The Amino Acid Sequence of the γ Chain of Bovine Fetal Hemoglobin*

Donald R. Babin, W. A. Schroeder, J. Roger Shelton, Joan Balog Shelton, and Barbara Robberson

ABSTRACT: The 145 residues of the γ chain of bovine fetal hemoglobin have been placed in sequence. For the investigation, the bovine fetal hemoglobin was isolated chromatographically from the blood of a newborn calf, the component polypeptide chains were separated by countercurrent distribution, and the peptides necessary for the determination of sequence were produced by tryptic and chymotryptic digestion.

The bovine γ chain has one residue less than the human γ chain and less than the human and horse β chains. When the sequences of six β and γ chains from human, horse, and bovine hemoglobins are compared, certain lengths of sequence show great variation from chain to chain, whereas others are virtually invariant. Examination of the three-dimensional model of hemoglobin shows that the invariant sections enclose the heme group and suggest that they are important in the proper functioning of hemoglobin. On the other hand, two variable sections near the heme group may influence such properties as oxygen affinity in which the various hemoglobins differ. Further ramifications from this examination are also discussed.

The human infant at birth has a hemoglobin (hemoglobin F) that differs pronouncedly in its properties from those of the hemoglobin (hemoglobin A) in the normal adult: the amino acid sequences in the two types are known (see Schroeder and Jones, 1965, for a compilation of data and references). Of the four polypeptide chains in both hemoglobins, the pair of α chains in each is identical in sequence. The β chains of hemoglobin A, on the other hand, differ greatly in sequence from the γ chains of hemoglobin F. Much of our knowledge of the adult and fetal forms of animal hemoglobins stems from the investigations of Muller (1961). In cattle, sheep, and goats, the same situation obtains as in man because one pair of chains is identical and the other is not. The subunit structure of an adult hemoglobin can then be represented as $\alpha_2\beta_2$ and of a fetal hemoglobin as $\alpha_2\gamma_2$.

Many abnormal adult human hemoglobins differ from human hemoglobin A in the substitution of one or another residue in either the α or β chains. Although these abnormal hemoglobins are naturally occurring "derivatives," their properties usually differ little from hemoglobin A except electrophoretically because of altered charge. In any attempt to correlate structure and function, they are, for the most part, less valuable than a comparison of hemoglobins A and F which are very different both in structures and properties. The animal hemoglobins, of course, offer still more "derivatives" for comparison and frequently show interesting correlation or lack of it with human hemoglobins. Thus, human hemoglobin A is very easily denatured by alkali, and human hemoglobin F is very resistant to such denaturation. However, the reverse situation obtains in adult and fetal bovine hemoglobins (Brinkman and Jonxis, 1936; Haurowitz *et al.*, 1954). Quantitatively, the two bovine hemoglobins differ in rate of denaturation by about a factor of two (van der Helm, 1958). Both are, therefore, very much like human hemoglobin F (Jonxis and Huisman, 1956), but the three are very different from human hemoglobin A whose rate is about 100-fold faster. Presumably, these characteristics are reflections of the structures of the molecules. Without entering into the question of the utility of any property to a given organism, it is appropriate to ask whether or not the sequence and structure of bovine hemoglobin F is more similar to that of human hemoglobin F than to human hemoglobin A. The results of Perutz and collaborators (Cullis *et al.*, 1962) on the three-dimensional structure of horse hemoglobin and of Kendrew and colleagues (1960) on that of sperm whale myoglobin suggest that four-chain hemoglobins probably will be very similar in three-dimensional conformation. Yet the sequence may very well influence the finer details of the structure; this is apparent from the differences in the α and β chains. Before some of these questions can be answered, complete sequences of more hemoglobins must be available. As yet, the only animal hemoglobin chain from which the complete sequence has been reported is the horse α chain (Matsuda *et al.*, 1963), although Smith (1964) has described about 80% of the sequence of the horse β chain.

The bovine hemoglobins because of the properties that have been mentioned above are attractive subjects for further investigation. Not only are the fetal and adult hemoglobins distinctive, but two kinds of adult

* Contribution No. 3333 from the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91109. *Received January 20, 1966.* This investigation was supported in part by grants (GM-04276 and HE-02558) from the National Institutes of Health, U. S. Public Health Service.

bovine hemoglobin also occur (Bangham, 1957). Ozawa and Satake (1955) reported that Val-Leu and Met-Gly are N terminal in adult bovine hemoglobin. Braunitzer (1958) substantiated the first sequence but isolated DNP[1]-Met-Leu rather than DNP-Met-Gly in an investigation of the N-terminal residues. Satake and Sasakawa (1963) have presented a partial sequence of the adult bovine α chain which disagrees in some aspects with very fragmentary data of Muller (1961). The present investigation was begun with the determination of the sequence of the γ chain of bovine fetal hemoglobin as its primary objective. In conjunction with the sequence of the bovine α chain as presented by Satake and Sasakawa, one then could make comparisons as suggested above. The complete sequence of the bovine γ chain is presented below. However, preliminary experiments with the α chain of bovine fetal hemoglobin have caused us to question some data in the literature. The results differ from those of Satake and Sasakawa on the α chain of adult bovine hemoglobin: either the α chains from adult and fetal bovine hemoglobins may be different or hemoglobin from different breeds of cattle may be unlike. This matter will be the subject of a separate communication.

Experimental Section

Source of Blood Sample. About 60 ml of blood from a 1-hr-old Holstein bull calf was mixed with an equal volume of 3.2% sodium citrate dihydrate to prevent coagulation. The sample was cooled in ice during conveyance to the laboratory where the preparation of hemoglobin was begun immediately.

Preparation of Hemoglobin Solutions. A solution of oxyhemoglobin was prepared from the blood by methods that did not depart significantly from those of Clegg and Schroeder (1959). After all cell debris had been removed by two centrifugations at 25,000*g*, the hemoglobin was then saturated with carbon monoxide and dialyzed against the chromatographic developer.

Chromatographic Isolation of Bovine Fetal Hemoglobin. From exploratory chromatograms on columns of Amberlite IRC-50 with adult bovine hemoglobin, it was learned that developers for human hemoglobins (Allen *et al.*, 1958; Clegg and Schroeder, 1959; Schnek and Schroeder, 1961) essentially eluted the bovine hemoglobin. A satisfactory developer for bovine fetal hemoglobin is no. 6b of Jones (1961) (27.6 g of monosodium phosphate dihydrate and 2.60 g of potassium cyanide in 4 l.; pH 6.20 ± 0.05). For isolative purposes, the procedure was identical with that for the isolation of human hemoglobin F (Schroeder *et al.*, 1963). About 2 g of calf hemoglobin was chromatographed on a 3.5 × 35 cm column with developer no. 6b at 6° until three minor zones had emerged from the column and the main zone (hereafter designated hemoglobin F_{Bov}) was about to emerge. Under these conditions, the adult bovine hemoglobins are strongly fixed. Consequently, when the hemoglobin F_{Bov} had reached the bottom of the column, the chromatogram was stopped, and the section with the adult hemoglobin was removed. The remaining column which now contained only hemoglobin F_{Bov} was warmed to 40° for 20 min after which time the flow of developer was started again at *ca.* 750 ml/hr to elute the F_{Bov} rapidly into chilled containers.

Concentrating of Hemoglobin Solutions. The hemoglobin F_{Bov} in the effluent was concentrated by centrifugation (Vinograd and Hutchinson, 1960), and the concentrated solution was dialyzed against distilled water at 4°.

Preparation of Globin. Globin was prepared in the manner already described (Schroeder *et al.*, 1963). However, after the globin had been thoroughly washed with acetone, it was not dissolved and dialyzed but rather dried over solid potassium hydroxide overnight in a desiccator. The product so prepared appears to dissolve more readily in the solvents for countercurrent distribution than does lyophilized material.

Separation of the α and γ Chains. Countercurrent distribution according to the method of Hill *et al.* (1962) with simplified apparatus was used to separate the chains of bovine fetal globin. The simplified apparatus gives satisfactory separations in relatively few transfers because of the fortunate relationship of the distribution coefficients of the hemoglobin chains. A 4-ft metal rod which was supported by bearings provided an axle to which were attached 18 buret clamps alternately on opposite sides of the rod for balance. The clamps held either glass-stoppered graduates of appropriate size or 125-ml cylindrical separatory funnels depending upon the volume to be used. The axle could be rotated by motor at 35 rpm through a belt and pulleys. Thus, the sample was dissolved in lower phase in graduate or funnel no. 1 and an equal volume of upper phase was added. The phases were then equilibrated by rotating the axle for 2 min. If small volumes (about 10 ml of each phase) were used, it was convenient to use graduates and to move the upper phases by syringe. With larger volumes, separatory funnels and the movement of the lower phases were better. Equilibration, separation of phases, and movement of one phase was continued until transfers had been completed to the end of the train.

