A Partial Amino Acid Sequence in the Heavy Chain of a Rabbit Antibody to Group C Streptococcal Carbohydrate*

Julian B. Fleischman†

ABSTRACT: Amino acid sequence studies in certain regions of rabbit IgG heavy chains have been hampered by heterogeneity of the material. To overcome this, an antibody with uniform properties to group C streptococcal carbohydrate was isolated from the serum of a single rabbit. The antibody was mainly of a2 allotype; it showed restricted electrophoretic mobility and was more than 90% precipitable with group C carbohydrate. A partial sequence of 65 residues was obtained in the N-terminal part of its heavy chain, including a section hitherto unreported in any rabbit IgG. Since the quantities of peptides studied were very small, no final conclusion about sequence homogeneity could be made. A comparison of the a2 sequence to those of the a1 and a3 allotypes revealed that about half of the first 34 residues are identical in all three allotypes.

Understanding the structural basis of antibody specificity demands a knowledge of the amino acid sequences of antibody polypeptide chains. Sequences are known for many myeloma globulin chains (Edelman and Gall, 1969) and for substantial regions of normal rabbit IgG heavy chains (Hill *et al.*, 1967; Wilkinson, 1969; Fruchter *et al.*, 1970). However, the sequences from residues 47 to 65 and from 100 to 115 in the Fd portion of the rabbit heavy chain are still unknown. The section from 110 to 115 may be important for binding antigen (Porter, 1970; Fruchter *et al.*, 1970). Peptides from this latter segment have not been recovered in good yield probably because normal IgG and most antibodies are quite heterogeneous and contain mixtures of different sequences in this section.

Certain rabbit antibodies to bacterial polysaccharides, formed in large amounts, are considerably more uniform than normal rabbit IgG (Krause, 1970; Haber, 1970). This uniformity also extends to the N-terminal amino acid sequences of their light chains (Eichmann *et al.*, 1970; Jaton *et al.*, 1970; Hood *et al.*, 1970). This type of antibody may therefore be better than normal rabbit IgG for sequence studies in the highly variable parts of the heavy chain.

There is a second reason for these studies. Certain substitutions in the variable region of rabbit IgG heavy chains appear to be correlated with the a-locus allotypes (Wilkinson, 1969; Koshland *et al.*, 1968, 1969). It is important to establish this point. If true, it would be difficult to reconcile with germ-line models for antibody variability (Hood and Talmage, 1970). Partial sequences have been determined for the a1 and a3 allotypes (Wilkinson, 1969), but none so far for the a2 allotype. Such a sequence would help identify the allotype-related residues.

To approach these problems, a relatively homogeneous antibody to the specific carbohydrate of group C streptococci was purified from the serum of a single rabbit. The antibody was predominantly of a2 allotype. Sequence studies were begun on the heavy chains in the hope of recovering better yields of peptides in the variable sections than could be obtained from normal rabbit IgG. This paper describes the preparation of the antibody, the partial sequence of the first 65 amino acid residues in the N-terminal portion of its heavy chain, and compares this sequence (allotype a2) with those of allotypes a1 and a3.

Materials and Methods

Immunization. A single New Zealand red rabbit (No. R26-90), allotype a2, a3, b4, A11, A12, was injected intravenously three times weekly for four weeks with group C streptococcal vaccine (Osterland *et al.*, 1966; Braun and Krause, 1968). After a 6-months rest, it was again injected intravenously three times weekly for 3 weeks. It was then bled three times, on alternate days, during week 4.

Purification of the Antibody. The major antibody component in the serum was isolated by preparative electrophoresis in Agarose gels (Seakem, 0.5%) (Eichmann *et al.*, 1970). This method generally permits higher recoveries of antibody than does isolation from specific immune precipitates (Braun and Krause, 1968). Serum (5 ml) was loaded in each gel. The total protein eluted from each gel fraction was determined on a Technicon Autoanalyzer programmed for Lowry protein determinations (Eichmann *et al.*, 1970). The peak fractions corresponding to the major antibody component (Figure 1) were pooled and concentrated by pressure dialysis in collodion membranes. The antibody was further purified by a second cycle of preparative electrophoresis as described above, and the peak fractions from each gel were again pooled and concentrated. The serum and purified antibody fractions were characterized by zone electrophoresis on cellulose acetate membranes (Osterland *et al.*, 1966). Affinity chromatography of the serum on an immunoadsorbent column is described by Eichmann and Greenblatt (1971).

Immunochemical Methods. Quantitative precipitin analysis and adsorption of serum with group C carbohydrate followed procedures of Osterland *et al.* (1966), Braun and Krause (1968), and Eichmann *et al.* (1970). The a- and b-locus allotype determinations are described by Kindt *et al.* (1970c) and the A11, A12 quantitations by Kindt *et al.* (1970b).

Preparation of Heavy Chains. The antibody (14.5 mg/ml) was mildly reduced with 0.01 M dithiothreitol in 0.5 M Tris-

* From the Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, England. *Received December 21, 1970.* Supported in part by Grants AI-09846 and AI-08429 from the National Institutes of Health.

† Present address: Department of Microbiology, Washington University School of Medicine, St. Louis, Mo. 63110. Reprint requests should be sent to this address. The work was done during the tenure of an Established Investigatorship of the American Heart Association.

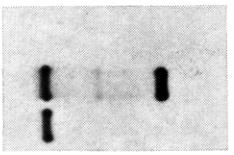

FIGURE 1: Microzone electrophoresis patterns of R26–90 immune serum (top frame), and of the purified antibody (bottom frame). Anode is at the right.

HCl buffer (pH 8.0), containing 0.002 M EDTA, for 1 hr at room temperature. It was alkylated with 50 mg of iodoacetic acid at 0° for 45 min, and the reagents were removed on a Sephadex G-25 (coarse) column in 1 M propionic acid. Heavy chains were separated from light chains on a Sephadex G-100 column in 1 M propionic acid (Fleischman *et al.*, 1963) and lyophilized.

Cyanogen Bromide Cleavage. Heavy chains (275 mg) were dissolved in 10 ml of 100% formic acid, and diluted to 14 ml with water. CNBr (1.4 g) was added and the solution kept at 25° for 24 hr. The protein was lyophilized, dissolved in 10 ml of 6 M urea–sodium formate buffer (pH 3.3), and applied to a Sephadex G-100 column (3 × 141 cm) in the same buffer (Fruchter *et al.*, 1970). The fractions corresponding to CA, CB, and C-1 (Figure 4 in Fruchter *et al.*, 1970) were pooled together and desalted in two batches on a Sephadex G-25 (coarse) column (3 × 61 cm) in 1 M acetic acid. The desalted protein batches, designated "C1," were combined and lyophilized.

