturers that have developed in-house software that constrains the users rather than opening the way to the immense possibilities offered by the technique.

TD-NMR has disadvantages as well. Firstly a strong magnetic field is required. This is a technological problem that can be solved even near process lines, for example by one-sided NMR instruments. Secondly, instrument initialization, pulse settings and data acquisition give infinite choices, which in the first place would appear as an advantage, but in practice is a hindrance. Signals from TD-NMR instruments are of relatively low resolution and a few pre-constructed and pre-optimized ready-to-use pulse sequences will be able to capture all necessary NMR information. Thirdly, because TD-NMR is relatively low resolution, the few functional elements or proton populations in the relaxation decays all have to be found within the same narrow time-constant range. This fact makes data interpretation difficult and tendentious. A good example is the interpretation of TD-NMR signals from meat, which despite relative simplicity and numerous studies still are not fully elucidated[1-3].

The quantitative aspect of TD-NMR data analysis has and still is suffering from over-simplified data analysis, mainly because the complexity is too high for a deterministic approach and because it is conceptually difficult to cope with a high number of variables. This simplified tradition has resulted in the widespread normative use of TD-NMR: if you keep only the part of the data that you think is relevant, you will confirm what you already know is important and thus reduce your chances of innovation.

We believe that NMR scientists in industry and academia can learn from the successful progress of Near-infrared (NIR) spectroscopy, which in the last decade has revolutionized quality control in practically all areas of primary food and feed production. Near-infrared spectroscopy has been implemented for monitoring quality of millions of samples of cereals, milk and meat with unprecedented precision and speed. The key to this success is the extraordinary synergy that lies in the merging of a rapid non-invasive spectroscopic sensor and the new data technology called chemometrics. In this paper we have therefore set out to investigate the explorative potential of TD-NMR via the application of multivariate data analysis (chemometrics). Using this approach engineers in the food industry are now able to explore their processes using so-called soft models and spectroscopic sensors to complement traditional deterministic models and univariate measurements. This dramatic development appears to transform the food industry from a traditionally low-technology industry into a high-technology one, alongside quality control developments in the pharmaceutical industry. This technology leap, while still in its onset, has by now matured to a stage where the exploratory TD-NMR approach should also be included in the toolbox of food researchers and scientists.

2 EXPLORATORY DATA ANAYLYSIS

Perhaps because NMR was developed by physicists and chemists, it has primarily been used for identification, authenticity determination and simple quantification purposes. Its explorative potential has largely remained unexploited. The first report on the use of chemometrics on NMR data appeared in 1983 by Johnels *et al.*[4], but it was not before the early nineties that new developments of applying multivariate data analysis to high-resolution NMR signals occurred, with the impetus being the analysis of complex 1D and 2D NMR structures of complex biological and pharmaceutic/therapeutic matrices[5,6]. The use of chemometric data analysis to TD-NMR which is still in its infancy, was first adapted in the late nineties[7-9].

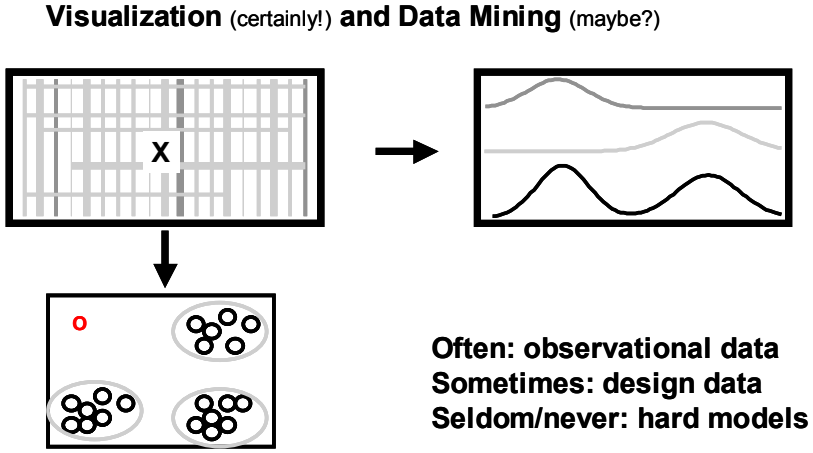


Figure 1 *Exploratory Data Analysis*

One of the main advantages of chemometric data analysis is the possibility of projecting multivariate data into a few dimensions via a graphical representation. Chemometrics is also called “statistics without tears”, because it makes it possible to

handle large data sets and deals efficiently with real-world multivariate data, taking advantage of the previously feared co-linearity of spectral data. With chemometrics it is even possible to analyse whole spectra in real time. The chemometric methods are based on the calculation of underlying latent data structures using a two-dimensional data strategy, i.e. measuring a series of samples and finding common latent data structures (Figure 1). In this paper we will demonstrate the application of multivariate data analysis to a number of different food systems and through those evaluate the current status of the use of chemometrics for the analysis of time domain NMR data obtained from static and dynamic food systems, including advanced multi-way chemometric approaches such as SLICING[10].

3 PRINCIPAL COMPONENT ANALYSIS

The basis of most chemometric algorithms is Principal Component Analysis (PCA) which has already celebrated its 100-year anniversary[11], still in splendid shape. PCA can be considered as the first amendment in exploratory analysis due to its extraordinarily robust data reduction and data presentation capabilities. In PCA the multivariate (spectral) data set is resolved into orthogonal components whose linear combinations approximate the original data set in a least squares sense. Common to bi-linear models is that an entire matrix, with each row being the measurement from one sample, must be acquired. The algorithm works on this entire set instead of on one sample at a time.

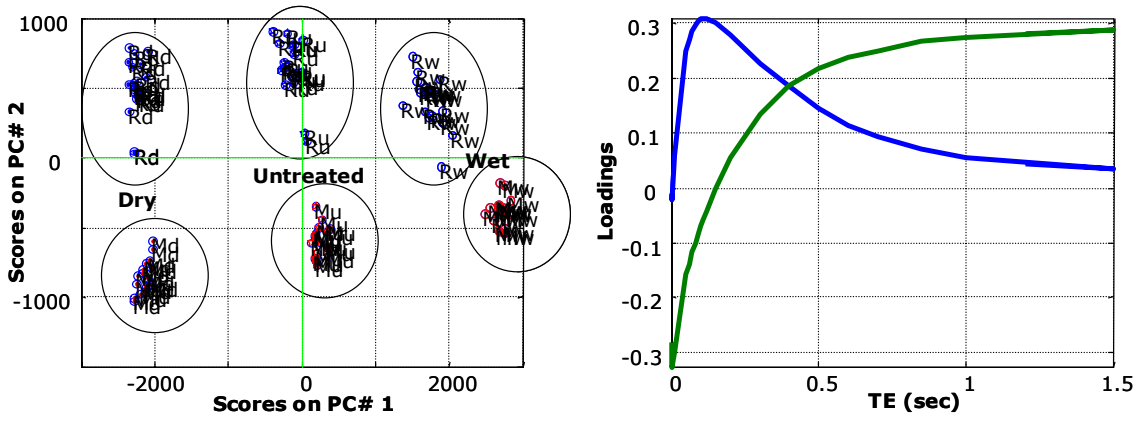


Figure 2 *Principal Component Analysis of inversion recovery data measured on oil seeds*

PCA remains the primary tool for initial investigations of large bi-linear data structures for the study of trends, grouping and outliers. Despite the fact that TD-NMR data rarely satisfies the bi-linear condition (underlying latent structures have a tendency to change with concentration), PCA has been proven very robust to this type of data[12]. Figure 2 shows an example of the application of PCA to TD-NMR Inversion Recovery data of oil seeds[13]. The reduction of the data matrix (117 samples × 23 inversion delays) is readily translated into two latent structures. The sample *scores* are grouped into 6 subclasses: dry (left), normal and wet (right) seeds. The upper parts are rapeseeds and the lower clusters are mustard seeds. It is also observed that two samples of the rapeseeds are atypically placed in the direction of the mustard seeds. The principal components or loadings indicate on the parts in the inversion recovery curves that are important. In this case the first loading (corresponding to the horizontal axis in the score plot) has the shape of a general inversion recovery curve, while the second loading gives the amount of modification. The degree of explanation for the total variation in the data for the first two principal components is above 99%. Thus, this highly compressed visualization of the *whole* data set gives an almost perfect window into the total variation in the data. More subtle details may be revealed by monitoring additional principal components.

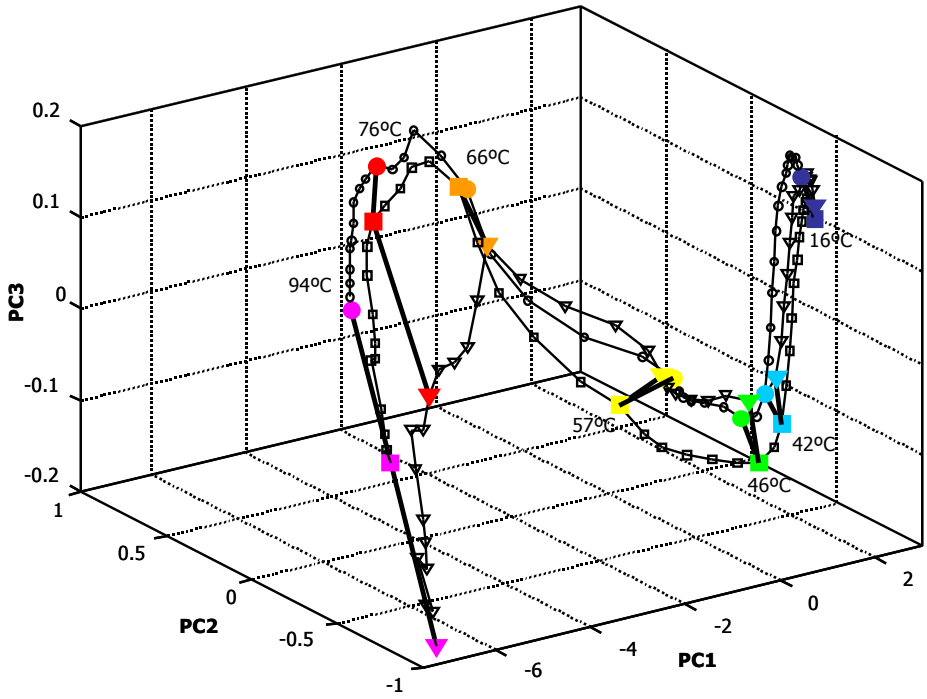


Figure 3 *Principal Component Analysis of CPMG relaxation data of three replicates of cooking meat inside the NMR magnet*

Figure 3 shows the PCA score plot of three replicates of CPMG data from a dynamic food process, the cooking of meat inside an NMR magnet. Each replicate consisting of 43 relaxation curves of 4000 data points[14]. Although the relaxation data as a function of cooking temperature is not a truly bilinear data structure (the T_2 distribution of characteristic relaxation times changes with temperature and the NMR signal decreases with temperature due to less favourable Boltzmann distribution), PCA analysis of TD-NMR data has proven to be a robust and valuable data reduction and enhancement method. In Figure 3 each point represents a full relaxation decay curve with 4000 data points. It demonstrates in full that LF-NMR measurements can serve as a window into a food matrix, as changes can be observed in a non-invasive manner. In the 3D PCA score plot the temperatures where major changes in the states of water in meat occur are readily identified; sharp bends in the scores plot, corresponding to significant changes in the properties of the water, can be identified at 46 ºC and 66 ºC, while softer bends are found at approximately 42 ºC, 57 ºC and 76 ºC. These transition temperatures, identified by the PCA, can be explained by changes in the main constituents of meat during cooking: denaturation and shrinkage of the myofibrillar proteins and the connective tissue, which affects the water-holding capacity/distribution of water in meat. The lowest transition temperature found in the PCA score plot (42 ºC) corresponds to the onset of myosin denaturation. The denaturation of sarcoplasmic and myofibrillar proteins begins at lower temperatures, but the rate and amount of denaturation increases significantly above this temperature. The transverse shrinkage of myofibrils starts at 45 ºC, which is where one of the major changes is seen in the PCA score plot. The transition of the states of water at 57 ºC can be explained by the fact that shrinkage parallel to the myofibrils starts at this temperature, but the onset of collagen denaturation beginning somewhere between 50 and 60°C could also contribute to this transition. The longitudinal contraction intensifies at 64-65 ºC, which is where a major change is found in the NMR data, as shown by the PCA analysis. The intensified contraction is related to a major peak in thermal denaturation of the endomysial, perimysial and epimysial collagen around these temperatures. The slight transition observed around 76 ºC may reflect further

gelatinisation of collagen, which occurs up to approximately 80 ºC, but it is more likely due to the onset of actin denaturation, which has been reported to happen around 80 ºC. This straightforward interpretation of the transitions in the PCA results on the basis of denaturation of major components of meat goes a long way to confirm that the information contained in TD-NMR data obtained during cooking of meat is indeed related to important changes of the meat matrix due to heating[14].

If it is desirable to further compress the data into only two dimensions - the optimum for human perception - PCA can be further reduced using a so-called topology-preserving non-linear Sammon projection[15]:

$$\min \quad E = \frac{1}{\sum_{i<j}^{N} D_{ij}} \sum_{i<j}^{N} \frac{\left(d_{ij} - D_{ij}\right)^2}{D_{ij}}$$

where the distances between objects d_{ij} in new space (e.g. 2D) are optimized to mimic as closely as possible the distances D_{ij} in high-dimensional space (fixed). This type of technique is exactly what is used by geographers when they project parts of the globe onto a flat map. If we do the exercise in the meat NMR-cooking example, the result appears as displayed in Figure 4.

We now have the entire variation in the cooking data mapped onto a 2D surface largely maintaining the information about the transition temperatures: 42 ºC, 46 ºC, 57 ºC, 66 ºC and 76 ºC already identified by the PCA. However, the directional information is lost, for which reason the Sammon map scores cannot be used in for example regression problems.

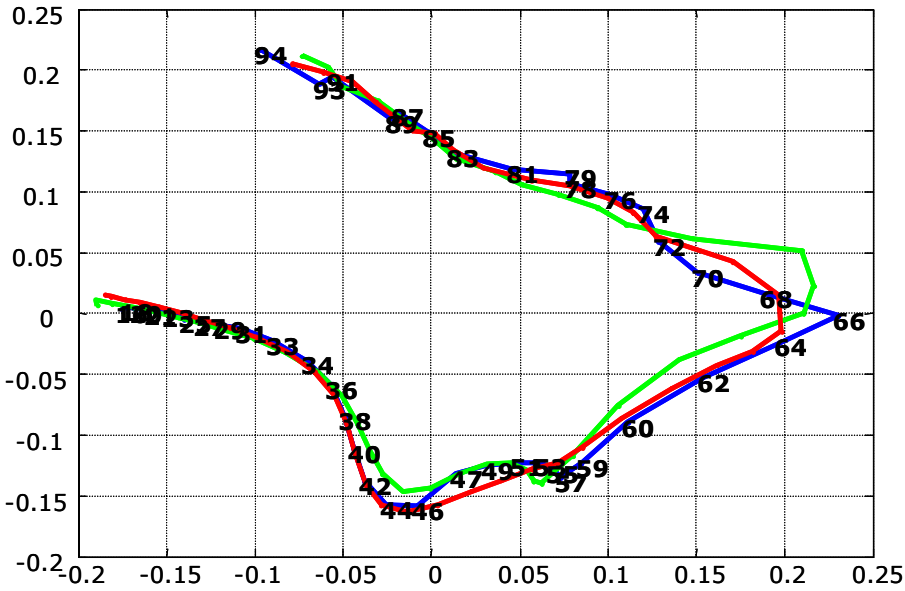


Figure 4 *Sammon map of the NMR cooking*

We have also studied a more exotic use of PCA namely Principal Phase Correction (PPC)[10]. TD-NMR signals are normally recorded as quadrature data that are magnitude-transformed to correct for phase errors in the acquisition. This procedure represents a simple and robust transformation independent of detection of the phase angle. However, it introduces a bias in terms of non-exponentiality in the magnitude-corrected data. This poses a serious problem to all data analytical algorithms based on underlying exponential

structures. To solve this problem we have proposed a new, simple and non-iterative procedure for performing phase correction for TD-NMR CPMG data called Principal Phase Correction (PPC). It is based on rank reduced singular value decomposition (SVD), a technique equivalent to PCA. If **a** and **b** are two vectors of length *J* (Figure 5) holding the two quadrature channels, then PPC-rotation is performed by a singular value decomposition (SVD) on the matrix holding **a** and **b** in its columns. Ideally, the second score value is zero, if there are no changes in phase throughout the measurements. In practice, though, minor differences are observed and the second principal component represents noise. The product of the loading times the score of the first principal component provides an optimal representation of the phase-rotated measurements and the influence of this noise is reduced. An example of PPC-correction of a CPMG relaxation curve is shown in Figure 5. PPC extracts the optimal single-channel representation of the data, obtained in a least squares sense, taking into account that errors appear in both original axes.

