Research Paper

Estimation of stereospecific fatty acid distribution in vegetable oils from liquid chromatography data

Helena Sovová[1], Miroslav Lísa[2], Michal Holčapek[2]

[1] Institute of Chemical Process Fundamentals, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[2] Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Pardubice, Czech Republic

The fatty acid distribution in triacylglycerols (TG), the major components of vegetable oils, affects the physical properties of oils and their physiological effects as components of the human diet. A computational scheme is proposed for estimating the stereospecific distribution of fatty acids in TG on the basis of the data from HPLC analysis, using a model for TG synthesis. A new model for the synthesis is proposed which, in contrast to the widely used model published independently by Vander Wal and Coleman, enables us to simulate the asymmetrical fatty acid distribution at the *sn*-1 and *sn*-3 stereochemical positions. The computational scheme combined with either literature models or the new model was validated by its application to several TG profiles and comparison of the resulting fatty acid distribution at stereochemical positions with experimental data from the literature. The stereospecific composition of TG was calculated using the new model for nine vegetable oils.

Keywords: Vegetable oils, Triacylglycerol, Stereospecificity, HPLC, Diacylglycerol.

Received: April 27, 2007; accepted: September 15, 2007

DOI 10.1002/ejlt.200700121

1 Introduction

The major components of vegetable oils are triacylglycerols (TG), which are made up of three fatty acids (FA) esterified to a glycerol backbone. Both FA composition and their position on the glycerol backbone (stereospecific distribution) determine the physical, chemical and nutritional characteristics of oils. The carbon atoms of the glycerol part are numbered 1–3. In a Fischer projection, the secondary hydroxyl group is shown to the left of C-2. The carbon atom above this then becomes C-1, that below becomes C-3, and the prefix *sn* (stereochemical numbering) is placed before the stem name of the compound [1].

Different chromatographic techniques have been successfully used for TG analyses including gas chromatography (GC), high-performance liquid chromatography (HPLC), thin-layer chromatography (TLC), and supercritical fluid chromatography. Until recently, the chromatographic methods did not resolve TG according to the FA distribution on the glycerol backbone, with the exception of silver-ion chromatography, where only limited resolution was possible. For complete structural analysis of triacyl-*sn*-glycerol, complementary enzymatic and chemical methods have been applied [2–5]. With the development of reversed-phase (RP)-HPLC and especially with the improvement of the stationary phase efficiency, however, complete resolution of dipalmitoyl monooleoyl TG isomers [6] and complete regioselective separation of five pairs of isomeric dipalmitoyl polyalkenoyl glycerols with two to six double bonds in the unsaturated acyl residues [7] were achieved. Also the application of silver-ion packed-column supercritical fluid chromatography with mass spectrometric detection gives promising results on the molecular association of FA in intact TG [8].

Molecular-level information of the regioisomeric structures of individual TG in vegetable and animal fats and oils was obtained by chemical ionization and collision-induced dissociation tandem mass spectrometry (MS) [9]. Only the most frequently occurring molecular weight species were selected for the analysis, and the resulting concentrations of

Correspondence: Helena Sovová, Institute of Chemical Process Fundamentals, Rozvojová 135, 165 02 Prague 6, Czech Republic.
E-mail: sovova@icpf.cas.cz
Fax: +420 220 920661

TG regioisomers in the mixture were subjected to statistical analysis, showing a wide range of uncertainties of individual data. Several ionization methods were used with liquid chromatography to quantify the regioisomers of the most abundant TG in two vegetable oils and lard [10].

Thus, the complete quantification of TG stereoisomers in vegetable oils will probably be possible in the future, but presently the method of choice for accurate quantitative analysis of the TG composition in vegetable oils remains RP-HPLC, the method that does not distinguish among the TG stereoisomers [11].

Different theories based on the available analytical data have been proposed to explain the stereochemical distribution of FA in the TG molecule and to calculate the composition of TG species in vegetable oils. Of these, the 1,3-random-2-random theory based on the FA composition in the oil and at the *sn*-2 position, proposed by Vander Wal [12] and Coleman [13], has become the most widely accepted. The theory assumes that FA are randomly distributed at the *sn*-2 and at the *sn*-1,3 positions, and implies the identity of FA distributions at the *sn*-1 and *sn*-3 positions.

Later analyses of the three *sn* positions in TG of some oils have, however, shown that a difference exists between the FA compositions at the *sn*-1 and *sn*-3 positions. Martínez-Force *et al.* [14] characterized the relative distribution of saturated FA between the external *sn*-1 and *sn*-3 positions in TG by the coefficient of asymmetry. Their model is based on the assumption of random FA distribution at each *sn* position in TG and enables the calculation of the coefficient from the compositions of oil FA, *sn*-2 FA, and TG. Van Vliet and van Kempen [15] presented a mathematical framework relating the stereospecific TG distribution to any measurement results obtained using different analytical techniques and their combinations (*e.g.* the data on FA composition in TG + FA composition at the *sn*-2 position of TG + profile of TG species, where the isomers are not distinguished). How the FA were distributed on the glycerol backbone was not defined *a priori*. As the number of TG is always much higher than the number of data points, an unequivocal calculation of TG from the measurement results alone is impossible. To obtain a unique solution, either a model for putting FA together to form TG must be included in the calculation or an algorithm for finding one of many equally good solutions should be defined.

Having at our disposal sets of reliable RP-HPLC data on TG composition (combined isomers) of different vegetable oils [11], we have adopted a computational scheme where the stereospecific TG composition is estimated using the model, the positional isomers of identical chemical composition are grouped to give TG fractions, and the difference between the measured and calculated TG profiles is minimized. In this paper, models from the literature (models A, B) and a new model simulating the specific action of enzymes in the biosynthesis of oils (model C) are tested in the computational scheme, and the new model is applied to calculate the TG positional isomers in nine vegetable oils.

