The use of clustering techniques in the elucidation or confirmation of metabolic pathways

Application to the branched-chain fatty acids present in the milk fat of lactating goats

Anne-Marie MASSART-LEËN*† and Désiré L. MASSART‡

*Department of Physiology, Veterinary Faculty of the University of Gent, Casinoplein 24, B-9000 Gent, Belgium, and ‡Pharmaceutical Institute, University of Brussels, Laarbeeklaan 103, B-1090 Brussels, Belgium

(Received 5 January 1980/Accepted 8 January 1981)

The aim of this paper is 2-fold. (1) To propose the use of a group of mathematical techniques, called clustering, in the elucidation of complex metabolic relationships. (2) To apply clustering for the identification of related groups of saturated fatty acids having a common metabolic pathway for their biosynthesis in the milk fat of lactating goats. In this way, four groups of branched-chain fatty acids and two groups of straight-chain fatty acids are identified; the odd-numbered iso-, the even-numbered iso-, the anteiso-acids and the branched-chain fatty acids with methyl substitution in the chain, the odd-numbered straight-chain and the even-numbered straight-chain fatty acids. The long-chain fatty acids are not part of any group. The different metabolic pathways for their biosynthesis are discussed. From the results, it is concluded that clustering is indeed a potentially useful tool in the study of complex metabolic relationships.

The branched-chain fatty acids present in milk fat of lactating goats were identified and quantified (Massart-Leën *et al.*, 1981). To study the causes and the extent of the variation in the composition of those branched-chain fatty acids, milk samples from 10 goats were taken on three successive weeks and subjected to analysis. One of the principal objects of the present work was to investigate the relationships between the several fatty acids. We wanted to know which acids behave in the same way, for example, which are the acids of which the content increases or decreases together from one sample to another and, more fundamentally, which acids follow the same biosynthetic pathway. This is a fairly simple matter when there are only a few substances to be compared. One then plots the content in different samples of the substances two by two and calculates a correlation coefficient. If the correlation is significant, one concludes there is a relationship. However, in the problems investigated by us, one wants to study the relationships between 14 branched-chain fatty acids and the relationships of these acids with the straight-chain acids. Of course, a correlation coefficient matrix can be computed (Gray, 1973), but it is quite another matter to

distinguish groups of related fatty acids in such a matrix. In the present paper, we report mathematical methods for carrying out such grouping or classification procedures and apply them to the study of branched-chain fatty acids.

A clustering procedure is used. Clustering has been applied in very diverse fields of science such as biological or microbiological taxonomy, the selection of analytical procedures, mathematical diagnosis of illness, geochemistry, the study of biological activity of drugs, phylogenetic studies of proteins etc., but not, as far as we know, in metabolic studies (Anderberg, 1973; Sneath & Sokal, 1973; Kowalski, 1975; Massart *et al.*, 1978). Because in our opinion clustering may be of general use in the study of complex metabolic relationships, it is discussed in some detail in the present paper.

Results

Application of the model to branched-chain fatty acids

Data. The data consist of the percentage distribution of branched-chain fatty acids present in the milk fat of 10 goats (Massart-Leën *et al.*, 1981). The data can also be presented in terms of mol% (because of the way in which scaling is carried out and the distances

† To whom reprint requests should be sent.

compared, this has no effect on the eventually obtained clustering). These branched-chain fatty acids represent about 2% of the milk fat. The 14 branched-chain fatty acids found in milk fat are numbered from 1 to 14 and their respective normalized value, equivalent chain length and the fatty acid identified by mass spectrometry are shown in Table 1. Since 14 branched-chain fatty acids are investigated, the original data constitute a matrix $x_{i,j}$ ($i = 1, 14; j = 1, 10$). Each entry $x_{i,j}$ in this matrix is the mean of three determinations on successive weeks.

