ANALYSIS AND MINIMUM MOLECULAR WEIGHT OF
β-LACTOGLOBULIN

BY ERWIN BRAND AND BEATRICE KASSELL

(From the Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York)

(Received for publication, June 15, 1942)

In 1934 Palmer (1), in R. K. Cannan’s laboratory, succeeded in obtaining from the plasma of cow’s milk a crystalline protein with the general properties of a globulin, which was subsequently designated as β-lactoglobulin (2). According to Palmer (1) and to Sørensen and Sørensen (3) from 50 to 60 per cent of the whey proteins can be obtained in the form of this globulin. Detailed physicochemical investigations (solubility (1), electrophoretic mobility (4), ultracentrifuge studies (2, 4), dissociation tendency (5)) indicate the homogeneity of β-lactoglobulin by criteria at present available. Pedersen’s ultracentrifuge investigations on skim milk (2) show that β-lactoglobulin is a protein native to milk and is not an artifact produced during isolation and crystallization.

This paper deals with the determination of cysteine, cystine, methionine, tyrosine, tryptophane, arginine, and threonine in β-lactoglobulin. From these data a value for the minimum molecular weight ($M_{\text{min.}}$) is obtained which is in close agreement with the molecular weight in solution computed from ultracentrifuge data (4). The number of amino acids per mole of β-lactoglobulin is ascertained with the aid of additional information (amid N and basic amino acid content (6, 7)). The rôle of the hydroxyl groups is discussed.

Micro- or semimicromethods for the accurate determination in proteins of total sulfur (8, 9) and in protein hydrolysates of sulfate sulfur (10), cysteine (11, 12), cystine (11, 12), cystine + cysteine (10), methionine (10), tyrosine (13), tryptophane (13), and arginine (14) have been described in recent publications from this laboratory. A system of analysis has been developed in which most of these constituents are determined by two independent methods. Hydrolysis is carried out in an inert atmosphere with HI in presence of hypophosphite; with HCl, and with HCl in presence of urea (12); with NaOH, and with NaOH-SnCl$_2$ in sealed tubes containing a minimum of enclosed air (13).

In HI digests, we determine methionine as volatile iodide and as homo-

---

1 We are indebted to Professor Cannan for a sample of recrystallized β-lactoglobulin and for information in advance of publication (6, 7).

