

Amino Acid Sequence of the *N*-Terminal 139 Residues of Light Chain Derived from a Homogeneous Rabbit Antibody

By JEAN-CLAUDE JATON

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland

(Received 3 October 1973)

The amino acid sequence of the *N*-terminal 139 residues of the L (light) chain derived from a homogeneous rabbit antibody (designated BS-1) to type III pneumococci was determined. A combination of methods involving tryptic cleavage restricted to the 2 arginine residues of the molecule and mild acid hydrolysis of a labile peptide bond between the V (variable) and C (constant) regions of the L chain (Fraser *et al.*, 1972) allowed the isolation of two large peptides comprising the entire V region (residues 1–109); these peptides were suitable for automated Edman degradation. The complete sequence analysis of the V region was carried out with only 4 μ mol of L chain. This material was homogeneous, although minor variant sequences, if present at the 10% value, would not have been detected. The L chain contains 3 intrachain disulphide bridges, whose pairing was established by diagonal electrophoresis: there is one V-region bridge between positions 23 and 88 and one C-region bridge between positions 134 and 194; the third one connects V and C domains between positions 80 and 171. When compared with the basic sequence of human κ chains, rabbit L chain BS-1 appears to be more similar to the $V_{\kappa I}$ prototype sequence than to $V_{\kappa II}$ or $V_{\kappa III}$ sequences, where $V_{\kappa I}$, $V_{\kappa II}$ and $V_{\kappa III}$ represent subgroups I, II and III respectively of V regions of κ light chains. The V regions of rabbit heavy and light chains are homologous to each other. The presence of two clusters of 3 glycine residues in positions 94–96 and 99–101 respectively is remarkable. Residues 94–96 may be related to antibody complementarity whereas residues 99–101 function probably as a pivot permitting the combining region of the L chain to make optimal contact with the antigenic determinant (Wu & Kabat, 1970).

Amino acid sequence studies of a large number of myeloma proteins have shown that both the heavy (H)* and light (L) chains of immunoglobulins comprise a variable *N*-terminal half (V region) of 106–110 residues and an invariant *C*-terminal half (C region) (for review, see Edelman & Gall, 1969).

Homogeneous antibodies stimulated by bacterial polysaccharide antigens are suitable material for study of the primary structure of both H (Fleischman, 1971, 1973) and L chains (Hood *et al.*, 1970; Jaton *et al.*, 1971; Kindt *et al.*, 1972; Braun & Jaton, 1973). On the basis of the available data, sequence variability between anti-polysaccharide antibodies was observed within sections 31–34 and 47–62 of the H chains (Jaton & Braun, 1972). In contrast, section 80–94 of a_1 , a_2 and a_3 homogeneous H chains showed little variation (Jaton *et al.*, 1973).

* Abbreviations IgG, immunoglobulin G; H and L chains, heavy and light chains of IgG; V and C regions, variable and constant regions of IgG respectively; $V_{\kappa I}$, $V_{\kappa II}$ and $V_{\kappa III}$, subgroups I, II and III respectively of V regions of κ light chains; V_L and C_L , V and C regions of L chains; Fab', pepsin fragment which consists of a complete L chain and of the *N*-terminal half of the H chain.

CmCys, \square Glu (in amino acid sequences and Tables), *S*-carboxymethylcysteine and pyrrolidonecarboxylic acid respectively.

No complete sequences of V regions of rabbit L chains have yet been determined; however, partial sequences from several pneumococcal antibody L chains, as well as from non-immune pooled chains, have been reported (Strosberg *et al.*, 1972). Similarly, half-cystine-containing peptides representing the intrachain disulphide bonds of an L chain derived from a restricted anti-azobenzoate antibody have been described (Appella *et al.*, 1971). The knowledge of the primary and tertiary structures of rabbit L chain is of great interest not only for the structure–function relationship but also for the location of the group b allotypic markers which, in contrast with their counterpart H chains, seem to be present in the C region (Appella *et al.*, 1969; Frangione, 1969; Kindt *et al.*, 1972), even though it was suggested that the *N*-terminal sequence also correlates with each L-chain allotype (Hood *et al.*, 1971).

In the present investigation, I report the amino acid sequence of the V region and of the beginning of the C domain of the L chain derived from a homogeneous rabbit antibody to type III pneumococci. This antibody, of allotypes a_2 , b_4 , designated BS-1, is specific for the simple antigenic determinant, cellobiuronic acid, the repeating disaccharide of the capsular pneumococcal polysaccharide type III.

Experimental

Materials

Trypsin treated with L-1-chloro-4-phenyl-3-tosyl-aminobutan-2-one, carboxypeptidases A and B treated with di-isopropyl fluorophosphate, α -chymotrypsin (three times crystallized) and thermolysin were purchased from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Dithiothreitol (A grade) was from Calbiochem, Los Angeles, Calif., U.S.A., iodo[2-¹⁴C]acetic acid (specific radioactivity 30mCi/mmol) was from The Radiochemical Centre, Amersham, Bucks., U.K. and hydriodic acid (65% v/v), stabilized with 0.03% (w/v) H₃PO₄, was from BDH Chemicals Ltd., Poole, Dorset, U.K. 4-Sulpho-phenyl isothiocyanate (sodium salt, monohydrate, sequal grade) was obtained from Pierce Chemical Co., Rockford, Ill., U.S.A. and citraconic anhydride was from Aldrich Chemical Co., Milwaukee, Wis., U.S.A. Sequencer-grade reagents and solvents used in the automated Edman degradation were purchased from Beckman International, Geneva, Switzerland. All other reagents and solvents were of the best grade available.

Methods

Antibody production and purification. Rabbit BS-1 was immunized with type III pneumococcal vaccine by a scheme described by Kimball *et al.* (1971). The antibody which showed mainly a single band in the γ -globulin region was isolated with an immunoadsorbent by the method of Jaton *et al.* (1971). It was repurified by preparative agarose block electrophoresis to eliminate 10–15% of a fast-moving component. Details of the method of purification and preparation of H and L chains have already been described (Jaton & Braun, 1972). The antibody BS-1 carried a₂, b₄ allotypes.

Complete reduction of the light chain. The light chain (100mg) was fully reduced and alkylated with iodo[2-¹⁴C]acetic acid by the technique of O'Donnell *et al.* (1970).

Citraconylation of the reduced and [2-¹⁴C]alkylated L chain. This was done essentially by the method of Gibbons & Perham (1970). The freeze-dried L chain (100mg) was dissolved in 7M-guanidine hydrochloride (10ml) and a 25-fold molar excess of citraconic anhydride with respect to lysine residues was added gradually within 45min; the pH of the reaction mixture was maintained between 8.5 and 9.0 by dropwise addition of 5M-NaOH solution. Excess of reagents and salts were removed by dialysis against water and the protein was freeze-dried. The extent of modification of amino groups was determined by the trinitrobenzenesulphonic acid test (Habeeb, 1966).

Dilute acid hydrolysis. Large citraconyl-fragment A (1–2 μ mol) derived from L chain BS-1 was subjected to acid hydrolysis in either 70% (v/v) formic acid

(2ml), or in 7M-guanidine hydrochloride containing 10% (v/v) acetic acid, pH2.5 (2ml), for 120h at 37°C as suggested by Fraser *et al.* (1972).