Scaling of the data. Because the contents of the fatty acids are very different, it is necessary to scale the data. If the original data were used as such, fatty acid no. 13 (equivalent chain length 16.70), which comprises about 20% of the branched-chain fatty acid fraction, would have a much higher effect on the classification than fatty acid no. 4 (equivalent chain length 12.55), with a concentration of only about 0.5% (Table 1). To avoid mathematical artefacts resulting from these differences, we carried out a scaling procedure. For each fatty acid i , we computed the mean $\bar{x}_i$ and the standard deviation, s_i , of the results $x_{i,j}$ ($i = 1, 14; j = 1, 10$) in the ten samples. We then computed $z_{ij} = (x_{i,j} - \bar{x}_i)/s_i$ and used the $z_{i,j}$ values in further calculations. Instead of the raw data, we used results compared with the mean and expressed in standard-deviation units. Samples with a higher content than the mean will bear a positive sign. In Table 2, the standardized data are shown.

Dimension of the model. In Fig. 1, the standardized data (Table 2) for goats 9 and 10 are shown. The points marked with numbers 1–14 are related to the branched-chain fatty acids summarized in Table 1. Fatty acids that are found close to each other in such a Figure behave in the same way in the two samples and are said to form a cluster. For example, acids with numbers 1, 6, 7 and 11 (respective equivalent-chain-length values 10.40, 14.31, 14.40, 16.39; see Table 1) form a cluster, because they all have very low contents in goat 9 compared with the mean for all the goats, and somewhat lower contents than the mean for goat 10. If all the goats are used, a ten-dimensional space results and formal procedures are necessary to discern the clusters. This is discussed further below.

Distance matrix. The clustering procedure is initiated by measuring the distance between all pairs of fatty acids. The smaller this distance, the more the behaviour of the acids is the same in the 10 goats investigated. The distances between acids 1 and 7, 1 and 6, 1 and 11, 6 and 7, 7 and 11 and 6 and 11 are much lower than, for instance, the distance between acids 3 and 8 (Fig. 1). Therefore, one concludes that acids 1, 6, 7 and 11 behave in a more similar

Fig. 1. Graph of the standardized values of the branched-chain fatty acids for two goats

The standardized values $z_{i,j}$ ($i = 1, 14$) for goats 9 and 10 ($j = 9, 10$), as given in Table 2, are shown. The numbers in the Figure identify the branched-chain fatty acids as defined in Table 1.

manner than acids 3 and 8. One can show that the distance (Δ) between acids 1 and 6 is given by:

$$\Delta_{1,6} = \sqrt{(z_{1,9} - z_{6,9})^2 + (z_{1,10} - z_{6,10})^2}$$

This distance is called the Euclidian distance. When all 10 goats are considered, i.e. in the 10-dimensional case, the distance is given by:

$$\Delta_{i,k} = \sqrt{\sum_{j=1}^{10} (z_{i,j} - z_{k,j})^2 / 10}$$

where $\Delta_{i,k}$ is the distance between fatty acids i and k . The division factor 10 is used, when some values are missing in the original data Table. If, for example, $x_{7,4}$ was not determined, then all distances involving fatty acid 7 would have been measured for nine goats and $\Delta_{i,7}$ would be computed by a summation of only nine terms and divided by 9.

These distances are summarized in a distance matrix such as the one given in Table 3. It is clear that, for example, the distance (Δ) between the branched fatty acid 7 (equivalent chain length 14.40) and the acids with equivalent chain lengths 10.40 (Δ 1.31), 12.40 (Δ 1.99), 14.31 (Δ 1.17), 14.40 (Δ 0) and 16.39 (Δ 1.46) are small compared with the distances with acids of equivalent chain lengths 13.55 (Δ 5.49), 14.55 (Δ 5.09), 14.71 (Δ 5.81) etc. and it is assumed that fatty acids with small

Table 1. *Distribution of the branched-chain fatty acids as a percentage of the total branched-chain fatty acids in milk fat of goats*
The branched-chain fatty acids were analysed by capillary gas chromatography-mass spectrometry. The mass spectra and the equivalent chain length are used to identify the acids. The data are for ten goats and are means of samples taken on three successive weeks. Abbreviation used: ECL, equivalent chain length.