2 Also Brand and Kassell, unpublished results.

365
### Table I

**Analysis of β-Lactoglobulin**

<table>
<thead>
<tr>
<th>Determination No.</th>
<th>Constituent</th>
<th>Hydrolysis</th>
<th>Method (bibliographic reference Nos.)</th>
<th>Per cent</th>
<th>Concentration (10^6 \times C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>1*</td>
<td>Cysteine</td>
<td>HCl-urea</td>
<td>(11, 12)</td>
<td>1.10</td>
<td>9.2</td>
</tr>
<tr>
<td>2*</td>
<td>Cystine</td>
<td></td>
<td>(11, 12)</td>
<td>2.29</td>
<td>19.1</td>
</tr>
<tr>
<td>3*</td>
<td>Cysteine +</td>
<td>HI</td>
<td>(10)</td>
<td>3.39</td>
<td>28.3</td>
</tr>
<tr>
<td></td>
<td>cystine</td>
<td></td>
<td></td>
<td>3.22</td>
<td>21.6</td>
</tr>
<tr>
<td>4</td>
<td>Methionine</td>
<td></td>
<td>(10)</td>
<td>0.905</td>
<td>28.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.692</td>
<td>21.6</td>
</tr>
<tr>
<td>5 = 3 + 4</td>
<td></td>
<td></td>
<td></td>
<td>1.597</td>
<td>49.9</td>
</tr>
<tr>
<td>6</td>
<td>Sulfate</td>
<td>HI</td>
<td>As H_2_S (10)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Total S</td>
<td></td>
<td>Pregl (8, 9)</td>
<td>1.60</td>
<td>50.0</td>
</tr>
<tr>
<td>8 = 7 - 6</td>
<td>Protein S</td>
<td></td>
<td></td>
<td>1.60</td>
<td>49.95</td>
</tr>
<tr>
<td>9 = 5 + 8 _ 2</td>
<td>Average pro-</td>
<td></td>
<td></td>
<td>1.60</td>
<td>49.95</td>
</tr>
<tr>
<td></td>
<td>tein S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Tyrosine</td>
<td>NaOH-SnCl_2</td>
<td>(13)</td>
<td>3.78</td>
<td>20.9</td>
</tr>
<tr>
<td>11</td>
<td>Tryptophane</td>
<td></td>
<td>(13)</td>
<td>1.94</td>
<td>9.5</td>
</tr>
<tr>
<td>12</td>
<td>Arginine</td>
<td>HCl</td>
<td>(14)</td>
<td>2.87</td>
<td>16.5</td>
</tr>
<tr>
<td>13†</td>
<td>Histidine†</td>
<td></td>
<td>Cf. (6)</td>
<td>2.0</td>
<td>13.0</td>
</tr>
<tr>
<td>14†</td>
<td>Lysine</td>
<td></td>
<td>“(6)</td>
<td>10.7</td>
<td>73.0</td>
</tr>
<tr>
<td>15</td>
<td>Threonine</td>
<td></td>
<td>“(16, 17)</td>
<td>5.85</td>
<td>49.1</td>
</tr>
<tr>
<td>16</td>
<td>Total N</td>
<td></td>
<td>Pregl-Dumas</td>
<td>15.60</td>
<td>1113.6</td>
</tr>
<tr>
<td>17†</td>
<td>Amino “</td>
<td></td>
<td>Titration, Van Slyke (cf. (6))</td>
<td>1.23</td>
<td>88.0</td>
</tr>
<tr>
<td>18†</td>
<td>Amide “</td>
<td>HCl, NaOH</td>
<td>Cf. (6, 7)</td>
<td>1.08</td>
<td>77.0</td>
</tr>
<tr>
<td>11a</td>
<td>Indole “</td>
<td></td>
<td></td>
<td>0.13</td>
<td>9.5</td>
</tr>
<tr>
<td>12a</td>
<td>Guanidino N</td>
<td></td>
<td></td>
<td>0.69</td>
<td>49.2</td>
</tr>
<tr>
<td>13a</td>
<td>Imidazole N</td>
<td></td>
<td></td>
<td>0.36</td>
<td>26.0</td>
</tr>
<tr>
<td>14a</td>
<td>N α-Amino N</td>
<td></td>
<td></td>
<td>1.02</td>
<td>73.0</td>
</tr>
<tr>
<td>19 = 18 + 11a +</td>
<td>Non-α-N</td>
<td></td>
<td></td>
<td>3.28</td>
<td>234.2</td>
</tr>
<tr>
<td>12a + 13a + 14a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 = 16 - 19</td>
<td>α-N</td>
<td></td>
<td></td>
<td>12.32</td>
<td>879.4</td>
</tr>
<tr>
<td>21 = 17 - 14a</td>
<td>Terminal α-N</td>
<td></td>
<td></td>
<td>0.21</td>
<td>13.0</td>
</tr>
<tr>
<td>22 = 20 - 21</td>
<td>Peptide N</td>
<td></td>
<td></td>
<td>12.11</td>
<td>864.6</td>
</tr>
<tr>
<td>23</td>
<td>Average residue weight = 1/C_α-N = 113.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23a</td>
<td>“ “ “ = 1/C_peptide N = 115.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Calculated as half cystine (mol. wt. = 120).
† Determinations by Cannan et al. (6, 7).
cysteine, cysteine + cystine as cysteine, sulfate sulfur as H$_2$S (10). In HCl hydrolysates, preferably in the presence of urea (12),$^2$ cysteine and cystine are determined separately by our photometric method (11). The formation of acid-insoluble humin interferes with the cysteine determination, since the precipitate may contain appreciable amounts of cysteine, as was first shown by Lugg (15). By carrying out the HCl hydrolysis in the presence of urea, the formation of an acid-insoluble humin precipitate can be almost entirely prevented with certain carbohydrate-containing proteins such as egg albumin and lactalbumin, although the hydrolysate may be dark brown owing to acid-soluble humin. With carbohydrate-free proteins, the HCl-urea hydrolysates are usually perfectly clear and colorless, or at most slightly yellowish (12).$^2$ In such (urea-containing) HCl hydrolysates, satisfactory results for both cysteine and cystine are obtained. In NaOH and NaOH-SnCl$_2$ hydrolysates, tyrosine is quantitatively separated from tryptophane, which is isolated as an insoluble mercury compound. The Millon reaction is used for the photometric determination of both these amino acids (13).

The analytical results are reported in Table I. The average values for the various constituents are given in per cent in Column 5 and in terms of their concentration in moles (or atoms) per gm. in Column 6, where for convenience the actual values are multiplied by $10^5$. The molar concentration ($C_i$) of an individual amino acid ($i$) is defined by Equation 1,

$$C_i = \frac{\%_i}{100 \times M_i}$$  \hspace{1cm} (1)

where $M_i$ is the molecular weight of an individual amino acid ($i$) and $\%_i$ the content of ($i$) in per cent by weight in the dry, ash-free and salt-free protein. It is customary to report the results of protein analysis as per cent of amino acid, notwithstanding the fact the amino acids are present in peptide linkage as amino acid residues (minus H$_2$O). From Equation 1 it is obvious that the molar concentration is the same, irrespective of whether it is calculated from the per cent amino acid divided by the molecular weight or from the per cent residue divided by the residue weight. This also holds for the calculation of $M_{\text{min}}$, according to Equations 3 and 4.

