Partial Amino Acid Sequences of Chicken and Turkey Immunoglobulin Light Chains. Homology with Mammalian λ Chains*

J. Andrew Grant,† Bob Sanders,‡ and Leroy Hood§

ABSTRACT: The light chains of normal (heterogeneous) immunoglobulins were isolated from individual chickens and turkeys. A new approach is described for determining the number and approximate ratio of light-chain types present in these species. More than 95% of the chicken and turkey light chains are of a single type that appear to be homologous to mammalian λ chains. On the basis of this homology and the presence of both κ and λ chains in mammals it seems likely that κ and λ chains diverged more than 250 million years ago—prior to the separation of the mammalian and avian evolutionary lines. Peptide studies and Edman degradations of intact normal light chains indicate in both species that there is a single predominant variable region sequence. About one-third of the chicken light-chain sequence has been determined. The partial sequences of peptides isolated from turkey light chains indicate that they differ from the corresponding peptides in chicken light chains in only 2 of a total of 74 residues.

Remarkable advances have been made toward an understanding of the evolution of immunoglobulins through the amino acid sequence analysis of homogeneous immunoglobulins (Putnam, 1969; Hood and Talmage, 1970; Hood *et al.*, 1970b) and through the study of normal (heterogeneous) immunoglobulins (Hood *et al.*, 1967; Hood *et al.*, 1970a,b; Milstein *et al.*, 1969; Grey, 1969). Such studies have shown that the immunoglobulin molecule is composed of light (~23,000) and heavy (~50,000–75,000) chains. Detailed sequence analysis of immunoglobulin chains has been confined primarily to mammals most of which have two types of light chains, λ and κ. The ratio of light chain types can vary widely in different species from >95% κ (mouse) to >95% λ (horse) (Hood *et al.*, 1967, 1970b). Light chains of both types can be divided into a variable (V) region (~residues 1–107) and a constant (C) region (~residues 108–214).[1] Each of these regions seems to be encoded by separate germ-line genes (Hood *et al.*, 1967; Milstein, 1967).

Preliminary studies on chicken light chains (Hood *et al.*, 1967) suggested that they are predominately of the κ type in that a 60% yield of alanine was obtained at the N terminus and the predominant C-terminal residue was half-cystine. Most mammalian κ chains have a free N-terminal α-amino group (in contrast to their λ counterparts which have predominately a blocked pyrrolidonecarboxylic acid residue) and a C-terminal half-cystine residue (most mammalian λ chains have C-terminal serine) (Hood *et al.*, 1967).

We were interested in extending these studies to determine the number of light-chain types, and the nature and extent of homology between fowl and mammalian V and C regions.

Materials and Methods

Proteins. Pooled chicken immunoglobulins were purchased from Pentex Inc. Plasma from individual birds was generously donated by Dr. M. Blaise (chicken) and the U. S. Agricultural Station at Beltsville, Md. (turkey). Human Bence-Jones proteins (κ-126 and λ-7) were precipitated from urine with half-saturated $(NH_4)_2SO_4$, dialyzed, and lyophilized. κ-126 was generously donated by Dr. C. Baglioni.

Enzymes. Trypsin treated with L-(tosylamido-2-phenyl)ethyl chloromethyl ketone was purchased from Worthington Biochemical Corp. and stored as a 1% solution in 0.001 N HCl at −20°. Thermolysin was purchased from California Biochemical Corp. and freshly prepared as a 1% solution in distilled water before each digestion.

Edman Reagents. Pyridine, ethyl acetate, and butyl acetate (reagent grade) were redistilled over ascorbic acid prior to use. Trifluoroacetic acid (Sequanal grade) was purchased from Pierce Chemical Corp. Phenyl isothiocyanate was obtained from Eastman Chemical Co. and redistilled *in vacuo*. All reagents were stored at −20° under N_2.

Sequencer reagents and solvents were purchased from Beckman Instruments Co.

PTH[2]-amino acids were purchased from Mann Research Laboratories.

Dansyl Reagents and Materials. Dansyl-Cl and dansylamino acids were obtained from Pierce Chemical Co. Polyamide thin-layer chromatography sheets were purchased from Cheng Chin Trading Co., Ltd. (Gallard-Schlesinger, Chemical Manufacturing Co.). The formic acid, benzene, glacial acetic acid, ethyl acetate, and methanol used in dansylamino acid thin-layer chromatography were reagent grade.

Other Reagents. Dithiothreitol (California Biochemical Corp.) was purified by sublimation. Iodoacetamide (Pierce

* From the Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014, and the University of Texas at Austin, Austin, Texas 78712. *Received February 19, 1971.*

† Present address: Johns Hopkins University School of Medicine, Division of Clinical Immunology, Baltimore, Md. 21212; to whom reprint requests should be addressed.

‡ Department of Zoology, The University of Texas at Austin.

§ Department of Biology, California Institute of Technology, Pasadena, Calif. 91109.

[1] The abbreviations used for immunoglobulin chains are those proposed by the World Health Organization; see *Bull. W. H. O. 41*, 975 (1969).

[2] Abbreviations used are: PTH, phenylthiohydantoin derivative; dansyl, 1-dimethylaminonaphthalene-5-sulfonyl.

TABLE I: Amino Acid Composition of Immunoglobulin Light Chains of an Individual Chicken.

Amino Acid	Moles/Mole of Light Chain[a]
Lys	7.9
His	2.1
Arg	6.6
CM-Cys	4.9
Asp	17.3
Thr	24.4
Ser	24.5
Glu	20.6
Pro	13.1
Gly	19.1
Ala	15.4
Val	14.4
Met	0.8
Ile	9.9
Leu	13.4
Tyr	9.6
Phe	6.5
Total	210.5

[a] Molar ratios were determined by assuming 210 amino acid residues per light chain. Tryptophan content was not determined.

Chemical Corp.) was recrystallized three times from ethanol. [^{14}C]Iodoacetamide was purchased from New England Nuclear Corp. and recrystallized with [^{12}C]iodoacetamide to a final specific activity of about 1.5 μCi/mg.

Ethylenimine was purchased from Matheson Coleman & Bell. Tris (Ultra Pure) was obtained from Schwarz Bio-Research. Guanidine-HCl (Ultra Pure) was obtained from Mann Research.

Purification of Immunoglobulins. Pooled chicken immunoglobulins from Pentex Inc. were used without further purification. Immunoglobulins from individual birds were isolated as described by Benedict (1967).

Isolation of Light Chains. Immunoglobulins were dissolved in a 0.1 M Tris-HCl (pH 8.6) buffer at a concentration of *ca.* 30 mg/ml. Interchain disulfide bonds were cleaved by reduction with 5 mM dithiothreitol at 25° for 1 hr. The solution was then dialyzed against 1 N acetic acid for 1 hr and heavy and light chains were separated by gel filtration chromatography (Sephadex G-100, 3.5 × 150 cm) in 1 N acetic acid.

Purification of Bence–Jones Proteins. DEAE-cellulose chromatography was used as previously described (Grant and Hood, 1971).