Complete Reduction of Fragment C1. C1 was dissolved to a final concentration of 10 mg/ml in 6 M guanidine-HCl–0.5 M Tris-HCl buffer (pH 8.2), containing 0.002 M EDTA. It was reduced with DTT (0.0035 M, 2 hr at 37°), cooled to room temperature, and the protein thiol groups were alkylated with [^{14}C]iodoacetic acid (O'Donnell *et al.*, 1970). The reduced-alkylated C1 was applied directly to a Sephadex G-100 column (3 × 141 cm) in 6 M urea–sodium formate buffer (pH 3.3). The elution diagram is shown in Figure 3. The fractions in the major peak (C1) were pooled, desalted on a Sephadex G-25 (coarse) column (3 × 63 cm) in 1 M acetic acid, and lyophilized.

Citraconylation and Tryptic Digestion of Completely Reduced-Alkylated C1. The ϵ-NH_2 groups of lysine residues in C1 were blocked with citraconic anhydride (Dixon and Perham, 1968) by the following procedure (L. E. Mole, 1970, personal communication): 60 mg of lyophilized C1 was suspended in 4 ml of distilled water, and the suspension was brought to pH 8.0 with 5 N NaOH. Citraconic anhydride (0.32 ml, 400 mg) was added in four aliquots of 0.08 ml each, with vigorous stirring, allowing each aliquot to dissolve completely before adding the next. The pH was maintained carefully at 7.5–8.5 by the manual addition of 5 N NaOH through a micrometer syringe. The final clear solution was dialyzed against two changes of 1 l. of 0.5% NH_4HCO_3 (pH 7.8) at 5°.

The citraconylated C1, in 0.5% NH_4HCO_3, was digested with trypsin (1:50 by weight) for 4 hr at 37°. The trypsin was added in four equal aliquots at 0.5-hr intervals during the first 2-hr digestion. The pH remained at 7.8–8.2 throughout. Following digestion, the protein was lyophilized. The citraconyl groups were removed by suspending the lyophilized digest in 5 ml of pyridine–acetate buffer (pH 3.5) (Bennett, 1967) for 24 hr at room temperature. The digest was again lyophilized, redissolved in 100% formic acid, applied to a Sephadex G-50 (fine) column (2 × 230 cm), equilibrated with 50% formic acid, and eluted with 50% formic acid.

Characterization of Radioactive Peptides. The peptide fractions containing [^{14}C]*S*-carboxymethylcysteine were characterized according to O'Donnell *et al.* (1970). Aliquots of each fraction (0.01–0.02 μmole) were digested with trypsin plus chymotrypsin (10 μg each) in 0.1 ml of 1% NH_4HCO_3 at 37°. After 4 hr, an additional 10 μg of each enzyme was added and the mixture incubated for another 4 hr. The digests were lyophilized, redissolved in distilled water, relyophilized twice more to ensure the removal of NH_4HCO_3, and electrophoresed on paper at pH 3.5 or 6.5. Autoradiography was carried out for 24–48 hr.

Other Methods. Chromatography of peptide fractions on DEAE-Sephadex A-25 is described in the text. Isolation of N-terminal Glp peptides by Pronase digestion and Dowex 50 chromatography has been described by Wilkinson *et al.* (1966). Tryptic fragments D and F were digested with trypsin or chymotrypsin in 1% NH_4HCO_3 according to Wilkinson (1969). The digests were fractionated first on a G-50 (fine) column (1.2 × 142 cm) in 0.05 M NH_4OH, and then by paper electrophoresis at pH 3.5 or 6.5. Analytical and preparative high-voltage paper electrophoresis (Press *et al.*, 1966) were performed on Whatman No. 1 paper. The hypochlorite–starch–iodide stain is described by Pan and Dutcher (1956), ninhydrin–collidine by Bennett (1967), and specific stains for arginine, tyrosine, and tryptophan by Easley (1965). Peptides were eluted from paper with 1 M acetic acid according to Edstrom (1968). Amino acid analyses were performed as described by Wilkinson (1969). Radioactivity was measured in vials containing 5 ml of scintillation fluid (Kinard, 1957) in a Unilux II (Nuclear-Chicago) scintillation counter. Amino acid sequences were determined by the dansyl-Edman method (Gray, 1967), identifying the dansyl derivatives by thin-layer chromatography on polyamide sheets (Woods and Wang, 1967). Peptides were digested with carboxypeptidases A and B (Ambler, 1967) in 0.1 ml of 1% NH_4HCO_3 at 37°. Digestion was terminated at the desired time by the addition of 0.6 ml of amino acid analyzer sample–dilutor buffer (pH 2.19) and the sample was loaded directly onto the analyzer. Peptide F1 (30 nmoles) was digested with papain (16 μg) in 0.08 ml of 0.05 M pyridine–acetate buffer (pH 4.3) plus mercaptoethanol (10 mM) and EDTA (1 mM), for 6 hr at 37°. Tryptophan was determined spectrophotometrically (Crumpton and Wilkinson, 1963), and qualitatively on paper by staining with Ehrlich's reagent (Easley, 1965). Assignment of free and amidated side-chain carboxyl groups was based on the molecular weight and electrophoretic mobility of the peptide at pH 6.5 (Offord, 1966).

Materials. The following enzymes were obtained from Worthington Biochemicals, Freehold, N. J.: trypsin L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone, α-chymotrypsin (three-times crystallized), papain (twice crystallized), CPA-DFP[1] (three-times crystallized), and CPB (lot COBC-8DA). Pronase-P was a product of Kaken Chemical Co. Ltd., Japan; DTT, A grade, was obtained from the Calbiochem Co., [^{14}C]iodoacetate from The Radiochemical Centre, Amersham, and citraconic anhydride from British Drug Houses, Ltd.

[1] Abbreviations used are: CPA, carboxypeptidase A; CPB, carboxypeptidase B; DTT, dithiothreitol; CMC, *S*-carboxymethylcysteinyl residue; Glp, pyrrolid-2-one-5-carboxyl residue. Other abbreviations are given in *Biochemistry 5*, 1445 (1966).

TABLE I: Allotype Determinations of the R26-90 Antibody.

	% Pptn by Antisera				
	a2	a3	b4	A11[c]	A12[c]
Sample 1[a]	50	18	98	25	75
Sample 2[b]	54	3	100	<10	>90

[a] Material used for sequence analysis, prepared by agarose gel electrophoresis. [b] Aliquot prepared by affinity chromatography and agarose gel electrophoresis. [c] Determined according to Kindt *et al.* (1970b).

Results

R26-90 Antibody. Three bleedings from rabbit R26-90 yielded 68 ml of antiserum containing a total of 2666 mg of precipitable antibody to group C carbohydrate. About 1500 mg was in the major electrophoretically monodisperse IgG component as calculated from a densitometric tracing of the serum microzone pattern. The purified antibody from this peak, after two cycles of preparative electrophoresis, contained 741 mg of protein in 23.7 ml of 0.05 M Veronal buffer (pH 8.6). This preparation was more than 90% precipitable by group C streptococcal carbohydrate.