Unblocking of citraconyl-peptides. The citraconyl-fragment B was incubated in 1–2ml of acetate buffer, pH4.2, for 6h at 37°C, similarly to the procedure of Habeeb & Atassi (1970); the solution was then dialysed for 12h against several changes of water and the content of the bag freeze-dried. No appreciable losses of material occurred during dialysis.

Analytical methods. Preparative high-voltage paper electrophoresis at pH6.5 was done as described by Press *et al.* (1966). Neutral peptides were rerun at pH3.5. Diagonal electrophoresis of peptic digest of L chain BS-1 (15mg) at pH6.5 was carried out by the method of Brown & Hartley (1966): peptides were eluted from the paper with 0.02M-NH₃ after light staining with ninhydrin solution in acetone (0.05% w/v). Peptides were detected by the ninhydrin-cadmium stain (Dreyer & Bynum, 1967), and those containing tyrosine, tryptophan and arginine by specific stains (Smith, 1960). Radioactive peptides were revealed by radioautography for 16h by using Kodak Royal Blue Medical X-ray film.

Enzymic digestions of 50 to 300nmol of peptides with trypsin and α -chymotrypsin were carried out for 3h at 37°C, by using 40–100 μ g of enzyme in the presence of 0.4–1.5ml of 1% (w/v) NH₄HCO₃. Peptic digestion was done in 5% (v/v) formic acid for 16h at 37°C and thermolysin digestion was done at an enzyme/protein ratio of 1:500 in 1% (w/v) NH₄HCO₃ for 1h at 25°C. Hydrolyses with carboxypeptidases were conducted as recommended by Ambler (1967).

Amino acid analyses on a Beckman model 121 Amino Acid Analyzer and the determination of sequences of small peptides by the 'dansyl-Edman' procedure (Gray, 1967) were as described by Jaton & Braun (1972). Whenever possible, amide residues were assigned on the basis of electrophoretic mobility at pH6.5 (Offord, 1966).

The determination of the amino acid sequence of sections of L chain and of large peptides was done in the Beckman sequencer model 890B, equipped with an undercut cup and N₂ flush. Protein or peptides (250–600nmol) were dissolved in 0.4ml of 1M-acetic acid and degraded by using the conventional Quadrol programme (Edman & Begg, 1967). Alternatively, large tryptic peptides (50–60 residues) were first manually treated with an excess of 4-sulpho-phenyl isothiocyanate to prevent losses during automated sequential degradation (Braunitzer *et al.*, 1970). In a typical experiment, the peptide (600 nmol) was dissolved in 400 μ l of dimethylallylamine buffer and incubated for 1h at 63°C in the presence of 50 μ l of 3% (w/v) 4-sulphophenyl isothiocyanate (sodium salt) in water. The reaction mixture was then transferred to the cup of the sequencer and dried down; a peptide programme utilizing the volatile buffer

dimethylallylamine was started at the cleavage step with heptafluorobutyric acid. The programme was adapted from the original one designed by Beckman, Spinco Division. The identification of the phenylthiohydantoin-amino acid derivatives was accomplished by: (1) g.l.c. (Pisano & Bronzert, 1969) and (2) amino acid analysis after conversion of the phenylthiohydantoin derivatives into free amino acids by hydrolysis with 65% (w/v) HI for 20h at 125°C (Smithies *et al.*, 1971), as described by Jatton & Braun (1972). The overall repetitive yield in sequencing experiments varied from 90% for fragments A and A₃ to 94% for fragments A₁ and B respectively, and the absolute yield of the phenylthiohydantoin-amino acid derivatives at the first step ranged between 22 and 56%.

Radioactivity measurements were carried out in vials containing 10ml of scintillation liquid (Bray, 1960) in a Nuclear-Chicago liquid-scintillation counter.

Results

Isolation and characterization of peptides from a tryptic digest of [2-¹⁴C]carboxymethylated and citraconylated L chain BS-1

L chain BS-1 contains only 2 arginine residues/molecule (Table 1). Advantage was taken of this fact by cleaving the L chain with trypsin after lysine

residues had been blocked with citraconic anhydride. The components of the tryptic digest of [2-¹⁴C]-carboxymethylated and citraconylated L chain were separated by gel filtration into two fractions, A and B (Fig. 1). A portion of radioactive material under peaks A and B was not dialysable. Fractions A and B were thus dialysed overnight against water to remove excess of salts and freeze-dried. The amino acid composition (Table 1) indicates that fragment A contains about 150 amino acids including 5 half-cystine residues, whereas fragment B has 61 residues among which is 1 half-cystine. Both have arginine at the C-terminal position. Three fractions are expected to occur in a tryptic digest of citraconylated L chain containing 2 arginine residues. Two arginine peptides were accounted for; the third fraction is likely to contain the C-terminal tripeptide Gly-Asp-Cys in positions 212–214 (Appella *et al.*, 1969), which escaped detection under our experimental conditions. On the basis of 4.3 μmol of L chain, peptides A and B were recovered in 80% and 82% yield respectively. Citraconyl groups were removed from fragment B and the resulting deblocked peptide (600nmol) was subjected to automated sequential degradation after prior treatment with 4-sulphophenyl isothiocyanate (see under 'Methods'). The amino acid sequence of the first 37 residues of fragment B (Fig. 2) is identical with the N-terminal sequence of L chain BS-1 determined by Braun & Jatton (1973). It is

Table 1. Amino acid composition of the citraconyl peptides isolated from a tryptic digest of reduced and carboxymethylated L chain BS-1

Values are residues/molecule of peptide. Values in parentheses are integral values confirmed by sequence analysis. Fragments A₁ and A₃ were released from citraconyl fragment A after treatment in dilute acid at 37°C (see the text). N.D., not determined.

Amino acid	Fragment A	Fragment A ₁	Fragment A ₃	Fragment B	Carboxymethylated L chain
Lys	6.1	4.0	2.0 (2)	4.2 (4)	10.5
His	1.0	0.9	—	—	1.5
Arg	1.1	0.8	—	1.2 (1)	2.2
CmCys	5.6	3.2	1.8 (2)	1.2 (1)	6.6
Asp	16.6	13.8	3.3 (3)	1.1 (1)	17.9
Thr	23.0	17.5	6.0 (6)	4.8 (5)	27.8
Ser	14.2	7.9	3.9 (4)	9.1 (9)	22.5
Glu	13.8	8.9	4.2 (4)	7.0 (7)	20.0
Pro	6.3	6.2	0.3 (0)	5.2 (5)	11.0
Gly	14.3	3.1	9.8 (11)	4.6 (5)	18.8
Ala	9.9	6.0	3.0 (3)	5.1 (5)	13.8
Val	13.8	11.0	3.2 (3)	5.6 (6)	20.1
Met	—	—	—	0.9 (1)	1.2
Ile	3.6	2.8	1.1 (1)	3.1 (3)	7.2
Leu	6.1	3.6	2.0 (2)	4.2 (4)	10.1
Tyr	6.7	3.8	2.9 (3)	3.1 (3)	10.1
Phe	7.8	3.6	3.8 (4)	— (0)	8.0
Trp	N.D.	N.D.	—*	+ (1)	+ †
Total residues	148.8	97.1	47.3	60.4	209.3

* Measured with Ehrlich's reagent.

