Rancid Flavor of Milk: Relationship of Acid Degree Value, Free Fatty Acids, and Sensory Perception

S.E. DUNCAN, G.L. CHRISTEN, and M.P. PENFIELD

ABSTRACT

Acid degree value (ADV) as determined by the standard method, free fatty acid profile, and sensory analyses by a trained panel (n = 5) using magnitude estimation were used as measures of intensity of rancid flavor in milk. The correlation between ADV and rancidity scores was 0.13 (p = 0.16). Laboratory-prepared rancid samples were assigned higher rancidity scores than farm-collected milk samples with similar ADV. Rancidity scores and ADV increased with storage time. Major free fatty acid concentrations increased as ADV increased (r = 0.93, p = 0.0001) for farm milk samples but correlation was low (r = 0.27, p = 0.40) for laboratory-prepared rancid samples.

INTRODUCTION

FREE FATTY ACIDS (FFA) are responsible for development of lipolytic flavor in milk (Walstra and Jenness, 1984). Lipoases, which may be inherent in milk or produced by psychrotrophic microorganisms, liberate FFA from glycerides. Increased concentration of FFA contributes a rancid, bitter, unpleasant taste to milk that is objectionable to many consumers. FFA of shorter chain length, primarily C4-C12, are important in characterizing rancid flavor but long-chain FFA are not associated with rancidity (Al-Shabibi et al., 1964; Scanlan et al., 1965).

Rancid flavor in milk is frequently determined by acid degree value (ADV) (Richardson, 1985). Senyk et al. (1985) suggested that raw milk with an ADV of 1.0 meq/100 g fat or greater may be rancid or near to becoming rancid and should be discarded. An ADV of 1.5 meq/100 g fat or greater was reported as “unsatisfactory (extremely lipolyzed)” (Richardson, 1985). Many researchers have found farm milk samples with ADV greater than 1.0 meq FFA/100 g fat that did not rancid to trained panelists (Earley and Hansen, 1982; Rerkrai et al., 1987); thus, the accuracy of ADV in predicting sensory detection of rancidity is questionable. Many investigations relating rancid flavor and ADV (Thomas et al., 1955) or identifying fatty acids responsible for rancid flavor (Al-Shabibi et al., 1964; Kintner and Day, 1965; Patton, 1964; Scanlan et al., 1965) have used laboratory-prepared rancid samples for evaluation.

ADV is frequently used to detect rancidity with no concurrent sensory evaluation of milk samples. If ADV is high, the milk sample may be discarded though it may not have a detectable rancid flavor. Therefore, it is important to find if ADV, as currently used, is an accurate predictor of rancid flavor development. The objectives of our research were to determine the effectiveness of ADV as a quality control method for detecting rancid flavor (as perceived by sensory evaluation using a trained panel) in fresh and stored milk and to determine which fatty acids contributed to the sensory detection of rancid flavor and to ADV.

MATERIALS & METHODS

Collection and processing of samples

Raw milk samples (2250 mL) were collected from bulk storage tanks of 13 East Tennessee farms and stored at 4°C until processed. Samples (n = 59) were homogenized (2700 psi) and pasteurized for 15 sec at 72°C in a tubular indirect heat exchanger consisting of preheating coil, pasteurizing coil and cooling coil (Wadsworth and Bas- scite, 1985). ADV was determined on raw milk samples within 24 hr of sample collection. Microbiological quality testing, Sharer photophase test, ADV (all according to Richardson, 1985) and FFA concentrations (Deeth et al., 1983) were determined for processed samples. Sensory evaluation was completed within 48 hr. Duplicate samples were stored at 4°C for 12 days after processing and then analyzed as before.

Preparation of rancid flavor standards

For laboratory-prepared rancid samples (LPRS), 35 mL raw milk were added to 750 mL commercial Vitamin D milk and incubated at 4°C for 24–36 hr. Milk was batch-pasteurized at 66°C for 3.5 min and cooled rapidly (Bandler and Barnard, 1981). Aliquots (35 mL) of raw milk were frozen at –20°C at onset of the study to standardize lipase source.