It can be seen from Table I that $\beta$-lactoglobulin contains 3.39 per cent of cysteine + cystine, one-third being cysteine and two-thirds cystine.$^3$

$^2$ Native $\beta$-lactoglobulin has no reactive SH groups, since we find that the nitroprusside test is negative (in the presence of ammonia and NaOH) and no oxidation

---

**Table I—Concluded**

‡ Average value; cf. foot-note 4.

§ These values result from the further evaluation of the data in Table II, Column 4.
Since sulfate S is absent (Determination 6), all of the total S, determined by elementary analysis, is protein S (Determinations 7, 8). Methionine (3.22 per cent) together with cysteine and cystine accurately accounts for the protein sulfur (Determination 8). The average value for protein sulfur (Determination 9) has, therefore, a high degree of accuracy, the agreement being better than 1 per cent. The tyrosine and tryptophane content are 3.78 and 1.94 per cent respectively. The arginine content is 2.87 per cent by the method described in the preceding paper (14). Threonine (5.85 per cent) was determined by Winnick's adaptation to Conway vessels (16) of the method of Shinn and Nicolet (17). The value for total N (15.60 per cent) was obtained by the Pregl-Dumas method.

Included in Table I are values for lysine, histidine, α-amino N, and amide N recently reported by Cannan, Palmer, and Kibrick (6) and by Warner and Cannan (7). The nitrogen of the side chains (i.e. indole, guanidino, imidazole, and ε-NH₂-N) was calculated from the percentages of the corresponding amino acids (Determinations 11a to 14a). These values, together with the amide N, constitute the non-α-nitrogen (Determination 19), which subtracted from the total N gives the α-N (Determination 20 = 12.26 per cent N); i.e., the average α-N content of the constituent amino acids, whether they are present in N peptide linkage or not. An estimate of the non-peptide (terminal) α-N can be obtained from the difference between the amino N (determined by titration and the Van Slyke procedure, Determination 17) and the ε-N of lysine (Determination 14a), on the assumption that neither proline nor hydroxyproline occupies a terminal position. This difference (cf. "Discussion" (6)) is rather large (Determination 21) and amounts to about 20 per cent of the ε-N of the lysine; it would indicate four to six such terminal groups for minimum molecular weights of from 30,000 to 40,000. Since the results for lysine by isolation procedures are apt to be low and since the value for amino N is probably accurate (titration and Van Slyke values check), the value for terminal α-N is prob.

by porphyrindin takes place (at 0° and pH 7). The heat-denatured protein, however, gives a strong nitroprusside test. The appearance of CyS—H groups upon heat denaturation and after acid hydrolysis may be due to the presence in the native protein of unreactive CyS—H or of CyS—X groups or both (for discussion cf. (12)). The preliminary experiments with porphyrindin also indicate that the phenolic groups are unreactive in the native protein.

4 A histidine content of 2.3 per cent was obtained by Cannan et al. (6) by the nitranilate method and from titration data; however, these authors could isolate only 1.7 per cent of histidine as flavianate (R. K. Cannan, private communication). In Table I (Determination 13) an average value of 2.0 per cent is given, corresponding to 5 residues in Table II. The effect of this uncertainty in the histidine content on Determinations 19, 20, 22, and 23 in Table I is slight, resulting in a possible error of ±1 residue in the corresponding values in Table II.
ably too high. Therefore the value for α-N (Determination 20) represents a maximum value and that for peptide N (Determination 22, i.e. α-N less terminal α-N) a minimum. The average residue weight (ARW) of the constituent amino acids is the reciprocal of the concentration (in atoms per gm.) of the α-N.

\[ ARW = \frac{1}{C_{α-N}} \]  

The reciprocal of the concentration of peptide N will closely approximate the average residue weight if the number of terminal α-N groups is small. In view of the uncertainty in the terminal α-N, a definite value for the average residue weight of β-lactoglobulin (Determinations 23 and 23a) cannot be established, the range according to the present data being from 113.5 to 115.6 (for these figures the further evaluation of the data in Table II, Column 4, is taken into consideration).