† Excluding tryptophan.

concluded that fragment B is the *N*-terminal 61-residue peptide of the L chain. Consequently, the second arginine-containing peptide, A, should follow fragment B and extend between positions 62 and 211.

Citraconyl-fragment A (500nmol) was subjected directly to automated Edman degradation by using the Quadrol programme. A single amino acid derivative was found at each of the *N*-terminal 28 positions (Fig. 2). An analysis of yield at each step is shown in Fig. 3. The absolute yield of phenylthiohydantoin-phenylalanine at the first step was 56%. The drastic fall in yield of phenylthiohydantoin-amino acids of 90% over 27 residues prevented further degradation of peptide A. The completion of the determination of the sequence of peptide A₃ by conventional methods is described below. The *N*-terminal sequence of fragment A is Phe-Lys-Gly-Ser-Gly-Ser-Gly-Thr, which is characteristic of most κ light chains and occurs after arginine in position 61 (Putnam *et al.*, 1967). Of special interest is the identification of 2 half-cystine residues in positions 19 and 27 (Fig. 3), i.e. in corresponding positions 80 and 88 respectively of the L chain (Fig. 6). The significance of the cysteine residue at position 80 is discussed below.

Partial sequence of fragment A

Fragment A was further subjected to dilute acid hydrolysis, as it has been shown that a specific

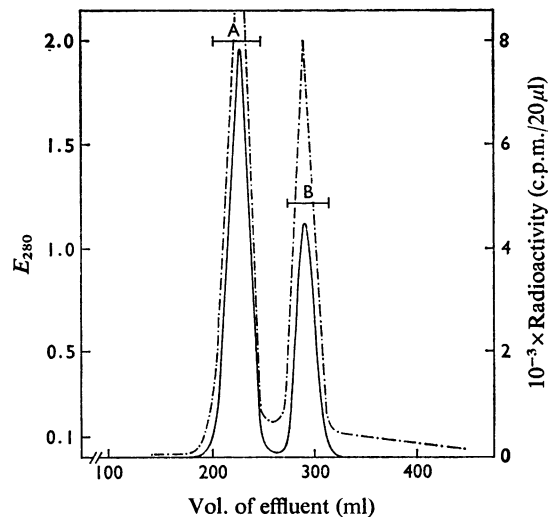


Fig. 1. Elution profile of a tryptic digest of [2-¹⁴C]carboxymethylated and citraconylated L chain BS-1

The digest (100 mg) was loaded on a column (2 cm × 200 cm) of Sephadex G-50 (fine grade) and developed in 6M-guanidine hydrochloride containing 0.1M-Tris-HCl buffer, pH 8.0. See the text for details. —, E₂₈₀; - - -, radioactivity (c.p.m./20 μl of eluate).

Residue position	1	10	20
L-chain BS-1*	Asp-Val-Val-Met-Thr-Gln-Thr-Pro-Ala-Ser-Val-Ser-Glu-Pro-Val-Gly-Gly-Thr-Val-Thr		
Fragment B	-----		
Fragment A	70 80		
Fragment A ₃	Phe-Lys-Gly-Ser-Gly-Ser-Gly-Thr-Glu-Phe-Thr-Leu-Thr-Ile-Ser-Asp-Leu-Glu-Cyst-Ala		
Fragment A ₁	110 120		
Residue position	30	35	
L-chain BS-1*	Ile-Lys-Cyst-Gln-Ala-Ser-Gln-Ser-Ile		
Fragment B	----- Tyr-Ser-Gly-Leu-Ala-Trp-Tyr-Glx		
Fragment A	Asx-Ala-Ala-Thr-Tyr-Phe-Cyst-Glx		
Fragment A ₃	-----		
Fragment A ₁	130 Val-Thr-Ile-Val-Cyst-Val-Ala-Asn-Lys-Tyr		

Fig. 2. Automated sequential degradations of L chain BS-1 and fragments A, A₁, A₃ and B derived thereof

L chain BS-1 (250nmol), fragments B (600nmol), A (500nmol), A₃ (300nmol) and A₁ (450nmol) respectively were used for sequence determination. A conventional Quadrol programme was used for L chain BS-1, fragments A and A₁, whereas a volatile buffer system and sulphophenyl isothiocyanate were employed for peptides B and A₃ (see under Methods for details). The solid line indicates identical sequence with that shown immediately above. The numbering of the residues in the fragments indicates their position in the complete sequence.

* From Braun & Jaton (1973).

† Determined by counting a portion of the phenylthiohydantoin-S-carboxymethylcysteine derivative for radioactivity (see also Fig. 3).

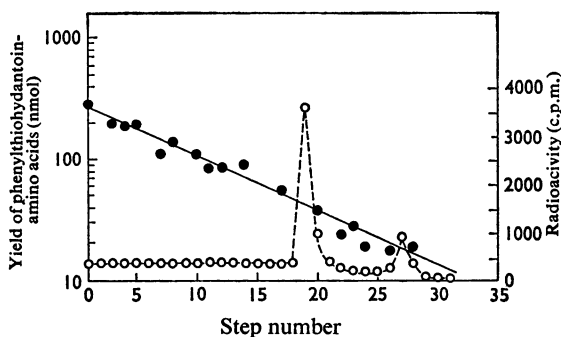


Fig. 3. Quantitative yields of phenylthiohydantoin-amino acids obtained at each step of the automated Edman degradation of fragment A

Yields were computed by comparing peak heights of samples with those of relevant standard derivatives by g.l.c. (●) and c.p.m. of a 5% sample of the phenylthiohydantoin-amino acid at each step (○). Radioactivity at steps 19 and 27 indicates a *S*-carboxymethylcysteine residue. The absolute yield of phenylthiohydantoin-phenylalanine at the first step was 56%, on the basis of 500 nmol of fragment A.

cleavage of the labile bond Asp₁₀₉-Pro₁₁₀ between V and C domains of rabbit L chains can occur (Fraser *et al.*, 1972). The separation of the cleavage products is depicted in Fig. 4. Two major fractions designated A₁ and A₃ were obtained. Fractions A₀ and A₂ probably contain aggregated and/or uncleaved material, present in small amounts, and were not analysed further. Peptides A₁ and A₃ were dialysed overnight against several changes of water and freeze-dried. No significant loss of material was found after dialysis.

Peptide A₃ comprises 48 amino acids among which are 2 half-cystine residues but no arginine (Table 1). On the basis of 3 μmol of fragment A, peptide A₃ was recovered in 65% yield. Its *N*-terminal sequence (Fig. 2) is identical with that of fragment A from which it was derived, indicating that fraction A₃ contains the V-region peptide extending from positions 62-109.

