Comparison of the Amino Acid Sequences of the Variable Regions of Light Chains Derived from Two Homogeneous Rabbit Anti-Pneumococcal Antibodies

By JEAN-CLAUDE JATON

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

(Received 30 November 1973)

The amino acid sequence of the *N*-terminal 139 residues of the L (light) chain derived from a homogeneous rabbit antibody to type III pneumococci was determined. This L chain, designated BS-5, exhibits a greater degree of homology with the basic sequence of human κ chains of subgroup I (72%) than with subgroups II and III. L-chain BS-5 differs from another L chain (BS-1), also derived from an antibody to type III pneumococci (Jaton, 1974), by eight amino acid residues, even though the chains are identical within the *N*-terminal 30 residues. Six of these eight substitutions are located within the three hypervariable sections of the variable half: Asn/Ser in position 31, Glu/Ala in position 55, Asx/Thr, Thr/Gly, Thr/Gly and Val/Tyr in positions 92, 94, 96 and 97 respectively. The two anti-pneumococcal L chains BS-1 and BS-5 are much more similar to each other than to an anti-azobenzoate L chain (Appella *et al.*, 1973), from which they differ by 30 and 29 residues respectively. Of these interchanges 13–15 are confined to the three hypervariable sections, and 11 occur within the *N*-terminal 27 positions. The three chains have an identical sequence from residue 98 to residue 139, except for a possible inversion of two residues in positions 130–131 of the anti-azobenzoate chain.

The light (L)* chains of immunoglobulins comprise a variable *N*-terminal half of about 110 residues and an invariant *C*-terminal half (Hilschmann & Craig, 1965; Putnam, 1969). The constant half (C region) is identical for human κ light chains except for an amino acid substitution at position 191, which differs according to the allotype. In contrast, the variable halves (V regions) of any two human κ chains differ by a minimum of about eight residues and a maximum of 60 residues (Hilschmann, 1969). Three subgroups of human κ chains have been defined on the basis of similarity in amino acid sequence of their V regions (Milstein, 1967).

Homogeneous rabbit antibodies stimulated by bacterial polysaccharide antigens appear to be ideal material for structural studies (Hood *et al.*, 1970; Jaton *et al.*, 1971; Jaton & Braun, 1972; Strosberg *et al.*, 1972; Braun & Jaton, 1973). The amino acid sequence of the *N*-terminal 139 residues of the L chain derived from a homogeneous antibody to type III pneumococci has been completed (Jaton, 1974).

* 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;

CmCys, γ -Glu (in amino acid sequences and Tables), *S*-carboxymethylcysteine and pyrrolidonecarboxylic acid respectively. V_{KI} , V_{KII} and V_{KIII} , subgroups I, II and III respectively of V regions of κ L chains; the nomenclature for subgroups of human κ chains is that used by Milstein & Pink (1970).

Among mammalian species that have been studied so far, rabbit κ chains are unique in possessing three intrachain disulphide bridges: one V-region bridge between positions 23 and 88, one C-region bridge between positions 134 and 194, and the third one connects V and C regions between positions 80 and 171. When compared with the basic sequence of human κ chains, this L chain (designated BS-1) appears to be more similar to V_{KI} prototype sequence than to V_{KII} or V_{KIII} sequences.

It is well documented that the structure of the variable regions of both H and L chains determines the specificity of the antibody (Porter, 1971). The availability of structurally different antibodies raised to the same antigen (e.g. type III pneumococcal polysaccharide) from different individuals may reveal which sections of the V region of L and H chains are related to antibody complementarity, and what is the specific nature of sequence variability relative to the structure of the antigenic determinant. This system may also enable us to estimate the number of alternate solutions to the binding of a single antigenic determinant (Haber, 1971). In an attempt to determine the complete sequence of the V regions of a few antibodies of defined specificity, I describe here the amino acid sequence of the V region and of the beginning of the C domain of the L chain derived from another homogeneous rabbit antibody to type III pneumococci, designated BS-5. The amino acid sequence of the *N*-terminal 69 residues of the H

chain derived from antibody BS-5 has been reported previously (Jaton & Braun, 1972).

Materials and Methods

Enzymes, radiochemicals and reagents used for the isolation of peptides and for the sequence determination were described previously (Jaton, 1974).