From the data in Table I the minimum molecular weight (\( M_{\text{min.}} \)) of β-lactoglobulin can be calculated on the assumption that the material analyzed is a pure chemical individual and that its indicated molecular composition must therefore involve integral numbers of residues of the constituent amino acids. \( M_{\text{min.}} \) of a protein is given by Equation 3,

\[ M_{\text{min.}} = \frac{M_i \times R_i}{(\%)_i} \times 100 \]  

where \( M_i \) and \( (\%)_i \) are as previously defined and \( R_i \) is an integer and represents the number of residues of \( (i) \) per \( M_{\text{min.}} \). Combining Equations 1 and 3 leads to Equation 4,

\[ M_{\text{min.}} = \frac{R_i}{C_i} \]  

If a single amino acid \( (a) \) has been determined, \( M_{\text{min.}} \) is obtained from Equation 3 or 4 by taking \( R_a = 1 \). If two amino acids \( (a) \) and \( (b) \) have been determined, minimum values for \( R_a \) and \( R_b \) are obtained by setting up two simultaneous equations for \( M_{\text{min.}} \), yielding Equation 5,

\[ \frac{C_a}{C_b} = \frac{R_a}{R_b} \]  

Conversion of this common fraction into the smallest simple fraction (i.e. one in which both numerator and denominator are integers) yields minimum values for \( R_a \) and \( R_b \). For \( R_a \) a value of 1, 2, 3, etc., is taken successively until for \( R_b \) the corresponding smallest integer is found, from which the experimental data do not differ significantly. If additional amino acids \( (c) \), \( (d) \), etc., have been determined, Equation 5 is set up for individual pairs and the combined minimum ratio established. E.g., if

\[ \frac{R_a}{R_b} = \frac{3}{4} \]  
\[ \frac{R_c}{R_d} = \frac{4}{9} \]  
\[ \frac{R_a}{R_d} = \frac{1}{1} \]
the smallest integral ratio \( a:b:c:d \) is 9:12:4:9. From the values for \( R \) so obtained and the corresponding concentrations, \( C, M_{\text{min.}} \) is calculated according to Equation 4.

Cystine (CyS—SCy, mol. wt. = 240) is the equivalent of two amino acids; it has, therefore, to be considered in terms of half cystine residues (CyS—, mol. wt. = 120); so that by definition the number of half cystine residues \( R_{\text{half cystine}} \) in a protein must always be an even integer. Cysteine (CyS—H, mol. wt. = 121) is best calculated in per cent of half cystine (mol. wt. = 120), and the number of cysteine residues in a protein molecule \( (R_{\text{cysteine}}) \) may obviously be even or odd.

At present, owing to experimental limitations on the accuracy of the analytical methods, significant results by this method of calculation can be obtained only for values of \( M_{\text{min.}} \) up to about 50,000. For higher values of \( M_{\text{min.}} \) such calculations become less and less significant, but under special conditions the limit may be extended to about 70,000. Amino acids present in very low or very high concentrations are not suitable for the calculations; the useful range of \( C_i \) is from 1.5 to 60 \( \times \) 10\(^{-5} \) mole per gm. and the optimum is 10 to 40 \( \times \) 10\(^{-5} \) mole per gm. The calculations must be restricted to amino acids which can be determined with a sufficient degree of accuracy (about 2 per cent). In the calculation of the integral values for the pairs of residues according to Equation 5, all values which differ from the nearest integers by more than about 2 per cent should be eliminated, while any value within 2 per cent of the nearest integers should be considered as a possible fit.

It is generally recognized that some uncertainty is attached to all amino acid determinations, since they are usually carried out in hydrolysates. The study of the behavior and destruction of the free amino acids under varying conditions (cf. (12) and foot-note 2) and the establishment of correction factors overcome this difficulty to some extent. There always remains the possibility that in a specific protein structure an amino acid is unusually labile and subject to increased hydrolytic destruction and other reactions. For example, the cystine in insulin is sensitive towards HI hydrolysis, whereas it is quite stable during HCl hydrolysis (18); also tryptophane in chymotrypsinogen is more extensively decomposed during hydrolysis by alkali alone than it is by alkaline stannite (12). However, in the case of the sulfur amino acids (and iodo amino acids), the analysis of hydrolysates may yield unequivocal evidence. If the protein sulfur, which can be determined by elementary analysis with an error of less than 1 per cent, is fully accounted for by methionine, cysteine, and cystine in hydrolysates, then we have conclusive evidence that the amount of the sulfur amino acids actually present in the protein has been determined. For the following proteins we have accurately accounted for the total...
sulfur (usually within 1 per cent), casein, lactalbumin, reduced lactalbumin, crystalline egg albumin (8); thyroglobulin (19); cattle globin, cattle blood fibrin;\textsuperscript{2} chymotrypsinogen (12); horse serum albumin B (20); horse serum albumin A, human serum albumin, ribonuclease, pepsin, α-, β-, and γ-chymotrypsin, trypsinogen, and trypsin.\textsuperscript{2}

In all these cases the number of sulfur atoms per \( M_{\text{min}} \) equals the sum of the residues of the sulfur amino acids as given in Equation 6,

\[
R_{\text{sulfur}} = R_{\text{methionine}} + R_{\text{cysteine}} + R_{\text{half cystine}}
\]  

(6)

The values for the sulfur amino acids are quantitatively more significant than those obtained for other amino acids and afford a favorable basis for the calculation of \( M_{\text{min}} \) according to Equations 3 to 6.