The other fragment, A₁, has about 100 amino acids on the basis of 1 arginine residue (Table 1). Its yield was 63%. Arginine was determined to be the *C*-terminal residue by carboxypeptidase B digestion and proline to be the *N*-terminal residue by the dansylation technique. This suggests that the mild acid cleavage of fragment A occurred at the bond between aspartic acid 109 and proline 110. The purity of the large fragment A₁ is indicated by the single sequence found in its first *N*-terminal 30 positions (Fig. 2). This sequence, homologous to human κ chains (Putnam *et al.*, 1967) and identical with other rabbit L chains (Strosberg *et al.*, 1972), proceeds

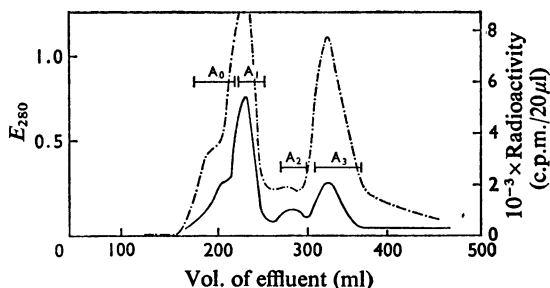


Fig. 4. Separation of the cleavage products of fragment A subjected to dilute acid hydrolysis

Fragment A was treated in 7M-guanidine hydrochloride containing 10% acetic acid, pH 2.5, for 120 h at 37°C. The digest (50 mg) was loaded on a column (2 cm × 200 cm) of Sephadex G-50 (fine grade) and developed in 5M-guanidine hydrochloride, pH 5.5. See the text for details. —, E₂₈₀; - - - -, radioactivity (c.p.m./20 μl of eluate).

starting with residue 110 at the beginning of the constant region. Fragment A₁ therefore comprises the constant domain (residues 110-211) with the exception of the three *C*-terminal residues of the L chain.

The ordering of the citraconyl fragments and of the peptides derived from them is summarized in Fig. 5. The entire V region comprises two peptides: fragments B and A₃, whereas the C domain is made up of the large fragment A₁ plus the *C*-terminal tripeptide.

Sequence determination of fragment A₃

Fragment A₃ was briefly digested with thermolysin (see under 'Methods'). Five major peptides designated Th₁, Th₂, Th₃, ThN₁ and ThN₂ were isolated by paper electrophoresis in satisfactory yield. Their characterization and sequence are described in Table 2 and Fig. 6.

Peptide ThN₁ is a tetrapeptide with the sequence Phe-Thr-Leu-Thr.

The *N*-terminal sequence of the nonapeptide ThN₂ was Phe-Lys-Gly-Ser; this and its amino acid composition indicate that peptide ThN₂ is the *N*-terminal peptide of fragment A₃.

Peptide Th₁ is a dodecapeptide rich in glycine and valine, which contains 1 lysine residue. After tryptic cleavage two peptides were obtained: the *N*-terminal decapeptide T-Th_{1b} and the *C*-terminal dipeptide Gly-Asp. The sequence of peptides Th₁ and T-Th_{1b} is shown in Fig. 6.

Peptide Th₂ has the *N*-terminal sequence Phe-CmCys-Glx-Gly-Ser-Thr-Tyr. The position of *S*-carboxymethylcysteine in this 11-residue peptide was confirmed by the removal of 85% of the radioactivity in the peptide after the second degradation

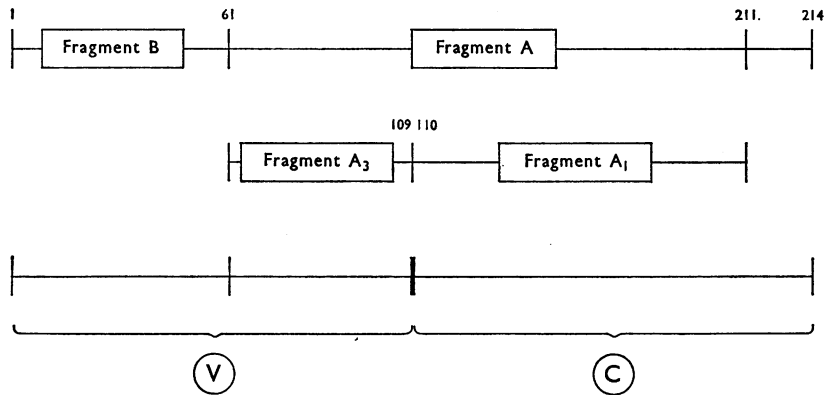


Fig. 5. Alignment of the citraconylated peptides derived from L chain BS-1

Fragments A and B were isolated from a tryptic digest of carboxymethylated and citraconylated L chain. Fragments A₁ and A₃ were derived from fragment A after dilute acid hydrolysis (see the text). V and C denote the variable and constant halves of the L chain.

Table 2. Amino acid composition and properties of constituent peptides of the fragment A₃ of L chain BS-1

Values are residues/molecule of peptide. Values in parentheses are integral values confirmed by sequence analysis. Thermolysin and chymotryptic peptides were derived from fragment A₃. Peptides Ch-Th_{2a} and Ch-Th_{2b} were derived from a chymotryptic digest of peptide Th₂. Peptides T-Th_{1a} and T-Th_{1b} were derived from peptide Th₁ by trypsin digestion. Mobilities are expressed relative to Asp (= -1.0). Yields are based on 1100nmol of fragment A₃.

Amino acid	Thermolysin peptides					Chymotryptic peptides		Tryptic peptides	
	Th1	Th2	Th3	ThN ₁	ThN ₂	Ch-Th _{2a}	Ch-Th _{2b}	T-Th _{1a}	T-Th _{1b}
Lys	1.1	—	—	—	1.0	—	—	—	1.0 (1)
Arg	—	—	—	—	—	—	—	—	—
CmCys*	—	0.8 (1)	0.9	—	—	—	0.7 (1)	—	—
Asp	1.0	—	2.1	—	—	—	—	1.0 (1)	—
Thr	1.1	0.8 (1)	0.9	2.2 (2)	1.2	—	0.8 (1)	—	0.9 (1)
Ser	—	1.2 (1)	1.1	—	1.9	—	1.2 (1)	—	—
Glu	1.0	1.1 (1)	1.0	—	1.3	—	1.2 (1)	—	1.0 (1)
Pro	—	—	—	—	—	—	—	—	—
Gly	3.6	4.0 (4)	—	—	2.9	2.9 (3)	0.8 (1)	0.9 (1)	2.8 (3)
Ala	—	—	2.9	—	—	—	—	—	—
Val	2.8	—	—	—	—	—	—	—	2.9 (3)†
Ile	—	—	0.9	—	—	—	—	—	—
Leu	—	—	1.0	0.8 (1)	—	—	—	—	—
Tyr	—	1.8 (2)	1.0	—	—	1.0 (1)	0.9 (1)	—	—
Phe	1.0	0.9 (1)	—	1.0 (1)	0.9	—	1.1 (1)	—	0.9 (1)
Total residues	12	11	12	4	9	4	7	2	10
Mobility at pH 6.5	-0.22	-0.48	-0.72	0	0	0	-0.68	-0.79	0
Yield (%)	29	31	35	29	35	29	25	12	15

* Value not corrected for losses during hydrolysis.

† Value after 96 h hydrolysis.

step. Peptide Th₂ was further fragmented with α -chymotrypsin and gave rise to two peptides: Ch-Th_{2a} and Ch-Th_{2b}. The sequence of peptide Ch-Th_{2b} is identical with that of peptide Th₂ in its N-terminal section. The tetrapeptide Ch-Th_{2a} should therefore

be the C-terminal part of the peptide Th₂; it has the sequence Gly-Gly-Gly-Tyr.

