Available online at www.sciencedirect.com
SCIENCE DIRECT
Thermochimica Acta 439 (2005) 94–102
thermochimica acta
www.elsevier.com/locate/tca

Use of first and second derivatives to accurately determine key parameters of DSC thermographs in lipid crystallization studies

Laziz Bouzidi [1], Marc Boodhoo, Kerry L. Humphrey, Suresh S. Narine *

Agri-Food Materials Science Centre, University of Alberta, Edmonton, Alta., Canada T6G 2P5

Received 21 June 2005; received in revised form 5 September 2005; accepted 10 September 2005
Available online 17 October 2005

Abstract

In this study, an objective method to provide sufficiently unbiased analysis of raw DSC data, using the calculated first and second derivatives in combination with error analysis is described and applied to analyze thermograms of lipid samples. A statistical method based on a closer study of the residuals of fit and data correlation techniques with adequate criterion for the goodness of fit has proven to be valuable in evaluating the experimental error. The derivatives were utilized to define unambiguously the key signal characteristics, such as minima, maxima, as well as end and start points of thermal events, with uncertainties directly related to the actual noise of each signal. By estimating the pure experimental error and data correlation length, we have been able to have a better appreciation of the quality of the experimental data and detailed information on our DSC system. The calculated errors using 15 runs for the same sample were consistent with the estimated standard deviations generated by the statistical analysis.

Keywords: Fitting to data; Residuals of fit; First derivatives; Second derivatives; DSC thermograms of lipids

1. Introduction

Use of the DSC as an analytical, diagnostic and research instrument is ubiquitous in the lipid literature. It is an efficient tool widely and commonly used, for example, for "fingerprinting" and in materials research. Information, such as peak positions, widths, areas, heights and shapes of the peaks are routinely extracted from DSC thermograms and used as appropriate. A DSC thermogram is often complex, because various modes of crystallization, solid-state transformations and melt-mediated transformations could be involved and contribute to a single peak. Users of DSC usually rely on the software which accompanies the purchase of the equipment for the handling of the output data. However, the determination of some key elements is visually (manually) obtained for subsequent calculations. As an example, in the most recent TA Instruments analysis software currently in use, the "TA Universal Analysis 2000", all the key starting points for analysis are dependent on operator choice. Visual examination is inherently subjective, fraught with possible errors of interpretation, and thus, different analysts (even the same analyst analyzing the data twice) might come to different conclusions from the same data plot. Evidently, an unbiased determination of the main features of a DSC thermograph should start by eliminating the variability introduced by the operator. A more rigorous approach in dealing with output data and aimed at the reduction of the variability on the data analysis, such as that proposed by Foubert et al. [1], who have developed a calculation algorithm for the start and end points of a thermal event, to determine the integration limits of DSC crystallization signal, is a necessary alternative.

In this paper, we will describe an objective method to provide sufficiently unbiased analysis of raw DSC data, using the calculated first and second derivatives in combination with error analysis. Furthermore, we make use of statistical analysis to extract information on the errors associated with the data, define unambiguously the signal characteristics, such as minima, max-

* Corresponding author. Tel.: +1 780 492 9081; fax: +1 780 492 8855.
E-mail addresses: laziz.bouzidi@ualberta.ca (L. Bouzidi), Suresh.narine@ualberta.ca (S.S. Narine).
[1] Tel.: +1 780 492 2871.

0040-6031/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.tca.2005.09.013

ima, as well as end and start points, and apply the method to analyze thermograms of lipid samples.

2. Experimental method

The crystallization and melting experiments were performed on a "TA 2920 Modulated DSC" system (TA Instruments, New Castle, DE, USA). Dry nitrogen was used to purge the thermal analysis system. A relatively pure (97%) 1,3-dimyristoyl-2-stearoylglycerol (MyStMy) triacylglyceride (TAG) was extensively investigated for the application of the method. Myristic and lauric binary systems used in our laboratory by Boodhoo and Narine [2] for a study of phase behavior of binary lipid systems have also been used to illustrate the method outlined in this paper. The lauric binary systems are a mixture of 1,3-dilauroyl-2-stearoylglycerol (LaStLa) and 1,2-dilauroyl-3-stearoylglycerol (LaLaSt) TAGs and the myristic binary systems are a mixture of MyStMy and 1,2-dimyristoyl-3-stearoylglycerol (MyMySt) TAGs. The purity of the individual TAGs was greater than 97%.

The samples were hermetically sealed into aluminum pans and an empty pan was used as a reference. The DSC thermogram of an empty pan was recorded prior to the sample measurements at the same conditions as the sample. The baselines were recorded before every new set of experiments, at the start of the day if running experiments overnight and after a new scan rate. The mass of each sample was within 13.0 ± 0.1 to 14.0 ± 0.1 mg. The data sampling and temperature control procedures were fully automated and controlled by the "TA Instrument Control" software program. For all the samples, if not stated otherwise, the procedure was as follows. Initially, the sample was kept at 20 °C for 20 min to reach steady state and then was heated to 90 °C at a rate of 5 °C/min and kept there for at least 5 min to erase its thermal history. To record the crystallization curve, the sample was cooled down at constant rate (depending on the experiment) to −5 °C and kept at this temperature for 20 min to allow for the completion of crystallization. The sample was then heated to 90 °C at a constant rate of 5 °C/min to record the melting curve. MyStMy (13.4 ± 0.1 mg) was run 15 times to test for reproducibility (it was equilibrated at 90 °C, then crystallized with a rate of 3 °C/min to −5 °C, where it was kept for 20 min and then melt at a rate of 5 °C/min up to 90 °C).

In this paper, the exothermic signals from DSC are reported in the upward direction.

3. The analysis method

To extract accurate information from a raw signal, we first assess the level of "error" associated with the data by calculating its standard deviation (S), compute the derivatives and then use them to find the signal characteristics, such as the maxima, minima and inflexion points, and find the starting and ending points of the event.

3.1. Data analysis, modeling limits and real spectrum

The scientific literature is replete with models and theoretical developments of thermodynamic and kinetic behavior of materials (including lipids) incorporating DSC experimental data but with a few exceptions, not much attention has been paid to the quantitative description of DSC thermograms, an essential step if the models are to be used predictively [3–6].

