Electronic Supplementary Data

Principal component analysis (PCA), partial least squares (PLS) were performed in XLStat Pro version 2013.6.03 (Addinsoft, Paris, France). PCA, PLS and PCR were performed on unweighted and mean centred variables. The ECL data-sets were examined in a raw format by MVS (PCA and PLS) with the RMSE and coefficients of multiple determination providing a measure of accuracy and model fitting. A second visual approach for structural identification, based upon a plot of the gradient (from a line-of-best fit of ECL variation with Temperature) *vs* calculated mid-point ECL was generated with MS Excel. The objective interpretation of this derived data-set was similarly performed using the XlStat Pro software with RMSE and R^2 values for accuracy and model fitting quality as determinants of efficacy.

Identification PUFA: Innowax PCA

The ECL values were set as variables and the fatty acids as objects for in the generation of a PCA plot. As can be seen in **Fig. 6** (and even slightly better with the Vf-23 and SP-2560 columns), there exists a trend with the FAMEs separated along Factor-1 and Factor-2, associated with degree of unsaturation and chain-length (**see Fig. 6**). However, in this plot it is clearly the case that the fatty acids in the reference mixtures do not align very well from straight lines, (in both factor planes). A possible explanation is due to the fact the Factor-1 explains 100.00% of the variation, with Factor-2 of negligible importance (<0.00% variation). Rather than a visual interpretation of this PCA plot we have qualified the assertion of this column for structural attribute interpretation with PLS-regression results of Table 1.

Identification PUFA: Vf-23 PCA

A similar PCA plot generated as for the SP-2560 indicates that this column polarity changes substantially more with temperature with the respective F1-99.98%, and F2-0.02% (see **Fig. 7**) indicating the greater impact of double-bond number to interact with the stationary phase, and separation based upon differences in volatility is lesser importance in comparison to the Innowax. The lines on the Fig. 5 appear very linear and clearly demarcated by structural features in comparison to the Innowax. We conclude that this is due to the fact the individual variability of ECL (%RSD) for a given FAME was tighter relative to the rate-of-change for double-bond structure. This is borne out by the fact a PLS regression model is better able to accurately predict the chain-length on the Vf-23 in comparison to the Innowax.

Identification PUFA: SP-2560 PCA

An example PCA plots on this data-set is shown in **Fig. 8**, whilst both **Fig. 6 and 7** reveal structural trends for fatty acids there is a greatly improved separation of Factor-2, with only a slight diminution in the explanation of the data-set for both factor components. **Fig. 8** has a F1 explaining 99.96% and F2 explaining 0.04% of the total variation (*vs*. 99.98 and 0.02 correspondingly for the Vf-23). The number of double-bonds is well separated on F2 with the majority of variation on F1 explained by chain-length. The positions of the analytes are especially well aligned along trend-lines in both chain-length and especially double-bond number.

As is the case with the Innowax and Vf-23 columns, there is a small but constant difference evident between the *n*-3, *n*-6 and *n*-9 structural isomers. Fatty acids with similar values along Factor-1 1 will have similar ECL values in most cases, and may therefore be confused when the identifications are based on retention characteristics alone. We found a significant number of analytes overlapped on F1 alone e.g. C20:1n-9 with C19:2, C24:1n-9 with C22:4n-6, C22:1n-9 with C20:3n-3, but all are well separated along Factor 2.

The trends seen in **Fig. 6 to 8** are similar to the trends seen when retention data from two columns with different polarity are compared and with **Fig. 3 to 5**. The unsaturated fatty acids can be identified visually by their positions in **Fig. 3 to 5** relative to the fatty acids in the reference mixture. The identifications have been controlled by mass spectrometry [19,20]. We trialled the PCA model to identify unknown FAMEs. We spiked a series of unknowns into our lipid mixture (fish-oil), and correctly identified these as C17:1, C19:2, C21:2*n-6* and C22:5*n-6*. The total ion chromatogram of the spiked analytes and separate runs of these analytes independently, confirmed the identities of the peaks and calculated ECL's and associated shifts.