In a typical experiment, 500 mg of globin F_{Bov} was dissolved in 50 ml of lower phase with constant agitation over at least 3 or 4 hr. After the sample had dissolved and the upper phase had been added, the two were equilibrated and then allowed to separate overnight. However, in most cases an emulsion was still present and usually had to be broken by centrifugation or with occasional success by warming to 60° for a few minutes.[2]

Emulsion was less after each transfer and usually

[1] Abbreviations used in this paper: DNP, dinitrophenyl; PTH, phenylthiohydantoin; AE, aminoethyl; and AEC, β-aminoethylcysteine.

[2] The stable emulsions that were observed are not entirely a consequence of the vigorous conditions of equilibration. Some globins give little or no emulsion.

became inconsequential after a few transfers. Examination of the phases at the end of the distribution and subsequent isolation of the products followed the procedure of Hill *et al.* (1962).

Reaction with Ethylenimine. The *S*-β-aminoethylcysteinyl (AEC) derivative was formed from ethylenimine (Raftery and Cole, 1963) except that the conditions of Crestfield *et al.* (1963) were used prior to the addition of ethylenimine to reduce any sulfhydryl groups that might have been oxidized during preceding operations.

For 300 mg of γ chain, four times the reagents recommended by Crestfield *et al.* (1963) for a 50-mg scale were added. A different experimental arrangement was used. The reaction flask had a side inlet through which nitrogen was passed and a stopper with a small opening through which the nitrogen was exhausted. After thorough flushing with nitrogen, the other reagents were added to the protein and solid urea in the flask. During the next few minutes, nitrogen was still passed through the flask during shaking to dissolve the protein. Then, nitrogen passage was stopped, the opening in the stopper was closed, and the solution was maintained at room temperature for 4 hr with occasional shaking. At the end of this time, 0.8 ml of ethylenimine (Matheson Coleman and Bell) was added while nitrogen was again passed over the solution. After reaction with ethylenimine for 1.5 hr at room temperature, about 50 ml of water was added, and the solution was thoroughly dialyzed against distilled water and finally lyophilized to dryness.

Oxidation Procedure. The oxidation procedure of Hirs (1956) was followed with minor modification only (Schroeder *et al.*, 1963).

Enzymatic Hydrolyses. Both oxidized and unoxidized bovine γ chains were hydrolyzed with trypsin and the β-aminoethyl derivative with chymotrypsin. These hydrolyses were made at room temperature under nitrogen for 24 hr at pH 8 using a pH-Stat (Radiometer). For the tryptic hydrolyses, trypsin equal to 0.5% of the weight of the substrate was added at 0, 4, and 10 or 16 hr. During the chymotryptic hydrolysis, enzyme equal to 1% of the substrate was added at 0, 2, and 6.5 hr. At the completion of a hydrolysis, the pH was reduced to 6.5, and any insoluble portion was removed by centrifugation. The supernatant solution was lyophilized to dryness.

Isolation of Peptides. The initial separation of peptides in enzymatic hydrolyses was achieved by chromatographing on Dowex 50-X2 with volatile pyridine–acetic acid developers. Further purification of peptides was carried out by rechromatography on Dowex 1-X2. The chromatographic procedures are described by Schroeder *et al.* (1962). Prior to the addition of the sample to a column of Dowex 50, the sample was dissolved in water and the pH was reduced to about 2 with hydrochloric acid; solution in pyridine–acetic acid as described earlier (Schroeder *et al.*, 1962) has been discontinued.

In a typical chromatogram on Dowex 50-X2, the soluble peptides from 200 mg of bovine γ chain were

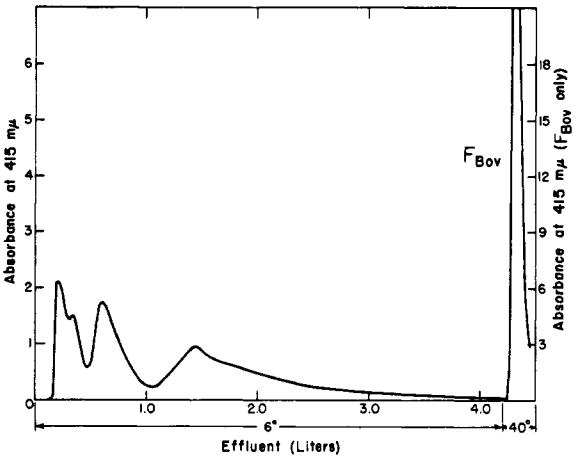

FIGURE 1: The isolation of hemoglobin F_{Bov} on Amberlite IRC-50 with developer no. 6b. Conditions: 1 g of total hemoglobin on a 3.5 × 30 cm column.

chromatographed on a 1 × 100 cm column at 38°. The gradient for developer was produced as described (Schroeder *et al.*, 1962) with 333 ml of pH 3.1 buffer in the mixer and 666 ml of pH 5.0 buffer in the reservoir. Fraction size was 1.2 ml, and from alternate fractions, 0.1 ml was taken for alkaline hydrolysis (Hirs *et al.*, 1956) and examination by ninhydrin. Zones from Dowex 50 were rechromatographed on 0.6 × 60 cm columns of Dowex 1-X2 (Babin *et al.*, 1964). In some of the chromatograms on Dowex 1, an improved gradient has been used (Schroeder and Robberson, 1965).

Edman PTH Procedure. The modified Edman PTH procedure (Schroeder *et al.*, 1963) has been relied on to arrange the amino acid residues in sequence within tryptic and chymotryptic peptides. The method has continued to give excellent results. In instances, where the sequence has been determined from both tryptic and chymotryptic peptides, the results of one have always confirmed the other.

Determination of Amino Acid Composition. Complete hydrolysis with hydrochloric acid and the determination of amino acid composition of polypeptide chains and peptides was carried out as described (Schroeder *et al.*, 1963).

Tryptophan was determined by the method of Noltmann *et al.* (1962) under modified chromatographic conditions. The sample, after drying by rotary evaporation, was dissolved in pH 2.2 buffer, and the pH was adjusted to 2.2 with acid. The appropriate volume of sample was applied to a 58-cm column of spherical resin, and analysis was begun with pH 3.25 buffer under the conditions of Benson and Patterson (1965). Because of the destruction of aspartic acid, serine, threonine, and glutamic acid during the alkaline hydrolysis, reaction of the effluent with ninhydrin and recording need not be started until shortly before proline emerges. The chromatogram was continued in the normal way until the base line was reached after leucine. At this time, the

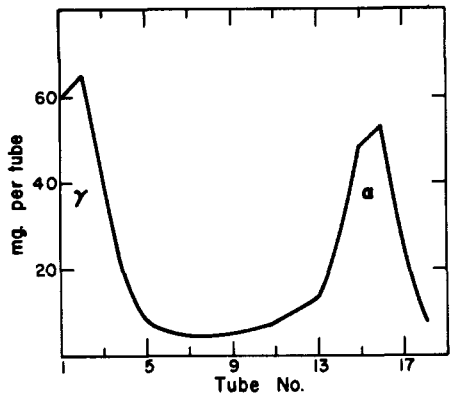

FIGURE 2: Countercurrent distribution of globin F_{Bov} after 18 transfers. Conditions: 500 mg of globin, 50 ml each phase, lower phase moved.

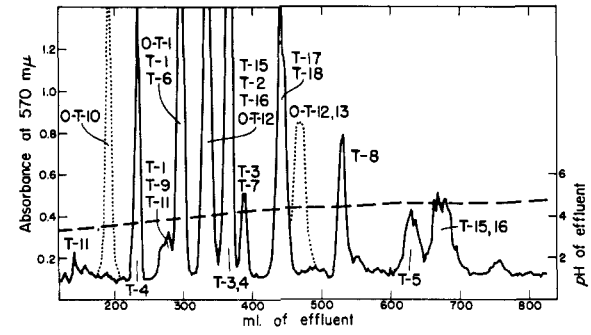

FIGURE 3: The separation of the tryptic peptides of the bovine γ chain on a 1 × 100 cm column of Dowex 50-X2 at 38°. – – – –, pH; . . . ., additional peaks in the hydrolysate of oxidized protein only. Alkaline hydrolysis preceded reaction with ninhydrin.

developer was changed to a solution of pH 5.28 buffer (for the short column) and 1-propanol (9:1 v/v). The change of developer takes effect after the base line has been reached following phenylalanine, and the tryptophan emerges as a well-defined peak somewhat less than 1 hr later. In such an alkaline hydrolysate, proline, valine, methionine, isoleucine plus alloisoleucine, leucine, phenylalanine, and tyrosine have been found to be in the same proportions as in an acidic hydrolysate and have been used as references to calculate the amount of tryptophan. This modified chromatographic procedure requires no more time than if both the short and long columns are used; it has the advantage that tryptophan and the reference amino acids are determined on the same chromatogram.