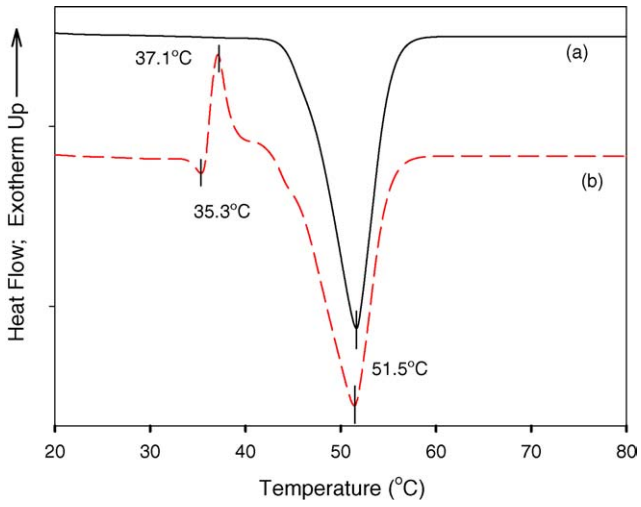

Fig. 1. Melting curves of MyStMy:MyMySt (40:60, w/w) binary system obtained with a melting rate of 5 °C/min after crystallization at different rates. Curve (a) after a crystallization rate of 0.1 °C/min and curve (b) after a crystallization rate of 3 °C/min.

A DSC thermogram is a complex combination of superposed thermal phenomena. DSC thermograms are particularly complex in the case of lipids even for simple triglyceride mixtures, where melt-mediated transformations between metastable phases can complicate the melting event. Such melt-mediated transformations themselves being dependent on the thermal history of the sample as illustrated in Fig. 1. Fig. 1 shows the melting curves of the binary system MyStMy:MyMySt (60:40, w/w) obtained with a melting rate of 5 °C/min after processing with two different crystallization rates. For a crystallization rate of 0.1 °C/min (curve (a)), only a single melting peak centered at 51.5 ± 0.1 °C appeared, whereas at a rate of 3 °C/min (curve (b)), two distinct melting peaks (35.3 ± 0.1 and 51.5 ± 0.1 °C, respectively) appeared separated by a melting-mediated crystallization at 37.1 ± 0.1 °C. In general and apart from the single and well defined line of a pure element in its most stable polymorphic state (or of an element not demonstrating polymorphism), a clear definition–description of discriminated events of the signal is not obvious.

The background-baseline which is due to complex contributions of the thermal behavior of the material and the apparatus is an unavoidable component of the signal. However, functions describing the pre- and post-transition baselines are very efficient tools which can help to assess the overall noise as illustrated in Fig. 2. Fig. 2(a) shows the melting curve of the pure MyStMy sample, crystallized at a rate of 1 °C/min and melted at 5 °C/min with its pre- and post-transition baselines (y_1 and y_2, respectively) fitted with straight lines (R^2 was 0.998 and the standard error of estimates was 0.002 for both sections). Fig. 2(b) highlights the post-baseline (60–90 min interval) and highlights the very good fit obtained. Fig. 2(c) shows the residuals which are less than three times the standard deviation.

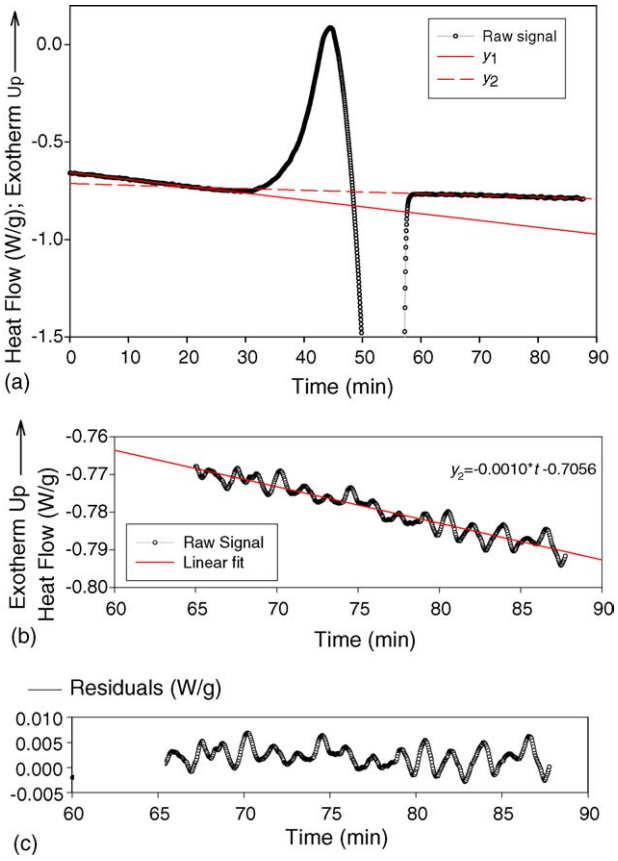

Fig. 2. (a) Melting curve of the pure MyStMy sample, crystallized at a rate of 1 °C/min and melted at 5 °C/min with its pre- and post-baselines (y_1 and y_2, respectively) fitted with straight lines; (b) zoom on the post-baseline (60–90 min interval); (c) residuals of fit.

Ideally, a pure signal starts at time zero with a zero amplitude value and increases (or decreases, depending on whether it is an endotherm or exotherm) as time increases, reaches its maximum (minimum) and dies out at infinity. Practically, the range where the "event" is measured is finite and the extent of its tails depends on the experimental conditions and DSC system performances. A signal without the baseline is considered a real thermal event if it is larger than the noise associated with the signal. To reduce the variability on the data and accurately define the start and end points, the analysis should start and end as far as possible from the extrema of the peaks.