Complete Reduction and Alkylation. Purified light chains or Bence–Jones proteins were dissolved in 7 M guanidine-HCl–0.2 M Tris-HCl (pH 8.6) at a concentration of 20 mg/ml. Intrachain disulfide bonds were reduced with 50 mM dithiothreitol at 37° for 90 min. An equal volume of 0.11 M iodoacetamide dissolved in 7 M guanidine-HCl–0.2 M Tris-HCl (pH 7.3) was added and the reaction was carried out at 25° for 15 min. The solution was then dialyzed against repeated changes of water at 4°. Light chains were alkylated with [^{14}C]-iodoacetamide under identical conditions.

Aminoethylation. Aminoethylation was carried out as described by Cole (1967).

Enzyme Digestions. Light chains which had been completely reduced and alkylated or aminoethylated were suspended in 0.1 M NH_4HCO_3 (10 mg/ml) and digested with 1% (w/w) trypsin at 37° for 90 min. Generally a second equal aliquot of trypsin was added and the solution was incubated again for 90 min.

Completely reduced and alkylated light chains were digested in 0.1 M NH_4HCO_3 with thermolysin (200:1, w/w) at 37° for 1 hr.

Isolation of Tryptic and Thermolysin Peptides. Peptide maps were prepared as described by Katz *et al.* (1959). Peptides were eluted as described by Bennett (1967).

Ehrlich Stain. Ehrlich stain was prepared and used as described by Smith (1953).

C-Terminal Peptide Isolation. Because the C-terminal tryptic peptide of light chains is small and has a half-cystine residue and no basic amino acid, it can readily be isolated after performic acid oxidation by electrophoresis at pH 6.5 as described by Hood *et al.* (1967).

Radioautography of Half-Cystine Peptides. Radioautography was carried out on trypsin and thermolysin fingerprints of ^{14}C-labeled light chains exposed to GAF-X film (GAF Corp.) for 2 weeks. Labeled peptides were then eluted as described above.

Dansyl-Edman Peptide Degradation. The dansyl-Edman procedure of Gray (1967) was followed. Dansylamino acids were identified by thin-layer chromatography on polyamide sheets (Woods and Wang, 1971; Crowshaw *et al.*, 1967) as modified by Kaplan and Metzger (1969).

To confirm the residue identification, the butyl acetate extract (containing the phenylthiocarbamylamino acid) was evaporated to dryness with N_2 and hydrolyzed to the free amino acid in an evacuated (<20 mTorr) glass tube containing 0.2 ml of 6 N HCl at 100° for 18 hr. The amino acid was identified qualitatively by high-voltage paper electrophoresis (Dreyer and Bynum, 1967).

Hydrazinolysis. To 10–40 nmoles of peptide was added 0.05 ml of anhydrous hydrazine. The tube was flushed with N_2, sealed, and incubated at 100° for 6 hr (Fraenkel-Conrat and Tsung, 1967). The seals were broken and the contents were dried under vacuum in a desiccator in the presence of NaOH and P_2O_5. The residue was dissolved in 0.4 ml of 0.02 M HCl and the carboxy-terminal residue determined on the amino acid analyzer.

Amino Acid Analysis of Peptides. Peptides were hydrolyzed in evacuated tubes (<20 mTorr) in 6 N HCl at 110° for 18 hr and examined by high-voltage paper electrophoresis (Dreyer and Bynum, 1967) or on the amino acid analyzer.

Edman Sequencer Analysis. About 5 mg of each protein was examined in the Beckman 890 protein sequencer using a program similar to that described by Edman and Begg (1967). The resulting PTH amino acids were examined on a Varian Model 1840 gas chromatograph (Pisano and Bronzert, 1969). The PTH-amino acids were hydrolyzed to free amino acids (6 N HCl, 150°, 20 hr) (Van Orden and Carpenter, 1964) and analyzed on a Beckman 120 amino acid analyzer.

Results

Amino Acid Composition. In Table I is given the amino acid composition of individual chicken light chains. As with most mammalian species, five half-cystines are present (see following section).

Number of Light-Chain Types. The half-cystine residues in most mammalian light chains are highly conserved: two are

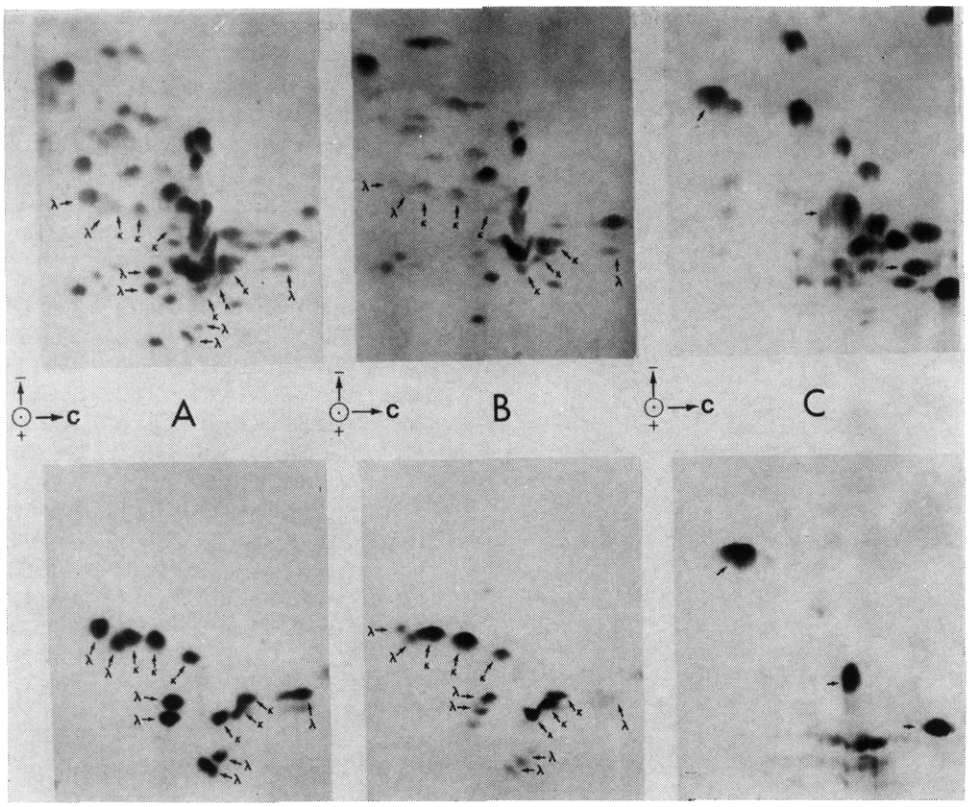

FIGURE 1: Thermolysin fingerprints of mixtures of human λ and κ myeloma light chains and chicken light chains. In the upper row fingerprints are stained with collidine ninhydrin and in the lower row radioautographs of the same fingerprint are presented for the proteins alkylated with [^{14}C]iodoacetamide. A is a λ to κ mixture of 1:1 (~1 mg:1 mg); B a λ to κ mixture of 1:20 (~0.1 mg:2 mg); and C is individual chicken light chains (~2 mg). Arrows indicate the Cys peptides. C indicates the direction of chromatography.

positioned in the V region and three in the C region. In normal (heterogeneous) light chains of a single type (λ or κ), each C region half-cystine is part of an invariant sequence whereas the V region half-cystines are found in regions of extensive sequence variation (Wu and Kabat, 1970). Thus after digesting normal light chains with an appropriate enzyme, three distinct C region peptides should be found for each light-chain type present whereas the more heterogeneous V region peptides should appear as multiple peptides in low yield. Half-cystine peptides can be readily localized on fingerprints after alkylation with [^{14}C]iodoacetamide and radioautography. We were interested in determining the number of light-chain types (C-region peptides) in chicken light chains, after first carrying out studies with the appropriate controls to determine the level at which a second light-chain type could be detected.

