Regressional Models that Describe Oil Absolute Viscosity

J.F. Toro-Vazquez* and **R. Infante-Guerrero**

Centro de Investigación y de Estudios de Posgrado de la Facultad de Ciencias Químicas, Universidad Autónoma de San Luis Potosí, Zona Universitaria, 78210, San Luis Potosí, México

Equations that describe the temperature dependence (298–338°K) of absolute viscosity (μ) of 21 oils and oil-liquid fat mixtures were obtained based on two different approaches. Fitting each particular viscosity profile to a quadratic extension of the Andrade equation provided the best predictive models ($R^2 > 0.96$). However, the coefficients associated with temperature effect did not have a physical-chemical meaning. In contrast, the multiple variable regressional approach fitted, in just one equation, the μ of all 21 oil systems ($R^2 \approx 0.93$). This equation included terms associated with structural parameters of acylglycerides, namely the degree of unsaturation (*i.e.*, iodine value) and chainlength (*i.e.*, saponification value) of the fatty acids. The models described effects of the *cis* double bonds and fatty acid chainlength on the acylglycerides' interactions that determine both the μ of the system and its capability to crystallize. Therefore, multiple variable regressional analysis might be an excellent tool to better understand the quantitative structure-functional property relationships in lipids systems.

KEY WORDS: Absolute viscosity, Andrade equation, crystallization, iodine values, nucleation, physical properties, regression coefficients, saponification values, triglycerides, vegetable oils.

The flow properties (*i.e.*, rheology) of food components are important factors in determining the texture, overall quality and stability of food systems. Rheological parameters are required in the design of several operations utilized during food processing (1). For example, oil viscosity and its dependence on temperature is an important parameter utilized in monitoring the adequacy of oil/fat utilized in food frying. It is also important for estimating distillation column efficiency for separation of fatty acids, controlling oil hydrogenation and interesterification, as well as designing the piping for oil transport (2–4).

In general, viscosity measures the resistance of a liquid to flow. From a molecular point of view, viscosity is an indirect measurement of the internal friction between the molecules that constitute the fluid and which oppose liquid movement. Therefore, viscosity must have significant correlations with structural parameters of the fluid molecules. For instance, on the basis of a review on published viscosity data, Dutt and Prasad (3) suggested the following trends: (i) The oil unsaturation degree, directly associated with the iodine value (IV), has an inverse relationship with the viscosity of the oil and (ii) the chainlength of the fatty acids that constitute the acylglycerides, indirectly correlated with the saponification value, has a direct relationship with the oil viscosity.

However, the specific nature of such relationships, as well as their interactions as a function of temperature, was not established in that work. To understand the lipid structure-oil viscosity relationships, models of such correlations must be established. The quantitative structure-property relationship approach (QSPRA) has been used in different studies to mathematically correlate a given set of molecular structural properties to certain physical parameters of different types of molecules (3,5–7).

In this paper, we present the equations and correlation constants that describe the variation of the absolute viscosity of oils and oil-liquid fat mixtures. The regression analyses utilized were based on two different approaches. For each oil and oil mixture, the viscosity was fitted to a quadratic extension of the Andrade equation (8), which considers only the temperature as an explanatory variable. In the second method, multiple-variable regressional modeling (9), the IV and the saponification value (SV) of the oils investigated were included, in addition to temperature, as explanatory variables. These last two variables are associated with structural characteristics (*e.g.*, degree of unsaturation and chainlength of the fatty acids, respectively) of the acylglycerides that constitute the fats and oils.

MATERIALS AND METHODS

Viscosity measurements. The vegetable oils utilized included pure oils and mixtures (60:40 and 40:60, vol/vol) of the oils with butters, vegetable butters or shortenings in the liquid state. The oils investigated included different lots of sunflower oil, corn oil, sesame oil and canola oil processed by different companies from Mexico and the United States.

To develop the oil-liquid fat mixtures, the appropriate proportions of oil and fat, previously melted at 60–70°C, were mixed in a beaker, and the oil phase was utilized for the viscosity measurements. To avoid the interference of solid particles during viscosity measurements, when a solid phase developed at a given temperature of viscosity measurement, the mixture was kept at that temperature for 2–3 h and the developed solid was separated. The oil phase was used in the viscosity determination. The purpose of these mixtures was to provide a wider interval of lipid structural properties to be investigated in oil systems.

