Biochem. J. (1971) **124**, 301–318
Printed in Great Britain

Allotypically Related Sequences in the Fd Fragment of Rabbit Immunoglobulin Heavy Chains

By L. E. MOLE, S. A. JACKSON, R. R. PORTER and J. M. WILKINSON*
Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

(*Received* 1 *April* 1971)

The sequence has been completed of the *N*-terminal 94 residues of the variable section of the Fd fragment of heavy chains from rabbit immunoglobulin G (IgG) of allotype Aa1. Most of the sequence of the same section from IgG of allotype Aa3 is also reported. These results, in conjunction with a substantial sequence of the variable region of allotype Aa2 reported elsewhere (Fleischman, 1971), show the presence of 16 positions (including six consecutive positions) in which the residue present correlates with the allotype. No allotype-related sequence variation has been found in the constant section of the Fd fragment. This evidence supports the view that two genes code for the heavy chain and it can be used as evidence in favour of somatic mutation as the origin of the variability in the sequence of the *N*-terminal section. The evolutionary origin of the 'a' locus allotypes of rabbit immunoglobulins remains obscure.

It has been shown that, within one species, the sequence of the *C*-terminal half of the light chains of immunoglobulin is constant for a given type and a given allele. In the *N*-terminal half the sequences differ between each myeloma protein (Hilschmann & Craig, 1965) and presumably for those contained in the many different molecules of immunoglobulin in the blood of any normal animal. Similarly, in the heavy chains of myeloma proteins, the *N*-terminal 115 residues differ, whereas the *C*-terminal 330 residues are constant for each allelic form of each subclass (Press & Hogg, 1969; Edelman *et al.* 1969).

It appears that there are very many different sequences occurring in these *N*-terminal sections of both heavy and light chains and there are two opposing views as to the genetic origin of such variability. One is that there are corresponding numbers of genes in the germ line which code for all the sequences that arise. A substantial part of the DNA of the cell would be required for this purpose. The alternative view is that there is in the germ line only a small number of genes coding for the sequences of the heavy and light chains and that extensive diversification arises by somatic mutation or recombination early in embryonic life.

Study of the antigenic specificities of myeloma proteins and immunoglobulins of normal animals has shown that inherited variants occur and subsequent structural studies have shown that most of these allotypic specificities can be correlated with sequence differences in the amino acid residues in the constant sections of both heavy and light chains. An exception to this, however, occurs with the allotypes of the heavy chains of rabbit immunoglobulins, where amino acid analysis suggested that related differences were in the *N*-terminal section (Inman & Reisfeld, 1968; Koshland, 1967; Koshland, Reisfeld & Dray, 1968; Koshland, Davis & Fujita, 1969), and sequence studies showed correlation with residues in several positions in the *N*-terminal 34-residue section (Porter, 1967; Wilkinson, 1969*a*). The apparent presence of a genetic marker in the variable section was clearly a major advantage in investigations of the origin of the variability. The known sequence has now been extended from position 35 to include most of the variable sections of the heavy chains of both the Aa1 and Aa3 allotypes. A further striking correlation of sequence and allotype has been found. The sequence of the first 65 residues of a relatively homogeneous rabbit antibody of allotype Aa2 is reported elsewhere (Fleischman, 1971).

MATERIALS AND METHODS

IgG† was prepared, as described by Prahl & Porter (1968), from rabbits of known allotype (kindly typed by Professor P. G. H. Gell). Enzymes and other materials

* Present address: Department of Biochemistry, University of Birmingham, P.O. Box 363, Birmingham 15, U.K.

† Abbreviations: IgG, immunoglobulin G; Fd fragment, *N*-terminal half of the heavy chain of IgG; CMCys, Hsr, pGlu (in amino acid sequences and tables), *S*-carboxymethylcysteine, homoserine and pyrrolid-2-one-5-carboxylic acid.

were used as previously described (Cebra, Steiner & Porter, 1968). The C-1 fragment of the γ-chain was prepared (Fruchter, Jackson, Mole & Porter, 1970) and fully reduced and [^{14}C]alkylated (O'Donnell, Frangione & Porter, 1970).

The conditions of enzymic digestions of peptides were as follows: tryptic digestion (300μg of trypsin/μmol of peptide) in 1% (w/v) NH_4HCO_3 buffer, pH8.2, at 37°C for the C-Ts4 peptides (digestion time 1½h), the C-GR peptides (digestion time 4h) and Aa3 fragment C5 (digestion time 8h); peptic digestion of peptide C-Ts4-2 (300μg of pepsin/μmol of peptide) in 5% (v/v) formic acid at 37°C for 2h; chymotryptic digestion of peptide C-GR-3 (100μg of chymotrypsin/μmol of peptide) in 1% (w/v) NH_4HCO_3 buffer, pH8.2, at 37°C for 4h; papain digestion of peptide Aa3-Ts4 (300μg of papain/μmol of peptide) in 0.05M-pyridine acetate–0.01M-2-mercaptoethanol, pH4.3, at 37°C for 5h.

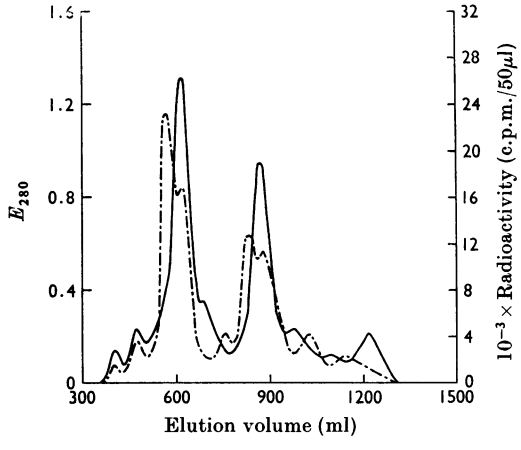

Fig. 1. Fractionation of a tryptic digest of the Aa1 heavy chain C-1 fragment blocked at the ϵ-amino groups of lysine by citraconic anhydride. After removal of the blocking group the digest was put on a Sephadex G-50 column (220cm × 3.0cm) in 50% formic acid. ——, E_{280}; ·-·-, radioactivity (c.p.m./50μl of eluate).

High-voltage paper electrophoresis at pH6.5, 3.5 and 1.9 was carried out as described by Cebra *et al.* (1968). Amino acid analyses were performed as described by Wilkinson (1969*b*), radioautography as described by O'Donnell *et al.* (1970) and isolation of methionine-containing peptides as described by Wilkinson (1969*b*).

Citraconylation and tryptic digestion of citraconylated C-1 fragment. Citraconic anhydride (BDH Chemicals Ltd., Poole, Dorset, U.K.) was used to block lysine residues (Dixon & Perham, 1968). The [^{14}C]carboxymethylated C-1 fragment (10mg/ml) was dialysed against water. The suspension was adjusted to pH8 by addition of 1M-NaOH and a 20-fold molar excess of citraconic anhydride over the lysine content was added in 50μl portions with vigorous stirring at room temperature, each portion being allowed to dissolve before addition of the next. Throughout the procedure 5M-NaOH was added manually through a micrometer syringe to maintain the pH at 7.5–8.5. The clear solution was then dialysed against several changes of dil. NH_3, pH9.0, before digestion with trypsin.

The enzyme (1%, w/w) was added in four portions at 15min intervals and 0.2M-NaOH was added by autotitrator to maintain the pH at 8.3. After complete digestion (1.5–2h) the solution was freeze-dried, taken up in 0.04M-pyridine–acetate buffer, pH3.5, and left at room temperature for 24h to remove citraconyl groups. The suspension was freeze-dried and dissolved in 98–100% formic acid for fractionation on a Sephadex G-50 column (220cm × 3.0cm) in 50% (v/v) formic acid.

RESULTS

Sequences for the heavy chains of Aa1 and Aa3 allotype rabbit IgG have been reported for residues 1–34 (Wilkinson, 1969*a*), and for pooled IgG from about residues 120–240 (Cebra *et al.* 1968). Our objective was to establish the sequence for Aa1 and Aa3 γ-chains from residue 35 to the hypervariable section commencing at about position 95.

Table 1. *Major radioactive peptides isolated from a tryptic–chymotryptic digest of reduced and* [^{14}C]*carboxymethylated fragment C-1 (inter-chain disulphide bonds not labelled) by paper electrophoresis at* pH3.5 *(O'Donnell* et al. 1970)

Mobilities are expressed relative to aspartic acid (−1), the origin being taken as the position of ϵ-DNP-lysine.