2 Calculations

2.1 Symbols for TG components and models for TG synthesis

The composition of FA is expressed in moles. The mole fraction of the *i*th FA in TG is x_i and the mole fractions of the *i*th FA at the *sn*-1, *sn*-2 and *sn*-3 positions are $x_{i,1}$, $x_{i,2}$ and $x_{i,3}$, respectively. They are constrained from definition by

$$\sum_{i=1}^{n} x_i = 1,\ x_i \geq 0 \text{ for } i = 1, 2,..., n \tag{1}$$

$$\sum_{i=1}^{n} x_{i,1} = 1,\ x_{i,1} \geq 0 \text{ for } i = 1, 2,..., n \tag{1a}$$

$$\sum_{i=1}^{n} x_{i,2} = 1,\ x_{i,2} \geq 0 \text{ for } i = 1, 2,..., n \tag{1b}$$

$$\sum_{i=1}^{n} x_{i,3} = 1\ x_{i,3} \geq 0 \text{ for } i = 1, 2,..., n \tag{1c}$$

where n is the number of FA in the TG. The mole fraction of the TG stereoisomer containing the acyl of the *i*th FA at the *sn*-1 position, the acyl of the *j*th FA at the *sn*-2 position and the acyl of the *k*th FA at the *sn*-3 position is x_{ijk}, and the sum of the mole fractions of TG is

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} x_{ijk} = 1 \tag{2}$$

The total number of TG consisting of n FA is n^3. The mole fraction of the TG stereoisomer created by a random combination of FA at the *sn*-1, *sn*-2 and *sn*-3 positions is the product

$$x_{ijk} = x_{i,1}x_{j,2}x_{k,3} \tag{3}$$

The simplest possible scheme for the synthesis of TG would assume an identical FA distribution at all *sn* positions,

$$x_i = x_{i,1} = x_{i,2} = x_{i,3} \tag{4}$$

with the result

$$x_{ijk} = x_i x_j x_k. \tag{5}$$

Such a model of an interesterified oil or fat has only $n - 1$ parameters, namely the mole fractions of FA 1, 2,..., $n - 1$; the last fraction, x_n, is calculated from Eq. (1). In native fatty oils, however, the FA composition at the *sn*-2 position is known to differ from the composition at the primary positions (*sn*-1 and *sn*-3). In TG of seed oils, saturated FA are concentrated almost entirely at the primary positions while the *sn*-2 position is greatly enriched in polyunsaturated FA [1].

2.1.1 Model A

The model derived by Vander Wal and Coleman has $2(n-1)$ parameters, the total FA composition and the FA composition at the *sn*-2 position. The composition at the primary positions is calculated according to

$$x_{i,1} = x_{i,3} = 0.5\left(3x_i - x_{i,2}\right). \tag{6}$$

After substitution of Eq. (6) to Eq. (3), the mole fractions of TG stereoisomers are obtained:

$$x_{ijk} = 0.25x_{j,2}\left(3x_i - x_{i,2}\right)\left(3x_k - x_{k,2}\right). \tag{7}$$

The model is most widely accepted, although improved analytical techniques have revealed an asymmetry between the *sn*-1 and *sn*-3 positions of the TG in some oils.

2.1.2 Model B

An asymmetric model was derived from model A by Martínez-Force *et al.* [8], who distinguished between the FA compositions at the *sn*-1 and *sn*-3 positions:

$$x_{i,1} = \left(3x_i - x_{i,2}\right)\alpha_i,\ x_{i,3} = \left(3x_i - x_{i,2}\right)(1-\alpha_i) \tag{8}$$

$$x_{ijk} = \alpha_i(1-\alpha_k)\left(3x_i - x_{i,2}\right)x_{j,2}\left(3x_k - x_{k,2}\right) \tag{9}$$

The model has $3(n-1)$ parameters, namely the total FA composition x_i, $i = 1, 2,..., n-1$, the FA composition at the *sn*-2 position $x_{i,2}$, $i = 1, 2,..., n-1$, and α_i, $i = 1, 2,..., n-1$, where α_i is the coefficient of asymmetry for the ith FA. When $\alpha_i = 0.5$ for all FA, Eq. (8) is reduced to Eq. (6).

2.1.3 Model C

Models A and B are based on a random combination of FA at all three positions. However, the synthesis of oils in plants is catalyzed by enzymes that discriminate among the FA, and thus it is not a completely random process. The major biosynthetic pathway for TG is the α-glycerophosphate or *sn*-glycerol-3-phosphate pathway in which *sn*-glycerol-3-phosphate is acylated in turn at the positions *sn*-1 and *sn*-2 by specific transferases to form phosphatidic acid, the enzyme phosphatidate phosphatase removes the phosphate group and the resulting *sn*-1,2-diacylglycerol (DG) is acylated by a further acyltransferase to form an *sn*-TG [16]. Model C simulates the specific formation of *sn*-1,2-DG followed by their random combination with the remaining free FA to TG. To describe the enzymatic synthesis of *sn*-1,2-DG, the specificity factors f_{ij} are introduced:

$$x_{ij} = f_{ij}x_i x_j,\ f_{ij} \geq 0 \text{ for } i = 1, 2,..., n \text{ and } j = 1, 2,..., n. \tag{10}$$

The composition of the free FA that will be esterified at the *sn*-3 position in the second step is calculated from the mass balance of FA as

$$x_{i,3} = 3x_i - \sum_{j=1}^{n}\left(x_{ij} + x_{ji}\right) \tag{11}$$

and the mole fractions of TG combined at random from *sn*-1,2-DG and free FA are

$$x_{ijk} = x_{ij}x_{k,3}. \tag{12}$$

The number of parameters of the model in the form given above would be $n(n+1)-2$, namely $n-1$ for the FA composition of oil and n^2 for the specificity factors, which are bound by the condition

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = 1. \tag{13}$$

2.2 Computational scheme

The above-mentioned models were tested in a computation scheme (Fig. 1) where the FA composition and the model parameters were adjusted to minimize the difference of the calculated TG profiles from the experimental TG profiles of vegetable oils. The enantiomers are not distinguished in the experimental TG profiles, which are expressed in mass fractions, $W_{ijk\,exp}$, or mole fractions, $X_{ijk\,exp}$. Thus, the calculated molar fractions of TG stereoisomers were summed up to give mole fractions of TG species X_{ijk} according to the equation

$$X_{ijk} = x_{ijk} + x_{ikj} + x_{jik} + x_{jki} + x_{kij} + x_{kji} \text{ for } i \neq j \neq k,\ X_{iij} = x_{iij} + x_{iji} + x_{jii} \text{ for } i \neq j,\ X_{iii} = x_{iii}. \tag{14}$$

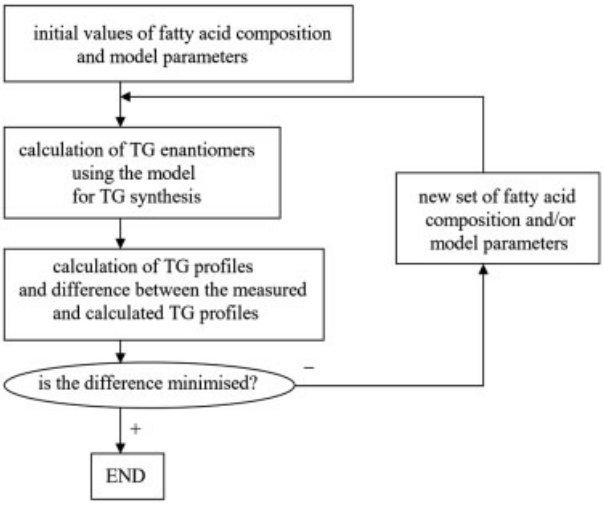

Figure 1. Computational scheme for the synthesis of TG.