The microzone electrophoresis patterns of the immune serum and of the purified antibody are shown in Figure 1. The purified antibody is restricted to a single band corresponding to the main band in the immune serum. Figure 2 shows the microzone patterns of the immune serum before and after absorption at equivalence with group C streptococcal carbohydrate. Absorption removed the band corresponding to the main component of antibody demonstrating that this component is antibody directed to the group C streptococcal antigen.

The allotypes of the purified antibody are shown in Table I. Sample 1 is the main preparation and is principally a2 but contains some a3; 100% precipitation is rarely attainable with a-locus allotypes (T. J. Kindt, 1970, personal communication), so that the relative amounts of each allotype are only an approximation. Sample 2 was purified from an aliquot of R26-90 immune serum by affinity chromatography (Eichmann and Greenblatt, 1971) and subsequent preparative electrophoresis. It contained more than 60% of the antibody in the original aliquot of serum. Its microzone electrophoresis pattern was identical with that of sample 1 (see Figure 1), but it contained only 3% of a3 molecules. This confirmed the evidence that the principal antibody component is a2.

N-Terminal Peptide of R26-90 Heavy Chain. The R26-90 heavy chains were separated from light chains by mild reduction, alkylation with unlabeled iodoacetate, and gel filtration, as described under Methods. Pronase digestion of 10 mg of completely reduced and alkylated R26-90 heavy chains and subsequent Dowex 50 chromatography yielded 1.1 moles of the Glp-Ser peptide per mole of heavy chain. Only trace amounts of the other N-terminal peptides found in normal rabbit IgG heavy chains (Wilkinson, 1969) were detected. This suggested that the primary structure of R26-90 heavy chain, at least in the N-terminal region, was more uniform that that of normal rabbit IgG heavy chain.

C1 Fragment. The C1 fragment was prepared from unlabeled heavy chain by cyanogen bromide cleavage (see Methods) and was then completely reduced and alkylated with [^{14}C]iodoacetic acid as described under Methods. Figure 3 shows the final purification of C1 after complete reduction and alkylation. Approximately 2.5 μmoles of C1 was recovered from 5 μmoles of heavy chain. The radioactivity in C1 probably represents [^{14}C]CMC residues I_{22}, II_{97}, V_{146}, and VI_{201}.[2] The other cysteine residues had been previously reduced under mild conditions and had been alkylated with unlabeled iodoacetate (O'Donnell *et al.*, 1970). If the labeling were uniform, the radioactivity would correspond to 2.5×10^6 cpm/μmole of labeled cysteine.

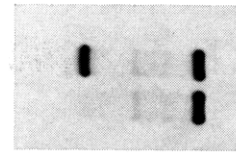

FIGURE 2: Microzone electrophoresis patterns of R26-90 immune serum (top frame) and of the serum after absorption with group C carbohydrate (bottom frame). Anode is at the right.

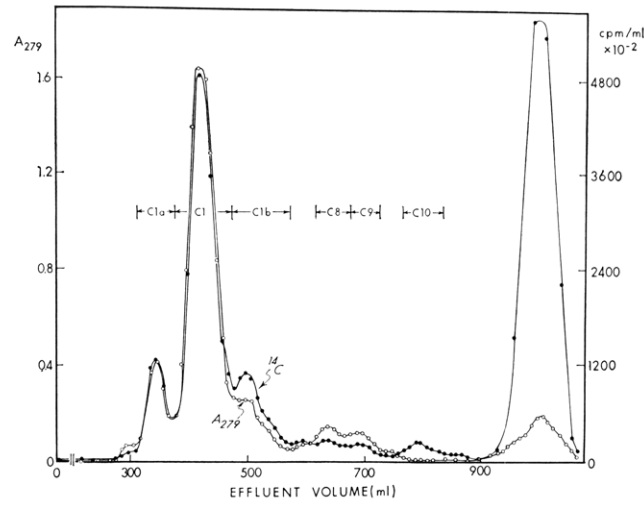

FIGURE 3: Elution diagram of reduced-alkylated C1 on a Sephadex G-100 column (3 × 141 cm) in 6 M urea–sodium formate buffer (pH 3.3). (O–O) Absorbancy$_{279}$; (●–●) ^{14}C counts. The large radioactive peak at the right is excess [^{14}C]iodoacetate eluted at total column volume.

The amino acid analysis of C1 is shown in Table II. There are 1.2 residues of homoserine/mole of C1, assuming 17 leucine residues (Cebra *et al.*, 1968). This is substantially less than the 1.8 residues of homoserine reported for C1 from normal rabbit IgG (Cebra *et al.*, 1968). The latter value was attributed by Cebra *et al.* (1968) to the presence of fractional methionyl residues in peptides remaining disulfide bonded to or aggregated with the main C1 fragment. Digestion of 5 mg of R26-90 C1 with Pronase yielded 0.8 mole of Glp-Ser/mole of C1. The R26-90 C1 fraction thus appears to contain mainly one N-terminal sequence, and one homoserine at the C terminus.

Fragments from a Tryptic Digest of Citraconylated C1. Citraconylated C1 was digested with trypsin. After blocking groups were removed, the digest was fractionated on a column of Sephadex G-50 (Figure 4). Aliquots of every second fraction were scanned for [^{14}C]CM-cysteine-containing peptides by digestion, paper electrophoresis of the digests at pH 3.5,

[2] The nomenclature for half-cystinyl residues in immunoglobulins is given in *Biochemistry 9*, 3469 (1970).

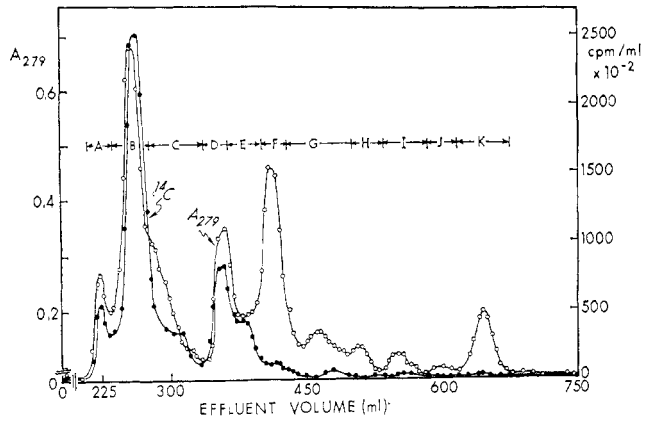

FIGURE 4: Elution diagram of a tryptic digest of citraconylated C1 on a Sephadex G-50 (fine) column (2 × 230 cm) in 50% formic acid. (O–O) Absorbancy$_{279}$; (●–●) ^{14}C counts.

and autoradiography as described under Methods. The radioactive peptides (O'Donnell *et al.*, 1970) characteristic of cysteines V_{146} and VI_{201} were found in the digests of fraction B, and of those characteristic of cysteine I_{22} were found in the digests of fractions D and E. A small amount of a second peptide derived from cysteine V_{146} was also found in the digests of fraction E. Only trace amounts of the CMC-Ala-Arg peptide from cysteine II_{97} were detected, suggesting that the sequence around cysteine II_{97} may be different in the R26-90 γ chain.