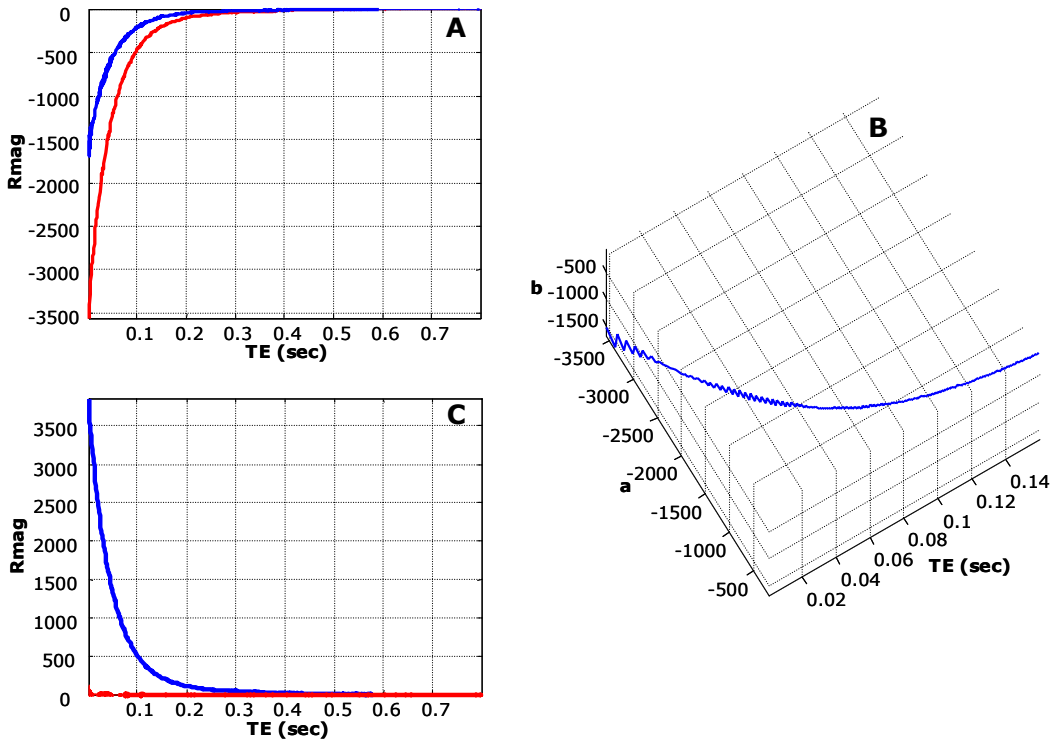


Figure 5 *Principal Phase Correction. The two channels **a** and **b** (A), **a** versus **b** (shown in the time interval 0-0.16 sec) (B) and the PPC corrected relaxation curve (C)*

4 PARTIAL LEAST SQUARES REGRESSION

Partial Least Squares Regression (PLSR), based on the PCA concept is its counterpart for regression analysis. While PCA can be compared to shopping (in the data) without a shopping list (i.e. the data analysis is performed without the use of *a priori* knowledge), PLSR is like shopping *with* a specific shopping list. PLSR is a predictive two-block regression method based on estimated latent variables and applies to the simultaneous analysis of two data sets on the same objects (e.g. spectra and physical/chemical tests). The purpose of PLSR is to build a linear regression model that enables prediction of a desired characteristic from a measured multivariate signal. It has been and still is common practice in TD-NMR to design calibration models from univariate signals collected from intrinsically multivariate signals. The most prominent examples are solid

fat determination based on the ratio between the signal at 11 μs and the signal at 70 μs and fat determinations based, for example, on the amplitude of a single gradient simulated echo. Such calibrations are, from a data technological point of view, the worst approach and it is most surprising that it is still common practice in the new millennium. The main argument for not using such univariate calibration approaches is that there are no means to automatically detect outlying measurements caused by instrument drift and erroneous handling of the sample. It is, however, surprising that the univariate practice has persisted for so long and indicates that either TD-NMR instruments are extraordinarily stable over time, that re-calibrations are performed frequently or that the TD-NMR instruments can easily achieve the required precision without making use of all the measured information.

Fats are important food constituents and are measured in a number of foodstuffs to assure product quality. TD-NMR is sensitive to proton mobility, which is why the technique is useful for solid-fat determinations in the edible oils industry and for swift and accurate measurement of total fat content in dry foods. In food the liquid phase is usually dominated by two proton-rich components: water (H_2O) and fat (*poly*-CH_2), which in practice has proven to be difficult to differentiate by standard pulsed NMR experiments, as the relaxation rates of the two types of protons only differ by a factor of two in many products. The established standard NMR approach to measuring total fat in dried food (including meat) is a univariate approach where only one point is acquired from each sample, namely the top-point of a spin-echo[16]. Simple linear regression is used for calibration. This procedure requires that the samples must be dried prior to analysis in order to remove the interfering signal contribution from water protons. Furthermore, the samples are heated to ensure that the entire fat phase is liquid upon NMR analysis to obtain the total fat signal. The second evolution in total fat determination of foods by TD-NMR occurred when pulsed field gradients[17-19] were introduced to suppress the water proton signal, leaving mainly the fat signal. In principle, the pulsed field-gradient experiment consists of two gradient pulses of which the first "labels" the protons according to their local magnetic field and the second creates a gradient-stimulated echo. In this manner the NMR relaxation is sensitized to proton diffusion, as all protons that have diffused to a new magnetic environment will not be refocused.

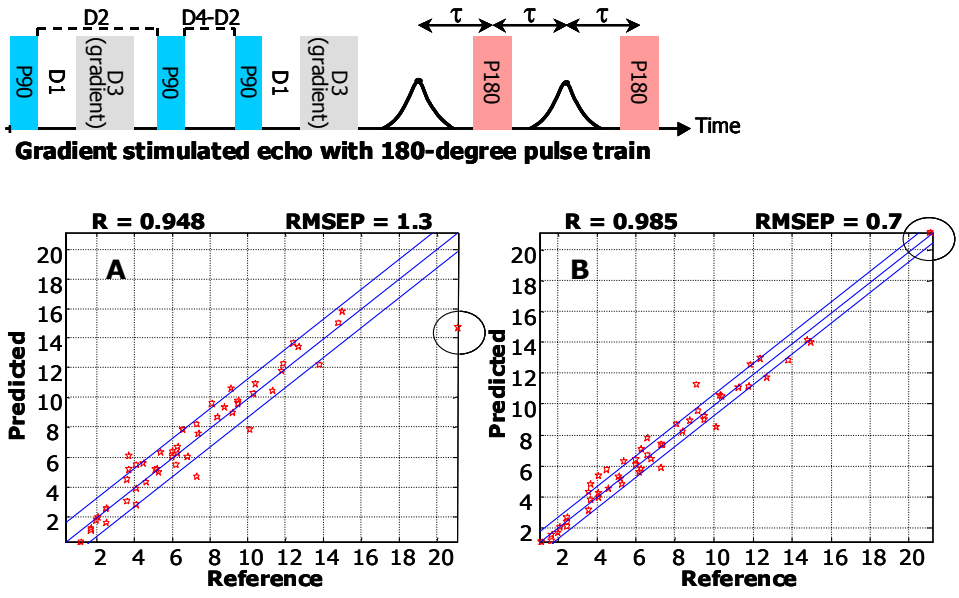


Figure 6 *Partial least squares regression. Predicted versus measured plots based on Linear regression of the gradient-stimulated echo (A) and the 1PC PLSR model of the gradient stimulated echo followed by a 180 degree pulse train (B)*

Since water protons diffuse more readily than fat protons, the water proton signal can be suppressed gradually (beginning with the most mobile) by a pulsed gradient of appropriate length and strength. By using field gradients to discriminate between the water and fat protons, the samples need not be dried prior to analysis, reducing the time of analysis significantly. However, the samples must still be heated in order to ensure that all fat components in the sample are liquid. In this method again only one point is acquired, as in the spin-echo experiment, namely the top-point of the gradient stimulated echo.

Figure 6 demonstrates clearly how a simple multivariate upgrade of the gradient-stimulated echo experiment can enhance calibration performance[20]. To the left is a calibration of the gradient-stimulated echo amplitude of 51 fresh meat samples regressed on conventional SBR (Schmid, Bondzynski and Ratzlaff) solvent extraction method. The cross-validated calibration, results in a correlation of r=0.948 and a root mean square error of cross-validation (RMSECV – average prediction error) of 1.3 %(w/w) fat, and it clearly demonstrates that the linear relationship between the echo amplitude and the SBR fat determination is broken above 20%(w/w) fat. To the right the corresponding calibration is shown for a multivariate upgrade of the gradient-stimulated echo pulse experiment in which the gradient echo is simply followed by a 180 degree pulse train. The cross-validated calibration based on only one PLS component has now improved the calibration performance to only one half the error of the traditional approach with r=0.986 and RMSECV=0.69%(w/w) fat. Moreover, this simple PLSR model is now able to linearly extrapolate measurements above 20%(w/w) fat. This type of performance enhancement is the rule rather than the exception when comparing uni- or oligo-variate approaches with multivariate approaches. For this reason it is intriguing that official standard methods for fat determination are still based on univarite reasoning – we can predict that these dinosaurs soon will be overtaken by new improved company standards, as has been the case with NIR models.

At one point in time we have also identified an exotic use of PLSR which at first appeared to be an extraordinary good idea. If we consider the partial PLSR optimization criterium:

$$\left\| y - Gb \right\|^2 = 0$$

and let the response vector **y** be a TD-NMR CPMG relaxation curve and the matrix **G** be a matrix containing a larger number of pure mono-exponential curve (see Figure 7). Then the regression coefficient vector **b** should in principle carry the T_2 distributed spectrum. While the simplicity of this method is appealing the results produced are incorrect.

Figure 7 shows a typical result of applying PLSR to the distributed exponential problem. In this case the experimental data is the CPMG curve measured on meat. The main variability in this PLSR model is the number of pure exponentials to include in the G matrix and the selection of these to obtain optimal condition number of PLSR components to include in the model. However, independent of the choices made it is not possible to obtain reasonable T_2 domain spectra. The reasons for this negative result of a model-free approach are multiple. In reality the problem is a so-called ill-posed matrix inversion problem – the matrix G which ideally should have full rank is in practise of low rank. A number of techniques have been developed to solve this type of numerical problem most of which introduces constraints such as regularisation, smoothing and non-negativity[21,22]. Obviously the condition number of the G matrix is crucial to the quality/validity of the solution and optimization of the condition number of the G matrix can be increased by, for example, logarithmic sampling of pure mono-exponentials.

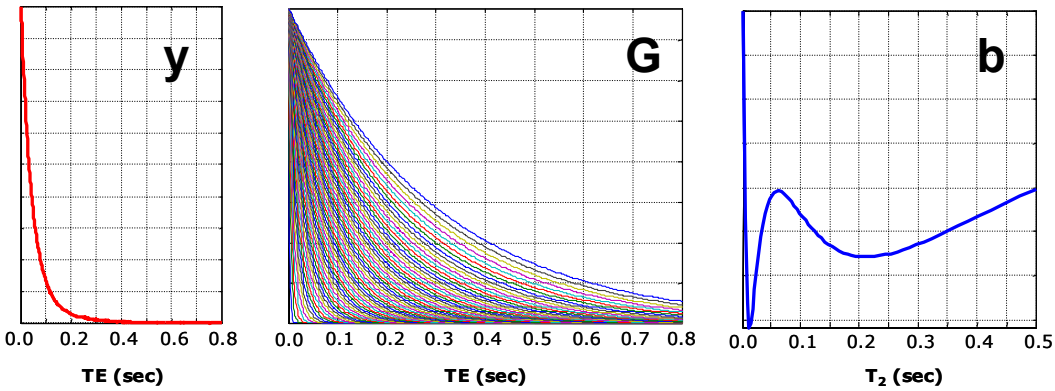


Figure 7 *Partial Least Squares Regression coefficients distribution. Non-successful 5-component PLS regression model fitting pure mono-exponentials to an experimental CPMG relaxation curve measured on fresh meat*

Despite this negative outcome, we have decided to show the result, since it includes some useful observations. Firstly, the resulting regression vector, **b**, has a peak at about 50 ms which is to be expected from a meat spectrum. In scrutinising the **b**-vector, a minor shoulder component can be observed just below 200 ms, also in good accordance with our TD-NMR relaxation understanding of meat. Secondly, although PLSR is not optimized to give interpretable regression coefficients (it focuses on the prediction of **y**), it is encouraging that the result gives a smooth T_2 distribution with reasonable trends, but clearly needs additional constraints such as non-negativity.

While distributed exponentials are often preferred to multi-exponential data analysis, because they theoretically reflect the underlying physics better, the existing algorithms are indeed most labile and the results retrieved strongly dependent on the numerical settings and choices required for the analysis. In addition, the distributed analysis is strongly dependent on proper phase correction of the data. Figure 8 displays the distribution analysis of meat cooking inside the TD-NMR magnet when analysed by standard software implementing the regularisation technique[21].

In this case the distributed analysis is relatively robust and consistent, as the sample remained in the magnet throughout the experiment and "time-slices" of T_2 distributions display a continuous development. Data has been pre-processed by phase rotation (A) and magnitude transformation (B), and the result of discrete fitting has been superimposed on the plots. The data presented turned out to be problematic to analyse using discrete multi-exponential fitting, as the system slowly changes with temperature from a two-component to a three-component system, much to the displeasure of the discrete fitting algorithm. The residuals using a two-component solution increase from around 40 ºC, but a valid three-component model cannot be calculated before 59 ºC has been reached (where the distribution of water is affected by the onset of collagen denaturation and lateral contraction of myofibrils, as described earlier in the paper). Thus, as the actual rank of the system is difficult to determine, the results of bi- and tri-exponential fitting overlap in the temperature interval 40-60 ºC in Figure 8. Since no assumptions are made regarding the number of components in the system of distributed exponentials this approach should presumably be advantageous in the case described.

The effect of pre-processing on the result of distributed exponentials is, however, obvious from Figure 8. The results from phase-rotated data and magnitude-transformed data are very different, the problem being that they may lead to different conclusions about the changes in the meat system upon cooking. According to the distributed fitting of the phase-rotated data (A), a third water population with slower relaxation characteristics than the water in raw meat is developed around 30 ºC. The relaxation time and magnitude of this population increases with temperature, which leads to the "easy" interpretation that the new population is water expelled from the myofibrils and the intermyofibrillar space to the outside of the meat matrix.