To compare with mass fractions, the profiles X_{ijk} are converted to mass fractions W_{ijk}. The optimization criterion is calculated as the sum of squares of differences between the measured and calculated mass fractions:

$$\delta_{TG} = \left[\frac{1}{n_{TG}}\sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=j}^{n}\left(KW_{ijk} - W_{ijk\,exp}\right)^2\right]^{0.5},$$

$$n_{TG} = \frac{n(n^2+3n+2)}{6} \quad (15)$$

where n_{TG} is the maximum possible number of TG species (mixtures of enantiomers) in the oil containing n FA. For TG profiles expressed in mole fractions, X_{ijk} and $X_{ijk\,exp}$ would substitute W_{ijk} and $W_{ijk\,exp}$ in Eq. (15).

The optimized coefficient K was introduced in Eq. (15) to account for a possible difference of the sum of all $W_{ijk\,exp}$ from 1, as the sum of the calculated W_{ijk} is equal to 1. The optimization was carried out using the Excel Solver. Several initial estimates of model parameters were applied to increase the chance of reaching a global minimum.

The model for an interesterified oil with random FA distribution according to Eq. (5) requires only the knowledge of $n-1$ mole fractions of FA. Whatever their initial estimates were, either homogeneous distribution (equal FA fractions $x_i = 1/n$ for $i = 1, 2,..., n$) or the experimental FA composition of the given oil or any other composition, the result was always the same optimum FA distribution and TG profile.

Model A needs according to Eq. (7), in addition to the total FA composition, also the FA composition at the *sn*-2 position in TG. As their initial estimate, we used either two homogeneous distributions or two experimental FA compositions of the oil, or one of these compositions for the total FA in the oil and the other for the *sn*-2 position. In the iterations, the total FA composition together with parameter K and the composition at the *sn*-2 position were adjusted in turn; only close to the minimum, all variables were adjusted at once. Model B was incorporated in the computation scheme similarly to model A, adjusting repeatedly three sets of FA distributions at the *sn*-1, *sn*-2 and *sn*-3 positions in turn.

Model C requires the knowledge of specificity factors, in addition to the FA composition. As the calculated fractions of stereoisomers are summed to calculate TG profiles, the specificity factors f_{ij} and f_{ji} cannot be distinguished in the sum $x_{ij} + x_{ji} = (f_{ij} + f_{ji})x_i x_j$. Thus, the parameters $F_{ij} = f_{ij} + f_{ji}$ are optimized instead of f_{ij}, f_{ji} and the number of adjusted specificity parameters is reduced from n^2 to $n(n+1)/2$. The initial estimates of the total FA composition were the same as for model A. The initial estimate of F_{ij} values was either that corresponding to random distribution ($F_{ij} = 1$ for $i = j$ and $F_{ij} = 2$ for $i \neq j$) or any arbitrarily chosen distribution. In the first case, the input data fulfilled the condition given by Eq. (13). Otherwise, the F_{ij} values were adjusted automatically in each step of iteration to fulfill the conditions as the equation was used as boundary condition.

3 Validation of calculation procedure using literature data

As the difference between the experimental and the calculated TG profiles decreases with increasing number of model parameters, the ability of the models to simulate the real oil composition must be judged according to other criteria. One of them is the difference between the experimental and the calculated FA compositions in the oil. The experimental TG profiles do not include the least abundant TG that contribute to the experimental FA composition determined after TG decomposition, but the calculated TG profiles include them. Another, more important criterion is the agreement between the calculated and the experimental distributions of FA in TG. For this comparison, literature data comprising both TG profiles and FA distributions in TG were used.

3.1 Averaged locations of FA in TG

Reske *et al.* [17] examined the changes in composition of oils from genetically modified sunflower and soybean seeds. They measured TG profiles using RP-HPLC, the FA composition of transmethylated crude oils by capillary GC, and average FA compositions at the *sn* positions of the glycerol backbone by partial hydrolysis of oils to DG, HPLC separation of their *sn*-1,2- and *sn*-2,3 derivatives, and analysis of the two groups for FA. In addition, the FA composition at the *sn*-2 position was determined after enzymatic hydrolysis of oils. The modified oils were denoted according to the high (H) or low (L) levels of specified FA, compared to standard oil (Lsat = low saturated acids, LLn = low linolenic acid, HP = high palmitic acid). Initially, we varied the number of FA included in the calculation and found that increasing the number above $n = 5$ had little effect on the results because the concentration of most TG species containing the sixth acid was too low to significantly affect the minimized standard deviation δ_{TG}. Thus, all calculations were carried out with $n \leq 5$ major FA. The abbreviations used for the FA are given in Table 1.

Table 1. Common FA in vegetable oils characterized by their carbon numbers and double bond numbers.

Trivial name	Abbreviation	CN:DB
Oleic acid	O	18:1
Linoleic acid	L	18:2
Linolenic acid	Ln	18:3
Palmitoleic acid	Po	16:1
Palmitic acid	P	16:0
Stearic acid	S	18:0

CN, carbon number; DB, double bond number.

Analogously to TG composition, the difference between the experimental and calculated FA compositions of the oil, x_i, was characterized by standard deviation

$$\delta_{FA} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - x_{i\,exp}\right)^2\right]^{0.5} \quad (16)$$

and the difference between the experimental and calculated FA compositions at the stereochemical positions was quantified by standard deviation

$$\delta_{FA,1-3} = \left[\frac{1}{3n}\sum_{i=1}^{n}\sum_{j=1}^{3}\left(x_{i,j} - x_{i,j\,exp}\right)^2\right]^{0.5} \quad (17)$$

In the case of models B and C, the calculation according to the computation scheme does not yield a complete pattern of TG stereoisomers, and additional information is needed. We utilized the fact that saturated FA in vegetable oils only rarely occur at the *sn*-2 position. In the asymmetric model B, where the question of allocation of three resulting FA compositions arises, the composition with the lowest content of saturated FA was assigned to the *sn*-2 position. As there was no rule to distinguish between the *sn*-1 and *sn*-3 positions, both combinations were compared with experimental data.