Peptide	Position of half-cystine residue	Mobility at pH3.5	Mobility at pH6.5	Sequence implied from amino acid composition
A1	146	0.73	0	Gly-CMCys-Leu-Val-Lys
A2	146	0.65	0	Related to peptide A1
A3	97	0.56	0	(CMCys,Ala,Arg)
B	201	−0.08	−0.31	Ser-Glx-Pro(Pro,Ser)Thr-CMCys-Asn-Val
C1	22	−0.15	−0.29	Thr-Leu-Thr-CMCys-Thr-$\frac{\text{Val}}{\text{Ala}}$-Ser-Gly-Phe
D1a	22	−0.27	0.34	Thr-CMCys-Thr-$\frac{\text{Val}}{\text{Ala}}$-Ser-Gly-Phe
D2	133–134	−0.27	−0.60	Ala-Pro-Ser-Val-Phe-Pro-Leu-Ala-Pro-CMCys-CMCys-Gly-Asp-Thr-Pro-Ser-Ser-Thr-Val-Thr-Leu

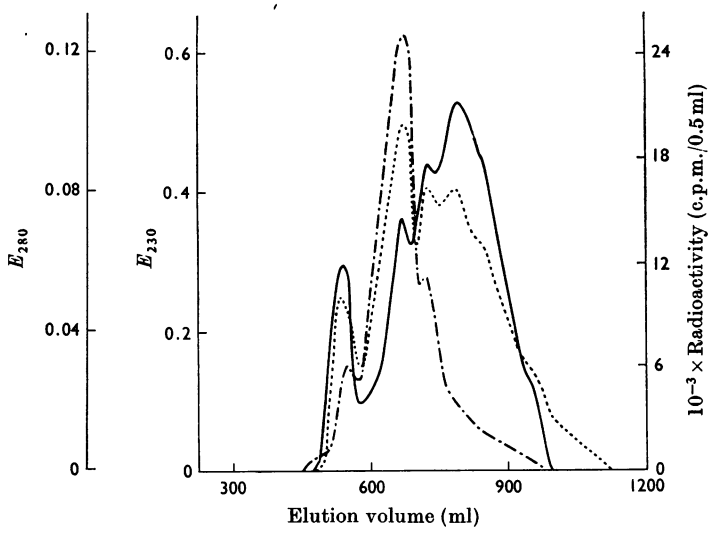

Fig. 2. Re-fractionation of material eluted between 795 and 930 ml from the Sephadex G-50 column in 50% formic acid. The material was fractionated on a Sephadex G-50 column (220 cm × 3.0 cm) in 0.05 M-NH_3. ——, E_{280}; . . . . . , E_{230}; –·–·, radioactivity (c.p.m./0.5 ml of eluate).

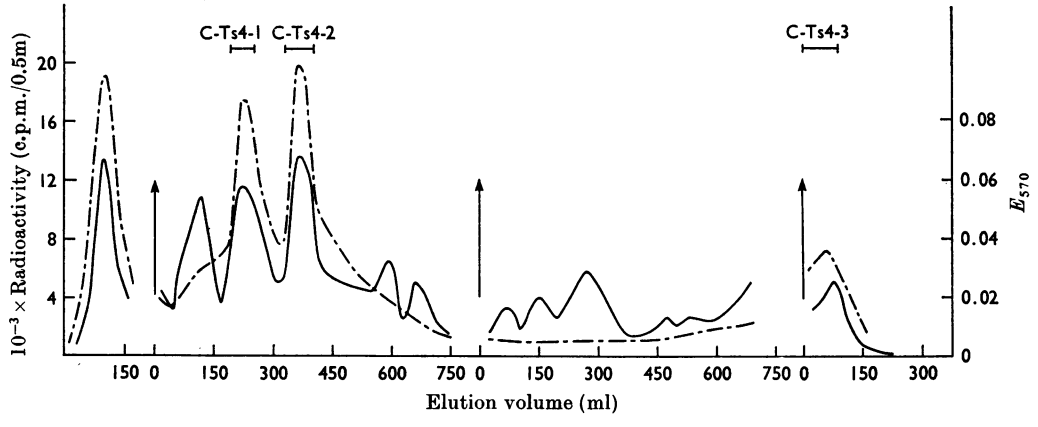

Fig. 3. Chromatography of the C-Ts4-containing material (from the Sephadex G-50 column in 0.05 M-NH_3) on a Dowex 50(X2) column (30 cm × 1.5 cm). The starting buffer was 0.2 M-pyridine acetate, pH 3.1, and the first linear gradient was to 2 M-pyridine acetate, pH 5.0, with 370 ml in each open chamber; the second linear gradient was from 2 M-pyridine acetate, pH 5.0, to 4 M-pyridine acetate, pH 6.5, with 330 ml in each chamber; the third linear gradient was from 4 M-pyridine acetate, pH 6.5, to 8.5 M-pyridine acetate, pH 6.5, with 100 ml in each chamber. The arrows show the commencement of the gradients. ——, Ninhydrin colour (E_{570}); –·–·, radioactivity (c.p.m./0.5 ml of eluate).

Sequence of the section 35–94 for Aa1 γ-chain. O'Donnell *et al.* (1970) showed that the T1 fraction from a tryptic digest of C-1 fragment in which the lysine residues had been treated with *S*-ethyl trifluorothioacetate was an aggregate derived from several sections of the peptide chain. Dissociation could not be effected satisfactorily, so the alternative technique of blocking the lysine residues with citraconic anhydride before tryptic digestion was used, the citraconyl groups being removed before fractionation on a Sephadex G-50 column in 50% formic acid (see the Materials and Methods section). The elution diagram (Fig. 1) shows that no major aggregate peak equivalent to fraction T1 is present.

The eluate was scanned for the distribution of the radioactive *S*-carboxymethylcysteine-containing peptides after tryptic–chymotryptic digestion (O'Donnell *et al.* 1970). The characteristics of these peptides are given in Table 1. The eluate was also scanned for peptides ending in Gly-Arg by tryptic digestion of samples and electrophoresis on paper at pH 6.5.

The presence of peptide A1 at 615ml and 750ml elution volume and peptide B at 570ml showed

Table 2. *Amino acid composition of peptides C-Ts4-1, C-Ts4-2 and C-Ts4-3*

	Composition (mol of amino acid residue/mol of peptide)		
Peptide ...	C-Ts4-1	C-Ts4-2	C-Ts4-3
Lys	1.4	2.1	2.06
Arg	1.0	1.0	1.0
CMCys	0.5	0.5	0.95
Asp	2.7	2.5	2.34
Thr	8.0	8.4	8.6
Ser	3.3	3.6	3.5
Glu	1.4	1.5	1.3
Pro	1.6	1.5	1.3
Gly	1.7	1.5	0.6
Ala	2.6	2.5	2.2
Val	1.7	1.6	1.2
Ile	1.4	1.9	2.1
Leu	2.2	1.9	1.3
Tyr	1.0	1.2	1.1
Phe	1.6	2.2	2.2
Hsr	0.2	0.1	0

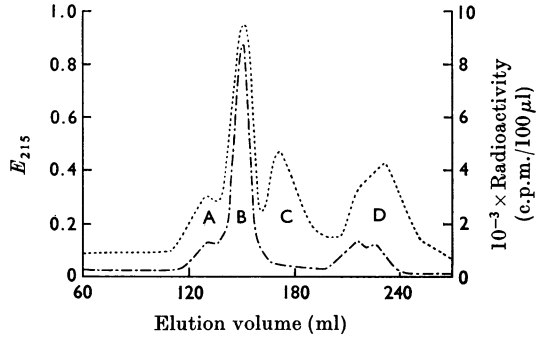

Fig. 4. Fractionation of a tryptic digest of peptide fraction C-Ts4-2 on a column (220cm × 1.2cm) of Sephadex G-50 in 0.05M-NH_3. ····· E_{215}; ·–·, radioactivity (c.p.m./100µl of eluate).

Table 3. *Isolation and amino acid composition of constituent peptides of peptide C-Ts4-2*

Elution volumes (with 0.05M-NH_3, from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to lysine +1, aspartic acid −1 and ε-DNP-lysine 0.

Peptide	*N*-Terminal amino acid	Relative elution vol.	Paper electrophoresis pH	Paper electrophoresis Mobility
C-Ts4-2	Phe/some Leu			
Tryptic peptides isolated from peptide C-Ts4-2				
C-Ts4-2B1	Ile	1.56	6.5	−0.35
C-Ts4-2B2	Ile	1.56	6.5	−0.41
C-Ts4-2C1	Thr	1.80	3.5	+0.40
C-Ts4-2C2	Thr	1.80	3.5	+0.07
			6.5	−0.37
C-Ts4-2D1	Phe	2.25	3.5	+0.55
C-Ts4-2D2	Phe/CMCys	2.25	3.5	+0.25
Peptic peptides isolated from peptide C-Ts4-2				
C-Ts4-2Pe1	Lys		3.5	+0.45
C-Ts4-2Pe2	Phe		3.5	+0.35

	Composition (mol of amino acid residue/mol of peptide)							
	Tryptic peptides						Peptic peptides	
Peptide ...	C-Ts4-2B1	C-Ts4-2B2	C-Ts4-2C1	C-Ts4-2C2	C-Ts4-2D1	C-Ts4-2D2	C-Ts4-2Pe1	C-Ts4-2Pe2
Lys	—	—	1.0	1.0	1.0	—	1.0	1.0
Arg	0.92	—	—	—	—	1.0	—	—
CMCys	0.52	—	—	—	—	0.6	—	—
Asp	1.1	1.1	1.15	1.2	—	—	—	1.0
Thr	4.2	4.2	2.85	1.7	1.0	—	2.3	3.4
Ser	1.0	1.0	1.2	1.6	1.1	0.3	1.3	1.2
Glu	1.0	1.0	—	—	—	—	0.6	—
Pro	0.97	1.0	—	—	—	—	0.9	—
Gly	0.20	0.2	—	0.7	0.2	0.3	0.2	—
Ala	1.85	1.0	—	—	—	1.1	—	—
Val	—	—	1.0	1.0	—	—	—	0.8
Ile	0.75	0.75	—	—	1.0	—	0.9	0.9
Leu	0.15	0.16	1.1	1.1	—	—	0.3	0.8
Tyr	0.95	0.91	—	—	—	—	—	—
Phe	0.95	0.62	—	—	1.0	0.4	—	0.8