In all cases the locations of the unknowns were found to agree precisely on the PCA plot as expected. In comparison to the approach Mjos used, the operation of the column under isothermal conditions is, as we expected, to be more consistent since it difficult for the GC to control a temperature ramp reproducibly. Furthermore, ECL's have been suggested to be more accurately determined by **Equation 1**, under isothermal conditions. The GC instrument we used is also a later member of the series of Agilent instruments and the electronic flow control is likely to be better.

PLS Determination of Chain Length and Double-Bond Number

We chose two PLS approaches:

1. PLS on the entire four isothermally derived datasets including the saturated FAMEs
2. PLS on the gradient of best-fitting lines obtained from plots of ECL change with temperature (δECL/δT) as variable 1 and calculated ECL (195 C) as variable 2.

The first approach is discussed within Section: PLS on four isothermal oven temperature derived ECLs), whilst the second is discussed in Section: Plots of δECL/δT vs Single Temperature ECL.

PLS on four isothermal oven temperature derived ECLs

PLS regressions[26] were performed with chain length and number of double bonds as dependent variables, and the ECL data as independent variables. We used the datasets found in Table 1 applying the PLS to generate models that gave correct predictions for the C18–C22 PUFA structure (double-bonds and chain-length) evaluating the quantitative efficacy of each model for each column to achieve this task. Not all calibration standards were included in the PLS regression model. Within Xlstat the automated component selection was selected, and on the basis of results for both chain length and double bonds in each of the three column cases, it was found that a model based on two PLS components would give the most accurate predictions for all three columns (little to no improvement in model prediction was found by increasing the component number). The calibration comprised of all fatty acids in the standards and reference mixture(s) and root mean squared error (RMSE) was found to vary

depending upon the polarity of the column used (particularly in relation to the bond-length prediction). The RMSE is synonymous with the standard deviation of the residuals. In all three columns tested, there was a random distribution of the residuals observed for chain-length and double-bond number, indicative of a normal distribution of the ECL data. The PLS model for both chain-length and double-bond number, provided non-integer values, hence the predicted values are rounded to the nearest integer. The incorrect assignation of a structural feature would occur when the RMSE exceeds 0.5.[13]. Based upon this index of error, then it can be inferred that the predictive accuracy, can be estimated from the root mean square error (RMSE)[27] of the validation results. The prediction of chain-lengths of mid- to long-chain mono – and PUFA are listed in **Table's 10, 11 and 12** for the Innowax, Vf-23 and SP-2560 columns respectively.

Chain-Length Prediction

For the variable chain-length the accuracy of prediction was 100% for the Innowax (RMSE=0.229, R^2=0.990), SP-2560 (RMSE=0.145, R^2=0.995) and Vf-23 (RMSE=0.141, R^2=0.996) columns. This result was unexpected as the rate of change in ECL's was highest with the SP-2560 column. Only a fortuitous increase in the numbers of overlapping FAMEs in the SP-2560 column, can explain the decreased accuracy of the model in comparison to the Vf-23, to explain this unexpected result. Nevertheless, the results indicate that for the chain-length variable, all three columns can be used. All chain lengths in the reference mixture were correctly predicted; the chain lengths of the analytes in test samples agreed with the interpretation of **Fig's. 3 to 5 and 6 to 8** and information from the mass spectra.

Double-bond number Prediction

The accuracy of predictions of the model to infer the number of double bonds for a given FAME, was very high for both the Vf-23 and SP-2560 cyanoalkyl based columns (and exceeded the accuracy achieved by Mjos[13]. The accuracy for double-bond prediction was ~100% for the Vf-23 (RMSE=0.096, R^2=0.996) and ~100% for the SP-2560 (RMSE=0.105, R^2=0.996). In the case of the Innowax the predictive accuracy is reduced to 94% (R^2=0.980, RMSE=0.214, incorrectly predicting DHA to have 5 double-bonds), and hence is unsuitable for determination of double-bond length **(Table 10)**. The model predicted correct values also when applied on monounsaturates, but negative values (≤0.5) when applied on the saturated fatty acids. Mjos predicted that the polarity of PEG columns is practically unaltered by temperature[12, 13], and indeed the changes in ECL values are small, these results indicate however the Innowax column type may be used to predict the double-bond number, with the notable exception of C22:6*n-3*. **Tables 10, 11 and 12** contain details on the ability of double-bond prediction for the Innowax, Vf-23 and SP-2560 columns respectively.