In model C, the optimized F_{ij} factors do not allow distinguishing between the *sn*-1 and *sn*-2 positions in DG formed in the first step. For a combination of saturated and unsaturated FA, the saturated FA is always assumed to be at the *sn*-1 position. When two different saturated acids or two unsaturated acids were combined, equal occurrence of both *sn*-1,2-DG isomers is assumed. Mathematically,

$$f_{Sat1U1} = F_{Sat1U1}, f_{U1Sat1} = 0, f_{U1U2} = f_{U2U1} = F_{U1U2}/2, f_{Sat1Sat2} = f_{Sat2Sat1} = F_{Sat1Sat2}/2. \quad (18)$$

where Sat1 and Sat2 denote two different saturated FA and U1, U2 denote two different unsaturated FA.

The results obtained for sunflower oils are summarized in Table 2. As expected, the deviations of the calculated TG profiles from the experimental ones decrease from the random FA distribution (interesterified oil) in the order of the models A-B-C. The agreement of the FA distribution in oil improves in the same order for most oils.

Out of two experimental data for the FA composition at the *sn*-2 position, those determined by enzyme hydrolysis were simulated better by all models, and therefore they were used in the calculations. Model A yields a better agreement than the random model, except for the HL oil. A further slight improvement is observed for the better of two results of model B. Model C, however, shows large differences from the experimental data, especially for the oils HS/HO and HP/HO where the largest deviations between the experimental and the calculated compositions are observed at the *sn*-3 position. One possible explanation is that increasing the number of adjustable parameters above a certain limit leads rather to increasing the effect of experimental error than improving the simulation of the real oil composition, and this limit was exceeded in model C. Therefore, we reduced the number of model parameters by 3, taking into account that saturated FA rarely occur at the *sn*-2 position and putting the F_{ij} factors for the 1,2-DG PP, PS and SS equal to zero. The resulting model C1 acceptably simulates the stereospecific distribution of FA.

As an example, the results of calculation using model C1 for the standard sunflower seed oil are listed in Table 3. A good agreement between the calculated and the experimental TG profiles is evident in Fig. 2; the differences in the FA

Table 2. Deviations between the calculated and the experimental oil compositions for sunflower oils [17]; TG profile: δ_{TG}; FA composition: δ_{FA}; stereospecific FA distribution: $\delta_{FA,1-3}$; FA: n = 4 (O, L, P, S) except for HP/HO, HP/HL oils with n = 5 (Po added).

		Standard	HO	HL	HS/HO	HP/HO	HP/HL	Average
δ_{TG}	Random	0.0041	0.0020	0.0038	0.0046	0.0136	0.0090	0.0062
	Model A	0.0037	0.0016	0.0031	0.0029	0.0093	0.0071	0.0046
	Model B	0.0036	0.0013	0.0028	0.0022	0.0075	0.0069	0.0040
	Model C	0.0010	0.0002	0.0013	0.0010	0.0062	0.0050	0.0025
	Model C1	0.0012	0.0013	0.0022	0.0023	0.0075	0.0055	0.0033
δ_{FA}	Random	0.015	0.008	0.008	0.024	0.047	0.004	0.024
	Model A	0.017	0.008	0.011	0.016	0.030	0.037	0.020
	Model B	0.018	0.009	0.011	0.016	0.031	0.037	0.020
	Model C	0.012	0.009	0.010	0.015	0.031	0.024	0.017
	Model C1	0.012	0.009	0.005	0.015	0.031	0.021	0.015
$\delta_{FA,1-3}$	Random	0.059	0.020	0.026	0.087	0.157	0.115	0.079
	Model A	0.064	0.021	0.074	0.062	0.073	0.054	0.058
	Model B[§]	0.041	0.029	0.055	0.069	0.049	0.060	0.051
		0.097	0.039	0.085	0.069	0.131	0.092	0.086
	Model C	0.044	0.055	0.070	0.158	0.270	0.060	0.109
	Model C1	0.045	0.035	0.072	0.059	0.048	0.059	0.053

[§] The two results correspond to two possible allocations of the two resulting FA compositions at the *sn*-1 and *sn*-3 positions.

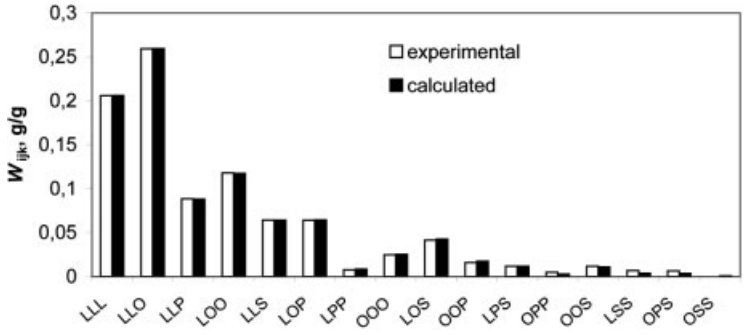

Figure 2. Experimental [17] and calculated compositions of TG in standard sunflower seed oil.

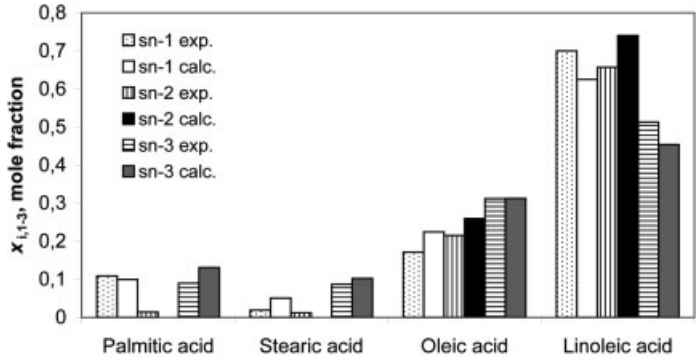

Figure 3. Experimental [17] and calculated FA distributions at the *sn*-1, *sn*-2 and *sn*-3 positions for standard sunflower seed oil.

stereospecific distribution are shown in Fig. 3. Although the agreement is not as good as for the total FA distribution, the asymmetry between the *sn*-1 and the *sn*-3 position for stearic, oleic and linoleic acids was correctly predicted.

The computational scheme was further applied to the data on soybean oils from the same literature source [17]. The deviations between the experimental and the calculated data for TG profiles, total FA distribution and stereospecific FA distribution are listed in Table 4. The best fit in TG profiles and, to some extent, in the total FA distribution was again achieved with model C, due to its best flexibility. The average fit of the stereospecific FA distribution improves in the order random-B-A-C-C1. The difference between the models C and C1 is not as large as for sunflower oil, most probably because the limit for the optimum number of adjustable parameters is higher as soybean oil consists of five major FA and thus the TG profile consists of more non-zero components.