Control studies with mixtures of human myeloma (Bence–Jones) light chains of the λ and κ types show that [^{14}C]carboxamidomethyl-Cys peptides can be clearly detected at the 5% level. A thermolysin digest of these mixtures is shown in Figure 1A,B. The upper maps are stained with collidine ninhydrin, whereas the lower maps represent radioautographs. Figure 1A is the fingerprint resulting from a mixture of 2 mg of κ and 2 mg of λ chain. Six κ and seven λ major peptides are present. The extra κ and two extra λ peptides probably represent overlapping thermolysin cleavage points for each of the respective chains, as may other fainter spots (peptides). Figure 1B is the fingerprint resulting from a mixture of 2 mg of κ and 0.1 mg of λ chain. Thus λ half-cystine peptides can clearly be detected at the 5% level by radioautography even though many of these peptides cannot be visualized in the ninhydrin stain. No radioactivity remained at the origin in these thermolysin fingerprints. Tryptic fingerprints, in contrast, left labeled, insoluble peptides at the origin.

A thermolysin fingerprint of 2 mg of light chains isolated from a single individual chicken is given in Figure 1C. The three major peptides (arrows) are from the C region (peptides C8, C9, and C11 in Figures 2 and 3). The smear of labeled peptides in the neutral band was isolated in low yield and proved to be extremely heterogeneous by dansyl end-group analysis and by amino acid analysis. Thus they probably represent the heterogeneous V region half-cystines peptides. A few other faint spots represent peptides which are present at less than the 5% level (compare to 1A and 1B). These fainter spots may represent V region peptides, peptides derived from overlapping cleavages, or those derived from a second light-chain type which is present at the level of a few per cent. Thus individual chicken light chains have a single light-chain type at the 95% level or greater. Labeled turkey light chains gave a thermolysin fingerprint virtually identical with chicken light chains. We then sought to determine whether this chicken (and turkey) light-chain type was homologous with either of the mammalian light-chain types by peptide and automatic sequence analysis.

Automatic Sequence Analysis of Avian Light Chains. Results obtained from the sequential Edman degradation of individual chicken and turkey light chains are given in Table II. The quantitative results are those derived by column chromatography (Beckman Model 120 amino acid analyzer) of free amino acids after acid hydrolysis of the PTH-amino acids. Since PTH-Ser and PTH-Thr are destroyed by the harsh conditions of hydrolysis, quantitation of these derivatives was obtained by gas chromatography. Lys and Arg could not be

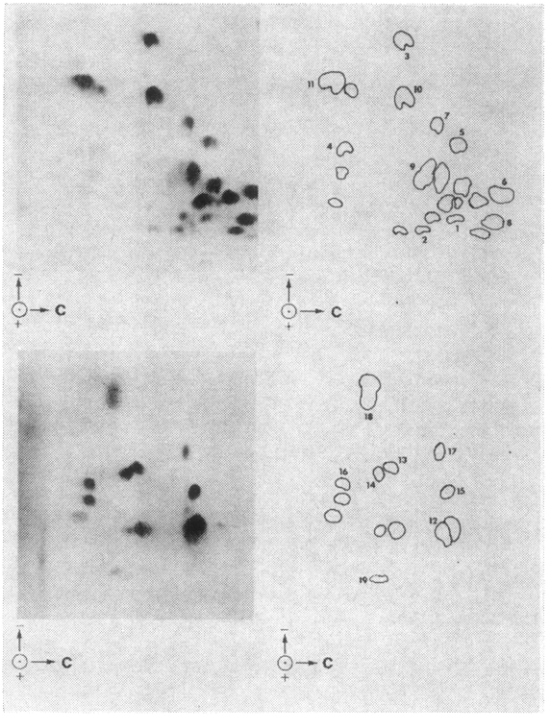

FIGURE 2: Fingerprints (2 mg) of carboxamidomethylated chicken light chains digested with thermolysin (upper map) and aminoethylated light chains digested with trypsin (lower map). Peptide numbering is that used in subsequent tables and figures.

detected in our analyses as neither is seen on gas chromatography and the amino acid samples were only analyzed on the acid–neutral column of the amino acid analyzer. In general the results from gas chromatography and amino acid analysis were in close agreement.

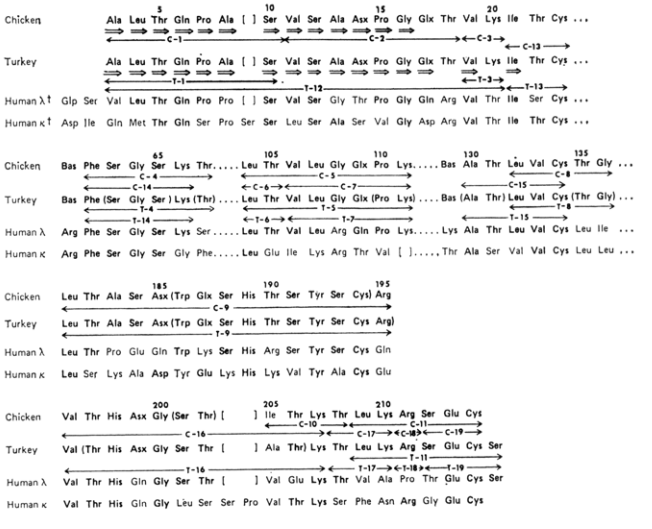

FIGURE 3: Sequence homologies between avian and human immunoglobulin light chains. Peptides are numbered according to Figure 2. Human λ and κ chain residues in **boldface** are identical with their avian counterparts. [ ] indicates gaps which are necessary for chain alignment. Parentheses indicate that a sequence has not been determined for the enclosed residues. → designates a residue identified by the protein sequencer. Bas indicates Arg or Lys. *, alignments and numbering are that of Dayhoff (1969). †, amino acid sequences of human λ and κ Bence–Jones proteins *Ha* and *Ag*, respectively.

