JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

High-Resolution ^{13}C Nuclear Magnetic Resonance Spectroscopy Pattern Recognition of Fish Oil Capsules

MARIT AURSAND,*,† INGER B. STANDAL,†,‡ AND DAVID E. AXELSON†,§

SINTEF Fisheries and Aquaculture, N-7465 Trondheim, Norway, Department of Biotechnology, Norwegian University of Natural Science and Technology, N-7491 Trondheim, Norway, and MRi_Consulting, 8 Wilmot Street, Kingston, Ontario, Canada K7L4V1

^{13}C NMR (nuclear magnetic resonance) spectroscopy, in conjunction with multivariate analysis of commercial fish oil-related health food products, have been used to provide discrimination concerning the nature, composition, refinement, and/or adulteration or authentication of the products. Supervised (probabilistic neural networks, PNN) and unsupervised (principal component analysis, PCA; Kohonen neural networks; generative topographic mapping, GTM) pattern recognition techniques were used to visualize and classify samples. Simple PCA score plots demonstrated excellent, but not totally unambiguous, class distinctions, whereas Kohonen and GTM visualization provided better results. Quantitative class predictions with accuracies >95% were achieved with PNN analysis. Trout, salmon, and cod oils were completely and correctly classified. Samples reported to be salmon oils and cod liver oils did not cluster with true salmon and cod liver oil samples, indicating mislabeling or adulteration.

KEYWORDS: Fish oil capsules; ^{13}C NMR spectroscopy; mono-, di-, triacylglycerols, positional distribution, multivariate data analysis

INTRODUCTION

In recent years, fish and fish oil have received growing interest due to their nutritional benefits. This interest has been focused on fish and fish oils' high levels of the long-chain *n*-3 fatty acids *cis*-5,8,11,14,17-eicosapentaenoic (20:5*n*-3) and *cis*-4,7,10,13,16,19-docosahexaenoic (22:6*n*-3), which are believed to play a preventive role in cardiovascular disease and in the alleviation of other health problems (*1*). Dietetic research has shown that most people do not have enough *n*-3 fatty acids in their diet (*2*). The health properties of the products are expected to be dependent on the molecular nature of the agents (e.g., as ethyl esters or triacylglycerols and their positional distribution in the triacylglycerol molecule) and the relative concentration of the *n*-3 fatty acids.

New fish oil products based on natural fish oil or their derivatives are constantly being introduced on the international market as health foods (capsules) or medicines. Microencapsulated fish oil has been introduced for enrichment of foodstuff including bread, infant formulas, baby food, soups, and prepared food, such as pizza. The constituents of these "fish oil products" can either be natural fish oil, in which the initial glycerol structure is intact; *n*-3 acid triacylglycerols, obtained by transesterification of concentrated or purified *n*-3 acids with glycerol; and methyl or ethyl esters and/or mono- and diacylglycerols. Most of the fish oil capsules on the market today are blends of natural fish oil and their derivatives. During the past few years, marine oil capsules have been introduced, specified as tuna oil, salmon oil, etc.

The quality of the fish oil products may vary significantly according to the quality of the raw material and how they have been produced or manufactured. Demand for product specification and label requirements varies significantly, depending upon whether the product is a health food, a food ingredient, a remedy, or an ingredient in a remedy. In general, precise directives for fish oil products on the health food market are lacking with respect to both national and international regulations. There are European monographs for salmon oil (*3*), cod liver oil, and fish oil rich in *n*-3 acids, *n*-3 acid ethyl esters, and *n*-3 acid triacylglycerols (*4*). However, the need for quality assessment and authentication work on fish oil products will intensify as more products with detailed descriptions related to species, etc., are introduced to the market. National and international (EU) regulations existing today include, to a large extent, methods meant to be used for product identification, which are valuable only for general quality control. Authorities have expressed an immediate need for standard methods to determine composition, quality, and authenticity of a wide variety of marine lipids. A broad range of methods used in basic lipid research is available for authentication purposes. For fish oil products, "fingerprint" methods [nuclear magnetic resonance (NMR), gas chromatography (GC), and gas chromatography isotope ratio mass spectrometry (GC-IRMS)], which make it possible to minimize the manipulation of the product before examination, will be

* Corresponding author: tel +47 48200158; fax +47 93270701; e-mail Marit.Aursand@sintef.no.
† SINTEF Fisheries and Aquaculture.
‡ Norwegian University of Natural Science and Technology.
§ MRi_Consulting.

10.1021/jf061754l CCC: $37.00
Published on Web 12/14/2006

valuable tools with respect to authentication of fish oil capsules, foods containing microencapsulated fish oil, and other related products.

The ^{1}H and ^{13}C NMR spectra of the most important saturated and mono- and polyunsaturated fatty acids found in fish lipids have been interpreted (*5*−*7*) and the ^{2}H spectra of pure fatty acids, including 20:5*n*-3 and 22:6*n*-3, have been published (*8*). ^{2}H SNIF (site-specific natural isotope fractionation) NMR in combination with data from GC and IRMS have been used to study commercial fish oil and lipid extracted from muscle of wild and farmed salmon from Norway and Scotland. It was possible to differentiate between wild and farmed salmon and between farmed salmon from Norway and Scotland (*9*, *10*). Multicomponent analysis of encapsulated cod liver oil by ^{1}H and ^{13}C NMR has been carried out by Siddiqui et al. (*11*).

The positional distribution of mono-, di-, and polyunsaturated fatty acids in the triacylglycerol molecule seems to be unique for each fish species, and ^{13}C NMR spectra give us information about such distribution (*7*).

Pattern recognition techniques have been frequently and successfully applied to a variety of applications related to food composition and authentication (*12*), and species differentiation has been reported by multivariate analysis of phospholipids from canned Atlantic tuna (*13*).

The aim of this study was to test the possibility of using ^{13}C NMR in quality assessment of fish oil (commercial fish oil and capsules) with emphasis on authentication. We demonstrate that the use of ^{13}C NMR high-resolution solution spectra of fish oils allows one to develop classification and quantification schemes and also that, due to the highly diagnostic properties of these complex spectra, a wide variety of classification methodologies can be successfully applied.

MATERIALS AND METHODS

Fish Oil Products. Commercial fish oil capsule products were bought at different supermarket, retail pharmacy, and health food stores. Fish oil samples were prepared either by extraction of liver (farmed cod from Norway and wild cod from Iceland and Barents Sea) or muscle lipids (farmed trout and salmon from Norway) (*14*, *15*) or delivered by commercial producers of fish oil (fish oil products from different manufacturing steps). Sample numbers and comments are given in **Table 1**.

NMR Parameters. Fish oil (0.1 mL) in $CDCl_3$ (0.6 mL, 99.8% purity) (Isotec Inc., Matheson) was placed in 5-mm NMR tubes, and proton-decoupled ^{13}C spectra were recorded on a Bruker Avance instrument at 125.77 MHz. A semiquantitative approach was chosen, due to the fact that quantitative measurements require considerably longer experimental time. The following experimental conditions were applied: spectral width 25252.5 Hz, pulse angle 30°, relaxation delay 1.0 s, dwell time 19.80 μs, acquisition time 1.298 s, and 32 768 data points. Number of scans (NS) was set to 10 240 for the spectra used in the multivariate data analysis. Prior to Fourier transformation, a line-broadening factor of 0.5 Hz was applied to minimize noise, but not at the expense of resolution among significant closely spaced resonances. Chemical shifts were referenced to the $CDCl_3$ peak at 77.08 ppm.