This approach has been satisfactory in the case of chymotrypsinogen (12) for which from the consideration of the sulfur distribution alone a value for \( M_{\text{min}} \) was obtained which was practically identical with the molecular weight deduced from osmotic pressure (21). In the case of β-lactoglobulin, other amino acids have to be taken into consideration to obtain a significant figure for \( M_{\text{min}} \).

Using the concentrations in Table I, Column 6, for the calculation of \( R_{\alpha} \) and \( R_{\beta} \) according to Equation 5, we find the following relationships for methionine and cysteine + cysteine (the per cent deviation from the nearest integers is given in parentheses),

\[
\frac{R_{\text{methionine}}}{R_{\text{cysteine + half cystine}}} = \frac{1}{1.31} (31%), \frac{2}{2.62} (13%), \frac{3}{3.93} (2%), \frac{4}{5.24} (5%)
\]

This series covers values for \( M_{\text{min}} \) up to 50,000. It can be seen that for 3, 6, 7, 9, 10, and 11 residues of methionine, the corresponding residues of cysteine + half cystine (i.e. 4, 8, 9, 12, 13, and 14) are within 2 per cent of the nearest integer. The number of residues of cysteine + half cystine must be consistent with the number of residues of cysteine (\( R_{\text{cysteine}} \)) and of half cystine (\( R_{\text{half cystine}} \)) derived from the separate determinations of these constituents (Table I, Determinations 1 and 2). Calculated\textsuperscript{5} according to Equation 5, \( R_{\text{cysteine}} : R_{\text{half cystine}} = 1:2 \); so that only values which are multiples of 3 can be considered for \( R_{\text{cysteine + half cystine}} \). This eliminates all but 9 and 12 from the series of figures given above. From the distribution of the sulfur amino acids, therefore, two alternative values are obtained for \( M_{\text{min}} \) of β-lactoglobulin and both these values are equally

\textsuperscript{4} The ratio is 1/2.08; the deviation from the integer is 4 per cent; indeed a greater deviation would not necessarily be significant, because any cysteine that has been oxidized is accounted for as cystine.
consistent with the analytical data: (a) $M_{\text{min.}} = 42,000$ with 9, 12, and 21 residues of methionine, cysteine + half cystine, and sulfur, respectively; (b) $M_{\text{min.}} = 32,000$ with 7, 9, and 16 residues, respectively.

As pointed out above, the most accurate determination is that of the protein sulfur; so that the decision between 16 and 21 atoms of sulfur per $M_{\text{min.}}$ must be made with the aid of additional data. Calculated according to Equation 5, $R_{\text{sulfur}} : R_{\text{tryptophane}}$ is 21:3.99 and 16:3.05, respectively; no decision can be made, since both figures are within 2 per cent of the nearest integers. For tyrosine we calculate (a) 21:8.78 (2.4 per cent) and (b) 16:6.70 (4.3 per cent); again no decision can be made, since both figures for the number of tyrosine residues differ by more than 2 per cent from integers. In view of the accurate accounting of the sulfur and of the close agreement of the tryptophane determination with the sulfur partition, we conclude that the tyrosine value is insufficiently accurate and is probably too low.

Next we try arginine and calculate $R_{\text{sulfur}} : R_{\text{arginine}}$ as 21:6.93 (1.0 per cent) and 16:5.29 (5.8 per cent). This favors (a).

In Table II the number of residues of the various constituents is given for $M_{\text{min.}} = 42,000$. The experimentally found residue numbers (Column 3) are corrected to the nearest integers in Column 4, the deviation from these integers being noted in Column 5.

The total number of N atoms per mole is 468. This figure is accurate within a few N atoms, since it is based on the determination of the total N (15.60 ± 0.02 per cent) which can be considered as accurately established.

Calculation of the total number of amino acid residues per mole ($R_{\text{total AA}}$, Determination 20 = No. 16 less No. 19) yields 370 including six terminal amino acids (terminal AA). Because of the uncertainty in the lysine and histidine values (cf. discussion of Table I) these figures represent maximum values; so that $R_{\text{total AA}}$ in $\beta$-lactoglobulin = 364 (+3) + one to six terminal AA.

According to certain current views on protein structure terminal amino groups may be interpreted as indicating polypeptide chains. It is therefore possible that $\beta$-lactoglobulin may contain more than one polypeptide chain per mole, but this question cannot be decided until the lysine content has been accurately established.

