ing the over-all preparation and distillation of the derivatives. For alanine, a specific rotation of 6.1 in 2 *N* hydrochloric acid was obtained. This is 42% of the expected value,[10] and indicates that 71% of the material was still in the L-form and 29% had been transformed to the D-form. For valine, a specific rotation of 22.8 in 2 *N* hydrochloric acid was obtained. This is 84% of the expected value,[11] and indicates that 92% of the material was still in the L-form and 8% has been converted to the D-form. Leucine had a specific rotation of 9.9 in 6 *N* hydrochloric acid. This is 65% of the expected value,[11] and indicates that 83% of the material was still in the L-form and 17% had been converted to the D-form. These small conversions to the D-form make it appear probable that the process could be made to yield the L-forms almost exclusively.

Most proteins contain an appreciable amount of glutamic acid, and therefore, the higher boiling fractions have a tendency to be predominantly pyrrolidone carboxylic acid ethyl ester, and the relative proportions of the other high boiling esters are small. We have, therefore, been unable to accumulate sufficient volumes of these higher boiling esters to permit an efficient fractionation study of them. A promising source of amino acids containing only small amounts of glutamic acid is the dried amino cakes obtained as by-products of the glutamic acid industry. Experiments with them have produced large quantities of the mono-amino-mono-carboxylic acid derivatives.

For large-scale production of amino acids by this method, preliminary separation of the low-boiling and high-boiling materials in a stripping column should be practical, because there is a 15 to 20° difference in the boiling points between these two groups of derivatives.

Further work on the preparation of the mixed acetylated amino acid ethyl esters may increase the yield of these materials. Distillation of these derivatives would then be useful for producing the naturally occurring amino acids in large quantities.

PHILADELPHIA, PA.

(10) P. J. Fodor, V. E. Price and J. P. Greenstein, *J. Biol. Chem.*, **178**, 503 (1949).

(11) M. A. Nyman and R. M. Herbst, *J. Org. Chem.*, **15**, 108 (1950).

[CONTRIBUTION FROM THE LABORATORIES OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH]

The Amino Acid Compositions of Insulins Isolated from Beef, Pork and Sheep Glands

BY ELIZABETH J. HARFENIST

RECEIVED MAY 25, 1953

Samples of beef, pork and sheep insulin which had previously been purified by countercurrent distribution have been analyzed for their amino acid contents. It has been found that the main components of the three insulins differ in their contents of threonine, serine, alanine, glycine, valine and isoleucine. The A and B components of beef insulin appear to differ only by one amide group.

Previous publications from this Laboratory indicate that the major component of beef insulin behaves as a single solute when studied by countercurrent distribution and that it has a molecular weight of the order of 6,000.[1,2] However, a molecular weight of 12,000 has been indicated by several physical methods.[3] Either molecular weight would be possible with the amino acid formula for insulin which Sanger has derived from his extensive structural studies,[4] but, on the other hand, the published figures for the amino acid composition of insulin[5] are in disagreement with his formula as far as certain residues are concerned. Several of the values also are incompatible with a molecular weight as low as 6,000. However, the analytical figures from the different laboratories show considerable discrepancies when compared with each other and give good reason to suspect either errors in the analyses or the use of samples which were not sufficiently homogeneous. We have, therefore, re-examined the quantitative amino acid composition of beef insulin in the hope that the major component isolated by countercurrent distribution would give an analytical result for every amino acid which would be fully compatible with the lower molecular weight indicated by the method of partial substitution and with the number of residues required by the peptide sequence proposed by Sanger.

Distribution studies with different insulin preparations[1] always revealed a major component (A) and varying amounts of a second component (B) present in smaller amount. The smaller component was found to have physical, chemical and biological properties very similar to those of the major component. An obvious approach to the question of the differences between these two active components was through comparison of their amino acid contents. Purified B component has accordingly been included in the analytical study. It appeared of considerable interest also to include in the study an investigation of the A components of pork and sheep insulins. Although it has been found that the major components of beef, pork and sheep insulins have very similar if not identical partition ratios in the system 2-butanol/1% aqueous dichloroacetic acid, observations concerning certain species differences have been reported.[6]

(1) E. J. Harfenist and L. C. Craig, THIS JOURNAL, **74**, 3083 (1952).

(2) E. J. Harfenist and L. C. Craig, *ibid.*, **74**, 3087 (1952).