3.2. Assessing the "noise"

In any experimental bivariate set of data $\{x,y\}$, noise is manifest in the spectrum on x or y alone by random errors, drifts or systematic errors. The first step of the analysis method is to assess accurately the noise. "Fitting" procedures are a common means of extracting it from raw data. When we approximate two-dimensional data containing n experimental points $(x_i, y_i)_{i=1,\ldots,n}$ by a smooth curve $[x,f(x)]$ to the unknown reality $g(x)$, the measured data are supposed to contain a noise component δ_i in y_i so that:

$$f_i = g(x) + \delta_i \tag{1}$$

The fitting procedures are primarily based on computing of the sum of squares of the residuals which are the difference between the data and the fitting function. The classical fitting strategy consists of minimizing the so-called Chi-square (χ^2) [7]:

$$\chi^2 = \sum_{i=1}^{n} \mathrm{d}_i^2 \tag{2}$$

where the normalized residual d_i is:

$$d_i = \frac{y_i - f(x_i)}{\sigma_i} \tag{3}$$

σ_i is the standard deviation of the uncertainty δ_i associated with the experimental data y_i. The standard deviation S is defined as the root-mean-square of the residuals between observed y_i and calculated y_i^{calc} values:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - y_i^{\mathrm{calc}})^2}, \text{ with } i = 1, \cdots, n \tag{4}$$

The statistical analysis of the residuals and the standard deviation associated with a data set leads to a quantification of its associated experimental error and of the correlation between data points [8] as in the so-called Durbin–Watson (DW) statistic [9–11]. The DW statistics is the basis of several good algorithms in physical data treatment and is very effective. It also has the considerable advantage of its relative insensitivity to the prior knowledge of the experimental uncertainties. The DW statistics outcome is much more insensitive to user misjudgments of the noise magnitude than if the χ^2 is used, as the χ^2 value depends on the estimate of the measurement errors [12]. Sophisticated approaches to analysis of residuals combining the DW statistic with other statistical tests can be found in the literature (see, for example [13–15]).

To fit the data, we have used a powerful method developed by Thijsse et al. and first published in 1998 [16]. In their approach, the x-interval is divided up in a certain number of intervals and uses piecewise polynomial functions (splines) to construct a smooth curve through data points containing noise. The break points (or knots) distribution is optimized and used to control the flexibility. The latest version of their software, including the program source, executable and user manual, can be freely downloaded from Thijsse's website [17]. Their approach assumes that the correlation function is a decaying exponential. It uses the so-called generalized DW statistics, defined as:

$$Q_m(\xi) = \frac{n-1}{n-m}\frac{\sum_{i=1}^{n-m}(d_{i+m} - d_i)^2}{\sum_{i=1}^{n}(d_i)^2} + \frac{2(n-1)}{n(n-m)}\sum_{i=1}^{n-m} e^{-(x_{i+m}-x_i)/\xi} \tag{5}$$

ξ is the assumed correlation length in the data and m is an integer characterizing the extent of imposed correlation. m is taken large enough so that d_i and d_{i+m} are not correlated, and therefore, can account for the correct noise component.

There are two possibilities to work with the spline program. One is to do an automatic search of autocorrelation using a statistic $\bar{Q}(\xi)$ that is m-averaged over the relevant part of the autocorrelation:

$$\bar{Q}(\xi) = \frac{1}{m_{\max}} \sum_{m=1}^{m_{\max}} Q_m(\xi) \tag{6}$$

with

$$m_{\max} = \left[\frac{3(n-1)\xi}{x_n - x_1} + 3 \right]. \tag{7}$$

The other possibility is to fix m and ξ at any value during the fit and use $Q_m(\xi)$ as statistic.

Both possibilities have been used to fit the data without experimental uncertainty values as they were not available. Due to the great flexibility of the polynomial function, however, a significant "overfitting" has been experienced, and therefore, led to an underestimation of the noise with all the DSC data analyzed for this study when an automatic search for correlation was used. The generalized DW statistics $Q_m(\xi)$ with a fixed m and a careful choice of the correlation length were necessary to achieve a good estimation of the experimental error. Care was taken to ensure that the algorithm recognizes the correct noise component and that the generalized DW parameter was as close as possible to its theoretical value of 2.00. All the DSC data collected by our system were satisfactorily fit using $m = 3$ and appropriate ξ values. Physically meaningful derivatives with relatively very few artifacts were evaluated using this last procedure.

3.3. Calculating the derivatives

Several methods have been developed to calculate the derivatives of noisy data. The most popular methods use the so-called Savitzky–Golay convolution functions (the original Savitzky–Golay paper [18] was cited 3885 times). However, the computation increases the standard deviation of the noise contribution and depending on how the derivatives are calculated, they could generate artifacts without any relationships with the true spectral features [19]. Moreover, the value of the second derivative is roughly an order of magnitude smaller than the first derivative which induces a degradation of its signal (i.e. the value of the derivative here) over noise by the same factor. If a random variable is multiplied by a constant (c, say), then the variance of the product is increased by a factor of c^2 (Mark and Workman [20]). Since the second derivative calculation is equivalent to using coefficients 1, −2 and 1 as multipiers for three data points at the desired x-spacing, the standard deviation of the noise contribution to a second derivative is $\sqrt{6}$ greater than the noise of the spectrum. Similarily, the noise contribution to a first derivative is $\sqrt{2}$ greater than the noise of the spectrum [21].

To calculate the derivatives, we have used both Tijsse's software and SigmaPlot software V9 for Windows (SPSS Inc., Chicago, IL, USA).

3.4. Utilizing the derivatives of thermograms in association with S

The first and the second derivatives are used according to the level of noise associated with them. Strictly, a measurement minus the baseline is considered an actual event (i.e. the feature is not "noise") when its value is larger than the standard deviation attached to the signal. In order to accommodate for unforeseen sources of errors and ensure that no spurious event is mistaken for a real feature, the signal is considered as "noise" or "zero event" when its value is less than $2S$. Accordingly, the values of first and second derivatives are assumed greater than zero when they are greater than $S' = 2\sqrt{2} \times S$ and greater than $S'' = 2\sqrt{6} \times S$, respectively. This is reasonable since the residuals have always been found less than two times the calculated standard deviation in the end and start of the event regions. $2S$, S' and S'' will be called the "departure value" (D_v) from the DSC signal, first derivative signal and second derivative signal, respectively. Note that the baseline was subtracted when using the DSC signal. In the case of DSC thermograms, second derivatives are data treatment of choice because they flatten to zero the baseline at the start and at the end of the cycle. However, second derivatives have reduced signal over noise compared to first derivative, give rise more frequently to spurious features and are hard to interpret in regions where the signal changes rapidly. The second derivative is therefore used in conjunction with the first derivative.