Palmer (1) originally reported a total N content of 15.3 per cent for $\beta$-lactoglobulin. Later, however, he found 15.6 per cent by using a copper-mercury catalyst in the Kjeldahl determination (5). The Pregl-Dumas determination yields the same value (Table I, Determination 16), which accurately represents the N content of dry, ash- and salt-free $\beta$-lactoglobulin since it has been consistently found with a number of preparations in several laboratories (personal communication from R. K. Cannan).
β-Lactoglobulin has been studied in great detail (more than 70 runs are reported) in the ultracentrifuge in Svedberg’s laboratory (2, 4). From sedimentation velocity and diffusion data and a specific volume of 0.751 (a value which we confirm), Pedersen computes an average value for $M$.

**Table II**

Amino Acid Residues and Nitrogen and Sulfur Atoms per Mole of β-Lactoglobulin

$M = M_{\text{min.}} = M_s = 42,000$.

<table>
<thead>
<tr>
<th>Determination No., Table 1</th>
<th>Constituent</th>
<th>No. of residues or atoms ($R$) per mole</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>Found (3)</td>
</tr>
<tr>
<td>0</td>
<td>Average protein S</td>
<td>20.98</td>
</tr>
<tr>
<td>3</td>
<td>Cysteine + half cystine</td>
<td>11.89</td>
</tr>
<tr>
<td>4</td>
<td>Methionine</td>
<td>9.07</td>
</tr>
<tr>
<td>11</td>
<td>Tryptophane</td>
<td>3.99</td>
</tr>
<tr>
<td>12</td>
<td>Arginine</td>
<td>6.93</td>
</tr>
<tr>
<td>1</td>
<td>Cysteine</td>
<td>3.86</td>
</tr>
<tr>
<td>2</td>
<td>Half cystine</td>
<td>8.92</td>
</tr>
<tr>
<td>10</td>
<td>Tyrosine</td>
<td>8.78</td>
</tr>
<tr>
<td>13‡</td>
<td>Histidine</td>
<td>5.46</td>
</tr>
<tr>
<td>14‡</td>
<td>Lysine</td>
<td>30.7</td>
</tr>
<tr>
<td>15</td>
<td>Threonine</td>
<td>20.7</td>
</tr>
<tr>
<td>16</td>
<td>Total N</td>
<td>467.8</td>
</tr>
<tr>
<td>17‡</td>
<td>Amino “ “</td>
<td>37.0</td>
</tr>
<tr>
<td>18‡</td>
<td>Amide “ “</td>
<td>32.3</td>
</tr>
<tr>
<td>19</td>
<td>Non-α-N</td>
<td>98.3</td>
</tr>
<tr>
<td>20</td>
<td>α-N</td>
<td>369.3</td>
</tr>
<tr>
<td>21</td>
<td>Terminal α-N</td>
<td>6.3</td>
</tr>
<tr>
<td>22</td>
<td>Peptide N</td>
<td>363.3</td>
</tr>
</tbody>
</table>

Total No. of amino acid residues per $M_{\text{min.}} = 364 (\pm 3) + 1$ to 6 terminal AA

* Deviation $= \frac{(4) - (3)}{(4)} \times 100$.
† Cf. foot-note 5.
‡ Determinations by Cannan et al. (6, 7).
§ Cf. foot-note 4.

of 41,600 and from sedimentation equilibrium an average for $M_s$ of 38,000 (cf. (4, 22)). Pedersen’s value for $M_s$ is practically identical with $M_{\text{min.}}$.

It can be concluded that the minimum molecular weight and the molecular weight in solution of β-lactoglobulin are identical (factor = 1), or, in other words, that β-lactoglobulin is monomolecular in solution, and that its molecular weight is close to 42,000.
The close agreement between $M_{\text{min.}}$ and $M_*$ is significant, since it probably indicates that some of the theoretical objections raised against $M_*$ are either not valid or are of no practical consequence, at least in the case of $\beta$-lactoglobulin. Both $M_{\text{min.}}$ and $M_*$ refer to the molecular weight of the unhydrated molecule, $M_*$ for reasons pointed out by Svedberg and Pedersen (22) and $M_{\text{min.}}$ because our determinations are carried out on air-dried material corrected for moisture (cf. (12) and experimental part). Consistent results for the moisture content of powdered, air-dried proteins can be obtained by drying to constant weight in an oven at 110° or, preferably, in vacuo over P$_2$O$_5$ at 100° (12). Sometimes it is questioned whether such results correspond to the actual moisture content of the protein, since, it is argued, some moisture might have been retained or some anhydride formation might have occurred. However, our moisture determinations would appear to give reliable results, since in preliminary experiments with Dr. D. Rittenberg we have found the same moisture content in crystalline egg albumin: (a) by drying at 100° in vacuo over P$_2$O$_5$ and (b) at room temperature by determining the water content by the isotope dilution method (23), water containing an excess of the heavy oxygen isotope O$^{18}$ (cf. (24)) being used as indicator.