Results

Isolation of Bovine Hemoglobin F. The hemoglobin of a new-born calf contains several hemoglobin components as is true of the human infant and as Muller (1961) also found in the hemoglobin of a calf fetus. The use of developer no. 6b on IRC-50 columns gives an excellent separation of bovine adult and fetal hemoglobins and also allows for separation of some minor components from the major fetal hemoglobin. Figure 1 presents a typical chromatogram in which hemoglobin F_{Bov} was isolated.

Separation of the α and γ Chains. The separation of the chains of globin F_{Bov} by countercurrent distribution was excellent (Figure 2). The γ chain was isolated from tubes 1–5 inclusive and the α chain from tubes 12–18. A PTH degradation of the γ chain showed methionine as the sole N-terminal amino acid. In the subsequent investigation of tryptic and chymotryptic hydrolysates, there was no trace of peptides from the α chain. The bovine γ chains from countercurrent distribution were purer than the human γ chains from the chromatographic method of Wilson and Smith (1959) or the fractional precipitation method of Hayashi (1961).

Amino Acid Composition of the Bovine γ Chain. The amino acid composition of bovine γ chain which is presented in Table I compares satisfactorily with the composition as it is now known from the sequence:

TABLE I: Amino Acid Composition of the Bovine γ Chain.

Amino Acid	Residues/145-Residue Chain: 22-hr Hydrolysate	Residues/145-Residue Chain: 70-hr Hydrolysate	Residues in Tryptic Peptides
Lys	10.50	10.58	11
His	4.92	4.92	5
Arg	5.90	5.77	6
Asp	11.79	11.84	12
Thr[a]	2.88	2.66	3
Ser[a]	9.55	7.56	11
Glu	14.17	14.24	14
Pro	4.13	4.15	4
Gly	10.92	10.94	11
Ala	14.94	15.03	15
Cys/2[b]	1.65	0.92	2
Val	13.98	15.98	16
Met	0.79	0.75	1
Ileu	1.00	1.02	1
Leu	19.25	19.69	20
Tyr	1.90	1.71	2
Phe	9.73	9.71	10
Try		0.92, 0.92[c]	1
			145

[a] Extrapolation to zero time gives a value of 2.98 residues of threonine and 10.46 residues of serine. [b] The content of cysteic acid in an oxidized sample verified the presence of two Cys/2 residues. [c] From separate 64-hr alkaline hydrolysates.

TABLE II: Amino Acid Composition[a] of Tryptic Peptides of the Bovine γ Chain.

Amino Acid	O-T-1	T-2	T-3[b]	T-3, 4[c]	T-4	T-5	T-6	T-7[b]	T-8	T-9	O-T-10	T-11	O-T-12	O-T-12, 13	T-14	T-15	T-15, 16	T-16	T-17	T-18
Try						0.17														
Lys	0.96	1.00	0.98	1.01			1.03	0.98	1.02	1.00	1.10	1.00	1.01	1.03	0.09		1.02	1.00		
His									1.00				0.81	1.96					1.01	1.00
Arg				0.96	0.99	1.05								0.97	1.02	1.00	0.83		1.05	
$CySO_3H$											0.78		0.63	0.93						
Asp				0.99	1.00		2.93				1.18	1.97	0.95	2.86	1.05				0.98	
$MetSO_2$	0.90																			
Thr		0.93				0.97									0.11				0.93	
Ser	0.90	0.87					2.60				0.89		1.74	1.69	0.21		2.69	2.63		
Glu	2.10			2.10	2.00	1.07	1.13				1.18	1.06	1.37	2.11	0.24		4.19	4.15		
Pro						1.01	1.00							0.97	0.12		0.92	0.96		
Gly				2.99	2.95		2.01		1.00		0.93		0.98	1.08	1.28		1.19	1.01	1.05	
Ala	1.00	2.99		1.02	1.01		2.02		0.97		0.21		1.81	2.06	1.14		1.12	1.03	2.98	
Val		1.00	1.02	3.00	1.96	1.69		1.02			0.93			1.08	3.19				2.49	
Met																				
Ileu							0.99													
Leu	0.99	1.00		1.05	0.99	1.96	2.01				2.02	2.00	1.94	2.95	3.51		1.08	0.99	0.98	
Tyr						0.93														0.77
Phe		1.03					2.89				0.93		1.17	1.87	0.12		2.90	2.84		

[a] The data are in terms of residues per peptide. Data are omitted if less than 0.05 residue was present. [b] T-3 and T-7 are identical and, of course, are distinguishable only in terms of the total sequence. [c] This type of designation indicates that the bond between the component tryptic peptides is still intact.

TABLE III: Amino Acid Composition of Chymotryptic Peptides of the Bovine γ Chain.[a]

	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10	C-11	C-12	C-13	C-14
Try						0.78								
Lys		1.00	0.07	1.92							1.98	1.02		2.00
AEC														
His	0.11											0.98	0.91	
Arg					1.09		1.15							
Asp	0.18			0.99					2.00	1.03	1.02			0.28
Thr		0.93					0.88							
Ser	0.20	0.89	0.96					0.92	1.60	1.76				
Glu	0.15	2.03	0.13	2.05			0.98	1.04	0.26					
Pro						1.01					0.99			
Gly	0.06			2.00	0.94		0.21	0.07	1.14		0.99			0.99
Ala	0.14	3.01	0.13	1.99					2.01	2.00		1.01	1.10	0.29
Val		1.03		2.99		1.47			0.24		1.02	0.99		0.98
Met	0.39													
Ileu									0.92	0.98				
Leu	1.00	0.16	1.05	1.02	1.96		0.26	0.08	1.99	1.01			0.12	1.01
Tyr						0.99								
Phe			0.98				0.91	0.99						

[a] See footnote *a* of Table II. [b] This peptide was isolated as the DNP derivative. The analysis is that of the hydroly-

$Lys_{11}His_5Arg_6Asp_{12}Thr_3Ser_{11}Glu_{14}Pro_4Gly_{11}Ala_{15}(Cys/2)_2Val_{16}Met_1Ileu_1Leu_{20}Tyr_2Phe_{10}Try_1$.

Separation of Peptides. The separation of the tryptic peptides of the bovine γ chain by chromatography on Dowex 50 is shown in Figure 3. The 200–250 mg of tryptic hydrolysate that was chromatographed on a 1 × 100 cm column is about five times the load that has previously been used (Schroeder *et al.*, 1962). The separations have been entirely satisfactory at this load. The numbers attached to each zone in Figure 3 designate the peptides that were isolated from the zone by rechromatographing on Dowex 1-X2. In this and other figures and tables, the peptides are numbered consecutively from the N terminus. Thus, all peptides are identified with a distinguishing prefixed letter to designate the source of the peptide from a tryptic or chymotryptic hydrolysate, for example, T-1 or C-10. If the cysteine or methionine in the peptide has been oxidized or if the aminoethyl derivative of cysteine has been formed, the letters "O" or "AE" are also prefixed. The amino acid composition of the tryptic peptides is presented in Table II.

The chymotryptic peptides were isolated in an analogous manner. Figure 4 shows the separation on the Dowex 50-X2 column, and Table III records the amino acid composition of the chymotryptic peptides after chromatography on Dowex 1-X2 columns. These peptides contained generally a higher percentage of contaminating material than the tryptic peptides. No doubt this results from the greater complexity of the chymotryptic hydrolysate because more peptides are present due to partial cleavage at various bonds.

Amino Acid Sequence of Bovine γ Chain. Figure 5 shows the amino acid sequence of bovine γ chain. We shall dispense with a residue-by-residue discussion of the sequence as it is presented in this figure. The data there and that in Tables II and III are sufficient to show the evidence for the sequence of most of the chain. Thus, the numbers of the tryptic peptides (T-1, etc.) above the sequence and of the chymotryptic peptides (C-1, etc.) below the sequence identify these portions in Tables II and III and Figures 3 and 4. We shall not point out repeatedly that certain chymotryptic peptides, on the basis of amino acid composition alone, can derive only from the positions to which they are assigned; thus, C-6, C-11, C-19, C-21, etc. The arrows above the sequence indicate PTH results from tryptic peptides and arrows below represent results from chymotryptic peptides.[3] A dotted arrow designates a probable but less definite identification from the degradation. The beginning of a PTH degradation is marked >. In the discussion below, special comment will be made when these data are insufficient to establish a sequence.

