PHYSICO-CHEMICAL STUDIES ON HAEMOLYSIN. PART V. ON THE COMPOSITION OF HAEMOLYSIN

By S. S. De

The molecular weight of haemolysin has been determined to be 35,200 from diffusion experiments. The estimation of the different amino acids was carried out and it was found that the different amino acids like arginine, cysteine, histidine, lysine, methionine, tyrosine and tryptophane comprise 1/8, 1/9, 1/12, 1/24, 1/36 and 1/96 respectively of all the amino acid residues. On examining the ratio and the frequency of the different amino acids it is found that haemolysin molecule must contain at least 288 amino acid residues.

The minimum molecular weight of haemolysin as obtained by multiplying the number of amino acid residues by its average residue weight (115·2 g.) is found to be 33,200.

The difference in molecular weight obtained by the two methods is thus about 4 per cent.

Bergmann and Niemann (*J. Biol. Chem.*, 1936, **115**, 77; 1937, **118**, 301; 1937, **122**, 577) have conducted a series of analysis of cattle haemoglobin, egg albumin, cattle fibrin and precursor of gelatin. It has been found by them that there exists a simple numerical relationship among the main amino-acids of these proteins. The total number of amino acid residues was deduced from the amino-acid composition of the protein by applying the periodicity hypothesis. It was found that cattle globulin contained 576 or $2^6 \times 3^2$ units, egg albumin 288 or $2^5 \times 3^2$ units and cattle fibrin 576 or $2^6 \times 3^2$ amino acid residues. From their results they have concluded that the total number of amino acid residues (N_t) and the number of individual amino-acid residues (N_i) and the frequency of the individual amino acid residues ($F_i = N_i/N_i$) which are contained in a protein molecule can be expressed by the following equations:

(*a*) $N_t = 2^n \times 3^m$ where n and m are positive whole numbers.

(*b*) $N_i = 2^{n'} \times 3^{m'}$ where n' and m' are either zero or positive whole numbers.

(*c*) $F_i = 2^{n''} \times 3^{m''}$ where n'' and m'' are either zero or positive whole numbers.

(*d*) $n = n' + n''$.

(*e*) $m = m' + m''$.

(*f*) $Nt = Ni' + Ni'' + Ni''' + Nii' + \ldots\ldots\ldots\ldots + N_i^x$.

From the determination of the number of individual amino-acid residues and their respective frequencies the total number of amino-acid residues has been found. From the analysis of the above proteins it has also been concluded that genuine protein molecules contain $n \times 288$ units, where n is a whole number other than zero.

The same authors obtained the minimum molecular weight of cattle globulin, cattle fibrin and gelatin by multiplying the total number of amino-acid in the protein by its average residue weight. The average residue weight has been obtained by subtracting the molecular weight of water from the molecular weight, averaged with respect to amount of the amino acids recovered from acid hydrolysis. As the average residue weight of egg albumin was estimated to be 123·9, its molecular weight was found to be 35,700 ($123{\cdot}9 \times 2^5 \times 3^2$) which is within the range of values (31,000–41,000) considered as the Svedberg unit. So also the molecular weight of cattle fibrin and cattle globulin were found to be 69,300 (*i.e.*, $120 \times 2^6 \times 3^2$) and 66,500 (*i.e.*, $115{\cdot}5 \times 2^6 \times 3^2$) or equivalent to two Svedberg units.

As full theoretical recovery of the amino-acids has not been obtained after hydrolysis, the average was adjusted in favour of the amino acids of lower molecular weight, because these cannot be determined quantitatively. Moreover, the amount of carbohydrate, lipoid,

inorganic ions or prosthetic group must be known before the calculated molecular weight is compared with that obtained by physical methods.

The average residue weight has been estimated by determining by titration, the increase in the number of equivalents of carboxyl and amino plus imino groups obtained after hydrolysis by acid. From the increase in the number of equivalents of carboxyl or amino groups, the number of peptide bonds present in a given weight of protein can be determined. Consequently the average weight of protein or average residue weight containing one mole of peptide bond can be calculated. The method followed was that of Hotchkiss (*J. Biol. Chem.*, 1939, **131**, 387), which he applied in the case of lactoglobulin. The average peptide weight as estimated here is equal to the average residue weight as determined by Bergmann *et al* (*loc. cit.*), only when amino acids are the main constituents. As haemolysin has been found to contain only very small quantities of inorganic ions and carbohydrates, the above method has been considered suitable for determining the weight per peptide bond.