Table 10: PLS predictions of chain length and number of double bonds based on the ECL values in Table 1 for Innowax column

Innowax Column PLS Predictions for chain-length and double bonds				
	Predicted	**Rounded**	**Predicted**	**Rounded**
C16:1*n-9**	15.941	16	1.061	1
C17:1*n-9*≠	16.975	17	0.972	1
C20:1*n-9**	19.840	20	1.158	1

C22:1*n-9**	21.865	22	1.099	1
C24:1*n-9**	23.886	24	1.063	1
C18:2*n-6**	18.016	18	1.962	2
C20:2*n-6**	20.004	20	1.998	2
C22:2*n-6*≠	22.018	22	1.985	2
C18:3*n-6**	17.993	18	2.885	3
C18:3*n-3*≠	18.416	18	2.644	3
C20:3*n-6**	19.825	20	3.227	3
C20:3*n-3**	20.376	20	2.762	3
C22:3*n-3**	22.389	22	2.794	3
C20:4*n-6**	19.689	20	4.196	4
C22:4*n-6**	21.681	22	4.232	4
C20:5*n-3*≠	19.934	20	5.302	5
C22:5*n-3**	21.986	22	5.184	5
C22:6*n-3**	22.164	22	5.476	5

Table 11: PLS predictions of chain length and number of double bonds based on the ECL values in Table 1 for Vf-23 column

Vf-23 Column PLS Predictions for chain-length and double bonds				
Sample	**Chain Length**		**Double Bonds**	
	Predicted	**Rounded**	**Predicted**	**Rounded**
C14:1n-9*	14.046	14	0.956	1
C16:1n-9*	15.922	16	1.026	1
C18:1n-9*	17.975	18	0.992	1
C19:1n-9*	18.882	19	1.005	1
C20:1n-9*	19.920	20	0.915	1
C22:1n-9*	21.924	22	1.016	1
C24:1n-9*	23.973	24	0.939	1
C18:2n-6*	18.092	18	2.005	2
C19:2n-6*	19.161	19	1.946	2
C20:2n-6*	19.920	20	2.103	2
C21:2n-6*	21.130	21	2.020	2
C18:3n-6*	18.111	18	2.957	3
C20:3n-3*	20.038	20	2.977	3
C20:3n-6*	19.952	20	3.114	3
C22:3n-3*	22.300	22	3.024	3
C20:4n-6*	19.837	20	4.062	4
C22:4n-6*	21.835	22	4.100	4
C22:6n-3*	21.984	22	5.843	6
C17:1n-9≠	16.978	17	0.907	1
C22:2n-6≠	22.087	22	1.955	2

C18:3n-3≠	18.389	18	3.017	3
C20:5n-3≠	20.110	20	5.303	5

Table 12: PLS predictions of chain length and number of double bonds based on the ECL values in Table 1 for SP-2560 column

SP-2560 Column PLS Predictions for chain-length and double bonds				
Sample	**Chain Length**		**Double Bonds**	
	Predicted	**Rounded**	**Predicted**	**Rounded**
C16:1n-9*	15.838	16	1.080	1
C18:1n-9*	17.892	18	1.051	1
C18:2n-6*	18.069	18	2.025	2
C18:3n-6*	18.110	18	2.788	3
C20:1n-9*	19.881	20	0.981	1
C20:2n-6*	20.035	20	2.079	2
C20:3n-3*	20.217	20	3.120	3
C22:1n-9*	22.089	22	0.932	1
C20:3n-6*	20.149	20	2.927	3
C20:4n-6*	20.042	20	3.950	4
C24:1n-9*	23.836	24	1.084	1
C22:4n-6*	22.108	22	3.910	4
C22:5n-3*	22.052	22	4.880	5
C22:6n-3*	21.681	22	6.194	6
C17:1n-9≠	16.885	17	1.051	1
C18:3n-3≠	18.482	18	2.783	3
C22:2n-6≠	21.867	22	2.179	2
C20:5n-3≠	20.052	20	4.919	5