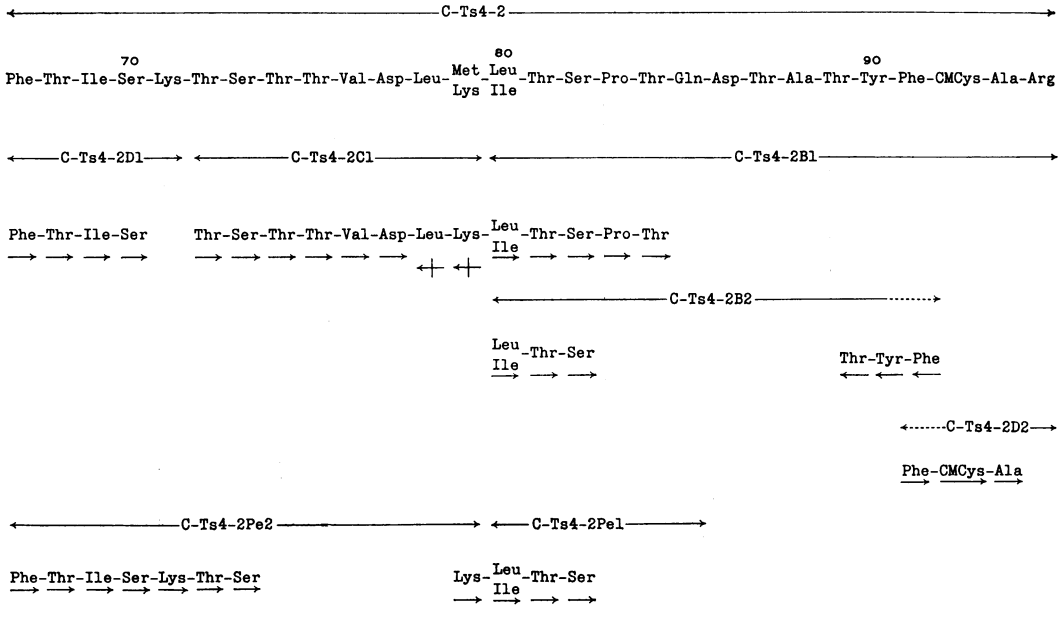

Fig. 5. Amino acid sequence of tryptic peptide C-Ts-4 showing the peptides derived from it. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →, and from the *C*-terminal end by hydrolysis with either carboxypeptidase A, shown by ←, or by a mixture of carboxypeptidases A and B, shown by ←+. The phenylalanine residue preceding the carboxymethylcysteine residue was present as an alternative residue in the peptides C-Ts4-2B2 and C-Ts4-2D2, and is indicated by a dotted line. The amino acid sequence of residues 85–88 is assumed from that of the peptide Ts4 (O'Donnell *et al.* 1970), which has properties identical with those of peptide C-Ts4-2B1.

that the material eluted between 500 and 750ml was predominantly from peptides T2 and T3, the constant *C*-terminal region of the C-1 section of the heavy chain (Cebra *et al.* 1968). Peptide A3 (CMCys-Ala-Arg) was found in peaks eluted at 825ml, 1020ml and 1140ml. Peptides C1 and D1a derived from sequences around the cysteine residue at position 22 were also present in the peak eluted between 825 and 900ml. It can be concluded that peptides ending in CMCys-Ala-Arg were present as three fragments differing substantially in size. The smallest proved to be free CMCys-Ala-Arg. The peptide of intermediate size corresponded to the Ts4 pentadecapeptide isolated by O'Donnell *et al.* (1970) and the large peptide, named C-Ts4, appeared to be about double this size. A peak with high u.v. absorption at 280nm was eluted at 870ml in the same region and this contained a peptide, named C-GR, terminating in Gly–Arg, which by homology with the γ-chain of human myeloma proteins (Press & Hogg, 1969; Edelman *et al.* 1969) was thought to be the other major peptide accounting for the unknown section in the variable region.

Thus the fraction eluted from 795 to 930ml contained the peptides of major interest and they were fractionated further on a Sephadex G-50 column in 0.5M-ammonia, since it was expected that peptide C-GR, which has a high content of aromatic amino acids, would be retarded on a Sephadex G-50 column in a solution of low ionic strength. When the eluate (Fig. 2) was scanned for radioactive peptides as described above, peptide A3 was in the fraction eluted between 600 and 705ml. Peptides C1 and D1a were present in the eluates from 550 to 720ml. The C-GR peptide had been retarded and was eluted after 750ml and had therefore been separated from peptide C-Ts4.

The first fraction (600–705ml), containing peptide C-Ts4, was freeze-dried, dissolved in water, adjusted to pH2.0 and fractionated further by chromatography on a Dowex 50 (X2) column in 0.2M-pyridine acetate, pH3.1. Gradient elution gave the profile shown in Fig. 3 and a radioautographic scan showed that peptides containing CMCys-Ala-Arg sequences were present in the three radioactive peaks; the corresponding peptides were called C-Ts4-1, C-Ts4-2 and C-Ts4-3. Peptides C-Ts4-1 and C-Ts4-2 were both contaminated by

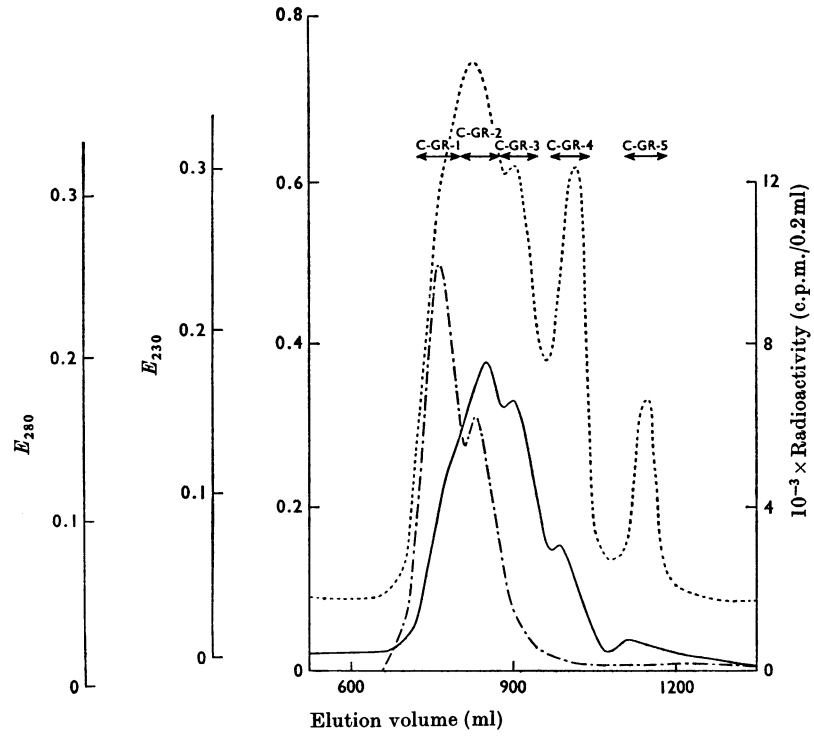

Fig. 6. Refractionation of the C-GR peptide-containing material from a Sephadex G-50 column (220 cm × 3.0 cm) in 0.05 M-NH_3 on the same column. ——, E_{280}; ....., E_{230}; -·-·, radioactivity (c.p.m./0.2 ml of eluate).

material from the cysteine-22 region, as judged by analysis. There was about 20% of such contaminating material in the first peak, and about 10% in the second. Peptide C-Ts4-3 was pure on isolation (see Table 2).

It was known from the work of Wilkinson (1969*a*) that the cysteine-22 material contained no trypsin-sensitive bonds, and so it was expected that tryptic digestion of material from the first two peaks would give in each case peptides easily separable from the small amount of contaminating material. This proved to be so, as tryptic digestion and subsequent fractionation on a Sephadex G-50 column resulted in the localization of all the contaminating material in peak A in both cases (Fig. 4).

Table 3 shows the isolation procedure for the tryptic peptides of C-Ts4-2 together with their analyses. The sequence results are summarized in Fig. 5. The order of the peptides was confirmed by a peptic digestion and the subsequent isolation of peptides 2-Pe1 and 2-Pe2. Peptide C-Ts4-2C2, whose mobility at both pH 3.5 and pH 6.5 was inconsistent with its analysis, may be an artifact of citraconylation but, as it was split by trypsin, the lysine residue was presumably unblocked. It was present in only one-sixth molar yield relative to peptide C-Ts4-C1, but does not appear to be derived from a variant sequence.

The presence of a methionine residue at position 79 in about 20% of the Aa1 γ-chain molecules was deduced from the isolation of a cyanogen bromide fragment equivalent to fragment C5 of Aa3 γ-chain (see below). The only other variant recognized in this section is the leucine-for-isoleucine replacement at position 80 in 15–20% of the molecules. The presence of both phenylalanine and carboxymethylcysteine as *N*-terminal in peptide C-Ts4-2D2 arises from chymotryptic-like hydrolysis on either side of the phenylalanine residue, as shown in Fig. 5.

Peptide C-Ts4-3, on hydrolysis with trypsin, gave peptides identical with those obtained from peptide C-Ts4-2, but the fraction containing peptide C-Ts4-1 contained only small amounts of the *N*-terminal Phe-Thr-Ile-Ser-Lys sequence; it is likely that incomplete reaction with citraconic anhydride led to partial hydrolysis of the lysine–threonine bond at position 71–72 during the tryptic digestion of C-1 fragment. The presence of small

Table 4. *Amino acid composition of peptides C-GR-2 and C-GR-3*

Peptide ...	Composition (mol of amino acid residue/mol of peptide) C-GR-2	C-GR-3
Lys	1.2	1.80
Arg	1.0	1.0
CMCys	0.2	—
Asp	1.1	1.26
Thr	2.2	2.0
Ser	2.2	2.4
Glu	1.6	2.2
Pro	1.6	1.8
Gly	3.3	5.0
Ala	1.6	2.7
Val	1.0	0.7
Ile	1.3	2.1
Leu	1.8	1.6
Tyr	1.3	2.1
Phe	0.5	0.5
Hsr	0.2	—

amounts of Phe-Thr-Ile-Ser-Lys in fraction C-Ts4-1 was probably due to the proximity of the peak containing peptide C-Ts4-2. The difference between peptides C-Ts4-2 and C-Ts4-3 was not due to a difference in primary sequence (Mole, 1971).

