Computer-aided determination of the protein composition of extended meat products

H. OLSMAN, P. SLUMP and J. THISSEN

Division for Nutrition and Food Research TNO, Zeist, The Netherlands

1. Introduction

For the purpose of analysing specific protein components in food products, it is sometimes possible to use a more or less uncommon amino acid as a quantitative index of the protein it forms part of. The best example may be hydroxyproline which makes up some 12 to 14 % of the amino acids in collagen and gelatine, whereas it hardly occurs in any other food protein. It is a rather old idea to derive information about the identity of the proteins in a food product from the latter's overall amino acid composition. Wheat gluten, for instance, shows an unusually high glutamic acid content; the proline content is also comparatively high whilst, on the other hand, the levels of aspartic acid and lysine are rather low. Each protein or complex of proteins has its own characteristic mutual proportion of contents of constituting amino acids. It is only because of the complexity of the calculations involved, that verification of the viability of this idea had to wait for the availability of adequate computer facilities.

Lindqvist et al. (1) earned some success with this method in analysing dairy products and baby foods. They made use of an IBM Fortran program for stepwise regression. At the latest meeting of Meat Research Workers, Martens et al. (2) demonstrated the feasibility of the approach for the determination of the soya protein content in meat products. Our results obtained along this line have been mentioned earlier (3, 4), but will now be reported more in detail.

2. Materials and methods

2.1 Basic proteins

Meat proteins and non-meat proteins, which might come into consideration to be used in the manufacture of meat products, are given in table 1. These proteins provide the collection A of amino acid patterns on which the calculations in this study are based. Lean beef and pork, liver, sinews and rind, used for amino acid analysis, were representative samples of the raw materials used in the meat products studied (see 2.3). Likewise, except for caseinate, the amino acid composition of each of the non-meat proteins and the blood plasma was obtained with a sample of the same batch as used for preparing the meat products. The blood plasma was a spray-dried product containing 10.86 % N, supplied by Ruitenberg B.V., Amersfoort, the Netherlands. The nature and origin of the non-meat proteins were:

Hen's egg white: spray-dried, 12.83 % N, produced by NIVE, Harderwijk, the Netherlands.
Whey protein concentrate: spray-dried, 12.30 % N, obtained by ultrafiltration by the Neth. Inst. f. Dairy Research at Ede, the Netherlands
Wheat gluten: 12,78 % N, produced by Latenstein, B.V., Nijmegen, the Netherlands
Soya protein: texturate ADM-180, 8.29 % N, produced by Archer, Daniëls and Midlands
Sodium caseinate: spray-dried, type EM-6, 14.40 % N, DMV-Veghel, the Netherlands
Potato protein: heat-coagulated, experimental sample, 12.92 % N, Avebe, Veendam, the Netherlands

2.2 Amino acid analysis

Three hydrolysates were made per sample:
a. The material was boiled in 6 molar hydrochloric acid; test portions were 200 - 400 mg for the dry protein powders and 600 - 1000 mg for the raw meat materials and the meat products.
b. The same amounts were oxidized for 16 h at 0 °C in 30 ml performic acid reagent, containing 3 % H_2O and 88 % HCOOH, and subsequently hydrolysed as under a; methionine and cystine were thus transformed into methione sulfon and cysteic acid respectively.
c. A test portion of 1 - 2 g was autoclaved for 8 h at 130 °C in a bariumhydroxide solution.

Complete chromatographic separations were conducted with hydrolysates a and b, using an automated ion-exchange single column system. In hydrolysate c only tryptophan was determined, as described by Slump and Schreuder (5). In the meat ingredients and in the meat products, hydroxyproline was determined manually in a separate fourth hydrolysate, obtained by treating 4 g sample with 100 ml 6 m hydrochloric acid as specified under a.

2.3 Preparation of meat products

Four different products were made from the raw meat materials mentioned in 2.1, pork fat trimmings and two of the non-meat proteins of table 1. Lean meats and liver were freed as much as practicable from adhering connective tissue, cut into cubes 1 - 2 cm sidelength, passed through a meat grinder, and homogenized in a vacuum mixer. Sinews and rind were separately canned with water in the proportion 6 : 5 and heated at 70 °C. After cooling, the contents of each can were homogenized in a laboratory cutter. The pork fat trimmings, which were comminuted and mixed in a bowl chopper, were the only meat ingredient not subjected to amino acid analysis, as they were supposed to contain the same protein as rind or sinews. The batch of each ingredient, so obtained, was sufficient for the four products and the sample for laboratory analysis. The meat products furthermore contained the usual additives (water, salt, nitrite, phosphate, ascorbate). The actual protein compositions of each product, as calculated from the amounts of the ingredients and the results of their proximate analysis, are given in table 2.