Antibody production and purification

Rabbit BS-5 was immunized with type III pneumococcal vaccine according to a scheme described by Kimball *et al.* (1971). Details of the method of purification and preparation of heavy and light chains have been described by Jaton & Braun (1972). The antibody BS-5 carried a_1 , b_4 allotypes (kindly determined by Dr. A. S. Kelus, Basel Institute for Immunology, Basel, Switzerland).

Complete reduction of the L chain

The light chain (75 mg) was fully reduced and alkylated with iodo[2- 14 C]acetic acid by the technique of O'Donnell *et al.* (1970).

Citraconylation of the reduced and [2- 14 C]alkylated L chain

Blocking of lysine residues (Gibbons & Perham, 1970), tryptic digestion of the fully citraconylated L chain and separation of the components of the digest by gel filtration on Sephadex G-50 were done exactly as described previously (Jaton, 1974).

Dilute acid hydrolysis

Citraconylated fragment A (2 μ mol) derived from L chain BS-5 was subjected to acid hydrolysis in 7 M-guanidine hydrochloride containing 10% (v/v) acetic acid, pH 2.5 (2 ml), for 120 h at 37°C (Fraser *et al.*, 1972).

Unblocking of citraconylated peptides

The citraconylated peptides were incubated in 2 ml of pyridine-acetate buffer, pH 3.5, containing 5 M-guanidine hydrochloride, for 4 h at 37°C; the solution was then dialysed against water overnight and the peptide freeze-dried.

Analytical methods

Preparative high-voltage paper electrophoresis, detection of peptides, enzymic digestion, amino acid analyses and the determination of sequences of small peptides by the dansyl-Edman procedure (Gray,

1967) were done as described by Jaton (1974). Amide residues were assigned on the basis of electrophoretic mobility at pH 6.5 (Offord, 1966). The determination of the sequence of large peptides was carried out in a Beckman Sequencer, model 890B, equipped with an undercut cup and N₂ flush. For every experiment, peptides (250–400 nmol) were dissolved in 1 M-acetic acid and degraded by using either the Quadrol programme (Edman & Begg, 1967) or the volatile-peptide programme (dimethylallylamine buffer) after prior treatment of the peptides with 4-sulphophenyl isothiocyanate to prevent losses during automated sequential degradation (Braunitzer *et al.*, 1970). The phenylthiohydantoin-amino acid derivatives were identified by g.l.c. (Pisano & Bronzert, 1969) and by amino acid analysis after conversion of the phenylthiohydantoin derivatives into free amino acids by hydrolysis with 65% (w/v) HI for 20 h at 125°C (Smithies *et al.*, 1971). The overall repetitive yield in sequencing experiments varied from 90% to 94% and the absolute yield of the phenylthiohydantoin-amino acid derivatives at the first step ranged between 25 and 40%.

Results

Characterization of peptides from a tryptic digest of [2- 14 C]carboxymethylated and citraconylated L chain BS-5

L chain BS-5 contains 2 arginine residues/molecule (Table 1). By analogy with the known structure of the V region of L chain BS-1, it was expected that one of the two arginine residues would be located at position 61 in the V region and the second must occupy the fourth position from the C-terminal end of the chain (Appella *et al.*, 1969; Jaton, 1974). L chain BS-5 was therefore cleaved by trypsin after all lysine residues had been blocked with citraconic anhydride. The arginine-containing peptides of the tryptic digest of [2- 14 C]carboxymethylated and citraconylated L chain were separated into two fractions, A and B, by gel filtration on Sephadex G-50 in the presence of 5 M-guanidine hydrochloride-Tris buffer, pH 8.0. The elution profile is similar to that obtained previously for L chain BS-1 (see Fig. 1 of Jaton, 1974). The amino acid analyses of the fragments indicate that fragment A contains about 150 amino acids, whereas fragment B has 61 residues, among which is one half-cystine (Table 1). On the basis of 3.4 μ mol of L chain BS-5, fragments A and B were recovered in 79% yield, which is in good agreement with that obtained for corresponding peptides derived from L chain BS-1 (Jaton, 1974).