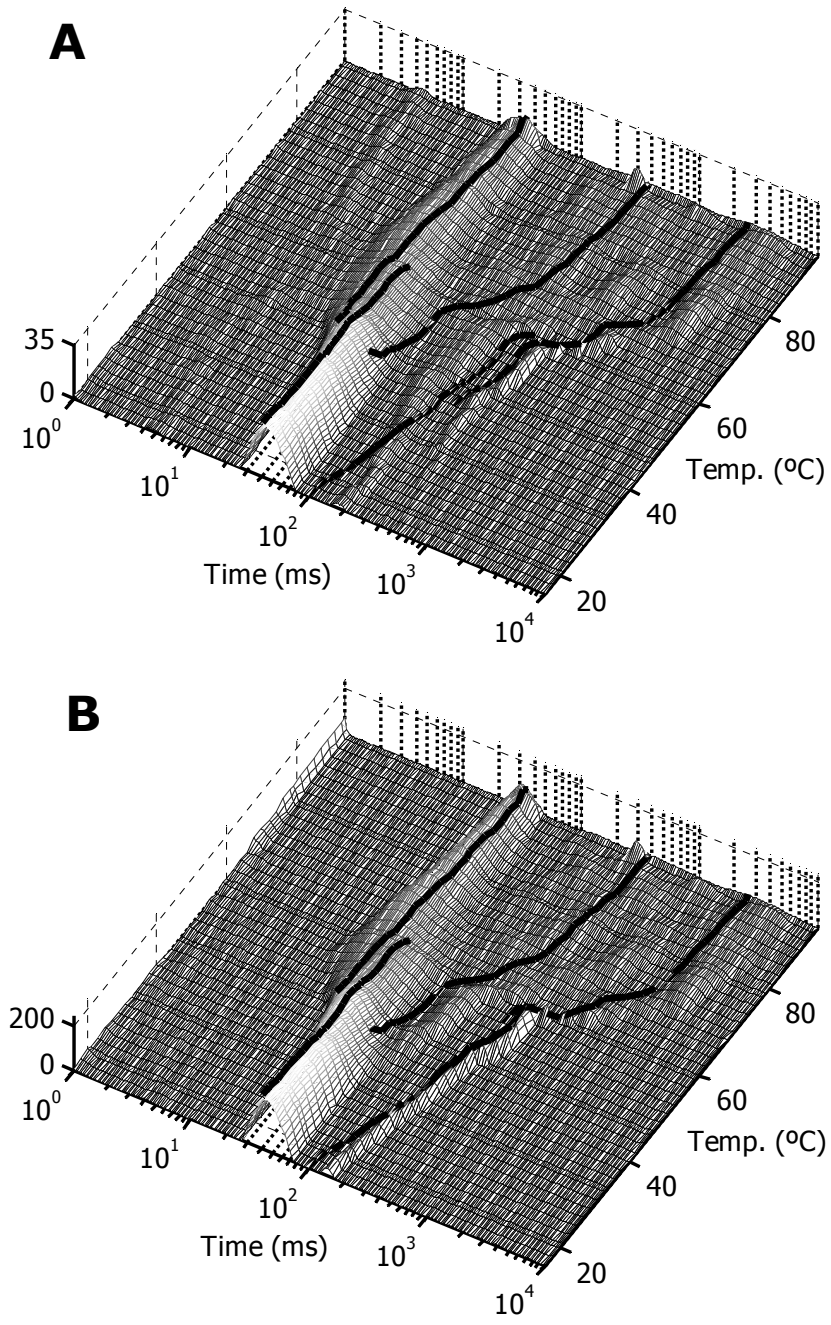


Figure 8 *Cooking meat inside the TD-NMR magnet[14]. Mesh plots of distributed exponentials calculated from the NMR transverse relaxation of meat as a function of cooking temperature*

The distribution analysis of the magnitude-transformed data, however, indicates that the third component emerges around 42 ºC with an intermediate relaxation time between the two original water populations. Thus, the water in the new water population developed is relatively free compared to strongly bound water within the myofibrils (T_{21} in raw meat), but it is isolated (on the NMR time scale) from the slowly relaxing water between the myofibrils (T_{22} in raw meat). As the new population develops when myosin denaturation starts at 42 ºC it is possible that the third water population represents water emerging in small compartments developed inside the myofibrils between denatured myosin units where it is trapped in a gel-like structure. Furthermore, it can be seen from

Figure 8 that the algorithm of distributed exponentials breaks down above 59 ºC when the proton relaxation is incomplete (the signal does not reach 0), whereas the discrete fitting using three components can be calculated.

Most surprisingly, the discrete fitting of the same data appeared almost independent of the pre-processing and displayed the best agreement with the distributed analysis of the magnitude transformed data. This fact has led us to a dual hypothesis concerning the new water component (compartment) arising during the meat-cooking experiment. Either the new water component with an intermediate average T_2 time arises due to the formation of a porous myosin gel or is it simply just expulsion from the myofibrillar lattice[14].

5 TRILINEAR CHEMOMETRIC MODELS

Where bi-linear chemometric models provide a tremendous advantage to univariate models, tri-linear data offers an additional unique resolution of underlying components in the mathematical sense[23]. Approximate tri-linear data follow the model:

$$x_{ijk} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} + e_{ijk}, \; i = 1,..,I; j = 1,..,J; k = 1,..,K$$

Tri-linear models extract common loadings for all samples, as is the case with bi-linear models, but they require a cube of data, in contrast to the table or matrix of data in the bi-linear case. Several different algorithms are available for tri-linear modelling including *Generalised Rank Annihilation Method* (GRAM)[24], *Direct Trilinear Decomposition* (DTLD)[25], which is a generalisation of GRAM, and *Parallel Factor Analysis* (PARAFAC)[26] all essentially fitting the model above. While the eigen-based algorithms GRAM and DTLD providing fast, analytical solutions, they do not give the least squares solution. However, for large data sets with high signal-to-noise ratio and little model error, the difference is often insignificant. PARAFAC[27] is an alternative iterative procedure that tends to be relatively slow, but has certain advantages, including a least squares solution to the problem under investigation. Providing that approximately tri-linear data can be measured, tri-linear models provide so-called computer chromatography which is a highly attractive alternative to actually performing physical chromatography. This has for instance been shown for fluorescence excitation-emission landscapes[28].

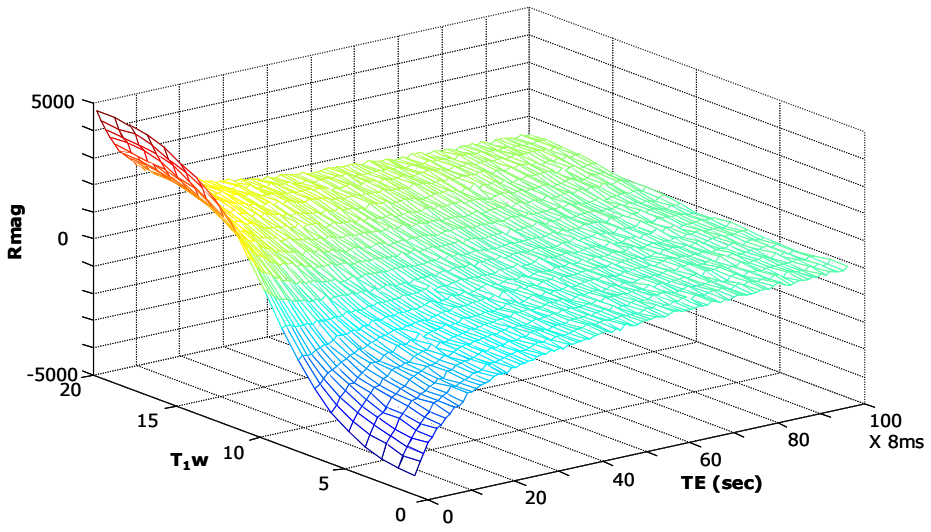


Figure 9 *2D TD-NMR. T_1-weigthed CPMG landscape of a butter sample*

Similar to fluorescence spectroscopy, it is possible to record TD-NMR data as a function of two or more variables. One such example is T_1-weighted CPMG data which have the dimension (samples × T_1 delay × T_2 relaxation). Figure 9 shows an example of a butter sample measured with a T_1 weighted CPMG pulse sequence[29].

Figure 10 shows the result of the PARAFAC decomposition of 32 samples of different spreads. The solution shows that the 2D TD-NMR data are approximately tri-linear, as loading **1** contains the *pure* CPMG profiles of three components and loading **2** contains the *pure* inversion recovery profiles. However, the third component of loading **2** has the shape of a compensation profile. In loading **3** (called the sample *score*) the concentrations of the three components of the spreads are obtained. These scores can be directly interpreted, as they require no further regression, but just scaling to represent concentrations. The scatter plot of the score of component **1** to a reference determination of total water content gave a direct correlation of r=-0.94. This preliminary investigation shows that it is indeed possible to generate approximate tri-linear data by TD-NMR and the results are most promising for future applications.

Another approach for generation of tri-linear data with TD-NMR is the so-called SLICING technique which has recently been described in detail by Pedersen *et al.*[10]. The novel approach is to upgrade a one-dimensional relaxation curve to become a pseudo two-dimensional structure and thus facilitate the unique advantages offered by trilinear models. The method is based on the fact that two different time "slices" of a given multi-exponential decay curve consist of the same underlying features (***quality:*** characteristic decay times), but in a new linearly related combination of amounts (***quantity:*** concentrations or magnitudes).

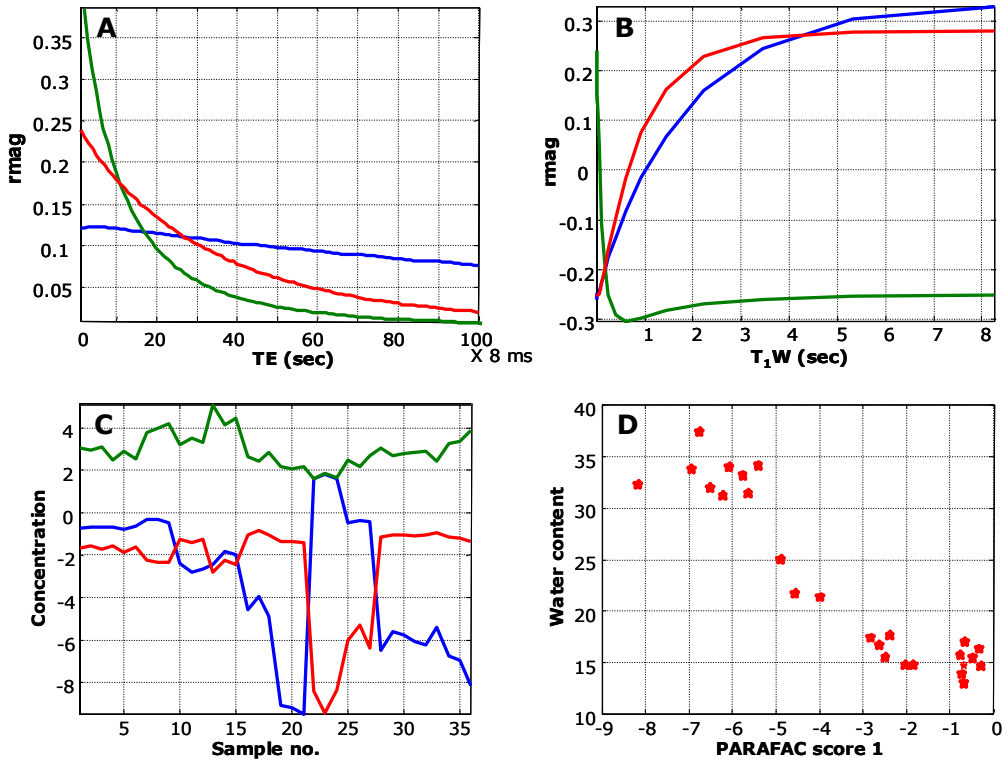


Figure 10 *Tri-linear TD-NMR data. A three-component PARAFAC solution to T_1-weighted CPMG data of 32 butters and spreads: A) loading **1** corresponding to the CPMG profiles , B) loading **2** corresponding to the T_1 profiles, C) loading **3** or scores corresponding to the concentrations, D) a simple scatter plot between water content of the samples and the score of component 1*

Windig and Antalek[30] originally conceived the idea and proposed a fast alternative to the tri-linear least squares solution, which they called Direct Exponential Curve Resolution Algorithm (DECRA). In other words, the new approach is based on the linear relationship between exponentials:

$$\exp(\frac{-\mathbf{t}}{T_{2n}}) \propto \exp(\frac{-\mathbf{t}+\Delta\mathbf{t}}{T_{2n}})$$

By a simple pre-operation, multi-exponential transverse relaxation curves can be pseudo upgraded to become tri-linear data structures. In the simplest case one relaxation curve can be translated one data point, called *lag* 1, and added in a new direction called *slab* (*slab* 2), creating a data matrix with the dimension two in the *slab* direction and the dimension N-1 in the lag direction. Hence, the major part of the relaxation curves will be nearly identical, but shifted "horizontally" by a fixed amount. The idea of "cutting" data into a number of overlapping slices has given rise to the name selected for this approach: SLICING. If this operation is performed on a series of samples, it is possible to obtain a tri-linear structure that can be analysed by, for example, PARAFAC. The result of this procedure will be exactly the same as that of a discrete exponential fitting algorithm, fitting common characteristic time constants to a series of samples[10]. Extensive simulations of TD-NMR data revealed that the SLICING approach was comparable, but not superior to a robust classical numerical approach. However the SLICING approach is very appealing, because this algorithm utilising highly redundant information and requiring no initial value guesses provides non-iterative and unique solutions with perfect mono-exponential loadings. In practice, the dramatically improved speed (independent of number of components extracted) and somewhat improved diagnostics (unique solutions) of this algorithm should be used as pre-processing (super-qualified initial guesses) to traditional numerical curve resolution algorithms.

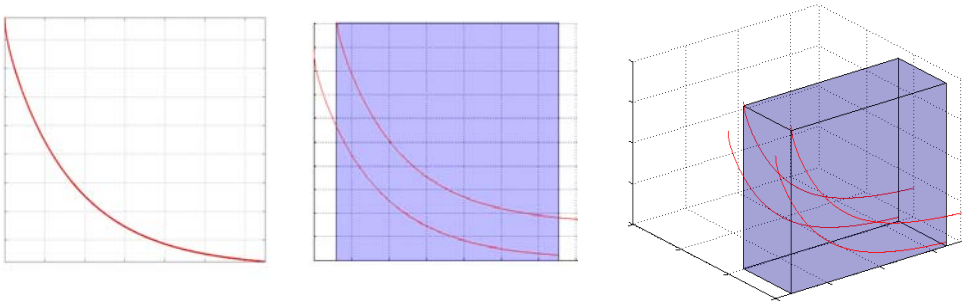


Figure 11 *DOUBLESLICING. The pseudo-upgrade of a single TD-NMR relaxation curve to become a tri-linear data structure*

It is possible to perform the SLICING pre-transformation twice (or more) and thus from a single relaxation curve generate a tri-linear data cube. We have recently conducted a series of preliminary experiments using this approach. The data pre-processing strategy called DOUBLESLICING provides a means to take advantage of tri-linear models when decomposing a single TD-NMR relaxation decay. Figure 11 shows in schematic form the simple principle behind DOUBLESLICING. In the first SLICING operation a number of *pseudo*-samples are created which all have the same underlying relaxation components, but in different relative amounts (see proportionality equation in the previous section). The number of *slabs* (*pseudo*-samples) must be equal to or larger than the number of components one wishes to extract. This *pseudo*-sample matrix is then sliced again using the SLICING pre-transformation to obtain a *pseudo* tri-linear upgrade of the *pseudo*-sample matrix. When this DOUBLESLICING is performed, the tri-linear data structure can be analysed using multi-way techniques such as PARAFAC.

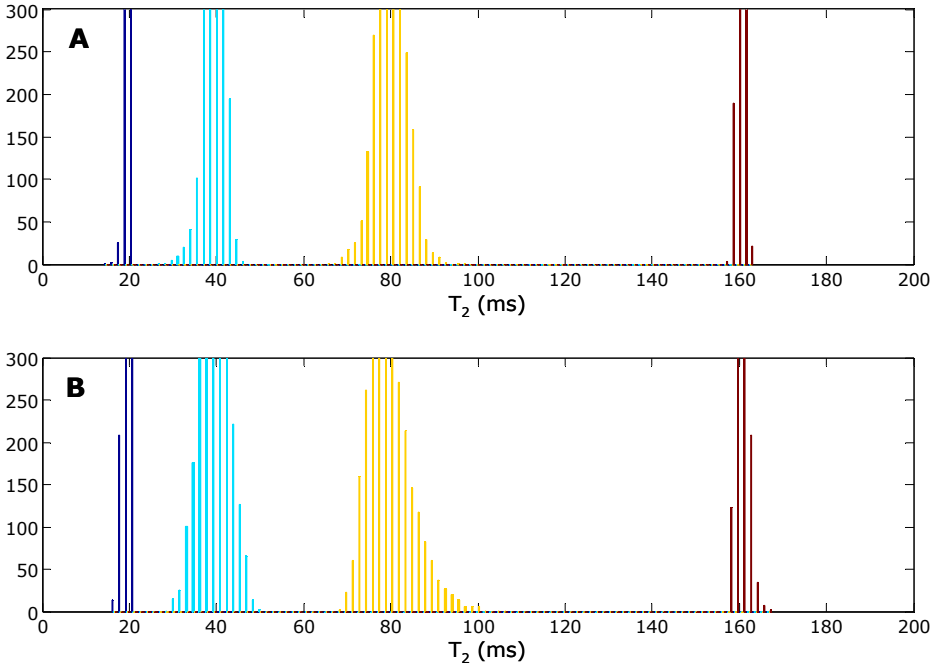


Figure 12 *Benchmark of the* DOUBLESLICING *approach on theoretical data. Model generation on 3000 replicates containing different (randomized) amounts of four T_2 components at 20, 40, 80 and 160 ms and added 0.1 % random noise. The dispersion of the solutions from the A) discrete exponential fit and from B)* DOUBLESLICING

Admitted, it may appear to be an unnessesary complication to create a 3-way data matrix with highly redundant information from a single relaxation curve to solve the simple discrete exponential fitting problem. However, with suitable resampling DOUBLESLICING produces practically the same results as traditional exponential fitting, but with significantly shorter computation times, especially when more than two components are extracted. Figure 12 shows an example of repeated deconvolution of a simulated data curve based on four exponential decays with random noise added. In this case DOUBLESLICING performs similarly to the exponential fitting, but with a computation time of less than one-fifth of the computation time of the exponential fitting. We iterate that the perspectives of this technique are primarily to be seen in cases where many oligo-exponential equations are to be solved or to serve as pre-processing (super-qualified and relatively un-biased initial guesses) to traditional numerical curve resolution algorithms.

