ELSEVIER

0144-8617(95)00130-1

Carbohydrate Polymers **28** (1995) 187–193

0144-8617/95/$09.50+.00

Characterization of partially *N*-acetylated chitosans by near infra-red spectroscopy

Kjell M. Vårum[a], Bjørg Egelandsdal[b] & Marit R. Ellekjær[b]

[a]*Norwegian Biopolymer Laboratory (NOBIPOL), Department of Biotechnology, The Norwegian Institute of Technology (NTH), University of Trondheim, N-7034 Trondheim, Norway*
[b]*MATFORSK, Oslovn. 1, N-1430-AAS, Norway*

(Received 20 June 1995; revised version received 6 October 1995; accepted 12 October 1995)

Six samples of fully water-soluble chitosans with fractions of *N*-acetylated units (F_A) from 0.01 to 0.51 (degrees of acetylation from 1 to 51%) were converted to four different salt forms (formate, acetate, propionate and chloride) in addition to the free amine form. Near infra-red (NIR) and visible reflectance spectra (400–2500 nm) of the dry samples of chitosans were recorded and data analysis was performed using multivariate calibration methods in order to relate the spectral data to F_A and salt forms/free amine form. About 95% of the variation in the NIR spectra could be explained by salt form and F_A, using a partial least squares method with 10 components. Reflectance data in the 1100–2500 nm wavelength range were also sufficient to determine F_A. Information about the intrinsic viscosities of the chitosans may also be obtained from the high-energy part of the spectra. A partial least squares model of nine components based on all samples (six chitosans with different F_A which were each prepared in four different salt forms and the free amine form) was able to predict F_A with a root mean square prediction accuracy of 0.023 for chitosans varying in F_A from 0 to 0.6.

INTRODUCTION

Chitin is the structural polysaccharide in the phylum Arthropoda (animals with an outer skeleton). Its chemical structure is similar to cellulose, where the hydroxyl group at carbon 2 has been replaced by an acetamido group. Chitosan, partially *N*-deacetylated chitin, is much less abundant in nature than chitin, and is distinguished from chitin by its solubility in dilute aqueous acid solutions (Roberts, 1992). Chitosan is a binary heteropolysaccharide of $\beta(1\rightarrow4)$-linked 2-acetamido-2-deoxy-β-D-glucopyranose (GlcNAc - **A**-unit) and 2-amino-2-deoxy-β-D-glucopyranose (GlcN - **D**-unit) of varying composition. It has been shown that fully water-soluble chitosans can be made with fractions of acetylated units (F_A) from 0 to 0.6 (Vårum *et al.*, 1991a, b; Anthonsen *et al.*, 1993; Nordtveit *et al.*, 1994).

A range of methods to determine F_A have been described (see Roberts, 1992 for a review). Most methods for determining F_A on chitosans rely on solubilizing the chitosan prior to analysis (i.e. IR-spectroscopy of chitosan films, NMR and UV spectroscopy, titration), which may result in erroneous F_A values when extending the range of F_A of chitosan towards more chitin-like materials.

Here we report on the use of near infra-red reflectance for determining F_A. Near infra-red (NIR) is a rapid technique which works well on dry materials. The technique is highly reproducible, robust and with good penetration depth into the sample. NIR usually works well when the components to be estimated are present in significant amounts ($\geq 2\%$). As some promising results recently appeared in the literature about the use of NIR on chitosans (Rathke & Hudson, 1993), work was initiated attempting to determine F_A on a broader range of chitosans.

Although chitosan is usually prepared in its free amine form ($-NH_2$), it can also be prepared in different salt forms with acids (i.e. HCl, HCOOH, CH_3COOH). These chitosan salts are frequently produced because they are readily soluble in pure water without addition of any acid.