A quick glance at Figure 5 shows that parts of the amino acid sequence have been degraded at least twice. The PTH procedure established the entire sequence of tryptic peptides T-1, T-2, T-3, T-7, T-8, and T-18.

[3] One of the referees has pointed out that this manner of representing the PTH results in no way permits the reader to judge the validity of the conclusions. The point is well taken. However, neither the authors nor the Editor have deemed it possible to reproduce the approximately 400 photographs that record the results of the approximately 200 degradations. The original photographs are available in the laboratory notebooks for inspection by any reader.

C-15	AE-C-16	C-17	C-18	C-19	AE-C-20	C-21	C-22	C-23[b]	C-24	C-25	C-26	C-27	C-28
2.13		2.11	2.00		1.05		0.36						0.97
	0.67				0.64								
					0.18	0.89							
							0.91		1.09	0.99			
0.90	0.13	1.87	1.93		0.99	2.00	0.13	1.02					0.97
								0.10					0.86
0.82				1.76				0.26			0.89	1.73	
0.28	1.00	2.13	1.06	1.02		1.05		0.26			1.02	2.00	1.00
						1.00		0.09				0.97	
1.21	1.00	2.15	1.05				0.23	0.42	0.25		1.00		1.05
	0.13	1.08	1.06	1.01			0.14	0.16	0.92			1.00	1.03
0.94						0.98	0.20	0.94					2.56
								0.07					
1.21	1.05	2.83	1.96	2.00	1.00		2.10	1.00	0.24			1.02	
0.84		0.87	0.95			0.98		0.10		1.01	1.00	0.96	

sate after extraction of the N-terminal DNP-amino acid.

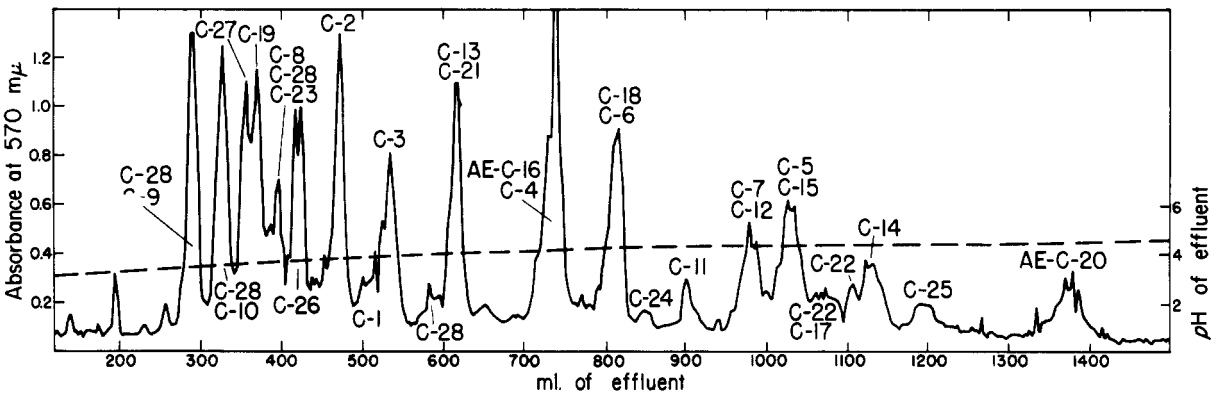

FIGURE 4: The separation of the chymotryptic peptides of the bovine γ chain on a 1.4 × 100 cm column of Dowex 50-X2 at 38°. – – – –, pH.

Let us consider the sequence beginning at the N terminus. Peptide T-1 must be placed at the N terminus of the bovine γ chain because the Edman degradation of the intact chain established the sequence "Met-Leu-Ser-Ala-Glu" as N terminal. The order of the first four tryptic peptides was conclusively established by C-2 and C-4. Because T-3 was only partially cleaved from T-4 during tryptic hydrolysis, T-3,4 could be isolated, and the ordering of the identical peptides T-3 and T-7 was greatly simplified; chymotryptic peptides C-4 and C-11 also show the existence of these identical peptides. The sequence of T-4 could not be completed with the PTH procedure alone. When a portion of T-4 was hydrolyzed with chymotrypsin, and the hydrolysate chromatographed on a Dowex 1-X2 column, the peptide Gly-Arg was isolated and established that detail of the sequence. The sequence Glu-Ala in T-4 is placed mainly by analogy to sequences in human β and γ chains.

The sequence of T-5 was determined by degrading not only the tryptic peptide as such but also C-7 whose unique composition relates it to T-5. Peptide C-5 joins T-4 to T-5. It is true that the "Leu-Leu" portion of C-5 may derive from the N terminus either of T-5 or of T-14. However, if T-14 were placed in the position of T-5, the great similarity between this chain, the γ and β chains of human hemoglobin, and the β chain of horse

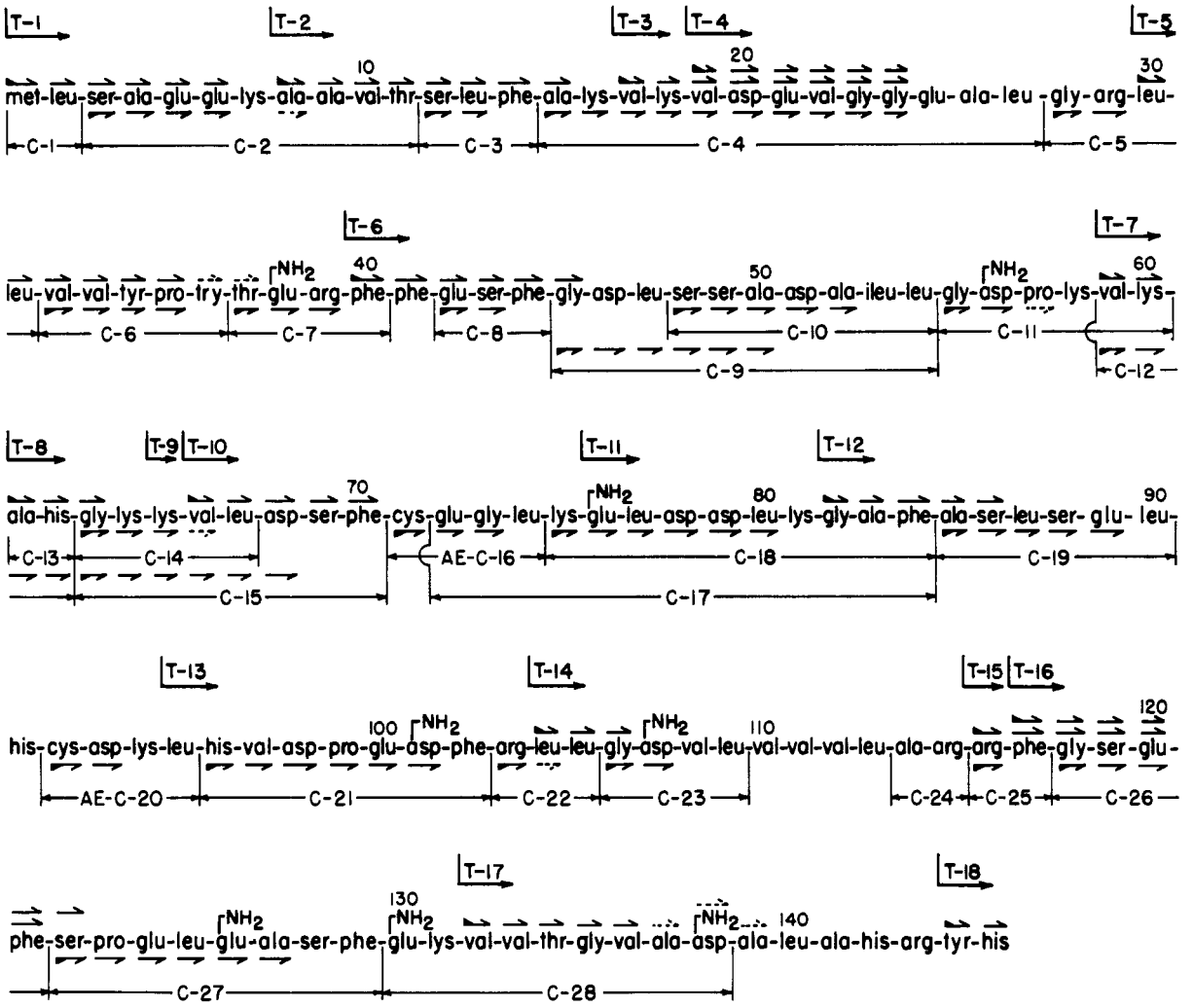

FIGURE 5: The amino acid sequence of the bovine γ chain. See text for details.

hemoglobin would be lost. Indeed, one could not reasonably place peptides T-5 to T-13 in what would become the order T-1 to T-4, T-14 to T-18, and T-5 to T-13.