After decitraconylation of fragment B, the de-blocked peptide was subjected to 36 cycles of automated Edman degradation. The amino acid sequence of the N-terminal section of peptide B,

Table 1. *Amino acid composition of the citraconylated peptides isolated from a tryptic digest of reduced and carboxymethylated L chain BS-5*

Values are residues/molecule of peptide. Values in parentheses are integral values confirmed by sequence analysis. Fragments A₁ and A₃ were released from citraconylated 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	6.0	4.2	2.0 (2)	4.3 (4)	10.7
His	1.0	1.2	—	—	1.3
Arg	1.1	0.9	—	1.1 (1)	2.0
CmCys	5.7	3.4	1.8 (2)	1.1 (1)	6.8
Asp	16.6	12.2	4.5 (5)	2.1 (2)	18.5
Thr	24.0	18.4	6.8 (7)	4.5 (5)	30.0
Ser	14.2	8.3	3.9 (4)	7.7 (8)	20.5
Glu	13.8	9.3	3.3 (3)	8.2 (8)	21.0
Pro	6.3	6.5	0.3 (0)	6.3 (6)	12.0
Gly	14.3	5.3	8.8 (9)	4.6 (4)	19.5
Ala	9.9	6.3	3.0 (3)	3.9 (4)	14.8
Val	14.9	11.6	4.2 (4)	5.8 (6)	20.5
Met	—	—	—	0.9 (1)	1.2
Ile	3.6	2.9	1.1 (1)	2.8 (3)	6.7
Leu	6.0	3.8	2.0 (2)	4.1 (4)	10.2
Tyr	7.2	4.0	1.8 (2)	3.0 (3)	9.3
Phe	7.7	4.1	3.8 (4)	— (0)	7.3
Trp	N.D.	N.D.	—*	+ (1)*	+
Total residues	152.3	102.4	48.0	60.4	212.3

* Tryptophan measured with Ehrlich reagent.

Residue position	1	10	20
L-chain BS-5*	Asp-Val-Val-Met-Thr-Gln-Thr-Pro-Ala-Ser-Val-Ser-Glu-Pro-Val-Gly-Gly-Thr-Val-Thr		
Fragment B	7080		
Fragment A	Phe-Lys-Gly-Ser-Gly-Ser-Gly-Thr-Asp-Phe-Thr-Leu-Thr-Ile-Ser-Asp-Leu-Glu-Cys†-Ala		
Fragment A ₁	110120Pro-Val-Ala-Pro-Thr-Val-Leu-Ile-Phe-Pro-Pro-Ala-Ala-Asp-Gln-Val-Ala-Thr-Gly-Thr		
Residue position	30		
L-chain BS-5*	Ile-Lys-Cys†-Gln-Ala-Ser-Gln-Ser-Ile		
Fragment B	Tyr-Ser-Asn-Leu-Ala-Trp-Tyr		
Fragment A	Asp-Ala-Ala-Thr-Tyr-Phe-Cys†-Glx		
Fragment A ₁	130Val-Thr-Ile-Val-Cys†-Val-Ala-Asn-Lys-Tyr		

Fig. 1. *Automated sequential degradation of fragments A, A₁ and B derived from L chain BS-5*

Fragments B (400 nmol), A (400 nmol) and A₁ (300 nmol) respectively were used for sequence determination. A conventional Quadrol programme was used for fragments A and A₁, whereas a volatile buffer system and sulphophenyl isothiocyanate was employed for peptide B (see the Materials and Methods section for details). The solid line indicates a sequence identical with that shown immediately above. The numbering of residues in the fragments indicates their position in the complete sequence.

* Braun & Jaton (1973).

† Determined by counting a portion of the phenylthiohydantoin-S-carboxymethylcysteine derivative for radioactivity.

shown in Fig. 1, is identical with the N-terminal sequence of L chain BS-5 previously described (Braun & Jaton, 1973). Fragment B is thus the N-terminal 61-residue peptide of the L chain. Citra-

conylated fragment A was subjected to automated Edman degradation. A single amino acid derivative was found at each of the N-terminal 28 positions (Fig. 1). By homology with the known sequence of

L chain BS-1, peptide A follows fragment B and extends between positions 62 and 211.