Next, calculations were carried out simulating the data on different vegetable oils published in ref. [14]. The experimental data consist of TG profiles determined by gas-liquid chromatography (GLC), FA compositions determined by GC after methyl esterification of oil, and FA compositions in TG at the *sn*-2 position of oil determined by the same method after enzymatic hydrolysis of purified TG and separation of the obtained mixture by TLC. Thus, instead of comparing the FA distribution at all *sn* positions, the standard deviation between the calculated and the experimental FA distributions at the *sn*-2 position was calculated:

$$\delta_{FA,2} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(x_{i,2} - x_{i,2\,exp}\right)^2\right]^{0.5} \quad (19)$$

For unresolved TG peaks OLLn and LLL in the experimental TG profiles, we assumed equal fractions of both species.

The results for models A, B and C1 are summarized in Table 5. According to the average values, the best agreement with experimental data was again obtained with model C1. The assumption of equal distribution of oleic and linoleic acids between the *sn*-1 and *sn*-2 positions in OL is possibly too strict. For example, a deviation as low as $\delta_{FA,2} = 0.021$ is obtained for olive oil when all oleic acid in the compound is at the *sn*-1 position and all linoleic acid is at the *sn*-2 position. The high average value of $\delta_{FA,2}$ for model A is caused by a high deviation for walnut oil. If the result for walnut oil is omitted, the average deviation $\delta_{FA,2}$ would be comparable with that of model C1.

3.2 Regioisomers of TG

Finally, the models were compared with experimental percentages of regioisomers of major TG in three vegetable oils and one animal fat, published by Kallio and coworkers [9, 10]. As the *sn*-1 and *sn*-3 positions in TG were not distinguished experimentally, the mole fractions of regioisomers were calculated from the results of the computation scheme according to

Table 3. The results of the calculation for standard (unmodified) sunflower seed oil [17] using model C: experimental and calculated FA composition, experimental and calculated composition of TG, calculated composition of 1,2-DG (X_{ij}) and FA at the positions *sn*-1, *sn*-2 and *sn*-3, adjustable parameter K, and standard deviations of the calculated compositions from the experimental ones.

FA	x_{iexp}	x_i		ij	X_{ij}
L	0.587	0.607		LL	0.487
O	0.278	0.265		LO	0.276
P	0.070	0.077		OO	0.086
S	0.048	0.051		SL	0.041
Sum	0.983	1.000		SO	0.010
				PL	0.074
TG	W_{ijkexp}	KW_{ijk}		PO	0.026
LLL	0.2061	0.2062		Sum	1.000
LLO	0.2594	0.2594			
LLP	0.0886	0.0883		$x_{i,1exp}$	$x_{i,1}$
LOO	0.1180	0.1175	L	0.700 ± 0.013	0.625
LLS	0.0644	0.0642	O	0.171 ± 0.018	0.224
LOP	0.0643	0.0645	P	0.109 ± 0.018	0.100
LPP	0.0077	0.0085	S	0.019 ± 0.018	0.051
OOO	0.0249	0.0253	Sum	0.999	1.000
LOS	0.0415	0.0427			
OOP	0.0160	0.0176		$x_{i,2exp}$	$x_{i,2}$
LSP	0.0119	0.0118	L	0.578 ± 0.038	0.741
OPP	0.0049	0.0030	O	0.363 ± 0.041	0.259
OOS	0.0119	0.0111	P	0.041 ± 0.003	0
LSS	0.0069	0.0040	S	0.017 ± 0.001	0
OSP	0.0066	0.0035	Sum	0.999	1.000
OSS	0.0000	0.0009			
Sum	0.933	0.928		$x_{i,3exp}$	$x_{i,3}$
			L	0.513 ± 0.025	0.454
K	0.928		O	0.312 ± 0.023	0.313
			P	0.090 ± 0.000	0.131
δ_{TG}	δ_{FA}	$\delta_{FA,1-3}$	S	0.087 ± 0.003	0.102
0.001	0.012	0.045	Sum	1.002	1.000

$$X_{regi,j,k} = x_{i,j,k} + x_{k,j,i} \text{ for } i \neq k,\ X_{regi,j,k} = x_{i,j,k} \text{ for } i = k \qquad (20)$$

Comparing all experimental and calculated regioisomers, the standard deviation was calculated:

$$\delta_{reg} = \left[\frac{1}{n_{reg}} \sum_{i=1}^{n_{reg}} \left(X_{reg,i} - X_{exp\,reg,i}\right)^2\right]^{0.5} \qquad (21)$$

The number of regioisomers determined in oil or fat, n_{reg}, ranged from 11 to 16.

The results are summarized in Table 6. In contrast to the previously examined vegetable oils, the data on palm oil [9] indicate the presence of saturated acids at the *sn*-2 position. Therefore, the use of model C1 for palm oil was not appropriate, and model C gave the best results, together with model B. The difference in the results of individual models is evident in Fig. 4 where the experimental and calculated compositions of regioisomers in palm oil are compared.

To show the potential of the computation scheme to simulate the composition of animal oils, the TG profiles of lard [10] were fitted by the models. As palmitic acid in lard occupies preferably the *sn*-2 position in TG, the additional characteristics in models B and C had to be reversed. In model B, the resulting FA composition richest in saturated acids was selected for the *sn*-2 position. In model C, in combinations of saturated and unsaturated FA in *sn*-1,2-DG, the saturated FA was always assumed to be at the *sn*-2 position. The results obtained with models B and C are satisfactory, but also model A, which does not require any additional information, simulates the regioisomer composition well.

The compositions of vegetable oils from rapeseed and sunflower seeds [10] were equally well simulated with model C and its version C1.

4 Application of model C1 to nine vegetable oils

Holčapek *et al.* [11] identified and quantified 133 TG containing 22 FA in nine vegetable oils (walnut, hazelnut, cashew nut, almond, poppy seed, yellow melon, mango, fig and date seed) using HPLC/MS with a response factor approach. The FA composition in seven oils was determined after their methyl esterification using GC/FID, showing that five common FA (O, L, Ln, P, S) represent 97.9–99.3 wt-% of FA. According to the estimate based on TG composition of the remaining two oils, the five common FA represent 98.6 wt-% of FA in mango oil, but the set of major FA in date seed oil is completely different, as discussed later. The computation scheme including model C1 was applied to estimate the stereospecific composition of TG in the oils. Good agreement between the experimental and calculated compositions of TG with δ_{TG} ranging between 0.001 and 0.004 was achieved. The results are shown in Table 7.