Peptide T-6 contains the only isoleucyl residue of the bovine γ chain. The sequence of T-6 was obtained without problem except for the positioning of the isoleucyl and leucyl residues. This uncertainty was resolved by hydrolyzing a portion of T-6 with dilute acetic acid (Schroeder *et al.*, 1963) and isolating the peptide Ala-Ileu-Leu-Gly-AspNH_2-Pro-Lys from the hydrolysate on a Dowex 50-X2 column. The Edman degradation readily defined the position of the isoleucyl residue.

The data in Figure 5 with many overlaps of tryptic and chymotryptic peptides as well as the degradative results place peptides T-5 to T-10 in the positions shown. T-10 contains a cysteinyl residue and was not isolated from the soluble tryptic peptides of unoxidized bovine γ chain. However, when oxidized bovine γ chain was hydrolyzed with trypsin, T-10 was isolated in good yield. The Edman procedure gave the sequence of five residues as shown. After a portion of T-10 had been hydrolyzed with dilute acetic acid, Ser-Phe-$CySO_3H$-Glu-Gly-Leu-Lys was isolated and degraded through five cycles. Peptide AE-C-16 from the chymotryptic hydrolysis of aminoethylated γ chains was degraded without difficulty. Peptide C-17 suggests an unexpected cleavage C-terminal to aminoethylcysteine by chymotrypsin. The analysis of C-17 gave no evidence of aminoethylcysteine, and there is no reason to expect complete destruction of this amino acid during hydrolysis (compare the analyses of peptides AE-C-16 and AE-C-20). The PTH procedure gave unsatisfactory results with C-17. This is attributed to the N-terminal glutamic acid of the peptide because peptides with N-terminal glutamic acid occasionally give unsatisfactory results possibly due to cyclization to pyrrolidonecarboxylic acid.

When the Edman degradation of T-11 failed to yield a detectable PTH-amino acid, it was suspected that an N-terminal glutaminyl residue had cyclyzed to pyrroli-

donecarboxylic acid during chromatography. It will be noted that T-11 was isolated from widely separated zones (Figure 3) as would be expected if partial cyclization had occurred prior or during chromatography. C-28 (Figure 4) shows the same behavior. The conclusion was substantiated by the degradation of peptide C-18.

Peptide T-12 contains the second cysteinyl residue in the bovine γ chain and, like T-10, was isolated only from a tryptic hydrolysate of the oxidized γ chain. The tryptic cleavage between peptides T-12 and T-13 was slight; the major peptide was O-T-12,13, and O-T-12 was isolated in small yield while T-13 was not found at all. The Edman procedure gave the sequence of O-T-12,13 through five amino acids, and additional sequences were obtained from chymotryptic peptides C-19, AE-C-20, and C-21. The data of Figure 5 allow no other arrangement of peptides T-10 through T-13.

T-14 comprises the major portion of the so-called insoluble "core" after tryptic hydrolysis. Attempts to isolate the pure peptide have been only partially successful, and, therefore, the sequence of this section is based on less definitive data than that of the other parts. In the description below of the investigation of this peptide, unsuccessful experiments will be presented briefly in the hope that the information will be of some value to others. At the completion of a tryptic digest at pH 8, only a slight amount of insoluble material was apparent in the hydrolysate, and the bulk of the core precipitated when the pH was lowered to 6.5. Because the initial developer for Dowex 1 chromatograms has a pH higher than 8, it was thought that if a tryptic hydrolysate were placed directly on a Dowex 1 column, T-14 might be successfully chromatographed. This idea was tested with a 10-hr tryptic hydrolysate of AE-bovine γ chains which, after lyophilization, was dissolved and chromatographed on Dowex 1. The chromatogram is shown in Figure 3 of Schroeder and Robberson (1965). It proved to be impossible to detect T-14 in any of the peaks. Likewise, the core alone under similar chromatographic conditions behaved anomalously. As an apparently basic peptide, T-14 should emerge during the first part of the chromatogram. Actually, it emerged (without obvious purification as evidenced by amino acid analysis) in poor yield late in the chromatogram as though it had merely been dissolved by the increased acetic acid in the developer. When the chromatographic procedure of Goldstein *et al.* (1961) for the purification of the human β-chain core was applied to the bovine γ-chain core, it failed (as it does also for the human γ-chain core) and a small amount of material was recovered only when 8.5 M pyridine developer (Konigsberg and Hill, 1962) was used.[4] The most complete purification was achieved when the samples that were recovered from the above two experiments were passed through a 2.2 × 130 cm column of Bio-Gel P-10 (Bio-Rad Laboratories, Richmond, Calif.) in 50% acetic acid.[5] One of the zones from this experiment gave the analysis for T-14 in Table II. A 70-hr hydrolysate had a valine content equivalent to 3.94 residues. Although the contamination is undesirably high, the composition of T-14 appears to be (Arg, Asp, Gly, Ala, Val_4, Leu_4) and is identical with the remaining residues when the total amino acid composition of the other tryptic peptides is subtracted from the amino acid composition of the bovine γ chain.

Another preparation was made in which the core was redigested with trypsin but not chromatographed on Dowex 1 or Dowex 50 before passage through the Bio-Gel column. The product actually was somewhat more contaminated, and the analysis suggested that the contaminants were about 30% T-5 and 5% T-16. When a portion of this sample was subjected to the Edman degradation, the sequence was Leu-Leu-Gly but some valine in the third position again suggested T-5 as the main contaminant. Peptides C-22, C-23, and C-24 are assumed to derive from T-14. The sequence Val-Val-Val-Leu is assigned to T-14 on the basis of the amino acid composition of T-14 and the specificity of chymotrypsin. We have remarked above that the placement of T-5 and T-14 is not unequivocal on the basis of the data.

Peptide T-15, which is a single arginyl residue, was isolated both as free arginine and in combination with T-16. Peptides T-15,16, T-16, and C-27 established most of the sequence of T-16. The sequence Ser-Phe in C-27 is based on the specificity of chymotrypsin. The placing of $GluNH_2$-Lys in this sequence involves the specificity of trypsin. That a glutaminyl residue rather than a glutamyl residue is present in C-28 is based on indirect evidence. As in peptide T-11, no N-terminal residue could be detected when C-28 was subjected to the Edman procedure. By analogy to the human β and γ chains, a glutaminyl residue might be expected in this position.

The sequence of T-17 was completed by the isolation of Ala-Leu and Ala-His-Arg from a chymotryptic hydrolysate of T-17. Peptide T-18 was readily shown to be Tyr-His. Peptide C-28 clearly links T-16 and T-17, but no chymotryptic peptide was found to provide an overlap between T-17 and T-18. Because T-18 is not C terminal in lysine or arginine, it is logically placed at the C-terminal end of the chain, the more so because human β and γ chains and horse β chains have such a sequence at the C terminus.

Discussion

Methodology of the Investigation. The methods that

[4] When the core was chromatographed on Dowex 50 with commonly used developers (Schroeder *et al.*, 1962) to which 1-propanol had been added, nothing was eluted. 1-Propanol appears to be more effective in altering chromatographic behavior under these conditions when aromatic amino acids are part of the peptide.

[5] The sample was dissolved by suspending in glacial acetic acid and then adding water drop by drop to the appropriate volume. Solution in 50% acetic acid itself was slow and uncertain.

have been used in the determination of this sequence have, for the most part, differed little from those that were used in the study of the γ chain of human hemoglobin F. Only toward the end of the investigation was an improved gradient for the chromatography of peptides applied (Schroeder and Robberson, 1965). The obvious contamination that is present in some of the chymotryptic peptides probably could have been reduced by rechromatographing after Dowex 1 chromatography again on Dowex 50 with a different gradient (Schroeder, 1966). The difficulty in purifying peptide T-14 has required the investigation of several systems. Because of the insolubility and intractability of this peptide as well as the fact that the contaminants were of similar size and behavior, a completely satisfactory method of purification has not been devised.