The absolute viscosity was determined at 298, 308, 318, 328 and 338°K in a Brookfield DV-II viscosimeter (Brookfield, Co., Stoughton, MA), which basically is a Couette instrument. A 600-mL beaker containing 550 mL of the sample was used in all determinations. Temperature control of the sample (±0.1°C) was achieved with a TC-200 Brookfield constant-temperature water bath. The shear stress was determined at shear rates of 0.066, 0.132, 0.33, 0.66, 1.32, 2.64, 6.60 and 13.20 s^{-1}. The absolute viscosity, in dynes·s/cm^2, was calculated from the slope of a plot of shear stress *vs.* shear rate. Using different samples of the same oil lot, duplicate determinations were made, and the mean of the absolute viscosity (μ) at each temperature was utilized in the regressional analysis.

Regression analysis. For each of the oil systems investigated, the variation in absolute viscosity was fitted to an extension of the Andrade equation (8). The form of the equation was:

---

*To whom correspondence should be addressed at Av. Fuente de Hércules 234, Balcones del Valle, San Luis Potosí, 78280, México.

$$\mathrm{Ln}\ (\mu) = \mathrm{A} + \mathrm{B}/\mathrm{T}^\circ + \mathrm{C}/(\mathrm{T}^\circ)^2 \qquad [1]$$

where Ln (μ) is the natural logarithm of the absolute viscosity; A, B and C are constants to be determined; and T° is the temperature of the system in °K. The constants for each oil system were determined by using the nonlinear estimation procedure in the STATISTICA program (StatSoft, Tulsa, OK) (10). The Rosenbrock algorithm was utilized to fit the experimental data to the equation.

Additionally, the IV and the SV of each of the oil systems investigated were determined by following the Association of Official Analytical Chemists' procedures (11). Multiple regression analysis was used to develop equations that describe the Ln (μ) variability among the different oil systems investigated by utilizing as explanatory variables the temperature of viscosity measurement (in °K), the IV and SV values for each oil system, their linear, reciprocal and quadratic effects, as well as their respective linear, reciprocal and quadratic interactions. The variable selection for the multiple regression modeling was performed by using the "forward" methodology available in the multiple regressional procedures in the STATISTICA program (10). The criteria utilized for model selection were the residual analysis, the magnitude of the determination coefficient (R^2) ($R^2 > 0.90$) and the magnitude of the "tolerances" (TOL) for the estimated β-regression coefficients of the explanatory variables in the equation. The TOL values measure the correlation of a particular explanatory variable with the rest of the explanatory variables in the equation (9,10). Thus, the higher the TOL value for the estimated β-regression coefficient, the higher the independence among the explanatory variables in the equation (*i.e.*, multicollinearity) and the lower the standard error of the estimate. The statistical basis of this procedure has been described previously (12).

RESULTS

The plots of shear stress *vs.* shear rate in the determination of oil absolute viscosity provided regression coefficients >0.997 (Fig. 1). Therefore, the oils investigated followed Newtonian behavior (*i.e.*, the oils presented a constant value of μ for all conditions of shear stress and shear rate) (13).

Typical absolute viscosity behavior as a function of temperature is shown in Figure 2. Absolute viscosity decreased as temperature increased; however, the extent and pattern of this decrease was dependent on the type of oil.

The statistics that described the number of replicates, as well as the magnitude of the variation in the variables used, in the regression analyses are shown in Table 1. The regression constants for some of the models developed on the basis of the Andrade equation are given in Table 2. This regressional approach considered each type of oil as an independent system and the absolute temperature as the only explanatory variable. Thus, the magnitude of the equations' coefficients varied considerably among the oil systems studied. The R^2 obtained under this approach was always greater than 0.96, *i.e.*, the temperature explained at least 96% of the total variability observed in Ln (μ).