TABLE II: Amino-Terminal Residue Alternatives in Avian Light Chains.[a]

Residue Position	Residue (nmoles Recovered)[b]					
	Chicken			Turkey		
	1	2	3	1	2	3
1	Ala 125	Gly 3	Glx 1	Ala 69	Val 4	Leu 4
2	Leu 189	Val 24	Ala 9	Leu 60	Val 17	Gly 7
3	Thr 52	Gly 11	Leu 6	Thr 63	Leu 8	Val 6
4	Glx[c] 107	Gly 10	Leu 5	Glx[c] 31	Leu 9	Gly 4
5	Pro 34	Glx 11	Gly 9	Pro 22	Glx 21	Leu 16
6	Ala 38	Gly 9	Leu 6	Ala 32	Leu 9	Pro 3
7	Ser 67	Ala 18	Gly 13	Ser 76	Gly 13	Val 8
8	Val 89	Gly 27	Ala 15	Val 22	Gly 9	Ala 4
9	Ser 48	Val 26	Gly 17	Ser 51	Gly 18	Val 9
10	Ala 78	Ser 32	Gly 17	Ala 39	Gly 10	Val 5
11	Asx 61	Ala 22	Ser 15	Asx 26	Gly 18	Ala 14
12	Pro 16	Leu 13	Asx 10	Pro 17	Val 17	Asx 8
13	Gly 70	Leu 12	Asx 10	Gly 24	Val 11	Leu 9
14				Glx 20	Gly 13	Val 7
15				Ala 11	Val 9	Glx 8
16				Val 27	Gly 9	Ala 5
17				Val 8	Glx 8	Ala 6
18				Ile 13	Val 8	Gly 7

[a] Sequential Edman degradation was carried out on the Beckman Model 890 protein sequencer. [b] Recoveries were determined by amino acid analysis of the hydrolyzed PTH derivatives. Ser and Thr which are destroyed by this procedure were obtained from gas chromatography. The three residues present in highest yield are given for each step. [c] Gln was noted at this step on gas chromatography.

The automatic Edman procedure and the harsh conditions of acid hydrolysis used to convert the PTH-amino acids to free amino acids each contribute "noise" which can be confused with low level residue alternatives present in the heterogeneous population of normal light chains. The Edman procedure produces "noise" by the random hydrolysis of the polypeptide chain during the acid cleavage of the phenylthiocarbamyl-protein derivative and by incomplete cleavage at a step leading to trailing residues in the subsequent steps. Acid hydrolysis generates noise by the breakdown of complex amino acids to simpler ones (*i.e.*, 2–5% of PTH-Val is converted into Gly by the conditions used). In these analyses, trailing of the major residue from the preceding step (see chicken light-chain steps 8–12) and large amounts of Gly (see turkey light-chain steps 7–11) seem to be the major sources of "noise." The Gly "noise" comes, in part, from the hydrolysis procedure as Gly was seen in smaller quantities at these steps on gas chromatography.

The overall yield of PTH-Ala at the first step ranged between 25 and 50% for repetitive runs on the same sample. The striking finding present in these data is that there is a single major residue at most of the positions examined for both chicken and turkey light chains. The major residues for each position in these light chains are shown in Figure 3 (with a double arrow).

Positions 5, 15, and 17 in turkey light chains do not show a predominant residue. Pro was designated the predominant

residue at position 5 as the second residue, Glx, is probably the trailing residue from position 4. The mixture of amino acids seen at positions 15 and 17 is consistent with the results expected from a position that has a hydroxyamino acid (Ser or Thr) or a basic amino acid (Lys or Arg), namely no predominant residue is present (PTH-Ser and -Thr are destroyed by acid hydrolysis and PTH-Lys and -Arg would not be seen on the acid–neutral amino acid analyzer column). Position 2 in both species had a minor residue (Val). Because of the noise problem, it is difficult to determine the extent and nature of heterogeneity at other positions (see Grant and Hood (1971) for a more thorough discussion of the noise and yield problems).

Peptide Characterization. Peptides were isolated from chicken and turkey light chains by the preparative peptide map procedure. Tryptic peptides were derived from aminoethylated chains and thermolysin peptides from carboxymethylated chains. Peptides were prepared from chains derived from individual birds and from pooled chicken light chains (Pentex Inc.). Representative peptide maps and drawings with numbered peptides are given for thermolysin and trypsin digests of chicken light chains in Figure 2. The peptide maps of various enzymatic digests of turkey light chains were very similar to those of their chicken counterparts. The peptide maps prepared from individual and pooled chicken light chains were also virtually identical.

The amino acid compositions of these peptides are given in Tables III and IV for chicken and turkey chains, respectively. The numbering used is that given in Figure 2. Peptide yields ranged from 10 to 35%. Sequence analysis of these peptides using the dansyl-Edman procedure is given in Tables V and VI.

Variable Region Peptides. The alignment of the avian peptides with human λ and κ chains is given in Figure 3 and suggests that a gap of two residues exists at the amino terminus of the avian chains. Because of the size difference, the numbering system of the human light chains given in Figure 3 will be used for comparing human and avian chains. Tryptic peptides T-12 and T-13 (T = turkey) appear to represent positions 3–23. The T-12 peptide is 17 residues in length, has N-terminal Ala, and has a composition identical with the major amino-terminal 17 residues as determined by the protein sequencer—assuming that the residues at positions 18 and 20 are Thr and Lys, respectively. Thus T-12, which admittedly is obtained in low yield, seems to represent a major amino-terminal sequence. The peptide T-13 can be placed at positions 21–23 because the protein sequencer has identified an overlapping Ile at position 21 and the three other half-cystine peptides isolated come from the C region (see next section). Furthermore, T-13 is clearly homologous to the corresponding half-cystine peptide at position 23 in human light chains and not homologous to the only remaining half-cystine peptide at position 88 (compare with human λ and κ chains in Figure 3). Thermolysin peptide T-1 is identical in sequence with the major residues at positions 3–10 in the intact turkey light chain and T-13 may be placed at positions 19–20. Thus a major sequence seems to exist in the V region of turkey light chains for positions 3–23.[3]

The same predominant amino-terminal sequence can be derived for the chicken light chain by similar arguments. Thermolysin peptides C-1 (C = chicken) and C-2 are consistent with the major residues obtained from the protein sequencer at positions 3–18. C-3 is identical with T-3 and C-13 to T-13. These last two peptides can only be aligned by homology with their turkey counterparts. One additional point is of interest regarding the heterogeneity present in this region. Peptide C-12 is identical in its amino-terminal five residues with the major amino-terminal sequence present in chicken light chains. Yet this peptide differs from the remainder of the amino-terminal 17 residues (that were determined on the intact protein by automatic sequence analysis) in that there is an extra Ser and one less Ala. These differences may represent one or more additional V region subgroups.

Peptides were isolated from one additional portion of the V region. Thermolysin (C-4 and T-4) and tryptic (C-14 and T-14) peptides were found homologous to positions 62–67 in mammalian light chains (Figure 3). This portion of the V region is highly conserved in the light chains from ten mammals (T. Stanton, R. Woods, D. Capra, and L. Hood, in preparation) and may play a critical but as yet unknown role in determining light chain structure.

"Switch" Region Peptides. Thermolysin peptides (C-5, C-6, C-7 and T-5, T-6, and T-7) are very similar to human λ switch peptides at positions 104–111 (Figure 3). The switch peptide is that which joins the V and C regions. Identical switch peptides have been isolated from the normal light chains of five mammals including human (Hood *et al.*, 1970b)—thus this is another highly conserved region.

Common Region Peptides. Peptides were isolated which could be aligned against their human counterparts for positions 130–136 and 181–214 or 215. Tryptic peptides C-15 and T-15 were identical with the human λ chain at positions 130–134. Peptides C-9 (T-9) and C-16 (T-16) are clearly homologous to human light chains at positions 181–207 although no peptide overlaps were obtained for these two large peptides (Figure 3). Sufficient thermolysin overlap peptides (C-10, C-11) do exist to align the C-terminal tryptic peptides C-16, C-17, C-18, and C-19. Thus peptide C-9 (T-9) must be adjacent to peptide C-16 (T-16) if one is to preserve the invariant Cys at position 194 seen in all mammalian light chains (Smith *et al.*, 1971; Hood *et al.*, 1970b).