The Validity of the Sequence. Because the complete sequences of the human β and γ chains are known and show obviously homologous structure, it might be expected to be and, indeed, is relatively simple to fit the tryptic peptides of the bovine γ chain into proper sequence once their amino acid composition is known. This type of sequence determination is somewhat less definite in the case of the bovine γ chain than it is for human hemoglobin variants. If we compare the potential tryptic cleavages of the bovine γ chain with the human γ chain, two more would be expected: when we consider the final sequence of the bovine γ chain, it will be observed that not only are there the additional expected cleavages but also that alteration in position of some of the other lysyl and arginyl residues relative to the human γ chain has altered the length of some of the tryptic peptides.

One immediately obvious feature of the sequence is its length. With 145 residues, it is one residue shorter than the human γ and β chains, the horse β chain (Smith, 1964), the camel β chain (Banerjee and Bhown, 1965), and probably some primate β chains (Bradshaw *et al.*, 1965). It is apparent that the shortening occurs near the N terminus of the bovine γ chain because peptide T-1 contains only seven residues in contrast to eight for the β and γ chains mentioned above as well as for pig, llama, and rabbit β chains on the basis of the amino acid composition of the tryptic peptides (Braunitzer *et al.*, 1964). However, it is probable that the β chains of sheep hemoglobins A and B likewise have only seven residues in peptide T-1 (T. H. J. Huisman, 1965, personal communication). Peptide T-1 of the bovine γ chain does not contain the histidyl residue that is present in the second position of the human β and γ chains. Braunitzer (1958) reported a sequence Met-Leu from adult bovine hemoglobin which is in contradiction to the Met-Gly sequence described by Ozawa and Satake (1955) and to the sequence Met-His-Leu-Ser-Pro-Glu-Glu-Lys as given for the bovine β chains by Sasakawa (1963). In order to resolve this uncertainty and to determine whether the bovine γ chain alone was shorter, an investigation of the β chains of adult bovine hemoglobins A and B was undertaken; in both instances, the sequence of the N-terminal portion was Met-Leu-Thr-Ala-Glu-Glu-Lys,[6] and was very similar to that of the bovine γ chain (threonine for serine in the third position). The results are in accord with those of Braunitzer.

The data that have been presented above on the tryptic and chymotryptic peptides of the bovine γ chain, in general, provide good overlaps to establish the sequence as it is presented in Figure 5. In only a few instances is the evidence indirect for short sequences within tryptic peptides; for example, the sequence of residues 25 and 26, 110–113, and 128–130. Frequently, the sequence is substantiated by PTH degradation of both tryptic and chymotryptic peptides. The most tenuous portion of the sequence, of course, involves peptide T-14 which has never been obtained in a satisfactorily pure state. It is of interest that Braunitzer *et al.* (1964) report a peptide of identical composition from the β chain of llama hemoglobin. The chymotryptic overlaps do not unequivocally place peptides T-5 and T-14 because both have an N-terminal Leu-Leu sequence. Were they interchanged in position, however, it would be difficult to propose a satisfactory sequence. Finally, it should be pointed out that T-18 is placed at the C-terminus because of analogy and the absence of basic amino acids inasmuch as no overlap to T-17 was isolated. Despite some shortcomings in the data, we believe that the sequence of the bovine γ chain as presented in Figure 5 is probably correct.

The Sulfhydryl Groups. The different arrangement of the sulfhydryl groups in the bovine γ chains is one of the more interesting results of the determination of the sequence. A cysteinyl residue at position 92 of the bovine γ chain is equivalent (because of the shorter chain length) to the cysteinyl residue at position 93 in the human β and γ chains. The sulfhydryl group in residue 93 is the so-called "reactive" sulfhydryl because it reacts, for example, with iodoacetamide in undenatured human hemoglobins A (Goldstein *et al.*, 1961) and F (Schroeder *et al.*, 1963), whereas the sulfhydryl groups at residue 104 in the human α chains and at 112 in the human β chains do not (the single sulfhydryl group of the human γ chain is in residue 93). The second cysteinyl residue of the bovine γ chain is in residue 71 (equivalent to 72 in human β and γ chains). An examination of the three-dimensional structure of hemoglobin as presented by Cullis *et al.* (1962) suggests that residue 71 of the bovine γ chain is in helix E on the exposed surface of the molecule. In contrast, the two unreactive sulfhydryl groups of human hemoglobin A are buried in the interior of the undissociated molecule. Cecil and Thomas (1965) have suggested that the unreactivity of the sulfhydryl groups is a result of hydrophobic interaction between the group itself and other groups of the same or another chain. In undenatured bovine hemoglobin F, both sulfhydryl groups of the γ chain react with iodoacetamide (Schroeder and Hartzman, 1965, unpublished). Whether this reactivity is simply a reflection of greater accessibility or of an environment in which hydrophobic interactions do not occur easily is

[6] Further sequences of the bovine β^A and β^B chains will be the subject of another communication.

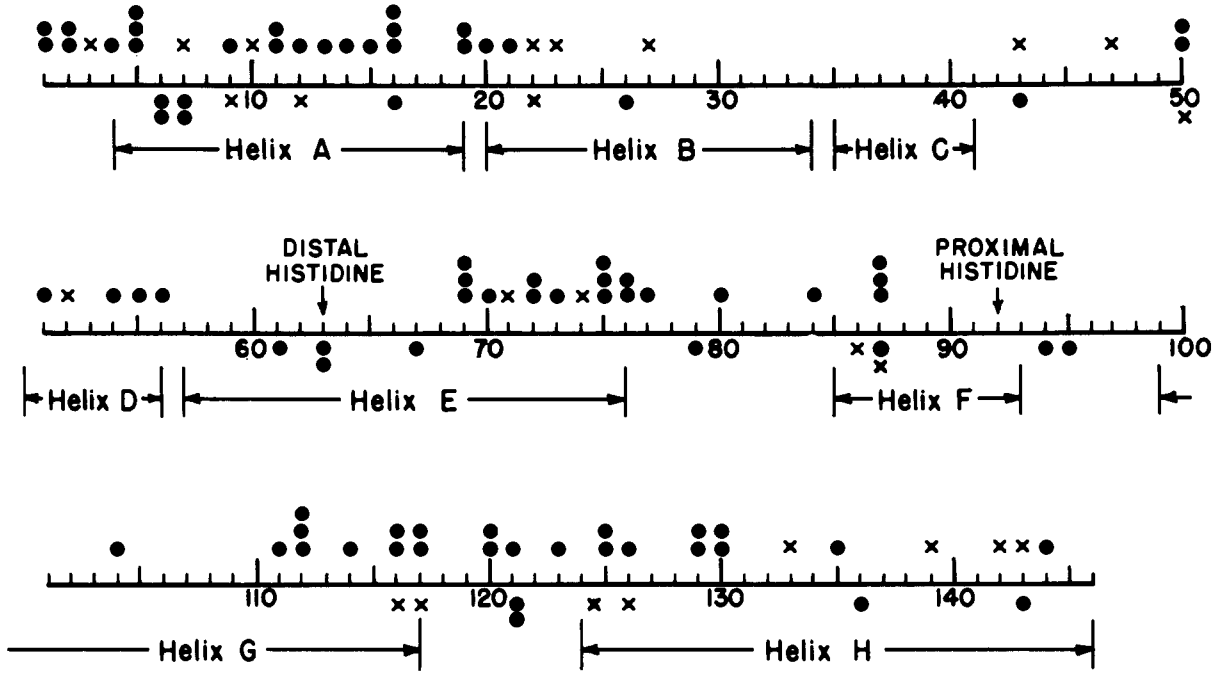

FIGURE 6: A comparison of the positions of differing sequences in the human β and γ; the bovine β^A β^B and γ; and the horse β chains. See text for explanation.

unknown. It should be mentioned that the human β chain alone, of the six chains that will be discussed below, has a cysteinyl residue at position 112. Cecil and Thomas (1965) may well be correct in their conclusion that the sulfhydryl group may not be an invariably active group in a protein as has so often been suggested.

Comparison with Other Sequences and a Correlation of Sequence and Function. By actual count, the bovine γ chain differs from the human γ chain in amino acid sequence at 40 sites and from the human β chain at 32 sites. Such numerology is of limited meaning because it tells not at all whether the differing residues are as similar as leucine and valine or as diverse as lysine and aspartic acid. Actually, comparison can also be made with the horse β chains (Smith, 1964) and the β chains of bovine hemoglobins A and B about which a preliminary report has been made (Schroeder and Jones, 1965). Approximately three-fourths of the sequence of these three chains has been established. At the present state of knowledge, a residue-by-residue comparison is pointless but a broad consideration of these six chains from six hemoglobins of rather different properties points up some meaningful correlations. There is such constancy of sequence in certain sections of the chains that they would appear to be of critical importance in the proper functioning of hemoglobin as an oxygen carrier.