The oil systems investigated provided wide intervals of unsaturation and chainlength in the fatty acids that constitute the acylglycerides (Table 1). The best equation obtained through the multiple regression analysis was:

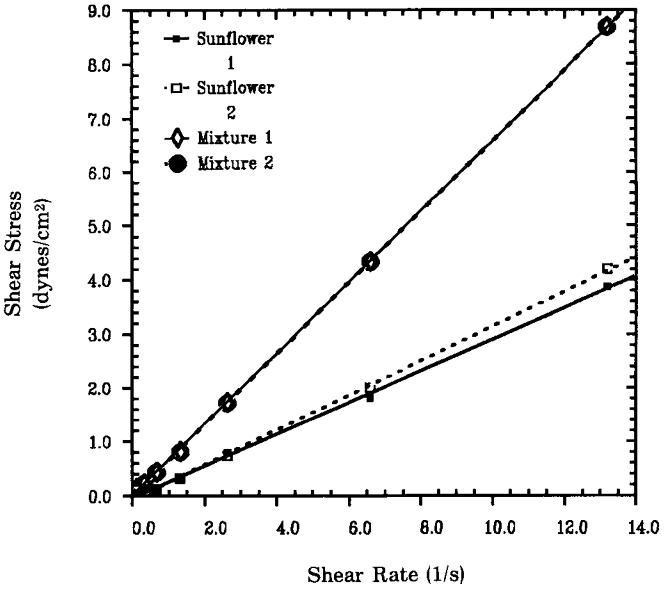

FIG. 1. Shear stress as a function of shear rate for sunflower oil and a 60:40 oil-liquid fat mixture (vegetable butter/sunflower oil). The numbers 1 and 2 represent independent determinations in the same oil.

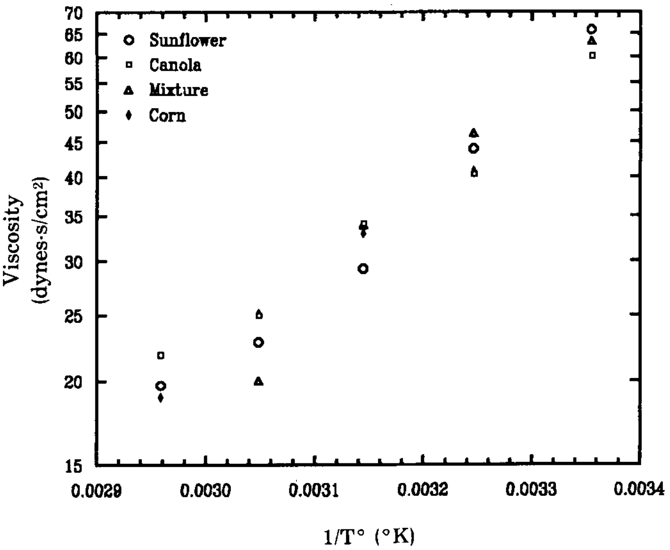

FIG. 2. Absolute viscosity as a function of temperature for three vegetable oils and one oil-liquid fat mixture (60:40, vegetable butter/sunflower oil).

$$\mathrm{Ln}\ (\mu) = -4.7965 + 2525.92962\ (1/\mathrm{T}^\circ) + 1.6144\ (\mathrm{SV})^2/(\mathrm{T}^\circ)^2 - 101.06 \times 10^{-7}(\mathrm{IV})^2 \qquad [2]$$

All equation coefficients (*e.g.*, β-regression coefficients) were significant at $P < 0.10$, and the standard error of the estimates obtained through this equation was 0.1064. In this approach, the data were analyzed independently of the oil source. This assumption generalized the inference

TABLE 1

Statistics That Describe the Experimental Conditions of the Variables Investigated

	n[a]	Mean	SD	Minimum	Maximum
Oils	9	n.a.[b]	n.a.	n.a.	n.a.
Mixtures[c]	12	n.a.	n.a.	n.a.	n.a.
δ[d] (dynes•s/cm^2)	110	34.62	14.01	17.33	74.94
T° (°K)	5	318.00	15.81	298.00	338.00
IV[e]	21	104.86	8.14	90.36	120.11
SV[f]	21	170.59	6.81	155.55	183.45

[a] Number of replicates. Each replicate done by duplicate.
[b] Not applicable.
[c] Vegetable oil-liquid fat in 60:40 or 40:60 (vol/vol) ratios.
[d] Absolute viscosity at the different temperatures.
[e] Iodine value.
[f] Saponification value.