The N-terminal sequence Ile-Ser-Asp-Leu-Glu-Cys of the dodecapeptide Th₃ indicates that it is the cysteine-80-containing peptide of fragment A₃.

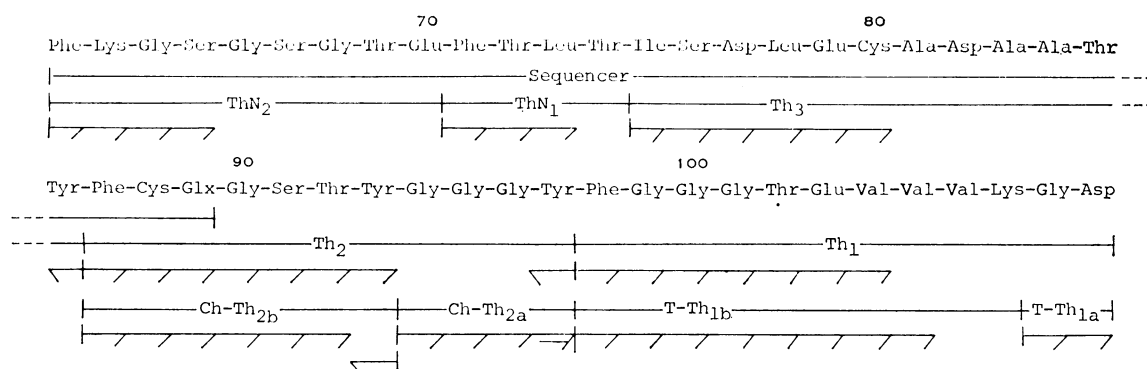


Fig. 6. Amino acid sequence of the fragment A_3 of L chain BS-1 (residues 62-109)

For experimental details see the text and Table 2. T, Th and Ch are tryptic, thermolysin and chymotryptic peptides respectively. \rightarrow , Amino acid residues identified by the 'dansyl-Edman' procedure; \Rightarrow identified as dansyl residue without hydrolysis; \leftarrow , identified by carboxypeptidase A or B. Residues identified with the sequencer are also shown.

Table 3. Amino acid composition and properties of constituent peptides of the N-terminal fragment B of L chain BS-1

Values are residues/molecule of peptide. Values in parentheses are integral values confirmed by sequence analysis. Chymotryptic peptides, ChT₃₋₁, ChT₃₋₂ and ChT₃₋₃, were derived from peptide T₃; and ChT₄₋₁ and ChT₄₋₂ from peptide T₄, respectively. Mobilities are expressed relative to Asp (= -1.0) or to Arg (= +1.0). Yields are based on 2.25 μ mol of starting material (N-terminal fragment B).

Amino acid	Tryptic peptides						Chymotryptic peptides						
	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	Ch-1	Ch-2	ChT ₃₋₁	ChT ₃₋₂	ChT ₃₋₃	ChT ₄₋₁	ChT ₄₋₂
Lys	1.1	1.2	2.1	—	1.2 (1)	1.1 (1)	1.0	2.1	—	—	2.1 (2)	—	—
Arg	—	—	—	1.2 (1)	—	—	1.1	—	—	—	—	—	0.9 (1)
CmCys*	—	0.7	0.8	—	—	—	—	—	0.7 (1)	—	—	—	—
Asp	1.0	—	0.3	—	—	—	—	—	—	—	—	—	—
Thr	3.9	—	—	1.1 (1)	—	—	1.1	—	—	—	—	1.1 (1)	—
Ser	2.1	2.9	2.8	4.2 (4)	—	—	4.2	—	1.7 (2)	1.1 (1)	—	0.8 (1)	3.1 (3)
Glu	2.0	3.8	5.0	—	1.0 (1)	—	—	3.0	2.1 (2)	—	3.0 (3)	—	—
Pro	2.1	0.2	3.4	—	2.8 (3)	—	—	3.1	—	—	2.9 (3)	—	—
Gly	2.1	1.0	1.9	1.1 (1)	1.3 (1)	—	1.0	1.0	—	1.0 (1)	0.9 (1)	—	1.2 (1)
Ala	0.9	1.9	2.1	2.1 (2)	—	—	2.1	—	1.0 (1)	0.9 (1)	—	1.0 (1)	1.0 (1)
Val	4.7	—	—	1.2 (1)	—	—	0.9	—	—	—	—	—	0.8 (1)
Met	0.9	—	—	—	—	—	—	—	—	—	—	—	—
Ile	1.1	1.0	1.1	—	—	0.9 (1)	—	—	0.9 (1)	—	—	—	—
Leu	—	1.2	1.2	1.0 (1)	—	2.0 (2)	0.9	2.2	—	1.1 (1)	—	1.1 (1)	—
Tyr	—	1.8	2.1	—	—	1.1 (1)	—	—	1.1 (1)	0.9 (1)	—	—	—
Trp	—	+†	+†	—	—	—	—	—	—	+† (1)	—	—	—
Total residues	22	17	23‡	11	6	5	12	11	8	6‡	9	4	7
Mobility at pH 6.5	-0.16	0	+0.10	+0.28	+0.37	+0.41	+0.24	+0.54	-0.61	0	—	0	+0.33
Yield (%)	15	10	19	27	9	44	30	32	10	12	19	15	16

* Value not corrected for losses during hydrolysis.

† Measured with Ehrlich reagent.

‡ Tryptophan assumed to be 1 residue/molecule of peptide.

Fig. 6 shows that the sequence of the N-terminal 28 residues of fragment A_3 overlaps peptides ThN₂, ThN₁, Th₃ and Th₂ in this order. The peptide Th₁ must therefore be the C-terminal peptide of fragment A_3 .

Sequence determination of tryptic fragment B

After removal of citraconyl groups, fragment B was digested with trypsin and the mixture fractionated by high-voltage paper electrophoresis at pH 6.5. Six peptides T₁-T₆ were isolated; peptides T₂ and T₃