2.4 Multivariate analysis

The solution of the following series of equations must be found:

$$y_1 = b_1x_{1;1} + b_2x_{1;2} + b_3x_{1;3} + \ldots\ldots\ldots + \underline{e}_1$$

$$y_2 = b_1x_{2;1} + b_2x_{2;2} + b_3x_{2;3} + \ldots\ldots\ldots + \underline{e}_2$$

$$\vdots$$

$$y_{19} = b_1x_{19;1} + b_2x_{19;2} + b_3x_{19;3} + \ldots\ldots\ldots + \underline{e}_{19},$$

where: $x_{i;j}$ is the content of amino acid i in the basic protein j

y_i is the content of amino acid i in the sample
(maximally 19 amino acids were determined: the 18 most common ones and hydroxyproline (Hyp)).

- b_j is the fraction of protein j in the sample; the coefficient is subjected to the constraints $0 \leq b_j \leq 1$ and $\sum_j b_j = 1$
- e_i is the difference between the contents of amino acid i, as determined in the sample and as calculated from the protein fractions b_j of the sample.

Multiple regression analysis is performed for the complete collection of basic proteins A. The coefficients b_j are estimated by minimalising the residual sum of squares $\sum_i e_i^2$. The proteins corresponding to those b_j-estimates, that exceed a pre-set treshold value, constitute the subcollection B of A. Next, the regression problem is solved for each combination of proteins of B. From the corresponding residual sum of squares (RSS), the residual variance is calculated:

$$s_e^2 = \frac{RSS}{n-p}$$

where n and p are the numbers of proteins in A and B respectively. The combination with the lowest s_e^2 will give the most probable protein composition. Usually, there are a number of combinations producing solutions of almost equal probability. The coefficients b_j in the corresponding regression equations are the estimates of the fractions of the basic proteins that make up the protein complex of the sample. Generally, the various protein compositions of almost equal probability differ only slightly in quantitative respect.
The amino acid contents of all basic proteins (table 1) and of the sample were offered to the computer in a normalised form: for each amino acid, the mean value of its contents in all proteins of collection A was calculated, and subsequently all amino acid contents - including those of the sample - were divided by the corresponding mean value. In this way, allowance could be made for the fact the some amino acids (tryptophan, cysteine, methionine etc.) always occur in lower quantities than others in proteins. Otherwise, the valuable information provided by these amino acids would not adequately affect the final result of the multivariate analysis. The reverse is true for amino acids that are generally abundant in proteins (e.g. glutamic acid + glutamine).

3. Results and discussion

The results of the multivariate analyses for the four meat products, each containing two non-meat proteins, are summarized in table 2. Three sets of solutions (calculated compositions) are given. For the first (I), the data of all 19 amino acids have been used. In the second set (II) tryptophan was excluded from the calculations in order to determine whether the information provided by the tryptophan data was worth the effort of the extra alkaline hydrolysis; aspartic acid was left out because in the chromatographic separation a small interfering peak, due to a contaminant in the elution buffer, might have affected the results. The third series of solutions was obtained by deleting tyrosine, phenylalanine and histidine as well. These amino acids are partly destroyed during hydrolysis b. As this is the only reason for hydrolysis a to be performed, it seemed to be worthwhile to determine the importance of the information provided by the three aromatic amino acids.
Solutions II do not appear to be significantly different from solutions I. With products 1 and 4, omission of Asp and Try even resulted in lower residual variances. The exclusion of still more amino acids (solutions III) had surprisingly little effect on the calculated protein composition as well as the residual variance. However, these observations, being based on the analysis of only four meat products, do not yet allow firm conclusions about the amount of amino acid data required to obtain satisfactory results. It will certainly also depend on the particular non-meat protein(s) in the sample; some have a "more exceptional" amino acid pattern than others.