Thus no evidence of heterogeneity of sequence in this section has been found except for the methionine-for-lysine and leucine-for-isoleucine replacements mentioned above.

Fraction C-GR. Re-fractionation of fraction C-GR on a Sephadex G-50 column in 0.05M-ammonia gave the elution pattern shown in Fig. 6. When the fractions shown were analysed and scanned for radioactive peptides, fraction C-GR-1 contained predominantly cysteine-22 peptides, fraction C-GR-4 had only a trivial content of peptide material in spite of the high absorption at 230nm and most of the C-GR peptides were found in fractions C-GR-2 and C-GR-3. The amino acid analyses (Table 4) suggest that fraction C-GR-2 may have 30–40% contamination by

Table 5. *Isolation and amino acid composition of tryptic peptides of peptide C-GR-3*

Elution volumes (with 0.05M-NH_3, from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to lysine +1 and ε-DNP-lysine 0. Tryptophan was detected by the Ehrlich stain.

Peptide	*N*-Terminal amino acid	Relative elution vol.	Paper electrophoresis pH	Mobility
C-GR-3	Glu			
C-GR-3TA	Gly	1.6	—	—
C-GR-3TB	Gly	1.8	—	—
C-GR-3TCA	Glu	2.2	3.5	0.65
C-GR-3TCB	—	2.2	3.5	0.19
C-GR-3TDA	Gly	2.5	3.5	0.89
C-GR-3TDB	Ala	2.5	3.5	0.59
C-GR-3TDC	Tyr	2.5	3.5	0.48

	Composition (mol of amino acid residue/mol of peptide) Column fractions		Peptides				
	C-GR-3TA	C-GR-3TB	C-GR-3TCA	C-GR-3TCB	C-GR-3TDA	C-GR-3TDB	C-GR-3TDC
Lys	0.8	0.2	1.0	1.0	—	1.0	1.0
Arg	—	—	—	—	1.0	—	—
Asp	1.2	1.2	—	—	—	—	—
Thr	1.2	1.2	—	—	—	—	—
Ser	1.7	1.3	—	—	—	1.2	1.5
Glu	1.2	1.2	1.1	1.0	—	—	—
Pro	0.3	0.4	1.0	0.9	—	—	—
Gly	2.7	2.9	1.1	1.1	1.2	0.4	0.6
Ala	1.6	1.0	0.9	1.0	—	1.7	1.7
Val	0.4	0.6	—	—	—	—	—
Ile	1.6	1.6	—	—	—	—	—
Leu	1.0	1.0	—	—	—	—	—
Tyr	1.4	0.8	—	—	—	—	0.7
Phe	0.2	—	—	—	—	—	—
Trp	+ve	+ve	—	—	—	+ve	+ve

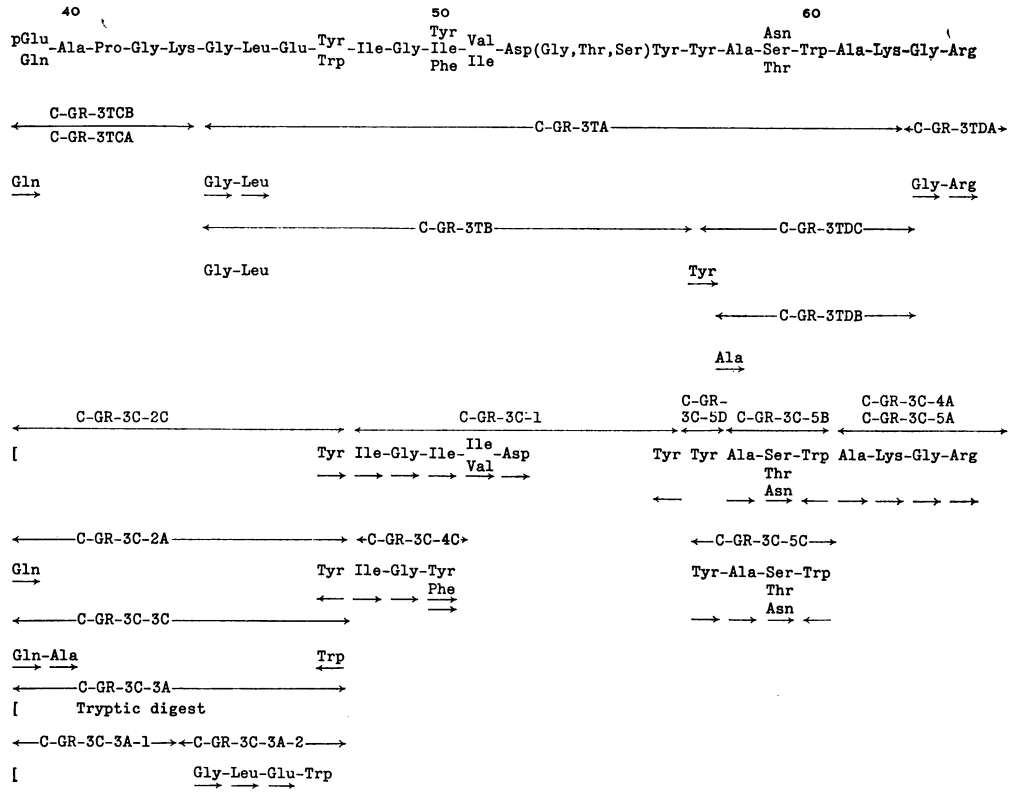

Fig. 7. Amino acid sequence of peptide C-GR-3 showing the peptides derived from it. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →, and from the *C*-terminal end by hydrolysis with carboxypeptidase A, shown by ←. The presence of *N*-terminal pyrrolid-2-one-5-carboxylic acid (pGlu), resulting from cyclization of glutamine, is indicated by [.

cysteine-22 peptides as judged by homoserine content, whereas fraction C-GR-3 has less than 5% of such contamination.

The results on the constituent tryptic peptides of peptide C-GR-3 are summarized in Table 5. Although identical on analysis, peptides C-GR-3TCA and C-GR-3TCB differ in mobility owing to cyclization of the *N*-terminal glutamine residue.

Difficulty in isolating a unique peptide C-GR-3TA suggested considerable heterogeneity in this section and the subsequent sequence studies shown in Fig. 7 confirmed that alternative residues existed in at least four positions. Chymotryptic-like splits during tryptic digestion gave rise to peptides C-GR-3TDC and C-GR-3TDB, which were isolated in low yields.

Chymotryptic digestion of peptide C-GR-3 gave the peptides listed in Table 6 and 7. The presence of three alternative residues in position 59 was deduced from the analyses of peptides C-GR-3C-5B and C-GR-3C-5C and subsequent sequence determination of both, which showed the same complexity in the third position of peptide C-GR-3C-5C and the second position of peptide C-GR-3C-5B; asparagine rather than aspartic acid is allocated to one of the variants on the basis of mobility results. Similar considerations led to the assignment of three alternative residues to position 50 and two to position 51, and also tyrosine and tryptophan to position 47. The blocked *N*-terminal pentapeptide C-GR-3TCA was assumed to have the same amino acid sequence as the corresponding peptide (C-GR-2TCB) isolated from peptide C-GR-2 (see below).

Peptide C-GR-2 could not be purified further. After tryptic digestion of this peptide similar peptides to those obtained from peptide C-GR-3 were isolated (Table 8) and their sequences partially determined (Fig. 8). The amino acid sequence of the blocked peptide C-GR-2TCB was deduced in

Table 6. *Isolation of chymotryptic peptides of peptide C-GR-3*

Elution volumes (with 0.05M-NH_3, from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to lysine +1, aspartic acid −1 and ε-DNP-lysine 0.

Peptide	*N*-Terminal amino acid	Relative elution vol.	Paper electrophoresis	
			pH	Mobility
C-GR-3C-1	Ile	1.64	6.5	−0.24
C-GR-3C-2A	Gln	1.82	3.5	+0.42
C-GR-3C-2C	—	1.82	3.5	+0.11
C-GR-3C-3A	—	1.94	6.5	+0.21
C-GR-3C-3C	Gln	1.94	6.5	+0.07
C-GR-3C-4A	Ala	2.12	3.5	+1.0
C-GR-3C-4C	Ile	2.12	3.5	+0.21
C-GR-3C-5A	Ala	2.46	3.5	+1.0
C-GR-3C-5B	Ala	2.46	3.5	+0.23
C-GR-3C-5C	Tyr	2.46	3.5	+0.17
C-GR-3C-5D	—	2.46	3.5	+0.04
C-GR-3C-3A1	—	—	6.5	0
C-GR-3C-3A2	Gly	—	6.5	−0.39

Table 7. *Amino acid composition of chymotryptic peptides of peptide C-GR-3*

Composition (mol of amino acid residue/mol of peptide)

Peptide ...	C-GR-3C-1	C-GR-3C-2A	C-GR-3C-2C	C-GR-3C-3A	C-GR-3C-3C	C-GR-3C-3A-1	C-GR-3C-3A-2
Lys	—	1.0	1.0	1.0	1.0	1.0	—
Arg	—	—	—	—	—	—	—
Asp	1.0	—	—	—	—	—	—
Thr	1.1	0.2	—	—	—	—	—
Ser	1.2	0.4	0.3	—	0.3	0.2	—
Glu	0.3	1.9	2.2	2.0	2.2	1.1	1.0
Pro	—	1.2	1.2	1.0	1.0	1.2	—
Gly	1.9	1.9	2.4	2.0	2.3	1.3	1.1
Ala	0.4	0.7	1.1	0.9	1.1	1.0	—
Val	0.3	0.6	—	—	—	—	—
Ile	1.8	—	—	—	—	—	—
Leu	0.2	1.0	0.9	0.9	1.0	—	1.0
Tyr	0.8	0.8	0.8	—	—	—	—
Phe	0.2	—	—	—	—	—	—
Trp	—	—	—	+ve	+ve	—	+ve