Multivariate Data Analysis. For the multivariate data analysis, in each spectrum the peak positions and corresponding intensities were obtained for all resonances with relative intensities greater than 1.5% of the maximum peak intensity (excluding any solvent resonances). Spectra were normalized, with the maximum peak in each spectrum scaled to a value of 100.

Principal component analysis (PCA**)** (*16*), Kohonen neural network analysis (*17*), and generative topographic mapping (GTM) (*18*) were applied as unsupervised multivariate analyses, whereas probabilistic neural networks (PNN) (*19*) provided quantitative, supervised classification. General regression neural networks (GRNN) (*20*) were applied to the calculation of the components of blends containing systematically varying proportions of salmon oil and adulterant.

Table 1. Fish Oil Products, Lipid Extracts,[a] Fish Oil Products from Different Manufacturing Steps, and Commercial Fish Oil Capsules

number	comments or reported composition[b]
1, 2	fish oil capsule product
3	fish oil capsule product, salmon oil
4	fish oil capsule product rich in omega-3
5	fish oil, reported as salmon oil
6, 7	cod liver oil before processing
8, 9	cod liver oil first processing step
10, 11	cod liver oil second processing step
12, 13	cod liver oil third processing step
14, 15	cod liver oil fourth processing step
16, 17	cod liver oil end product
18, 19	cod liver oil, Hustøy, Norway
20, 21	cod liver oil, Andenes, Norway
22, 23	cod liver oil, Iceland
24, 25	cod liver oil, Varøy, Norway
26, 27	fish oil product, trout oil from farmed fish, Norway
28–31	fish oil product, salmon oil from farmed fish, Norway
32, 33	health food/fish oil (supposedly salmon)—not
34, 35	standard commercial fish oil, South America
36	capsule, natural fish oil
37	capsule, free fatty acids
38, 39	capsule, mono-, di-, and triacylglycerols
40	capsule, ethyl esters
41	65% omega-3, capsule, mono-, di-, and triacylglycerols
42	plant oils mixed with fish oils
43	cod liver oil, natural triacylglycerols
44	*cis*-5,8,11,14,17-eicosapentaenoic, GLA, capsule, plant/fish oils
45	capsule, mono-, di- and triacylglycerols
46	capsule, natural fish oils
47	fluid, natural fish oils
48	capsule, ethyl esters
49	capsules, natural fish oils
50	capsules, mono-, di-, and triacylglycerols
51	capsules, cod liver oil (natural triacylglycerols)
52	capsules, mono-, di-, and triacylglycerols
53	natural cod liver oil
54	shark liver oil (special oil containing squalene/sterols, etc.)
55–63	wild cod, Iceland, extract
64–66	farmed cod, Norway, extract
67–76	wild cod, Barents Sea, extract
77	commercial fish oil, salmon oil
78–89	capsule, labeled as salmon oil
90–92	blend 0% South American oil–100% salmon oil
93–95	blend 100% South American oil–0% salmon oil
96–98	blend 5%
99–101	blend 10%
102–104	blend 20%
105–107	blend 40%
108–112	commercial cod liver oil

[a] Farmed salmon and trout and cod liver from farmed and wild cod. [b] For commercial products, the product declaration is reported.

RESULTS AND DISCUSSION

Classification of the Fish Oil Product: (A) Principal Component Analysis. The results of the PCA analysis for all 112 samples (**Table 1**) are shown in **Figure 1**. The first two principal components account for about 37% of the variance. Sample groups were defined as follows: samples 1−5 (fish oil capsules); samples 6−25 (cod liver oils, known origin); samples 26 and 27 (commercial trout oil, known origin); samples 28−31 (commercial salmon oils, known origin); samples 32−54 (miscellaneous capsules); samples 55−63 (wild cod, Iceland, extracts); samples 64−66 (farmed cod, Norway, extracts); samples 67−76 (wild cod, Barents Sea, extracts); samples 77−

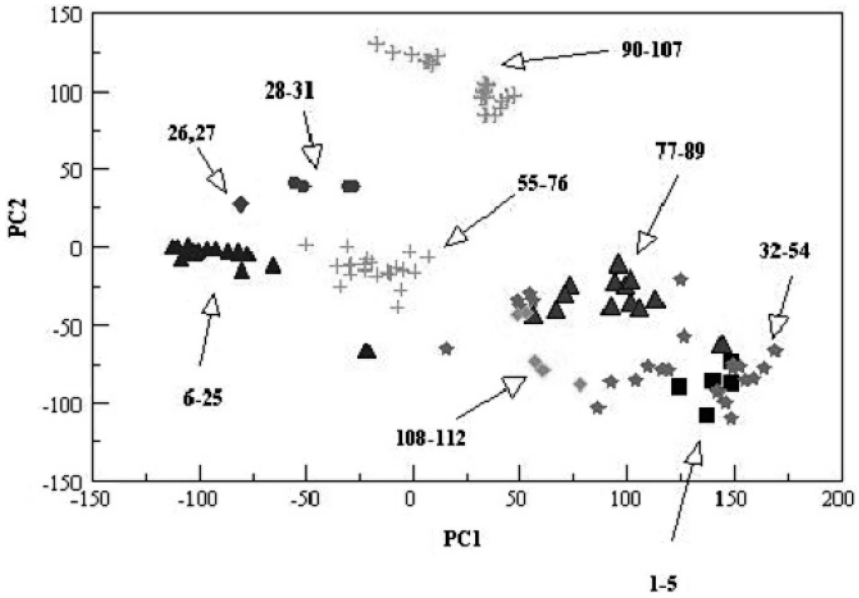

Figure 1. PCA plot of the 112 fish oil samples.

89 (various capsules, described as natural salmon oil); samples 90−107 (blends, commercial salmon oil with South American oil (5−40%)); and samples 108−112 (described as commercial cod liver oil). Unambiguous discrimination among all classes is not possible for this data set, although substantial separation of most classes is achieved. The method discriminated between lipids from farmed and wild cod and between wild fish from Icelandic and Barents seas (data not shown). The samples described as natural salmon oil (32, 33, and 78−89) are clustered together with the commercial South American fish oils and encapsulated fish oil products. The samples described as commercial cod liver oil are also clustered in the same region of the PCA plot, indicating mislabeling and/or adulteration.

(B) Kohonen Neural Network Analysis. Results for a 20 × 20 2D Kohonen map are given in **Figure 2A**. Classification was based on the same categories defined previously for the PCA analyses. There are no assumptions regarding class membership. Samples are grouped together on the basis of their similarity. The method is primarily used for the examination of data sets for which no, or little, a priori knowledge concerning the internal relationships is available. Once the network has been trained, each unit in the Kohonen map might be associated with an object class and then the map may be used for visualization and classification purposes. Samples belonging to apparently related clusters are outlined. As long as interclass variations in chemical shift intensities are greater than experimental (and intraclass) variability, then this approach is quite effective.

Compared to the PCA analysis, there was better differentiation of the defined classes when the Kohonen network analysis was applied. We were able to discriminate between cod liver oil from wild and farmed cod in addition to geographical origin (Icelandic Sea and Barents Sea). Crude and processed cod liver oil of known origin were clustered together. However, the samples described as commercial cod liver oil (108−112) were not clustered together with either the crude oil or the processed cod liver oil samples with known origin. This indicates that the samples could be mixtures of different fish oil or oil from other species than cod. In this data treatment, the commercial salmon oils (two different batches produced from byproducts from the salmon processing industry) are clustered together. Several of the capsules and health foods described as natural salmon oil (32, 33, and 79−89) were not clustered together with the commercial salmon oil of known origin. In the PCA plot (**Figure 1**) these samples are close to the capsules and the commercial South American fish oil. The basis for many of the fish oil capsules on the market is 18/12 oil from South America (the designation 18/12 means that the oil contains 18% 20:5*n*-3 and 12% 22:6*n*-3, where anchovy or sardines are the raw material for the production). Shark liver oil was not clustered with any of the other fish oil groups; this marine oil has a unique composition containing high amounts of squalene that gives rise to unique signals in the ^{13}C NMR spectrum.