Arbitrary fatty acid no.	ECL	Identity of methyl ester	Goat no. ...	Distribution (%)									
				1	2	3	4	5	6	7	8	9	10
1	10.40	Methyl 4-methyldecanoate	6.55	8.57	6.58	6.48	8.58	5.14	5.34	4.51	1.88	4.27	
2	11.20	Methyl 6-methyldecanoate	0.93	0.54	0.77	0.56	3.43	2.35	2.52	2.21	1.16	0.99	
		Methyl 4,8-dimethyldecanoate											
		Methyl 6,8-dimethyldecanoate											
3	12.40	Methyl 4-methyldodecanoate	2.82	3.29	2.62	2.85	3.71	2.53	2.34	2.48	0.88	1.69	
4	12.55	Methyl 11-methyldodecanoate (iso)	0.43	0.62	0.84	0.63	0.89	0.50	0.50	0.53	0.80	1.05	
5	13.55	Methyl 12-methyltridecanoate (iso)	2.69	2.93	3.24	2.70	2.58	3.02	5.28	3.69	3.82	3.16	
6	14.31	Methyl 6-methyltetradecanoate	4.01	4.53	2.85	2.93	3.39	2.23	1.84	1.59	0.83	2.44	
7	14.40	Methyl 4-methyltetradecanoate	3.87	3.71	2.77	3.25	3.86	2.86	2.26	2.48	1.22	2.63	
8	14.55	Methyl 13-methyltetradecanoate (iso)	8.06	6.64	8.38	7.59	8.45	9.54	9.02	6.31	9.71	10.30	
9	14.71	Methyl 12-methyltetradecanoate (anteiso)	15.44	14.12	18.45	15.40	14.11	17.35	17.55	21.32	21.20	18.41	
10	15.55	Methyl 14-methylpentadecanoate (iso)	10.53	13.75	11.30	14.03	9.37	10.32	11.25	13.06	14.90	9.29	
11	16.39	Methyl 4-methylhexadecanoate	3.84	3.71	2.10	2.14	2.73	2.13	1.92	1.86	1.08	2.20	
12	16.55	Methyl 15-methylhexadecanoate (iso)	16.82	17.12	16.75	17.36	16.41	19.22	17.50	13.03	18.61	19.14	
13	16.70	Methyl 14-methylhexadecanoate (anteiso)	17.07	15.31	20.02	20.20	19.27	20.01	19.68	24.46	21.57	22.34	
14	17.52	Methyl 16-methylheptadecanoate (iso)	1.27	0.98	1.13	1.51	1.29	1.49	1.76	1.23	1.65	1.16	

Table 2. *Standardized ($z_{i,j}$) values of the data given in Table 1.*

The mean $\bar{x}_i$ and standard deviation s_i of the data $x_{i,j}$ (fatty acid $i = 1, 14$; goat number $j = 1, 10$) in Table 1 were calculated. The standardized value $z_{i,j} = (x_{i,j} - \bar{x}_i)/s_i$ was then obtained. Abbreviation used: ECL, equivalent chain length.

Arbitrary fatty acid no.	ECL	Goat no. ...	$z_{i,j}$									
			1	2	3	4	5	6	7	8	9	10
1	10.40		0.37	1.37	0.39	0.34	1.37	-0.32	-0.22	-0.63	-1.93	-0.75
2	11.20		-0.62	-1.00	-0.78	-0.98	1.88	0.80	0.97	0.66	-0.39	-0.57
3	12.40		0.42	0.91	0.21	0.45	1.35	0.11	-0.08	0.06	-1.61	-1.81
4	12.55		-1.22	-0.29	0.79	-0.24	1.03	-0.87	-0.87	-0.73	0.59	1.81
5	13.55		-0.77	-0.47	-0.09	-0.76	-0.91	-0.36	2.45	0.47	0.63	-0.19
6	14.31		1.20	1.66	0.17	0.24	0.65	-0.39	-0.73	-0.96	-1.63	-0.20
7	14.40		1.18	0.99	-0.15	0.43	1.17	-0.04	-0.76	-0.50	-2.02	-0.32
8	14.55		-0.26	-1.35	-0.01	-0.62	0.04	0.88	0.48	-1.61	1.01	1.46
9	14.71		-0.73	-1.23	0.43	-0.74	-1.23	0.01	0.08	1.53	1.48	0.41
10	15.55		-0.66	0.98	-0.23	1.12	-1.19	-0.72	-0.26	0.64	1.55	-1.23
11	16.39		1.74	1.58	-0.32	-0.27	0.42	-0.28	-0.53	-0.60	-1.53	-0.20
12	16.55		-0.21	-0.04	-0.25	0.09	-0.44	1.14	0.17	-2.35	0.80	1.10
13	16.70		-1.14	-1.83	0.01	0.08	-0.28	0.01	-0.12	1.74	0.61	0.91
14	17.52		-0.31	-1.49	-0.88	0.66	-0.23	0.58	1.67	-0.47	1.23	-0.76