Partial sequence of fragment A

The large fragment A was further subjected to mild acid cleavage in order to break the Asp₁₀₉-Pro₁₁₀ bond between V and C domains of rabbit κ chains (Fraser *et al.*, 1972). Fig. 2 illustrates the typical separation of the components into two major fractions, designated A₁ and A₃, as previously described for L chain BS-1 (Jaton, 1974). The small amount of material in fraction A₂ was not analysed. The amino acid composition of peptides A₁ and A₃, obtained in 73% and 66% yield respectively, is given in Table 1. Fragment A₁, which contains one arginine and three half-cystine residues, was subjected to 30 cycles of Edman degradation; a unique sequence was found (Fig. 1), which is identical with that of the homologous peptide obtained from the C region of L chain BS-1 (residues 110-211). Peptide A₃ contains 48 amino acids, among which are two half-cystine residues, but neither arginine nor tryptophan. Its N-terminal sequence Phe-Lys-Gly-Ser, determined by the dansyl-Edman procedure, indicates that fragment A₃ is the N-terminal peptide of fragment A and extends from positions 62 to 109.

Sequence determination of fragment A₃

Fragment A₃ was digested with thermolysin for 1 h at 25°C and five peptides were isolated by paper electrophoresis at pH 6.5 and 3.5 (Fig. 3). Their amino acid compositions and mobilities are sum-

marized in Table 2. The neutral peptides ThN₁ and ThN₂ were repurified at pH 3.5. Their N-terminal sequences are given in Fig. 3.

Peptide Th₁ is a dodecapeptide which contains one lysine residue. Tryptic digestion of peptide Th₁ yielded the C-terminal dipeptide Gly-Asp and the N-terminal decapeptide T-Th_{1b}, whose partial sequence was Phe-Gly-Gly-Gly-Thr-Glu-Val. The C-terminal sequence of this peptide was found by digestion with carboxypeptidases B and A to be

Fig. 2. 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 (30 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. —, E_{280} ; ---, radioactivity (c.p.m./20 μ l of eluate).

Fig. 3. Amino acid sequence of the fragment A₃ of L chain BS-5 (residues 62-109)

For experimental details, see the text and Table 2. T, Th, and Ch are tryptic, thermolysin and chymotryptic peptides respectively. → Amino acid residues identified by the dansyl-Edman procedure; ⇒ identified as dansyl residue without hydrolysis; ← identified by carboxypeptidase A and/or B. Residues identified with the sequencer are also shown.

Table 2. Amino acid composition and properties of constituent peptides of the fragment A₃ of L chain BS-5

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 700 nmol of fragment A₃.

Amino acid	Thermolysin peptides					Chymotryptic peptides			Tryptic peptides	
	Th ₁	Th ₂	Th ₃	ThN ₁	ThN ₂	Ch-Th _{2a}	Ch-Th _{2b}		T-Th _{1a}	T-Th _{1b}
Lys	1.1	—	—	—	1.2	—	—		—	1.0 (1)
Arg	—	—	—	—	—	—	—		—	—
CmCys*	—	0.8 (1)	0.9	—	—	—	0.7 (1)		—	—
Asp	0.9	1.0 (1)	2.0	—	1.0	—	—		1.0 (1)	—
Thr	1.0	1.8 (2)	0.9	2.1 (2)	1.2	1.9 (2)	0.9 (1)		—	0.9 (1)
Ser	—	1.2 (1)	1.2	—	1.9	—	1.2 (1)		—	—
Glu	1.1	1.1 (1)	1.0	—	—	—	1.1 (1)		—	1.0 (1)
Pro	—	—	—	—	—	—	—		—	—
Gly	3.9	1.7 (2)	—	—	3.0	1.1 (1)	0.9 (1)		0.9 (1)	2.8 (3)
Ala	—	—	2.7	—	—	—	—		—	—
Val	2.7†	0.9 (1)	—	—	—	0.9 (1)	—		—	2.9 (3)†
Ile	—	—	0.9	—	—	—	—		—	—
Leu	—	—	1.1	0.9 (1)	—	—	—		—	—
Tyr	—	0.8 (1)	0.9	—	—	—	1.0 (1)		—	—
Phe	1.1	0.9 (1)	—	1.2 (1)	1.0	—	1.1 (1)		—	0.9 (1)
Total residues	12	11	12	4	9	4	7		2	10
Mobility at pH6.5	-0.22	-0.62	-0.72	0	0	0	-0.68		-0.79	0
Yield (%)	25	33	33	29	30	27	23		13	16

* Value not corrected for losses during hydrolysis.
† Value after 96h hydrolysis; after 24h the value found was 2.2.