(3) P. Doty, M. Gellert and B. Rabinovitch, *ibid.*, **74**, 2065 (1952); H. Gutfreund, *Biochem. J.*, **50**, 564 (1952); J. L. Oncley, E. Ellenbogen, D. Gitlin and F. R. N. Gurd, *J. Phys. Chem.*, **56**, 85 (1952).

(4) (a) F. Sanger and H. Tuppy, *Biochem. J.*, **49**, 463, 481 (1951); (b) F. Sanger and E. O. P. Thompson, *ibid.*, **53**, 353, 366 (1953).

(5) See G. R. Tristram, *Advances Protein Chem.*, **5**, 83 (1949); G. L. Mills, *Biochem. J.*, **50**, 707 (1952); C. Fromageot, *Cold Spring Harbor Symposia on Quantitative Biol.*, **14**, 49 (1950).

(6) (a) F. Sanger, *Nature*, **164**, 529 (1949); (b) J. Lens and A. Evertzen, *Biochim. Biophys. Acta*, **8**, 332 (1952).

Therefore it was highly important to learn the precise nature of these differences, not only as a contribution to the insulin problem, but also for a proper evaluation of the limitations of the methods of fractionation now at hand.

Experimental

The samples of insulin used in this study were those which had been purified previously by countercurrent distribution as already described.[1] They included the A and B components of beef insulin (Lilly), the A component of pork insulin (Nordisk) and the A component of sheep insulin (Boots).

Samples for hydrolysis were dried for three hours at 100° *in vacuo* before weighing. They were suspended in sufficient 5.7 *N* hydrochloric acid (glass distilled) to give 1% solutions of the protein and heated at 110° in sealed evacuated tubes for 24 hours except in the case of the A component of beef insulin. Here 48- and 96-hour hydrolyses were carried out in addition to the 24 hours.

The amino acids were estimated by the chromatographic procedure described by Moore and Stein.[7] The neutral and acidic acids were determined on a 100-cm. column of the ion-exchange resin, Dowex 50. Then a shorter column (15 cm.) was used for the basic amino acids and ammonia. The color yields and recoveries reported by Moore and Stein were used except in the case of ammonia which was found to give a color yield of 0.90, and histidine which gave a recovery of only 87% on a synthetic mixture. In our hands the determination of proline at 440 mμ as suggested by Moore and Stein[8] proved to be unreliable because of the presence of the large, partially overlapping band of glutamic acid. Therefore alternate tubes in the proline–glutamic acid region were analyzed by the Chinard acid ninhydrin method[9] for proline. For this purpose 0.2 ml. of 6 *N* phosphoric acid was added to each tube to give a *p*H of 1 or 2 before the addition of the ninhydrin reagent. The absorption was determined at 490 mμ.

As an independent check on the ammonia values from the chromatographic procedure, amide nitrogen was determined directly on the four samples by the method of Rees.[10] For these determinations the samples were weighed directly without drying. The results were corrected for the ash and moisture content found in the C, H, N and S analyses. For the latter analyses all samples were dried at 100° *in vacuo* for three hours. No change in nitrogen value was found when the drying temperature was 110°.