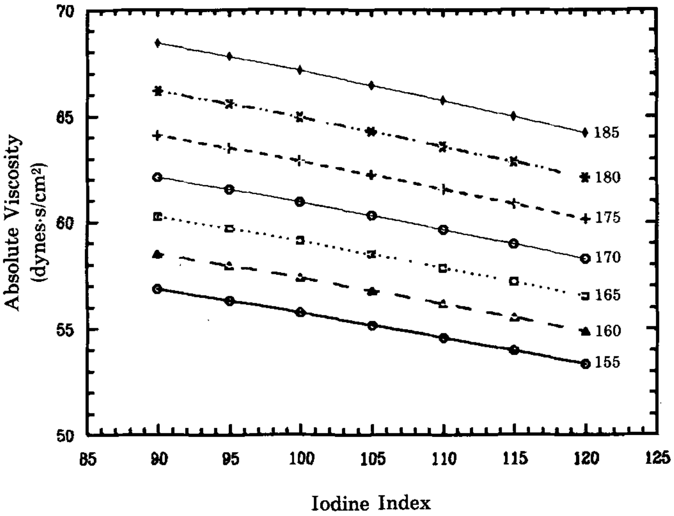

FIG. 3. Absolute viscosity as a function of iodine value and saponification value. The temperature was 298°K. The viscosity was calculated with the equation obtained through the multiple-regressional analysis.

obtained from statistical analysis, because the applicability of the equation was not restricted by the oil source, *i.e.*, the equation obtained was a generalized model. The R^2 value for this equation indicates that $\approx$93% of the total variability in Ln (μ) observed among the different oils was described by the variables present in Equation 2 [*e.g.*, $1/\mathrm{T}^\circ$, $\mathrm{SV}^2/(\mathrm{T}^\circ)^2$, and IV^2]. According to this equation, Figure 3 shows the graphic behavior of μ as a function of IV and SV at 298°K. In the same way, Figure 4 shows the graphic behavior of μ as a function of SV at different temperatures at a fixed IV of 90.

As in the models obtained from the Andrade equation, μ decreased as a quadratic function of temperature. The extent of this decreasing effect of the temperature was independent of the IV of the oil system. However, the temperature effect on μ was significantly affected by the length of the fatty acid chains (*i.e.*, SV) present in the acylglycerides of the oil.

The *cis* configuration of the double bonds in fatty acids is the naturally occurring configurational form. This particular geometry produces a kink in the natural "linear" zig-zag organization of the hydrocarbon chain in the saturated fatty acid molecule. As a result, as the concentration of double bonds increases in the oil system, the intermolecular interactions among the acylglycerides are hindered, decreasing the temperature of oil solidification, as well as the oil viscosity (Fig. 2). The multiple regression equation indicated that, in the oil systems investigated, the concentration of double bonds decreased Ln (μ) in a quadratic way. This effect (*i.e.*, IV^2) was independent of both the chainlength of the fatty acids (*i.e.*, SV) and the temperature of the system (Fig. 3).

The effect of SV on the absolute viscosity depended on the temperature of the system (*i.e.*, $\mathrm{SV}^2/\mathrm{T}^2$). At the lower temperature interval (<318°K), an increase on the length of the fatty acid chains that constituted the acylglycerides of the oil (*i.e.*, a decrease in SV) produced a quadratic decrease in Ln (μ). However, as the temperature increased, the fatty acids' chainlength effect was less apparent (Fig. 3). These results were not in agreement with the SV-to-oil viscosity relationship suggested by Dutt and Prasad (3). It seemed that the model developed was capable of describing the interactions among the acylglycerides that were prone to crystallize at the low temperature interval (<318°K). According to Hartel (14), the onset of nuclei formation during the process of crystallization is preceded by a decrease in molecular diffusivity and a substantial

TABLE 2

Regression Equations for the Natural Logarithm of Absolute Viscosities (dynes•s/cm^2) for Some of the Oils Investigated[a]

Vegetable oils/oil mixtures	Constants			Data points	R^2
	A	B	C		
Sunflower (Mexico)	−6.2361	3063.722	7493.786	5	0.988
Sunflower (United States)	−6.4381	3137.235	−1035.168	4	0.975
Corn (United States)	−4.9536	2677.526	−311.844	4	0.992
Canola (United States)	−4.4068	2501.109	5098.307	5	0.979
Sesame (Mexico)	−4.2709	2485.478	−4907.076	5	0.962
VB/Sunflower[b] (60:40)	−5.9858	3017.555	3370.815	5	0.979
B/Sunflower[c]	−6.7180	3276.498	2399.552	5	0.982
B/Corn[d]	−3.6952	2262.795	667.986	4	0.968

[a] The experimental data were fitted to the extension of the Andrade equation (Ref. 8; Equation 1).
[b] Mixture of melted vegetable butter/sunflower oil in a ratio 60:40 (vol/vol).
[c] Mixture of melted butter/sunflower oil in a ratio 60:40 (vol/vol).
[d] Mixture of melted butter/corn oil in a ratio 40:60 (vol/vol).