In this work we have prepared and characterized fully water-soluble chitosan samples with a broad range of F_A values, with varying chain lengths, and in different salt forms. Their NIR spectra were recorded and data analysis was performed using multivariate regression methods in order to relate the spectral data to the salt form and the physical–chemical characteristics of the chitosans. It was found that a model based on all samples was able to predict F_A with an acceptable prediction accuracy for chitosans varying in F_A from 0 to 0.6.

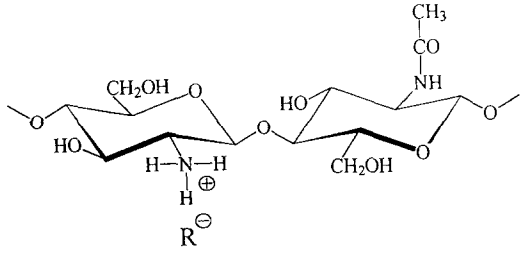
MATERIALS AND METHODS

Preparation of chitosan samples

Six chitosans with different fractions of acetylated units (F_A) were used in this study (Table 1). All chitosans were prepared by heterogeneous deacetylation, and the acid-soluble fraction was prepared as previously described (Vårum *et al.*, 1992). The chitosans were converted into four different salt forms (chitosan-HCl, chitosan-acetate etc., see Fig. 1) in addition to the free amine form. Thus, a total of 30 different chitosan samples were prepared. The different salt forms were prepared by dissolution in dilute acetic acid, three times dialysis against 0.2 M concentration of the sodium salt of the corresponding acid (NaCl, Na-acetate etc.) and further dialysis against distilled water. Before lyophilization, the pH of the chitosan solution was adjusted to ~5 with the acid corresponding to the chitosan salt that was prepared. The chitosan salts were filtered and lyophilized. The lyophilized samples varied in 'lumpiness' and were thus unsuitable for direct NIR measurements. The samples were made more homogeneous by wetting in a mortar with acetone–water (6:1), drying and finally grinding in a hammer mill to pass a 1 mm sieve.

Table 1. Characterization of the chitosan samples used for calculations in Figs 1–5

Sample	F_A	Intrinsic viscosity of HCl form (ml g^{-1})	$\overline{M}_n$	$\overline{DP}_n$
1	0.01	830	260,000	1300
2	0.09	210	330,000	165
3	0.15	1100	340,000	1700
4	0.16	1640	570,000	2900
5	0.35	960	240,000	1200
6	0.51	1270	290,000	1400



R⁻ = COO^- (formate) (A)
= CH_3COO^- (acetate) (B)
= $CH_3CH_2COO^-$ (propionate) (C)
= Cl^- (chloride) (D)
Free amine form (NH_2) (E)

Fig. 1. Chemical structure of partially *N*-acetylated chitosans in different salt forms.

Characterization of chitosan samples

The fraction of acetylated units were determined by high-field proton NMR spectroscopy as previously described (Vårum *et al.*, 1991a). Intrinsic viscosities ([η]) of chloride salts of the chitosans were determined as previously described (Draget *et al.*, 1992). The number-average molecular weights ($\overline{M}_n$) were estimated from the Mark–Houwink–Sakurada equation as reported by Anthonsen *et al.* (1993), and the number-average degree of polymerization ($\overline{DP}_n$) calculated by dividing $\overline{M}_n$ by the molecular weight of the monomer.

Near infra-red (NIR) spectroscopy

Reflectance measurements (400–2500 nm) were made by the scanning spectrophotometer Model 6500 from NIRSystems Inc. (Silver Springs, USA). Samples (0.7 g) were put into the microsample cup 7076 and the sample (and reference) spectra were collected at room temperature by averaging 24 scans. As reference for the sample measurements, the white bottom plate of the sample cup was used. The reference was scanned between every sample measurement. The 30 'original' samples (Table 1) were scanned in a randomized order. The samples prepared with increasing amounts of sodium acetate (Table 2), used for model testing, were measured in a non-random order 1 month later. Each sample was measured in three repacked replicates, and the means of the spectra obtained were used in the regressions.