The start and end points of thermal events were determined by tracking the "departure value" (D_v) starting from the zero (of the DSC signal, the first or second derivative). The start (or end) of the event was the average of the range where the signal is in absolute value strictly larger than zero $+D_v$ and strictly smaller than zero $+2D_v$. The error associated with its determination is the extent of the range itself.

The derivatives are also used as a practical procedure of assessing the instrument resolution by quantitatively identifying resolved events and locating shoulders which could not be discriminated by other means, such as the van Ekeren resolution criterion of identifying caloric events [22]. The van Ekeren resolution criterion (or resolution factor) R as redefined and used by Marti et al. [23] does not account for weak shoulders even if they are apparent. It is defined as:

$$R = \frac{h_{\text{peak}}}{h_{\min}} \tag{8}$$

where $h_{\min}$ is the minimum between two caloric events and h_{peak} is the peak height of the second (weak) transition. Fig. 3 contains three different cases of overlapping contrived DSC curves. The first case (Fig. 3(a)) can be quantitatively described with $R = 1.25$. The second case (Fig. 3(b)) with a prominent shoulder gives $R = 1$ and is the limit which can be quantitatively described with the parameter R. The quantitative determination of the resolution factor according to Eq. (7) is not possible in the last case (Fig. 3(c)) where overlapping phenomena are clearly visible but without a minimum in between.

The first and second derivatives define precisely the positions of the resolved single peaks and partially separated peaks.

A point in the signal is a minimum (maximum) at the position where the first derivative intersects with the x-axis and where the second derivative has a maximum (minimum). For the weakly resolved peaks, the first derivative shows with an inflexion point, observable between the succession of a minimum and a maximum, and the second derivative shows with an extremum between them.

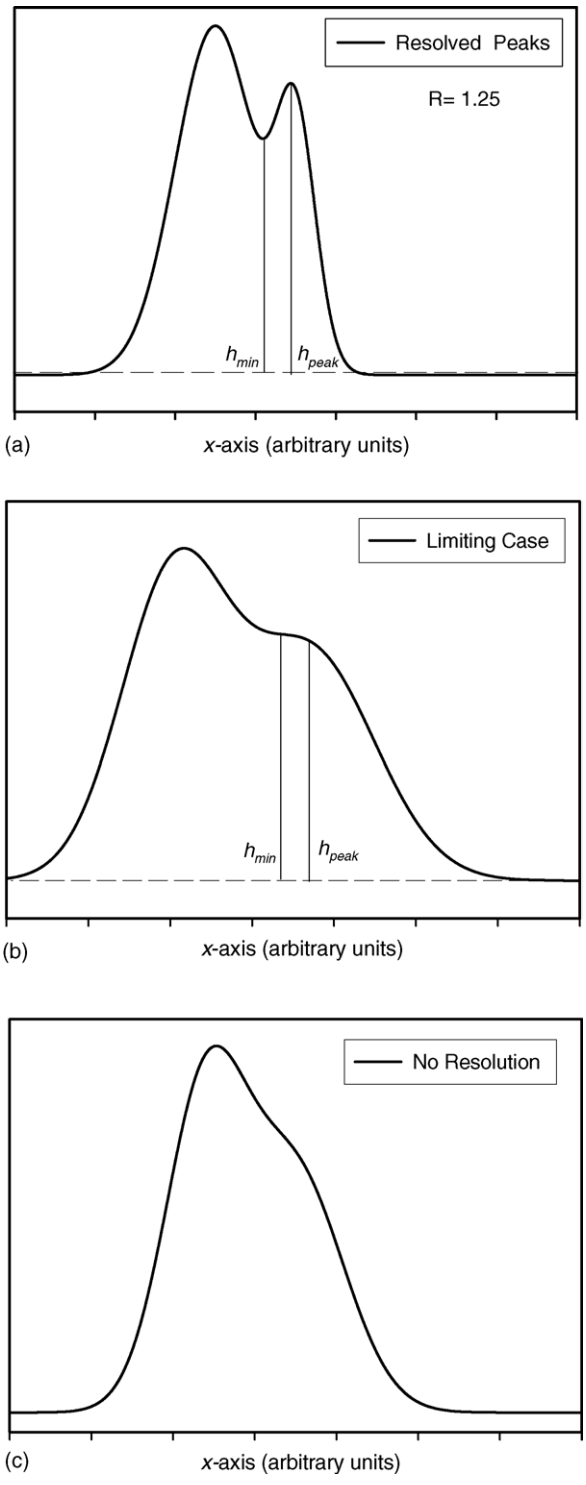

Fig. 3. Three different cases of overlapping contrived DSC curves: (a) case where resolution can be quantitatively described with parameter R; (b) case with a prominent shoulder which can be quantitatively described with the limit $R = 1$; (c) case where overlapping phenomena are clearly visible but without a minimum in between. The quantitative determination of the resolution factor according to Eq. (7) is not possible in this case.

The x-accuracy region (time or temperature) of such determined points is taken as the domain where its value comprised within an absolute error of S. All signal characteristics determined with the two above definitions are unique. For example, the starting temperature of a thermal event will be determined as the average between the temperature at which the signal starts strictly to depart from "zero" ($T_0(0)$) and the temperature at which the signal is zero plus a standard deviation $T_s(0 + S)$ with an accuracy equal to half the difference between $T_0(0)$ and $T_s(0 + S)$. Section 4 demonstrates that this is amply justified.

3.5. Synopsis of the method

Starting from the raw DSC Data spline was run both

- with automatic search of autocorrelation;
- using $Q_m(\xi)$ as statistic: fixe m (in our case, $m = 3$) and ξ at values which give the best DW statistics, starting with the average response time of the system.

The results obtained were:

1. the "best" value of ξ, the DW statistics, the rms (S as no experimental uncertainties were supplied);
2. the spline approximation data set;
3. the extrema and inflexion points which were evaluated with the evaluation function of the software;
4. the first and second derivative data sets.