Table 3. Amino acid composition and properties of constituent peptides of the N-terminal fragment B of L chain BS-5

Values are residues/molecule of peptide. Values in parentheses are integral values confirmed by sequence analysis. ChT₃₋₃, 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 1.5 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	ChT ₃₋₁	ChT ₃₋₂	ChT ₃₋₃	ChT ₄₋₁	ChT ₄₋₂	
Lys	1.1	1.1	2.1	—	1.0 (1)	1.1 (1)	—	—	2.1 (2)	—	—	
Arg	—	—	—	1.1 (1)	—	—	—	—	—	—	0.9 (1)	
CmCys*	—	0.8	0.8	—	—	—	0.7 (1)	—	—	—	—	
Asp	1.0	1.0	1.2	—	—	—	—	1.0 (1)	—	—	—	
Thr	4.1	—	—	1.1 (1)	—	—	—	—	—	1.1 (1)	—	
Ser	2.1	2.9	2.9	3.2 (3)	—	—	1.8 (2)	1.1 (1)	—	1.0 (1)	1.9 (2)	
Glu	2.0	3.9	5.0	1.1 (1)	1.0 (1)	—	2.2 (2)	—	3.0 (3)	—	1.0 (1)	
Pro	2.1	—	3.2	0.9 (1)	2.8 (3)	—	—	—	2.9 (3)	—	0.9 (1)	
Gly	2.0	—	1.1	1.1 (1)	1.3 (1)	—	—	—	0.9 (1)	1.2 (1)	—	
Ala	1.1	2.1	2.2	1.1 (1)	—	—	1.0 (1)	0.9 (1)	—	—	—	
Val	4.8	—	—	1.0 (1)	—	—	—	—	—	—	0.8 (1)	
Met	0.9	—	—	—	—	—	—	—	—	—	—	
Ile	1.0	1.0	1.1	—	—	0.9 (1)	0.9 (1)	—	—	—	—	
Leu	—	1.1	1.1	1.0 (1)	—	2.0 (2)	—	1.1 (1)	—	1.0 (1)	—	
Tyr	—	1.9	2.0	—	—	1.1 (1)	1.1 (1)	0.9 (1)	—	—	—	
Trp	—	++	++	—	—	—	—	++ (1)	—	—	—	
Total residues	22	17	23‡	11	6	5	8	6‡	9	4	7	
Mobility at pH6.5	-0.15	0	+0.10	0	+0.38	+0.41	-0.61	0	—	0	0	
Yield (%)	15	12	23	25	10	50	17	16	20	12	13	

* Value not corrected for losses during hydrolysis.
† Measured with Ehrlich reagent.
‡ Tryptophan assumed to be 1 residue/molecule of peptide.

	1	10	20	25
BS-5	Asp-Val-Val-Met-Thr-Gln-Thr-Pro-Ala-Ser-Val-Ser-Glu-Pro-Val-Gly-Gly-Thr-Val-Thr-Ile-Lys-Cys-Gln-Ala			
BS-1				
2717	Val-Glu-----Leu-----30	Ser-Pro-----Ala-Ala-----40	Ser-----Ser-----50	
BS-5	Ser-Gln-Ser-Ile-Tyr-Asn-----Gly-Ileu-Ala-Trp-Tyr-Glx-Glx-Lys-Pro-Gly-Gln-Pro-Pro-Lys-Leu-Ile-Tyr-Lys			
BS-1				
2717	Thr-Lys-----Asx-Asx-Tyr-----60		Ala-----Ala-----70	Thr-----Thr
BS-5	Ala-Ser-Thr-Leu-Glu-Ser-Gly-Val-Pro-Ser-Arg-Phe-Lys-Gly-Ser-Gly-Ser-Gly-Thr-Asp-Phe-Thr-Leu-Thr-Ile			
BS-1				
2717	-----Ala-----Ser-----80	-----Thr-----90	Glu-----Glx-----100	Leu-----Leu
BS-5	Ser-Asp-Leu-Glu-Cys-Ala-Asp-Ala-Ala-Thr-Tyr-Phe-Cys-Glx-Gly-Ser-Asx-Tyr-Thr-Gly-Thr-Val-Phe-Gly-Gly			
BS-1				
2717	-----Val-----Asp-----101	-----Tyr-----Gly-----Ala-----110	-----Gly-----Gly-Tyr-----120	-----Tyr-Ser-----125
BS-5	Gly-Thr-Glu-Val-Val-Lys-Gly-Asp-Pro-Val-Ala-Pro-Thr-Val-Leu-Ile-Phe-Pro-Pro-Ala-Ala-Asp-Gln-Val			
BS-1				
2717				
BS-5	Ala-Thr-Gly-Thr-Val-Thr-Ile-Val-Cys-Val-Ala-Asn-Lys-Tyr			
BS-1				
2717	-----Thr-Val-----130			-----139