Data analysis

Multiplicative scatter correction (MSC) (Martens & Næs, 1989) was performed on the spectral data prior to calibration. MSC was developed in order to correct for light scattering variations in NIR reflectance spectra. This is mainly an empirical method although the technique was derived from the fact that scattered light can be differentiated from adsorbed light by the use of data obtained from many wavelengths. Previous work has reported improved performance by applying MSC to

Table 2. Predicted F_A for samples with added sodium acetate (NaAc)

% added NaAc	Original model (30 samples)	New model (33 samples)
0	0.14	0.14
10	0.14	u.c.
20	0.17	0.14
30	0.21	u.c.
40	0.27	0.16
50	0.32	u.c.

The value determined by proton NMR spectroscopy for F_A is 0.16 for all samples.
u.c., used for calibration.

NIR reflectance data (1100–2500 nm) (Geladi *et al.*, 1985; Isaksson & Næs, 1988), NIR transmittance data (850–1050 nm) (Ellekjær & Isaksson, 1992; Isaksson & Næs, 1992) and visible reflectance spectra (380–760 nm) (Iversen & Palm, 1985) prior to multivariate calibration. MSC gives a better linear fit between spectral data and chemical composition, better spectral interpretability and improved prediction results compared to uncorrected data.

The multivariate calibration methods partial least squares (PLS1 and PLS2) regression (Martens & Næs, 1989) were used to relate the spectral data (X-data) to the fraction of acetylated units and counterions (Y-data) of chitosan. These are linear methods which first reduce the spectral data to a few factors which express the main sources of variation in the X-data which are relevant for the Y-data, and then use a number of these estimated components as X-variables (regressors). The difference between PLS1 and PLS2 is that PLS1 makes one model for each Y-variable, whereas PLS2 makes one common model for all three Y-variables. PLS1 was preferred to the more well known method, principal component regression, because simpler models are obtained by PLS.

Cross-validation (Martens & Næs, 1989) was used to validate the calibration models. The prediction accuracy of the calibration model is described by the root mean square of prediction (RMSEP) (Martens & Næs, 1989) which is defined as

$$\text{RMSEP} = \sqrt{\frac{1}{I}\sum_{i=1}^{I}(\hat{y}_i - y_i)^2}$$

where y_i is the known and $\hat{y}_i$ is the predicted value for sample number i, and I is the number of cross-validations (which in this case, equals the number of samples).

The MSC, PLS1 and PLS2 regressions were performed by the Unscrambler Software Version 5.5 (Camo A/S, Trondheim, Norway). To validate the models obtained, full cross-validation was used, i.e. each sample was used to test the model prepared from all of the other samples.

RESULTS AND DISCUSSION

Chitosan samples

The characterization of the chitosan samples with increasing F_A from 0.01 to 0.51 is given in Table 1. The chitosans were of varying intrinsic viscosities [η], but no attempt was made to select samples representative of this parameter. In addition, each of the six chitosans was converted to four different salt forms and the free amine form, as seen in Fig. 1. Thus, an array of 30 samples was prepared, and their NIR spectra were recorded.

Main predictive information in the spectra of chitosan

PLS2 was performed on the spectra from 400 to 2500 nm towards the three properties counterion, F_A and [η] simultaneously, in order to identify properties describing the main information in the spectra of the chitosans. F_A and counterion described the main variation in the chitosan spectra (Fig. 2a). The first two PLS components explained 60% of the variation in F_A whereas a model with three PLS components explained 70% of the variation in counterions. This made up the main information in the spectra as 85% of the variation in the spectra (X-data) was described by the three first PLS components. Totally, the spectra described about 95% of the variation in counterions and F_A using a PLS model with 10 components (Fig. 2a).