(C) Generative Topographic Mapping. The GTM map provides a useful complementary method for visualization of class relationships (*18*), obviates some of the possible deficiencies associated with the Kohonen algorithm, and also gives an opportunity to easily compare spectra from various classes in a systematic way (**Figure 2B**). The definition of a class is largely a functional assignment based on the available empirical evidence. Thus, while trout and salmon and cod may be clearly identified as separate groups, it is also clear that relevant subgroups may be of considerable importance for classification purposes. As an example we note that the cod class in our present data set can actually comprise several significant subsets based on the geographical origin of the samples in question (i.e., Iceland vs Barents Sea). In addition, the wild versus farmed distinction can also be identified from these approaches. Similarly, cod samples obtained at different stages within the processing of the samples can be viewed as a significant cluster when viewed as a group, but may also be further subdivided within the cluster based on the exact stage of processing.

Clusters identified as capsules 1, 2, and 3 have been identified as single groupings (denoted C, H, and J, respectively) in **Figure 2B**, but in reality, these samples are associated with a distribution of chemical shift patterns. Thus, while they have many features in common (hence the clustering tendencies), close inspection of the data and of the origins of these samples indicate subclassifications that appear consistent among certain samples. For instance, in the capsules 2 cluster (denoted as H) we have identified three subgroups (designated I, II, and III). Group I samples were obtained from a South American source, and despite the varied claimed compositions of the samples, they form a relatively tight grouping indicative of their close similarity in composition. Group II appears similar and is

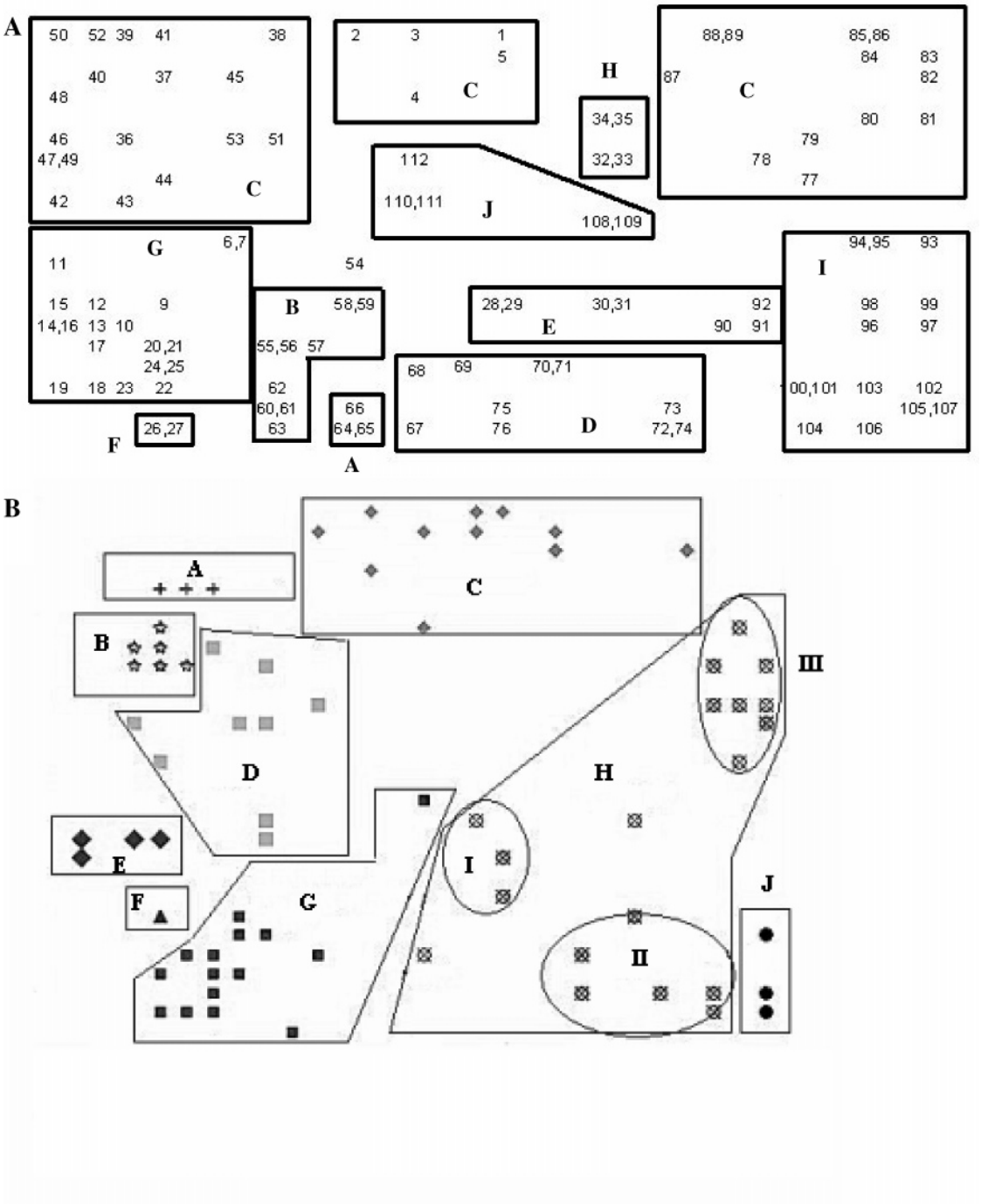

Figure 2. (**A**) 20 × 20 2D Kohonen map: A, farmed cod, Norway; B, wild cod, Iceland; C, capsules 1; D, wild cod, Barent's Sea; E, salmon; F, trout; G, cod liver oils, different processing stages; H, capsules; I, salmon oil blends; J, commercial cod liver oil. (**B**) GTM data analysis. Boundaries drawn are not meant to be quantitative but rather to draw attention to the nature and extent of the clustering. A, farmed cod, Norway; B, wild cod, Iceland; C, capsules 1; D, wild cod, Barent's Sea; E, salmon; F, trout; G, cod liver oils, different processing stages; H, capsules 2, with subsets I, II, and III; J, capsules 3.

characterized by sample descriptions that are also similar (natural fish oils and triacylglycerols). Conversely, Group III samples are described as mixtures of mono, di-, and triacylglycerols, consistent with their clustering tendencies in these plots. Samples from a given vendor may form a single cluster while specific products within the vendor's product line may form significant subgroups. Unsupervised classification indicates significant and consistent groupings of samples of apparently similar origins.

The PCA score plots and Kohonen and GTM maps provide clear opportunities for visualization of clusters within a complex array of sample compositions. While they can be applied as unsupervised multivariate analyses, they can also be used in a supervised manner to make predictions for unknown samples acquired.

(D) Probabilistic Neural Networks. The principles discussed above apply equally to the use of probabilistic neural network analyses. To be consistent, we performed calculations using the same group definitions as for the PCA and Kohonen analyses.