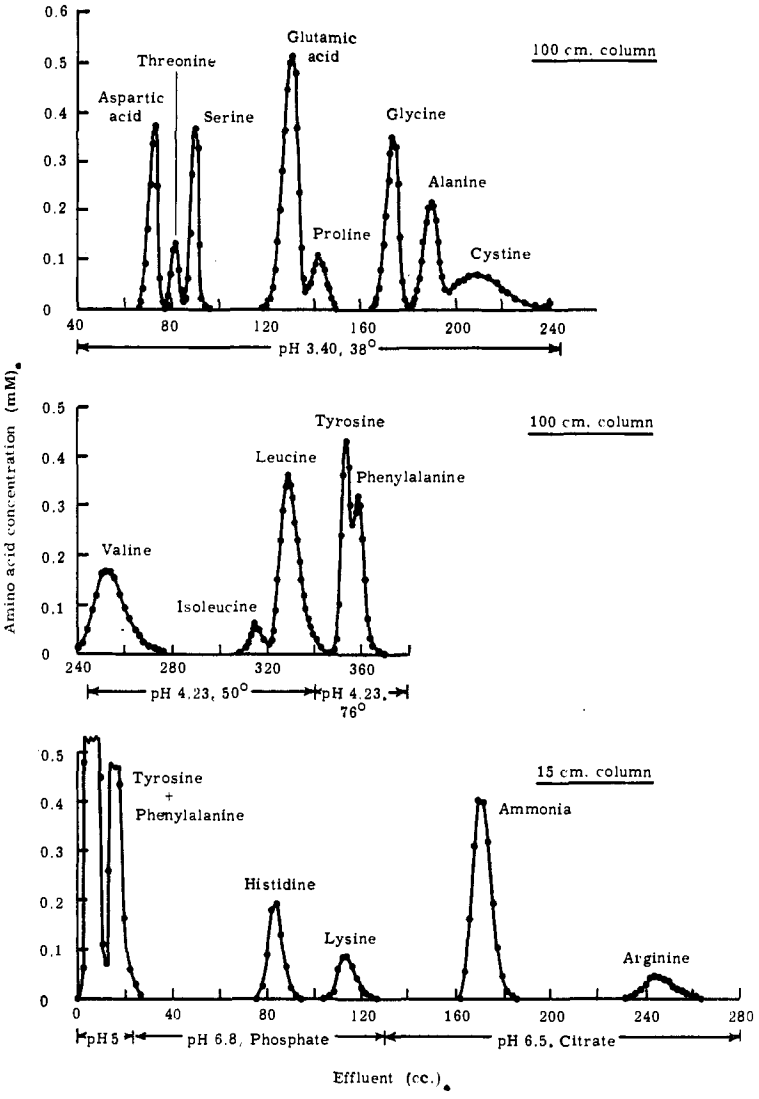

Fig. 1.—Typical effluent curves for 24-kv. hydrolysate of beef insulin.

Results

The data obtained are given in the form of tables except for a typical effluent pattern shown in Fig. 1. All values are calculated to an ash-free, dry weight basis. Tables I–IV give the results obtained with 24-hour hydrolysates of the four insulin samples. The values represent the average of at least two determinations except in a few cases, and in these cases there seems to be no doubt about the essential correctness of the results. The degree of precision can be seen from the variations indicated after the values. These variations indicate the actual range of values obtained and do not represent statistical errors. In no case are the deviations greater than 6% from the average value and in most cases they are within 3%. The recoveries of both weight and total nitrogen are in all cases almost quantitative, varying between 97 and 101%.

The threonine and serine values are those based on the 24 hour hydrolysis time and corrected for decomposition according to the estimates of Rees.[10] Therefore corrections for decomposition of 5% for threonine and 10% for serine have been applied to the values obtained directly from the column. The cystine values were determined from the total sulfur of the samples since there are no other sulfur-containing residues in insulin and since the cystine peak is smeared and overlaps both alanine and valine. There was insufficient separation of tyrosine and phenylalanine to warrant use of the method proposed by Stein and Moore for calculating overlapping bands,[11] and so the values for these two amino acids were taken as 4/7 and 3/7, respectively, of the combined peaks. This appeared to be the best estimate of the relative

(7) S. Moore and W. H. Stein, *J. Biol. Chem.*, **192**, 663 (1951).
(8) S. Moore and W. H. Stein, *ibid.*, **176**, 367 (1948).
(9) F. P. Chinard, *ibid.*, **199**, 91 (1952).
(10) M. W. Rees, *Biochem. J.*, **40**, 632 (1946).
(11) W. H. Stein and S. Moore, *J. Biol. Chem.*, **176**, 337 (1948).