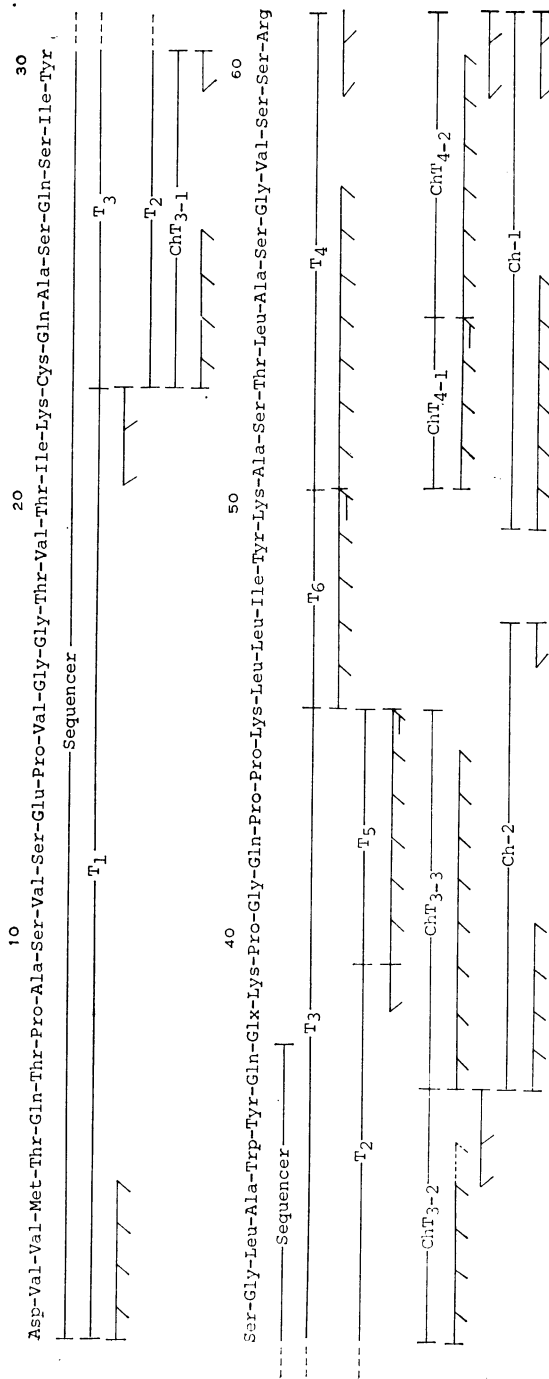


Fig. 7. Amino acid sequence of the N-terminal fragment B of L chain BS-1 (residues 1-61)

For experimental details see the text and Table 3. T and ChT are tryptic and chymotryptic peptides, respectively. Symbols as in Fig. 6.

Table 4. Peptic peptides isolated from L chain BS-1 by diagonal electrophoresis

Mobilities at pH 6.5 are expressed relative to Asp = -1. A loss of about 50% of the tyrosine residues in cysteine acid-containing peptides d₁, d₂, e and f was found as a result of performic acid oxidation. The numbered positions of the cysteine residues 23, 80, 88 and 134 are deduced from a comparison of the amino acid composition and partial N-terminal sequence of the cysteine acid-containing peptides with the extended sequence of L chain BS-1 (Fig. 9). The positions of the cysteine residues 171 and 194 are assigned by homology with human κ chains (Putnam *et al.*, 1967). The N-terminal sequence was established by the dansyl-Edman procedure.

Position of cysteine residues	Peptic peptide	Mobility at pH 6.5	Amino acid composition	N-terminal sequence
23	c	-0.20	(Ile, Lys, Cys, Glx, Ala, Ser)	Ile-Lys-Cys
88	d ₁	-0.43	(Tyr _{0.5} , Phe, Cys, Glx, Gly, Ser, Thr)	Tyr-Phe-Cys
88	d ₂	-0.47	(Tyr _{0.5} , Phe, Cys, Glx, Gly, Ser, Thr)	Tyr-Phe-Cys
80	b	-0.91	(Glx, Cys, Asx, Ala, Ala, Thr)	Glx-Cys-Ala
171	a	-0.11	(Ser, Lys, Thr, Pro, Glx, Asx, Ser, Ala, Asx, Cys, Thr)	Ser-Lys-Thr
134	f	-0.13	(Ile, Val, Cys, Val, Ala, Asx, Lys, Tyr _{0.5} , Phe, Pro, Asx, Val, Thr)	Ile-Val-Cys-Val
194	e	0	(Thr, Ser, Thr, Glx, Tyr _{0.5} , Asx, Ser, His, Lys, Glx, Tyr _{0.5} , Thr, Cys, Lys, Gly, Thr, Val)	Thr

The basic peptide T₅ contains 3 proline residues and has the sequence Pro-Gly-Gln-Pro-Pro-Lys. It is assigned as the C-terminal hexapeptide of fragment T₃. This suggests that the tryptic peptide T₂ arose from peptide T₃ by partial tryptic cleavage of the Lys₃₉-Pro₄₀ bond (Fig. 7).

Finally, the basic pentapeptide T₆ was isolated in good yield and has the sequence Leu-Leu-Ile-Tyr-Lys. Two peptides, Ch-1 and Ch-2, which were isolated from a chymotryptic digest of fragment B, overlap peptide T₆, which therefore should extend in positions 46-50 between peptides T₃ and T₄. The sequence determination of the N-terminal fragment B is summarized in Fig. 7.

The amino acid composition determined for fragments B and A₃, which represent the entire V domain, is in good agreement with that computed from the sequence (Table 1). The amino acid sequence of the N-terminal 139 residues of L chain BS-1 is shown in Fig. 8.

Intrachain disulphide bridges of L chain BS-1

L chain BS-1 contains 6.6 residues of S-carboxymethylcysteine/molecule, which indicates that this protein belongs to the K_B type, characterized by three disulphide bridges (Rejnek *et al.*, 1969). Poulsen *et al.* (1972) have demonstrated that the third, i.e. the 'extra' disulphide bridge connects the V and C domains. However, the position of the half-cystine within the V region has not yet been reported. The sequence of the V region (the present paper) and the partial sequence data from the C region (Appella *et al.*, 1971; Strosberg *et al.*, 1972) enabled us to study the pairing of the cysteine residues. Diagonal peptide 'mapping' was done at pH 6.5 with a peptic digest of the L chain. After performic acid oxidation and

electrophoresis in the second direction, three sets of two cysteine acid-containing peptides (a and b, c and d₁+d₂, and e and f respectively) were separated as outlined in Fig. 9 and characterized as summarized in Table 4.

Three cysteine acid-containing peptides, c, d₁ and d₂, constitute the first intrachain disulphide bridge 23-88 of the V region. Peptides d₁ and d₂ are identical in N-terminal sequence and probably in amino acid composition; they differ however slightly in electrophoretic mobility. Peptides b and a form the 'extra' disulphide bond of the rabbit K_B chain: 1 cysteine residue in one of these can be convincingly placed at position 80 in the sequence of BS-1 L chain (Fig. 8), whereas its partner is the cysteine-171 as deduced by homology with human and rabbit κ -chain sequences from the C region. The amino acid composition and partial sequence of the cysteine acid-containing peptides f and e indicate that they form the disulphide bridge 134-194 of the constant region in agreement with the results of Appella *et al.* (1971); the position of the cysteine residue-134 is deduced from the BS-1 sequence (Fig. 8) and that of the cysteine-194 is based on homology with human κ chains (Putnam *et al.*, 1967).