Probabilistic neural network predictions were made for 86 samples that were randomly selected for the training set, while 26 were held back for the validation calculations. Training results (leave-one-out cross-validation results) showed 98.8%

($n = 85$) classified correctly and 1.2% ($n = 1$) classified incorrectly. Validation results were also excellent [96.2% ($n = 25$) classified correctly; 3.8% ($n = 1$) classified incorrectly]. Furthermore, samples for each group (training, test, and validation) were randomly selected several times and calculations were made with systematic variations in the number of validation samples. Typically about >96% of the predictions were correct, the majority of errors being associated with the clearly unique samples (e.g., shark oil) that belong in their own distinct classifications. In the full leave-one-out cross-validation calculation using all samples, 111 of 112 samples were correctly classified; sample 54 (shark oil) was not classified, as opposed to being misclassified, by the PNN, as its characteristic spectrum was considered too different from any other defined class to be assigned. This points out the need for optimizing neural networks in such a way as to set an acceptable threshold for a prediction to be made to minimize errors. More robust training sets may also be designed by taking original data sets and systematically varying parameters known to vary in routine data acquisitions (including peak heights and peak positions) to create new training samples.

Both chemical shift intensity data and scores from PCA were used as input for the PNN analyses with approximately equal success. Less than 10 scores, chosen from calculations of variable relevance during the PNN calculations, were adequate to achieve >96% prediction accuracy. Similarly good predictions were achieved through the use of as few as 13 chemical shifts as input, these shifts also being selected from inspection of measures of variable relevance made during the neural network modeling process.

Blend Analysis. We have already demonstrated that the blend samples occupy a distinct group in the unsupervised analyses. In order to determine whether neural networks could be used to quantify the relative amount of a material used to adulterate true salmon oil, we prepared a series of samples with systematically varying proportions of real salmon oil and diluent oil (commercial South American fish oil). Three sets of samples were prepared at four different concentrations of diluent: 5%, 10%, 20%, and 40%. Multiple samples at the same concentration were prepared as a way of assessing the sensitivity of the calculations to preparation procedures and unavoidable variations in acquisition conditions. It would appear that there are, for instance, minor variations in relative signal intensities for some resonances as a consequence of differential signal overlap, minor solvent effects, and the inherently high resolution of the spectra. Generalized regression neural networks (GRNN) work in a similar fashion to PNNs but perform regression (*20*) rather than classification tasks. The results indicate that the degree of adulteration in a mixture can be quantified to within about ±1% (absolute); the limits to which this can be taken will ultimately depend on the number of samples used in the training set and the signal-to-noise ratios readily attainable (*21*, *22*). Principal components regression (PCR) was also applied, yielding a root-mean-square error of about 1.8%.

While unsupervised visualization, limited by being 2D representations of multidimensional data, may indicate that samples of one class may be somewhat confused with samples of another class, the actual differences may be more clearly delineated by PNN analyses. For instance, if one deliberately mislabels any of the blend samples and then repeats the PNN analysis, the sample that was given the inaccurate target class assignment was always predicted to belong to the blend class. The inherent characteristics of the shift intensities associated with the mixtures are in no way confused with the assignment of those samples that are known to be of true salmon oil origin.

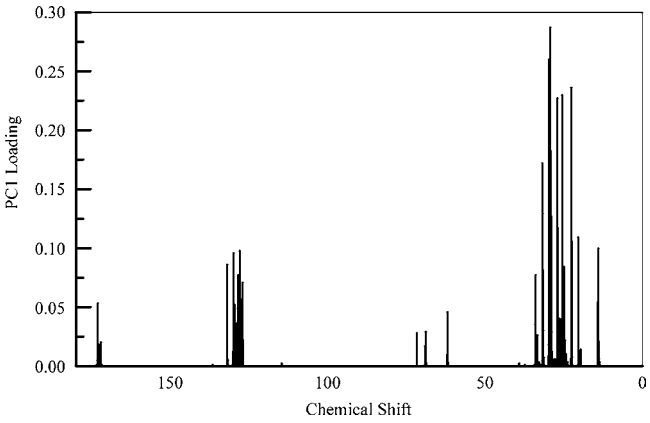

Figure 3. Principal component analysis loadings for all chemical shifts for first principal component, PC1.

Interpretation of ^{13}C NMR Spectra. ^{13}C NMR spectroscopy provides a fingerprint of the specific lipids and oils analyzed. This process resulted in an initial data matrix, that is, unique chemical shifts, of approximately 282 points for the 112 samples investigated. Detailed examination of the original data revealed small variations in resonance positions of comparable lines in different samples. Variations in the positions may arise from differences in relative concentrations, temperature effects, pH, magnetic field homogeneity variations, and shimming effects. Corrections were made, as required, in all samples to optimize consistency among the peak positions. In the present case each spectrum was inspected visually and resonances were assigned and modified by hand, although recent advances in automated peak alignment algorithms reduce the necessity for our procedure for future work (*23*). The composition and quality of the fish oil studied varied significantly according to the raw material and how it was produced or manufactured. Many of the samples in this study are clearly composed of a few clearly identifiable components (e.g., pure triacylglycerols), whereas others are mixtures of mono-, di-, and triacylglycerols (commercial health food products with high concentration of 20:5*n*-3 and 22:6*n*-3), all of which exhibit characteristic peaks in the NMR spectra. Similarly, many samples contain ethyl esters and related functional groups that, although diagnostic, may also overlap and even obscure other components. While characteristic resonances may be observed for various classes or groupings of samples, such visual inspection is generally insufficient to unambiguously identify any given class. Complex, nonlinear relationships among metabolite concentrations are inherently present in biological systems and necessitate the use of methods such as those used in this study. In order to address the question of the possible existence of specific markers associated with any defined class, we inspected loading plots from PCA and variable relevance plots from the PNN analyses.

Figure 3 shows the PCA loading plot for the first principal component. Although the aliphatic carbons exhibit relatively higher loadings than the aromatic carbons, the majority of the resonances could be considered as significant. Although the score plot exhibits reasonable differentiation of various classes, the origins of the positions in the score plot are complex; it appears that the positions of the samples in this plot depend on a large number of shifts with similar loadings. The potential importance of applying more sophisticated approaches for biomarker identification is provided by inspection of the most significant variables arising from the PNN analysis (which, by