TABLE I

AMINO ACID COMPOSITION OF THE A COMPONENT OF BEEF INSULIN DETERMINED FROM A 24-HOUR HYDROLYSATE

Amino acid	G. amino acid/100 g. protein	G. amino acid residue/ 100 g. protein	N as % of total N	No. of residues/mole of insulin
Aspartic acid	6.71 ± 0.14	5.80 ± 0.12	4.58 ± 0.09	2.89 ± 0.06
Threonine	2.00 ± .07	1.70 ± .06	1.53 ± .05	0.97 ± .03
Serine	5.28 ± .02	4.38 ± .02	4.57 ± .02	2.89 ± .01
Glutamic acid	17.27 ± .28	15.16 ± .25	10.68 ± .18	6.74 ± .11
Proline[a]	2.13	1.80	1.69	1.07
Glycine	5.17 ± .10	3.93 ± .08	6.26 ± .12	3.95 ± .08
Alanine	4.59 ± .04	3.66 ± .03	4.67 ± .03	2.95 ± .02
Cystine	12.18	10.36	9.22	2.91
Valine	9.65 ± .02	8.17 ± .02	7.49 ± .02	4.73 ± .01
Isoleucine	1.51 ± .03	1.30 ± .03	1.04 ± .03	0.66 ± .02
Leucine	13.45 ± .32	11.61 ± .28	9.27 ± .28	5.85 ± .18
Tyrosine	12.58 ± .31	11.33 ± .28	6.31 ± .15	3.98 ± .10
Phenylalanine	8.60 ± .20	7.66 ± .18	4.73 ± .09	2.98 ± .07
Histidine	5.41 ± .16	4.78 ± .14	9.50 ± .27	2.00 ± .06
Lysine	2.58 ± .11	2.26 ± .10	3.20 ± .13	1.01 ± .04
Arginine	3.10 ± .13	2.78 ± .12	6.46 ± .28	1.02 ± .05
Ammonia	.....	.....	8.95 ± .39	5.65 ± .24
Total		97.20	100.15	

TABLE II

AMINO ACID COMPOSITION OF THE B COMPONENT OF BEEF INSULIN DETERMINED FROM A 24-HOUR HYDROLYSATE

Aspartic acid	7.00	6.05	4.78	3.01
Threonine	2.10	1.78	1.61	1.01
Serine	5.44	4.51	4.72	2.97
Glutamic acid	17.65 ± 0.15	15.49 ± 0.13	10.94 ± 0.09	6.88 ± 0.06
Proline	2.13 ± .11	1.80 ± .09	1.70 ± .08	1.06 ± .05
Glycine	5.51 ± .01	4.19 ± .01	6.70 ± .02	4.20 ± .02
Alanine	4.81 ± .06	3.84 ± .05	4.91 ± .05	3.08 ± .04
Cystine	12.53	10.65	9.50	2.98
Valine	9.62 ± .26	8.14 ± .22	7.48 ± .21	4.70 ± .13
Isoleucine	1.39 ± .03	1.20 ± .03	0.97 ± .02	0.61 ± .01
Leucine	13.38 ± .12	11.54 ± .10	9.28 ± .08	5.83 ± .05
Tyrosine	12.19 ± .14	10.98 ± .13	6.13 ± .07	3.84 ± .05
Phenylalanine	8.35 ± .10	7.44 ± .09	4.60 ± .05	2.89 ± .04
Histidine	5.27 ± .05	4.66 ± .04	9.29 ± .08	1.95 ± .02
Lysine	2.51 ± .05	2.20 ± .04	3.12 ± .05	0.98 ± .02
Arginine	3.13	2.81	6.57	1.03
Ammonia	.....	.....	8.12 ± .05	5.12 ± .02
Total		97.82	100.42	

[a] Determined on a 96-hour hydrolysate.

amounts of the two acids and was in agreement with the ultraviolet absorption at 277 mμ due to tyrosine. The proportions so determined agreed with those obtained when the Stein and Moore method was used except in the cases where longer hydrolysis periods had been used. Here there was an indication that tyrosine was slowly decomposing. In order to calculate the number of residues of each amino acid per mole of insulin the best minimum molecular weights were determined from the amino acid compositions. These values are given in Table VII and have been used since they are in the range of the molecular weight determined by the method of partial substitution.

It is at once evident from Tables I–IV that except for valine and isoleucine all the amino acid constituents of the different insulins are present in approximately whole molar ratios for molecular weights of about 6,000. Incomplete hydrolysis could cause the values for these two amino acids to be too low, especially since according to Sanger and Thompson[4b] the peptide sequence isoleucylvaline is present in the A chain of beef insulin, and resistance of this peptide to acid hydrolysis is a phenomenon predictable from steric considerations[12] and also from previous experience with valylvaline.[13] No support for such a possibility could be derived from the 24-hour hydrolysis mixture, however, probably because isoleucylvaline gives considerably less color with ninhydrin than an amino acid and therefore does not show up in the effluent. This prediction is based on the finding that valylvaline gives a low color yield with ninhydrin.[14] Because of these considerations it appeared highly desirable to study the effect of longer hydrolysis periods with at least one of the preparations. This was done with the A component of beef insulin. The results are given in Table V. It will be seen from the

(12) H. S. Levenson and H. A. Smith, THIS JOURNAL, **62**, 1556 (1940); M. S. Newman, *ibid.*, **72**, 4783 (1950).