The most common NIR instruments use the spectral region from 1100 to 2500 nm, and we, therefore, compared the predictive information in this region with the predictive information achieved when using the spectral region from 400 to 2500 nm. F_A and counterion were equally well explained when only the lower energy region was used (Fig. 2a and b). Totally, the spectra from 1100 to 2500 nm described about 95% of the variation in both counterion and F_A (PLS model with

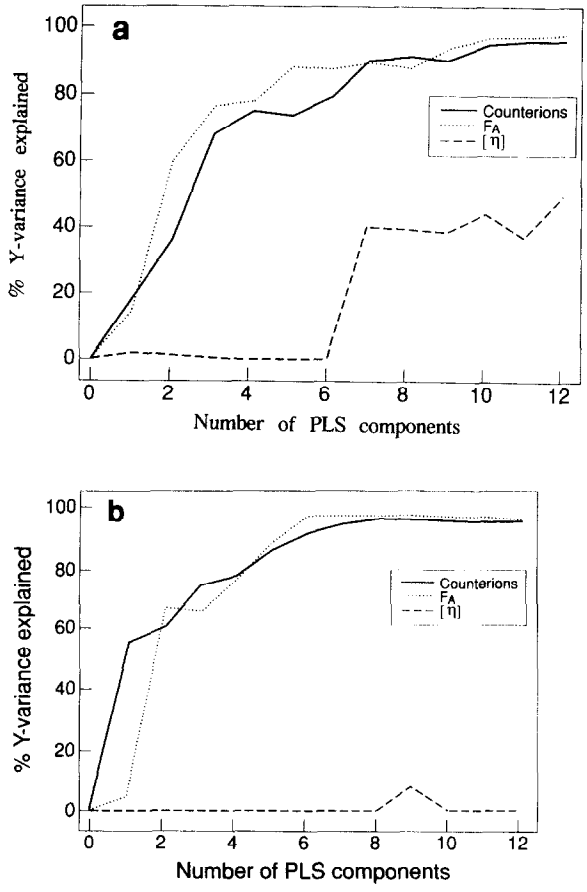


Fig. 2. Explained variance (%) for F_A, counterion and intrinsic viscosity as a function of the number of components used in the PLS2 analysis. (a) 400–2500 nm; (b) 1100–2500 nm.

eight components). It was thus possible to both differentiate between different counterions and determine the F_A of the chitosan samples by using the spectral region from 1100 to 2500 nm. The first PLS component (Fig. 2b) explained 60% of the variation in counterion, whereas the second PLS component explained 70% of the variation among the samples in F_A. These two PLS components could, therefore, be used as indicators for the chemistry behind the predictive information in the region from 1100 to 2500 nm related to the salt form and F_A, respectively.

In the spectral region from 1100 to 2500 nm, 0.2% of the spectral variation contained predictive information about intrinsic viscosity (Fig. 2b). For the whole spectral region from 400 to 2500 nm (Fig. 2a) less than 1% of the spectral variation contained predictive information about intrinsic viscosity; still a small figure. Thus the variation in intrinsic viscosity inherent in the samples here was of no problem for the determination of F_A. However, work is in progress which will scrutinize the possibility of extending the NIR reflectance method to perform a crude classification of samples in terms of intrinsic viscosity or degree of polymerization ($\overline{DP}_n$).

The first PLS component (Fig. 2b) which described nearly 60% of the variation in counterions and about 50% of the spectral (X) variation, described mainly the difference between the free amine forms of the chitosans (E) with high positive scores and the organic acid chitosan salts (formate (A), acetate (B) and propionate (C)) with high negative scores (Fig. 3a). This was particularly pronounced for samples with low F_A, which is as expected since a low F_A means a high content of counterions. From the scores in Fig. 3a it is seen that as F_A increases the scores of PLS-component 1 (attributed mainly to variation in counterions) decrease.

The major change in scores within samples, however, appeared upon changing from the free amine form (E) to an HCl salt (D), but in general, also from E samples to either A, B, C or D form, as seen in Fig. 3a. Thus, the chitosans containing deacetylated units in the neutral amine form ($-NH_2$) have spectroscopically striking differences from all ammonium ion forms ($-NH_4^+$).