Peptide ...	C-GR-3C-4A	C-GR-3C-4C	C-GR-3C-5A	C-GR-3C-5B	C-GR-3C-5C	C-GR-3C-5D
Lys	1.0	—	1.0	—	—	—
Arg	1.0	—	1.0	—	—	—
Asp	—	—	—	0.2	0.2	—
Thr	—	—	—	0.3	0.3	—
Ser	0.3	0.5	—	0.7	0.8	0.4
Glu	—	—	—	—	—	—
Pro	—	—	—	—	—	—
Gly	1.4	1.1	1.1	0.1	0.2	—
Ala	1.1	—	1.0	1.0	1.0	
Val	—	—				
Ile	—	1.0	—	—	—	—
Leu	—	—	—	—	—	—
Tyr	—	0.6	—	—	1.0	1.0
Phe	—	0.4	—	—	—	—
Trp	—	—	—	+ve	+ve	—

the following manner. After Pronase digestion and subsequent passage down a column of Dowex 50 (X2; H^+ form) (Wilkinson, Press & Porter, 1966), a peptide of composition (Glx,Ala,Pro) was obtained. No *N*-terminal residue could be detected, but hydrazinolysis established the *C*-terminal residue as

Table 8. *Isolation and amino acid composition of tryptic peptides of peptide C-GR-2*

Elution volumes (with 0.05M-NH_3, from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to lysine +1 and ϵ-DNP-lysine 0. Tryptophan was detected by the Ehrlich stain.

Peptide	*N*-Terminal amino acid	Relative elution vol.	Paper electrophoresis Mobility at pH 3.5
C-GR-2TCA	Glu	2.2	+0.65
C-GR-2TCB	—	2.2	+0.18
C-GR-2TDA	Gly	2.5	+0.87
C-GR-2TDB	Ala	2.5	+0.59
C-GR-2TDC	Tyr	2.5	+0.48

	Composition (mol of amino acid residue/mol of peptide)				
Peptide ...	C-GR-2TCA	C-GR-2TCB	C-GR-2TDA	C-GR-2TDB	C-GR-2TDC
Lys	1.0	1.0	—	1.0	1.0
Arg	—	—	1.0	—	—
Asp	—	—	—	0.2	0.2
Thr	—	—	—	0.2	0.3
Ser	—	—	—	1.1	1.4
Glu	1.1	1.0	—	—	—
Pro	1.0	0.9	—	—	—
Gly	1.0	1.1	1.1	0.2	0.3
Ala	1.0	1.0	—	1.9	1.8
Tyr	—	—	—	—	0.8
Trp	—	—	—	+ve	+ve

C-GR-2TCA Gln-Ala-Pro-Gly-Lys

C-GR-2TCB pGlu-Ala-Pro-Gly-Lys

C-GR-2TDA Gly-Arg

C-GR-2TDB Ala-Ser(Trp,Ala)Lys

C-GR-2TDC Tyr-Ala(Ser,Trp,Ala)Lys

Fig. 8. Amino acid sequence of tryptic peptides of peptide C-GR-2. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →, and from the *C*-terminal end by hydrazinolysis, shown by ←. The position of cleavage of peptide C-GR-2TCB by Pronase is shown by ↑.

proline. On elution of the Dowex column with M-ammonia a peptide of composition (Gly,Lys) was obtained; it had glycine as the *N*-terminal residue. Thus the sequence of peptide C-GR-2TCB is pGlu-Ala-Pro-Gly-Lys. As further sequence results could not be obtained, peptide C-GR-2 probably contained sections of even greater heterogeneity than did peptide C-GR-3.

Overlap peptides and ordering of the Aa1 peptides within the variable region. Arginine-containing peptides were isolated from a chymotryptic digestion of fragment C-1 as described by Fruchter *et al.* (1970). The C-1 fragment had been treated with maleic anhydride and then passed through a column of Dowex 50 (X2; H^+ form). The fraction that was retained and subsequently eluted with M-ammonia was greatly enriched in the arginine-containing peptides.

Two of these, namely peptides C-C-1-6d and C-C-1-5b, were isolated as shown in Table 9. The sequence of peptide C-C-1-6d was established directly as shown in Fig. 9. Peptide C-C-1-5b was digested with trypsin and its sequence determined (Fig. 10). Peptide C-C-1-6d overlaps peptides C-GR and C-Ts4, whereas peptide C-C-1-5b overlaps the *N*-terminus of peptide C-GR. A tryptic digest of C-1 fragment yielded peptides Ser-Trp-Val-Arg and Gly-Trp-Val-Arg, which extend the *N*-terminal section and give a sequence of the Aa1 γ-chain between positions 34 and 94 (Fig. 10).

A methionine-containing peptide F1b was isolated, as described by Wilkinson (1969*b*), by labelling the methionine residues of the Aa1 Fd fragment with iodo[^{14}C]acetamide at pH 3.5 and subsequently digesting the fragment with trypsin and chymotrypsin. Fractionation of this digest by Sephadex G-50 gel filtration followed by electrophoresis on paper at pH 3.5 and pH 6.5 gave a peptide of composition: serine 0.3, glycine 0.6, alanine 1, tryptophan 1, carboxymethyl methionine 1, with alanine as the *N*-terminal residue. Boiling in water for 2h cleaved the peptide (Tang & Hartley, 1967), and three ninhydrin-reactive peptides were separated by paper electrophoresis at pH 1.9. Two of these peptides also gave positive

Table 9. *Isolation and amino acid composition of overlap peptides from fragment C-1*

Elution volumes (with 0.05M-NH_3 from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to arginine +1, aspartic acid −1 and ε-DNP-lysine 0. Tryptophan was detected by the Ehrlich stain and carboxymethylmethionine (CM-Met) was positive on amino acid analysis.

Peptide	*N*-Terminal amino acid	Relative elution vol.	Paper electrophoresis pH	Paper electrophoresis Mobility
C-C-1-6d	Ala	2.8	6.5	+0.85
C-C-1-5b	Val	2.5	3.5	+0.66
C-C-1-5b-1	Gly		3.5	Neutral
C-C-1-5b-2	Glx		3.5	+0.56
C-C-1-5b-3	Val		3.5	+0.81
F1b	Ala	2.8	6.5	−0.47
T8a	Ser		6.5	+0.40
T8b	Gly		6.5	+0.50

Composition (mol of amino acid residue/mol of peptide)

	Chymotryptic peptides						Tryptic peptides	
Peptide ...	C-C-1-6d	C-C-1-5b	C-C-1-5b-1	C-C-1-5b-2	C-C-1-5b-3	F1b	T8a	T8b
Lys	1.0	0.7		0.9				
Arg	1.0	1.0			1.2		1.0	1.0
Asp								
Thr								
Ser	0.4					0.3	0.9	
Glu		1.9	1.4	0.9				
Pro		1.2		1.2				
Gly	1.1	1.6	1.4	1.1		0.6		1.0
Ala	1.0	1.0		1.0		1.0		
Val		0.9			1.0		1.0	1.0
Ile								
Leu		0.9	1.0					
Tyr								
Phe	0.8							
CM-Met						+ve		
Trp		+ve	+ve			+ve	+ve	+ve

Peptide C-C-1-6d Ala-Lys-Gly-Arg-Phe

Rabbit Aa1 γ-chain Ala-Lys-Gly-Arg-Phe-Thr-Ile-Ser-Lys (C-GR | C-Ts4)

Human γ1-chain Eu Phe-Gln-Gly-Arg-Val-Thr-Ile-Thr-Ala

Human γ1-chain Cor Leu-Glu-Thr-Arg-Leu-Thr-Ile-Ser-Lys

Fig. 9. Amino acid sequence of peptide C-C-1-6d. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →. Peptide C-C-1-6d overlaps peptides C-GR and C-Ts4 as shown by the homology with human γ1-chains Eu (Edelman *et al.* 1969) and Cor (Press & Hogg, 1969). Residues identical in the human and rabbit sequences are underlined.

reactions with the Ehrlich stain and hence contained tryptophan. These two peptides had glycine and serine as their respective *N*-terminal residues and hence had the sequences Gly-Trp and Ser-Trp. The other peptide was Ala-Hsr and hence the sequence must be Ala-Met-$\frac{\text{Gly}}{\text{Ser}}$-Trp, establishing the overlap at position 34.

All the above sequence results are summarized in Fig. 10.

Sequence of the variable region of Aa3 γ-chain. After citraconylation and tryptic digestion of fragment C-1 from Aa3 γ-chain, fractionation of the peptides on Sephadex G-50 in 50% formic acid gave the elution diagram shown in Fig. 11. A radioautographic scan showed that the CMCys-Ala-Arg-containing peptides were in peaks eluted at 855ml, 1050ml and 1150ml. The peak eluted at 1050ml was expected to contain the peptide corresponding to peptide Ts4 from Aa1 γ-chain, and this was isolated by further fractionation on a Sephadex G-50 column in 0.05M-ammonia.