Table 2 shows good agreement between the actual and the calculated composition for products 1 and 2. For product 4 the differences between both compositions was larger (esp. muscle and potato protein), but the protein components were identified correctly. Product 3 is the only one, in which a protein component (whey protein) was not recovered in all of the solutions; on the contrary, hen's egg white was wrongly found to be present in considerable amounts in all and sodium caseinate in nearly all solutions. That rind was sometimes mistaken for sinews (product 3 and 4) should not be taken too seriously. Discrimination between both connective tissue protein sources was anticipated to be very difficult because of their closely similar amino acid patterns (see table 1). Therefore, when the contribution of connective tissue is involved, the quantitative outcome for rind and sinew protein should rather be taken together. The amino acid patterns of lean beef and pork were almost identical, and the composition for muscle protein in table 1 is the average of both.
If the solution showing the lowest number of protein components is chosen, it appears that the right four were designated as the protein constituents in three of the four samples. The residual variance s_e^2 reflects the goodness of fit of the observed amino acid levels and those derived from the calculated protein composition. On the basis of long-term experience with the method, it would be possible to establish a treshold value for the residual variance, above which a solution is rejected as not adequately fitting the sample data. Thus, faulty results would be precluded. The possibility that two or more completely different protein compositions with equal probability and sufficiently low residual variance would result from the calculations is only theoretical, as long as the number of protein components in the sample remains well below the number of amino acids taken into

account for the numerical characterisation of basic proteins and sample.
the laboratory analysing the amino acid composition of the sample in order to determine its protein composition,
will also have to analyse the amino acid composition of all basic proteins that could possibly have been used in
manufacturing the product. This is necessary because of the existence of systematic between-laboratories
differences in the results; there is no generally accepted reproducible standard procedure available for amino
acid analysis.
Considering the results of this study, it should be borne in mind that the multivariate analysis was based on
the amino acid data of material from the very batches of the 11 proteins ingredients (except caseinate) that
were used in the manufacture of the four meat products. In this way optimal results can be expected which are
not adversely affected by small variations in amino acid pattern that always exist due to differences in origin
(genetic variation, climatical conditions, etc.) and processing of the raw materials.
Commercial meat products, which may contain protein hydrolysates, require a preliminary removal of low molecular
weight N-containing compounds.

5. Conclusion

While the viability of multivariate analysis for the determination of protein compositions is now established,
its ultimate success in practice will largely depend on three factors:
1. The reproducibility of the results of amino acid analysis.
2. The extent to which the amino acid pattern varies according to its origin and its processing and storage
history.
3. The extent to which the amino acid pattern of a protein component, actually present in the sample, differs
from that of the other proteins pre-selected as possible components of that sample.
In the continuation of the project special attention is paid to these aspects.

6. Literature

1. Lindqvist, B., Östgren, J. and Lindberg, I.
A method for the identification and quantitative investigation of denatured proteins in mixtures based on
computer comparison of amino acid patterns
Z. Lebensm. Unters. - Forsch. 24 (1975) 15 - 22
2. Martens, H., Hildrum, K.I., Bakker, A., Jensen, S.A., Lea, P., Eskeland, B., Vold, E. and Russwurm, H.
"Unscrambling" multivariate data from mixtures: Proceed. 26th Eur. Meet. of Meat Res. Workers, Colorado
Springs, Col. USA, Aug. 31 - Sept. 5, 1980
3. Olsman, W.J.
Methods for detection and determination of vegetable proteins in meat products
J. Am. Oil. Chem. Soc. 56 (1979) 285 - 7
4. Olsman, W.J. and Hitchcock, C.
Detection and determination of vegetable proteins in meat products; in: Developements in Food Analysis
Techniques - 2 (ed. R.D. King) pp. 225 - 60; Appl. Sci. Publ. Ltd., Barking, Essex, England, 1980
5. Slump, P. and Schreuder, H.A.W.
Determination of tryptophan in foods
Anal. Biochem. 27 (1969) 182 - 6

Table 1 Contents of 19 amino acids in 11 basic protein sources (collection A), in g per 16 g N