6 CONCLUSIONS

The potential advantages of implementing spectroscopic sensors for quality control directly in the food process will create a continuous quest for still more informative and multivariate sensors to be developed. High-resolution nuclear magnetic resonance is probably the most successful and versatile spectroscopic technique yet developed and although its implementation as an on-line monitoring tool is severely hampered by the requirement of a strong homogeneous magnetic field, we foresee that this technique will also invade the more advanced segments of the food and medical industries for quality control.

The success of implementing NMR to process control is, however, dependent on qualified data analysis. With the use of chemometric techniques such as PCA, PLS and PARAFAC the multivariate nature of the NMR data can be fully utilized, increasing the stability of calibrations, making outlier detection possible and allowing for the measurement of multiple quality parameters simultaneously. We have demonstrated that TD-NMR has the capability to provide complex multivariate and multi-way information on food samples that allows application of tri-linear data analytical methods to recover pure analyte concentrations and to explore the co-variances with food quality.

In this paper we have illustrated the advantage of using multivariate data in the exploratory analysis of TD-NMR measurements on oil seeds and meat being cooked inside the instrument. We have presented an exotic use of PCA namely the Principal Phase Correction of quadrature data. We have demonstrated the advantage of using **multiway** multivariate TD-NMR data, for example by a pseudo upgrade of bi-linear transverse relaxation data to become tri-linear in an algorithm called SLICING. This method is much faster than traditional methods in extracting characteristic relaxation times and the method has for example great potential for **mathematical contrasting** of MRI images.

Perhaps most promising we have identified that T_1-weighted CPMG data can have approximate tri-linear structure and thus can be mathematically separated into its pure T_1 and T_2 components and the related concentration that require no further regression except for scaling to **one** know standard. If successful, further developments of this **mathematical chromatography** approach have the potential to separate components that are very close in characteristic relaxation times.

In this paper we have chosen to focus on the versatility of applying chemometrics to TD-NMR data. For this reason we have not mentioned the important issue of method validation here. However, the importance of validation - i.e. replicate measurements, re-sampling schemes and robust prediction techniques cannot be overestimated and is together with other important considerations such as the statistical properties, numerical stability and ease of interpretation, a continuous topic in our investigations.

Acknowledgements

This research was sponsored by the Danish Veterinary and Agricultural Research Council (SJVF), Advanced Quality Monitoring (AQM), and the Danish Centre for Advanced Food Studies (LMC). Gilda Kischinovsky is acknowledged for proofreading the manuscript.

References

1. C.F.Hazlewood, B.L.Chang, B.L.Nichols, and N.F.Chemberlain. *Nature*, 1969, **222,** 747
2. M.A.Borisova and E.F.Oreshkin. *Meat Sci.*, 1992, **31,** 257
3. H.C.Bertram, A.H.Karlsson, M.Rasmussen, O.D.Petersen, S.Dønstrup, and H.J.Andersen. *J. Agric. Food Chem.*, 2001, **49,** 3092
4. D.Johnels, U.Edlund, H.Grahn, S.Hellberg, M.Sjostrom, S.Wold, S.Clementi, and W.J.Dunn. *J. Chem. Soc. Perkin Trans. 2*, 1983, **863**
5. H.Grahn, F.Delaglio, M.A.Delsuc, and G.C.Levy. *J. Magn. Reson.*, 1988, **77,** 294
6. R.E.Hoffman and G.C.Levy. *Progr. NMR Spectroscopy*, 1991, **23,** 211
7. L.G.Thygesen. *Holzforschung*, 1996, **50,** 434

8. A.Gerbanowski, D.N.Rutledge, M.H.Feinberg, and C.J.Ducauze. *Sci. Aliment.*, 1997, **17,** 309
9. S.M.Jepsen, H.T.Pedersen, and S.B.Engelsen. *J. Sci. Food Agric.*, 1999, **79,** 1793
10. H.T.Pedersen, R.Bro, and S.B.Engelsen. *J. Magn. Reson.*, 2002, **157,** 141
11. K.Pearson. *Phil. Mag.*, 1901, **2,** 559
12. I.E.Bechmann, H.T.Pedersen, L.Nørgaard, and S.B.Engelsen. Comparative Chemometric Analysis of Transverse Low-field ^{1}H NMR Relaxation Data. *In* Advances in Magnetic Resonance in Food Science. *Edited by* P.S.Belton, B.P.Hills, and G.A.Webb. Cambridge, UK. 1999. p. 217.
13. H.T.Pedersen, L.Munck, and S.B.Engelsen. *J. Am. Oil Chem. Soc.*, 2000, **77,** 1069
14. E.Micklander, B.Peshlov, P.P.Purslow, and S.B.Engelsen. *Trends Food Sci. Technol.*, 2002, in press.
15. J.W.Sammon. *IEEE Transactions on Computers*, 1969, **18,** 401
16. E.L.Hahn. *Phys. Rev.*, 1950, **80,** 580
17. E.O.Stejskal and J.E.Tanner. *The Journal of Chemical Physics*, 1965, **42,** 288
18. J.E.Tanner and E.O.Stejskal. *The Journal of Chemical Physics*, 1968, **49,** 1768
19. C.Beauvallet and J.P.Renou. *Trends Food Sci. Technol.*, 1992, **3,** 241
20. H.T.Pedersen, H.Berg, F.Lundby, and S.B.Engelsen. *Innov. Food Science Emerg. Tech.*, 2001, **2,** 87
21. J.P.Butler, J.A.Reeds, and S.V.Dawson. *SIAM Journal of Numerical Analysis*, 1981, **18,** 381
22. S.W.Provencher. *Comp. Phys. Comm.*, 1982, **27,** 229
23. N.D.Sidiropoulos and R.Bro. *J. Chemom.*, 2000, **14,** 229
24. E.Sánchez and B.R.Kowalski. *Anal. Chem.*, 1986, **58,** 499
25. E.Sánchez and B.R.Kowalski. *J. Chemom.*, 1990, **4,** 29
26. R.A.Harshman. *UCLA working papers in phonetics*, 1970, **16,** 1
27. R.Bro. *Chemom. Intell. Lab. Syst.*, 1997, **38,** 149
28. R.Bro. *Chemom. Intell. Lab. Syst.*, 1999, **46,** 133
29. D.N.Rutledge and A.S.Barros. *Analyst*, 1998, **123,** 551
30. W.Windig and B.Antalek. *Chemom. Intell. Lab. Syst.*, 1997, **37,** 241

PAPER V

Quantification of Nitrogen Concentration in Perennial Ryegrass and Red Fescue Using Near-infrared Reflectance Spectroscopy (NIRS) and Chemometrics

R. Gislum, E. Micklander and J.P. Nielsen

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Field Crops Research 88 (2004) 269–277
Field Crops Research
www.elsevier.com/locate/fcr

Quantification of nitrogen concentration in perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and chemometrics

R. Gislum[a,*], E. Micklander[b], J.P. Nielsen[b]

[a]*Department of Plant Biology, Research Centre Flakkebjerg, Danish Institute of Agricultural Sciences, DK-4200 Slagelse, Denmark*
[b]*Food Technology, Chemometrics Group, Department of Dairy and Food Science, Centre for Advanced Food Studies, The Royal Veterinary and Agricultural University, DK-1958 Frederiksberg C, Denmark*

Received 10 September 2003; received in revised form 23 December 2003; accepted 10 January 2004

Abstract

Measurements of nitrogen (N) concentrations in plant samples are increasingly being used to support the development of novel N application strategies, which are based on actual plant N concentration as well as introduction of N budgets in agriculture. In order to meet the increasing demands for N measurements, the development of a fast and cheap, but still reliable technique is required. In the present study it was accordingly investigated whether near-infrared spectroscopy (NIRS) can be implemented for measurement of N concentration in grass samples. From 2000 to 2002 a total of 837 plant samples were collected from different field trials on 12 sampling sites in Denmark. The sample set consisted of 17 cultivars of red fescue (*Festuca rubra* L.) and perennial ryegrass (*Lolium perenne* L.) with a range in N concentration from 0.6 to 6.26% N. Visual-NIRS measurements (400–2498 nm) were performed on the dried, ground samples and plant N concentrations were measured using the Dumas method. Partial least squares regression models were developed on the near-infrared (NIR) spectra (1100–2498 nm) and the N concentrations in the dry grass samples with the aim of predicting the N concentration in samples not contained in the models. Models on raw and scatter corrected spectra gave root mean square error of prediction, RMSEP = 0.19−0.24% N and correlation coefficients, $R = 0.97-0.98$, when tested on an independent test set of samples from all harvest years, whereas models tested on samples from a harvest year not included in the calibration gave RMSEP = 0.23−0.35% N and $R = 0.95-0.99$. The prediction error is higher than the reproducibility of the Dumas method, but the NIRS method developed can still be used for measuring the N concentration in samples of perennial ryegrass and red fescue with sufficient precision and accuracy for practical use. Studies of the year effect showed that samples from more years needs to be included in the calibration data in order to increase the robustness of the model.

Keywords: Perennial ryegrass; Red fescue; Nitrogen; Near-infrared reflectance spectroscopy (NIRS); Chemometrics; Partial least squares regression (PLSR); Scatter correction

1. Introduction

Measurements of nitrogen (N) concentration in grass samples are an integral part of novel N application strategies in grass seed production (Sicard, 1995;

* Corresponding author. Tel.: +45-58-113300; fax: +45-58-113301.
E-mail address: rene.gislum@agrsci.dk (R. Gislum).

0378-4290/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2004.01.021

Rowarth et al., 1998), and the analyses of N concentrations in plant material are fundamental for making N budgets on farms and on a national level (Olsthoorn and Fong, 1998; Watson and Atkinson, 1999). The need for analyses of N concentration in grass samples has therefore increased in the laboratories working within these areas and if the novel N application strategies and N budgets are successfully implemented an increase in N determinations will follow in practical grass seed production.

Determination of N concentrations in grass samples using the routine methods such as Kjeldahl distillation (AOAC, 1990) or Dumas (Hansen, 1989) are widespread even though they are time-consuming and expensive. If the increasing demand for measurements of N concentrations in plant samples is to be met, a faster and cheaper, but still reliable and accurate technique is needed. Near-infrared reflectance spectroscopy (NIRS) is an obvious candidate by being fast, cheap and environmental friendly since no chemicals are used. NIRS has become a widely used method in analysis of a range of agricultural products, as extensively reviewed (Williams and Norris, 1987; Osborne et al., 1993). NIRS measurements of protein, water and carbohydrates are based on the absorption of light in the near-infrared region (1100–2500 nm) which corresponds to overtones and combinations of fundamental mid-infrared vibrational transitions of, in particular, C–H, O–H and N–H. In NIR spectra the different constituents have broad overlapping peaks, and therefore NIRS measurements have to be calibrated against samples with known chemical composition in order to extract the desired information, e.g. N concentration. In this study calibrations have been performed using partial least squares regression (PLSR) (Martens and Næs, 1993), which is a powerful chemometric method for multivariate calibration of data consisting of many variables like the visual-NIR spectra. One of the advantages of PLSR is that all spectral data measured are included in the calibration models, which means that all the information contained in the spectra is utilised.

NIRS calibration models have been successfully developed for analysis of plant N concentrations in fresh as well as dry grass samples grown in green houses (McClure et al., 2002), in grass silage (Park et al., 1998), in timothy (*Phleum pratense* L.) and meadow fescue (*Festuca pratense* L.) (Paul and Schoberlein, 1991) and from semiarid grasslands (Garcia-Ciudad et al., 1999). In order to obtain an NIRS method that can be used on future samples, samples covering the maximum variation in the predicted parameter (% N) must be included in the calibration set. Depending on the plant species, development stage, and organ, the plant N concentration required for optimal growth is in the range of 2–5% N of the plant dry weight (Marschner, 1995), decreasing during the growing season as a result of dilution (Greenwood et al., 1990, 1991). To obtain a great variation in plant N concentration sampling should therefore be performed during plant ontogeny. Including plants with different availability for N will help to secure that the extreme N concentrations are covered. Furthermore, samples covering other types of variation should be included in the calibration set so the resulting spectral variation can be modelled or corrected for. This will secure the robustness and wide applicability of the method as future predictions of N concentrations will be reliable in spite of the presence of interfering variations in the NIR spectra.

The objective of the present study was to investigate the possibilities of substituting the routine Dumas method for measurement of N concentration in perennial ryegrass and red fescue with NIRS. The robustness of the calibration models developed was tested by predicting the N concentration in samples from a year that was not included in the calibration model in order to simulate the use of the method on samples from future years.

2. Materials and methods

2.1. Materials

A total of 837 plant samples were collected from different field trials with various experimental designs on 12 sampling sites in Denmark in the period from 2000 to 2002 (Table 1). The purpose of the field trials was to test different N application strategies on seed yield. N application rates therefore varied from 0 to 60 kg N ha^{-1} applied in the autumn and from 0 to 200 kg N ha^{-1} applied at the initiation of spring growth or divided between the initiation of spring growth and stem elongation. N application rate was

Table 1
Total number of samples (*n*), mean and range of plant N concentration (% N) in all samples and within years and species

Parameter 1	Parameter 2	*n*	Mean (% N)	Range (% N)
All samples		837	2.81	0.60–6.26
Year	2000	428	2.70	1.27–6.05
	2001	311	2.95	1.60–5.73
	2002	98	2.81	0.60–6.26
Species	Red fescue	244	2.18	1.43–3.15
	Perennial ryegrass	593	3.06	0.60–6.26

dependent on grass species. The sample set consisted of 17 cultivars of red fescue (f) (*Festuca rubra* L.) and perennial ryegrass (p) (*Lolium perenne* L.). Plant samples were cut at soil surface in the period from the initiation of spring growth until approximately one month before seed harvest. After drying at 80 °C for 48 h or until the samples were completely dry, the samples were ground to pass through a 1 mm screen prior to determination of plant N concentration on aliquots of the samples using Dumas (Hansen, 1989).

2.2. VIS-NIRS measurements

Reflectance spectra of the ground grass samples were obtained using an NIRSystems 6500 (Foss NIRSystems, Silver spring, MD, USA). The samples were packed as uniformly as possible in mini sample cups with a depth of 1 cm and a diameter of 3.9 cm and measured using a spinning sample module. Spectra were collected in the visual and near infrared (VIS-NIR) range from 400 to 2498 nm with data collection at every 2 nm. Each sample was measured once, where the spectrum from a sample was an average of 16 sub-scans. The spectra are reported as $\log(1/R)$. Using this VIS-NIRS procedure, approximately 30 samples can be analysed per hour.

2.3. Multivariate data analysis

Principal component analysis (PCA) (Wold et al., 1987) and partial least squares regression (PLSR) (Martens and Næs, 1993) was performed using The Unscrambler version 7.6 SR-1 (CAMO A/S, Norway). Both raw NIR spectra, multiplicative signal corrected (MSC) NIR spectra (Geladi et al., 1985) and NIR spectra pre-processed by the extended inverted signal correction (EISC) method (Pedersen et al., 2002; Martens et al., 2003) were used. EISC was carried out using MatLab Version 6.5 (The Mathworks Inc., Natick, MA). The EISC is based on inverted scatter correction (Helland et al., 1995), which is the inverted version of MSC, with physical and chemical extensions (Martens and Stark, 1991). In this experiment only the extension with squared spectra is included in the EISC as this turned out to give the best and most simple scatter correction. This means that the spectra were not only corrected for additive and multiplicative scatter as in the traditional MSC method but also for squared terms of the spectra.