(13) See for example: W. A. Schroeder, *ibid.*, **74**, 5118 (1952); R. L. M. Synge, *Biochem. J.*, **39**, 351 (1945); *ibid.*, **44**, 542 (1949); H. N. Christensen, *J. Biol. Chem.*, **151**, 319 (1943).

(14) L. C. Craig, J. D. Gregory and G. T. Barry, *Cold Spring Harbor Symposia on Quantitative Biol.*, **14**, 24 (1950).

TABLE III

AMINO ACID COMPOSITION OF THE A COMPONENT OF PORK INSULIN DETERMINED FROM A 24-HOUR HYDROLYSATE

Amino acid	G. amino acid/100 g. protein	G. amino acid residue/100 g. protein	N as % of total N	No. of residues/mole of insulin
Aspartic acid	6.73 ± 0.14	5.82 ± 0.12	4.55 ± 0.10	2.93 ± 0.05
Threonine	3.66 ± .20	3.11 ± .17	2.76 ± .14	1.77 ± .10
Serine	5.01 ± .21	4.15 ± .17	4.30 ± .17	2.76 ± .11
Glutamic acid	17.93 ± .17	15.74 ± .15	10.97 ± .11	7.05 ± .07
Proline	2.11 ± .06	1.78 ± .05	1.65 ± .05	1.06 ± .03
Glycine	5.14 ± .07	3.91 ± .05	6.18 ± .08	3.97 ± .05
Alanine	3.32 ± .16	2.65 ± .13	3.36 ± .17	2.16 ± .10
Cystine	12.60	10.71	9.44	3.03
Valine	7.50 ± .18	6.35 ± .15	5.76 ± .14	3.70 ± .09
Isoleucine	3.51 ± .08	3.03 ± .07	2.40 ± .06	1.54 ± .04
Leucine	13.62 ± .21	11.75 ± .18	9.34 ± .15	6.00 ± .10
Tyrosine	12.49 ± .34	11.25 ± .31	6.21 ± .17	3.99 ± .11
Phenylalanine	8.54 ± .24	7.61 ± .21	4.65 ± .12	2.99 ± .08
Histidine	5.30 ± .25	4.68 ± .22	9.20 ± .45	1.96 ± .10
Lysine	2.49 ± .02	2.18 ± .02	3.06 ± .02	0.98 ± .01
Arginine	2.70 ± .09	2.42 ± .08	5.57 ± .18	0.89 ± .03
Ammonia	.....	.....	9.14 ± .56	5.88 ± .35
Total		97.66	98.54	

TABLE IV

AMINO ACID COMPOSITION OF THE A COMPONENT OF SHEEP INSULIN DETERMINED FROM A 24-HOUR HYDROLYSATE

Aspartic acid	7.00 ± 0.03	6.05 ± 0.03	4.73 ± 0.03	3.00 ± 0.02
Threonine	2.00 ± .07	1.70 ± .06	1.52 ± .05	0.96 ± .04
Serine	3.79 ± .07	3.14 ± .06	3.26 ± .06	2.06 ± .04
Glutamic acid	17.90 ± .05	15.71 ± .04	10.96 ± .02	6.95 ± .01
Proline	2.23 ± .04	1.88 ± .03	1.76 ± .02	1.10 ± .02
Glycine	6.17 ± .05	4.69 ± .04	7.41 ± .06	4.69 ± .04
Alanine	4.66 ± .04	3.72 ± .03	4.71 ± .03	2.99 ± .02
Cystine	13.05	11.10	9.78	3.10
Valine	9.81 ± .05	8.30 ± .04	7.54 ± .04	4.78 ± .03
Isoleucine	1.58 ± .05	1.36 ± .04	1.08 ± .03	0.69 ± .02
Leucine	13.74 ± .42	11.85 ± .36	9.43 ± .27	5.98 ± .18
Tyrosine	12.94 ± .33	11.65 ± .30	6.42 ± .15	4.08 ± .11
Phenylalanine	8.84 ± .24	7.88 ± .21	4.82 ± .12	3.06 ± .08
Histidine	5.18 ± .06	4.58 ± .05	9.02 ± .11	1.90 ± .03
Lysine	2.60 ± .08	2.28 ± .07	3.20 ± .10	1.02 ± .03
Arginine	2.88 ± .07	2.58 ± .06	5.92 ± .13	0.94 ± .02
Ammonia	.....	.....	9.69	6.14
Total		99.00	101.25	