Sensory evaluation

Training and evaluation. Five female panelists experienced in identifying common milk off-flavors (oxidized, cooked, feed, rancid) in laboratory-prepared samples (Bandler and Barnard, 1984) were trained to detect low levels of rancid flavor using LPRS diluted to different intensities with fresh milk. Panelists’ ability to differentiate among LPRS of varying intensity [0 (fresh milk), 25, 65, 85, and 100% rancid milk] using magnitude estimation was determined with ANOVA (SAS, 1985) and replicated for statistical evaluation of panelist performance. Samples were presented as for data collection. Results showed that panelists could replicate performance and differentiate among samples. Panel performance was again tested after 33 samples with similar results.

Sample evaluation. Thirty mL milk samples were presented in random order in amber glass bottles and coded with randomly chosen 3-digit numbers. Panelists tasted samples directly (14°C) from the bottle or used plastic cups, selecting their portion size from the 30-ML sample. Panelists waited 30 sec between samples and rinsed with water between samples. Panelists evaluated each sample before proceeding to the next. Evaluations were completed in individual booths. Testing was completed on 3-6 farm samples per session. A labeled LPRS was presented first so the panelist could orient the palate to the rancid taste. The samples for evaluation then were presented to each panelist in random order. Two randomly coded LPRS were included within the sample set for statistical evaluation of panelist performance.

Panelists developed their scales (free modulus), assigning values to each sample to reflect rancid flavor intensity in relation to the first coded sample. Zero represented absence of rancid flavor. Numerical values also were assigned to concepts for intensity levels of unpalatable, very rancid, moderately rancid, and slightly rancid by each panelist using the same scale they used for the samples (Moskowitz, 1977).

Data transformation. Magnitude scores were normalized so comparisons among panelists could be made. A pivot number was determined by averaging the values given the four concept scores for each panelist for each calibrating scale. The pivot number for each panelist was divided into each sample magnitude estimate score and concept score, yielding a normalized sample and concept score for each panelist (Moskowitz, 1977). A constant 1 was added to all scores prior to logarithmic transformation to avoid any 0 scores (Butler et al., 1988).

Authors Christen and Penfield are with the Dept. of Food Technology & Science, The Univ. of Tennessee, P.O. Box 1071, Knoxville, TN 37901-1071. Author Duncan’s Current address: Dept. of Food Science and Technology, Virginia Polytechnic Institute & State Univ., Blacksburg, VA 24061.
Table 1—Mean log rancidity scores for fresh and stored milk samples in 5 ADV groups and for laboratory-prepared rancid samples (LPRS)

<table>
<thead>
<tr>
<th>ADV group</th>
<th>ADV range</th>
<th>Fresh</th>
<th>Stored</th>
<th>N</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt; 1.26</td>
<td>0.12</td>
<td>23</td>
<td>0.14</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1.26-2.0</td>
<td>0.13</td>
<td>11</td>
<td>0.13</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>2.01-2.3</td>
<td>0.18</td>
<td>14</td>
<td>0.18</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>2.31-2.75</td>
<td>0.18</td>
<td>2</td>
<td>0.14</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>&gt; 2.76</td>
<td>0.14</td>
<td>6</td>
<td>0.14</td>
<td>23</td>
</tr>
<tr>
<td>LPRS</td>
<td>2.61a</td>
<td>0.24</td>
<td>30</td>
<td>0.24</td>
<td>30</td>
</tr>
</tbody>
</table>

* Scores among ADV groups within fresh and stored samples did not differ (p > 0.05). Mean of 30 samples; range of ADV for LPRS was 1.46-3.62 meq FFA/100 g fat with one sample ADV at 7.87.

Logarithmic (base 10) transformation of the normalized scores was completed to normalize the distribution since a small sample size (<200 samples) was evaluated (Butler et al., 1988). Means for log transformed scores across all panelists (log rancidity score) for each sample were used for statistical analyses.