The wavelengths 1526 and 2026 nm had high positive loadings (Fig. 3b). Those wavelengths can primarily be related to the NH stretch in the free amine forms (E), but information from the acetylated units (C=O stretch, NH stretch) is also contained at these or closely related wavelengths as well as at 1928 nm (Shenk *et al.*, 1992). Thus, neutral samples with a high F_A would be expected to adsorb most strongly at the above wavelengths, in agreement with Fig. 3a. This is quite evident from the spectra already before processing as shown in Fig. 4, where the neutral amine forms (E) show pronounced peaks at 1526 and at 2026 nm. Such distinct peaks were missing in the acetate (B) and chloride (D) forms. The wavelengths 1692–1694 nm, 1854–1856 nm and 2220–2222 nm which are related to the absorbance of C=O stretch in (amino acid) ionized carbonyls (Murray & Williams, 1987) had high negative loadings (Fig. 3b). This means that the E-forms of the chitosan samples had high absorbance at the wavelengths 1526 and 2026 nm and lower absorbances at the wavelengths 1692–1694 nm, and vice versa for the A, B and C forms of chitosan samples.

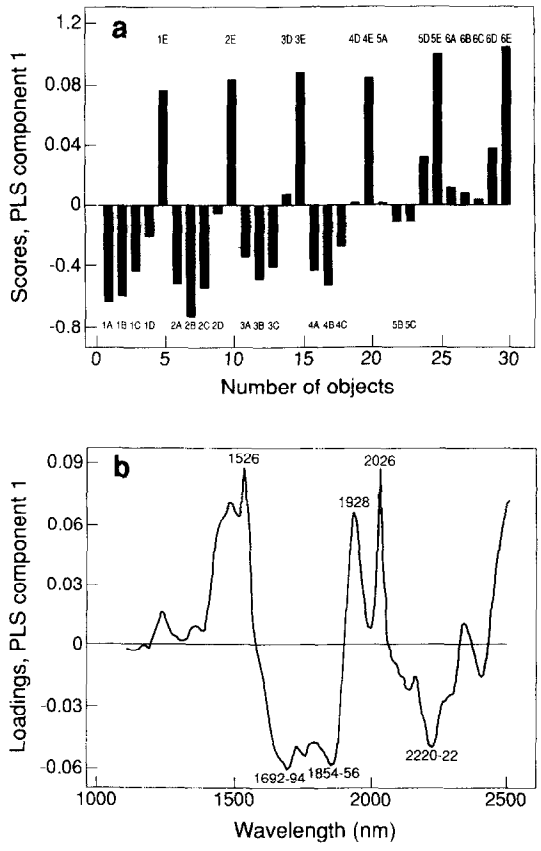


Fig. 3. Relationship of the spectral data (1100–2500 nm) to the main information about counterions visualized by the (a) scores; and (b) spectral loadings for PLS component 1.

The second PLS component which explained 70% of

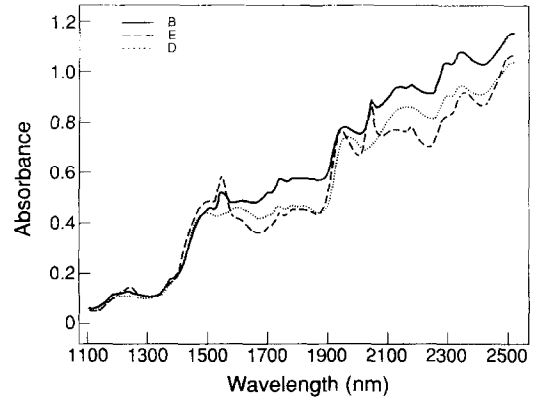


Fig. 4. Absorption spectra for the two different counterions and the neutral form. The spectra are the arithmetic mean spectra for the different F_A.

the variation in F_A but much less of the spectral variation (about 25%), differentiated mainly between the chitosan with the lowest F_A (sample 1, see Table 1), and the chitosans with the highest F_A (samples 5 and 6, see Table 1), as seen from Fig. 5a. Again we can see that the type of counterion is not so influential for samples with a high F_A (Fig. 5a). However, for the samples with a low F_A the scores are highly influenced by the type of counterion.