In order to aid in the comparison and correlation, Figure 6 has been prepared. In this figure, the position of each residue is indicated along the line.[7] The dots above the line designate the number of observed differences among the six chains: thus, no dot, no difference; one dot, one difference (which could mean that three chains have one type of residue and the other three a second type); and so on. For example, at residue 5 no less than four kinds have been observed. An "x" above the line denotes positions at which the human γ chain alone has been observed to vary. The dots below the line show those residues where alterations occur in the human hemoglobins that are β chain variants, whereas an "x" below the line refers to the human δ chain (see Schroeder and Jones, 1965, for a compilation). The residues in helical configuration are also indicated. The placement of residues in helices is slightly different from that of Cullis *et al.* (1962) and is based upon the data of Edmundson (1965) for sperm whale myoglobin (Schroeder and Jones, 1965).

Examination of a three-dimensional model of hemoglobin is of value in understanding the following discussion. If none is available, Figure 23 of Cullis *et al.* (1962) will be helpful or the reader may wish to construct a paper model (Schroeder and Jones, 1965). In Figure 23 (Cullis *et al.*, 1962) the residue designated as "gly 56" should be "gly 46."

In the first 23 residues from the N terminus, 19 of the positions are occupied by different residues in one or the other chain. Some are "conservative"[8] (serine for

[7] The counting is based on the 146 residues of the human chains. Thus, the N terminus of the bovine γ chain in Figure 5 would be at residue 2.

[8] Terms such as these have been used in the sense of similarity in structure by Margoliash and Smith (1965) for similar comparisons.

threonine), whereas others are quite "radical" (alanine, glycine, glutamic acid, and proline all occupy position 5). All, however, have lysyl residues at positions 8 and 17. Nevertheless, considerable variation appears to be possible within these 23 residues.

Between residues 22 and 49 inclusive, five of the chains are identical. The human γ chain alone differs at positions 22, 23, 27, 43, and 47 with relatively conservative differences. In three-dimensional structure, these 28 residues comprise most of helix B, and all of helix C and the interhelical region CD. Among the abnormal human hemoglobins, hemoglobin E in position 26 and hemoglobin $G_{Galveston}$ in position 43 are altered in this part of the sequence. Helix D (residues 50–56) appears to be subject to considerable variation.

From residues 57 to 68 the six chains have the same sequence, and on the basis of amino acid composition alone, other animal β chains may be identical with the exception of llama (Braunitzer *et al.*, 1964). Within this sequence at residue 63 is the so-called "distal" histidine which is in the general site of the sixth coordination position of the heme iron. The importance of position 63 in the functioning of hemoglobin is clearly attested by the various hemoglobins M which are among the few abnormal hemoglobins to show some definite physiological manifestation. The alteration in residue 61 that is observed in hemoglobin Hikari apparently is benign. This section is mainly the N-terminal half of helix E.

Although every residue between 69 and 77 inclusive is different in one or the other of the six chains, minor variation only occurs between 78 and 86. Noteworthy is the fact that four kinds of residues have been detected at position 87, whereas from 88 to 110 there is no alteration in the six chains except interchange of lysine and arginine at residue 104. In spatial terms, 69–76 (variant) lie in the C-terminal portion of helix E toward the N-terminal end of which the distal histidine at position 63 is located. Residues 77–86 (invariant) are mainly in the interhelical region EF, whereas 87 (variant) is about a turn and a half in helix F from the "proximal" histidine at residue 92, but the interhelical region FG and the N-terminal two-thirds of helix G (residues 88–110) are identical except for the interchange of basic amino acids in position 104. The potential significance of these data may be more apparent if we consider next the sequence from residues 131 to 146 which comprise the C-terminal two-thirds of helix H. We note at residue 135 either threonine or alanine and at residue 144 either lysine or arginine but, in four other instances, the human γ chain alone is different; thus, residues 131–146 are relatively invariant.

The amino acid compositions are known for the tryptic peptides of the β chains of llama and pig hemoglobin (Braunitzer *et al.*, 1964) and of sheep hemoglobins A, B, and C (T. H. J. Huisman, 1965, personal communication). If sequences for these chains are arranged with minimum differences, the above conclusions about variant and invariant sections are generally unaltered. The alterations in the variant human hemoglobins, in general, as well as in the human δ chain also follow much the same pattern.

It is not without interest to note that in 15 positions the human γ chain alone of the six chains is responsible for a variation. If it were not included in the comparison, the sequences from 22 to 49 and from 131 to 146 are much more invariant.

Consider then that the most invariant portions of sequence as deduced from these six chains are (1) helices B and C and interhelical CD, (2) the N-terminal half of helix E, (3) interhelical EF and helix F except for one residue, (4) interhelical FG and the N-terminal two-thirds of helix G, and (5) the C-terminal two-thirds of helix H.

The plane of the heme group lies between helix E on one side and helices F and B on the other. But if helices E, F, and B form two sides of the crevice which surrounds the heme group, then the N-terminal two-thirds of helix G and the C-terminal two-thirds of helix H enclose the bottom of the crevice while the interhelical EF at one end and helix C and part of interhelical CD at the other end block the two ends of the crevice. Apparently only minimal alteration is acceptable in the immediate vicinity of the heme group. If the sequences in the α chains of human, horse, and bovine hemoglobins are examined, much the same conclusions can be drawn although, of course, the sequences are somewhat different. The environment of the heme group, thus, may require a constancy greater than that of those residues in immediate contact with the heme group (Kendrew, 1962).

What is the relation and possible significance of these "invariant" and "variable" portions of the β and γ chains to the three-dimensional structure and function of the whole molecule? It should recalled that the distances between iron atoms in the hemoglobin molecule are not identical and are altered by oxygenation and deoxygenation (Muirhead and Perutz, 1963). If each iron atom is assigned to an α or β chain and designated as α_1, α_2, β_1, and β_2, the various distances would be α_1–α_2, α_1–β_1, α_1–β_2, and β_1–β_2. For this discussion, α_1–β_1 and α_1–β_2 are important, and α_1–β_1 is assigned to the shorter distance. The heme groups involved in the distance α_1–β_1 have roughly parallel planes, their edges are relatively close together, and, in the volume between, lie several of the "invariant" sequences both in the α and β or γ chains although the contact between chains is relatively limited. (See Figure 10 of Cullis *et al.*, 1962.) Furthermore, this distance is unaltered by oxygenation or deoxygenation. On the other hand, the other hemes (α_1–β_2) have more or less roughly perpendicular planes and are separated by a considerable mass (equivalent to about two α helices). The chains have considerable areas of contact which involve the variable parts of helices G and H and the interhelical GH (residues 111–130) about which nothing has been said above.

Although they are difficult to visualize without a careful study of a three-dimensional model, the following relationships seem to hold. The "invariant" sections enclose the heme group, are present between the closer heme groups of different chains, and make points of contact between identical chains (parts of the G and H helices). The "variable" sections usually are more

distant from the heme group, participate in larger areas of contact between unlike chains, and are more on the surface of the molecule (actually some "invariant" portions are on the surface but more of the variable portion is).

If the invariant sections are most intimately involved with the primary function of hemoglobin, the variable portions presumably are responsible for properties in which the various hemoglobins differ, oxygen affinity, rate of alkali denaturation, etc. Two variable parts (the C-terminal half of helix E as well as residue 87 in helix F) are in close proximity to the heme and, in some way as yet impossible to speculate, may be especially important in altering oxygen affinity or other properties. But the three-dimensional model shows that other variable sections despite their separation along the chain in the primary sequence are relatively close in space. On one side of variable helix A and more or less parallel to it lies the variable half of helix E while on the other side are those variable residues in helices G and H and interhelical GH. The possibility is clear that the varied residues because of their proximity can interact to produce the altered properties.

Such considerations do not tell us how the apparently constant portions of the sequence confer upon the molecule its particular properties nor how the variation influences the properties. Yet, as more sequence data become available and more structural data are procured, an understanding of the function of hemoglobin in relation to its structure may indeed become a reality.

Acknowledgment

It is a pleasure to express our thanks to Mr. Ed Stueve of the Alta-Dena Dairy, City of Industry, Calif., for obtaining for us the calf blood which was used in this investigation.

References

Allen, D. W., Schroeder W. A., and Balog, J. (1958), *J. Am. Chem. Soc. 80*, 1628.