Fig. 2. Graph of hypothetical values of a two-dimensional clustering problem

The $z_{m,u}$ and $z_{m,v}$ values are the normalized concentrations of acids $m = A-I$ and O in the milk fat of goats u and v . Acids A, I, B and C, D, G and E and H and F form clusters and O is an outlier.

distances have the same behaviour. These results can be grouped by using the clustering technique.

The clustering procedure. To understand the clustering procedure, we need to return to the two-dimensional case and suppose that the nor-

malized results of Fig. 2 were obtained. Clearly, the behaviour of A, B, C and I, D, E and G , and H and F are similar. One therefore says that A, B, C and I 'form a cluster'. This is also the case for D, E and G and for H and F . On the other hand, O behaves in a special way and is called an 'outlier'. In fact, one has created three groups in this way. In the language of clustering, one says that three clusters have been 'created'. In operational-research language, one says that a $p = 3$ median has been computed. If O is also taken into account, a $p = 4$ median is obtained. One can extend this argument until each component forms its own cluster. When different compounds belong to the same cluster at different p values, one concludes that these compounds are highly related to each other. A cluster consisting of one point only, appearing at low p values, is considered to be an outlier. In a 10-dimensional space (and even in a realistic two-dimensional case such as in Fig. 1), it is of course not so easy to select these clusters and outliers. Therefore, a mathematical method is necessary. The one reported in the present paper was developed by Massart & Kaufman (1975) and is part of a family of methods called centroid clustering methods. The basis of the method used in this article is an operational-research problem, called the location problem.

How to carry out the computations in practice has been described previously (Massart & Kaufman, 1975; De Clercq *et al.*, 1976). The clustering model can be put into practical use by a computer program termed MASLOC.

Table 4 shows the clusters formed for $p = 2-14$. The numbers 1, 3, 6, 7 etc. refer to the fatty acids

Table 3. Distance matrix of all pairs of branched-chain fatty acids

From the results summarized in Table 2, the Euclidian distances between all pairs of branched-chain fatty acids were computed using the equation:

$$\Delta_{i,k} = \sqrt{\sum_{j=1}^{10} (z_{i,j} - z_{k,j})^2 / 10}$$

$\Delta_{i,k}$ is the distance between the fatty acids i and k . For other symbols see Table 2. When the distance between two acids is small, this is interpreted as an indication of analogous behaviour. Abbreviations used: ECL, equivalent chain length.

Arbitrary fatty acid no.	Arbitrary fatty acid no.	...	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	ECL		10.40	11.20	12.40	12.55	13.55	14.31	14.40	14.55	14.71	15.55	16.39	16.55	16.70	17.52
1	10.40															
2	11.20		0													
3	12.40		4.09	0												
4	12.55		1.47	3.62	0											
5	13.55		4.42	4.40	5.02	0										
6	14.31		5.15	3.65	4.95	4.66	0									
7	14.40		1.45	4.79	2.45	4.43	5.40	0								
8	14.55		1.31	4.24	1.99	4.57	5.49	1.17	0							
9	14.71		5.20	3.97	5.45	3.13	3.81	5.05	5.09	0						
10	15.55		5.81	4.22	5.54	4.02	2.98	5.80	5.81	3.76	0					
11	16.39		4.78	5.11	4.43	4.70	3.89	4.72	5.00	5.07	3.59	0				
12	16.55		2.06	4.64	2.69	4.78	5.22	1.01	1.46	5.05	5.65	4.79	0			
13	16.70		4.71	4.68	5.10	3.57	4.30	4.35	4.63	1.84	4.55	4.62	4.47	0		
14	17.52		5.52	3.69	5.29	3.61	3.55	5.71	5.43	3.79	1.84	4.15	5.68	4.72	0	
			5.22	3.46	4.77	4.80	2.72	5.44	5.29	3.20	3.63	3.83	3.56	3.54	3.68	0