The high number of samples available allowed for true test set validation. Models that included all years were validated using a test set consisting of every fourth sample which were selected prior to development of the calibration model, but after the samples were sorted according to plant N concentration. Models developed to simulate prediction of future samples were calibrated on all samples from 2 years and validated on the samples from the third year. MSC, EISC and PLSR models were thus developed and internally cross-validated on a calibration set and afterwards applied and tested on an independent test set. The PLSR results are presented as correlation coefficients (R), root mean square error of cross-validation (RMSECV) (calibration models) and root mean square error of prediction (RMSEP) (test set validations). Other authors present the prediction performance of regression models as bias and standard error of performance (SEP) where bias is a measure of the accuracy of the model, and SEP expresses the precision of the results corrected for bias. The relationship between RMSEP, SEP and bias is: $\text{RMSEP}^2 \approx \text{SEP}^2 + \text{bias}^2$, hence if the bias is zero RMSEP = SEP (Esbensen, 2000).

3. Results

3.1. PCA on VIS-NIR spectra

VIS-NIR spectra (400–2498 nm) of all samples showed great differences in the visual region (400–800 nm) and an offset was seen throughout the spectra (Fig. 1A). These variations were reflected in a PCA on the raw spectra. A score plot of the two first principal components (PC) revealed a large group of samples (II) and two smaller outlying clusters to the lower left (I) and the lower right (III) (Fig. 1B). The two smaller clusters consisted of all the samples from two different sampling sites in two different years. Cluster I had a low absorption in the visual region indicating small amounts of chlorophyll. This was supported by a visual examination, which revealed that the samples were very pale and coarse (Fig. 1C, cluster I) compared to the samples in the main group (Fig. 1C, cluster II), whereas cluster III consisted of darker samples with a coarser structure (Fig. 1C, cluster III).

In a PCA on only the NIR region of the spectra from 1100 to 2498 nm (Fig. 2A) the pale samples (cluster I in Fig. 1B) joined the main group leaving only cluster III separate (Fig. 2B) underlining that cluster I mainly differed in colour. Scatter correction of the NIR spectra in Fig. 2A using the EISC method, clearly reduced the offset in the spectra accentuating the chemical bands (Fig. 2C). A PCA on the scatter corrected spectra resulted in one homogenous group of samples with the red fescue samples located as a subgroup (f) in the upper right part of the main group (Fig. 2D).

3.2. PLSR models on NIR spectra and N concentration

The correlation between the NIR spectra and the N concentrations was non-linear, but a linear relationship could be obtained by taking the square root of the N concentrations prior to PLSR modelling. The square root of the N concentrations was consequently used in all PLSR models.

The performance of PLSR models on raw and scatter corrected NIR spectra and N concentrations are compared in Table 2. An overall good calibration

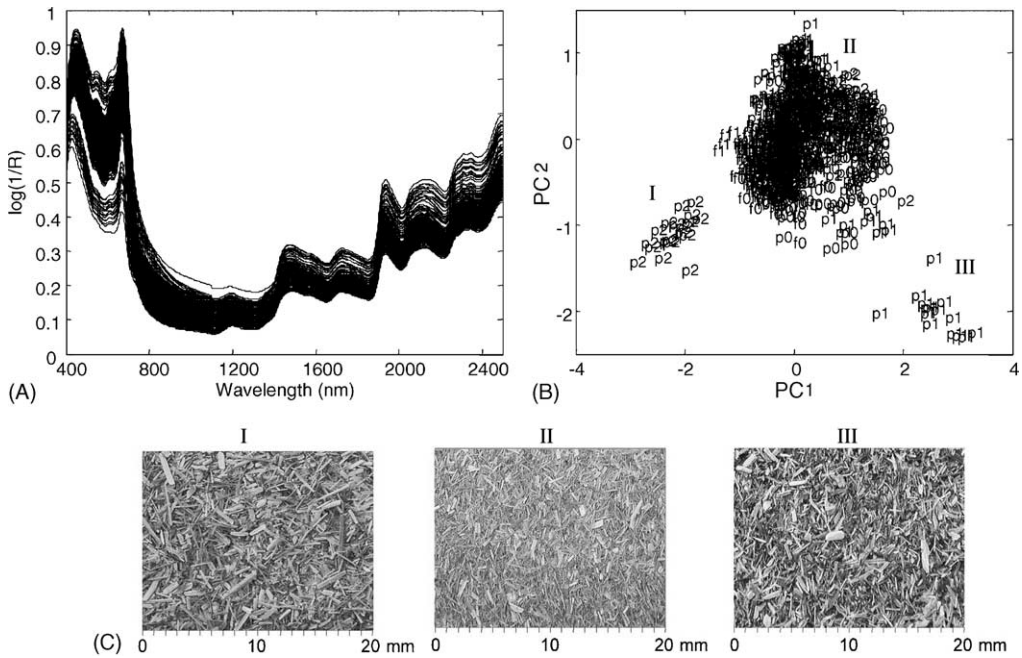


Fig. 1. Visual-NIR spectra (400–2498 nm) of all samples represented as (A) spectra and (B) PCA scores (PC 1 vs. PC 2). Perennial ryegrass samples are abbreviated as 'p' and red fescue samples as 'f'. Sampling year is abbreviated as follows: 2000 (0), 2001 (1) and 2002 (2). (C) Images of representative samples of the three clusters in B.

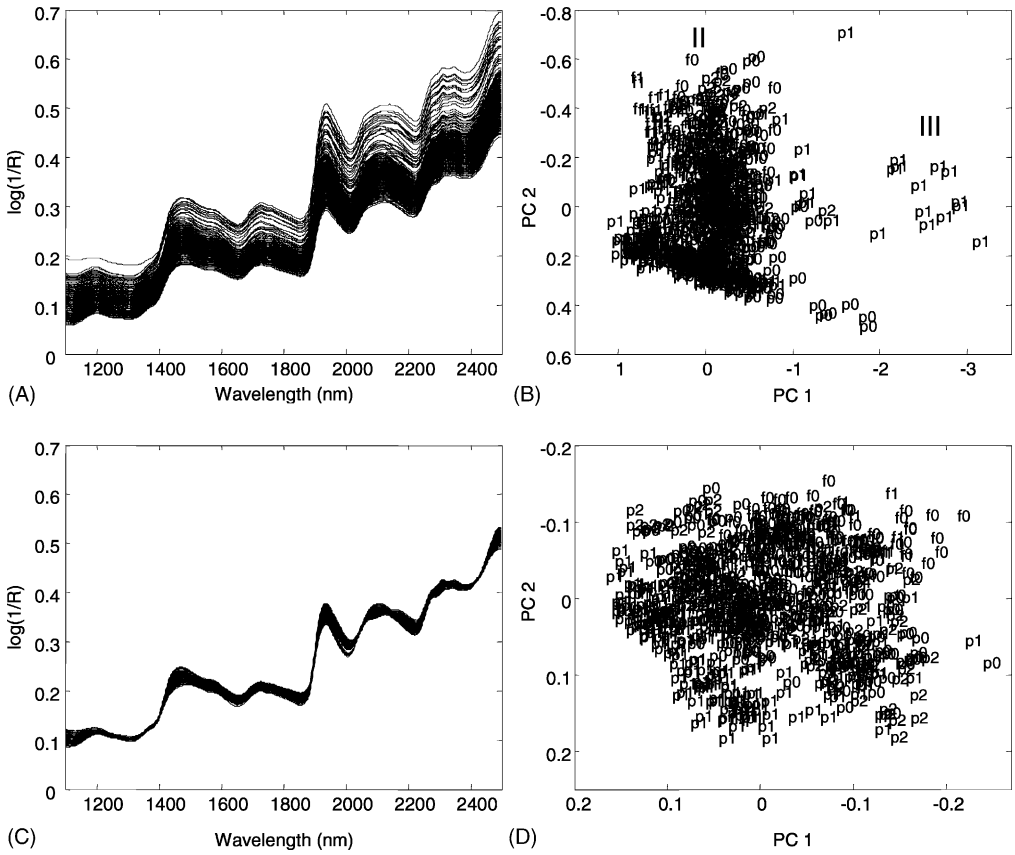


Fig. 2. NIR spectra (1100–2498 nm) of all samples represented as (A) spectra and (B) PCA scores (PC 1 vs. PC 2). NIR spectra of all samples scatter corrected using EISC represented as (C) spectra and (D) PCA scores (PC 1 vs. PC 2). Perennial ryegrass samples are abbreviated as 'p' and red fescue samples as 'f'. Sampling year is abbreviated as follows: 2000 (0), 2001 (1) and 2002 (2).

model could be achieved using four PLSR components on the raw NIR spectra of samples from all years with RMSECV = 0.18% N and $R = 0.98$ for the calibration and RMSEP = 0.19% N and $R = 0.98$ for the independent test set. The prediction error of the models developed on the raw spectra of samples from 2 years and validated on the third year was dependent on which year that was left out of the calibration model. RMSECV was 0.18–0.23% N, whereas the RMSEP increased to 0.27–0.35% N when the models were test set validated. Scatter correction of the spectra resulted in simpler models based on only one to two PLSR components for MSC spectra and two to three components for EISC spectra resulting in slightly higher RMSECV values, but RMSEP values in the same range as the models on the raw data. The biases of the predictions of selected samples were close to zero and the slopes close to unity, whereas the predictions of samples from individual years generally were heavy biased and the slopes ranged from 0.87 to 1.09.

Prediction of N concentrations in the two investigated species using the overall models gave a bias in the range −0.01 to 0.04% N and a slope of 0.90–1.02 for red fescue (64 validation samples) depending on pre-treatment and a bias in the range −0.01 to −0.04% N and a slope of 0.99 for perennial ryegrass (146 validation samples).

4. Discussion

The current sample set consisted of different species, cultivars, growth stages, sampling sites and years. This secured a large variation in plant N concentration,

Table 2
Calibration and validation statistics of PLSR models on NIR spectra (1100–2498 nm) for determination of N concentration in perennial ryegrass and red fescue applying different scatter corrections

Calibration						Test set validation						
Samples	*n*	Range (% N)	#PLSR comp	*R*	RMSECV (% N)	Samples	*n*	Range (% N)	*R*	RMSEP (% N)	Slope	Bias (% N)
Raw NIR spectra												
Sorted selection	627	0.60–6.26	4	0.98	0.18	Sorted selection	210	0.80–5.92	0.98	0.19	0.99	−0.01
2000 + 2001	739	1.27–6.10	4	0.98	0.18	2002	98	0.60–6.26	0.99	0.28	1.07	−0.18
2000 + 2002	526	0.60–6.26	4	0.98	0.18	2001	311	1.60–5.73	0.97	0.27	0.87	0.13
2001 + 2002	409	0.60–6.26	3	0.97	0.23	2000	428	1.27–6.05	0.96	0.35	1.01	0.21
MSC NIR spectra												
Sorted selection	627	0.60–6.26	1	0.97	0.24	Sorted selection	210	0.80–5.92	0.97	0.24	0.99	0.02
2000 + 2001	739	1.27–6.10	2	0.97	0.22	2002	98	0.60–6.26	0.99	0.28	1.02	−0.16
2000 + 2002	526	0.60–6.26	2	0.98	0.22	2001	311	1.60–5.73	0.97	0.23	0.91	0.03
2001 + 2002	409	0.60–6.26	2	0.98	0.20	2000	428	1.27–6.05	0.96	0.32	1.09	0.17
EISC NIR spectra												
Sorted selection	627	0.60–6.26	3	0.98	0.21	Sorted selection	210	0.80–5.92	0.97	0.23	1.00	−0.01
2000 + 2001	739	1.27–6.10	3	0.97	0.21	2002	98	0.60–6.26	0.99	0.23	1.00	−0.10
2000 + 2002	526	0.60–6.26	3	0.98	0.20	2001	311	1.60–5.73	0.97	0.23	0.95	0.06
2001 + 2002	409	0.60–6.26	2	0.97	0.25	2000	428	1.27–6.05	0.95	0.31	1.08	0.08

The table includes total number of samples used in the calibration or validation set (*n*), number of partial least squares regression components (#PLSR comp), regression coefficient (*R*), root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP), slope and bias.

but it also induced a great variation in colour and particle size and shape which was reflected in the PCA plot in Fig. 1B where clustering due to both physical and chemical variation was seen. The three clusters differed in average N concentration (cluster I: 0.96% N; cluster II: 2.80% N; cluster III: 4.30% N), however the main reason for the clustering was that the samples varied in colour and coarseness, which could be seen from the raw spectra and inspection of the samples. McClure et al. (2002) suggested that a decrease in the 672 nm chlorophyll band could be caused by drying the grass samples. In any case the absorbance in the visual region was not connected to N concentration. Accordingly, the visual part of the spectra was left out of the calibrations, as it would only contribute with noise to the PLSR models of N concentration.

The N concentrations were transformed prior to PLSR modelling as PLSR is a linear method which cannot handle non-linear data. Taking the square root of the N concentration was the optimal transformation when the PLSR models were developed on the samples from 2000 to 2002. However at present relatively few samples with extreme N concentrations are available when the large span in other variations is considered. The ends of the regression curve are therefore not well described and the choice of transformation is therefore made on a relatively small basis. When the NIRS grass database is extended the optimal transformation may turn out to be slightly different. Another approach would be to use a non-linear method like non-linear artificial neural network.

The RMSEP of the model on the NIR region of the spectra was 0.19–0.24% N depending on pre-treatment, when all years were represented in the calibration model. This corresponds to an average error of 6.8–8.5% relative on an average sample (2.81% N). The reproducibility of the Dumas method has been reported to be 0.1% N for concentrations below 0.7% N and 4% relative above 0.7% N (personal communication). Thus the average error of the NIRS method is higher than the error of the Dumas method. It was expected that the error of the NIRS method would be higher than the error of the Dumas method as the RMSEP of an NIRS method under most circumstances will be greater than the reproducibility of the reference method. Only if there is no noise in the NIR spectra

and there is no model error, the prediction error will be equal to the error of the reference method. In this study the noise and modelling errors are complemented by a sampling error, as the Dumas measurements and the NIRS measurements were not performed on exactly the same fraction of each sample.

An RMSEP in the range of 0.19–0.24% N for predictions of N concentrations in the range 0.60–6.26% N is an acceptable result as the NIRS method developed is sufficiently accurate and precise to be used instead of the Dumas method for determinations of N concentration when the N status in a field and N budgets are to be evaluated. In these situations where the plant N concentrations are measured at field level a lower precision and accuracy on the N measurements is acceptable as other sources of errors, such as the within field variation and sampling, may be much larger.

The model developed is superior to McClure et al. (2002) who obtained an SEP of 0.29% N for a PLSR model on NIR spectra of fescue grown in green houses. The variation in N concentration was comparative to the present study, but other sources of sample variation was smaller, as samples from a green house study are grown under controlled growing conditions. The fact that the model from the present study developed on samples from the field exhibits a lower prediction error demonstrates that a large number of samples can make up for increased sample variation and give a broad but still reliable model. It is furthermore promising that the current model could be applied on two grass species without problems as illustrated by the small biases on the prediction of the species individually.

More species can presumably be included in the model, however, it can not be expected that a global model for all grass species and cultivars can be made with the same prediction ability as was found for two species. SEP may increase due to increased sample variation and some species may introduce a considerably bias. If the latter is the case it will usually be preferable to develop local models on the deviating species. The work of Ruano-Ramos et al. (1999) illustrates well how prediction ability may decrease when more sample variation is introduced. They found an SEP of 0.073–0.098% N depending on pre-treatment of the NIR spectra when the N contents in 32 grass samples were predicted using a model calibrated on grass samples only. When the N contents of the same 32 grass samples were predicted using a PLSR model calibrated on grasses, legumes and forbs together the SEP was 0.88–1.37% N. Thus, the precision on the predicted N contents were lower when several types of plant material were contained in the model than when a local grass model was used. The bias of the grass samples was not increased when the total herbage model was used for prediction, thus the accuracy of the predictions was maintained.

The prediction errors for grass obtained in the work by Ruano-Ramos et al. (1999) were considerably lower than found in the present study. Measurements were performed on samples from semiarid grassland including different species, sampling sites and different topographical positions of the sampling sites. The work was extended with samples from several years by Garcia-Ciudad et al. (1999) where calibration models developed on botanical material from 1 year were validated on samples from subsequent years. This resulted in SEP values between 0.07 and 0.10% N and $R = 0.98$. The range in N concentration in the experiments was limited to 0.47–2.52% N. Thus the higher RMSEP obtained in the present study is presumably due to a significantly larger range in N concentration included in the models.