The accurate estimation of the moisture content of proteins enters into the calculation of their molecular weight from x-ray measurements of the unit cell volume. Crowfoot's latest estimate (25) of the molecular volume of air-dried tabular crystals of $\beta$-lactoglobulin is 52,000 cu. Å, but, in the absence of specific data on the moisture content (presumably about 7 per cent) of the crystals and of their density, no estimate of the molecular weight was made. A molecular weight of 36,500 (not corrected for residual moisture) had previously been deduced by Crowfoot and Riley (26) on the basis of a unit cell volume smaller than that recently reported (25), and on the assumption that the density of the crystals was the same as that of insulin (measured in an organic medium; cf. (27)).

Attention should be called to the large number (at least forty-five) of hydroxyl groups per mole of $\beta$-lactoglobulin. About fifteen of these are in serine$^7$ and twenty-one in threonine, as estimated by the excellent methods of Nicolet and Shinn (17, 28, 29); nine are in tyrosine and an unknown further number in hydroxyproline. This is not unusual$^8$ since our data$^2$

---

$^7$ On the basis of the value reported by Nicolet and Shinn (28).
$^8$ Our values for serine are still approximate, but the threonine content of the following pure crystalline proteins can be considered as fairly accurate: horse serum albumin A and B (35 residues per mole), human serum albumin (33 residues per mole), chymotrypsinogen (35 residues per mole; cf. (12)), $\alpha$-chymotrypsin (11.3 per cent), $\beta$-chymotrypsin (10.4 per cent), $\gamma$-chymotrypsin (10.6 per cent), trypsinogen (5.2 per cent), trypsin (5.8 per cent), ribonuclease (9.0 per cent), swine pepsin (denatured, 9.5
on crystalline horse serum albumin A and B and on crystalline human serum albumin indicate a minimum of 80 hydroxyl groups per mole of 70,000. It is uncertain whether in the native protein these groups form hydrogen bridges or whether they exist in combined form and are set free only on hydrolysis.

In 1906, Emil Fischer (30) warned that experimental progress might be hindered by too exclusive an emphasis on the peptide linkage and pointed out that the hydroxyls of the hydroxyamino acids are by no means "in-different" groupings in the protein molecule. Fischer also suggested (30) that these hydroxyl groups might exist in proteins in the form of esters or ethers. Very few such groups have been found. Only one ether linkage is known: the thyroxine in thyroglobulin. The thio ethers lanthionine (cf. (31)) and cystathionine ((32), cf. (33)) have as yet not been found in native proteins. Phosphoric acid esters exist; e.g., in casein where some of the aliphatic hydroxyl groups are so linked. The occurrence of esters of sulfuric acid, in view of the sulfate content of many purified proteins, has to be considered as possible. Their presence is strictly excluded only if a protein has been obtained free from sulfate as, for example, β-lactoglobulin, egg albumin, insulin, pepsin, and others. Esters with carboxylic acids are still purely hypothetical; a small number may exist, but the pertinent data are as yet not accurate enough to exclude or to establish this type of linkage in any protein.9

With regard to hydrogen bonding, it would seem from the phrases cited above that Fischer had anticipated some such effect. Like the sulfhydryl groups, the phenolic hydroxyl groups of native β-lactoglobulin (similar to egg albumin (34)) are highly unreactive. They cannot be titrated (6), they react with Folin's phenol reagent only in strongly alkaline solution, and they are not oxidized by porphyrindin.3 This contrast to the behavior of free tyrosine may reasonably be ascribed to intramolecular hydrogen bonding, akin to chelation. To judge from the available evidence for other compounds (cf. (35–37)) the aliphatic hydroxyl groups in the native proteins may also be involved in hydrogen bonding and thus contribute to the cohesion of the molecule, particularly by hydrogen bridges through water molecules. Such water molecules would be expected to exchange

per cent), purified elastin (about 2.7 per cent). We are indebted to Dr. H. Neurath for horse serum albumin A, to Mr. Manfred Mayer from this Department for B, to Dr. W. H. Stein and Dr. E. G. Miller, Jr., for the elastin, to Dr. F. E. Kendall for the human serum albumin, to Dr. R. M. Herriott for pepsin, and for the other crystalline enzymes to Dr. M. Kunitz.

9 At the present state of our knowledge caution should be exercised in assuming that free carboxyl groups are equal to the difference between total dicarboxylic acids and amide groups.
readily with water containing the heavy isotope $^{18}O$ (cf. above). Obliteration of hydrogen bridges may be related to denaturation by dehydrating agents. The hypothesis of the rôle of the hydroxyl groups in the binding of water is in harmony with Perutz's view on hydration; *i.e.*, "the disposition of water between structural units of the molecule in such a way as to cause an internal expansion" (38).