Table 4. *Sequence of cluster formation of the branched-chain fatty acids present in the milk fat of goats*

The distance matrix of the 14 branched-chain fatty acids, presented in Table 3, were partitioned into two to 14 clusters ($p = 2-14$) of fatty acids. The numbers refer to the fatty acids as identified in Table 1. For instance, if 3 clusters ($p = 3$) are required, the best clusters consist of respectively acids 1, 3, 6, 7 and 11, acids 10, 5, 9 and 13 and acids 2, 4, 8, 12 and 14. The significance of the clusters is derived in a manner explained by Massart *et al.* (1980). Clusters are indicated by parentheses. Significant clusters are italicized.

	Clusters														
p															
2	(1, 3, 6, 7, 10, 11)	(2, 4, 5, 8, 9, 12, 13, 14)													
3	(1, 3, 6, 7, 11)	(10, 5, 9, 13)	(2, 4, 8, 12, 14)												
4	(1, 3, 6, 7, 11)	(10, 5, 9, 13)	(2, 4, 8, 12, 14)												
5	(1, 3, 6, 7, 11)	(5, 9, 13)	(10)	(2, 4, 8, 12, 14)											
6	(1, 3, 6, 7, 11)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)										
7	(1, 3, 6, 7, 11)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)										
8	(1, 3, 6, 7, 11)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)										
9	(1, 6, 7, 11)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)									
10	(1, 6, 7, 11)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)									
11	(1, 6, 7, 11)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)									
12	(1, 6, 7, 11)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)									
13	(1, 6, 11)	(7)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)								
14	(1, 6, 11)	(7)	(3)	(9, 13)	(5, 14)	(10)	(2, 4, 8, 12)								

summarized in Table 1. The results of Table 4 are interpreted in the following way. 1. Substances that are found in groups consisting only of one substance at low p values are considered to be outliers. This is the case for fatty acid 2 (equivalent chain length 11.20), which is separated from all the other substances at $p = 4$, and for acid 10 (equivalent chain length 15.55), which is separated at $p = 5$. Equivalent chain length 11.20 is the only equivalent chain length value of $x.20$ among the 14 acids studied and it is therefore not surprising to find it is an outlier. This confirms its special character in the group of 14 fatty acids. The equivalent chain length 15.55 is discussed below. 2. Substances that are found together for a long time (i.e. until high p values are reached) are considered to form a cluster and therefore to have common characteristics. The following clusters are detected: acids 1 (equivalent chain length 10.40), 3 (12.40), 6 (14.31), 7 (14.40) and 11 (equivalent chain length 16.40), i.e. all the $x.40$ equivalent chain lengths; acids 8 (equivalent chain length 14.55), 12 (16.55), and to a lesser degree acid 4 (equivalent chain length 12.55), i.e. all the odd-numbered branched fatty acids with $x.55$ (the fact that acid 4 is somewhat separated from the two others can be explained by the very low concentration of this acid, which should lead to higher variance of the analytical results); acids 9 (equivalent chain length 14.72) and 13 (16.70), i.e. both the $x.70$ acids; acids 5 (equivalent chain length 13.55) and 14 (17.52), i.e. two of the three even-numbered $x.55$ equivalent chain lengths. The other acid is no. 10 (equivalent chain length 15.55), which is detected as an outlier. This is in fact the only acid that is not placed in a logical group.

The same procedure was carried out on a second data set concerning the same branched fatty acids except fatty acids 4, 11 and 14, but containing also the data on straight-chain fatty acids. The following clusters were found: equivalent chain lengths 10.00, 12.00 and 14.00, i.e. three straight-chain even-numbered fatty acids; equivalent chain lengths 11.00, 13.00 and 15.00, i.e. three straight-chain odd-numbered fatty acids; equivalent chain lengths 10.40, 12.40, 14.31 and 14.40, i.e. four branched-chain fatty acids that are not part of the iso- or anteiso-group; equivalent chain lengths 14.71, and 16.70, i.e. the two anteiso-branched-chain fatty acids; equivalent chain lengths 14.55 and 16.55, i.e. the two odd-numbered iso-branched-chain fatty acids; equivalent chain lengths 16.00 and 17.00, i.e. two straight-chain fatty acids. Of these clusters, the first to fall apart is the one consisting of equivalent chain lengths 16.00 and 17.00. Outliers are of equivalent chain lengths 11.20, 13.55, 18.00 and 15.55. Of these outliers those of equivalent chain lengths 11.20 and 15.55 were found to be outliers also in the clustering of branched-chain fatty acids, whereas the acid with equivalent chain length 13.55 is the only member of a cluster of two present in the first data set, which was included in the second data set.