Large quantities of the avian C-terminal peptides were isolated in their cysteic acid form (see Methods). The C terminus of the turkey chain is one residue longer (Ser) than its chicken counterpart both by dansyl-Edman sequence analysis and by hydrazinolysis of the tryptic peptides C-19 and T-19. These hydrazinolysis data are presented in Table VII. This is consistent with previously published results on the hydrazinolysis of whole chicken and turkey light chains (see Hood *et al.*, 1967, Table 6)[4] One other difference is noted at position 205 where an Ile-Ala interchange occurs.

Discussion

The radioautographs of chicken [^{14}C]carboxamidomethyl-Cys peptides (Figure 1) in conjunction with the sequence analysis of the labeled peptides (Figures 2 and 3) demonstrate that the chicken and the turkey have one predominant light-chain type. Control radioautographs with mixtures of labeled myeloma light chains suggest that a second light-chain type, if it exists at all, must comprise less than 5% of the normal

[3] This alignment is further supported by observations on the pheasant light chain. This chain has a major sequence which is identical with that derived for the turkey over residues 3–23 based on quantitative studies with the automatic sequencer (R. Acton, G. Leslie, and L. Hood, in preparation).

[4] The sequence of the C-terminal tryptic peptide of turkey was incorrectly reported in this paper as Ser-Glu-Cys due to a low yield of the Ser residues on amino acid analysis.

TABLE III: Amino Acid Composition of Thermolysin and Tryptic Peptide Derived from Chicken Light Chains.[a]

Enzyme	Thermolysin											Trypsin							
Peptide	1[b]	2	3[b]	4	5	6[b]	7	8	9	10[b]	11	12	13[b]	14[b]	15[b]	16	17[b]	18[b]	19
Lys			1	1.1 (1)	0.9 (1)		0.8 (1)			1	1.0 (1)	0.9 (1)		1		0.9 (1)	1		
His									0.9 (1)							0.9 (1)			
Arg									1.1 (1)		1.3 (1)							1	
AE-Cys or CM-Cys[c]								0.7 (1)	0.4 (1)		0.4 (1)		1		1				0.9 (1)
Asp		1.0 (1)							1.0 (1)			1.0 (1)				1.0 (1)			
Thr	1	1.2 (1)		0.9 (1)	0.8 (1)	1		0.7 (1)	1.5 (2)	2		2.0 (2)	1		1	2.7 (3)	1		
Ser	1	0.9 (1)		1.6 (2)					3.3 (4)		0.8 (1)	2.9 (3)		2		1.1 (1)			0.8 (1)
Glu	1	1.1 (1)			1.1 (1)		1.1 (1)		1.0 (1)		0.7 (1)	2.1 (2)							1.0 (1)
Pro	1	0.8 (1)			0.9 (1)		1.0 (1)					2.2 (2)							
Gly		1.2 (1)		1.1 (1)	1.1 (1)		1.2 (1)	1.0 (1)				1.2 (1)		1		1.2 (1)			
Ala	2	1.0 (1)							1.0 (1)			1.9 (2)			1				
Val		0.6 (1)[f]	1		0.9 (1)		0.7 (1)[f]	0.9 (1)				2.2 (2)			1	0.7 (1)[f]			
Met										1			1			1.0 (1)			
Ile																			
Leu	1				1.3 (2)[f]	1	0.9 (1)	1.0 (1)	0.6 (1)[f]		0.7 (1)[f]	0.9 (1)			1		1		
Tyr									0.8 (1)										
Phe				0.8 (1)										1					
Trp									(1)[e]										
Total residues	7	8	2	6	8	2	6	5	15	4	6	17	3	5	5	10	3	1	3
Yield[d] (%)	25	35	30	25	25	20	15	30	30	30	25	10	25	20	35	20	25	35	35

[a] Values reported are moles of amino acid residue per mole of peptide. Amino acids present at a level of less than 0.2 residue are omitted. Values in parentheses are nearest integral numbers of residues. [b] Molar ratios of amino acids were estimated from paper electrophoretograms. [c] AE represents aminoethylcysteine and CM carboxymethylcysteine. [d] Yields are based on micromoles of peptides isolated compared with total micromoles of light chain digested with the enzymes and rounded off to the nearest 5%. [e] Trp was determined by Ehrlich strain. [f] These residues are probably low because they are N terminal and have partly been destroyed by the ninhydrin used for peptide detection.

TABLE IV: Amino Acid Composition of Thermolysin and Tryptic Peptides Derived from Individual Turkey Light Chains.[a]

Enzyme	Thermolysin								Trypsin							
Peptide	1[b]	3[b]	4	5	6[b]	7	8	9	11	12	13[b]	14[b]	15[b]	16	17[b]	18[b]	19
Lys		1	0.9 (1)	0.9 (1)		1.0 (1)			1.0 (1)	0.9 (1)		1		0.9 (1)	1		
His								0.7 (1)						1.1 (1)			
Arg								1.0 (1)	1.4 (1)							1	

AE-Cys or CM-Cys[c]							0.5 (1)	0.5 (1)	0.4 (1)		1		1				0.8 (1)
Asp								1.1 (1)		1.2 (1)				1.1 (1)			
Thr	1		0.9 (1)	0.9 (1)	1		1.0 (1)	2.1 (2)		1.8 (2)	1		1	2.6 (3)	1		
Ser	1		1.6 (2)					3.4 (4)	1.5 (2)	1.8 (2)		2		0.8 (1)			1.4 (2)[f]
Glu	1			1.0 (1)		1.1 (1)		1.2 (1)	1.0 (1)	1.8 (2)							1.0 (1)
Pro	1			1.0 (1)		0.9 (1)				2.3 (2)							
Gly			1.0 (1)	1.1 (1)		1.0 (1)	0.9 (1)			1.2 (1)		1	1	1.0 (1)			
Ala	2							0.8 (1)		2.3 (3)[f]			1	1.1 (1)			
Val		1		1.0 (1)		0.8 (1)	1.0 (1)			1.8 (2)				1.0 (1)			
Met																	
Ile											1						
Leu	1			1.4 (2)[f]	1	1.2 (1)	0.6 (1)[f]	0.5 (1)[f]	0.4 (1)[f]	0.9 (1)			1		1		
Tyr								0.8 (1)									
Phe			0.8 (1)									1					
Trp								(1)[e]									
Total residues	7	2	6	8	2	6	5	15	7	17	3	5	5	10	3	1	4
Yield[d] (%)	20	30	20	20	20	15	20	25	25	10	25	30	25	20	30	30	35

[a] See Table III for an explanation of *a–f*.