To simulate the composition of date seed oil, its FA were represented by five pseudo-components with FA grouped according to the number of double bonds and with similar carbon numbers: The first pseudo-component was a sum of caprylic, capric and lauric acids; the second one was a sum of oleic, linoleic and gadoleic acids; the third one was a sum of arachidic, behenic and lignoceric acids; the fourth one was a sum of palmitic and stearic acids; and the fifth pseudo-component was a sum of myristic and margaric acids. Thus, the second pseudo-component consisted of unsaturated FA and the rest were saturated FA. The last two pseudo-components replaced P and S in the computational scheme, and their combination in 1,2-DG was therefore forbidden. The calculated composition of 1,2-DG is in agreement with the expectation that saturated FA are preferably combined with unsaturated acids (mole fractions X_{LnL}, X_{SL}, X_{PL} in the last column of Table 7) while the probability of mutual combinations of saturated acids is low

Table 4. Deviations between the calculated and the experimental oil compositions for soybean oils [17]; TG profile: δ_{TG}; FA composition: δ_{FA}; stereospecific FA distribution: $\delta_{FA,1-3}$; FA: n = 5 (O, L, Ln, P, S).

		Standard	Lsat	LLn	HP	HS	Hsat	Lsat/LLn	HP/LLn	Average
δ_{TG}	Random	0.009	0.013	0.012	0.007	0.012	0.018	0.013	0.009	0.012
	Model A	0.009	0.011	0.011	0.006	0.011	0.013	0.013	0.008	0.010
	Model B	0.009	0.011	0.011	0.006	0.011	0.012	0.013	0.008	0.010
	Model C	0.005	0.004	0.004	0.004	0.009	0.010	0.002	0.003	0.005
	Model C1	0.009	0.011	0.012	0.006	0.011	0.013	0.013	0.008	0.010
δ_{FA}	Random	0.030	0.022	0.035	0.027	0.027	0.032	0.028	0.033	0.029
	Model A	0.030	0.022	0.034	0.021	0.024	0.026	0.027	0.026	0.026
	Model B	0.031	0.021	0.034	0.021	0.021	0.028	0.027	0.024	0.026
	model C	0.016	0.022	0.010	0.022	0.023	0.030	0.018	0.021	0.020
	Model C1	0.030	0.022	0.035	0.019	0.024	0.026	0.027	0.022	0.025
$\delta_{FA,1-3}$	Random	0.061	0.054	0.074	0.113	0.111	0.0136	0.084	0.098	0.092
	Model A	0.083	0.060	0.072	0.041	0.031	0.060	0.127	0.042	0.065
	Model B[§]	0.069	0.048	0.073	0.067	0.063	0.078	0.084	0.062	0.068
		0.070	0.066	0.087	0.069	0.063	0.087	0.109	0.067	0.077
	Model C	0.052	0.063	0.031	0.101	0.066	0.079	0.058	0.038	0.061
	Model C1	0.039	0.044	0.041	0.044	0.038	0.060	0.077	0.039	0.048

[§] The two results correspond to two possible allocations of the two resulting FA compositions at the *sn*-1 and *sn*-3 positions.

Table 5. Deviations between the calculated and the experimental oil compositions for different vegetable oils [14]; TG profile: δ_{TG}; FA composition in oil: δ_{FA}; FA composition at *sn*-2: $\delta_{FA,2}$; FA: n = 5 (O, L, Ln, P, S) except for sunflower oils where n = 4 (O, L, P, S) and olive oil where n = 5 (O, L, P, S, Po).

		Rice	Soybean	Walnut[†]	Hazelnut	Olive	Sunflower RHA-274	Sunflower CAS-3	Average
δ_{TG}	Model A	0.010	0.013	0.012	0.006	0.008	0.007	0.012	0.010
	Model B	0.010	0.012	0.009	0.006	0.007	0.007	0.005	0.008
	Model C1	0.005	0.007	0.010	0.001	0.004	0.001	0.002	0.004
δ_{FA}	Model A	0.008	0.021	0.034	0.021	0.013	0.025	0.021	0.020
	Model B	0.008	0.020	0.020	0.022	0.012	0.026	0.020	0.018
	Model C1	0.006	0.012	0.029	0.004	0.020	0.021	0.013	0.015
$\delta_{FA,2}$	Model A	0.020	0.041	0.198	0.068	0.047	0.018	0.056	0.064
	Model B	0.020	0.078	0.090	0.081	0.026	0.019	0.068	0.055
	Model C1	0.022	0.058	0.054	0.071	0.062	0.029	0.027	0.046

[†] The experimental mole percentage of linoleic acid in walnut oil at the *sn*-2 position was assumed to be equal to 71.0 (its value was misprinted in Ref. [14], Table 4).

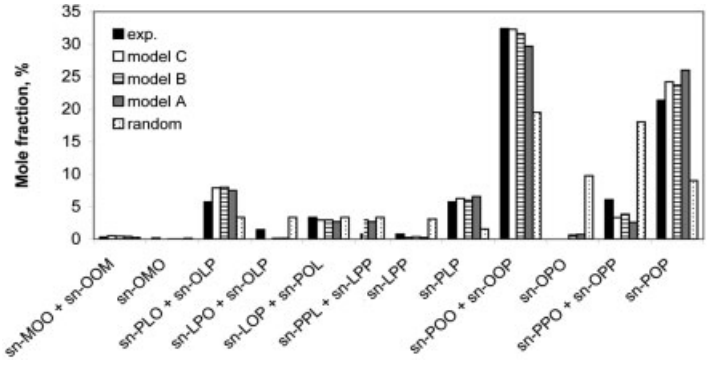

Figure 4. Experimental [9] and calculated compositions of regioisomers of palm oil major TG.

Table 6. Deviations between the calculated and the experimental oil compositions for palm oil [9] and lard, rapeseed and sunflower oils [10].

		Palm oil	Lard†	Rapeseed	Sunflower	Average
δ_{TG}	Random	0.0016	0.0028	0.0015	0.0009	0.0017
	Model A	0.0008	0.0011	0.0015	0.0009	0.0011
	Model B	0.0004	0.0007	0.0015	0.0008	0.0008
	Model C	0.0003	0.0006	0.0001	0.0003	0.0003
	Model C1	0.0006	–	0.0002	0.0005	0.0004
δ_{FA}	Random	0.027	0.020	0.041	0.016	0.026
	Model A	0.032	0.016	0.044	0.020	0.028
	Model B	0.025	0.014	0.044	0.025	0.027
	Model C	0.023	0.015	0.015	0.006	0.015
	Model C1	0.033	–	0.015	0.007	0.018
$\delta_{sn\text{-}TG}$	Random	0.073	0.061	0.026	0.018	0.044
	Model A	0.021	0.010	0.045	0.012	0.022
	Model B	0.013	0.008	0.024	0.012	0.014
	Model C	0.014	0.008	0.014	0.013	0.013
	Model C1	0.027	–	0.015	0.015	0.019

† Saturated FA were assumed at the *sn*-2 position in 1,2-DG.

or equal to zero. The grouping into five pseudo-components gives the possibility to simulate the stereospecific composition of oils containing larger numbers of FA.