Thus, fractions D and E were identified as being derived from the N-terminal section. Fraction F contained little radioactivity, a high absorption at 279 nm, and by size and composition resembled the section from residues 41 to 68 in the Daw and Cor heavy-chain sequences (Press and Hogg, 1970). The major fragments in these fractions were isolated and sequenced.

Chromatography of Fractions D and E. Fractions D and E were concentrated, dried, redissolved in 0.005 M NH_4HCO_3, combined (total volume 3.0 ml), clarified by centrifugation at 8000 rpm for 0.5 hr, and applied to a DEAE-Sephadex A25

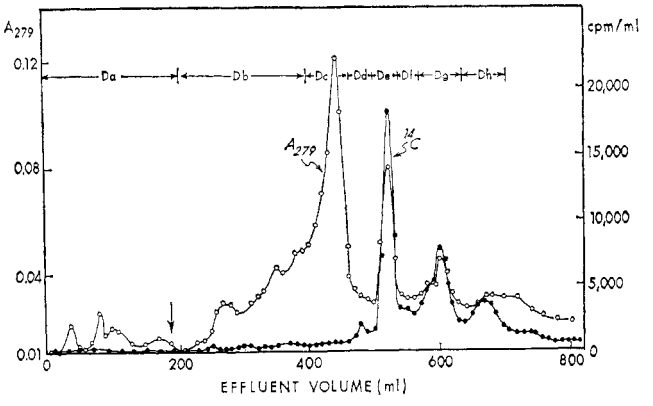

FIGURE 5: Elution diagram of the combined fractions D and E on a DEAE-Sephadex A-25 column (1.8 × 20 cm). The sample was applied in 3.0 ml of 0.005 M NH_4HCO_3 (pH 8.5) to the column equilibrated in the same buffer. After elution with 190 ml of 0.005 M NH_4HCO_3 (pH 8.5) a linear gradient was begun (arrow) to 0.5 M NH_4HCO_3 (pH 8.5) with 350 ml in each open chamber. (O–O) Absorbancy$_{279}$; (●–●) ^{14}C counts. Fractions were also read at 225 mμ. The profile (not shown) was similar to that at 279 mμ.

TABLE II: Amino Acid Compositions of R26-90 Heavy Chain and Fragments:[a]

Amino Acid	Heavy Chain	C1	Fragment De (Fragment D)	Fragment F
Lys	22	10	2.5 (2)	0.9 (1)
His	4.8	1.0		
Arg	15	5.4	1.1 (1)	0.8 (1)
CMC	n.d.	6.8	0.7 (1)	
Asp	34	16	2.5 (2)	2.5 (2)
Thr	50	31	4.6 (5)	2.1 (2)
Ser	55	26	5.4 (6)	2.6 (2)
Glu	34	10	3.5 (3)	2.5 (2)
Pro	38	19	1.9 (1)	2.4 (1)
Gly	31	20	4.2 (4)	4.7 (5)
Ala	22	13	0.6	3.6 (4)
Val	36	19	3.5 (4)	
Met	5.3			
Ile	15	5.8	1.0 (1)	2.0 (2)
Leu	31	17	3.9 (4)	1.4 (1)
Tyr	12	7.4	0.9 (1)	2.2 (3)
Phe	12	5.3	1.1 (1)	
Hsr		1.2		
Trp[b]	n.d.	n.d.	(1)	1.5 (2)
Total residues[c]	417	214	37	28
Per cent yield[d]		50	15	13

[a] Numbers are residues per mole of peptide. Values less than 0.1 are omitted. Numbers in parentheses are integral values either assumed or confirmed by sequence analysis. Values for C1 are based on 17 residues of leucine (Cebra *et al.*, 1968). [b] Assumed to be 1 in fragment D from sequence data; determined spectrophotometrically (Crumpton and Wilkinson, 1963) in fragment F. [c] Excluding cysteine and tryptophan in heavy chain; excluding tryptophan in C1. [d] Yield of C1 based on 5 μmoles of heavy chain digested with CNBr; yields of fragments D and F based on 1 μmole of citraconylated C1 digested with trypsin. n.d. = not determined.

column equilibrated in 0.005 M NH_4HCO_3 (pH 8.5). The conditions for gradient elution and the elution pattern are given in Figure 5. The major symmetrical radioactive peak, De, was pooled and concentrated. It had an amino acid composition (Table II) similar to that of the N-terminal fragment of C1 from a2 heavy chains (Prahl and Porter, 1968). Since it was the radioactive peptide recovered in largest yield (0.14 μmole; 37% of the radioactivity applied to the column) from the combined D and E fraction, it is probably the main N-terminal fragment of C1. The pooled fractions of peak Dg had an amino acid composition similar to that of De, but peak Dg was asymmetric and was recovered in smaller yield (0.09 μmole). Neither Dc nor any of the other peaks contained more than 0.05 μmole of peptide by analysis. Sequence studies were accordingly begun on De, henceforth called fragment D.

Purification of Fraction F. Fraction F from the tryptic digest of C1 (Figure 4) was concentrated and further purified on a Sephadex G-50 (fine) column (2 × 230 cm) in 0.05 M NH_4OH. A major peak and a faster moving shoulder were resolved,

TABLE III: Amino Acid Compositions and Properties of Peptides from Fragment D.[a]

Amino Acid	Peptide										
	D1	D2[b]	D3	D4	D5	D6	D7	D8	D9	D10	D11
Lys	0.9 (1)		1.0 (1)		0.6 (1)						
Arg			0.8 (1)								1.0 (1)
CMC			+ (1)			0.8 (1)		1.0 (1)	0.7 (1)		
Asp			2.4 (2)		0.8 (1)	1.3 (1)		1.3 (1)	1.1 (1)		0.3
Thr			4.8 (5)		1.6 (2)	2.7 (3)	1.0 (1)	2.0 (2)	1.5 (2)		
Ser	1.2 (1)		4.8 (5)	1.4 (1)	0.6	2.8 (3)		2.9 (3)	1.6 (1)	1.5 (1)	0.6
Glu	1.1 (1)		2.5 (2)	2.0 (2)	0.3	0.4		0.4	0.7		0.5
Pro			2.8 (1)		1.3 (1)						
Gly	0.6		4.5 (4)	2.3 (2)	0.5	1.5 (1)		1.6 (1)	1.6 (1)	1.6 (1)	0.7
Ala					0.2	0.2		0.2	0.3		0.3
Val	1.0 (1)		2.5 (3)			1.0 (1)		1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)
Ile			1.0 (1)			1.0 (1)		1.0 (1)	0.9 (1)		
Leu			4.1 (4)	1.0 (1)	1.0 (1)	2.1 (2)	1.0 (1)	1.2 (1)	1.3 (1)		
Tyr			0.8 (1)			0.9 (1)		0.8 (1)			
Phe			1.2 (1)	0.9 (1)							
Trp[c]			+ (1)							+ (1)	
Total residues	4		33	7	6	14	2	12	9	4	2
Per cent yield[d]	57		48	24	57	7	9	22	7	14	20
G-50 elution position[e]	0.79		0.45	0.45	0.55	0.42	0.70	0.42	0.42	0.89	0.89
Mobility[f]											
pH 6.5				−0.59		−0.35		−0.41	−0.50		
pH 3.5		−0.74								+0.26	+0.80
pH 1.9							+0.50				
Staining properties[g]											
Ninhydrin–collidine		−		Blue		Blue	Blue	Blue	Blue	Yellow	Blue
Hypochlorite–starch–iodide		+		+		+		+	+		
Tryptophan										+	−
Arginine										−	+
Tyrosine				−				+	−		
Autoradiography[h]				−		+		++	+		