TABLE V: Amino Acid Sequences of Chicken Thermolysin and Tryptic Peptides.[a]

Peptide	
C-1	Ala-Leu-Thr-Glx(Pro,Ala,Ser)
C-2	Val-Ser-Ala-Asx-Pro-Gly-Glx-Thr
C-3	Val-Lys
C-4	Phe-Ser-Gly-Ser-Lys-Thr
C-5	Leu-Thr-Val-Leu-Gly-Glx-Pro-Lys
C-6	Leu-Thr
C-7	Val-Leu-Gly-Glx(Pro-Lys)
C-8	Leu-Val-Cys-Thr-Gly
C-9[b,c]	Leu-Thr-Ala-Ser-Asx(Trp,Glx,Ser,His,Thr,-Ser,Tyr,Ser,Cys)Arg
C-10	Ile-Thr(Lys,Thr)
C-11	Leu-Lys-Arg-Ser-Glu-Cys
C-12	Ala-Leu-Thr-Glx-Pro(Ala,Ser,Val,Ser,Ser,-Asx,Pro,Gly,Glx,Thr,Val)Lys
C-13	Ile-Thr-Cys
C-14	Phe-Ser-Gly-Ser-Lys
C-15	Ala-Thr-Leu-Val-Cys
C-16	Val-Thr-His-Asx-Gly(Ser,Thr,Ile,Thr)Lys
C-17	Thr-Leu-Lys
C-18	Arg
C-19[b]	Ser-Glu-Cys

[a] C-1 through C-11 are thermolysin peptides; C-12 through C-19 are tryptic peptides. [b] C-terminal residue determined by hydrazinolysis. [c] Trp determined by Ehrlichs stain.

light-chain population. This approach appears to be a powerful method for determining the number and approximate ratio of light-chain types.

The predominant avian light-chain type is more closely related to human λ than to human κ chains in both the C and V regions. Gaps (or insertions) are shared by the avian and human λ chains at positions 9, 111, and 203–204 (Figure 3). These shared gaps (or insertions) constitute striking evidence for homology as (1) deletions are much rarer events than single base substitutions, (2) deletions seem to be relatively irreversible events in protein evolution, and (3) it is very unlikely that parallel gaps would be generated in separate genes (Nolan and Margoliash, 1968). This homology is also supported by a comparison of amino acid substitutions. At the 31 positions that can be compared in the V region (the V region extends over residues 1–108 in Figure 3), 22 identities are shared by the avian and the human λ chain, whereas only 15 residues are shared with the κ chain. In the C region 43 residues are compared and 29 are shared with the human λ chain and only 18 with the human κ chain. A somewhat more extensive series of comparisons is shown in Figure 4 in which the method of Sokal and Sneath (1963) is used to determine an approximate phylogenetic tree (see legend to Figure 4 for details). This figure demonstrates that both the V and C regions of avian light chains are more closely related to λ than to other immunoglobulin chains. Since the peptides chosen for sequence analysis were, in part, chosen because of obvious homology with mammalian light chains, these comparisons certainly overestimate the degree of homology which exists between mammalian and avian light chains. Also unanswered in this study is the question of whether or not the avian species

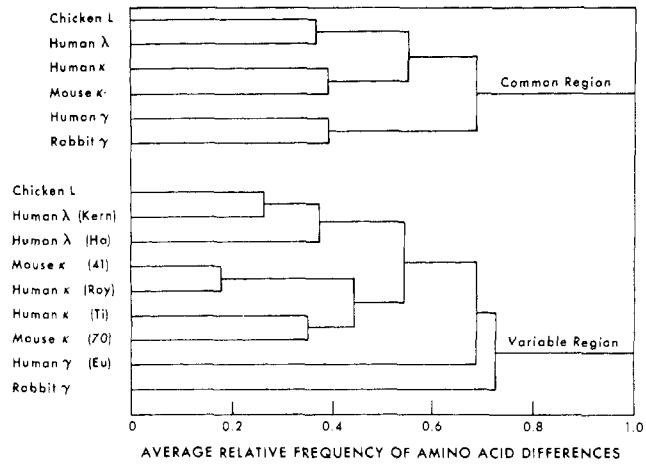

FIGURE 4: Average relative frequency of amino acid differences between fowl and mammalian immunoglobulins. The sequence of chicken light chain given in Figure 3 is compared to mammalian immunoglobulin light and heavy chains over the regions 1–23, 130–136, and 181–214 (Figure 3). The mammalian chains have also been compared over these regions. Sequences have been aligned to give the best homologies. Sequences were obtained from the following sources: Kern, Ha, Roy, Ti, 41, and 70 (Dayhoff, 1969); Eu (Edelman *et al.*, 1969); and rabbit γ chains (Fruchter *et al.*, 1970; Wilkinson, 1969). A matrix was calculated for amino acid differences between each two chains. Amino acid interchanges and gaps were treated equally. The above hierarchial arrangement of average amino acid differences was then determined according to the technique of Sokal and Sneath (1963).

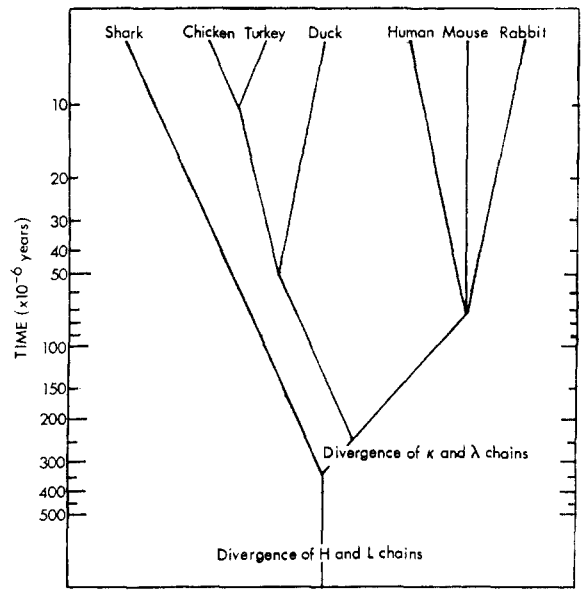

FIGURE 5: An evolutionary tree for various vertebrate species based on fossil data summarized by Harland (1967), Brodkorb (1964; 1970, personal communication) and Yapp (1965). Clearly λ and κ genes could have diverged prior to the separation of the lines leading to contemporary mammals and sharks *or* at some point after the divergence of sharks and before the divergence of birds from the mammalian line (as shown).

have a set of λ chains with a blocked α-amino group as is found in most mammalian λ chains.

As this paper was being written, Kubo *et al.* (1970) reported an automatic sequence analysis of the amino-terminal 17 residues of chicken light chains (positions 3–20 in Figure 3). The amino-terminal six residues are identical in both studies, but we have different major residue alternatives at six of the last eleven positions. It is difficult to believe these discrepancies are due to genetic polymorphisms in view of the virtual identities observed at the amino terminus of pheasant, turkey, and chicken light chains. Perhaps these differences are due, in part, to the selection by differing antigenic environments for different chicken V region pools.