table that, although the values for valine and isoleucine have increased by the end of 48 hours, the value for isoleucine is still too low to account for a whole molar ratio. Accordingly, a hydrolysis time of 96 hours was also studied. This increased the value for isoleucine to a figure fully compatible with a molecular weight in the range of 6,000. The values for none of the other amino acids were raised significantly by this long hydrolysis time, but there were decided decreases in the values for serine and threonine as was to be expected. Only one other amino acid, tyrosine, showed any signs of decomposition during the prolonged time of heating. All the other amino acids displayed a remarkable stability under our conditions of hydrolysis.

In Table VI the values for the four insulin preparations are expressed as numbers of residues per minimum molecular weights. These values have been rounded off to the nearest whole numbers. Although studies of longer hydrolysis periods have been reported only for the A component of beef insulin, a 48-hour hydrolysate of sheep insulin gave parallel results and it has been assumed that in all the samples the valine and isoleucine values would increase correspondingly. This has been taken into account in calculating the values for Table VI.

TABLE V

EFFECT OF PROLONGED HYDROLYSIS TIME ON THE AMINO ACID COMPOSITION OF THE A COMPONENT OF BEEF INSULIN

Amino acid	Number of residues 24 hr.	48 hr.	96 hr.
Aspartic acid	2.89	2.91	3.00
Threonine[a]	0.92	0.92	0.81
Serine[a]	2.60	2.36	1.87
Glutamic acid	6.74	6.86	7.11
Glycine	3.95	4.03	3.90
Alanine	2.95	3.13	2.96
Valine	4.73	4.97	5.14
Isoleucine	0.66	0.80	0.97
Leucine	5.88	5.86	5.99
Tyrosine[b]	3.98	3.91	3.59
Phenylalanine[b]	2.98	2.96	3.05

[a] Uncorrected for decomposition.
[b] Calculated by the method for overlapping bands.

The ammonia values are derived from the experimental amide values given in Table VII, since the ammonia figures from the columns were not reproducible.

TABLE VI

AMINO ACID COMPOSITIONS OF INSULINS FROM DIFFERENT SOURCES

Amino acid	Number of residues Beef-A	Beef-B	Pork	Sheep
Aspartic acid	3	3	3	3
Threonine	1	1	2	1
Serine	3	3	3	2
Glutamic acid	7	7	7	7
Proline	1	1	1	1
Glycine	4	4	4	5
Alanine	3	3	2	3
Cystine	3	3	3	3
Valine	5	5	4	5
Isoleucine	1	1	2	1
Leucine	6	6	6	6
Tyrosine	4	4	4	4
Phenylalanine	3	3	3	3
Histidine	2	2	2	2
Lysine	1	1	1	1
Arginine	1	1	1	1
Ammonia	6	5	6	6

The empirical formulas for the different insulins, calculated from Table VI, are given in Table VII. Exact molecular weights for these formulas and the C, H, N, S and amide nitrogen analytical data are also given.

is the first time that analytical data have been obtained which clearly permit the lower molecular weight figures calculated by the method of partial substitution.[2] The values for beef insulin are in complete agreement with the number of residues required by Sanger's postulated sequences in the peptide chains. Insulin had previously been analyzed several times,[5] in some cases with large variations among the results of different investigators. Comparison shows that only in the cases of glycine and valine do the present results differ markedly from most of the former values. The new, higher values are entirely in line with the other recent chemical studies on insulin.[4]

From the data of Table VI it appears that the only difference between the A and B components of beef insulin lies in their amide content as shown by the ammonia values. The A component has 6 amide groups while the B component has only 5. Since during the isolation of insulin from the pancreas it is exposed to alcoholic hydrochloric acid, it is not improbable that the B component arises during this isolation and therefore it is not necessarily present in the original insulin. This would explain the reason for the variation in the proportion of the B component among the samples originally investigated by countercurrent distribution.