The wavelengths 1526–1528 and 2026 nm which above were related to the NH stretch of the free amine forms (E) and to the C=O and NH stretch in acetylated units were shown to be important for the differentiation between samples with different F_A, as seen from the loadings vs wavelength plot of PLS component 2 in Fig. 5b. Sample 1 (A–E) with $F_A = 0.01$ was positively associated with these wavelengths; i.e. these samples adsorb strongly at these wavelengths while the most acetylated samples 5–6 (A–D) were highly negatively associated.

The scores and spectral loadings in Fig. 5a and b, respectively, were not so easy to interpret in terms of chemical information. This was presumably due to the fact that the signal from the deacetylated units was more dominant than the signal from the acetylated units. Thus, F_A values seem to be determined by counting the number of deacetylated units, and not by counting the number of acetylated units. The neutral amine forms (E) of the chitosan samples seem to interfere with the prediction of F_A because the information related to the free amine form (E) and to F_A is found at practically the same wavelengths. It was, therefore, of interest to compare the prediction accuracy of F_A with and without the E-forms in the calibration model (see 'The robustness of the model when predicting F_A').

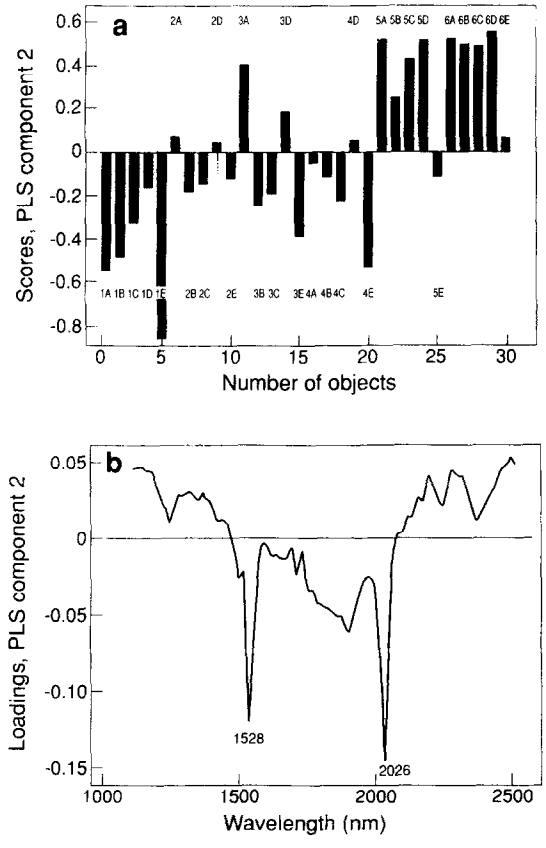


Fig. 5. Relationship of the spectral data (1100–2500 nm) to the main information about F_A visualized by the (a) scores; and (b) spectral loadings for PLS component 2.

Determination of F_A

A separate calibration model (PLS1) was made for the prediction of F_A in order to optimize the prediction accuracy of this property. A root mean square prediction accuracy of 0.023 was achieved with a PLS model of nine components, when using the whole spectrum from 400 to 2500 nm. The prediction accuracy for F_A in the reference method (proton NMR spectroscopy) was determined by preparing five independent samples of a chitosan with $F_A = 0.35$ (sample 5 in Table 1) for NMR analysis. The mean F_A value was determined as 0.354 with a standard deviation of 0.012, suggesting that the NIR method at best determines F_A with comparable precision to the reference (proton NMR) method, as the error of the reference is contained in the prediction error of the NIR method.