The avian V regions are remarkably homogeneous at their N terminus in contrast to their mammalian counterparts which show extensive heterogeneity in this region due to the major subsets (subgroups) into which mammalian V_λ and V_κ regions can be divided (Hood and Talmage, 1970; Hood *et al.*, 1970a,c). Perhaps the reason for the homogeneity seen in the amino-terminal sequences of chicken and turkey light chains is due to the fact that they each have just one unblocked V_λ subgroup. About 20% of normal human λ chains are of the $V_{\lambda III}$ subgroup which has an unblocked N-terminal Tyr

TABLE VI: Amino Acid Sequences of Turkey Thermolysin and Tryptic Peptides.[a]

Peptide	
T-1	Ala-Glx-Thr-Glx(Pro,Ala,Ser)
T-3	Val-Lys
T-4	Phe(Ser,Gly,Ser,Lys,Thr)
T-5	Leu-Thr-Val-Leu(Gly,Glx,Pro,Lys)
T-6	Leu-Thr
T-7	Val-Leu-Gly-Glx(Pro,Lys)
T-8	Leu-Val-Cys(Thr,Gly)
T-9[b,c]	Leu-Thr-Ala-Ser-Asx(Trp,Glx,Ser,His,Thr,-Ser,Tyr,Ser,Cys,Arg)
T-11	Leu-Lys-Arg-Ser-Glu-Cys-Ser
T-12	Ala(Leu,Thr,Glx,Pro,Ala,Ser,Val,Ser,Ala,-Asx,Pro,Gly,Glx,Thr,Val)Lys
T-13	(Ile,Thr)Cys
T-14	(Phe,Ser,Gly,Ser)Lys
T-15	(Ala,Thr,Leu,Val)Cys
T-16	Val(Thr,His,Asx,Gly,Ser,Thr,Ala,Thr)Lys
T-17	(Thr,Leu)Lys
T-18	Arg
T-19[b]	Ser-Glu-Cys-Ser

[a] T-1 through T-11 are thermolysin peptides; T-12 through T-19 are tryptic peptides. For an explanation of *b* and *c* see Table V.

TABLE VII: Hydrazinolysis of Tryptic Peptides C-19 and T-19.[a,b]

	Chicken	Turkey
SO_3-Cys[c]	55	8
Ser	6	32
Glu	<1	<1

[a] Per cent yield of free amino acid after hydrazinolysis as compared with the yield of an identical amount of the peptide hydrolyzed in 6 N HCl. [b] Certain hydrazides (*i.e.*, Ser, Ala, and SO_3-Cys) are unstable under the hydrolysis conditions used and break down slowly to free amino acids. [c] Cysteic acid.

(Grant and Hood, 1971). If one could examine normal human λ chains by automatic sequence analyis, a single major sequence would probably be obtained which would be identical with the prototype sequence of the $V_{\lambda III}$ proteins (Grant and Hood, 1971).

Clear evidence for avian chain heterogeneity exists at positions 2 (see Table II) and in the C-12 peptide (see Results section) which may reflect the presence of two or more minor V region subgroups. In addition, the thermolysin and tryptic peptides from the Cys-88 region are also extremely variable.

There are two certain differences between the chicken and turkey light chains over 74 positions of comparison (see Figure 3, positions 205 and 215). Both species are in the same order (Galliforms) and the divergence of species lines probably occurred about 10 million years ago (P. Brodkorb, personal communication) (see Figure 5). A preliminary analysis of amino terminal 10 residues of duck light chains (order Anseriform) suggests that they are similar to the Galliform chains but that there is far more V region heterogeneity (J. A. Grant and L. Hood, unpublished results).

The evolutionary tree depicted in Figure 5 gives the approximate divergence times for mammals, avians, and sharks. Mammals have both λ and κ light-chain types (Hood *et al.*, 1967, 1970b). Avians have predominately chains of the λ type. Preliminary studies on the amino-terminal 21 residues of the Nurse shark demonstrate a population of light chains which show a striking homology with human κ chains (L. Hood and W. Clem, in preparation; also see Suran and Papermaster, 1967). There are to our knowledge no available data to answer the question of whether or not sharks also have a blocked population of λ chains. It will, of course, be important to extend the sequence data in the shark to strengthen the homology with mammalian κ chains; nevertheless it seems likely from the data cited above that the ancestor to sharks and mammals must have had κ genes. The ancestor to avians and mammals must have had λ genes. Thus *λ and κ genes must have diverged prior to the divergence of the avian and mammalian lines (about 250 million years ago)*. To argue otherwise would require an unlikely series of parallel divergences.

The criteria for determining light-chain type in new species suggested by Hood *et al.* (1967) obviously must be used with caution. We suggested then that κ chains were characterized at their amino terminus by a free α-amino group (λ chains generally have a blocked Glp residue) and at their carboxy terminus by a half-cystine residue (λ chains generally have Ser). The chicken light chain fulfills both of the κ criteria, yet is definitely a λ type chain.

There is evidence to support the conclusion that the V and C regions are encoded by separate germ line genes (Smith *et al.*, 1971; Milstein and Pink, 1970; Hood and Talmage, 1970). Apparently the V and C genes of a given light-chain type have been associated throughout vertebrate evolution since V_λ regions in mammals and avians are associated with homologous C_λ regions.

Finally, there is an extremely intriguing question which relates to the expression of λ and κ light-chain types. Just as with the avians, many mammals express a preponderance of one light-chain type. For example, rats and mice have >95% κ chains whereas most artiodactyls and perissodactyls express >95% λ chains (Hood *et al.*, 1967, 1970b). The ancestors for mammals and avians must at one time have had both λ and κ genes. Have contemporary avians lost the structural genes for κ chains or does the predominance of λ chains reflect the presence of a control mechanism for the selective expression of a single light-chain type?

Acknowledgments

We thank Dr. David Alling for his help with the statistical analyses and Mr. Matthew Miller for his help with the amino acid analysis.

References

Benedict, A. A. (1967), *in* Methods in Immunology and Immunochemistry, Williams, C. A., Ed., Vol. 1, New York, N. Y., Academic Press, p 229.
Bennett, J. C. (1967), *Methods Enzymol. 11*, 300.
Brodkorb, P. (1964), *Bull. Fla. State Museum Biol. Sci. 8*, 195.
Cole, R. D. (1967), *Methods Enzymol. 11*, 315.
Crowshaw, K., Jessup, S. J., and Ramwell, P. W. (1967), *Biochem. J. 103*, 79.
Dayhoff, M. O. (1969), Atlas of Protein Sequence and Structure, Vol. 4, Silver Spring, Md., National Biomedical Research Foundation.
Dreyer, W. J., and Bynum, E. (1967), *Methods Enzymol. 11*, 32.
Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauer, A., and Waxdal, M. J. (1969), *Proc. Nat. Acad. Sci. U. S. 63*, 78.
Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.
Fraenkel-Conrat, H., and Tsung, C. M. (1967), *Methods Enzymol. 11*, 151.
Fruchter, R. G., Jackson, S. A., Mole, L. E., and Porter, R. R. (1970), *Biochem. J. 116*, 249.
Grant, J. A., and Hood, L. (1971), *Immunochemistry 8*, 63.
Gray, W. R. (1967), *Methods Enzymol. 11*, 139.
Grey, H. M. (1969), *Advan. Immunol. 10*, 51.
Harland, W. B. (1967), The Fossil Record, London, Geological Society.
Hood, L., Eichmann, K., Lackland, H., Krause, R., and Ohms, J. (1970a), *Nature (London) 228*, 1040.
Hood, L., Grant, J. A., and Sox, H. C. (1970b), *in* Developmental Aspects of Antibody Formation and Structure, Sterzl, J., and Riha, I., Ed., Prague, Academia Publishing House of the Czechoslovak Academy of Sciences, Vol. I, p 283.
Hood, L., Gray, W. R., Sanders, B. G., and Dreyer, W. J. (1967), *Cold Spring Harbor. Symp. Quant. Biol. 32*, 133.
Hood, L., Potter, M., and McKean, D. (1970c), *Science 170*, 1207.
Hood, L., and Talmage, D. W. (1970), *Science 168*, 325.
Kaplan, A. P., and Metzger, H. (1969), *Biochemistry, 8*, 3944.
Katz, A. M., Dreyer, W. J., and Anfinson, C. B. (1959), *J. Biol. Chem. 234*, 2897.
Kubo, R. T., Rosenblum, I. Y., and Benedict, A. A. (1970), *J. Immunol. 105*, 534.
Milstein, C. (1967), *Nature (London) 216*, 330.
Milstein, C., Milstein, C. P., and Feinstein, A. (1969), *Nature (London) 221*, 151.
Milstein, C., and Pink, J. (1970), *Progr. Biophys. Mol. Biol. 21*, 211.
Nolan, C., and Margoliash, E. (1968), *Annu. Rev. Biochem. 37*, 727.
Pisano, J., and Bronzert, T. (1969), *J. Biol. Chem. 244*, 5597.
Putnam, F. W. (1969), *Science 163*, 633.
Smith, G., Hood, L., and Fitch, W. (1971), *Annu. Rev. Biochem. 40*, in press.
Smith, I. (1953), *Nature (London) 171*, 43.