Discussion

The clustering of the branched-chain fatty acids present in the milk fat of goats shows that these acids can be divided into several groups of related fatty acids. The group with equivalent chain lengths 10.40, 12.40, 14.31, 14.40 and 16.39 all have methyl

substitution in the chain as indicated by mass-spectrometric analysis (Massart-Leën *et al.*, 1981) (i.e. not in the iso- or anteiso-location) (Table 1). Since these acids form a cluster, it must be assumed that they follow a common metabolic pathway. It has been proposed that they may be formed by the incorporation of methylmalonyl-CoA instead of malonyl-CoA in the chain-lengthening process of fatty acids synthesized with acetyl-CoA as the primer unit. The incorporation of [^{14}C]methylmalonyl-CoA into branched fatty acids of intermediate chain length was shown in chicken liver (Cardinale *et al.*, 1970) and adipose tissue of sheep (Scaiffe & Garton, 1975; Scaiffe *et al.*, 1978) and ox (Wahle & Paterson, 1979). The absence of all these branched-chain fatty acids in the milk fat of cows, although they are all present in the milk fat of goats, confirms that they have a common metabolic route (Massart-Leën *et al.*, 1981). The observation of a cluster confirms this conclusion.

The other branched-chain fatty acids are divided into several groups, showing different metabolic pathways for their biosynthesis. Iso-acids are divided into two clusters, those with an even number of total C atoms [equivalent chain lengths 13.55 (acid 5) and 17.52 (acid 14)] (Table 4) and those with an odd-number of total C atoms [equivalent chain lengths 12.55 (acid 4), 14.55 (acid 8) and 16.55 (acid 12)]. It has been proposed that even- and odd-numbered iso-acids can be synthesized by oxidative deamination of the branched-chain amino acids leucine and valine respectively (El-Shazly, 1952; Lennarz, 1961). Only the fatty acid with equivalent chain length 15.55 (acid 10) was not included in the cluster to be expected from this reasoning. The anteiso-acids with equivalent chain lengths 14.71 (acid 9) and 16.70 (acid 13) form another small cluster. Both are anteiso-acids with an odd number of C atoms. It is predicted that isoleucine can form anteiso-acids by oxidative deamination and chain elongation. In summary, we conclude that, by using a clustering procedure, it is indeed possible to group branched-chain fatty acids having a common metabolic pathway for their biosynthesis. This is confirmed by the clustering of the second data set. Taking into account that three branched-chain fatty acids of the first data set were not included the same clusters regarding the branched-chain fatty acids were formed. Except for the cluster consisting of $\text{C}_{16:0}$ and $\text{C}_{17:0}$ acids, which is also the first to split up, and therefore not very significant, the odd-numbered fatty acids are separated from the even-numbered ones, as it should be, since the first have propionyl-CoA as their precursor and the others acetyl-CoA. It is noteworthy that $\text{C}_{10:0}$, $\text{C}_{12:0}$ and $\text{C}_{14:0}$ acids form a cluster separated from $\text{C}_{16:0}$ and from $\text{C}_{18:0}$ fatty acids. This agrees with the results of Gray (1973).