Gas chromatographic method

FFA were extracted from 22 fresh milk samples and the corresponding stored sample with cold HCl/methanol containing C2:0 internal standard (Duncan, 1983). The sample was centrifuged 5,000 rpm for 10 min at 0°C. Isolation was completed by extracting FFA from the lipid onto a small alumina column (12 cm x 5 mm i.d.) and removing triglycerides with hexane-dichethyl ether. Adsorbed FFA were dried on alumina with vacuum and stored under nitrogen in a capped glass centrifuge tube at 4°C overnight. FFA were removed from alumina using diisopropyl ether containing 6% formic acid. The sample then was centrifuged for 3 min at 13,600 x g in a microcentrifuge (Fisher Scientific, Pittsburgh, PA).

A 4 µL diisopropyl ether sample containing FFA was injected directly into a Shimadzu Mini GC-2 gas chromatograph (Sakakibara Ltd., Kyoto, Japan) equipped with 1 m x 3 mm glass column packed with 10% SP-216-PS on 100/120 Supelcoport (Supelco Inc., Bellefonte, PA), a Shimadzu Chromatopak E-1A data printer, an OmniScribe Recorder (Houston Instruments, Austin, TX), and a flame ionization detector, providing resolution of major short-, medium-, and long-chain FFA. The initial temperature was 110°C programmed to 195°C at 8°C/min with 25 min duration. Carrier gas was nitrogen with a flow rate 55 ml/min. Injector/detector temperatures were 230°C.

Concentrations were calculated relative to the area of the C5 internal standard, assuming complete recovery of the internal standard. Tentative identification of the fatty acids present in milk samples was based on retention times of fatty acid standards injected into the gas chromatograph under the same conditions (Duncan, 1989). Relative molar response (RMR) of each fatty acid was determined with respect to the internal standard (Karger et al., 1973).

Statistical analyses

General Linear Model (GLM) and ANOVA (SAS, 1985) were applied to mean log rancidity score data. Farm milk samples (n=59) were placed into 5 groups depending on ADV at the time of sensory evaluation (Table 1). Group means were separated with the Student Newman Keuls (SNK) multiple range test when appropriate. Paired t-tests were used to evaluate individual panelist performance on LPRS through the study, to compare mean log rancidity scores and ADV of fresh and stored samples, and to compare ADV of raw and fresh samples. A correlation was determined for rancidity scores and ADV. Significance was pre-established at α = 0.05.

Correlations between ADV and concentration of 10 major fatty acids (C2:0, C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:2, C18:3) and between ADV and the sum of concentrations of shorter chain fatty acids (C2:0-C8:0) and between ADV and the total concentration of palmitic, stearic, oleic, and linoleic acids. Uncorrected (C2:0-C8:0) and between ADV and total fatty acids (C2:0-C20:4) were obtained for 44 farm milk samples and 12 LPRS. Correlations also were calculated for log rancidity scores and the concentration of fatty acid. General Linear Model (SAS, 1985) was used to analyze individual fatty acid concentrations data. Least-squares means (LS Means) for FFA were calculated and differences (p < 0.05) noted.

RESULTS & DISCUSSION

ADV and sensory scores

There were significant differences (p = 0.0001) in ADV between paired raw farm milk samples and freshly pasteurized samples and between fresh and stored farm milk samples indicating FFA concentration had increased, probably due to lipase activity (Table 2). The difference between log rancidity scores for paired fresh and stored samples was small (Table 2) but significantly different (p = 0.0012). Scores were interpreted as "slightly rancid" for both fresh and stored samples.

Means of log rancidity scores for farm milk samples were low across all ADV groups (Table 1). Flavor descriptors may be assigned to the mean scores from the calibration scale concerning concept scores included in Table 2. There were no significant differences among mean log rancidity scores for either fresh or stored farm samples though the ADV range was very broad. Samples in each ADV group would be described as "slightly rancid" (0.10) or "moderately rancid" (0.18). The LPRS would be described as "moderately rancid" (Table 1).