Discussion

L chain BS-1 contains 2 arginine residues. One of these is the fourth residue from the C-terminal end (Appella *et al.*, 1969) and the other is assumed to be located at position 61 in the V region (Freedlender & Haber, 1972). This suggested that tryptic cleavage restricted to the arginine residues would offer a method for obtaining large V-region fragments suitable for automated Edman degradation. Indeed, two large arginine-containing peptides, A (residues

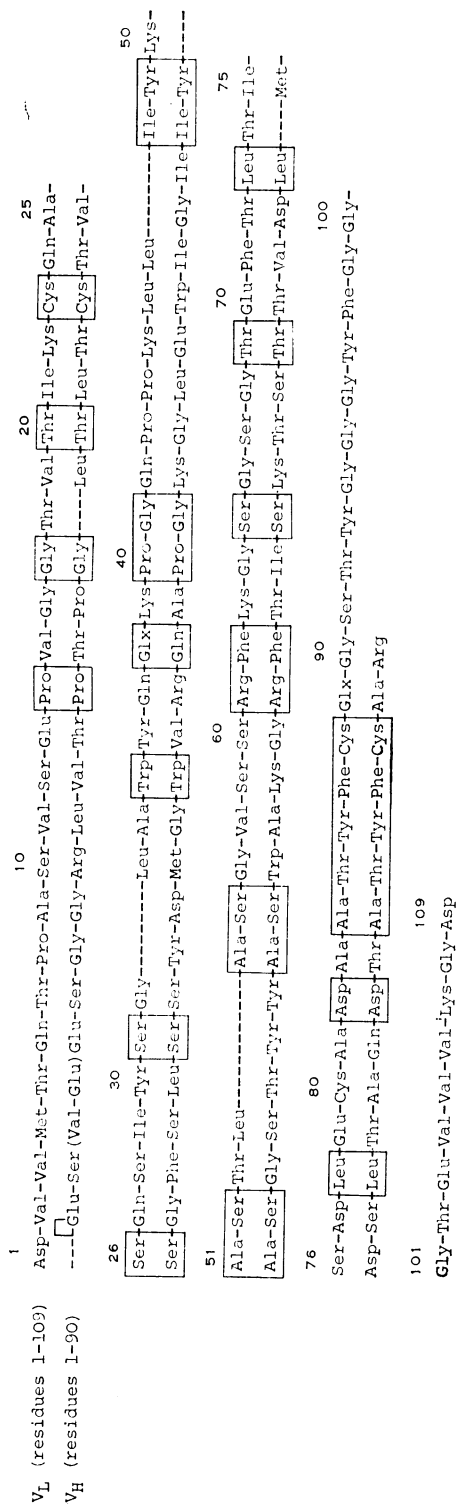


Fig. 10. Comparison of the V region of L chain BS-1 with the V region of anti-pneumococcal H chains

Identical residues are framed in boxes. Deletions (---) are introduced to maximize homology. The V_H region sequence of residues 1-65 is derived from the anti-pneumococcal type III antibody BS-5 (Jaton & Braun, 1972); the sequence of residues 65-75 is from a₁ pooled normal H chains (Mole *et al.*, 1971) and that of residues 76-90 is from the corresponding H chain BS-1 (Jaton *et al.*, 1973).

62-211) and B (residues 1-61), were isolated in high yield. Fragment A was further cleaved at the acid-labile Asp₁₀₉-Pro₁₁₀ bond (Fraser *et al.*, 1972) into two major peptides, A₁ and A₃. The entire V region is made up of only two large peptides, B and A₃, and fragment A₁ contains most of the C region (Fig. 5). Automated Edman degradation of peptides B and A₃ enabled us to determine 60% of the sequence (Fig. 2) and this knowledge is very valuable for ordering small tryptic, chymotryptic and thermolysin peptides. It should be pointed out that the procedures used in this work required only 4 μmol of L chain, an amount which is substantially smaller than that used with the traditional sequencing methodology.

Homogeneity of the V region of the L chain is shown by (1) the unique sequence determined with the protein sequencer in sections 1-37 and 62-89, (2) the high yield of the N-terminal fragment B (82%) and of peptide A₃ (65%), and (3) the recoveries of small peptides isolated from the hypervariable sections of fragments A₃ and B by paper electrophoresis, which compare favourably with those reported for human and mouse myeloma κ chains (Tables 2 and 3). However, minor variant peptides, if present at the 10% value, might have escaped detection by the techniques used here.

The observation of the partial tryptic cleavage at the Lys₃₉-Pro₄₀ bond is somewhat unexpected (Fig. 7). That this cleavage is a result of the acid treatment of fragment B during decitraconylation is unlikely, as only one N-terminal sequence could be determined when a portion of decitraconylated fragment B was subjected to automated Edman degradation (Fig. 2). The partial tryptic cleavage of this bond was also noted in the study of two other rabbit L chains (J.-C. Jaton, unpublished work) and was reported to occur, in homologous positions, in murine κ chains (McKean *et al.*, 1973).

When compared with the basic sequences of human κ chains (Milstein & Pink, 1970), the V-region sequence of L chain BS-1 exhibits 66%, 53% and 54% homology respectively with K_I, K_{II} and K_{III} sequences (the comparison does not include hypervariable positions). Sections 20-30, 34-41, 44-49 and 60-88 of L chain BS-1 are almost identical with their human counterparts. Thus L chain BS-1 appears to be more similar to V_{KI} prototype sequence than to V_{KII} and to V_{KIII} basic sequences. The degree of homology between the L chain V region subgroups of different species is similar to that found between the H chain V regions of various species (Ray & Cebara, 1972; Kehoe & Capra, 1972; Bourgois *et al.*, 1972; Jaton & Braun, 1972). The V region of L chain BS-1 was compared with the V region of anti-pneumococcal H chains (Fig. 10). Homology of these V regions is evident, as of 90 residues there are 28 identical residues in homologous positions. This confirms the original finding of Edelman (1970) on protein Eu and

suggests that V-region sequences from both H and L chains evolved from a common precursor.

On the basis of the great sequence variability observed in the N-terminal 20–30 positions of rabbit L chains, the existence of at least 6 V_K region subgroups was proposed (Braun & Jaton, 1973). Anti-pneumococcal and streptococcal L chain N-terminal sequences obtained to date distribute between several subgroups as do also anti-hapten L chains (Appella *et al.*, 1971; Freedman *et al.*, 1972). This suggests that there is no obligatory association between antibody specificity and L chain subgroups, even though preferential expression of a given subgroup has been noted for other antigens (Potter *et al.*, 1970; Capra *et al.*, 1972). Further, the L chains of a parent and of an offspring rabbit immunized with the same bacterial vaccine were found identical within their N-terminal 22 amino acid residues. These and additional data suggest preferential expression of inherited V-region genes for the synthesis of antipolysaccharide antibodies (Braun & Jaton, 1973).

Among mammalian species which have been studied so far, rabbit κ chains are unique in possessing three intrachain disulphide bridges (Rejnek *et al.*, 1969). One of them has been shown to link V and C domains (Poulsen *et al.*, 1972). The interdomain connecting bond is formed by a half-cystine in position 80 and its partner is found in position 171. These results (Fig. 9, Table 4) confirm and extend those recently reported by Lamm & Frangione (1972). The 'domain' hypothesis (Edelman *et al.*, 1969) predicts that V and C regions each containing one disulphide bond are folded independently as compact domains linked by loose connecting regions. This is quite consistent with the X-ray crystallographic model of a human Fab' fragment (Poljak *et al.*, 1972). As the disulphide bonds are known to be an important factor for maintaining and stabilizing the conformation of polypeptide chains, one can wonder whether alteration in the number and in the arrangement of the disulphide bridges would create a distortion of the three-dimensional structure of the molecule. The presence of the interdomain-connecting disulphide bond 80–171 suggests that the region around cysteine-80 should come very close to the cysteine-171 region. This indeed was demonstrated by the X-ray analysis of a Bence-Jones protein dimer which lacks this disulphide bond (Schiffer *et al.*, 1973). The inspection of the X-ray model at 0.28 nm (2.8 Å) resolution (Poljak *et al.*, 1973) suggests that the presence of that bridge would not impose a conformational constraint on the folding of the V and C domains of the L chain.