[a] Numbers are residues per mole of peptide. Values less than 0.1 are omitted. Numbers in parentheses are integral values either assumed or confirmed by sequence analysis. [b] Peptide D2, derived from D1 by Pronase digestion, was not analyzed but was identified as Glp-Ser by its staining properties and mobility at pH 3.5 (Wilkinson, 1969). [c] Tryptophan content was assumed to be 1 for D3 and for chymotryptic peptides staining with Ehrlich's reagent on paper. [d] Yields of D1 and D3 based on 145 nmoles of fragment D digested with trypsin; yields of D4–D11 based on 69 nmoles of peptide D3 digested with chymotrypsin. [e] Gel filtration on Sephadex G-50 in 0.05 M NH_4OH. Numbers are fractions of total column volume. [f] Calculated according to Offord (1966). Expressed at pH 6.5 relative to Asp (= −1.00); at pH 3.5 relative to Glp (= −1.00) for D2 and relative to methyl green (= +1.00) for other peptides; and at pH 1.9 relative to methyl green (= +1.00). [g] + = positive reaction; − = no reaction; no entry means that the test was not done. [h] Symbols indicate the relative intensities of spots on the film.

both of which had identical amino acid compositions. That of the major peak, called fragment F, is given in Table II. Fragment F was pooled, concentrated, and used for subsequent sequence studies.

Amino Acid Sequences of Fragments D and F. Fragment D was digested with trypsin and fractionated on Sephadex G-50 (fine) in 0.05 M NH_4OH. The two major products were peptides D1 and D3. Part of D1 was digested with Pronase (Wilkinson, 1969) and another part with carboxypeptidases B and A. D3 was digested with chymotrypsin, refractionated on Sephadex G-50 (fine) in 0.05 M NH_4OH, and the products (D4–D11) were isolated by paper electrophoresis. The digestion products and their amino acid compositions and sequence are given in Table III and Figure 6.

Fragment F was digested with chymotrypsin, fractionated on Sephadex G-50 in 0.05 M NH_4OH, and the products were isolated by paper electrophoresis. Their amino acid compositions and sequences are given in Table IV and in Figure 7. Peptide F1 was isolated in a blocked N-terminal (Glp) form. Its partial sequence was determined by digesting it with papain at pH 4.3 (see Methods) (Press, 1967), producing peptides F2–F5.

There was insufficient material to determine overlapping sequences. Figures 6 and 7 indicate the proposed alignments

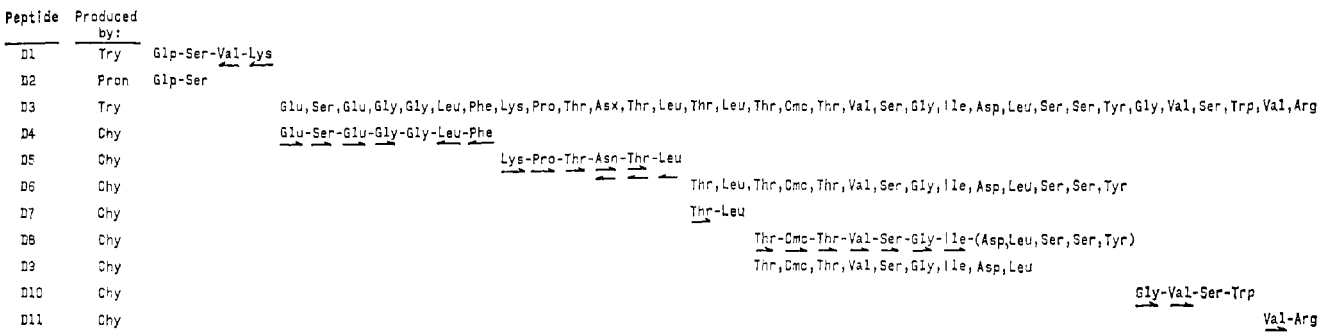

FIGURE 6: Amino acid sequence determinations of peptides derived from fragment D. Next to each peptide is listed the protease which released it (Try = trypsin, Pron = Pronase, Chy = chymotrypsin). The amino acid compositions and properties of the peptides are given in Table III. → indicates residue identified by dansyl-Edman or subtractive-Edman method. ⇀indicates residue released by carboxypeptidase A or B. The position of CMC in D8 was established by the removal of 75% of the ^{14}C counts from the peptide after the second Edman step. The horizontal alignment of the peptides is based on homology with other sequences (see text).