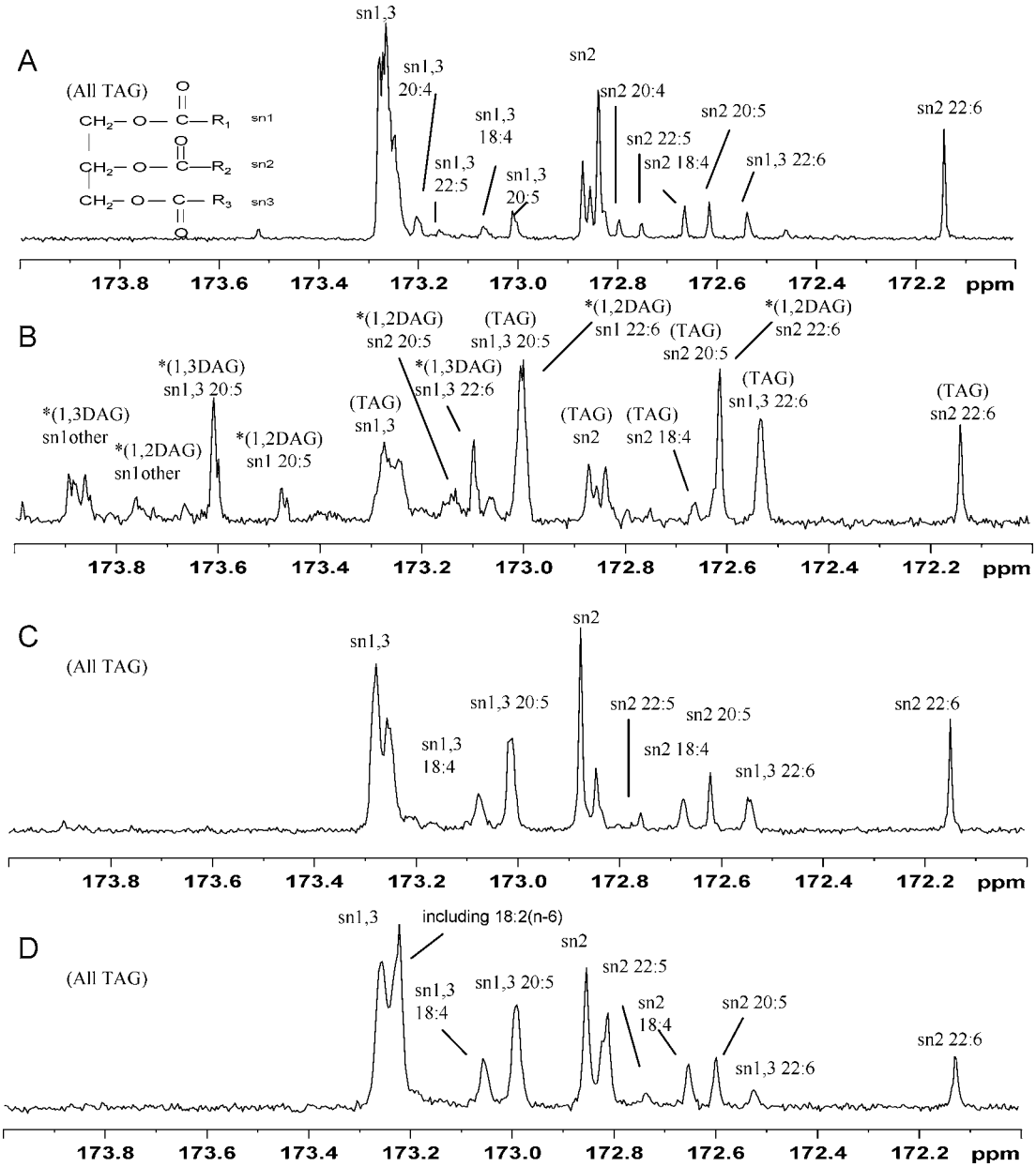

Figure 4. Carbonyl region of ^{13}C NMR spectra of extracted salmon lipids (**A**) and three commercial fish oil products labeled as follows: concentrated fish oil (**B**), fish oil from deep-water fish (**C**), and 20:5*n*-3-rich fish oil with evening primrose oil (**D**). Unique signals from fatty acids in triacylglycerols (TAG) and diacylglycerols (DAG) are assigned. The position of the fatty acids (R) in acylglycerols is designated *sn*-1, *sn*-2, or *sn*-3. Asterisks indicate tentative assignment of the individual fatty acids in diacylglycerols.

use of the genetic analysis algorithm, explicitly takes into account nonlinear relationships among variable contributions); as few as 13 chemical shifts (172.54, 172.50, 172.46, 129.66, 128.06, 127.93, 129.72, 68.85, 34.13, 33.80, 31.87, 29.71, and 29.64 ppm) account for all of the defined classifications. However, the nature of the genetic analysis requires that multiple runs be made, from which the most frequently occurring variables (chemical shifts) yielding the best predictions/models be identified. Many different nonlinear combinations of resonance intensities may yield accurate predictions. While successful classification models employing neural network methods do not require that known assignments be made, we provide an overview of several systematic chemical shift variations that can be identified within several regions of interest.

The interpretation of the ^{13}C spectra of some of the products examined is given in **Figures 4**−**6** and is based on published data (*5*, *7*, *11*, *24*−*26*). A ^{13}C spectrum of fish oil gives multicomponent information about the fatty acid profile (*27*), content of polar lipids (*28*), lipid class composition (*24*), and positional distribution of fatty acids in the triacylglycerols/phospholipids (*7*, *28*). The various regions of the spectra contain different information regarding the lipid composition.

Carbonyl Region. In this region (174.0−172.0 ppm), signals correspond to different carbonyl carbons from fatty acids present in the sample (**Figure 4**). In general, the chemical shift in this region depends on the type of glycerol ester (i.e., triacylglycerols, diacylglycerols, or monoacylglycerols), the stereospecific conformation [fatty acids in α (*sn*-1,3) or β (*sn*-2) position in acylglycerols), and the distance (△) of the nearest double bond to the carbonyl carbon. Since the positional distribution of fatty acids in triacylglycerols seems to be unique for each species (*7*), the carbonyl region can be of value for fish species identification. Additionally, it may be possible to discover if

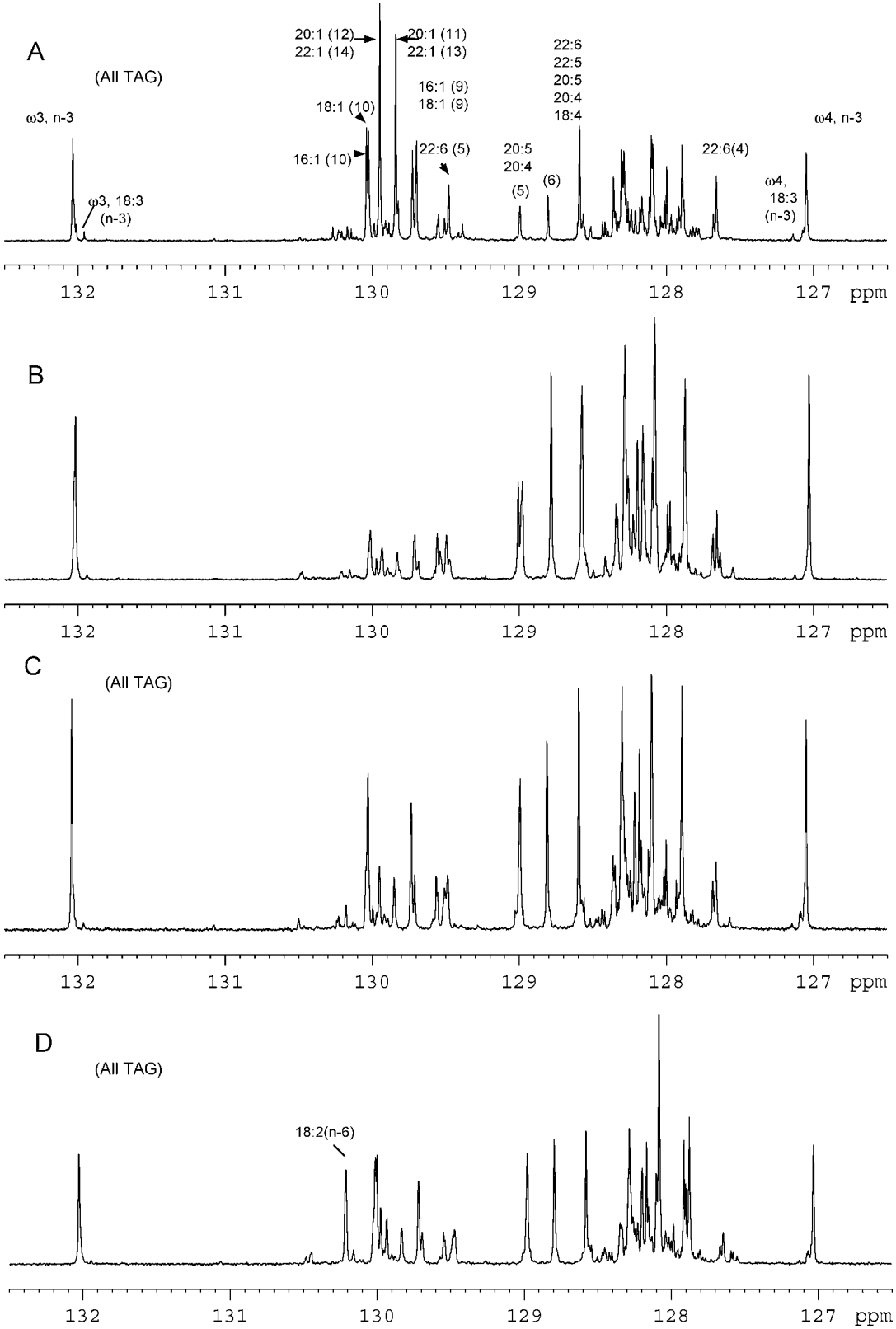

Figure 5. Olefinic region of the ^{13}C NMR spectra of extracted salmon lipids (**A**) and three health oil products labeled as follows: concentrated fish oil (B), fish oil from deep-water fish (**C**), and 20:5*n*-3-rich fish oil with evening primrose oil (**D**). ω; carbon number from the acyl chain.

natural oil has been chemically modified or oil added from another species from this region of the spectra.