Fig. 5. Amino acid sequence of the V_L regions of three rabbit antibody L chains

L chains BS-5 (this paper) and BS-1 (Jaton, 1974) are derived from antibodies to type III pneumococci. Their sequence comprises the N-terminal fragment B (residues 1-61), fragment A₃ (residues 62-109) and part of the fragment A₁, subjected to automated Edman degradation (residues 110-139). L chain 2717 was obtained from an anti-azobenzene antibody (Appella *et al.*, 1973). The solid line indicates a sequence identical with that shown immediately above. Deletions (----) are introduced to maximize homology.

Val-Lys (Fig. 3); hence the third valine residue in peptide T-Th₁₆ was assumed on the basis of the amino acid analysis which indicates 2.9 residues of valine/molecule of peptide (Table 2).

Peptide Th₂ contains 11 amino acids, among which is one *S*-carboxymethylcysteine residue. Its *N*-terminal sequence, Phe-CmCys-Glx-Gly-Ser, indicates that peptide Th₂ extends between positions 87 and 97, a region known to be hypervariable. Peptide Th₂ was cleaved by chymotrypsin into two parts: peptide Ch-Th_{2b} was shown to be its *N*-terminal fragment (Fig. 3), and the tetrapeptide Ch-Th_{2a}, which has the sequence Thr-Gly-Thr-Val, is its *C*-terminal peptide.

Peptide Th₃ is the other radioactive peptide of the digest. Its *N*-terminal sequence Ile-Ser-Asp-Leu-Glu-CmCys indicates that it is the cysteine-80-containing peptide of fragment A₃ (Fig. 3). The sequence of the first 28 residues of fragment A analysed by the protein sequencer overlaps peptides ThN₂, ThN₁, Th₃ and Th₂, in this order. Peptide Th₁ is therefore the *C*-terminal peptide of fragment A₃.

Sequence of fragment B

After removal of citraconyl groups, fragment B was subjected to trypsin digestion and the components of the mixture were separated by high-voltage paper electrophoresis at pH6.5 and 3.5. The characterization of the 6 tryptic peptides isolated in this way is given in Table 3 and Fig. 4.

Peptide T₁ contains 22 amino acids, among which is the unique methionine found in L chain BS-5 in position 4 from the *N*-terminus (Braun & Jaton, 1973). Peptide T₁ is therefore the *N*-terminal fragment of fragment B.

Peptides T₂ and T₃ have the same *N*-terminal sequence CmCys-Gln-Ala, which characterizes the region containing the half-cystine in position 23. Peptide T₃ is a 23-residue peptide, whereas peptide T₂ is shorter by six amino acids, and therefore they extend between positions 23 and 39 and 23 and 45 respectively. Peptides T₂ and T₃ were pooled and subjected to chymotrypsin digestion. Three major peptides, ChT₃₋₁, ChT₃₋₂ and ChT₃₋₃, were isolated in satisfactory yield. Their amino acid compositions and mobilities were identical with those found for the homologous chymotryptic peptides derived from L chain BS-1, except for one interchange in peptide ChT₃₋₂, which did not affect its mobility. The *N*-terminal sequence of the octapeptide ChT₃₋₁, CmCys-Glx-Ala, and its amino acid analysis, indicate that this peptide extends between positions 23 to 30. The hexapeptide ChT₃₋₂ is neutral and contains one tyrosine and one tryptophan residue; its *N*-terminal sequence is Ser-Asn-Leu-Ala(Trp,Tyr). Tryptophan and tyrosine were identified in positions 35

and 36 respectively, when fragment B was subjected to automated Edman degradation. Six cycles of dansyl-Edman degradation established the *N*-terminal sequence of peptide ChT₃₋₃ to be Glx-Glx-Lys-Pro-Gly-Glx. This peptide contains two lysine residues, one of which is the *C*-terminal residue of the peptide, and the other is the internal lysine-39 adjacent to proline-40 (Fig. 4). The peptide bond between positions 39 and 40 has been shown to be particularly resistant to trypsin digestion (Jaton, 1974). The partial split of this bond is suggested by the isolation of peptide T₅, which has the sequence Pro-Gly-Gln-Pro-Pro-Lys (Fig. 4).