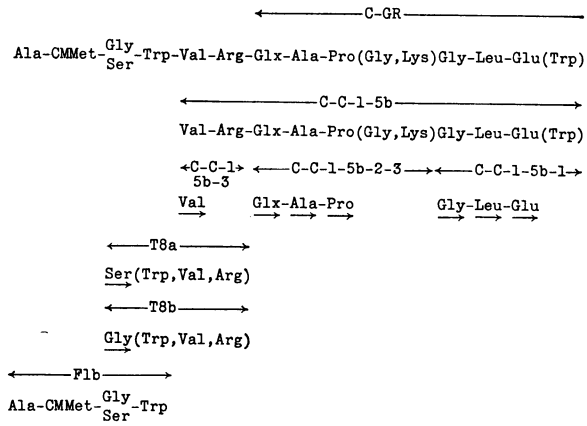

Fig. 10. Partial amino acid sequence of residues 33–47 of Aa1 heavy chain, with numbering from the *N*-terminus. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →. The derivation of the sequence of peptide F1b is described in the text. Abbreviation: CMMet, carboxymethylmethionine.

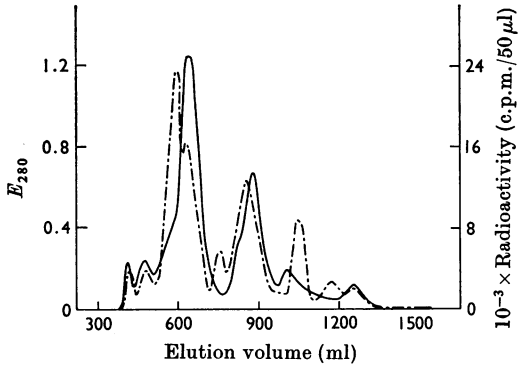

Fig. 11. Fractionation of a tryptic digest of the Aa3 heavy chain C-1 fragment blocked at the ε-amino group of lysine by citraconic anhydride. After the removal of the blocking group the digest was put on a Sephadex G-50 column (220 cm × 3.0 cm) in 50% formic acid. ——, E_{280}; –·–·, radioactivity (c.p.m./50 µl of eluate).

Only one fraction contained radioactivity and this was subjected to electrophoresis on paper at pH 6.5, after which two radioactive peptides were isolated. Their analyses (Table 10) were identical and it is considered that peptide Ts4-A, which was isolated in low yield, may contain a second carboxymethylcysteine residue (O'Donnell *et. al.* 1970). The sequence of the major peptide was established by papain digestion as shown in Fig. 12, the analyses and mobilities of the peptides being given in Tables 11 and 12. Peptide P3B was alanylalanine as judged by electrophoretic mobility at pH 1.9 and amino acid analysis of the unhydrolysed peptide, which gave no peak in the position of alanine. Peptides P1 and P9 were shown to be free amino acids in a similar manner.

Table 10. *Isolation and amino acid composition of peptides Aa3-Ts4-A and Aa3-Ts4-B*

Gel filtration was on Sephadex G-50 in 0.05M-NH_3. Elution volumes are relative to exclusion volume = 1.0. Paper electrophoresis was at pH 6.5; mobilities are given relative to aspartic acid −1 and ε-DNP-lysine 0.

Peptide ...	Aa3-Ts4-A	Aa3-Ts4-B
Relative elution vol.	1.85	1.85
Mobility on paper electrophoresis	−0.36	−0.19
Amino acid composition (mol/mol of peptide)		
Arg	1.0	1.0
CMCys	0.6	+ve
Asp	1.1	1.1
Thr	3.8	3.9
Ser	1.1	1.1
Glu	0.2	0.3
Gly	0.4	0.3
Ala	4.0	3.9
Val	0.1	0.2
Leu	1.1	1.0
Tyr	1.0	1.0
Phe	1.0	1.0

The larger C-Ts4 peptide could be separated from the cysteine-22-containing peptides on a Sephadex G-50 column (220 cm × 3.0 cm) in 0.05M-ammonia but a persistent homoserine-containing

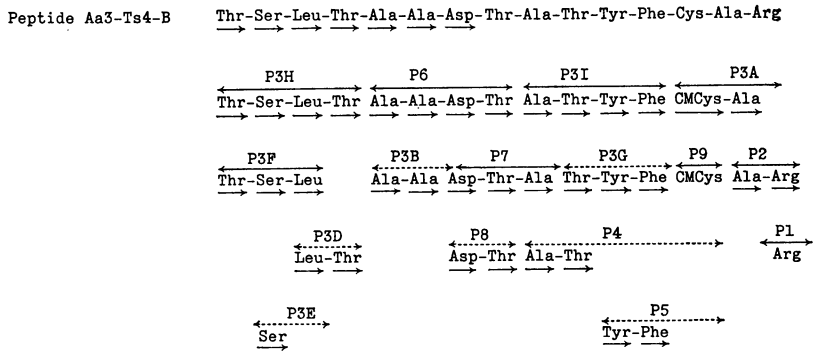

Fig. 12. Amino acid sequence of peptide Aa3-Ts4-B. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →. Peptides obtained in low yield are shown by dotted lines.

Table 11. *Isolation procedures for component papain peptides of peptide Aa3-Ts4-B*

Mobilities are given relative to lysine +1, aspartic acid −1 and ε-DNP-lysine 0. The material with zero mobility at pH 6.5 was resewn and separated at pH 1.9.

Peptide	*N*-Terminal amino acid	Paper electrophoresis Mobility at pH 6.5
Ts4-B-P1	—	+0.94
Ts4-B-P2	Ala	+0.75
Ts4-B-P3	—	0
Ts4-B-P4	Ala	−0.36
Ts4-B-P5	Tyr	−0.48
Ts4-B-P6	Ala	−0.57
Ts4-B-P7	Asp	−0.66
Ts4-B-P8	Asp	−0.87
Ts4-B-P9	—	−1.0
Peptide	***N*-Terminal amino acid**	**Paper electrophoresis Mobility at pH 1.9**
Ts4-B-P3A	—	+0.77
Ts4-B-P3B	Ala	+0.65
Ts4-B-P3C	Thr, Ala	+0.51
Ts4-B-P3D	Leu	+0.39
Ts4-B-P3E	Ser	+0.26
Ts4-B-P3F	Thr	+0.23
Ts4-B-P3G	Thr	+0.18
Ts4-B-P3H	Thr	+0.14
Ts4-B-P3I	Ala	+0.07

contaminant peptide remained. Tryptic digestion of this C-Ts4 peptide did not release a peptide corresponding to the peptide Ts4 isolated from Aa1 γ-chain, suggesting that in Aa3 γ-chain no lysine residue is present in position 79.

Fragment C-5 from cyanogen bromide digestion of Aa3 γ-chain. In the fractionation of the cyanogen bromide digest (Fruchter *et al.* 1970) of the Aa3 γ-chain it was observed that there was a relatively high content of fragment C-5 (Fig. 13). From analysis and subsequent sequence studies it became apparent that this fragment was derived from splitting of methionine residues at both positions 34 and 79. On a molar basis the yield was about 16%. The peptide was difficult to handle as it was rather insoluble in 1% ammonium hydrogen carbonate buffer, pH 8.2, but tryptic digestion for 8 h gave a soluble product that could be fractionated on a Sephadex G-50 column in 0.05 M-ammonia. The characteristics of the peptides produced are given in Table 13 and their sequences in Fig. 14. Peptides Gln-Ala-Pro-Gly-Lys and Phe-Thr-Ile-Ser-Lys are identical with the corresponding peptides isolated from Aa1 γ-chain. The *C*-terminal homoserine-containing peptide:

Thr(Ser,Thr,Thr,Val,Glu,Leu)Hsr

is identical with the equivalent sequence from Aa1 γ-chain except that there appears to be a glutamic acid-for-aspartic acid replacement.

DISCUSSION

The use of citraconic anhydride rather than *S*-ethyl trifluorothioacetate to block the lysine residues of C-1 fragment before tryptic digestion was of major advantage, as the tryptic hydrolysis went to completion, and aggregated or partial hydrolysis products comparable to fraction T1 were not obtained. This made possible a more accurate estimation of the relative recoveries of peptides from different parts of the variable section, this information being necessary for assessing the significance of the sequences that have been obtained. All the available results for the sequence

Table 12. *Amino acid composition of papain peptides of peptide Aa3-Ts4-B*

	Composition (mol of amino acid residue/mol of peptide)								
Peptide ...	Ts4-B-P1	Ts4-B-P2	Ts4-B-P4	Ts4-B-P5	Ts4-B-P6	Ts4-B-P7	Ts4-B-P8	Ts4-B-P9	Ts4-B-P3A
Arg	Free Arg	1.0							1.0
CMCys			0.4	+ve				Free CMCys	+ve
Asp				0.4	1.0	1.0	1.0		
Thr			1.1	0.4	1.0	0.9	0.9		
Ser			0.4	0.3		0.2	0.2		
Gly			0.4	0.2		0.3			
Ala		1.0	1.1	0.5	1.9	1.1	0.5		1.1
Leu									
Tyr			0.9	1.0					
Phe			1.0	1.0					

Peptide ...	Ts4-B-P3B	Ts4-B-P3C	Ts4-B-P3D	Ts4-B-P3E	Ts4-B-P3F	Ts4-B-P3G	Ts4-B-P3H	Ts4-B-P3I
Arg								
CMCys								
Asp								
Thr		1.0	1.0		0.9	0.9	1.7	0.8
Ser				1.0	1.2		1.0	
Gly								
Ala	+ve	1.8						1.0
Leu			1.0	1.0	1.0		1.0	
Tyr						0.7		0.7
Phe						1.0		1.0

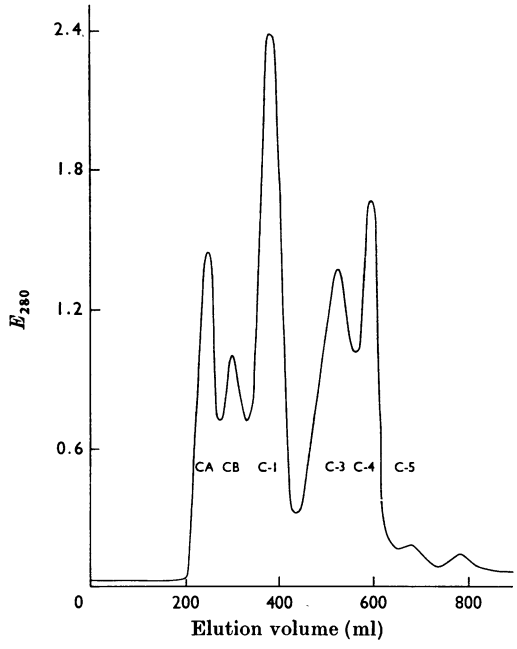

Fig. 13. Fractionation of a CNBr digest of rabbit Aa3 heavy chain on a column (140 cm × 3.0 cm) of Sephadex G-100 in 6 M-urea–0.2 M-sodium formate, pH 3.2. ——, E_{280}. Naming of the fractions is as shown.

of the variable region of the three allotypes are summarized in Fig. 15. The positions at which residues are considered to correlate with allotype have been boxed. At other positions, e.g. 3 and 4, where serine or leucine or valine respectively are common, the correlation appears uncertain and the residues have not been boxed.