Figure 4 shows an expansion of the carbonyl region of the ^{13}C spectrum of salmon oil (**A**) and three commercial fish oil products labeled as concentrated fish oil (**B**), fish oil from deep-water fish (**C**), and 20:5*n*-3-rich fish oil with evening primrose oil added (**D**). Characteristic peaks from the main *n*-3 fatty acids in *sn*-1,3 and *sn*-2 positions of the triacylglycerols are identified in the carbonyl region for all samples. The ^{13}C spectrum of oil extracted from farmed salmon gives a signal profile (**Figure 4A**) that is in accordance with the ^{13}C NMR profile given in European Pharmacopoeia 2005 (*3*). The commercial concen-

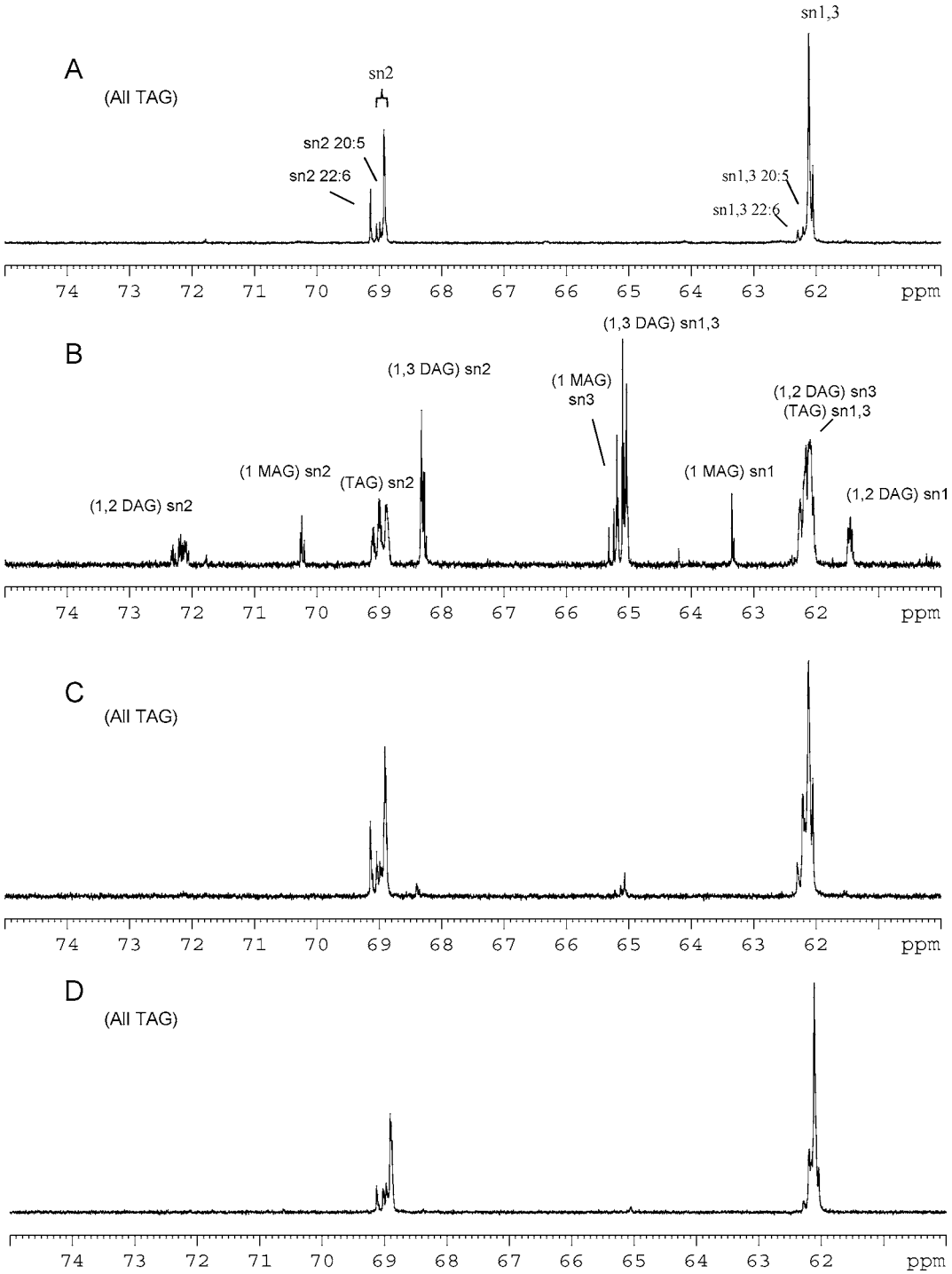

Figure 6. Glycerol region of the ^{13}C NMR spectra of extracted salmon lipids (**A**) and three commercial fish oil products labeled as follows: concentrated fish oil (**B**), fish oil from deep-water fish (**C**), and 20:5*n*-3-rich fish oil with evening primrose oil (**D**).

trated fish oil product (**Figure 4B**) displays a number of peaks (173.1 and several peaks in the region 173.4−174.0 ppm) assigned to monoacylglycerols and diacylglycerols (*25*, *26*). These findings confirm that this oil has gone through chemical modifications to reach high concentration of *n*-3 fatty acids. Concentrated fish oil products are usually prepared by esterification (*29*) or transesterification (*30*) with subsequent physicochemical purification processes. In addition, ethyl esters can be formed during concentration of fish oil, and from previous ^{13}C NMR studies the ethyl esters of 20:5*n*-3 and 22:6*n*-3 have been identified at 60.06 and 60.25 ppm (assigned to CH_3CH_2-atoms) (*11*). However, in this study the absence of these signals in the concentrated product **B** excludes the presence of 20:5*n*-3 and 22:6*n*-3 ethyl esters in this oil. Furthermore, free fatty acids, which display signals in the region 181−176 ppm (*5*), were not detected in the spectra obtained.

By comparison of the carbonyl profiles given in **Figure 4**, product **B** has gone through chemical treatments and product **D** is a mixture of oils (contains evening primrose oil), while both products **A** and **C** are based on natural fish oil with the species-authentic positional distribution of fatty acids (*3*, *4*, *7*). However, the carbonyl profile of the salmon oil differs from the profile of the fish oil products **C** and **D**, especially for the positional distribution of 20:5*n*-3 and 22:6*n*-3, where 20:5*n*-3

is preferentially esterified in the *sn*-1,3 position in samples **C** and **D** [also verified by information obtained from the *C*2 carbon region of 20:5*n*-3 (33.4−33.7 ppm; data not shown)]. In general, stereospecific analysis of triacylglycerols has shown that the positional distribution of fatty acids varies among plant and animal fat (*31*, *32*). In vegetable oils, linoleic acid is mainly present in the *sn*-2 position, while saturated fatty acids are concentrated in the *sn*-1 and *sn*-3 positions (*33*). In animals (and fish), the stereospecificity varies among species (*7*, *31*) and among parts of the same animal. Studies of fish has shown that 22:6*n*-3 are preferentially esterified at the *sn*-2 position, while saturated and monounsaturated fatty acids are preferably esterified at the *sn*-1,3 positions (*31*, *34*).