The neutral peptide T₄ is the only arginine-containing peptide of the digest. It is therefore the *C*-terminal peptide of fragment B. Peptide T₄ was further fragmented with chymotrypsin and two peptides, ChT₄₋₁ and ChT₄₋₂, were isolated by electrophoresis at pH3.5. Peptide ChT₄₋₁ has the sequence Ala-Ser-Thr-Leu, whereas six Edman degradations of the peptide ChT₄₋₂ and cleavage of the *C*-terminal residue by carboxypeptidase B established its sequence to be Glu-Ser-Gly-Val-Pro-Ser-Arg.

The basic peptide T₆ was isolated in good yield. Four cycles of Edman degradation and digestion with carboxypeptidases A and B established its sequence to be Leu-Leu-Ile-Tyr-Lys. By homology with the sequence of L chain BS-1 (Jaton, 1974), peptide T₆ extends between peptide T₃ and the *C*-terminal peptide T₄.

The sequence of the *N*-terminal 36 residues of fragment B analysed by the protein sequencer overlaps peptides T₁, T₂ and T₃ and establishes their order. The determination of the sequence of the constituent peptides of fragment B is summarized in Fig. 4. The amino acid composition of the two V-region peptides A₃ and B agrees well with that computed from the sequence of the fragments (Table 1). The amino acid sequence of the *N*-terminal 139 residues of L chain BS-5 is presented in Fig. 5.

Discussion

Rabbit L chain BS-5 was selected for sequence determination because (1) it is derived from an antibody to type III pneumococcal polysaccharide, thus offering a direct comparison with the V region of another anti-pneumococcal type III L chain recently completed (Jaton, 1974), and (2) because of the presence of a single arginine residue in the V region which greatly facilitates the determination of the primary structure. The procedures used here are identical with those described for another rabbit L chain (Jaton, 1974). Only 3 μ mol of L chain BS-5 was needed to establish the sequence of its *N*-terminal 139 residues.

A comparison of the V_K regions of three rabbit L chains is presented in Fig. 5. L Chains BS-5 and BS-1, which were isolated from two distinct antibodies to type III pneumococci, are identical within their *N*-terminal 30 positions. They belong to the same rabbit κ -chain subgroup (Braun & Jaton, 1973) and are very similar throughout their V regions: indeed 101 of 109 residues are identical (93% homology). They differ from each other, however, by eight residues in positions 31, 55, 59, 70, 92, 94, 96 and 97 (Fig. 5). All these substitutions occur within the three recognized hypervariable sections, i.e. in positions 24–34, 50–56 and 89–97 (Wu & Kabat, 1970) except for 2 residues, in positions 59 (Pro/Ser) and 70 (Asp/Glu).

The complete sequence of the L chain isolated from a restricted rabbit antibody (designated 2717) to the azobenzoate group has been reported (Appella *et al.*, 1973). The V region of L chain 2717 exhibits 29 amino acid interchanges when compared with L chain BS-5, among which 13 (approx. 50%) are located within the three hypervariable sections (Wu & Kabat, 1970). L chain 2717 differs from L chain BS-1 by 30 residues, of which half belong also to the hypervariable sections (Fig. 5). Within their *N*-terminal 27 positions, the anti-pneumococcal and anti-azobenzoate L chains differ by 11 residues; it has been suggested that sequence variability in this region is not related to antibody specificity (Hood *et al.*, 1970; Braun & Jaton, 1973; Schiffer *et al.*, 1973). By analogy with the occurrence of characteristic linked groups of residues in the *N*-terminal sections of human κ chains (Niall & Edman, 1967; Milstein, 1967; Hilschmann, 1969), some of the amino acid substitutions in rabbit L-chain *N*-termini are likely to be subgroup-specific (Hood *et al.*, 1970). The finding of additional L chains belonging to the same subgroup as the anti-azobenzoate L chain 2717 will help clarify this point. The anti-azobenzoate L chain is characterized by an insertion of two aspartic acid or asparagine residues in the first hypervariable section. This is suggestive evidence that it belongs to a different subgroup of rabbit κ chains, perhaps similar to human V_{KII} chains, which exhibit an insertion of four to six residues compared with V_{KI} or V_{KIII} chains (Milstein & Pink, 1970). Beyond residue 97, the three L chains have an identical sequence, until residue position 139, except for a possible inversion of Thr-Val in positions 130–131 of the anti-azobenzoate chain. More sequence data are required before we can establish the invariance of the C-terminal end of the V region of rabbit L chain.