Studying seasonal variations in the composition of New Zealand butter fat, we also noted similar behaviour for $\text{C}_{10:0}$, $\text{C}_{12:0}$ and $\text{C}_{14:0}$ acids, different from $\text{C}_{16:0}$ and $\text{C}_{18:0}$ acids. Both findings once more confirm the general concept, reviewed by Patton & Jensen (1976), that the lower fatty acids ($\text{C}_{4:0}$ – $\text{C}_{14:0}$) and long-chain fatty acids ($\text{C}_{16:0}$ – $\text{C}_{18:0}$) of the milk fat of ruminants are derived from a different source. In short, acetate and β -hydroxybutyrate supply carbon for the synthesis of the ruminant milk fatty acids $\text{C}_{4:0}$ – $\text{C}_{14:0}$ (Popjak *et al.*, 1951; McCarthy & Smith, 1972; Smith *et al.*, 1974). Palmitic acid is partly synthesized in this way, whereas the other part comes from blood lipid precursor. Stearic acid of the milk fat originates completely from the bloodstream (Palmquist *et al.*, 1969; Bickerstaffe & Annison, 1974; Glascock & Welch, 1974; Peeters *et al.*, 1979). Therefore clustering appears to be a promising new tool in the study of complex metabolic relationships and probably can be applied in other fields of biochemistry.

We thank Professor G. Peeters for helpful discussion and reading the manuscript and Dr. D. Coomans for making available the computer programs used in the present study. We are grateful for the technical assistance of Mrs. Liliane De Neve-Van Lancker. This work was carried out with the aid of grants from the I.R.S.I.A. Foundation.

References

- Anderberg, M. R. (1973) in *Cluster Analysis for Applications*, Academic Press, London and New York
- Bickerstaffe, R. & Annison, E. F. (1974) *J. Agric. Sci.* **82**, 71–85
- Cardinale, G. J., Carthy, T. J. & Abeles, R. H. (1970) *J. Biol. Chem.* **245**, 3771–3775
- De Clercq, H., Despontin, M., Kaufman, L. & Massart, D. L. (1976) *J. Chromatogr.* **122**, 535–551
- El-Shazly, K. (1952) *Biochem. J.* **51**, 647–653
- Glascock, R. F. & Welch, V. A. (1974) *J. Dairy Sci.* **57**, 1364–1370
- Gray, I. K. (1973) *J. Dairy Res.* **40**, 207–214.
- Kowalski, B. R. (1975) *Anal. Chem.* **47**, 1152A–1162A
- Lennarz, W. J. (1961) *Biochem. Biophys. Res. Commun.* **6**, 112–116
- Massart, D. L. & Kaufman, L. (1975) *Anal. Chem.* **47**, 1244A–1253A
- Massart, D. L., Dijkstra, A. & Kaufman, L. (1978) in *Evaluation and Optimization of Laboratory Methods and Analytical Procedures*, pp. 361–384, Elsevier, Amsterdam
- Massart, D. L., Kaufman, L. & Coomans, D. (1980) *Anal. Chim. Acta* **122**, 347–354
- Massart-Leën, A. M., De Pooter, H., Decloedt, M. & Schamp, N. (1981) *Lipids* in the press
- McCarthy, S. & Smith, G. H. (1972) *Biochim. Biophys. Acta* **260**, 185–196

- Palmquist, D. L., Davis, C. L., Brown, R. E. & Sachan, D. S. (1969) *J. Dairy Sci.* **52**, 633–638
- Patton, S. & Jensen, R. G. (1976) in *Biomedical Aspects of Lactation* (Patton, S. & Jensen, R. G., eds.), pp. 55–56, Pergamon Press, Oxford and New York
- Peeters, G., Houvenaghel, A., Roets, E., Massart-Leën, A. M., Verbeke, R., Dhondt, G. & Verschooten, F. (1979) *J. Anim. Sci.* **48**, 1143–1153
- Popjak, G., French, T. H., Hunter, G. D. & Martin, A. J. P. (1951) *Biochem. J.* **48**, 612–618
- Scaiffe, J. R. & Garton, G. A. (1975) *Biochem. Soc. Trans.* **3**, 1011–1012
- Scaiffe, J. R., Wahle, K. W. J. & Garton, G. A. (1978) *Biochem. J.* **176**, 799–804
- Smith, G. H., McCarthy, S. & Rook, J. A. F. (1974) *J. Dairy Res.* **41**, 175–191
- Sneath, P. H. A. & Sokal, R. R. (1973) in *Numerical Taxonomy*, Freeman, San Francisco
- Wahle, K. W. J. & Paterson, S. M. (1979) *Int. J. Biochem.* **10**, 433–437