TABLE I

Ratio of amino-acids in haemolysin after hydrolysis

Amino-acids.	Mol. wt.	Weight.	Gram mole per 100 g. protein	Ratio.	Frequency.
Arginine	174	18·94%	0·1088	36	8
Cysteine	121	11·60	0·0959	32	9
Histidine	155	11·25	0·0726	24	12
Lysine	146	5·24	0·0359	12	24
Methionine	149	3·62	0·0243	8	36
Tyrosine	181	4·34	0·0240	8	36
Tryptophane	204	1·81	0·0089	3	96

The percentage composition of crystalline haemolysin (De, *Ann. Biochem. Expt. Med.*, 1944, **4**, 45), has been determined for those amino acids for which reliable analytical methods are available and has been presented in column 2 of Table I. The values were recalculated on a gm. molecular basis and are given in column 3. The ratios of the amino acids given in column 4 were obtained from column 3. As the average residue weight of haemolysin is found to be 115·2 g., 100 g. of haemolysin must give approximately 0·868 g. molecule of the hypothetical average amino acid on complete hydrolysis. From the values given in column 3 it is observed that the various amino acids comprise 1/8, 1/9, 1/12, 1/24, 1/36, 1/36 and 1/96 of all the amino acid residues, and this frequency is given in column 5. On examining the ratios in column 4 and the fractional values in column 5, it can be concluded that haemolysin must contain at least 288 amino acid residues or a whole multiple thereof. When this number of amino acid residues is multiplied by the average residue weight (115·2), it is found that haemolysin has a minimum molecular weight of 33,200.

On analysis haemolysin was found to contain C, 48·66%, H, 6·21%, N, 15·92% and S, 3·88%.

The sulphur present in the form of cysteine and methionine accounts for the 3·85% of sulphur which is in fair agreement with the total sulphur content 3·88%. So it can be assumed that all the sulphur in haemolysin is present in the form of cysteine and methionine. The ultimate ratio of cysteine and methionine is found to be 32·8 (*vide*

column 4, Table I). So the minimum number of sulphur atoms present in haemolysin is 40. From this the ultimate composition of haemolysin has been found to be

$$C_{1247}H_{2028}O_{526}N_{378}S_{40}$$

The minimum molecular weight of haemolysin from this ultimate composition is thus found to be 33,200. But the molecular weight determined by the diffusion method is 31,900. So the variation in molecular weight as determined by the two methods is not more than 6%.

Estimation of the Basic Amino-Acids.—The method employed for the determination of the basic amino-acids is that of Block (*J. Biol. Chem.*, 1934, **106**, 457) with slight modifications.

Estimation of Methionine and Cysteine.—The salt-free haemolysin was extracted with petroleum ether and dried in a vacuum desciccator over P_2O_5, since traces of alcohol interfere with the estimation of cysteine and methionine.

The method followed for the estimation of these two amino-acids is that of Baernstein (*J. Biol. Chem.*, 1936, **116**, 25, 33), as modified by Kassell *et al* (*ibid.*, 1938, **125**, 145).

Estimation of Tyrosine and Tryptophane.—The method employed is that of Folin and Marenzi (*J. Biol. Chem.*, 1929, **83**, 89).

Determination of Average residue weight.

Crystalline haemolysin (salt free 486 mg.) was boiled under reflux condenser in 15 c.c. of 6 *N* hydrochloric acid for 18 hours. The hydrolysate which contained no filtrable humin was made up to 50 c.c. From aliquot portions nitrogen and ammonia were determined. And a 40 c.c. portion was concentrated *in vacuo* three times for removing the major portion of hydrochloric acid. The residue was made up to 10 c.c. and from aliquot portions chloride and nitrogen were estimated. Increase in amino group was determined in 90% acetone solution by titration with 90% alcoholic 0·05 *N* hydrochloric acid using naphthyl red as indicator. Increase in carboxyl group was determined in 90% alcoholic solution by titration with 90 per cent alcoholic potassium hydroxide using thymolphthalein as indicator. The results recorded below are the average of three sets of experiments.

TABLE II

0·05*N* acid or alkali per mg. of nitrogen.

	Acid groups.	Basic groups.	Mean.
Groups in original protein	0·11 c.c.	0·13 c.c.	—
Groups in hydrolysed protein	1·32	1·32	—
Increase not corrected for ammonia ($-NH_2$ and $-COOH$ groups)	1·21	1·19	1·20 c.c.
Ammonia	—	—	0·11
Peptide bond	—	—	1·09

From this the weight per peptide bond or average residue weight is found to be 115·2 g.

My best thanks are due to Dr. B. N. Ghosh for his kind encouragement and advice.

UNIVERSITY COLLEGE OF SCIENCE AND TECHNOLOGY, CALCUTTA. *Received April 25, 1944*