Using only the spectral region from 1100 to 2500 nm resulted in a prediction accuracy of 0.030 with six PLS components in the model. In Fig. 6, the predicted F_A when using the spectra from 1100 to 2500 nm was plotted against the actual F_A (determined from proton NMR spectroscopy). The correlation coefficient for the straight line in Fig. 6 was 0.985.

The robustness of the model when predicting F_A

The effect of the E-forms on the prediction accuracy of F_A, as indicated above, was evaluated. In Fig. 7 the prediction accuracy (expressed as RMSEP) is presented

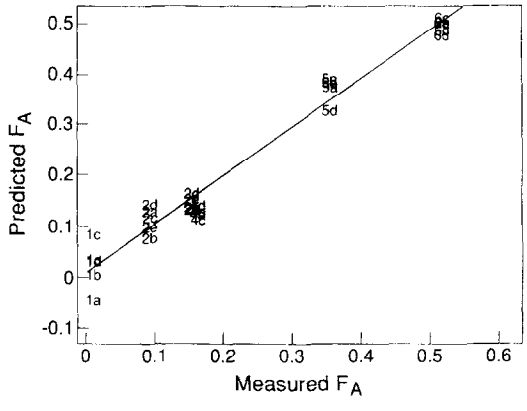


Fig. 6. Predicted vs measured F_A for the chitosan samples in Table 1. PLS1 on spectral data from 1100 to 2500 nm with six components in the model.

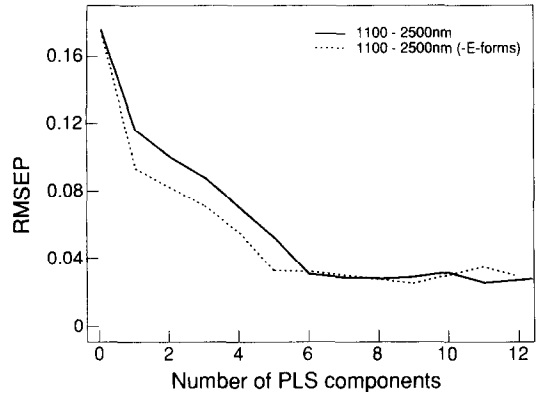


Fig. 7. Prediction errors (RMSEP; see text) as a function of the number of components in the model when determining the F_A of the chitosan samples in Table 1.

for a PLS model based on all 30 samples and a PLS model with the E-forms removed from the sample set. After five components, the model with the E-forms excluded had a lower prediction error. Practically the same prediction accuracy was, however, achieved with and without E-forms in the model when six components were included. Thus, by combining the information at several wavelengths it was eventually possible to compensate for the interference that the free amine forms caused by inducing overlapping of important signals.

It should be possible to improve the prediction accuracy for F_A by the NIR method, when applying the method for the determination of only one form of chitosan and/or for a more narrow range of F_A values. A separate calibration model for only the free amine form or one particular salt form of the chitosans would presumably reduce RMSEP for both types of sample. However, the present data show that it is also possible to determine F_A with an acceptable accuracy without knowledge about the salt form of the chitosan. A separate, cross-validated model for only one chitosan form could not be built here as only six samples were available.

It was surprising that such a good model could be established using widely different chitosan salts, especially since one of them (the acetate) is chemically very similar to the acetyl groups in chitosans, which determine the F_A. We, therefore, evaluated the model established with the 30 samples referred to above on samples with an increasing amount of sodium acetate (NaAc). An unacceptable error in F_A was seen for samples containing more than 20% added sodium acetate (Table 2). However, if three of these samples were included in the calibration test set, i.e. to establish a new model, the systematic tendency for the determined F_A to increase with NaAc addition was reduced, and F_A was well predicted for the remaining samples (Table 2).