Sokal, R. R., and Sneath, P. H. (1963), Principles of Numerical Taxonomy, San Francisco, Calif., W. H. Freeman.
Suran, A. A., and Papermaster, B. (1967), *Proc. Nat. Acad. Sci. U. S. 58*, 1619.
Van Orden, H. O., and Carpenter, F. H. (1964), *Biochem. Biophys. Res. Commun. 14*, 399.
Wilkinson, J. M. (1969), *Biochem. J. 112*, 173.
Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.
Wu, T. T., and Kabat, E. (1970), *J. Exp. Med. 132*, 211.
Yapp, W. B. (1965), Vertebrates Their Structure and Life, New York, N. Y., Oxford University Press, p 94.

Sequence of a Segment of Muscle Glycogen Phosphorylase Containing the Pyridoxal 5′-Phosphate Binding Site*

Arden W. Forrey,† Cynthia L. Sevilla,‡ John C. Saari,§ and Edmond H. Fischer||

ABSTRACT: This article describes the amino acid sequence of the pyridoxal 5′-phosphate binding site of rabbit muscle glycogen phosphorylase (EC 2.4.1.1). Limited chymotryptic and tryptic hydrolysis of $NaBH_4$-reduced phosphorylase yielded three phosphopyridoxyl peptides (33, 29, and 26 residues long), each terminating with the same N^{ϵ}-(P-pyridoxyl)Lys-Phe sequence at their carboxyl end. Peptic hydrolysis of these gave rise to a number of smaller fragments whose sequence was determined by a combination of Dansyl-Edman degradation, carboxypeptidases A and B, and leucine aminopeptidase digestions. Partial acid hydrolysis of the parent peptides provided the necessary overlaps from which the peptic fragments could be ordered. Cyanogen bromide degradation of $NaBH_4$-reduced phosphorylase gave rise primarily to a phosphopyridoxyl tripeptide, indicating that the sequence N^{ϵ}-(P-pyridoxyl)Lys-Phe occurs between two methionyl residues. Fortuitously, however, the methionyl peptide-bond distal to the P-pyridoxyllysyl residue was not quantitatively cleaved and an undecapeptide containing a single amino terminus and two homoseryl residues was also isolated and partially characterized. The following sequence for the pyridoxal 5′-phosphate binding site of phosphorylase was established

1 5 10 15
Arg-Val-Ser-Leu-Ala-Glx-Lys-Val-Ile-Pro-Ala-Ala-Asp-Leu-Ser-
20 25
Glx-Glx-Ile-Ser-Thr-Ala-Gly-Thr-Gln-Ala-
(5′-P-Pxy)
30 | 35
Ser-Gly-Thr-Gly-Asp-Met-Lys-Phe-Met-Gly-
42
Arg-Thr-Leu-(Glx,Asx,Thr)-Met . . .

Chymotryptic digestion of $NaBH_4$-reduced rat muscle phosphorylase also produced a 29-residue peptide identical in composition with that obtained from the rabbit, further stressing the structural analogies between the two enzymes.

Since the finding over 10 years ago (Fischer *et al.*, 1958a) that sodium borohydride reduction of muscle phosphorylase (EC 2.4.1.1) would covalently fix pyridoxal-P to the protein with little loss of enzymatic activity, several attempts were made to isolate and characterize a sizeable portion of the cofactor binding site. This task proved more difficult than originally anticipated. Exhaustive chymotryptic digestion of the reduced enzyme yielded only a substituted dipeptide identified as N^{ϵ}-(P-pyridoxyl)lysylphenylalanine (Fischer *et al.*, 1958a; Nolan *et al.*, 1964); no other phosphopyridoxyl peptide was detected, strongly supporting the assumption that $NaBH_4$ reduction had specifically fixed the cofactor at a unique site. Disappointingly, as discussed in this article, cyanogen bromide cleavage of the reduced enzyme led only to the isolation of a tripeptide, indicating that the dipeptide mentioned above was "sandwiched" between two methionyl residues.

For several reasons, it appeared essential to pursue this problem and determine the structure of a larger fragment of the cofactor binding site. First, all phosphorylases so far investigated possess stoichiometric amounts of pyridoxal phosphate, strongly suggesting that this compound is directly involved in catalysis and, therefore, that its binding site may be part of the active site of the enzyme. Second, since the exact mode of binding of P-Pyr[1] to the native molecule has not been established, characterization of the phosphopyridoxyl peptide might shed some light on this problem and explain some of the unusual spectral properties of the enzyme. Third, all glycogen phosphorylases consist of a basic subunit of very similar size (mol wt *ca.* 100,000; Seery *et al.*, 1967, 1970) suggesting that they might all have originated from the same ancestral gene. Information regarding the comparative amino acid sequence of the pyridoxal phosphate site could be of considerable in-

* From the Department of Biochemistry, University of Washington, Seattle, Washington 98105. *Received March 1, 1971.* This work was supported by grants from the National Science Foundation (GB 20482) and the National Institutes of Health (AM 07902).

† Present address: Department of Medicine, Clinical Research Center, University of Washington, Harborview Medical Center, Seattle, Wash. 98104.

‡ Present address: Department of Biological Sciences, University of Southern California, Los Angeles, Calif. 90007.

§ Present address: Physiologie Microbienne, Institut Pasteur, Paris, France.

|| To whom to address correspondence.

[1] Abbreviations used are: Hse, homoserine; Asx, aspartic acid or asparagine; Glx, glutamic acid or glutamine; DNS or dansyl, dimethylaminonaphthalene-5′-sulfonyl; Pxy, pyridoxyl; P-Pxy, pyridoxyl 5′-phosphate.