Figure 1 illustrates the random scattering of log rancidity scores for samples (fresh, stored and LPRS) within ADV group. Fresh samples clustered in the region of low ADV (1, 2, and 3) with slight to moderate log rancidity scores. Most stored farm samples were in the higher ADV groups with log rancidity scores ranging from slight to moderate. The correlation between ADV and log rancidity scores for fresh and stored samples was low (r = 0.22, p = 0.09; r = -0.12, p = 0.35, respectively). In contrast, the correlation between ADV and log rancidity scores for LPRS was significant (r = 0.17, p = 0.0023). The LPRS has a more intense rancid flavor ("moderately rancid" to "unpalatable") than found in farm milk samples with similar ADV. One farm sample had an ADV of 0.76, a "borderline (indefinite)" rancid flavor (Richardson, 1985) but had a log rancidity score of 0.22, a "moderately rancid" score, based on evaluation by the trained panel. Some samples with ADV as high as 3.54 meq/100 g fat had "not at all rancid" to "slightly rancid" flavor yet would be classified as "insatisfactory (extremely lipolyzed)" if rancid flavor intensity were determined based on standard methods interpretation of ADV (Richardson, 1985). Table 3 provides examples of individual samples with similar ADV but very different mean log flavor scores.

A panel that has undergone extensive training and performance evaluation provides more reliable perception of rancid flavor detection and intensity than an inexperienced panel and is better suited to the task of rancid flavor evaluation due to the many flavors that may occur in milk (Connolly et al., 1979). The methods we used for training yielded a panel able to provide reliable and accurate scores for rancid flavor (Table 4). Still, correlations between ADV and log rancidity scores for LPRS for most panelists were low and not significant (Table 4).

Fatty acid concentrations, acid degree values, and sensory scores

The relationship between ADV and fatty acid concentration was weak (r = 0.33, p < 0.0001) for all individual and groups
samples and correlation for each panelist between ADV rancidity score for laboratory-prepared rancid sample and FFAs equally (Duncan and Christen, 1990). Short-chain FFAs classified into ADV groups 4 or 5, and significant differences in mean individual fatty acid concentrations were found for only C16:0 and C18:1. ADV does not extract and measure all FFAs equally (Duncan and Christen, 1990). Correlations for C4 and C14-C16 were similar to those for total FFA (C4-C22) and between ADV and individual fatty acid with significant correlation was C6 (r = 0.59, p = 0.06). This implies that ADV increased as concentration of most fatty acids increased. The correlations between fatty acid concentration and ADV were more variable for LPRS. The group of shorter chain fatty acids (C10-C14) had a positive moderate correlation (r = 0.53, p = 0.07) but the correlations between ADV and longer chain fatty acids (C15-C18) were not significant. The only individual fatty acid with significant correlation was C8 (r = 0.59, p = 0.04); the correlation for C9 was approaching significance (r = 0.55, p = 0.06). Ikinci et al. (1988) reported high positive correlations of 0.74 and 0.75 between concentration and ADV for total FFA (C4-C22) and longer chain fatty acids (C15-C18) in cheese. Correlations for C10 and C12-C14 were similar to those found for LPRS in our study.

The concentrations of C6, C8, C10, C12, C13, C15, C16, and C17 in farm samples increased with ADV (Table 5). LPRS were classified into ADV groups 4 or 5, and significant differences in mean individual fatty acid concentrations were found for only C16:0 and C18:1. ADV does not extract and measure all FFAs equally (Duncan and Christen, 1990). Short-chain FFAs (C4-C6) are partly distributed in the aqueous phase and are not completely quantified by ADV, but, these may contribute to rancid flavor. ADV does not measure and identify specific fatty acids, making it impossible to find which fatty acids are contributing to a high titration value. Correlations between log rancidity score and individual fatty acids or groups of fatty acids were not significant (data not presented). The mean concentrations (LS means) of total fatty acids for farm samples and laboratory-prepared rancid samples (LPRS) in each ADV group increased with ADV group (Table 5). LPRS were found for LPRS in our study.