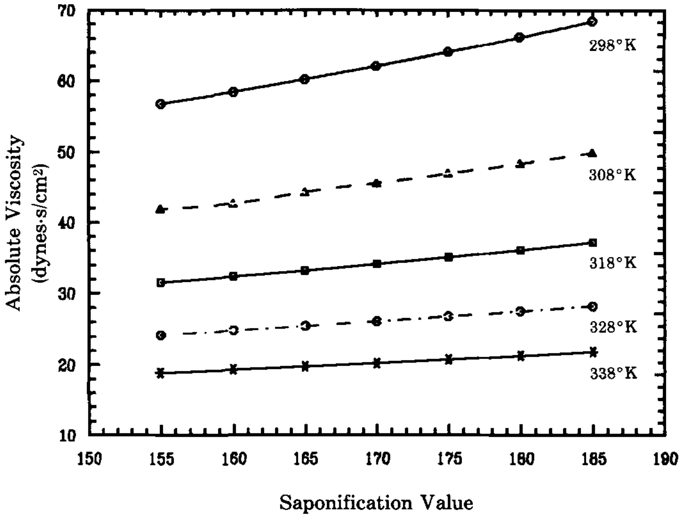

FIG. 4. Absolute viscosity as a function of saponification value and absolute temperature. The iodine value was kept constant to a value of 90. The viscosity was calculated with the equation obtained through the multiple-regressional analysis.

increase in viscosity. These families of acylglycerides, likely constituted by the shortest fatty acids present in the oils, should be in higher concentrations in oils with the highest SV (*i.e.*, oils with the acylglycerides with the lowest average of fatty acid chainlength).

From the statistical point of view, the analysis of the contribution of each variable to the prediction of Ln (μ) was achieved through the determination of the so-called partial and semipartial correlation coefficients (Table 3). These coefficients were accessible in the options menu for the multiple-regressional procedure in the STATISTICA program (10). The basis of this analysis is briefly described. The multiple-regression analysis develops an equation based on a decomposition of each observation into two parts: the predicted portion, which is assigned to explanatory variables [*e.g.*, $1/T°$, $SV^2/(T°)^2$, and IV^2] that significantly affect the response variable [*e.g.*, Ln (μ); the other portion, the residual portion, is assumed to be due to random variability (9), and its magnitude results from the cumulative differences between the experimental observation and the predicted value. The particular effect of a given explanatory variable on the response variable is given by regressional coefficients (*i.e.*, slopes), which might be expressed in several forms with different statistical/physical-chemical meaning, depending on the type of variation in the response variable fitted by the coefficient. Thus, the magnitude of the partial regression coefficient represents the unique contribution of the respective independent variable to the prediction of Ln (μ). The squared partial correlation coefficient is an indicator of the percent of residual variability (*e.g.*, variability obtained with the equation without the respective explanatory variable) uniquely accounted by the effect of the respective explanatory variable (*e.g.*, the explanatory variable adjusted for its correlations with all other independent variables present in the equation). Thus, $1/T°$ had the largest unique contribution to the prediction of Ln (μ), accounting for >82% of the residual variability. The variables $SV^2/(T°)^2$ and IV^2 accounted for ≈11% and ≈2.6% of the residual variability, respectively (Table 3).

TABLE 3

Partial and Semipartial Regression Coefficients of the Variables Present in Equation 2[a]

Independent variable	P.C.[b]	(P.C.)^2	S.P.C.[c]	(S.P.C.)^2
$1/T°$	0.9069	0.8225	0.5868	0.3443
$SV^2/(T°)^2$	0.3334	0.1112	0.0964	0.0093
IV^2	−0.1616	0.0261	−0.0446	0.0020

[a] Squared coefficients are also shown. Abbreviations as in Table 1.
[b] Partial correlation coefficient.
[c] Semipartial correlation coefficient.