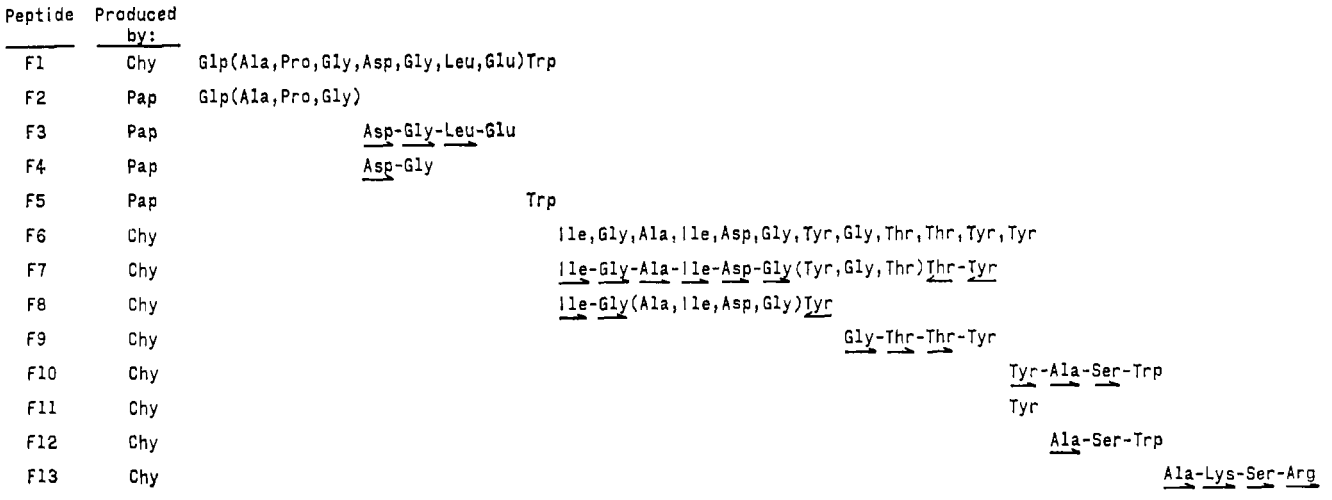

FIGURE 7: Amino acid sequence determinations of peptides derived from fragment F. Next to each peptide is listed the protease which released it (Chy = chymotrypsin, Pap = papain). The amino acid compositions and properties of the peptides are given in Table IV. →indicates residue identified by dansyl-Edman or subtractive-Edman method. ⇀indicates residue released by carboxypeptidase A. The horizontal alignment of the peptides is based on homology with other sequences (see text).

of the peptides. The alignment of peptides D1–D9 is based on their homology with the sequences of Wilkinson (1969). Peptide D11, containing C-terminal arginine, is assigned to the C terminus of fragment D. Peptide F13 is likewise placed at the C terminus of fragment F. Peptides D10, D11, and F1–F5 were aligned by analogy with sequences of tryptic and chymotryptic peptides derived from normal rabbit C1 which overlap the N-terminal cyanogen bromide fragment (S. A. Jackson and J. M. Wilkinson, 1970, personal communication), and with sequences of human heavy chains (Press and Hogg, 1970; Cunningham *et al.*, 1970). Peptide F6, containing three tyrosine residues, aligns peptides F7 and F10. The partial sequence of the 65 residues comprising both fragments D and F is shown in Figure 8.

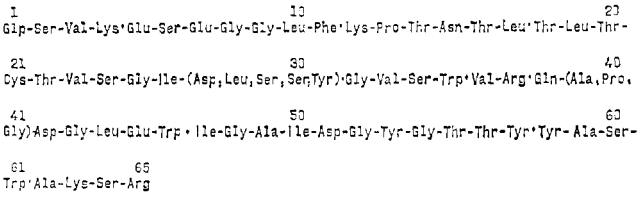

FIGURE 8: Partial sequence of the first 65 amino acid residues in the R26-90 heavy chain, comprising fragments D (1–37) and F (38–65). Dots between the peptides indicate that there was insufficient material to determine overlaps, and that the alignment is based on homology with other sequences (see text). The Cys residue at position 21 corresponds to Cys-22 in sequences where a deletion has been introduced at position 2 (see Figure 9).

Discussion

The R26-90 antibody, though not entirely homogeneous, was much more uniform than normal rabbit IgG and had certain properties resembling those of myeloma proteins. It migrated as a single band on microzone electrophoresis (Figure 1) and it selectively expressed the a2 and A12 allotypes (Table I). The relative yields of individual amino acid residues in the N-terminal sequence of the R26-90 light chain was comparable to those obtained from Bence-Jones proteins (Hood *et al.*, 1970; L. Hood, 1970, personal communication). In the studies reported here, uniform structures were found at the N termini of the heavy chain and of the C1 fragment. Peptides were isolated in good yield from very small amounts of protein, and a defined sequence was elucidated in the C1

TABLE IV: Amino Acid Compositions and Properties of Peptides from Fragment F.[a]

Amino Acid	Peptide												
	F1	F2	F3	F4	F5[b]	F6	F7	F8	F9	F10	F11[h]	F12	F13
Lys													0.8 (1)
Arg													1.0 (1)
Asp	1.0 (1)		0.9 (1)	1.0 (1)		1.6 (1)	1.0 (1)	1.0 (1)	0.3				
Thr						2.0 (2)	1.7 (2)	0.2	2.0 (2)				
Ser	0.1	0.2	0.5	0.3	0.5	1.3	0.3	0.4	0.4	1.0 (1)	0.4	1.0 (1)	1.1 (1)
Glu	2.1 (2)	1.1 (1)	1.1 (1)	0.3		0.7	0.2	0.2					
Pro	1.7 (1)	1.1 (1)											
Gly	2.1 (2)	1.0 (1)	1.4 (1)	1.5 (1)	0.3	3.9 (3)	2.8 (3)	1.8 (2)	1.2 (1)	0.3	0.4	0.2	0.3
Ala	1.0 (1)	1.0 (1)	0.3			1.4 (1)	1.0 (1)	1.0 (1)	0.2	1.0 (1)		1.0 (1)	1.0 (1)
Ile						2.0 (2)	1.8 (2)	1.5 (2)					
Leu	1.0 (1)	0.2	1.0 (1)										
Tyr						2.7 (3)	1.8 (2)	0.7 (1)	1.3 (1)	0.9 (1)	1.0 (1)		
Trp[c]	+ (1)				+ (1)					+ (1)		+ (1)	
Total residues	9	4	4	2	1	12	11	7	4	4	1	3	4
Per cent yield[d]	37	38	16	14	10	4	29	11	7	17	12	17	22
G-50 elution position[e]	0.57					0.53	0.53	0.53	0.66	0.89	0.89	0.89	0.87
Mobility[f]													
pH 6.5	−0.75	−0.59	−0.85	−0.81	+0.05	−0.17	−0.21	−0.29					
pH 3.5									+0.23	+0.21	+0.05	+0.27	+1.00
pH 1.9					+0.11								
Staining properties[g]													
Ninhydrin–collidine	−	−	Blue	Blue	Gray	Blue	Blue	Blue	Yellow	Blue	Brown	Blue	Blue
Hypochlorite–starch–iodide	+	+											
Tryptophan	+	−	−	−	+	−	−	−	−	+	−	+	−
Arginine										−	−	−	+
Tyrosine									+	±	±	−	−

[a] Numbers are residues per mole of peptide. Values less than 0.1 are omitted. Numbers in parentheses are integral values either assumed or confirmed by sequence analysis. [b] Peptide F5 had mobilities at pH 6.5 and pH 1.9 and staining properties identical with those of free tryptophan. [c] Tryptophan content was assumed to be 1 for chymotryptic peptides staining with Ehrlich's reagent on paper. [d] Yields of F1 and F6–F13 based on 125 nmoles of fragment F digested with chymotrypsin; yields of F2 through F5 based on 30 nmoles of F1 digested with papain. [e] Gel filtration on Sephadex G-50 in 0.05 M NH_4OH. Numbers are fractions of total column volume. [f] Calculated according to Offord (1966). Expressed at pH 6.5 relative to Asp (= −1.00); and at pH 3.5 and 1.9 relative to methyl green (= +1.00). [g] + = positive reaction; − = no reaction; no entry means that the test was not done. [h] Both hydrolyzed and unhydrolyzed preparations of F11 had the same amino acid composition.

fragment. It is difficult, however, to draw an absolute conclusion about the sequence homogeneity of this material, since with the small quantities studied (145 nmoles of fragment D; 125 nmoles of fragment F), trace amounts of peptides containing sequence variations would have been missed.