The high biases on the predictions of individual years show that variation specific for each year is present in the data. The predictions of N concentrations are affected by this year specific variation, which is why the RMSEP values of the models validated on individual years are higher than the prediction error of the model calibrated on all years. The fact that the models do not perform as well on extrapolated years as on years contained in the calibration show that they are not robust. This means that spectra of samples from other (future) years might contain variation not described by the model, which may affect the prediction of the N concentration. In order to increase the robustness, samples from more years need to be incorporated in the calibration models. As the number of samples in the calibration set increases and variation due to different growing conditions in different years are described better, the RMSEP should approach the RMSEP of the overall model. Even when a robust model has been reached the model should continuously be improved by including new samples to secure the robustness.

PCA on scatter corrected spectra indicated that pre-treatment of the spectra could compensate for the offset induced in the spectra due to differences in coarseness. Accordingly, simpler models based on fewer PLSR components could be obtained by scatter correction of the spectra. It usually gives more robust models when scatter is modelled separately than when the scatter is modelled as an integrated part of the PLSR model. This is also the case for the present data which is illustrated by the fact that when scatter correction is used RMSEP on the individual years is closer to the overall RMSEP. As a robust model is at least as important as an accurate and precise model, scatter corrected data are to be preferred to raw data.

In the present experiment the grass samples were dried before NIRS measurements. The main reason for this was that McClure et al. (2002) found that NIRS calibration models for N concentration developed on dry grass samples were more precise (SEP = 0.29% N) than models based on fresh grass samples (SEP = 0.97% N). The advantage of using dry grass samples is furthermore that the NIRS measurements do not need to be performed within a certain time after the sampling time, which is convenient when samples are collected from different sampling sites. The disadvantage is that pre-processing of the samples is incorporated in the method, thus future samples will have to be dried and ground. Extension of the method to on-line or ''in-the-field'' measurements will therefore require implementation of totally new calibrations on measurements of fresh samples. On-line measurements will, however, involve different measuring equipment and consequently new calibrations under all circumstances.

5. Conclusion

A method for measuring N concentration in perennial ryegrass and red fescue based on NIRS was developed. N concentrations in the range from 0.6 to 6.26% N could be predicted with an RMSEP of 0.19–0.24% N using PLSR models on raw and scatter corrected NIR spectra. However, samples from more years need to be included in the calibration data in order to increase the robustness of the models. The RMSEP corresponds to a higher measuring error than the reproducibility of the Dumas method, but the NIRS method developed is sufficiently accurate and precise to replace Dumas for evaluation of the plant N status in a field in practical seed production. Accordingly the method developed can be of great value for fast and cheap determinations of N concentration in the future, when novel N application strategies based on measurement of the plant N concentration are developed and when N budgets are introduced in grass seed production.

Acknowledgements

The second author wishes to thank the Danish Ministry of Science, Technology and Innovation for financial support through the centre contract: ''Centre for New Sensor Systems for the Measurement of Food Quality''. Lars Nørgaard is acknowledged for critical review of the manuscript.

References

AOAC, 1990. Nitrogen (Total) in Fertilizers: Kjeldahl Method, 15th ed., No. 978 02. AOAC, pp. 18–19.

Esbensen, K.H., 2000. Multivariate Data Analysis–In Practice, 4th ed. CAMO, Oslo, 598 pp.

Garcia-Ciudad, A., Ruano, A., Becerro, F., Zabalgogeazcoa, I., Vazquez de Aldana, B.R., Garcia, C.B., 1999. Assessment of the potential of NIR spectroscopy for the estimation of nitrogen content in grasses from semiarid grasslands. Anim. Feed Sci. Technol. 77, 91–98.

Geladi, P., McDougall, D., Martens, H., 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. Appl. Spectrosc. 39, 491–500.

Greenwood, D.J., Gastal, F., Lemaire, G., Draycott, A., Millard, P., Neeteson, J.J., 1991. Growth rate and % N of field grown crops: theory and experiments. Ann. Bot. 67, 181–190.

Greenwood, D.J., Lemaire, G., Gosse, G., Cruz, P., Draycott, A., Neeteson, J.J., 1990. Decline in percentage N of C3 and C4 crops with increasing plant mass. Ann. Bot. 66, 425–436.

Hansen, B., 1989. Determination of nitrogen as elementary N, an alternative to Kjeldahl. Acta Agric. Scand. 39, 113–118.

Helland, I., Næs, T., Isaksson, T., 1995. Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data. Chemom. Intell. Lab. Syst. 29, 233–241.

Marschner, H., 1995. Mineral Nutrition of Higher Plants, 2nd ed. Academic Press, London, 889 pp.

Martens, H., Nielsen, J.P., Engelsen, S.B., 2003. Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. Anal. Chem. 75 (3), 394–404.

Martens, H., Næs, T., 1993. Multivariate Calibration. Wiley, New York, 419 pp.

Martens, H., Stark, E., 1991. Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. J. Pharm. Biomed. Anal. 9, 625–635.

McClure, W.F., Crowell, B., Stanfield, D.L., Mohapatra, S., Morimoto, S., Batten, G., 2002. Near infrared technology for precision environmental measurements. Part 1. Determination of nitrogen in green-and dry-grass tissue. J. Near Infrared Spectrosc. 10, 177–185.

Olsthoorn, C.S.M., Fong, N.P.K., 1998. The anthropogenic nitrogen cycle in The Netherlands. Nutr. Cycl. Agroecosyst. 52, 269–276.

Osborne, B., Fearn, T., Hindle, P.H., 1993. Practical NIR Spectroscopy with Applications in Food and Beverage Analysis, 2nd ed. Longman Scientific and Technical, Harlow, UK, 227 pp.

Park, R.S., Agnew, R.E., Gordon, F.J., Steen, R.W.J., 1998. The use of near infrared reflectance spectroscopy (NIRS) on undried samples of grass silage to predict chemical composition and digestibility parameters. Anim. Feed Sci. Technol. 72, 155–167.

Paul, C., Schoberlein, W., 1991. Using near infra-red reflectance spectroscopy (NIRS) to monitor the N supply situation in grass seed multiplication stands. Wirtschaftseigene Futter 37, 201–217.

Pedersen, D.K., Martens, H., Nielsen, J.P., Engelsen, S.B., 2002. Near-infrared absorption and scattering separated by extended inverted signal correction (EISC): analysis of near-infrared transmittance spectra of single wheat seeds. Appl. Spectrosc. 56 (9), 1206–1214.

Rowarth, J.S., Boelt, B., Hampton, J.G., Marshall, A.H., Rolston, M.P., Sicard, G., Silberstein, T.B., Sedcole, J.R., Young, W.C., 1998. The relationship between applied nitrogen, nitrogen concentration in herbage and seed yield in perennial ryegrass (*Lolium perenne* L.). I. Cv. Grasslands Nui at five sites around the globe. J. Appl. Seed Prod. 16, 105–114.

Ruano-Ramos, A., Garcia-Ciudad, A., Garcia-Criado, B., 1999. Near infrared spectroscopy prediction of mineral content in botanical fractions from semi-arid grasslands. Anim. Feed Sci. Technol. 77, 331–343.

Sicard, G., 1995. Nitrogen fertilization, nitrogen uptake and seed yield in perennial ryegrass. In: Proceedings of the Third International Herbage Seed Conference, Halle, Germany, 18–23 June 1995.

Watson, C.A., Atkinson, D., 1999. Using nitrogen budgets to indicate nitrogen use efficiency and losses from whole farm systems: a comparison of three methodological approaches. Nutr. Cycl. Agroecosyst. 53, 259–267.

Williams, P., Norris, K., 1987. Near Infrared Technology in the Agricultural and Food Industries, 2nd ed. American Association of Cereal Chemists, Minnesota, USA, 296 pp.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemom. Intell. Lab. Syst. 2, 37–52.

PAPER VI

Testing of an On-line Ultrasound Method for Early Discrimination of High and Low Water-holding Capacity in Pig Carcasses – A Study Including LF-NMR

E. Micklander, H.C. Bertram, H. Marnø, L.S. Bak, H.J. Andersen,
S.B. Engelsen and L. Nørgaard

Testing of an on-line ultrasound method for early discrimination of high and low water-holding capacity in pig carcasses – a study including LF-NMR

E. Micklander[a*], H.C. Bertram[b], H. Marnø[c], L.S. Bak[d], H.J. Andersen[b], S.B. Engelsen[a], L. Nørgaard[a]

[a] *Centre for Advanced Food Studies, Food Technology, Department of Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark*

[b] *Department of Food Science, Danish Institute of Agricultural Sciences, Research Centre Foulum, PO Box 50, DK-8830 Tjele, Denmark*

[c] *SFK Technology A/S, Transformervej 9, DK-2730 Herlev, Denmark*

[d] *Danish Technological Institute, Division of Applied Food Technology, Holbergsvej 10, DK-6000 Kolding, Denmark*

[*] *Corresponding author*

Abstract

A newly developed ultrasound equipment was tested for on-line determination of water-holding capacity (WHC) early post mortem (p.m.). The experiment was based on measurements of muscle from ten pigs, five of which were treated to result in pork with a high WHC and five to result in pork with a low WHC. Ultrasound measurements were carried out on the carcass of each pig every tenth minute from 15 to 85 min p.m. at three locations of the M. longissimus dorsi (LD) and low field nuclear magnetic resonance (NMR) measurements were concomitantly performed on an excised sample. Drip loss was measured by Honikel's bag method and NMR measurements were carried out 24 h p.m. The ultrasound experiment demonstrated that classification of meat with high and low WHC was possible, but only at 85 min p.m. and only by measurements obtained at the posterior end of the LD. NMR measurements underlined that there are structural differences in the meat according to pre-slaughter treatment prior to 85 min p.m., however, the properties that determine the WHC are apparently not expressed sufficiently to be registered by ultrasound before 85 min p.m.

Keywords: Water-holding capacity; ultrasound; on-line, NMR relaxation

1. Introduction

Superior water-holding capacity (WHC) is one of the most important quality parameters of meat primarily because of reduced weight loss during cutting and storage and improved ability of the meat to retain water during processing. Furthermore the WHC may influence the eating quality of the meat (Lawrie, 1998). The WHC of meat is amongst other things

determined by the progression in temperature and pH in the carcass post mortem (p.m.) (Schäfer, Rosenvold, Purslow, Andersen & Henckel, 2002). A high temperature combined with a fast decrease in pH leads to increased myofibrillar shrinkage and protein denaturation, resulting in low WHC (Wismer-Pedersen, 1959; Offer & Trinick, 1983; Offer, 1991).

In order to adjust the production according to the quality of the individual carcasses a classification of the carcasses based on WHC early post mortem is desired. Measurements of pH at 45 min p.m. (pH_{45}) have commonly been used to obtain an early indication of the final meat quality (Bendall & Swatland, 1995). However, while pH_{45} can be used for detecting pale, soft and exudative (PSE) and dark, firm and dry (DFD) meat, several studies have shown that the general correlation between pH_{45} and WHC is poor in pork of "normal" quality (Warris & Brown, 1987; Kauffman et al., 1993; Garrido, Pedauyé, Bañón & Laencina, 1993). Consequently, much effort has been devoted to finding alternative methods suitable for on-line determination of WHC early p.m. For example, methods based on conductivity and internal light scattering and absorption have been developed for early detection of PSE meat, as reviewed by Swatland (1995). However, studies including several different techniques have shown that none of the available methods for measurements early p.m. are usable for WHC classification of individual carcasses (Kauffman et al., 1993; van Oeckel, Warnants & Boucqué, 1999). Multivariate vibrational spectroscopy seems to be the most promising technique tested thus far. In a study in a research slaughterhouse Fourier transform infrared spectroscopy (FT-IR) carried out at-line 45 min p.m. gave a correlation of 0.89 with drip loss and a validation of the method under industrial conditions gave a correlation of 0.79 (Pedersen, Morel, Andersen & Engelsen, 2003). Equally good correlations were obtained using near infrared spectroscopy 30 min p.m., but the prediction error of the WHC was higher (Forrest, Morgan, Borggaard, Rasmussen, Jespersen & Andersen, 2000). Implementation of the FT-IR on-line method will, however, require development of instrumentation that can handle the humid conditions in slaughter plants, as the technique at present is vulnerable to water.

Ultrasound is a spectroscopic method which is better suited for non-invasive on-line measurements under humid conditions. Accordingly, ultrasound equipment for on-line measurements of fat and lean meat content has been developed (Brøndum, Egebo, Agerskov & Busk, 1998). However, the possibilities of using on-line ultrasound equipment for the measurement of WHC have not yet been investigated. Ultrasound spectroscopy uses longitudinal mechanical waves which have frequencies that are too high to be detected by the human ear (> 20 kHz). Ultrasound requires an elastic medium and the wave propagation velocity depends on the medium, which is entirely in contrast to electromagnetic waves (e.g. near infrared) that require no medium and have almost constant velocity (speed of light). In the pulse-echo technique, a pulse of an appropriate frequency, duration and

amplitude is applied to a sample through a transducer. It propagates through the sample until it reaches a change of material, at which point it is partly reflected and partly transmitted, and the echoes returned to the ultrasound transducer are measured as a function of time. Sound propagating through a carcass encounters several transitions between tissues, and the signal collected is therefore relatively complex. Chemometric data analytical methods like principal component analysis (PCA) (Hotelling, 1933) are well-suited data analytical tools for extraction of information from such multivariate complex data. Despite the multivariate nature, the chemometric methods cannot be used directly on raw ultrasound data, because the harmonics in ultrasound are contracted or expanded, representing a shifted pattern for a bilinear model like PCA. Therefore, the oscillating signal has to be transformed before data analysis, for example, by calculation of frequency spectra using Fourier transformation. Frequency analysis using Fourier transformation has been successful in laboratory studies on excised muscles for determination of intramuscular fat and sensory attributes of (Whittaker et al., 1992; Park et al., 1994).

Low field nuclear magnetic resonance (NMR) is another spectroscopic method which has potential for measurement of meat quality (Trout, 1988; Micklander, Peshlov, Purslow & Engelsen, 2002). NMR requires an external magnetic field and existing NMR equipment is not yet suitable for on-line measurements (Hills, 1998); however, laboratory studies have shown strong correlations between transverse relaxation, T_2, and WHC. Correlations of 0.74-0.75 between drip loss and entire relaxation decays from Carr-Purcell-Meibom-Gill (CPMG) experiments obtained 24-54 h p.m. have been obtained using partial least squares regression (PLSR) (Brown, Capozzi, Cavani, Cremonini, Petracci & Placucci, 2000; Brøndum, Munck, Henckel, Karlsson, Tornberg & Engelsen, 2000). Multivariate data analysis of the NMR data is well suited for predictive and explorative purposes, but interpretation of the NMR data is enhanced when the T_2 relaxation is analysed as a sum of exponential decays corresponding to different water components. There is some dispute with regard to the total number of water components in meat, but it is generally agreed that there are two major water components with relaxation constants around 35-50 ms and 100-250 ms which may be designated to either intra-cellular and extra-cellular water (Tornberg, Andersson, Göransson & von Seth, 1993) or intra-myofibrillar and extra-myofibrillar water (Bertram, Karlsson, Rasmussen, Pedersen, Dønstrup & Andersen, 2001b). Continuous NMR measurements early p.m. have shown that differences in the pH development are reflected in the T_{21} and T_{22} times and populations within a few hours p.m. (Tornberg, Wahlgren, Brøndum & Engelsen, 2000; Bertram, Whittaker, Karlsson, & Andersen, 2003.)

In the present study it was investigated whether ultrasonic measurements using a modified version of the on-line instrument Ultrafom-300 (SFK Technology A/S, Herlev, Denmark) could be used for early discrimination of low and high WHC in pigs. NMR measurements

were carried out concomitantly to the ultrasound measurements to provide information about the progression in the state of water in the samples.