The results were used as follows:

- sort out maxima, minima and inflexion points by comparing to the actual thermogram;
- locate the end and start of an event using the computed derivatives by:

  Defining the linear trend of signal: use of the first derivative to locate the baselines left and right of the peak, it is where the first derivative of the linear tails of the signal are less than or equal to $2\sqrt{2} \times S$ off its constant value.

  Locating where the event departs strictly from zero (the start or the end of the event) with the second derivative, it is where the second derivative departs strictly from $D_v(= 2\sqrt{6} \times S)$. If not possible with the second derivative, the first derivative is used the same way with $D_v = 2\sqrt{2} \times S$. If not possible with the first derivative, the data set of the heat flow minus the baseline is used with $D_v = 2 \times S$.

  Defining the temperature range where the absolute value of the second derivative (first derivative or heat flow signal) in the range from D_v (time t_1 and temperature T_1) to $D_v + D_v/2$ (time t_2 and temperature T_2). The temperature was taken as the average temperature of the temperature range and the error attached to it equal to half the magnitude of the temperature range $(T_2 - T_1)/2$.

- the spline was used:
  to integrate the signal above the sigmoidal baseline, using the start and end points;
  to obtain the integral the peak height, the full width at half maximum (FWHM) of the peak and eventually the onset of the event.

4. Results and discussion

The baseline recorded with two empty pans for our system was a straight line in the domain where the thermal events of interest are studied. It has slightly different slopes for the heating process and the cooling process ($0.006 \pm 1.0 \times 10^{-4}$ W/(g °C) against $0.004 \pm 1.0 \times 10^{-4}$ W/(g °C), respectively). The subtraction of the baselines did not introduce any measurable distortion on peak characteristics in any of our thermograms.

Fig. 4 presents an actual DSC crystallization curve of 1,2-distearoyl-3-palmitoylglycerol (PaStSt) TAG (97% pure) crystallized at a rate of 1 °C/min with a visible overlapping phenomenon which cannot be treated with *R* and the use of the derivatives is shown to discriminate them. The first derivative locates accurately the apparent maximum where it intersects the *x*-axis (arrow (1)) and evidences the inflexion point between a maximum and a minimum in the first derivative (arrow (2)). The second derivative precisely locate the inflexion point (arrow (3) at 41.7 °C) and discriminates between two possible different overlapping thermal events (arrow (4) at 43.7 °C and arrow (5) at 45.1 °C). The first and second derivatives of the signal show perfectly the positions of major inflexions in the signal and locate obvious overlapping contributions. The horizontal dashed lines in the figure are a guide to help locate the zeros of the derivatives.

Fig. 5(b) showing the fit of a melting curve of a 20% LaStLa in LaLaSt (w/w) sample crystallized at a rate of 0.1 °C/min and melted at a rate of 5 °C/min illustrates the effectiveness of the spline software. In this case, the automatic search of autocorrelation was used, produced an approximating curve of 180 polynomials (179 knots) of order 3 with a DW statistics of 2.01, close to the theoretical value. The fit followed closely the raw signal, particularly, dramatically in the region where

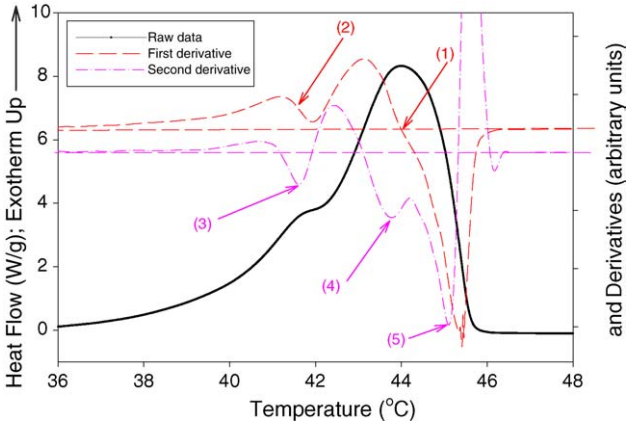

Fig. 4. Actual DSC crystallization curve of a pure PaStSt TAG crystallized at a rate of 1 °C/min with a visible overlapping phenomenon which cannot be treated with *R*. The first derivative is shown to discriminate them.

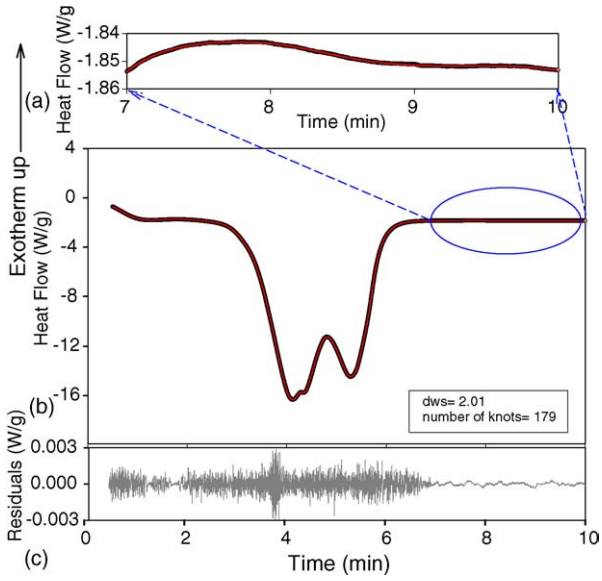

Fig. 5. Melting curve of a 20% LaStLa/LaLaSt sample crystallized at a rate of 0.1 °C/min and fit with polynomials (splines) functions, classical DW statistics (Q_1) and an automatic search of autocorrelation. (a) Enlarged region [7–10 min] of the curve where the variations of the heat flow with time were small. The approximating curve was obtained using 180 polynomials (179 knots) of order 3. (b) The main curve. (c) The residuals of fit. The signal over noise ratio (S/N) is greater than 10^{4+}. The horizontal dashed lines in the figure are a guide to help locate the zeros of the derivatives.

the variations of the heat flow with time were small as seen in the enlarged region [7–10 min] ((Fig. 5(a)) circled and shown above the main curve). The residuals as shown in Fig. 5(c) are very small and the subsequent calculated signal over noise ratio (S/N) is still very large (greater than 10^{4+}) even for the region where the heat flow varies very rapidly with time.