When compared with the basic sequences of human κ chains (Milstein & Pink, 1970), L chain BS-5 exhibits 72% homology with the prototype sequence of V_{KI} (i.e. 68 identical residues in 95 positions), and 55 and 57% homology with V_{KII} and V_{KIII} basic sequences respectively. With L chain 2717,

the degree of homology is 61, 54 and 54% respectively. Sequence variations within the V regions of L chains may correlate, at least in part, with alterations of the biological properties of the native antibodies. It will be of interest to test whether the eight amino acid differences between the V regions of L chains BS-1 and BS-5 exert a marked influence on the binding specificity of the corresponding antibodies. Moreover, preliminary observations (J. Haimovich & J.-C. Jaton, unpublished work) suggest that amino acid substitutions in the V regions of the two anti-pneumococcal L chains BS-1 and BS-5 have a marked effect on the expression of the idiotype determinant(s) of the antibody molecule. Indeed, the work of Carson & Weigert (1973) shows that a few amino acid substitutions occurring in hypervariable sections of L chains may significantly alter the idotype.

Note Added in Proof (Received 19 April 1974)

The sequence of L chain 2717 contains typographical errors in positions 130–131. The corrected sequence should read Val₍₁₃₁₎–Thr₍₁₃₁₎ (E. Appella, personal communication). Therefore L chains BS-1, BS-5 and 2717 have identical sequences between positions 98 and 139.

The author thanks Dr. D. G. Braun for his critical review of the manuscript and Mr. W. Hunkeler for his excellent technical assistance.

References

- Appella, E., Rejnek, J. & Reisfeld, R. A. (1969) *J. Mol. Biol.* **41**, 473–477
- Appella, E., Roholt, O. A., Chersi, A., Radzimski, G. & Pressman, D. (1973) *Biochem. Biophys. Res. Commun.* **53**, 1122–1129
- 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
- Carson, D. & Weigert, M. (1973) *Proc. Nat. Acad. Sci. U.S.* **70**, 235–239
- Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
- Fraser, K. J., Poulsen, K. & Haber, E. (1972) *Biochemistry* **11**, 4974–4977
- Gibbons, I. & Perham, R. N. (1970) *Biochem. J.* **116**, 843–849
- Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475
- Haber, E. (1971) *Ann. N. Y. Acad. Sci.* **190**, 285–305
- Hilschmann, N. (1969) *Naturwissenschaften* **56**, 195–205
- Hilschmann, N. & Craig, L. C. (1965) *Proc. Nat. Acad. Sci. U.S.* **53**, 1403–1409
- Hood, L., Eichmann, K., Lackland, H., Krause, R. M. & Ohms, J. J. (1970) *Nature (London)* **228**, 1040–1044
- Jaton, J. C. (1974) *Biochem. J.* **141**, 1–13
- 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

- Kimball, J. W., Pappenheimer, A. M., Jr. & Jaton, J. C. (1971) *J. Immunol.* **106**, 1177–1184
- Milstein, C. (1967) *Nature (London)* **216**, 330–332
- Milstein, C. & Pink, J. R. L. (1970) *Progr. Biophys. Mol. Biol.* **21**, 209–263
- Niall, H. D. & Edman, P. (1967) *Nature (London)* **216**, 262–263
- 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
- Porter, R. R. (1971) *Harvey Lect.* **65**, 157–174
- Putnam, F. W. (1969) *Science* **163**, 633–644
- Schiffer, M., Girling, R. L., Ely, K. R. & Edmundson, A. B. (1973) *Biochemistry* **12**, 4620–4631
- Smithies, O., Gibson, D., Fanning, E. M., Goodfliesch, 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