Table VI shows clearly that the A components of beef, pork and sheep insulins differ in their amino acid contents although they appear identical by countercurrent distribution, crystal form, ele-

TABLE VII

ELEMENTARY COMPOSITIONS OF INSULINS FROM DIFFERENT SOURCES

Beef-A: $C_{254}H_{377}N_{65}O_{75}S_6$, mol. wt., 5733; beef-B: $C_{254}H_{378}N_{64}O_{76}S_6$, mol. wt., 5732; pork: $C_{256}H_{381}N_{65}O_{76}S_6$, mol. wt., 5777; sheep: $C_{253}H_{375}N_{65}O_{74}S_6$, mol. wt., 5703.

Insulin	Analyses, % Carbon Calcd.	Found	Hydrogen Calcd.	Found	Nitrogen Calcd.	Found	Sulfur Calcd.	Found	Amide N Calcd.	Found
Beef-A	53.19	53.00	6.62	6.68	15.88	15.40	3.36	3.25	1.46	1.42 ± 0.04
Beef-B	53.20	53.16	6.61	6.63	15.63	15.36	3.36	3.34	1.22	1.16
Pork	53.20	52.97	6.64	6.62	15.75	15.56	3.33	3.36	1.45	1.45
Sheep	53.26	53.04	6.63	6.74	15.96	15.55	3.37	3.48	1.47	1.45 ± .01

Discussion

The most striking feature of the experimental data is the close approximation of the values for the amino acid residues to molar proportions for molecular weights in the range of 6,000 which contain only single residues of several of the amino acids. This lends strong support to the thesis of purity of the samples prepared by countercurrent distribution. It should be pointed out that the results from only the 24-hour hydrolysis time were insufficient to obtain these data, and it appears from the experience with insulin that for a critical amino acid analysis of a protein at least two hydrolysis times should be studied. In this case the prolonged heating time was necessary to release all of the valine and isoleucine, and it also showed that the rates of decomposition of threonine and serine agree quantitatively with the observations of Rees,[10] thus giving justification to the use of his values for the corrections applied to these two amino acids. While the quantitative amino acid data do not distinguish between molecular weights in the range of 6,000 and those in the range of 12,000, this

mentary analysis and biological activity. There is no doubt that beef and pork insulins differ by one residue each of threonine, alanine, valine and isoleucine, and that beef and sheep insulins differ by one residue each of serine and glycine. These differences have not been clearly demonstrated and placed on a quantitative basis heretofore, although Sanger[6a] and Lens and Evertzen[6b] indicated certain qualitative differences in the amino acid compositions. Thus far immunological and other biological studies[15] have failed to show differences among insulins from different species. It is quite evident that the effect of the small differences found among the amino acid compositions, both on physical properties and on physiological activity, must indeed be small.

In a previous paper[1] it was noted that the partition ratios (K) of the A components from beef, pork and sheep insulins were nearly the same. It can therefore be inferred that the different

(15) D. A. Scott, *J. Biol. Chem.*, **92**, 281 (1931); D. A. Scott and A. M. Fisher, *Trans. Roy. Soc. Canada*, **34**, 137 (1940); P. Wasserman and I. A. Mirsky, *Endocrinology*, **31**, 115 (1942).

insulins would be separable with great difficulty, if at all, in the system used. This is an interesting observation in trying to assess from a structural standpoint the upper limit of separability. All the differences are small and are in neutral side chains which afford no changes in charge. However, covering or releasing one carboxyl or amino group in a weight of 6,000 causes sufficient differences in physical properties for resolution to be achieved. This has been shown in the separation of the A and B components. These differ only in that B has one more free carboxyl group than A. This conclusion is also supported by the experience with the reaction with fluorodinitrobenzene where covering one free amino group with the dinitrophenyl residue increases the K fivefold. Further unpublished data indicate that the methylation of one carboxyl group in 6,000 reduces the K by a factor of two. The separation of molecules of the size of insulin where there are only very small variations in the neutral amino acid residues, with resulting small changes in the physical properties, would be expected to be very difficult. In the future systems may be found with greater resolving power than that of the 1% aqueous dichloroacetic acid/2-butanol system used for all the work with insulin. However, it should be pointed out that in spite of the possible failure of this system in separating such closely related compounds, considerable purification has been achieved by countercurrent distribution. Several minor impurities have been removed and the A and B components which had not been detected by electrophoresis, the ultracentrifuge, the solubility method, or partition chromatography have been clearly demonstrated and separated.