2. Materials and methods

2.1. Animals, treatment

Ten pigs, five pairs of female littermates, of the crossbreed between Danish Duroc boars and Danish Landrace x Large White sows were used in the study, which was carried out in the experimental abattoir at Research Centre Foulum, Denmark. The animals had an average live weight of 93 kg. They were exposed to two different pre-slaughter treatments based on adrenaline injections and exercise combined with electrical stunning, which are known to give rise to two very different pH developments in the muscle post mortem (Henckel, Karlsson, Oksbjerg & Petersen, 2000) and hereby a large span in drip loss from the obtained pork. Five animals (one from each pair of littermates) were given a subcutaneous injection of 0.2 mg adrenaline/kg live weight in the evening prior to the day of slaughter and were stunned by 80% CO_2 for 3 minutes prior to exsanguination. The other five animals were exercised on a treadmill for approximately 20 minutes prior to electrical stunning (220 V) and subsequently exsanguinated. Cleaning, evisceration, splitting and weighing of the carcasses was completed within approximately 30 min p.m. after which the carcasses were transferred to a room with a temperature of 12ºC. At 60 min p.m. they were placed in a storage room at 4ºC.

2.2. Ultrasound measurements

Ultrasound measurements were performed with a modified version of an Ultrafom-300 (SFK Technology A/S, Herlev, Denmark). The instrument is developed for on-line measurement of fat and muscle thickness and consists of a handheld measuring pistol with an array of 64 transducers covering an area of approximately 4.4 cm in width. All transducers in the test equipment were using a frequency of 3.1 MHz. In the present study the signal from the three outermost transducers on each side of the array were discarded in order to eliminate noise due to edge errors. The measurements were performed on the skin of the back of the carcasses corresponding to the anterior, the centre, and the posterior of the right LD (approximately at rib number 5, 10, and 17). The backs of the carcasses were not singed, since this has shown to result in air bubbles in the skin and hereby unreliable ultrasound measurements, as ultrasound cannot penetrate air. Measurements were performed at 15, 25, 35, 45, 55, 65, 75 and 85 min p.m.

2.3. pH measurements

pH was measured at the posterior of the left LD at the time of exsanguination and at the centre of the right LD at 15, 25, 35, 45, 55, 65, 75, 85 min p.m. and at 24 h p.m. A pH-

meter (Metrohm AG CH 9101 Herisau, Switzerland) equipped with an insertion glass electrode (LL glass electrode, Methrom) was used. At the 15, 25, 35, 45, 55, 65, 75, 85 min p.m. measurements the electrode was calibrated at a temperature of 35ºC, while at 24 h p.m. the calibration temperature used was 4ºC. A two-point calibration was performed with calibration buffers of pH 7.000 and 4.005 at 25ºC.

2.4. NMR measurements

A sample from the centre of the left LD was cut out after exsanguination and immediately transported to the laboratory for NMR measurements. A Maran benchtop pulsed NMR Analyser (Resonance Instruments, Witney, U.K.) operating at 23.2 MHz and equipped with an 18 mm variable temperature probe head was used. A sub-sample of approximately 3.5 cm and 1 x 1 cm in cross-sectional area was cut out along the fibres using a scalpel and placed in a cylindrical glass tube with the fibre direction parallel to the tube wall in the instrument. The tube was sealed with a lid and placed in the NMR instrument inside an NMR test tube. The temperature of the instrument was set to follow the temperature decrease in a carcass undergoing the slaughtering and cooling process in the experimental abattoir. CPMG experiments were performed at 15, 25, 35, 45, 55, 65, 75 and 85 minutes p.m., corresponding to the temperatures 34.2, 34.2, 33.2, 32.3, 31.4, 30.4, 29.5, and 28.5ºC, using relaxation delay = 2 s, number of scans = 16, tau = 150 μs, number of echoes = 4096 (only the even-numbered echoes were used in the data analysis). At 24 h p.m. a sample was taken from the centre of the right LD (5 cm below where ultrasound measurement had been performed) for replicate CPMG measurements at 25º C using the same experimental parameters as for the other measurements.

2.5. Water-holding capacity

Measurement of drip loss was performed using the Honikel bag method (Honikel, 1998). Samples from the centre of the right LD were taken 24 hours p.m. and the drip loss was measured after the samples had been hanging in nets in inflated plastic bags for 48 h at 4º C. Measurements were carried out in replicate and the results are expressed as percentage weight loss.

2.6. Intramuscular fat

Intramuscular fat (IMF) was measured on samples from the posterior of the left LD according to Nordic Committee on Food Analysis (NMKL) method No. 160, 1998 (NMKL, 1998). The samples were stored at -18ºC prior to measurements.

2.7. Data analysis

PCA was carried out in MatLab version 6.5 (The MathWorks Inc., Natick, MA, USA) using the PLS toolbox version 2.0.1c (Eigenvector Research Inc., Manson, WA, USA). Discrete and distributed exponential fitting of CPMG decay curves were performed in MatLab using in-house scripts. The NMR data were phase rotated using principal phase correction (Pedersen, Bro & Engelsen, 2002) and the first five points of the CPMG curves were removed from data prior to exponential fitting, as this gave more robust results. The remaining data were maximum normalized.

3. Results

Figure 1 displays the pH development and the final pH at 24 h p.m., which differed significantly between the adrenaline-administered pigs and the pigs that were exercised and electrically stunned. In addition, the two treatments resulted in significantly different WHC of the LD ($P<0.001$), with a mean drip loss from the adrenaline pigs of 0.75% (std. dev. 0.12%) and a mean drip loss of 13.18% (std. dev 1.43%) from the treadmill pigs.

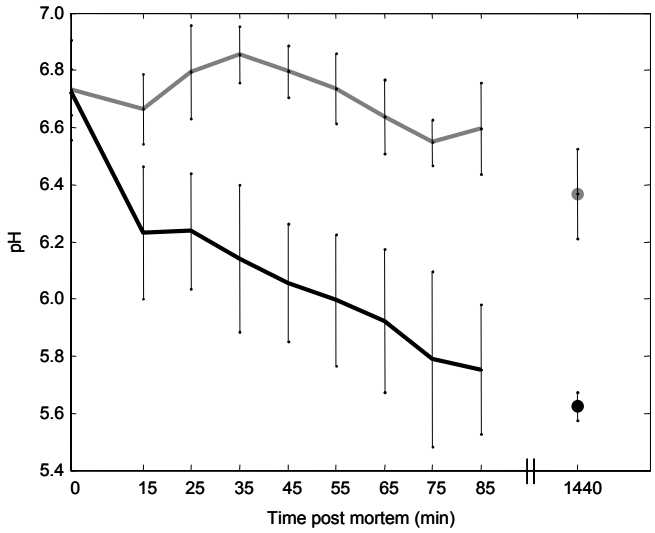


Fig. 1. pH development in adrenaline pigs (grey) and treadmill pigs (black) 0-24 h p.m. The curves represent averages of five pigs. Error bars indicate ± standard deviation.

A typical ultrasound signal from LD is shown in Fig. 2A. In the signal from one transducer, transitions between tissues appear as highly positive or negative intensities of the signal. The spatial understanding of the ultrasound data is enhanced when the spectra from the 58 transducers are looked upon as an image (Fig. 2B). Transitions between different types of tissues are seen as bands with extreme pixel values (white or black pixels). The first peak is

due to reflectance of the ultrasound as it enters the sample. After that follows a broad band corresponding to transitions between the different fat layers. The subsequent relatively dark area with smaller signal amplitudes is the meat followed by the transition to the ribs/the membrane between ribs and the final peak corresponds to the transition to air.

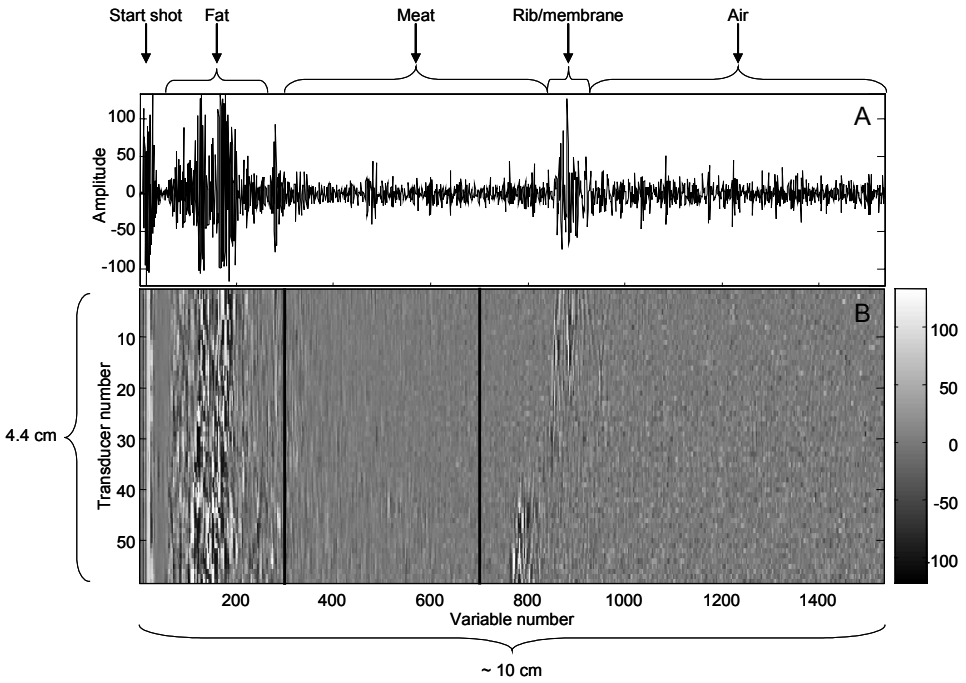


Fig. 2. Ultrasound data from A) one transducer B) 58 transducers making up an image. Width of transducer array and approximate travelling distance of the ultrasound are indicated. The area between the two vertical lines in 2B indicates the part of the images which has been used for data analysis (variables 301-700).

Histograms of the pixel values in the meat part (variables 301-700, Fig. 2B) of the images obtained at the posterior of the LD 85 min p.m. show a clear difference between adrenaline and treadmill pigs (Fig. 3D). A similar separation was not obtained for the images obtained earlier p.m., but a comparison between the histograms from 15, 65, 75, and 85 min p.m. (Fig. 3A-D) shows the development towards better separation with time during the last 20 minutes of measurement.

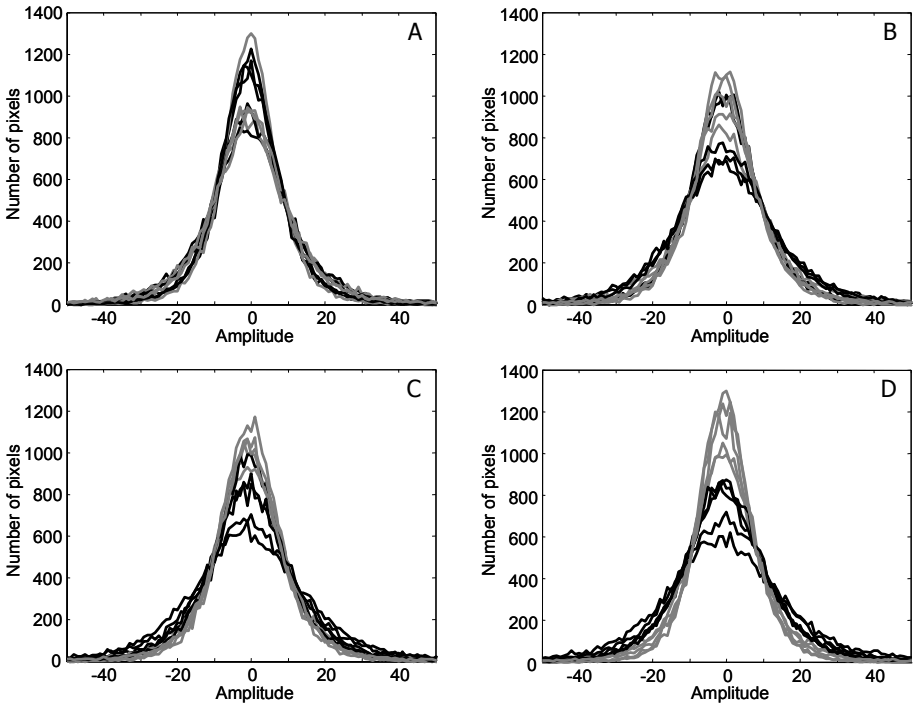


Fig. 3. Histograms of pixel values in the meat part (variables 301-700) of ultrasound images from five adrenaline pigs (grey) and five treadmill pigs (black) obtained at the posterior of the LD at A) 15 min p.m., B) 65 min p.m., C) 75 min p.m., and D) 85 min p.m. Values varied between -150 and 150, only values between -50 and 50 are shown.

Bi-exponential fitting of the NMR relaxation profiles resulted in a fast relaxing component with a relaxation time, T_{21}, of 35-50 ms (Fig. 4A) and a slow relaxing component with a relaxation time, T_{22}, in the range of 100-200 ms (Fig. 4B). The residuals from discrete exponential fitting of the CPMG curves showed that the data could best be described by a two-component system. At all measurements from 15 min to 85 min p.m., the T_{21} time constant was higher in muscles from the treadmill pigs. The proportion of protons in the fast relaxing population, M_{21}, made up approximately 95% of the water 15 min p.m., which slightly decreased to 93% at 85 min p.m. (Fig. 4C). Accordingly, M_{22} made up 5% 15 min p.m. and increased to 7% 85 min p.m. (Fig. 4D). The situation is completely reversed at 24 hours p.m. where the T_{21} time constant is longer in pork from the adrenaline-administered pigs (Fig. 4A), which was also characterized by a shorter T_{22} time constant (Fig. 4B), higher relative amount of the T_{21} population (M_{21}) (Fig. 4C) and concomitantly a lower amount of the T_{22} population (M_{22}) (Fig. 4D) compared to the same relaxation characteristics in pork from the treadmill pigs.

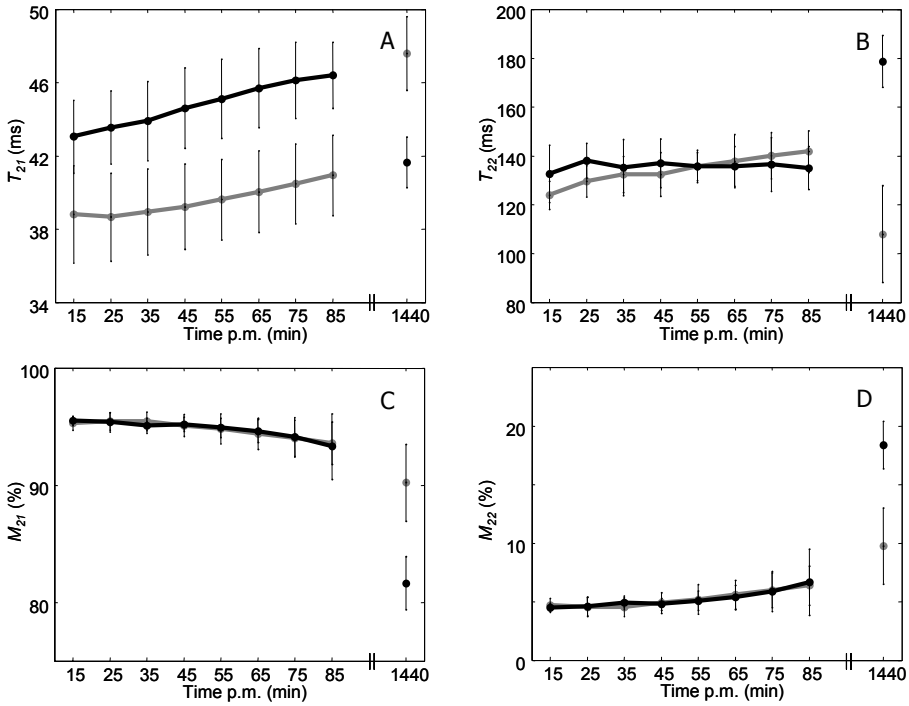


Fig. 4. Transverse time constants (T_2) and amplitudes (M_2) early p.m. and 24 h p.m. estimated by bi-exponential fitting of CPMG decay curves obtained on adrenaline pigs (grey) and treadmill pigs (black). The curves represent averages of five pigs. Error bars indicate ± standard deviation.

Distributed exponential fitting of the CPMG curves confirmed that the T_2 relaxation is characterized by two populations with time constants and population sizes comparable to those found by bi-exponential fitting (Fig. 5). In addition, the distributed analysis revealed that the muscles from the treadmill pigs exhibited a very broad distribution of T_{21} and T_{22} 24 h p.m., leading to an overlap of the T_{21} and T_{22} populations, which was not observed early p.m. or in muscles from the adrenaline-administered pigs (Fig. 5).