As shown in Fig. 6, the crystallization curve of a pure MyStMy sample (crystallized at a rate of 5 °C/min), when fit using the automatic search, generated a very small correlation length ($\xi = 0.001$ °C) and a large optimized number of knots of 446 (Fig. 6(a)). The calculated standard deviation (0.0003) was visibly not a correct assessment of the experimental noise as shown in the 30 min "flat" portion of the curve (Fig. 6(b)). The portion was fitted with 179 knots and it is obvious that even the noise has been fitted. When the generalized DW statistic with fixed $m = 3$ and ξ ($\xi = 0.05$ °C) statistics was used as shown in Fig. 6(c) for the same portion of the curve as in Fig. 6(b), a smaller number of knots (32) led to a correct fit of the data. As can be seen, the fit-curve follows the "true baseline" behavior instead of tracking the noise. The peak itself was fit with residuals less than three times the calculated standard deviation.

The standard deviation associated with temperature measurement was estimated by fitting the temperature versus time ($T(t)$) curves for isothermal portions of the DSC experiments as well as for constant heating or cooling rates. It was also calculated from heat flow versus time curves in the regions with no thermal event. Fig. 7 shows a typical example of a $T(t)$ curve of a MyStMy sample heated to a final set-temperature of 90 °C at a rate of 5 °C/min and left at this temperature for 30 min. The temperature was effectively stabilized in about 18 min (arrow

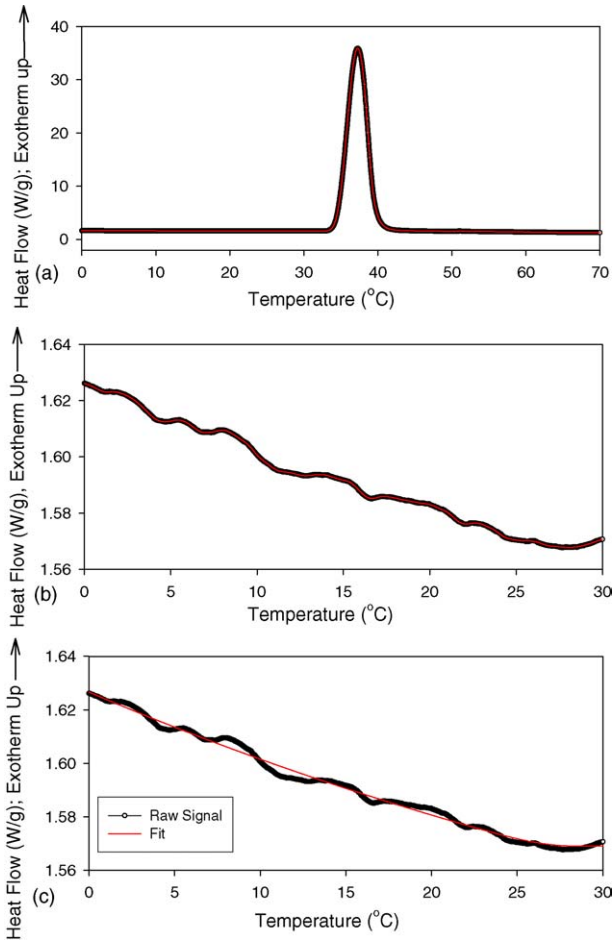

Fig. 6. (a) Crystallization curve of a pure (97%) MyStMy sample crystallized at a rate of 5 °C/min fit using Q_1. The correlation length is $\xi = 0.001$ and the optimized number of knots is 446. (b) The 30 min portion of the curve fitted using Q_1 statistics. Number of knots is 179 knots and the correlation length $\xi = 0.001$. (c) Same flat portion as in (b) fitted using Q_3 statistics. Number of knots is 32 knots and the correlation length $\xi = 0.05$.

(1)). It dropped by only 0.2 °C and the peak-to-peak noise was equal to 0.02 °C underlining the effectiveness of the temperature controller. The residuals had a nearly periodical pattern and were evenly distributed around zero with extrema less than $2S$ as pointed out by the dashed lines in Fig. 7(b). The calculated first derivative departs from zero (as outlined in Sections 3 and 4) exactly where the value of T is above half the point-to-point (ptp) noise (arrow (2)). The average standard deviations and correlation lengths obtained for the flat parts of the DSC thermograms are listed in Table 1.

Fig. 8 examplifies the treatment of start and end points of thermal events. It shows an enlarged region of the melting curve of

Table 1
Correlation length, standard deviation and point-to-point noise (ptp) for the temperature and heat flow in the regions of the thermograms where there were no peaks

	Correlation length (min)	Standard deviation	ptp noise
$T(t)$	0.01	0.008	0.02
$H(t)$	0.07	0.01	0.04

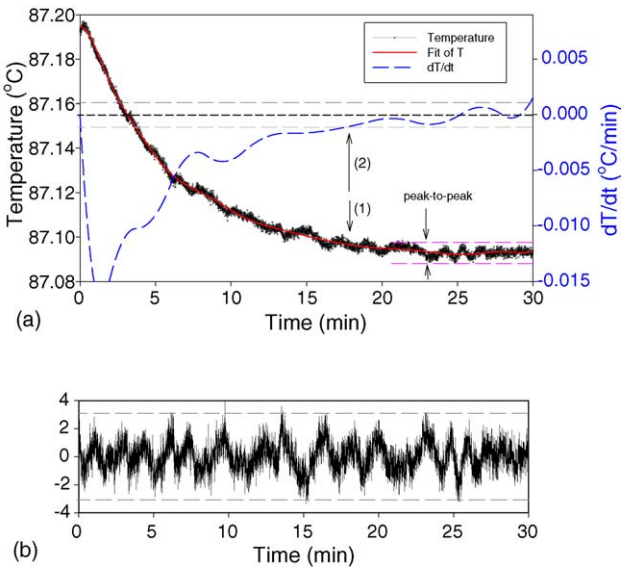

Fig. 7. (a and b) Temperature vs. time $T(t)$ curve and its first derivative for a MyStMy sample heated to a final set-temperature of 90 °C at a rate of 5 °C/min and left at this temperature for 30 min. The residuals are plotted below in standard deviation units. The dashed lines are a guide for the eye to locate the zero of the derivative and the peak-to-peak noise associated with the temperature in the flat region [20 and 30 °C].

a pure (97%) MyStMy sample crystallized at a rate of 3 °C/min and melt at a rate of 5 °C/min and its first derivative (the entire melting curve and its first derivative is shown in the upper left corner of the figure), the arrow (1) directs to the point where the first derivative departs from "zero", i.e. $1D_v$ and the arrow (2) directs to the end of the temperature determination range, i.e. $2D_v$.