amino acid	basic proteins										
	egg wh.+	bl. pl.+	whey	wh. gl.+	muscle	liver	rind	sinews	Na-cas.+	soya	pot.+
Asp	12.9	11.2	10.3	3.7	9.9	9.9	7.1	7.1	7.3	15.4	12.8
Thr	4.9	6.7	6.6	2.5	4.6	4.4	2.0	2.4	4.7	4.0	6.0
Ser	7.7	6.7	5.1	5.1	4.1	4.7	3.6	3.5	6.2	5.6	5.8
Glu	14.3	14.6	16.1	34.6	16.3	12.3	10.1	10.2	23.1	18.4	11.3
Pro	3.7	5.6	5.6	12.4	3.4	4.2	12.0	9.4	11.1	5.0	4.3
Gly	3.5	3.4	1.9	3.1	3.9	5.1	19.8	16.2	1.8	4.0	4.8
Ala	6.4	5.2	4.7	2.5	5.6	5.6	7.6	7.1	3.1	4.3	5.1
Cys	3.0	3.6	2.3	2.1	1.17	1.75	0.27	0.48	0.40	1.51	1.73
Val	8.0	7.7	5.9	4.2	5.2	6.2	2.9	3.1	7.1	5.5	7.5
Met	3.8	0.89	1.50	1.78	2.8	2.6	1.11	1.45	2.7	1.43	2.4
Ile	6.0	4.0	6.0	3.8	4.9	4.7	1.57	2.2	5.6	5.0	6.0
Leu	8.6	9.4	9.0	6.8	7.8	8.8	3.4	4.0	9.5	7.6	10.1
Tyr	4.0	5.2	2.5	3.3	3.5	3.7	0.98	1.26	5.8	3.7	5.6
Phe	6.1	5.4	2.8	4.8	3.9	4.9	2.3	2.4	5.2	5.0	6.1
Lys	7.1	8.3	6.6	1.62	8.7	7.6	4.0	4.2	9.5	6.3	8.1
His	2.4	3.0	1.61	1.98	3.9	2.7	1.00	1.38	3.0	2.6	2.2
Arg	5.7	5.6	2.4	3.3	6.2	5.4	7.5	6.8	3.6	7.3	5.0
Try	1.42	1.91	1.50	0.93	1.19	1.49	0.11	0.28	1.30	1.47	1.46
Hyp	0.00	0.00	0.00	0.00	0.57	0.39	10.55	8.73	0.00	0.00	0.00

+ egg wh. = hen's egg white; bl. pl. = blood plasma protein; wh. gl. = wheat gluten; Na-cas. = sodium caseinate; pot. = potato protein

Table 2 Actual and most probable calculated protein compositions of four pasteurised meat products, each containing two non-meat proteins

protein	actual prot. comp. %	calculated protein compositions, in %											
		I all AA's included				II Try and Asp excluded					III Try, Asp, Tyr, Phe and His excluded			
1.													
muscle	13.8	17	15	17	17	15	15				16	15		
rind	41.1	40	39	40	40	40	40				40	40		
soya	28.9	24	22	23	23	26	26				25	25		
egg wh.	16.2	19	17	19	18	19	18				19	18		
liver	-	-	7	-	-	-	-				-	-		
sinews	-	-	-	1	-	-	-				-	-		
pot.	-	-	-	-	2	-	1				-	2		
res. var. S_e^2 (x 10^4)		8.7	8.5	9.0	9.1	6.1	6.5				7.4	8.0		
2.														
muscle	62.2	61				63	64				57			
rind	6.0	3				3	4				-			
liver	19.4	25				23	19				16			
Na-cas.	12.4	11				11	10				11			
sinews	-	-				-	-				5			
pot.	-	-				-	3				11			
res. var. S_e^2 (x 10^4)		15.0				15.0	15.9				12.4			
3.														
muscle	52.9	44	38	47		45	46	43	45	47	60	51	54	46
rind	12.3	-	-	12		-	-	-	-	-	11	12	12	-
sinews	6.3	15	15	-		15	15	15	15	15	-	-	-	15
wh. gl.	16.3	12	12	12		12	15	12	16	18	17	13	14	14
whey	12.2	8	6	9		7	-	6	-	-	-	5	-	-
egg wh.	-	13	10	13		13	16	12	12	9	12	12	13	11
Na-cas.	-	8	8	7		8	8	8	5	-	-	7	7	7
pot.	-	-	-	-		-	-	-	7	11	-	-	-	7
liver	-	-	11	-		-	-	4	-	-	-	-	-	-
res. var. S_e^2 (x 10^4)		22.7	22.1	23.0		18.2	24.1	19.6	23.1	24.1	24.9	16.9	18.7	17.5
4.														
muscle	33.2	25	22	24		26	24	26	26	25	26	25	25	
rind	12.7	-	-	-		-	-	-	-	-	-	-	-	
sinews	32.6	46	47	47		47	47	47	47	47	47	47	47	
bl. pl.	10.1	9	8	8		8	8	8	7	8	8	8	6	
pot.	11.4	20	18	19		19	19	18	18	19	19	18	18	
whey	-	-	-	2		-	-	-	2	-	-	-	4	
soya	-	-	5	-		-	2	-	-	-	-	-	-	
wh. gl.	-	-	-	-		-	-	-	-	1	-	-	-	
Na-cas.	-	-	-	-		-	-	1	-	-	-	2	-	
res. var. S_e^2 (x 10^4)		16.5	16.0	17.2		8.9	9.4	9.2	8.8	9.4	6.9	6.6	6.4	