Olefinic Region. From the olefinic region (132.5−126.5 ppm) of the ^{13}C NMR spectrum (**Figure 5**), information about unsaturated fatty acid can be obtained (*5*). Since fatty acid composition varies among species and geographical origin, addition of different types of oil can be seen in this region. Olefinic carbon atoms in *n*-3 fatty acids display characteristic peaks because of the influence on the chemical shift from any methyl end group in the vicinity. The two resonances at 132.0 (ω3 carbon) and 127.0 ppm (ω4 carbon), respectively, are unique for fatty acids with a double bond at *n*-3 (ω3) position, and these peaks allow quantification of the relative concentration of *n*-3 fatty acids in lipid mixtures (*5*). Signals from ω3 and ω4 carbons in 18:3*n*-3 are separated from the rest and display peaks at 131.9 and 127.1 ppm. The salmon oil (**Figure 5A**) has higher concentrations of long-chain monounsaturated fatty acids (20:1, 22:1) than the commercial fish oil products (**Figure 5B-D).** The commercial products, on the other hand, are richer in polyunsaturated fatty acids. This can be seen by comparing the relative intensity of peaks arising mainly from polyunsaturated fatty acids (128.6−127.8 ppm) to the intensity of the peaks in the region of monounsaturated fatty acids (130.1−129.8 ppm) (*5*, *11*). Commercial South American fish oil (mainly sardine/anchovy oil) has low levels of monounsaturated fatty acids (*35*). In general, the fatty acid composition of fish (as liver or muscle) reflects the lipid supplied in the diet (*36*−*38*). The enrichment of evening primrose oil in health oil **D** can be verified by the higher level of 18*n*-6 fatty acid (130.2 ppm); evening primrose oil is rich in this fatty acid (*39*).

Glycerol Carbon Region. Figure 6 shows an expansion of the region between 61 and 72 ppm where resonances from glycerol carbons (in mono-, di-, and triacylglycerols) appear. As for the carbonyl region, the glycerol region gives valuable information about chemical modifications a natural crude oil sample might have gone through. The glycerol carbon resonances are influenced by the stereospecific conformation and the nature of the esterified fatty acid, namely, the distance to the nearest double bond. For triacylglycerols, the resonances from glycerol carbons esterified at *sn*-1,3 and *sn*-2 positions are separated by approximately 7 ppm. Peaks from *n*-3 fatty acids are separated from the other fatty acids to a certain degree; however, only signals from 22:6*n*-3 are fully resolved to allow quantification of the *sn*-1,3 and *sn*-2 positions (*5*). Consistent with the findings from the carbonyl regions, observations in this region confirm that several glycerol ester species in addition to triacylglycerols are present in the concentrated product (**B**); specifically 1-monoacylglycerols (63.3, 65.3, and 70.3 ppm), *sn*-1,2-diacylglycerols (61.4, 62.2, and 72.2 ppm), and 1,3-diacylglycerols (65.1 and 68.3 ppm) (*24*). This spectrum is more complex than for pure triacylglycerols, as chains in *sn*-1 and *sn*-3 positions can be distinguished (i.e., in 1,2- diacylglycerols). In "symmetrical" molecules (2-monoacylglycerols, 1,3- diacylglycerols, and triacylglycerols), ^{13}C NMR cannot distinguish between fatty acids in *sn*-1 and *sn*-3 positions (*31*).

In conclusion, unsupervised multivariate analyses, Kohonen neural networks, and generative topographic mapping provide excellent visual discrimination among fish oil classes and would also be suitable for supervised classification. Trout, salmon, and cod oils can be completely and correctly classified. In addition, different cod liver oil samples and the corresponding compositions observed as a function of processing conditions are readily differentiated from other oils studied to date. Samples reported to be salmon oils, but which are not, are clearly represented as being different from true salmon oils. Products from selected vendors may be consistently identified through the use of these methodologies, and products that have been inappropriately labeled can be readily identified. Neural network models can be designed (1) from whole spectra, (2) from selected spectral regions, (3) from selected chemical shifts chosen by genetic algorithms, or (4) from PCA scores derived from the original variables. Several diverse classification approaches work well for these data, which also suggests that employing a combination of predictive methods (consensus analysis) might provide a robust approach for practical implementation of these methods.

LITERATURE CITED

(1) Jude, S.; Roger, S.; Martel, E.; Besson, P.; Richard, S.; Bougnoux, P.; Champeroux, P.; Le Guennec, J. Dietary long-chain omega-3 fatty acids of marine origin: A comparison of their protective effects on coronary heart disease and breast cancers. *Prog. Biophys. Mol. Biol.* **2006**, *90*, 299−325.

(2) Simopoulos, A. P. The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomed. Pharmacother.* **2002**, *56*, 365−379.

(3) Council of Europe, European Directorate for the Quality of Medicines (EDQM), Salmon oil, farmed (05/2005:1910), European Pharmacopoeia, 2005.

(4) Council of Europe, European Directorate for the Quality of Medicines (EDQM), Cod liver oil 4.8 (04/2003:1192 and 1193), Fish oil rich in omega-3-acids 4.7 (01/2003:1912), Omega-3-acid ethylesters 4.7 (01/2003:2063 and 1250) omega-3-acid triacylglycerols 4.7 (01/2003:1352), European Pharmacopoeia, 2003.

(5) Aursand, M.; Grasdalen, H. Interpretation of the ^{13}C-NMR spectra of omega-3 fatty acids and lipid extracted from the white muscle of Atlantic Salmon (*Salmo salar). Chem. Phys. Lipids* **1992**, *62*, 239−251.

(6) Aursand, M.; Rainuzzo, J. R.; Grasdalen, H. Quantitative high-resolution ^{13}C and ^{1}H nuclear magnetic resonance of omega-3 fatty acids from white muscle of Atlantic salmon (*Salmo salar). J. Am. Oil Chem. Soc.* **1993**, *70*, 971−981.

(7) Aursand, M.; Jørgensen, L.; Grasdalen, H. Positional distribution of n-3 fatty acids in marine lipid triacylglycerols by high-resolution ^{13}C nuclear magnetic resonance spectroscopy. *J. Am. Oil Chem. Soc.* **1995**, *72*, 293−297.

(8) Aursand, M.; Mabon, F.; Martin, G. J. High-resolution ^{1}H and ^{2}H NMR spectroscopy of pure essential fatty acids from plants and animals. *Magn. Reson. Chem.* **1997**, *35*, S91−S100.

(9) Aursand, M.; Mabon, F.; Martin, G. J. Characterization of farmed and wild salmon (*Salmo salar*) by combined use of compositional and isotopic analysis. *J. Am. Oil Chem. Soc.* **2000**, *77*, 659−666.

(10) Aursand, M.; Axelson, D. Origin recognition of wild and farmed salmon (Norway and Scotland) using ^{13}C NMR spectroscopy in combination with pattern recognition techniques. In: *Magnetic Resonance in Food Science: A View to the Future*; Webb, G. A., Belton, P. S., Gil, A. M., Delgadillo, I., Eds.; The Royal Society of Chemistry: Cambridge, U.K., 2001; pp 227−231.