The temperatures of start and of end of thermal event for the individual runs were determined within an average of ±0.05 °C accuracy for crystallization and of ±0.08 °C accuracy for melt. The maxima and minima of the signals have been determined with average uncertainties of 0.1 °C. The start and end points for the 15 identical runs of the crystallization and the melting of the MyStMy system were determined using the method outlined

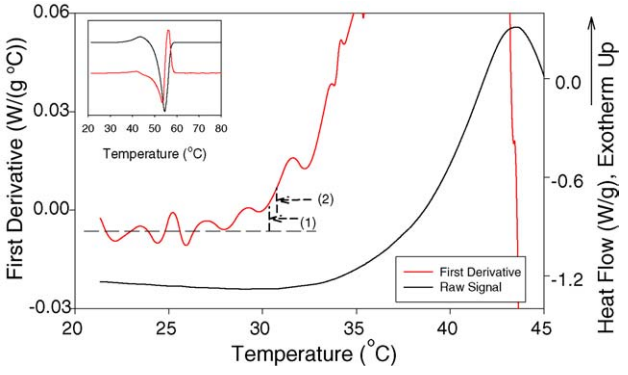

Fig. 8. Enlarged region of the melting curve of a MyStMy sample crystallized at a rate of 3 °C/min and melt at a rate of 5 °C/min and its first derivative. The entire melting curve and its first derivative are shown in the upper left corner of the figure. The long-dashed line is a guide for the eye to locate the zero of the derivative.

Table 2a
Characterization of the crystallization peak

15 runs	T_{max} (°C)	FWHM (°C)	Height (W/g)	Area (J/g)	Start of crystallization (°C)	End of crystallization (°C)
Mean	32.437	2.783	2.635	170.614	23.43	36.34
Standard deviation	0.034	0.018	0.009	1.623	0.15	0.12
Standard error	0.009	0.005	0.002	0.434	0.04	0.03

above. The average correlation length used to fit the data was 70 s which is roughly equal to the average response time of 65 s of the DSC pans.

The areas under the signals were calculated using the TA Universal 2000 software and the SigmaPlot software V9 for Windows (SPSS Inc., Chicago, IL, USA) feeding them with the starting and ending points determined with this method. The integrations under curves were performed using the trapezoidal rule. A sigmoidal baseline initially is calculated as a straight line from peak start to peak end. It is then recalculated for each data point between the peak limits as the weighted average between the tangent baselines at peak start and end. The weighting factors for a given point are: (1) one minus the fraction reacted (alpha) times the initial baseline and (2) alpha times the final baseline. The area is then recalculated with the new baseline. If the new area differs from the previous area by more than 1%, the area is recalculated and the sigmoidal curve shifted repeatedly until two consecutive calculations of the area differ by no more than 1%.

As expected, the two calculations yielded exactly the same results for the same curves. The results are reported in Table 2a for the crystallization peak, Table 2b for the melting peak and Table 2c for the crystallization peak mediated by melt. The individual start and end temperatures were determined with accuracies comparable to the repeatability errors listed in the tables and individual determination of the minima and maxima with errors smaller than those related to repeatability. Near the extrema of the peaks where the slopes are the steepest, the residuals are in average less than four times S and have alternating signs with relatively the same periodicity. The signal over noise ratio is still very large (approximately 10^4). This is probably not due to poor "piecing" of the curves into polynomials but to a stronger dependency of the sensitivity of the system on the scale of the thermal event.

The constant heating and cooling rates (temperature versus time curves) are another powerful tool to determine the start and end of a caloric event. When endothermic or exothermic processes occur, the instrument cannot follow exactly its programmed temperature but adjusts more or less depending on the rate itself and other experimental factors, such as the type of instrument, the mass of the sample and the purge gas and

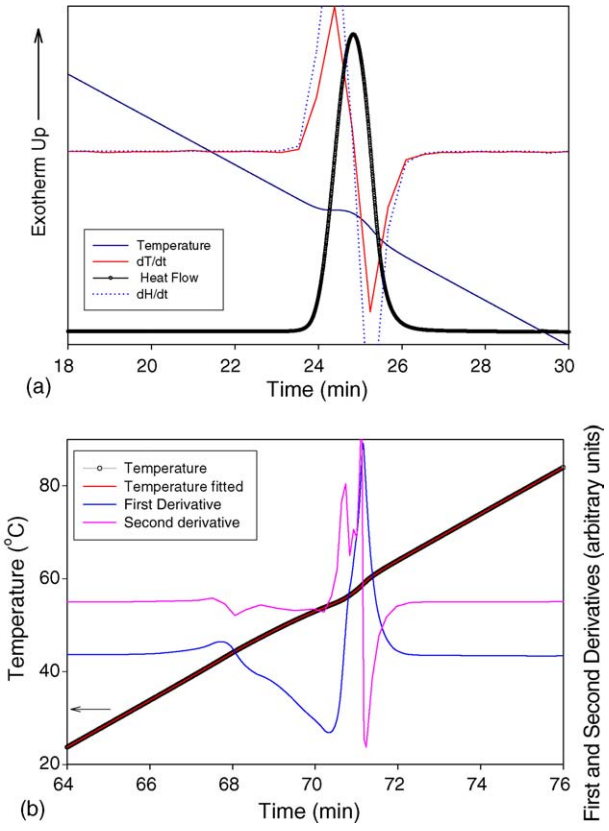

Fig. 9. (a) Temperature vs. time $T(t)$ and heat flow vs. time of the exothermic process occurring for the MyStMy sample crystallized with a programmed rate of 3 °C/min, and their respective first derivatives with respect to time. (b) Temperature vs. time $T(t)$ of endothermic process at a programmed heating rate of 5 °C/min of the same sample as in (a) as well as the first and second derivatives curves of the temperature with respect to time.