In spite of the inevitable losses during fractionation, it appears clear that peptides containing the CMCys-Ala-Arg sequence (residues 92–94) are common to all or almost all the molecules in fragment C-1 from both the Aa1 and Aa3 γ-chains. The fact that this sequence is found in three peptides of various lengths is explained by cleavage at the methionine residues which occur in some cases at position 79 in both γ-chains, and by partial tryptic hydrolysis of the immediately preceding Tyr–Phe and Phe–CMCys bonds. No other peptides from this section (residues 67–94) were found in the fully characterized Aa1 γ-chain, and as far as could be judged from a comparison of the radioactivity of these peptides with that of the carboxymethylcysteine-containing peptides from the constant region, the yield in the crude fractions was nearly 100%. It therefore follows that section 67–94 is exceptionally stable with only two positions (79 and 80) showing recognizable replacements.

It is more difficult to assess the recoveries of peptides from the region of positions 34–66 but it is apparent that there is much more variability, particularly between positions 50 and 60. A replacement has also been identified at position 35.

The comparison of the sequences of the variable regions of the γ-chains of the Aa1, Aa2 and Aa3 allotypes (Fig. 15) confirms that allotypically related differences in sequence occur in positions 10, 13, 15, 16, 17, 27, 28, 29 and 33 in the *N*-terminal 34-residue section. In the region from positions 35 to 66 there is considerable variability between positions 50 and 60, with apparently two insertions

Table 13. *Isolation and amino acid composition of tryptic peptides of fragment C5*

Elution volumes (with 0.05M-NH_3, from Sephadex G-50) are given relative to exclusion volume = 1.0. Mobilities are given relative to lysine +1 and ϵ-DNP-lysine 0.

Peptide	*N*-Terminus	Relative elution vol.	Paper electrophoresis pH	Paper electrophoresis Mobility
C5-T7A	Gly	3.0	6.5	+0.80
C5-T6A	Glu	2.6	6.5	+0.49
C5-T6B	Phe	2.6	6.5	+0.42
C5-T6C	—	2.6	6.5	Neutral
			1.9	+0.14
C5-T4A	Thr	2.0	1.9	−0.05

	Composition (mol of amino acid residue/mol of peptide)				
Peptide ...	C5-T7A	C5-T6A	C5-T6B	C5-T6C	C5-T4A
Lys		1.0	1.0	1.0	
Arg	1.0				
Asp					
Thr			0.9		3.5
Ser			1.2		1.8
Glu		1.0		1.1	0.9
Pro		1.0		1.1	
Gly	1.1	1.0		1.2	0.5
Ala		0.9		0.9	
Val					1.0
Ile			1.0		
Leu					1.0
Tyr					
Phe			0.9		
Hsr					0.8

Peptide	Sequence
C5-T7A	$\overrightarrow{\text{Gly-Arg}}$
C5-T6A	$\overrightarrow{\text{Gln-Ala}}$(Pro,Gly)Lys
C5-T6B	$\overrightarrow{\text{Phe-Thr-Ile-Ser-Lys}}$
C5-T6C	([Glx,Ala,Pro,Gly,Lys)
C5-T4A	$\overrightarrow{\text{Thr}}$($\text{Thr}_{2.5}$,$\text{Ser}_{1.8}$,Glx,Leu,Val)Hsr

Fig. 14. Amino acid sequence results of the tryptic peptides isolated from fragment C-5. Sequence determination from the *N*-terminal end was by the dansyl–Edman technique, shown by →. The presence of *N*-terminal pyrrolid-2-one-5-carboxylic acid (pGlu), resulting from cyclization of glutamine, is indicated by [.

in the Aa2 sequence of a homogeneous antibody (Fleischman, 1971) relative to the Aa1 sequence. The variability makes it unlikely that allotypically related changes occur but there may be such a change in position 43. As the sequence of Aa1 and Aa3 heavy chains from pooled IgG is the same, the difference in Aa2 sequence may be idiotypic rather than allotypic.

Between positions 68 and 94 only two positions (79, 80) with alternative residues have been detected and for residues 81–94 in both Aa1 and Aa3 sequences no variability was found, and yields based on the recovery of the radioactive marker at position 92 appeared to be high. It was between positions 80 and 85 that the most striking allotypically related difference was found. No peptides deriving from position 95 to about position 115 have been recognized, and it is believed that this is due to the exceptional variability and hence low yield of any one peptide in this section. An attempt was made to assess this complexity by radioactive 'mapping' of chymotryptic peptides, all of which had the *N*-terminal sequence CMCys-Ala-Arg (S. A. Jackson, unpublished work). More than 30 such peptides could be distinguished and it seems clear that the variability of this section is much higher than is found in positions 30–35 and 50–60, where, although replacements occur, some coherent sequence could be established.

The constant section from about position 115 to beyond the extent of the Fd fragment is accounted for by two large peptides, T2 and T3, and a small peptide T7. The constituent peptides of T3 have been isolated from both Aa1 and Aa3 γ-chains and have been found to be identical in electrophoretic mobility, amino acid analysis and *N*-terminal assay. Peptides T2 and T7 from both allotypes have the

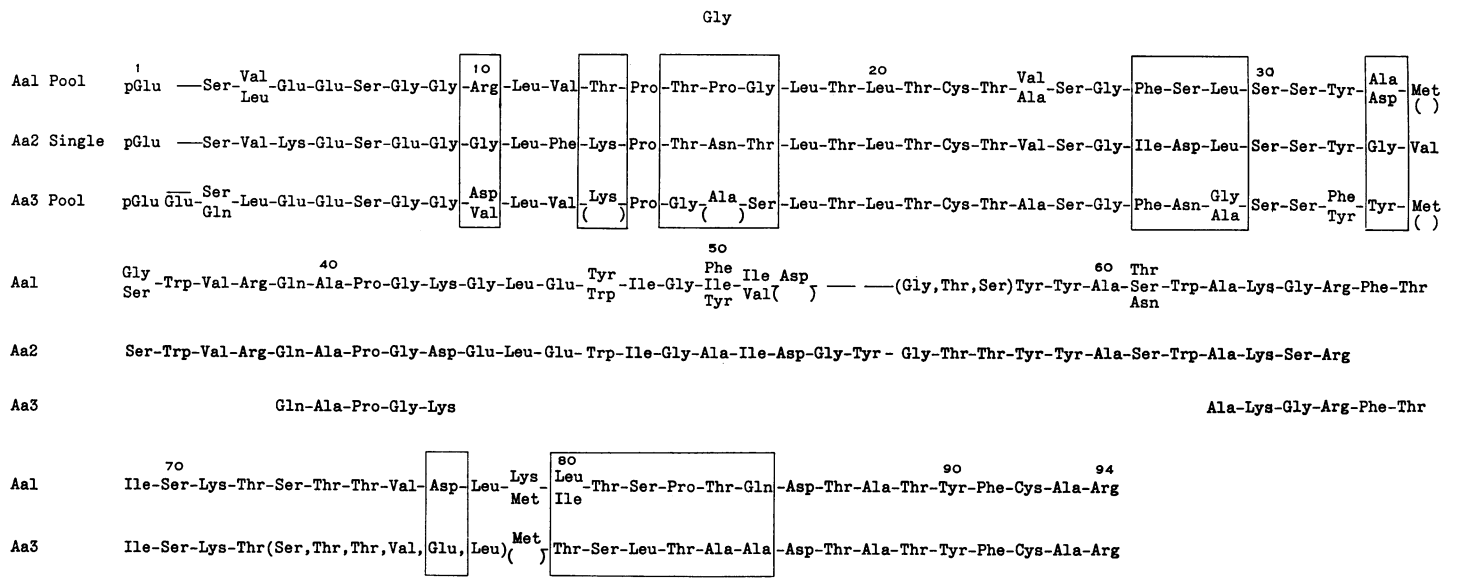

Fig. 15. Comparison of the variable-region sequences of Aa1, Aa2 and Aa3 heavy chains. Residues in boxes are those that vary between the Aa1, Aa2 and Aa3 sequences, and those in parentheses are minor variants in each sequence. The bar at positions 2, 53 and 54 indicates a presumed deletion. The Aa2 sequence is taken from Fleischman (1971).

same analysis, as has also been reported by Koshland (1967), and hence it appears that all the allotypically related differences in sequence are found in the variable section. Support for this conclusion is given by comparing the differences between Aa1 and Aa3 sequences with analysis results reported by other laboratories (Table 14). The three sets of results show good agreement. The methionine–threonine replacement at about position 220 reported by Prahl & Porter (1968) as correlating with the Aa3 and Aa1 allotypes was shown subsequently to show a close relation with the newly discovered A11 and A12 allotypes (Prahl, Mandy & Todd, 1969). By chance the rabbits of Aa3 allotype used by Prahl & Porter (1968) were also of the then unknown A11 allotype, whereas the Aa1 and Aa2 allotypes were A12. The association of the methionine-threonine change with the 'a' locus allotype was fortuitous. Thus the allotypically related differences in sequence are confined to the variable region and occur there in eight positions scattered near the *N*-terminus, in a block in positions 80–85 and possibly also at positions 43 and 77. That there are multiple changes associated with the rabbit allotypic specificities has always been probable because they have been detected by precipitating antisera, rather than inhibition methods, implying that several antigenic sites are involved.