(11) Siddiqui, N.; Sim, J.; Silwood, C. J. L.; Toms, H.; Iles, R. A.; Grootveld, M. Multicomponent analysis of encapsulated marine oil supplements using high-resolution ^{1}H and ^{13}C NMR techniques. *J. Lipid Res.* **2003**, *44*, 2406−2427.
(12) Rezzi, S.; Axelson, D. E.; Héberger, K.; Reniero, F.; Mariani, C.; Guillou, C. Classification of olive oils using high throughput flow ^{1}H NMR fingerprinting with principal component analysis, linear discriminant analysis and probabilistic neural networks. *Anal. Chim. Acta* **2005**, *552*, 13−24.
(13) Medina, I.; Aubourg, S. P.; Martin, R. P. Species differentiation by multivariate analysis of phospholipids from canned Atlantic tuna. *J. Agric. Food Chem.* **1997**, *45*, 2495−2499.
(14) Bligh, E. G.; Dyer, W. A rapid method to total lipid extraction and purification. *J. Can. Biochem. Physiol.* **1959**, *37*, 911−917.
(15) Aursand, M.; Bleivik, B.; Rainuzzo, J.; Mohr, V. Lipid distribution and composition of commercially farmed Atlantic salmon (*Salmo salar*). *J. Sci. Food Agric.* **1994**, *64*, 239−248.
(16) Jolliffe, I. T. *Principal Component Analysis*; Springer-Verlag: New York, 1986.
(17) Kohonen, T. *Self-Organizing Maps*, 2nd ed.; Springer-Verlag: Berlin, Germany, 1997.
(18) Bishop, C. M.; Svensen, M.; Williams, C. K. I. *GTM: The generative topographic mapping*; Technical Report NCRG/96/015; Neural Computing Research Group, Aston University: Birmingham, U.K., 1996.
(19) Specht, D. F. Probabilistic neural networks and the polynomial adaline as complementary techniques for classification. *IEEE Trans. Neural Network* **1990**, *1*, 111−121.
(20) Specht, D. A generalized regression neural network. *IEEE Trans. Neural Networks* **1991**, *2*, 568−576.
(21) Amendolia, S. R.; Dopplu, A.; Ganadu, M. L.; Lubinu, G. Classification and quantitation of ^{1}H NMR spectra of alditols binary mixtures using artifical neural networks. *Anal. Chem.* **1998**, *70*, 1249−1254.
(22) Defernez, M.; Colquhoun, I. J. Factors affecting the robustness of metabolite fingerprinting using ^{1}H NMR spectra. *Phytochemistry* **2003**, *62*, 1009−1017.
(23) Forshed, J.; Torgrip, R. J. O.; Aberg, K. M.; Karlberg, B.; Jacobsson, S. P. A comparison of methods for alignment of NMR peaks in the context of cluster analysis. *J. Pharm. Biomed.* **2005**, *38*, 824−832
(24) Gunstone, F. D. ^{13}C NMR studies of mono-, di- and triacylglycerols leading to qualitative and semiquantitative information about mixtures of these glycerol esters. *Chem. Phys. Lipids* **1991**, *58*, 219−224.
(25) Medina, I.; Sacchi, R. Acyl stereospecific analysis of tuna phospholipids via high resolution ^{13}C NMR spectroscopy. *Chem. Phys. Lipids* **1994**, *70*, 53−61.
(26) Zamora, R.; Gómez, G.; Hidalgo, F. J. Classification of vegetable oils by high resolution ^{13}C NMR spectroscopy using chromatographically obtained oil fractions. *J. Am. Oil Chem. Soc.* **2002**, *79*, 267−272.
(27) Zamora, R.; Hidalgo, F. J. Determination of fatty acid composition and oxidation in fish oils by high resolution nuclear magnetic resonance spectroscopy. In *Modern Magnetic Resonance*; Webb, G. A., Ed.; Springer-Verlag: Berlin, 2006.
(28) Falch, E.; Størseth, T. R.; Aursand, M. Multi-component analysis of marine lipids in fish gonads with emphasis on phospholipids using high-resolution NMR spectroscopy. *Chem. Phys. Lipids* **2006** (in press).
(29) Torres, C. F.; Garcia, H. S.; Ries, J. J.; Hill, C. G. Esterification of glycerol with conjugated linoleic acid and long-chain fatty acids from fish oil. *J. Am. Oil Chem. Soc.* **2001**, *78*, 1093−1098.
(30) Haraldsson, G. G.; Kristinsson, B.; Sigurdardottir, R.; Gudmundsson, G.; Breivik, H. The preparation of concentrates of eicosapentaenoic and docosahexaenoic acids by lipase-catalyzed transesterification of fish oil with ethanol. *J. Am. Oil Chem. Soc.* **1997**, *74*, 1419−1424.
(31) Gunstone, F. D.; Seth, S. A study of the distribution of eicosapentaenoic acid and docosahexaenoic acid between the α and β glycerol chains in fish oils by ^{13}C NMR spectroscopy. *Chem. Phys. Lipids* **1994**, *72*, 119−126.
(32) Ando, Y.; Ota, T.; Matsuhira, Y.; Yazawa, K. Stereospecific analysis of triacyl-*sn*-glycerols in docosahexaenoic acid-rich fish oils. *J. Am. Oil Chem. Soc*. **1996**, *73*, 483−487.
(33) Mannina, L.; Luchinat, C.; Emanuele, M. C.; Segre, A. Acyl positional distribution of glycerol tri-esters in vegetable oils: a C-13 NMR study. *Chem. Phys. Lipids* **1999**, *103*, 47−55.
(34) Ando, Y.; Satake, M.; Takahashi, Y. Reinvestigation of positional distribution of fatty acids in docosahexaenoic acid-rich fish oil triacyl-*sn*-glycerols. *Lipids* **2000**, *35*, 579−582.
(35) Menoyo, D.; Lopez-Bote, C. J.; Bautista, J. M.; Obach, A. Herring vs. anchovy oils in salmon feeding. *Aquat. Living Resour.* **2002**, *15*, 217−223.
(36) Bell, J. G.; Henderson, R. J.; Tocher, D. R.; McGhee, F.; Dick, J. R.; Porter, A.; Smullen, R. P.; Sargent, J. R. Substituting fish oil with crude palm oil in the diet of Atlantic salmon (*Salmo salar*) affects muscle fatty acid composition and hepatic fatty acid metabolism. *J. Nutr.* **2002**, *132*, 222−230.
(37) Kirsch, P, E.; Iverson, S. J.; Bowen, W. D.; Kerr, S. R.; Ackman, R. G. Dietary effects on the fatty acid signature of whole Atlantic cod (*Gadus morhua). Can. J. Fish. Aquat. Sci.* **1998**, *55*, 1378−1386.
(38) Hemre, G. I.; Sandnes, K. Effect of dietary lipid level on muscle composition in Atlantic salmon (*Salmo salar*). *Aquacult. Nutr.* **1999**, *5*, 9−16.
(39) Peiretti, P. G.; Palmegiano, G. B.; Masoero, G. Chemical composition, organic matter digestibility and fatty acid content of evening primrose (*Oenothera paradoxa)* during its growth cycle. *Anim. Feed Sci. Technol.* **2004**, *116*, 293−299.

Received for review June 22, 2006. Revised manuscript received October 27, 2006. Accepted October 31, 2006. This work was carried out as a part of a project supported by the Norwegian Research Council (NFR Projects 115147/112 and 146932/130).

JF061754L