its flow rate [24]. During the event, the time dependence of the experimental heating or cooling rate follows very closely the time dependence of the heat flow. The derivatives of the $T(t)$ curves are easier to compute and any deviation to the set linear trend appears clearly. Fig. 9(a) shows the heat flow curve and its first derivative with respect to time for the MyStMy sample crystallized with a programmed rate from 3 to 20 °C/min with the influence of the exothermic process on the cooling rate. Fig. 9(b)

Table 2b
Characterization of the melting peak

15 runs	T_{min} (°C)	FWHM (°C)	Height (W/g)	Area (J/g)	Start of melt (°C)	End of melt (°C)
Mean	55.517	5.952	2.345	180.55	31.67	68.83
Standard deviation	0.040	0.023	0.007	0.756	0.17	0.20
Standard error	0.016	0.009	0.003	0.308	0.07	0.073

Table 2c
Characterization of the crystallization peak induced by melt

15 runs	T_{max} (°C)	FWHM (°C)	Height (W/g)	Area (J/g)	Onset of crystallization (°C)
Mean	44.432	5.432	0.269	18.594	38.99
Standard deviation	0.037	0.062	0.004	0.306	0.04
Standard error	0.015	0.025	0.002	0.125	0.02

shows an example of the influence of endothermic process on the experimental temperature at a programmed heating rate of 5 °C/min of the same sample and shows the plot of the first and second derivatives of the temperature with respect to time. In both cases, the derivatives of the temperature change and of the heat flow show the same trend and yielded the same values for the start and end of the caloric event within ±0.05 °C. The detailed results of the study of the DSC temperature behavior during a caloric event are to be published elsewhere.

5. Conclusion

The method based on analysis of the residuals of fit and the use of first and second derivatives used on DSC thermogram of lipid system was very efficient and yielded accurate and unbiased results. The statistical method based on a closer study of the residuals of fit and associating data correlation techniques with adequate criterion for the goodness of fit has proven to be useful in evaluating both the experimental error and the model bias. By estimating relatively accurately the pure experimental error and data correlation length, we have been able to have a better appreciation of the quality of the data and detailed information on our DSC system. The errors inherent to the DSC system were determined unequivoquely and the values of key parameters, such as start and end of the thermal event, apparent extrema and FWHMs were determined accurately with uncertainties directly related to the actual noise of each signal. The calculated errors using 15 runs for the same sample were consistent with the estimated standard deviations generated by the statistical analysis. This will enable us to compare data from different experiments.

Acknowledgements

The authors acknowledge the technical contributions of Mr. Ereddad Kharraz. The financial support of NSERC, Bunge Corp., AVAC Ltd. and Archer Daniels Midland are gratefully acknowledged.

References

[1] I. Foubert, P.A. Vanrolleghem, K. Dewettinck, A differential scanning calorimetry method to determine the isothermal crystallization kinetics of cocoa butter, Thermochim. Acta 400 (2003) 131–142.
[2] M. Boodhoo, S. Narine, J. Am. Oil Chem. Soc., submitted for publication.
[3] V.J. Griffin, P.G. Laye, in: E.L. Charsley, S.B. Warrington (Eds.), Thermal Analysis, Royal Society of Chemistry, Cambridge, 1992, pp. 17–30.
[4] D. Dollimore, S. Lerdkanchanaporn, Thermal analysis (review), Anal. Chem. 70 (1998) 27R–35R.
[5] D. Dollimore, P. Phang, Thermal analysis (review), Anal. Chem. 72 (2000) 27R–36R.
[6] S. Vyazovkin, Thermal analysis (review), Anal. Chem. 76 (2004) 3299–3312.
[7] N.R. Draper, H. Smith, Applied Regression Analysis, Wiley, New York, 1981.
[8] P.R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw Hill, New York 1969 (new edition with D.K. Robinson, McGraw Hill, 1992).
[9] J. Durbin, G.S. Watson, Testing for serial correlation in least squares regression. 1, Biometrika 37 (1950) 409–428.
[10] J. Durbin, G.S. Watson, Testing for serial correlation in least squares regression. 2, Biometrika 38 (1951) 159–177.
[11] J. Durbin, G.S. Watson, Testing for serial correlation in least squares regression. 3, Biometrika 58 (1971) 1–19.
[12] P. Minguzzi, Some practical tools for everyday least squares, J. Mol. Spectrosc. 209 (2001) 169–177.
[13] J. Bechhoefer, Curve fits in the presence of random and systematic error, Am. J. Phys. 68 (2000) 424–429.
[14] J.L. Femenias, Goodness of fit: analysis of residuals, J. Mol. Spectrosc. 217 (2003) 32–42.
[15] J.L. Femenias, Fitting models to correlated data (large samples), J. Mol. Spectrosc. 224 (2004) 73–98.
[16] B.J. Thijsse, M.A. Hollanders, J. Hendrikse, A practical algorithm for least-squares spline approximation of data containing noise, Comput. Phys. 12 (1998) 393–399.
[17] http://dutsm183.stm.tudelft.nl and can be accessed from the CIP web site at http://www.aip.org/cip/source.htm. Last accessed April 20, 2005.
[18] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627.
[19] H. Mark, J. Workman, Derivatives in spectroscopy, Spectroscopy 18 (2003) 32–37.
[20] H. Mark, J. Workman, Spectroscopy 3 (1988) 13–15.
[21] H. Mark, J. Workman, Derivatives in spectroscopy. Part II—the true derivative, Spectroscopy 18 (2003) 25–28.
[22] P.J. van Ekeren, C.M. Holl, A.J. Witteveen, A comparative test of differential scanning calorimeters, J. Therm. Anal. 49 (1997) 1105–1114.
[23] E. Marti, E. Kaisersberger, W.D. Emmerich, New aspects of thermal analysis. Part I: resolution of DSC and means for its optimization, J. Therm. Anal. Calorim. 77 (2004) 905–934.
[24] R. Keuleers, J. Janssens, H.O. Desseyn, Thermochim. Acta 333 (1999) 67–71.