The 1,2-DG composition X_{ij} calculated on the basis of the TG composition as determined by HPLC could serve as vegetable oil characteristics. Figure 5 compares the calculated X_{ij} profiles for different samples of walnut, sunflower seed, soybean and hazelnut oils. The content of individual FA, x_i, in oils of the same plant species may vary (for example, the walnut oil analyzed by Holčapek *et al.* [11] contains more linolenic, oleic and palmitic acids and less linoleic acid than the walnut oil analyzed by Martínez-Force *et al.* [14]) but the shapes of their X_{ij} profiles are similar. This is, however, not true for genetically modified oils where the differences in FA composition are too large.

5 Number of FA in the model

The number of FA included in the computational scheme should correspond to the information contained in the experimental TG profile. According to Table 8, the number of

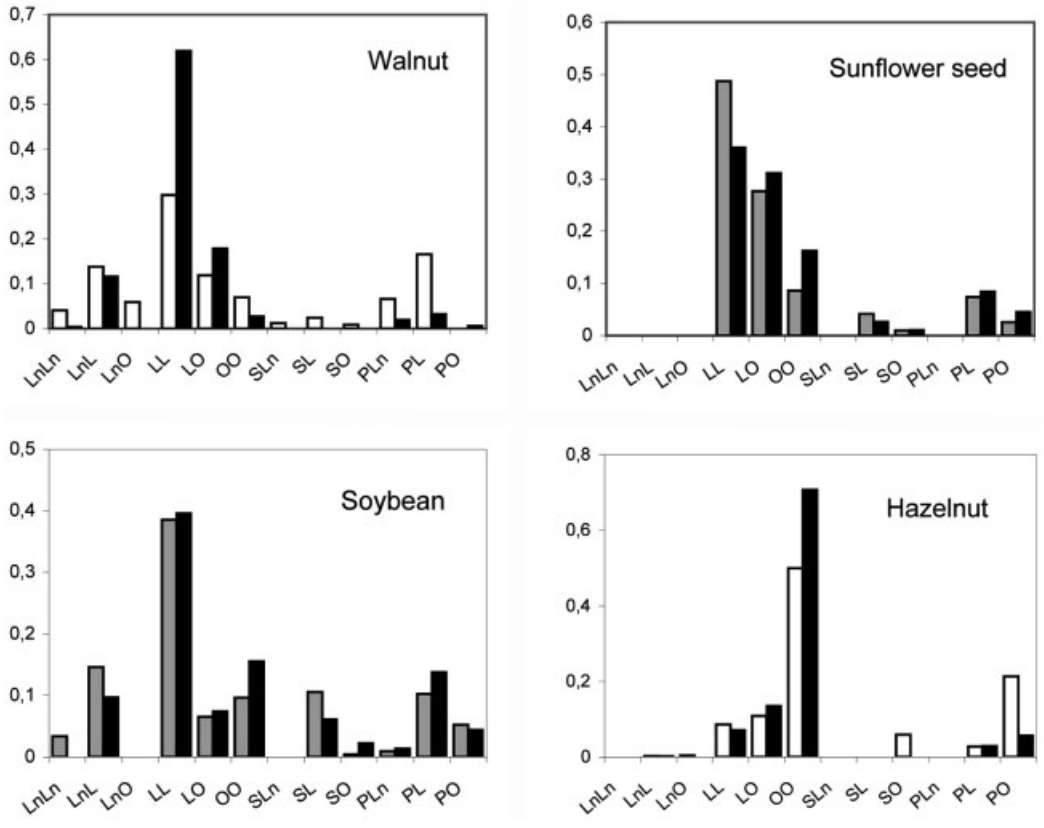

Figure 5. Comparison of the calculated compositions of 1,2-DG for pairs of samples of vegetable oils (X_{ij}, mole fraction). Sources of experimental TG profiles: (■) Reske *et al*. [17], (□) Holčapek *et al*. [11], (■) Martínez-Force *et al*. [14].

Table 7. Calculated compositions of 1,2-DG, FA in oil, and FA at the *sn*-3 position for vegetable oils analyzed by HPLC [11].

	Walnut	Hazelnut	Cashew	Almond	Poppy seed	Yellow melon	Mango stone	Fig seed	Date seed[†]
X_{LnLn}	0.040	0.000	0.000	–	0.000	0.000	0.000	0.162	0.016
X_{LnL}	0.138	0.003	0.000	–	0.032	0.003	0.000	0.079	0.706
X_{LnO}	0.059	0.004	0.000	–	0.007	0.000	0.000	0.154	0.000
X_{LL}	0.298	0.086	0.143	0.149	0.374	0.554	0.014	0.105	0.063
X_{LO}	0.119	0.109	0.098	0.150	0.152	0.263	0.000	0.123	0.000
X_{OO}	0.070	0.500	0.401	0.413	0.070	0.104	0.202	0.052	0.000
X_{SLn}	0.012	0.000	0.000	–	0.000	0.000	0.007	0.041	0.007
X_{SL}	0.024	0.000	0.050	0.003	0.045	0.020	0.042	0.022	0.048
X_{SO}	0.009	0.059	0.143	0.027	0.008	0.004	0.558	0.000	0.000
X_{PLn}	0.066	0.000	0.002	–	0.012	0.003	0.003	0.160	0.017
X_{PL}	0.165	0.028	0.044	0.119	0.268	0.049	0.026	0.087	0.144
X_{PO}	0.000	0.021	0.119	0.139	0.032	0.000	0.148	0.015	0.000
x_{Ln}	0.167	0.002	0.001	–	0.017	0.002	0.006	0.397	0.347
x_{L}	0.533	0.174	0.235	0.269	0.666	0.594	0.056	0.288	0.443
x_{O}	0.185	0.678	0.519	0.612	0.179	0.238	0.422	0.194	0.007
x_{P}	0.093	0.116	0.128	0.104	0.119	0.123	0.117	0.098	0.120
x_{S}	0.022	0.030	0.117	0.015	0.019	0.043	0.399	0.023	0.083
$x_{Ln,3}$	0.145	0.000	0.003	–	0.000	0.000	0.008	0.432	0.278
$x_{L,3}$	0.556	0.213	0.226	0.236	0.753	0.339	0.074	0.344	0.305
$x_{O,3}$	0.231	0.649	0.395	0.695	0.200	0.239	0.155	0.184	0.022
$x_{P,3}$	0.047	0.108	0.218	0.053	0.043	0.318	0.174	0.033	0.200
$x_{S,3}$	0.021	0.030	0.158	0.016	0.004	0.104	0.589	0.007	0.195
δ_{TAG}	0.002	0.001	0.004	0.002	0.001	0.002	0.004	0.003	0.003
δ_{FA}	0.037	0.049	0.035	0.031	0.033	0.015	–	0.011	–

[†] The FA Ln, L, O, S, and P are substituted by the following pseudo-components: caprylic + capric + lauric acids for Ln, oleic + linoleic + gadoleic acids for L, arachidic + behenic + lignoceric acids for O, palmitic + stearic acids for S, and myristic + margaric acids for P.