Table 3—Examples of individual farm samples with similar ADV but different log rancidity scores in each ADV group.

Table 4—Paired t-tests for individual panelist performance on duplicate samples and correlation (r) for each panelist between ADV and log rancidity score for laboratory-prepared rancid samples.

Table 5—Least-squares mean concentrations for fatty acids for farm samples and laboratory-prepared rancid samples (LPRS) in each ADV group.

Table 6—Least-squares mean concentrations of total fatty acids from farm and laboratory-prepared rancid samples with different intensities of rancidity.
flavor intensity provided no farm samples with "very rancid" or "unpalatable" scores nor any LPRS with "slightly rancid" scores.

The importance of shorter chain fatty acids to rancid flavor perception has been shown in LPRS (Al-Shabibi et al., 1964; Patton, 1964; Scanlan et al., 1965). The distribution of fatty acids and pH of the system in which they are found affects perception and recovery of fatty acids (Bills et al., 1963). A low pH, between 1.3-3.0, is needed to recover short-chain (C<sub>2</sub>-C<sub>5</sub>) fatty acids (Kintner and Day, 1965). As pH decreases, more fatty acid enters the fat phase (Walstra and Jensen, 1984). Chemical methods that change the pH of the system may provide greater recovery of the short-chain fatty acids and have a stronger correlation to rancid flavor than the ADV method.

The relationship between concentration of fatty acids in LPRS to ADV was different from that found for farm samples. The differences observed in individual fatty acid concentrations with increasing ADV may have been caused by different lipase sources. The hydrolysis of fatty acids from glycerides in LPRS was primarily caused by milk lipase from the raw milk added to the homogenized milk. Milk lipase is a nonspecific lipase, releasing various fatty acids in nearly the same proportion that the acids are present in the intact fat (Nelson, 1972). The high correlations between individual fatty acid concentrations and ADV for farm samples may be relatively to the activity of lipases from psychrotrophic bacteria. These bacteria (i.e., *Pseudomonas* spp.) have opportunity to proliferate during bulk storage of the milk. Many of these organisms produce heat-resistant lipases with varying specificities. ADV and intensity of rancid flavor increased with storage, suggesting that FFA concentration increased over time, probably due to heat-resistant lipase from psychrotrophic microorganisms.

The relationship between concentration of fatty acids and flavor was difficult to explain. In farm samples, a numerical increase in concentration of shorter chain fatty acids as ADV increased was observed (Table 5) although rancid flavor score did not increase as ADV increased (r = 0.13, p = 0.16). Possibly varying concentrations of individual long-chain fatty acids overpowered the flavor strength of the short-chain fatty acids. That is, as concentration of long-chain fatty acids increased (and ADV increased), a greater concentration of short-chain fatty acids may have been needed to impart rancid flavor. The relationship between fatty acid concentration and rancidity scores observed for LPRS and farm samples was similar both in direction of change and magnitude of change. In contrast, the concentration of most major fatty acids in farm samples highly correlated with ADV while in LPRS only the concentration of short-chain fatty acids slightly correlated with ADV. This relationship may have contributed to the more intense rancid flavor found in these samples (and the significant correlation between ADV and log rancidity score) as compared to farm samples. These observations supported the conclusions of early studies (Al-Shabibi et al., Kintner and Day, 1965; Patton, 1964; Scanlan et al., 1965; Thomas et al., 1955) that used LPRS for ADV and rancid flavor evaluation. More work is needed to determine the relationship among fatty acids with respect to rancid flavor and to determine the differences in flavor perception between LPRS and farm samples. ADV, as a measure of rancid flavor intensity, has limitations although it indicates changes in total FFA concentrations. Caution should be applied when interpreting ADV in relation to detection of rancid flavor. Sensory evaluation by trained panelists should be completed to decide if the milk flavor quality is acceptable.

**REFERENCES**


Ms received 4/6/90; revised 9/12/90; accepted 10/2/90.