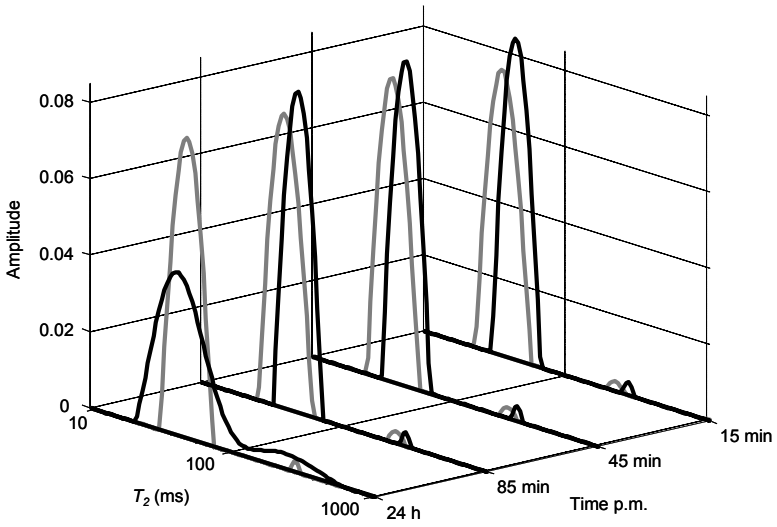


Fig. 5. Distribution of T_2 relaxation times at 15 min, 45 min, 85 min and 24 h p.m. estimated by distributed exponential fitting of CPMG decay curves obtained on samples from an adrenaline pig (grey) and a treadmill pig (black).

4. Discussion

The extent of post mortem glycolysis and thereby pH decrease is mainly determined by the amount of glycogen present in the muscle at the time of slaughter, which can be reduced through adrenaline administration (Bendall & Lawrie, 1962). In accordance with this, a high ultimate pH and consequently a high WHC in the pork was obtained using pre-slaughter adrenaline administration, as reported previously (Henckel et al., 2000). In contrast, the combination of pre-slaughter exercise and electrical stunning resulted in inferior WHC, which is expected due to the resulting fast pH decrease in combination with high temperatures (Wismer-Pedersen, 1959; Offer, 1991). Consequently the WHC of the meat produced made up two groups corresponding to the extreme values of WHC which might be registered under commercial conditions. This is a sound basis for the preliminary testing of the possibility of using the ultrasound equipment for early prediction of WHC despite the fact that the extreme WHC's are generated artificially.

It was initially tested whether usable information about the velocity and attenuation of the ultrasound signal in the meat could be obtained indirectly by calculating frequency spectra of the meat part of the images, but separation according to WHC could not be made based on this transformation in combination with PCA. The best pre-treatment of the ultrasound images was calculation of histograms of the pixel values. Using this transformation, information concerning the WHC could successfully be retrieved from the ultrasound

images 85 min p.m. at the posterior of the LD. Consequently, it seems as if the differences in meat early p.m., which lead to different WHC, manifest themselves in different attenuation of the ultrasound signal, leading to higher amplitudes in treadmill pigs than in adrenaline pigs. This is a positive result, but there are some points of concern.

It is problematic that a separation according to WHC was only obtained at 85 min p.m. The result could be coincidental, but the fact that the histograms display an increased separation with time especially between 65 min p.m. to 85 min p.m. (Fig. 3B-D) indicates that 85 min p.m. is simply the earliest time at which the differences due to different WHC are expressed sufficiently in the meat to be detected by ultrasound. NMR measurements revealed differences in the T_{21} time constant already 15 min p.m. The higher T_{21} time constant in samples with a fast post mortem pH-decrease is probably partly a result of the lower pH affecting the myofibrillar structures, but it may also be reflecting a different degree of contraction. The faster T_{21} relaxation in muscle from the adrenaline-administered pigs could reflect higher degree of longitudinal contraction, as the T_{21} time constant has been shown to be positively correlated to myofibrillar cross-sectional area (Brøndum et al., 2000) and to the sarcomere length (Bertram, Purslow & Andersen, 2002). This is in agreement with Tornberg et al. (2000) who found that myofibrils in meat with a slow pH decrease p.m. are shortened to a greater extent early p.m. than myofibrils in meat with a fast pH decrease. The different degree of shortening could presumably cause a difference in the propagation of the ultrasound, as the difference in T_{21} reflects that the molecular mobility differs and/or the macromolecular spacing differs. However, the effect on the ultrasound is apparently too subtle before 85 min p.m. Thus, a safe classification/quantification of WHC cannot be made from the ultrasound measurements until 85 min p.m., which is interesting, but not sufficiently early for on-line use in an abattoir. It is, however, possible that (invasive) high resolution ultrasound equipment with sweeping frequency may lower the detection limit significantly.

Of greater concern is the fact that a separation according to WHC is only found from the ultrasound images obtained at the posterior of the LD. An explanation for this could be that the images from the anterior and the centre of the LD are noisier i.e., the preceding tissue layers seem to affect the signal in the meat part of the images to a greater extent. At least for the measurements obtained at the anterior of the LD additional interference from the preceding tissue layers could be expected, since the ultrasound had to pass through the spinalis dorsi muscle as well as the fat layers before it reached the LD. Differences in the accuracy of ultrasound measurements along the LD have been observed in an other study where prediction of meat yield using ultrasound exhibited increased accuracy from the cranial positions towards the last rib (Fortin et al., 2003). In any case, it is a weakness of the method that the ultrasound is affected by the transition through the fat layers before it reaches the meat. Accordingly, the frequency and velocity of the ultrasound that is incident

on the meat is inconsistent between measurements and maybe even between transducers during the same measurement depending on the barriers the sound has met on its way before it reaches the meat. Measuring directly on the meat could eliminate the interference from the fat layers, however, this would require development of invasive ultrasound equipment.

The NMR measurements showed that major structural transformations occur from 85 min p.m. to 24 h p.m. as the muscle water characteristics changed considerably in this period. In addition to the changes in the proportions of the populations, which have been reported in previous studies performed 24-72 h p.m. on pigs with different pH decreases p.m. (Tornberg et al., 1993; Bertram, Karlsson & Andersen, 2001a), the T_{21} and T_{22} time constants changed markedly from 85 min p.m. to 24 h p.m. In the early measurements (15-85 min p.m.) the treadmill pigs were characterized by the highest T_{21} time constant, whereas at the 24 h measurement the adrenaline pigs were characterized by the highest T_{21} time constant. This drastic change should probably be ascribed to the fact that early post mortem the higher degree of contraction reduces the T_{21} time constant in the muscles from adrenaline-administered pigs, whereas at 24 h p.m. the higher degree of electrostatic repulsion between the myofilamentous proteins caused by the elevated pH becomes the dominating factor with a resultant higher T_{21} time constant in these muscles. Distributed analysis of the CPMG data acquired 24 h p.m. revealed very broad T_{21} and T_{22} populations in meat from the treadmill pigs 24 h p.m., whereas the T_{21} and T_{22} populations were much more well-defined in meat from adrenaline-administered pigs. The broad distribution of relaxation times in meat from the treadmill pigs is most probably due to a larger variation in protein-water interactions induced by the increased protein denaturation, giving rise to more dramatic structural changes. Thus, the distributed exponential fitting of the NMR decay curves indicates that protein denaturation is more severe 24 h p.m. in pigs with a low WHC than in pigs with a high WHC. The broadening of the water populations was not noticeable in the measurements obtained 85 min p.m. or earlier p.m. in this study. However, Bertram et al. (2003) registered a similar broadening within the first two hours p.m. in rabbit muscles exposed to PSE conditions. The fact that Bertram et al. (2003) found that the degree of protein denaturation was reflected in the width of the water populations when measurements were continued slightly longer than 85 min p.m. underlines that 85 min p.m. is approximately the point in time when variations due to different treatment of the carcasses begin to manifest.

IMF may influence ultrasound measurements as well as NMR. However, the variation in IMF did not correlate with either of the measurements. Thus, IMF is not believed to influence the conclusions of the present investigation.

5. Conclusion

It is a general problem for the on-line methods developed for the prediction of WHC that the biochemical and physical properties that determine the WHC are not fully expressed early p.m. (Kauffman et al. 1993). However, the good predictions of WHC which have been obtained with NIR and IR measurements 30-45 min p.m. (Pedersen et al., 2003; Forrest et al., 2000) show that certain methods are relatively sensitive to the subtle differences early p.m., but this does not seem to be the case for ultrasound, as demonstrated in the present study.

The tested ultrasound method proved to be able to differentiate high and low WHC. However, classification was not obvious before 85 min p.m., which is not sufficiently early for on-line use in a commercial abattoir. NMR measurements underlined that there are structural differences in the meat according to pre-slaughter treatment prior to 85 min p.m.; however, the properties that determine the WHC are apparently not expressed sufficiently to be registered by ultrasound before 85 min p.m. Distributed T_2 analysis of the NMR relaxation data revealed a broadening of the T_2 components for the low WHC meat 24 h p.m. As a similar broadening of the T_2 components has been observed already within two hours p.m. in rabbit muscles exposed to PSE conditions (Bertram et al., 2003), it is likely that 85 min. p.m. is approximately when the differences in the two WHC groups begin to manifest.

Acknowledgements

This investigation was sponsored by The Danish Ministry of Science, Technology and Innovation through the centre contract: “Centre for New Sensor Systems for the Measurement of Food Quality”. We thank Gilda Kischinovsky for proofreading of the manuscript.

References

Bendall, J. R., & Lawrie, R. A. (1962). The effect of pre-treatment with various drugs on post-mortem glycolysis and the onsert of rigor mortis in rabbit skeletal muscle. Journal of Comparative Pathology, 72, 118-130.

Bendall, J. R., & Swatland, H. J. (1988). A review of the relationships of pH with physical aspects of pork quality. Meat Science, 24, 85-126.

Bertram, H. C., Karlsson, A. H., & Andersen H. J. (2001a): Comparative study of low-field NMR relaxation measurements and two traditional methods in the determination of water holding capacity of pork. Meat Science, 57, 125-132.

Bertram, H. C., Karlsson, A. H., Rasmussen, M., Pedersen, M. D., Dønstrup, S., & Andersen, H. J. (2001b). The origin of multi-exponential T_2 relaxation in muscle myowater. Journal of Agricultural and Food Chemistry, 49, 3092-3100.

Bertram, H. C., Purslow, P. P., & Andersen, H. J. (2002). Relationship between meat structure, water mobility, and distribution: a low-field nuclear magnetic resonance study. Journal of Agricultural and Food Chemistry, 50, 824-829.

Bertram, H. C., Whittaker, A. K., Karlsson, A. H., & Andersen, H. J. (2003). pH-dependence of the progression in NMR T_2 relaxation times in post mortem muscle. Journal of Agricultural and Food Science, 51, 4072-4078.

Bouton, P. E., Carroll, F. D., Fisher, A. L., Harris, P. V., & Shorthose, W. R. (1973). Effect of altering ultimate pH on bovine muscle tenderness. Journal of Food Science, 38, 816-820.

Brown, R. J. S., Capozzi, F., Cavani, C., Cremonini, M. A., Petracci, M., & Placucci, G. (2000). Relationships between ^{1}H NMR relaxation data and some technological parameters of meat: A chemometric approach. Journal of Magnetic Resonance, 147, 89-94.

Brøndum, J., Egebo, M., Agerskov, C., & Busk, H. (1998). On-line pork carcassgrading with the Autofom ultrasound system. Journal of Animal Science, 76, 1859-1868.

Brøndum, J., Munck, L., Henckel, P., Karlsson, A., Tornberg, E., & Engelsen S. B. (2000). Prediction of water-holding capacity and composition of porcine meat with comparative spectroscopy. Meat Science, 55, 177-185.

Forrest, J. C., Morgan, M. T., Borggaard, C., Rasmussen, A. J., Jespersen, B. L., & Andersen, J. R. (2000). Development of technology for the early post mortem prediction of water holding capacity and drip loss in fresh pork. Meat Science, 55, 115-122.

Fortin, A., Tong, A. K. W., Robertson, W. M., Zawadski, S. M., Landry, S. J., Robinson, D.J., Liu, T., & Mockford, R. J. (2003). A novel approach to grading pork carcasses: computer vision and ultrasound. Meat Science, 63, 451-462.

Garrido, M. D., Pedauyé, J., Bañón, S., & Laencina, J. (1993). Objective assessment of pork quality. Meat Science, 37, 411-420.

Henckel, P., Karlsson, A., Oksbjerg, N., & Petersen, J. S. (2000). Control of post mortem pH decrease in pig muscles: experimental design and testing of animal models. Meat Science, 55, 131-138.

Hills, B. P. (1998). Magnetic resonance imaging in food science, New York: John Wiley & Sons, Inc.

Honikel, K. O. (1998). Reference methods for the assessment of physical characteristics of meat. Meat Science, 49, 447-457.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components, Journal of Educational Psychology, 24, 417-441.

Kauffman, R. G., Sybesma, W., Smulders, F. J. M., Eikelenboom, G., Engel, B., van Laack, R. L. J. M., Hoving-Bolink, A. H., Sterrenburg, P., Nordheim, E. V., Walstra, P., & van der Wal P. G. (1993). The effectiveness of examining early post-mortem musculature to predict ultimate pork quality. Meat Science, 34, 283-300.

Lawrie, R. (1998). Meat Science (6th ed.). Cambridge, UK: Woodhead Publishing Ltd.

Micklander, E., Peshlov, B., Purslow P. P., & Engelsen S.B. (2002). NMR cooking: Monitoring the changes in meat during cooking by low-field ^{1}H-NMR, Trends in Food Science and Technology, 13, 341-346.

Nordic Committee on Food Analysis (1998). NMKL method no. 160, 1998: Fat determination in foods.

Offer, G. (1991). Modelling of the formation of pale, soft and exudative meat: effects of chilling regime and rate and extent of glycolysis. Meat Science, 30, 157-184.

Offer, G., & Trinick, J. (1983). On the mechanism of water holding in meat: The swelling and shrinking of myofibrils. Meat Science, 8, 245-281.

Park, B., Whittaker, A.D., Miller, R.K., & Bray, D.E. (1994). Measuring Intramuscular fat in beef with ultrasonic frequency analysis. Journal of Animal Science, 72, 117-125.

Pedersen, D. K., Morel, S., Andersen, H. J., & Engelsen, S. B. (2003). Early prediction of water-holding capacity in meat by multivariate vibrational spectroscopy. Meat Science, 65, 581-592.

Pedersen, H. T., Bro, R., & Engelsen, S. B. (2002). Towards rapid and unique curve resolution of low-field NMR relaxation data: Trilinear SLICING versus two-dimensional curve fitting. Journal of Magnetic Resonance, 157, 141-155.

Schäfer, A., Rosenvold, K., Purslow, P. P., Andersen, H. A., & Henckel, P. (2002). Physiological and structural events post mortem of importance for drip loss in pork. Meat Science, 61, 355-366.

Swatland, H. J. (1995). Online evaluation of meat. Lancaster, PA: Technomic Publishing Company, Inc.

Tornberg, E., Andersson, A., Göransson, Å., & von Seth, G. (1993). Water and fat distribution in pork in relation to sensory properties. In E. Puolanne, & D. I. Demeyer, Pork quality: Genetic and metabolic factors. (pp. 239-258). CAB International.

Tornberg, E., Wahlgren, M., Brøndum, J., & Engelsen, S. B. (2000). Pre-Rigor conditions in beef under varying temperature- and pH-falls, studied with rigometer, NMR and NIR. Food Chemistry, 69, 407-418.

Trout, G. R. (1988). Techniques for measuring water-binding capacity in muscle foods – a review of methodology. Meat Science, 23, 235-252.

van Oeckel, M. J., Warnants, N., & Boucqué, C. V. (1999). Comparison of different methods for measuring water holding capacity and juiciness of pork versus on-line screening methods. Meat Science, 51, 313-320.

Warris, P. D., & Brown, S. N. (1987). The relationships between initial pH, reflectance and exudation in pig muscle. Meat Science, 20, 65-74.

Watanabe, A., Daly, C. C., & Devine, C. E. (1996). The effects of the ultimate pH of meat on tenderness changes during ageing. Meat Science, 42, 67-78.

Whittaker, A.D., Park, B., Thane, B.R., Miller, R.K. & Savell, J.W. (1992). Principles of ultrasound and measurement of intramuscular fat. Journal of Animal Science, 70, 942-952.

Wismer-Pedersen, J. (1959). Quality of pork in relation to rate of pH change post mortem. Food Research, 24, 711-727.