Babin, D. R., Jones, R. T., and Schroeder, W. A. (1964), *Biochim. Biophys. Acta 86, 136.*

Banerjee, S., and Bhown, A. S. (1965), *Biochim. Biophys. Acta 100*, 503.

Bangham, A. D. (1957), *Nature 179*, 467.

Benson, J. V., Jr., and Patterson, J. A. (1965), *Anal. Chem. 37*, 1108.

Bradshaw, R. A., Rogers, L. A., Hill, R. L., and Buettner-Janusch, J. (1965), *Arch. Biochem. Biophys. 109*, 571.

Braunitzer, G. (1958), *Z. Physiol. Chem. 312*, 72.

Braunitzer, G., Hilse, K., Rudloff, V., and Hilschmann, N. (1964), *Advan. Protein Chem. 19*, 1.

Brinkman, R., and Jonxis, J. H. P. (1936), *J. Physiol. 88*, 162.

Cecil, R., and Thomas, M. A. W. (1965), *Nature 206*, 1317.

Clegg, M. D., and Schroeder, W. A. (1959), *J. Am. Chem. Soc. 81*, 6065.

Crestfield, A. M., Moore, S., and Stein, W. H. (1963), *J. Biol. Chem. 238*, 622.

Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G., and North, A. C. T. (1962), *Proc. Roy. Soc. (London) A265*, 161.

Edmundson, A. B. (1965), *Nature 205*, 883.

Goldstein, J., Guidotti, G., Konigsberg, W., and Hill, R. J. (1961), *J. Biol. Chem. 236*, PC77.

Haurowitz, F., Hardin, R. L., and Dicks, M. (1954), *J. Phys. Chem. 58*, 103.

Hayashi, H. (1961), *J. Biochem. (Tokyo) 50*, 70.

Hill, R. J., Konigsberg, W., Guidotti, G., and Craig, L. C. (1962), *J. Biol. Chem. 237*, 1549.

Hirs, C. H. W. (1956), *J. Biol. Chem. 219*, 611.

Hirs, C. H. W., Moore, S., and Stein, W. H. (1956), *J. Biol. Chem. 219*, 623.

Jones, R. T. (1961), Ph.D. Thesis, California Institute of Technology.

Jonxis, J. H. P., and Huisman, T. H. J. (1956), *Blood 11*, 1009.

Kendrew, J. (1962), *Brookhaven Symp. Biol. 15*, 216.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. W., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960), *Nature 185*, 422.

Konigsberg, W., and Hill, R. J. (1962), *J. Biol. Chem. 237*, 2547.

Margoliash, E., and Smith, E. L. (1965), *in* Evolving Genes and Proteins, V. Bryson and H. J. Vogel, Ed., New York, N. Y., Academic Press Inc., 1965, p 221.

Matsuda, G., Gehring-Mueller, R., and Braunitzer, G. (1963), *Biochem. Z. 338*, 669.

Muirhead, H., and Perutz, M. F. (1963), *Nature 199*, 633.

Muller, C. J. (1961), Molecular Evolution. A Comparative Study on the Structure of Mammalian and Avian Hemoglobins, The Netherlands, Van Gorcum, Assen.

Noltmann, E. A., Mahowald, T. A., and Kuby, S. A. (1962), *J. Biol. Chem. 237*, 1146.

Ozawa, H., and Satake, K. (1955), *J. Biochem. (Tokyo) 42*, 641.

Raftery, M. A., and Cole, R. D. (1963), *Biochem. Biophys. Res. Commun. 10*, 467.

Sasakawa, S. (1963), *J. Biochem. (Tokyo) 53*, 188.

Satake, K., and Sasakawa, S. (1963), *J. Biochem. (Tokyo) 53*, 201.

Schnek, A. G., and Schroeder, W. A. (1961), *J. Am. Chem. Soc. 83*, 1472.

Schroeder, W. A. (1966), *Methods Enzymol. 7* (in press).

Schroeder, W. A., and Jones, R. T. (1965), *Fortschr. Chem. Org. Naturstoffe 23*, 113.

Schroeder, W. A., Jones, R. T., Cormick, J., and McCalla, K. (1962), *Anal. Chem. 34*, 1570.

Schroeder, W. A., and Robberson, B. (1965), *Anal. Chem. 37*, 1583.

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Cormick, J., and Jones, R. T. (1963), *Biochemistry 2*, 992.

Smith, D. B. (1964), *Can. J. Biochem. Physiol. 42*,

755.
van der Helm, H. J. (1958), Doctoral Thesis, State University of Groningen, The Netherlands, p 33 (communicated by T. H. J. Huisman).
Vinograd, J., and Hutchinson, W. D. (1960), *Nature 187*, 216.
Wilson, S., and Smith, D. B. (1959), *Can. J. Biochem. Physiol. 37*, 405.

The Optical Rotatory Dispersion of Ribosomes and Their Constituents*

P. McPhie and W. B. Gratzer

ABSTRACT: Optical rotatory dispersion measurements have been performed on solutions of ribosomes from yeast, *Escherichia coli*, and rabbit reticulocytes. The ultraviolet Cotton effect of ribonucleic acid (RNA) is observed in the same position as in the free molecule, and an additional feature is present in the 233-mμ region, which arises from the presence of α-helical structure in the ribosomal protein. All the ribosomes obey a single-term Drude law toward longer wavelengths. Dissociation of yeast ribosomes, and their disorganization by high concentrations of chelating agent, are found to have essentially no effect on their optical rotatory dispersion. By subtracting the contribution of the RNA, the optical rotatory dispersion curve of the protein *in situ* is obtained, and values of $[m']_{233}$ and the Moffitt constant, b_0, are reported. The helix content of the ribosomal protein *in situ* is some 60% of the maximal value attained by the extracted protein in 2-chloroethanol. The extracted proteins in aqueous solution contain no detectable proportion of α-helix. The nature of the specificity of the protein–RNA complex in the ribosome is discussed. The optical rotatory dispersion curves of the three species of ribosome are quantitatively similar, except insofar as they reflect the differences in RNA:protein ratio. Analytical applications are suggested.

The ribosome is a compact structure, made up of ribonucleic acid (RNA) and a seemingly wide range of basic proteins (*e.g.* Waller, 1964; Leboy *et al.*, 1964). It is highly specific in nature, and has eluded attempts at reconstitution from the dissociated nucleic acid and protein components. The native form is stabilized by poorly understood factors, involving among other things magnesium ions and probably diamines (Petermann, 1964). The RNA has been shown, at least in the case of *Escherichia coli* ribosomes (Schlessinger, 1960), to have the same degree of helical structure in the ribosome as in the free state. The conformation of the proteins has not however been hitherto examined. As a first step toward the investigation of the morphology of the ribosome, and the nature and specificity of the relationship between the protein and RNA, we have examined the optical rotatory dispersion under varying conditions of three types of ribosomes of widely different RNA–protein ratio. We have related the results to the characteristics of the RNA and proteins alone. These findings are described below.

Experimental Section

The greater number of our experiments were performed on ribosomes from yeast, *Saccharomyces fragilis*. The ribosomes were prepared in the manner previously described (McPhie *et al.*, 1966). The RNA was extracted with phenol (Tissières *et al.*, 1959), and three times precipitated with ethanol. All experiments with the intact ribosomes were performed in a buffer of 1 mM Tris, 2 mM magnesium chloride, 0.1 M potassium chloride, pH 7.2. All preparations were screened by sedimentation in the analytical ultracentrifuge (Spinco Model E). Ribosome concentrations were measured spectrophotometrically, taking the specific absorptivity as $E^{1\%}_{1cm} = 113$ (Yin, 1960). The molar residue absorptivity of the RNA was determined by phosphorus analysis by the procedure of Jones *et al.* (1951); it was found to be $\epsilon(P) = 7400$. For a mean residue weight of 317 (from the base composition given by Maeda, 1960), neglecting counterions, this leads to $E^{1\%}_{1cm} = 233$ for the RNA.

The proteins were prepared by extraction with acetic acid (Waller, 1961), with guanidinium hydrochloride (Cox and Arnstein, 1961), and with lithium chloride (Curry and Hersh, 1962). The first of these methods was found to give the best yield of protein and the most stable preparations, and was routinely used. The acetic acid extract was dialyzed at 4° into dilute acetic acid, pH 4, adjusted to pH 4.5 with alkali, and passed through a column of Sephadex G-75 to remove nucleotides, which

* From the Department of Biophysics, King's College, Drury Lane, London W.C.2., England. *Received November 23, 1965.*