The presence of two clusters of 3 glycine residues in a row in positions 94–96 and 99–101 is particularly striking. The glycine residues in positions 99–101 are essentially invariant residues in human L chains. They may provide flexibility of the polypeptide chain

to allow a better fit for the V_L region in contact with V_H region, once an antigenic determinant reacts with the antibody-combining site (Wu & Kabat, 1970). Glycine residues at positions 94–96 are unique to this protein and are located in the hypervariable area. They may therefore be related to the specificity of this antibody. In other antibody L chains of similar but not necessarily identical specificity, glycine residues 94–96 are substituted by other residues, whereas those in positions 99–101 are invariant. (Jaton, 1974, and unpublished work). The X-ray-diffraction data, in conjunction with amino acid-sequence data, are expected to reveal which amino acid substitutions are permissible within the V regions without markedly altering the three-dimensional structure of the antibody molecule.

I thank Dr. D. G. Braun and Dr. J. R. L. Pink for helpful discussions and Mr. W. Hunkeler for valuable technical assistance.

References

- Ambler, R. P. (1967) *Methods Enzymol.* **11**, 135–166
 Appella, E., Rejnek, J. & Reisfeld, R. A. (1969) *J. Mol. Biol.* **41**, 473–477
 Appella, E., Chersi, A., Roholt, O. A. & Pressman, D. (1971) *Proc. Nat. Acad. Sci. U.S.A.* **68**, 2569–2573
 Bourgois, A., Fougereau, M. & De Preval, C. (1972) *Eur. J. Biochem.* **24**, 446–455
 Braun, D. G. & Jaton, J. C. (1973) *Immunochemistry* **10**, 387–395
 Braunitzer, G., Schrank, B. & Ruhfus, A. (1970) *Hoppe-Seyler's Z. Physiol. Chem.* **351**, 1589–1590
 Bray, G. A. (1960) *Anal. Biochem.* **1**, 279–285
 Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228
 Capra, J. D., Kehoe, J. M., Williams, R. C., Jr., Feizi, T. & Kunkel, H. G. (1972) *Proc. Nat. Acad. Sci. U.S.A.* **69**, 40–43
 Dreyer, W. J. & Bynum, E. (1967) *Methods Enzymol.* **11**, 32–39
 Edelman, G. M. (1970) *Biochemistry* **9**, 3197–3204
 Edelman, G. M. & Gall, W. E. (1969) *Annu. Rev. Biochem.* **38**, 415–466
 Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, V. & Waxdal, M. J. (1969) *Proc. Nat. Acad. Sci. U.S.A.* **63**, 78–85
 Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
 Fleischman, J. B. (1971) *Biochemistry* **10**, 2753–2760
 Fleischman, J. B. (1973) *Immunochemistry* **10**, 401–407
 Frangione, B. (1969) *FEBS Lett.* **3**, 341–342
 Fraser, K. J., Poulsen, K. & Haber, E. (1972) *Biochemistry* **11**, 4974–4977
 Freedlender, E. F. & Haber, E. (1972) *Biochemistry* **11**, 2362–2370
 Freedman, M. H., Guyer, R. B. & Terry, W. D. (1972) *J. Biol. Chem.* **247**, 7051–7061
 Gibbons, I. & Perham, R. N. (1970) *Biochem. J.* **116**, 843–849
 Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475

- Habeeb, A. F. S. A. (1966) *Anal. Biochem.* **14**, 328-336
- Habeeb, A. F. S. A. & Atassi, M. Z. (1970) *Biochemistry* **9**, 4939-4944
- Hood, L., Eichmann, K., Lackland, H., Krause, R. M. & Ohms, J. J. (1970) *Nature (London)* **228**, 1040-1044
- Hood, L., Waterfield, M. D., Morris, J. E. & Todd, Ch.W. (1971) *Ann. N.Y. Acad. Sci.* **190**, 26-35
- Jaton, J. C. (1974) *Biochem. J.* **141**, 15-25
- Jaton, J. C. & Braun, D. G. (1972) *Biochem. J.* **130**, 539-546
- Jaton, J. C., Waterfield, M. D., Margolies, M. N. & Haber, E. (1971) *Biochemistry* **10**, 1583-1587
- Jaton, J. C., Braun, D. G., Strosberg, A. D., Haber, E. & Morris, J. E. (1973) *J. Immunol.* **111**, 1838-1843
- Kehoe, J. M. & Capra, J. D. (1972) *Proc. Nat. Acad. Sci. U.S.* **69**, 2052-2055
- Kimball, J. W., Pappenheimer, A. M., Jr. & Jaton, J. C. (1971) *J. Immunol.* **106**, 1177-1184
- Kindt, T. J., Seide, R. K., Lackland, H. & Thunberg, A. L. (1972) *J. Immunol.* **109**, 735-741
- Lamm, M. E. & Frangione, B. (1972) *Biochem. J.* **128**, 1357-1359
- McKean, D., Potter, M. & Hood, L. (1973) *Biochemistry* **12**, 749-759
- Milstein, C. & Pink, J. R. L. (1970) *Progr. Biophys. Mol. Biol.* **21**, 209-263
- Mole, L. E., Jackson, S. A., Porter, R. R. & Wilkinson, J. M. (1971) *Biochem. J.* **124**, 301-318
- O'Donnell, I. J., Frangione, B. & Porter, R. R. (1970) *Biochem. J.* **116**, 261-268
- Offord, R. E. (1966) *Nature (London)* **211**, 591-593
- Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597-5607
- Poljak, R. J., Amzel, L. M., Avey, H. P., Becka, L. N. & Nisonoff, A. (1972) *Nature (London) New Biol.* **235**, 137-140
- Poljak, R. J., Amzel, L. M., Avey, H. P., Chen, B. J., Phizackerley, R. P. & Saul, F. (1973) *Proc. Nat. Acad. Sci. U.S.* **70**, 3305-3310
- Potter, M., Lieberman, R., Hood, L. & McKean, D. J. (1970) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **29**, 437
- Poulsen, K., Fraser, K. J. & Haber, E. (1972) *Proc. Nat. Acad. Sci. U.S.* **69**, 2495-2499
- Press, E. M., Piggot, P. J. & Porter, R. R. (1966) *Biochem. J.* **99**, 356-366
- Putnam, F. W., Titani, K., Wikler, M. & Shinoda, T. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32**, 9-29
- Ray, A. & Cebra, J. J. (1972) *Biochemistry* **11**, 3647-3657
- Rejnek, J., Appella, E., Mage, R. G. & Reisfeld, R. A. (1969) *Biochemistry* **8**, 2712-2718
- Schiffer, M., Girling, R. L., Ely, K. R. & Edmundson, A. B. (1973) *Biochemistry* **12**, 4620-4631
- Smith, I. (1960) in *Chromatographic and Electrophoretic Techniques* (Smith, I., ed.), vol. 1, pp. 96-97, Interscience, New York
- Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. C. & Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912-4921
- Strosberg, A. D., Fraser, K. J., Margolies, M. N. & Haber, E. (1972) *Biochemistry* **11**, 4978-4985
- Wu, T. T. & Kabat, E. A. (1970) *J. Exp. Med.* **132**, 211-250