Other technical difficulties arose from working with small amounts of material. Nonstoichiometric compositions are evident in the amino acid analyses of several peptides (Tables II, III, and IV). These may have arisen either from contaminating amino acids eluted from chromatography paper, or from smaller amounts of homologous peptides derived from other heavy chains in the preparation. Serine and glycine are prominent contaminants in certain peptides (D9, D10, D11, F3, F5, F6, and F11), suggesting that the contamination arose by elution from paper. Since the quantities of the peptides are small, the contaminants are proportionately high. Nevertheless, during the sequence determinations, dansyl derivatives of the contaminating amino acids were not detected in significant amounts. In most cases, the level of the contaminants was markedly reduced after the first Edman step, suggesting that they were present as free amino acids, rather than as peptides.

The analytical values for proline in certain cases were higher than those found by sequence analysis (fragments D and F in Table II; peptides D3 and F1 in Tables III and IV). There was a small but persistent distortion in the 440-mμ analyzer trace, close to the elution position of proline. This did not affect proline values in larger amounts of peptides, but it sometimes caused spuriously high proline values in the small samples analyzed in this work.

The main component in the antibody preparation had allotypes a2 and A12. The a3 and A11 molecules were probably contaminants unrelated to the main component. This conclusion is supported by the recovery of the main component essentially free of a3 and A11 molecules after affinity chromatography on an immunoadsorbent (sample 2, Table I). The a2

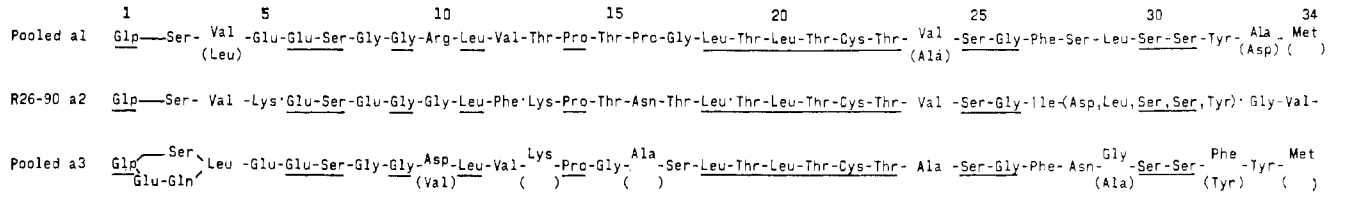

FIGURE 9: A comparison of the amino acid sequences of the three a-locus allotypes from residues 1 to 34 in the rabbit IgG heavy chain. The numbering and the a1 and a3 sequences are from Wilkinson (1969). Residues underlined are identical in all three sequences. Those in parentheses in a1 and a3 are minor or unidentified variants. The dash at position 2 is a deletion to maximize homology.

molecules probably also carry the A12 determinants since (a) the proportions of a2 and A12 are similar (Table I), and (b) the a2 and A12 markers are linked in the colony of rabbits from which R26-90 was obtained (T. J. Kindt and K. Eichmann, 1970, personal communication; Kindt *et al.*, 1970a,b).

The main C1 fragment (Figure 3), isolated by cyanogen bromide cleavage of the heavy chains and subsequent complete reduction, was probably nearly free of fragments derived from the a3 heavy chains. Pooled, heterogeneous a3 and a2 heavy chains have fractional methionyl residues amino-terminal to the C-terminal residue of C1 (Prahl and Porter, 1968; Wilkinson, 1969). There are four methionyl residues in the constant region of A12 heavy chains, and no additional methionyl residues were found amino-terminal to the C-terminal residue of the main R26-90 C1 fragment. However, the original R26-90 heavy-chain preparation had 5.3 methionyl residues (Table II). The total homoserine content of the smaller peaks in Figure 3 (C1b, C8, C9, C10) accounted for most of the extra 1.3 methionyl residues. Also, shorter N-terminal fragments were found in peaks C8, C9, and C10 (Figure 3) by Pronase digestion and Dowex 50 chromatography. Thus the smaller peaks in Figure 3 may contain cyanogen bromide fragments split at fractional methionyl residues in the N-terminal portions of heterogeneous a3 and a2 chains contaminating the preparation of R26-90 heavy chains. Such fragments could account for some of the extra methionine found in the heavy-chain composition (Table II), but the correct assignment of all the methionyl residues in the R26-90 heavy chain must await complete characterization of all fragments produced by cyanogen bromide.

Figure 9 compares the partial sequence of the R26-90 (a2) heavy chain with those of pooled a1 and a3 heavy chains from positions 1 to 34 (Wilkinson, 1969). Sixteen residues (47%) are identical in all three allotypes and thus cannot be allotype related. The residues which are not underlined could be either allotype or idiotype related. Some of them agree with amino acid composition differences between the a2 and other allotypes in this part of the heavy chain (Prahl and Porter, 1968; Koshland *et al.*, 1969). This is consistent with the view that allotype-related sequence variations may be present in the variable region. Identification of such residues must await further sequences in the Fd regions of pooled, allotypically uniform heavy chains.

The relative uniformity of the R26-90 heavy chain raised the hope that it will be useful for further sequence studies in the Fd region. Of particular interest is the section from residue 100 to 115. This section has been difficult to isolate, probably due to its high variability in normal IgG. It may also play an important role in the antibody combining site (Fruchter *et al.*, 1970; Porter, 1970). It is now being investigated in the C1 fragment isolated from the R26-90 heavy chain.

Acknowledgments

I am very grateful to Dr. Richard M. Krause for providing the R26-90 immune serum and the laboratory facilities for purifying the antibody, and to Dr. K. Eichmann for his assistance in this work. The allotype determinations were kindly performed by Drs. T. J. Kindt and C. W. Todd. I am deeply grateful to Professor R. R. Porter for providing space in his laboratory at Oxford and for his guidance and encouragement during the sequence work; to Drs. E. M. Press and L. E. Mole for their help with sequencing methods, and to Mr. T. Gascoyne for the amino acid analyses.