On the other hand, the squared semipartial correlation coefficient is an indicator of the percent of total variability uniquely accounted for by the effect of the respective explanatory variable (*e.g.*, the explanatory variable, adjusted for its correlations with all other explanatory variables present in the equation). Accordingly, $1/T°$ accounted for ≈34% of the total variability in Ln (μ), and $SV^2/(T°)^2$ and IV^2 accounted for ≈0.9 and ≈0.2% of the total variability, respectively (Table 3). The rest of the variation (≈64%) was accounted for by the magnitude of the intercept (−4.7965) in the equation.

Within this framework and in terms of practical significance, the absolute temperature by itself explained most of the variability observed in μ among the oils investigated. However, additional considerations must be made. If the magnitude of the semipartial coefficient is small relative to the magnitude of the partial coefficient, the respective variable might be explaining a unique portion of variability in the response variable [*i.e.*, Ln (μ)] that was not accounted for by other explanatory variables in the equation (10). Therefore, the magnitude of the ratio partial to semipartial coefficients indicated the relevance of this factor. Thus, $SV^2/(T°)^2$ and IV^2 with ratios of 3.46 and 3.62, respectively, were considered important explanatory variables. Although these two variables represented, in comparison with $1/T°$, a small portion of the total and residual variability observed in Ln (μ), the absolute temperature had a partial-to-semipartial coefficient ratio of only 1.55. This result indicates that the absolute temperature was not capable of describing "information" associated with the description of the magnitude in μ, and that it was explained by variables associated with structural characteristics of the lipids and their interactions with temperature conditions [*e.g.*, $SV^2/(T°)^2$ and IV^2].

The equation obtained through the multiple-variable regression approach fitted the absolute viscosities of all 21 oil systems investigated. In comparison, the equations obtained based on the Andrade equation (8) also fitted the viscosities for each particular oil system. Although this last approach provides better predictive equations ($R^2 > 0.96$) than the multiple variable equation ($R^2 \approx 0.93$), the coefficients associated with the temperature effect do not have a particular physical-chemical meaning. In contrast, the multiple-variable regressional approach provided an excellent tool to better understand the quantitative structure–property relationships in lipid systems. Research is now in progress to use multiple-variable regression methodologies in the investigation of the events involved in the determination of oil viscosity at sub-zero temperatures, and its effect on acylglyceride crystallization.

ACKNOWLEDGMENT

The present work was supported by CONACYT.

REFERENCES

1. Steffe, J.F., and F.A. Osorio, *Food Technol. 41*:72 (1987).
2. Blumenthal, M.M., *Ibid. 45*:68 (1991).
3. Dutt, N.V.K., and D.H.L. Prasad, *J. Am. Oil Chem. Soc. 66*:701 (1989).
4. Noureddini, H., B.C. Teoh and L.D. Clements, *Ibid. 69*:1189 (1992).
5. Dutt, N.V.K., Y.V.L. Ravi Kumar and H. Sumathi Vedanayagam, *Ibid. 69*:1263 (1992).
6. Toro-Vazquez, J.F., and J.M. Regenstein, *J. Food Sci. 54*:1177 (1989).
7. Kato, A., and S. Nakai, *Biochem. Biophys. Acta 624*:13 (1980).
8. Andrade, E.N., *Nature 125*:309 (1930).
9. Allen, D.M., and F.B. Cady, *Analyzing Experimental Data by Regression,* Lifetime Learning Publications, Belmont, 1982.
10. STATISTICA User's Guide, Version 3.1, StatSoft, Tulsa, 1992.
11. *Official Methods of Analysis of the Association of Official Analytical Chemists,* 13th edn., Association of Official Analytical Chemists, Washington, D.C.
12. Toro-Vazquez, J.F., *J. Food Sci. 56*:1648 (1991).
13. Jackson, A.T., and J. Lamb, *Calculations in Food and Chemical Engineering,* The Macmillan Press Ltd., Hong Kong, 1981.
14. Hartel, R.W., in *Physical Chemistry of Foods,* edited by H.G. Schwartzberg, and R.W. Hartel, Marcel Dekker, New York, 1992, pp. 47–81.

[Received May 25, 1993; accepted August 15, 1993]