The surprisingly large number of hydroxyl groups which can be recognized in the side chains of protein molecules is bound to play an important part in the future development of hypotheses of protein structure and enzyme action.

**EXPERIMENTAL**

The sample of $\beta$-lactoglobulin employed in these studies was obtained from Dr. Cannan as a suspension in water, preserved with toluene. Most of the water was removed by centrifugation and the crystalline mass was dried over $P_2O_5$ in a vacuum desiccator at room temperature. The dry cake was finely powdered in an agate mortar and equilibrated in air until it reached constant weight (9 days). The air-dried material was used for all analytical procedures, the weights being corrected for moisture content. The moisture determinations were carried out as described previously (12); average moisture content = 7.01 per cent. The analytical data are reported in Table I.

**Total Nitrogen**—We are indebted to Mr. W. Saschek for this determination (No. 16), carried out by the Pregl-Dumas method.

**Total Sulfur**—We are indebted to Mr. W. Saschek for this determination (No. 7), carried out by a modification of the Pregl method (8, 9, 12). The ash was negligible, being less than 0.1 per cent.

**Methionine**—The value reported in Table I (Determination 4) is the average of four determinations with about 250 mg. of protein each; the same results were obtained by the volatile iodide and homocysteine titrations (10).

**Cysteine, Cystine, and Cysteine + Cystine**—For the separate determination of cysteine and cystine by the photometric method (11) about 100 mg. of protein were hydrolyzed for 16 hours with $HCl$-urea in an inert atmosphere, as described previously (12); the hydrolysates were practically colorless and there was no solid humin. The values reported (Determinations 1 and 2) are the average of three determinations; hydrolysis for 8 hours gave low results.

Cysteine + cystine (Determination 3) was determined in HI hydrolysates simultaneously with methionine (10).

For the cystine determination by the Sullivan method 200 mg. of protein were hydrolyzed under nitrogen gas for 16 hours with $6 N HCl$. Owing to
the presence of cysteine in the hydrolysate, the results were high (for discussion cf. (12)), the "cystine" content being 3.89 per cent. This is additional evidence (cf. (12)) that the difference between total S and methionine S is accounted for by cysteine and cystine and not by any other sulfur-containing compound.

Tyrosine and Tryptophane—The values reported (Determinations 10 and 11) are the average of three determinations with 30 to 35 mg. of protein in NaOH-SnCl₂ hydrolysates.

Arginine—The value reported (Determination 12) was obtained by the method described in the preceding paper (14).

Threonine—About 40 mg. of protein were hydrolyzed under nitrogen for 16 hours at 130° with 2 cc. of 6 N HCl. The value reported (Determination 15) is the average of three separate determinations by Winnick's modification of Shinn and Nicolet's method. Winnick (16) reports a threonine content of 5.36 per cent for his preparation of β-lactoglobulin, which, however, contained only 14.35 per cent of nitrogen. Nicolet and Shinn (28) find only 4.72 per cent of threonine but state that, because of the small amount of material available for analysis, their value may be less reliable than some of their other determinations.

Specific Volume—The determination was carried out in capped pycometers (about 10 cc.). The concentration was 1.816 per cent of β-lactoglobulin in 0.5015 per cent NaCl (d₄⁰/₂₀ = 0.99960). The density of the protein solution was d₄⁰/₂₀ = 1.00410. From these data we compute V₂₀ = 0.754, corrected V₂₀ = 0.750, which is in close agreement with Pedersen's (4) value of V₂₀ = 0.7514.

SUMMARY

For β-lactoglobulin a minimum molecular weight of 42,000 was obtained from the distribution of the sulfur amino acids and from the arginine content.

The minimum molecular weight is practically identical with the molecular weight in solution (M₄ = 41,600). Therefore, M = Mₘᵢₙ = M₄. The total number of amino acid residues per molecule of β-lactoglobulin is estimated to be 364 (±3) + one to six terminal amino acids.

1 molecule of β-lactoglobulin contains the following residues: cysteine 4, half cystine 8 (i.e. 4 S–S linkages), methionine 9, tryptophane 4, tyrosine 9, arginine 7, threonine 21, serine about 15, amide groups 32, histidine 4 to 6, and lysine 31 to 36.

Attention is called to the large number of hydroxyl groups which can be recognized in the side chains of β-lactoglobulin and of other proteins. The possible contribution of the hydroxyl groups to the cohesion of the molecule by hydrogen bridges through water molecules is discussed.
BIBLIOGRAPHY

38. Perutz, quoted by Crowfoot (26).
ANALYSIS AND MINIMUM
MOLECULAR WEIGHT OF $\beta$
-LACTOGLOBULIN
Erwin Brand and Beatrice Kassell


Access the most updated version of this article at http://www.jbc.org/content/145/2/365.citation

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/145/2/365.citation.full.html#ref-list-1