References

Ambler, R. P. (1967), *Methods Enzymol. 11*, 155, 436.
Bennett, J. C. (1967), *Methods Enzymol. 11*, 330.
Braun, D. G., and Krause, R. M. (1968), *J. Exp. Med. 128*, 969.
Cebra, J. J., Givol, D., and Porter, R. R. (1968), *Biochem. J. 107*, 69.
Crumpton, M. J., and Wilkinson, J. M. (1963), *Biochem. J. 88*, 228.
Cunningham, B. A., Rutishauser, V., Gall, W. E., Gottlieb, P. D., Waxdal, M. J., and Edelman, G. M. (1970), *Biochemistry 9*, 3161.
Dixon, H. B. F., and Perham, R. N. (1968), *Biochem. J. 109*, 312.
Easley, C. W. (1965), *Biochim. Biophys. Acta 107*, 386.
Edelman, G. M., and Gall, W. E. (1969), *Annu. Rev. Biochem. 38*, 415.
Edstrom, R. D. (1968), *Anal. Biochem. 26*, 204.
Eichmann, K., and Greenblatt, J. (1971), *J. Exp. Med. 133*, 424.
Eichmann, K., Lackland, H., Hood, L., and Krause, R. M. (1970), *J. Exp. Med. 131*, 207.
Fleischman, J. B., Porter, R. R., and Press, E. M. (1963), *Biochem. J. 88*, 220.
Fruchter, R. G., Jackson, S. A., Mole, L. E., and Porter, R. R. (1970), *Biochem. J. 116*, 249.
Gray, W. R. (1967), *Methods Enzymol. 11*, 469.
Haber, E. (1970), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 29*, 66.
Hill, R. L., Lebovitz, H. E., Fellows, R. E., and Delaney, R. (1967), *Nobel Symp. 3*, 109.
Hood, L., Eichmann, K., Lackland, H., Krause, R. M., and Ohms, J. J. (1970), *Nature (London) 228*, 1040.
Hood, L., and Talmage, D. W. (1970), *Science 168*, 325.
Jaton, J. C., Waterfield, M. D., Margolies, M. N., and Haber, E. (1970), *Proc. Nat. Acad. Sci. U. S. 66*, 959.
Kinard, F. E. (1957), *Rev. Sci. Instrum. 28*, 293.
Kindt, T. J., Mandy, W. J., and Todd, C. W. (1970a), *Biochemistry 9*, 2028.

Kindt, T. J., Mandy, W. J., and Todd, C. W. (1970b), *Immunochemistry 7*, 467.
Kindt, T. J., Todd, C. W., Eichmann, K., and Krause, R. M. (1970c), *J. Exp. Med. 131*, 343.
Koshland, M. E., Davis, J. J., and Fujita, N. J. (1969), *Proc. Nat. Acad. Sci. U. S. 63*, 1274.
Koshland, M. E., Reisfeld, R. A., and Dray, S. (1968), *Immunochemistry 5*, 471.
Krause, R. M. (1970), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 29*, 59.
O'Donnell, I. J., Frangione, B., and Porter, R. R. (1970), *Biochem. J. 116*, 261.
Offord, R. E. (1966), *Nature (London) 211*, 591.
Osterland, C. K., Miller, E. J., Karakawa, W. W., and Krause, R. M. (1966), *J. Exp. Med. 123*, 599.
Pan, S. C., and Dutcher, J. D. (1956), *Anal. Chem. 28*, 836.
Porter, R. R. (1970), *Harvey Lect.* (in press).
Prahl, J. W., and Porter, R. R. (1968), *Biochem. J. 107*, 753.
Press, E. M. (1967), *Biochem. J. 104*, 30c.
Press, E. M., and Hogg, N. M. (1970), *Biochem. J. 117*, 641.
Press, E. M., Piggott, P. J., and Porter, R. R. (1966), *Biochem. J. 99*, 356.
Wilkinson, J. M. (1969), *Biochem. J. 112*, 173.
Wilkinson, J. M., Press, E. M., and Porter, R. R. (1966), *Biochem. J. 100*, 303.
Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.

Retention of Common Nucleotide Sequences in the Ribosomal Deoxyribonucleic Acid of Eukaryotes and Some of Their Physical Characteristics*

John H. Sinclair† and Donald D. Brown†

ABSTRACT: The rRNA from *Xenopus laevis* has been hybridized with DNA from over 50 organisms, including prokaryotes and eukaryotes.

In all cases the DNA was prefractionated by preparatory centrifugation in CsCl. Analysis of hybridization profiles yielded the following information. (1) Without exception, the DNA of all eukaryotes analyzed had some molecular homology with *X. laevis* rRNA; no homology was found between *X. laevis* rRNA and the DNA of 8 prokaryotes. (2) There is a general tendency for rDNA to have a higher G + C content than bulk DNA from the same organism. The approximate buoyant density of rDNA from 50 eukaryotes has been tabulated and from this value the base composition of each rDNA has been estimated. (3) High levels of redundancy exist for ribosomal genes in all species examined. In plants and invertebrates a larger fraction of genome codes for rRNA than in vertebrates; in mammals this fraction is the smallest observed. (4) Clustering of ribosomal genes as evidenced by sharp banding of rDNA in CsCl occurs in most of the species examined.

The segment of the eukaryote chromosome termed the "nucleolar organizer" contains a unique set of genes which determine the structure of ribosomal RNA (rRNA).[1] In *Xenopus laevis* several hundred genes for the 18S and 28S rRNAs are clustered together on a single allele of the 18 haploid chromosomes (Wallace and Birnstiel, 1966). Individual 18S and 28S genes are adjacent and separated from the next pair by a "spacer" region of DNA (Brown and Weber, 1968b). The DNA which contains these three repeating nucleotide sequences (18S, 28S, and spacer DNAs) is termed rDNA, and it has a high G + C content relative to bulk *X. laevis* DNA and bands at a higher buoyant density in CsCl. This separation of rDNA from bulk DNA by CsCl centrifugation, first described by Wallace and Birnstiel (1966), adds considerable information and specificity to molecular hybridization studies. Since the rDNA is separated from bulk DNA, contamination of radioactive rRNA preparations with DNA-like RNA does not interfere with the measurement of hybridization to rDNA. From the buoyant density of rDNA its approximate base composition can be calculated. Finally, the distribution of rDNA in the gradient gives an indication of the extent to which these redundant genes are clustered within a genome.

These experiments were designed to analyze some physical properties of rDNA from a variety of eukaryotes. The combined techniques of equilibrium centrifugation of DNA followed by hybridization with radioactive rRNA yielded a variety of information from a single preparation of DNA. From examination of these gradients, it has been possible to unequivocally demonstrate homology between the rRNA of *X. laevis* and the rDNA of distantly related eukaryotes, and to estimate the buoyant density (and consequently the base composition) of these ribosomal genes. Results also indicate that the ribosomal genome is highly redundant

* From the Department of Embryology, Carnegie Institution of Washington, Baltimore, Maryland 21210. *Received February 19, 1971.*
† Present address: Department of Zoology, Indiana University, Bloomington, Ind. 47401; supported by Grant NSF GB19450. To whom to address correspondence.

[1] Abbreviations used are: rRNA, ribosomal RNA, including 18S RNA and 28S RNA; rDNA, ribosomal DNA, the portion of bulk DNA which contains gene sequences homologous to 18S and 28S rRNA; MAK, methylated albumin kieselguhr; SSC, standard saline citrate (0.15 M NaCl–0.015 M sodium citrate).