We also checked if a chitosan with F_A larger than 0.51 could be predicted by the original model. For a water-soluble chitosan sample with $F_A = 0.61$, the model was still able to predict F_A within acceptable limits. However, when a water-insoluble chitin sample with $F_A = 0.89$ (determined by solid-state (CPMAS) NMR spectroscopy (Ottøy *et al.*, 1994)) was included in the calibration data set, the model was extensively destabilized. One possible explanation may be the higher crystallinity in the chitin sample compared to the chitosan samples.

CONCLUSIONS

It was possible to determine F_A of dry samples of chitosans with an RMSEP of 0.023 using an array of chitosans (30 samples) consisting of six chitosans with F_A from 0 to 0.5 which were each prepared in the free amine form and four different salt forms. The free amine form of the chitosans had characteristic adsorption peaks, not present in the salt forms of chitosans which interfered with the determination of F_A. Improvement in the prediction error should be achieved in the future by establishing different calibration models for the free amine form and the salt forms. The model which was established seemed reasonably robust when applied on slightly different chitosan samples to those used for calibration.

ACKNOWLEDGEMENTS

We thank Ingrid Aune and Siv Moen for skilful technical assistance. Financial support has been provided by the Research Council of Norway.

REFERENCES

Anthonsen, M.W., Vårum, K.M. & Smidsrød, O. (1993). *Carbohydr. Polym.*, **22,** 193–201.

Draget, K.I., Vårum, K.M., Moen, E., Gynnild, H. & Smidsrød, O. (1992). *Biomaterials,* **13,** 635–638.

Ellekjær, M.R. & Isaksson, T. (1992). *J. Sci. Food Agric.*, **59,** 335–343.

Geladi, P, MacDougall, D. & Martens, H. (1985). *Appl. Spectrosc.*, **39,** 491–500.

Isaksson, T. & Næs, T. (1988). *Appl. Spectrosc.*, **42,** 1273–1284.

Isaksson, T. & Næs, T. (1992). In *Proc. 4th Int. Conf. Near Infrared Spectroscopy*, eds I. Murray & I.A. Cowe. Ian Michael, Chichester, UK, pp. 140–146.

Iversen, A. & Palm, T. (1985). *Appl. Spectrosc.*, **39,** 641–646.

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

Murray, I. & Williams, P.C. (1987). In *Near-Infrared Technology in the Agricultural and Food Industries*, eds P. Williams & K. Norris. American Assoc. Cereal Chem., St.Paul, Minnesota, 17–34.

Nordtveit, R.J., Vårum, K.M. & Smidsrød, O. (1994). *Carbohydr. Polym.*, **23,** 253–260.

Ottøy, M.H., Vårum, K.M. & Smidsrød, O. (1996). *Carbohydr. Polym.*, in press.

Rathke, T.D. & Hudson, S.M. (1993). *J. Polym. Science. Part A: Polymer Chemistry*, **51**, 749–753.

Roberts, G.A. (1992). *Chitin Chemistry*. Macmillan, Hong Kong, p. 116.

Shenk, J.S., Workman, J.J., Jr & Westerhaus, M.O. (1993). In *Handbook of Near-Infrared Analysis*, eds D.A. Burns & E.W. Ciurczak. Marcel Dekker, New York, pp. 383–432.

Vårum, K.M., Anthonsen, M.W., Grasdalen, H. & Smidsrød, O. (1991*a*). *Carbohydr. Res.*, **211**, 17–23.

Vårum, K.M., Anthonsen, M.W., Grasdalen, H. & Smidsrød, O. (1991*b*). *Carbohydr. Res.*, **217**, 19–27.

Vårum, K.M., Ottøy, M.H., Anthonsen, M.W., Grasdalen, H. & Smidsrød, O. (1992). In *Advances in Chitin and Chitosan*, eds C.J. Brine, P.A. Sandford & J.P. Zikakis. Elsevier Science Publishers, London, pp. 127–136.