Table 8. Number of TG (*sn* position not distinguished) and adjusted parameters in a computational scheme with *n* FA.

n	TG $n(n^2 + 3n + 2)/6$	Model A $2(n-1)$	Model B $3(n-1)$	Model C $n(n+3)/2-2$
3	10	4	6	7
4	20	6	9	12
5	35	8	12	18
6	56	10	15	25
7	84	12	18	33
8	120	14	21	42
9	165	16	24	52
10	220	18	27	63

TG obtained by combination of *n* FA is always higher than the number of adjusted model parameters. However, the experimental TG profile is not as rich in information as the table indicates, because some TG species occur at very low concentrations and they either are not detectable or cannot be quantified. Therefore, the number of FA should be reduced, neglecting those whose concentration in the oil is below a certain limit, *e.g.* 1 wt-%.

Another possibility of reduction of the number of FA is grouping of similar acids. This is described in Section 4 where 13 FA of date seed oil were reduced to five pseudo-components. The reduction was probably too severe as the experimental TG profile consisted of 55 concentrations and thus up to eight to nine FA could be safely included in model C according to Table 8. The sensitivity of the results to the number of included FA will be examined in the near future.

Acknowledgments

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic (project COST D30.001) and by the Czech Science Foundation (project No. 203/04/0120). M.H. and M.L. acknowledge the support of grant project No. 203/06/0219 provided by the Czech Science Foundation. The authors gratefully acknowledge valuable comments of the manuscript reviewers.

Conflict of interest statement

The authors have declared no conflict of interest.

References

[1] M. Buchgraber, F. Ulberth, H. Emons, E. Anklam: Triacylglycerol profiling by using chromatographic techniques. *Eur J Lipid Sci Technol.* 2004, **106**, 621–648.

[2] S. H. Fatemi, E. G. Hammond: Glyceride structure variation in soybean varieties: I. Stereospecific analysis. *Lipids.* 1977, **12**, 1032–1036.

[3] P. Laakso, W. W. Christie: Chromatographic resolution of chiral diacylglycerol derivatives: Potential in the stereospecific analysis of triacyl-*sn*-glycerols. *Lipids.* 1990, **25**, 349–353.

[4] T. Takagi, Y. Ando: Stereospecific analysis of triacyl-*sn*-glycerols by chiral HPLC. *Lipids.* 1991, **26**, 542–547.

[5] P. Damiani, L. Cossignani, M. S. Simonetti, B. Campisi, L. Favretto, L. Gabrielli Favretto: Stereospecific analysis of the triacylglycerol fraction and linear discriminant analysis in a climatic differentiation of Umbrian extra-virgin olive oils. *J Chromatogr A.* 1997, **758**, 109–116.

[6] S. Momchilova, K. Tsuji, Y. Itabashi, B. Nikolova-Damyanova, A. Kuksis: Resolution of triacylglycerol positional isomers by reversed-phase high-performance liquid chromatography. *J Sep Sci.* 2004, **27**, 1033–1036.

[7] S. Momchilova, Y. Itabashi, B. Nikolova-Damyanova, A. Kuksis: Regioselective separation of isomeric triacylglycerols by reversed-phase high-performance liquid chromatography: Stationary phase and mobile phase effects. *J Sep Sci.* 2006, **29**, 2578–2583.

[8] P. Sandra, A. Medvedovici, Y. Zhao, F. David: Characterization of triglycerides in vegetable oils by silver-ion packed column supercritical fluid chromatography coupled to mass spectroscopy with atmospheric pressure chemical ionization and coordination ion spray. *J Chromatogr A.* 2002, **974**, 231–241.

[9] H. Kallio, K. Yli-Yokipii, J.-P. Kurvinen, O. Sjövall, R. Tahvonen: Regioisomerism of triacylglycerols in lard, tallow, yolk, chicken skin, palm oil, palm olein, palm stearin, and a transesterified blend of palm stearin and coconut oil analyzed by tandem mass spectrometry. *J Agric Food Chem.* 2001, **49**, 3363–3369.

[10] H. Leskinen, J.-P. Suomela, H. Kallio: Quantification of triaclglycerol regioisomers in oils and fat using different mass spectrometric and liquid chromatographic methods. *Rapid Commun Mass Spectrom.* 2007, **21**, 2361–2373.

[11] M. Holčapek, M. Lísa, P. Jandera, N. Kabátová: Quantitation of triacylglycerols in plant oils using HPLC with APCI-MS, evaporative light-scattering, and UV detection. *J Sep Sci.* 2005, **28**, 1315–1333.

[12] R. J. Vander Wal: Calculation of the distribution of the saturated and unsaturated acyl groups in fats, from pancreatic lipase hydrolysis data. *J Am Oil Chem Soc.* 1960, **37**, 18–20.

[13] M. H. Coleman: The structural investigation of natural fats. *Adv Lipid Res.* 1963, **1**, 1–64.

[14] E. Martínez-Force, N. Ruiz-López, R. Garcés: The determination of the asymmetrical stereochemical distribution of fatty acids in triacylglycerols. *Anal Biochem.* 2004, **334**, 175–182.

[15] M. H. van Vliet, G. M. P. van Kempen: Computational estimation of the triacylglycerol composition of vegetable fats from gas and liquid chromatography data. *Eur J Lipid Sci Technol.* 2004, **106**, 697–706.

[16] R. A. Coleman, D. P. Lee: Enzymes of triacylglycerol synthesis and their regulation. *Prog Lipid Res.* 2004, **43**, 134–176.

[17] J. Reske, J. Siebrecht, J. Hazebroek: Triacylglycerol composition and structure in genetically modified sunflower and soybean oils. *J Am Oil Chem Soc.* 1997, **74**, 989–998.