The pattern of residue changes in the variable section, consisting of one hypervariable region (positions 95–115), two less variable regions (positions 30–35 and 50–60) and fewer changes elsewhere, is in general agreement with that found on comparison of sequences of the heavy chains of myeloma proteins (Press & Hogg, 1969; Edelman *et al.* 1969; Wikler, Köhler, Shinoda & Putnam, 1969). The same pattern is also apparent from comparison of the much more extensive results on the variable sections of the light chains (see Wu & Kabat, 1970).

In our work, if the allotypically related residues do in fact determine the allotypic specificity, it would be expected that there would be absolute constancy in 16 positions in the variable section of the γ-chain of a given allotype. In five of these positions, however, alternative residues have been detected and it may be that they reflect the further subdivision of the allotypic specificities that has been reported (Oudin, 1960).

Numerous attempts to explain the genetic origin of this complex pattern of variability have been made, most recently by Milstein & Pink (1970), Gally & Edelman (1970), Wu & Kabat (1970), Weigert, Cesari, Yonkovitch & Cohn (1970), Hood, Eichmann, Lackland, Krause & Ohms (1970), Baglioni (1970) and Mäkelä & Cross (1970). There appears to be general agreement that there are two separate genes coding for the variable and constant sections of both heavy and light chains. The present confirmation that the Aa1, Aa2 and Aa3 allotypic specificities are determined by sequences in the variable region, when considered together with the observation that these specificities are found in any combination with the A11 and A12 (Mandy & Todd, 1968) and A14 and A15 (Dubiski, 1969) specificities in the constant section of the γ-chain, and also of μ- (Todd, 1963), α- (Feinstein, 1963) and ϵ-chains (Kindt & Todd, 1969), supports this view. In breeding experiments (Mage, Young-Cooper & Alexander, 1971), one recombinant of the 'a' locus gene with A14/15 has been detected among 151 progeny of 42 different litters. Other genetic markers suggest that the parentage of the recombinant was correct and hence, if this rate of recombination is confirmed, supports the view that distinct genes do control the synthesis of the variable and constant sections of the γ-chain. More extensive results will be necessary to establish the map distance between these genes. In earlier experiments no somatic recombination had been observed, but the methods available gave an accuracy of only about 5% (Landucci-Tosi, Mage & Dubiski, 1970; see also Kindt, Mandy & Todd, 1970).

Table 14. *Comparison of differences in amino acid content between Fd fragment of Aa1 and Aa3 allotypes by analysis and by sequence results*

For further details see the Discussion section.

Sequence results		Analysis results	
This paper, and Wilkinson (1969*a*)		Koshland (1967)	Inman & Reisfeld (1968)
Arg	−1	−1	−1
Thr	−2	−2	−2
Glu	+1	+1	+1
Pro	−2	−1	−1
Ala	+3	+2 (→3)	+2
Ile	−1	−1	−1
Val	−1	−1	−1 (→2)
Phe	+0.6	+1	+1

Wu & Kabat (1970) have made the additional suggestion that there may be an episome coding for the short hypervariable section and this inserts into the two fusing genes. Since our results emphasize the quite exceptional variability in the region of residues 95–115 they are compatible with some such unique mechanism. However, as Mäkelä & Cross (1970) noted, the observed pattern of one hypervariable region, two less variable regions, fewer changes in some positions and stability in others is similar to that found in the analysis of spontaneous mutations in the rII region of the T4 bacteriophage (Benzer, 1961). There seems no reason to invoke any special mechanism such as an episome, though the origin of spontaneous mutation is unknown. The finding of allotype-related sequence changes throughout the variable region suggests that there is only one or a small number of copies of each allelic gene, as it is difficult to envisage the method of conserving the allotype determinants if there are very many copies. If this is correct, exceptional rates of somatic mutation, probably associated with cell division, must occur in the precursors of the immunocompetent cells.

The origin of the postulated allotypic determinants remains obscure. In some positions, the residue changes necessitate a double nucleotide-base change in the codon, implying that frequent mutation was occurring during the divergence of the allotypes and yet the allotypes are now stable. One suggestion (W. F. Bodmer, unpublished work) would be that controlling rather than structural allelic genes are responsible for the allotypes. The function of the controlling gene would be the prevention or elimination of mutations in certain sections of a highly mutable structural gene coding for the variable region, particularly in those sections coding for the allotypic determinants. Such a system could reconcile the wide sequence differences between the allotypic sequences with their present stability. The controlling gene might also influence the sequence changes determining the antibody-combining specificity and hence lead to some relationship between the allotype and the range of combining specificity. However, there is no simple dependence of recognizable specificities on sequence (Porter, 1971), so the testing of such a hypothesis does not seem feasible.

We thank Professor P. G. H. Gell for typing of the rabbit sera, Professor W. F. Bodmer, Mr C. E. Fisher and Dr D. B. Roberts for much valuable discussion and Mr T. Gascoyne and Mr A. C. Willis for their excellent technical help.

REFERENCES

Baglioni, C. (1970). *Proc. R. Soc.* B, **176**, 329.
Benzer, S. (1961). *Proc. natn. Acad. Sci. U.S.A.* **47**, 403.
Cebra, J. J., Steiner, L. A. & Porter, R. R. (1968). *Biochem. J.* **107**, 79.
Dixon, H. B. F. & Perham, R. N. (1968). *Biochem. J.* **109**, 312.
Dubiski, S. (1969). *J. Immun.* **103**, 120.
Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, U. & Waxdal, M. J. (1969). *Proc. natn. Acad. Sci. U.S.A.* **63**, 78.
Feinstein, A. (1963). *Nature, Lond.*, **199**, 1197.
Fleischman, J. B. (1971). *Biochemistry, Easton* (in the Press).
Fruchter, R. G., Jackson, S. A., Mole, L. E. & Porter, R. R. (1970). *Biochem. J.* **116**, 249.
Gally, J. A. & Edelman, G. M. (1970). *Nature, Lond.*, **227**, 341.
Hilschmann, N. & Craig, L. C. (1965). *Proc. natn. Acad. Sci. U.S.A.* **53**, 1430.
Hood, L., Eichmann, K., Lackland, H., Krause, R. M. & Ohms, J. J. (1970). *Nature, Lond.*, **228**, 1040.
Inman, J. K. & Reisfeld, R. A. (1968). *Immunochemistry*, **5**, 415.
Kindt, T. J., Mandy, W. J. & Todd, C. W. (1970). *Biochemistry, Easton*, **9**, 2028.
Kindt, T. J. & Todd, C. W. (1969). *J. exp. Med.* **130**, 859.
Koshland, M. E. (1967). *Cold Spring Harb. Symp. quant. Biol.* **32**, 119.
Koshland, M. E., Davis, J. J. & Fujita, N. J. (1969). *Proc. natn. Acad. Sci. U.S.A.* **63**, 1274.
Koshland, M. E., Reisfeld, R. A. & Dray, S. (1968). *Immunochemistry*, **5**, 471.
Landucci-Tosi, S., Mage, R. G. & Dubiski, S. (1970). *J. Immun.* **104**, 641.
Mage, R. G., Young-Cooper, G. O. & Alexander, C. (1971). *Nature New Biol., Lond.*, **230**, 63.
Mäkelä, O. & Cross, A. M. (1970). *Prog. Allergy*, **14**, 145.
Mandy, W. J. & Todd, C. W. (1968). *Vox. Sang.* **14**, 264.
Milstein, C. & Pink, J. R. L. (1970). *Prog. Biophys. molec. Biol.* **21**, 211.
Mole, L. E. (1971). D.Phil. Thesis: University of Oxford.
O'Donnell, I. J., Frangione, B. & Porter, R. R. (1970). *Biochem. J.* **116**, 261.
Oudin, J. (1960). *J. exp. Med.* **112**, 107.
Porter, R. R. (1967). *Biochem. J.* **105**, 417.
Porter, R. R. (1971). *Harvey Lect.* (in the Press).
Prahl, J. W., Mandy, W. J. & Todd, C. W. (1969). *Biochemistry, Easton*, **8**, 4935.
Prahl, J. W. & Porter, R. R. (1968). *Biochem. J.* **107**, 753.
Press, E. M. & Hogg, N. M. (1969). *Nature, Lond.*, **223**, 807.
Tang, J. & Hartley, B. S. (1967). *Biochem. J.* **102**, 593.
Todd, C. W. (1963). *Biochem. biophys. Res. Commun.* **11**, 170.
Weigert, M. G., Cesari, I. M., Yonkovitch, S. J. & Cohn, M. (1970). *Nature, Lond.*, **228**, 1045.
Wikler, M., Köhler, H., Shinoda, T. & Putnam, F. W. (1969). *Science, N.Y.*, **163**, 75.
Wilkinson, J. M. (1969*a*). *Biochem. J.* **112**, 173.
Wilkinson, J. M. (1969*b*). *FEBS Lett.* **4**, 170.
Wilkinson, J. M., Press, E. M. & Porter, R. R. (1966). *Biochem. J.* **100**, 303.
Wu, T. T. & Kabat, E. A. (1970). *J. exp. Med.* **132**, 211.