Acknowledgment.—The author wishes to acknowledge the kind suggestions and advice of Dr. L. C. Craig, Dr. S. Moore and Dr. W. H. Stein, and the expert technical assistance of Miss E. A. Jacobs and Miss D. M. McNamara.

NEW YORK 21, N. Y.

[CONTRIBUTION FROM THE BANTING AND BEST DEPARTMENT OF MEDICAL RESEARCH, UNIVERSITY OF TORONTO]

Phosphatide Analogs. The Synthesis of Glycollecithins and Bis-(glycol)-phosphatidic Acids[1]

BY ERICH BAER

RECEIVED JUNE 17, 1953

A new class of phosphatides and a generally applicable method for the synthesis of its members is described. The new phosphatides, which can be regarded as analogs of both the lecithins and lysolecithins, and which differ from these substances only in that they contain glycol instead of glycerol, have been assigned the generic name "glycollecithin."[2] The glycollecithins possess neither positional- nor stereoisomers. The synthesis of two representative members of the glycollecithins, namely, stearoyl- and palmitoylglycollecithin, is described and their infrared spectra and other physical data are reported. The stearoyl- and palmitoylglycollecithins were found to be highly soluble in water and to possess a strong hemolytic activity. The new phosphatides because of their structural similarity to the glycerollecithins should prove of interest as substrates upon which to conduct systematic zymological and other biochemical and physiological studies. A general procedure is described for the synthesis of bis-(glycol)-phosphatidic acids which, in the form of their phenyl esters, are obtained also as by-products in the synthesis of the glycollecithins.

The naturally occurring phosphatides have been identified as esters of either glycerol, sphingosine or inositol. The isolation of propylene glycol phosphate from sea urchin eggs[3] and from cattle brain[4] seems to indicate the presence of a fourth polyhydric alcohol in phosphatides, namely, propylene glycol. Others, without doubt, will be found in time. Theoretically it is possible to visualize an almost unlimited number of phosphatide analogs by varying the polyhydric component. Although no experimental evidence has been reported suggesting the natural existence of ethylene glycol-containing phosphatides, it occurred to the author that phosphatides of this type, because of their structural simplicity, would be of interest to both the chemist and biochemist. The synthesis of glycol analogs of lecithin, cephalin and phosphatidyl serine is being undertaken in this Laboratory. The present communication, the first in this series, describes the preparation of glycollecithins[5] and of bis-(glycol)-phosphatidic acids.

The synthesis of the glycollecithins follows in general the procedure developed in this Laboratory for the synthesis of the α-lecithins[6] except that monoacyl glycols are used as starting materials. Since the naturally occurring glycerollecithins contain mainly fatty acids with 16 and 18 carbon atoms, preference has been given to the synthesis of palmitoyl and stearoyl glycollecithin. The required starting materials, namely, monopalmitoyl and monostearoyl glycol until now have been obtained by heating mixtures of either glycol and fatty acid,[7,8] or of glycol, fatty acid, camphor-β-sulfonic acid and phenol[9] to 180° for several hours and separating the monoacyl and diacyl glycols by fractional crystallization. Both procedures, however, have the disadvantage of yielding many

(1) An account of this work was included in a lecture presented before the American Chemical Society at its 75th Anniversary meeting, New York, N. Y., September, 1951.

(2) The prefix "glycol" should not be confused with the prefix "glyco" in glycolipids. The latter prefix indicates the presence of a carbohydrate in the lipid molecule.

(3) O. Lindberg, *Ark. Kemi. Mineral. o. Geol.*, **A16**, No. 15, 1 (1943).

(4) O. Lindberg, *ibid.*, **A23**, 1 (1946).

(5) In naming the new phosphatides the names of their glycerol analogs are retained but modified by the prefix glycol.

(6) E. Baer and M. Kates, THIS JOURNAL, **72**, 942 (1950).

(7) R. F. Ruttan and J. R. Roebuck, *Trans. Roy. Soc. Can., III*, [3] **9**, 1 (1915).

(8) I. Bellucci, *Chem. Z.*, **35**, 669 (1911).

(9) T. P. Hilditch and J. G. Rigg, *J. Chem. Soc.*, 1774 (1935).